Received: 23 January 2024 | Revised: 7 May 2024 | Accepted: 9 June 2024

DOI: 10.1002/eng2.12950

RESEARCH ARTICLE

Engineering Reports Open Access WILEY

Fault detection and classification in overhead transmission lines through comprehensive feature extraction using temporal convolution neural network

Nadeem Ahmed Tunio[1,2] | Ashfaque Ahmed Hashmani[2] | Suhail Khokhar[3] | Mohsin Ali Tunio[1] | Muhammad Faheem[4]

[1]Department of Electrical Engineering, Mehran University of Engineering & Technology Shaheed Zulfiqar Ali Bhutto Campus Khairpur Mir's, Khairpur, Sindh, Pakistan

[2]Department of Electrical Engineering, Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan

[3]Department of Electrical Engineering, Quaid-e-Awam University of Engineering, Science and Technology, Nawabshah, Sindh, Pakistan

[4]School of Technology and Innovations, University of Vaasa, Vaasa, Finland

Correspondence
Muhammad Faheem, School of Technology and Innovations, University of Vaasa, Vaasa, Finland.
Email: muhammad.faheem@uwasa.fi

Abstract

Faults in transmission lines cause instability of power system and result in degrading end users sophisticated equipment. Therefore, in case of fault and for the quick restoration of problematic phases, reliable and accurate fault detection and classification techniques are required to categorize the faults in a minimum time. In this work, 500 kV transmission line (Jamshoro-New Karachi), Sindh, Pakistan has been modeled in MATLAB. The discrete wavelet transform (DWT) has been used to extract features from the transient current signal for different faults in 500 kV transmission line under various parameters such as fault location, fault inception angle, ground resistance and fault resistance and time series data has been obtained for fault classification. Moreover, the temporal convolutional neural network (TCN) is used for fault classification in 500 kV transmission network due to its robust framework. From simulation results, it is found that faults in 500 kV transmission line are classified with 99.9% accuracy. Furthermore, the simulation results of the TCN model compared to bidirectional long short-term memory (BiLSTM) and Gated Recurrent Unit (GRU) and it has been found that TCN model is capable of classifying faults in 500 kV transmission line with high accuracy due to its ability to handle long receptive field size, less memory requirement and parallel processing due to dilated causal convolutions. Through this work, the meantime to repair of 500 kV transmission line can be reduced.

KEYWORDS

fault classification, fault detection, machine learning, temporal convolutional neural network, transmission line

1 | INTRODUCTION

Accurate detection and classification of faults in transmission lines can reduce power line renewal costs and increase the chances of power grid protection. Failure of transmission lines leads to power interruption to the customers.[1] Faults occur in grid-connected transmission lines as reported in the literature are line-to-line (LL), double line-to-ground (DLG), and three phase fault (LLL).[2] Faults in transmission lines cause various disturbances such as mechanical stress, unbalancing of power flow and overheating. Moreover, accurate Fault detection and classification (FDC) has important role in ensuring stability of the grid system. Although recovery from failure phases is human dependent and depends on detection and classification scheme used to identify type of failure along with its location throughout the network. This is crucial because a fast and accurate FDC enhances the chances of isolating faulty phases from the transmission system, ensures quick repair thus enhancing the power quality transient stability of the interconnected power network.[2]

In past literature, various feature extraction techniques and fault classification algorithms based on machine learning approaches have been studied for accurate fault identification and classification in transmission lines within minimum time.[3] Moreover, many tools such as S-transform (ST), short time Fourier transform (STFT), Fourier transform (FT) and discrete wavelet transform (DWT) have been used to extract features from the fault signal, evaluating to determine the health of the network. Since the fault information is obtained from voltage or current signals which include a lot of raw data, so fault categorization using raw data lacks regularity. This data may come under the influence of noise, making it difficult to differentiate between whether there has been a fault or otherwise. As faulty signal is transient in nature so it consists of various frequency components, which have a significant impact on the accurate fault classification, leading to inconsistent results.[4] The accuracy of fault identification and fault has therefore remained a concern and challenge for the researchers while transmitting power through extra high voltage transmission lines.

Studies reported in the literature have utilized multiple algorithms for the classification of faults in transmission lines, Artificial Intelligence (AI) and Machine Learning (ML) techniques are mostly used choices for fault classification due to their fast-learning ability, giving right output and recognizing learned patterns through input training data.[5] Fault classification techniques used in the past are categorized as prominent and modern techniques. The prominent techniques cover Wavelet Transform (WT) based analysis added with Artificial Neural Networks, fuzzy logic-based methods for fault classification also known as hybrid methods. Moreover, support vector machine (SVM), artificial intelligence (AI), phasor measurement units (PMU) and principal component analysis (PCA) are covered under modern fault classification techniques.[6]

The noise affects the fault signals, degrading its originality and affects the statistical values required in computation for analyzing the performance of fault detection and classification tools and techniques.[7] Therefore in order to de-noise the signal DWT is implemented for feature extraction and collection of training data for accurate classification of faults in transmission lines.[8] The DWT splits the transient signal into a sequence of components; these components represent high-frequency and low-frequency components obtained from the faulty signal using multi-resolution analysis (MRA). Referring to the simulation results, it is confirmed that high-frequency components of fault signal play important role for the fault detection in transmission lines.[9] Figure 1 show how detailed and approximate coefficients are obtained when the fault signal undergoes through the number of filters such as high pass and low pass filters for feature extraction from transient signal to segregate high-frequency and low-frequency components to obtain approximate and detailed coefficients using a digital filtering approach. For each signal x(t), the DWT is defined as in Equation (1)[10]:

$$DWT(x,m,n)=\frac{1}{\sqrt{a_0^m}}\Sigma_{\mathrm{m}}\Sigma_{\mathrm{n}}\,x(k)\psi\left(\frac{k-nb_0a_0^m}{a_0^m}\right), \tag{1}$$

where n, m and k, are integers, $\mathrm{nb_0a_0^m}$ translation (time shift) and $\mathrm{a_0^m}$ are dilation (scale) parameters. b_0 and a_0 are constants.

In this work, the Db4 has been adopted as the mother wavelet for the extraction of features from the fault signal due to its proven suitability for tackling transient data in MRA.[11]

Fault localization helps determine the precise location of the fault. This allows repair teams to be dispatched immediately to the correct location, reducing the recovery time. Less downtime means less disruption for consumers and businesses. Power frequency fault location method, traveling wave fault location method and artificial intelligence based fault location method are most commonly used in the literature. Accurate Fault Location by Impedance-Based Method requires successful extraction of Voltage and current Phasor quantities.

Traveling Wave Fault Location Method is Fast and Precise; however, it requires an Appropriate Measurement that is capable of Detecting the Voltage or Current Transient with very High Sampling Frequencies (e.g., 200 kHz). Artificial Intelligence for fault location accuracy attracts many researchers. Deep Learning is becoming increasingly important for Fault localization in transmission line with time series data. It can be efficiently implemented for fault localization in modern complex power networks.

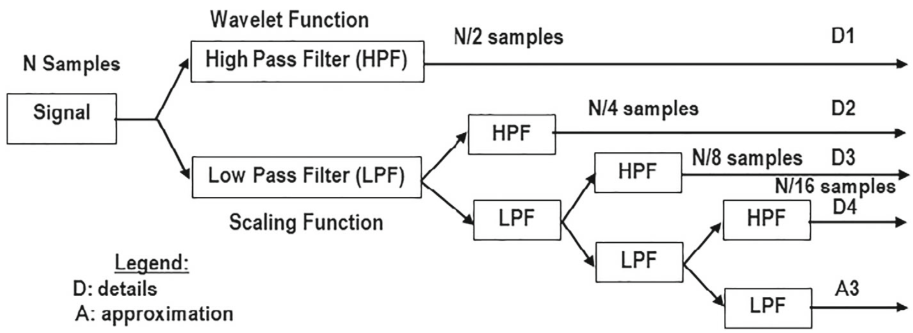

FIGURE 1 Multi-resolution analysis (MRA) in DWT for feature extraction.[10]

Due to the reliable capabilities of artificial neural networks, researchers have implemented them for classifying faults in overhead power lines, because of its reliability in real time and offline applications based on adaptive nature and generalization property.[11] The Back propagation Neural Network (BPNN) is also efficiently used for pattern recognition and fault classification in power transmission network, where the error is controlled through its feedback properties.[12] Moreover, the limitation of BPNN is proper selection of neurons in each hidden layers, higher number of layers and neurons, results in enlarged and slower training process.[13]

The Probabilistic Neural Network (PNN) is major type of ANN, consisting of input, hidden, summation and output layers, possess inbuilt characteristics of pattern recognition.[14] It has been found that the ability of PNN is 10% higher than forward neural network (FNN) while classifying faults in transmission lines. Furthermore, PNN learns fast but is limited to training time and memory requirement as the network size increases, making the selection of number of layers and neurons uncertain.[15]

The Feed Forward Neural Network (FFNN) stands out as the most basic and uncluttered architecture among all ANN models. FFNN employees supervised learning techniques used to modify the weights and biases. Once initialized with pseudo-random values, has input, hidden and an output layer characterized with single or multi-layer perception and back-propagation learning algorithm. Applications of back propagation (BP) with FNN have been found in selection of faulty transmission line with limitation of an output error.[16]

Multilayer Perception (MLP) is one of the variants of feed-forward ANN, having input layer, hidden, and output layer and requires large number of training samples and long training time to achieve high performance. This type of ANN is also sensitive to the initial weights and the learning rate of the network which affect its convergence and accuracy.[17] Radial Basis Function Neural Network (RBFNN) has been used in designing fault models integrated with wavelet features. This extensively used algorithm is expensive in computation with more memory requirement than other neural networks, under high dimension data.[18]

Extreme learning machine (ELM) and Feed-Forward Neural Network belong to same family of ANN, contains only a single hidden layer[19] and uses analytical approach toward network parameters compared to weights-based gradient descent training procedure as used in MLP.[19] This learning machine does not require tuning of its hidden layer and has been found faster with better performance for fault classification in the power network compared to conventional algorithms.[20] Fuzzy logic has gained great applications for fault identification and classification in power transmission networks. Its ability to reason and make decisions is based on "if-then" rules, proves particularly valuable in handling uncertainties inherent in complex systems.[21] In addition, adaptive neuro-fuzzy inference system (ANFIS) has high processing speed, and found more feasible for real time fault identification and classification in transmission lines.[22]

Decision Tree (DT) for data mining practice has been used in past for high dimensional pattern classification which has tree-like graphs, capable of making decisions.[23] The DT algorithm actively identifies inter-grid fault types through determination of the exact time of fault inception and analyzing the odd harmonics present within the measured signal to reveal fault-specific patterns. DT algorithm has been used in FACTs controlled transmission lines for the classification of all shunts faults.[24] Studies have demonstrated that SVM based fault diagnosis method is highly accurate and very encouraging for fault classification in transmission lines.[25] This diagnostic method has ability to classify the hidden data, based on the model parameters learned during training phase.[26] However, it is limited to noisy data, over fitting and computational complexity.[27]

