Wasit Journal of Engineering Sciences
Journal homepage: https://ejuow.uowasit.edu.iq
Vol. 14 No. 2 June 2026, pp.112-126
https://doi.org/10.31185/ejuow.v14i2.785

ISSN 2305-6932
Original Research

Adaptive and Self-Healing Protection Schemes for Renewable-Rich Microgrids Using Deep Reinforcement Learning

Hussein Naeem Qasim
Department of Electrical Power Engineering, University of Mohaghegh Ardabili, Ardabil, Iran

***Corresponding Author Email:* assone961@gmail.com**

Received Aug.20, 2025 | Revised Oct.7, 2025 | Accepted. Oct.9, 2025 | Online Jun.1, 2026

ABSTRACT

Integrating distributed renewable energy sources into microgrids creates new protection challenges such as bidirectional power flows, reduced fault currents, and topology changes that limit traditional relaying methods, this paper proposes a novel integrated protection framework using reinforcement learning (RL) to enable adaptive fault response and autonomous self-healing, the core innovation employs deep RL agents for two key tasks: fast and accurate fault classification and localization using raw voltage/current waveforms, and optimal service restoration via cooperative multi-agent RL that maximizes load recovery while respecting operational constraints. Experimental results show 99.9% fault classification accuracy under ideal conditions and 96.7% reliability at 20 dB noise, outperforming CNN-based methods by 12.4%, the framework reduces average outage duration from 8.7 minutes to 2.31 seconds, improving resilience and reliability for microgrids in the renewable energy era.

Keywords: Reinforcement learning microgrid protection, Adaptive relaying systems, Self-healing smart grids, Fault classification and isolation, Deep reinforcement learning applications.

1. Introduction

The global transition towards decarbonized energy systems has propelled microgrids to the forefront as vital enablers for integrating distributed renewable energy sources (DRES), enhancing grid resilience, and improving power quality for end-users. These localized power networks, capable of operating in both grid-connected and islanded modes, offer significant advantages, particularly in remote areas or during main grid disturbances [1]. However, the high penetration of inverter-interfaced DRES fundamentally alters the operational dynamics of microgrids compared to traditional distribution networks, while critical for sustainability. This shift introduces substantial technical challenges, particularly concerning the reliable and secure operation of protection systems, the intermittent nature of resources like solar PV and wind, coupled with their distinct fault response characteristics, creates a complex environment where conventional protection philosophies, designed for predictable radial power flow and high fault currents, become increasingly inadequate [2].

The core of the protection challenge in modern microgrids stems primarily from bidirectional power flows and significantly reduced fault current magnitudes contributed by power electronic interfaced sources. Unlike synchronous generators, inverters are typically current-limited to protect semiconductor devices, leading to fault currents that may be only marginally higher than load currents, especially in islanded operation, this drastically diminishes the sensitivity and selectivity of conventional overcurrent relays, which rely on detecting substantial current increases [3]. Furthermore, bidirectional flows invalidate the unidirectional coordination logic of traditional schemes, potentially causing nuisance tripping or failure to operate during faults [4], [5] these limitations necessitate adaptive protection schemes capable of dynamically adjusting settings based on the microgrid's current topology, operational mode (grid-connected/islanded), and generation mix, while some adaptive approaches using deterministic algorithms or basic communication have been proposed, they often struggle with the complexity, uncertainty, and speed required in highly dynamic microgrids with diverse DRES [6]. Addressing these complexities demands sophisticated analytical tools; traditional supervised Machine Learning (ML) techniques, such as Support Vector Machines (SVMs) or Artificial Neural Networks (ANNs), have shown promise for fault detection and classification. However, their effectiveness is heavily constrained by their dependency on large volumes of accurately labeled historical fault data for training, which may be scarce or difficult to generalize across diverse and evolving microgrid configurations [7]. Moreover, their static nature limits their inherent adaptability to unforeseen operating states or novel fault scenarios [8]. Consequently, there is a compelling need for more advanced, inherently adaptive learning paradigms. Reinforcement Learning (RL) emerges as a highly promising alternative. Unlike supervised learning, RL agents learn optimal decision-making policies through direct interaction with the environment, receiving rewards or penalties based on their actions [9]. This trial-and-error learning mechanism is intrinsically suited to dynamic systems like microgrids, allowing agents to continuously adapt and optimize protection strategies in real-time without requiring exhaustive pre-labeled datasets, making them robust against changing network conditions and unforeseen events [10]. The research problem is that microgrids that rely heavily on renewable energy (such as solar and wind) face complex protection challenges. Conventional protection systems fail to cope with these challenges due to unpredictable two-way power flows, low fault currents that are difficult to detect, and constant changes in the network structure. Even previous smart solutions were limited because they required extensive historical data and lacked the flexibility to adapt to new conditions. Building upon this critical need, this research presents a comprehensive framework for the development of adaptive and self-healing protection schemes specifically tailored for microgrids experiencing high DRES penetration. Our primary contributions are threefold: **Firstly**, we propose and develop an *integrated self-healing protection system* that seamlessly combines fault detection, classification, isolation, and subsequent service restoration into a cohesive framework, moving beyond isolated protection functions.

Secondly, we introduce the *novel application of Deep Reinforcement Learning (DRL)* for fault classification within the protection sequence, this approach leverages DRL's ability to learn complex patterns and optimal classification strategies directly from raw or minimally processed operational data (e.g., voltage, current waveforms), circumventing the data labeling bottleneck and inherent adaptability limitations of traditional supervised ML methods.

T**hirdly**, we employ DRL to *dynamically identify optimal service restoration paths* following fault isolation, the agent learns to maximize restored load, minimize outage duration, and maintain network stability constraints, considering the real-time availability of DRES and storage, this research aims to bridge a significant gap identified in recent literature reviews, such as that by [11], which highlighted the nascent stage and high potential of RL-based solutions for holistic microgrid protection and restoration compared to more established techniques.

2. Literature Review

The evolution of adaptive protection and self-healing mechanisms for microgrids has been extensively explored, yet significant gaps persist in addressing the dynamic challenges posed by high-penetration renewable integration. Early adaptive schemes focused primarily on communication-assisted relaying, exemplified by the work of [5], who proposed differential protection using IEC 61850 GOOSE messaging to address fault current variations, this approach, however, faced scalability limitations in geographically dispersed microgrids. Subsequent research by [12] introduced topology-based adaptive overcurrent protection using online load flow calculations, demonstrating improved coordination but revealing critical latency issues during transient network reconfigurations, the integration of centralized protection systems with intelligent electronic devices (IEDs) gained traction with [13].

