IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 13 June 2023, accepted 7 July 2023, date of publication 10 July 2023, date of current version 13 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3294093

RESEARCH ARTICLE

Fault Detection for Medium Voltage Switchgear Using a Deep Learning Hybrid 1D-CNN-LSTM Model

YASEEN AHMED MOHAMMED ALSUMAIDAEE[1,2], JOHNNY KOH SIAW PAW[3], CHONG TAK YAW[3], SIEH KIONG TIONG[3], (Senior Member, IEEE), CHAI PHING CHEN[4], TALAL YUSAF[5,6], FOO BENEDICT[7], KUMARAN KADIRGAMA[8,9,10], TAN CHUNG HONG[3], AND AHMED N. ABDALLA[11]

[1]College of Graduate Studies (COGS), Universiti Tenaga Nasional (The Energy University), Jalan Ikram-Uniten, Kajang, Selangor 43000, Malaysia
[2]Department of Computer Science, Imam A'adhum University College, Baghdad 10001, Iraq
[3]Institute of Sustainable Energy, Universiti Tenaga Nasional (The Energy University), Jalan Ikram Uniten, Kajang, Selangor 43000, Malaysia
[4] Department Electrical and Electronics Engineering, Universiti Tenaga Nasional (The Energy University), Jalan Ikram-Uniten, Kajang, Selangor 43000, Malaysia
[5] School of Engineering and Technology, Central Queensland University, Brisbane, QLD 4009, Australia
[6]College of Engineering, Almaaqal University, Basra 61003, Iraq
[7]Enhance Track Sdn. Bhd., Puchong 47120, Malaysia
[8]Advance Nano Coolant-Lubricant (ANCL), College of Engineering, Universiti Malaysia Pahang, Pekan 26600, Malaysia
[9]Faculty of Mechanical and Automotive Engineering Technology, Universiti Malaysia Pahang, Gambang 26300, Malaysia
[10]Centre for Research in Advanced Fluid and Processes, Universiti Malaysia Pahang, Pekan 26600, Malaysia
[11]Faculty of Electronic Information Engineering, Huaiyin Institute of Technology, Huai'an 223003, China

Corresponding authors: Johnny Koh Siaw Paw (johnnykoh@uniten.edu.my) and Chong Tak Yaw (chongty@uniten.edu.my)

This work was supported by Tenaga Nasional Berhad (TNB) and Universiti Tenaga Nasional (UNITEN) through the BOLD Refresh Publication Fund under Grant J510050002-IC-6 BOLDREFRESH2025-Centre of Excellence, and in part by the AAIBE Chair of Renewable Energy (ChRE) for providing all out-laboratory support under Grant 202101KETTHA.

ABSTRACT Medium voltage (MV) switchgear is a vital part of modern power systems, responsible for regulating the flow of electrical power and ensuring the safety of equipment and personnel. However, switchgear can experience various types of faults that can compromise its reliability and safety. Common faults in switchgear include arcing, tracking, corona, normal cases, and mechanical faults. Accurate detection of these faults is essential for maintaining the safety of MV switchgear. In this paper, we propose a novel approach for fault detection using a hybrid model (1D-CNN-LSTM) in both the time domain (TD) and frequency domain (FD). The proposed approach involves gathering a dataset of switchgear operation data and pre-processing it to prepare it for training. The hybrid model is then trained on this dataset, and its performance is evaluated in the testing phase. The results of the testing phase demonstrate the effectiveness of the hybrid model in detecting faults. The model achieved 100% accuracy in both the time and frequency domains for classifying faults in Switchgear, including arcing, tracking, and mechanical faults. Additionally, the model achieved 98.4% accuracy in detecting corona faults in the TD. The hybrid model proposed in this study provides an effective and efficient approach for fault detection in MV switchgear. By learning spatial and temporal features simultaneously, this model can accurately classify faults in both the TD and FD. This approach has significant potential to improve the safety of MV switchgear as well as other industrial applications.

INDEX TERMS Energy, fault detection, arcing fault, hybrid model, medium voltage switchgear, power system safety, deep learning.

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

I. INTRODUCTION

In order to ensure the safe and reliable distribution of electricity, medium voltage (MV) switchgear is utilized.

However, problems with switchgear can lead to costly downtime and potential dangers. Early fault detection is crucial to avoid these issues [1]. Failures in switchgear can be categorized into mechanical faults and electrical faults [2]. Mechanical faults include wear and tear, corrosion, misalignment, and component failure [3]. These faults can result in contact erosion, overheating, and tripped circuit breakers [4], [5], [6], [7].

Electrical faults, on the other hand, encompass insulation failure, short circuits, and overloads [8], [9], [10]. These failures can lead to arcing [11], [12], [13], corona [14], [15], and Tracking [16], [17]. MV switchgear failures can progress from initial stages of deterioration to more complex stages involving arcing, corona, and tracking. Arcing occurs when an electrical current jumps across a gap between two conductors due to problems like contamination, wear, or contact erosion. It can result in elevated temperatures, pressure, and potential catastrophic failure [18].

Corona, another type of electrical discharge, happens when two conductors with a significant electrical potential difference ionize the surrounding air, causing visible discharge. Corona can produce hazardous ozone and damage insulation [15]. Tracking, the deterioration of insulation surfaces, occurs when moisture or contaminants are present, creating a conductive path for electrical current to flow across the surface.

To address these issues, it is essential to detect and resolve switchgear problems as early as possible through maintenance and testing programs [19]. Various methods and tools such as visual inspection, electrical testing, partial discharge monitoring, condition monitoring, and remote monitoring have been implemented to effectively recognize and monitor faults in switchgear [20], [21], [22], [23], [24]. While mechanical and electrical faults are the most common types, other factors like environmental conditions and human error can also cause failures [25], [26].

In addition, there are studies that provide valuable insights into fault diagnosis methodologies and their application in various domains. These studies encompass different methods, such as Digital twin-driven fault diagnosis, optimal sensor placement, and diagnosing faults in closed-loop control systems [27], [28], [29].

For instance, Yang et al. in this study [27] proposes a fault diagnosis method that combines virtual and real data through digital twins. By utilizing virtual models to simulate physical systems and comparing them with real-time data, this approach enhances the accuracy and effectiveness of fault diagnosis. Furthermore, another study by Kong et al. [28] focuses on optimal sensor placement methodology for fault diagnosis in hydraulic control systems. Its objective is to determine the best sensor locations that facilitate effective fault detection and diagnosis. Additionally, Kong et al. in [29] introduces a fault diagnosis methodology for redundant closed-loop feedback control systems, employing a subsea blowout preventer system as a case study. This methodology combines analytical redundancy, model-based diagnosis, and data-driven techniques to detect and diagnose faults in the control system. The effectiveness of this approach is demonstrated through the case study. These studies significantly contribute to enhancing our understanding of fault detection and diagnosis in the field.

Deep learning techniques, specifically deep learning (DL) and convolutional neural networks (CNNs), have shown promising results in fault detection in various domains, including power systems [30]. DL techniques can accurately classify faults by automatically learning features from data, making them useful for fault diagnosis in switchgear systems [31], [32]. Traditional fault diagnosis methods relying on expert knowledge can be time-consuming and prone to errors [33].

On the other hand DL approaches automate the diagnosis process by learning from large datasets of sensor measurements and maintenance records, leading to more precise and timely fault detection [34], [35]. CNNs have been widely used in switchgear fault diagnosis, as they can analyze both time-series data, such as sensor measurements, and perform image recognition tasks [36], [37]. Additionally, long short-term memory (LSTM) networks excel at identifying long-term dependencies in time-series data, making them particularly suitable for analyzing switchgear measurements [38].

The hybrid 1D-CNN-LSTM model combines the strengths of both CNNs and LSTMs, where the 1D CNN extracts relevant spatial features from the input data, and the LSTM layer captures temporal dependencies [39], [40], [41]. The hybrid 1D-CNN-LSTM model has been successfully applied in various domains, including speech recognition, image and video analysis, power grids, and wind turbines [42], [43], [44], [45]. Its capacity to simultaneously learn spatial and temporal features enhances fault detection accuracy. The LSTM layer identifies temporal dependencies, while the CNN layer extracts relevant spatial features. The objective of this study is to detect faults in medium voltage switchgear quickly using soft computing and ultrasonic inspection systems. The aim is to provide governments and relevant businesses with the necessary tools to act preemptively and reduce the likelihood of incidents and potential losses. The objectives and contributions of the research are summarized in the following points:

i. Ensuring a stable power supply, and traditional fault detection methods can be time-consuming and inaccurate. So, this study presented a hybrid model (1D-CNN-LSTM) that uses DL techniques to efficiently and accurately detect faults in MV switchgear. whereas the effectiveness of the hybrid model has been demonstrated in quickly detecting arcing faults and distinguishing them from other types of defects. In general, the hybrid model is considered the best model for detecting arcing and other faults in both the time and frequency domains. Based on the results that were obtained and compared with other studies.
ii. The hybrid model incorporates the strengths of both the 1D-CNN and LSTM models, making it possible to

identify faults in a manner that is both more accurate and more productive.

iii. This work stands out since it is the first of its kind to examine the detection of faults in both the time and frequency domains in switchgear. This is a novel approach that has not been explored in previous studies that used similar DL techniques. The proposed model's effectiveness in detecting switchgear faults can help prevent equipment damage, power outages, and potential safety hazards.

iv. The use of DL techniques in fault detection has shown promise in various applications, and this study aims to extend its effectiveness to the context of MV switchgear.

II. PROPOSED METHODOLOGY

The purpose of this research is to use DL methods for fault and defect detection in switchgear in the TD and FD. The particular sorts of electrical and mechanical problems were accurately determined by analyzing ultrasonic data recorded during the faults in an audio format. The fault classification system's digital signal processing (DSP) and hybrid model flowchart is shown in Figure 1. The power utility company (PUC) in Peninsular Malaysia gathered raw distribution data from seven states: Kedah, Kuala Lumpur, Melaka, Selangor, Perak, Negeri Sembilan, and Johor. MATLAB was used to read all of the data. For the DL model, the hybrid model was chosen as the most suitable technique. The MATLAB program was used for data collection and pre-processing, whereas the Google Collab environment was used for model programming. The raw data collected from AUT equipment was recorded in audio formats such as mp3, MPEG, or wav. To prepare the data for the DL algorithm, data transformation, a pre-processing technique, was utilized to consolidate the data into a format suitable for the MATLAB software to work with, i.e., a matrix format. The pre-processing steps for the dataset and the method used are explained in the following sections, as shown in Figure 1.

