ISSN: 2723-9535

Available online at www.HighTechJournal.org

HighTech and Innovation Journal

Vol. 7, No. 2, June, 2026

Fault Diagnosis in Electrical Power Networks Using a Novel Hybrid Deep Learning Model

Al-Hussein M. Alturf [1*], Shahad Mohammed Radhi [1], Sadeq D. Al-Majidi [1], Mohammed Kh. Al-Nussairi [1]

[1] *Department of Electrical Engineering, College of Engineering, University of Misan, Amarah 62001, Iraq.*

Received 03 April 2026; Revised 26 May 2026; Accepted 28 May 2026; Published 01 June 2026

Abstract

Faults in high-voltage transmission lines (HVTLs) represent a significant challenge to the stability and reliability of power grids. Therefore, fault detection (FD) and fault classification (FC) in HVTLs are crucial for ensuring rapid power supply restoration and preventing electrical energy losses. In this paper, a novel hybrid deep learning model combining Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) techniques with the Harris Hawks Optimization (HHO) algorithm is proposed for accurate fault detection and classification. Three power grids, rated at 132 kV and 50 MW and extending over a total length of 150 km, were designed in the MATLAB environment to generate the training data. The data were then processed using Python to train the proposed model. Subsequently, HHO was employed to determine the optimal parameters of the LSTM-GRU model, including the number of neurons, learning rate, dropout rate, and mean squared error. To evaluate the model's performance, four statistical measures were used: the confusion matrix, precision, recall, and F1-score. The results show that the LSTM-GRU-HHO model achieved the highest FD and FC accuracies of 99.90% and 99.84%, respectively, outperforming the LSTM, GRU, and LSTM-GRU models, as well as related models reported in previous studies.

Keywords: Deep Learning; Fault Diagnosis; Transmission Line; Power System; Optimization Algorithm.

1. Introduction

In general, the stability and reliability of modern power grids are regarded as critical factors in ensuring the continuous and efficient operation of industrial, commercial, and residential sectors, thereby minimizing potential disruptions and economic losses. On the other hand, one of the major challenges affecting power-grid stability and reliability is the occurrence of electrical faults in high-voltage transmission lines (HVTLs) [1, 2]. Such faults typically occur for various reasons, including transmission-line insulation breakdown, thunderstorms, equipment failure, human and animal interference, falling trees, and unusual weather conditions [3, 4]. These faults may lead to power outages, increased maintenance costs, equipment damage, and recovery expenses over time. In practical applications, conventional protection devices based on impedance measurement, such as relays and circuit breakers, are used to isolate and recover from fault conditions. However, these devices are often unable to address such problems accurately and promptly. Therefore, faults must be detected and cleared immediately to ensure the stability and reliability of the power grid [5, 6]. Recently, several researchers have developed fault-diagnosis models based on intelligent algorithms to address fault detection (FD) and fault classification (FC) in HVTLs. Theoretical and practical fault data, including

* Corresponding author: al_hussein.m@uomisan.edu.iq

https://doi.org/10.28991/HIJ-2026-07-02-02

voltage and current measurements from the grid, have been used to train and develop these diagnostic models for new fault cases [7]. These models are then integrated into intelligent power-system networks based on machine-learning approaches to enable early fault detection.

In recent years, numerous studies have been conducted on FD and FC methodologies. Accordingly, researchers developed an effective fault-detection system for electrical transmission lines using an Extreme Learning Machine (ELM). Two HVTL models, TL-1 and TL-2, were implemented to collect fault-training data, and the proposed ELM classification model achieved accuracy rates of 99.18% and 99.09% for TL-1 and TL-2, respectively [6]. The researchers further optimized the weighted ELM algorithm using Particle Swarm Optimization (PSO) to adjust the hidden-layer weights and input biases of the ELM, thereby improving overall performance and reducing errors. The optimized model achieved an accuracy of 99.85% across various fault types [8]. In another study, the authors presented a new FC technique based on a Hierarchical Ensemble Extreme Learning Machine (HE-ELM) to identify different types of HVTL faults. This method addressed the limitations of conventional approaches such as Support Vector Machines (SVMs) and ELMs, achieving an accuracy of approximately 99.91% in identifying internal faults under various conditions [9]. Meanwhile, Chen et al. [10] proposed an integrated framework based on the Wave-ELM Grouping (SW-ELM) algorithm and extended it to include Gaussian Grouping (SG-ELM), incorporating feature extraction directly into the activation layer. The proposed framework was then applied to three different transmission-line configurations. Ultimately, very low error rates of less than 2% were recorded in estimating fault locations and types across all three configurations.

In the study reported by Akmaz et al. [11], the researchers presented a hybrid approach that first relies on Fast Fourier Transform (FFT) analysis of the waveforms and then refines the results using an ELM regression model to minimize errors caused by impedance. FFT was applied to the transient currents to extract the principal frequencies, (f_i), with values $\geq$ 244 Hz, and the equation ($x = v/(2f_i)$) was used to calculate the approximate distance from the measurement point to the fault. The results showed that the average distance estimation error was approximately 5.31% for the test data before refinement. After refinement, the average error decreased to 1.33% for both the training and test datasets.

In another study, Long Short-Term Memory (LSTM) networks were used to classify transmission-line faults and were compared with traditional feature extraction methods. The LSTM networks achieved 100% accuracy under noise-free conditions and maintained high accuracies of 99.77% and 99.55% under Gaussian noise levels of 30 dB and 20 dB, respectively, outperforming traditional methods [12]. In 2024, researchers compared the performance of three deep learning models: an Artificial Neural Network (ANN), an LSTM network, and an LSTM with window regression (LSTM-WR). In the experiments, the ANN achieved an accuracy of only 42.98%, whereas the LSTM performed exceptionally well, achieving an accuracy of 99.98% on the test datasets. The LSTM-WR slightly outperformed the LSTM, achieving an accuracy of 99.99% across different tests [13].

Minh et al. [14] proposed an integrated framework that first employs multiple machine learning algorithms, including Support Vector Machine (SVM), Decision Tree, Logistic Regression, XGBoost, and ANN, to classify fault types. It then uses deep learning techniques, including a Convolutional Neural Network (CNN), LSTM, and a hybrid CNN-LSTM model, to estimate fault distances. The training data were collected through various tests conducted at different locations to optimize the models. Consequently, the results showed that XGBoost achieved a classification accuracy of 99.82%, while the hybrid CNN-LSTM model achieved a mean absolute error in the location ratio of less than 1% and a mean absolute distance error of less than 0.16 km.

Furthermore, Altaie & Abderrahim [15] identified conventional HVTL fault classes by proposing a new framework that utilizes a single-ended sensor array to capture phase currents and voltages. The recorded data were transformed into color scalograms through a stepwise continuous wavelet transform (S-CWT), and classification was performed using pretrained Convolutional Neural Networks (VGG19 and VGG16). The framework was evaluated using 61,226 images across 10 fault types and achieved 100% classification accuracy in both noise-free and high-noise environments, demonstrating its robustness and scalability without requiring algorithmic adjustments.

Another study introduced a hybrid data-driven method for fault identification in shunt-compensated transmission lines, addressing the limitations of traditional signal-processing methods. Using MATLAB/Simulink, the researchers modeled a STATCOM-compensated system. They applied the Daubechies-4 Discrete Wavelet Transform to analyze three-phase fault-current waveforms and extract characteristics such as standard deviation and power. These features were used to train multiple classifiers, among which the Naive Bayes (NB) classifier demonstrated superior performance in terms of accuracy and other statistical metrics, proving effective for identifying fault types in shunt-compensated transmission lines [16].

Finally, Leh et al. [17] developed fault detection (FD) and fault classification (FC) frameworks based on a feedforward Artificial Neural Network (ANN). The FD model was trained using the Levenberg–Marquardt (trainlm) algorithm, while the FC model was trained using the Scaled Conjugate Gradient (trainscg) algorithm. The FD stage

achieved a mean error of 5.5614×10^{-8}, a correlation coefficient of 1, and an accuracy of 100%. In contrast, the FC stage achieved a mean error of 0.43699, a correlation coefficient of 0.83955, and an accuracy of 70%.

The researchers also developed a predictive method for power transmission-line disconnection faults using LSTMs and Support Vector Machines (SVMs). The LSTMs were employed to extract temporal characteristics from multisource data, while the SVMs provided robust classification for fault prediction. Experimental data from the Wanjiang substation of the China Southern Power Grid, comprising 5,120 samples collected between 2012 and 2014, were used for the simulation. The input measurements included current, voltage, and active power. The model achieved an average accuracy of 97.7%, outperforming conventional LSTM-only methods by approximately 10% [18].

In another study, the researchers proposed a novel and adaptive methodology for fault detection and localization in three-phase power transmission lines based on a Long Short-Term Memory Autoencoder (LSTM-AE) model. The proposed model was trained and tested using a large dataset comprising more than 50,000 simulated fault cases generated in the MATLAB/Simulink environment, in addition to experimental validation using 1,000 real fault cases obtained from an open-source data repository. The results showed that the proposed LSTM-AE model achieved the highest classification accuracy of 98%, outperforming both the conventional LSTM model and the one-dimensional Convolutional Neural Network (1D-CNN) model. The F1-score values for the LSTM and 1D-CNN models were approximately 95.2% and 96.9%, respectively, reflecting an effective balance between precision and recall; however, neither achieved the superior performance demonstrated by the proposed model [19].

As noted, most previous studies have employed various Deep Neural Network (DNN) models to implement power-system protection schemes. However, many existing approaches still suffer from limited optimization capabilities and reduced classification accuracy under complex operating conditions. Furthermore, most previous studies have focused separately on either fault detection or fault classification rather than integrating both tasks within a unified framework. In addition, insufficient attention has been paid to combining hybrid deep learning architectures, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, with modern optimization algorithms to improve model robustness and convergence speed.

Therefore, this study develops a hybrid deep learning approach for both the FD and FC stages using an LSTM-GRU model optimized through the Harris Hawks Optimization (HHO) algorithm. The proposed framework is designed to accurately diagnose HVTL faults within a single unified model, thereby improving computational efficiency and reducing the complexity associated with using separate models for each task. As a result, the proposed LSTM-GRU-HHO framework provides a more robust and accurate fault-classification solution, achieving an accuracy of 99.84% compared with conventional standalone deep learning methods and recent approaches reported in the literature.

The LSTM layers are integrated with the GRU model to improve the accuracy of fault-data processing, while the HHO algorithm is employed to optimize the main parameters of the LSTM-GRU model, as illustrated in Figure 1. The proposed framework is divided into three stages. In the first stage, the processed data are initialized and divided into training and testing datasets. In the second stage, the LSTM-GRU model layers are constructed according to the fitness functions. Finally, the HHO algorithm is used to update and determine the optimal parameters of the proposed LSTM-GRU model. The main contributions of this work are summarized in the following bullet points:

- A comprehensive dataset was simulated in the MATLAB environment based on various fault scenarios for a 50 MW, 132 kV, 150 km high-voltage transmission line.
- An innovative framework was proposed by combining LSTM and GRU deep learning models in the Python environment. The resulting model has a simple structure with substantial memory capacity.
- The HHO algorithm was applied to tune and optimize the parameters of the LSTM-GRU model, including the learning rate, number of neurons per layer, dropout rate, and mean squared error.
- A comparative analysis of several deep learning models, including LSTM, GRU, and LSTM-GRU, was conducted using the confusion matrix, precision, recall, and F1-score. The results show that the proposed LSTM-GRU-HHO model achieved the highest fault-diagnosis accuracy across the various test scenarios.

The paper is organized into seven main sections. Section 2 explains faults in HVTLs, while Section 3 presents the proposed FD and FC framework. Section 4 describes the methods and their mathematical formulations. Section 5 presents the proposed LSTM-GRU model optimized using the HHO algorithm. Section 6 discusses the main results of the fault tests. Finally, Section 7 summarizes the key conclusions of the study.

