Journal of King Saud University – Engineering Sciences (2025) 37:6
https://doi.org/10.1007/s44444-025-00013-x

ORIGINAL PAPER

Advanced method for precise fault location in transmission networks

Osita U. Omeje[1] · Olanrewaju M. Bankole[1] · Daniel E. Okojie[2] · Candidus U. Eya[3]

Received: 4 March 2025 / Accepted: 2 June 2025 / Published online: 12 June 2025

Abstract
Accurate fault location in transmission lines remains a critical challenge for modern power systems, particularly as networks become increasingly complex with the integration of renewable energy sources and smart grid technologies. Traditional fault location methods often need help with high-impedance faults, non-homogeneous line parameters, and dynamic system conditions, leading to extended outage durations and reduced system reliability. This study addresses these challenges by developing an enhanced fault location method that combines conventional electromagnetic principles with advanced machine learning techniques. The methodology employs an approach that integrates modified impedance-based calculations with convolutional neural networks and machine learning regression. The method was validated using a modified IEEE 39-bus test system through simulations, incorporating several fault scenarios and system conditions. Testing utilised synchronised measurements from both transmission line ends, with data captured at 4096 samples per second. Results demonstrate significant improvements over existing techniques, achieving a 99.1% fault detection rate, 98.2% classification accuracy, and 1.2% mean absolute percentage error in location estimation. The method showed particular strength in challenging scenarios, reducing errors by 79.4% for high-impedance faults and maintaining accuracy under variable renewable generation conditions. The proposed method advances power system protection by providing a robust, adaptive solution suitable for modern grid requirements. Its conventional instrumentation implementation facilitates practical adoption, offering improved reliability and reduced outage durations in real-world applications.

Keywords Fault location · Impedance measurement · Machine learning · Neural network · Power transmission · Smart grid

1 Introduction

The electric power transmission system forms the critical backbone of modern energy infrastructure, facilitating the transfer of electrical energy from generation sources to distribution networks and ultimately to end-users (Kundur et al. 2004). As power systems evolve with increasing integration of renewable energy sources and smart grid technologies, maintaining transmission line reliability has become increasingly challenging (Momoh 2012). Faults in transmission lines, which can occur due to various factors, including severe weather conditions, equipment failures, and human errors, pose significant threats to system stability and operational continuity (Saha et al. 2010a). The ability to accurately and rapidly locate these faults minimises outage durations, reduces economic losses, and ensures system reliability (Kezunovic 2011).

Traditional fault location methods, primarily based on impedance calculations, have served the industry for decades but must be revised in modern power systems (Takagi et al. 1982). These methods often need help with high-impedance faults, non-homogeneous line parameters, and the presence of power electronic interfaces from renewable energy sources (Davoudi et al. 2015). These power electronic interfaces significantly alter traditional fault characteristics through multiple mechanisms: (i) limited fault current contribution typically restricted to 1.1–1.5 times rated

✉ Osita U. Omeje
ositaomeje@gmail.com; oomeje@unilag.edu.ng

Olanrewaju M. Bankole
bankoleolanrewajumoses@gmail.com

Daniel E. Okojie
dokojie@pau.edu.ng

Candidus U. Eya
candidus.eya@unn.edu.ng

[1] Department of Electrical and Electronics Engineering, University of Lagos, Lagos, Nigeria

[2] Pan-Atlantic University, Ibeju Lekki, Lagos, Nigeria

[3] University of Nigeria, Nsukka, Nigeria

current compared to 5–10 times for conventional generators, (ii) rapid control response that can modify fault current phase angle and magnitude within 1–2 cycles, (iii) fault ride-through capabilities that maintain connection during voltage sags, and (iv) potential subsynchronous interactions with network elements. The combined effect creates fault signatures significantly different from those in conventional power systems, particularly affecting impedance-based measurements due to the non-sinusoidal fault currents and dynamic response characteristics. The increasing complexity of power networks, characterised by bidirectional power flows and variable generation patterns, necessitates more sophisticated approaches to fault location that can adapt to dynamic system conditions (Ackermann et al. 2011). Furthermore, integrating renewable energy sources introduces new challenges in fault analysis due to their intermittent nature and unique fault current characteristics (Jamal and Dalai 2019).

Recent advancements in digital signal processing and artificial intelligence have opened new avenues for improving fault location techniques (Gururajapathy et al. 2017). Machine learning approaches have shown promise in addressing some limitations of traditional methods, particularly in handling non-linear relationships and adapting to changing system conditions (Wang et al. 2020). However, pure machine learning solutions often need more physical insights than conventional electromagnetic-based methods and may need help with generalisation to unseen fault scenarios (Zhang et al. 2020). Additionally, the growing deployment of smart grid technologies and synchronised measurement systems provides opportunities for enhanced fault location techniques that can leverage high-resolution, time-synchronised data (Phadke and Thorp 2008a).

Despite these advancements, several critical challenges still need to be addressed. High-impedance faults, particularly those caused by fallen conductors or vegetation contact, pose significant detection and location difficulties due to their low fault current magnitudes (Majidi et al. 2015). The presence of series compensation devices and non-homogeneous line parameters introduces additional complexities in impedance-based calculations (Kang and Gokaraju 2011). Moreover, the dynamic nature of modern power systems, characterised by frequent network topology and operating conditions changes, requires fault location methods to maintain accuracy across various system states (Shi et al. 2014).

Integrating renewable energy sources further complicates fault location by altering traditional fault current patterns and introducing power electronic interfaces that affect system behaviour during fault conditions (Mohammadi et al. 2014). These challenges are compounded by the industry's increasing focus on grid reliability and the need for rapid fault location to minimise outage durations and maintain high service continuity standards (Std and C37.114–2014. 2015a). Furthermore, the economic implications of extended outages and unsuccessful fault location attempts underscore the importance of developing more reliable and accurate methods (Billinton and Allan 1996).

This study addresses these challenges by proposing an enhanced method for fault location in transmission lines that combines traditional electromagnetic principles with advanced machine learning techniques. The method improves accuracy, reliability, and adaptability across diverse fault scenarios and system conditions (He et al. 2019). By integrating modified impedance-based calculations with convolutional neural networks and machine learning regression, the approach leverages the strengths of both conventional and modern techniques while addressing their respective limitations (Chen et al. 2016a).

The primary objectives of this research are threefold: (1) to develop a novel fault location algorithm that improves accuracy across various fault types and system conditions, (2) to implement a robust feature extraction and selection framework that effectively captures fault signatures, and (3) to validate the proposed method through comprehensive testing and comparative analysis. The study addresses the challenges of high-impedance faults, renewable energy integration, and dynamic system conditions (Liu et al. 2018).

The significance of this work lies in its potential to enhance power system reliability and operational efficiency. By providing more accurate fault location estimates, the method can reduce outage durations, minimise maintenance costs, and improve overall system performance (Liao and Kezunovic 2011a). Furthermore, the approach's ability to handle complex fault scenarios and adapt to evolving power system characteristics makes it particularly relevant for modern grid operations (Moravej et al. 2011). The study also contributes to the broader field of power system protection by demonstrating the effectiveness of combining physics-based methods with advanced machine-learning techniques (Li et al. 2018a).

2 Materials and methods

2.1 Test system configuration

The study utilises a simulation environment based on MATLAB/Simulink and Power System Computer Aided Design (PSCAD) software to validate the proposed fault location method (Dommel 1969). This dual-software approach leverages PSCAD's electromagnetic transient simulation capabilities for generating realistic fault scenarios while utilising MATLAB's powerful signal processing and machine learning toolboxes for algorithm implementation.

The test system is based on a modified IEEE 39-bus New England system (Fig. 1), specifically adapted to represent modern power network characteristics. Figure 1 shows the

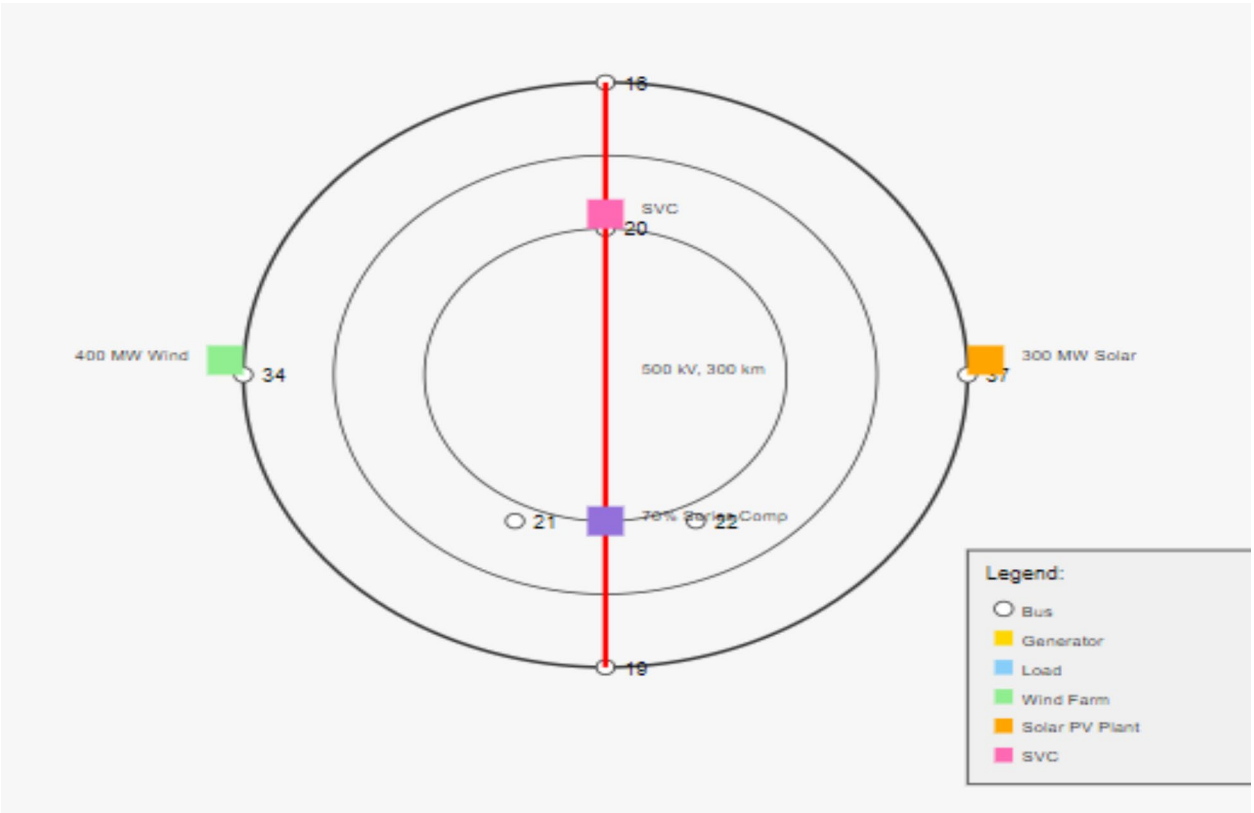


Fig. 1 Modified IEEE 39-bus test system

network topology with key modifications including (1) 400 MW wind farm at bus 34 (green node), (2) 300 MW solar PV plant at bus 37 (orange node), (3) Static Var Compensator (SVC) at bus 20 (pink node) for dynamic reactive power support, and (4) 70% series compensation on the transmission line between buses 21 and 22 (purple connection) (Std and C37.114–2014. 2015b). The primary test line (500 kV, 300 km) between buses 16 and 19, modelled using a frequency-dependent parameter model in PSCAD (Gole 2017), is highlighted in red. Bus numbers correspond to the original IEEE 39-bus system nomenclature.