A. DATA COLLECTION

The samples underwent an analysis of sound waves in the switchgear to develop a multidimensional categorization of faults in both the TD and FD. The classification process involved differentiating between mechanical and electrical faults and further analysing the pattern of electrical faults to identify specific types such as corona, arcing, or tracking. To minimize errors in decision-making, an intelligent solution was developed to emulate expert analysis given the possibility of varying interpretations of the data by different engineers. To enable MATLAB R2017a to analyze the ultrasound data, the raw data had to be retrieved using the ''audioread'' and ''audioinfo'' functions, which respectively analyze the WAVE file and obtain information about the audio file. Table 1 presents the extracted sample audio information.

TABLE 1. Fundamental details regarding the sample sound.

Audio File	Background Information
Arc.wav	No of Channels: 1 Sample of Rate: 11025 Total Samples: 68900 Duration: 6.2494 Bits of Per-Sample: 16
Corona.wav	No of Channels: 1 Sample of Rate: 8000 Total Samples: 53991 Duration: 6.7489 Bits of Per-Sample: 8
Tracking.wav	No of Channels: 1 Sample of Rate: 8000 Total Samples: 52000 Duration: 6.5000 Bits of Per-Sample: 8
Good Bearing.wav	No of Channels: 1 Sample of Rate: 11025 Total Samples: 48551 Duration: 4.4037 Bits of Per-Sample: 16
Bad Bearing.wav	No of Channels: 1 Sample of Rate: 11025 Total Samples: 55301 Duration: 5.0160 Bits of Per-Sample: 16

It can be observed from Table 1 that the files in the dataset contain various sampling rates, which include:

i. 44,100 bits per second.
ii. 22,050 bits per second.
iii. 16,000 bits per second.
iv. 11,025 bits per second (Selected as base frequency)

So that the ultrasound sample could be analysed and put into groups in a consistent way, the time-domain ultrasound data were changed to a sampling rate of 11,025 bps. This was a step that had to be done before the analysis could begin, because data with a higher sampling rate would need smaller sample sizes for the data timeline to be in sync. In this case, a multiplier of four was needed to get data with a sampling rate of 44,100 bps. Once the data sampling was complete, the entire time domain data was collected, and the first part of the data was trimmed based on the highlighted area in Figure 2. This was necessary because the initial audio information was deemed useless as it only captured the ultrasound equipment being turned on without being properly positioned. Moreover, the ultrasound pattern repeated itself within the allocated time frame. Therefore, the extracted data started at the number 1000 and had a temporal frame of 10 Kbits, as shown in Figure 3.

All five sets of data, all of which were in the TD, were processed using the same steps:

i. Arching – 54 Sets.

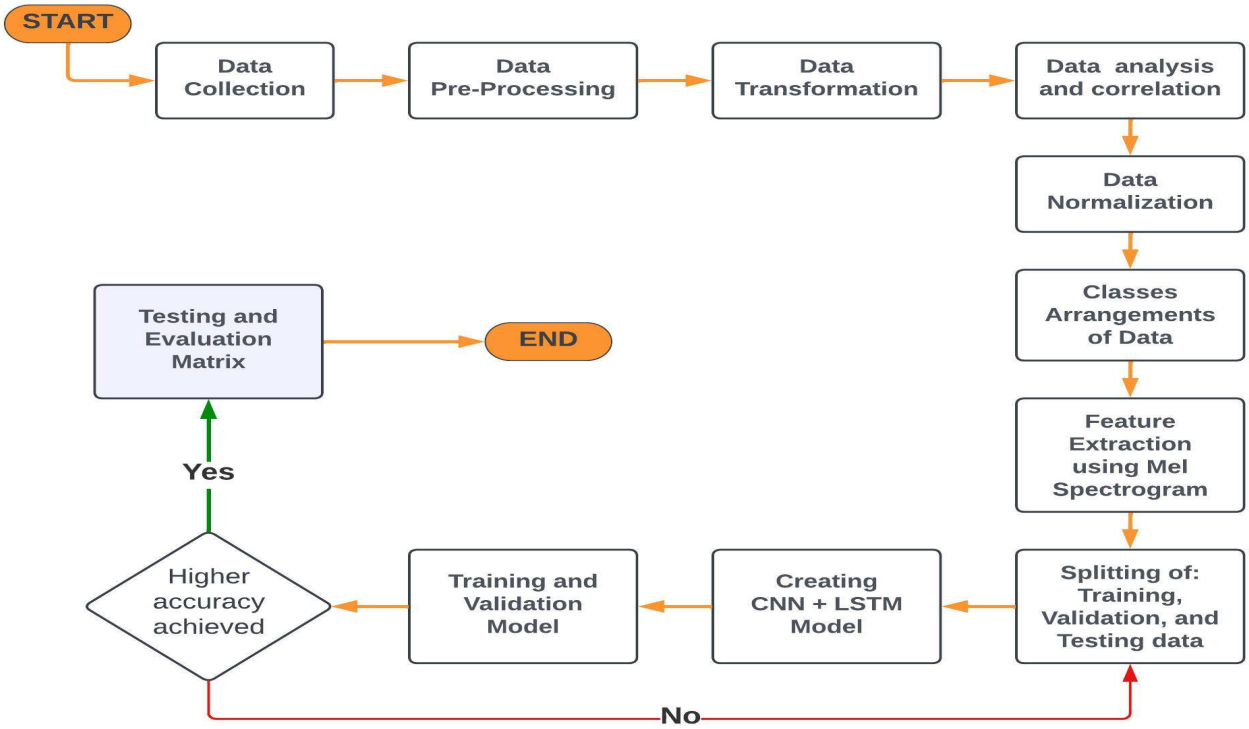

FIGURE 1. Flow-chart of research methodology.

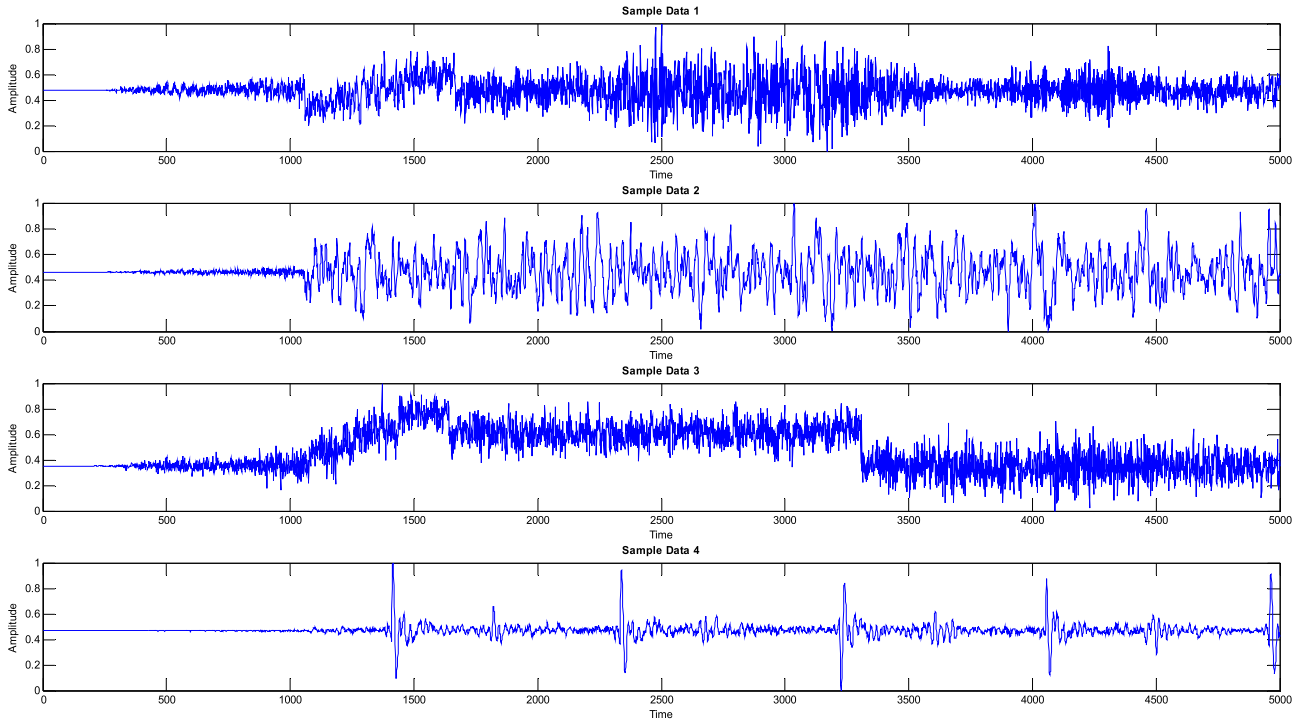

FIGURE 2. Data extracted of time domain.

ii. Corona – 41 Sets.
iii. Mechanical – 17 Sets.
iv. Tracking – 39 Sets (available 314 data sets for single channel wave file).
v. Normal – 13 Sets.