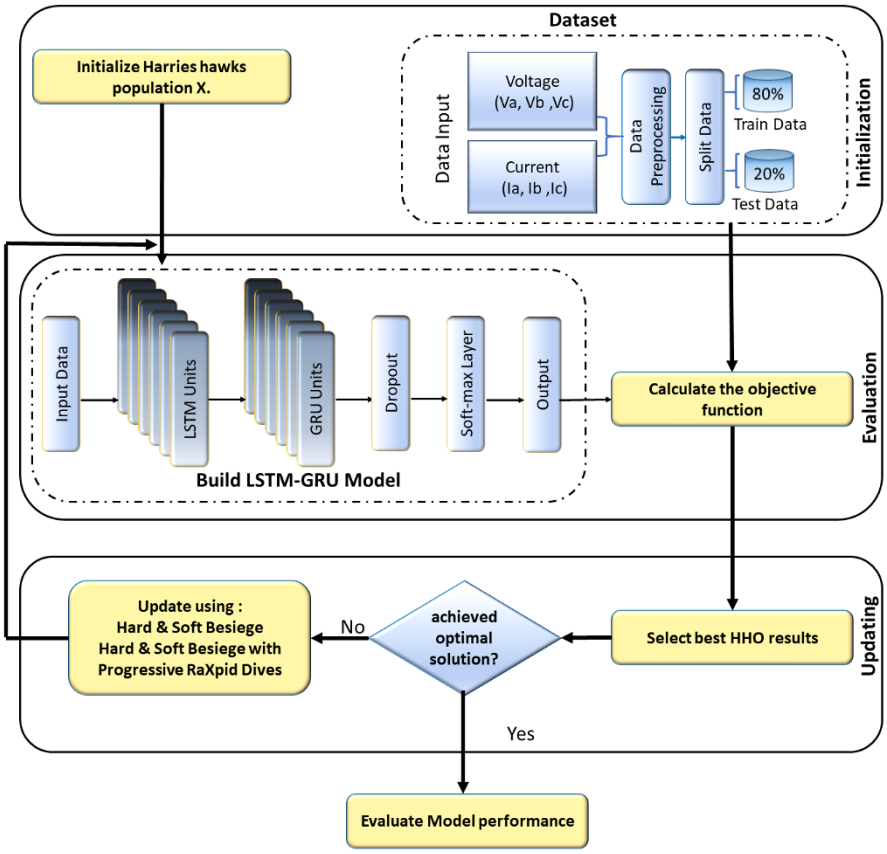


Figure 1. The structure of a proposed model

2. Fault in High Voltage Transmission Lines

Basically, a power system network consists of three main components: generation units, transmission lines, and distribution systems. High-voltage transmission lines (HVTLs) are a vital part of modern power grids because they transfer electrical energy from generation units to distribution systems. Ensuring the reliable operation of HVTLs is essential for maintaining the stability of power generation under different operating conditions. Although faults may occur in all components of a power system, faults in HVTLs are considered among the most serious challenges affecting the stability and reliability of modern power grids [20]. A fault is generally defined as an abnormal electrical condition caused by internal or external circumstances, resulting in deviations in grid current and voltage [21]. Under normal operating conditions, faults are cleared by protection devices such as circuit breakers and relays, which operate based on impedance-distance measurements. However, mechanical failure of protection devices, accidental short circuits, or overload conditions may lead to fault occurrence [14, 22]. In such cases, the grid current may increase, the grid voltage may decrease, and the grid power may fluctuate, potentially causing equipment damage or even a complete power-system shutdown [23].

HVTL faults are typically classified into two major categories: symmetrical and unsymmetrical faults, as shown in Figure 2. Symmetrical faults occur when all three phases of the power grid are equally affected and are divided into three-phase faults (L–L–L) and three-phase-to-ground faults (L–L–L–G). The second category consists of unsymmetrical faults, which occur when not all phases are affected by internal or external conditions. These faults are divided into three main types: single line-to-ground (L–G), line-to-line (L–L), and double line-to-ground (L–L–G) faults [24].

To address HVTL fault-related problems, three fundamental tasks are employed: detection, classification, and location. In the first stage, HVTL faults are detected by a digital relay through analysis of deviations in grid current and voltage. In the second stage, the fault type is classified using an intelligent model trained on HVTL fault data. In the final stage, locating the fault along the HVTL is essential for minimizing power-grid outages. Three main techniques are used for this purpose: impedance-based methods, traveling-wave methods, and data-driven methods.

The development of advanced FD and FC technologies is a crucial step toward building smarter, more efficient, and more reliable power grids. By integrating deep learning algorithms with intelligent optimization methods, faster response times and higher fault-diagnosis accuracy can be achieved, thereby reducing outage duration and improving the overall reliability and stability of the power system [25]. Therefore, this study employs a two-stage deep learning framework integrated with an optimization algorithm to develop FD and FC models for HVTLs.

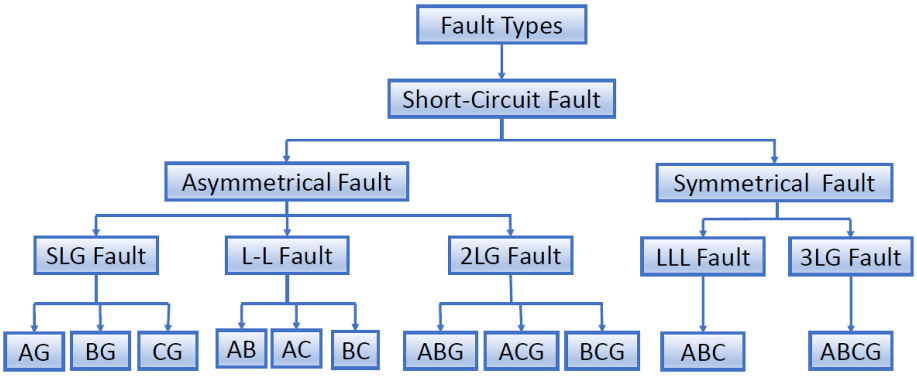


Figure 2. The classification of fault types in HVTL

3. Proposed Framework of Fault Detection and Classification

In this study, a novel framework for fault detection (FD) and fault classification (FC) is proposed, as illustrated in Figure 3. A power transmission system was simulated using MATLAB/Simulink to generate voltage and current data representing 10 different fault types. In total, 40,010 fault samples and 4,001 non-fault samples were collected. The dataset was then divided into training and testing subsets using an 80:20 ratio. This split is commonly adopted in FD and FC applications for power systems because it provides a practical balance between learning accuracy and evaluation reliability.

The training subset, comprising 80% of the data, was used to provide the intelligent model with a sufficient number of samples to learn the relationships between the input and output data. The remaining 20% was used to evaluate the model's generalization capability. A normalization process was subsequently applied to standardize the input variables and ensure that all values fell within the same range. Finally, the proposed hybrid LSTM-GRU-HHO model was trained using binary labels for the FD stage and multiclass labels for the FC stage. The proposed model was then compared with the LSTM, GRU, and LSTM-GRU models to analyze performance differences between conventional models and hybrid or optimized models, and to evaluate the effects of different deep network architectures on the accuracy and efficiency of FD and FC in the HVTL test system.

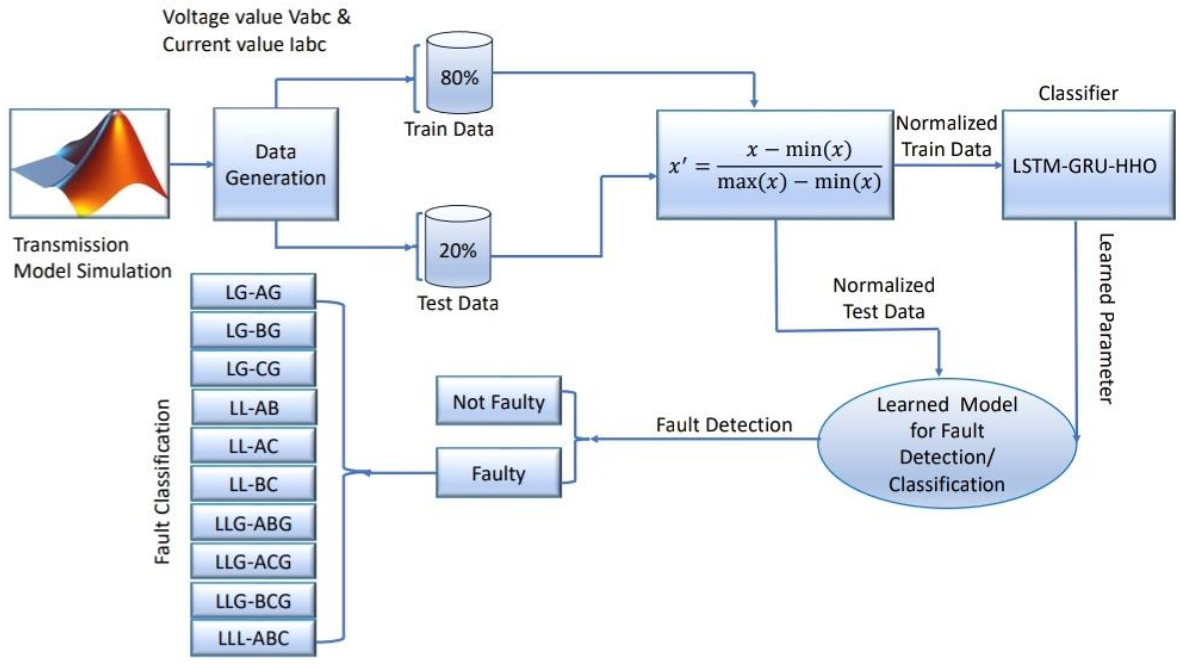


Figure 3. The proposed work of a detection and classification faults

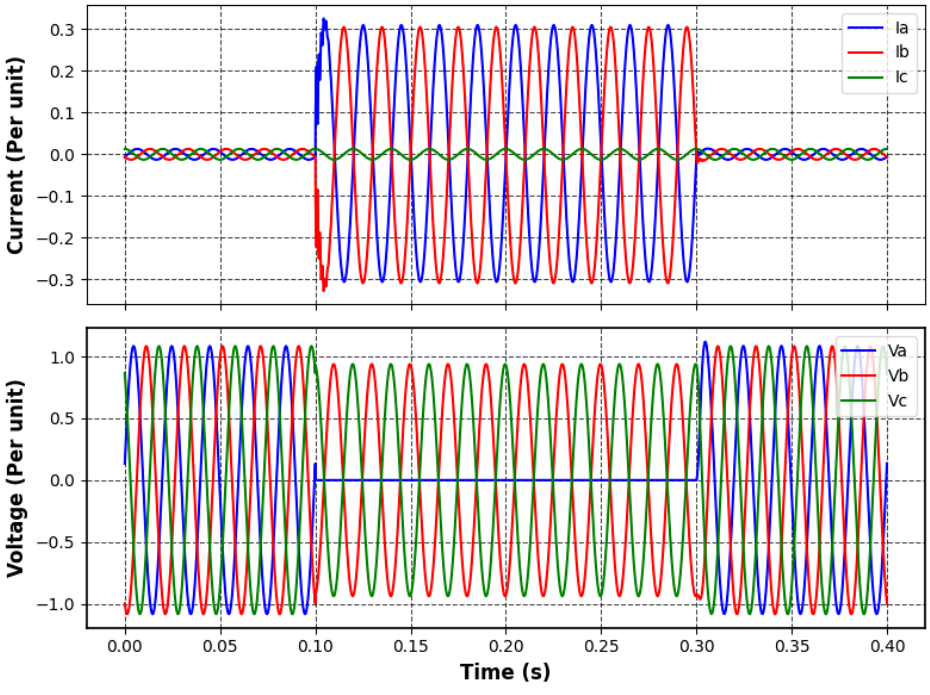
The transmission-line configuration used in this study consists of three interconnected grids modeled in MATLAB/Simulink R2021a to evaluate fault dynamics and the effectiveness of the proposed detection and classification method. The HVTL is rated at 132 kV and 50 MW and is connected through two transmission lines, each 75 km long, resulting in a total length of 150 km. Transformer T1 increases the voltage from 11 kV to 132 kV at the sending end, while transformer T2 reduces the voltage at the receiving end. The electrical parameters used in the simulation are summarized in Table 1.