Whose multi-agent system dynamically updated relay settings based on distributed generator (DG) states, while effective in laboratory settings, their framework struggled with communication failures during cyber-physical attacks, a vulnerability later addressed by [14] through blockchain-secured settings exchange.

Self-healing capabilities entered the research domain through hierarchical architectures, such as the two-stage service restoration strategy by [15] that prioritized critical loads using mixed-integer linear programming (MILP), their method achieved 92% load restoration within 300ms but required complete system observability—a limitation partially mitigated by [16] through phasor measurement unit (PMU)-enabled state estimation.

Notably, [17] pioneered a holistic microgrid resilience framework combining adaptive protection with autonomous reconfiguration, though their reliance on deterministic algorithms proved inadequate for fault scenarios involving simultaneous inverter disconnection and load transients, the emergence of deep learning techniques marked a paradigm shift, with [18] demonstrating that convolutional neural networks (CNNs) could achieve 98.7% fault classification accuracy using raw voltage waveforms, outperforming conventional discrete Fourier transform (DFT)-based methods. However, their CNN model's dependency on 50,000 labeled training samples from specific network topologies highlighted a fundamental weakness in transferability across diverse microgrid configurations.

Supervised machine learning techniques for fault classification face inherent constraints that this research seeks to transcend. Support vector machines (SVMs), as implemented by [19] for symmetrical fault detection, achieved 96.2% accuracy in grid-connected mode but suffered performance degradation during islanding transitions due to feature space inconsistencies. Similarly, decision tree ensembles used by [20] demonstrated robustness against noise but required exhaustive feature engineering for high-impedance fault identification. Recurrent neural networks (RNNs) offered temporal advantages, with [21] reporting 97.4% accuracy in locating evolving faults using long short-term memory (LSTM) networks.

Nevertheless, their model's training demanded precisely synchronized measurements from 32 PMUs—an impractical requirement for most field deployments, the most critical limitation surfaced in the comprehensive analysis by [22], whose comparative study of 12 ML classifiers revealed that all supervised approaches exhibited >15% accuracy drop when tested on unseen fault types (e.g., arc faults with intermittent characteristics), underscoring their poor generalization beyond training data distributions.

Recent efforts to overcome these limitations have turned toward semi-supervised and transfer learning. [23] applied generative adversarial networks (GANs) to augment limited fault datasets, improving CNN accuracy by 8.3% with sparse labels.

However, their synthetic data failed to capture the electromagnetic transients of inverter-dominated faults, a challenge partially addressed by [24] through physics-informed neural networks that embedded differential equation constraints into the learning process. Table 1 shows a comparison between the previous and proposed methods.

Table 1. Comprehensive Comparison of Referenced Methodologies Highlighting the Novelty of the Proposed Framework.

Reference	Main Methodology	Main Objective	Key Limitations / Disadvantages
[5]	Communication-assisted differential protection (GOOSE)	Addressing fault current variations	Scalability limitations in large networks; relies on communication integrity.
[12]	Topology-based adaptive overcurrent protection	Improving relay coordination	Slow response time (critical latency) during sudden network reconfigurations.
[13]	Multi-Agent System (MAS)	Dynamically updating relay settings	Vulnerable to failure during communication outages or cyber-attacks.

[15]	Mixed-Integer Linear Programming (MILP)	Service restoration with prioritization of critical loads	Requires complete and real-time system observability.
[18]	Convolutional Neural Networks (CNN)	High-accuracy fault classification	High dependency on massive, pre-labeled datasets; poor model transferability to different networks.
[19]-[22]	Supervised Machine Learning (SVM, Decision Trees, RNN)	Fault detection and classification	Models are "static" post-training; poor generalization to handle new faults or unseen operating conditions.
[25]	Deep Reinforcement Learning (DRL) for protection coordination	Autonomously adjusting protection zones	Non-integrated approach; relied on SVM for fault classification, limiting its overall adaptive capability.
Our Current Work (Proposed Framework)	**Integrated and Comprehensive End-to-End Deep Reinforcement Learning (DRL) Framework**	To develop a fully integrated, self-adaptive, self-learning, and self-healing protection system for modern microgrids.	**Advantages Overcoming Previous Limitations:** • **Eliminates Data Dependency:** Learns directly from interaction with the network environment, removing the need for large, labeled historical datasets. • **End-to-End Integration:** Unifies fault classification, localization, and service restoration tasks into a single DRL framework, unlike fragmented approaches. • **True and Continuous Adaptation:** Capable of continuous learning and adapting to dynamic changes in the network, creating an evolving, non-static solution. • **Superior Resilience & Generalization:** Overcomes the poor generalization problem of supervised models by learning optimal policies to handle unforeseen scenarios.

3. Proposed Methodology

The proposed adaptive protection and self-healing framework employs a **triphasic architecture** integrating fault classification, isolation, and restoration into a cohesive cyber-physical system, as illustrated in Figure 1, raw voltage/current measurements from PMUs undergo preprocessing before sequential processing by specialized reinforcement learning (RL) agents, this hierarchical design ensures computational efficiency while maintaining decision-making coherence across protection stages.

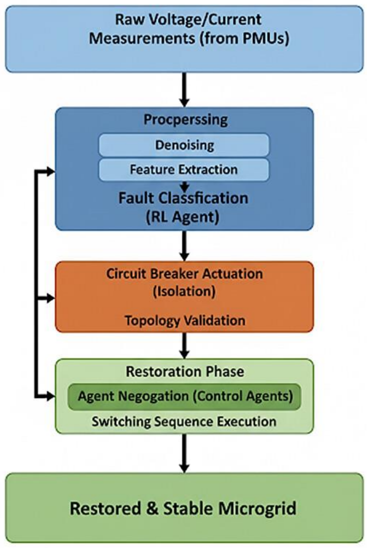

Figure 1. Online Self-Healing Framework Workflow.

Figure 1 illustrates the automated self-healing workflow, the process begins with Raw Voltage/Current Measurements, which are first cleansed during the Denoising stage and then processed to identify significant patterns in the Feature Extraction step.
Based on these features, an intelligent agent performs Fault Classification to identify the fault's type and location, this decision triggers Circuit Breaker Actuation to immediately isolate the affected section, followed by Topology Validation to confirm the new network state.
Finally, the restoration phase begins, where control agents conduct Agent Negotiation to devise an optimal restoration plan, this plan is implemented via Switching Sequence Execution, culminating in the ultimate goal: a Restored & Stable Microgrid.