The time-domain data sets were employed to train time-domain models through DL methods. Subsequently, a one-dimensional Fast Fourier Transform (FFT) in

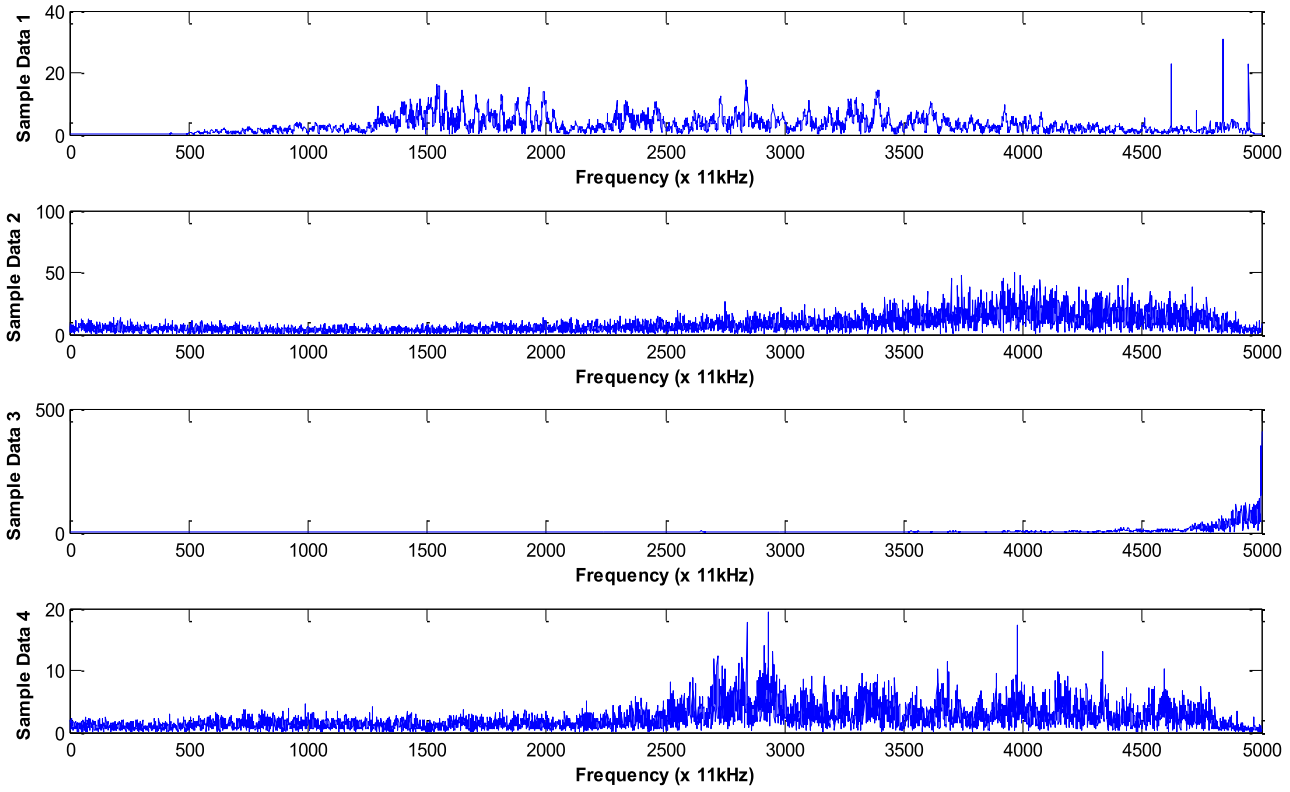

FIGURE 3. Final data extracted of time domain.

MATLAB was applied to transform TD functions into FD representations. The transformed data is illustrated in Figure 4. In the frequency domain, the dataset for the faults after the last process is:

i. Arching – 53 Sets.
ii. Corona – 39 Sets.
iii. Mechanical – 17 Sets.
iv. Tracking – 40 Sets.
v. Normal – 12 Sets.

B. DATA PRE-PROCESSING

This research involved the use of four different types of equipment for diagnosing the health of switchgear using ultrasound. The data collected from the various ultrasonic test devices, including:

i. Ultra_TEV Plus.
ii. Ultra_TEV Plus2.
iii. Ultra_Probe 9,000.
iv. Ultra_Probe 10,000.

This equipment was utilized to diagnose the health of switchgear using ultrasound. During the data collection process, two common audio formats were used, which are MPEG audio layer-3 (mp3) and waveform audio (wav). The use of different audio formats is important because it affects the quality and quantity of data obtained. The choice of audio format depends on the specific equipment used, and in this study, both mp3 and wav formats were used. Table 2 provides a comprehensive list of the equipment used, as well as the corresponding sampling rates for each format. By utilizing different types of equipment and audio formats, the study was able to gather a diverse set of data samples, which can improve the accuracy and reliability of the analysis.

TABLE 2. Units for magnetic properties.

Name of Device	Rate of Sample (Bits/second)	File Format
Ultra_TEV_Plus	11,025 16,000 22,000	mp3
Ultra_TEV_Plus 2	11,025 16,000 22,000 44,100	wav
Ultra_probe 9000	8,000 11,025 16,000 22,000	wav
Ultra_probe 10000	8,000 11,025 16,000 22,000 44,100	wav

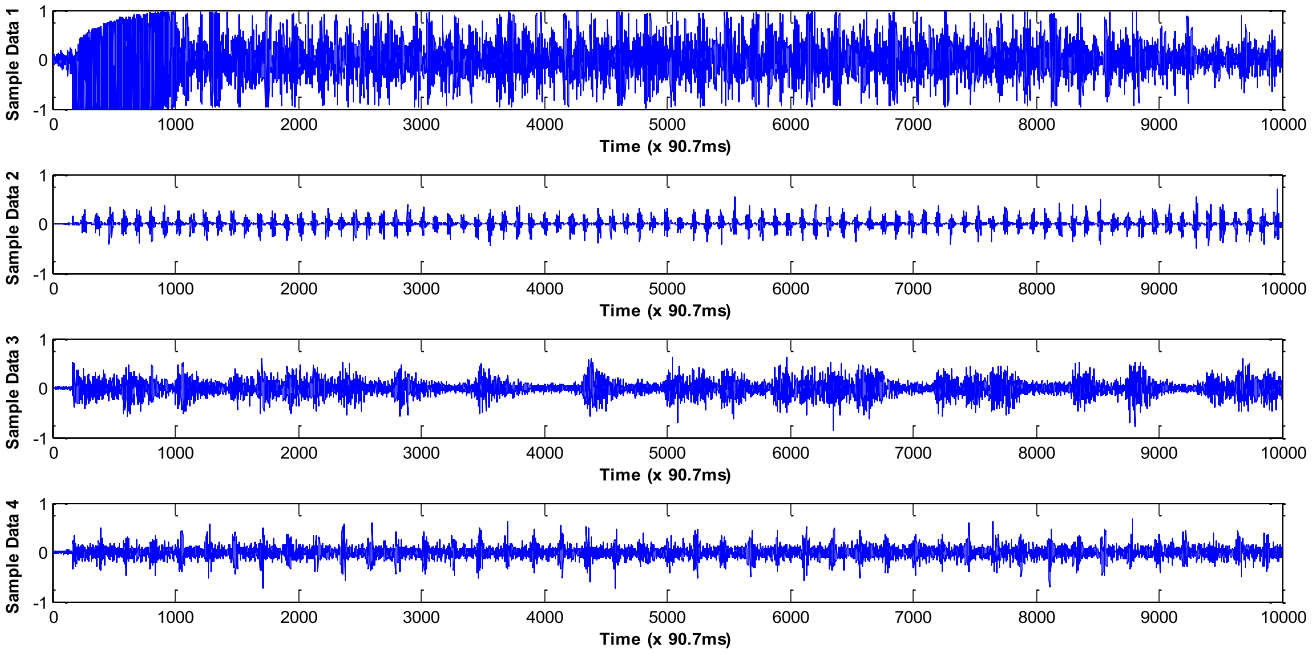

FIGURE 4. Frequency domain data representation after transforms.

To address the importance of the ultrasonic sensors utilized in our testing work, we will now provide a comprehensive discussion on their key characteristics as shown in Table 3. Understanding these characteristics is crucial for comprehending the effectiveness of the method in detecting faults in medium voltage switchgear.

By providing detailed discussions on the characteristics of these ultrasonic sensors, we aim to highlight their specific functionalities and capabilities, whereas combination of high sensitivity, wide frequency range, real-time data acquisition.

C. DATA TRANSFORMATION

The ultrasound data obtained in either the WAV or MP3 file formats has to be converted into a matrix format in order to be compatible with MATLAB programming and other DL methods. The data was able to be transformed in a way that made it suitable for additional analysis, including TD and FD analyses of the data, which were made available in the resulting. Mat file format. The TD data had a high sampling rate of 17.5 megabytes, which meant that a large amount of data was collected over a relatively short period of time. In order to analyses this data, it was necessary to convert it into the FD using a Fourier transform. By allowing the data to be represented in terms of its frequency components, this transformation made it simpler to spot patterns and other features in the data. The FD data had a smaller size of 11.3 megabytes after the Fourier transform was applied to the TD data, which made it easier to handle for additional analysis. In order to be compatible with DL algorithms, the ultrasound data that was gathered in WAV or MP3 format was converted into a matrix format. The data then had a high sampling rate of 17.5 megabytes in the TD and was accessible in both the TD and FD. The data was transformed into the FD using the Fourier transform, resulting in a smaller size of 11.3 megabytes.

D. DATA ANALYSIS AND CORRELATION

Finding the best sampling rate for the data is crucial during the data analysis process. In this study, the ideal sampling rate was established after a thorough analysis of the data that had been gathered. Following data analysis, it was determined that the baseline sampling rate for the ultrasound sample data should be 11,025 bits per second. To ensure that all data sets are compatible with one another, a data normalization procedure is necessary when the data come from different sampling rates. This procedure involves modifying the data to create a connection with other sets of data and to make it possible to fairly compare various data sets. After collecting the original data in its entirety, the next step was to filter out any interference from the environment to obtain a clear representation of the ultrasound sound that accurately reflects the current situation. This was achieved by removing the initial sampling, which consisted of background noise and other irrelevant sounds that were recorded before the ultrasound equipment was properly positioned. The filtering process was necessary to obtain accurate and reliable results from the analysis.

E. DATA NORMALIZATION

Normalization is a method of data transformation that is comparable to scaling, with the exception that it scales ''individual samples to the unit norm.'' In this study, the input data was normalized, which also resulted in the values being scaled between 0 and 1. The main objective of the application

TABLE 3. Detailed of characteristics for the ultrasonic sensors.