The measured three-phase voltages and currents (Vabc, Iabc) were extracted from MATLAB and used as inputs to the proposed modeling framework. To evaluate the Simulink model used for fault-data generation, the grid voltage and current waveforms were recorded under different fault conditions, as shown in Figure 4. Figure 4(a) presents the voltage and current waveforms during a line-to-line fault between phases A and B. In this case, a direct connection occurs between the two phases without involving the ground, resulting in a sharp increase in current in both affected phases and a substantial voltage drop because the fault current is shared between them. This fault is classified as unsymmetrical because it does not affect all three phases equally.

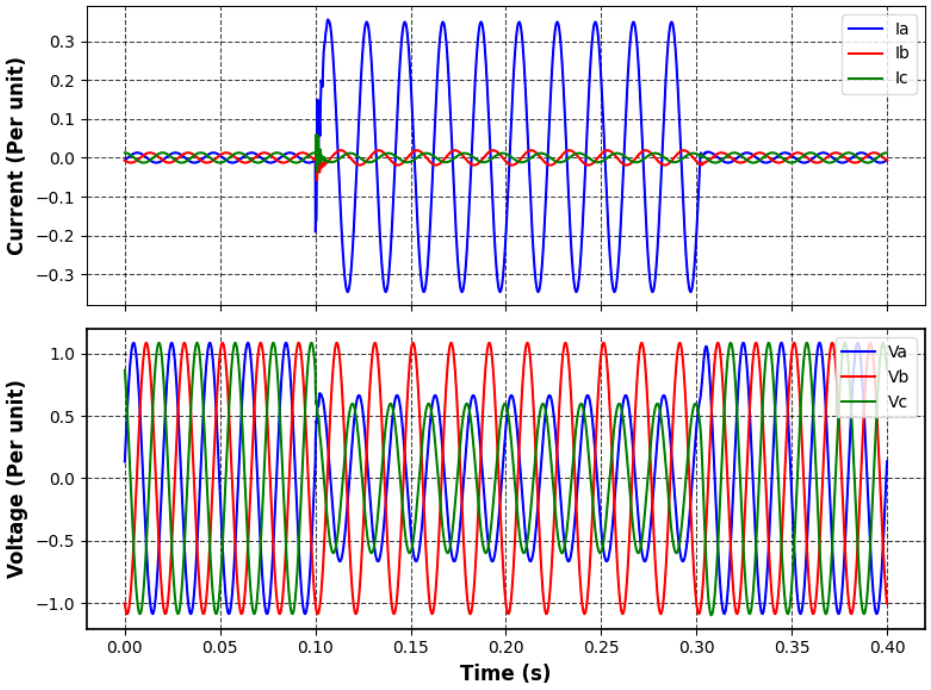
Figure 4(b) illustrates the voltage and current waveforms during a single line-to-ground fault. In this case, the voltage in phase A decreases to nearly zero, while the current in the same phase increases significantly because of the low-resistance path to ground. Practical studies indicate that line-to-ground (L–G) faults account for approximately 70–80% of all transmission-line faults, whereas line-to-line (L–L) faults represent approximately 10–15% [21].

Table 1. Transmission lines model parameters

Category Name	Quantity	Model
Overall system	Frequency (HZ)	50
	Voltage (kV)	132
Generator	Phase-to-phase RMS Voltage (kV)	132
	X/R ratio	7
	3-phase short-circuit level (VA)	50 e^{6}
Load	Active Power (W)	10 e^{4}
Transmission Line	Resistance (Ω/km) – Positive Sequence	0.0132
	Inductance (H/km) – Positive Sequence	0.9337e^{-3}
	Capacitance (F/km) – Positive Sequence	12.74e^{-9}
	Resistance (Ω/km) – Zero Sequence	0.132
	Inductance (H/km) – Zero Sequence	4.1264e^{-3}
	Capacitance (F/km) – Zero Sequence	7.751e^{-9}
Transformers	Transformer1	11 kV/132kV
	Transformer2	132kV/11kV
	Three-phase rated power (VA)	100 e^{6}
	X(p.u.)	0.05
Three Phase Fault	Fault resistance Ron (Ohm)	0.001



(a)



(b)

Figure 4. Voltage and current waveforms during fault conditions a)LL-AB fualt b)LG-AG fualt

4. Mathematical Modeling of Proposed Work

In this section, the mathematical modeling of each algorithm that is utilized in this work is divided into four deep learning models.

4.1. Long Short-Term Memory

Long Short-Term Memory (LSTM) is a specialized type of Recurrent Neural Network (RNN) designed to overcome the limitations of traditional RNNs in capturing long-term dependencies in sequential fault signals, making it effective for detecting high-resistance faults. Figure 5 illustrates the architecture of an LSTM unit. The LSTM unit consists of four main components: a memory cell (ct), a forget gate (ft), an input gate (it), and an output gate (ot). The memory cell serves as a long-term storage mechanism that retains important information over time. The input gate determines how much new information should enter the memory cell, while the forget gate controls which parts of the previous state should be discarded by assigning values between 0 (forget) and 1 (retain). Finally, the output gate determines which information is passed to the output [26]. For any given input, the sigmoid activation function, (\sigma), generates values between 0 and 1, as expressed in Equation 1, while Equation 2 defines the forget gate:

$$\sigma\,(x) = \frac{1}{1+e^{-x}} \tag{1}$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

where, W_f and b_f are stand for the forget gate's weight matrices and bias vector parameters, respectively. In contrast, h_{t-1} is represented the hidden state and x_t is represented the current input.

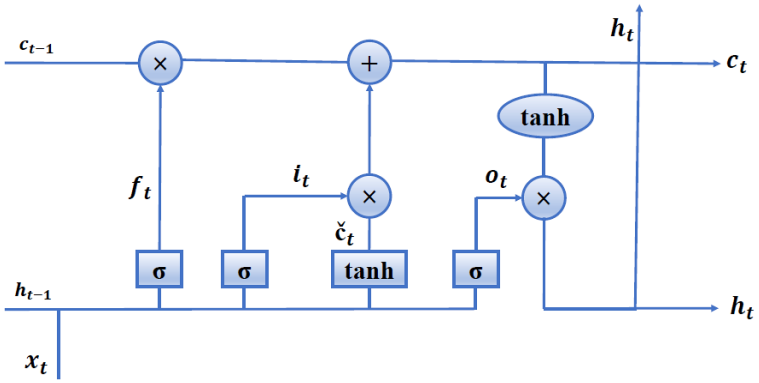


Figure 5. Architecture of LSTM unit

While Equation 3 describes a $tanh$ layer, and Equations 4 and 5 describe the input gate's constituent parts, including the weight matrices i and bi and the candidate cell state parameters candidate cell weight matrix Wc, candidate Cell State $\tilde{C}_t$, bias vector of input gate b_i, and bias vector b_c [26].

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{5}$$

The LSTM unit's output is controlled by the output gate (o_t) in Equation 6,

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

where the parameters W_o and b_o stand for learnt weight matrices and bias vectors during training, respectively [26]. The cell state is finally updated by the forget gate and input gates, as indicated in Equation 7. In other side, the hidden state h_t of a LSTM unit output is then updated by the updated cell state and the output gate, accordantly, as indicated in Equation 8 [26].

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \tag{7}$$

$$h_t = o_t * \tanh(c_t) \tag{8}$$

4.2. A Gated Recurrent Unit

The Gated Recurrent Unit (GRU) represents a novel RNN architecture, which was initially introduced in 2014. This architecture signifies a significant enhancement over the LSTM model due to its simplifying structure with maintaining comparable performance [27]. The inclusion of two distinct gates characterizes the GRU: the reset gate and the update gate, in addition to a singular hidden unit, which results in a reduced number of structural parameters relative to the triadic gates found in the LSTM. The update gate governs the degree to which information from the preceding state is preserved within the current state. Conversely, the reset gate dictates the extent to which the current state assimilates information from prior states. The fundamental configuration of the GRU is illustrated in Figure 6, wherein "*X*" denotes the matrix multiplication operation, "σ" and "tanh" represent activation functions, and "1−" indicates that the data propagated forward through the connection is represented by 1−zt [28].

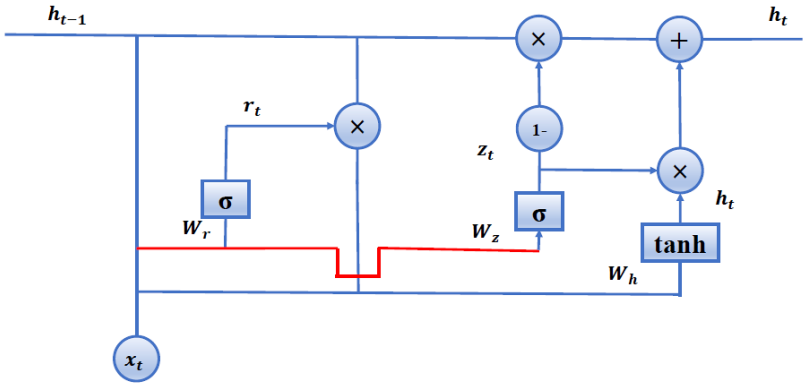


Figure 6. Unit structure of the GRU

In the equations below, z_t and r_t represent the update gate and the reset gate respectively; x_t is the input; W_z, W_r, and $W_{\tilde{h}}$ represent the weight matrices for the update gate, reset gate, and candidate hidden state, respectively; and h_t is the output of the hidden unit. h_t is calculated as follows [28]:

$$z_t = \sigma(W_z \cdot [x_t, h_{t-1}]) \tag{9}$$

$$r_t = \sigma(W_r \cdot [x_t, h_{t-1}]) \tag{10}$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [x_t, r_t \times h_{t-1}]) \tag{11}$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \tag{12}$$

4.3. A Harris Hawks Optimizer

The Harris Hawks Optimizer (HHO) is a recently proposed population-based optimization algorithm inspired by the cooperative hunting strategies of Harris hawks and developed by Heidari et al. [29]. The positions of the hawks represent potential solutions to the optimization problem, while the optimal solution corresponds to the target location under different conditions. In general, HHO consists of three main phases: exploration, the transition from exploration to exploitation, and exploitation [30]. The schematic representation of the HHO search mechanism is presented in Figure 7.

