IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 March 2024, accepted 22 April 2024, date of publication 26 April 2024, date of current version 3 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3394166

RESEARCH ARTICLE

Fault Detection and Classification in Ring Power System With DG Penetration Using Hybrid CNN-LSTM

AHMED SAMI ALHANAF[1], MURTAZA FARSADI[2], AND HASAN HÜSEYIN BALIK[3]
[1]Department of Computer Engineering, Yildiz Technical University (YTU), 34220 İstanbul, Turkey
[2]Department of Electrical Engineering, Faculty of Engineering, Istanbul Aydın University, 34295 İstanbul, Turkey
[3]Department of Computer Engineering, Faculty of Engineering, Istanbul Aydın University, 34295 İstanbul, Turkey

Corresponding author: Ahmed Sami Alhanaf (ahmed.alhanaf@std.yildiz.edu.tr)

ABSTRACT A modern electric power system integrated with advanced technologies such as sensors and smart meters is referred to as a ''smart grids'', aimed at enhancing electrical power delivery efficiency and reliability. However, fault location and prediction can become challenging when dynamic fault currents from renewable energy sources are present. To address these challenges, three unique deep learning models that make use of Deep Neural Networks (DNN) have been proposed. CNN, LSTM, and Hybrid CNN-LSTM are deep learning models. Line faulty identification (LF), fault classification (FC), and fault location estimate (FL) are the subjects on which they concentrate. These models analyze data gathered both pre and post faults occur in order to enhance decision making. Signals including the voltage and current were fed into these models from many different locations across the test networks. Once the 1D CNN has extracted characteristics from the gathered signals, LSTM uses these features to make accurate estimations and identify faults. Complex data are compatible with this method in terms of optimal outcomes. Using training and testing data from transmission line failure simulations, the proposed approaches were evaluated on the IEEE 6-bus and IEEE 9-bus systems. The tests encompassed a range of fault classes, locations, and ground fault resistances at various locations. Distributed Generator (DG) resources were additionally included in the system architecture and changes in the topology of the networks were considered in terms of location and number of DG resources. The results demonstrated that the proposed algorithms outperformed contemporary technologies in terms of detection, classification, and location accuracy. They demonstrated high accuracy and robustness in their performance.

INDEX TERMS Deep learning, smart grids, fault detection, fault classification and location, CNN, LSTM, hybrid CNN-LSTM.

I. INTRODUCTION

A. RESEARCH MOTIVATION

Contemporary Energy Systems Witness Growing Penetration of Renewable Resources and Heightened Complexity Across Distribution, Transmission, and Generation Components. All aimed at fulfilling the escalating energy requirements. Over the past few years, numerous developed nations have adopted smart grids, leading to the comprehensive transformation of traditional energy grids into adaptable, intelligent, and collaborative systems. This transition facilitates the seamless integration of various distributed energy resources such as solar, wind, and tidal energy along with advanced metering and communication infrastructures. In transmission lines, occurrences of power outages are significantly attributed to unpredictable and irregular faults [1], [2], [3], [4], [5]. Power system faults are unavoidable and cannot be ignored. The fault detection and classification are paramount importance for maintaining the stability of both conventional and smart power grids.

Faults, especially transmission line faults and equipment failures, can cause significant disruptions to the power system, leading to power outages and equipment damage. Therefore, accurate and timely fault detection and

The associate editor coordinating the review of this manuscript and approving it for publication was Arash Asrari.

classification are essential for maintaining stable and safe operation of the smart grid. A smart grids is a complex and dynamic system that requires continuous monitoring and maintenance to ensure reliability and efficiency. Fault detection and classification are crucial tasks in the operation and management of the smart grids. In recent years, machine learning techniques have shown significant promise in fault detection and classification in various domains, including fault detection in smart grids [6], [7]. As stated in [8]and [9], a majority of the faults within the transmission segment of the power system manifest in transmission lines. Short-circuit faults are prevalent and regarded as the most severe type, posing substantial risks to transmission lines. These risks include diminishing the operational lifespan of components, elevating power losses, cable heat, and insulators damage.

B. LITERATURE REVIEW

In this literature review, we discussed the use of CNN and LSTM for fault detection and classification in smart grids (SGs). In recent years, machine learning techniques have shown promising results in fault detection and classification in various domains. In particular, recurrent neural networks (RNN) have been widely used in time series analysis and have shown excellent performance in various applications such as speech recognition, natural language processing, and financial forecasting. Long short-term memory (LSTM) networks, a special version of RNN, are particularly suitable for modeling temporal data with long-term dependencies, making them ideal for detecting and classifying faults in the smart grid. Moreover, a smart grid is a complex and dynamic system that requires continuous monitoring and maintenance to ensure reliability and efficiency.

Fault detection and classification are crucial tasks in the operation and management of the smart grid. Contemporary power systems are transitioning into digital and data-intensive environments. Numerous parameters can now be measured across different terminals using phasor measurement units (PMUs), involving various metrics such as current, voltage, and frequency. These measures serve as valuable input for intelligent detection and classification algorithms. Additionally, the deployment of PMUs has prompted new projects aimed at addressing the issues of fault localization, categorization, and identification in transmission lines. This is achieved through the utilization of diverse measurements collected from various locations within the power system. The rise of using PMUs has led to the availability of extensive datasets, facilitating the broader implementation of data-driven methodologies within power systems, as demonstrated in [10] and [11].

In recent years, machine learning techniques have shown significant promise for fault detection and classification in various domains, including smart grids [12], [13]. Recurrent neural networks, especially long short-term memory neural networks, have demonstrated exceptional performance in time series analysis. LSTM networks are particularly suited for modeling long and short-term dependencies in sequential data, making them ideal for modeling the complex and dynamic behavior of smart grids. Several studies have employed LSTM networks for fault detection and classification in smart grids [14], [15]. In [16], LSTM was demonstrated to be effective for automatically extracting features for accurate fault diagnosis in photovoltaic arrays. Likewise, [17] established the superiority of LSTM in diagnosing faults within wind turbines using multivariate time series data.

In [18], the authors introduced a hybrid architecture that combining CNN and LSTM. This model was trained to estimate the distance to the fault location within a 220 km long 2-bus single-line test system, utilizing voltage and current measurements. The effectiveness of this approach surpassed that of alternative methods for accurately classifying faults within transmission lines. In [19], the authors introduced an amalgamation of LSTM with a calibration training filter to classify faults in transmission lines.This approach was evaluated on a 2-bus single-line test system 300 km in length. The authors in [20] presented a CNN model to classify and locate fault in 2-bus system in different architectures of CNN with the penetration of DGs resources obtaining high performance in terms of computation complexity and testing time. A support vector machine technique accompanied by wavelet transform was presented in [21] for feature extraction to efficiently locate and classify the shunt faults on transmission lines networks. This approach was tested on a single 69 kV line spanning 29.4 km.

A transmission line fault classification technique based on support vector machines (SVM) with different types of training models was proposed by [22]. This method performs effectively in fault classification when examined on a transmission line. Nevertheless, previous methodologies have been implemented within the confines of two-bus, single-line power systems. However, large-scale multi-machine power systems possess a broader scope and increased intricacy, necessitating the formulation of novel models to overcome these challenges. in this study, the proposed models were designed to accomplish the tasks of identifying the faulty section, classifying the fault type, and determining the fault location distance within the affected region.

In recent years, a variety of techniques have been proposed for fault detection and classification, including the most effective and widely used theories such as wavelet transform (WT), artificial neural networks (ANN), and fuzzy logic in addition to the combined of two or more algorithms such as (WT-ANN) and neuron-fuzzy. Thus, wavelet transform (WT) is commonly utilized for feature extraction across different frequency ranges. For example, authors in [23] and [24] offered a WT-based strategy to record high-frequency traveling waves for fault detection, classification and phase selection of faults.

A feed-forward neural network, combined with a back-propagation algorithm, was employed for fault detection and classification. This involves using three-phase voltages and currents as inputs to the neural networks [25]. In [26],

a fuzzy system was introduced to enhance detection and classification performance, and combined techniques have been extensively used. The use of the two algorithms overcomes different drawbacks and improves the detection and classification accuracy in addition to the benefits of using each algorithm and combining them together. Additionally, these algorithms can overcome the system challenges represented by the complexity of the designed power system network.

However, the combination of ANN and wavelet transform was presented in [13] and [27] for transmission-line fault detection and classification. Furthermore, fuzzy logic with WT was developed to identify various faults in an electrical distribution system [28]. Deep learning has been categorized as an end-to-end model, demonstrating superior performance in handling large datasets compared to traditional models, which rely on manually created features. Study in [29] offered different promising new models named as DBN (deep belief network), LSTM, and CNN, for fault classification tasks.

A novel self-attention convolutional neural network (SAT-CNN) model was introduced for the detection and classification of transmission line faults [3], [25]. Their notable advantage lies in their ability to implicitly focus on output information from hidden layers, thereby enhancing the classification accuracy of the proposed system. In addition, authors in [30] utilized a convolutional neural network (CNN) as a classifier for utilizing bus voltages to locate the faulted line. Despite the effectiveness of deep learning-based models in fault detection and classification, practical application can be challenging owing to the limitations inherent in their main structures.

However, LSTM is particularly effective in capturing the temporal correlations present in time-series data. For instance, authors in [32] introduced DWT-LSTM for fault detection in insulated overhead conductors of transmission lines. However, LSTM struggles with the extraction of the spatial features. Therefore, this challenge was addressed using a combination of CNN and LSTM structures which offers a solution for extracting spatial and temporal features simultaneously.

A combination of CNN and LSTM deep networks was introduced for sentiment analysis of Twitter datasets owing to its ability to analyze large datasets [33]. As a result, a model integrating CNN and bidirectional LSTM networks was proposed for network intrusion detection. While the fusion of CNN and LSTM architectures effectively extracts spatial and temporal features, they often underutilized the concept of multi-channel input, where diverse features or regions are simultaneously fed into deep neural networks in a parallel multi-channel manner. Therefore, Incorporating the multi-channel approach facilitates multi-angle and multi-directional feature extraction and expression, thereby enhancing the classification accuracy.

Additionally, it is worth noting that CNN-LSTM serves as an assumption-free algorithm capable of effectively managing complex non-linear dynamics within higher-dimensional noisy spaces. A CNN-LSTM hybrid model was presented by the authors in [34], with the purpose of identifying unusual events in solar power generation. Although other machine learning techniques encountered challenges, the model performed exceptionally well. With the extraction of relevant features from the PV power output-influencing factors, the hybrid model forecasts PV power generation efficiently.

A comparative analysis was conducted against several machine learning techniques to showcase the efficacy of proposed models. In summary, the use of CNN and LSTM neural networks for fault detection and classification in smart grids has shown significant promise. LSTM networks are particularly suited for modeling the complex and dynamic behavior of a smart grid, making them ideal for fault detection and classification. The proposed methods have achieved high accuracy in fault detection and classification, outperforming the traditional methods.

C. RESEARCH GAP

Protection algorithms in power systems can be broadly categorized into two approaches: (1) the utilization of ‘‘conventional’’ protection devices and (2) the application of AI-based algorithms. Traditional methods encounter limitations owing to the fault resistance and signal distortion caused by noise. The motivation for seeking another solution lies in the potential to expedite power supply restoration to customers by accurately pinpointing fault locations. Conventional techniques entail intricate calculations and may result in errors when estimating fault locations. These challenges can be effectively addressed by using AI-algorithms [35].

Recently, a significant amount of research has been directed towards investigating protection challenges within ring grids. A solution involving artificial neural networks (ANN) with communication among protective devices (PDs) was introduced in [36] to address this issue. Addressing the same concern, [37] and [38] focused on the utilization of directional relays. Furthermore, [39] and [40] and explored the using of multi-agent systems (MAS) to identify faulty sections within the grid. It is noteworthy that the aforementioned studies did not emphasize the on penetration of distributed generation (DG)-based inverters.

Recent developments have led to the establishment of diverse methodologies that specifically consider the presence of DG-based inverters. In [41] and [42], a fault detection approach utilizing the discrete wavelet transform (DWT) and artificial neural networks (ANN) was introduced. Despite this, the presence of extensive transient data required for ANN training results in low accuracy, which calls for enhancement. Regarding [43] and [44], a protective system that depends on communication was established following the guidelines of the IEC 61850 standard. Nevertheless, additional information is required to thoroughly assess the dynamic behavior and inherent communication attributes of the system. In [45] and [46], a protection approach for an

isolated medium voltage direct current (MVDC) micro-grid was examined.

A communication-based DC directional overcurrent protection method is implemented to isolate the faulted section of the grid. In the context of [47] and [48], the feasibility of Wireless Fidelity (Wi-Fi) in a communication-oriented protection coordination strategy is explored. Nevertheless, considerations regarding the security and susceptibility to interference of the Wi-Fi protocol must be addressed, followed by experimental verification. Previous research has primarily focused on solutions reliant on communication, often neglecting considerations of potential communication breakdowns.

Within the context of [49], an examination was conducted regarding the protection requisites for substantial photovoltaic (PV) integration. The outlined protection strategy relies on communication between overcurrent protections to detect faults, yet the findings are constrained to situations involving symmetrical faults. The authors in [50] focused on fault location in radial system using KNN and SVM based on the voltage of the faulty section and DG sources. In this study, the proposed approach can detect, classify and identify faulty lines in an entire ring system without requiring communication facilities or traditional protections relays. Subsequently, the impact of distributed generation (DG) penetration on conventional protection schemes is investigated considering the topological change during the occurrence of faults in terms of the presence of renewable energy sources in the various parts of the network. Finally, a comprehensive representation of the proposed protection scheme involving various fault scenarios was provided.

D. CONTRIBUTIONS

This study aims to detect, classify, and localize faults that occur in power transmission lines and distributed system. To address the challenge of handling vast amounts of data encompassing diverse fault scenarios, a different optimized deep neural network approach is proposed. The fault scenarios were simulated within the MATLAB environment, utilizing the two IEEE test networks. This evaluation demonstrates the superior efficacy of the proposed method compared to existing approaches. Prior to this study, earlier research in the realm of ring systems and large-scale multi-machine power system did not encompass the simultaneous incorporation of the fault class (FC), line fault (LF), and fault location (FL) models simultaneously.

This study aims to introduce and evaluate various models, with a particular focus on the precision of responses including line fault identification, classification of fault types and fault location estimation. The proposed models introduces an innovative algorithm that is appropriate for identifying the faulty line, analyzing the fault, and estimating precise location of fault scenarios. These algorithms provide higher accuracy and robustness by monitoring and modeling the entire sequence, which captures the transient phase and reflects a distinct system reaction to a fault. This study made significant advances in the following areas:

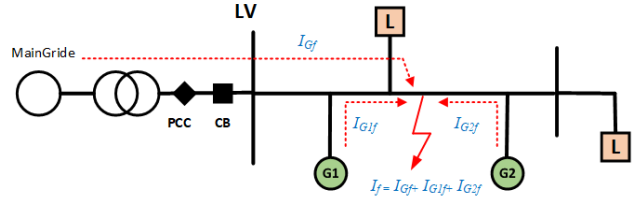
- The proposed algorithm (hybrid CNN-LSTM) model combines the advantages of our customized CNN, which is responsible for extracting relevant fault features from the signal, and then serves as an input for our manipulated LSTM, which results in high accuracy.
- The proposed method has proven its immunity and robustness in terms of network topologies (grid-connected,island-modes) and operation conditions that include fault scenarios and network specifications.
- The proposed methodology overcomes the computational burden by reducing the computational complexity according to mathematical equations to practically reach the optimal hyperparameters of the proposed model.
- To validate and assess the performance of the proposed algorithm, a meticulous evaluation of the model accuracy was conducted and discussed according to metric factors including Precision, Recall, F1-score, Sensitivity and Dependability.

