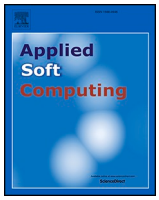
Applied Soft Computing Journal 170 (2025) 112667

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc



A data-driven approach to microgrid fault detection and classification using Taguchi-optimized CNNs and wavelet transform

Paul Arévalo [a,b], Antonio Cano [b], Olena Fedoseienko [c,d], Francisco Jurado [b,*]

[a] Department of Electrical Engineering, Electronics and Telecommunications (DEET), University of Cuenca, 010107 Balzay Campus, Cuenca, Azuay, Ecuador
[b] Department of Electrical Engineering, University of Jaen, Jaen, Spain
[c] Electric Power Engineering Department, Warsaw University of Technology, Poland
[d] Electric Power Transmission Department, National Technical University Kharkiv Politechnik Institute

HIGHLIGHTS

- Introduces "Taguchi-CNN" method for efficient fault detection in microgrid clusters.
- Achieves 99.13 % accuracy in less than 10 ms, exhibiting adaptability to diverse fault scenarios.
- Outperforms both conventional approaches and state-of-the-art models in fault detection.
- Recommends integrating the method into PMUs/IEDs and exploring real-time fault detection..

ARTICLE INFO

Keywords:
Fault detection and localization
Taguchi method
Wavelet
Convolutional neural network
Microgrid cluster

ABSTRACT

The integration of microgrids into the bulk power system introduces inherent uncertainties that challenge conventional protection systems, encompassing factors such as low fault currents, operational modes, penetration levels of renewable sources, load variations, and network topology. These uncertainties significantly impact the overall reliability of the electrical system. In the event of a fault occurrence within or external to the microgrid, swift disconnection from the primary grid is imperative. This disconnection is facilitated through the immediate operation of a static switch positioned proximate to the common coupling point. Such rapid action is essential to mitigate potential damages and expedite the restoration of electrical services. To ensure the delivery of reliable and high-quality energy to end consumers while alleviating stress on the utility grid, this paper introduces a novel methodology for the efficient detection, classification, and localization of faults in a microgrid cluster connected to the external grid. The proposed system addresses diverse irregular conditions, including conventional faults, high-impedance faults, islanding scenarios, and adverse events, covering several zones within the microgrid cluster and the external electrical grid. The proposed approach is based on a fusion of the Taguchi methodology and the discrete Wavelet transform. This combination enables the optimization of convolutional neural network training using scalograms generated from the fault signals. The results demonstrate the model's high performance, achieving 99.25 % accuracy in fault localization and 99.13 % in fault detection and classification, all within less than 10 ms. In comparison, traditional methods like support vector machine and decision trees require over 16 ms with lower accuracy, underscoring the superior speed and precision of the proposed approach.

1. Introduction

The evolution of modern electrical grids has incorporated clusters of microgrids (MG) to enhance system flexibility and resilience during the transition to renewable energy sources (RES). These interconnected MGs are crucial for the reliable and sustainable distribution of energy [1]. The cluster includes MGs with diverse RES and varying load profiles. Despite dedicated research on the protection of individual MGs, the security and reliability of MG clusters remain critical concerns [2–4]. A key challenge is the detection and characterization of high-impedance faults (HIF) in systems with multiple RES, presenting as disruptive

* Corresponding author.
E-mail address: fjurado@ujaen.es (F. Jurado).

https://doi.org/10.1016/j.asoc.2024.112667
Received 21 June 2024; Received in revised form 13 December 2024; Accepted 19 December 2024
Available online 25 December 2024

Nomenclature

Acronyms

AB	PhaseA-PhaseB
ABC	PhaseA-PhaseB-PhaseC
ABCG	PhaseA-PhaseB-PhaseC-Gound
AC	PhaseA-PhaseB
ACG	PhaseA-PhaseC-Ground
AG	PhaseA-Gound
BC	PhaseB-PhaseC
BCG	PhaseA-PhaseC-Ground
BG	PhaseB-Ground
CG	PhaseC-Ground
PV	Photovoltaic
BS	Battery Storage
W	Wind turbine
H_2	Hydrogen storage
HK	Hydrokinetic
ANN	Artificial Neural Networks
DWT	Discrete Wavelet Transform
FT	Fourier Transform
GTO	Gorilla Troops Optimization
GPR	Gaussian Process Regression

IED

Tanh	Hyperbolic Tangent
IoT	Internet of Things
HIF	High Impedance Faults
MG	MicroGrid
NNRBF	Neural Network Radial Basis Function
CCP	Common connection point
DL	Dropout Layer
PCQ-WT	Pseudo-Continuous Quadrature Wavelet Transform
PMU	Phasor Measurement Units
RES	Renewable energy system
ReLU	Rectified Linear Unit
STFT	Short-Time Fourier Transform
SMX	Softmax
WT	Wavelet Transform

Parameters and decisions variables

$x_{c,\ w}$	Current sign
ψ	Wavelet mother
a,b	Scale Factor and temporal displacement
$c_{j,k}$	Approximation coeficients
$d_{j,k}$	Approximation coeficients
$\phi_{j,k}$	Scaling Function
$\psi_{j,k}$	Discrete Wavelet
S_{DWT}	Scalogram discrete Wavelet
m_{Tg}	Factorial Combinations Factors
n_{Tg}	Factorial Combinations Levels
x_{Tg}	Number of experiments
y_{Tg}	Factors
z_{Tg}	Levels
α_{Tg}	Load level
β_{Tg}	Wind speed
γ_{Tg}	Short circuit capacity
δ_{Tg}	Solar irradiance
ε_{Tg}	River speed
ζ_{Tg}	Hydrogen flow
ξ_{Tg}	Ambient temperature
S_{CWT}	Scalogram Continous Wavelet
$\mathscr{I}_{sr,CNN}$	weighted input
Υ, Γ	convolution filters
$\varpi_{sr,CNN}$	activation of the characteristic in the position
$\omega_{\nu\rho}$	Weights
$\varsigma_{s,r}$	Biases
$\widehat{y}$	One-hot prediction
y	One-hot label vector

anomalies [5]. HIFs pose a threat to the stability and security of MG clusters [6]. Detecting faults, especially in islanded mode with low fault currents, is a challenge [7]. While sophisticated algorithms could consider all conditions, the computational time and effort would be critical. It is essential to develop an efficient fault detection method capable of identifying and locating different types of faults considering various operational conditions of RES. This is the central focus of this research.

In the literature, conventional protection methods employ relays and sensors to monitor electrical parameters in MG, such as short circuits and overloads, activating automatic switches. With the incorporation of the internet of things (IoT) and MG management systems, smart protection uses networks and sensors to monitor and collect data. Machine learning algorithms and artificial intelligence analyze the information to respond to faults and anticipate problems. In fault detection, especially HIF faults, specialized devices such as phasor measurement units (PMUs) and protection relays are required. These devices collect voltage and current data to identify HIF through algorithms. PMUs are crucial for real-time fault detection, allowing a quick response and disturbance management in the electrical system. For example, in [8], a robust state estimation approach is proposed for fault location using optimization task modeling with voltage and current measurements obtained from PMUs. This method considers errors in network parameters, improving the accuracy of the process. Additionally, in [9], a method is presented to identify the location of HIF in the electrical power distribution based on the estimation of fault impedance using synchronized voltage and current phasors of the third harmonic collected by μPMUs. In [10], the authors present a fault location algorithm for observability using PMUs in the presence or absence of zero-injection buses. Regarding specific devices, the device [11] that detects and recognizes HIF in 12/7 kV distribution networks uses two main real-time algorithms, monitoring the energies of non-fundamental components and detecting wide variations in the waveform of the signal and its spectral components.

Simulations provide a controlled environment to explore and validate fault detection methods without compromising the stability of real electrical systems. For simulations, selecting an appropriate model of HIF is crucial. In [12], the authors propose the enhanced Emanuel model to accurately simulate HIF under different working conditions. Furthermore, a field-tested model representing the two characteristics of HIF current, nonlinearity, and asymmetry, is presented in [5]. Study [13] suggests the use of two variable fault resistances to mimic the randomness of HIF faults. The use of techniques such as the Wavelet transform (WT), especially the discrete Wavelet transform (DWT), plays a fundamental role in extracting and analyzing relevant parameters from simulated fault signals [14]. The authors in [15] have explored different fault scenarios and have shown that the Wavelet packet transform is more robust in the presence of signal noise compared to the wavelet transform. Studies like [16,17] have demonstrated that this hybrid approach with DWT is reasonably accurate and presents minimal estimation error in determining fault location in power systems. Specifically, [18] proposes a fault detection method based on the DWT, using decomposition to analyze traveling wave signals in details and

approximations. The results indicate that this fault detection algorithm is reliable in its assessments of fault presence in the electrical system.

In addition to WT-based techniques, there has been a growing interest in leveraging the capabilities of convolutional neural networks (CNN) for fault detection and characterization. Recognized for their efficiency in image processing, CNNs have demonstrated notable effectiveness in handling scalograms generated by DWT, scalograms are visual representations of time-frequency information generated through WT analysis, capturing the transient characteristics of fault signals in a 2D matrix. This capability makes them instrumental for achieving accurate fault detection and localization within clusters of electrical MG. Experimental results from various studies, including [17,19], illustrate the success of this approach, achieving ultra-fast detection with a high accuracy rate, even in the presence of noise. In studies such as those presented in [20], a hybrid technique is proposed that combines DWT, artificial neural networks (ANN), and Gaussian process regression (GPR) for HIF diagnosis. The results of these studies indicate that this hybrid approach is reasonably accurate and presents minimal estimation error in determining fault location. Furthermore, in another study [21], an algorithm based on a neural network with a radial basis function (NNRBF) is suggested for fault detection. Additionally, CNN-based models optimized using the gorilla troop optimization (GTO) have shown promising accuracy rates [18].

Despite extensive research on rapid fault detection, classification, and localization in MG-based electrical systems, several critical gaps remain unaddressed. Most conventional methods lack adaptive learning mechanisms, relying on empirical data alone, which reduces detection accuracy when key parameters—such as fault resistance, renewable source power, or HIF types—fluctuate [2,4–6,9]. This dependence on fixed data limits the flexibility of these systems, making them less effective in dynamically adjusting to real-world variations that are common in MG. In addition, many AI-based models, including ANNs and Decision Trees, require large training datasets to reach acceptable accuracy, which greatly increases computational requirements [8,17–19]. Such computational demands restrict these models' feasibility for real-time applications in rapidly changing environments like MGs, where fast processing is critical to ensure protection and resilience.

Furthermore, a large portion of existing fault detection research is focused on transmission systems rather than distribution systems, creating a gap in solutions tailored for MG clusters. MG clusters present specific challenges: fault currents are significantly lower than in the main grid, which makes conventional protection systems less effective in this context. Some studies [10,18,22,23] recommend deploying multiple sensors to improve detection and localization accuracy across distributed MG networks. However, the use of numerous sensors raises issues of system vulnerability and cost; the failure of any one sensor could reduce the effectiveness of the protection system and increase operational expenses. Then, the variability of fault currents under changing power levels—driven by renewable resource availability, especially during islanded operations—presents another challenge. Addressing this variability requires fault detection methods that can effectively adapt to different load and generation scenarios in MG clusters [13–15]. Failing to accommodate this variability can compromise system reliability, highlighting the need for approaches like the Taguchi-CNN that provide flexible and robust detection across diverse operating conditions.

Given these identified limitations in conventional methods, this article introduces the innovative Taguchi-CNN method. This approach uniquely combines the capabilities of DWT [19,23,24], data optimization using the Taguchi method [20,25], and pattern recognition through CNN [13,18,19] to address the persistent challenges in MG-based fault detection systems. Specifically, it offers an optimized, adaptable solution to enhance detection accuracy under varying operating conditions, overcoming the drawbacks of static, data-heavy approaches.

The Taguchi method enables a reduction in the volume of input data required without compromising accuracy by using orthogonal arrays for optimization. This statistical optimization approach is crucial for managing data efficiently while capturing system variability through minimal experimental runs. By integrating DWT, the proposed method achieves effective multiscale decomposition, extracting key features from fault signals to enhance detection accuracy.

The primary goal of this method is to improve detection, classification, and localization capabilities across a wide range of fault types in MG and distribution systems, including HIF, islanding events, and variations in renewable generation. Unlike conventional methods, this approach is resilient under diverse operational scenarios, ensuring reliable performance in complex MG environments.

The key contributions of this article are summarized as follows:

- Development of a multiscale decomposition approach using CWT and DWT near the static switch at the point of common coupling (PCC), eliminating the need for complex sensors and integrating easily into existing protection relays or intelligent electronic devices (IEDs) with PMU-based monitoring.
- Implementation of a fault pattern recognition system combining CNN and DWT, reducing reliance on peak maxima and providing greater robustness in fault pattern identification compared to conventional methods.
- Efficient use of an orthogonal array through the Taguchi method to create a concise, representative training dataset for the CNN. This approach accommodates renewable source variations and load conditions, minimizing the need for extensive random data and significantly reducing CNN training time.
- A comprehensive approach addressing essential functions in fault detection, classification, and localization, covering HIF anomalies, islanding, and switching events both within and beyond the MG cluster, ensuring full and effective protection for the interconnected electrical system.

The remaining structure of this paper is organized as follows. Section 2 presents the proposed methodology in the study, followed by the introduction of the innovative Taguchi-CNN in Section 3. Section 4 thoroughly addresses the case study, while in Section 5, the results are presented and discussed. Finally, Section 6 summarizes the conclusions derived from this paper.