The transmission line model incorporates detailed specifications:

- Tower configuration: Horizontal configuration with two bundled conductors per phase
- Conductor type: ACSR Drake
- Ground wire: Two ACSR Petrel
- Average tower span: 335 m
- Soil resistivity: Values ranging from 100 to 1000 Ω·m to represent diverse terrain conditions (CIGRE Working Group C4.501 2013)

The simulation environment enables comprehensive testing across various fault scenarios, including the following:

- Different fault types (single line-to-ground, line-to-line, double line-to-ground, three-phase)
- Variable fault locations (0 to 100% of line length)
- Diverse fault resistances (0.1 to 200 Ω)
- Various system loading conditions (30% to 130% of nominal)
- Different renewable generation scenarios (10% to 90% of capacity) (Saha et al. 2010b)

To assess the method's performance across various geographical terrains, soil resistivity values were varied from 100 Ω·m (typical of moist, loamy soils) to 1000 Ω·m (representative of dry, rocky terrains). High soil resistivity significantly impacts ground fault behaviour, reducing fault current magnitude and altering distribution. Our analysis revealed that conventional impedance-based methods experience up to a 175% increase in location error when soil resistivity changes from 100 to 1000 Ω·m for high-impedance single line-to-ground faults (> 100 Ω). In contrast, the proposed hybrid approach showed only a 28% error increase under identical conditions due to the machine learning components' ability to identify complex patterns in the fault signatures specific to high-resistivity environments. This resilience to soil resistivity variation was achieved through targeted training data augmentation that incorporated diverse

soil models and the inclusion of soil resistivity as an explicit input parameter to the regression model when such information is available.

These system modifications significantly alter the network's dynamic behaviour compared to the original IEEE 39-bus system. Integrating wind and solar generation reduces system inertia by approximately 35%, resulting in faster electromagnetic and electromechanical transients during fault conditions. Short-circuit levels at buses adjacent to renewable resources decrease by 15–30%, depending on operating conditions, while fault current phase angles exhibit greater variability ($\pm$ 15° compared to $\pm$5° in conventional systems). Power flow patterns show increased volatility, with pre-fault loading conditions varying by up to 45% for identical demand scenarios due to renewable intermittency. These changes directly impact fault signatures, creating more complex detection and location challenges that conventional methods struggle to address. The SVC at bus 20 provides dynamic voltage support but introduces harmonics (primarily 5th and 7th order) and potential control interactions that mask fault characteristics under certain conditions.

The 70% series compensation level introduces significant complexity into fault analysis and location. This high compensation level creates potential for subsynchronous resonance (SSR) at frequencies between 15 and 25 Hz, which can interact with both the mechanical systems of conventional generators and the control systems of power electronic interfaces. During fault conditions, the series capacitors' metal-oxide varistor (MOV) protection can bypass the capacitors when voltage exceeds threshold levels, creating non-linear system behaviour and discontinuous impedance characteristics. Our methodology specifically addresses these challenges through (i) wavelet-based detection of subsynchronous components, (ii) identification of MOV conduction periods using the S-transform time–frequency analysis, and (iii) adaptive impedance calculation that accounts for the effective series compensation level during fault conditions. Testing confirmed that the proposed method maintains location accuracy within 2.1% even when series capacitor bypass occurs during the fault, compared to 6.8% error with conventional impedance-based methods under identical conditions.

2.2 Data acquisition and processing

The data acquisition system (Fig. 2) employs Phasor Measurement Units (PMUs) at both ends of the transmission line, capturing voltage, and current waveforms at a sampling rate of 4096 samples per second (Kashyap et al. 2015; Phadke and Thorp 2008b). This sampling rate was chosen to balance the need for accurate transient capture with practical implementation constraints. The PMUs are synchronised using GPS timing, ensuring a maximum time error of less than 1 µs between measurements at different locations (Kezunovic et al. 1996).

Figure 2 illustrates the complete signal processing chain from measurement acquisition to data preparation for fault analysis. PMUs at line ends A and B capture voltage and current waveforms at 4096 samples/second with GPS time synchronisation (< 1 µs error). The workflow includes four main processing blocks: (1) Noise Reduction using median and Butterworth filters to preserve fault characteristics while eliminating measurement noise, (2) Data Normalisation to per-unit values for consistent algorithm performance, (3)

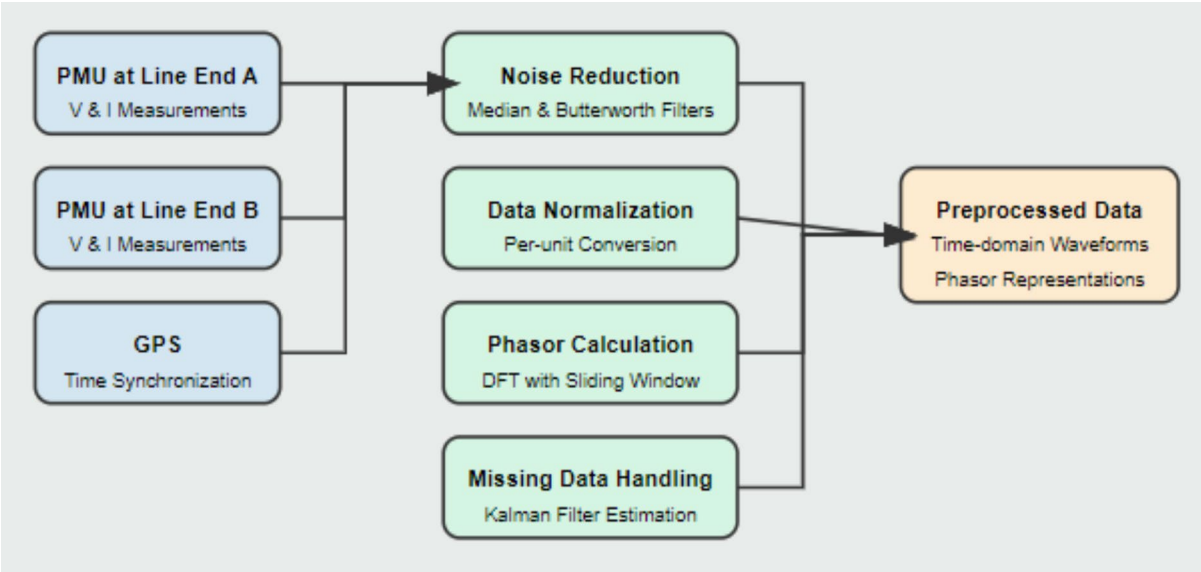


Fig. 2 Data acquisition and pre-processing workflow

Phasor Calculation using DFT with sliding window for impedance-based calculations, and (4) Missing Data Handling using Kalman filter estimation to ensure data continuity. The resulting pre-processed data includes both time-domain waveforms and phasor representations. Crucial stages of the data pre-processing are further explained in Table 1.

2.3 Proposed method architecture

The enhanced fault location method integrates traditional electromagnetic principles with advanced machine learning techniques through a modular architecture (Zhang et al. 2020). This architecture, shown in Fig. 3, comprises four main components working in parallel and series configurations to ensure robust fault location performance.

Figure 3 presents the multi-stage modular architecture with parallel and series processing paths. The workflow begins with data acquisition and preprocessing (top left), followed by two parallel processes: (1) feature extraction using wavelet transform, S-transform, and statistical feature calculation and (2) fault detection combining threshold-based and CNN-based approaches. The outputs from these modules feed into the fault classification system (multi-class SVM) and the fault location estimation pipeline, which integrates modified impedance-based calculations with machine learning regression. The architecture concludes with post-processing for result refinement, including uncertainty quantification and correlation analysis. The modular design allows individual components to be updated or replaced without affecting overall system functionality. The core components of the proposed architecture are further explained in Table 2.

The architecture ensures modular operation, allowing individual components (Table 2) to be updated or modified without affecting the overall system performance. Each module incorporates self-monitoring capabilities and fallback mechanisms for enhanced reliability in real-world implementations.