K-nearest neighbors (k-NN) is one of the most basic machine learning algorithms based on the supervised learning, is accessible and uses similarity to categorize data into new.[28] New data are further classified into well-suited class using k-NN algorithm.[29] Furthermore, Genetic algorithm (GA) has the property of offline training required for online applications.[30] This fault classification method does not depend on fault impedance, fault starting point angle and distance from relaying points.[31]

A deep learning based convolution neural network (CNN) has been used extensively for fault classification in modern days, CNN performs relatively complex tasks with images, sounds, texts and videos requiring a huge number of data sets for training.[32] It exhibits better classification performance due to its deep architecture and characteristics of learning high-level features.[20] CNNs leverage their layered structure to progressively unveil hidden patterns and extract highly distinctive features through deeper representation to achieve high efficiency.[21] Moreover, the requirement for large amount of labeled data, computational time and susceptibility to noise are certain limitations of CNN for accurate fault classification in transmission lines.[33] Self-attention (SAT) integrated with the CNN is also been studied. SAT is a learning mechanism that enables the neural network to learn to focus on different segments of time series data that are necessary for fault classifying in transmission lines. The wavelet transform is proposed to improve the noise immunity of SAT-CNN model. The digital filtering technique is used by the DWT to denoise the signals and a vector of hidden activations propagating over time is maintained by recurrent networks.[34]

CNNs found well in recognizing patterns and objects in images with some limitations while dealing with sequential data such as limited receptive field, accuracy and generalization. Time series forecasting witnessed a shift toward recurrent neural networks (RNNs) as fully connected networks, despite their success in CNNs, couldn't handle the sequential nature of time series data, therefore RNN emerged as the preferred deep Learning architecture for this task. Since it is notoriously difficult to train basic RNN designs, long short term memory (LSTMs) and gated recurrent unit (GRU), with their intricate design, emerge as frequent replacements for less sophisticated algorithms.[35] Furthermore, both LSTM and RNN are types of recurrent layers that can be used in conjunction with CNNs to process sequential data within the network architecture.[22] It has been found from the literature that LSTM and RNN can limit the scalability of the network, making training and inference slower, especially for large-scale models and complex data sets. The ability of LSTM and RNN to remember and propagate information over long distances also degrades over time.[33] The sequential nature of RNNs and LSTMs makes it difficult to parallelize computations across time steps.[36] In addition, the simplest RNN networks also suffer from gradient vanishing. As the information moves from one node to another, the gradient decreases.[37] Furthermore, RNNs cannot have temporal dependencies as we increase the number of hidden input windows.[38] When dealing with small data sets, they may not be able to learn meaningful patterns or perform as expected.

To prevent overfitting, LSTMs often need to use regularization techniques such as dropping out or weight decay.[39] LSTMs can also be difficult to interpret because of their black-box nature, which limits their interpretability.[40] LSTMs also require a lot of training data to be able to generalize well, which can be difficult when the data set is small.[41] Additionally, LSTMs have multiple hyperparameters that must be fine-tuned for optimal performance, such as number of layers, memory cell count, learning rate, and so forth. Selecting the appropriate hyperparameters can take a lot of time and require a lot of trial and error.[42]

Recognizing the limitations of recurrent neural networks (RNNs) in handling long-range dependencies and vanishing gradients, temporal convolutional networks (TCNs) have emerged as a compelling alternative for processing time series data. The comparison between various fault classification algorithms with their respective advantages and limitations are mentioned in Table 1.

This research paper consists of five sections. The background and the review of various algorithms are presented in Section 1. Section 2 reports the system model with time series data obtained for different fault types. Section 3, suggests TCN model for fault classification. In Section 4, we look at how the proposed model performs under different conditions. In Section 5, we look at the conclusion of this paper.

2 | SYSTEM MODELING AND DATA PREPARATION

This research considers 500 kV transmission line which is 155 km connecting 500 kV substation at New Karachi from the 500 kV transmission line network via Jamshoro (Sindh, Pakistan) for modeling and simulation. Table 2 shows the specification of the 500 kV transfer line under investigation. The actual positive and zero sequence resistances, inductance and capacitance of the transmission lines of the network are $(0.0185 + 0.231)\ \Omega/\text{km}$, $\left(j1.0767 * 10^{-3} + j1.889 * 10^{-3}\right)$ H/km and $\left(j12.74 * 10^{-9} + j7.751 * 10^{-9}\right)$ F/km, respectively.

MATLAB has been used to model the transmission line so that fault simulation can be performed and time series data can be obtained at 20 kHz sampling frequency. This generated data is used for the detection and classification of different

TABLE 1 Weaknesses and strengths of different techniques used for fault classification in transmission lines.

Technique	Strength	Weakness
ANN technique	ANN exhibits accuracy in diagnosing exact fault type. Easily implemented through adjustment of a few parameters. Its applications are numerous in real life. ANN learns and no need for reprogramming the algorithm.	For high-dimensional data, the training process is time-consuming and complicated. For non-linear separable pattern classification problems, gradient-based back-propagation method is used. ANN provides slow BP convergence because it depends on the initial value of weight constraints associated with the network.
PNN technique	It does not require learning process. Initial weights of the network are not required. PNN is converged with Bayesian classifier to enhance learning time. No connection between learning process and the recalling process.	It requires high processing time for large networks. Difficulty in determining requirement of layers and neurons. Enough memory space requirement to save PNN model.
Fuzzy method	Uncertainty problems are solved using "if-then" relation.	Large training data requires mandatory experts, which help to determine membership function and fuzzy rules. Robustness is not reported.
ANFIS technique	Combined or hybrid learning rules are required for tuning of parameters. Convergence in AFNIS faster. The space size required for searching search is also low. Adaptability of ANFIS is proven.	ANFIS is very complicated to process.
SVM technique	SVM is a very precise method, SVM performs well even for data that is not linearly separable in the base feature space. The chances of misclassifying the data are very low.	For testing and training, it needs more space and speed. It has a high degree of complexity in classification. It needs a large memory.
Deep Learning (DL)	Deep learning (DL) excels in solving problems in many domains. It eliminates feature engineering because it takes up most of the time compared to machine learning training. Its adaptive nature allows it to solve new problems quickly and easily with the help of neural networks such as convolutional, recurrent, and long-term memory.	DL is computationally expensive.
Decision Tree	It is easy to be interpreted and understood. Compatible with other available decision methods, Rules can be easily set.	Calculating is time-consuming when there is a high degree of uncertainty or when there are multiple possible results. Overfitting is an issue. Information gain is biased against features which have more levels.
ELM technique	Contains only one layer that is optimize hidden layer Hidden layer does not require tuning. Adjustment of Weight and bias values is not needed.	Local minima issue Easy overfitting Difficult to find the optimal solution.
Wide-area fault location technique	This technique is similarly adopted for monitoring and control operation purposes.	Deployment of PMU in power network is tedious job.
GSM	It is cost effective, rigid, robust and high-speed communication solution covers longer geographical area and distance. Real time monitoring system is actualized	Bandwidth limitation & security

TABLE 1 (Continued)

Technique	Strength	Weakness
k-NN	k-NN has a high speed of calculation, is easy to comprehend, can be used for regression and classification, and has high accuracy.	In case of large data accuracy issue of accuracy matters, require high memory therefore computationally expensive.
Hidden Markov Model (HMMs)	The flexibility of the Markov model comes from the fact that the Markov model is based on the idea of memory-free property, that is, the change from state to state is dependent only on the current state	Expensive, both in terms of memory and computing time. As compared to Markov models, HMM requires extensive training on a set of seed sequences and also needs a large seed.
CNN	In a deep learning algorithm, inputs are mapped to pre-trained data sets to produce an exact result. Requires long short-term memory. No need for labeling of data. It possesses adaptive nature characteristics to new problems. Reduced dimensionality	Limited capture of long-range dependencies. Fixed-length input Computational cost
RNN	Suitable for sequential data Simpler architecture	Vanishing gradient problem Limited memory
LSTM	Addresses vanishing gradient problem Better suited for complex tasks	More complex architecture Overfitting, especially with limited data
TCN	Process information sequentially Address the vanishing and exploding gradient problems Flexible Receptive Field Handling Sequences of Varying Lengths Superior performance compared to other models	

TABLE 2 Specifications of 500 kV transmission line.

System parameters	Values
Circuit name	500 kV transmission line
Line length	155 (km)
Configuration of transformer	Delta star(D11-Yn)
Loading capacity/circuit	1700 MW
Thermal loading	2750 MW

TABLE 3 Fault nomenclature and parameters used for simulation.

System parameters	Types of values
Types of faults	AG, BG, CG, ABG, BCG, ACG, AB, BC, AC, ABC, ABCG.
Fault distance (km)	15, 30, 45, 60, 75, 90,105,115,125,135,145,160,170
Fault resistance	2,5,7,8,11,13,16,17,19,20,22,24,28,30,32,33,35,38, 42
Ground resistance	0,1,2,6,9,11,13,15,17,19,21,23,25,27,29,31,33,35,37
Fault inception angle	0,15,25,35,45,55, 65, 75, 85, 95, 115, 125, 135, 145, 215, 225, 265
Positive and zero sequence resistance (Ω/km)	Positive sequence resistance R1 = 0.0185 (Ω/km) Zero sequence resistance R0 = 0.231 (Ω/km)
Positive and zero sequence inductance (H/km)	Positive sequence inductance L1 = 1.0767(mH/km) Zero Sequence Inductance L0 = 1.889 (mH/km)
Positive and zero sequence capacitance (F/km)	Positive sequence capacitance C1 = 12.74 (nF/km) Zero sequence inductance C0 = 7.75 (nF/km)
Phase to phase voltage	500 kV
Base voltage	25 kV
Base power	100 MVA
System frequency	50 Hz
X/R ratio of an alternator	7

types of faults in transmission line, including the data for no-fault conditions under various system parameters as shown in Table 3. Figure 2 illustrates the full process of feature extraction from the fault signal using DWT and TCN training through time series data for classification of transmission line faults.

