RESEARCH ARTICLE ISSN 2738-9898 (Print) ISSN 2738-9901 (Online)

A Deep Neural Network Based Approach for Fault Detection and Localization in Power System Network

Krishna Shrestha* and Ananta Adhikari

School of Engineering, Faculty of Science and Technology, Pokhara University, Nepal

* Author to whom correspondence should be addressed; E-Mail: *sthakrishna101@gmail.com*

Received : 15 November,2024; Received in revised form : 05 December 2024; Accepted : 10 December, 2024; Published: 31 January, 2025

Abstract

Ensuring reliable operation in electricity transmission networks requires efficient fault detection, classification, and localization. However, the integration of distributed generators and the dynamic nature of these systems pose challenges to traditional relaying devices in managing fault currents. This study investigates the application of deep learning techniques to address these challenges by autonomously extracting fault characteristics from three-phase voltage and current signals. Using Artificial Neural Networks (ANNs) and one-dimensional Convolutional Neural Networks (1D-CNNs), fault detection, classification, and localization are performed on the IEEE 9-bus system. Simulated fault data is generated in MATLAB/Simulink, and deep learning models are trained using Python libraries such as scikit-learn and TensorFlow. Results indicate high accuracy, with 1D-CNN achieving 99.87% for faulty line identification, 92.42% for fault classification, and 96.95% for fault location. Similarly, the ANN model attained 99.54%, 92.35%, and 96.24%, respectively. To optimize the cost and complexity of phasor measurement unit (PMU) deployment, a selective feature reduction strategy was implemented, focusing on critical buses (5, 6, and 8), demonstrated that fault analysis can be effectively performed with reduced data inputs, while minimizing the required PMUs. Additionally, transfer learning for N-1 contingency scenarios allowed the pre-trained models to efficiently adapt to new cases, enhancing fault diagnosis performance. These findings highlight the potential of deep learning to improve the accuracy and reliability of fault diagnosis in power transmission systems, supporting future real-time implementation.

Keywords

Fault detection, fault classification, fault localization, deep learning, ANN, CNN, transfer learning, N-1 contingency

1. Introduction

The reliable operation of power systems depends on the rapid detection, classification, and localization of faults. Faults, whether caused by environmental factors, equipment failures, or unexpected operational conditions, pose a significant threat to grid stability and require swift action to prevent cascading failures and minimize service interruptions [1]. In power transmission systems, faults occur when system parameters, such as voltage and current, exceed their threshold values due to abnormal conditions. This is particularly common in overhead transmission lines, which are more exposed to environmental factors compared to underground cables, making them more susceptible to faults. These faults can be broadly classified into series (open conductor) and shunt (short circuit) types, with shunt faults further divided into asymmetrical (e.g., line-to-ground, line-to-line) and symmetrical (e.g., three-phase) categories [2]. Figure 1 shows classification of fault.

There have been several studies to determine the best methods for fault detection and classifications. The available methods can be divided into three main types; prominent technique, hybrid technique, and modern technique. The prominent approaches include wavelet which involves the wavelet transformations, Artificial Neural Network (ANN), and fuzzy logic. The hybrid approaches apply a combination of more than one approach to detect and classify faults, and includes hybrid methods of neuro and fuzzy techniques, wavelet and ANN, Wavelet and fuzzy logic, and wavelet and neuro-fuzzy technique. The third type is modern techniques including the recently used approaches such as Support Vector Machine (SVM), genetic algorithms, decision tree technique, deep learning technique, pattern recognition technic to name a few [3].

Over the years, researchers have explored various methodologies to enhance fault detection and classification techniques, aiming to mitigate downtime and prevent cascading failures in power networks. This report concentrates on the literature related to Fault detection, classification and localization in power distribution systems and power transmission systems using traditional method, machine learning as well as deep learning algorithms [4]. Fault classification and localization techniques have traditionally included Support Vector Machines (SVM) and Fuzzy Logic Systems (FLS), offering significant improvements over rule-based methods by enhancing fault identification accuracy and reliability [5],[6],[7]. While machine learning-based approaches such as SVM, K-Nearest Neighbor (KNN), and fuzzy inference systems provide benefits for specific power system configurations, these models often require extensive feature engineering and manual adjustments, limiting their adaptability to the increasingly dynamic nature of modern power networks [8],[9] and also KNN response slow in high-dimension problems [10].

A method for identifying and categorizing power system failures on transmission lines was presented in [11]. To detect and categorize defects, it employed a rule-based decision tree and a Stockwell transform-based multi-resolution analysis of current signals. In [12] proposed a protection fault scheme that uses a discrete wavelet transform (DWT), genetic algorithm (GA) and neural network (NN). To improve the performance of the classifier, the GA technique is used to determine the optimal parameters of the NN scheme. The scheme is test on a practical network and a high accuracy is obtained for fault determination. Addressing transmission line faults exposed to environmental conditions, [13] proposed an artificial neural network (ANN) approach utilizing a feed-forward ANN with backpropagation for fault detection and classification. This model significantly improves fault detection accuracy, contributing to power system quality and stability. Furthermore, in [15], a customized convolutional neural network (CNN) was presented for fault classification in distributed networks with distributed generation (DG) systems. This deep learning model requires no pre-processing, demonstrating effective fault classification with high cross-validation accuracy, though fault localization was designated for future work. Expanding on deep learning applications, [16] introduced a method using both ANN and 1D-CNN models for fault detection, classification, and location in smart grids, though its applicability may be constrained due to potential noise sensitivity and limited validation on the IEEE 6-bus system. In [17] a paper presented an advanced real-time short-term voltage stability (STVS) assessment using a hybrid Temporal Convolutional Neural Network (TempCNN) and Long Short-Term Memory (LSTM) model. A key highlight of this study incorporated the N-1 contingency scenario, with transfer learning enabling the model to adapt effectively to unseen conditions with minimal labeled data, highlighting its robustness in real-time applications.

Given the critical need for resilient and adaptable fault detection methods, this study investigates the use of deep learning techniques, specifically ANN and CNN models, for fault detection, classification, and localization within the IEEE 9-bus system. By simulating faults under both normal and fault conditions, this research evaluates the effectiveness of these models in managing diverse fault scenarios. Additionally, it explores the potential of transfer learning in N-1 contingency cases and selective feature reduction to enhance model robustness, improving adaptability to unforeseen disturbances and providing a cost-effective solution for practical deployment in real-world power systems.

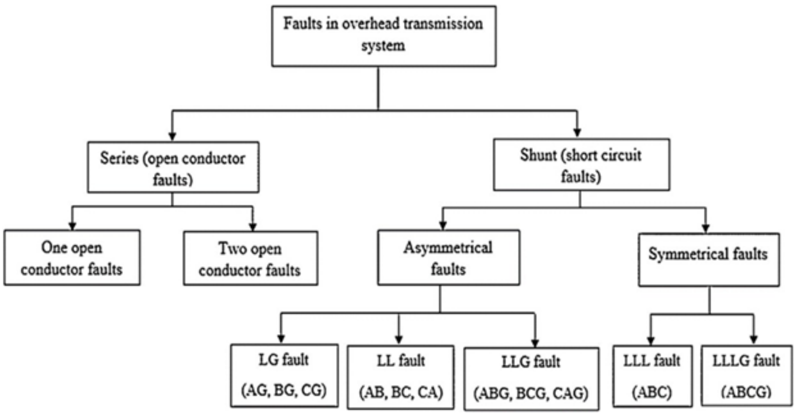

Figure 1: Classification of fault [2]

2. Methodology

2.1 Research framework

The framework shown in Figure 2 presents the overall workflow for fault detection, classification, and localization in a power system using deep learning models. It has three main stages: Data Generation Base Case, Offline Training, and Online Implementation, with an additional section dedicated to Transfer Learning for N-1 Contingency Case. During the Offline Training phase, the generated dataset is divided into Training and Test Sets or subjected to Cross-validation using 10-fold splits. Deep learning models, such as ANN and 1D-CNN, are trained on this data to perform fault detection and localization. After training, the models are evaluated based on their performance in the base case (without any contingency), and the results are documented.

The trained deep learning models are saved for future use and possible transfer learning scenarios. The Online Implementation phase involves deploying the trained models for real-time fault monitoring at a control center. Time-series voltage and current data, obtained from a Wide-Area Monitoring System (WAMS), is continuously fed into the models. These deployed models monitor the system for fault conditions in real-time and identify the Fault Class, Faulty Line, and Fault Location when faults occur. This system operates in an online environment, providing ongoing fault detection for the power system network. Finally, the

Transfer Learning for N-1 Contingency Case adapts the pre-trained base case models to account for changes in network topology due to a single component failure. A new dataset, representing the N-1 contingency scenario, is generated, following a similar process as in the base case. The pre-trained model is fine-tuned using this new dataset, allowing it to handle system conditions with altered configurations. The model's performance is then re-evaluated for the N-1 contingency case, and the newly trained model is saved for future deployment in online systems.

The proposed methodology aims to build a comprehensive fault detection, classification, and localization system for power networks using deep learning models. By leveraging a combination of Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Transfer Learning, the system addresses the core tasks of Faulty Line Identification (FLI), Fault Class Type (FCT) determination, and Fault Location Estimation (FLE). Each of these deep learning techniques contributes uniquely to optimizing model performance, as described in the following sections 2.3 and 2.4.