II. THE IMPACT OF DG PENETRATION LEVEL ON THE DS PROTECTION

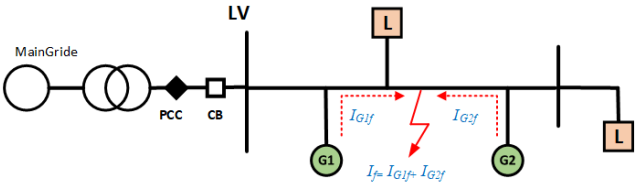
As renewable energy (RE) penetration levels increase over time, the complexity of hybrid power systems also increases. This, in turn, complicates the design, operation, and control of these power systems. Employing suitable methods is crucial to distinguish between normal and abnormal operating conditions. The grid-connected might become disrupted owing to abnormal circumstances, resulting in an island-mode scenario where the distributed generator sources continue to provide power to local loads. Similarly, fault disturbances can arise from various types of defects occurring between the phases and the ground. There are many protections issue in smart grids that should be taken care to prevent fault current to flow from the micro-grid side to the grid sides as follows:

- Dynamics in level of fault current
  Distributed generators have a significant effect on the fault current levels. Furthermore, the magnitude of the fault current fluctuates based on the operating modes of the micro-grid (MG). Figure 1 a illustrates that in the event of a fault, indicated by the red line, the fault current is contributed by both G1 and G2 renewable resources. Consequently, the fault current I_f is the sum of I_{Gf}, I_{G1}, and I_{G2}. This scenario pertains to the grid-connected mode. When the grid disconnects from the rest of the network because of a circuit breaker opening, meaning that the two DG sources disconnect from the main grid, the fault current becomes equal to the sum of only these DG sources, as shown in Figure 1 b. Thus, the fault current changes according to the operation mode. Furthermore, depending on the type of DG integration, inverter-based DG limits the fault current to two per unit of the rated current in the case of a fault inception. Conversely, for asynchronous-based DG, the fault current increases to five times the rated

current, which reduces the sensitivity of the protection devices.



(a) During Grid connected mode



(b) During island mode

FIGURE 1. Magnitude of fault current in micro-grid.

- Bi-directional fault current
  The integration of DGs sources alters direction and magnitude of fault current. In Figure 2, when fault F_1 occurs on the lower-voltage side, as indicated by the red line, the fault current is contributed by G1 in a left-to-right direction. However, if a fault occurs at the LV side or grid-side denoted as F_2, G1 contributes to the fault current in the right-to-left direction. This implies that the direction of the fault current contributed by a particular DG varies based on the fault's location relative to the DG position. Overcurrent relays are commonly used in distributed system. Therefore, traditional unidirectional overcurrent relays cannot adequately protect the safety of MG. For mesh networks, this issue has been addressed by employing directional overcurrent relays, that consider the fault current direction. Currently, with the integration of renewable energy into the distribution network, relying solely on traditional overcurrent relays is insufficient. It is necessary to incorporate directionality aspects into existing relays or utilize AI algorithms to address the complexities of smart grids protection.

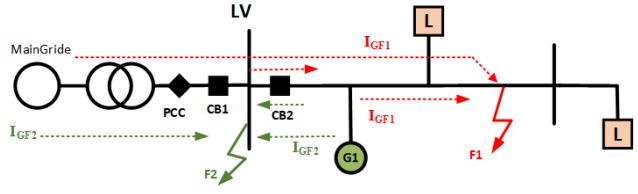


FIGURE 2. Bi-directional fault current in micro-grid.

- False Tripping
  Another issue of concern is false tripping, which refers to the unnecessary interruption of power to loads connected to a healthy feeder.This issue arises when a distributed generator in a healthy feeder causes a malfunction in a nearby feeder. To further illustrate, consider a simple grid with two feeders as shown in Figure 3, each equipped with its own protection relay. For instance, when a fault occurs in feeder2, G1 contributes current to the fault point, and the grid also supplies some current to this fault point ($I_F = I_{grid} + I_{G1}$). When the fault current contribution from G1 exceeds the current setting, Relay 1 (R1) is tripped before the operation of the relay in the faulted feeder, Relay 2 (R2). Consequently, this results in an unnecessary power interruption for loads connected to healthy feeder1.

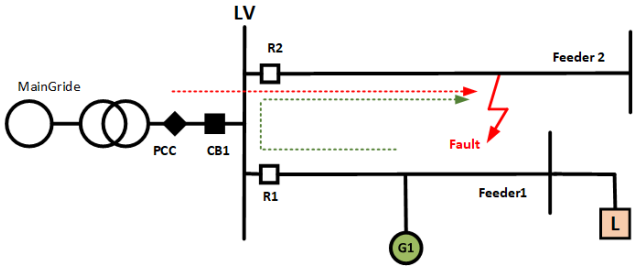


FIGURE 3. False tripping in micro-grid.

- Blinding of Protection
  In a micro-grid system integrated with numerous Distributed Generators along the feeder, there is a notable effect on the fault current level. Consequently, the contribution of utility grid to the fault current decreases when DG sources participate, and the feeder relay may no longer detect the fault state. This type of occurrence is usually referred to as "blinding of protection," and it is represented in Figure 4. I_{GF} represents the fault current contributed by the main grid, and I_{GF1} represents the fault current contributed by DGs sources. Initially, assuming the absence of DGs, I_{GF} is equal to zero, indicating that the fault is entirely supplied by the grid through I_{GF}.
  However, when DGs is present, the fault current is contributed by both the DGs and the grid ($I_F = I_{GF} + I_{GF1}$). The settings of the particular relay are based on I_{GF}, which represents the fault current contributed by the grid system. However, owing to the integration or penetration of DG renewable sources, the fault current remains relatively constant and falls below the set value of the overcurrent relay. This causes the relay do not respond to the fault because it cannot detect it, thus leading to the phenomenon of "Blinding of Protection".

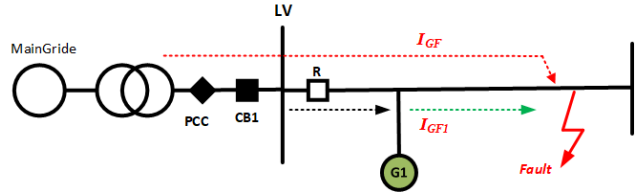


FIGURE 4. Blinding of protection in micro-grid.

III. MATERIAL AND METHODOLOGY

Initially, we present an overview of the created dataset, followed by a description of the data preprocessing

method. Finally, we describe the theoretical foundations and architecture of the proposed hybrid deep learning models.

A. DATA COLLECTION AND TRAINING PROGRESS

In this study the data set will be generated through MATLAB simulation for two test network IEEE 6-bus and IEEE 9- bus test system. Figure 5 and 6 show the single-line diagram of the IEEE 6-bus and IEEE 9-bus test systems to simulate different fault scenario in a transmission line system with DGs penetration in different location. We consider the IEEE 6-bus with four machines, seven transmission lines and three loads [51]. A standard IEEE 6-bus test system equipped with two types of DGs, a PV inverter-based with 20 mw and wind turbine synchronous with 30 mw was selected and located on (Bus5 and Bus6) respectively [52], as depicted in Figure 5.While in the IEEE 9-bus test system equipped with three types of DGs, two PV inverter-based with 30 mw which located on (Bus6 and Bus8) respectively and wind turbine synchronous with 50 mw was selected and located on Bus5 as shown in Figure 6.

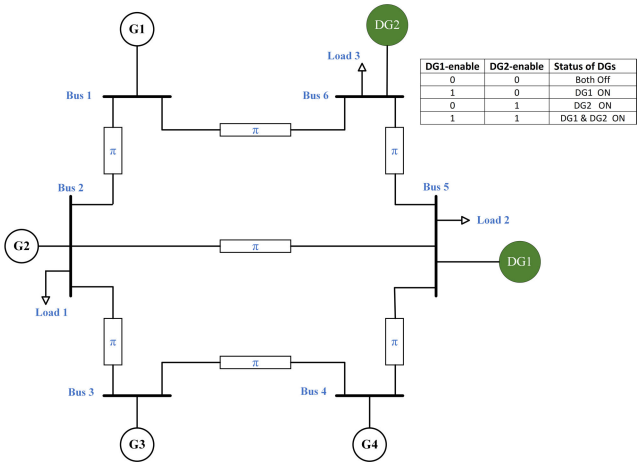


FIGURE 5. IEEE 6-bus System with DG penetration.

Furthermore, the popular western system Coordinating Council (WSCC) or (IEEE 9-bus) 3-machine, 9-bus system. The system consists of three generators, three transformers, three loads, and six transmission lines. The IEEE 9-bus system was employed as an electric distribution system. In this study The IEEE 9-bus transmission setup has been adapted to serve as a distribution system. Data generation during the simulation is as a result of the time sequence of the fault scenarios. These values greatly help us in analyzing the Fault occurred. The timer determines the time span considered for capturing fault data. The feature data (voltage and current signal) and its corresponding classes are exported from the workspace as a CSV file to use them in training progress without any further transform process.

The training data set (MATLAB simulation data) have a total of 115220 × 60 instances for the IEEE 6-bus network and 51840 × 45 instances for the IEEE 9-bus network. The collected features were fed into various types of deep learning algorithm to predict fault class (FC), line faulty

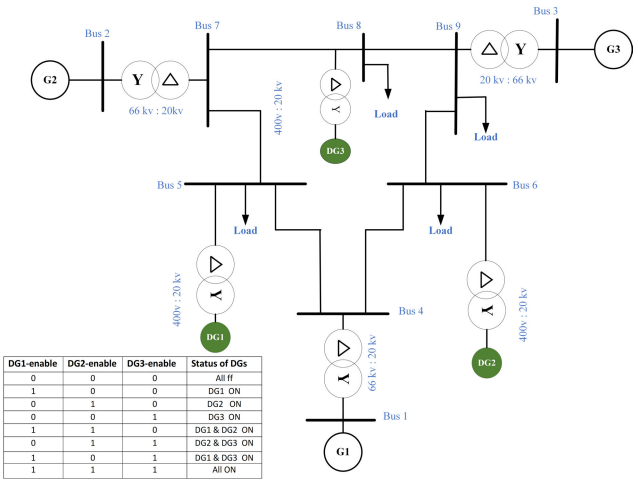


FIGURE 6. IEEE 9-bus System with DG penetration.

(LF) and determination the fault location (FL) as a response to proposed algorithms. During the data collection process, various fault resistance (10 value) was considered as shown in Table 1 and 2 to simulate setup for fault and non-fault scenarios with the total number of cases in each class, this includes both low and high settings.

Fault resistance and the fault inception angle, holds significant importance in fault detection within electrical systems, as they experience alterations during fault incidents. These changes result in voltage and current fluctuations within the affected section. Thus, integrating diverse fault resistance values into the data collection process plays a crucial role in elevating the accuracy of fault identification techniques. Table 3 and 4 illustrate the detail for each class and number of symbols generated for each class.

The input vector of features consists of numerous training and validating samples,each consisting of a long sequence of high-dimensional vectors produced from the real measurements of 3-phase voltages as well as currents signals obtained at different locations as illustrated in Figure 7. This study presents three novel classifiers. The primary objective of the first classifier is to detect the fault types. Subsequently, for each given location, a classifier is built to categorize faulty lines based on fault types, aiding in fault diagnosis within that area. Finally, the third classifier model is used to estimate the location percentage and calculate the distance from which the fault occurred inside the predefined region.

The diagram in Figure 8 illustrates a comprehensive flowchart of the suggested fault identification, classification, and estimated location methodology. This algorithm comprises three key stages: identifying the fault region, classifying its type, and predicting the fault distance.

B. DATA PREPROCESSING

The preprocessing of data involved collecting and storing data in a data frame with specific variable names such as fault type, faulty line, fault location, fault parameters, sample number, and measurements of three-phase voltage

TABLE 1. Simulation setup for fault scenarios with the total number of cases in each class.

Parameter	Possible configuration	No. of cases
Fault type	AG, BG, CG, ABG, ACG, BCG, AB, AC, BC, ABC	10
Line Faulty section	Line1-2, Line2-3, Line3-4, Line5-6, Line6-7, Line7-8, Line8-9 (6-bus)	7
	Line4-5, Line4-6, Line7-6, Line7-8, Line8-9, Line9-6 (9-bus)	6
Fault location	10%,20%,30%40%,50%,60%,70%,80% ,90%	9
Fault resistance in Ω	0.01, 0.05, 0.2, 0.5, 2.5, 10, 20, 50, 75, 100	10
DG1, DG2(6-bus)	00,01,10,11 (0-means Off ,1-means ON)	4
DG1, DG2, DG3 (9-bus)	000,001,010,011,100,101,110,111	8

TABLE 2. Simulation setup for non-fault scenarios with the total number of cases in each class.

Parameter	Possible configuration	No. of cases
Load-1(L1)	± (0, 10, 20, 30, 40, 50) % of change in load	11
Load-2(L2)	± (0, 10, 20, 30, 40, 50) % of change in load	11
Load-3(L3)	± (0, 10, 20, 30, 40, 50) % of change in load	11
DG1, DG2 (6-bus) DG1, DG2, DG3 (9-bus)	With and without DG	2

TABLE 3. Total No. of samples for IEEE 6-bus network.

Fault Class (FC)	No. of distribution	Faulty Line (FL)	No. of distribution	Fault location (FL)	No. of distribution
Normal	15120	Normal	15120	Normal	15120
AG	10080	Line1-2	14400	10% LineLentgh	11200
BG	10080	Line2-3	14400	20% LineLentgh	11200
CG	10080	Line3-4	14400	30% LineLentgh	11200
ABG	10080	Line4-5	14400	40% LineLentgh	11200
ACG	10080	Line5-6	14400	50% LineLentgh	11200
BCG	10080	Line6-1	14400	60% LineLentgh	11200
ABC	10080	Line2-5	14400	70% LineLentgh	11200
AB	10080	-	-	80% LineLentgh	11200
AC	10080	-	-	90% LineLentgh	11200
BC	10080	-	-	-	-

TABLE 4. Total No. of samples for IEEE 9-bus network.