Name of Device	The Characteristics
Ultra_TEV_Plus	•Time of Flight (TOF) technology: Enables precise measurement and analysis of transient earth voltage (TEV) signals. •High sensitivity: Allows for the detection of even minor electrical discharges within the switchgear. •Wide frequency range: Covers a broad spectrum of frequencies associated with different fault conditions. •Real-time data acquisition: Provides instantaneous data capture for efficient fault detection. •Multiple component monitoring: Supports simultaneous monitoring of various switchgear components, enhancing the overall diagnostic capabilities.
Ultra_TEV_Plus 2	•Upgraded signal processing algorithms: Improves the accuracy and reliability of TEV signal analysis. •Enhanced noise reduction techniques: Minimizes the impact of external interference on signal detection. •Superior sensitivity: Enables the detection of subtle electrical discharges, enhancing fault identification. •Robust design: Ensures reliable operation in challenging electrical environments. •Compatibility with medium voltage switchgear: Specifically designed for fault detection in this application.
Ultra_probe 9000	•Piezoelectric transducer technology: Converts high-frequency acoustic signals generated by partial discharges into electrical signals for analysis. •Compact size: Allows for flexible sensor placement and easy integration into the switchgear system. •High sensitivity: Enables the detection of low-intensity partial discharges. •Excellent frequency response: Captures a wide range of frequencies associated with different fault types. •Precise localization: Facilitates the identification of the exact location of partial discharge sources within the switchgear.
Ultra_probe 10000	•Advanced signal processing algorithms: Enhances the accuracy and reliability of partial discharge analysis. •Extended frequency range: Enables the detection of a broader spectrum of partial discharge activities. •Robust construction: Provides durability and resistance to environmental factors. •High resistance to electromagnetic interference: Ensures reliable performance in electrically noisy environments. •Versatility: Suitable for various applications requiring precise partial discharge detection.

was to facilitate feature selection and extraction for identifying waveforms based on the TD and frequency domain. For reducing unnecessary noise caused by interference during the setup of the test equipment, the system employed a noise suppression technique by removing the first 1000 bits of data. This initial part of the data is typically useless because it only contains background noise generated when the equipment is turned on and the technician has not positioned it in the appropriate spot. By eliminating this noise, the accuracy of the analysis is improved, and only the relevant data is used for analysis.

F. DATA CLASSIFICATION

We will explain each step-in detail, as shown in Figure 1.

1) USE THE MEL SPECTROGRAM FOR FEATURE EXTRACTION

The Mel spectrogram is a popular feature extraction technique used in speech recognition and audio processing [46]. It is based on the Mel scale, which is a logarithmic scale that is used to convert frequency into a perceptual scale that is more in line with human hearing. The Mel spectrogram shows the power spectral density of the audio signal, but with the frequency axis changed to the Mel scale. In this article on fault detection for medium voltage switchgear, a Mel spectrogram has been used to extract features from the ultrasound data. This can be done by converting the TD data into a Mel spectrogram using a spectrogram function. The spectrogram function splits the audio signal into a number of brief, overlapping frames, and then each frame receives a unique window function. The short-time Fourier transform (STFT) is then used to calculate the power spectral density of each frame. A Mel spectrogram is created by turning each frame's power spectral density into one. Time is plotted on the x-axis and frequency is plotted on the y-axis in the resulting Mel spectrogram, which is a two-dimensional representation of the audio signal. Each point in the Mel spectrogram corresponds to a frequency band's energy at a specific time. The Mel spectrogram can be viewed as a picture of the audio signal, with various features represented as patterns. These features have been used as inputs to DL models, including the hybrid 1D-CNN-LSTM model employed in this study. where the presence or absence of switchgear faults can be used to classify the spectrograms by the model. The model can be trained to recognize patterns and features that are suggestive of faults in the switchgear by exposing it to a large dataset of labeled spectrograms. Using a window size of 25 ms and a hop length of 10 ms, the Mel spectrogram was produced from the pre-processed audio signals. As a result, the audio signals were represented as time-frequency data that could be used as input into the hybrid model. Convolutional and recurrent operations are then used by the hybrid 1D-CNN-LSTM model to extract and learn the pertinent features from the audio signals using the Mel spectrogram as input.

2) SPLITTING DATASET TO: TRAINING, VALIDATION, AND TESTING PHASES

During the testing phase, when predictions are made using data that was not used for model training, the train-valid-test split approach is frequently used to assess the performance of a DL model. The dataset is divided into three subsets, with 70% of the data designated for training, 15% for validation, and the remaining 15% for testing and

performance evaluation. With the help of this technique, it is possible to test the model's generalization capabilities on previously unreported data, which is essential for evaluating the model's predictive power. In order to get the best performance on the test set, the model's hyperparameters, like the learning rate and batch size, are adjusted on the validation set. This method offers a more accurate simulation of the model's performance on new data, making it possible to estimate the model's accuracy more precisely than when the entire dataset is used for training.

3) HYBRID 1D-CNN-LSTM MODEL

The hybrid 1D-CNN-LSTM model is a DL approach that combines both CNNs and LSTM Models for fault detection in MV switchgear. The model takes as input the Mel spectrogram of the sound signal obtained from the switchgear, which is a 2D array of intensities that represent the frequency content of the signal over time. The Mel spectrogram is computed by applying a filterbank of Mel-scale triangular overlapping windows to the signal. The model has two main components: the 1D-CNN and the LSTM. The 1D-CNN is responsible for extracting local features from the Mel spectrogram. It applies a set of learnable filters to the Mel spectrogram to generate a feature map that highlights certain frequency patterns. The filters are learned through backpropagation during training. The output of the 1D-CNN is a 3D tensor that represents the learned local features of the spectrogram. The LSTM component processes the sequence of local features extracted by the 1D-CNN. It takes as input a sequence of feature maps and uses a set of learnable gates to selectively remember or forget certain features based on their importance to the task at hand. The LSTM also has a hidden state that is updated at each time step and serves as a memory of the past inputs. The output of the LSTM is a vector that summarizes the sequence of local features into a global feature representation. The output of the LSTM is then fed into a fully connected layer that maps the global feature representation to the output classes (arcing or non-arcing and the same process for another faults like corona and tracking). The model is trained using the categorical cross-entropy loss function, which measures the difference between the predicted and actual output classes, and backpropagation is used to update the model parameters. The equations for the 1D-CNN and LSTM components of the model are as follows:

$$x_{o,fl}^{l} = f\left(\sum_{im} x_i^{l-1} * k_{io,fl}^{l} + b^{l}\right) \tag{1}$$

$$x_{o}^{l} = f\left[max\left(\sum_{im} x_i^{l-1}\right) + b^{l}\right] \tag{2}$$

$$x_{0}^{l} = f(x_i^{l-1} * d_{io}^{l} + b^{l}) \tag{3}$$

The input to the model is a one-dimensional matrix represented by x, with n elements. An activation function, denoted as f, is applied to the output of each layer of the model. The kernel filter for each layer is denoted as $k_{io,fl}^{l}$ and has a dimension of $k \times 1$. The model has l convolutional layers and F filters. The output of the lth convolutional layer is represented as $x_{o,fl}^{l}$. The bias vectors for each layer are denoted as b, and the learnable parameters are denoted as d. These parameters are used in the computation of the output of each layer of the model.

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{4}$$

where i_t is the input gate activation, σ is the sigmoid function, W_i is the weight matrix for the input gate, h_{t-1} is the previous hidden state, x_t is the current input, and b_i is the bias vector for the input gate.

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{5}$$

where f_t is the forget gate activation, σ is the sigmoid function, W_f is the weight matrix for the forget gate, h_{t-1} is the previous hidden state, x_t is the current input, and b_f is the bias vector for the forget gate.

$$O_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{6}$$

where O_t is the output gate activation, σ is the sigmoid function, W_o is the weight matrix for the output gate, h_{t-1} is the previous hidden state, x_t is the current input, and b_o is the bias vector for the output gate.

$$C_t = \sigma\left(f_t \odot C_{t-1} + i_t \odot \check{C}_t\right) \tag{7}$$

where C_t is the current cell state, f_t is the forget gate activation, C_{t-1} is the previous cell state, i_t is the input gate activation, and $\odot$ is the candidate cell state.

$$\check{C}_t = tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{8}$$

where $\check{C}_t$ is the candidate cell state, $tanh$ is the hyperbolic tangent function, W_c is the weight matrix for the candidate cell state, h_{t-1} is the previous hidden state, x_t is the current input, and b_c is the bias vector for the candidate cell state.

$$h_t = tanh\left(C_t\right) \odot O_t \tag{9}$$

where h_t is the current hidden state, O_t is the output gate activation, C_t is the current cell state, and $tanh$ is the hyperbolic tangent function. In addition, Figure 5 shows the architecture of a 1D-CNN model. The 1D-CNN model takes the input signal and uses a set of convolutional filters to pull out features.

These filters slide over the input signal, performing convolutional operations and producing a set of feature maps. As for Figure 6, it shows the architecture of an LSTM model. The LSTM model uses a set of gates to control the flow of information through the network. The gates selectively pass, or block information based on the current input and the previous state of the network. Whereas Figure 7 shows the architecture of the proposed hybrid 1D-CNN-LSTM model, which combines the feature extraction capabilities of the 1D-CNN model with the sequential modeling capabilities of the LSTM model. The input signal is first processed by the 1D-CNN model to extract a set of high-level features, which are then passed to the LSTM model for sequential modeling. The output of the LSTM model is then passed through a

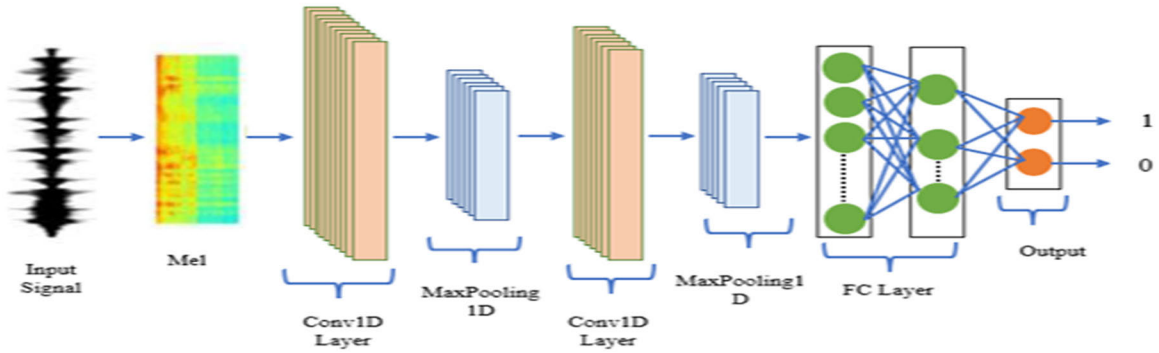

FIGURE 5. The architecture of a 1D-CNN model.