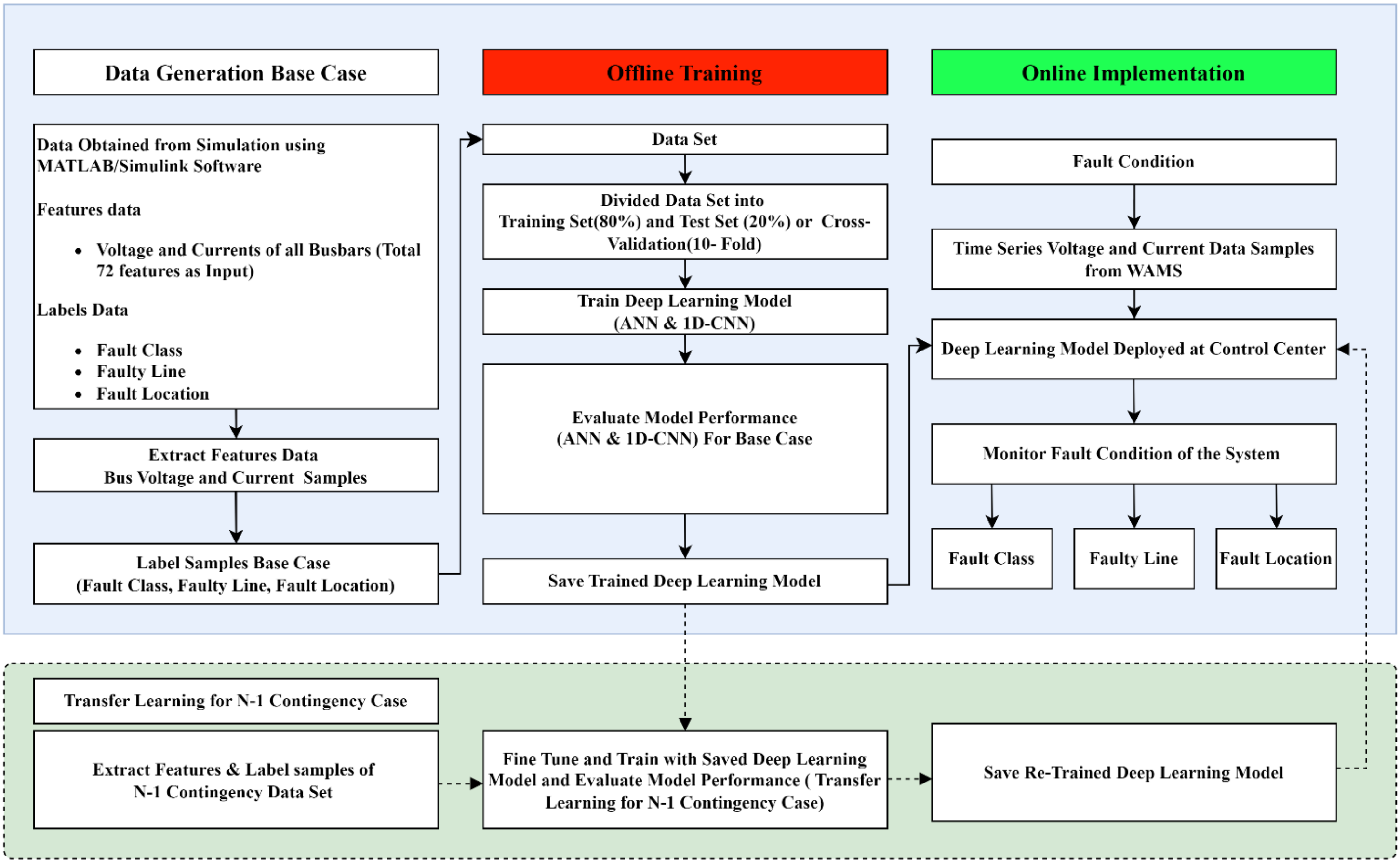

Figure 2: Proposed framework for fault condition assessment using deep learning

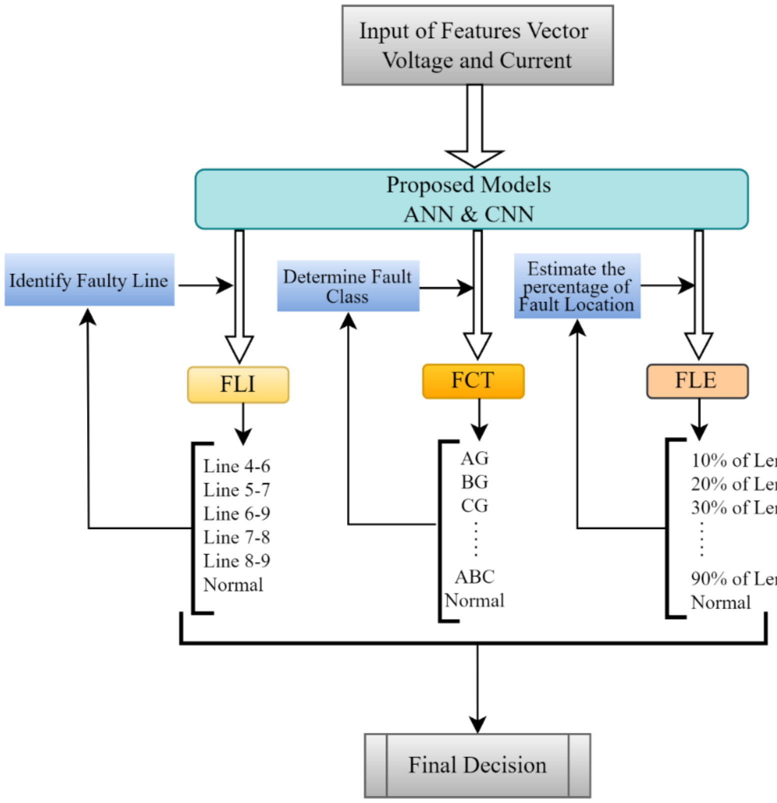

Figure 3: Flowchart of the Project based on ANN and CNN in general

As illustrated in Figure 3, the input features consist of voltage and current vectors obtained from various busbars in the power system. These features are fed into the proposed models, which are based on Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN) architectures. The models process the input data to perform three critical tasks Faulty Line Identification (FLI), Fault Class Type (FCT) and Fault Location Estimation (FLE).

2.2 System Simulation and Data Collection

This section outlines the process of generating and preparing the dataset for fault detection, classification, and localization using the IEEE 9-bus system. The IEEE 9-bus system is a standard test case used for power system analysis, providing a simplified model of a transmission network. The system consists of loads, transmission lines, and generators as shown in Table 1. The transmission lines are modeled as medium lines with three-phase pi section lines. Additionally, the system includes three loads that consume both active and reactive power at bus 5, bus 6, and bus 8.

Figure 4 shows the single-line diagram of IEEE 9-bus system, which comprises three traditional voltage sources with a voltage rating for transmission lines of 220 kV and a frequency of 60 HZ. The system has six transmission lines with a length of 100 km, 9 buses, 3 transformer and three general loads i.e. Load A, Load B and Load C.

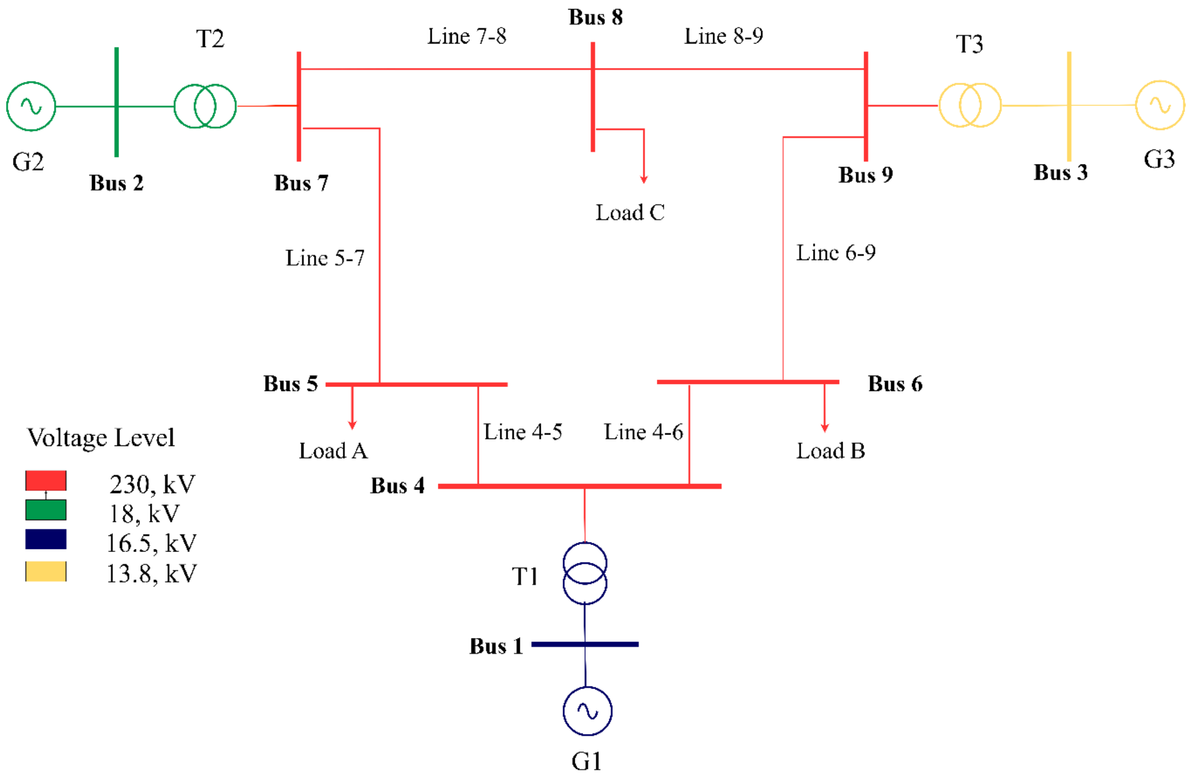

Figure 4: Single-line diagram of IEEE 9-Bus System

Table 1: IEEE 9-bus system, generator information and load data

Gen No.	S(MVA)	P(MW)	Q(MVAR)	V(pu)	Load at Bus No	P(MW)	Q(MVAR)
G1	247.5	71.6	27	1.04	At Bus 5	125	50
G2	192	163	6.7	1.03	At Bus 6	90	30
G3	128	85	-10.9	1.02	At Bus 8	100	35