Fault Class (FC)	No. of distribution	Faulty Line (FL)	No. of distribution	Fault location (FL)	No. of distribution
Normal	25920	Normal	25920	Normal	25920
AG	2592	Line4-5	4320	10% LineLentgh	2880
BG	2592	Line4-6	4320	20% LineLentgh	2880
CG	2592	Line7-8	4320	30% LineLentgh	2880
ABG	2592	Line8-9	4320	40% LineLentgh	2880
ACG	2592	Line7-5	4320	50% LineLentgh	2880
BCG	2592	Line9-6	4320	60% LineLentgh	2880
ABC	2592	-	-	70% LineLentgh	2880
AB	2592	-	-	80% LineLentgh	2880
AC	2592	-	-	90% LineLentgh	2880
BC	2592	-	-	-	-

and current amplitudes. The fault types included 10 different fault classes listed as AG, BG, CG, AB, BC, CA, ABG, BCG, CAG, and ABC/ABCG fault. The number of fault varies from zero to ten, where label 0 represents non-fault signals, and fault types one to ten represent different fault types on the transmission line. The fault location included nine different fault locations listed as 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% of the transmission line length. The faulty lines are different for each network type. For the IEEE 9-bus and 6-bus were (6) and (7) respectively. In addition to the aforementioned process,

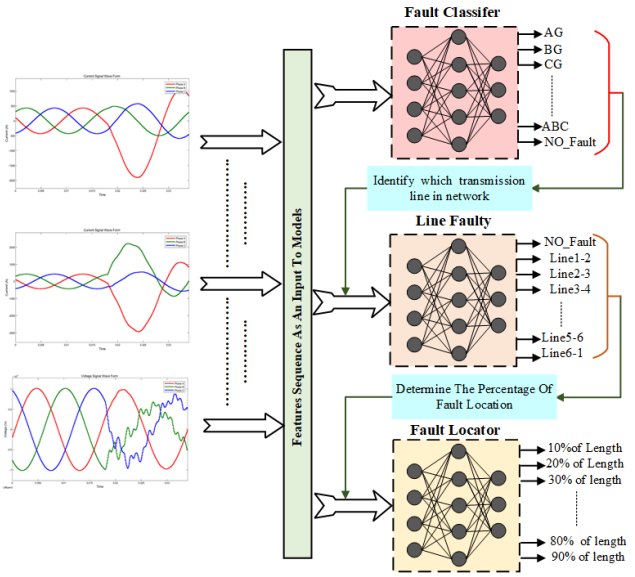


FIGURE 7. Protection scheme diagram for proposed methodology.

three-phase current and voltage amplitudes from all busbars were measured through the simulation. The final dataset, which included training, validation, and testing, was utilized to train the proposed algorithms.

For data preprocessing, we used two fundamental data preprocessing techniques: data conversion and data normalization which have been widely used. Data conversion involves transforming the features of the voltage and current signals from nominal to numeric format, ensuring all data are in a numerical state for compatibility with the deep learning model. In contrast, data normalization is implemented to mitigate the significant variance among features, narrowing down to a specific range of values [53]. During this process, null and overlapping values were eliminated, and we used a minimum-maximum scaling strategy, as shown in equation 1, to address the normalization of larger values and reduce their influence. This technique works with values between zero and

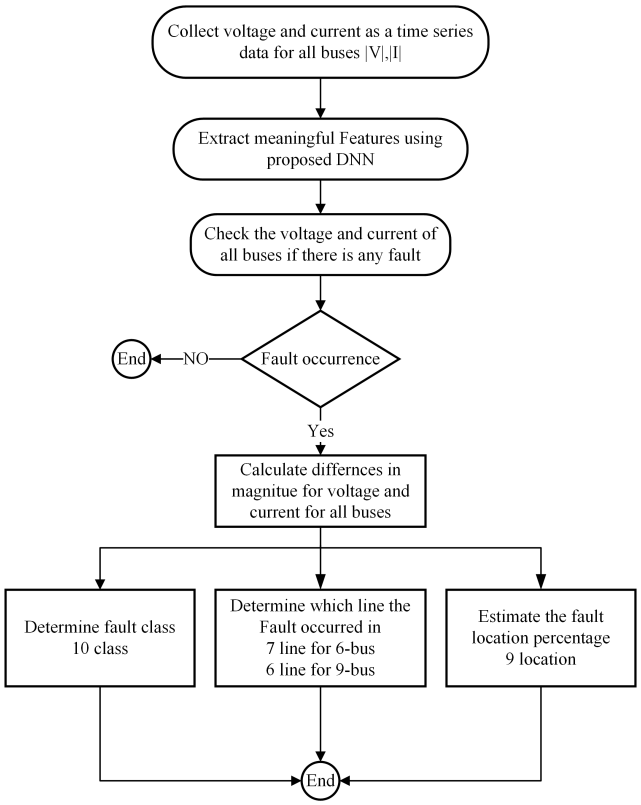


FIGURE 8. Flow chart of proposed protection scheme.

one.

$$x\_normalized = \frac{x - x\_min}{x\_max - x\_min} \tag{1}$$

where x is the original value, $x_{\min}$ is the minimum value, and $x_{\max}$ the maximum value. Sample fault signals for different fault types, including single line to ground fault (SLGF), line to line fault (LLF), and three line fault (LLL), which took place on the transmission line, are depicted in Figure 9.

C. LSTM

Recently, the advancement of deep learning techniques has established the recurrent neural network (RNN) as one of the most powerful models for classification tasks that involve sequential data. Moreover, the RNN is capable of establishing correlations between current and past information within the network. However, despite its ability to learn from sequences of any length, RNN are plagued by issues such as gradient exploding and vanishing. These issues can be addressed by employing a specialized type of RNN known as LSTM.

The fundamental concept that drives the LSTM model is the cell state, which distinguish it distinct from other recursive neural networks. LSTM is designed to learn long-term dependencies within a network. Based on the input X_t, LSTM creates a temporal feature vector h_t at each time-step by updating its cell state (temporal memory). The LSTM model is characterized by three primary gates: the forget gate, input gate, and output gate, as illustrated in Figure 10. Using the sigmoid activation function (σ), the forget gate (f_t) controls the amount of data removed from candidate memory cell state (c_t^*) and decides whether to utilize memory from the previous time-step. Its output ranges between zero and one, indicating portion of the input to be transmitted to the output.

The decision of whether and how much new information should be included in the current time-step is made by the "update gate," which is represented by (i_t). The memory must be updated at each step, which is accomplished using (c_t^*). The "output gate" (o_t) controls the quantity of information from the previous time-step that should be combined with the current time-step information and passed on to the subsequent time-step [34], [35], [36]. equations 2 and 3 below provide a mathematical description of how the LSTM model works.

$$\left\{\begin{array}{l} i_t = \sigma\left(W_i \odot \left[h_{(t-1)}, X_t\right] + b_i\right) \\ f_t = \sigma\left(W_f \odot \left[h_{(t-1)}, X_t\right] + b_f\right) \\ C_t^* = \sigma\left(W_C \odot \left[h_{(t-1)}, X_t\right] + b_C\right) \\ O_t = \sigma\left(W_O \odot \left[h_{(t-1)}, X_t\right] + b_O\right) \end{array}\right\} \tag{2}$$

where W_f, W_i, W_c, and W_o represent the input weight vectors.while b_f, b_i, b_c, and b_o representing the bias vectors. The memory cell value (C_t) and output (h_t) of the network are obtained using the following equations:

$$\left\{\begin{array}{l} C_t = C_{(t-1)} \odot f_t + C_t^* \odot i_t \\ y_t = O_t \odot \tanh\left(C_t\right) \end{array}\right\} \tag{3}$$

LSTM models were trained using the Python programming language with Keras on Google Collaboratory, which has GPU capability. In two case studies on the IEEE 6-bus and IEEE 9-bus system, the performance of the suggested structural technique for identifying, classifiying, and locating faults (FC, LF, and FL) in power systems was evaluated.It is crucial to evaluate performance of the proposed classifiers using testing data once the LSTM models have been trained. In general, the proposed LSTM model illustrated in Figure 11, were examined and meticulously evaluated using the F1-score, Accuracy, Recall, and Precision criteria. Custom splitting was used to train the LSTM models for each of the three scenarios' FC, LF and FL with 70% for the training set, 10% for the validation set and 20% for the testing set, the number of epochs was 100 and the learning rate equal to 0.001 and Adam optimizer was used. The hyper parameter configuration is presented in table 5.

To explore the computational complexity of a given model, we select the optimal input parameters and hyperparameters. Because of computational limitations, exhaustive testing of all parameter combinations is impractical, therefore, we focus on key parameters that are likely to have the greatest impact on the computational complexity of the system. Specifically, the selected hyperparameters for this study focused on the number of (convolution layers and dense layers), as well as the size of (pooling, filter, neuron, and convolution kernel). The most significant layers were convolutional, pooling, and dense layers. Therefore, we can mitigate time complexity by carefully selecting the number of convolutional and fully

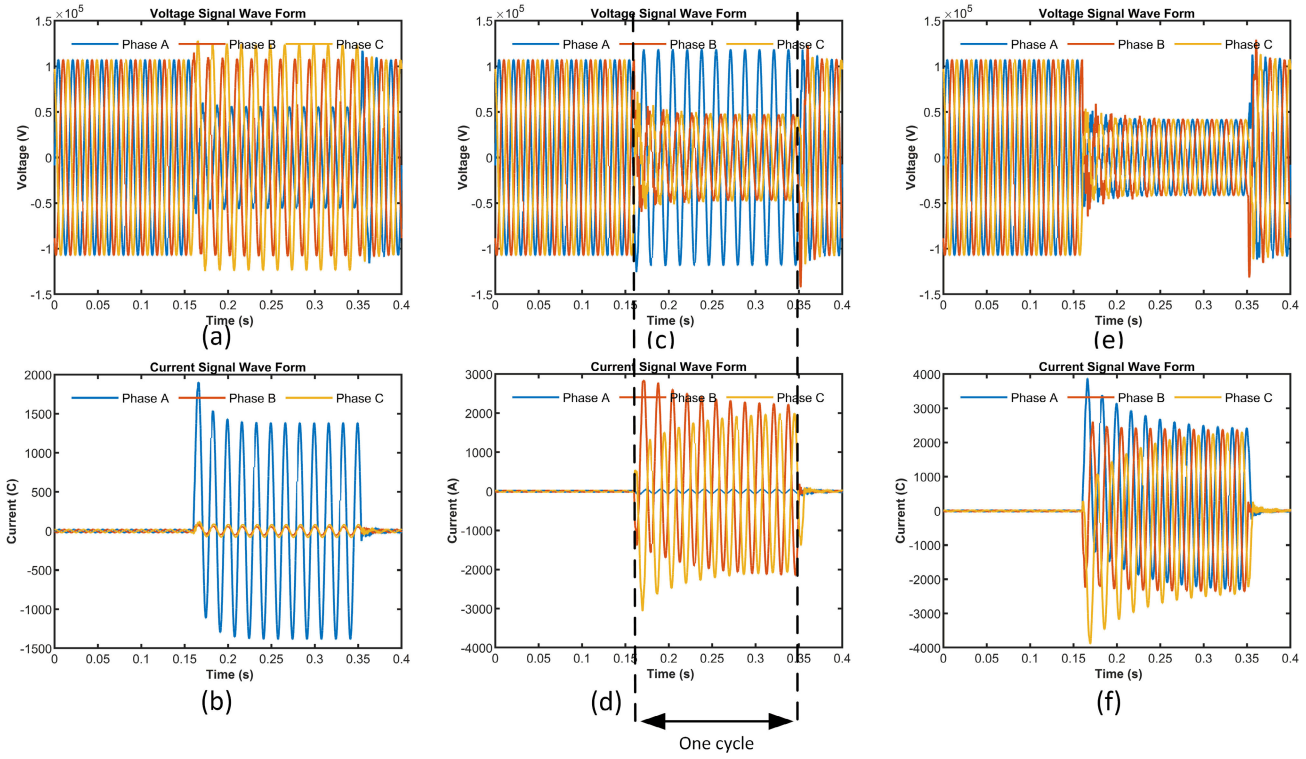


FIGURE 9. Fault Signal waveforms with different fault resistance. (a) Voltage of AG fault, (b) Current of AG fault, (c) Voltage of BCG fault (d) Current of BCG fault, (e) Voltage of ABCG fault, (f) Current of ABCG fault.

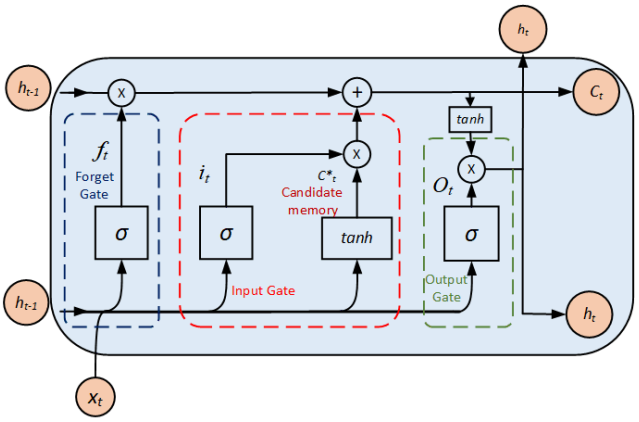


FIGURE 10. Architecture of single LSTM cell.

connected layers. The time complexity of the convolutional layer can be computed using equation 4:

$$\left(\sum_{n=1}^{d} k_{n-1} \cdot s_n^2 \cdot f_n \cdot l_n^2\right) \cdot r_1 \cdot b_1 \tag{4}$$

where, d represents the depth of the convolutional layer, l_n is the length of the output feature map, f_n denotes the number of filters in the n-th layer, S_n represents the length of the filter, k_{n-1} specifies the number of input channels in the $(n-1)$ layer, r_1 indicates the learning rate, and b_1 represents the batch size.

The batch size and learning rate are two critical parameters which significantly affect the computational complexity.

TABLE 5. Hyperparmater selection.