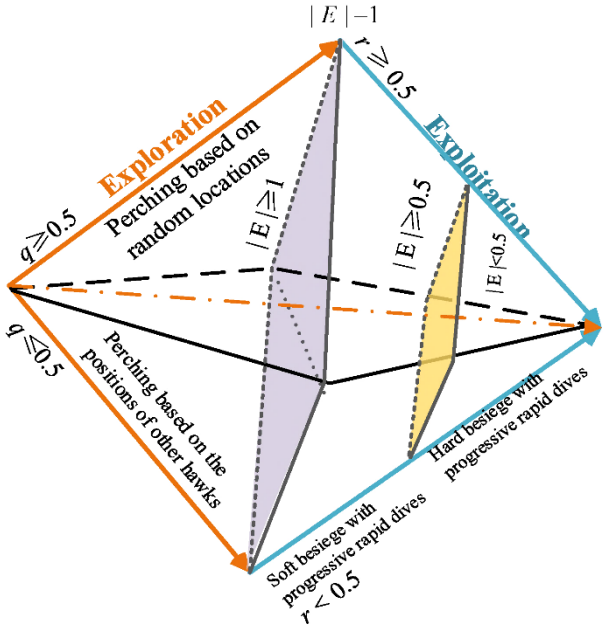


Figure 7. Schematic diagram of the HHO [31]

i. Exploration Phase

In the exploration phase of the HHO, Harris's hawks employ two strategies: they monitor and hunt prey in various random locations. The mathematical expression is as follows [31]:

$$X_i(k+1) = \begin{cases} X_{\text{rand}}(k) - r_1 \cdot |X_{\text{rand}}(k) - 2r_2 \cdot X_i(k)| & ,q \geq 0.5 \\ [X_{\text{prey}}(k) - X_m(k)] - r_3[Lb + r_4 \cdot (Ub - Lb)] & ,q < 0.5 \end{cases} \quad (13)$$

where, r_1, r_2, r_3, r_4, and q are stochastic numbers in the interval [0, 1], X_i, X_{prey} and $X_{rand}(t)$ denote the current position, best position and randomly chosen hawk location, respectively, and its mathematical expression is as follows [31]:

$$X_m(k) = \frac{1}{N}\sum_{i=1}^{N} X_i\,(k) \quad (14)$$

where, $X_i(t)$ refers to each hawk position and N is the maximum iteration number in training state [30].

ii. Transition from Exploration to Exploitation

To make the shift from exploration to exploitation easier, the HHO presents a model of the prey escape energy function, E. Whereas exploitation happens when $|E| < 1$, exploration happens when $|E| \geq 1$. This model is defined by the appropriate mathematical equation [31]:

$$E = 2E_0 \cdot (1 - \frac{k}{K}) \quad (15)$$

where, E_o and E are the initial energy and the escaping energy, respectively.

iii. Exploitation Phase

This phase is divided into three parts; (a). Soft besiege occurs $r \geq 0.5$ and $|E| \geq 0.5$ the mathematical expression is as follows [31]:

$$X_i(k+1) = \begin{Bmatrix} \Delta X(k) - E\left|J \cdot X_{\text{prey}}(k) - X_i(k)\right| \\ \Delta X(k) = X_{\text{prey}}(k) - X_i(k) \\ J = 2 \cdot (1 - r_5) \end{Bmatrix} \quad (16)$$

The position vector difference in each iteration is captured by ΔX(k), where r_5 represents a random number between 0 and 1 and J denotes the prey's jumping intensity during escape .While (b) is represented the hard besiege occurs when $r \geq 0.5$ and $|E| < 0.5$. This means the escape energy is very low because the prey has been on the run for a long time [31].

$$X_i(k+1) = X_{\text{prey}}(k) - E|\Delta X(k)| \tag{17}$$

Next, *(c)* is proposed based on the soft besiege with progressive rapid dives occurs when $r<0.5$ and $|E| \geq 0.5$. The following is the mathematical expression for this phase:

$$Y = X_{\text{prey}}(k) - E|J.X_{\text{prey}}(k) - X_i(k)| \tag{18}$$

The next hawk move is obtained by Equation 19:

$$Z = Y + S \times Lf(D) \tag{19}$$

where, *D* is to dimension; *S* is a random vector, and *Lf* is the Levy flight function, which may be written like this:

$$\text{Lf}(x) = 0.01 \cdot \frac{u \times \sigma}{|v|^{1/\beta}} \tag{20}$$

$$\sigma = \left(\frac{\sin(\pi\beta/2)\cdot\Gamma(1+\beta)}{\beta\cdot 2^{\left(\frac{\beta-1}{2}\right)}\cdot\Gamma\left(\frac{1+\beta}{2}\right)}\right)^{\frac{1}{\beta}} \tag{21}$$

where, β is a constant, which in this study is considered to be 1.5; u and v are random variables that were randomly acquired in [0, 1]. As a result, the following is the mathematical depiction of the Hawks' final position update during the soft siege phase [31]:

$$X_i(k+1) = \begin{cases} Y & \text{if } F(Y) < F\left(X_i(k)\right) \\ Z & \text{if } F(Z) < F\left(X_i(k)\right) \end{cases} \tag{22}$$

where, *F (·)* denotes the fitness function. Finally, part (*d*) is designed for hard besiege with progressive, rapid dives occurs when $r < 0.5$ and $|E|<0.5$. In this case, the location vectors (Y) and (Z) are calculated using Equations 23 and 24. Subsequently, the hawk positions are updated using the same position update rule given in Equation 22 [31]:

$$Y = X_{\text{prey}}(k) - E\left|J \cdot X_{\text{prey}}(k) - X_m(k)\right| \tag{23}$$

$$Z = Y + S \cdot \text{Lf}(D) \tag{24}$$

where, *Xm(k)* is the population's average position vector at iteration *k*.

4.4. Performance Metrics

In final stage, fitness functions are utilized to evaluate the performance of the proposed models, including; Precision, Recall, and F1-score. The following is a mathematical definition of these fitness metrics [32]. The Precision function quantifies the fraction of correctly identified positive instances out of all instances predicted as positive. A low number of false alarms is indicated by good accuracy, as presented in Equation 25:

$$\text{Precision} = \frac{\text{TP}}{\text{FP+TP}} \tag{25}$$

where an accurate forecast of the positive class is called a True Positive (TP), while an inaccurate result of the positive class is called a False Positive (FP). Then, the model's recall is employed, also known as sensitivity, that gauges how well it can detect every real positive case. Equation 26 provides the proportion of true positives that were correctly detected:

$$\text{Recall} = \frac{\text{TP}}{\text{FN+TP}} \tag{26}$$

When the model is unable to predict the positive class, it produces a False Negative (FN). Next, The F1-score is companied to compute the harmonic mean of prediction model, which is defined in Equation 27 and offers a single metric that strikes a balance between recall and precision:

$$\text{F1\_score} = \frac{2\times(\text{precision}\times\text{recall})}{(\text{precision}+\text{recall})} \tag{27}$$

Next, Equation 28 is applied to calculate the accuracy of the prediction model that is refer to the total percentage of accurate predictions the model masking across all classes;

$$\%\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Samples}} \tag{28}$$

Lastly, Mean Squared Error (MSE) is used to calculate the average squared difference between actual and predicted values, where better model performance is indicated by a lower MSE [33].

$$MSE = \frac{1}{M}\sum_{m=1}^{M}\left(y_p - y_c\right)^2 \quad (29)$$

where, m is an integer between 1 and M, where M is the total number of data points in the test set, and y_c and y_p are the test set's actual and predicted values, respectively. This fitness function is useful to evaluate the training state after find the major parameters of proposed model.

5. Proposed LSTM-GRU-HHO Model

The proposed methodology aims to develop an improved hybrid deep learning model for detecting and classifying faults in HVTLs by combining two neural network stages, namely LSTM and GRU, and enhancing the model's performance using the HHO algorithm. The primary role of the LSTM is to retain long-term information in sequential data, thereby supporting fault detection and enabling the model to understand fault evolution over time through its substantial memory capacity. In contrast, the GRU enables faster learning and reduces training time because of its simpler structure. By integrating LSTM and GRU layers, the hybrid model simultaneously benefits from the information-retention capability of the LSTM and the computational efficiency of the GRU. Consequently, this architecture improves accuracy, stability, and convergence speed when detecting and classifying both symmetrical and asymmetrical faults.

The model's performance is further enhanced using the HHO algorithm to determine the optimal values of the LSTM and GRU parameters, including the number of neurons, learning rate, and dropout rate. The strength of HHO lies in its effective balance between exploration and exploitation, which improves convergence behavior and minimizes the likelihood of the model becoming trapped in local optima. Moreover, HHO demonstrates superior computational efficiency compared with traditional optimization methods such as Particle Swarm Optimization (PSO), Genetic Algorithms (GA), and Bayesian Optimization.

Initially, the data were loaded and divided into two groups. The first group represented the fault-detection data and included the current and voltage measurements for each phase, together with an output value of 0 or 1, representing a healthy or faulty condition, respectively. The second group represented the fault-classification data and also included the current and voltage measurements for each phase, along with four columns used to identify the fault types. A total of 11 fault-condition categories were considered, including 10 faulty categories and one healthy category. These categories comprised phase-to-ground faults (AG, BG, and CG), line-to-line faults (AB, AC, and BC), double line-to-ground faults (ABG, ACG, and BCG), and three-phase faults (ABC), as shown in Figure 8. Thus, a total of 44,011 cases were analyzed. The dataset was perfectly balanced, with each fault category containing an equal number of samples generated under various operating conditions in the MATLAB/Simulink environment. This balancing strategy was intended to prevent model bias toward specific categories and enhance classification fairness. The data were also divided into training and testing sets to minimize the risk of overfitting.

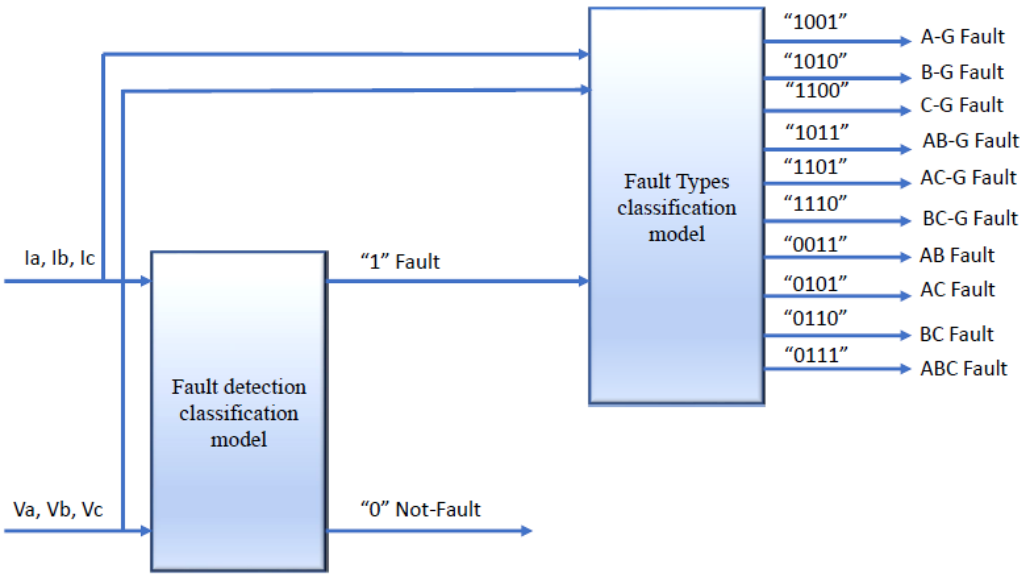


Figure 8. The methodology of faults description

Fault data are often noisy and imbalanced; therefore, in this research, it is essential to preprocess the input data before using it for training and testing, such as normalization, to ensure that all features are on a symmetric scale, resulting in enhancing the training stability. This is achieved using a minimum-maximum scale, which relates the original data to a fixed range between 0 and 1, as defined in Equation 30:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{30}$$

where, x is the original value, x' is the normalized value, min(x) is the minimum value, and max(x) is the maximum value of the feature. The data were then divided for the training and testing sets. Hence, they were randomly divided into 80% training sets and 20% test sets to ensure a fair evaluation of the model's performance on previously unseen data.

After data preparation, a population of Harris hawks (h_hawks) is generated, with each hawk representing a set of model parameters to be optimized: the number of neurons in the LSTM layer, the number of neurons in the GRU layer, the dropout rate, and the learning rate. A fitness value is then calculated for each hawk to represent the model's performance according to a set of evaluation metrics, including precision, recall, and F1-score. The hawk with the best performance is considered the current optimal solution.