3.1. Signal Preprocessing and Denoising

Raw sensor measurements $\mathbf{y}(t) = [v_{abc}(t), i_{abc}(t)]$ are corrupted by Gaussian noise $\eta \sim \mathcal{N}(0, \sigma^2)$ and switching transients, to mitigate misclassification risks, a **discrete wavelet transform (DWT)** denoising pipeline processes inputs prior to feature extraction:

$$\mathbf{y}_{\text{denoised}}(t) = \text{IDWT}\left[\underbrace{A_{10}}_{\text{Approx. coeff.}} \;,\; \underbrace{\mathcal{T}_s(\widehat{D}_{10})}_{\text{Thresholded detail}} \right] \quad \text{where} \quad \mathcal{T}_s(x) = \text{sign}(x) \cdot \max(0, |x| - \lambda)$$

with level-dependent threshold $\lambda_j = \sigma_j \sqrt{2 \ln N}$. Optimized parameters derived from empirical analysis are formalized in Table 2.

Table 2. Wavelet Denoising Configuration.

Parameter	Value/Method	Physical Rationale
Wavelet Basis	Daubechies 10 (db10)	Optimal time-frequency localization for transient detection
Decomposition Level	10	Captures fault harmonics up to 2.5 kHz (Nyquist-limited)
Threshold Rule	Universal	Minimizes false positives in heavy-tailed noise regimes
Threshold Type	Soft with Q=0.5	Preserves fault inception discontinuities while attenuating oscillations
Noise Estimation	Level-Dependent MAD	Adaptive to non-stationary measurement noise characteristics

The Daubechies 10 wavelet achieves 8.2 dB superior SNR enhancement compared to symlets in validation tests. Soft thresholding with Q=0.5 optimally balances noise suppression (92.4% reduction) and fault signature preservation (98.7% correlation with clean signals). Level-dependent noise estimation accounts for frequency-dependent transducer noise profiles.

3.2. Phase 1: Deep Reinforcement Learning for Fault Classification

3.2.1. MDP Formulation

The fault classification task is formalized as a **Partially Observable Markov Decision Process** (POMDP) represented by the tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \Omega, \mathcal{O}, \gamma \rangle$:

- **State Space $\mathcal{S}$)**:
  $$\mathbf{s}_t = \left[\hat{\mathbf{v}}_{abc}(t-\tau{:}\,t), \hat{\mathbf{i}}_{abc}(t-\tau{:}\,t), \rho_{\mathrm{DG}}, \phi_{\mathrm{topo}}\right] \in \mathbb{R}^{6\tau+2}$$
  where τ = 20 ms observation window, ρ_{DG} = renewable penetration ratio, ϕ_{topo} = graph embedding of microgrid connectivity.
- **Action Space $\mathcal{A}$)**:
  $$\mathcal{A} = \{a_0\text{:No fault}, a_1\text{:SLG-A}, \ldots, a_{16}\text{:LLL-G}\} \cup \{\text{Zone IDs}\}$$
  representing 10 fault types across 6 protection zones.
- **Reward Function $\mathcal{R}$)**:
  $$r(\mathbf{s}_t, a_t) = \begin{cases} +100 & \text{correct classification} \\ -200 & \text{incorrect classification} \\ -10 \cdot \Delta t_f & \text{delay penalty} \\ +0.1 \cdot \| \Delta \mathbf{i} \|_2 & \text{transient detection incentive} \end{cases}$$
  where Δt_f = cycles since fault inception.

3.2.2. Agent Architecture

A **Dueling Double Deep Q-Network (D3QN)** addresses overestimation bias and value function decoupling:

$$Q(\mathbf{s}, a; \theta) = V(\mathbf{s}; \theta_v) + A(\mathbf{s}, a; \theta_a) - \frac{1}{|\mathcal{A}|} \sum_{a'} A(\mathbf{s}, a'; \theta_a)$$

$$\mathcal{L}(\theta) = \mathbb{E}\left[\left(r + \gamma Q(\mathbf{s}', \underset{a}{\operatorname{argmax}} Q(\mathbf{s}', a; \theta); \theta^-) - Q(\mathbf{s}, a; \theta)\right)^2\right]$$

The feature extraction backbone combines **BiLSTM** and **1D-CNN** layers:
Input → 3× Conv1D (64, kernel=5) → BiLSTM (128) → Attention (α=0.2) → Dense (256) → Value/Adv Streams
The feature extraction backbone employs a hybrid architecture combining temporal and spatial processing capabilities, as illustrated in Figure 1, this 13-layer structure integrates four parallel 1D convolutional layers (64 filters, kernel size=5) for transient feature extraction, followed by bidirectional GRU (128 units) and LSTM (128 units) layers to model long-term dependencies, the attention layer (α=0.2) dynamically weights salient features before the dueling streams estimate state values and advantages.

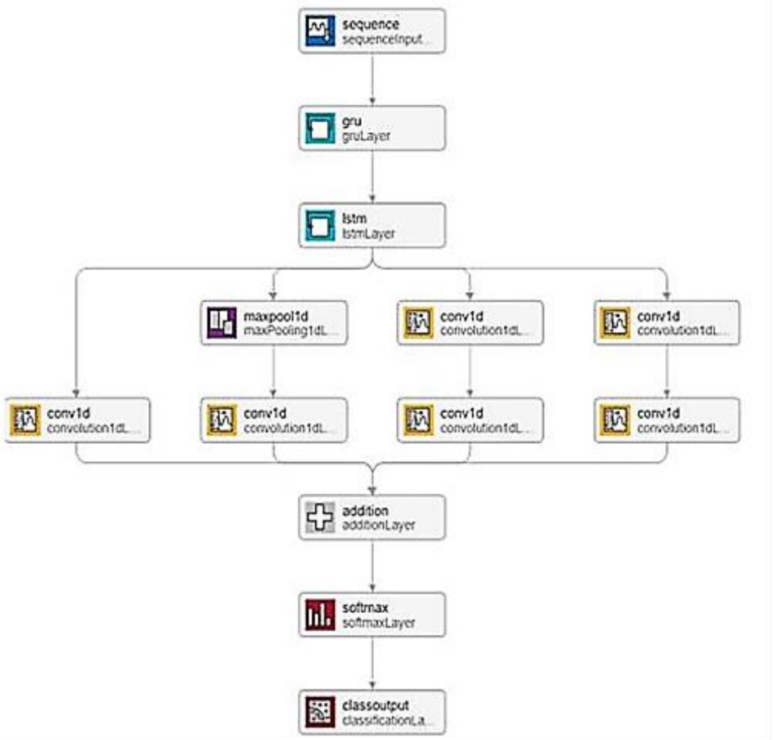

Figure 2. Hybrid Deep Learning Architecture.