Hyperparameter	1D-CNN, LSTM, CNN-LSTM
Epoch No.	100
Batch size	64
Optimizer	Adam
Initial learning rate	0.001
Learning rate adjustment	Monitor = Validation loss, patience=10,
Drop=0.1, epoch_drop=10	epoch(drop*epochs/epoch_drop)
Dropout	0.2
Activation function	tanh

Therefore, Increasing the batch size led to an increase in the computational burden of the model and vice versa. However, it is crucial to consider the product of these two parameters to measure the complexity of the model accurately. In a fully connected layer, every parameter is interconnected,thereby establishing a link between the higher and output layers. This layer comprises varying numbers of neurons that influence the output size. To assess the total complexity of all fully connected layers within the model, it is essential to multiply the parameters of each layer (including the height and width of the input), number of neurons, and input dimension. Subsequently, summing the complexities of all the layers yields the total complexity of the fully connected layers in the model. The time complexity of the fully connected layer can be determined using equation 5:

$$\left(\sum_{l=1}^{f} D \cdot W \cdot H \cdot N\right) \tag{5}$$

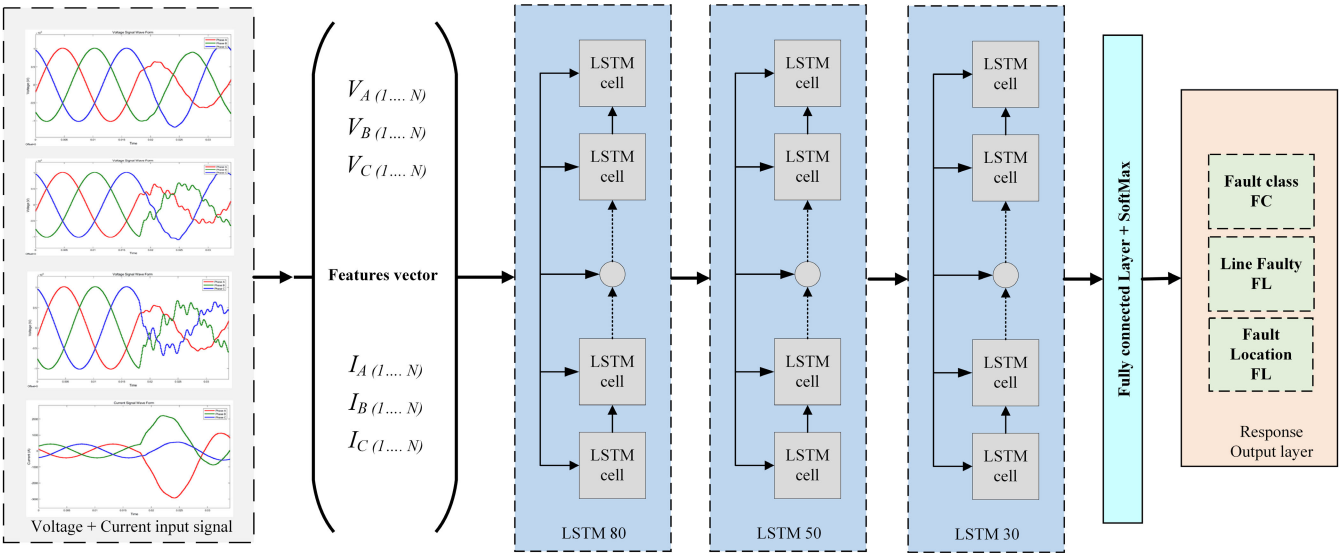


FIGURE 11. Block Diagram of Proposed LSTM.

where l represents the depth of the fully connected layer, and D, W, H, and N denote the dimensions of the input/output channel, width of the input, height of the input, and number of outputs, respectively.

The classification process occurs during the design stage of the network architecture. The input layer was constructed to receive exactly the same input signal as the order input layer at this point. The hidden layer of the network is composed of three LSTM hidden layers with 80, 50 and 30 cells. To prevent overfitting, a dropout layer was included between each LSTM layer, resulting in a total of three dropout layers.

Subsequently, for classification purpose, a fully connected layer with a size of 11 was placed at the final stage of the structure for both faulty and non-faulty cases for the FC classifier and a size of 10 for the FL classifier. While for the LF the fully connected layer has a size of 8 in IEEE 6-bus and a size of 7 in IEEE 9-bus. The fully connected layer is followed by a softmax layer and a classification layer, which predict the probability of each fault class, line faulty and fault location, respectively. Once the network settings are complete, the network training stage and validation process proceed. Finally, the network was tested using a testing dataset to classify the fault type and fault location on the transmission line, the results obtained are listed in Table 6.

D. CNN

Recently, advanced machine learning techniques, in particular, 1D convolution neural networks (CNNs) and recurrent neural networks (RNNs), have demonstrated remarkable performance in tackling demanding activity recognition and classification tasks. It's worth noting that these algorithms achieve such coups with little or no manual feature engineering, instead relying on feature learning from raw data. CNN architectures are versatile in their capability to accommodate various input data formats, including 1D, 2D, and even higher-dimensional representations. These data shapes typically involve a range of 1 to N channels. For the model proposed in this study, we specifically utilized the 1D input data format which is more suitable and effective for time series data.

Each CNN neuron's output is computed based on its inputs as well as the weights and biases of neurons in the layers of the network architecture that come before it. Accordingly, the weights and biases associated with each layer can be adjusted separately by applying the subsequent equation to the preceding layers of the network's architecture. Consequently, the weights and biases pertaining to each layer can be individually updated using the equations 6 and 7:

$$\Delta w_i(t+1) = -\frac{\alpha r}{n} w_i - \frac{\alpha}{n} \times \frac{\partial C}{\partial w_i} + m\Delta w_i(t) \quad (6)$$

$$\Delta b_i(t+1) = -\frac{\alpha}{n} \times \frac{\partial C}{\partial b_i} + m\Delta b_i(t) \quad (7)$$

where w_i and b_i represent the weight and bias of a neuron, respectively. The regularization parameter is indicated by α, and the learning rate is given by r. The numbers n and m represent the total number of samples used for the training and momentum respectively. The cost function is represented by C, whereas the updating step is indicated by t. Throughout the training process, these parameters are iteratively adjusted and fine-tuned to achieve optimal performance. Within the depths of a deep convolutional neural network (CNN), two discrete operations stand out: the convolution and pooling layers. These strata of the network are aptly labeled as convolutional and pooling layers as shown in Figure 12. A number of filters across the input were combined in the convolution layer to create a unique feature vector. In one-dimensional data, convolution has the

TABLE 6. Layer characteristic detail of proposed LSTM.

Layer (type)	Output Shape	Param #IEEE 6-bus	Param #IEEE 9-bus
lstm (LSTM)	(None, 60, 80)	26240	26240
dropout (Dropout)	(None, 60, 80)	0	0
lstm_1 (LSTM)	(None, 60, 50)	26200	26200
dropout_1 (Dropout)	(None, 60, 50)	0	0
lstm_2 (LSTM)	(None, 30)	9720	9720
dropout_2 (Dropout)	(None, 30)	0	0
dense (Dense)	(None, 11) for FC	341	341
	(None, 8) for LF	248	217
	(None, 10) for FL	310	310
Trainable params	FC	62501	62501
	LF	62408	62377
	FL	62470	62470

following mathematical definition in the equation 8:

$$y_i(k) = \sum_{n=0}^{N-1} x_i(n)h(k-n) \tag{8}$$

where n is the number of elements in the output vector y_i at position K, h is the filter, and N is the number of elements in the input vector X_i. Conversely, the pooling process (also known as down-sampling) at the pooling layer minimizes the output dimensions of the convolutional layer. This reduction aids in decreasing the computational complexity and mitigating overfitting concerns. A pivotal hyperparameter within CNN architecture is the number of filter maps. It is possible to explore a spectrum of values, such as (8, 16, 32, 64, 128, 256), based on the task's inherent complexity. Another vital hyperparameter for a 1D CNN is kernel size (3,5,7,..). This kernel size regulates the number of time steps considered during each "reading" of the input sequence, which is subsequently projected onto the feature map via a convolutional process. A larger kernel size implies a more relaxed data analysis, potentially yielding a broader and more generalized representation of the input. One-dimensional convolutional neural networks (1D CNNs) offer distinct advantages and are therefore favored over their two-dimensional counterparts when handling signals in a one-dimensional format. This preference is attributed to the following reasons [54], [55]:

- The computational complexities of the 1D and 2D convolutions exhibited a noteworthy distinction. Specifically, when convolving an image with dimensions NxN using a KxK kernel size, the computational complexity is approximately $O(N^2K^2)$ for 2D convolutions. In contrast, the complexity of the corresponding 1D convolution with the same dimensions (N and K) is nearly $O(NK)$. This indicates that the computational complexity of a 1D CNN is significantly lower than of a 2D CNN with the same configured network, and hyperparameter settings.
- In a broad analysis, particularly based on recent studies, it is notable that the majority of 1D CNN applications tend to adopt compact configurations, typically with fewer than 10,000 parameters and networks with 1-2 hidden CNN layers. On the other hand, almost all 2D CNN applications choose "deep" architectures, which frequently include more than one million parameters (typically more than ten million). Clearly, networks with shallower architectures are considerably more straightforward to train and implement.
- Deep 2D CNNs typically require a certain hardware configuration, such as GPU farms or cloud computing. In contrast, for compact 1D CNNs with a small number of neurons (less than 50) and hidden layers (two or fewer), training on a conventional computer with a CPU is both possible and relatively quick.
- Concise 1D CNNs have low computational burden requirements, which makes them ideal for real-time and cost effective applications.

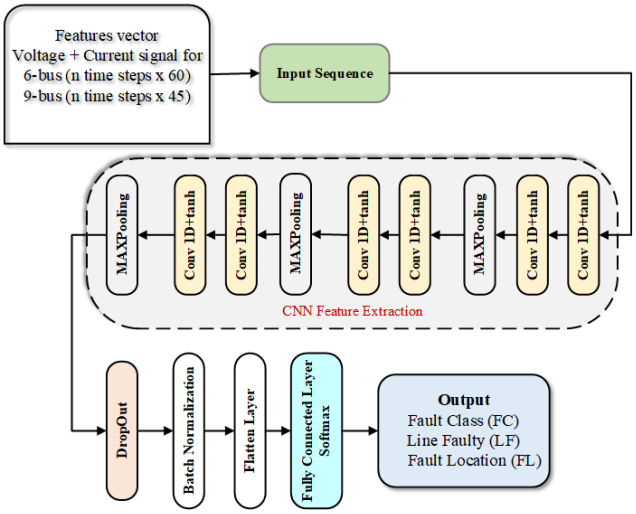


FIGURE 12. Layer details of proposed 1D-CNN.

E. HYBRID CNN-LSTM

In this study, CNN layers were employed to capture localized features from time-series data of voltage and current inputs across all lines within the test network, following a preliminary preprocessing stage. 1D CNNs are useful for time-series applications because they automatically extract hidden properties from data that might not be visible in the temporal dimension by using convolution

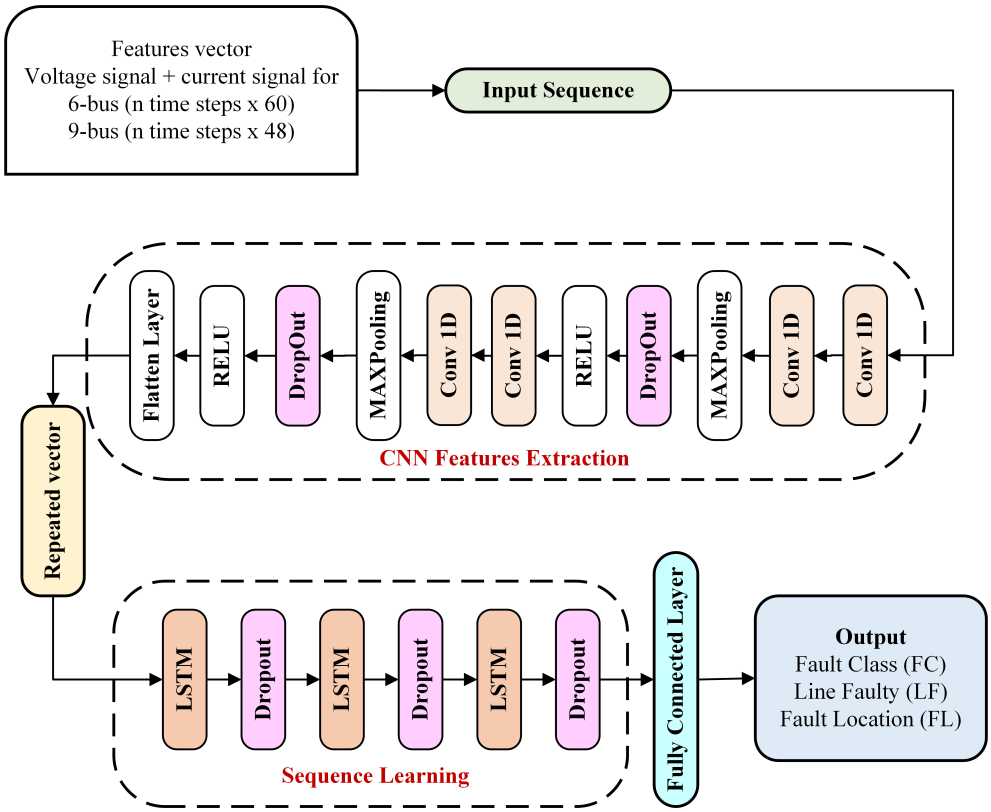


FIGURE 13. The proposed framework of hybrid deep CNN-LSTM.

kernels. Subsequently, the LSTM architecture use the features extracted by the encoder (CNN) as the input.

During the training process, both the training data and the numerous gates within the LSTM network were continuously adjusted, enabling the LSTM model to identify the relationships within the input and output sequences. The architecture of the proposed hybrid deep learning system, which incorporates CNN and LSTM layers, is shown in Figure 13. The CNN feature extraction block consists of two consecutive layers of 1D convolution. We have added a maxpooling layer, drop out layer, and rectified linear unit (ReLU) layer in between, these subsequent layers of convolution neural networks to prevent overfitting issue.

The input feature was passed through the kernel layer, which was is 3×3. Subsequently, two convolutional layers were utilized to extract the feature maps with a convolutional sizes of 32 and 64 accordingly.

Stacking a several convolutional layers inside a deep learning framework enables the initial layers to extract and capture low-level characteristics from the input data, which makes this configuration of convolutional layers very effective. It is worth mentioning that the convolutional layers maintain precise feature location data from the input sequence. Therefore, even with small changes in the placements of the input features,unique feature maps were obtained.

We add a pooling layer after the convolutional layer to overcome this limitation and improve the model's ability to learn complicated structures. A maxpooling layer (downsampling approach) was incorporated into the model to lower the spatial dimensions of the feature maps by a factor of 20%. This decrease in spatial dimensions contributes to a reduction in the overall computational burden. We also used the ReLU activation function, which is well-known for its ability to withstand the vanishing gradient issue. Researchers have used this activation function extensively to increase the trainability and learning capacity of the model by adding non linear aspect to increase the model complexity to solve more complex problems.

A useful method for decreasing the issue of overfitting in any deep learning model is to incorporate a dropout layer. During training, a portion of the neurons in this layer are randomly deactivated. To reduce the overfitting problem, we added a dropout layer after each CNN feature extraction layer as well as in the LSTM layers. The final output is created by connecting the results of the sequential learning segment to a dropout layer, and then to a fully connected layer. Three LSTM layers, 128, 64, and 32, each with a different number of neurons were used in the sequence learning component. For the first two LSTM layers, we set the parameter "return sequence" to the state "true". As a result, the output will be the entire sequence of hidden states without deleting anything. However, in the last LSTM layer, we set the "return sequence" to the state "False" to consider only the hidden state of the last time step.