2.1 | Data preparation

To train the TCN model, time series data is obtained from the features extracted from the fault signal, this data is actually raw in its form which requires proper segregation and smoothing necessary for fault diagnosis in transmission to avoid misclassification of faults due to the presence of noise in the raw data. This is a challenging task due to the lack of regularity. In this investigation, the impact of variations in system setting on the transient current waveforms has been examined. The specific values considered are listed in Table 3. These waveforms demonstrate that the key characteristics of the fault

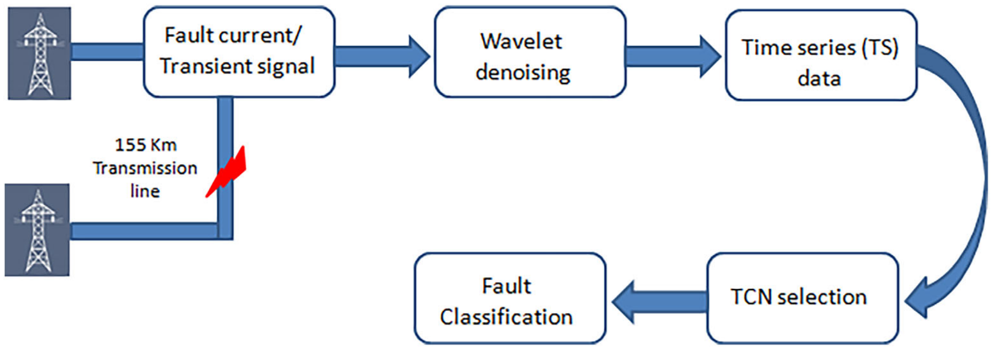

FIGURE 2 Fault detection and classification DWT-TCN.

signal can still be extracted precisely, even when the system settings vary. This paves the way for reliable fault identification. However, the accuracy of fault classification decreases when fault resistance is very high and fault inception angle is very low. This is due to the fact that these factors influence the fault signal itself as well as the amount of energy in the transient part of the fault.

3 | TEMPORAL CONVOLUTIONAL NETWORKS

TCN (temporal convolutional network) is a neural network that is specially designed for processing time-series data. Compared to LSTM and RNN architectures, TCN offers several advantages in terms of efficiency, parallelism, use of 1D convolution, capability to capture long-term dependencies, dilated convolutions, global and local context modeling, interpretability, shift-invariance, implementation, simplicity and scalability. TCN can be used to classify faults based on univariate or multivariate time series data depending on the objective of signal analysis. Further, TCNs use causal convolutions, ensuring that the past does not depend on the future and can handle input and output sequences of any length, similar to RNNs. Moreover, TCN can achieve very large effective history sizes by combining very deep networks. TCNs are different from other deep learning techniques due to its simple structure, making predictions based on previous data points, and having a long memory compared to WaveNet, for handling sequential data.

The TCNs avoid using gates in LSTM and have high receptive field ability with many benefits over LSTM and RNN models.[43] Furthermore, TCN based models are not used in the previous study for fault analysis in transmission lines. The detailed architecture of the temporal convolution neural network (TCN) is shown in Figure 3A–C.

3.1 | Sequence modeling

Sequence modeling is the task that utilizes any type of sequential data covering auto-regressive prediction where the output is the same as the input shifted by one time step. In this research work, the whole input sequence (including “future” states) is used to forecast each output. In order to understand network structure for sequence modeling, an input sequence $(X_0, \dots, X_T)$ is considered to produce related outputs $(Y_0, \dots, Y_T)$ at each time step. The main challenge is to produce the output Y_T for time “t,” the input is considered same as seen before $(X_0, \dots, X_T)$.[46]

Formally, the mapping of sequence model network is done through model function $f : X^{T+1} \rightarrow Y^{T+1}$ whereas function "f" maps a vector of $(T+1)$ for input element "X" to output vector "Y" elements as shown in Equation (2).[46]

$$L\left(y_0, y_1, y_2, \dots, y_{(T-1)}, y_T\right), f\left(x_0, x_1, x_2, \dots, x_{(T-1)}, x_T\right). \tag{2}$$

If model function $'f'$ meets the causal condition that y_T only depends on $x_0, \dots, x_t$ and not on any “future” inputs $x_{t+1}, \dots, x_t$. The objective of learning the sequence modeling setting is to determine network ability that reduces expected loss between the real outputs and prediction. In case the output time series is not matched with length of input time series without using padding, then the output length can be calculated by using Equation (3)[47]:

$$Tout = Tin - (k-1), \tag{3}$$

where k is the kernel size.

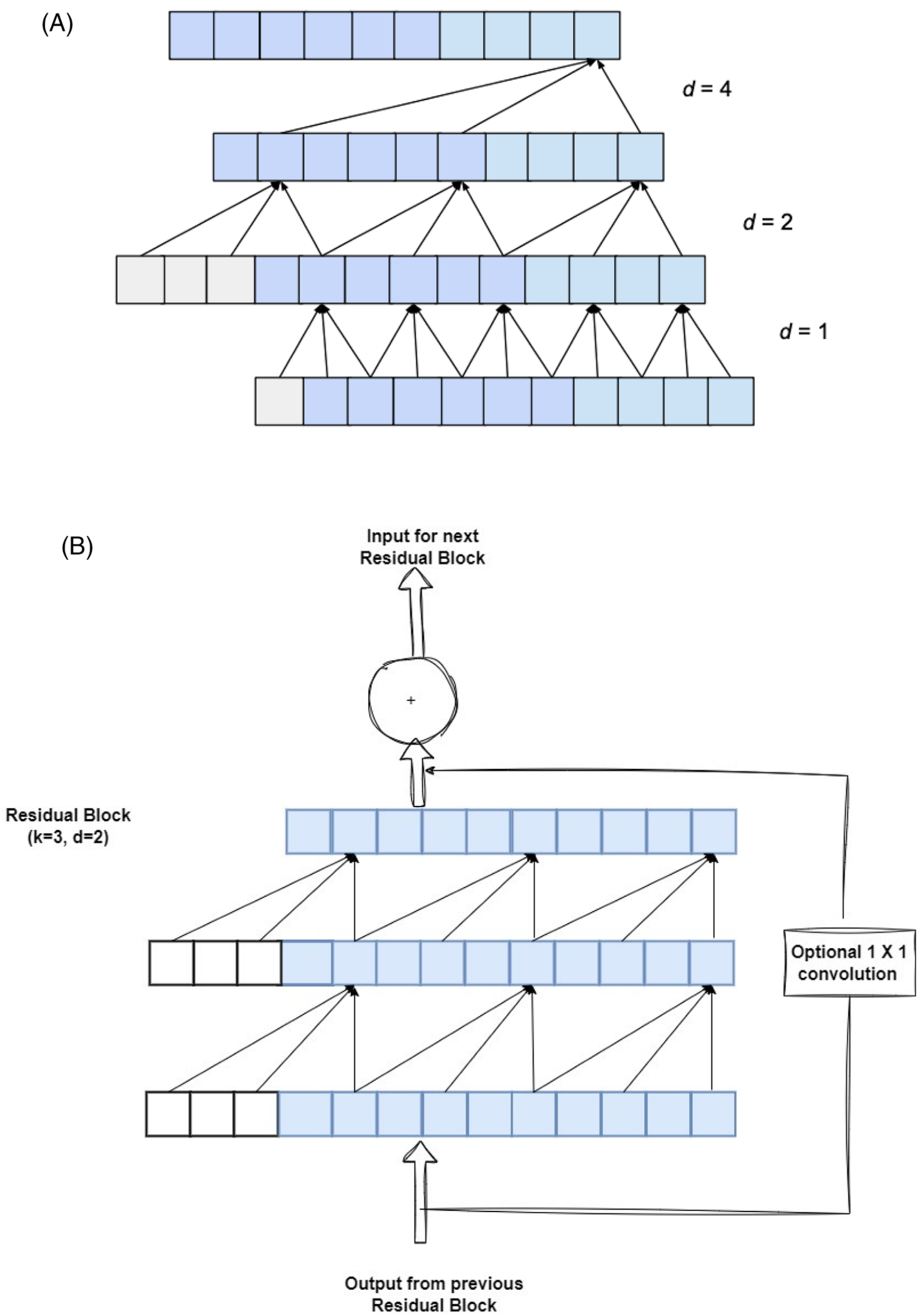

FIGURE 3 (A–C): Architecture of TCN.[44] (A) A dilated causal convolution with factors $d = 1, 2, 4$ and kernel size $k = 3$. (B) TCN Residual Connection. (C) TCN residual block.[45]

Further, if the length of both output and input time series data are matched then zero padding on the left-hand side of the input time series is added for achieving causal. The padding size is dependent on kernel size (k) and dilation factor (d) as given in Equation 4.[48] Since, Causal padding introduces an output time series having same length as input shown in Equation 5.[10]

$$\text{Padding} = (k - 1) \times d \tag{4}$$

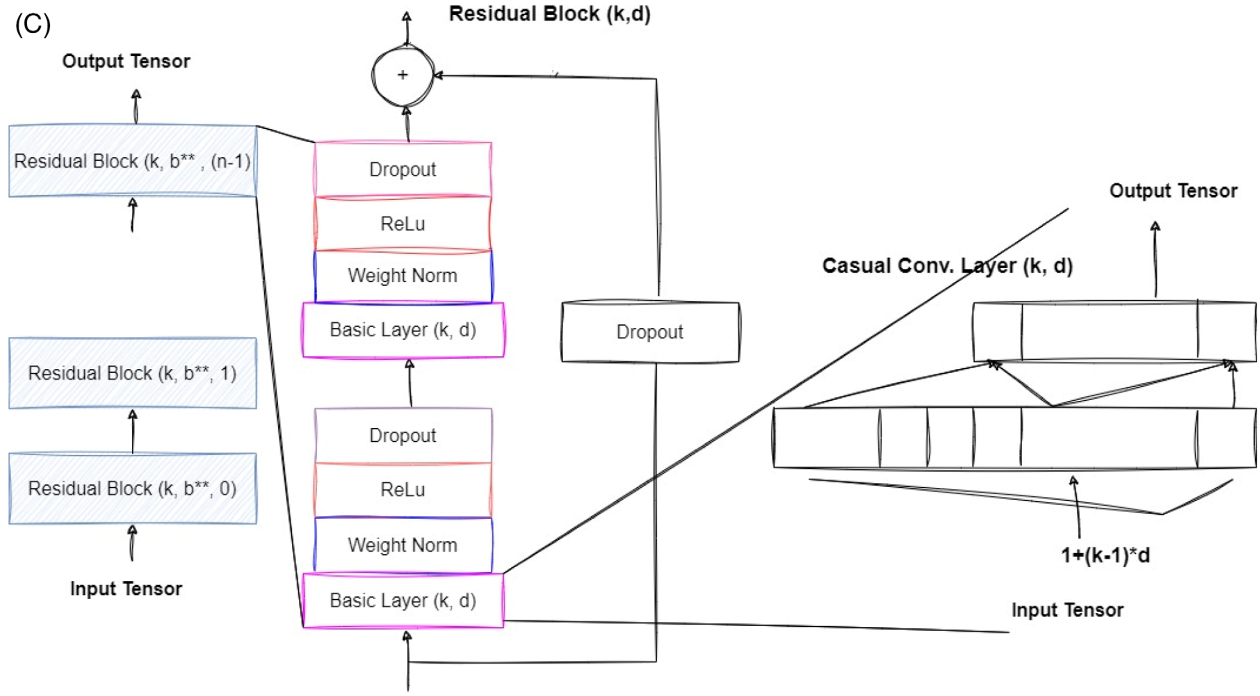

FIGURE 3 (Continued)

$$Tout = Tin + (k - 1) \times d. \tag{5}$$

Hence Tout = Tin.