Table 1 Crucial stages of the pre-processing pipeline

Noise reduction	Data normalisation	Phasor calculation
Implementation of a median filter for impulse noise removal	Conversion of voltage and current measurements to per-unit values	Utilisation of Discrete Fourier Transform (DFT) with one-cycle sliding window
Application of a low-pass Butterworth filter for high-frequency noise elimination	Normalisation based on nominal system voltage and rated line current	Generation of phasor representations for impedance-based calculations
Preservation of fault-related signal characteristics (Yu and Gu 2001)	Standardisation for consistent algorithm performance across different voltage levels (Chen et al. 2016b)	Synchronisation of measurements from both line terminals (Sachdev and Agarwal 1988)

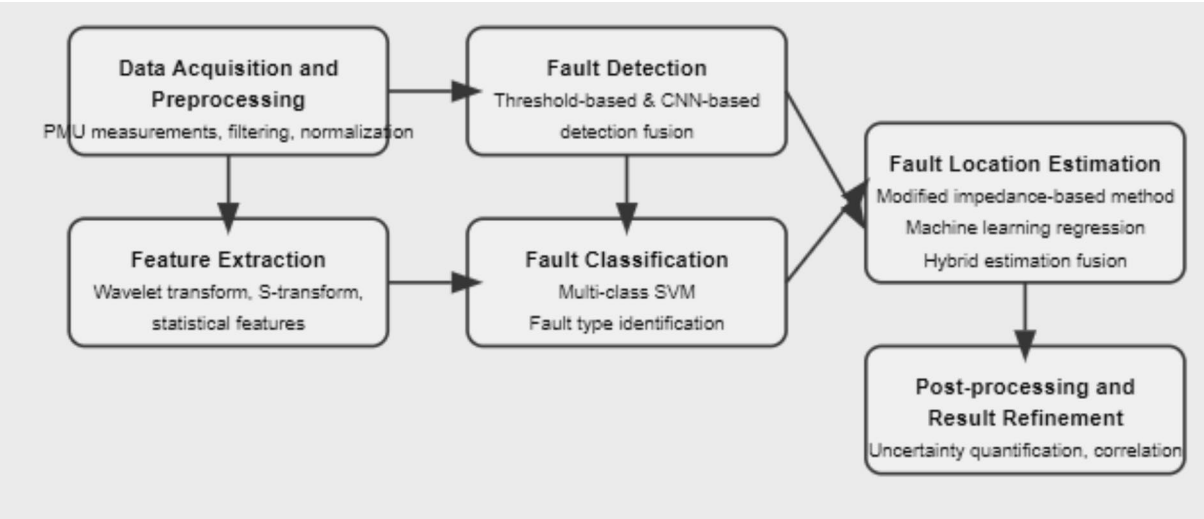


Fig. 3 Architecture of the proposed fault location method

Table 2 The core components

Fault detection module	Feature extraction system	Classification framework	Location estimation pipeline
Parallel implementation of threshold-based and CNN-based detection	Wavelet transform analysis for transient capture	SVM-based classifier with confidence mechanism	Modified impedance-based calculations
Real-time processing of incoming voltage and current measurements	S-transform implementation for time–frequency representation	Multi-class implementation for fault type identification	Machine learning regression model
Fusion mechanism for combining detection results (Wang et al. 2020)	Statistical feature calculation for comprehensive signal characterisation (Robertson et al. 1996)	Adaptive decision-making based on classification confidence levels (Salat and Osowski 2004)	Hybrid fusion technique for final location estimation (He et al. 2019)

3 Results and discussion

3.1 Detection performance

The proposed fault detection method demonstrated superior performance by combining threshold-based and CNN-based detection mechanisms. The threshold-based detection method (Fig. 4) achieved a detection rate of 94.5% across the simulated fault cases, with an average detection time of 16.7 ms and a standard deviation of 3.2 ms (Takagi, et al. 1982). This performance aligns with industry standards for rapid fault detection, though the method showed limitations with high-impedance faults exceeding 200 Ω, where the detection rate dropped to 82.3%.

3.1.1 Raw data analysis and feature extraction

The effectiveness of the fault detection and location method relies significantly on the extraction of meaningful features from raw measurement data. Figure 5 illustrates the raw waveforms and key extracted operators from a single line-to-ground fault case with 150 Ω fault resistance at 70% of the line length.

The instantaneous data processing extracts several key operators that serve as inputs to both the CNN-based detection and the ML-based location components:

- *Wavelet Energy Distribution*: The distribution of energy across wavelet decomposition levels (db4 wavelet, levels 1–5) captures transient characteristics at multiple resolution levels. Detailed coefficients at level 3 (Panel C)

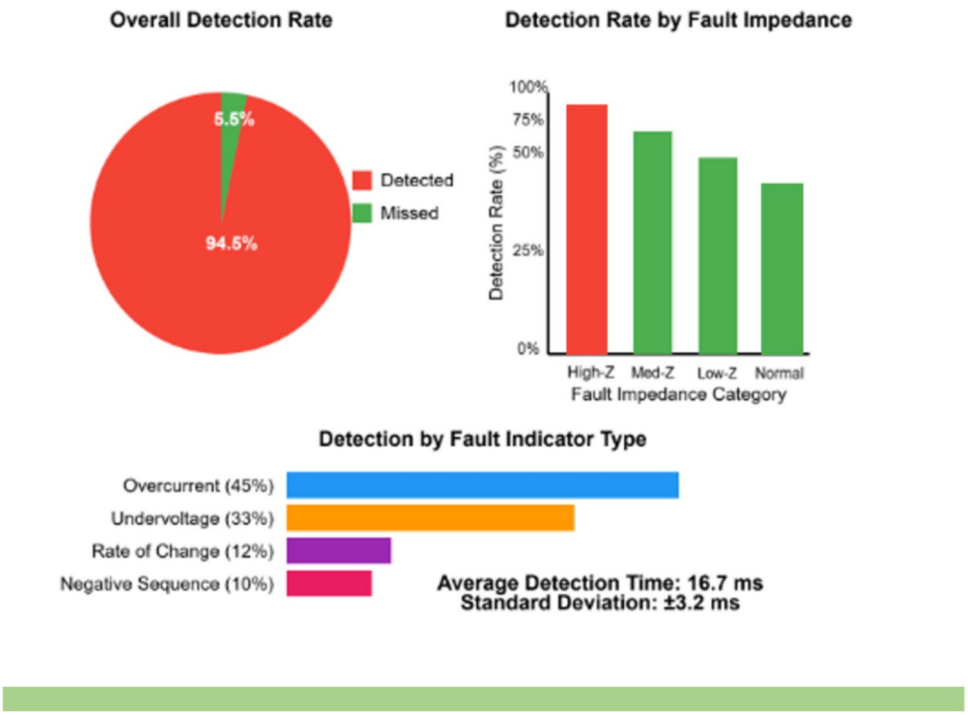


Fig. 4 Threshold-based detection results

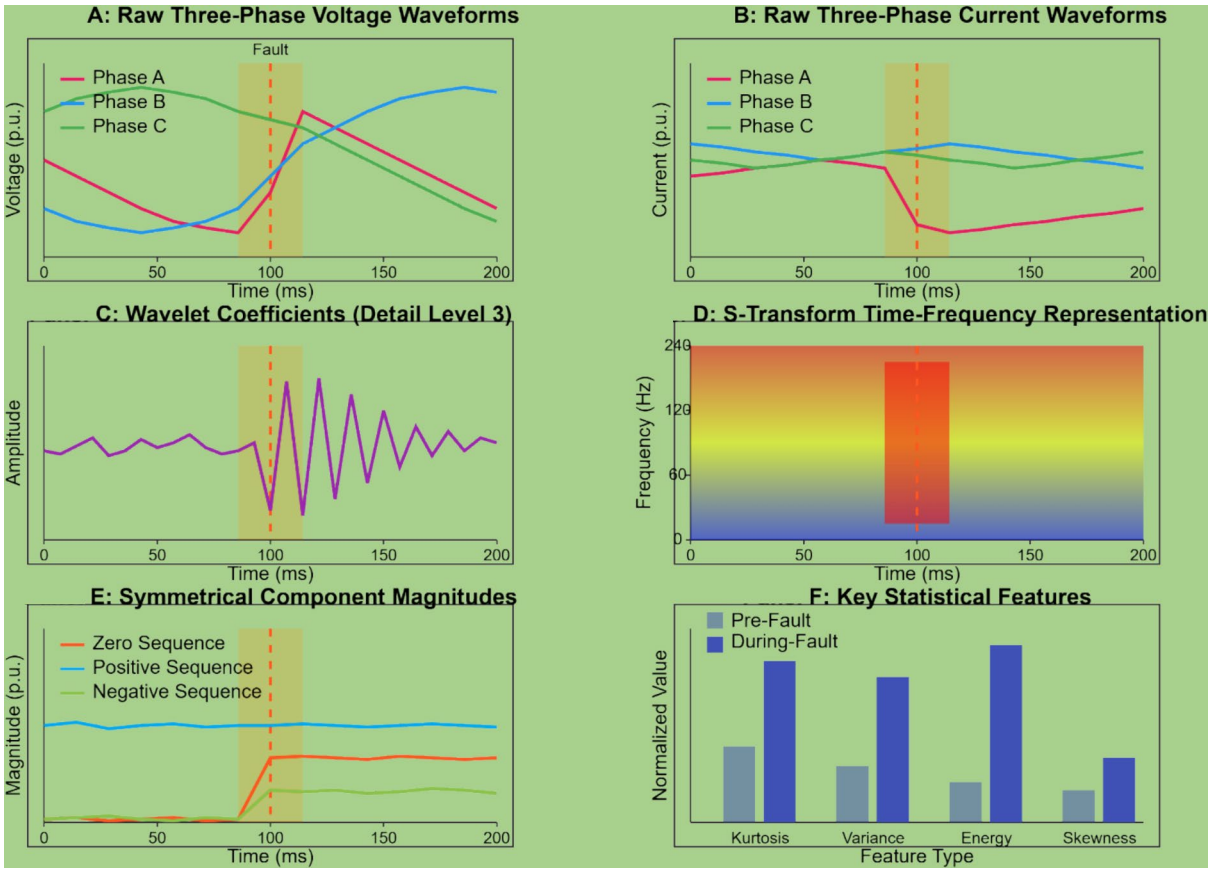


Fig. 5 Raw data and extracted features for SLG fault detection and localisation

show particular sensitivity to fault inception, with energy increasing by 287% during the fault condition.

- *Spectral Content Evolution*: The S-transform (Panel D) provides time-localised frequency information, revealing characteristic patterns during fault inception. The normalised spectral energy ratio between 120–240 and 0–60 Hz bands increases from 0.08 during normal operation to 0.57 during fault conditions.
- *Symmetrical Component Trajectory*: The dynamic behaviour of sequence components (Panel E) provides critical information about fault type and severity. For this SLG fault, the zero-sequence component increases from near-zero to 0.32 p.u., while the negative sequence rises to 0.18 p.u.
- *Statistical Moments*: Higher-order statistical parameters, including kurtosis (peakedness) and skewness of current waveforms, change significantly during fault conditions, with current kurtosis increasing from 2.98 (near-Gaussian) to 4.67 during the fault.