Fault simulations were performed using MATLAB/Simulink. The system allows the introduction of various fault types through a three-phase fault block, where faults could be introduced at different phases (A, B, C) and ground (G). The transmission lines were divided into two sections, enabling variation in resistance and inductance to simulate faults at different locations. The simulation system proposed in this study captures voltage and current signals in three phases, which are sampled at a frequency of 500 Hz, resulting in 100 samples in total over the 0.2 second simulation period. To obtain sufficient data for training and evaluating the proposed model's performance, various fault and non-fault scenarios are simulated by adjusting the system's parameters and settings. Table 2 presents detailed configurations for both fault and non-fault scenarios. These simulations are conducted to ensure the model's versatility and ability to detect and classify different types of faults that may occur in the system.

Table 2: Configuration for fault possibility cases during simulation

Parameters	Possible Configuration	No. of Cases
Fault Class	A-G, B-G, C-G, AB, BC, AC, AB-G, BC-G, AC-G, ABC, and Normal	11
Faulty Line	Line 4-5, Line 4-6, Line 5-7, Line 6-9, Line 7-8, line 8-9, and Normal	7
Fault Location	10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% of line length, and Normal (100 km)	10

The simulations were carried out using MATLAB/Simulink 2024a to generate time-series voltage and current signals for the IEEE 9-bus system. All experiments were conducted on a system equipped with an AMD Ryzen 5 7520U processor with Radeon Graphics (8 CPUs) operating at approximately 2.8 GHz, along with 8 GB of RAM. The operating system used was Windows 11 Home Single Language (64-bit, Build 22631). The training progress was implemented using the Python within a Jupyter notebook environment.

2.3 Artificial Neural Network (ANN)

Figure 5 outlines the proposed methodology through a schematic block diagram. The Artificial Neural Network (ANN) used in this research is a multi-layer feedforward network designed to process features extracted from the power system's operational data, specifically for detecting, classifying, and localizing faults in the IEEE 9-bus system. The ANN model is structured to enable it to learn complex relationships from the input data, such as voltage and current measurements during different fault conditions. The ANN architecture consists of three main parts: an input layer, hidden layers, and an output layer. The input layer receives feature sets from the dataset, and the hidden layers are composed of fully connected (Dense) layers with ReLU (Rectified Linear Unit) activation functions, which enhance the model’s capacity to capture non-linear relationships. Hyperparameter tuning is used to optimize the number of neurons in each hidden layer. The output layer utilizes a Softmax activation function, ideal for multi-class classification, as it outputs probability distributions over the different fault classes. During model compilation, the Adam optimizer with a learning rate of 0.001 is employed to adjust learning rates dynamically for efficient convergence. Categorical cross-entropy serves as the loss function, common for multi-class classification, with accuracy as the evaluation metric. The ANN model comprises five layers (L0–L4) with a total of 48,268 trainable parameters for fault classification (FCT), 55,240 for fault line identification (FLI), and 54,603 for fault location estimation (FLE), as detailed in Table 3.

The Artificial Neural Network (ANN) model follows a layered architecture where each dense layer performs a weighted sum of its inputs, adds a bias, and passes the result through an activation function. The mathematical operation for the dense layers in an Artificial Neural Network (ANN) can be expressed as:

$$Z_n = f[W_n . Z_{n-1} + b] \quad \text{... (1)}$$

Where, Z_n represents the output of the n-th layer, f denotes the activation function, W_n is the weight matrix, Z_{n-1} is the input to the layer (or the output of the previous layer), and b_n is the bias vector. Each dense layer in the ANN applies this operation sequentially to propagate information forward through the network.

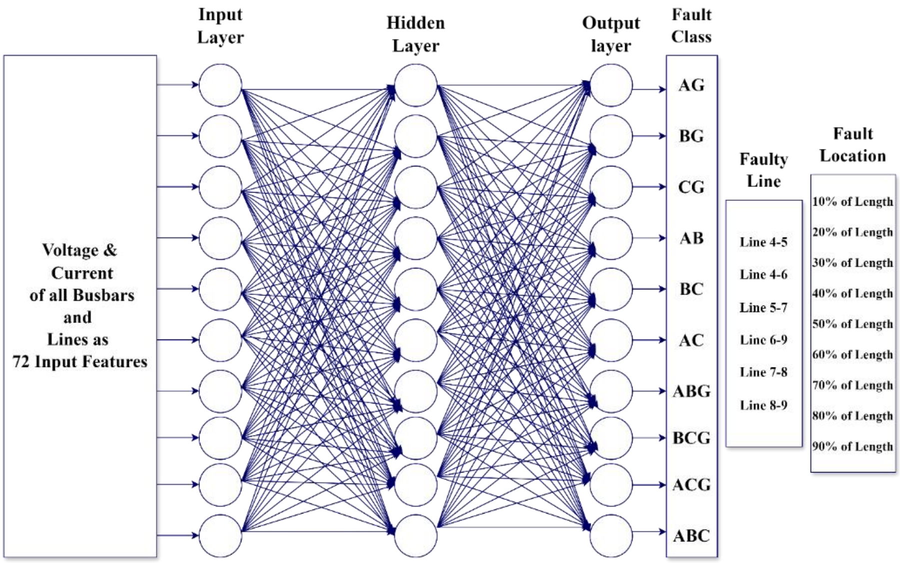

Figure 5: Proposed ANN architecture for fault classification, identification, and location

Table 3: Structures of ANN models

Layer No.	Layer Detail	Output Shape	Training Parameters
L0	Input data	(72 × 1)	-
L1	dense (Dense)	(None, 128)	9,344
L2	dense_1 (Dense)	(None, 256)	33,024
L3	dense_2 (Dense	(None, 512)	131,584
L4	dense_3 (Dense)	FCT (None, 11)	5,643
		FLI (None, 7)	3,591
		FLE (None, 10)	5,130

This equation illustrates the transformation from one layer to the next, beginning from the input layer $Z_0 = X$.

The forward propagation can be represented as:

1. Input Layer: $Z_0 = X$
2. Layer 1: $Z_1 = f_1(W_1.Z_0 + b_1)$
3. Layer 2: $Z_2 = f_2(W_2.Z_1 + b_2)$
4. Layer 3: $Z_3 = f_3(W_3.Z_2 + b_3)$
5. Output Layer: $Z = f_4(W_4.Z_3 + b_4)$

The output of the ANN model is given by the following compact equation:

$$Z = f_4[(W_4.f_3((W_3.f_2((W_2.f_1(W_1.\mathrm{X} + b_1) + b_2) + b_3) + b_4] \qquad \text{... (2)}$$

In this equation:

- X is the input feature vector.
- W_n and b_n are the weights and biases for the n-th layer, representing learnable parameters of the model.
- f_1, f_2, f_3 are the activation functions for the hidden first, second and third dense layers, typically ReLU or Tanh, which add non-linear transformations to the network.
- f_4 is the activation function for the output layer, often a Softmax function for classification tasks.

This sequential composition of transformations allows the ANN to map the input X to an output Z, learning the complex relationships in the data through its layered structure. Each layer extracts progressively higher-level features, culminating in an output that represents the probabilities for each class or the desired target values. This representation highlights the mathematical foundation of the ANN and its ability to model non-linear patterns in the data. In the proposed ANN model, three hidden layers with optimal configurations are utilized to extract complex patterns in the dataset. The model is optimized using the Adam optimizer with categorical cross-entropy loss. Custom train/test splits and 10-fold cross-validation methods are employed to ensure robust evaluation of the model's performance. This architecture enables the ANN to effectively detect, classify, and localize faults in the IEEE 9-bus system.

2.4 Convolution Neural Network

The Convolutional Neural Network (CNN) is another deep learning architecture utilized in this research, designed for fault classification, identification and localization using time-series data of voltage and current waveforms from the IEEE 9-bus system. CNN's are particularly effective in recognizing patterns in time-series data due to their ability to capture local spatial dependencies through convolutional filters. However, the study proposes a deep learning algorithm using One-Dimensional Convolutional Neural Network (1D-CNN) for Fault Classification (FCT), Line Faulty Detection (LFI), and Fault Location Estimation (FLE) based on raw data samples. 1D-CNN is well-suited for handling the sequential nature of power system data (such as voltage, current, or sensor readings over time). The 1D-CNN architecture consists of three primary layers: the convolutional layer, pooling layer, and fully connected layer. The architecture of the 1DCNN model is visually depicted in Figure 6.