To increase the receptive filed, TCN work in a very similar way with one additional factor "dilation" (d) and is calculated using Equation (6):

$$\text{Receptive Field Size (RFS)} = (k - 1) \times d + 1. \tag{6}$$

3.2 | Causal convolutions

The TCN uses 1D fully-convolutional network to achieve the causality as first step goal where length of hidden layer (each) = length of input layer with zero padding (Kernel size-1) is applied to keep this length consistency between layers. In order to complete second step goal, TCN uses causal convolutions where output at time "t" depends on elements from time "t" and before (previous layer) only. Equation (7) describes the major layers of TCN.

$$TCN = 1D\ FCN + \text{Causal Convolutions}. \tag{7}$$

3.3 | Dilated convolutions

The regular causal convolutions can only learn from a short history resulting in issue for sequence tasks that require large receptive field. However, dilated convolutions used in TCN based models can learn from an exponentially larger history, making them ideal for longer receptive field. The dilation of sequential data in TCN is defined in Equation (8):

$$xt - (k - 1)d, \ \ldots \ , xt - 2d, xt - d, xt. \tag{8}$$

In above equation "d" is the dilation factor and "k" is the filter size, which is shown in Figure 3A. Therefore, dilation is similar to adding a constant gap between every two neighboring filter elements. To cover broader range of inputs, higher value of “d” allows increased receptive field of a ConvNet. The receptive field of TCN is improved by choosing larger filter sizes“k”and high dilation factor “d.” As with extended networks “d” is increased exponentially with network depth expressed as $D = O\,(2i)$ at network level, where “i” is ith level of network. This ensures that any filter hits all effective history inputs and allows for very large receptive fields in deep networks as shown in Figure 3A.

3.4 | Residual connections

The residual block consists of a branch that leads to a series of transformations whose outputs are added to the input x of the block as shown in Equation (9):

$$\mathrm{O} = \mathrm{Activation}(\mathbf{x} + \mathrm{F}(\mathbf{x})), \tag{9}$$

where “O” is the sum of input blocks and “F” is series of block transformations on the input. This allows layers to learn identity map modifications instead of the entire transformation. The residual connection is essential to enable TCN to learn a long history for higher receptive field. The residual block of TCN as shown in Figure 3B, is used between its each layer to speed up convergence allowing the training of model at deeper level. The residual block consists of extended causal convolution, ReLU activation function, weight normalization and dropout layer to adjust the convolution filters.[44]

An extra 1×1 convolution is applied in the model to make the input and output widths compatible, so that the tensors can be added element-wise further, addition of normalization and dropout layers helps in adjusting TCN during training and testing of data set as shown in Figure 3B,C. The receptive field size (RFS) is increased as per input data size and can be calculated using Equation (10):

$$RFS = 1 + \left(2^{L} - 1\right)(k - 1)\, X\, d, \tag{10}$$

where “L” stands for number of residual blocks stacked on the top of each other, “k” is kernel size and “d” dilation factor.

4 | RESULTS AND DISCUSSIONS

This section presents all simulation results. In this study 500 kV transmission line model has been developed in MATLAB and time series data is obtained for various faults at different values of parameters such as fault inception angle, fault resistance, fault location and ground resistance. The deep learning has become increasingly important for fault localization in transmission lines using time series data which is one of the challenges in analyzing through traditional methods such as recurrent neural network (RNN) and LSTM. For accurate fault localization, TCN excel at learning intricate patterns and relationships within time series data. Furthermore, TCN is capable of automatically learning relevant features directly from raw time series data, leading to accurate fault detection and localization tasks where extracting informative features from the data is essential. TCN efficiently processes large volumes of time series data, analyzing extensive data sets collected from multiple sensors deployed across transmission networks. Moreover, continuously analyzing incoming time series data, TCN can quickly identify and localize faults, resulting in minimizing downtime and prevention from potential disruptions in power transmission network.

For protection of transmission line, the protective relays are located at both ends of a transmission line, is near the power system control bus (PSC), substations, and interconnecting points dividing into different protection zones. As the fault occurs within zone, the relay sends a tripping signal to the circuit breaker, isolating the faulty section. To provide complete backup protection across the entire power system, coordinating with overlapping protection zones.[49] There are various types of relays being used in power system of different make such as Siemens (Siprotec 5), ABB Relion, General Electric Company (GE Micom) and Schweitzer Engineering Laboratories (SEL 411L) and are deployed in 500, 220 and 132 kV transmission line network for the protection purpose in Pakistan. The Siprotec 5 relay is widely used for 500 kV transmission line protection due to their specific features as highlighted in Table 4.[50]

In this work, the case study of 500 kV transmission line from Jamshoro to New Karachi, Sindh, Pakistan has been taken for fault analysis, collection of time series data, and testing on the data set after proper training of the TCN. Table 5 illustrates the impact of fault location, fault inception angle, fault resistance, and ground resistance for Double Line to Ground (DLG) on the fault location. The data are presented as a percentage error and fault location results for specific

TABLE 4 Performance comparison between relays of SEL make, GE make, ERL make and ABB.[50]

Feature	SEL 411	GE Micom	Siemens (Siprotec)	ABB relion
Accuracy and speed	SEL relays are known for their high accuracy, using advanced algorithms and precise measurement techniques. They respond quickly to fault conditions due to their fast-processing capabilities.	GE make relays are generally accurate, reliable and respond swiftly to faults.	These relays are having high accuracy for selected protection functions. They can detect current and voltage signals up to the 50th harmonic with precis. They offer fast tripping times to ensure effective protection.	ABB relays are accurate, dependable and respond promptly.
Flexibility	SEL is well-regarded for its robust and reliable protection. They are known for their advanced functionality, flexibility, and ease of use. SEL relays often incorporate advanced communication capabilities and cybersecurity features, which are increasingly important in modern power systems.	These relays are known for their performance and adaptability to different system configurations. GE's relays often incorporate advanced features such as digital fault recording and disturbance analysis.	These relays are known for their accuracy and performance in demanding environments.	These relays are known for their high accuracy, reliability, and versatility.
Selectivity and sensitivity	SEL relays are known for their sensitivity. They use advanced algorithms and precise settings to detect faults accurately, even in complex network scenarios.	These are designed to coordinate effectively, ensuring that faults are isolated without unnecessary disconnection of healthy parts of the network.	Siemens relays combine selectivity and sensitivity. Their protection devices are designed to achieve both fault detection accuracy and coordinated operation during faults.	ABB emphasizes the importance of selectivity in low and medium voltage networks.
Cybersecurity	SEL relays are designed with strong focus on cybersecurity and designs secure systems for various power system elements including transmission.	GE relays utilize secure communication protocols. These protocols employ encryption and authentication mechanisms to prevent unauthorized access and tampering of data.	Siemens offers a wide range of network and automation components with integrated security functions.	ABB recognizes the importance of cyber resilience and understands how systems work or recover in case of cyber-attacks.

fault locations. Fault location along transmission line is considered at every 15 km along the length of transmission line. The absolute error is calculated using Equation (11) to evaluate the performance:

$$\text{Error}\ (\%) = \frac{\text{Actual Fault Distance} - \text{Calculated Fault Distance}}{\text{Total Section Length}} \times 100. \tag{11}$$

In order to detect faults in 500 kV transmission lines, Discrete Wavelet Transform (DWT) is used for feature extraction from transient signal and TCN is employed for fault classification. The extracted features are collected in the time series for training validation and testing of TCN. Hyperparameters in deep learning are the parameters that are defined prior to training a model and act like switches that control the learning process, learning rate, batch size, number of epochs, activation function and selection of right optimizer are few hyper parameters considered in deep learning. So, there must

TABLE 5 Fault location on every 15 km in 155 km long 500 kV transmission line between Jamshoro-New Karachi Sindh, Pakistan for double line to ground (DLG) fault.

Actual distance (km)	Calculated fault distance (km) TCN.	Fault resistance (Ω)	Fault inception angle (°)	Ground resistance (Ω)	Error (%)
15	14.929	0.5	0	1	0.047
30	30.096	1.5	15	2	0.32
45	44.893	2.2	25	2.5	0.237
60	60.097	3.0	35	4.0	0.161
75	74.991	4.5	45	6.5	0.012
90	90.243	5.0	55	0.01	0.27
105	104.897	2.0	65	0.5	0.098
120	119.892	10.5	75	10	0.09
135	134.981	3.0	85	12	0.014
150	150.310	0.01	95	13.5	0.206
155	154.543	0.6	115	15	0.29

TABLE 6 Hyperparameters considered for TCN model.

Framework	Temporal convolutional network
Temporal window size	0.1 s
Learning rate	0.001 s
Optimizer	Adam
Batch size	64
Epochs	30
No. of features/channels	4 (Ia, Ib, Ic, Ig)
Training samples (temporal)	2304
Testing samples (temporal)	768
Training testing split ratio	3:1

be an optimal selection and consideration of these parameters for achieving high performance of deep learning model. The hyperparameters considered in this study are detailed in Table 6.

In deep learning, learning rate is a crucial hyperparameter that determines the step size of algorithm while updating the weights of a neural network during the training process. Learning rate controls how the model adjusts its parameters in response to the error it makes during validation process. In this study the learning rate of TCN model is considered as 0.001 s. Adaptive moment estimation (Adam) has been employed as an optimizer in this work due to its fast and efficient convergence on effective solutions for time series data of 100,000 length of temporal length. It has been found that using this optimizer, the problem of overfitting during training is reduced beside the hyperparameter settings compared to other optimization algorithms is also minimized. The batch size consists of 64 subset of the data is considered for training of TCN processed during single iteration, a larger batch size can lead to smoother optimization but might require more computational resources.

Moreover, the training data set is divided into smaller batches instead of feeding the entire training data set at once. During each iteration, the TCN model has taken input data performed forward propagation for predicting the fault type along with calculation of loss (error) to further compare between the faults classified by TCN and actual data fed to the algorithm (truth labels). Furthermore, the number of epochs is 30 and each epoch take 100,000 samples of time series data is considered in this work. More epochs generally allow for better learning, but there is a point of diminishing returns where overfitting might occur.