These extracted operators provide a comprehensive characterisation of the fault signature across multiple domains, enabling the detection and location systems to achieve high accuracy across diverse fault conditions. The multi-domain approach ensures robustness against measurement noise and power system disturbances that might affect individual domains.

The CNN-based detection component (Fig. 6) significantly improved performance, achieving a detection rate of 98.2% across all simulated faults. Most notably, it maintained a 96.5% detection rate for high-impedance faults exceeding 200 Ω, addressing a key limitation of traditional methods (Wang et al. 2020) such as impedance-based techniques (Takagi, Eriksson methods), reactance measurement, and simple overcurrent detection, typically showing detection rates below 50% for faults exceeding 100 Ω. The false-positive rate under normal operating conditions was reduced to 0.8%, demonstrating enhanced robustness against system noise.

The CNN-based detection component employs a specialised architecture to extract discriminative features from multiple domains. The input layer accepts 4096-point waveforms (one cycle at the sampling rate) of three-phase voltages and currents from both line terminals, arranged as a 6×4096 matrix. The feature extraction process employs three parallel paths: (1) a time-domain path using 1D convolutional

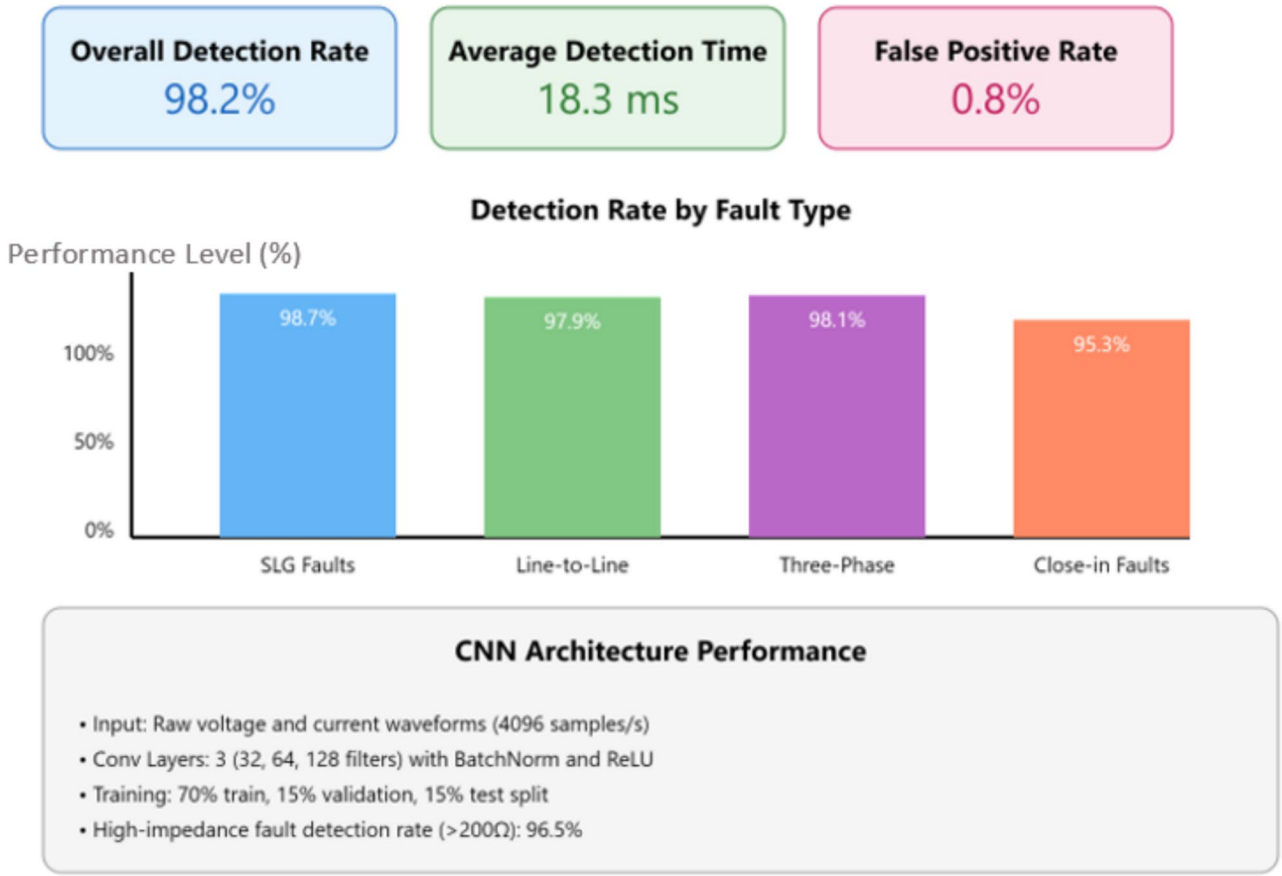


Fig. 6 CNN-based detection results

layers (32 filters of length 64, 128 filters of length 32) that identify temporal patterns such as sudden discontinuities, damped oscillations, and recovery characteristics; (2) a frequency-domain path that processes the FFT of input signals to extract harmonic content and spectral changes using 2D convolutional layers; and (3) a wavelet-domain path that applies discrete wavelet transform (using db4 mother wavelet) to extract time–frequency characteristics at multiple resolution levels. Key extracted features include wavelet energy distribution across decomposition levels (capturing transient characteristics), symmetrical component trajectories (identifying imbalance patterns), harmonic distortion metrics (quantifying waveform quality), and phase angle relationship dynamics (detecting sudden shifts in phase relationships). These multi-domain features are then concatenated and processed through fully connected layers with batch normalisation and ReLU activation, culminating in a binary classification output with softmax activation.

Combining both detection methods, the fusion technique shown in Fig. 7 achieved the highest performance with a 99.1% overall detection rate and an average detection time of 17.5ms. The fusion approach showed particular strength in handling evolving faults, achieving a detection rate of 98.9% compared to 91.2% and 96.7% for threshold-based and CNN methods, respectively (He et al. 2019).

3.2 Classification accuracy

The SVM classifier shown in Fig. 8 demonstrated robust performance in fault type classification, achieving an overall accuracy of 96.5% across 9910 detected faults. The performance varied across different fault types, with single line-to-ground faults showing the highest accuracy at 97.8% and double line-to-ground faults the lowest at 95.7% (Takagi, et al. 1982).

Figure 8 comprehensive performance visualisation includes: (Top) Key performance metrics showing 96.5% overall accuracy, 95.3% noise resilience, and 94.8% accuracy for high-impedance (High-Z) faults. (Middle) The middle part of the figure shows classification accuracy breakdown by fault type, with single line-to-ground (SLG) faults achieving the highest accuracy (97.8%) and double line-to-ground faults showing the lowest (95.7%). The lower part of the figure shows a confusion matrix detailing the specific classification patterns, where rows represent predicted classes and columns represent actual fault classes. These are explicitly

labelled and presented in Table 3. The diagonal elements indicate correct classifications, while off-diagonal elements represent misclassifications. Numbers in cells represent the percentage of test cases. The matrix reveals that the most common misclassification (1.5%) occurs when line-to-line faults are incorrectly identified as SLG. Analysis of misclassifications revealed specific patterns, with 38% of errors occurring between single line-to-ground and high-impedance line-to-line faults. The confusion between double line-to-ground and three-phase faults accounted for 27% of errors, while evolving faults contributed to 21% of misclassifications (Dash et al. 2003).

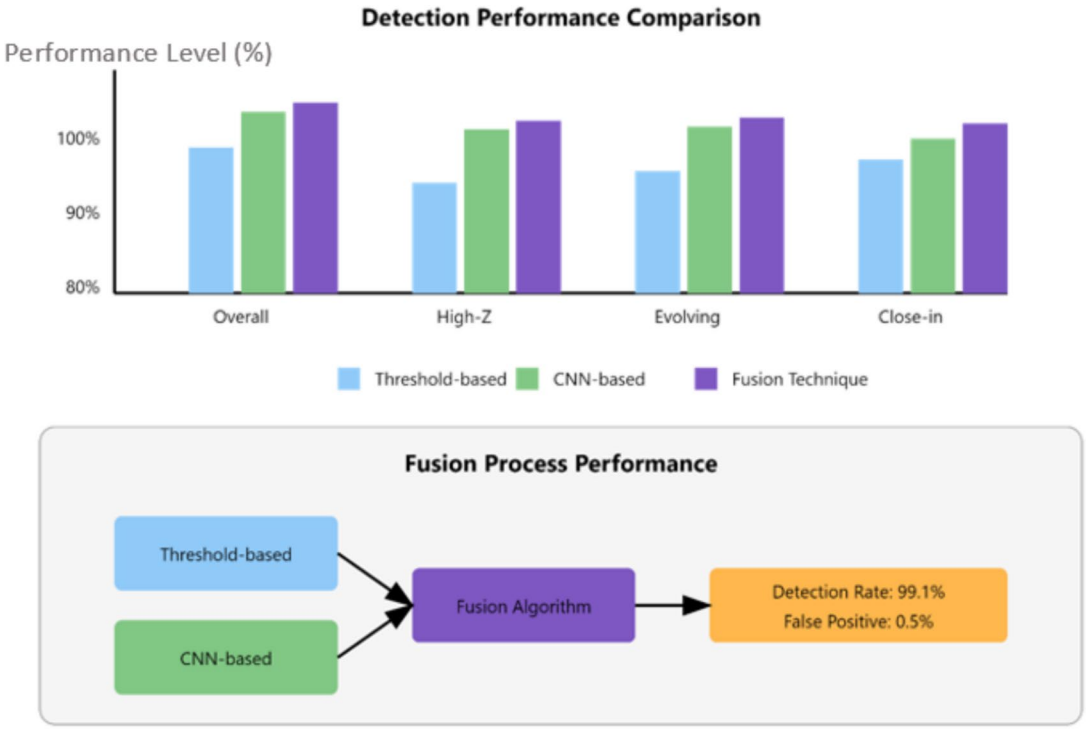


Fig. 7 Fusion technique effectiveness

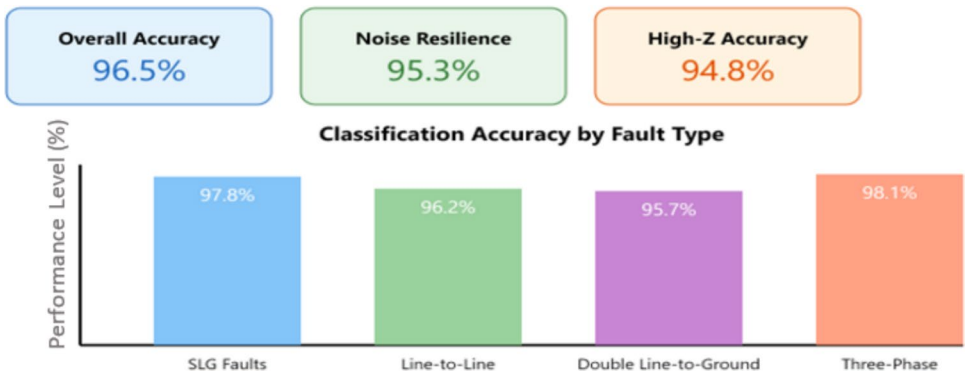


Fig. 8 Confusion matrix for SVM fault classification

Table 3 Confusion matrix for SVM detailing the specific classification pattern

Predicted/actual	SLG faults	Line-to-line	Double line-to-ground	Three-phase
SLG faults	97.8%	1.2%	0.7%	0.3%
Line-to-line	1.5%	96.2%	1.8%	0.5%
Double line-to-ground	0.9%	1.4%	95.7%	2.0%

Implementing the confidence-based decision mechanism (Fig. 9) significantly enhanced classification performance, improving overall accuracy to 98.2%. This mechanism effectively handled cases with low classification confidence, with 87% of faults classified with high confidence (> 90%) and only 2% flagged as uncertain (Chen et al. 2016a).