The hawk positions are subsequently updated based on the escape energy value (E), which determines whether the algorithm is in the exploration phase, where it searches for new solutions, or the exploitation phase, where it focuses on improving existing promising solutions. The values of (E) and the random number (r) control the search behavior. If (|E| < 1), the algorithm operates in the exploration phase. If (|E| < 1), it moves to the exploitation phase using different strategies, such as hard besiege, soft besiege, or progressive rapid dives. Each strategy represents a different method for improving result accuracy and accelerating convergence toward the optimal solution.

The positions of all hawks are then updated, the iteration counter is increased ($t = t + 1$), and the process continues until the maximum number of iterations ((max_itr)) is reached. In this way, the HHO algorithm optimizes the parameters of the LSTM and GRU models, including the learning rate, number of hidden units, and dropout rate. The values of the learning rate, number of neurons, and dropout rate before and after optimization are presented to highlight the effect of the HHO algorithm on improving model performance and adjusting its parameters more efficiently.

Finally, after the optimal parameter set has been identified using the HHO algorithm, the hybrid LSTM-GRU model is trained using these values. The LSTM layer captures long-term dependencies in the current and voltage signals, while the GRU layer captures short-term and rapidly changing patterns, thereby improving temporal accuracy and reducing training time. After training, the model is evaluated using the testing dataset, which was not used during training, to assess its performance under unseen conditions.

The evaluation is conducted using three fundamental statistical measures: precision, which determines the accuracy of fault predictions; recall, which measures the model's ability to detect actual faults; and F1-score, which represents the balance between precision and recall. The flowchart of the proposed model is illustrated in Figure 9.

In this study, MSE was used as the loss function because the model initially experienced difficulty identifying patterns due to the complexity of the data, which significantly affected the training process. This resulted in fluctuations in the MSE while the model learned the underlying structure of the data. As the number of training epochs increased, the model parameters gradually adapted to better fit the training data. As illustrated in Figure 10, MSE was evaluated during the FD stage before and after applying the HHO algorithm. Before optimization, the MSE reached 0.000426 at epoch 97. As shown in Figure 10(a), high MSE values were recorded during the initial training stage. As the number of epochs increased, the error gradually decreased and reached a steady state at epoch 97. Figure 10(b), which presents the MSE curve after applying the optimization algorithm, shows that the model reached a near-perfect steady state at epoch 83, with an MSE value of 0.000172. At this stage, the error values remained low, and the training and testing curves converged.

Figure 11 presents the mean squared error (MSE) for both the training and testing datasets during the FC stage before and after applying the HHO algorithm. The proposed model demonstrates stable training behavior and strong generalization capability under this scenario. Figure 11(a) shows the MSE curve before applying the optimization algorithm, which reached a value of 0.000341 at epoch 49. However, Figure 11(b), which presents the MSE after applying the optimization algorithm, shows that the error decreased further to 0.000250 at epoch 50. This improvement is attributed to the effective role of the optimization algorithm in tuning the model's hyperparameters.

These results clearly indicate that the proposed model reached a steady-state error while maintaining strong generalization capability. Overall, the findings demonstrate that the selected model architecture and hyperparameters, including the numbers of LSTM and GRU units, dropout rate, and learning rate, were appropriate and contributed to reliable and stable performance during the FC and FD stages in electrical power systems. In the following section, the optimized model is tested and compared under various scenarios to demonstrate its accuracy in diagnosing faults. Table 2 presents a comparative review of the proposed model parameters under conventional and optimized training conditions.

Table 2. The parameters of the LSTM and GRU methods optimized using the HHO algorithm

Parameters	Before the optimization algorithm	After the optimization algorithm
Learning Rate	0.001	0.003763
Dropout Rate	0.3	0.3286
Number of Neurons	LSTM (32), GRU (32)	LSTM (67), GRU (175)
(MSE) Detection)	0.000426	0.000172
MSE (Classification)	0.000341	0.000250

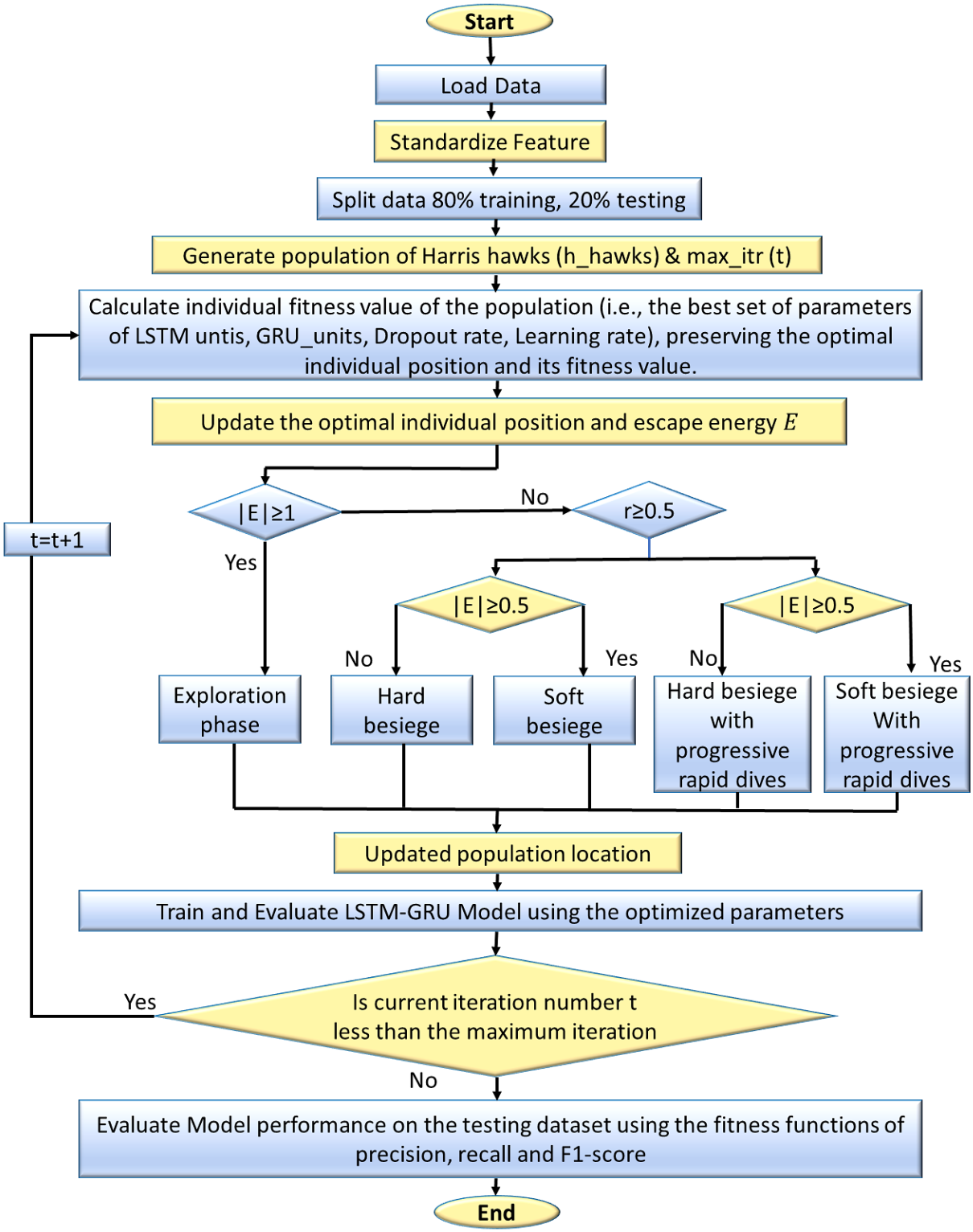
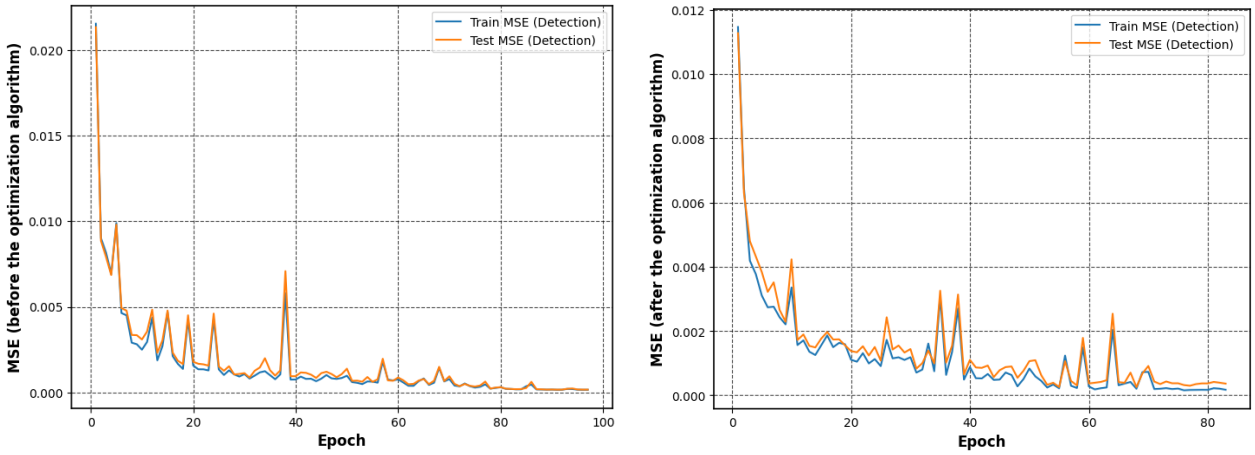
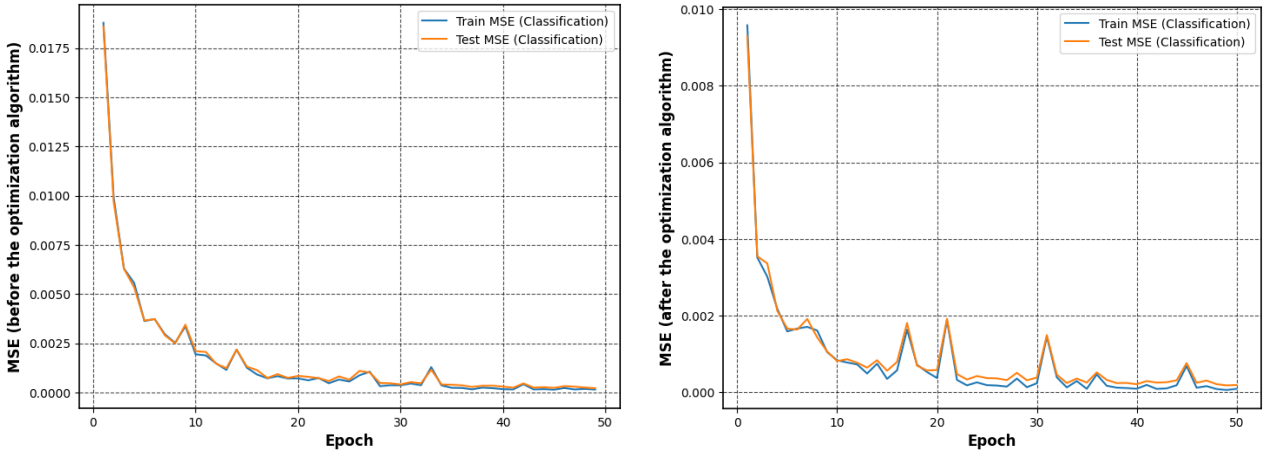


Figure 9. The flowchart of a proposed model



(a) (b)

Figure 10. Mean squared error performance for FD, a) before optimization b) after optimization



(a) (b)

Figure 11. Mean squared error performance for FC, a) before optimization b) after optimization

6. Results and Discussion

In this study, a novel deep learning framework combining GRU and LSTM models was employed for the detection and classification of transmission-line faults. Ten types of electrical faults were investigated, including AG, BG, CG, AB, BC, AC, ABG, BCG, ACG, and ABC, in addition to one fault-free category. Initially, the data were collected through a transmission-line simulation using MATLAB and then pre-processed in Python. Subsequently, the dataset was divided into training and testing subsets to validate the proposed model. During the training phase, the proposed model was trained using the pre-processed data to reduce the effects of noise. The HHO algorithm was then used to fine-tune the parameters of the LSTM and GRU models in order to achieve the highest possible prediction accuracy, as explained in Section 5. This section discusses the performance of the proposed model in comparison with conventional methods and related state-of-the-art studies on the detection and classification of transmission-line faults. Finally, the model performance was evaluated using appropriate statistical measures to verify the effectiveness of the proposed framework.