2. Methodology

The proposed methodology for fault detection in a MG cluster is illustrated in Fig. 1. Waveforms of fault currents in different zones of the MG cluster are generated through simulations conducted in DigSILENT Power Factory software. These waveforms are measured at the PCC, located in proximity to the static switch. Subsequently, employing the Taguchi method, an optimization of the quantity of simulation-derived data is performed. This process aims to preserve the accuracy and robustness of the data, which are crucial for generating scalograms using the WT. The resulting scalograms serve as a training set for a CNN. The primary function of this CNN is to recognize patterns that enable the identification not only of the type, phase, and location of faults in the MG cluster but has also been extended to include the detection of HIF, line opening and closing events, and islanding situations.

2.1. HIF modeling

High Impedance Faults represent a complex phenomenon characterized by nonlinear behavior and the generation of electric arcs, often challenging traditional protection systems due to their low current magnitude. HIF exhibits three distinctive traits—accumulation, nonlinearity, and asymmetry—manifesting as a gradual increase in current over cycles, the presence of odd harmonics, and asymmetric characteristics [11]. While modern technology enables HIF detection using micro-PMUs and high-resolution IEDs, implementing such solutions

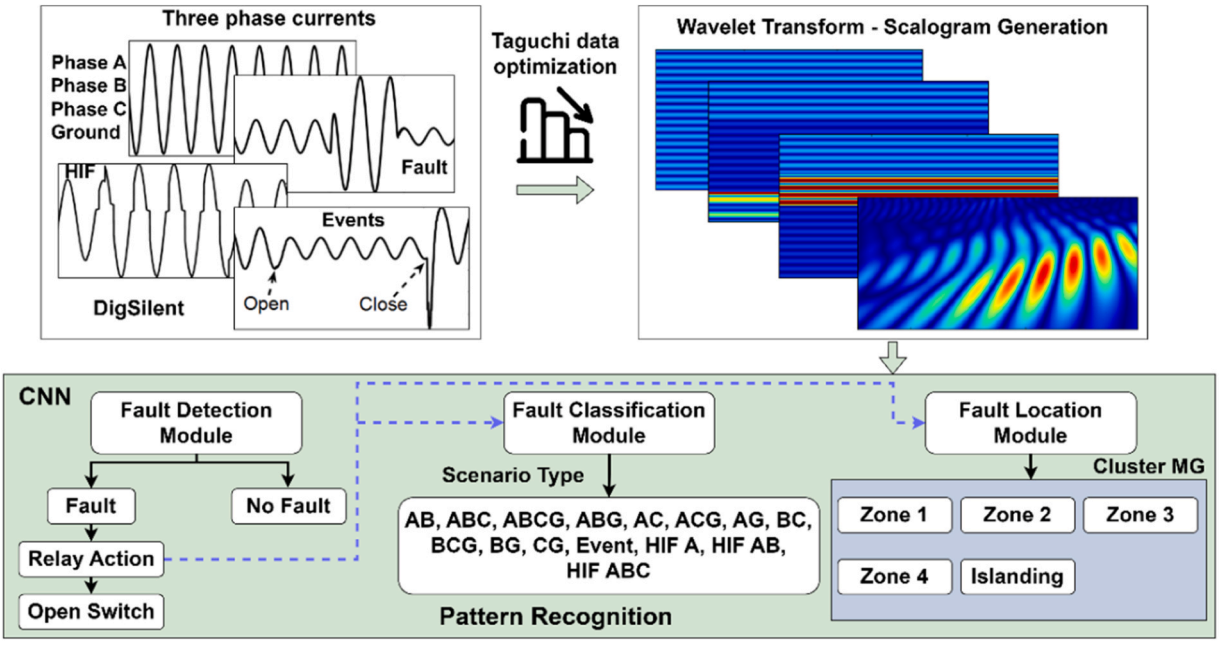


Fig. 1. Schematic representation of the proposed research.

remains costly and complex [26].

To replicate these arcs in automatic switches, the Emanuel arc model was selected as the primary focus of this investigation [19]. Emanuel's model is a representative HIF model comprising variable resistors, diodes, and DC voltage sources connected in antiparallel. In this study, voltage sources and resistors simulate the stochastic nature of arc voltage and resistance fluctuations [19]. The schematic structure of the proposed arc model is depicted in Fig. 2. In the context of the model, V_{FP} and V_{FN} denote the positive and negative arc voltages during HIF events, capturing the nonlinear characteristics of HIFs. Fault resistance and inductance parameters are assigned to the positive and negative branches. For simplicity, Z_{FP} and Z_{FN} refer to the positive and negative arc impedances, respectively.

2.2. Wavelet transform

The Fourier transform (FT) does not provide direct information about oscillating signals and is more suitable for the analysis of problems in a steady state. In contrast, the short-time Fourier transform (STFT) divides the entire time interval into small equal intervals, each analyzed individually by the FT. However, the STFT is not effective in detecting signals of very short duration and high frequency. On the other hand, the WT has been extensively used in the analysis of transient signals due to its diverse window function in the time domain. The WT overcomes the limitations of both FT and STFT [27,28]. In this work, both CWT and DWT are employed to generate scalograms and train the CNN, and the better option between them will be selected.

Let WT be the set of wavelet values for this experiment, and N the set of natural numbers. The CWT of a signal $x_{c,\ w}(t)$ (current wave signals) with respect to a mother wavelet $\psi(t)$ (Daubechies-20 has been heuristically selected in this case), at a scale a, and a shift b is defined as shown in Eq. (1) [29]:

$$CWT_{a,b}[x_{c,\ \ w}(t)] = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} x_{c,\ \ w}(t).\psi(t)^{*}\left(\frac{t-b}{a}\right)dt; \forall CWT \in WT \cup \mathbb{N} \tag{1}$$

where a, b represents the scale factor and temporal displacement, respectively.

The DWT is an extension of the WT designed for discrete-time signals. It decomposes the signal using low-pass and high-pass filters. When applied to the signal $x_{c,\ w}(t)$, the result is given by Eq. (2) [29]:

$$DWT_{a,b}[x_{c,\ \ w}(t)] = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} x(t).\psi(t)\left(\frac{t-b}{a}\right)dt; \forall DWT \in WT \cup \mathbb{N} \tag{2}$$

Calculation of the coefficients $c_{j,k}$ and $d_{j,k}$ involves convolution with DWT and scaling functions, as expressed by Eqs. (3) and (4). Employing a three-level DWT for the specific signal characteristics, this process provides three resolution levels for both approximation and detail coefficients. Each stage results in signal reduction by a factor of 2, [29]:

$$c_{j,k} = \langle x(t),\ \ \phi_{j,k}(t)\rangle = \frac{1}{\sqrt{a_j}}\sum\nolimits_n x(n).\phi\left(\frac{n-k}{a_j}\right); \forall DWT \in WT \cup \mathbb{N} \tag{3}$$

$$d_{j,k} = \langle x(t),\ \ \psi_{j,k}(t)\rangle = \frac{1}{\sqrt{a_j}}\sum\nolimits_n x(n).\psi\left(\frac{n-k}{a_j}\right); \forall DWT \in WT \cup \mathbb{N} \tag{4}$$

where $\phi_{j,k}(t)$ and $\psi_{j,k}$ denote the discrete scaling and wavelet functions, respectively, with n representing the discrete time variable. For a more comprehensive understanding of the signal and its wavelet coefficients, along with an enhanced visualization of signal energy, scalograms have been generated. This process involves mapping the 1-D signal into a 2-D matrix. The resulting matrix, analyzed in multiresolution with a sampling frequency of 200 Hz, provides valuable insights. For the signal $x_{c,\ w}(t)$ and its DWT decomposition up to level J, the scalogram $S_{DWT}(j,k)$ is constructed using detail coefficients $d_{j,k}$ obtained at each decomposition level, as expressed in Eq. (5) [30].

$$S_{DWT}(j,k) = \left|d_{j,k}\right|^2; \forall \quad S \in WT \cup \mathbb{N} \tag{5}$$

In the case of CWT, the scalogram $S_{DWT}(a,b)$ is built by squaring the

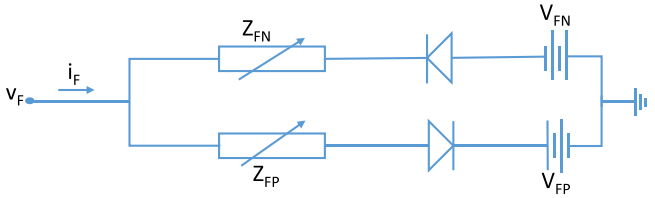


Fig. 2. Diagram illustrating the Emanuel arc model.

magnitude of CWT coefficients at various scales a and positions b, as illustrated in Eq. (6) [30].

$$S_{CWT}(a,b) = \left|CWT_{a,b}[x(t)]\right|^2; \forall \quad S \in WT \cup \mathbb{N} \tag{6}$$

2.3. Taguchi method

The Taguchi method is an advanced experimental design strategy used to optimize industrial systems and processes by determining the ideal configuration of input parameters. In the context of fault detection in a MG cluster, this method is applied to efficiently manage simulation data acquisition, minimizing the number of experiments required without compromising result quality.

Instead of testing all possible combinations of input variables as in traditional full factorial designs, the Taguchi method utilizes orthogonal arrays in a fractional factorial design. These arrays allow exploration of a reduced set of experiments that capture the essential variability of the system. This approach is particularly crucial in studying complex systems like MG clusters, where factors such as short-circuit capacity, solar irradiation, wind speed, among others (detailed in Table 1), significantly influence system response under various operating conditions and fault events.

In this study, various fault simulations were conducted by varying conditions of renewable sources and electrical loads within the MG cluster, using DigSILENT Power Factory software. These simulations generated fault current waveforms in different zones of the MG cluster, measured at the PCC near the static switch. The resulting data were used to create scalograms using WT, serving as training sets for a CNN. To highlight data optimization in this study, each of the eighteen experiments (L1-L18) includes four input variables and 16 noise variables, resulting in a total of 1152 fault scenarios and 18 normal scenarios. In contrast, opting for a full factorial design without Taguchi would have required 16,640 scenarios to cover all possible combinations ($4^4 \times 4 \times 16 + 4^4$). This represents a significant reduction in dataset size ($\mathbb{C}TH$) by approximately 83.44 %. This optimized methodology not only minimizes computational resources and time but also ensures a thorough and efficient evaluation of system response to various conditions and fault events in MG clusters.

In other words, employing full factorial combinations with m_{Tg} factors, each with n_{Tg} levels, would result in an impractical number of required experiments, calculated as ${n_{Tg}}^{m_{Tg}}$. In Taguchi notation $L_{x_{Tg}}({y_{Tg}}^{z_{Tg}})$, where x_{Tg}, y_{Tg}, z_{Tg} represent the number of experiments, levels, and factors respectively, the orthogonal matrix $L_{18}(3^7)$ covers eighteen orthogonal experiments (L1 - L18) to encompass all possible combinations [25]. Specifically, the Taguchi method in this study involves five essential elements:

- MG cluster at PCC. This component, connected to the main grid, serves as the system response generator.
- Fault or even. This factor occurring in the four zones acts as the system input. Isolation occurs if the fault is outside the MG cluster.
- System response. Represented by generated scalograms, this output reflects the system's reaction.
- Types of faults. Whether balanced, unbalanced, or HIF, these faults impact the output response, resulting in a total of 16 types or noises, as illustrated in Fig. 1.
- Other factors. Additional factors affecting system output, detailed in Table 1 $\left(\alpha_{Tg},\ \beta_{Tg},\quad \gamma_{Tg},\ \delta_{Tg}, \varepsilon_{Tg}, \zeta_{Tg},\ \xi_{Tg}\quad\right)$.

Table 1
Factors that affect system behavior.

Factors	Level 1	Level 2	Level 3
α_{Tg}: Load level (MVA)	6	4	8
β_{Tg}: Wind speed (m/s)	7	0	13
γ_{Tg}: Short circuit capacity (MVA)	800	600	1200
δ_{Tg}: Solar irradiance (W/m^2)	700	0	1100
ε_{Tg}: River speed (m/s)	3	0	4
ζ_{Tg}: Hydrogen flow (m^3/h)	15	10	20
ξ_{Tg}: Ambient temperature (°C)	15	5	25

2.4. Convolutional neural network

Convolutional neural networks are instrumental in efficiently managing substantial data volumes while minimizing computational costs, making them pivotal in addressing various classification challenges [31–33]. In the present study, CNNs are employed to classify scalograms $S_{CWT}(a,b)$. The configuration of the CNN used is outlined in Fig. 3, illustrating how scalograms are processed as raw time series (images) of currents by the input layer. These images have dimensions of (656 ×875) pixels and consist of three-color channels (RGB).

The CNN architecture begins with basic components and evolves through systematic sensitivity analysis, as detailed later, to optimize its performance. Initially, 2-D filters are applied in the convolutional layers to sample the input images, thereby transforming them into new matrices. The selection of filter numbers and sizes, such as the implementation of three convolutional layers $(32 \times 3 \times 3); (32 \times 5 \times 5); (32 \times 5 \times 5)$, is determined by the spatial range of neurons within the input matrix. Following convolution, the rectified linear unit (ReLU) activation function is applied to introduce non-linearities and enhance the model's training efficiency [34]. The mathematical formulations for the convolutional layers with ReLU are precisely articulated by Eqs. (7) and (8), with CNN representing the amalgamation of all experiments within this framework. This foundational architecture is further refined through comprehensive sensitivity analysis, elucidated subsequently, which rigorously examines the impact of key hyperparameters on CNN performance. This approach aims to derive optimal values for these parameters, ensuring robustness and efficacy in classifying scalograms $S_{CWT}(a,b)$ [34].

$$\mathscr{I}_{sr,CNN}{}^{(\iota)} = \sum_{\nu=1}^{\Upsilon}\sum_{\rho=1}^{\Gamma}\left(\omega_{\nu\rho}^{(\iota)}.S_{WT(a,b),\quad s+\nu-1,\quad r+\rho-1}\right) + \varsigma_{sr}^{(\iota)} \forall \mathscr{I} \in WT \cup \mathbb{C}TH \cup CNN \tag{7}$$

$$\varpi_{sr,CNN}^{(\iota)} = ReLU({\mathscr{I}_{sr,CNN}}^{(\iota)}); \forall \varpi \in WT \cup \mathbb{C}TH \cup CNN \tag{8}$$

where ${\mathscr{I}_{sr,CNN}}^{(\iota)}$ is defined as the weighted input before applying the activation function, Υ and Γ are the dimensions of the convolution filters, $\varpi_{sr,CNN}^{(\iota)}$ is the activation of the feature at position (s,r) in layer (ι). Additionally, $\omega_{\nu\rho}^{(\iota)}$ represents the weights, and $\varsigma_{s,r}^{(\iota)}$ represents the biases associated with that activation.