TABLE 7. Performance metrics of LSTM for each classifier (FC, LF and FL) in IEEE 6-bus.

classifier Name	Class Name	Precision	Recall	F1-Score	Accuracy	n-actual	n-classified (TP)	FP
FC	AG	1.00	1.00	1.00	99.97	2073	2071	2
	BG	1.00	0.99	1.00	99.93	1981	1979	2
	CG	1.00	0.98	0.99	99.84	1961	1958	3
	AB	0.99	1.00	1.00	99.91	2025	2012	13
	AC	1.00	1.00	1.00	99.94	2016	2007	6
	BC	1.00	0.99	0.99	99.95	1956	1955	1
	ABG	1.00	1.00	1.00	99.95	2075	2066	9
	ACG	1.00	0.99	1.00	99.94	2035	2030	5
	BCG	0.98	1.00	0.99	99.96	2040	2030	10
	ABC	1.00	0.98	0.99	99.79	1957	1915	36
	Normal	1.00	1.00	1.00	100	3065	3065	0
Average					**99.93%**			
LF	Line1-2	1.00	0.99	0.99	99.79	3577	3564	13
	Line2-3	1.00	0.99	0.99	99.86	3585	3569	16
	Line3-4	1.00	1.00	1.00	99.94	3596	3591	5
	Line4-5	1.00	0.99	0.99	99.82	3573	3568	5
	Line5-6	0.99	1.00	0.99	99.87	3614	3577	37
	Line6-1	0.99	1.00	1.00	99.93	3605	3586	19
	Line2-5	0.99	0.99	0.99	99.72	3638	3558	80
	Normal	1.00	1.00	1.00	99.79	3791	3781	10
Average					**99.84%**			
FL	10%_Length	0.96	0.97	0.97	99.23	2195	2107	88
	20%_Length	0.95	0.96	0.95	98.63	2269	2103	166
	30%_Length	0.97	0.96	0.96	98.81	2247	2106	141
	40%_Length	0.97	0.97	0.97	98.97	2223	2123	100
	50%_Length	0.94	0.97	0.96	98.95	2265	2123	142
	60%_Length	0.97	0.95	0.96	98.96	2253	2120	133
	70%_Length	0.95	0.96	0.95	99.05	2121	2036	85
	80%_Length	0.97	0.95	0.96	98.59	2340	2147	193
	90%_Length	0.98	0.98	0.98	98.98	2228	2139	89
	Normal	1.00	1.00	1.00	99.97	3043	3036	7
Average					**99.02%**			

IV. SIMULATION RESULT AND EXPERIMENTAL VALIDATION

A. CASE STUDY 1: IEEE 6-BUS TEST SYSTEM

The effectiveness of the proposed hybrid deep learning approach was assessed by conducting a series of simulations aimed at classifying three distinct types of faults: fault Class, faulty Lines, and fault Location. For these simulations, we utilized a standard IEEE 6-bus test system equipped with distributed generators (DGs). This network consist of six bus-bars, seven line segments, three loads, and two DGs, located at bus-bars 5 and 6 near the load section. The DG at bus-5 had a power rating of 20 MW, while the DG at bus-6 had a power rating of 30 MW, both with an X/R ratio of 10. Despite the relatively short distances between nodes in distribution networks We utilized the pi-line model to accommodate the capacitive characteristics of the transmission lines.

This enable us to collocate a more accurate training dataset. Simulink in MATLAB 2022b is used as the simulation environment. The proposed models were developed using Python, utilizing the Keras library with the TensorFlow backend engine. MATLAB Simulation was used to create time-series data for the voltage and current signals. An Intel CORE i7-10510U CPU running at a 2.2 GHz base clock and 2.30 GHz turbo boost, paired with 16 GB of RAM, was used for all experiments. The generated dataset was divided into three subsets: 70% for training, 10% for validation, and 20% for testing. For the testing phase, 20% of the data from each class were used.

The performance evaluation of the proposed classifiers for FC, LF, and LF was demonstrated by using LSTM, CNN, and hybrid CNN-LSTM models, respectively. n-actual refer to the number of actual cases in a class while n-classified mean number of cases belonging to a class. The performance of the proposed model was evaluated using various metrics, such as Accuracy, Precision, Recall and F1-Score. Together, these metrics offer a comprehensive evaluation of the accuracy of classification model, the ability to detect relevant instances, and the balance in minimizing both false positives and false negatives. These metrics prove particularly valuable in scenarios where the costs or impacts of classification errors vary significantly.

- **Precision** evaluates the precision of the positive predictions generated by a model by determining the ratio of true positives to the total number of predicted positives. A high precision value indicates that the model's positive predictions are likely to be accurate. Precision is particularly beneficial in scenarios where the cost associated with false positives is significant because it reduces the probability of erroneous positive predictions.

TABLE 8. Performance metrics of CNN for each classifier (FC, LF and FL) in IEEE 6-bus.

classifier Name	Class Name	Precision	Recall	F1-Score	Accuracy	n-actual	n-classified (TP)	FP
FC	AG	1.00	1.00	1.00	99.97	2059	2053	6
	BG	1.00	1.00	1.00	99.97	1975	1970	5
	CG	0.99	1.00	0.99	99.89	1971	1953	18
	AB	1.00	1.00	1.00	99.98	2063	2059	4
	AC	1.00	1.00	1.00	99.97	2015	2012	3
	BC	1.00	0.99	1.00	99.98	2081	2077	4
	ABG	0.99	1.00	1.00	99.95	2011	1999	12
	ACG	1.00	1.00	1.00	99.94	1998	1986	12
	BCG	1.00	0.99	1.00	99.97	1961	1959	2
	ABC	1.00	1.00	1.00	99.92	1976	1974	2
	Normal	1.00	1.00	1.00	100	3092	3092	0
Average					**99.96%**			
LF	Line1-2	1.00	1.00	1.00	99.79	2880	2880	0
	Line2-3	1.00	1.00	1.00	99.86	2971	2969	2
	Line3-4	1.00	1.00	1.00	99.94	2875	2875	0
	Line4-5	1.00	1.00	1.00	99.82	2876	2865	11
	Line5-6	1.00	0.99	1.00	99.87	2879	2878	1
	Line6-1	1.00	1.00	1.00	99.93	2850	2850	0
	Line2-5	1.00	1.00	1.00	99.72	2858	2851	7
	Normal	1.00	1.00	1.00	99.79	2995	2991	4
Average					**99.97%**			
FL	10%_Length	0.96	0.97	0.97	99.66	2195	2107	88
	20%_Length	0.95	0.96	0.95	99.38	2269	2103	166
	30%_Length	0.97	0.96	0.96	99.22	2247	2106	141
	40%_Length	0.97	0.97	0.97	99.18	2223	2123	100
	50%_Length	0.94	0.97	0.96	99.52	2265	2123	142
	60%_Length	0.97	0.95	0.96	99.49	2253	2120	133
	70%_Length	0.95	0.96	0.95	99.51	2121	2036	85
	80%_Length	0.97	0.95	0.96	99.41	2340	2147	193
	90%_Length	0.98	0.98	0.98	99.65	2228	2139	89
	Normal	1.00	1.00	1.00	100	3043	3036	7
Average					**95.85%**			

- **Recall**, alternatively referred to as the Sensitivity or True Positive Rate, quantifies the capacity of a model to accurately detect all relevant instances within datasets. It computes the ratio of the true positives to the total number of actual positives. A high recall score indicated the effectiveness of the model in capturing the majority of positive instances in the datasets. This is essential in situations where the omission of positive instances carries a significant cost or is intolerable.
- **F1-Score**: The F1-Score is the harmonic mean of precision and recall. It provides a balanced assessment of model performance by considering both false positives and false negatives. The F1-Score is particularly useful when there is an imbalance between classes or when one wants to strike a balance between precision and recall. It is often used as a single metric to compare the overall classification performance of models.

In this context, true positive (TP) refers to the count of correctly identified positive class instances, true negative (TN) represents accurately identified negative class instances, false positive (FP) indicates the count of incorrectly identified positive class instances, and false negative (FN) represents the count of incorrectly identified negative class instances. These values are determined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{10}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{11}$$

$$F1 - \text{score} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{12}$$

Tables 7, 8 and 9 provide a detailed evaluation of different models or configurations for the IEEE 6-bus system using LSTM, CNN and Hybrid CNN-LSTM respectively, demonstrating the performance of each model in term of precision, recall, F1-Score, and accuracy across different classes and conditions.

The performance metrics of the proposed LSTM model are listed in Table 7.The LSTM model is employed as a fault classifier, encompassing FC, LF, and FL. Regarding the FC, the model exhibited superior accuracy for the "BC" class, contrast to the lower accuracy for the "ABC" class under fault conditions. Conversely, the model achieved 100% accuracy in identifying the no-fault conditions. In the context of line faulty classification, the model demonstrated

TABLE 9. Performance metrics of hybrid CNN-LSTM for each classifier (FC, LF and FL) in IEEE 6-bus.

classifier Name	Class Name	Precision	Recall	F1-Score	Accuracy	n-actual	n-classified (TP)	FP
FC	AG	1.00	1.00	1.00	99.97	99.97	2039	2
	BG	1.00	1.00	1.00	100	100	2046	0
	CG	1.00	1.00	1.00	99.93	99.93	1980	8
	AB	1.00	1.00	1.00	99.98	99.98	2058	1
	AC	1.00	1.00	1.00	99.99	99.99	2028	2
	BC	0.99	1.00	0.99	99.9	99.9	2001	22
	ABG	0.99	1.00	1.00	99.94	99.94	1914	15
	ACG	1.00	0.99	1.00	99.93	99.93	1950	2
	BCG	1.00	1.00	1.00	99.98	99.98	1950	0
	ABC	1.00	0.99	0.99	99.88	99.88	2039	6
	Normal	1.00	1.00	1.00	100	100	3120	0
Average					**99.95%**			
LF	Line1-2	1.00	1.00	1.00	99.96	2959	2954	5
	Line2-3	1.00	1.00	1.00	99.96	2825	2822	3
	Line3-4	1.00	1.00	1.00	100	2903	2903	0
	Line4-5	1.00	1.00	1.00	99.99	2828	2825	3
	Line5-6	1.00	1.00	1.00	99.98	2828	2824	4
	Line6-1	1.00	1.00	1.00	99.97	2829	2827	2
	Line2-5	1.00	1.00	1.00	99.97	2184	2183	1
	Normal	1.00	1.00	1.00	100	2160	2160	0
Average					**99.98%**			
FL	10%_Length	0.99	0.98	0.99	99.75	2286	2275	11
	20%_Length	0.97	0.98	0.97	99.44	2331	2256	75
	30%_Length	0.97	0.97	0.97	99.42	2174	2110	64
	40%_Length	0.97	0.96	0.97	99.4	2183	2127	56
	50%_Length	0.95	0.97	0.96	99.25	2333	2222	111
	60%_Length	0.97	0.97	0.97	99.41	2199	2123	76
	70%_Length	0.98	0.97	0.98	99.52	2237	2183	54
	80%_Length	0.97	0.97	0.97	99.41	2228	2160	68
	90%_Length	0.98	0.99	0.98	99.68	2304	2261	43
	Normal	1.00	1.00	1.00	100	2930	2930	0
Average					**99.53%**			

minimal accuracy for ''lines 2-5'' and elevated accuracy for ''lines 2-3'' and ''3-4''. Concerning fault location accuracy presented as a percentage, the model attains a maximum accuracy of 99.23% for 10% of the line length, while registering a minimum accuracy of 99.02% for 50% of the line length.

Table 8 lists the performance metrics of the 1D-CNN model across three response categories: FC, LF, and FL. In the domain of fault class identification, the model attains a commendable accuracy range of (99.89% - 99.97%) for fault scenarios, while achieving perfect accuracy (100%) for no-fault conditions. Regarding line faulty classification, the model demonstrated an average accuracy range of (99.72% - 99.94%). Additionally, in fault location identification, the model demonstrates an average accuracy range of (99.18% - 99.66%), considering the percentage of line length.

Table 9 presents the performance evaluation of the proposed hybrid CNN-LSTM deep learning model, across three response categories: Fault Class, Line Faulty, and Fault Location. The results indicate a noteworthy performance enhancement across all the output responses. This improvement is attributed to a reduction in the number of false positives (FP) within each class, thereby enhancing the overall quality of the proposed model and contributing to an elevated level of model accuracy.

B. CASE STUDY 2: IEEE 9-BUS SYSTEM

To validate the proposed algorithms experimentally, We tested by creating a simulated IEEE 9-bus network on a MATLAB environment. subsequently, we used a typical IEEE 9-bus test system integrated with Distributed Generators (DGs). This network comprised nine buses, six-line segments, four loads, and three DGs. Tables 10, 11 and 12 offer a comprehensive evaluation of various models or configurations for 1DCNN, LSTM, and Hybrid CNN-LSTM in the test network IEEE 9-bus system.

Table 10 outlines the performance evaluation of the LSTM model, showing commendable levels of accuracy across various categories. Specifically, the model achieved an average accuracy range of (98.34% - 99.99%) for fault class identification (FC), (99.74% - 99.93%) for Line Faulty classification (LF), and (98.59% - 99.02%) for Fault Location classification (FL). It is noteworthy that, despite a slightly elevated number of false positives(FP) resulting from the utilization of the percentage of line length, the overall accuracy remains within acceptable limits.

TABLE 10. Performance metrics of LSTM for each classifier (FC, LF and FL) in IEEE 9-bus.

classifier Name	Class Name	Precision	Recall	F1-Score	Accuracy	n-actual	n-classified (TP)	FP
FC	AG	1.00	1.00	1.00	99.98	517	517	0
	BG	0.99	1.00	1.00	99.95	530	525	5
	CG	1.00	1.00	1.00	99.99	565	565	0
	AB	0.99	1.00	1.00	99.97	526	523	3
	AC	1.00	1.00	1.00	99.99	509	509	0
	BC	0.99	1.00	1.00	99.96	508	504	4
	ABG	1.00	1.00	1.00	99.99	523	523	0
	ACG	0.97	0.94	0.96	99.81	518	504	14
	BCG	0.96	0.96	0.96	98.34	534	512	22
	ABC	0.96	0.95	0.96	99.58	518	497	21
	Normal	0.99	1.00	1.00	99.17	5118	5093	25
Average					**99.72%**			
LF	LINE 4-5	0.99	0.99	0.99	99.85	851	846	5
	LINE 4-6	0.99	1.00	0.99	99.85	886	882	4
	LINE 7-8	0.99	1.00	0.99	99.91	811	807	4
	LINE 8-9	0.98	0.99	0.99	99.77	874	869	5
	LINE 7-5	0.99	0.98	0.99	99.74	927	909	18
	LINE 9-6	1.00	1.00	1.00	99.93	835	831	4
	Normal	1.00	0.99	1.00	99.6	5184	5154	30
Average					**99.81%**			
FL	10%_Length	0.88	0.9	0.89	98.74	593	522	71
	20%_Length	0.89	0.86	0.87	98.59	573	510	63
	30%_Length	0.88	0.9	0.89	98.81	568	498	70
	40%_Length	0.91	0.89	0.9	98.97	551	505	49
	50%_Length	0.93	0.88	0.9	98.95	544	505	39
	60%_Length	0.86	0.9	0.88	98.7	560	479	81
	70%_Length	0.90	0.88	0.89	98.75	578	518	60
	80%_Length	0.90	0.86	0.88	98.7	554	500	54
	90%_Length	0.90	0.92	0.91	99.02	576	516	60
	Normal	0.99	1.00	0.99	99.36	5269	5221	48
Average					**98.86%**			

Table 11 presents the performance metrics of three classifiers: FC, LF, and FL. The classifiers are evaluated based on accuracy, sensitivity,precision, F1-score,, true positives (TP), false positives (FP). The classifier demonstrated exceptional performance across all fault classes, consistently achieving high Accuracy, Recall, F1-score, and Precision. The distribution of instances in the classes was fairly balanced, with slightly higher misclassifications in the ''BCG'' and ''ABC'' classes. The LF classifier performed remarkably well for all line faulty scenarios, achieving perfect Accuracy in each case. Furthermore, the FL classifier exhibits commendable metrics performance across different percentages of line length. The slight increase in number of false positives in the FL classifier as a result of using percentage of line length rather than km-based, which indicates a trade-off between precision and sensitivity, but overall accuracy remains adequate.