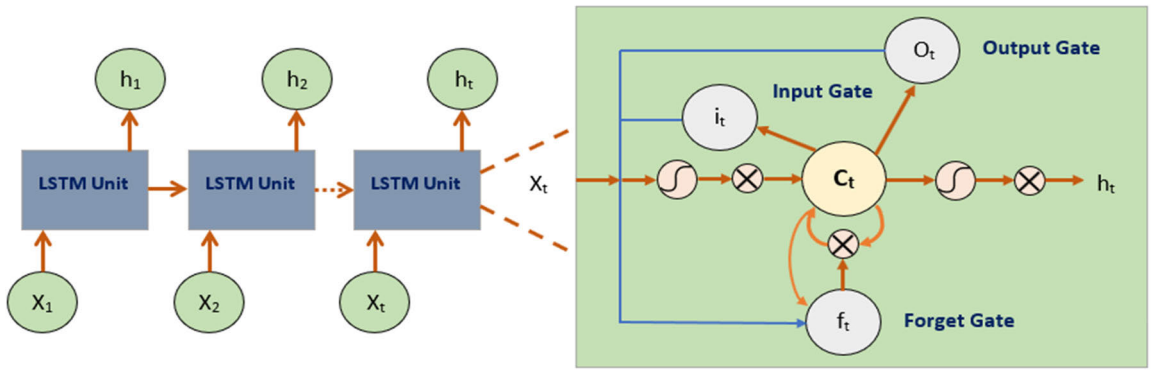

FIGURE 6. The architecture of a LSTM model.

fully connected layer for classification. By combining the strengths of the 1D-CNN and LSTM models, the hybrid 1D-CNN-LSTM model is able to effectively model time series data and extract meaningful features for classification tasks. This model has been shown to achieve high levels of accuracy in various applications, including fault detection in medium voltage switchgear.

4) TRAINING AND VALIDATION DATASET

The training and validation phases are critical steps in the development of a DL model, including the hybrid 1D-CNN-LSTM model for fault detection in medium voltage switchgear. In the training phase, the model is fed a dataset containing input features and corresponding output labels. The model uses this data to learn the relationship between the input features and output labels, adjusting its internal parameters through a process called backpropagation. The training phase typically involves a large portion of the available data; in this case, 70% of the dataset. Once the model has been trained, it is important to evaluate its performance on new data that it has not seen before. This is where the validation phase comes in. In this phase, a portion of the dataset, typically 15% in this case, is held out of the training data and used to assess the model's performance. The validation data is used to tune hyperparameters, such as learning rate, number of layers, and activation functions, to improve the model's accuracy. During the validation phase, the model is tested on the validation dataset, and metrics like accuracy and loss are used to change the model's hyperparameters. This process is repeated until the models. performance on the validation data is optimized. It is important to note that the validation phase is not used to train. the model. Instead, it is used to evaluate the model's performance and improve its hyperparameters. Once the model has been optimized using the validation data, it can be tested on the remaining 15% of the dataset to evaluate its generalization performance.

Finally, the decision-making stage, the result is assessed against specific criteria to determine its acceptance. In order to select the best-performing model, the validation data is used to compare the performance of different models. A model with an accuracy of above 90% is considered successful in the training phase, enabling the program to move on to the testing and prediction steps. However, if the accuracy is below the desired threshold, the entire process of splitting the data, training the model, and evaluating performance is repeated until the desired accuracy is achieved.

Testing plays a crucial role in evaluating the performance of a trained model. In this phase, the model is applied to unseen test data to determine its accuracy and ability to generalize to new data. The predicted values are compared to the actual values in the test data to evaluate the performance of the model. The confusion matrix and classification report are

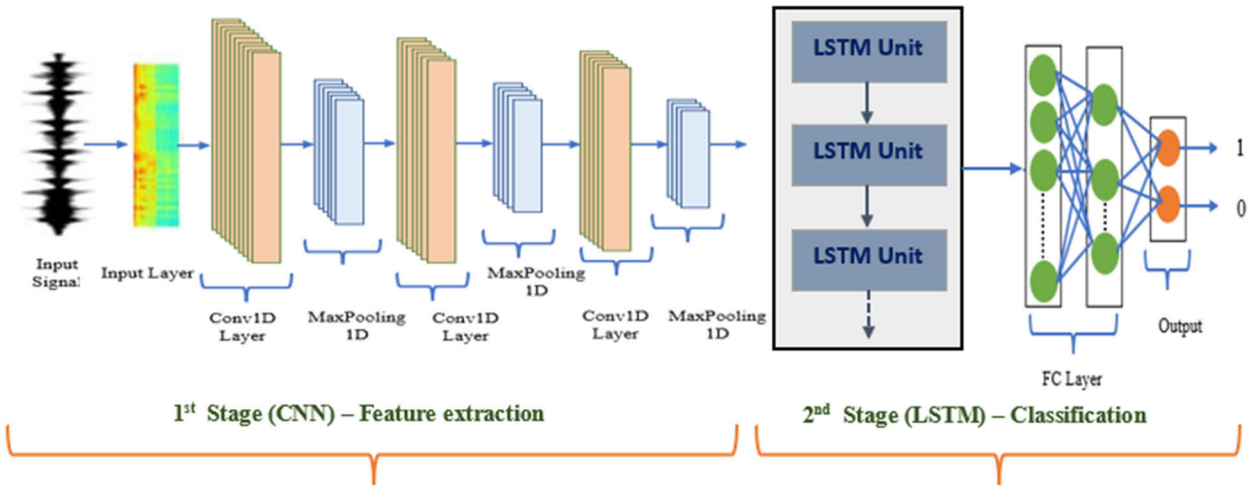

FIGURE 7. The architecture of a hybrid 1D-CNN-LSTM model.

commonly used metrics to assess the model's performance, including accuracy, precision, recall, and F1 score. By identifying the strengths and weaknesses of the model, the testing phase provides valuable insights into ways to further improve its performance.

III. RESULTS AND DISCUSSION

The development of the hybrid 1D-CNN-LSTM model involved configuring the number of features and output types for specific faults, including corona, arcing, tracking, normal cases, and mechanical faults. In this section, we will discuss the classification results of the hybrid model for errors in both the time and frequency domains. It is worth mentioning that the experiments were conducted on a high-performance deep learning workstation equipped with an Intel 12-core i7 GHz CPU, a powerful GTX1080TI GPU, a spacious 320 GB SSD, and 16 GB of ample memory. These exceptional specifications ensured efficient processing and analysis of the data. It is important to note that the calculation time of the algorithm in the time domain (TD) was approximately 40 seconds, while in the frequency domain (FD) it was around 29 seconds. This information emphasizes the computational efficiency of the proposed algorithm and its ability to handle large datasets within reasonable time frames.

A. ARCING FAULTS

To identify arcing faults, the hybrid 1D-CNN-LSTM model underwent training and evaluation using two distinct types of domains, namely TD analysis and FD analysis, in order to perform fault classification.

1) TIME DOMAIN ANALYSIS

Simulations were conducted to analyze the TD for both arcing and non-arcing faults using a dataset of 438 samples, of which 54 were arcing and the rest were non-arcing faults such as corona, tracking, mechanical, and normal cases. The samples were randomly divided into three sets, where 70% were used for training, 15% for model validation, and 15% for model testing. resulting in a total of 306 samples used for training. The obtained results were evaluated using a confusion matrix, which demonstrates the accuracy percentage and error rate for each phase. The testing and validation phases each consisted of 66 samples. Table 4 shows the results of the TD for 1D-CNN-LSTM classification outcomes for arcing faults. The results are based on several trials, and the best outcomes are presented in the form of a confusion matrix. The confusion matrix indicates the percentage of accurate detection and error rate for each phase. The testing and validation phases both had 66 samples. The accuracy of the model was calculated by comparing the predicted values to the actual values in the test data. The confusion matrix provides valuable insights into the performance of the model and helps identify areas where it needs improvement.

TABLE 4. The classification results of the 1D-CNN-LSTM model in the time domain for arcing faults.

	Training	Validation	Testing
Number of Sample	306	66	66
Accuracy Rate	%100	100%	100%
Error Rate	0%	0%	0%
Feature Number		20001	
Output Number		2	

The arcing fault classification model in the TD was able to achieve perfect accuracy of 100% and a 0% error rate,

demonstrating the high effectiveness of the 1D-CNN-LSTM model in detecting switchgear arcing faults, as evidenced by the performance outcomes presented in Table 5. When the arcing TD classifier was used to classify the test dataset, all data occurrences were categorized as either positive or negative, resulting in four classification outcomes: ''true positive,'' ''true negative,'' ''false positive,'' and ''false negative,'' as illustrated in Tables 5. This indicates the ability of the model to correctly identify true instances of arcing faults whereas also minimizing the occurrence of false positives and negatives, which are crucial for effective fault detection in practical applications.

TABLE 5. The output matrix for the training, validation, and testing phases of the arcing fault classification in time domain analysis.

Hybrid 1D-CNN-LSTM Model		
Training Phase		
	Arcing	Non-Arcing
Actual Arcing	34	0
Actual non-Arcing	0	272
Validation Phase		
	Arcing	Non-Arcing
Actual Arcing	10	0
Actual non-Arcing	0	56
Testing Phase		
	Arcing	Non-Arcing
Actual Arcing	10	0
Actual non-Arcing	0	56

The results of a binary classifier during the training phase are presented below:

True Positive (TP): 34 instances were correctly classified as positive.

False Positive (FP): 0 instances were incorrectly classified as positive.

True Negative (TN): 272 instances were correctly classified as negative.

False Negative (FN): 0 instances were incorrectly classified as negative.