First, a confusion matrix was used to validate the performance of the proposed model by comparing the predicted and actual fault classes for both the FD and FC stages. Figure 12 presents the confusion matrices, where the rows represent the actual fault categories and the columns represent the categories predicted by the proposed model. The diagonal elements indicate correctly classified samples, while the off-diagonal elements represent misclassified samples. Figure 12(a) shows the confusion matrix for FD before optimization, whereas Figure 12(b) presents the confusion matrix for FD after optimization using the HHO algorithm. Figures 12(c) and 12(d) show the confusion matrices for FC before and after optimization, respectively. In Figure 12(c), the AG and No-Fault categories achieved the highest classification performance with the fewest misclassifications. In contrast, slight overlap was observed between the AC–CG and AC–C fault classes due to the similarity of their transient signal characteristics. In Figure 12(d), the AB, AG, BCG, LLL, ABG, and No-Fault classes achieved the best classification performance, with very

low misclassification rates, as demonstrated by the high diagonal values of 802, 832, 833, 809, 818, and 833 samples, respectively. These results highlight that the HHO algorithm improves the detection and classification performance and accuracy of the LSTM-GRU model.

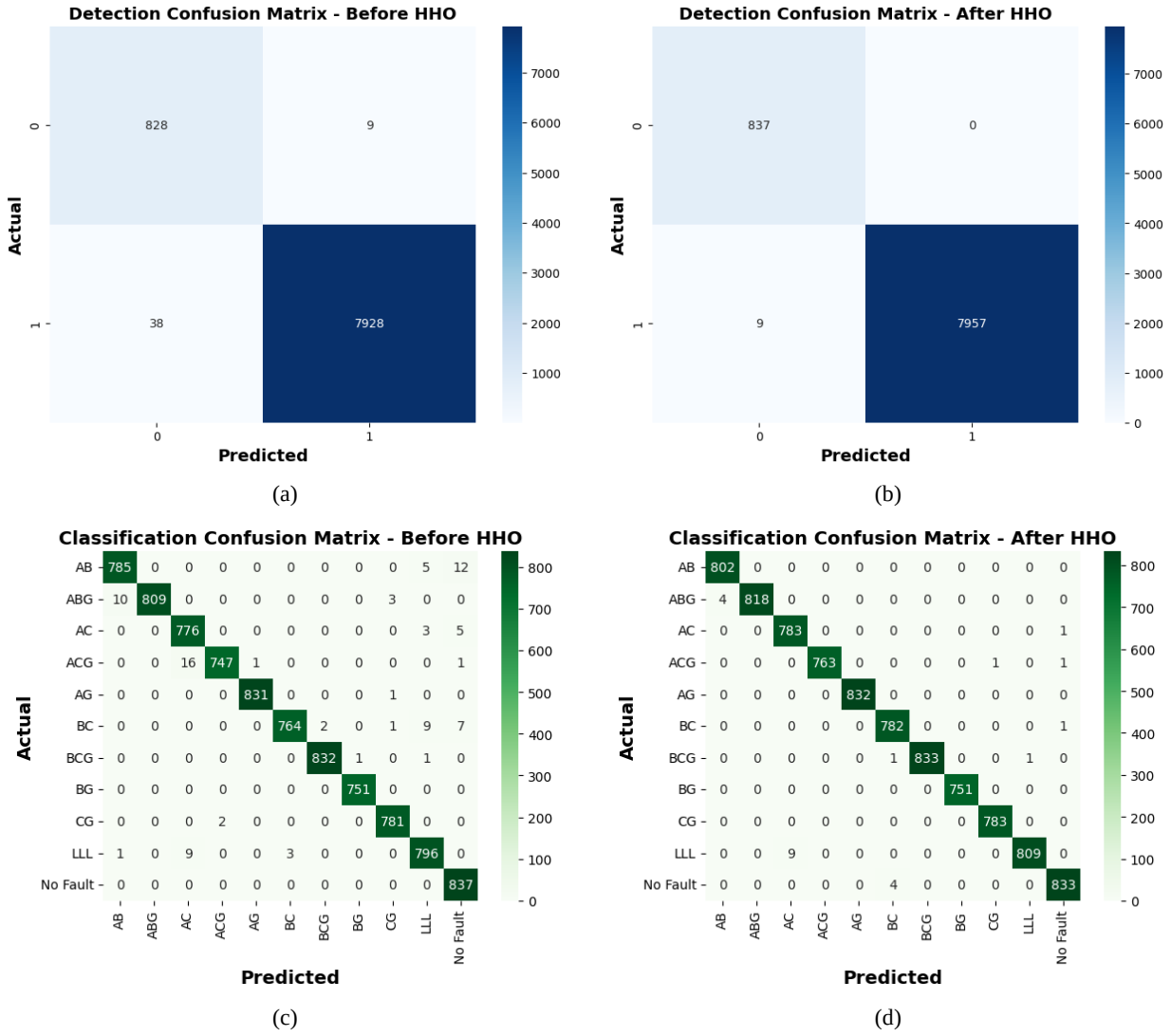


Figure 12. Confusion matrix a) fault detection before HHO, b) fault detection after HHO, c) fault classification before HHO, and d) fault classification after HHO

In the next step, the proposed model was evaluated using three widely used performance metrics: precision, recall, and F1-score. The obtained results showed high values across all evaluation metrics, demonstrating the effectiveness of the proposed framework in detecting and classifying different fault categories. Figure 13 presents a comparison between the optimized model and the conventional models based on the precision metric for the FD and FC stages. Figure 13(a) shows the results for the FD stage. The conventional LSTM model achieved a precision of 0.835, the GRU model achieved 0.9745, and the hybrid LSTM-GRU model achieved 0.9561, whereas the improved LSTM-GRU-HHO model achieved the highest precision of 0.9894 for the no-fault condition. For fault cases, the proposed LSTM-GRU-HHO model achieved an approximately 100% detection rate, outperforming all other methods and confirming its superior fault-detection capability compared with the conventional models. Using the same evaluation approach, the proposed model also demonstrated its ability to identify fault types by achieving precision values of approximately 99% across various fault categories, as shown in Figure 13(b).

Figure 14 presents the comparative evaluation of the FD and FC stages based on the recall metric. As shown in Figure 14(a), the proposed model and the conventional models achieved similar fault-detection performance, with recall values of approximately 99%. However, in the FC stage, the conventional LSTM model showed the weakest performance across different fault types, with recall values ranging from 87% for the BC fault to 99% for the LLL fault, as illustrated in Figure 14(b). At the same time, the hybrid model achieved the best performance in identifying

all fault types under different scenarios. This is because the LSTM-GRU-HHO model was trained using optimized layers and a reduced learning rate, thereby decreasing the processing time while maintaining accurate classification performance.

The final metric used to compare the proposed model with the conventional models was the F1-score. The proposed model demonstrated high and balanced performance in terms of precision and recall, confirming its effective and reliable classification capability under various operating conditions. As shown in Figure 15(a), the FD performance of the LSTM-GRU model and the proposed model was very similar, with both achieving values above 99%. However, in the FC stage, the LSTM-GRU model deviated from the optimal performance and achieved values below 97% compared with the proposed model and the other conventional methods, as shown in Figure 15(b). This result reflects the effectiveness of the HHO algorithm in improving the performance of the proposed model and significantly increasing its accuracy, as explained in the previous section.

Finally, Table 3 shows that the proposed LSTM-GRU-HHO model consistently outperformed the conventional LSTM and GRU models, as well as the hybrid LSTM-GRU model, across all fault categories. Consequently, the optimized hybrid model achieved the highest average F1-score of 99.82%, highlighting its strong generalization capability and reliability for practical fault-diagnosis applications.

Table 3. Comparison of average performance for all fault classes

Method	Avg. Precision (%)	Avg. Recall (%)	Avg. F1-score (%)
LSTM	95.05	95.2	95.24
GRU	99.03	99.01	99.03
LSTM–GRU	98.18	98.13	98.15
LSTM–GRU–HHO (Proposed)	99.79	99.86	99.82

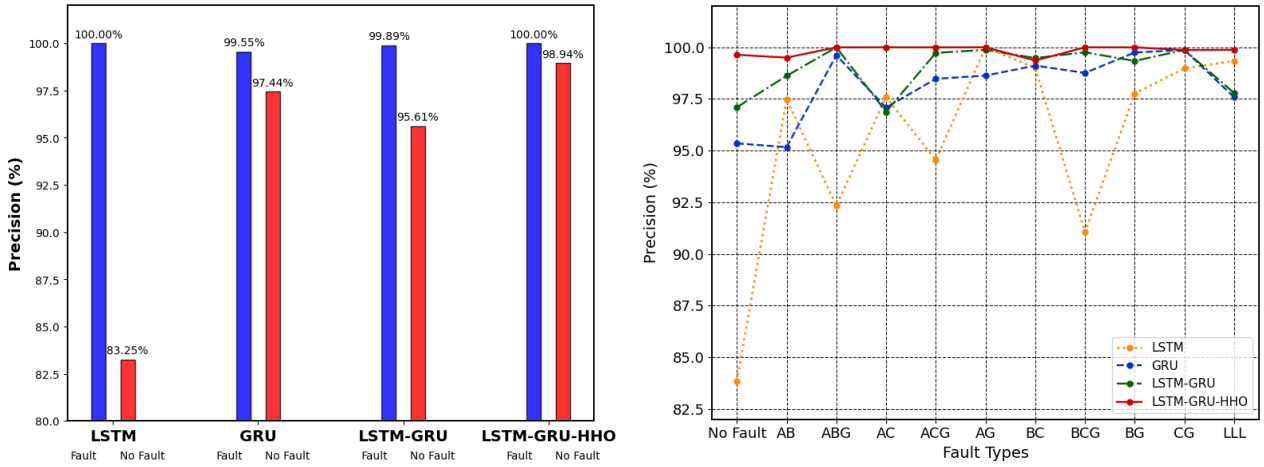
In the final step, the average accuracies of the FD and FC scenarios were calculated to compare the performance of the proposed model with that of the conventional models. As shown in Figures 16(a) and 16(b), the LSTM model exhibited the lowest average accuracies of 98.17% for FD and 95.26% for FC because of its limited modeling capability. In contrast, the GRU model demonstrated a noticeable improvement, achieving average accuracies of 99.36% for the FD scenario and 98.10% for the FC scenario. This improvement can be attributed to its simpler architecture and effective handling of temporal dependencies. The performance of the hybrid LSTM-GRU model was further enhanced, achieving accuracies of 99.47% in the FD stage and 98.93% in the FC stage. These results confirm that combining LSTM and GRU leverages the strengths of both architectures and provides more effective temporal feature representation. Through the integration of the HHO algorithm, the proposed hybrid model achieved the best performance, with the highest and most closely matched accuracies of 99.90% and 99.84% for the FD and FC stages, respectively. This is because the HHO algorithm effectively enhances the capability of the LSTM model and optimizes the main parameters of the GRU model.