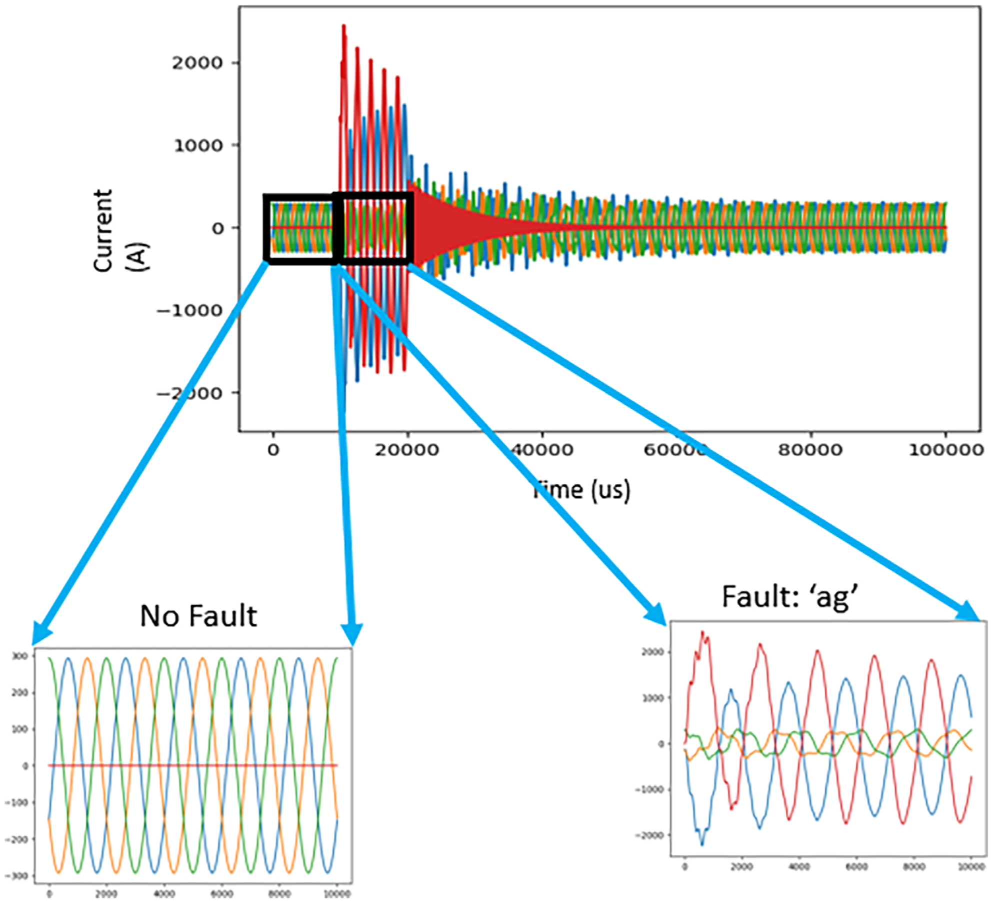

FIGURE 4 State of current signal in time domain.

For the protection of 500 kV Transmission lines, the digital relays are used, their operation is based similar to analog relays, requiring fundamental frequency components (current and voltage). However, when a fault occurs, high frequency components are present in the current signal which travels farther as traveling wave. In order to detect the fault, the information contained in these traveling waves during fault window is extracted which is 0.1 s as given in Table 6, for obtaining the temporal data for TCN training required for fault classification in 500 kV transmission line. Furthermore, the state of the current signal such as pre-fault during fault and after fault in time domain to analyze the behavior of high-frequency content in the transient signal which is necessary for fault detection and classification as shown in Figure 4.

Figure 5 shows the complete process regarding data acquisition, preparation of data to separate data from irrelevant or redundant information, removing noise and fluctuations that can lead to misinterpretations and modeling of TCN to learn and identify patterns associated with different fault types.

The TCN architecture used in this research is shown in Figure 6. Four inputs as fault current, that is, the three phases including ground current (I_a I_b, I_c and I_g) used for the training of TCN. In this work, the 100,000 data points are considered in each sample for phase A, B, C and ground with temporal window size of 0.1 s, the window size has significant impact on the performance of the TCN for accurate fault classification in 500 kV transmission line.

The flow of input data through various layers of temporal convolution network including residual, dense and activation blocks as shown in Figure 7. In this work, the number of samples for each phase and ground current obtained from the fault signal are processed through various layers of TCN, particularly the residual block to distinguish between the normal and faulty states of the transmission line and further to classify the type of fault. Furthermore, the residual block also contains the dilated causal convolution to learn long-range dependencies in the input data, while the filter size (k) controls the width of the convolution kernels. The detailed size of each TCN layer used specific to this work is mentioned in Table 7.

Based upon the input time series data, the model has classified various faults in 500 kV transmission line, therefore it is important to validate the results through a confusion matrix as an assessment tool that has been used in deep learning

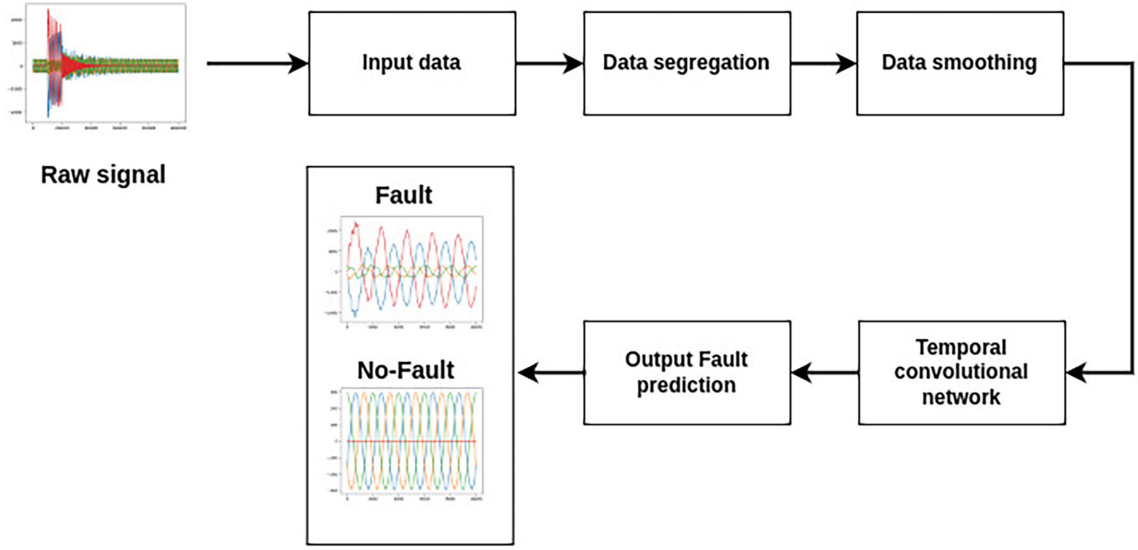

FIGURE 5 Data flow for TCN training.

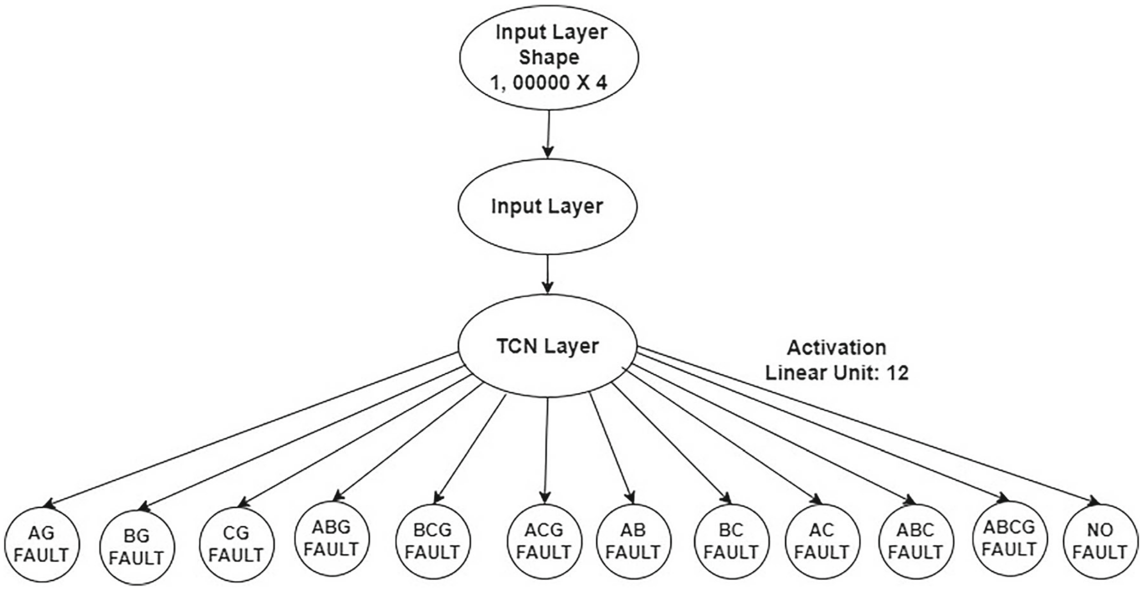

FIGURE 6 TCN architecture for classification of 12 different faults.

to determine the accuracy for each fault type. It is essential to understand the confusion matrix in order to evaluate the performance of the TCN model for fault classification in transmission line. The confusion matrix gives a detailed description of the model's predictions against the ground truth label. Also shows how many true positives, true negatives, false positives, and false negatives are correctly classified as faults, false positives, false negatives, and so on. Knowing the distribution of the error types helps to determine where the model can improve. It further allows optimizing model parameters such as threshold, regularization techniques or class weights. The trade-offs between performance metrics can be analyzed and the model parameters can be adjusted accordingly based on the specific needs of transmission line monitoring system requirements.

Additionally, the confusion matrix is one of the useful tools to assess how well machine learning models classify transmission line faults. They can be used to calculate the accuracy, precision, recall, and F1 score for each type of fault. The accuracy of a model is the percentage of faults that it correctly classifies, the proportion of all instances that were correctly predicted, regardless of their class. In this research, a batch of size 64, has been used, as seen from the matrix,

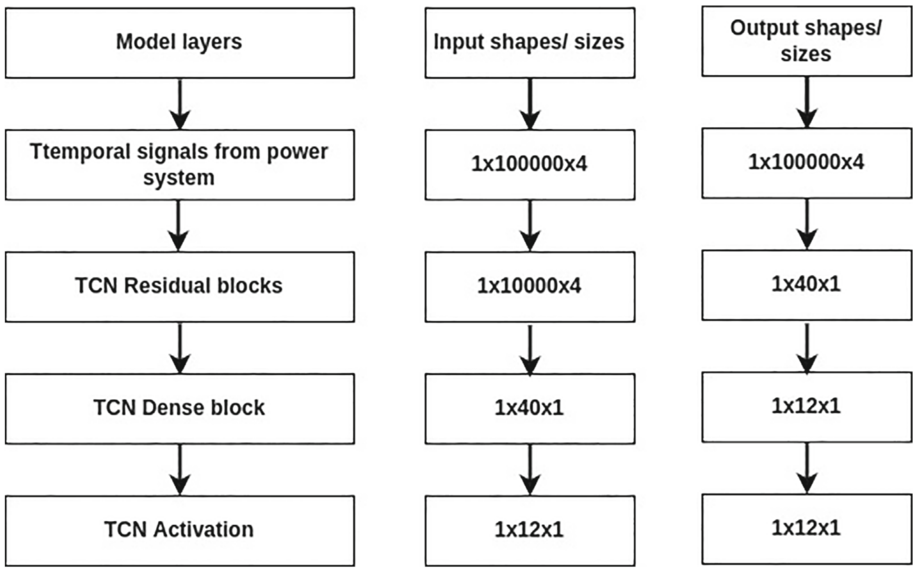

FIGURE 7 TCN blocks for fault classification 500 kV transmission lines.

TABLE 7 Size details of TCN layers.