The evaluation of the proposed model's performance involved the use of a confusion matrix, which represented arcing and non-arcing scenarios as 0 and 1, respectively. The matrix utilized indices provided in Equations (10)-(14) to assess the model's reliability, dependability, sensitivity, categorization, and F1-measure. Precision was defined as the percentage of identified samples that were relevant, whereas recall measured the proportion of relevant samples correctly identified out of all relevant samples. Furthermore, cross-entropy loss was used to evaluate how well the model's predictions matched the target data specified in Equation (13). Minimizing the distance between the arcing and non-arcing conditions of two probability distributions was a crucial performance requirement, and binary classification techniques were utilized to achieve this. The model's classification accuracy, or arcing fault detection accuracy, was calculated by dividing the total number of correctly classified instances (TP + TN) by the total number of points in the dataset (P + N). This equation was used to evaluate the accuracy of the model's performance for Arcing.

$$\begin{aligned} Acc &= \frac{TP + TN}{TP + TN + FN + FP} \times 100\% \\ &= \frac{TP + TN}{P + N} \times 100\% \\ &= \frac{34 + 272}{34 + 272} \times 100\% = 100\% \end{aligned} \tag{10}$$

The error rate (ERR) is derived by dividing the total number of false classifications by the total number of entire datasets by using the equation as follows:

$$\begin{aligned} ERR &= \frac{FP + FN}{TP + TN + FN + FP} \\ &\quad \times 100\% \\ &= \frac{FP + FN}{P + N} \times 100\% \\ &= \frac{0 + 0}{0 + 0} \times 100\% = 0\% \end{aligned} \tag{11}$$

$$\begin{aligned} Recall\ (Sensitivity)\ (\%) &= \frac{TP}{TP + FN} \times 100\% \\ &= \frac{34}{34 + 0} \times 100\% = 100\% \end{aligned} \tag{12}$$

$$\begin{aligned} Precision\ (Dependability) &= \frac{TP}{TP + FP} \times 100 \\ &= \frac{34}{34 + 0} \times 100 = 100\% \end{aligned} \tag{13}$$

$$\begin{aligned} F1\ measure\ (\%) &= 100 \times 2 \\ &\quad \times \frac{(Precision \times Recall)}{(Precision + Recall)} \\ &= 100 \times 2 \times \frac{100 \times 100}{100 + 100} \\ &= 100\% \end{aligned} \tag{14}$$

The confusion matrix was used in Table 5. to evaluate the performance of the model in identifying arcing and non-arcing scenarios. The recall and precision metrics were used to measure the model's ability to correctly identify relevant samples and the percentage of identified samples that were relevant, respectively.

2) FREQUENCY DOMAIN ANALYSIS

For creating the classification model for arcing faults in a FD analysis, 160 data samples were utilized and divided into three sets for training, validation, and testing, with the training set comprising 70% and the validation and testing

sets containing 15% each, as demonstrated in Table 6. The accuracy and error rates of the classification model were found to be 100% and 0%, respectively, indicating the high effectiveness of the model. The performance of the model was further evaluated using a confusion matrix, which was constructed for the training, validation, and testing phases and is presented in Table 7. The confusion matrix illustrated the model's ability to correctly classify the data into true positives, true negatives, false positives, and false negatives in each phase, thereby providing valuable insights into the model's overall performance.

TABLE 6. The classification results of the 1D-CNN-LSTM model in the frequency domain for arcing faults.

	Training	Validation	Testing
Number of Sample	112	24	24
Accuracy Rate	%100	100%	100%
Error Rate	0%	0%	0%
Feature Number		10001	
Output Number		2	

TABLE 7. The output matrix for the training, validation, and testing phases of the arcing fault classification in time domain analysis.

Hybrid 1D-CNN-LSTM Model		
Training Phase		
	Arcing	Non-Arcing
Actual Arcing	35	0
Actual non-Arcing	0	77
Validation Phase		
	Arcing	Non-Arcing
Actual Arcing	9	0
Actual non-Arcing	0	15
Testing Phase		
	Arcing	Non-Arcing
Actual Arcing	9	0
Actual non-Arcing	0	15

To thoroughly assess the reliability of our hybrid model, we conducted a comprehensive analysis using various performance measures, including accuracy, loss, sensitivity, dependability, and F1-measure. The detailed results of this evaluation can be found in Table 8, which provides valuable insights into the model's performance. By considering multiple performance metrics, we gain a holistic understanding of the model's ability to accurately detect and classify faults, making it highly applicable in real-world scenarios.

Furthermore, in both the TD and FD, we utilized the Adam optimizer for weight optimization during the training process. The RELU activation function was employed to introduce non-linearity, and a learning rate of 0.0001 was chosen to control the optimization process. In addition, a SoftMax activation function was utilized for the final classification layer. For both TD and FD, a batch size of 16 was used, and the model was trained for a total of 60 epochs to ensure optimal convergence and performance.

Overall, the combination of the Adam optimizer, RELU activation function, SoftMax activation function, batch size of 16, and 60 training epochs contributed to the success of our hybrid model in achieving high accuracy and robust fault detection performance in both TD and FD.

TABLE 8. The assessment of metrics and performance for outcomes of the hybrid model during testing for the TD and FD in cases of arcing and non-arcing.

Time Domain				
Case Arcing Findings (0)				
Method	Accuracy	Sensitivity	Dependability	F1-Measure
Hybrid model	100	100	100	100
Case Arcing Findings (1)				
Method	Accuracy	Sensitivity	Dependability	F1-Measure
Hybrid model	100	100	100	100
Frequency Domain				
Case Arcing Findings (0)				
Method	Accuracy	Sensitivity	Dependability	F1-Measure
Hybrid model	100	100	100	100
Case Arcing Findings (1)				
Method	Accuracy	Sensitivity	Dependability	F1-Measure
Hybrid model	100	100	100	100

The study focused on developing a classification model for different switchgear faults, including corona, arcing, tracking, normal, and mechanical faults. Table 9 presents a minimum accuracy of 98.4% and a maximum accuracy of 100%, the results show a high level of accuracy in identifying and categorizing different types of switchgear defects. Maximum and minimum error rates were 0.7% and 0%, respectively. The accuracy for corona fault classification in the TD was marginally worse compared to arcing in the TD and other faults in the FD, according to the output matrix for all classification faults. Since it was impossible to confirm the accuracy of all submitted datasets with their labels, this may be attributable to the input data's cleanliness. The study provides evidence of the potential of hybrid models in accurately classifying different switchgear faults. The results suggest that the model could be useful in real-world applications to improve the reliability and efficiency of power systems

by detecting and diagnosing faults in a timely and accurate manner. However, further research is needed to address the limitations of the current study, such as the quality of the input data and the need for a larger dataset to test the generalizability of the models.

TABLE 9. The accuracy and error rates for various types of switchgear faults.

		Accuracy %			Error Rate %		
Fault	domain	Training	Validation	Testing	Training	Validation	Testing
Arcing	Time	100	100	100	0	0	0
	Frequency	100	100	100	0	0	0
Corona	Time	99.3	98.4	98.4	0.7	1.6	1.6
	Frequency	100	100	100	0	0	0
Mecha-nical	Time	100	100	100	0	0	0
	Frequency	100	100	100	0	0	0
Trackin-g	Time	100	100	100	0	0	0
	Frequency	100	100	100	0	0	0
	Min	99.3	98.4	98.4	0	0	0
	Max	100	100	100	0.7	1.6	1.6
	Average	99.9	99.9	99.9	0.08	0.2	0.2

TABLE 10. Comparison of the proposed algorithm with other approaches for diagnosing switchgear faults.

Methods	Accuracy (%)
1D-CNN-LSTM $_{ARCING}$ [13]	100
1D-CNN-LSTM $_{CORONA}$ [15]	98.4
1D-CNN-LSTM $_{TRACKING}$	100
1D-CNN-LSTM $_{MECHANICAL}$	100
ELM $_{ARCING}$ [47]	95.83
ELM $_{CORONA}$ [14]	87.5
ELM $_{MECHANICAL}$ [2]	91.67
ELM $_{TRACKING}$ [2]	95.33
Random Forest [48]	87.5
Decision Tree [48]	22.9
Federated Deep Learning [49]	95.61

On the other hand, Table 10 shows how well and how competitively the proposed algorithm works by comparing it to other methods like Extreme Learning Machine (ELM), Random Forest (RF), Decision Tree (DT), and Federated Deep Learning (FDL). The table shows the accuracy of the proposed algorithm along with other state-of-the-art methods. The results demonstrate that the proposed algorithm outperformed the other approaches in terms of accuracy, indicating its superior performance in detecting and classifying switchgear faults. Moreover, it is important to note that the proposed algorithm utilizes a 1D-CNN-LSTM architecture, which allows it to effectively capture temporal and frequency features of the signals, resulting in a more accurate classification. The use of DL techniques, such as CNN and LSTM, in fault detection and classification has been gaining popularity in recent years due to their ability to handle complex data and achieve high accuracy rates. Overall, the proposed algorithm was able to find and classify different types of switchgear faults with high accuracy and low error rates. This makes it a promising method for use in the real world.

IV. CONCLUSION

In this study, we proposed a novel approach for fault detection in medium voltage switchgear using a hybrid 1D-CNN-LSTM model. Our model demonstrated high accuracy in both TD and FD fault detection, outperforming traditional machine learning methods and other deep learning models, including Extreme Learning Machine, SVM, and Random Forest. The hybrid model successfully classified various fault types, including arcing, tracking, mechanical, and corona faults, with accuracy rates of 100% in TD and FD for arcing faults and other faults, and 98.4% accuracy in corona fault detection in TD. The key advantage of our hybrid model is its ability to simultaneously learn spatial and temporal features, enabling accurate fault detection in both TD and FD. By leveraging this capability, the model can quickly and accurately identify faults, reducing the likelihood of equipment failure and downtime. The proposed method has potential applications in various fields, such as wind turbines, power grids, and manufacturing processes, where accurate fault detection can enhance maintenance efficiency, equipment reliability, and overall safety.

In conclusion, our study has demonstrated the effectiveness and efficiency of the hybrid 1D-CNN-LSTM model for fault detection in medium voltage switchgear. By combining spatial and temporal feature learning, the model provides a comprehensive understanding of switchgear operation and contributes to the advancement of fault detection methods in power systems. We anticipate that our findings will contribute to the development of advanced fault detection techniques and promote the safe and reliable operation of power systems.

REFERENCES

[1] J. Kennedy, P. Ciufo, and A. Agalgaonkar, "A review of protection systems for distribution networks embedded with renewable generation," *Renew. Sustain. Energy Rev.*, vol. 58, pp. 1308–1317, May 2016, doi: 10.1016/J.RSER.2015.12.258.

[2] S. Ishak, C. T. Yaw, S. P. Koh, S. K. Tiong, C. P. Chen, and T. Yusaf, "Fault classification system for switchgear CBM from an ultrasound analysis technique using extreme learning machine," *Energies*, vol. 14, no. 19, p. 6279, Oct. 2021, doi: 10.3390/en14196279.

[3] R. Collacott, *Mechanical Fault Diagnosis and Condition Monitoring*. Berlin, Germany: Springer, 2012.