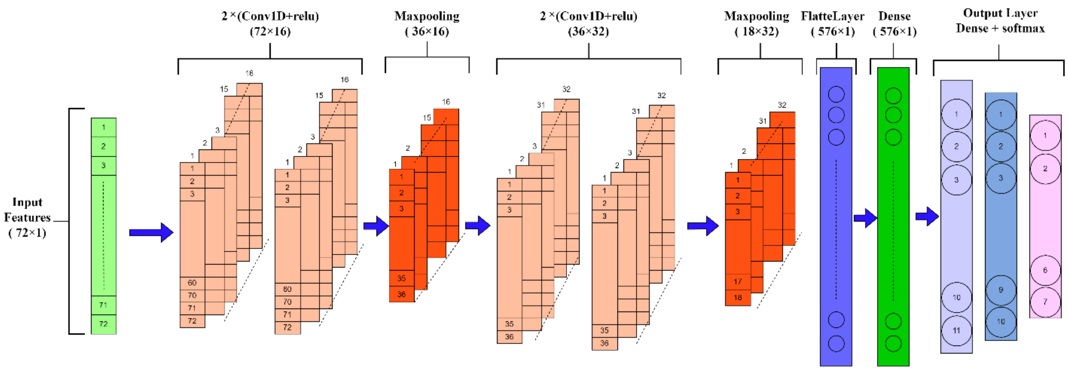

Figure 6: Architecture of the proposed 1D-CNN model

The 1D-CNN model's input layer is configured to accept time-series data shaped as (72, 1), where 72 timesteps correspond to voltage and current waveforms measured under faulted and normal conditions. Each convolutional layer employs a kernel size of 3, with 16 or 32 filters to capture local dependencies critical for fault localization. After each pair of convolution layers, MaxPooling1D layers are used to reduce data dimensionality, mitigating overfitting and accelerating learning. Batch Normalization layers are also applied to stabilize and expedite training by normalizing the outputs of the convolution layers. The output from these layers is then flattened into a one-dimensional vector and passed through a dense layer with 576 neurons, culminating in a SoftMax activation function for multi-class classification in the final dense layer. The CNN model is compiled using the Adam optimizer with a 0.001 learning rate, categorical cross-entropy as the loss function, and accuracy as the main evaluation metric. The 1D-CNN comprises 11 layers (L0–L10), including four dense or weighted layers, as detailed in Table 4.

Table 4: Architectural details of proposed 1D-CNN model

Layer No.	Layer Detail	Kernel Size	Output Shape	Training Parameters
L0	Input data	-	72 × 1	-
L1	Conv1d_2+Relu	1 × 3	72 × 16	64
L2	Conv1d_3+Relu	1 × 3	72 × 16	784
L3	MaxPooling1D	1 × 2	36 × 16	0
L4	Conv1d_4+Relu	1 × 3	36 × 32	1,568
L5	Conv1d_5+Relu	1 × 3	36 × 32	3,104
L6	MaxPooling1D	1 × 2	18 × 32	0
L7	BatchNormalization	-	18 × 32	128
L8	Flatten	-	576	0
L9	Dense	-	576	332,352
L10	Dense+SoftMax	-	11	6,347
		-	7	4,039
		-	10	5,770

The activation function, such as ReLU or Tanh, transforms inputs into non-linear outputs, functioning as a gate that controls information flow between input, hidden neurons, and output layers in the network. Each convolutional layer employs the rectified linear activation function (ReLU or Tanh), enhancing the network's ability to model non-linear relationships in the data.

The mathematical operation for the one-dimensional convolution is represented by the following equation:

$$C_j = f\left[\sum_{i=1}^{n} (x_{i+j} \times w_i) + b\right] \quad \text{... (3)}$$

where C_j, is the output of the convolution layer, f denotes the ReLU activation function, x and W represent the mini-batch of input data and the filters respectively, b is the bias term, and n is the dimension of the filter. Batch normalization (BN) is also applied to input x to stabilize the learning process. The model's 1D convolutional layers, positioned at layers L1, L2, L4, and L5, utilize learnable filters to capture features. The first stacked convolutional layer employs 16 filters, while the second stacked convolutional layer uses 32 filters. Both the 10-fold cross-validation method and custom train/test split were employed.

2.5 Hyperparameter and Training Process

Hyperparameter tuning is essential in optimizing the performance of the Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) models used in this study, as these models are sensitive to their hyperparameter settings. Hyperparameters, such as the number of neurons in each layer, activation function type, batch size, number of epochs, and learning rate, must be defined before training and significantly impact the model's ability to generalize. To select the optimal configuration, RandomizedSearchCV is employed. This method randomly samples a specified number of hyperparameter combinations, which are evaluated using cross-validation to identify the best-performing settings without the computational demands of exhaustive grid search. The finalized hyperparameters are outlined in Table 5 and Table 6.

For both ANN and CNN, the Adam optimizer is used with a learning rate of 0.001, categorical cross-entropy loss function, and a Softmax output layer for multi-class classification tasks.

Table 5 : Hyperparameter Search Values

Hyperparameter	**Search Values**
Neurons Layer 1	[64, 128, 256]
Neurons Layer 2	[64, 128, 256]
Neurons Layer 3	[128, 256, 512]
Activation Function	['ReLU', 'tanh']
Batch Size	[64, 128]
Epochs	[50, 100]
Learning Rate	[0.0001, 0.001, 0.01]
Optimizer	Adam (fixed)

Table 6: Hyperparameter Tuning Search Space

Hyperparameter	**ANN Fault Class**	**ANN Line Faulty**	**ANN Fault Location**
Optimizer	Adam	Adam	Adam
Neurons Layer 1	64	64	256
Neurons Layer 2	64	256	64
Neurons Layer 3	512	128	256
Activation Function	Tanh	ReLU	ReLU
Batch Size	128	128	128
Epochs	100	100	100
Learning Rate	0.001	0.001	0.001
Loss Function	categorical	categorical	categorical
Output Layer	Softmax	Softmax	Softmax

The training process is vital to developing robust deep learning models that can accurately predict and generalize to new data. Both ANN and CNN models are trained to handle fault classification, faulty line detection, and fault location tasks in power systems. The training follows two approaches: Normal Classification (NC) and Cross-Fold Validation (CV). In NC, 80% of the dataset is used for training, and 20% for testing, while CV applies a 10-fold cross-validation, exposing the model to the entire dataset and mitigating overfitting. During training, models iteratively optimize trainable parameters using gradient descent and backpropagation, running for up to 100 epochs with a batch size of 128. Once training is complete, the model is tested in an online mode, predicting fault classifications, fault locations, and line faults using the trained parameters without additional preprocessing, as illustrated in Figure 7. This structured workflow underscores the transition from training to real-time testing and demonstrates the model's effectiveness in addressing power system fault scenarios.

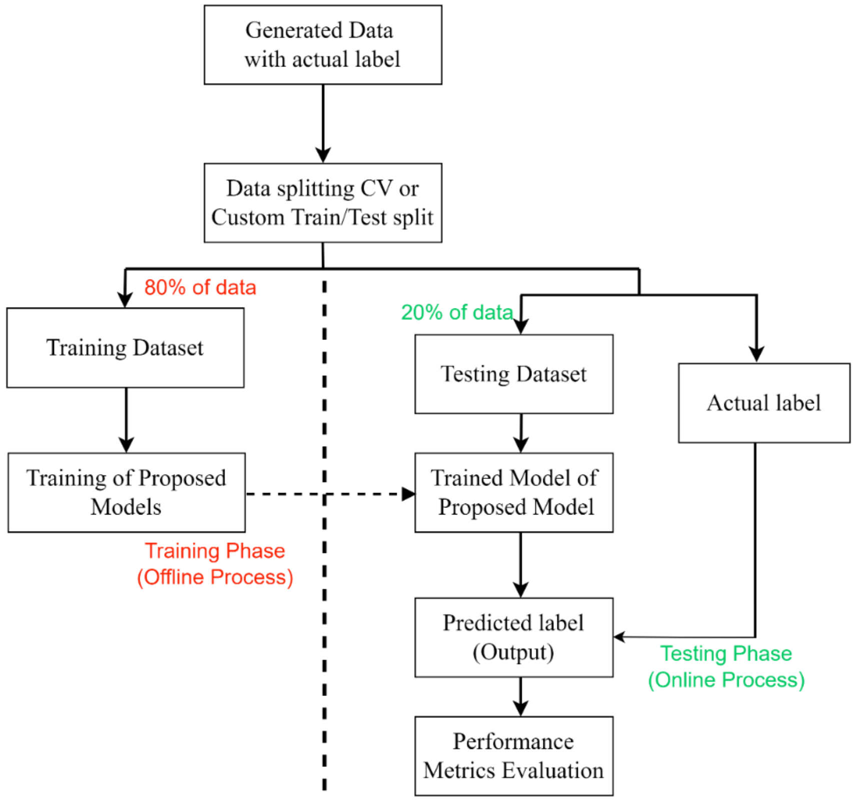

Figure 7: Flow chart of data training and testing based on proposed methods

2.6 Performance Evaluation

The evaluation of the proposed models is a critical aspect of determining their performance and reliability in fault detection, classification, and localization tasks within power systems. To evaluate the models' performance various performance metrics were used: Accuracy, Precision, Recall and F1-Score, based on based on True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN). Accuracy is used commonly to evaluate the performance of the deep learning and machine learning models for the assessment of fault condition in power systems. It is calculated from the confusion matrix using Equation (4). In the context of binary and multiclass classification, the abbreviations have the following meanings:

• TP (True Positives) represents the number of positive elements correctly predicted as positive.