S #	TCN layer	Size	Description
1	Temporal signals from power system	$1 \times 100000 \times 4$	1 represents sample of data, 100,000 lengths of sequence in sample, 4 indicates the number of features Ia, Ib, Ic and Ig.
2	TCN residual block	$1 \times 10000 \times 4$	1 describes batch size refers to number of data samples processed together, 10,000 is length of temporal sequence arranged in time series. 4 indicates no of variables/features in each data point.
3	TCN dense block	$1 \times 40 \times 1$	1 denotes batch size/input data consists of single sample. 40 is sequence length each sample contains. 1 describes no of variables in each data point.
4	TCN activation block	$1 \times 12 \times 1$	1 denotes batch size/input data consists of single sample. 12 is sequence length each sample contains. 1 describes no of variables in each data point.

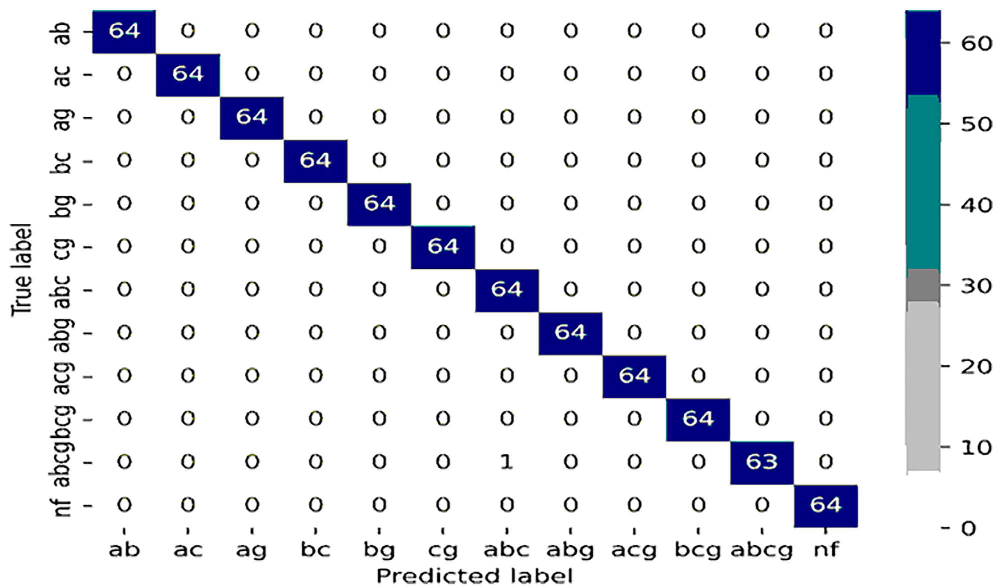

FIGURE 8 Confusion matrix of TCN.

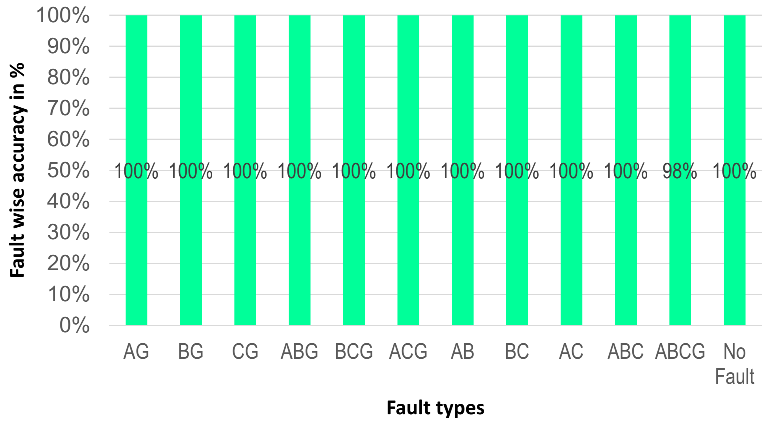

FIGURE 9 Accuracy of TCN model.

FIGURE 10 Loss and accuracy curves of TCN model.

that there is no variation between the true and the predicted labels for SLG, LL, DLG and LLL faults, which are rightly predicted. However, calculated result of ABCG fault has been obtained with an error of 1.57%. Whereas other fault types including no fault are classified with 100% accuracy. From results, it has been found that the obtained results confirm the robustness of the TCN model for analyzing the raw time series data for the fault classification in transmission lines. Figure 8 illustrates a confusion matrix, which shows the relationship between the true and the predicted labels for TCN training. The confusion matrix supports the robustness of the TCN model for fault classification in 500 kV transmission line based on time series data. The confusion matrix also indicates that the improvement of the TCN model for ABCG fault which is misclassified can be achieved through increasing the batch size.

Figure 9, revealed the performance evaluation of model through confusion matrix and loss & accuracy curves and confirms the performance and accuracy of TCN as fault classifier.

The importance of loss and accuracy curves in fault classification in transmission lines lies in their ability to provide valuable insights into model performance and the performance of classification algorithms. Loss curves, which typically represent the variation of a loss function (e.g., the loss function for a particular classification task) over time, are used to evaluate how well a model learns from training data. Whereas, accuracy curves used to track the trend of a model's classification accuracy over time. These curves (loss and accuracy) can be used to determine whether a model is overfitting, underfitting, or converging. Loss and accuracy curves can be used to prevent overfitting, saving computational resources and used to monitor the performance of a model during training and stop the training when the performance of the model on the validation set begins to deteriorate. The TCN model's loss and accuracy can be seen in Figure 10. The curve shows that the loss diminishes and the accuracy increases over time as TCN model matures with training data.

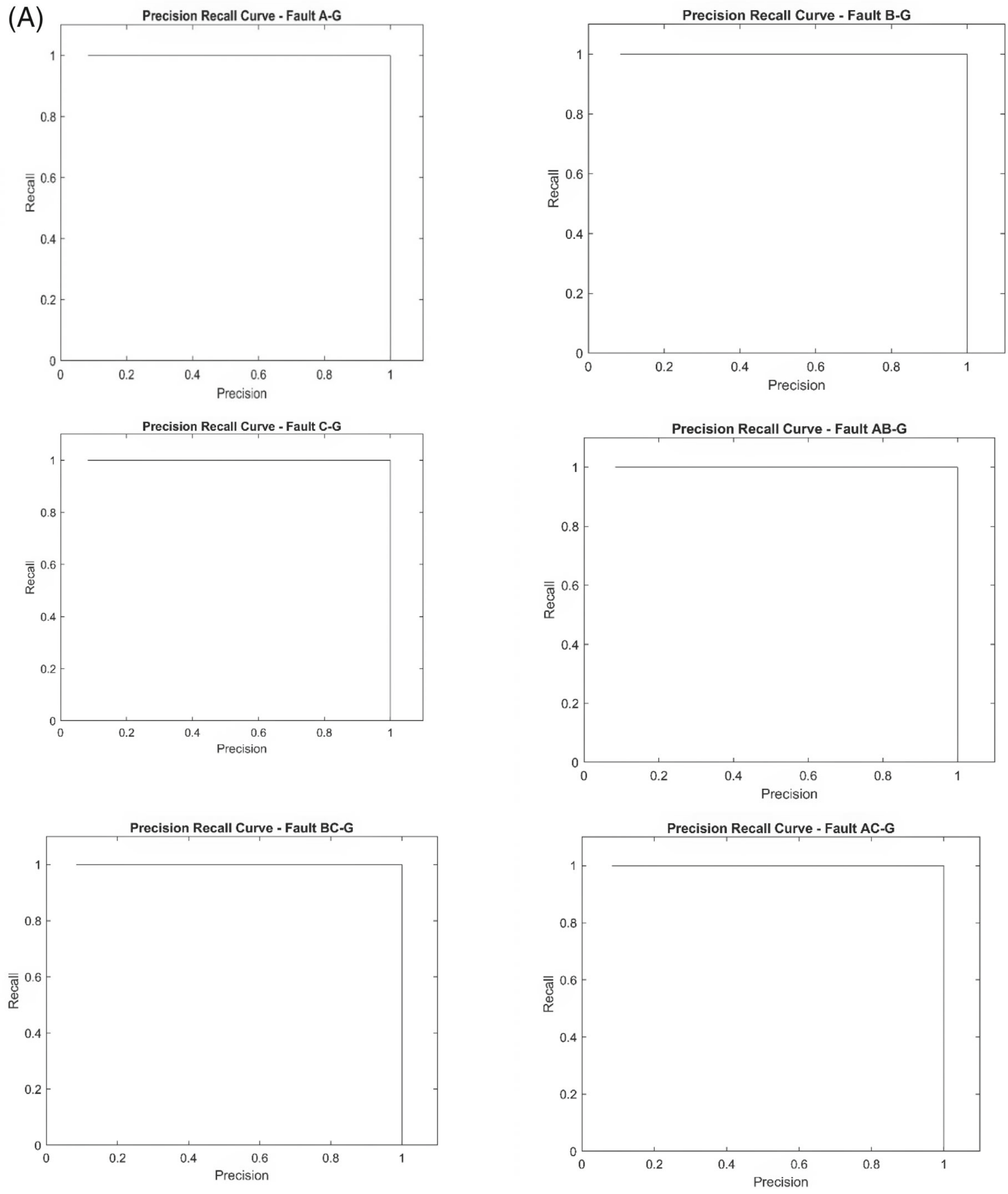

FIGURE 11 (A) Precision recall curves for different faults A-G, B-G, C-G, AB-G, BC-G, and AC-G occurred in 500 kV transmission line. (B) Precision recall curves for different faults A-B, B-C, A-C, A-B-C, A-B-C-G and no fault occurred in 500 kV transmission line.