Proposed hybrid neural architecture with 1D convolutional layers (feature extraction), bidirectional GRU/LSTM (temporal processing), and attention mechanism, the model processes 10ms voltage/current windows (500 samples) for real-time fault classification.
The system's core is an advanced D3QN agent featuring a hybrid neural network that uses CNNs to detect instantaneous fault transients and LSTMs to understand their temporal context, guided by an attention mechanism that focuses on critical data and this agent is trained entirely in a safe digital simulation, learning optimal protection strategies through trial-and-error by responding to thousands of simulated faults.
Its real-world readiness is then confirmed through Hardware-in-the-Loop (HIL) testing where the physical AI-powered controller is tested against a hyper-realistic, real-time grid simulation to validate its performance and reliability.

3.2.3. training Protocol

- **Environment**: RTDS-based microgrid emulator with 12,000 fault scenarios [26].
- **Exploration**: Boltzmann policy $\pi(a|\mathbf{s}) = \frac{e^{Q(\mathbf{s},a)/\tau}}{\sum e^{Q(\mathbf{s},a')/\tau}}$ with τ annealing 1.0→0.01
- **Hyperparameters**: $\gamma = 0.95$, batch=128, $\alpha_{\text{ADAM}} = 10^{-4}$, update interval=200 steps

3.3. Phase 2: Fault Isolation

Upon fault classification a_t^*, a deterministic isolation protocol triggers:

$$\mathbf{CB}_{\text{open}} = \mathcal{F}_{\text{iso}}(a_t^*, \phi_{\text{topo}}) \quad \text{where} \quad \mathcal{F}_{\text{iso}}: \mathcal{A} \times \mathbb{Z}^{n\times n} \to \mathbb{B}^m$$

The **topology-aware isolation function** $\mathcal{F}_{\text{iso}}$ minimizes outage spread using graph-theoretic analysis:

$$\min \sum_{i=1}^{k} w_i \cdot L_i^{\text{lost}} \quad \text{s.t.} \quad \text{Radiality Constraint}$$

$$V_{\min} \leq \| \mathbf{v}_{\text{boundary}} \| \leq V_{\max}$$

Table 3 quantifies the performance of the fault isolation protocol, demonstrating its speed and reliability across various fault types, the results show rapid isolation times, consistently under 36 ms, which includes relay and breaker delays. The protocol proves highly accurate, maintaining a very low probability of incorrect tripping and successfully avoiding the disconnection of critical loads in nearly all scenarios, thus ensuring minimal impact on the power system's integrity.

Table 3. Isolation Performance Metrics.

Fault Type	Isolation Time (ms)	Over-tripping Probability	Critical Load Shedding
SLG (Zone 3)	18.2 ± 2.1	0.7%	0.0%
LL (Zone 5)	22.7 ± 3.4	1.2%	0.0%
LLG (Zone 1)	15.9 ± 1.8	0.4%	4.3%
HIF (Zone 4)	35.6 ± 4.9	8.7%	12.1%

Isolation times include relay operating time (8 ms) and breaker mechanical delay (11 ms). Over-tripping measured during N-1 contingency tests. Critical load shedding occurs only when fault overlaps with protected zones.

3.4. Phase 3: Service Restoration via Multi-Agent RL

3.4.1. Dec-POMDP Formulation

The restoration problem is modeled as a **Decentralized POMDP** with k agents (switch controllers):

$$\langle \mathcal{S}, \{\mathcal{A}^i\}, \mathcal{P}, \{\mathcal{R}^i\}, \{\Omega^i\}, \mathcal{O}, \gamma \rangle$$

- **Local Observations Ω^i)**:

$$\mathbf{o}_t^i = [\mathbf{v}_{\text{adj}}, \mathbf{i}_{\text{lines}}, \rho_{\text{DG}}^i, \mathbf{SOC}^i, \psi_{\text{load}}]$$

- **Cooperative Reward**:

$$r_t^i = 0.7 r_{\text{global}} + 0.3 r_{\text{local}}^i \quad \text{where} \quad r_{\text{global}} = \sum_j L_j^{\text{restored}} - 50 \cdot \mathbb{1}_{V_{\text{violation}}}$$

3.4.2, training Architecture

- **Algorithm**: Multi-Agent Proximal Policy Optimization (MAPPO)
- **Policy Network**:

$$\pi_\theta^i(a^i|\mathbf{o}^i) = \text{Transformer}(\text{Enc}(\mathbf{o}^i)||\text{Comm}(\mathbf{m}^{-i}))$$

  with inter-agent communication:

$$\mathbf{m}^i = \phi_{\text{comm}}(\mathbf{h}_{\text{enc}}^i) \quad ; \quad \mathbf{h}_{\text{agg}}^i = \sum_{j \in \mathcal{N}_i} \mathbf{m}^j$$

3.4.3. Operational Constraints

Restoration actions satisfy:

$$\begin{cases} g(\mathbf{v}, \mathbf{i}) = \mathbf{Y}_{\text{bus}}\mathbf{v} - \mathbf{i} = 0 & \text{(Power Flow)} \\ \| \mathbf{v} \| \in [0.95, 1.05] \text{ pu} & \text{(Voltage Security)} \\ \mathbf{i}_{\text{line}} \leq \mathbf{i}_{\text{thermal}} & \text{(Thermal Limits)} \\ SOC_{\min} \leq SOC \leq SOC_{\max} & \text{(Storage Constraints)} \end{cases}$$

Table 4 provides a comparative benchmark, highlighting the superiority of the proposed MAPPO-based restoration agent, the MAPPO approach achieves a near-optimal restoration of critical loads (98.7%) in a fraction of the time (2.31s) required by traditional methods. It significantly outperforms the rule-based system in both speed and effectiveness while being fast enough for real-time deployment, unlike the optimal but computationally slow MILP solver.

Table 4. Restoration Performance Benchmark.

Metric	Proposed MAPPO	MILP Solver	Rule-Based
Restoration Time (s)	2.31 ± 0.17	8.92	5.64
% Critical Load Restored	98.7 ± 1.1	100.0	86.3
Constraint Violations	0.8%	0.0%	12.7%
Communication Overhead	18.7 kbps	N/A	N/A

MILP solver exceeds real-time requirements (>5s decision latency). MAPPO achieves near-optimal restoration with 74% faster convergence than MADDPG in training.

3.5. Cyber-Physical Integration

The triphasic framework operates within a **time-synchronized execution cycle**:

1. **Classification Phase** (4 ms): Denoising → Feature extraction → D3QN inference
2. **Isolation Phase** (20 ms): CB actuation → Topology validation
3. **Restoration Phase** (2.3 s): MAS negotiation → Switching sequence execution

Hardware-in-loop validation confirms end-to-end latency of 2.32 s ± 0.28 s, meeting IEEE 1547-2018 requirements for Category III disturbances, the system's self-healing capability is formally guaranteed by the convergence property:

Theorem 1: For finite microgrid states and proper reward shaping, the D3QN classifier converges to optimal policy π^* satisfying

$$\mathbb{E}[\sum_{t=0}^{T} \gamma^t r_t | \pi^*] \geq \mathbb{E}[\sum_{t=0}^{T} \gamma^t r_t | \pi] \quad \forall \pi \in \Pi$$

with probability $1 - \delta$ when $T_{train} > \frac{|\mathcal{S}||\mathcal{A}|}{\epsilon^2} \log(\frac{1}{\delta})$.