• FP (False Positives) denotes the number of negative elements incorrectly predicted as positive.

• FN (False Negatives) represents the number of positive elements incorrectly predicted as negative.

• TN (True Negatives) denotes the number of negative elements correctly predicted as negative.

The performance metric formulas are as follows:

$$\text{Accuracy} = \frac{TP + TN}{\text{Sum of Confusion matrix}} \tag{4}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

Accuracy is the ratio of correctly predicted instances to the total instances. Precision is the ratio of true positive predictions to the total number of instances predicted as positive. A high precision indicates that the model is reliable in making positive predictions. The recall measures how many positive class predictions were made out of all positive class predictions in the data set. Finally, the F1-score provides a single score that balances both the concerns of precision and recall in one number. It is the harmonic mean of precision and recall. The greater the F1-score, the better the classifier's performance.

2.7 Transfer Learning for N-1 Contingency Case

Transfer learning is a key approach used to improve the performance of deep learning models in scenarios where only a limited amount of new data is available. In this project, transfer learning is applied to adapt pre-trained ANN and CNN models for the N-1 contingency case on the IEEE 9-bus system, where one generator has been removed. In normal base case, the original deep learning models (both ANN and CNN) are pre-trained on data generated from normal operational conditions of the IEEE 9-bus system. The model learns to predict fault class, line faulty, and fault location using the dataset that reflects the entire system functioning with all components in place. But, For the transfer learning phase, the pre-trained models are fine-tuned on the new dataset, which is collected after simulating the failure of one generator (N-1 contingency). The fine-tuning process involves retraining only the final few layers of the network or adjusting the learning rates to ensure the model adapts without overfitting to the new data. Finally, the re-trained deep learning model at the control center can classify fault, identify faulty line and estimate the fault location of the power system.

2.8 Selective Feature Reduction for Cost-Effective PMU Deployment

In modern power systems, the deployment of phasor measurement units (PMUs) is essential for real-time monitoring, fault detection, classification, and location estimation. However, the cost of installing and maintaining PMUs at every bus in large networks can be prohibitive. To address this challenge, this study proposes a selective feature reduction approach aimed at minimizing the number of PMUs required while ensuring the effectiveness of fault analysis. By focusing on strategically selected buses, the study aims to maintain high monitoring accuracy with a reduced financial investment.

For this purpose, Buses 5, 6, and 8 were strategically chosen for monitoring based on their critical positions in the IEEE 9-bus system. These buses were selected due to their high connectivity and their placement in areas where faults are more likely to propagate through multiple branches. By focusing on these key buses, essential information about the power network's state can still be captured, allowing for effective fault analysis with fewer PMUs. The feature reduction process involved isolating data from these selected buses and using only this subset of features in model training. This approach not only reduces the cost associated with PMU installation and maintenance but also simplifies data processing by reducing the dimensionality of the input data. With this reduced feature set, the study aimed to train and validate that ANN and CNN models could maintain high levels of accuracy in fault detection tasks while relying on data from only a few critical locations within the network. This selective deployment of PMUs exemplifies a practical and cost-effective solution for utilities, enabling robust monitoring without the financial burden of comprehensive network-wide PMU coverage.

For model evaluation, the primary measure of success was the accuracy of the fault detection, classification, and location estimation tasks. This focus on accuracy is critical in demonstrating that the selective feature reduction did not compromise the effectiveness of the models. Emphasizing accuracy as the evaluation metric aligns with industry standards for fault detection, where timely and reliable identification of issues is paramount.

3. Results and Discussions

This section provides a detailed evaluation of the deep learning models developed for fault detection, classification, and localization within the IEEE 9-bus power system. The performance of the Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) models is analyzed across two scenarios: normal (base case) operation and an N-1 contingency case, wherein Generator 1 is removed. Fault data, simulated using MATLAB Simulink, was used to train and test the models in various fault diagnosis tasks. Evaluations focus on training and loss curves, confusion matrices, and 10-fold cross-validation results. A comparative analysis of model performance in the N-1 contingency case with and without transfer learning is also included, demonstrating the benefits of model fine-tuning on unseen fault data. This section ultimately seeks to validate the effectiveness and robustness of the models in accurately detecting, classifying, and localizing faults, providing a balance of complexity and performance.

3.1 Performance of the ANN Model

The performance of the ANN model across fault classification, line identification, and fault location tasks demonstrates its effectiveness in power system fault diagnosis. The model achieved high accuracy rates, with 92.35% for fault classification, 99.54% for line identification, and 96.24% for fault location estimation. These results underscore the ANNs ability to accurately predict fault classes, identify faulty lines, and estimate fault locations. Figures 8 and 9 further illustrate the model's performance. Figure 8 presents the ANNs training and loss curves, showing a steady increase in training and validation accuracy while both training and validation loss stabilized, indicating minimal overfitting. Figure 9 displays the confusion matrix for the Artificial Neural Network (ANN) model applied to fault detection, classification, and localization tasks in the power system, which highlights the model's strengths in correct classification across tasks.

It represents the predicted results in the matrix’s rows and the actual results in the columns. The matrix’s diagonal dominance reveals high precision, with few misclassifications, mainly among similar fault classes and adjacent lines. Additionally, 10-fold cross-validation further validated the robustness of the ANN model, confirming its generalization ability for fault classification. Overall, these results demonstrate the ANN model’s practical potential for real-time fault detection and diagnosis in power systems, effectively balancing accuracy and computational efficiency.

The performance of the proposed ANN model for fault classification was assessed through 10-fold cross-validation, and the outcomes are presented in Table 7 and Table 8. A detailed analysis of Fold 5 is provided as a representative sample for in-depth understanding. Additionally, Figure 10 displays the training curve for specific fold number 5 for fault classification using ANN model.