In subsequent stages, maximum pooling layers are utilized to downscale data, diminish dimensions, and attenuate redundant information. This process is crucial for averting overfitting and fortifying system resilience. Specifically, a (3×3) maximum pooling layer with a stride of 2 is applied to condense the spatial dimensions of the output (resulting in a $3 \times 3 \times 2$ matrix) for each convolutional layer. To further mitigate overfitting, a dropout layer (DL) with a 50 % dropout rate is incorporated. The mathematical operations governing these processes are formally delineated in Eqs. (9) and (10) [22].

$$\rho_{sr}^{(\iota)} = \max_{m,n}\left(\varpi_{(s-1,CNN)\times\Gamma+m,\quad (r-1)\times\Upsilon+n}^{(\iota)}\right); \forall \rho \in WT \cup \mathbb{C}TH \cup CNN \cup \mathbb{N} \tag{9}$$

$$DL = dropout({\mathscr{I}_{sr,CNN}}^{(\iota)}); \forall DL \in WT \cup \mathbb{C}TH \cup CNN \cup \mathbb{N} \tag{10}$$

Following the convolutional layers, fully connected layers are employed to facilitate pattern recognition. At this stage, each neuron is connected to all neurons in the preceding layer, combining the features

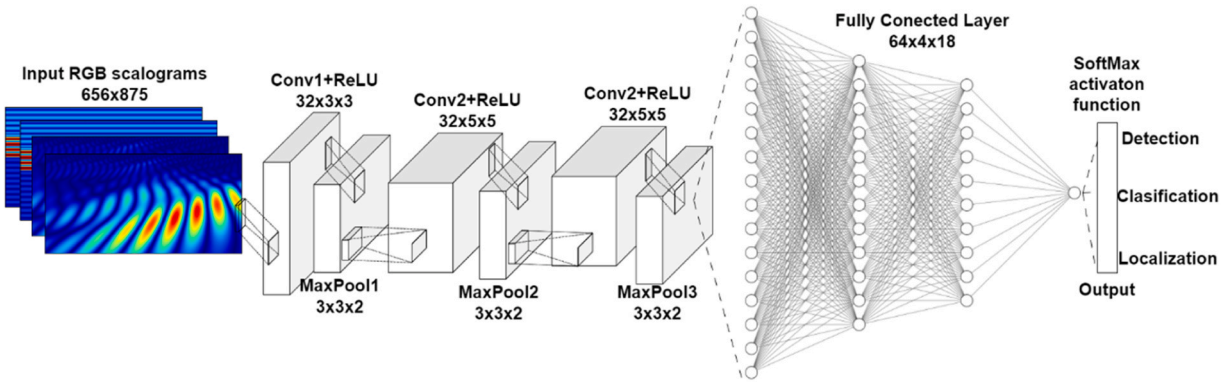


Fig. 3. Fundamental architecture of convolutional neural network.

extracted from previous stages. This configuration enables the classification of intricate patterns, with the number of output neurons corresponding to the distinct classes in the classification task [22]. In this study, the fully connected layer comprises 64 neurons, each dedicated to classifying specific fault types (16 different types across 4 zones) in each Taguchi experiment. The ReLU activation function is applied in these fully connected layers to introduce non-linearity, crucial for enhancing the network's discriminative capability. The mathematical expressions that define the operations of the fully connected layer are provided in Eqs. (11) and (12) [22].

$$\sigma_{\kappa,CNN}^{(\iota)} = \sum_{f=1}^{P}\sum_{e=1}^{Q}\left(\omega_{fe}^{(\iota)}.\varpi_{fe,CNN}^{(\iota-1)}\right) + \varsigma_{\kappa}^{(\iota)}; \forall\sigma \in WT \cup \mathbb{C}TH \cup CNN \cup \mathbb{N} \tag{11}$$

$$\varpi_{\kappa,CNN}^{(\iota)} = ReLU(\sigma_{\kappa,CNN}^{(\iota)}); \forall\varpi \in WT \cup \mathbb{C}TH \cup CNN \cup \mathbb{N} \tag{12}$$

where P and Q are the dimensions of the max-pooling layer activations. The output layer uses the softmax function (SMX) for multiclass classification, as shown in Eq. (13). For the classification task, the cross-entropy loss function is employed, as represented in Eq. (14) [35].

$$SMX = z_c^L; \forall \quad z \in WT \cup \mathbb{C}TH \cup CNN\mathbb{N} \tag{13}$$

$$CrossEntropyLoss(\widehat{y}, y) = \sum_{c} y_c.\log(\widehat{y}_c); \forall \quad \widehat{y}, \quad y \in WT \cup \mathbb{C}TH \cup CNN \cup \mathbb{N} \tag{14}$$

where $\widehat{y}$ is the predictions and y is the one-hot encoded label vector.

During the training setup, pivotal parameters are carefully selected. The Adam optimizer is employed with a maximum of 1000 epochs to optimize model weights iteratively. An initial learning rate of 0.0001 is set to control the magnitude of weight updates during training. Additionally, L2 regularization with a coefficient of 0.001 is applied to prevent overfitting by penalizing large weights in the model.

The model's efficacy is assessed using a separate validation set, ensuring unbiased evaluation of performance. Training progress is visually monitored to track convergence and assess potential issues like overfitting or underfitting. For robust training, the dataset is partitioned into 80 % for training and 20 % for validation, adhering to standard practices in machine learning. The Taguchi methodology is applied to optimize data acquisition in fault simulations within MG clusters. Each of the 18 Taguchi experiments ($L1 - L18$) explores unique combinations of 4 input variables and 16 noise variables, generating 1152 fault scenarios and 18 normal scenarios. These simulations encompass 11 distinct fault types, totaling 12,870 training data points. Compared to a full factorial design approach, which would require 16,640 scenarios to cover all possible combinations, the Taguchi methodology significantly reduces the required dataset size by 83.44 %. Without applying this method, the total number of scenarios would have been 77,742. This optimization conserves computational resources and ensures an efficient evaluation of system response to various conditions and fault events in MG clusters. On the other hand, in the study, the sampling frequency used in DigSILENT Power Factory to simulate the power system was 50 Hz, suitable for capturing variations in system currents and voltages. Subsequently, when processing the data in MATLAB 2021a, for CNN application and DWT analysis, a sampling frequency of 200 fs was employed. This facilitated a detailed representation of the time-frequency energy distribution of current signals, aiding in the identification of relevant patterns for event classification in the power system.

3. Proposed method

This section provides a detailed and chronological explanation of the proposed Taguchi-CNN method, as illustrated in both Fig. 4. The focus of this method is on fault detection, classification, and localization through a static switch at the PCC in a MG cluster. It is noteworthy that MG clusters are typically not owned by the main electrical company. In the event of a fault in the main grid or the MG cluster, it is crucial to activate the static switch on the MG cluster side, near the PCC, before conventional protective relays.

3.1. Taguchi method for fault current data acquisition

At the outset, input data is extracted following multiple computational simulations using DigSILENT PowerFactory 2020 software. In practical applications, these would be acquired through a standard PMU or a protection relay incorporating PMU capabilities. Subsequently, the Taguchi method, is employed to construct an orthogonal dataset specifically designed for training the CNN. This approach not only streamlines training procedures but also maintains the accuracy of CNN predictions.

3.2. DWT for feature extraction

Once the orthogonal is defined, a comparison is made between the accuracy and simulation time in each case, and the best option is chosen. In this case, DWT is selected as the optimal choice. The scalograms are matrices containing signal signatures corresponding to events in the MG cluster. Subsequently, the scalograms are quantized into digital images to compress the data size.

3.3. CNN architecture design and training

The next phase involves inputting the scalograms and training the CNN model. The CNN model is then implemented to detect, classify, and locate faults based on features extracted by the DWT, whether they occur inside or outside the MG cluster. Leveraging its pattern recognition capabilities, the CNN identifies fault patterns within the input data.

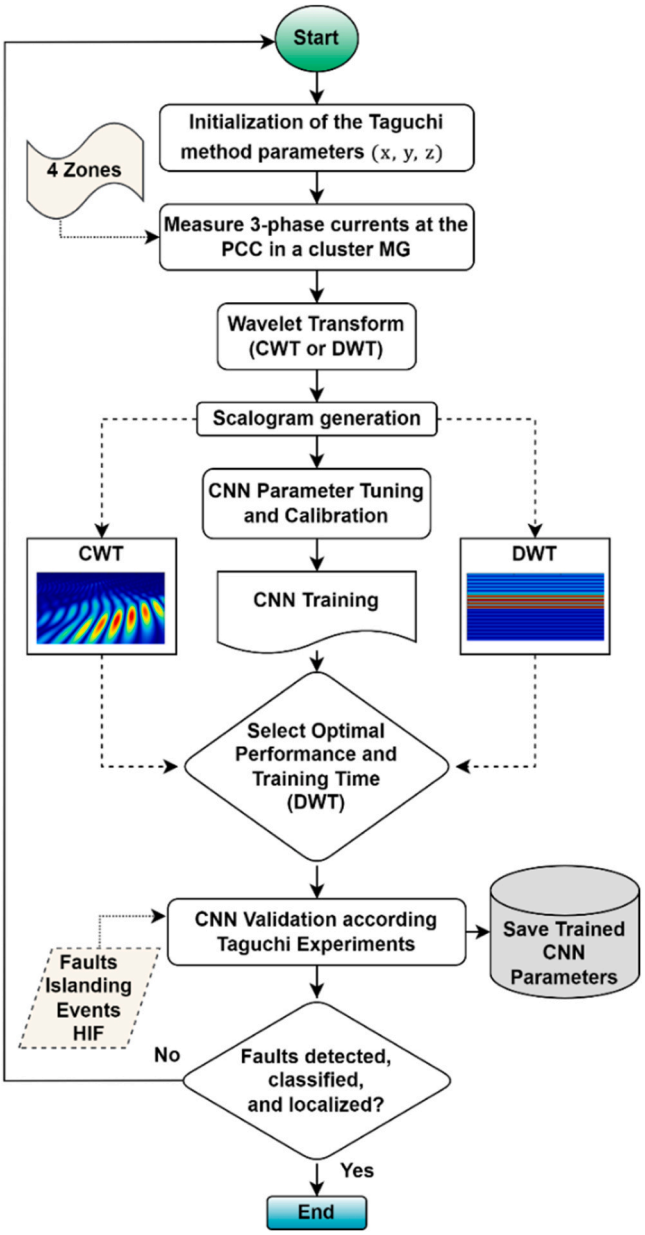


Fig. 4. Flowchart for fault detection, classification, and localization in clustered MG using Taguchi-CNN proposed method.

Finally, the detected event is reported to the static switch at the PCC.

In summary, the flowchart depicted in Fig. 4 explains the methodology employed in this study. The process begins with data preparation, where a set of training images is collected and labeled based on simulations conducted in DigSILENT Power Factory. Using the Taguchi method, the amount of representative data from simulations is reduced to generate scalograms using CWT and DWT. The optimal method is selected in terms of both temporal accuracy and precision efficiency from these scalograms generated for all proposed fault and event combinations in the study. Subsequently, CNN validation is carried out in MATLAB 2021a. Here, an extensive process of training and parameter tuning is conducted to optimize the model. Once the network is trained, it will be capable of detecting, locating, and classifying faults and events within and outside a MG cluster.

4. Case study and data simulation

In this paper, a study and simulation were conducted on a MG cluster connected to the main grid, as depicted in Fig. 5. Zones 1–3 consist of MG cluster, while zone 4 represents the external grid (outside the MG cluster). The key characteristics of the sources and loads forming each MG are outlined in Table 2. Fault locations are randomly chosen and marked on the unified diagram, with fault current measurements taken at the static switch at the PCC. A total of 11 fault types (AB, ABC, ABCG, ABG, AC, ACG, AG, BC, BCG, BG, CG) were analyzed, along with simulations of line opening and closing events (event), (HIF-A, HIF-AB, HIF-ABC), islanding and others events. This results in a total of 16 distinct scenarios or noise types. To gather more data and enhance accuracy, resistances for the 11 fault types, islanding, and line events were varied (R=0.01 Ω, 0.05 Ω, 0.1 Ω, 0.8 Ω, 1.5 Ω, and 3 Ω), considering the factors and levels from the Taguchi experiments that influence the system output. In this study, both symmetrical and asymmetrical faults are included to assess the robustness of the detection system. Symmetrical faults are those in which all phases are equally affected, such as three-phase faults (ABC). On the other hand, asymmetrical faults affect only one or two phases and are more common in distribution systems; examples include phase-to-ground faults (AG, BG, CG), phase-to-phase faults (AB, BC, AC), and phase-to-phase-to-ground faults (ABG, BCG, ACG). To ensure a comprehensive system analysis, an initial fault angle of 90° was considered in the waveform cycle, a point where the current reaches its peak value. This choice allows the evaluation of the model's ability to detect and classify faults under conditions of maximum system demand.

The proposed model was evaluated using a computational workstation with a Core i9 CPU at 3.90 GHz, 16 GB of RAM, and an NVIDIA GeForce RTX 3070 GPU. The Taguchi-CNN code was implemented using MATLAB 2021a.