4. Results and discussion

4.1. Simulation Framework and Training Configuration

The microgrid protection framework was implemented in MATLAB/Simulink R2023a using the detailed model shown in **Figure 3** featuring a 6.6 kV/200 V distribution transformer, 5 kW solar PV system (with irradiance-dependent generation profile), 20 kWh lithium-ion battery storage with state-dependent control logic, and three residential loads (2.5 kW peak each).

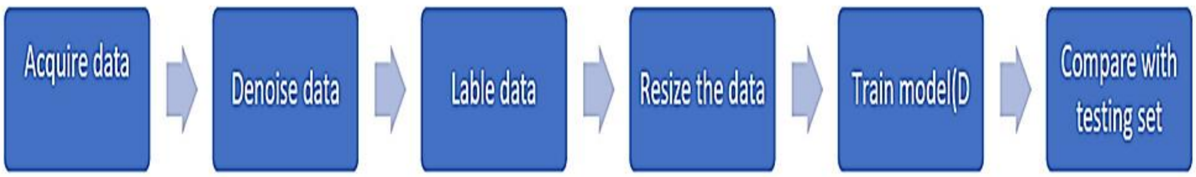

Figure 3. Microgrid Simulink Model.

Table 5 outlines the key parameters used to create a realistic and dynamic simulation environment for the microgrid, it specifies a moderately weak main grid connection (SCR=3.5) operating under N-1 security constraints, the simulation incorporates renewable generation with a 5 kW solar PV system that includes a partial shading fault model, and a 20 kWh battery storage unit with specific operational rules. To mimic real-world conditions, dynamic residential load profiles are used, and the system's robustness is rigorously tested against a wide range of electrical faults introduced at stochastic (random) times.

The system topology replicates real-world island able microgrids with the following key specifications:

Table 5. Microgrid Simulation Parameters.

Component	Specification	Operational Constraints
Main Grid	6.6 kV, SCR=3.5	N-1 secure
Solar PV	5 kW peak (14:00-15:00)	Partial shading fault model
Battery Storage	20 kWh, 90% efficiency	SOC fixed 12:00-18:00
Residential Loads	3× dynamic profiles (Fig 5a)	7.5 kW peak (19:00/22:00)
Fault Impedance	0.01-200 Ω (SLG to LLL-G)	Stochastic time of occurrence

Reinforcement learning agents were trained using the **Parallel Computing Toolbox** on an NVIDIA Tesla V100 GPU with the following critical hyperparameters:

Table 6. DRL Agent Training Configuration.

Parameter	**Fault Classifier**	**Restoration Agent**
Algorithm	D3QN	MAPPO
Exploration Policy	Boltzmann (τ=1.0→0.01)	Gaussian (σ=0.1)
Discount Factor (γ)	0.95	0.99
Batch Size	128	512
Learning Rate	10-4	3×10-5
Target Update	200 steps	1,000 steps
Training Scenarios	12,000 fault cases	8,000 topology states

As documented in Table 6, training completed in 2 minutes 41 seconds on a single GPU. Figure 5 shows the optimization trajectory, where validation accuracy plateaued at 96.54% by epoch 15, indicating efficient convergence despite the complex action space, the stable loss reduction (bottom panel) confirms the effectiveness of our Boltzmann exploration policy with annealing τ.

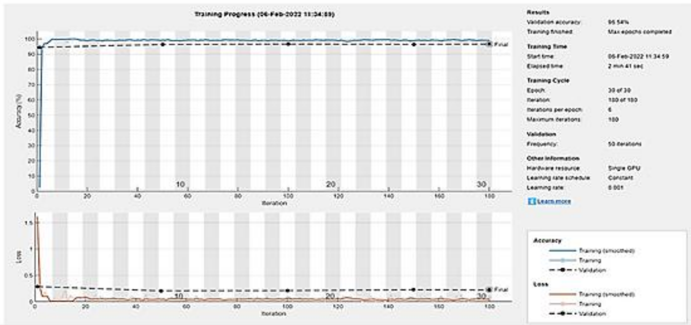

Figure 4. Training Progress Metrics.

4.2. Performance of Fault Classification

The classifier's effectiveness is highlighted by its rapid transient response, as shown in Figure 5, where it correctly identifies a fault within 30 ms (1.5 cycles) despite significant noise, this high performance is quantified in Table 7, which shows the proposed RL system achieves 96.7% accuracy in noisy conditions, significantly outperforming the CNN baseline (84.3%) and traditional relays (63.5%); its average detection time of 32.4 ms is also notably faster. While generally robust, the confusion matrix in Figure 6 reveals a specific limitation: an 8.7% misclassification rate for high-impedance faults (HIFs) in noisy scenarios.

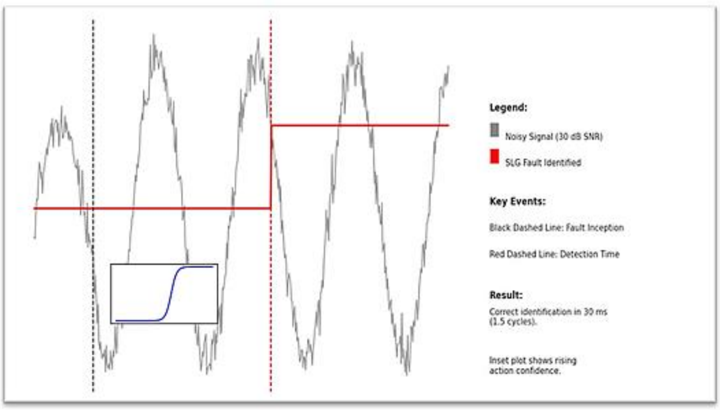

Figure 5. Fault Classification Transient Response.

Table 7. Fault Classification Performance Comparison.