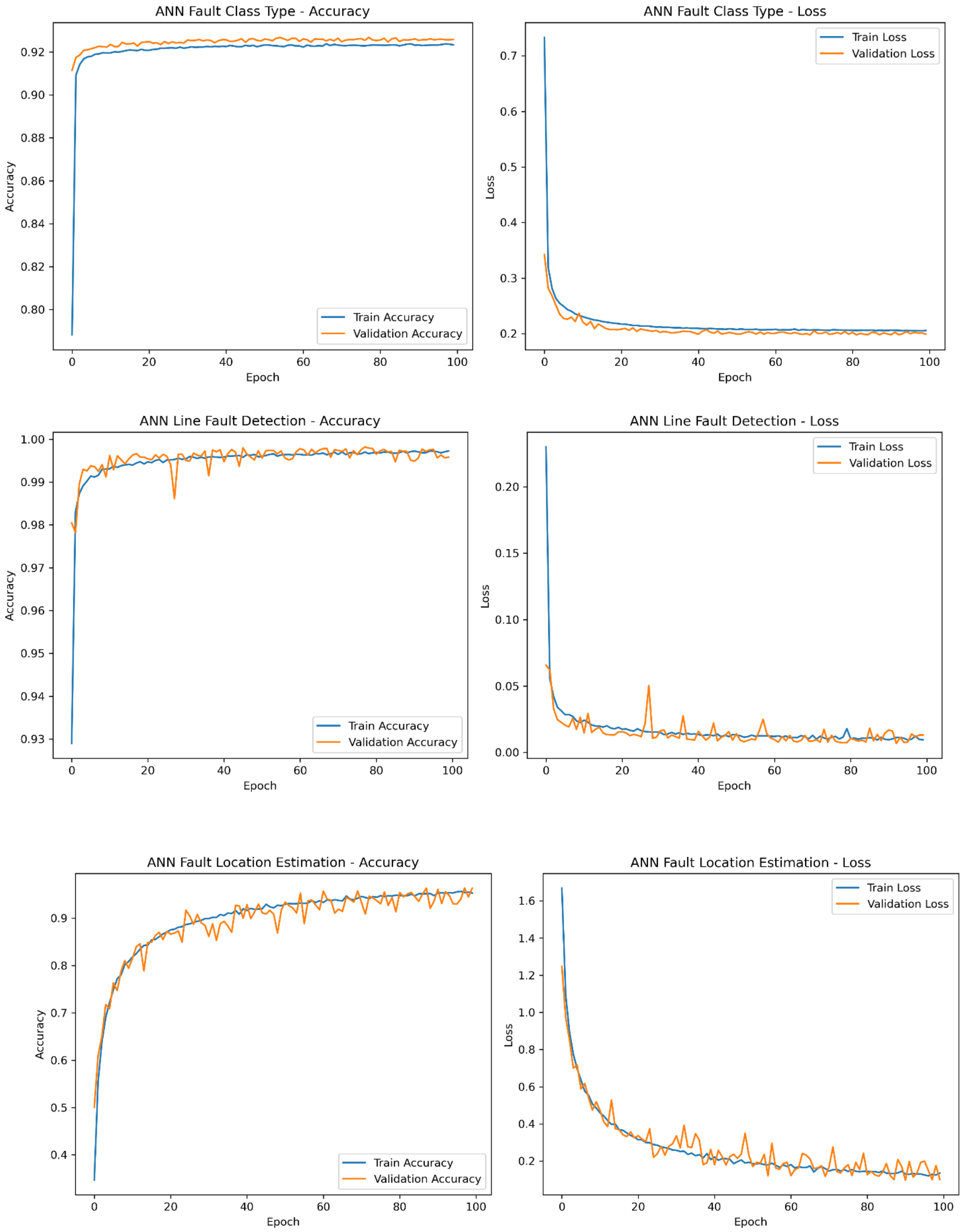

Figure 8: Training and loss curve of ANN for training progress

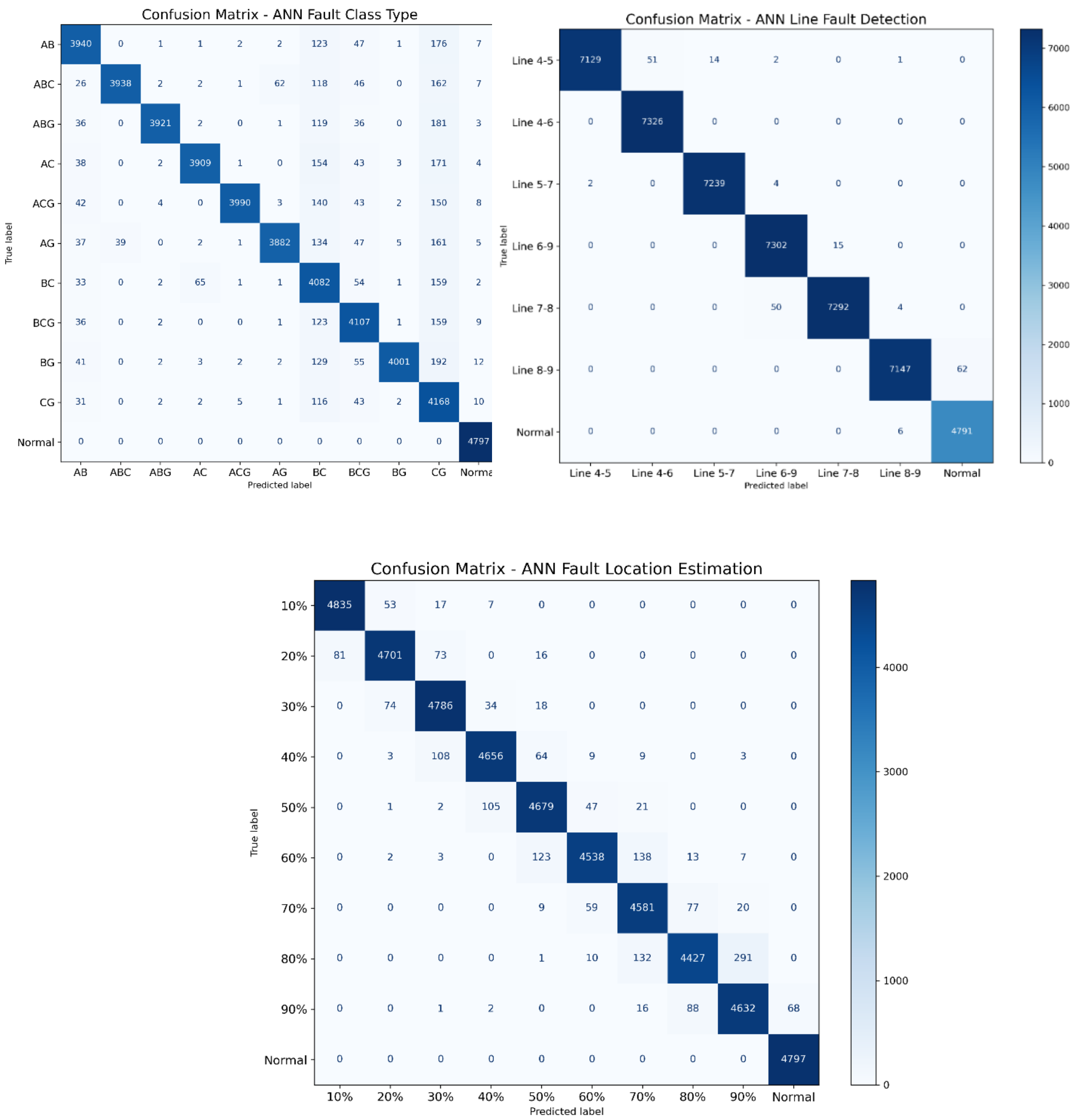

Figure 9: Confusion matrix of the proposed ANN model

Table 7: Performance result of the 10-fold cross-validation for fault classification using the ANN

Fold No.	Accuracy	Precision	Recall	F1-Score
Fold-1	0.921673	0.929185	0.920852	0.922697
Fold-2	0.924605	0.933116	0.924102	0.926091
Fold-3	0.923118	0.933005	0.922537	0.92483
Fold-4	0.924398	0.930619	0.923728	0.925241
Fold-5	0.924023	0.927426	0.923049	0.924103
Fold-6	0.921677	0.928896	0.921146	0.922638
Fold-7	0.920638	0.922866	0.920138	0.92052
Fold-8	0.921629	0.929223	0.920807	0.922786
Fold-9	0.925184	0.930937	0.924515	0.925716
Fold-10	0.924354	0.930823	0.923684	0.925195
Average	0.923129	0.929617	0.922456	0.923982

Table 8: Performance result detailed metrics of fold no. 5 for the proposed ANN

Class	Precision	Recall	F1-Score	Support
AB	0.99	0.91	0.95	2151
ABC	0.58	0.99	0.73	2205
ABG	1.00	0.90	0.95	2170
AC	0.98	0.91	0.94	2147
ACG	1.00	0.90	0.95	2232
AG	0.93	0.90	0.91	2117
BC	1.00	0.89	0.94	2211
BCG	1.00	0.91	0.95	2148
BG	0.97	0.90	0.94	2261
CG	1.00	0.92	0.96	2129
Normal	0.98	1.00	0.99	2448
Average	0.95	0.92	0.93	24219

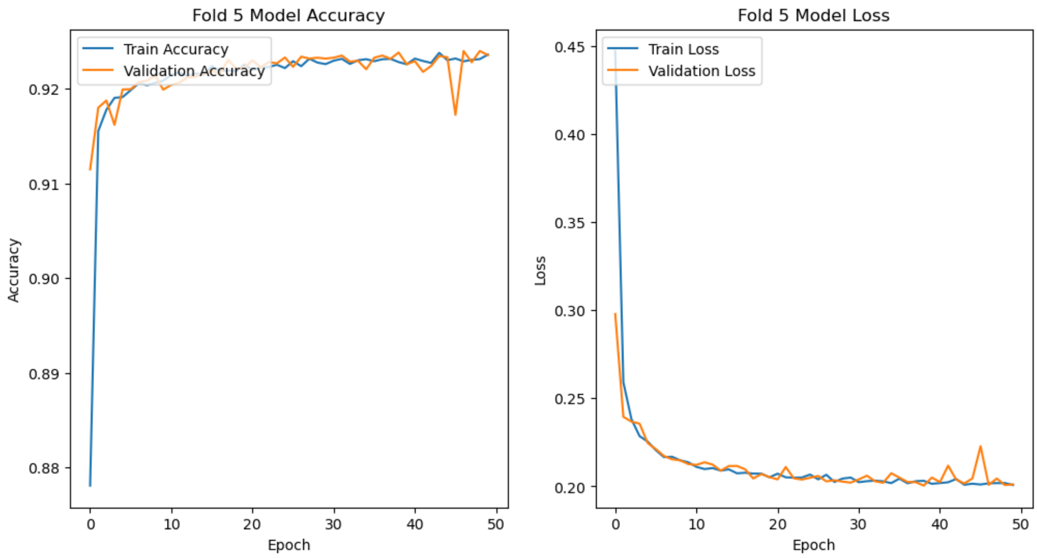

Figure 10: Training and loss curve of fold no. 5 for fault classification using ANN model

3.2 Performance of the CNN Model

The performance of the CNN model across fault classification, line identification, and fault location tasks highlights its strong capability in power system fault detection and diagnosis. The model achieved exceptional accuracy rates, with 92.42% for fault classification, 99.87% for line identification, and 96.95% for fault location estimation. These results demonstrate the CNN’s high precision in identifying fault types, pinpointing faulty transmission lines, and estimating fault locations. Figure 11 shows the training and loss curves for fault classification, revealing steady increases in training and validation accuracy, while both training and validation loss curves converge, suggesting strong generalization with minimal overfitting.