Moreover, the detailed specifications of each component within the microgrid system, including generation sources, transformers, converters, and the static switch are presented in Table 3. This Table provides a comprehensive overview of the capacities, voltage ratings, and functional descriptions of each element to enhance understanding of the system's operational capabilities and configuration.

5. Results and discussions

This section presents the results of the CNN simulations and their implications. It begins with a comparative analysis between DWT and CWT in terms of fault detection. Following this, the training and validation results of the Taguchi-CNN model are discussed, followed by a sensitivity analysis. Finally, a comparison with other established methods is provided.

5.1. Comparative DWT/CWT

This section provides a detailed examination of the results obtained from our CNN simulations utilizing DWT and CWT scalograms. Table 4 presents a comprehensive comparison highlighting the superior performance of DWT over CWT across various metrics including training time and accuracy for fault detection, event identification, HIF, and MG cluster isolation. Notably, DWT consistently outperforms CWT in fault localization, exhibiting higher accuracy and efficiency, thereby validating its selection as the preferred method in our proposed approach. These findings underscore the effectiveness of DWT in enhancing the performance of fault detection and localization tasks within our proposed methodology. The superior accuracy and efficiency offered by DWT over CWT can significantly contribute to improving the reliability and stability of power systems. By accurately identifying and localizing faults, such as HIF and MG cluster isolation, DWT empowers utility companies and system operators to promptly respond to electrical disturbances, thereby minimizing downtime and enhancing overall system resilience. Additionally, the reduced training time associated with DWT further emphasizes its practical utility, enabling faster deployment and implementation of fault detection systems in real-world scenarios. This

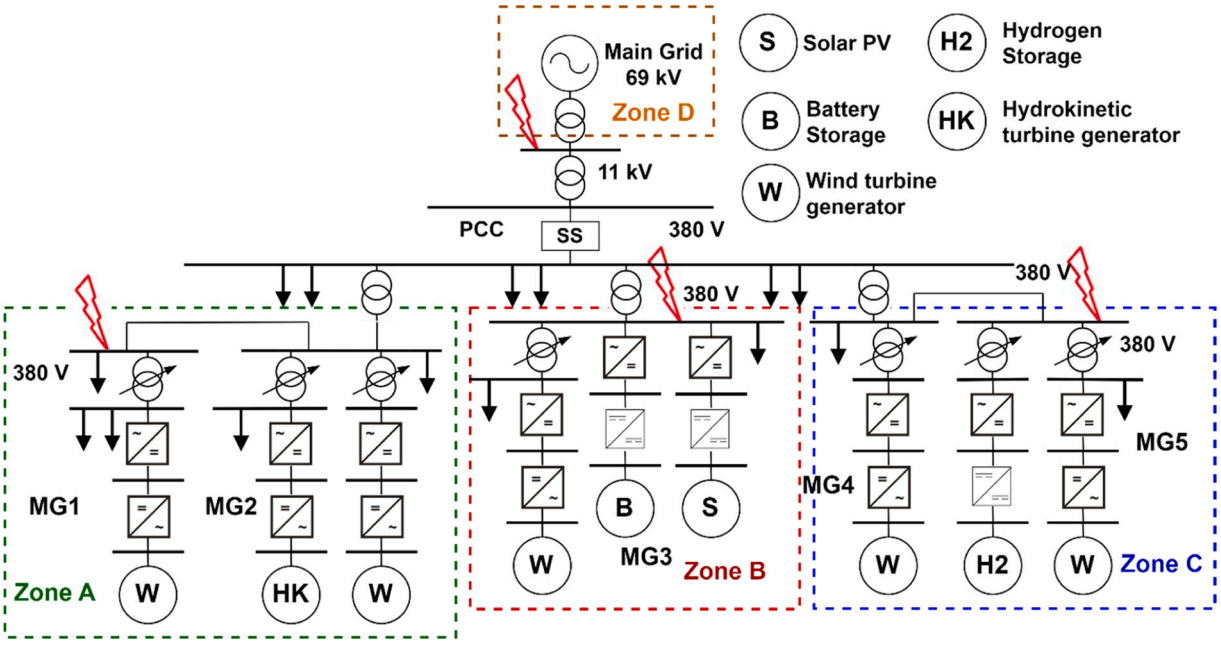


Fig. 5. One-line diagram of proposed cluster MG system.

Table 2
Details of RES and loads in the studied MG cluster.

MG	Source	Power (MW)	Power (MVA)	Load (MVA)
1	Wind	2.50	2.78	1.50
2	Wind, Hydrokinetic	9.10	12.78	8.34
3	Wind, Battery, PV	3.21	3.78	3.00
4	Wind	2.50	2.78	1.50
5	Wind, Hydrogen	2.62	3.08	2.00

reduction in detection time is particularly beneficial for HIF scenarios, where quick, precise responses are essential for maintaining stability in MG clusters.

Fig. 6 exhibits DWT-generated scalograms representing various faults and events within an MG cluster, to be utilized for training and validation of the CNN model. These scalograms enable accurate classification and localization by capturing signal characteristics. Each subplot illustrates distinct operational conditions, thereby contributing to the effective analysis and diagnosis of faults and events.

5.2. Taguchi-CNN training and validation

Fig. 7 provides a detailed insight into the accuracy and loss during the training and validation of the Taguchi-CNN method in four distinct zones of the MG cluster (zone A, zone B, zone C, and zone D). The training accuracy reflects the proportion of examples correctly classified by the model in the training set. It is observed that the training accuracy for all four zones increases over time, reaching the maximum value of 100 %. However, it is important to note that zone A requires 60 epochs to reach 100 %, in contrast to the other zones. This variation could be attributed to factors such as the size and complexity of the dataset in zone A, where there are several RES and energy storage, as seen in Fig. 5. Regarding the loss, both training and validation losses represent the average error of the model's predictions. In Fig. 7, a consistent decrease in both training and validation losses is observed across all zones over time.

Below, Fig. 8 presents the monitoring of training and validation of the Taguchi-CNN across three types of faults and a transmission line opening and closing event. Regarding fault location, it is observed that the training accuracy reaches an optimal level of 100 % after 32 epochs, accompanied by a training loss converging to zero. Subsequently, under a transmission line opening and closing event, it is highlighted that accuracy reaches 100 % after merely 6 epochs, with a corresponding decrease in training loss to zero. This suggests that the CNN is capable of accurately classifying different types of events after a short training period while maintaining minimal loss. In the third case of islanding, it is recorded that accuracy reaches 100 % after 17 epochs, showcasing the model's ability to identify islanding situations within the power system with moderate precision after a moderate number of training epochs, while maintaining minimal loss. Lastly, in the case of HIF, it is emphasized that the CNN achieves a perfect accuracy of 100 % after only 3 epochs, with a rapid convergence of training loss to zero. This result is significant as this type of fault detection often poses a challenge for conventional relays.

5.3. K-fold validation

The observed variations in accuracy within the confusion matrix (Fig. 9) stem from inherent similarities among fault classes, particularly notable in cases like (ACG) and (AC), as well as (BG) and (BCG) faults. These resemblances in electrical characteristics may lead to overlaps in signal features, influenced by factors such as sampling resolution, electrical noise, and variations in power grid conditions. Despite these subtle distinctions, accuracy rates of approximately 96.2 %, 95.6 %, and 94.2 % highlight the model's proficiency in accurately classifying fault types in electrical transmission systems, even when faced with inherent challenges.

Moreover, as depicted in Fig. 10, the CNN demonstrates adept fault localization across four distinct zones, achieving 100 % precision in most cases. Notably, in zone A, the accuracy is commendably high at 96.98 %. Despite this slight variability, the overall average accuracy of 99.25 % in fault localization across diverse areas reaffirms the model's robustness and effectiveness.

In the context of validating a CNN for global fault detection and localization, the results obtained from a systematic evaluation involving multiple folds of the dataset are presented in Fig. 11. These results showcase the model's performance across various metrics, reflecting its effectiveness in accurately detecting and localizing faults. The analysis involves a total of 15 folds, each representing a distinct evaluation scenario. The key performance metrics assessed include accuracy, F1-

Table 3
Technical specifications of microgrid components and system configuration.

Element	Description	Capacity / Limit	Rated Voltage	Comments
MG1 - Wind Turbine	Wind generation with adjustable power factor.	2.5 MW, 2.78 MVA	380 V	Operating under IEC IIA/IIIB wind class, power factor adjustable between 0.95 capacitive and inductive.
MG2 - Wind and Hydrokinetic Turbine	Wind and hydrokinetic energy sources.	9.1 MW, 12.78 MVA	380 V	Turbines with converters allowing reactive power adjustment.
MG3 - Wind, Battery, Solar PV	Combination of wind energy, battery storage, and solar PV.	3.21 MW, 3.78 MVA	380 V	Battery and PV include DC/DC and DC/AC converters for integration with AC systems.
MG4 - Wind Turbine	Additional wind generation.	2.5 MW, 2.78 MVA	380 V	High-efficiency wind generator designed for low wind speeds.
MG5 - Wind and Hydrogen Storage	Wind energy and hydrogen storage.	2.62 MW, 3.08 MVA	380 V	Hydrogen storage with converters for continuous integration in the microgrid.
Main Transformer	Voltage reduction from 69 kV to 11 kV at PCC.	Approximately 25 MVA	69 kV / 11 kV	ONAN transformer with on-load tap changer (OLTC) for voltage adjustment based on load.
Microgrid Transformers	Transformers in each zone to step down from 11 kV to 380 V.	15.56 MVA (Zone A), 3.78 MVA (Zone B), 5.86 MVA (Zone C)	11 kV / 380 V	Distribution transformers to stabilize and distribute voltage at each microgrid level.
Static Switch (SS)	Switch at PCC enabling microgrid disconnection from main grid in case of faults.	N/A	11 kV	Supports isolated operation of MG in emergencies, enhanced by PMUs for real-time monitoring and control under IEC 61850 standard.
DC/DC and DC/AC Converters	Used in batteries, PV, and hydrogen storage to adapt generated energy to AC current.	Bidirectional, automatic active and reactive power adjustment	N/A	Flexible and efficient converters facilitating renewable integration, optimizing power conversion and system stability.

Table 4
Comparative results between DWT and CWT.

Fault Modules	Training time (s)		Accuracy (%)		Epoch (unit)	
	DWT	CWT	DWT	CWT	DWT	CWT
Fault detection	48	51	99.13	98.36	30	23
Event	13	15	100	99.5	6	5
HIF	2.3	1.5	100	99.45	3	4
Islanding	10.5	16.8	100	98.43	17	15
Fault location	53	56.5	99.25	98.36	33	30

score, recall, and Matthews correlation coefficient, all of which are pivotal in gauging the model's robustness. Across these 15 folds, the CNN consistently demonstrates high levels of accuracy, with an overall mean accuracy of approximately 99.13 %.

This underscores the model's capability to correctly identify faults within the dataset. Furthermore, the F1-score, a metric that balances precision and recall, averages at around 98.99 %. This reflects the CNN's ability to effectively strike a balance between precise fault localization and comprehensive fault detection. The recall metric, which assesses the model's ability to accurately identify true positives among all actual positive instances, averages at approximately 98.70 %. This signifies the model's proficiency in capturing most of the actual fault instances present in the dataset. Lastly, the Matthews correlation coefficient, a measure of the quality of binary classifications, demonstrates an overall average of about 98.53 %. This metric reaffirms the model's competence in producing reliable classification outcomes, particularly in the context of fault detection and localization.

5.4. Sensibility analysis

5.4.1. With respect to CNN hyperparameters

The sensitivity analysis presented in Fig. 12 explores how key CNN hyperparameters impact the model's performance regarding accuracy and detection time. Starting with the number of convolutional layers (Fig. 12a), an optimal balance is achieved with three layers, resulting in 99.13 % accuracy and a detection time of 7.6 ms. When increasing beyond seven layers, the model's accuracy declines while detection time increases, indicating that additional layers add computational complexity without improving performance.

Examining the number of filters (Fig. 12b) reveals that using 32 filters achieves a stable performance, with an accuracy of 99.13 % and a detection time of 7.6 ms. Higher filter counts maintain accuracy levels but result in slightly extended detection times. In terms of filter size (Fig. 12c), a 3x3 configuration yields the highest accuracy at 99.13 % with a detection time of 7.8 ms. Smaller sizes, such as 1x1, show consistently strong performance, while larger sizes like 7x7 reduce accuracy to 95.77 % with similar detection times. The analysis of the fully connected layer (Fig. 12d) suggests that 56 neurons provide optimal accuracy (99.13 %) with a detection time of 7.6 ms. Increasing the neuron count in this layer leads to slower detection times, for instance, using 100 neurons reduces accuracy to 91.25 % and extends detection time to 29.36 ms.