The precision and recall curves for different faults in 500 kV transmission line are shown in Figure 11A,B. Precision-recall curves are used to assess the performance of a TCN model for fault classification, under imbalanced data or where the importance of correctly identifying positive samples is key consideration. Since, precision measures the accuracy of positive predictions made by the model, while recall measures the ability of the TCN model to find all the relevant cases in the data set. The confusion matrix as shown in Figure 8 gives a detailed description of the TCN model's predictions against the ground truth label. Showing how many true positives, true negatives, false positives, and false negatives are faults are correctly classified as. In this work, the precision and recall curves are calculated using Equation (12) and Equation (13) for all types of faults with 100 percent accuracy except ABCG. Since, there are no false positives and or false

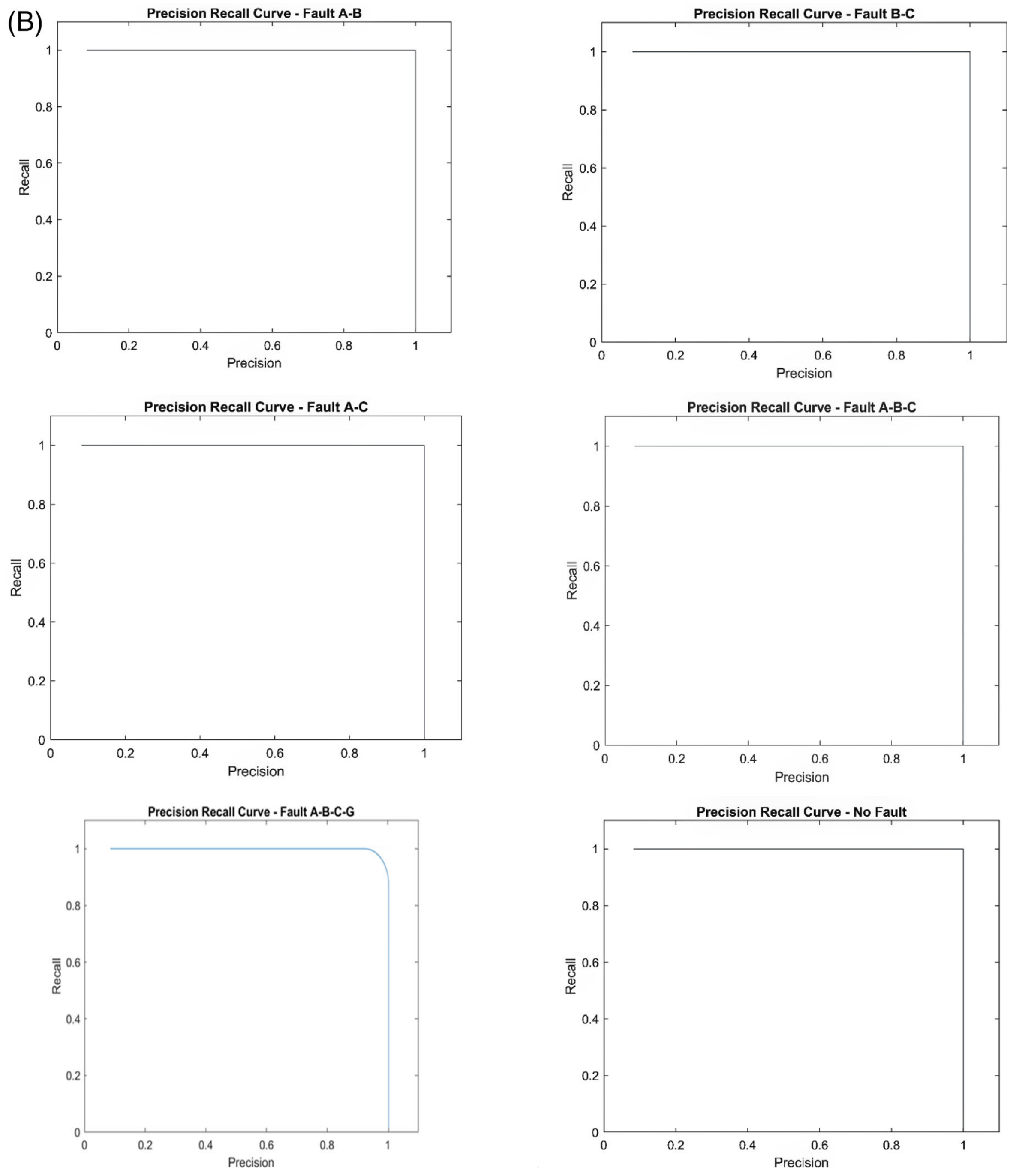

FIGURE 11 (Continued)

negatives whereas, for ABCG class, the precision is 1 and recall is 0984375 due to 1 false negative. Since, the accuracy of the TCN is driven by the overwhelming majority of data points in the most common fault class under unbalance data, and it has been confirmed that TCN performs well on the training set and categorize faults accurately except three phase to ground fault (ABC-G).

The summary of fault wise precision-recall curve of TCN model is shown in Figure 12. The F1 score of the TCN model of 500 kV transmission line is used to evaluate performance under imbalanced classification scenarios, considering both false positives and false negatives, showing the performance of model. Since, F1 score is threshold independent, this factor is useful in comparison with different models. The F1 score of all faults including ABCG which is misclassified is detailed and evaluated using Equations (12–14):

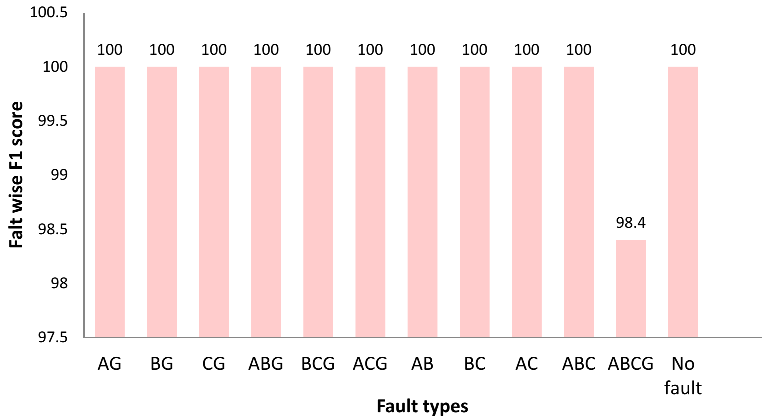

FIGURE 12 Fault-wise performance of TCN model in terms of precision-recall curves.

TABLE 8 Fault-wise performance metrics and system configuration of TCN model of 500 kV transmission line.

Fault type	Precision	Recall	F1 score	Training time (Hours)	System requirements for training and testing
AG	1	1	1	5:41	16 GB RAM, with 2.3 GHz 8-Core Intel Core i9 processor
BG	1	1	1		
CG	1	1	1		
ABG	1	1	1		
BCG	1	1	1		
ACG	1	1	1		
AB	1	1	1		
BC	1	1	1		
AC	1	1	1		
ABC	1	1	1		
ABCG	1	0.9843	0.9921		
No fault	1	1	1		

For ABCG class:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}. \tag{12}$$

$$\text{Precision} = \frac{63}{63+0} = 1.$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Psitive} + \text{False Negative}}. \tag{13}$$

$$\text{Recall} = \frac{63}{63+1} = 0.984375.$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{14}$$

TABLE 9 Performance metrics of TCN, BiLSTM and GRU for classification in transmission lines.

Algorithm	BiLSTM	GRU	TCN
Accuracy	0.9231	0.9527	0.996
Loss	0.0809	0.2272	0.0040
Precision	0.9838	0.9532	1
Recall	0.98307	0.9427	0.9986
F1 Score	0.98131	0.9424	0.999342

TABLE 10 Requirements for implementation of TCN model in the existing grid system.[51]

Sr. no.	System capability	Hardware/software requirement
1.	Sensor device	1. Current transformers (CTs) and voltage transformers (VTs) along the transmission lines to capture electrical signals. 2. Data Acquisition Systems will be required to collect and digitize analog signals from sensors. 3. Process the acquired data to extract relevant features and prepare it for training the TCN model. Preprocessing steps may include noise removal, signal normalization, and feature extraction techniques.
2.	Fault data collection	1. Including data from normal operating conditions of transmission lines to provide contrast for fault detection.
3.	TCN model development and training	1. Software framework: such as TensorFlow or PyTorch to implement and train the TCN model. 2. Hardware accelerators such as GPUs or TPUs to expedite the training process, especially for large-scale data sets.
4.	Model deployment and inference	1. Deployment of trained TCN models on edge computing devices installed at substations or along the transmission lines. 2. These devices should have sufficient computational resources to perform real-time inference and integration with SCADA
5.	Maintenance and monitoring	1. Implementation of remote monitoring capabilities to track the health and performance of deployed TCN models and cybersecurity measures to protect the integrity and confidentiality of data collected and processed by the fault classification system.

For all other classes, precision, recall and F1 score will be 1 since there are no false positives or false negatives.

$$F1 \text{ Score} = 2 \times \left[\frac{1 \times 0.984375}{1 + 0.984375}\right]$$

≈0.9921 is an overall F1 score of ABCG fault.

The performance metrics of TCN model mentioning the training time and system configuration is shown in Table 8:

The performance metrics of TCN model are validated through bidirectional long short-term memory (BiLSTM) and Gated Recurrent Units (GRU) which are trained and tested on similar data set and same hypermeters used for TCN. The comparative analysis of these models is presented in Table 9.

From the results outcome it is found that TCN outperforms compared to BiLSTM and GRU for analysis time series data and confirms through results that TCN model is robust in fault classification due to its long receptive field and dilated convolution. The proposed study can further be implemented for enhancing the performance of transmission lines. Furthermore, the requirements for implementation of TCN model in the existing grid system is presented in Table 10.

It is suggested to the power industry that implementing this algorithm, the performance of the existing 500 kV transmission lines may be improved by accurately classifying various faults within minimum time and this approach can also be considered for improving the design of protection devices and the reliability of the network. In addition, hybrid blockchain and machine learning models such as [52-57], could be essential for securely classifying faults in the smart grid.

5 | CONCLUSION

Researchers are working on improving fault classification in transmission lines to quickly and effectively detect and classify faults and minimizing interruption time. In this work, 500 kV transmission line section of 155 km length between Jamshoro-New Karachi, Sindh, Pakistan is used to simulate under various faults through different parameters such as fault inception angle, fault resistance, fault location and ground resistance. The discrete wavelet transform (DWT) has been employed as a signal-processing technique to analyze the behavior of the phase and ground current for feature extraction. The time series data obtained from the DWT is further used to train the Temporal Convolutional Network (TCN) for fault classification due to its long-term dependencies, high computation efficiency with lower memory requirements. The simulated model has classified faults in 500 kV transmission line with an accuracy of 99.9%. Moreover, the obtained results have also been validated through precision-recall curves, precision score, recall and F1 scores and compared with BiLSTM and Gated Recurrent Unit (GRU), confirming the performance of TCN model is found improved.

AUTHOR CONTRIBUTIONS

Nadeem Ahmed Tunio: Conceptualization; investigation; funding acquisition; writing – original draft; writing – review and editing; visualization; software. **Ashfaque Ahmed Hashmani:** Supervision; resources; project administration; writing – review and editing. **Suhail Khokhar:** Methodology; software; data curation. **Mohsin Ali Tunio:** Writing – original draft; formal analysis; visualization; supervision. **Muhammad Faheem:** Software; data curation; data validation, review and editing.

ACKNOWLEDGMENTS

The authors would like to thank to their affiliated universities for supporting this study.

CONFLICT OF INTEREST STATEMENT

The authors declare no known competing interests' financial interests or personal relationships that could have appeared to influence the work reported in this study.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/eng2.12950.