Finally, the proposed hybrid method was compared with related state-of-the-art studies to demonstrate its superiority, achieving a fault-classification accuracy of 99.84%. It was compared with the approach proposed by Zhang et al. [18], who developed a data-driven method for predicting power-outage faults in electrical power systems by combining an LSTM model with an SVM technique. Their approach was evaluated using real fault records from the Wanjiang substation in Guangdong Province, China, after being trained on historical data. As a result, the model achieved a power-outage prediction accuracy of approximately 97.7%. Hossain et al. [19] employed an LSTM-AE model for fault detection and localization in three-phase power transmission lines. The model was trained using more than 50,000 fault samples and validated with 1,000 real fault records, achieving a classification accuracy of 98%. This performance surpassed that of conventional LSTM and 1D-CNN models, which achieved F1-scores of approximately 95.2% and 96.9%, respectively.

In summary, the comparative analysis presented in Table 4 demonstrates the advantages of the proposed LSTM-GRU-HHO framework over recent studies, particularly those conducted by Zhang et al. [18] and Hossain et al. [19]. The proposed model outperforms these earlier approaches in terms of classification accuracy and robustness because of its hybrid design, which combines GRU and LSTM deep learning networks with the HHO algorithm. This integration not only improves feature extraction but also optimizes the network parameters and accelerates convergence. Furthermore, the framework demonstrates high stability across various fault classes and operating conditions, distinguishing it from conventional standalone models and optimization methods.

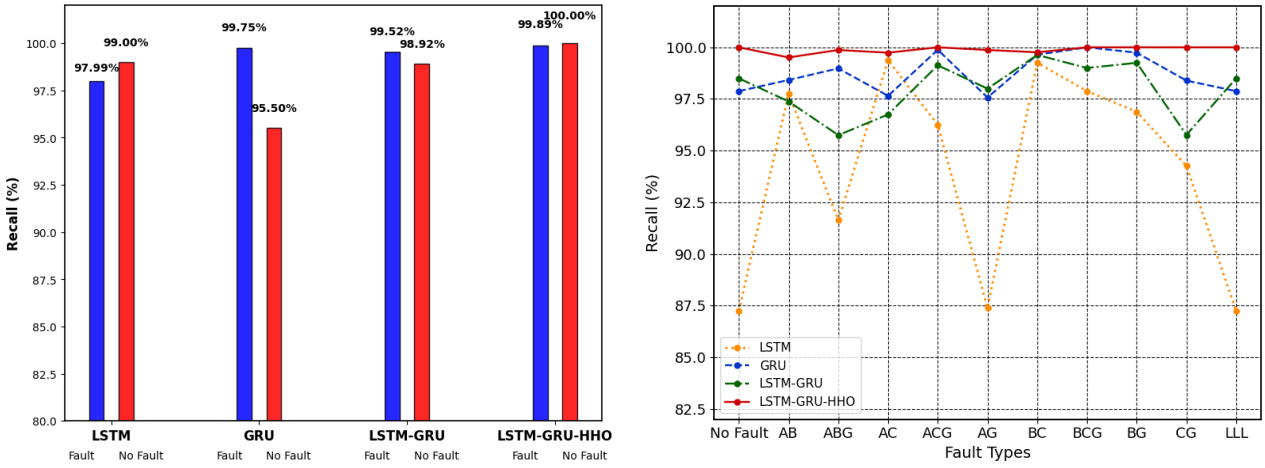
Table 4. A comparative summary of previous research

Method	Fault Analysis Type	Model	Optimization	Data Source	Learning Rate	Accuracy
Proposed Method	FD & FC	GRU-LSTM-HHO	HHO	Large-scale dataset of simulated signals	0.003763	99.84%
Zhang et al. [18]	Fault Prediction	LSTM + SVM	None	Real substation records (2011–2014)	0.001	97.7%
Hossain et al. [19]	FD & FL	LSTM Autoencoder	None	>50k simulated + 1k real faults	0.001	98.0%



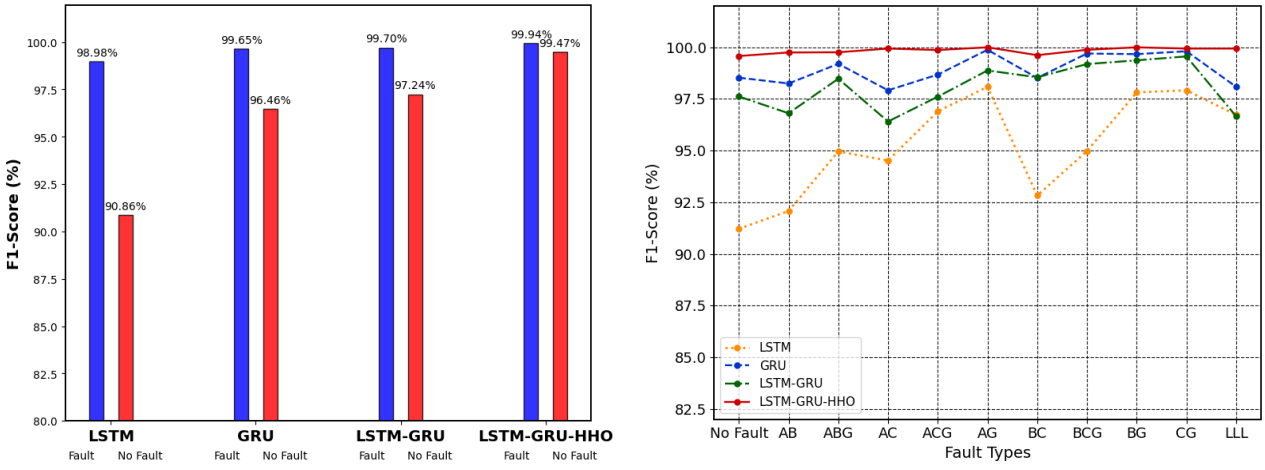
(a) (b)

Figure 13. The precision function compression based on (a) FD and (b) FC



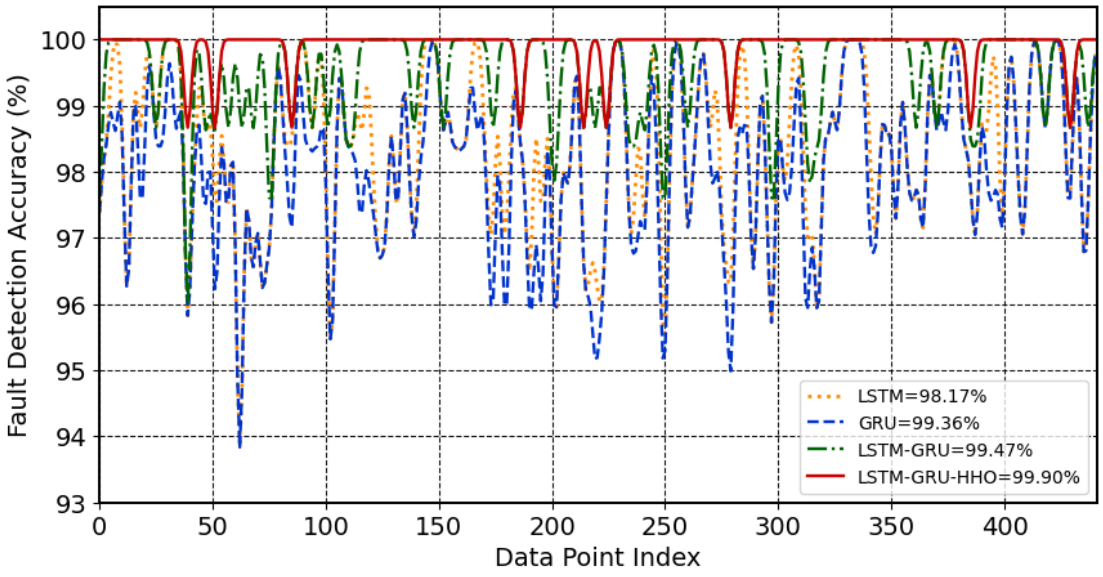
(a) (b)

Figure 14. The recall function compression based on (a) FD and (b) FC

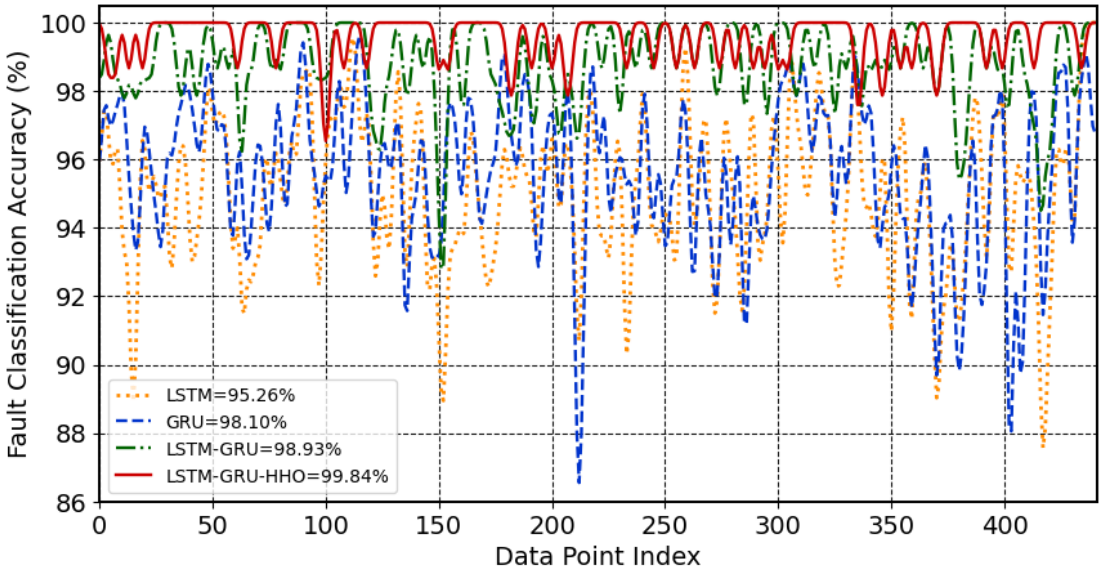


(a) (b)

Figure 15. The F1 function compression based on (a) FD and (b) FC



(a)



(b)

Figure 16. The accuracy function comparisons based on types of model stage: a) FC, b) FD

Finally, the results demonstrated that the LSTM-GRU-HHO model achieved exceptional performance in both the FD and FC stages for HVTLs. The numerical findings showed that the optimized model outperformed its unoptimized counterparts in terms of accuracy. Specifically, during the detection stage, the improvements in accuracy were 1.73% for LSTM, 0.54% for GRU, and 0.43% for LSTM-GRU. During the classification stage, the corresponding improvements increased to 4.58% for LSTM, 1.74% for GRU, and 0.91% for LSTM-GRU. These results highlight the effectiveness of the HHO algorithm in enhancing model performance through the careful selection of neuron numbers and the fine-tuning of network parameters.

After the LSTM-GRU-HHO model demonstrated a clear improvement in fault detection and classification compared with conventional models, the power system immediately transitioned to the operational response stage. In this stage, the affected section of the transmission line was isolated using circuit breakers to prevent the fault from propagating to the rest of the network, while the loads were redistributed to ensure continuity of the power supply. Simultaneously, fault information was transmitted to the control room, enabling the operation and maintenance teams to address the problem and restore service rapidly.

Furthermore, the proposed LSTM-GRU-HHO method is distinguished by its strong capability to process complex temporal data in electrical power systems. By combining LSTM and GRU, the model provides an effective balance between capturing short- and long-term patterns and responding to rapid changes in electrical signals. This capability facilitates the identification of differences among various fault conditions, including line-to-ground (LG), line-to-line

(LL), and line-to-line-to-ground (LLG) faults, thereby significantly improving classification accuracy. In addition, the incorporation of the HHO algorithm further optimizes and enhances the model by selecting the optimal numbers of layers and neurons and adjusting the training parameters to achieve the highest possible performance. This reduces errors and improves model stability across different fault types.

However, the LSTM-GRU-HHO method involves greater computational complexity than standalone LSTM and GRU models, resulting in longer training times because of the HHO optimization process. Increasing the number of hawks or optimization iterations also extends the training duration, requiring a balance between performance accuracy and available computational time. Despite these limitations, the proposed method is considered one of the most advanced and reliable techniques for developing FD-FC models for HVTLs because of its high accuracy.