Following the analysis of CNN hyperparameters, Fig. 13 further investigates how other configuration elements—such as optimization algorithms, activation functions, pooling layer sizes, and stride values—influence classification accuracy and detection time. Among the optimization algorithms tested (Adam, RMSprop, and SGD), Adam performed best, achieving an accuracy of 99.13 % with a detection time of 0.79 ms. RMSprop and SGD had slightly lower accuracies (95.09 % and 89.18 %) and marginally longer detection times (0.93 ms and 0.87 ms, respectively). The activation functions ReLU and Sigmoid both achieved 99.13 % accuracy with detection times of 0.79 ms and 0.75 ms, respectively, while Tanh algorithm reached 99.07 % accuracy with a detection time of 0.81 ms. Pooling layer sizes were also explored, with 3x3 providing the highest accuracy at 99.13 %, while 2x2 and

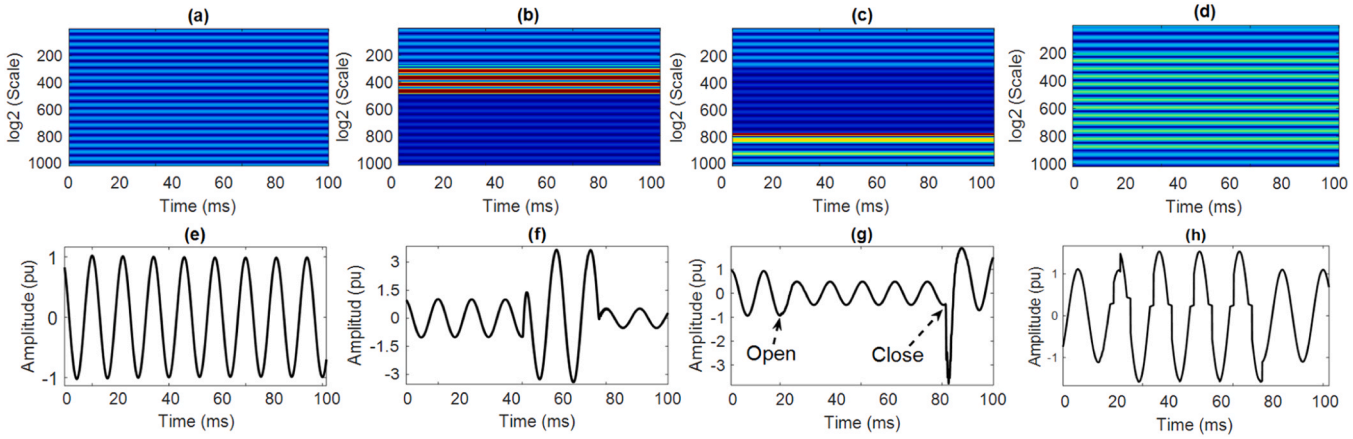


Fig. 6. Scalograms and amplitude waveforms for different conditions: (a, e) No fault, (b, f) Single-phase fault, (c, g) Line opening/closing event, (d, h) High Impedance Fault (HIF).

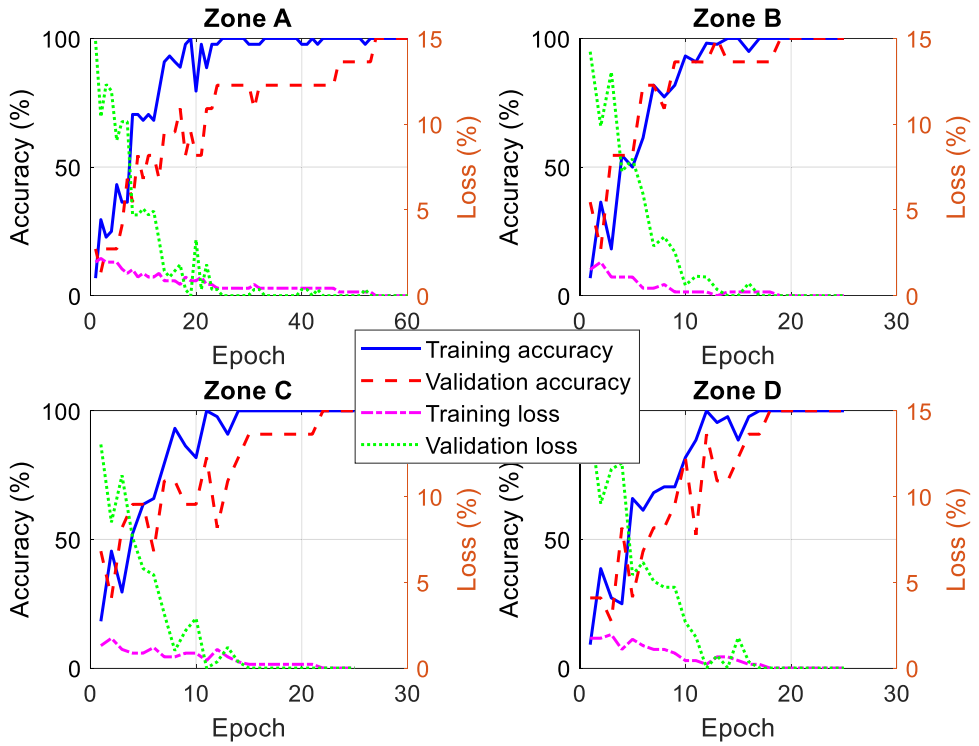


Fig. 7. Training and validation accuracy and loss of the Taguchi-CNN model across MG Cluster zones (A, B, C, D).

4*x*4 sizes reached accuracies of 97.31 % and 99.14 %. Detection times were slightly different, with 0.79 ms for both 2*x*2 and 3*x*3, and 0.93 ms for 4*x*4. Then, stride values of 1, 2, and 3 were tested. A stride of 1 achieved an accuracy of 99.09 % with a detection time of 0.88 ms, while strides of 2 and 3 reached accuracies of 99.13 % and 98.03 % with detection times of 0.79 ms and 0.78 ms, respectively.

This section further examines CNN hyperparameters, offering insights into optimizing performance in data classification. Fig. 14 explores four key categories: dropout rate, initial learning rate, L2 regularization, and the number of training epochs. In dropout rate analysis (Fig. 14a), values from 0.1 to 0.9 were tested, with an optimal rate of 0.5 achieving the highest accuracy of 99.13 %. Higher dropout rates reduce accuracy but improve computational efficiency. For initial learning rate (Fig. 14b), a rate of 0.00001 produces peak accuracy (99.13 %), while slightly higher rates, such as 0.0001 and 0.0002, maintain strong performance, suggesting the importance of precise tuning.

Regarding L2 regularization (Fig. 14c), an optimal value of 0.00002 achieves 99.13 % accuracy, whereas higher regularization levels reduce accuracy, illustrating the balance between regularization strength and model performance. Lastly, examining the number of training epochs (Fig. 14d), accuracy stabilizes at 99.13 % after 35 epochs, indicating convergence. Increasing epochs beyond this point extends training time without significant accuracy gains, highlighting the trade-off between computational efficiency and model precision.

5.4.2. *Sensibility analysis considering Taguchi parameters*

Fig. 15 presents the results of experiments evaluating system accuracy and detection time across various configurations of factors and levels within the Taguchi framework. In Fig. 15(a), we observe that the system's accuracy, displayed as a percentage, remains consistently high (around 99.13 %) even as the number of factors increases from 2 to 13

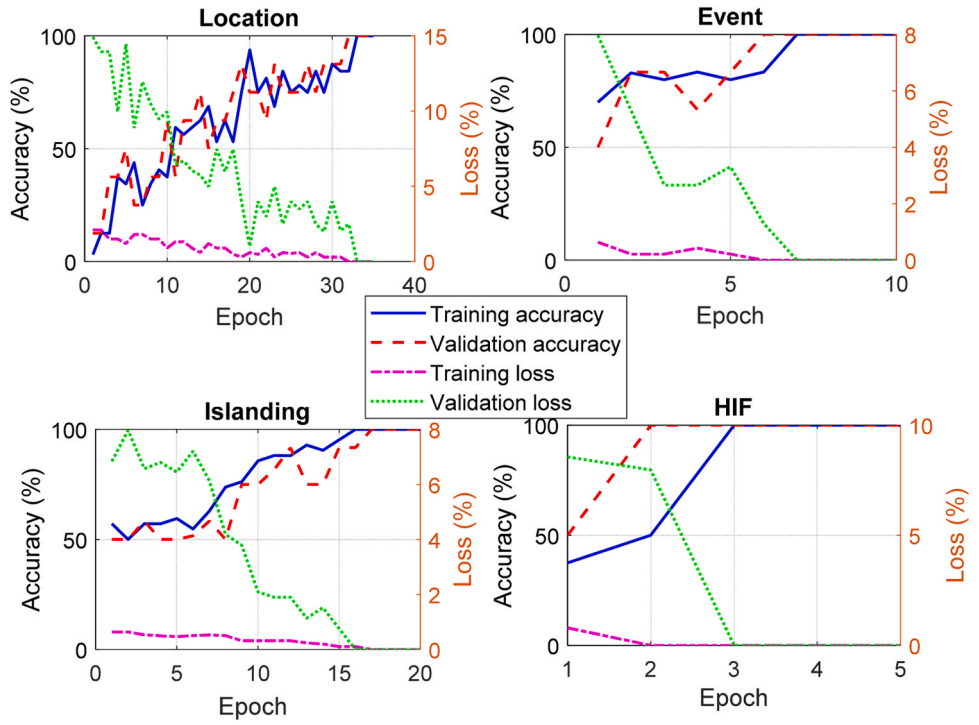


Fig. 8. Training and validation of the Taguchi-CNN across four zones under various event types.

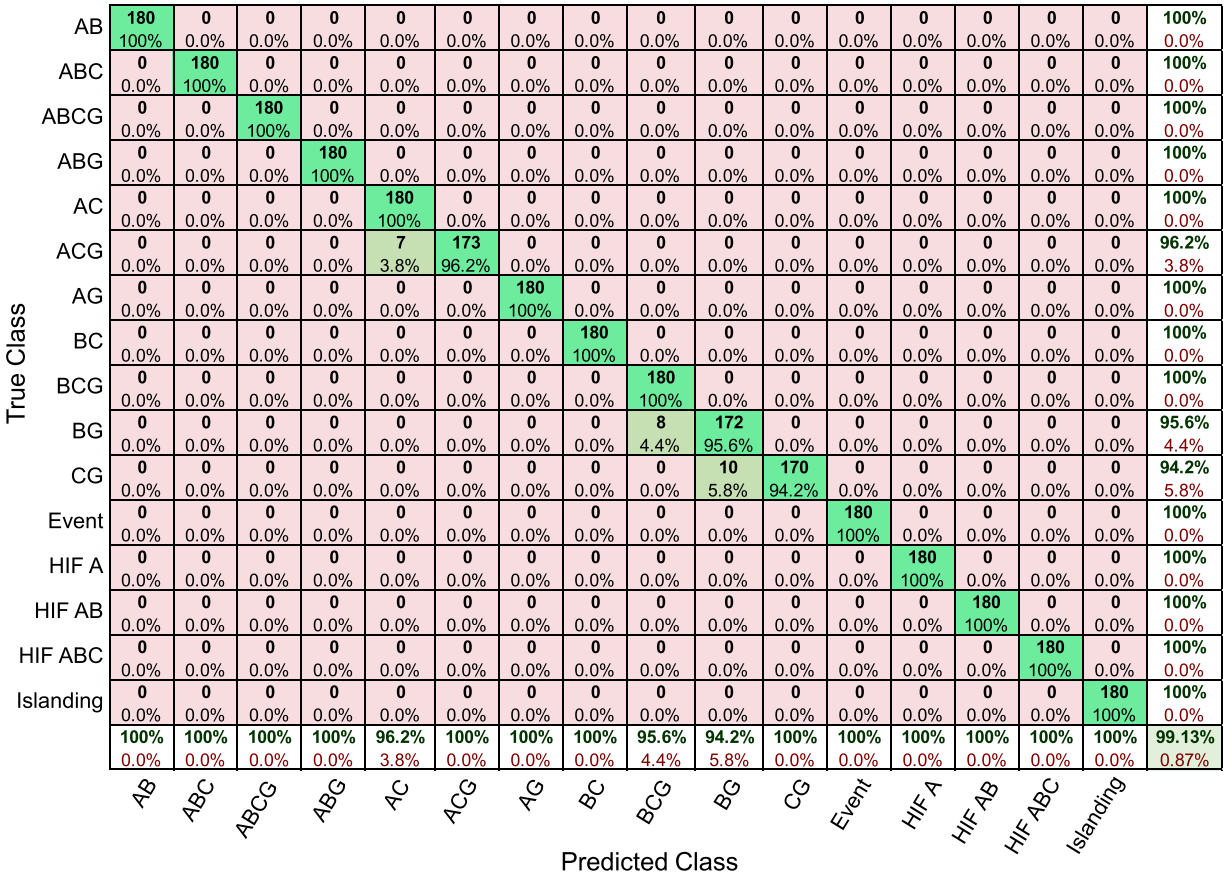


Fig. 9. Taguchi-CNN fault detection confusion matrix.

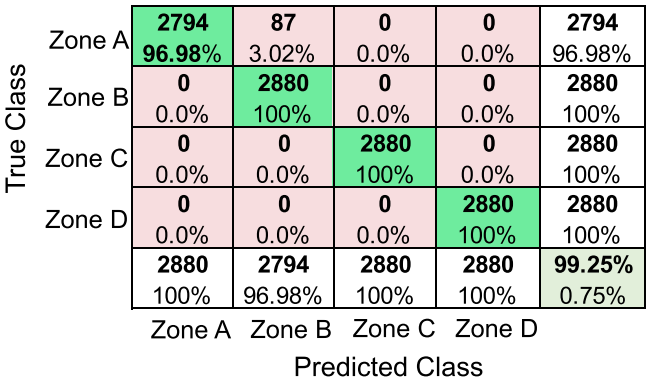


Fig. 10. Taguchi-CNN fault localization confusion matrix.

while keeping the levels constant. This stability suggests that additional factors do not significantly impact accuracy, demonstrating the model's robustness in maintaining performance despite increasing complexity.