Table 12 presents the performance of the hybrid CNN-LSTM model for the IEEE 9-bus system evaluated across three classifiers: FC, LF, FL. The FC classifier demonstrated excellent precision, recall, F1-score, and accuracy for most fault classes. The average accuracy of 99.92% suggests highly accurate identification of the fault classes. The LF classifier performs exceptionally well, achieving high precision, recall, F1-score, and accuracy for various line faulty scenarios. An average accuracy of 99.96% indicates a robust performance in identifying line faulty conditions. The FL classifier demonstrates high performance across different percentages of line length. The average accuracy of 99.4% indicates accurate fault location identification. The hybrid CNN-LSTM model demonstrates exceptional performance in fault classification, line faulty identification, and fault location for the 9-bus system. The high accuracy and low misclassification rates suggest its efficacy in real-world applications of power system fault analysis.

A confusion matrix is a structured table used to evaluate the performance of supervised learning techniques, particularly machine learning and statistical classification. This matrix offers a comprehensive breakdown of the classifier's performance. A confusion matrix is a vital tool in assessing the performance of a classifier. It displays the true classes along the vertical axis and the predicted classes along the horizontal axis. This arrangement facilitated the identification of correct and incorrect predictions. The diagonal of the matrix represents the number or percentage of true classes correctly identified by the classifier, offering insights into the accuracy of class detection across different categories in the training data. Figure 14 illustrates the classification confusion matrices

TABLE 11. Performance metrics of CNN for each classifier (FC, LF and FL) in IEEE 9-bus.

classifier Name	Class Name	Precision	Recall	F1-Score	Accuracy	n-actual	n-classified (TP)	FP
FC	AG	1.00	1.00	1.00	99.99	507	506	1
	BG	1.00	1.00	1.00	99.98	518	518	0
	CG	1.00	1.00	1.00	99.99	511	511	0
	AB	1.00	1.00	1.00	99.99	518	517	1
	AC	1.00	1.00	1.00	100	489	489	0
	BC	1.00	1.00	1.00	100	530	530	0
	ABG	1.00	1.00	1.00	99.99	507	507	0
	ACG	0.99	0.98	0.98	99.96	505	501	4
	BCG	0.96	0.97	0.96	99.63	510	488	22
	ABC	0.97	0.96	0.96	99.63	492	476	16
	Normal	1.00	1.00	1.00	99.87	5282	5271	11
Average					**99.92%**			
LF	LINE 4-5	1.00	1.00	1.00	99.97	873	870	3
	LINE 4-6	1.00	1.00	1.00	99.92	853	847	6
	LINE 7-8	1.00	0.99	0.99	99.98	883	883	0
	LINE 8-9	1.00	0.99	1.00	99.96	853	851	2
	LINE 7-5	1.00	1.00	1.00	99.98	801	799	2
	LINE 9-6	1.00	1.00	1.00	100	880	880	0
	Normal	1.00	1.00	1.00	99.93	5209	5209	0
Average					**99.96%**			
FL	10%_Length	0.92	0.89	0.91	99	527	483	44
	20%_Length	0.86	0.88	0.87	98.5	622	536	86
	30%_Length	0.88	0.91	0.89	98.8	593	522	71
	40%_Length	0.92	0.88	0.9	98.85	576	530	46
	50%_Length	0.85	0.92	0.89	98.76	592	506	86
	60%_Length	0.91	0.89	0.9	98.92	546	496	50
	70%_Length	0.94	0.89	0.91	99	570	538	32
	80%_Length	0.83	0.92	0.87	98.52	625	518	107
	90%_Length	0.96	0.85	0.91	98.99	512	492	20
	Normal	1.00	1.00	1.00	99.66	5204	5179	25
Average					**98.90%**			

generated during a single testing phase for the proposed models. Subfigures 14 a and 14 b provide the confusion matrices for the Fault class, employing hybrid LSTM, in the grid-connected scenario for both the IEEE 6-bus and 9-bus systems, respectively. Meanwhile, subfigures 14 c and 14 d depict the confusion matrices for Line faulty, utilizing hybrid CNN-LSTM, in the IEEE 6-bus and 9-bus systems, respectively. Finally, subfigures 14 e and 14 f show the prediction performance of Fault location using the hybrid CNN-LSTM.

V. RESULT AND PERFORMANCE ANALYSIS

Table 13 compares the features of the various methods with those of the proposed approach, facilitating a clearer comprehension of these studies. Three common deep learning (DL) models for fault location prediction (FLP), fault type classification (FTC), and fault region identification (FRI) in large-scale multi-machine power systems were introduced by the authors in [19]. These models are based on deep recurrent neural networks (DRNN). To produce reliable categorization and prediction findings, the models use fully transient data from pre and post-faults recorded using phasor measurement units (PMUs). the proposed algorithms performed better when evaluated in a two-area with four-Machine power system. the study focuses on testing the algorithms in a two-area four-machine power system, which may limit the generalizability of the results to other power system with configurations, Further research and testing on different power system setups and fault scenarios would be necessary to fully assess the performance and limitations of the proposed models.

The authors in [20] concentrate on fault detection in distribution networks incorporating distributed generators (DGs). With conventional relaying methods becoming inadequate due to fluctuating fault current levels, the authors proposed a CNN model for fault classification without pre-processing or feature engineering. The performance evaluation is conducted using 10-fold cross-validation, resulting in an accuracy of 99.92%. The proposed model surpasses the conventional approaches in terms of accuracy and computational burden.

The authors in [35] and [36] developed classification schemes utilizing ANN models for the IEEE 9-bus system. However, these studies have a notable limitation as they do not offer a comprehensive fault detection and location scheme, which is a critical element in smart grid networks. In another study [56], a deep learning approach based on a benchmark dataset was introduced to classify non-fault and faulty scenarios. Unfortunately, this study lacks a fault section identifier and fault location. Additionally, it does not investigate different modes of operation such as grid-connected and island-mode with high level of DG penetration.

TABLE 12. Performance metrics of hybrid CNN-LSTM for each classifier (FC, LF and FL) in IEEE 9-bus.

classifier Name	Class Name	Precision	Recall	F1-Score	Accuracy	n-actual	n-classified (TP)	FP
FC	AG	1.00	1.00	1.00	100	524	524	0
	BG	1.00	1.00	1.00	99.99	528	528	0
	CG	1.00	1.00	1.00	100	510	510	0
	AB	1.00	1.00	1.00	100	532	532	0
	AC	1.00	1.00	1.00	100	516	516	0
	BC	1.00	1.00	1.00	100	526	526	0
	ABG	1.00	1.00	1.00	100	543	543	0
	ACG	0.98	0.96	0.97	99.71	506	494	12
	BCG	0.98	0.92	0.95	99.52	475	465	10
	ABC	0.93	0.98	0.95	99.52	546	506	40
	Normal	1.00	1.00	1.00	99.72	5162	5144	18
Average					**99.86%**			
LF	LINE 4-5	1.00	1.00	1.00	99.97	885	885	0
	LINE 4-6	1.00	0.99	0.99	99.91	836	834	2
	LINE 7-8	0.99	0.99	0.99	99.87	882	874	8
	LINE 8-9	1.00	0.99	1.00	99.97	849	849	0
	LINE 7-5	0.99	0.99	0.99	99.82	859	859	0
	LINE 9-6	1.00	1.00	1.00	100	839	839	0
	Normal	0.99	1.00	1.00	99.59	5208	5181	27
Average					**99.88%**			
FL	10%_Length	0.97	0.94	0.96	99.5	571	554	17
	20%_Length	0.92	0.94	0.93	99.24	568	523	45
	30%_Length	0.95	0.94	0.94	99.34	615	584	31
	40%_Length	0.94	0.96	0.95	99.52	573	536	37
	50%_Length	0.96	0.96	0.96	99.63	593	577	16
	60%_Length	0.96	0.95	0.95	99.48	543	519	24
	70%_Length	0.93	0.93	0.93	99.23	566	526	40
	80%_Length	0.91	0.93	0.92	99.1	569	515	54
	90%_Length	0.95	0.95	0.95	99.44	580	553	27
	Normal	1.00	0.99	1.00	99.59	5180	5164	16
Average					**99.40%**			

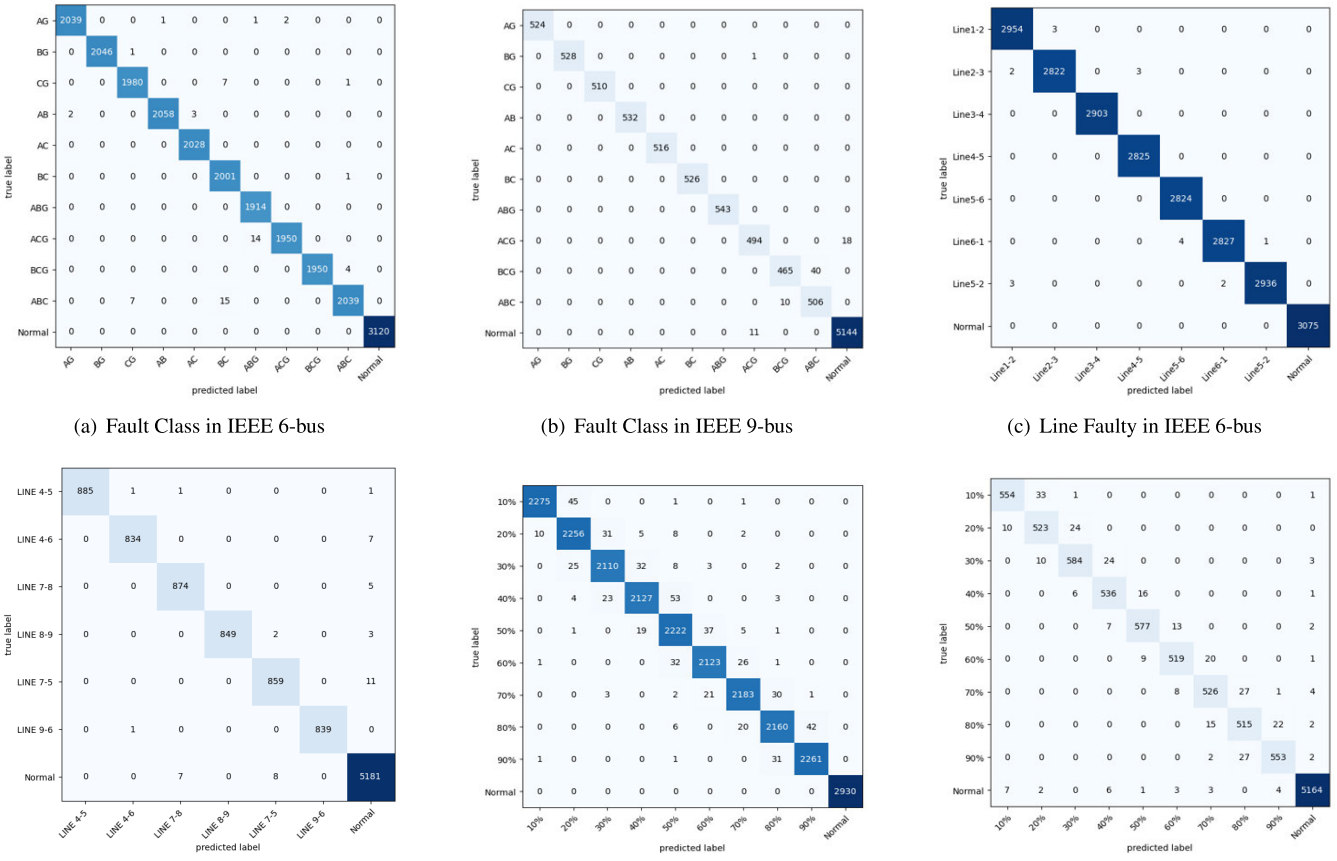
In [57], the authors proposed a shuffle attention using DNN. However, the primary limitation of this study is its failure to provide information on the faulty section, which could help reduce the isolated regions in the radial IEEE 6-bus system, especially compared to the mesh or ring system. Reference [58] evaluated fault classification in a radial two-machine system but did not consider the impact of DG penetration, utilizing only voltage signals. Similarly, in [59], a current signal-based with probabilistic neural network (PNN) approach was used in a radial system with grid-connected mode. However, the fault location information was not provided in either study. Finally, in [60], an RNN model was introduced to classify fault cases in single-machine radial and ring systems. Unfortunately, the study did not investigate the proposed model in large system networks and did not provide a fault location scheme. In this study, we present three deep learning models (1D-CNN, LSTM, Hybrid CNN-LSTM) for three classifiers (FC, LF, FL) evaluated in ring systems (IEEE 6-bus and 9-bus) with different modes of operation (grid-connected and Island-mode). Moreover, we considered the effect of DG penetration in both networks, to achieve optimal accuracy for each classifier.

The statistical metrics provided were utilized to assess the efficacy of the proposed methods for fault detection, classification, and localization, as outlined below.

- **Dependability**: This was defined as the ratio of the total number of predicted fault cases to the total number of actual fault cases.
- **Security**: The ratio of the total number of predicted no-fault cases to the total number of actual no-fault cases.
- **Accuracy**: This represents the ratio of the total number of correctly predicted cases (both fault and no-fault) to the total number of actual cases (both fault and no-fault).

In the context of fault detection and classification, evaluating dependability is crucial, as it directly assesses the scheme's ability to predict fault cases in comparison to the actual number of fault cases, which indicates the rate of misclassification. This metric serves as a measure of the reliability of a scheme for identifying faults. However, security quantifies false alarms, representing instances where non-fault events are incorrectly predicted as faults. As a result, the primary goal is to reduce misclassification, which is more important than lowering false alarms. Finally, accuracy measures the model's ability to predict all cases, considering both fault and no-fault cases.

A comparative analysis of the performances of the 1D-CNN, LSTM, and Hybrid CNN-LSTM models is illustrated in Figures 15, 16, 17, and 18. The utilized dataset encompasses mixed scenarios, incorporating faults occurring in both grid-connected and island-modes with



(a) Fault Class in IEEE 6-bus (b) Fault Class in IEEE 9-bus (c) Line Faulty in IEEE 6-bus

(d) Line Faulty in IEEE 9-bus (e) Fault Location in IEEE 6-bus (f) Fault Location in IEEE 9-bus

FIGURE 14. Confusion matrix for Hybrid CNN-LSTM in IEEE 6-bus & IEEE 9-bus.