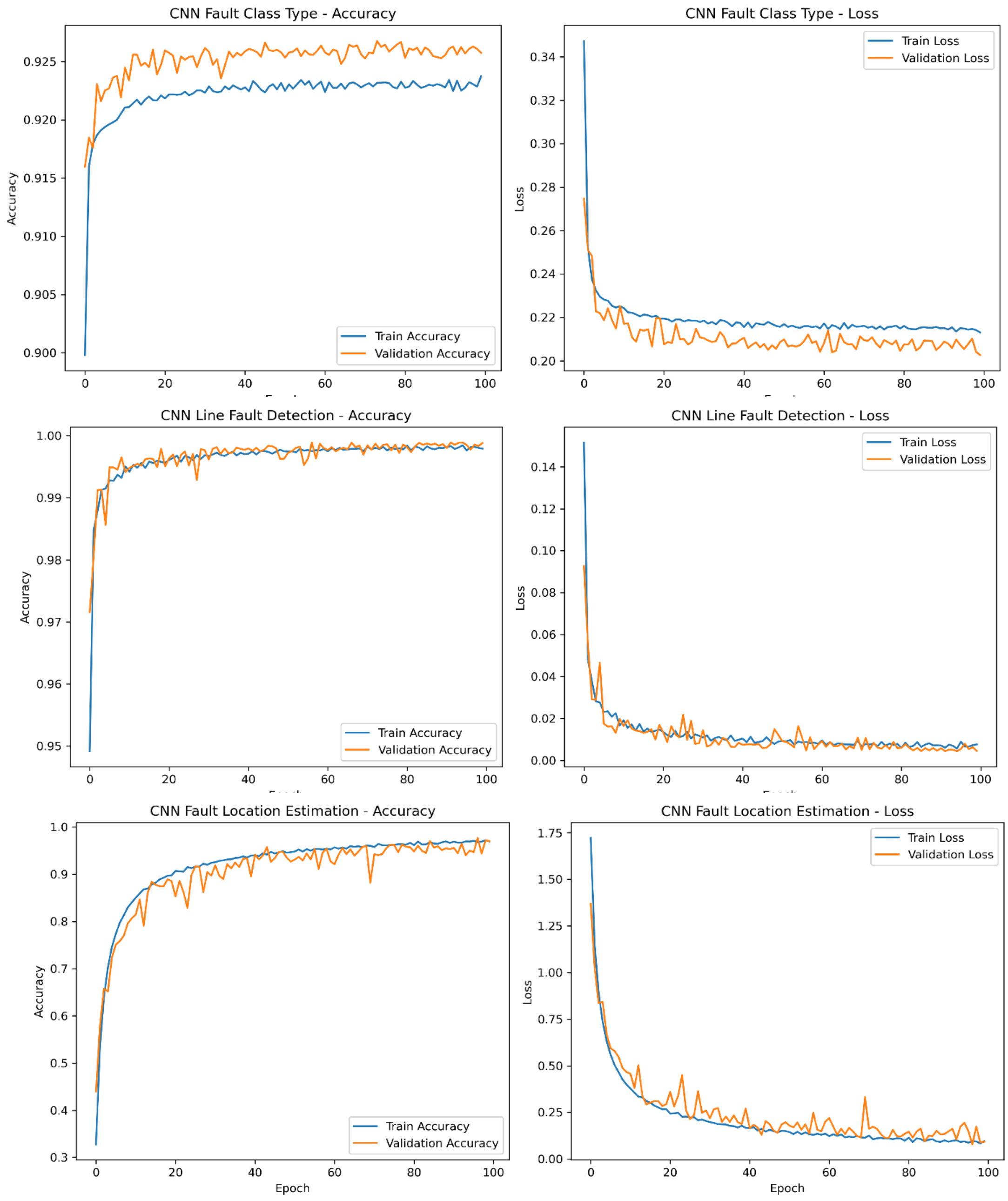

Figure 11: Training and Loss Curve of CNN for training progress

Figure 12 showcases the confusion matrix results for the CNN model across all three tasks. The matrix reveals high precision in fault classification, faulty line identification, and fault location estimation, with diagonal dominance showing correct classifications for the majority of instances. Minor misclassifications are observed between similar fault classes, but these are minimal and do not significantly impact the overall performance.

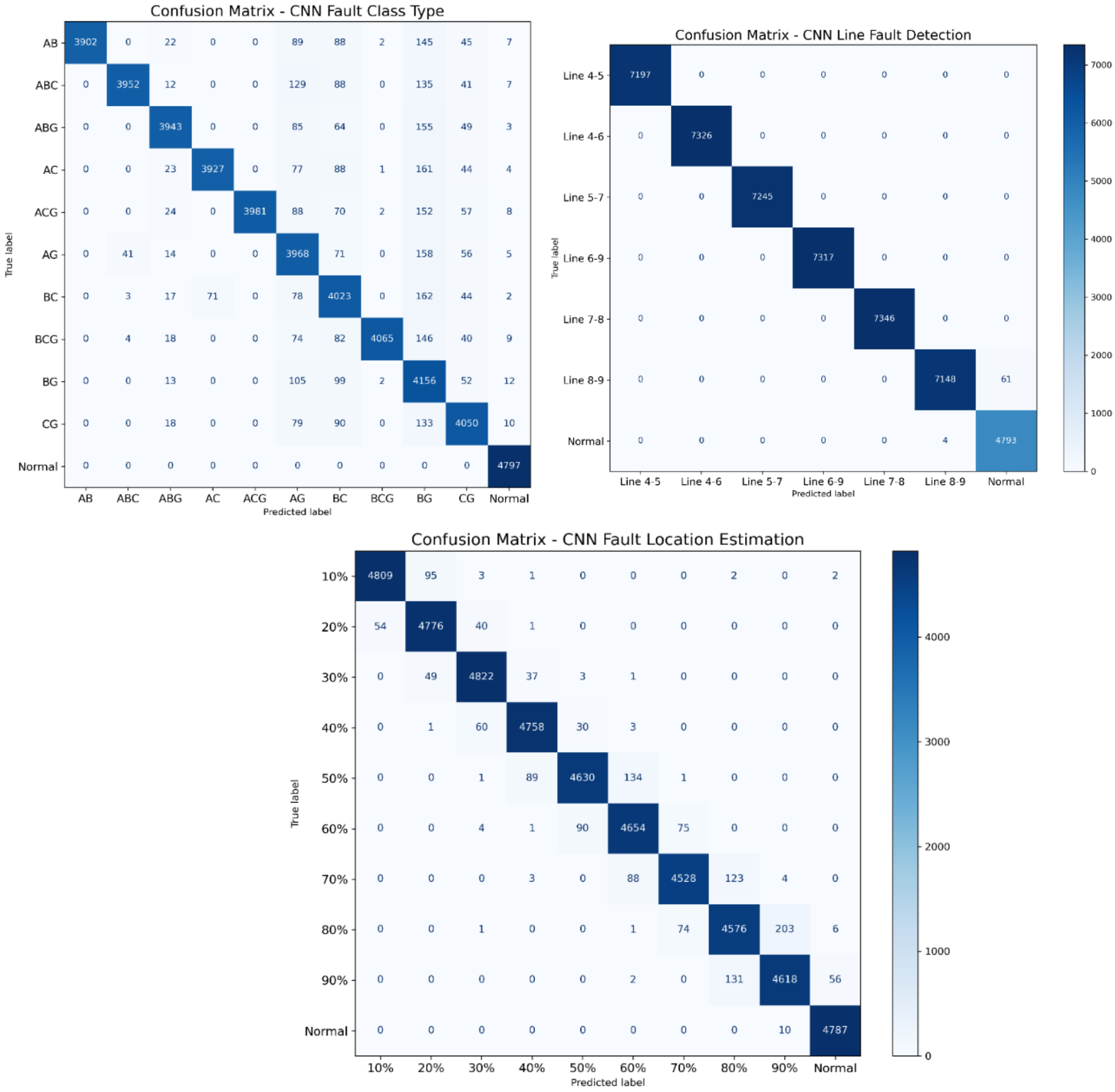

Figure 12: Confusion matrix of the Proposed CNN model

Furthermore, the CNN model was validated using 10-fold cross-validation, providing robust evidence of its generalization ability across different fault conditions. The cross-validation results further support the model's potential for real-time deployment in power system diagnostics, highlighting its efficiency and reliability in fault detection tasks. The performance of the proposed CNN model for fault classification was assessed through 10-fold cross-validation, and the outcomes are presented in Table 9 and Table 10. A detailed analysis of Fold 5 is provided as a representative sample for in-depth understanding. Additionally, Figure 13. displays the training curve for specific fold number 5 for fault classification using CNN model.

Table 9: Performance result of the 10-fold cross-validation for fault classification using the CNN

Fold No.	Accuracy	Precision	Recall	F1-Score
Fold-1	0.92143	0.93291	0.92061	0.92350
Fold-2	0.92436	0.94745	0.92381	0.93008
Fold-3	0.92279	0.94775	0.92232	0.92863
Fold-4	0.92427	0.94182	0.92368	0.92814

Fold-5	0.92336	0.93594	0.92233	0.92565
Fold-6	0.92105	0.94977	0.92062	0.92815
Fold-7	0.92126	0.93788	0.92038	0.92482
Fold-8	0.92126	0.93788	0.92038	0.92482
Fold-9	0.92473	0.93336	0.92412	0.92576
Fold-10	0.92473	0.94134	0.92407	0.92830
Average	0.92291	0.94104	0.92225	0.92689

Table 10: Performance result detailed metrics of fold no. 5 for the proposed CNN

Class	Precision	Recall	F1-Score	Support
AB	1.00	0.92	0.96	2157
ABC	0.86	0.92	0.89	2130
ABG	1.00	0.90	0.95	2194
AC	0.98	0.91	0.95	2159
ACG	0.61	0.98	0.75	2202
AG	0.98	0.89	0.94	2167
BC	1.00	0.90	0.95	2216
BCG	1.00	0.92	0.96	2184
BG	1.00	0.91	0.96	2146
CG	1.00	0.91	0.95	2262
Normal	0.99	1.00	0.99	2401
Average	0.95	0.93	0.93	24218