In Fig. 15(b), detection time, measured in ms, is displayed alongside accuracy. The results show that increasing both factors and levels leads to longer detection times, indicating a trade-off between complexity and response speed. For example, as the number of factors increases from 2 to 14, detection time rises from 6.55 ms to 11.5 ms, illustrating the added computational demands associated with a higher factor count. Continuing with Fig. 15(c), the analysis shifts focus to levels alone, isolating their effect on system accuracy and detection time. While accuracy remains consistently high as the number of levels rises, detection time increases noticeably. This result indicates that the model can sustain its accuracy across a wider range of levels, but each additional level contributes to increased processing time, impacting overall response efficiency. Then, Fig. 15(d) examines the impact of varying factors

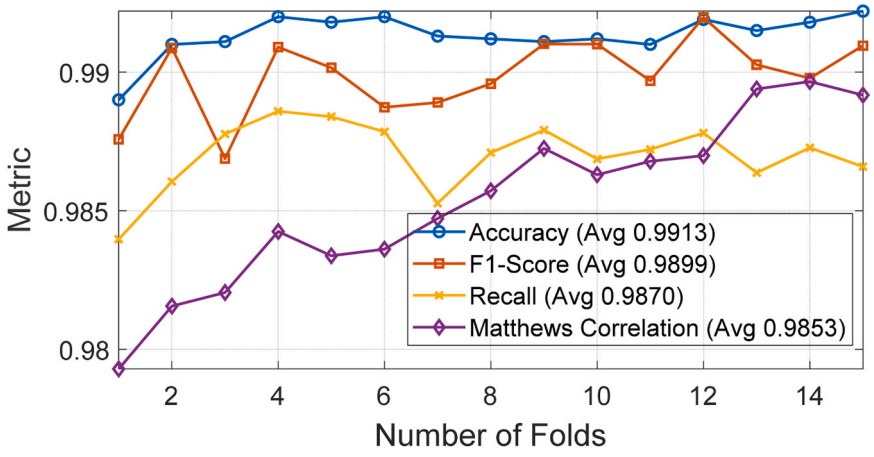


Fig. 11. K-Fold validation of the global CNN.

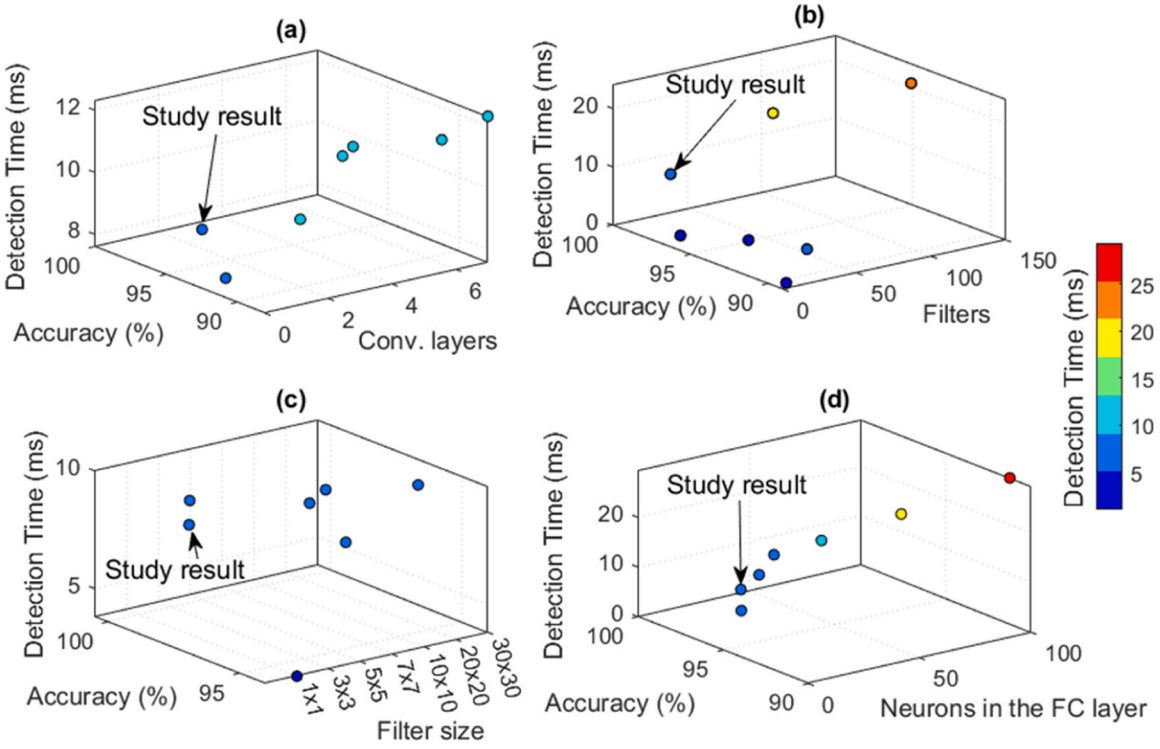


Fig. 12. Sensitivity result: (a) Convolutional layers. (b) Filters. (c) Filter size and (d) Neurons in the FC layer.

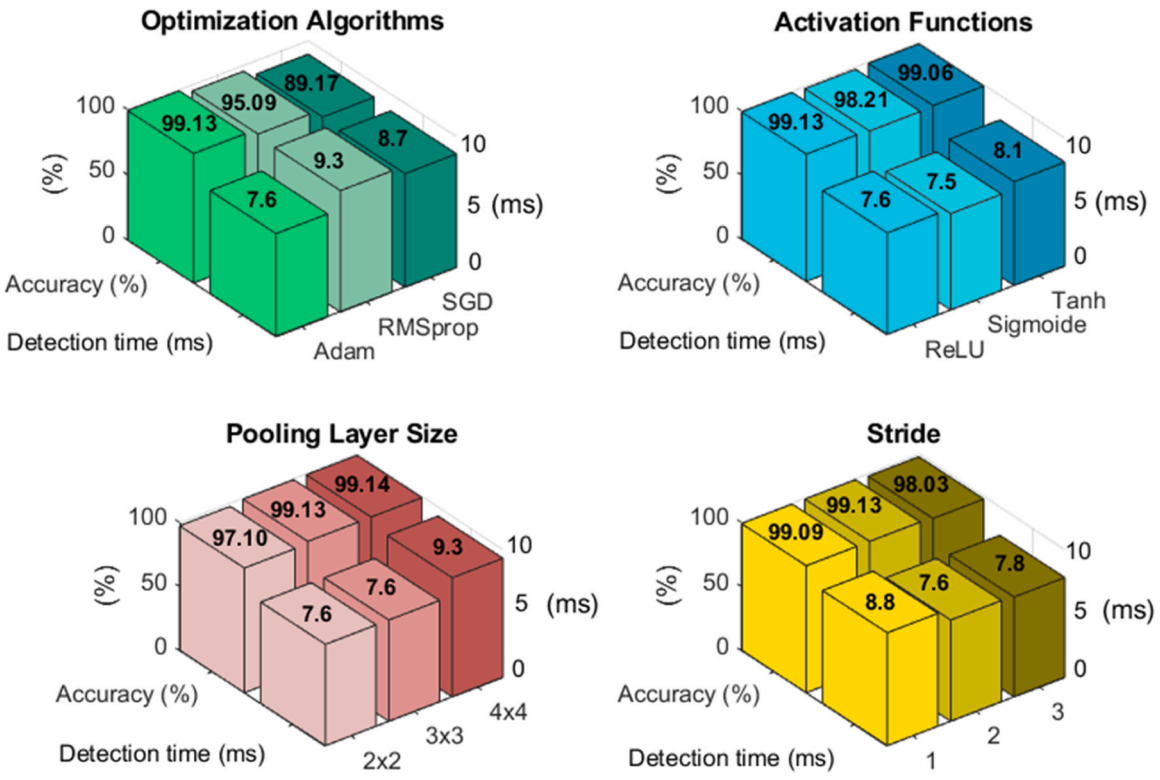


Fig. 13. Sensitivity analysis with respect to CNN parameters.

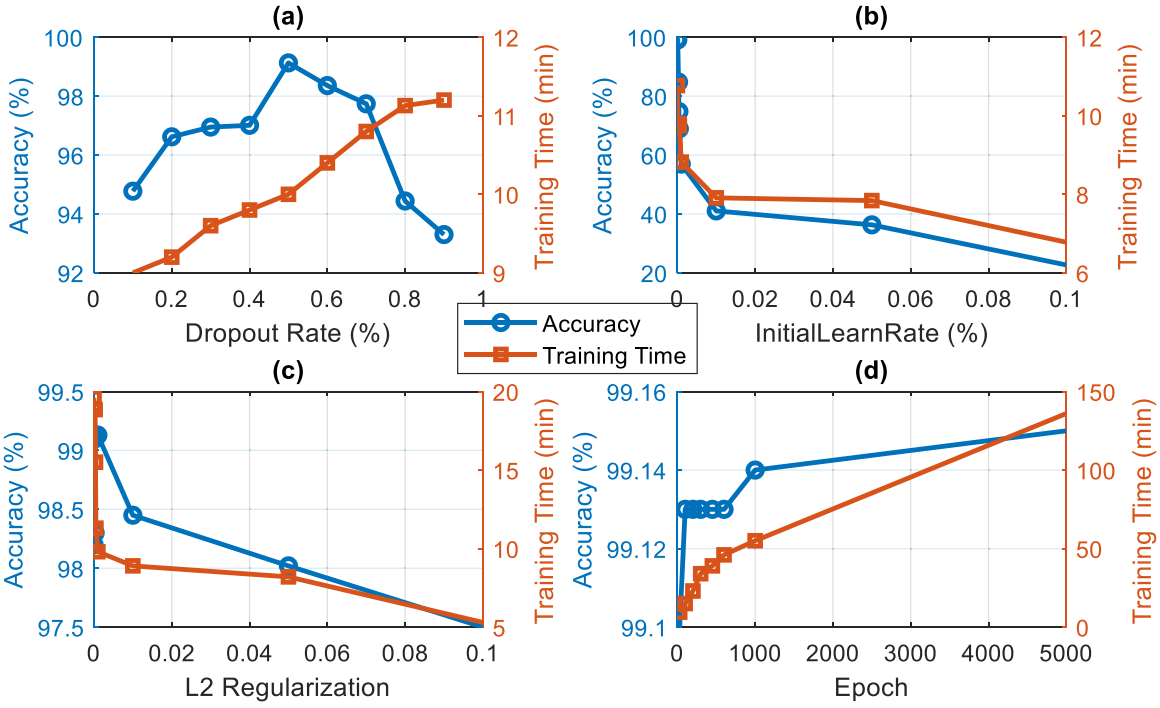


Fig. 14. Hyperparameter comparison for Taguchi-CNN model performance.

independently, showing how each factor level affects accuracy and detection time. Similar to the trends observed in Fig. 15(b), an increase in the number of factors correlates with an increase in detection time, further underscoring the relationship between model complexity and computational load.

5.4.3. Sensitivity analysis with respect to system events

This section provides an in-depth analysis of critical events that commonly occur in electrical power systems, including generator switching, capacitor switching, transformer energization, and load switching within the main grid. This analysis was conducted outside the MG cluster environment, as events within the MG cluster—defined by

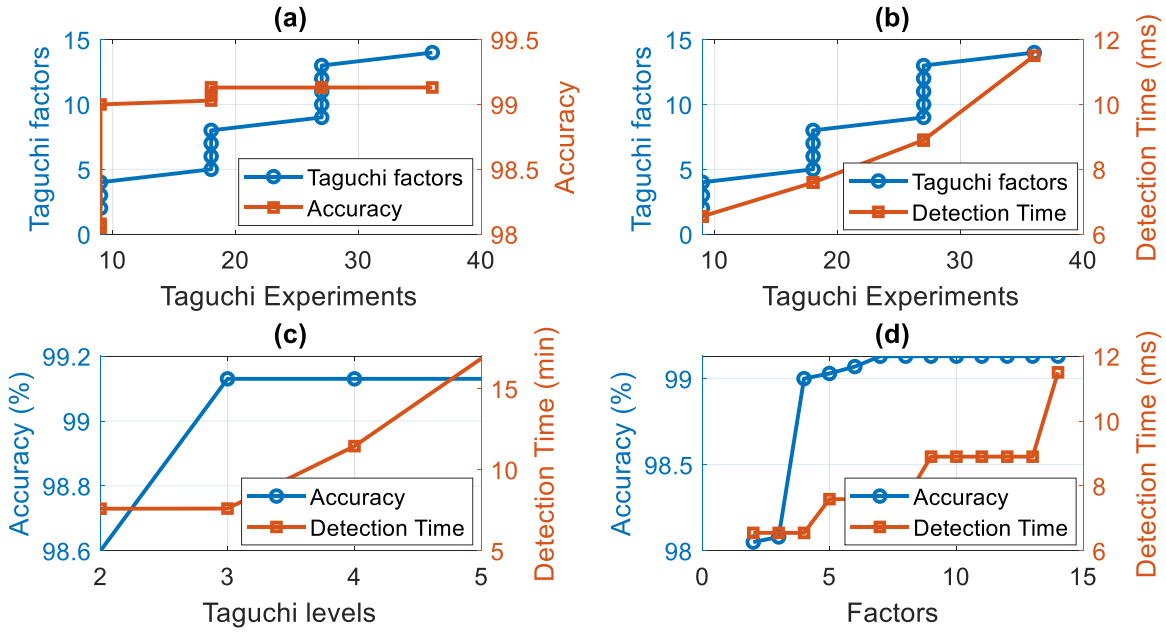


Fig. 15. Trends in system performance, accuracy stability amidst increasing factors and levels.

the SS—are less relevant due to the unique composition of MG, which predominantly rely on RES and storage systems.

To conduct this analysis, the events described in Section 5.3 were simulated using DIgSILENT software to generate the necessary temporal data, which were then input into the Taguchi-CNN model. The model effectively classified, located, and detected faults associated with generator switching, capacitor switching, transformer energization, and load switching. These events were also compared with HIF, single-phase faults, other transient events, and islanding scenarios. Fig. 16 shows the temporal waveforms of each new event, with separate plots for generator switching, capacitor switching, transformer energization, and load switching.

Fig. 17 presents the DWT scalograms generated from these waveforms, which were used to train the CNN and enable accurate classification and detection.

The Taguchi method was applied to streamline the data by adjusting renewable loads to three levels per source, as outlined in Table 2. The resulting data generated a confusion matrix shown in Fig. 18, which demonstrates the model's classification accuracy across different events. The model achieves an overall accuracy of 99.72 %, with a margin of error of 0.28 %.