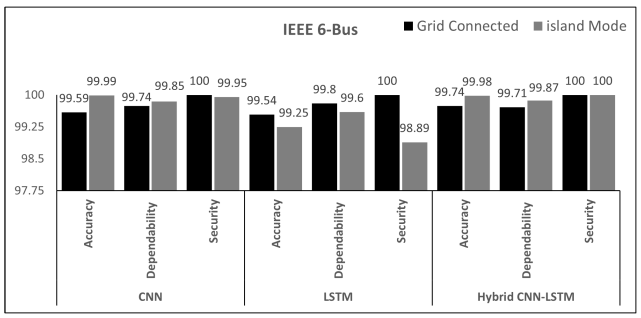


FIGURE 15. Performance of proposed methodology for fault classification, in grid and island-connected modes (Ring topology IEEE 6-bus).

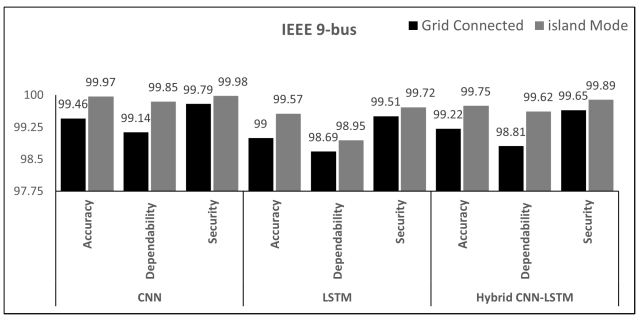


FIGURE 16. Performance of proposed methodology for fault classification, in grid and island-connected modes (Ring topology IEEE 9-bus).

ring network topologies. The development of deep learning models, namely 1D-CNN, LSTM, and Hybrid CNN-LSTM, were conducted using the open-source software Collab. Notably, the performance metrics exhibit proximity among the models, with dependability being a pivotal criterion, and all the proposed models showed significantly enhanced fault classification performance.

The performance comparisons, specifically for fault classification under different operating modes (grid-connected or island-mode), are presented in Figures 15 and 16. In Figures 17 and 18, the evaluations for fault location are similarly presented. Additionally, in the grid-connected mode, 1D-CNN achieves 99.59% accuracy (with 99.74% dependability and 100% security) for fault classification, whereas in the island-mode for the IEEE 6-bus network, it achieves 99.99% accuracy (with 99.85% dependability and 99.95% security). Similarly, the 1D-CNN shows 99.46% accuracy for the IEEE 9-bus network (with 99.14% dependability and 99.79% security). On the other hand, for the IEEE 6-bus network, LSTM demonstrated an accuracy of 99.25% in the island-mode (with 99.6% dependability and 98.89% security) and 99.54% accuracy (with 98.8% dependability and 100% security) for fault classification in the grid-connected.

With respect to the IEEE 9-bus network, LSTM attains 99% accuracy in the grid-connected with (98.69%

TABLE 13. Comprehensive comparison with previous studies.

Methods	Grid type	Test network	Mode of operation	Data	DG Equipped	Classifier Type FC	FL	LF
LSTM [19]	Radial	Four-Machine	Grid connected	V and I	Yes	97.73	99.36	87.50
CNN [20]	Radial	Two-Machine	Grid connected	V and I	Yes	99.92	99.00	Nil
ANN [35]	Ring/Radial	9-bus	Grid + Island mode	V and I	Yes	100.00	Nil	Nil
ANN [36]	Ring	9-bus	Grid connected	V and I	Yes	99.85	Nil	Nil
MCCNN-LSTM [56]	-	-	-	V	No	92.80	Nil	Nil
SAE–DNN [57]	Radial	6-bus	Grid + Island mode	V and I	Yes	99.52	Nil	99.56
FCLSTM [58]	Radial	Two-Machine	-	V	No	97.47	Nil	Nil
PNN [59]	Radial	Single machine	Gride	I	Yes	99.30	Nil	Nil
LSTM [59]	Ring	14-bus	Gide	V	Yes	98.40	90.52	Nil
RNN [60]	Radial	39-bus				98.00	92.24	Nil
Proposed CNN						99.59	95.41	99.89
Proposed LSTM	Ring	6-bus	Gide + Island mode	V and I	Yes	99.54	96.68	99.45
Proposed CNN-LSTM						99.74	97.60	99.93
Proposed CNN						99.46	94.52	99.72
Proposed LSTM	Ring	9-bus	Gide + Island mode	V and I	Yes	99.00	94.27	99.44
Proposed CNN-LSTM						99.22	96.94	99.54

*V refer voltage , I refer to current.

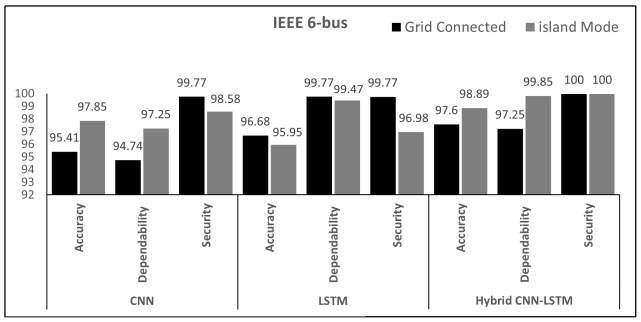


FIGURE 17. Performance of proposed methodology for Fault location, in grid and island-connected modes (Ring topology IEEE 6-bus).

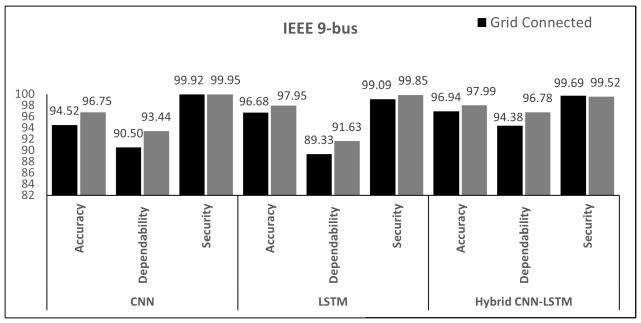


FIGURE 18. Performance of proposed methodology for fault location in grid and island-connected modes (Ring topology IEEE 9-bus).

dependability and 99.51% security) and 99.57% accuracy in the island-mode with (98.95% dependability and 99.72% security). Because of this, accuracy and security are only slightly impacted in the island-mode, whereas reliability is almost unchanged in both operating scenarios. As a result, both data-mining models exhibit comparable performance in grid-connected and island-mode operation.

Comparing the accuracy of the hybrid CNN-LSTM for fault classification in the grid-connected (with 99.71% dependability and 100% security) to its accuracy in island-mode (with 99.87% dependability and 100% security) for IEEE 6-bus, it is observed that the hybrid CNN-LSTM performs better in the earlier scenario as in Figure 15. In Figure 16, the hybrid model for IEEE 9-bus indicates an

TABLE 14. Analyzing the reliability of fault detection and classification methods in both operational modes for IEEE 6-bus.

IEEE 6-Bus topology					
		Mode of operation			
		Grid connected		Island mode	
Classifier	proposed model	Dependability %	Security	Dependability %	Security
FC	CNN	99.74	100	99.4	99.95
	LSTM	99.8	100	99.6	98.89
	CNN-LSTM	99.71	100	99.87	100
LF	CNN	99.9	99.87	98.99	98.65
	LSTM	99.31	99.74	99.45	97.84
	CNN-LSTM	99.97	100	99.78	99.95
FL	CNN	94.74	99.77	97.25	98.58
	LSTM	99.77	99.77	99.47	96.98
	CNN-LSTM	97.25	100	99.85	100

TABLE 15. Analyzing the reliability of fault detection and classification methods in both operational modes for IEEE 9-bus.

IEEE 9-Bus topology					
		Mode of operation			
		Grid connected		Island mode	
Classifier	proposed model	Dependability %	Security	Dependability %	Security
FC	CNN	99.14	99.79	99.85	99.98
	LSTM	98.69	99.51	98.95	99.72
	CNN-LSTM	98.81	99.65	99.62	99.89
LF	CNN	9.75	99.87	99.25	99.56
	LSTM	99.23	99.42	99.76	99.36
	CNN-LSTM	99.81	99.48	99.97	100
FL	CNN	90.5	99.92	93.44	99.95
	LSTM	89.33	99.09	91.63	99.85
	CNN-LSTM	94.38	99.69	96.78	99.52

accuracy of 99.22% with (98.81% dependability and 99.65% security) in the grid-connected mode, while in island-mode it has an accuracy of 99.75% with (99.62% dependability and 99.89% security), as depicted in the same scenario.

The Hybrid CNN-LSTM model is a promising choice, offering a well-balanced performance in terms of accuracy, dependability, and security across different operation modes. The selection of an optimal model may depend on the specific application requirements and priorities. In Figures 17 and 18, a comprehensive comparative analysis is presented, highlighting the performance metrics of accuracy,

dependability, and security across three distinct deep learning models: CNN, LSTM, and Hybrid CNN-LSTM. This evaluation is conducted under varying operational scenarios of grid-connected and island-mode specifically tailored for fault location.

Remarkably, the Hybrid CNN-LSTM model consistently displayed elevated performance levels across all metrics, exhibiting competitive or superior outcomes when juxtaposed with CNN and LSTM in both operational conditions. Notably, in the grid-connected, CNN has the lowest accuracy, whereas Hybrid CNN-LSTM excels in terms of accuracy, dependability, and security. Conversely, in the island mode, both LSTM and Hybrid CNN-LSTM demonstrate noteworthy performances, with the latter showing superior accuracy and dependability. The findings underscore the robust and versatile capabilities of the Hybrid CNN-LSTM model in fault location tasks across diverse operational conditions.

Table 14 presents an exhaustive analysis of the reliability metrics, specifically dependability and security percentages, pertaining to fault classification, line fault detection, and fault location across various proposed models. The evaluation encompasses both operational modes, namely ''grid-connected'' and ''island-mode,'' within the IEEE 6-bus system. This table compares the performance of all models in these distinct operational contexts. Notably, higher percentages indicate superior performance in terms of both dependability and security.

In addition, Table 15 extends the evaluation of the metric performance to the IEEE 9-bus system, encompassing both operational modes. These tables are used to make informed decisions regarding the most reliable fault detection and classification models based on their performance in different operational modes.

VI. CONCLUSION

Fault occurrences is inevitable in electrical networks. The precise and real-time detection of faulty sections in smart distribution networks, integrated with distributed generator (DGs), is imperative for fulfilling customer demand, preventing disruptions, and mitigating financial losses over the long term. In this study, we addressed the challenge of fault detection, diagnosis, and fault location in power systems using a data-driven approach within the context of ring grids or mesh topology. This approach was applied to both transmission and distributed systems, and comprehensive data validation was conducted across various scenarios using a two-IEEE system (comprising 6 and 9-bus configurations). The challenge revolves around a data-driven approach for identifying faulted zones, accurately classifying fault types, and precisely predicting fault locations, with a particular focus on distributed systems. This emphasis arises because of the absence of protection relays within distributed systems that are capable of determining fault locations.

We developed three innovative deep learning models for intelligent fault identification, classification, and location in the transmission lines of the two IEEE test systems. These models are based on a 1D-CNN, LSTM, and a Hybrid CNN-LSTM architecture. The new models take advantage of comprehensive transient data that includes measurements of both voltage and current signal obtained during pre- and post-fault cycles. These signals are acquired either directly or through phase measurement units (PMUs), which provide high-frequency sampling over very short time intervals. These data are collected from different bus-bares and are used as features for training deep neural network models. These models depend on the automated extraction of features directly from the input voltage and current patterns over a specific time interval, obviating the requirement for additional techniques such as transforming them into a different domain or utilizing image-based features for feature extraction.

The sequential learning algorithms presented in this context were designed to extract the utmost spatiotemporal information from these sequential features to effectively model the behavior of the system. Both fault class and line faulty classification models exhibit remarkable accuracy in fault detection and classification, with the highest precision achieved in identifying the faulty region. However, fault location accuracy is comparatively lower, primarily because of the utilization of a percentage-based approach instead of a kilometer-based method. Nevertheless, the accuracy achieved in fault location, even with the percentage method, is acceptable, and this approach offers a more versatile solution applicable to any network, irrespective of the length of the transmission lines.

The findings indicate that CNN-LSTM models are highly accurate, dependable, and exceptionally efficient in identifying, classifying, and locating faults in power system transmission lines. This contribution, involving the introduction of three novel deep learning models that work collaboratively, significantly enhances the effectiveness of maintenance strategies for ring power systems. Furthermore, the suggested models rely heavily on voltage and current data from all bus-bars, and these variables play a crucial role in enhancing the performance of the proposed classifiers. Moreover, this study considered the uncertainty associated with changes in topology, including variations in the penetration level of different DGs sources.In future study, we will evaluate the proposed approach on larger ring power networks.

ABBREVIATIONS

SGs	Smart Grids
DG	Distributed Generator
RE	Renewable Energy
MG	Microgrid
PV	Photovoltaic
DNN	Deep Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
FNNs	Feed forward Neural Network
LSTM	Long Short Term Memory

DS	Distributed System
KNN	K-Nearest Neighbors Network
SVM	Support Vector Machine
PNN	Probabilistic Neural Network
ANFIS	Adaptive Neuro-Fuzzy Inference System
FC	Fault Class
LF	Line Faulty
FL	Fault Location
WT	Wavelet Transform
PMU	Phasor Measurement Unit

ACKNOWLEDGMENT

Ahmed Sami Alhanaf, have received a Ph.D. scholarship from the Iraqi Ministry of Electricity. This research was carried out as part of the graduation thesis named "Artificial intelligent application in smart grid." It is being carried out under the framework of Yildiz Technical University.

REFERENCES

[1] H. Liang, Y. Liu, G. Sheng, and X. Jiang, "Fault-cause identification method based on adaptive deep belief network and time–frequency characteristics of travelling wave," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 5, pp. 724–732, Mar. 2019.

[2] W. Wang, H. Yin, C. Chen, A. Till, W. Yao, X. Deng, and Y. Liu, "Frequency disturbance event detection based on synchrophasors and deep learning," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3593–3605, Jul. 2020.

[3] Y. Xi, W. Zhang, F. Zhou, X. Tang, Z. Li, X. Zeng, and P. Zhang, "Transmission line fault detection and classification based on SA-MobileNetV3," *Energy Rep.*, vol. 9, pp. 955–968, Dec. 2023.

[4] L. Xiang, P. Wang, X. Yang, A. Hu, and H. Su, "Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism," *Measurement*, vol. 175, Apr. 2021, Art. no. 109094.

[5] J. Xie, A. P. S. Meliopoulos, and B. Xie, "Transmission line fault classification based on dynamic state estimation and support vector machine," in *Proc. North Amer. Power Symp. (NAPS)*, Sep. 2018, pp. 1–5.