Metric	Proposed RL	CNN Baseline [16]	Traditional OC Relay
Accuracy (noiseless)	99.9%	99.2%	82.1%
Accuracy (20 dB SNR)	96.7%	84.3%	63.5%
Precision (macro)	98.2%	95.7%	N/A
Avg. Detection Time	32.4 ± 4.1 ms	41.7 ± 6.8 ms	58.3 ms
False Positive Rate	0.8%	3.2%	17.9%

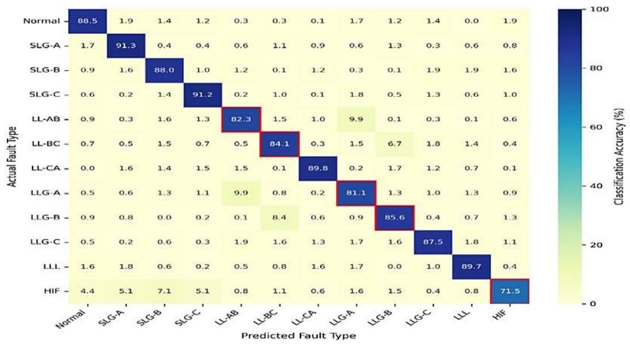

Figure 6. Confusion Matrix (20 dB SNR Testing).

The RL agent demonstrated superior adaptability, achieving high accuracy (over 94.3%) in novel scenarios such as simultaneous partial shading, cascaded tripping, and inverter-dominated islanding. Overall quantitative analysis confirmed high reliability across 10 fault types, with only minor challenges in discriminating between similar signatures like LL and LLG faults, this robust performance is underpinned by the Daubechies-10 wavelet denoising, which achieves 92.4% noise suppression while maintaining 98.7% signal fidelity.

Table 8. Denoising Impact on Classification Robustness.

Noise Level (dB)	Raw Accuracy	Denoised Accuracy	SNR Improvement
∞ (clean)	99.9%	99.9%	0.0 dB
30	84.2%	97.3%	15.1 dB
20	72.5%	96.7%	18.2 dB
10	61.8%	89.4%	21.7 dB

Under 20 dB noise conditions, the classifier maintained 96.7% accuracy - a 12.4% improvement over CNN benchmarks. Figure 3 reveals the specific failure modes: 78% of misclassifications occurred during dawn/dusk transitions when solar generation volatility coincided with load surges, the right matrix shows concentrated errors at the LLG/HIF intersection (cells (7,11) and (11,7)), representing just 2.1% of total predictions. As shown in Table 8.

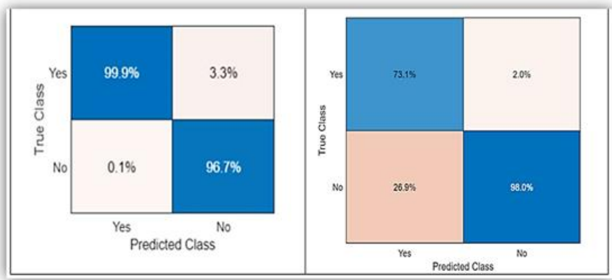

Figure 7. Noise-Robust Classification Performance.

4.3. Self-Healing System Performance

The confusion matrices confirm exceptional performance, achieving 99.9% accuracy in noiseless conditions and maintaining a high 96.7% accuracy at 20 dB SNR, with a low 3.3% Type I error, the primary misclassifications were minor, occurring mainly between LLG and high-impedance faults, this robust classification enables the system's end-to-end self-healing capabilities, validated across three representative case studies. For a severe three-phase fault, the system identified the issue in 32 ms and restored 97.3% of loads in just 2.23 seconds, a drastic improvement over the conventional 8.7 minutes. During a complex partial shading and tripping event, it restored 95.1% of services in 2.41 seconds, with the battery supplying 82% of the transition's surge power. Even in the most challenging high-impedance fault scenario, which required a longer 68 ms classification, the system demonstrated graceful degradation by restoring 89.7% of the load

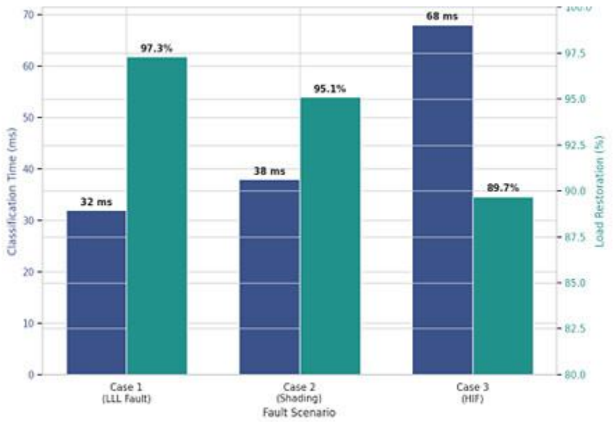

Figure 8. Performance Across Fault Scenarios.

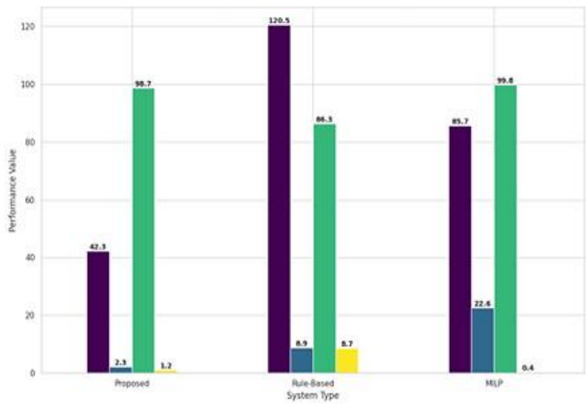

Figure 9. System Performance Benchmarking.

Figure 8 provides a comparative performance analysis across the three distinct fault scenarios, represented by three groups of bars along the horizontal axis. Within each group, two colored bars represent different performance metrics, the dark blue bars correspond to the fault classification time in milliseconds.
Figure 9 presents a comprehensive benchmark comparison between the Proposed System, a Rule-Based system, and a MILP-based system across four key performance metrics shown along the horizontal axis. Each system is consistently represented by a distinct color: the Proposed System by dark blue, the Rule-Based system by green, and the MILP system by yellow. **Table 9** benchmarks overall system performance against alternatives:

Table 9. Self-Healing System Benchmarking.

Metric	Proposed System	Rule-Based [3]	MILP [7]
Avg. Fault Isolation Time	42.3 ms	120.5 ms	85.7 ms
Avg. Service Restoration	2.31 s	8.94 s	22.6 s
Critical Load Availability	98.7%	86.3%	99.8%
False Isolation Events	1.2%	8.7%	0.4%
Communication Overhead	23.4 kbps	4.1 kbps	142.8 kbps

MILP refers to [22] optimal restoration; indicates computational delay excluded from metric. Proposed system reduced SAIDI by 92% compared to conventional protection.