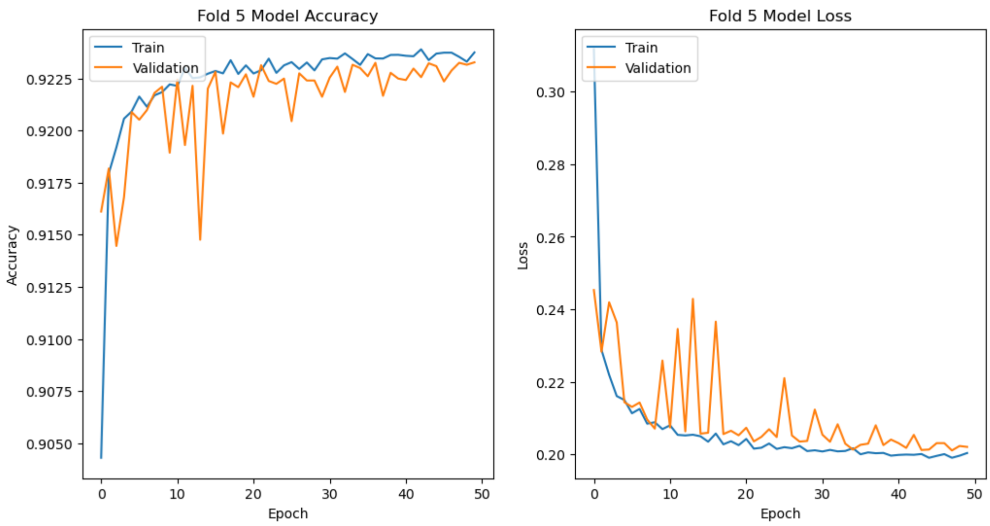

Figure 13: Training and loss curve of fold no. 5 for fault classification using CNN model

3.3 Performance Comparison

The performance comparison between the ANN and CNN models for fault detection, classification, and localization is summarized in Table 11.Both models demonstrate high accuracy across the three tasks, with CNN showing a slight improvement in performance over ANN.

Table 11: Performance Comparison Base Case

Metric	ANN Accuracy	CNN Accuracy
Fault Classification	92.35%	92.42%
Fault Line Identification	99.54%	99.87%
Fault Location Estimation	96.24%	96.95%

The performance differences, while not large, indicate that the CNN model is slightly better suited for this specific fault diagnosis application due to its ability to better capture complex patterns in the data. However, ANN still provides comparable performance, demonstrating its effectiveness as a simpler alternative. Table 12 demonstrates the significant improvements achieved through transfer learning (TL) in the N-1 contingency case, particularly for fault classification, faulty line identification and estimating fault location.

Table 12: Performance comparision using proposed transfer learning method for N-1 contingency case

Task	**ANN**		**CNN**	
	Test Accuracy (Without TL)	Test Accuracy (With TL)	Test Accuracy (Without TL)	Test Accuracy (With TL)
Fault Classification	53.039%	92.611%	59.679%	76.484%
Fault Line Identification	64.937%	100%	14.709%	95.557%
Fault Location	13.956%	95.787%	10.009%	27.373%

Here, the classification accuracy of the pre-trained model for the unseen N-1 contingency case is significantly low without transfer learning. However, it can be seen from the table that by applying transfer learning with fine-tuning, the pre-trained models were able to leverage the knowledge from the base case, improving their performance significantly.

The accuracy results for each task using selective feature reduction are summarized in the Table 13 below.

Table 13: Performance Comparision using Selective Feature Reduction

Metric	ANN Accuracy	CNN Accuracy
Fault Classification	92.233%	92.231%
Fault Line Identification	99.789%	99.818%
Fault Location Estimation	83.295%	85.047%

By limiting the dataset to features from Buses 5, 6, and 8, both the ANN and CNN models demonstrated strong performance in fault classification, line identification, and location estimation tasks indicate that selective feature reduction effectively balances cost efficiency and model performance in power system monitoring. These outcomes demonstrate that the strategic selection of monitored buses is sufficient for accurate fault analysis. However, the fault location estimation accuracy for both models revealed a minor trade-off, with ANN achieving 83.295% and CNN 85.047%. Although the selective feature approach resulted in a slight decrease in fault location accuracy, particularly in complex scenarios, this trade-off is justified by the significant reduction in PMU deployment costs.

4. Conclusions

This study presented a deep neural network (DNN)-based approach for fault detection, classification, and localization in power systems, using both ANN and CNN models within the IEEE 9-bus system. The results demonstrated that both models effectively performed fault detection tasks, with CNN exhibiting a slight advantage due to its ability to capture spatial correlations in the data.

The ANN model achieved notable accuracy rates, with 92.35% for fault classification, 99.56% for faulty line identification, and 96.27% for fault location estimation. Similarly, the CNN model delivered strong performance, achieving 92.42% for fault classification, 99.87% for faulty line identification, and 96.95% for fault location estimation. While both models performed well, the CNN model showed a marginal advantage in fault classification and location estimation, and both models were nearly flawless in identifying faulty lines, with CNN achieving slightly higher accuracy than ANN. This study also explored transfer learning to improve the models' ability to handle unseen data in N-1 contingency scenarios. Initially, without transfer learning, both models struggled to predict faults under these conditions. However, fine-tuning the models through transfer learning led to significant improvements. Additionally, a selective feature reduction strategy focusing on key buses (5, 6, and 8) allowed the models to maintain high diagnostic accuracy while reducing PMU requirements, offering a practical approach for cost-effective deployment.

In conclusion, CNN showed strong potential for fault condition assessment, with ANN providing a simpler yet effective alternative. Future research should focus on scaling these models to larger power systems, such as the IEEE 14-bus or 30-bus systems, and incorporating real-world fault data to increase model resilience against noise and load variations. Further development in real-time deployment frameworks and advanced transfer learning methods, such as for N-1-1 scenarios, will enhance these models’ adaptability and utility for comprehensive fault diagnosis in diverse grid environments.

Conflicts of Interest Statement

The authors declare no conflicts of interest for this study.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. S. Das, S. Santoso, and S. N. Ananthan, *Fault Location on Transmission and Distribution Lines: Principles and Applications*. John Wiley & Sons Ltd, 2021. doi: 10.1002/9781119121480.
2. A. Prasad, J. B. Edward, and K. Ravi, "A review on fault classification methodologies in power transmission systems: Part—I," *Journal of Electrical Systems and Information Technology*, vol. 5, no. 1, pp. 48–60, May 2018, doi: 10.1016/j.jesit.2017.01.004.
3. D. Mnyanghwalo, H. Kundaeli, E. Kalinga, and N. Hamisi, "Deep learning approaches for fault detection and classifications in the electrical secondary distribution network: Methods comparison and recurrent neural network accuracy

comparison," *Cogent Engineering*, vol. 7, no. 1, 2020, doi: 10.1080/23311916.2020.1857500.
4. N. Ketkar, *Deep Learning with Python*. Apress, 2017. doi: 10.1007/978-1-4842-2766-4.
5. P. Ray and D. P. Mishra, "Support vector machine based fault classification and location of a long transmission line," *Engineering Science and Technology, an International Journal*, vol. 19, no. 3, pp. 1368–1380, Sep. 2016, doi: 10.1016/j.jestch.2016.04.001.
6. M. Shafiullah, M. Ijaz, M. A. Abido, and Z. Al-Hamouz, "Optimized Support Vector Machine & Wavelet Transform for Distribution Grid Fault Location," 2017.
7. A. U. Q. Bhat, A. Prakash, V. K. Tayal, and P. Choudekar, "Three-Phase Fault Analysis of Distributed Power System Using Fuzzy Logic System (FLS)," in *Lecture Notes in Electrical Engineering*, vol. 721 LNEE, pp. 615–624, 2021.
8. M. Bhatnagar and A. Yadav, "Fault Detection and Classification in Transmission Line Using Fuzzy Inference System," in *2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, Dec. 2020, doi: 10.1109/ICRAIE51050.2020.9358386.
9. N. Swaraj, M. Yadnesh, M. Sahil, V. Jaydeep, and S. Bindu, "Classification of Power Transmission Line Fault Types and Location Using Machine Learning Techniques," [Online]. Available: https://ssrn.com/abstract=4107762.
10. H. Rezapour, S. Jamali, and A. Bahmanyar, "Review on Artificial Intelligence-Based Fault Location Methods in Power Distribution Networks," *MDPI Energies*, vol. 16, no. 12, Jun. 2023, doi: 10.3390/en16124636.
11. *2018 International Conference on Smart Electric Drives and Power System (ICSEDPS)*. IEEE.
12. K. Moloi and A. A. Yusuff, "Power Distribution System Fault Diagnostic Using Genetic Algorithm and Neural Network," in *2021 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa (SAUPEC/RobMech/PRASA)*, Jan. 2021, doi: 10.1109/SAUPEC/RobMech/PRASA52254.2021.9377241.
13. IEEE Control Systems Society Chapter Malaysia and Institute of Electrical and Electronics Engineers, *Proceedings, 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2020): Penang, Malaysia, 21-22 August 2020*.
14. V. N. Ogar, S. Hussain, and K. A. A. Gamage, "The use of artificial neural network for low latency of fault detection and localisation in transmission line," *Heliyon*, vol. 9, no. 2, Feb. 2023, doi: 10.1016/j.heliyon.2023.e13376.
15. P. Rai, N. D. Londhe, and R. Raj, "Fault classification in power system distribution network integrated with distributed generators using CNN," *Electric Power Systems Research*, vol. 192, Mar. 2021, doi: 10.1016/j.epsr.2020.106914.
16. 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, Nov. 2023, doi: 10.3390/en16227680.
17. A. Adhikari, S. Naetiladdanon, and A. Sangswang, "Real-Time Short-Term Voltage Stability Assessment Using Combined Temporal Convolutional Neural Network and Long Short-Term Memory Neural Network," *Applied Sciences*, vol. 12, no. 13, Jul. 2022, doi: 10.3390/app12136333.