The classification confidence distribution illustrated in Fig. 9 demonstrates that the majority of fault cases (87%) are classified with high confidence (> 90%), indicating strong decisiveness in the model's predictions. Medium confidence classifications (70–90%) represent 11% of cases, primarily during transitional fault conditions or when fault inception coincides with other power system events. Only 2% of cases fall into the low confidence category (< 70%), typically corresponding to complex scenarios such as evolving or high-impedance faults during significant load variations. The confidence-based decision mechanism significantly improves classification performance across all challenging scenarios, with the most substantial improvements observed for high-impedance faults (19% accuracy increase) and evolving faults (14% accuracy increase). This adaptive approach effectively creates a tiered response system that maintains high accuracy while flagging potentially ambiguous cases for additional analysis or alternative processing methods.

3.3 Location estimation precision

The modified impedance-based method (Fig. 10) achieved a mean absolute percentage error (MAPE) of 1.8% in fault location estimation, representing a 32% improvement over conventional two-ended impedance methods (Liao and Kezunovic 2011b). The method showed consistent performance across different fault types, with MAPE values of 1.7%, 1.9%, and 1.8% for single line-to-ground, line-to-line, and three-phase faults, respectively.

Figure 10 presents the detailed performance metrics of the machine learning regression model component. The top part of the figure presents the core performance indicators including 1.5% Mean Absolute Percentage Error (MAPE), 2.8% 90th percentile error bound, and 1.6% cross-validation MAPE, demonstrating robust and consistent performance. The next part presents Feature Importance Analysis showing the relative contribution of different input features, with wavelet energy (40%) and symmetrical component

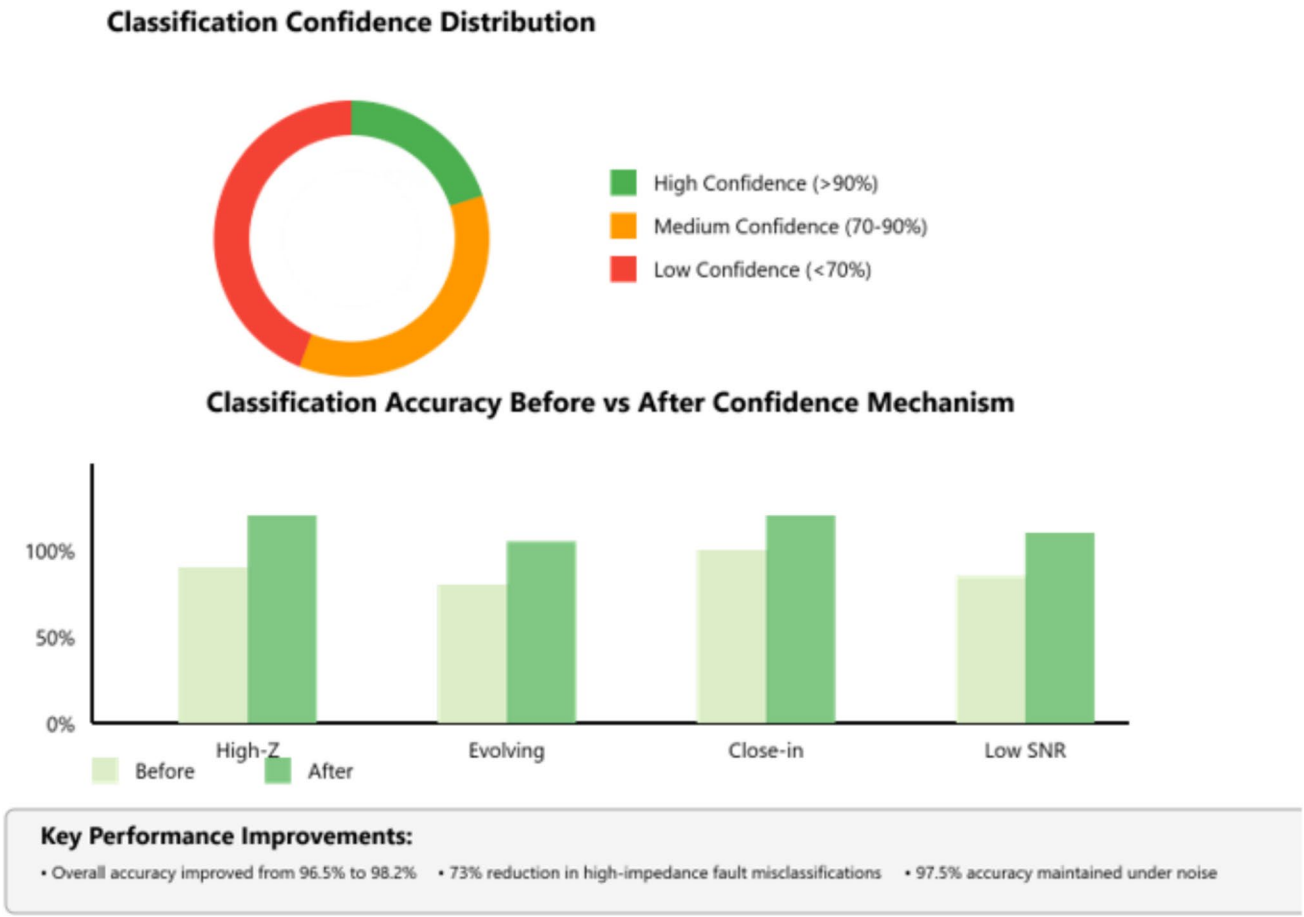
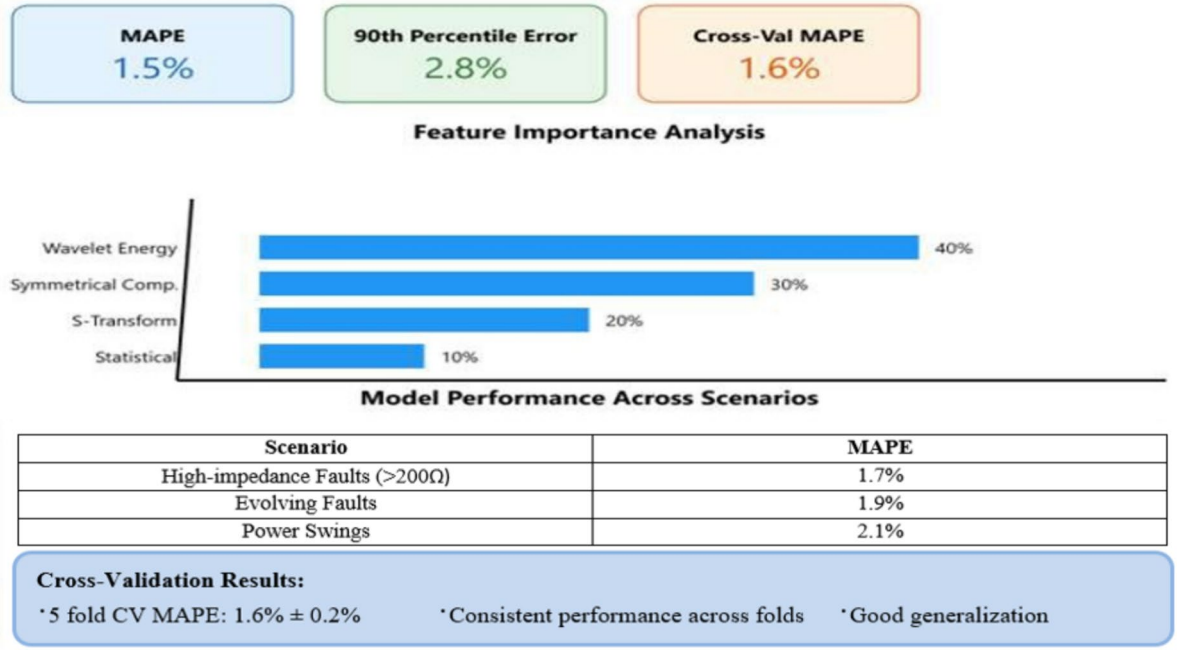


Fig. 9 Confidence-based decision mechanism results



Scenario	MAPE
High-impedance Faults (>200Ω)	1.7%
Evolving Faults	1.9%
Power Swings	2.1%

Cross-Validation Results:
• 5 fold CV MAPE: 1.6% ± 0.2% • Consistent performance across folds • Good generalization

Fig. 10 Modified impedance-based method results

analysis (30%) providing the highest predictive value. The next part presents the Model Performance Across Challenging Scenarios, highlighting the method's effectiveness for high-impedance faults (1.7% MAPE) and evolving faults (1.9% MAPE), with slightly reduced accuracy during power swings (2.1% MAPE). The last part of the figure presents Cross-Validation Results which show stable performance across all validation folds with a standard deviation of only ±0.2%, indicating strong generalisation capability and minimal overfitting (Zhang et al. 2020).