[4] L. Zhao, Y. Wu, X. Huang, G. Hong, J. Ren, W. Ning, L. Wang, T. Sun, and S. Yang, "Research on the temperature rise characteristics of medium-voltage switchgear under different operation conditions," *IEEJ Trans. Electr. Electron. Eng.*, vol. 17, no. 5, pp. 654–664, May 2022, doi: 10.1002/TEE.23553.

[5] S. O, A. Bimaridi, J. Park, W. Shin, U. Jo, R. Handito, S. Park, S. Choi, and H. Kang, "A study on the monitoring parameters for power asset management of switchgear," in *Proc. 9th Int. Conf. Condition Monitor. Diagnosis (CMD)*, Kitakyushu, Japan, Nov. 2022, pp. 305–308, doi: 10.23919/CMD54214.2022.9991476.

[6] M. Runde, "Failure frequencies for high-voltage circuit breakers, disconnectors, earthing switches, instrument transformers, and gas-insulated switchgear," *IEEE Trans. Power Del.*, vol. 28, no. 1, pp. 529–530, Jan. 2013, doi: 10.1109/TPWRD.2012.2220638.

[7] A. Kamaludin, H. Prasetia, and Y. Nugroho, "Implementation of GOOSE for overcurrent relays with non-cascade scheme in medium voltage switchgear as breaker failure and busbar protection system," in *Proc. Int. Conf. Technol. Policy Energy Electric Power (ICT-PEP)*, Bandung, Indonesia, Sep. 2020, pp. 179–182, doi: 10.1109/ICT-PEP50916.2020.9249907.

[8] Y. Xing, Z. Wang, L. Liu, Y. Xu, Y. Yang, S. Liu, F. Zhou, S. He, and C. Li, "Defects and failure types of solid insulation in gas-insulated switchgear: In situ study and case analysis," *High Voltage*, vol. 7, no. 1, pp. 158–164, Feb. 2022, doi: 10.1049/hve2.12127.

[9] S. A. Eroshenko, A. O. Egorov, M. R. Zagidullin, and M. D. Senyuk, "The indicators system for the short circuit currents levels assessment in the power systems," in *Proc. 15th Int. Conf. Elect. Mach., Drives Power Syst. (ELMA)*, Sofia, Bulgaria, 2017, pp. 144–148, doi: 10.1109/ELMA.2017.7955419.

[10] J. Erbrink, R. Vosse, M. Mosselaar, and T. van Rijn, "Practical verification of medium voltage switchgear thermal loadability based on the IEC62271 thermal model," in *Proc. Int. Conf. Exhib. Electr. Distrib.*, 2021, pp. 218–222.

[11] A. Baug, N. R. Choudhury, R. Ghosh, S. Dalai, and B. Chatterjee, "Identification of single and multiple partial discharge sources by optical method using mathematical morphology aided sparse representation classifier," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 6, pp. 3703–3712, Dec. 2017, doi: 10.1109/TDEI.2017.006398.

[12] G. Paoletti and M. Baier, "Failure contributors of MV electrical equipment and condition assessment program development," in *Proc. Conf. Rec. Annu. Pulp Paper Ind. Tech. Conf.*, Portland, OR, USA, 2001, pp. 37–47, doi: 10.1109/PAPCON.2001.952946.

[13] Y. A. M. Alsumaidaee, C. T. Yaw, S. P. Koh, S. K. Tiong, C. P. Chen, C. H. Tan, K. Ali, and Y. A. L. Balasubramaniam, "Detecting arcing faults in switchgear by using deep learning techniques," *Appl. Sci.*, vol. 13, no. 7, p. 4617, Apr. 2023, doi: 10.3390/app13074617.

[14] S. Ishak, S.-P. Koh, J.-D. Tan, S.-K. Tiong, and C.-P. Chen, "Corona fault detection in switchgear with extreme learning machine," *Bull. Electr. Eng. Informat.*, vol. 9, no. 2, pp. 558–564, Apr. 2020, doi: 10.11591/eei.v9i2.2058.

[15] Y. A. M. Alsumaidaee, C. T. Yaw, S. P. Koh, S. K. Tiong, C. P. Chen, T. Yusaf, A. N. Abdalla, K. Ali, and A. A. Raj, "Detection of corona faults in switchgear by using 1D-CNN, LSTM, and 1D-CNN-LSTM methods," *Sensors*, vol. 23, no. 6, p. 3108, Mar. 2023, doi: 10.3390/s23063108.

[16] Z. Jia, S. Fan, D. He, J. Wang, H. Dong, and Z. Tang, "Multi-physical tracking and sensing technology of switchgear partial discharge based on edge computing and smart gateway," *J. Phys., Conf. Ser.*, vol. 1952, no. 4, Jun. 2021, Art. no. 042101, doi: 10.1088/1742-6596/1952/4/042101.

[17] M. Yi, M. Pu, Z. Zhu, C. Gu, H. Su, and X. Wang, "Research on insulation aging of distribution switchgear," in *Proc. Int. Conf. Condition Monitor. Diagnosis (CMD)*, Xi'an, China, Sep. 2016, pp. 206–209, doi: 10.1109/CMD.2016.7757781.

[18] Y. Alsumaidaee, C. Yaw, S. Koh, S. Tiong, C. Chen, and K. Ali, "Review of medium-voltage switchgear fault detection in a condition-based monitoring system by using deep learning," *Energies*, vol. 15, no. 18, p. 6762, Sep. 2022, doi: 10.3390/en15186762.

[19] A. Subramaniam, A. Sahoo, S. S. Manohar, S. J. Raman, and S. K. Panda, "Switchgear condition assessment and lifecycle management: Standards, failure statistics, condition assessment, partial discharge analysis, maintenance approaches, and future trends," *IEEE Elect. Insul. Mag.*, vol. 37, no. 3, pp. 27–41, May 2021, doi: 10.1109/MEI.2021.9399911.

[20] N. Davies and D. Jones, "Testing distribution switchgear for partial discharge in the laboratory and the field," in *Proc. Conf. Rec. IEEE Int. Symp. Electr. Insul.*, Vancouver, BC, Canada, Jun. 2008, pp. 716–719, doi: 10.1109/ELINSL.2008.4570430.

[21] P. Gill, *Electrical Power Equipment Maintenance and Testing*. Boca Raton, FL, USA: CRC Press, 2008.

[22] C. Zhang, M. Dong, M. Ren, and W. Huang, "Partial discharge monitoring on metal-enclosed switchgear with distributed non-contact sensors," *Sensors*, vol. 18, no. 2, p. 551, 2018, doi: 10.3390/s18020551.

[23] I. Blokhintsev, J. Kozusko, B. Oberer, and D. Anzaldi, "Continuous and remote monitoring of partial discharge in medium voltage switchgear," in *Proc. IEEE Electr. Insul. Conf. (EIC)*, Baltimore, MD, USA, Jun. 2017, pp. 205–208, doi: 10.1109/EIC.2017.8004669.

[24] C. D. Horne, "Condition monitoring of power station primary electrical plant," in *Proc. IET Power Convention*, 2007, pp. 1–12.

[25] A. P. Purnomoadi, A. R. Mor, and J. J. Smit, "Health index and risk assessment models for gas insulated switchgear (GIS) operating under tropical conditions," *Int. J. Electr. Power Energy Syst.*, vol. 117, May 2020, Art. no. 105681, doi: 10.1016/j.ijepes.2019.105681.

[26] M. Ivanova, R. Dimitrova, and A. Filipov, "Analysis of power outages and human errors in the operation of equipment in power grids," in *Proc. 12th Electr. Eng. Fac. Conf. (BulEF)*, Varna, Bulgaria, Sep. 2020, pp. 1–5, doi: 10.1109/BulEF51036.2020.9326058.

[27] C. Yang, B. Cai, Q. Wu, C. Wang, W. Ge, Z. Hu, W. Zhu, L. Zhang, and L. Wang, "Digital twin-driven fault diagnosis method for composite faults by combining virtual and real data," *J. Ind. Inf. Integr.*, vol. 33, Jun. 2023, Art. no. 100469, doi: 10.1016/J.JII.2023.100469.

[28] X. Kong, B. Cai, Y. Liu, H. Zhu, Y. Liu, H. Shao, C. Yang, H. Li, and T. Mo, "Optimal sensor placement methodology of hydraulic control system for fault diagnosis," *Mech. Syst. Signal Process.*, vol. 174, Jul. 2022, Art. no. 109069, doi: 10.1016/j.ymssp.2022.109069.

[29] X. Kong, B. Cai, Y. Liu, H. Zhu, C. Yang, C. Gao, Y. Liu, Z. Liu, and R. Ji, "Fault diagnosis methodology of redundant closed-loop feedback control systems: Subsea blowout preventer system as a case study," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 3, pp. 1618–1629, Mar. 2023, doi: 10.1109/TSMC.2022.3204777.

[30] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, "A survey of deep learning and its applications: A new paradigm to machine learning," *Arch. Comput. Methods Eng.*, vol. 27, no. 4, pp. 1071–1092, Sep. 2020, doi: 10.1007/s11831-019-09344-w.

[31] G. Zhao, G. Zhang, Q. Ge, and X. Liu, "Research advances in fault diagnosis and prognostic based on deep learning," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Chengdu)*, Chengdu, China, Oct. 2016, pp. 1–6, doi: 10.1109/PHM.2016.7819786.

[32] C. Li, Z. Yu, and M. Zhuo, "Research on fault detection method of infrared thermal imaging for power equipment based on deep learning," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 714, no. 4, 2021, Art. no. 042045, doi: 10.1088/1755-1315/714/4/042045.

[33] A. Davies, Ed., *Handbook of Condition Monitoring: Techniques and Methodology*. Springer, 2012, doi: 10.1007/978-94-011-4924-2.

[34] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Deep learning algorithms for bearing fault diagnostics—A comprehensive review," *IEEE Access*, vol. 8, pp. 29857–29881, 2020, doi: 10.1109/ACCESS.2020.2972859.

[35] R. Iqbal, T. Maniak, F. Doctor, and C. Karyotis, "Fault detection and isolation in industrial processes using deep learning approaches," *IEEE Trans. Ind. Informatics*, vol. 15, no. 5, pp. 3077–3084, May 2019, doi: 10.1109/TII.2019.2902274.

[36] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6, doi: 10.1109/ICEngTechnol.2017.8308186.