DATA AVAILABILITY STATEMENT

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

ORCID

Nadeem Ahmed Tunio https://orcid.org/0009-0004-4124-3890
Muhammad Faheem https://orcid.org/0000-0003-4628-4486

REFERENCES

1. Rafique F, Fu L, Mai R. End to end machine learning for fault detection and classification in power transmission lines. *Electr Power Syst Res*. 2021;199:107430.
2. Hessine MB, Jouini H, Chebbi S. Fault detection and classification approaches in transmission lines using artificial neural networks. *MELECON 2014–2014 17th IEEE Mediterranean Electrotechnical Conference*. IEEE; 2014.
3. Mishra M, Panigrahi RR. Taxonomy of high impedance fault detection algorithm. *Measurement*. 2019;148:106955.
4. Arranz R, Muñoz A. Detection of the natural frequency of transmission power lines applying S-transform on transient recovery voltage. *Electr Power Syst Res*. 2020;182:106142.
5. Raza A, Benrabah A, Alquthami T, Akmal M. A review of fault diagnosing methods in power transmission systems. *Appl Sci*. 2020;10(4):1312.
6. Aleem SA, Shahid N, Naqvi IH. Methodologies in power systems fault detection and diagnosis. *Energy Syst*. 2015;6(1):85-108.
7. Adly AR, Aleem SHEA, Algabalawy MA, Jurado F, Ali ZM. A novel protection scheme for multi-terminal transmission lines based on wavelet transform. *Electr Power Syst Res*. 2020;183:106286.
8. Choudhury M, Ganguly A. Transmission line fault classification using discrete wavelet transform. *2015 International Conference on Energy, Power and Environment: towards Sustainable Growth (ICEPE)*. IEEE; 2015.

9. Chanda D, Kishore N, Sinha AJEPSR. Application of wavelet multiresolution analysis for identification and classification of faults on transmission lines. *Electr Power Syst Res*. 2005;73(3):323-333.
10. Ahmed N, Hashmani AA, Khokhar S, Tunio MA, Faheem M. Fault detection through discrete wavelet transform in overhead power transmission lines. *Energy Sci Eng*. 2023;11(11):4181-4197.
11. Aker EE, Othman ML, Aris I, Wahab NIA, Hizam H, Emmanuel O. Transmission line fault identification and classification with integrated FACTS device using multiresolution analysis and naïve bayes classifier. *Int J Power Electron Drive Syst*. 2020;11(2):907.
12. Abdullah A. Ultrafast transmission line fault detection using a DWT-based ANN. *IEEE Trans Ind Appl*. 2017;54(2):1182-1193.
13. Saha B, Patel B, Bera P. DWT and BPNN based fault detection, classification and estimation of location of HVAC transmission line. *2016 International Conference on Intelligent Control Power and Instrumentation (ICICPI)*. IEEE; 2016.
14. Klomjit J, Ngaopitakkul A. Comparison of artificial intelligence methods for fault classification of the 115-kV hybrid transmission system. *Appl Sci*. 2020;10(11):3967.
15. Li Z, Li Y, Zhao G, Zhang X, Xu W, Han D. A model for obstructive sleep apnea detection using a multi-layer feed-forward neural network based on electrocardiogram, pulse oxygen saturation, and body mass index. *Sleep Breath*. 2021;25:1-8.
16. Mirzaei M, Ab. Kadir MZA, Hizam H, Moazami E. Comparative analysis of probabilistic neural network, radial basis function, and feed-forward neural network for fault classification in power distribution systems. *Electr Power Compon Syst*. 2011;39(16):1858-1871.
17. Fahim SR, Sarker S, das S, Islam MR, Kouzani A, Mahmud MAP. A probabilistic generative model for fault analysis of a transmission line with SFCL. *IEEE Trans Appl Supercond*. 2021;31(8):1-5.
18. Shafiullah M, Abido MA. S-transform based FFNN approach for distribution grids fault detection and classification. *IEEE Access*. 2018;6:8080-8088.
19. Chen Y, Chen S, Zhang N, s, Jing H, Min G. LPR-MLP: a novel health prediction model for transmission lines in grid sensor networks. *Complexity*. 2021;2021:1-10.
20. Akmaz D, Mamiş MS, Arkan M, Tağluk ME. Transmission line fault location using traveling wave frequencies and extreme learning machine. *Electr Power Syst Res*. 2018;155:1-7.
21. Petrović I, Nikolovski S, Glavaš H, Relić F. Power system fault detection automation based on fuzzy logic. *2020 International Conference on Smart Systems and Technologies (SST)*. IEEE; 2020.
22. Faheem M, Kuusniemi H, Eltahawy B, Bhutta MS, Raza B. A lightweight smart contracts framework for blockchain-based secure communication in smart grid applications. *IET Gener Transm Distrib*. 2024;18:625-638.
23. Eboule PSP, Pretorius JHC, Mbuli N. Artificial neural network techniques apply for fault detecting and locating in overhead power transmission line. *2018 Australasian Universities Power Engineering Conference (AUPEC)*. IEEE; 2018.
24. Patel B, Bera P, Saha B. Wavelet packet entropy and RBFNN based fault detection, classification and localization on HVAC transmission line. *Electr Power Compon Syst*. 2018;46(1):15-26.
25. Shakiba FM, Azizi SM, Zhou M, Abusorrah A. Application of machine learning methods in fault detection and classification of power transmission lines: a survey. *Artif Intell Rev*. 2023;56(7):5799-5836.
26. Ekici S. Support vector machines for classification and locating faults on transmission lines. *Appl Soft Comput*. 2012;12(6):1650-1658.
27. Faheem M, Umar M, Butt RA, Raza B, Ngadi A, Gungor VC. Software defined communication framework for smart grid to meet energy demands in smart cities. *2019 7th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)*. IEEE; 2019.
28. Naik S, Koley E. Fault detection and classification scheme using KNN for AC/HVDC transmission lines. *2019 International Conference on Communication and Electronics Systems (ICCES)*. IEEE; 2019.
29. Mamuya YD, Lee YD, Shen JW, Shafiullah M, Kuo CC. Application of machine learning for fault classification and location in a radial distribution grid. *Appl Sci*. 2020;10(14):4965.
30. Moloi K, Yusuff A. 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). 2021.
31. Ray P, Mishra D. Application of extreme learning machine for underground cable fault location. *Int Trans Electr Energy Syst*. 2015;25(12):3227-3247.
32. Eissa M, Awadalla M, Sharaf A. A new high impedance fault detection technique using deep learning neural network. *Artificial Intelligence Applications in Electrical Transmission and Distribution Systems Protection*. CRC Press; 2021:179-195.
33. Mnyanghwalo D, Kundaeli H, Kalinga E, Hamisi N. Deep learning approaches for fault detection and classifications in the electrical secondary distribution network: methods comparison and recurrent neural network accuracy comparison. *Cogent Eng*. 2020;7(1):1857500.
34. Fahim SR, Sarker Y, Sarker SK, Sheikh MRI, das SK. Self attention convolutional neural network with time series imaging based feature extraction for transmission line fault detection and classification. *Electr Power Syst Res*. 2020;187:106437.
35. LeCun Y, Bengio Y, Hinton GJn. Deep learning. *Nature*. 2015;521(7553):436-444.
36. Ghosh PR, Mohanty SK, Mohapatra S. Fault zone identification and fault classification in a series compensated transmission line using decision tree. *Advances in Power Systems and Energy Management: Select Proceedings of ETAEERE 2020*. Springer; 2021.
37. Ramasamy MD, Periasamy K, Periasamy S, et al. A novel adaptive neural network-based Laplacian of gaussian (AnLoG) classification algorithm for detecting diabetic retinopathy with colour retinal fundus images. *Neural Comput & Applic*. 2024;36(7):3513-3524.
38. Kozakijevic S, Salb M, Elsadai A, et al. Seizure detection via time series classification using modified metaheuristic optimized recurrent networks. *Theor Appl Comput Intell*. 2023;1(1):82-94.
39. Raghavendran P, Ragul S, Asokan R, et al. A new method for chest x-ray images categorization using transfer learning and CovidNet_2020 employing convolution neural network. *Soft Comput*. 2023;27(19):14241-14251.

40. Gnanadesigan NS, Dhanasegar N, Devi RM, et al. An integrated network topology and deep learning model for prediction of Alzheimer disease candidate genes. *Soft Comput*. 2023;8:1-15.
41. Kannan E, Avudaiappan M, Kaliyaperumal S, et al. A novel single phase grid connected solar photovoltaic system for state of charge estimation using recurrent neural networks. *Energy Sources A: Recovery Util Environ Eff*. 2023;45(1):841-859.
42. Bennet MA, Mishra OP, Muthusamy S. Modeling of upper limb and prediction of various yoga postures using artificial neural networks. *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*. IEEE; 2023.
43. Wang S, Cao J, Yu PS. Deep learning for Spatio-temporal data mining: a survey. *IEEE Trans Knowl Data Eng*. 2020;34(8):3681-3700.
44. Cossu A, Carta A, Lomonaco V, Bacciu D. Continual learning for recurrent neural networks: an empirical evaluation. *Neural Netw*. 2021;143:607-627.
45. Bai S, Kolter JZ, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. 2018.
46. Gopali S, Abri F, Siami-Namini S, Namin AS. A comparison of TCN and LSTM models in detecting anomalies in time series data. *2021 IEEE International Conference on Big Data (Big Data)*. IEEE; 2021.
47. Gan Z, Li C, Zhou J, Tang G. Temporal convolutional networks interval prediction model for wind speed forecasting. *Electr Power Syst Res*. 2021;191:106865.
48. Benhaddi M, Ouarzazi J. Multivariate time series forecasting with dilated residual convolutional neural networks for urban air quality prediction. *Arab J Sci Eng*. 2021;46:3423-3442.
49. Sahoo B, Samantaray SR. System integrity protection scheme for enhancing backup protection of transmission lines. *IEEE Syst J*. 2020;15(3):4578-4588.
50. Gunasekera S. *Design and Development of Protection, Control and SCADA of Model Zone Substation Laboratory Based on the IEC61850 Communication Standard*. Victoria University; 2017.
51. Falih M, Fadhil A, Shakir M, Atiyah BT. Exploring the potential of deep learning in smart grid: addressing power load prediction and system fault diagnosis challenges. *AIP Conference Proceedings*. AIP Publishing; 2024.
52. Faheem M, Mahmoud Ahmad A-K. Multilayer cyberattacks identification and classification using machine learning in internet of blockchain (IoBC)-based energy networks. *Data in Brief* 2024;54:110461.
53. Faheem M, et al. Cyberattack patterns in blockchain-based communication networks for distributed renewable energy systems: a study on big datasets. *Data in Brief* 2024:110212.
54. Faheem M, Raza B, Bhutta MS, Madni SHH. A blockchain-based resilient and secure framework for events monitoring and control in distributed renewable energy systems. *IET Blockchain* 2024;1-15:110212.
55. Abubakar M, et al. High-precision identification of power quality disturbances based on discrete orthogonal s-transforms and compressed neural network methods. *IEEE Access* 2024;11(7):1-18.
56. Malik H, et al. Blockchain and internet of things in smart cities and drug supply management: open issues, opportunities, and future directions. *Internet of things* 2023:100860.
57. Zafar A, et al. Machine learning autoencoder-based parameters prediction forsolar power generation systems in smart grid. *IET SmartGrid*. 2024;7(3): 328-350.

How to cite this article: Tunio NA, Hashmani AA, Khokhar S, Tunio MA, Faheem M. Fault detection and classification in overhead transmission lines through comprehensive feature extraction using temporal convolution neural network. *Engineering Reports*. 2024;6(12):e12950. doi: 10.1002/eng2.12950