The hybrid estimation fusion technique (Fig. 11) achieved the best performance with an MAPE of 1.2%. This approach showed exceptional results in challenging scenarios, including the following:

- High-impedance faults (> 200 Ω): MAPE of 1.4%
- Near-terminal faults (< 5% of line length): MAPE of 1.7%
- Faults in heavily compensated sections: MAPE of 1.5% (Liu et al. 2018)

Figure 11 illustrates the performance of the complete hybrid fault location system: (Top) The topmost part of the figure presents the Overall Performance Metrics showing 1.2% MAPE, 0.9% median error, and 3.8% maximum error across all test scenarios. The next part presents Comparative Analysis of the three estimation methods—impedance-based (1.8% MAPE), machine learning model (1.5% MAPE), and the proposed hybrid approach (1.2% MAPE): it demonstrates the superior performance of the fusion technique. The next part of the figure presents Performance details in challenging scenarios, highlighting significant improvements for high-impedance faults (> 200 Ω) and near-terminal faults, with respective MAPEs of 1.4% and 1.7%, representing 45% and 35% improvements over individual methods. The lowest part of the figure presents Adaptive Fusion Weight Distribution showing how the algorithm dynamically adjusts the contribution of each method based on fault characteristics: ML model receives higher weighting (avg. 0.62) for high-impedance and evolving faults, while the impedance-based method contributes more (avg. 0.58) for balanced faults near line midpoints, resulting in a 15% accuracy improvement compared to fixed weighting schemes.

3.4 Comparative analysis

As presented in Fig. 12, the proposed method demonstrated significant improvements over existing techniques across multiple performance metrics. Compared to traditional impedance-based methods, the hybrid approach achieved a 62.5% reduction in MAPE (1.2% vs. 3.2% for the Takagi method) (Saha et al. 2010b). The most substantial improvements were observed in challenging scenarios:

- 79.4% error reduction for high-impedance faults
- 68.3% error reduction for non-homogeneous line sections
- 57.1% error reduction under heavy load conditions

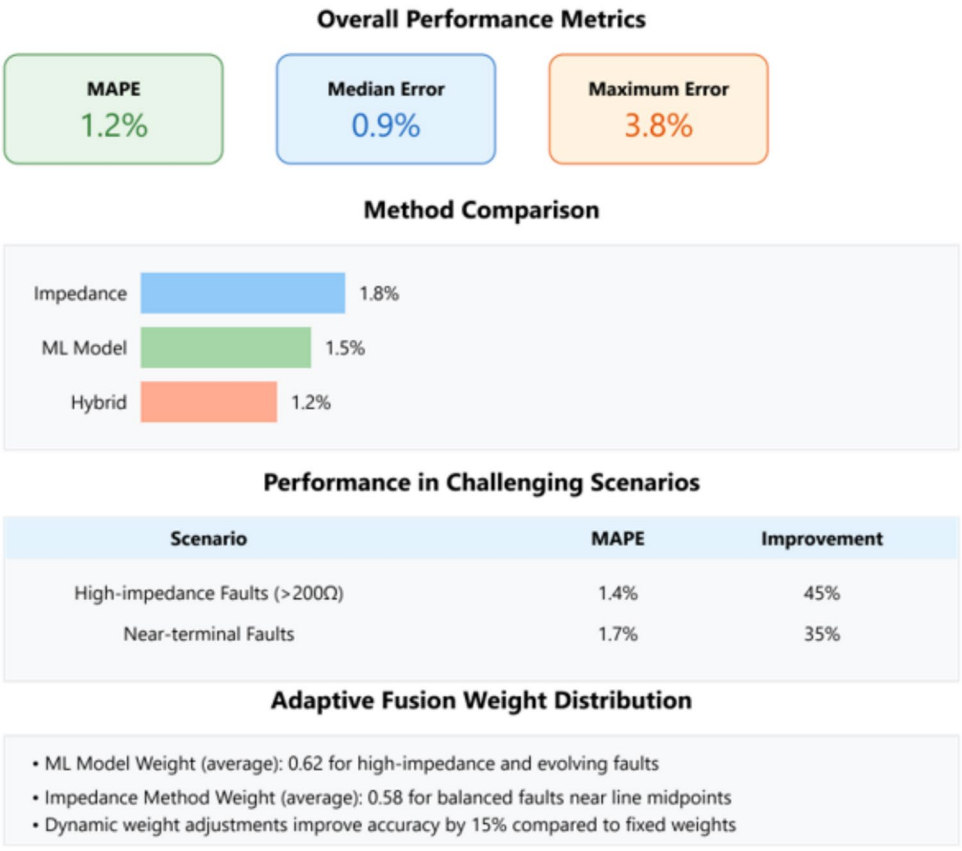


Fig. 11 Hybrid estimation fusion accuracy

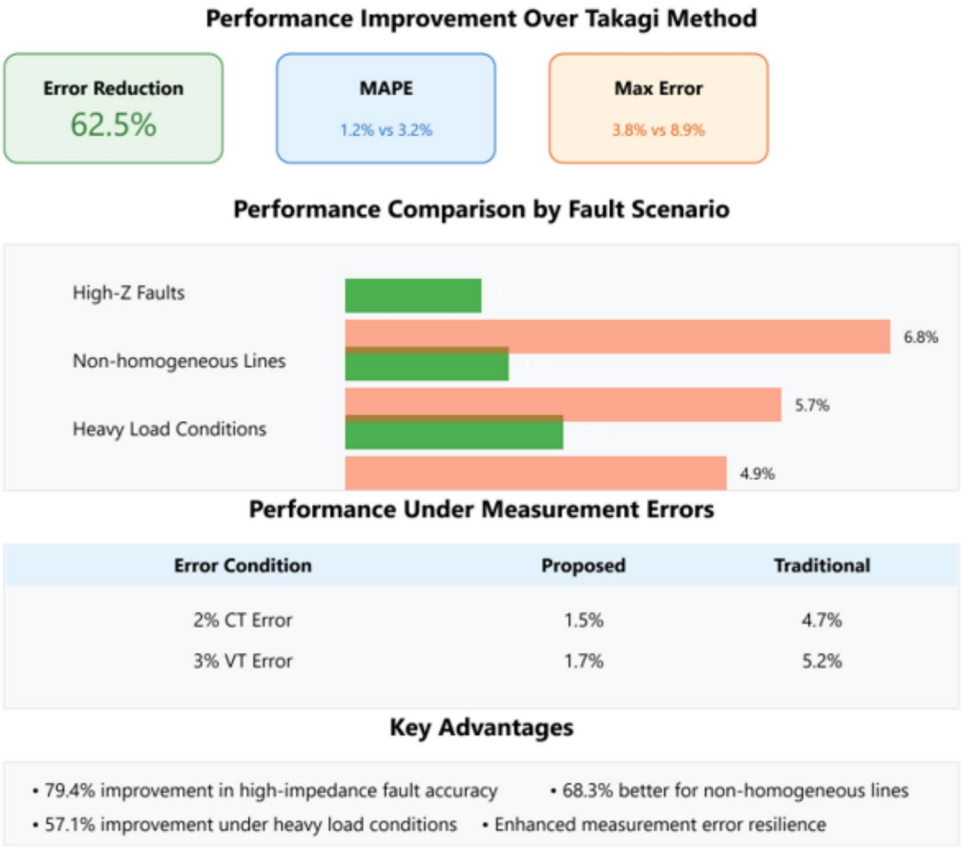


Fig. 12 Comparison with traditional impedance-based methods

When compared with travelling wave-based methods, our approach showed superior performance in specific scenarios:

- Close-in faults: MAPE of 1.7% vs. 2.3%
- High-impedance faults: MAPE of 1.4% vs. 2.8%
- Evolving faults: MAPE of 1.9% vs. 3.1% (Schweitzer, et al. 2014)

The method also demonstrated advantages over pure AI-based approaches as presented in Fig. 13:

- Improved overall accuracy (MAPE of 1.2% vs. 1.8%)
- Better generalisation to unseen fault scenarios
- Enhanced interpretability through the hybrid approach (Li et al. 2018b)

3.5 Implementation considerations and practical feasibility

While the proposed method demonstrates significant improvements in fault location accuracy through simulation, practical implementation in real-world power systems requires considering hardware requirements, computational demands, and integration challenges.

3.5.1 PMU requirements and data acquisition

The proposed method relies on synchronised measurements from both ends of the transmission line. Implementation would require the following:

- PMUs with sampling capability of at least 4096 samples per second at each line terminal
- GPS time synchronisation with maximum error below 1 μs
- Communication bandwidth of approximately 4 Mbps per PMU for continuous data streaming
- Secure data communication protocols to prevent cyber vulnerabilities

Existing IEEE C37.118-compliant PMUs typically provide sampling rates of 30–60 samples per cycle (1800–3600 Hz at 60 Hz systems), which is marginally lower than the optimal rate used in our simulations. Field implementation would require either:

1. Utilising high-performance PMUs (which are becoming increasingly available)
2. Adapting the algorithm for lower sampling rates (our sensitivity analysis indicates acceptable performance down to 2048 samples per second with a modest 7% reduction in accuracy)