7. Conclusion

In this study, a novel hybrid deep learning technique based on the LSTM-GRU-HHO algorithm is proposed to develop a fault-diagnosis model for an HVTL system. First, LSTM and GRU networks are combined to benefit from the simple structure of GRU and the large memory capacity of LSTM. The HHO algorithm is then employed to determine the optimal parameters of the hybrid LSTM-GRU network. Next, a 50 MW, 132 kV, and 150 km HVTL system is simulated using the MATLAB/Simulink environment to generate fault data. These data are divided into training and testing sets using Python for the proposed FD and FC frameworks under different fault conditions, including LG faults (AG, BG, and CG), LL faults (AB, AC, and BC), LLG faults (ABG, ACG, and BCG), and three-phase faults (LLL).

The study demonstrated that integrating LSTM with GRU significantly improves predictive accuracy in both the FD and FC stages. Furthermore, the HHO algorithm plays an important role in optimizing the model parameters. In the FD scenario, the proposed model achieved an exceptional accuracy of 99.90%, compared with 98.17% for LSTM, 99.36% for GRU, and 99.47% for LSTM-GRU. In the FC scenario, the model achieved an accuracy of 99.84%, compared with 95.26% for LSTM, 98.10% for GRU, and 98.93% for LSTM-GRU.

The results further demonstrate that the LSTM-GRU-HHO model provides an effective approach for developing intelligent and reliable protection systems for modern power networks because of its ability to handle multiple fault scenarios. Despite the positive results achieved by the proposed model, the present study relied on simulated datasets generated using MATLAB/Simulink under different operating conditions. Future research should focus on testing the model using real power-network data to verify its effectiveness in practical environments. Statistical significance analysis and confidence intervals should also be considered, and the proposed methodology may be extended to develop fault-location models.

8. Declarations

8.1. Author Contributions

Conceptualization, S.M.; methodology, A.M.; software, A.M.; validation, S.M. and S.D.; formal analysis, S.D.; investigation, S.D.; resources, M.K.; data curation, S.M.; writing—original draft preparation, A.M.; writing—review and editing, M.K. and S.D.; visualization, A.M.; supervision, S.D.; project administration, S.D. and M.K. All authors have read and agreed to the published version of the manuscript.

8.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

8.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

8.4. Institutional Review Board Statement

Not applicable.

8.5. Informed Consent Statement

Not applicable.

8.6. Declaration of Competing Interest

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

9. References

[1] Zhang, Q., Ma, W., Li, G., Ding, J., & Xie, M. (2022). Fault diagnosis of power grid based on variational mode decomposition and convolutional neural network. Electric Power Systems Research, 208, 107871. doi:10.1016/j.epsr.2022.107871.

[2] Moradzadeh, A., Teimourzadeh, H., Mohammadi-Ivatloo, B., & Pourhossein, K. (2022). Hybrid CNN-LSTM approaches for identification of type and locations of transmission line faults. International Journal of Electrical Power and Energy Systems, 135, 107563. doi:10.1016/j.ijepes.2021.107563.

[3] Aleem, S. A., Shahid, N., & Naqvi, I. H. (2015). Methodologies in power systems fault detection and diagnosis. Energy Systems, 6(1), 85–108. doi:10.1007/s12667-014-0129-1.

[4] Chen, K., Hu, J., Zhang, Y., Yu, Z., & He, J. (2020). Fault Location in Power Distribution Systems via Deep Graph Convolutional Networks. IEEE Journal on Selected Areas in Communications, 38(1), 119–131. doi:10.1109/JSAC.2019.2951964.

[5] Ajagekar, A., & You, F. (2021). Quantum computing based hybrid deep learning for fault diagnosis in electrical power systems. Applied Energy, 303, 117628. doi:10.1016/j.apenergy.2021.117628.

[6] Goni, M. O. F., Nahiduzzaman, M., Anower, M. S., Rahman, M. M., Islam, M. R., Ahsan, M., Haider, J., & Shahjalal, M. (2023). Fast and Accurate Fault Detection and Classification in Transmission Lines using Extreme Learning Machine. E-Prime - Advances in Electrical Engineering, Electronics and Energy, 3, 100107. doi:10.1016/j.prime.2023.100107.

[7] Chen, K., Huang, C., & He, J. (2016). Fault detection, classification and location for transmission lines and distribution systems: A review on the methods. High Voltage, 1(1), 25–33. doi:10.1049/hve.2016.0005.

[8] Harish, A., Prince, A., & Jayan, M. V. (2022). Fault Detection and Classification for Wide Area Backup Protection of Power Transmission Lines Using Weighted Extreme Learning Machine. IEEE Access, 10, 82407–82417. doi:10.1109/ACCESS.2022.3196769.

[9] Raichura, M. B., Chothani, N. G., & Patel, D. D. (2020). Identification of internal fault against external abnormalities in power transformer using hierarchical ensemble extreme learning machine technique. IET Science, Measurement and Technology, 14(1), 111–121. doi:10.1049/iet-smt.2019.0102.

[10] Chen, Y. Q., Fink, O., & Sansavini, G. (2018). Combined Fault Location and Classification for Power Transmission Lines Fault Diagnosis with Integrated Feature Extraction. IEEE Transactions on Industrial Electronics, 65(1), 561–569. doi:10.1109/TIE.2017.2721922.

[11] Akmaz, D., Mamiş, M. S., Arkan, M., & Tağluk, M. E. (2018). Transmission line fault location using traveling wave frequencies and extreme learning machine. Electric Power Systems Research, 155, 1–7. doi:10.1016/j.epsr.2017.09.019.

[12] Omar, A. M. S., Osman, M. K., Ibrahim, M. N., Hussain, Z., & Abidin, A. F. (2020). Fault classification on transmission line using LSTM network. Indonesian Journal of Electrical Engineering and Computer Science, 20(1), 231–238. doi:10.11591/ijeecs.v20.i1.pp231-238.

[13] Shukla, P. K., & Deepa, K. (2024). Deep learning techniques for transmission line fault classification – A comparative study. Ain Shams Engineering Journal, 15(2), 102427. doi:10.1016/j.asej.2023.102427.

[14] Minh, N. Q., Khiem, N. T., & Giang, V. H. (2024). Fault classification and localization in power transmission line based on machine learning and combined CNN-LSTM models. Energy Reports, 12, 5610–5622. doi:10.1016/j.egyr.2024.11.061.

[15] Altaie, A. S., Abderrahim, M., & Alkhazraji, A. A. (2024). Transmission Line Fault Classification Based on the Combination of Scaled Wavelet Scalograms and CNNs Using a One-Side Sensor for Data Collection. Sensors, 24(7), 2124. doi:10.3390/s24072124.

[16] Aker, E., Othman, M. L., Veerasamy, V., Bin Aris, I., Wahab, N. I. A., & Hizam, H. (2020). Fault detection and classification of shunt compensated transmission line using discrete wavelet transform and naive bayes classifier. Energies, 13(1), 243. doi:10.3390/en13010243.

[17] Leh, N. A. M., Zain, F. M., Muhammad, Z., Hamid, S. A., & Rosli, A. D. (2020). Fault Detection Method Using ANN for Power Transmission Line. Proceedings - 10th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2020, 79–84. doi:10.1109/ICCSCE50387.2020.9204921.

[18] Zhang, S., Wang, Y., Liu, M., & Bao, Z. (2017). Data-Based Line Trip Fault Prediction in Power Systems Using LSTM Networks and SVM. IEEE Access, 6, 7675–7686. doi:10.1109/ACCESS.2017.2785763.

[19] Hossain, M. I., Anonto, H. Z., Riyad, T., Shufian, A., Hossain, M. S., & Pathik, B. B. (2026). Adaptive fault diagnosis in power transmission lines using deep learning and LSTM autoencoders for enhancing grid reliability. International Journal of Electrical Power and Energy Systems, 174. doi:10.1016/j.ijepes.2025.111458.

[20] Belagoune, S., Bali, N., Bakdi, A., Baadji, B., & Atif, K. (2021). Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems. Measurement: Journal of the International Measurement Confederation, 177, 109330. doi:10.1016/j.measurement.2021.109330.

[21] Glover, J. D., Sarma, M. S., Overbye, T. J., & Padhy, N. P. (2012). Power system analysis and design. Course Technology, Boston, United States.

[22] Al-Mistarehi, B., Shtayat, A., Imam, R., & Abdallah, A. (2025). An Automated Assessment Technique for Pavement Defects Using a Laser Scanner and Deep Machine Learning. Civil Engineering Journal, 11(3), 1088–1105. doi:10.28991/CEJ-2025-011-03-015.

[23] Prasad, A., Belwin Edward, J., & Ravi, K. (2018). A review on fault classification methodologies in power transmission systems: Part—I. Journal of Electrical Systems and Information Technology, 5(1), 48–60. doi:10.1016/j.jesit.2017.01.004.

[24] Yadav, A., & Dash, Y. (2014). An Overview of Transmission Line Protection by Artificial Neural Network: Fault Detection, Fault Classification, Fault Location, and Fault Direction Discrimination. Advances in Artificial Neural Systems, 2014, 1–20. doi:10.1155/2014/230382.

[25] Ji, L., Tian, X., Wei, Z., & Zhu, D. (2025). Intelligent fault diagnosis in power distribution networks using LSTM-DenseNet network. Electric Power Systems Research, 239, 111202. doi:10.1016/j.epsr.2024.111202.

[26] Tende, I. G. (2025). Solar Radiation Forecasting Using LSTM and PeepHole LSTM Deep Learning Models: A Case of Kishapu District, Tanzania. East African Journal of Information Technology, 8(1), 472–488. doi:10.37284/eajit.8.1.3613.

[27] Wang, Y., Liao, W., & Chang, Y. (2018). Gated recurrent unit network-based short-term photovoltaic forecasting. Energies, 11(8), 2163. doi:10.3390/en11082163.

[28] Liu, Z., Teng, S., & Wang, S. (2025). Optimized GRU with Self-Attention for Bearing Fault Diagnosis Using Bayesian Hyperparameter Tuning. Algorithms, 18(9), 1–18. doi:10.3390/a18090576.

[29] Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., & Chen, H. (2019). Harris hawks optimization: Algorithm and applications. Future Generation Computer Systems, 97, 849–872. doi:10.1016/j.future.2019.02.028.

[30] Hussien, A. G., Abualigah, L., Zitar, R. A., Hashim, F. A., Amin, M., Saber, A., Almotairi, K. H., & Gandomi, A. H. (2022). Recent Advances in Harris Hawks Optimization: A Comparative Study and Applications. Electronics (Switzerland), 11(12), 1919. doi:10.3390/electronics11121919.

[31] Tang, Z., Liu, S., Qin, H., Zhang, Y., Zhu, X., Chen, X., & Ren, P. (2025). Chaos-Enhanced Harris Hawks Optimizer for Cascade Reservoir Operation with Ecological Flow Similarity. Sustainability (Switzerland), 17(19), 8616. doi:10.3390/su17198616.

[32] Mulissa, Y. G., Li, W., Kumar, A., & Wang, L. (2025). A Hybrid Deep Learning Model for IoT Network Anomaly Detection. Conference Proceedings - IEEE SOUTHEASTCON, 1370–1375. doi:10.1109/SoutheastCon56624.2025.10971453.

[33] Mohammed, A. J., Al-Majidi, S. D., Al-Nussairi, M. K., Abbod, M. F., & Al-Raweshidy, H. S. (2022). Design of a Load Frequency Controller based on Artificial Neural Network for Single-Area Power System. 57th International Universities Power Engineering Conference: Big Data and Smart Grids, UPEC 2022 - Proceedings, 1–5. doi:10.1109/UPEC55022.2022.9917853.