[37] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *J. Syst. Eng. Electron.*, vol. 28, no. 1, pp. 162–169, Feb. 2017, doi: 10.21629/JSEE.2017.01.18.

[38] S. Arshi, L. Zhang, and R. Strachan, "Prediction using LSTM networks," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Budapest, Hungary, Jul. 2019, pp. 1–8, doi: 10.1109/IJCNN.2019.8852206.

[39] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: Current status and future directions," *Int. J. Forecasting*, vol. 37, no. 1, pp. 388–427, Jan. 2021, doi: 10.1016/j.ijforecast.2020.06.008.

[40] A. K. Ozcanli and M. Baysal, "Islanding detection in microgrid using deep learning based on 1D CNN and CNN-LSTM networks," *Sustain. Energy, Grids Netw.*, vol. 32, Dec. 2022, Art. no. 100839, doi: 10.1016/j.segan.2022.100839.
[41] W. Zhang, G. Biswas, Q. Zhao, H. Zhao, and W. Feng, "Knowledge distilling based model compression and feature learning in fault diagnosis," *Appl. Soft Comput.*, vol. 88, Mar. 2020, Art. no. 105958, doi: 10.1016/j.asoc.2019.105958.
[42] J. Zhu, H. Chen, and W. Ye, "A hybrid CNN–LSTM network for the classification of human activities based on micro-Doppler radar," *IEEE Access*, vol. 8, pp. 24713–24720, 2020, doi: 10.1109/ACCESS.2020.2971064.
[43] K. Nishikawa, R. Hirakawa, H. Kawano, K. Nakashi, and Y. Nakatoh, "Detecting system Alzheimer's dementia by 1D CNN-LSTM in Japanese speech," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Las Vegas, NV, USA, Jan. 2021, pp. 1–3, doi: 10.1109/ICCE50685.2021.9427692.
[44] S. Li, Z. Chen, X. Li, J. Lu, and J. Zhou, "Unsupervised variational video hashing with 1D-CNN-LSTM networks," *IEEE Trans. Multimedia*, vol. 22, no. 6, pp. 1542–1554, Jun. 2020, doi: 10.1109/TMM.2019.2946096.
[45] T. M. R. Bashar, M. Munem, M. S. Islam, M. Hossain, T. B. Shawkat, and H. Rahaman, "Optimized hybrid neural network for wind speed forecasting," in *Proc. IEEE Elect. Power Energy Conf. (EPEC)*, Victoria, BC, Canada, Dec. 2022, pp. 284–289, doi: 10.1109/EPEC56903.2022.10000164.
[46] M. Dörfler, R. Bammer, and T. Grill, "Inside the spectrogram: Convolutional neural networks in audio processing," in *Proc. Int. Conf. Sampling Theory Appl. (SampTA)*, Tallinn, Estonia, Jul. 2017, pp. 152–155, doi: 10.1109/SAMPTA.2017.8024472.
[47] S. Ishak, S. P. Koh, J. D. Tan, S. K. Tiong, C. P. Chen, and C. T. Yaw, "Arcing faults detection in switchgear with extreme learning machine," *J. Phys., Conf. Ser.*, vol. 2319, no. 1, Aug. 2022, Art. no. 012007, doi: 10.1088/1742-6596/2319/1/012007.
[48] N. A. Muhamad, I. V. Musa, Z. A. Malek, and A. S. Mahdi, "Classification of partial discharge fault sources on SF6 insulated switchgear based on twelve by-product gases random forest pattern recognition," *IEEE Access*, vol. 8, pp. 212659–212674, 2020, doi: 10.1109/ACCESS.2020.3040421.
[49] Y. Wang, J. Yan, Q. Jing, J. Wang, and Y. Geng, "A novel federated deep learning framework for diagnosis of partial discharge in gas-insulated switchgear," *Meas. Sci. Technol.*, vol. 33, no. 9, Sep. 2022, Art. no. 095112, doi: 10.1088/1361-6501/ac7a09.

YASEEN AHMED MOHAMMED ALSUMAIDAEE received the B.Sc. degree in software engineering from Northern Technical University, in 2011, and the M.Sc. degree in electrical and computer engineering from Altinbas University, Turkey, in 2019. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Universiti Tenaga Nasional (UNITEN), Malaysia. His research interests include machine learning and deep learning.

JOHNNY KOH SIAW PAW received the bachelor's (Hons.), M.Sc., and Ph.D. degrees in electrical and electronic engineering from Universiti Putra Malaysia, in 2000, 2002, and 2008, respectively. He is currently a Professor with the Institute of Sustainable Energy, Universiti Tenaga Nasional. His research interests include machine intelligence, automation technology, and renewable energy.

CHONG TAK YAW received the bachelor's and master's degrees (Hons.) in electrical and electronics engineering and the Ph.D. degree in artificial neural network from Universiti Tenaga Nasional (UNITEN), Malaysia, in 2008, 2012, and 2019, respectively. He is currently a Postdoctoral Researcher with the Institute of Sustainable Energy, UNITEN. His research interests include artificial neural networks and renewable energy.

SIEH KIONG TIONG (Senior Member, IEEE) received the B.Eng. (Hons.), M.Sc., and Ph.D. degrees in electrical, electronic and system engineering from the National University of Malaysia (UKM), in 1997, 2000, and 2006, respectively. He is currently a Professor with the College of Engineering. He is also the Director of the Institute of Sustainable Energy (ISE), Universiti Tenaga Nasional. His research interests include renewable energy, artificial intelligence, data analytics, microcontroller systems, and communication systems. He is also a Professional Engineer registered with the Board of Engineers Malaysia (BEM).

CHAI PHING CHEN is currently a Senior Lecturer with Universiti Tenaga Nasional (UNITEN), Malaysia. She has been associated with technical education for more than ten years and actively participated in various UNITEN projects that involve machine learning. Her project field contains prediction for gas emission from power plant; prediction for short-term wind speed; health condition analysis for HV circuit breaker; fault detection for switchgear; fault detection for transformer; heat waste recovery system via thermoelectric generator; and non-technical losses detection.

TALAL YUSAF received the B.Sc. degree (Hons.) from UoT Baghdad, in 1987, the master's degree in engineering science from the National University of Malaysia (UKM), in 1994, the Ph.D. degree from USQ, Queensland, Australia, and the Ph.D. degree in renewable energy and biotechnology from UKM, in 1999. He commenced his role as the Pro Vice Chancellor of International and Partnership with Federation University Australia, in September 2018, and was the Executive Director (International and Research Partnership) with the University of Southern Queensland, from 2008 to 2018. He is currently an Executive Dean of the Higher Education and Emerging Technologies, Aviation Australia, Queensland Government, Australia, and an Adjunct Professor with Central Queensland University. His international university career has included teaching, research, and management positions in Southeast Asia, Middle East, the U.K., and Queensland for the last 30 years. He has a strong research background in renewable energy, hydrogen technologies, environmental pollution, and alternative fuels for IC engines. He has supervised over 50 Ph.D. and master's students and has written and contributed to over 200 journal articles and ten books. His professional memberships and affiliations include the Pro Vice-Chancellor/Deputy Vice Chancellor Forum (Australia), the Australian University International Director Forum (AUIDF), the Chair of the QLD International Director Forum, the Director of the International Sponsor Engagement and Research Partnerships (ISERP), and the Bio-Energy Research Group Leader of the National Centre for Engineering in Agriculture Queensland. He is the Associate Editor-in-Chief of *Sustainability*, *Applied Sciences*, *Energies*, and *Groundwater for Sustainable Development*.

FOO BENEDICT received the dual bachelor's degrees in mechanical engineering and the Doctor of Philosophy degree in engineering from Universiti Malaysia Pahang. He is currently the Managing Director of Enhance Track Sdn. Bhd. He is also the Founder of Enhance Track Sdn. Bhd. He has been managing the company for 17 years since its establishment in 2005. As the Managing Director of Enhance Track Sdn. Bhd., he has vast experience in many industries, such as oil and gas, renewable energy, education, and laboratory and testing equipment. In this role, he has also personally been involved with all the research and development projects that the company has undertaken and improving the outcome for clients in Malaysia, Australia, and Middle East. He is also active in associations and professional bodies, such as the Associated Chinese Chamber of Commerce and Industries of Malaysia (ACCCIM) and the Selangor and Kuala Lumpur Foundry and Engineering Association (SFEIA). He received the Graduate Certificate in Mechanical Engineering from the University of Southern Queensland, Australia. He is also a Professional Technologist registered with the Malaysia Board of Technologists (MBOT).

KUMARAN KADIRGAMA is currently a Professor and a Research Fellow of the Advanced Nano Coolant Lubricant Laboratory (ANCoL), Automotive Engineering Research Group, Faculty of Mechanical and Automotive Engineering Technology, University Malaysia Pahang (UMP). In research, he is involved in the supervision of postgraduate students at the master's and Ph.D. levels and has mentored numerous students to graduation. He has also published and presented various technical papers and journals at an international and national level. He has an H-index of 31 with 4199 citations. He has received grants totaling RM 6.02 million from various agencies and institutions. He has won gold medals in the International Invention, Innovation and Technology Exhibition (ITEX), Seoul International Invention Fair (SIFF), South Korea, and British Invention Show (BIS). He is a Professional Engineer registered under the Board of Engineers Malaysia (BEM); and a Charted Engineer (U.K.) under the Institution of Mechanical Engineers (IMechE).

TAN CHUNG HONG received the bachelor's (Hons.) and Ph.D. degrees in chemical engineering from the University of Nottingham Malaysia (UNM), in 2013 and 2020, respectively. He is currently a Postdoctoral Researcher with the Institute of Sustainable Energy, Universiti Tenaga Nasional (UNITEN). His research interests include renewable energy and machine learning.

AHMED N. ABDALLA received the B.Sc. and M.Sc. degrees from the University of Technology, Baghdad, Iraq, in 1997 and 2002, respectively, and the Ph.D. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2007. He was the former Dean of the Workshop and Training Center, University of Technology, Iraq. He is currently a Professor with the Huaiyin Institute of Technology, Huai'an, China. He has authored or coauthored numerous articles published in several SCI, EI, and Scopus-indexed journals with an impact factor. His research outcomes have been exhibited and have been bestowed high recognition internationally. His research interests include system modeling and parameter identification, sensor design, and the application of intelligent techniques.