The high classification accuracy is slightly affected by similarities between specific events. For example, generator switching and load switching show an accuracy of 94.5 % due to similar variations in current and voltage, which complicate differentiation. Likewise, capacitor switching and transformer energization yield an accuracy of 96.77 %, as both events generate transients that impact system signals in similar ways. Despite these minor challenges, the Taguchi-CNN model remains robust and effective in detecting and classifying these critical events, confirming its value in monitoring and analyzing complex electrical systems, including MG.

5.5. Comparison with other methods

In the comparison of fault detection and classification methods, Table 5 provides a summary. The Taguchi-CNN method stands out for its

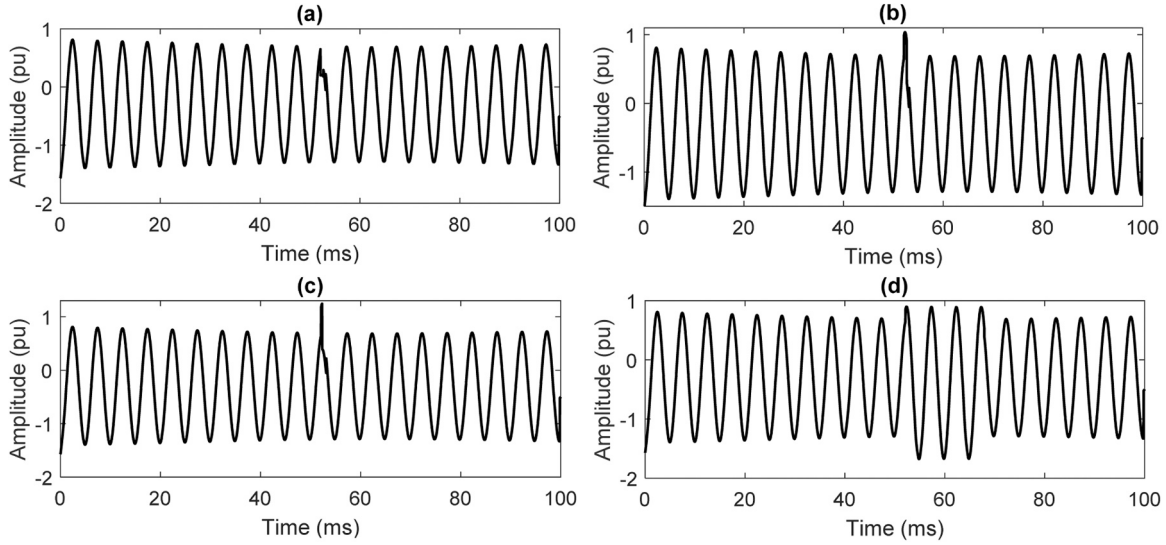


Fig. 16. Temporal waveforms of events. (a) Generator switching. (b) Capacitor switching. (c) Transformer energization. (d) Load switching.

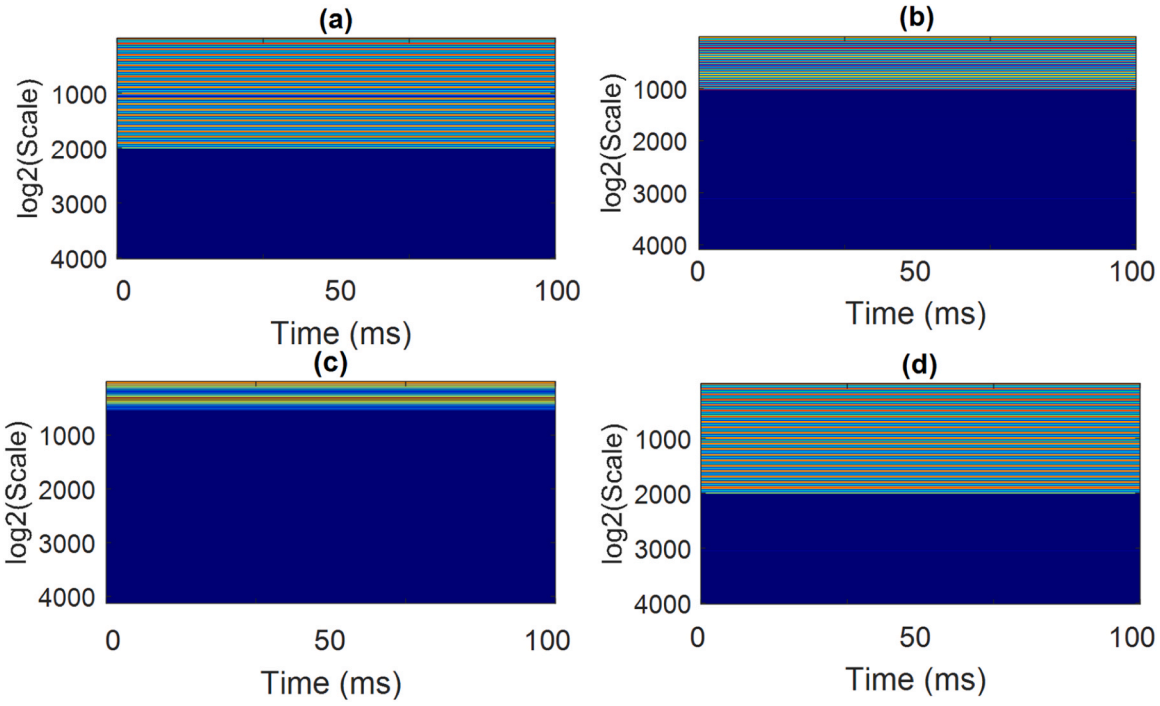


Fig. 17. DWT scalograms generated for CNN training from events. (a) Generator switching. (b) Capacitor switching. (c) Transformer energization. (d) Load switching.

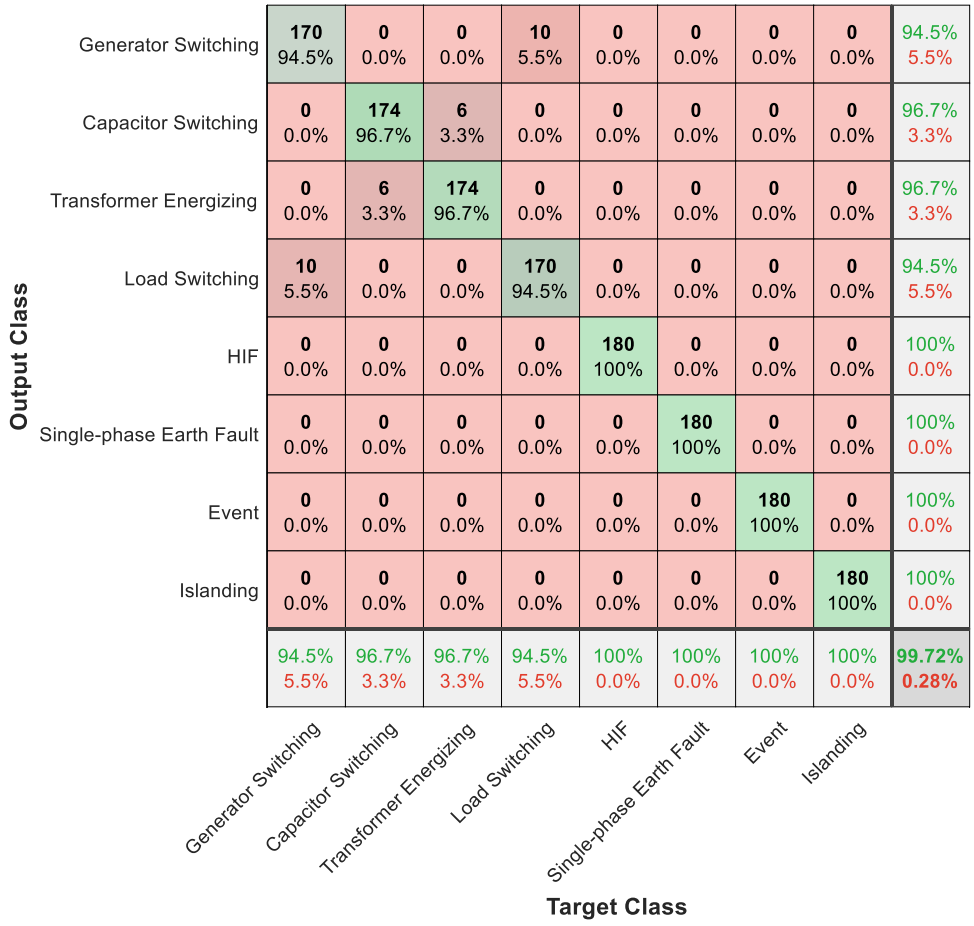


Fig. 18. Confusion matrix of the Taguchi-CNN model for event classification and additional faults in the system with respect to the MG cluster.

Table 5
Comparative analysis of fault detection, classification and location methods in MG systems.

Method	Accuracy (%)			Parameter classification
	Fault detection	Fault classification	Fault location	
Over-current relay [24]	56.00	–	–	–
Differential relay [24]	96.00	–	–	–
Decision Tree [24]	97.00	85.00	–	
Random Forest [24]	99.00	94.00	–	
SVM [18]	99.03	–	–	–
1-D CNN BiLSTM-Attention model	94.53	94.19	93.08	Phases
Xception transformer [19]	98.60	98.60	98.60	Phases, HIF
WT-CNN [23]	99.31	97.60	94.10	Phases, swith event
CNN-GTO [18]	99.36	99.00	98.20	Phases
AI-CNN [13]	99.95	99.95	–	HIF, load change
RFBNN [13]	99.99	99.99	–	Phases
ANFIS [36]	99.09	–	–	Phases
NARX [37]	99.03	–	–	Phases, swith event
FFNN [38]	98.96	98.89	–	Phases
FDI-MPPT [39]	95	95	–	unsymmetrical faults
DTE-ANN [40]	99.3	97.6	–	Voltage Signature Signal
WPT-Node power index [41]	99.32	99.5	–	Phases, swith event
Transient monitoring function [42]	99.39	99.5	–	Phases, swith event
Taguchi-CNN proposed	99.13	99.13	99.25	Phases, HIF, switch event, islanding

effectiveness. When compared with traditional methods like over-current relay and decision tree, as well as modern techniques like 1-D CNN BiLSTM-Attention and Xception transformer, Taguchi-CNN shows competitive accuracy across various parameters.

Notably, Taguchi-CNN achieves an accuracy of 99.13 % in fault detection, 99.13 % in fault classification, and 99.25 % in fault location. Moreover, it excels in classifying parameters such as phases, HIF, switch events, and islanding scenarios. Compared to other models, Taguchi-CNN consistently delivers strong performance, positioning it as a promising solution for fault detection in power systems. Among the methods not listed in Table 5, Taguchi-CNN outperforms recent approaches such as AI-CNN, RFBNN, ANFIS, NARX, FFNN, FDI-MPPT, DTE-ANN, WPT-Node power index, and transient monitoring function. These methods focus on various aspects like phases, HIF, switch events, voltage signature signals, and unsymmetrical faults. The Taguchi-CNN consistently demonstrates strong capabilities across these parameters, establishing its prominence as a robust solution for enhancing fault detection and classification in MG systems.

Beyond the highlighted accuracy figures, the Taguchi-CNN method demonstrates significant advantages in critical scenarios of fault detection and classification compared to traditional methods. The following sections describe three key areas where the proposed system significantly outperforms conventional methodologies:

5.5.1. Real-time detection and response in complex conditions

The Taguchi-CNN system enables fault detection and classification in under 10 ms, a critical speed for MG protection involving high levels of renewable energy penetration and variable load conditions. In contrast, traditional methods, such as those based on overcurrent or differential relays, require longer processing times due to inefficiencies in handling multiple data sources or HIF scenarios. This rapid response capability is essential to improve system stability and continuity, preventing disconnections or damage under complex fault conditions.

The Taguchi-CNN system enables fault detection and classification in under 10 ms, a critical speed for MG protection involving high levels of renewable energy penetration and variable load conditions. In contrast, traditional methods, such as those based on overcurrent or differential relays, require longer processing times due to inefficiencies in handling multiple data sources or HIF scenarios. This rapid response capability is essential to improve system stability and continuity, preventing disconnections or damage under complex fault conditions.

According to [24], machine learning techniques like decision tree and random forest require more than 24 ms to detect faults, limiting their effectiveness in real-time applications for MG environments. Additionally, Ref. [18] reports that SVM methods require over 16 ms for fault detection, highlighting the significant speed advantage of the Taguchi-CNN model, which achieves detection in less than 10 ms. In terms of accuracy and robustness, the Taguchi-CNN method demonstrates superior performance by incorporating a wide range of fault scenarios and renewable variability conditions. Unlike simpler machine learning models, such as SVM or Random Forest, which may struggle to adapt to diverse operational conditions, the Taguchi-CNN model leverages Taguchi optimization to efficiently train on a broad set of scenarios. This results in a highly adaptive and robust model capable of maintaining high accuracy across variable conditions, including HIF faults and renewable generation fluctuations. The combination of rapid detection times and high precision positions the Taguchi-CNN system as a comprehensive and efficient solution for fault detection and classification in MG.

5.5.2. Efficiency in identifying high-impedance faults

The proposed system excels in identifying HIF, a common limitation of conventional methods, which are insufficiently sensitive to detect these faults due to the low fault currents involved. Leveraging Wavelet-optimized CNN, the model can identify complex patterns of HIF, maintaining high accuracy under conditions where traditional methods might fail. This capability allows the proposed system to offer effective protection even in low-magnitude fault situations, substantially enhancing system safety.