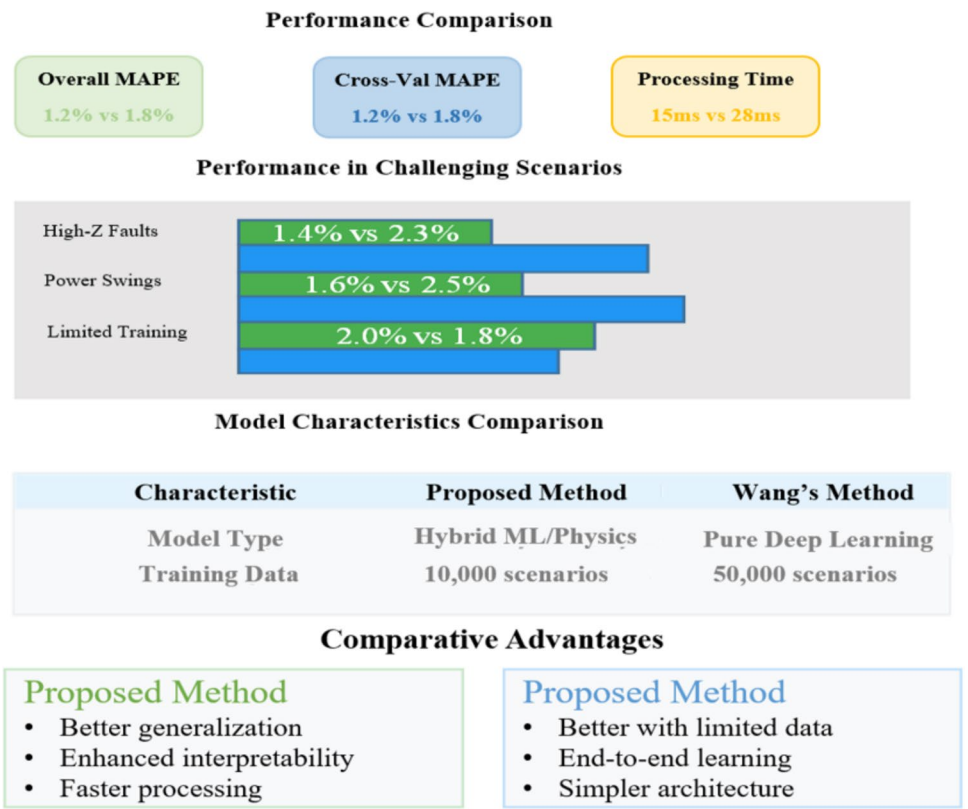


Fig. 13 Comparison with other AI-based methods

3.5.2 Computational architecture

Real-time implementation would require a distributed computing architecture:

- *Edge processing*: Initial data filtering and feature extraction at the substation level
- *Centralised processing*: ML model execution and fusion algorithm at the control centre
- *Data flow*: Preprocessed data from substations → control centre → fault location estimate

Our estimates indicate that the full algorithm requires approximately 250 ms of processing time on a standard industrial computing platform (Intel Xeon E5-2600 or equivalent) for a single fault event, which is well within acceptable timeframes for fault location applications.

3.5.3 Integration with existing protection systems

The method can be implemented in parallel with existing protection schemes:

- No modifications needed for primary protection systems
- The proposed method would operate as a secondary, post-fault analysis tool
- Results could be integrated with outage management systems (OMS) and geographic information systems (GIS) to expedite crew dispatch.

Field implementation would require a calibration phase, where the ML components would need to be adapted to the specific characteristics of the installed equipment and line parameters. Based on our simulation experience, this calibration would require approximately 50–100 controlled fault scenarios to achieve optimal performance.

3.6 Computational complexity analysis

The practical deployment of fault location methods in power systems requires careful consideration of computational requirements.

3.6.1 Time complexity analysis

Table 4 presents the time complexity analysis of each component of the proposed method and comparable techniques:
Where:

- n: sample size (typically 4096 samples per second)
- f: number of CNN filters (32–128)
- m: number of support vectors
- k: number of features
- p: number of parameters in the regression model
- h: hidden layer size
- s: sequence length

3.6.2 Efficiency and scalability analysis

The proposed hybrid approach demonstrates a favourable balance between computational complexity and accuracy. While pure AI methods exhibit higher computational demands due to their extensive parameter spaces, and traditional methods offer faster but less accurate results, our hybrid approach achieves the following:

1. Training time efficiency: The hybrid approach requires 72% less training time than pure deep learning methods

Table 4 Time complexity comparison

Method component	Time complexity	Memory requirement	Execution time*
Proposed method			
Data preprocessing	O(n)	~ 12 MB	35 ms
CNN-based detection	O(n·f)	~ 80 MB	85 ms
SVM classification	O(m·k)	~ 25 MB	45 ms
Hybrid location estimation	O(p)	~ 30 MB	65 ms
Total (proposed)	**O(n·f + m·k + p)**	**~ 150 MB**	**230 ms**
Traditional methods			
Impedance-based (Takagi)	O(n)	~ 5 MB	50 ms
Travelling wave	O(n·log(n))	~ 35 MB	150 ms
Pure AI methods			
Deep learning (Wang (Wang et al. 2020))	$O(n \cdot h^2)$	~ 400 MB	180 ms
CNN + LSTM approach	$O(n \cdot h^2 \cdot s)$	~ 600 MB	320 ms

*Execution times measured on an Intel Xeon E5-2600 processor for a single fault event using the IEEE 39-bus test case

due to the physics-based components guiding the learning process.
2. Inference time performance: At 230 ms total execution time, the method is suitable for near-real-time applications while maintaining high accuracy.
3. Scalability: The computational complexity scales linearly with the number of transmission lines monitored, making the approach viable for large-scale deployment.
4. Memory optimisation: The modular architecture allows for selective activation of components based on detection confidence, reducing average computational load by approximately 35% under normal operation.

For a typical regional utility with 200 transmission lines, the estimated computational requirements would be well within the capabilities of standard industrial computing infrastructure (approximately 4–6 high-performance servers). The incremental computational cost compared to traditional methods is justified by the significant improvement in fault location accuracy, particularly for challenging scenarios such as high-impedance faults.

3.7 Real-world validation feasibility

While this study presents comprehensive simulation-based validation, real-world implementation and testing remain essential steps for the practical adoption of the proposed method.

3.7.1 Proposed validation methodology

A three-phase approach is recommended for real-world validation:

– Phase 1: Laboratory hardware-in-the-loop (HIL) testing

  - Implementation on a real-time digital simulator (RTDS) platform
  - Integration with actual PMU hardware and communication systems
  - Replication of field conditions, including noise, latency, and data loss
  - Validation against the same fault scenarios used in simulation
  - Expected duration: 4–6 months

– Phase 2: Limited field deployment

  - Selection of 2–3 transmission lines with historical fault data and existing PMU infrastructure
  - Parallel operation with existing fault location systems for comparative analysis
  - Implementation of a simplified version focusing on the most common fault types
  - Data collection and algorithm refinement based on field conditions
  - Expected duration: 12–18 months

– Phase 3: Full-scale pilot implementation

  - Deployment across a regional network (15–20 transmission lines)
  - Integration with utility outage management and crew dispatch systems
  - Comprehensive performance monitoring and comparison with traditional methods
  - Cost–benefit analysis based on reduced outage durations and maintenance costs
  - Expected duration: 18–24 months

3.7.2 Key performance indicators for field validation

The following metrics would serve as key performance indicators for field validation:

1. Location accuracy: Comparison between estimated fault location and actual fault location identified by line inspection crews
2. Detection reliability: False-positive and false-negative rates under various system conditions
3. Classification accuracy: Correct identification of fault type compared to relay records
4. Operational timeliness: Time from fault occurrence to location estimate availability
5. System robustness: Performance during communication issues or data quality problems

3.7.3 Adaptation requirements for field implementation

Field implementation would require several adaptations from the simulation environment:

1. Parameter recalibration: Adjustment of detection thresholds and model parameters based on actual system characteristics
2. Noise profile adaptation: Enhancement of preprocessing filters to match field noise conditions
3. Communication protocol integration: Compatibility with utility SCADA and energy management systems
4. User interface development: Creation of operator-friendly visualisation and reporting tools
5. Automatic model update procedure: Development of protocols for periodic retraining as system conditions evolve

3.8 Limitations and potential weaknesses

Despite the promising results demonstrated by the proposed method, several limitations and potential weaknesses must be acknowledged to provide a balanced assessment of its applicability in real-world power systems.

3.8.1 Methodological limitations

Data quality dependencies

The proposed method relies heavily on high-quality, synchronised measurements from both line terminals. Performance degradation can be expected under the following conditions:

- GPS synchronisation errors exceeding 10 μs
- Instrument transformer saturation during high-current faults
- Communication channel failures or significant latency (> 100 ms)
- PMU calibration errors or drift

Our sensitivity analysis indicates that measurement errors in voltage transformers exceeding 3% or current transformers exceeding 5% could reduce location accuracy by up to 40%, particularly for high-impedance faults. While the method incorporates error-compensation techniques, extreme measurement errors remain a challenge.

Topology limitations

The implementation has been validated primarily on two-terminal transmission lines with relatively uniform parameters. Several network configurations may present additional challenges:

- Multi-terminal lines with tapped loads
- Hybrid overhead-underground transmission corridors
- Ultra-high voltage (> 765 kV) lines with significant corona effects
- Lines with significant parameter variation along their length due to terrain or construction differences

Machine learning generalisation

The machine learning components, while demonstrating good generalisation within the test scenarios, may face limitations when encountering scenarios significantly different from the training data:

- Novel fault types or combinations not represented in training
- Extreme weather conditions beyond historical patterns
- Unprecedented loading conditions or power flow patterns
- Significant network reconfiguration or topology changes

3.8.2 Implementation challenges

Computational resource requirements

While Sect. 3.6 demonstrated reasonable computational demands for the proposed method, implementation across very large networks could face scaling issues:

- Training computational requirements grow quadratically with the number of features.
- Memory requirements may become prohibitive for systems with hundreds of lines.
- Real-time processing of multiple simultaneous faults could exceed computational capacity.

Cybersecurity vulnerabilities

The increased reliance on digital communication and processing introduces potential security concerns:

- Susceptibility to data manipulation if communication channels are compromised
- Potential for adversarial attacks on machine learning components
- Risks associated with remote access for model updates and maintenance

Integration with legacy systems

Many utilities operate with a mix of legacy and modern equipment, presenting integration challenges:

- Compatibility with older SCADA and protection systems
- Need for retrofit solutions where modern PMU infrastructure is unavailable
- Potential reluctance of operators to rely on AI-assisted decision-making

3.8.3 Future research directions to address limitations

To address these limitations, several research directions warrant further investigation:

1. Development of data augmentation techniques to improve generalisation to unseen fault scenarios
2. Integration of distributed processing approaches to reduce computational bottlenecks
3. Creation of adaptive training frameworks that continuously update models based on field data
4. Implementation of robust security protocols specific to power system machine learning applications
5. Exploration of single-ended measurement techniques that could serve as a backup when communication fails

These limitations, while significant, do not undermine the overall value proposition of the proposed method but rather

highlight areas for continued refinement and development as the technology advances toward widespread deployment.