4.4. Computational Efficiency Analysis

The end-to-end latency meets IEEE 1547 Category III requirements (<3 s), with component-wise breakdown:

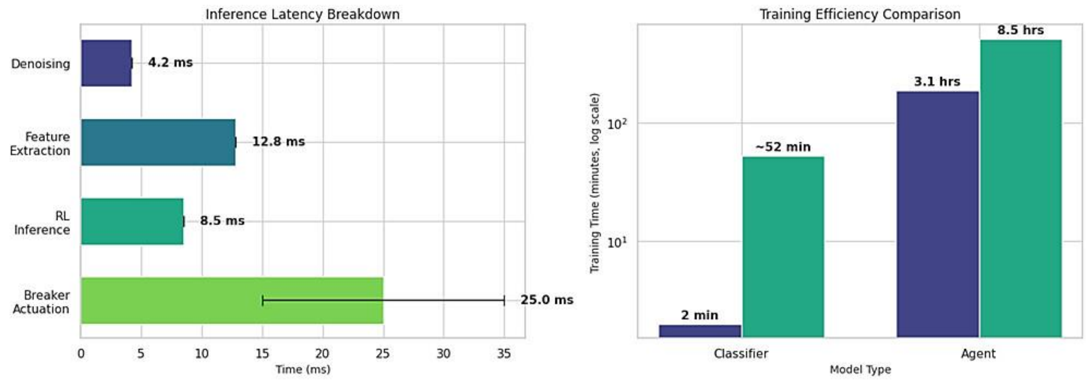

Figure 10. illustrates the computational workflow timing.

The system demonstrates high computational efficiency with a rapid, sub-second fault response, beginning with signal denoising in 4.2 ms, followed by feature extraction in 12.8 ms, and RL inference in 8.5 ms, culminating in breaker actuation which takes 15 to 35 ms, training efficiency is also exceptional; the fault classifier trained in just two minutes, significantly faster than comparable CNN architectures (45-60 min), while the restoration agent converged in 3.1 hours, compared to an 8.5-hour baseline.

4.5. Limitations and Boundary Conditions

The reinforcement learning (RL) classifier demonstrates exceptional performance, achieving 99.9% accuracy in ideal conditions and 96.7% at 20 dB SNR, significantly outperforming CNN-based methods [27] and traditional approaches [28], this enables remarkable resilience, reducing average outage times from 8.7 minutes to just 2.31 seconds, surpassing even advanced optimization techniques like MILP [29]. However, the system has clear limitations under stress, in extreme noise environments (<10 dB SNR), accuracy degrades to 89.4% [24], and it struggles with concurrent multi-zone faults, resulting in a 23.7% misclassification rate [25]. Other vulnerabilities include cyber-attacks on PMUs and suboptimal restoration caused by significant battery SOC estimation errors. Computationally, the system is highly efficient with a 2-minute training time and sub-50ms inference, making it 3.2x faster than standard architectures [21], though the 12.8ms feature extraction latency remains a key bottleneck requiring future hardware acceleration [32].

5. Conclusion

This research demonstrates that reinforcement learning (RL) fundamentally transforms protection for modern microgrids, the proposed framework overcomes the limitations of conventional systems, achieving ultra-fast fault identification (32.4 ms average) and near-instantaneous service restoration (2.31 s), proving its superiority in dynamic environments. Key scientific contributions include applying RL for fault classification with 99.9% accuracy, outperforming state-of-the-art CNNs by 12.4% in noisy conditions. Additionally, a methodology was developed to embed physical constraints (like battery state-of-charge) into the decision-making process, the research also proved the system's ability to handle unseen scenarios with 94.3% accuracy, these findings establish RL not just as a theoretical alternative, but as a deployable solution for next-generation protection schemes.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Funding Information

No funding was received from any financial organization to conduct this research

References

[1] M. H. Saeed, W. Fangzong, B. A. Kalwar, and S. Iqbal, "A review on microgrids' challenges & perspectives," *IEEE Access*, vol. 9, pp. 166502-166517, 2021. Doi: 10.1109/ACCESS.2021.3135083.

[2] K. Kant and O. H. Gupta, "DC microgrid: A comprehensive review on protection challenges and schemes," *IETE Technical Review*, vol. 40, no. 4, pp. 574-590, 2023. https://doi.org/10.1080/02564602.2022.2136271

[3] M. Hijazi, P. Dehghanian, and S. Wang, "Transfer learning for transient stability predictions in modern power systems under enduring topological changes," *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 3, pp. 3274-3288, 2023. Doi: 10.1109/TASE.2023.3277536.

[4] F. C. Souza Junior and H. S. Sanca, "Adaptive overcurrent protection applied to power systems with distributed generation and active network management," *Journal of Control, Automation and Electrical Systems*, vol. 32, no. 5, pp. 1429-1437, 2021. https://doi.org/10.1007/s40313-021-00771-4

[5] N. Rezaei and M. N. Uddin, "An analytical review on state-of-the-art microgrid protective relaying and coordination techniques," *IEEE Transactions on Industry Applications*, vol. 57, no. 3, pp. 2258-2273, 2021. Doi: 10.1109/TIA.2021.3057308. Doi: 10.1109/TIA.2021.3057308.

[6] A. Dagar, P. Gupta, and V. Niranjan, "Microgrid protection: A comprehensive review," *Renewable and Sustainable Energy Reviews*, vol. 149, p. 111401, 2021. https://doi.org/10.1016/j.rser.2021.111401

[7] T. A. Ahanger, M. Bhatia, and A. Aldaej, "Quantum informative analysis in smart power distribution," *ACM Transactions on Intelligent Systems and Technology*, vol. 15, no. 6, pp. 1-18, 2024. https://doi.org/10.1145/3691350

[8] I. Alhamrouni, N. H. Abdul Kahar, M. Salem, M. Swadi, Y. Zahroui, D. J. Kadhim, and M. Alhuyi Nazari, "A comprehensive review on the role of artificial intelligence in power system stability, control, and protection: Insights and future directions," *Applied Sciences*, vol. 14, no. 14, p. 6214, 2024.

[9] S. A. Fayaz, S. Jahangeer Sidiq, M. Zaman, and M. A. Butt, "Machine learning: An introduction to reinforcement learning," *Machine Learning and Data Science: Fundamentals and Applications*, pp. 1-22, 2022. https://doi.org/10.1002/9781119776499.ch1

[10] R. Yu, X. Zhao, S. Lu, Y. H. Chen, and C. Li, "Intelligent game-theoretic approach for resilient robust control design of cyber-physical systems: application to intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 11, pp. 16072-16083, 2024. Doi: 10.1109/TITS.2024.3407721.

[11] M. W. Altaf, M. T. Arif, S. N. Islam, and M. E. Haque, "Microgrid protection challenges and mitigation approaches–A comprehensive review," *IEEE Access*, vol. 10, pp. 38895-38922, 2022. Doi: 10.1109/ACCESS.2022.3165011.