[6] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mech. Syst. Signal Process.*, vol. 100, pp. 439–453, Feb. 2018.

[7] Q.-L. Zhang and Y.-B. Yang, "SA-Net: Shuffle attention for deep convolutional neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 2235–2239.

[8] T. A. Kawady, N. I. Elkalashy, A. E. Ibrahim, and A.-M.-I. Taalab, "Arcing fault identification using combined Gabor transform-neural network for transmission lines," *Int. J. Electr. Power Energy Syst.*, vol. 61, pp. 248–258, Oct. 2014.

[9] H. Fathabadi, "Novel filter based ANN approach for short-circuit faults detection, classification and location in power transmission lines," *Int. J. Electr. Power Energy Syst.*, vol. 74, pp. 374–383, Jan. 2016.

[10] M. R. Aghamohammadi and M. Abedi, "DT based intelligent predictor for out of step condition of generator by using PMU data," *Int. J. Electr. Power Energy Syst.*, vol. 99, pp. 95–106, Jul. 2018.

[11] Y. Zhou, Q. Guo, H. Sun, Z. Yu, J. Wu, and L. Hao, "A novel data-driven approach for transient stability prediction of power systems considering the operational variability," *Int. J. Electr. Power Energy Syst.*, vol. 107, pp. 379–394, May 2019.

[12] J. Hu and A. V. Vasilakos, "Energy big data analytics and security: Challenges and opportunities," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2423–2436, Sep. 2016.

[13] K. M. Silva, B. A. Souza, and N. S. D. Brito, "Fault detection and classification in transmission lines based on wavelet transform and ANN," *IEEE Trans. Power Del.*, vol. 21, no. 4, pp. 2058–2063, Oct. 2006.

[14] A. M. Saidina Omar, M. Khusairi Osman, M. Nizam Ibrahim, Z. Hussain, and A. Farid Abidin, "Fault classification on transmission line using LSTM network," *Indonesian J. Electr. Eng. Comput. Sci.*, vol. 20, no. 1, pp. 231–238, Oct. 2020.

[15] D. Wei, B. Wang, G. Lin, D. Liu, Z. Dong, H. Liu, and Y. Liu, "Research on unstructured text data mining and fault classification based on RNN-LSTM with malfunction inspection report," *Energies*, vol. 10, no. 3, p. 406, Mar. 2017.

[16] A. Y. Appiah, X. Zhang, B. B. K. Ayawli, and F. Kyeremeh, "Long short-term memory networks based automatic feature extraction for photovoltaic array fault diagnosis," *IEEE Access*, vol. 7, pp. 30089–30101, 2019.

[17] J. Lei, C. Liu, and D. Jiang, "Fault diagnosis of wind turbine based on long short-term memory networks," *Renew. Energy*, vol. 133, pp. 422–432, Apr. 2019.

[18] R. Fan, T. Yin, R. Huang, J. Lian, and S. Wang, "Transmission line fault location using deep learning techniques," in *Proc. North Amer. Power Symp. (NAPS)*, Oct. 2019, pp. 1–5.

[19] S. Belagoune, N. Bali, A. Bakdi, B. Baadji, and K. Atif, "Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems," *Measurement*, vol. 177, Jun. 2021, Art. no. 109330.

[20] P. Rai, N. D. Londhe, and R. Raj, "Fault classification in power system distribution network integrated with distributed generators using CNN," *Electr. Power Syst. Res.*, vol. 192, Mar. 2021, Art. no. 106914.

[21] S. Ekici, "Support vector machines for classification and locating faults on transmission lines," *Appl. Soft Comput.*, vol. 12, no. 6, pp. 1650–1658, Jun. 2012.

[22] S. Patil, A. K. Jalan, and A. M. Marathe, "Support vector machine for misalignment fault classification under different loading conditions using vibro-acoustic sensor data fusion," *Exp. Techn.*, vol. 46, no. 6, pp. 957–971, Dec. 2022.

[23] E. Aker, M. L. Othman, V. Veerasamy, I. B. Aris, N. I. A. Wahab, and H. Hizam, "Fault detection and classification of shunt compensated transmission line using discrete wavelet transform and naive Bayes classifier," *Energies*, vol. 13, no. 1, p. 243, Jan. 2020.

[24] M. Jalayer, C. Orsenigo, and C. Vercellis, "Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, fast Fourier and continuous wavelet transforms," *Comput. Ind.*, vol. 125, Feb. 2021, Art. no. 103378.

[25] S. R. Fahim, Y. Sarker, S. K. Sarker, M. R. I. Sheikh, and S. K. Das, "Self attention convolutional neural network with time series imaging based feature extraction for transmission line fault detection and classification," *Electr. Power Syst. Res.*, vol. 187, Oct. 2020, Art. no. 106437.

[26] S. Samonto, S. Kar, S. Pal, A. A. Sekh, O. Castillo, and G. Park, "Best fit membership function for designing fuzzy logic controller aided intelligent overcurrent fault protection scheme," *Int. Trans. Electr. Energy Syst.*, vol. 31, no. 5, May 2021, Art. no. e12875.

[27] N. Zhang and M. Kezunovic, "Transmission line boundary protection using wavelet transform and neural network," *IEEE Trans. Power Del.*, vol. 22, no. 2, pp. 859–869, Apr. 2007.

[28] M. Jamil, R. Singh, and S. K. Sharma, "Fault identification in electrical power distribution system using combined discrete wavelet transform and fuzzy logic," *J. Elect. Syst. Inf. Technol.*, vol. 2, no. 2, pp. 257–267, 2015.

[29] K. Chen, C. Huang, and J. He, "Fault detection, classification and location for transmission lines and distribution systems: A review on the methods," *High Voltage*, vol. 1, no. 1, pp. 25–33, 2016.

[30] W. Li, D. Deka, M. Chertkov, and M. Wang, "Real-time faulted line localization and PMU placement in power systems through convolutional neural networks," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4640–4651, Nov. 2019.

[31] N. Qu, Z. Li, J. Zuo, and J. Chen, "Fault detection on insulated overhead conductors based on DWT-LSTM and partial discharge," *IEEE Access*, vol. 8, pp. 87060–87070, 2020.

[32] M. Umer, I. Ashraf, A. Mehmood, S. Kumari, S. Ullah, and G. S. Choi, "Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model," *Comput. Intell.*, vol. 37, no. 1, pp. 409–434, 2021.

[33] M. M. Hassan, A. Gumaei, A. Alsanad, M. Alrubaian, and G. Fortino, "A hybrid deep learning model for efficient intrusion detection in big data environment," *Inf. Sci.*, vol. 513, pp. 386–396, Mar. 2020.

[34] A. Agga, A. Abbou, M. Labbadi, Y. E. Houm, and I. H. Ou Ali, "CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production," *Electr. Power Syst. Res.*, vol. 208, Jul. 2022, Art. no. 107908.

[35] M. Bakkar, S. Bogarra, F. Córcoles, and J. Iglesias, "Overcurrent protection based on ANNs for smart distribution networks with grid-connected VSIs," *IET Gener., Transmiss. Distrib.*, vol. 15, no. 7, pp. 1159–1174, Apr. 2021.
[36] R. F. Pujiantara, M. Pujiantara, A. Priyadi, and D. A. Asfani, "Protection coordination using zone selective interlocking method and neural network in IEEE 9 bus plan," in *Proc. Int. Conf. Inf. Commun. Technol. (ICOIACT)*, Mar. 2018, pp. 196–201.
[37] W. Al Hanaineh, J. Matas, M. Bakkar, J. El Mariachet, and J. M. Guerrero, "A robust THD based communication-less protection method for electrical grids with DGs," *Int. J. Electr. Power Energy Syst.*, vol. 155, Jan. 2024, Art. no. 109546.
[38] I. Xyngi and M. Popov, "An intelligent algorithm for the protection of smart power systems," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1541–1548, Sep. 2013.
[39] Q. Yang, J. A. Barria, and T. C. Green, "Communication infrastructures for distributed control of power distribution networks," *IEEE Trans. Ind. Informat.*, vol. 7, no. 2, pp. 316–327, May 2011.
[40] G. B. Costa, J. S. Damiani, G. Marchesan, A. P. Morais, A. S. Bretas, and G. Cardoso Jr., "A multi-agent approach to distribution system fault section estimation in smart grid environment," *Electr. Power Syst. Res.*, vol. 204, Mar. 2022, Art. no. 107658.
[41] Q. Chen, G. Nicholson, J. Ye, and C. Roberts, "Fault diagnosis using discrete wavelet transform (DWT) and artificial neural network (ANN) for a railway switch," in *Proc. Prognostics Health Manage. Conf. (PHM-Besançon)*, May 2020, pp. 67–71.
[42] A. S. Altaie, A. A. Majeed, M. Abderrahim, and A. Alkhazraji, "Fault detection on power transmission line based on wavelet transform and scalogram image analysis," *Energies*, vol. 16, no. 23, p. 7914, Dec. 2023.
[43] T. S. Ustun, C. Ozansoy, and A. Zayegh, "Modeling of a centralized microgrid protection system and distributed energy resources according to IEC 61850-7-420," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1560–1567, Aug. 2012.
[44] A. H. El-Hamrawy, A. A. M. Ebrahiem, and A. I. Megahed, "Improved adaptive protection scheme based combined centralized/decentralized communications for power systems equipped with distributed generation," *IEEE Access*, vol. 10, pp. 97061–97074, 2022.
[45] C. Yuan, M. A. Haj-ahmed, and M. S. Illindala, "Protection strategies for medium-voltage direct-current microgrid at a remote area mine site," *IEEE Trans. Ind. Appl.*, vol. 51, no. 4, pp. 2846–2853, Jul. 2015.
[46] L. C. Hunter, C. D. Booth, A. Egea-Alvarez, A. Dysko, S. J. Finney, and A. Junyent-Ferré, "A new fast-acting backup protection strategy for embedded MVDC links in future distribution networks," *IEEE Trans. Power Del.*, vol. 36, no. 2, pp. 861–869, Apr. 2021.
[47] R. B. Kannaian, B. B. Joseph, and R. P. Ramachandran, "An adaptive centralized protection and relay coordination algorithm for microgrid," *Energies*, vol. 16, no. 12, p. 4820, Jun. 2023.
[48] M. P. Nthontho, S. P. Chowdhury, S. Winberg, and S. Chowdhury, "Communication networks for domestic photovoltaic based microgrid protection," in *Proc. 11th IET Int. Conf. Develop. Power Syst. Protection (DPSP)*, Apr. 2012, pp. 1–6.
[49] N. D. Tuyen, N. S. Quan, V. B. Linh, V. Van Tuyen, and G. Fujita, "A comprehensive review of cybersecurity in inverter-based smart power system amid the boom of renewable energy," *IEEE Access*, vol. 10, pp. 35846–35875, 2022.
[50] H. Mirshekali, R. Dashti, A. Keshavarz, and H. R. Shaker, "Machine learning-based fault location for smart distribution networks equipped with micro-PMU," *Sensors*, vol. 22, no. 3, p. 945, Jan. 2022.
[51] M. Ubbarayudu and S. Abassum, "Implementation of active power filter in IEEE 6-bus system will non-linear load applications," *Int. J. Sci. Technol. Res.*, vol. 4, pp. 8722–8726, Oct. 2015.
[52] S. Kar, S. R. Samantaray, and M. D. Zadeh, "Data-mining model based intelligent differential microgrid protection scheme," *IEEE Syst. J.*, vol. 11, no. 2, pp. 1161–1169, Jun. 2017.
[53] K. Sakthidasan Sankaran and B.-H. Kim, "Deep learning based energy efficient optimal RMC-CNN model for secured data transmission and anomaly detection in industrial IoT," *Sustain. Energy Technol. Assessments*, vol. 56, Mar. 2023, Art. no. 102983.
[54] A. S. Alhanaf, H. H. Balik, and M. Farsadi, "Intelligent fault detection and classification schemes for smart grids based on deep neural networks," *Energies*, vol. 16, no. 22, p. 7680, Nov. 2023.
[55] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," *Mech. Syst. Signal Process.*, vol. 151, Apr. 2021, Art. no. 107398.
[56] Y. Xi, X. Tang, Z. Li, Y. Shen, and X. Zeng, "Fault detection and classification on insulated overhead conductors based on MCNN-LSTM," *IET Renew. Power Gener.*, vol. 16, no. 7, pp. 1425–1433, May 2022.
[57] M. Manohar, E. Koley, and S. Ghosh, "Enhancing the reliability of protection scheme for PV integrated microgrid by discriminating between array faults and symmetrical line faults using sparse auto encoder," *IET Renew. Power Gener.*, vol. 13, no. 2, pp. 308–317, Feb. 2019.
[58] M. Li, Y. Yu, T. Ji, and Q. Wu, "On-line transmission line fault classification using long short-term memory," in *Proc. IEEE 12th Int. Symp. Diag. Electr. Mach., Power Electron. Drives (SDEMPED)*, Aug. 2019, pp. 513–518.
[59] A. Mukherjee, K. Chatterjee, P. K. Kundu, and A. Das, "Probabilistic neural network-aided fast classification of transmission line faults using differencing of current signal," *J. Inst. Eng. (India): B*, vol. 102, pp. 1019–1032, Mar. 2021.
[60] M. R. Shadi, M.-T. Ameli, and S. Azad, "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.*, vol. 134, Jan. 2022, Art. no. 107399.

AHMED SAMI ALHANAF received the B.Sc. degree in computer engineering from Basra University, Basra, Iraq, in 2004, and the M.Sc. degree in network services and management engineering from the Polytechnical University of Bucharest, Bucharest, Romania, in 2017. He is currently pursuing the Ph.D. degree with the Department of computer Engineering, Yildiz Technical University, İstanbul, Turkey. He has been an Assistance Chief Engineer with General Company for Electric Power Transmission, Iraq, since 2006. His main working areas are power system protection (design, engineering, testing, and commissioning for protection and control), power system analysis, and transmission systems.

MURTAZA FARSADI was born in Khoy, Iran, in September 1957. He received the B.Sc. degree in electrical engineering, the M.Sc. degree in electrical and electronics engineering, and the Ph.D. degree in electrical engineering (high voltage) from Middle East Technical University, Ankara, Turkey, in 1982, 1984, and 1989, respectively. He is currently a Professor with the Department of Electrical Engineering, Istanbul Aydin University İstanbul, Turkey. His main research interests include high-voltage engineering, industrial power system studies and FACTS, high-voltage direct current transmission systems, DC/AC active power filters, renewable energy, hybrid and electrical vehicles, new control methods, microgrids, and DG.

HASAN HÜSEYIN BALIK is currently with the Faculty of Engineering, Department of Computer Engineering, Istanbul Aydın University, Florya Campus, Küçükçekmece, İstanbul, Turkey, in the fields of computer networks, computer organization and architecture, operating systems, sensors and sensor networks, biomedical networks, bioinformatics, distance learning, computer system design and management, forensics, computer and network security, electromagnetic fields and waves, and enterprise resource planning implementation and management.