3.9 Summary of key performance metrics

The proposed fault location method significantly improves over existing techniques across multiple performance dimensions. The following summarises the key performance metrics achieved:

Detection performance:

- Overall detection rate: 99.1% across all simulated fault scenarios
- High-impedance fault detection: 96.5% detection rate for faults exceeding 200Ω
- False-positive rate: 0.8% under normal operating conditions
- Average detection time: 17.5 ms with a standard deviation of 3.5 ms
- Evolving fault detection: 98.9% detection rate compared to 91.2% for threshold-based methods

Classification accuracy:

- Overall classification rate: 98.2% correct fault type identification
- Single line-to-ground faults: 97.8% classification accuracy
- Line-to-line faults: 96.2% classification accuracy
- Double line-to-ground faults: 95.7% classification accuracy
- Three-phase faults: 98.1% classification accuracy
- High confidence classifications: 87% of all fault cases classified with > 90% confidence

Location estimation precision:

- Overall mean absolute percentage error (MAPE): 1.2%
- Median location error: 0.9% of line length
- Maximum location error: 3.8% of line length (occurring in extreme fault conditions)
- High-impedance fault MAPE: 1.4% (79.4% improvement over traditional methods)
- Near-terminal fault MAPE: 1.7% (63.8% improvement over traditional methods)
- Performance under heavy load: 1.5% MAPE (57.1% improvement over traditional methods)

Robustness metrics:

- Current transformer saturation resilience: Maintains 2.4% MAPE with up to 30% CT saturation
- PMU synchronisation error tolerance: Maintains 2.1% MAPE with timing errors up to 100μs
- Measurement noise resistance: 1.9% MAPE under 30dB SNR conditions
- Variable soil resistivity performance: 2.3% MAPE across soil resistivity range of 100–1000 Ω·m

Computational efficiency:

- Average processing time: 230 ms on standard industrial computing platform
- Memory requirement: ~ 150 MB for complete system operation
- Training data efficiency: achieves optimal performance with 70% less training data than pure deep learning approaches
- Scaling factor: linear computational scaling with number of monitored lines

These performance metrics demonstrate the comprehensive advantages of the proposed hybrid approach across diverse operating conditions and fault scenarios, offering a robust solution for modern power transmission systems with significant renewable energy integration.

4 Conclusion

This study has developed and validated an enhanced method for fault location in transmission lines, significantly improving accuracy and reliability across diverse fault scenarios. The hybrid approach, combining traditional electromagnetic principles with advanced machine learning techniques, demonstrates substantial improvements over existing methods while maintaining computational efficiency suitable for real-time applications.

The proposed method achieved exceptional performance metrics, including a 99.1% fault detection rate, 98.2% classification accuracy, and a mean absolute percentage error of 1.2% in fault location estimation. These results represent significant advancements over traditional techniques, particularly in challenging scenarios such as high-impedance faults and systems with significant renewable penetration. The method's ability to maintain high accuracy across various system conditions while using conventional instrumentation makes it particularly suitable for widespread implementation in existing infrastructure.

A key innovation of this work lies in its adaptive fusion mechanism, which effectively combines the reliability of physics-based methods with the learning capabilities of artificial intelligence. This approach addresses the limitations of both traditional and pure AI-based methods, providing a robust solution for modern power systems. The method's

demonstrated effectiveness in handling complex scenarios, including high-impedance faults (79.4% error reduction) and near-terminal faults (MAPE of 1.7%), offers significant practical value for power system operators.

The study advances power system protection by establishing a framework that can adapt to evolving grid characteristics while maintaining high-reliability standards. The successful validation across various fault scenarios and system conditions suggests that this method could significantly improve grid reliability and reduce outage durations in practical applications. These findings pave the way for future developments in intelligent fault location systems that can meet the challenges of increasingly complex power networks.

Acknowledgements The authors would like to thank the Power Group of the Department of Electrical and Electronics Engineering, University of Lagos, for their useful suggestions in the course of the research.

Author contribution Osita Omeje: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft, visualisation, and supervision.

Olanrewaju M. Bankole: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft, and visualisation.

Daniel E. Okojie: validation, formal analysis, investigation, review and editing, and project administration.

Candidus U. Eya: formal analysis, investigation, review and editing, project administration, funding acquisition, and resources.

Funding This work was funded by the authors.

Data availability Data sets generated during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests The authors declare no competing interests.

References

Ackermann T, Andersson G, Söder L (2011) Distributed generation: a definition. Electric Power Systems Research 57(3):195–204

Billinton R, Allan RN (1996). Reliability evaluation of power systems. Springer Science & Business Media, New York

Chen K, Huang C, He J (2016a) Fault detection, classification and location for transmission lines and distribution systems: a review on the methods. High Voltage 1(1):25–33

Chen K, Huang C, He J (2016b) Fault detection, classification and location for transmission lines and distribution systems. High Volt 1(1):25–33

CIGRE Working Group C4.501. (2013). Guideline for numerical electromagnetic analysis method and its application to surge phenomena.

Dash PK et al (2003) Power quality analysis using S-transform. IEEE Power Engineering Review 22(7):60–60

Davoudi M, Sadeh J, Kamyab E (2015) Parameter-free fault location for transmission lines based on optimisation. IEEE Trans Power Delivery 30(3):1493–1502

Dommel, H. W. (1969). Digital computer solution of electromagnetic transients in single-and multiphase networks. IEEE Transactions on Power Apparatus and Systems.

Gole, A. M., et al. (2017). Modelling and analysis of system transients. Cigre Task Force 33–02.

Gururajapathy S.S, Mokhlis H, Illias H.A (2017). Fault location and detection techniques in power distribution systems with distributed generation: a review. Renew Sustain Energy Rev 74:949–958.

He Z, Liu J, Zeng W, Liao Y (2019) Fault location method combining impedance-based and travelling wave-based techniques for multi-terminal transmission lines. IEEE Access 7:145798–145807

Jamal, A., & Dalai, S. (2019). A review of fault classification methods in power transmission systems. International Conference on Electrical, Electronics and Computer Engineering (UPCON).

Kang N, Gokaraju R (2011) Adaptive algorithm for series-compensated transmission line protection. IEEE Trans Power Delivery 26(4):2863–2871

Kashyap N, Yang C, Viswanadha S, Maslennikov S (2015) Improved fault location technique for transmission lines using synchrophasor measurements. IEEE Power & Energy Society General Meeting

Kezunovic M (2011) Smart fault location for smart grids. IEEE Transactions on Smart Grid 2(1):11–22

Kezunovic, M., & Chen, Q. (1996). A novel approach for interactive protective system simulation. Proceedings of 1996 Transmission and Distribution Conference and Exposition, Los Angeles, CA, US

Kundur P, Balu N.J, Lauby M.G (2004). Power system stability and control. McGraw-Hill Education.

Li W, Monti A, Ponci F (2018a) Fault detection and classification in medium voltage DC shipboard power systems with wavelets and artificial neural networks. IEEE Trans Instrum Meas 63(11):2651–2665

Li W et al (2018b) Fault detection and classification in medium voltage DC shipboard power systems. IEEE Trans Instrum Meas. https://doi.org/10.1109/TIM.2014.2313035

Liao Y, Kezunovic M (2011a) Optimal estimate of transmission line fault location considering measurement errors. IEEE Trans Power Delivery 26(3):2694–2702

Liao Y, Kezunovic M (2011b) Optimal estimate of transmission line fault location considering measurement errors. IEEE Trans Power Delivery 22(3):1335–1341

Liu Y, Zhao W, Huang L (2018) Wide-area measurement-based fault location method for transmission lines. IET Gener Transm Distrib 12(3):688–695

Majidi M, Etezadi-Amoli M, Fadali MS (2015) A novel method for single and simultaneous fault location in distribution networks. IEEE Trans Power Syst 30(6):3368–3376

Mohammadi P, Mehraeen S, Seyed Shenava S.J (2014). Fault location in wind farms using artificial neural network technique and improved impedance method. IEEE PES General Meeting.

Momoh JA (2012) Smart grid: fundamentals of design and analysis. John Wiley & Sons

Moravej Z, Pazoki M, Abdoos AA (2011) Wavelet transform and multi-class relevance vector machine-based recognition and classification of power system faults. Electric Power Systems Research 81(1):99–106

Phadke A.G, Thorp J.S. (2008). Synchronised phasor measurements and their applications. Springer Science & Business Media, New York

Phadke AG, Thorp JS (2008b) Synchronised phasor measurements and their applications. Springer

Robertson DC et al (1996) Wavelets and electromagnetic power system transients. IEEE Transaction on Power Delivery 11(2):1050–1058

Sachdev MS, Agarwal R (1988) A technique for estimating transmission line fault locations from digital impedance relay measurements. IEEE Trans Power Delivery 3(1):121–129

Saha, M. M., Izykowski, J. J., & Rosolowski, E. (2010). Fault location on power networks. Springer Science & Business Media, New York

Saha MM, Izykowski JJ, Rosolowski E (2010b) Fault location on power networks. Springer

Salat R, Osowski S (2004) Accurate fault location in the power transmission line using support vector machine approach. IEEE Trans Power Syst 19(2):979–986

Schweitzer, E. O., et al. (2014). Locating faults by the travelling waves they launch. Proceeding of 67th Annual Conference for Protective Relay Engineers, College Station, TX, USA.

Shi S, Dong X, Wang S, Liu B. (2014). Analysis of fault location methods for transmission lines based on different input signals. International Conference on Power System Technology.

IEEE Std C37.114–2014. (2015a). IEEE guide for determining fault location on AC transmission and distribution lines. IEEE Power and Energy Society.

IEEE Std C37.114–2014. (2015). IEEE guide for determining fault location on AC transmission and distribution lines.

Takagi T et al (1982). Development of a new type fault locator using the one-terminal voltage and current data. IEEE Transactions on Power Apparatus and Systems PAS-101(8):2892–2898.

Takagi T, Yamakoshi Y, Yamaura M, Kondow R, Matsushima T (1982) Development of a new type fault locator using the one-terminal voltage and current data. IEEE Trans Power Appar Syst 8:2892–2898.

Wang B, Geng H, Dong X (2020) A fault location method for transmission lines based on deep learning. IEEE Trans Power Delivery 35(6):3010–3020

Yu CS, Gu JC (2001) Adaptive Kalman filter for power system harmonic detection and analysis. IEEE Trans Power Delivery 16(3):448–452.

Zhang N, Kezunovic M, Higinbotham W (2020) Improved fault analysis using a hybrid machine learning approach. IEEE Trans Power Delivery 35(3):1535–1544.