5.5.3. Robustness against changes in renewable power generation

Another advantage of the Taguchi-CNN approach is its robustness in maintaining high performance under varying levels of generation and load in the MG cluster. This is a critical challenge for traditional methods, which often rely on fixed operational scenarios and lack adaptability. By optimizing input data with the Taguchi method, the proposed system minimizes the amount of data required for training without sacrificing accuracy. This enables adaptability to variations in operating conditions, ensuring reliable protection in a variable renewable generation environment and thus improving response to disconnection events, load changes, and HIF. While the Taguchi-CNN methodology demonstrates high accuracy and efficiency in fault detection, classification, and localization within MG clusters, its design also shows promising adaptability for larger power networks. The method's optimization capabilities and computational efficiency suggest it could be effectively scaled to handle the increased data volume and complexity in larger systems, with potential for further refinement as required in broader applications

6. Conclusions

This paper presents the Taguchi-CNN methodology based on DWT for fault detection, classification, and localization in MG cluster systems, utilizing existing PMUs at the static switch located at the PCC. This

approach effectively addresses diverse fault types, transient events, and islanding scenarios, achieving a high detection accuracy of 99.25 % within less than 10 ms. The DWT-based multiscale decomposition is shown to be a critical component in maintaining accuracy levels of around 99.13 %, even as the number of factors and levels increases, demonstrating the robustness and stability of the system. The Taguchi-CNN model's speed and adaptability stand out compared to traditional methods, such as overcurrent and differential relays, as well as advanced machine learning models like SVM, Xception transformer, and WT-CNN, achieving over 99 % accuracy across configurations and improving response time efficiency by approximately 60 % over these traditional techniques. While the Taguchi-CNN model demonstrates strong fault detection and classification capabilities, the potential impact of high noise levels in measurements on performance remains an area for further exploration. Future work could incorporate advanced filtering techniques to enhance noise resilience and robustness in diverse operational environments. Additionally, exploring the integration of advanced methods, such as reinforcement learning or attention mechanisms, could further improve the model's adaptability and computational efficiency. Testing this approach in larger, more complex network architectures beyond the MG setting would also help validate its robustness across diverse scenarios. Expanding the performance analysis across various environmental conditions would provide deeper insights into the methodology's reliability and applicability.

CRediT authorship contribution statement

Francisco Jurado: Supervision, Software, Resources, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Paul Arévalo:** Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation. **Antonio Cano:** Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration. **Lena Fedoseenko:** Software, Resources, Project administration, Methodology, Investigation.

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.

Acknowledgment

The author (Paul Arévalo) thanks the Call for Grants for the Requalification of the Spanish University System for 2021–2023, Margarita Salas Grants for the training of young doctors awarded by the Ministry of Universities and financed by the European Union – Next Generation EU.

The authors acknowledge the support provided by the Thematic Network 723RT0150 "Red para la integración a gran escala de energías renovables en sistemas eléctricos (RIBIERSE-CYTED)'' financed by the call for Thematic Networks of the CYTED (Ibero-American Program of Science and Technology for Development) for 2022.

The authors thank Universidad de Cuenca (UCUENCA), Ecuador, for easing access to the facilities of the Micro-Grid Laboratory, Faculty of Engineering, for allowing the use of its equipment, to provide the academic support for the descriptive literature analysis included in this article.

Data availability

Data will be made available on request.

References

[1] N. Salehi, H. Martinez-Garcia, G. Velasco-Quesada, J.M. Guerrero, A comprehensive review of control strategies and optimization methods for individual and community microgrids, IEEE Access 10 (2022) 15935–15955, https://doi.org/10.1109/ACCESS.2022.3142810.

[2] S.N.V.B. Rao, N.R. Narayana, T.D. Prasanna, M.S.N.L.N. Rao, P.G. Chand, Fault detection in cluster microgrids of urban community using multi-resolution technique based wavelet transforms, Int. J. Renew. Energy Res. 12 (2022) 1204–1215, https://doi.org/10.20508/IJRER.V12I3.13129.G8505.

[3] X. Li, L. Guo, Y. Li, C. Hong, Y. Zhang, Z. Guo, et al., Flexible interlinking and coordinated power control of multiple DC microgrids clusters, IEEE Trans. Sustain Energy 9 (2018) 904–915, https://doi.org/10.1109/TSTE.2017.2765681.

[4] N. Bayati, H.R. Baghaee, M. Savaghebi, A. Hajizadeh, M. Soltani, Z. Lin, EMD/HT-based local fault detection in DC microgrid clusters, IET Smart Grid 5 (2022) 177–188, https://doi.org/10.1049/STG2.12060.

[5] A.E. Emanuel, D. Cyganski, J.A. Orr, S. Shiller, E.M. Gulachenski, High impedance fault arcing on sandy soil in 15kV distribution feeders: Contributions to the evaluation of the low frequency spectrum, IEEE Trans. Power Deliv. 5 (1990) 676–686, https://doi.org/10.1109/61.53070.

[6] Moloi K., Jordaan J.A., Hamam Y. High Impedance Fault Classification and Localization Method for Power Distribution Network. 2018 IEEE PES/IAS PowerAfrica, PowerAfrica 2018 2018:84–9. ⟨https://doi.org/10.1109/POWERAFRICA.2018.8520972⟩.

[7] E. Baharozu, S. Ilhan, G. Soykan, High impedance fault localization: a comprehensive review, Electr. Power Syst. Res. 214 (2023) 108892, https://doi.org/10.1016/J.EPSR.2022.108892.

[8] S.H. Mortazavi, M.H. Javidi, E. Kamyab, Robust wide area fault location considering network parameters error, IEEE Trans. Power Deliv. 34 (2019) 786–794, https://doi.org/10.1109/TPWRD.2019.2897402.

[9] Farajollahi M., Shahsavari A., Mohsenian-Rad H. Location identification of high impedance faults using synchronized harmonic phasors. 2017 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2017, Institute of Electrical and Electronics Engineers Inc.; 2017. ⟨https://doi.org/10.1109/ISGT.2017.8086048⟩.

[10] T.A. Alexopoulos, N.M. Manousakis, G.N. Korres, Fault location observability using phasor measurements units via semidefinite programming, IEEE Access 4 (2016) 5187–5195, https://doi.org/10.1109/ACCESS.2016.2602838.

[11] C.L. Banner, B. Don Russell, Practical high-impedance fault detection on distribution feeders, IEEE Trans. Ind. Appl. 33 (1997) 635–640, https://doi.org/10.1109/28.585852.

[12] H. Bai, B. Tang, T. Cheng, H. Liu, High impedance fault detection method in distribution network based on improved Emanuel model and DenseNet, Energy Rep. 8 (2022) 982–987, https://doi.org/10.1016/J.EGYR.2022.05.199.

[13] S. Wang, P. Dehghanian, On the use of artificial intelligence for high impedance fault detection and electrical safety, IEEE Trans. Ind. Appl. 56 (2020) 7208–7216, https://doi.org/10.1109/TIA.2020.3017698.

[14] W.C. Santos, F.V. Lopes, N.S.D. Brito, B.A. Souza, High-impedance fault identification on distribution networks, IEEE Trans. Power Deliv. 32 (2017) 23–32, https://doi.org/10.1109/TPWRD.2016.2548942.

[15] Ray P.K., Panigrahi B.K., Rout P.K., Mohanty A., Dubey H. Fault Detection in IEEE 14-Bus Power System with DG Penetration Using Wavelet Transform 2016. ⟨https://doi.org/10.13140/RG.2.2.32899.09763⟩.

[16] E. Hamatwi, O. Imoru, M.M. Kanime, H.S.A. Kanelombe, Comparative analysis of high impedance fault detection techniques on distribution networks, IEEE Access 11 (2023) 25817–25834, https://doi.org/10.1109/ACCESS.2023.3254923.

[17] H. Ren, Z.J. Hou, B. Vyakaranam, H. Wang, P. Etingov, Power system event classification and localization using a convolutional neural network, Front Energy Res 8 (2020) 607826, https://doi.org/10.3389/FENRG.2020.607826/BIBTEX.

[18] A.Y. Hatata, M.A. Essa, B.E. Sedhom, Adaptive protection scheme for FREEDM microgrid based on convolutional neural network and gorilla troops optimization technique, IEEE Access 10 (2022) 55583–55601, https://doi.org/10.1109/ACCESS.2022.3177544.

[19] J.B. Thomas, S.G. Chaudhari, K.V. Shihabudheen, N.K. Verma, CNN-based transformer model for fault detection in power system networks, IEEE Trans. Instrum. Meas. 72 (2023), https://doi.org/10.1109/TIM.2023.3238059.

[20] Y.Y. Hong, M.T.A.M. Cabatac, Fault detection, classification, and location by static switch in microgrids using wavelet transform and taguchi-based artificial neural network, IEEE Syst. J. 14 (2020) 2725–2735, https://doi.org/10.1109/JSYST.2019.2925594.

[21] Reda H.T., Mahmood A., Anwar A., Chilamkurti N. Adversarial Models Towards Data Availability and Integrity of Distributed State Estimation for Industrial IoT-Based Smart Grid 2022. ⟨https://doi.org/10.48550/arXiv.2206.06027⟩.

[22] S.B.A. Bukhari, C.H. Kim, K.K. Mehmood, R. Haider, M.S.U. Zaman, Convolutional neural network-based intelligent protection strategy for microgrids, IET Gener., Transm. Distrib. 14 (2020) 1177–1185, https://doi.org/10.1049/IET-GTD.2018.7049.

[23] J.J.Q. Yu, Y. Hou, A.Y.S. Lam, V.O.K. Li, Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks, IEEE Trans. Smart Grid 10 (2019) 1694–1703, https://doi.org/10.1109/TSG.2017.2776310.

[24] D.P. Mishra, S.R. Samantaray, G. Joos, A combined wavelet and data-mining based intelligent protection scheme for microgrid, IEEE Trans. Smart Grid 7 (2016) 2295–2304, https://doi.org/10.1109/TSG.2015.2487501.

[25] R.K. Roy, Design of Experiments and the Taguchi Approach, John Wiley & Sons, Inc, 2001.

[26] R.F. Nuqui, A.G. Phadke, Phasor measurement unit placement techniques for complete and incomplete observability, IEEE Trans. Power Deliv. 20 (2005) 2381–2388, https://doi.org/10.1109/TPWRD.2005.855457.
[27] B. JŽ, Comparative analysis of Hilbert Huang and discrete wavelet transform in processing of signals obtained from the cutting process: An intermittent turning example, FME Trans. 41 (2013) 343.
[28] S. Henry, IEEE Computer Society. ICCIMA 2007: International Conference on Computational Intelligence and Multimedia Applications: proceedings: 13-15 December, 2007, Sivakasi, Tamil Nadv, India, IEEE Computer Society; 2007.
[29] D. Liu, X. Lai, Z. Xiao, D. Liu, X. Hu, P. Zhang, Fault diagnosis of rotating machinery based on convolutional neural network and singular value decomposition, Shock Vib. 2020 (2020), https://doi.org/10.1155/2020/6542913.
[30] R. Arvanaghi, S. Danishvar, M. Danishvar, Classification cardiac beats using arterial blood pressure signal based on discrete wavelet transform and deep convolutional neural network, Biomed. Signal Process Control 71 (2022) 103131, https://doi.org/10.1016/J.BSPC.2021.103131.
[31] S. Veerapandiyan, V. Sugavanam, On-line fault identification, location, and seamless service restoration using transfer learning-based convolution neural network for low-voltage DC microgrid, Electr. Power Compon. Syst. 51 (2023) 785–808, https://doi.org/10.1080/15325008.2023.2183997.
[32] S. Karan, Classification of faults in microgrids using deep learning, California State University, 2020.
[33] Q. Zheng, R. Wang, X. Tian, Z. Yu, H. Wang, A. Elhanashi, et al., A real-time transformer discharge pattern recognition method based on CNN-LSTM driven by few-shot learning, Electr. Power Syst. Res. 219 (2023), https://doi.org/10.1016/J.EPSR.2023.109241.
[34] M. Manohar, E. Koley, S. Ghosh, D.K. Mohanta, R.C. Bansal, Spatio-temporal information based protection scheme for PV integrated microgrid under solar irradiance intermittency using deep convolutional neural network, Int. J. Electr. Power Energy Syst. 116 (2020) 105576, https://doi.org/10.1016/J.IJEPES.2019.105576.
[35] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (1989) 674–693, https://doi.org/10.1109/34.192463.
[36] T.C. Srinivasa Rao, S.S. Tulasi Ram, J.B.V. Subrahmanyam, An effective technique for fault detection and classification in distribution system with the aid of DWT and ANFIS, Int. J. Autom. Control 11 (4) (2017) 411–427, https://doi.org/10.1504/IJAAC.2017.087055.
[37] M. Geng, et al., Deep learning-based cooperative trail following, Int. Jt. Conf. Neural Netw. (IJCNN 2018 (2018).
[38] E. Reyes-Archundia, E.L. Moreno-Goytia, J.L. Guardado, An algorithm based on traveling waves for transmission line protection in a TCSC environment, Int. J. Electr. Power Energy Syst. 60 (2014) 367–377, https://doi.org/10.1016/j.ijepes.2014.03.022.
[39] S. Samal, S.R. Samantaray, A novel sequence component based fault detection index for microgrid protection, Electr. Power Syst. Res. 232 (2024) 110380, https://doi.org/10.1016/j.epsr.2024.110380.
[40] B.G. Basher, A. Ghanem, S. Abulanwar, M.K. Hassan, M.E.M. Rizk, Fault classification and localization in microgrids: Leveraging discrete wavelet transform and multi-machine learning techniques considering single point measurements, Electr. Power Syst. Res. 231 (2024) 110362, https://doi.org/10.1016/j.epsr.2024.110362.
[41] M.A. Ajith, R.M. Shereef, Islanding and fault detection of inverter based distributed generations using wavelet packet transform and ensemble, Electr. Power Syst. Res. 231 (2024) 110356, https://doi.org/10.1016/j.epsr.2024.110356.
[42] Deepika Chhetija, Dhiraj Khadka, Yusuf Gupta, Zakir Hussain Rather, Suryanarayana Doolla, Fault detection and classification scheme for power islands with inverter interfaced distributed generators, Electr. Power Syst. Res. 233 (2024) 110438, https://doi.org/10.1016/j.epsr.2024.110438.