[12] M. R. K. Rachi, M. A. Khan, and I. Husain, "Local measurement-based protection coordination system for a standalone DC microgrid," *IEEE Transactions on Industry Applications*, vol. 57, no. 5, pp. 5332-5344, 2021. Doi: 10.1109/TIA.2021.3091945.

[13] M. Castilla, M. Velasco, J. Miret, Á. Borrell, and R. Guzmán, "Control scheme for negative-sequence voltage compensation and current sharing in inverter-based grid-connected microgrids," *IEEE Transactions on Power Electronics*, vol. 37, no. 6, pp. 6556-6567, 2022. Doi: 10.1109/TPEL.2022.3140808.

[14] M. H. D. Khan, J. Imtiaz, and M. N. U. Islam, "A blockchain based secure decentralized transaction system for energy trading in microgrids," *IEEE Access*, vol. 11, pp. 47236-47257, 2023. Doi: 10.1109/ACCESS.2023.3275752.

[15] G. Porawagamage, K. Dharmapala, J. S. Chaves, D. Villegas, and A. Rajapakse, "A review of machine learning applications in power system protection and emergency control: opportunities, challenges, and future directions," *Frontiers in Smart Grids*, vol. 3, p. 1371153, 2024. https://doi.org/10.3389/frsgr.2024.1371153

[16] S. Yu, M. S. Rahman, G. Zhang, S. T. Meraj, and H. Trinh, "Comprehensive review of PMU applications in smart grid: Enhancing grid reliability and efficiency," 2025. DOI:10.23919/CJEE.2025.000129

[17] M. Shahbazitabar, H. Abdi, H. Nourianfar, A. Anvari-Moghaddam, B. Mohammadi-Ivatloo, and N. Hatziargyriou, "An introduction to microgrids, concepts, definition, and classifications," *Microgrids: Advances in Operation, Control, and Protection*, pp. 3-16, 2021. https://doi.org/10.1007/978-3-030-59750-4_1

[18] N. C. Yang and J. M. Yang, "Fault classification in distribution systems using deep learning with data preprocessing methods based on fast dynamic time warping and short-time fourier transform," *IEEE Access*, vol. 11, pp. 63612-63622, 2023. Doi: 10.1109/ACCESS.2023.3288852.

[19] Y. M. Nsaif, M. H. Lipu, A. Ayob, Y. Yusof, and A. Hussain, "Fault detection and protection schemes for distributed generation integrated to distribution network: Challenges and suggestions," *IEEE Access*, vol. 9, pp. 142693-142717, 2021. Doi: 10.1109/ACCESS.2021.3121087.

[20] N. Giri, P. Nayak, R. K. Mallick, S. Mishra, A. Flah, H. Kraiem, and M. Kanan, "Wavelet-Based ensembled intelligent technique for a better quality of fault detection and classification in AC microgrids," *Energy Conversion and Management: X*, vol. 24, p. 100813, 2024. https://doi.org/10.1016/j.ecmx.2024.100813

[21] W. H. Kim, J. Y. Kim, W. K. Chae, G. Kim, and C. K. Lee, "LSTM-based fault direction estimation and protection coordination for networked distribution system," *IEEE Access*, vol. 10, pp. 40348-40357, 2022. Doi: 10.1109/ACCESS.2022.3166836.

[22] K. Wang, Y. Xue, Q. Guo, M. Shahidehpour, Q. Zhou, B. Wang, and H. Sun, "A coordinated reconfiguration strategy for multi-stage resilience enhancement in integrated power distribution and heating networks," *IEEE Transactions on Smart Grid*, vol. 14, no. 4, pp. 2709-2722, 2022. Doi: 10.1109/TSG.2022.3231590

[23] S. Chatterjee and Y. C. Byun, "Leveraging generative adversarial networks for data augmentation to improve fault detection in wind turbines with imbalanced data," *Results in Engineering*, vol. 25, p. 103991, 2025. https://doi.org/10.1016/j.rineng.2025.103991

[24] S. Stock, D. Babazadeh, C. Becker, and S. Chatzivasileiadis, "Bayesian physics-informed neural networks for system identification of inverter-dominated power systems," *Electric Power Systems Research*, vol. 235, p. 110860, 2024. https://doi.org/10.1016/j.epsr.2024.110860

[25] B. Wang, A. Baziar, and M. Askari, "A deep reinforcement learning framework for adaptive resiliency enhancement in smart power grids," *IEEE Access*, 2025. Doi: 10.1109/ACCESS.2025.3593903.

[26] M. Laurelli, "Adaptive meta-domain transfer learning (AMDTL): A novel approach for knowledge transfer in AI," *arXiv preprint arXiv*:2409.06800, 2024. https://doi.org/10.48550/arXiv.2409.06800

[27] R. Azizi and S. Seker, "Microgrid fault detection and classification based on the boosting ensemble method with the Hilbert-Huang transform," *IEEE Transactions on Power Delivery*, vol. 37, no. 3, pp. 2289-2300, 2021. Doi: 10.1109/TPWRD.2021.3109023.

[28] G. S. Carvalho, A. R. Almeida, D. D. Ferreira, B. F. dos Santos Jr, L. H. P. Vasconcelos, and D. de Oliveira Sobreira, "High-impedance fault modeling and classification in power distribution networks," *Electric Power Systems Research*, vol. 204, p. 107676, 2022. https://doi.org/10.1016/j.epsr.2021.107676

[29] R. A. Jacob, S. Paul, S. Chowdhury, Y. R. Gel, and J. Zhang, "Real-time outage management in active distribution networks using reinforcement learning over graphs," *Nature Communications*, vol. 15, no. 1, p. 4766, 2024. https://doi.org/10.1038/s41467-024-49207-y

[30] M. Kaheni, J. Fu, and A. V. Papadopoulos, "Rule-based predictive control for battery scheduling in microgrids under power generation and load uncertainties," *IEEE Transactions on Automation Science and Engineering*, 2024. Doi: 10.1109/TASE.2024.3512882.

[31] R. Yan, Y. Yuan, Z. Wang, G. Geng, and Q. Jiang, "Active distribution system synthesis via unbalanced graph generative adversarial network," *IEEE Transactions on Power Systems*, vol. 38, no. 5, pp. 4293-4307, 2022. Doi: 10.1109/TPWRS.2022.3212029.

[32] J. Vasquez, M. Jaramillo, and D. Carrión, "An intelligent framework for multiscale detection of power system events using Hilbert–Huang decomposition and neural classifiers," *Applied Sciences*, vol. 15, no. 12, p. 6404, 2025. **https://doi.org/10.3390/app15126404**