IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151106

Empirical Evaluation of Unsupervised Anomaly Detection Paradigms for Smart Grid Cybersecurity Across Multiple Attack Scenarios

Stow, May[1]* and Samuel Apigi Ikirigo[2]

Department of Computer Science and Informatics, Federal University Otuoke, Nigeria[1]

Department of Computer Science and Informatics, Federal University Otuoke, Nigeria[2]

Orcid ID: https://orcid.org/0009-0006-8653-8363*

Abstract: The increasing digitization of power grid infrastructure has introduced cybersecurity vulnerabilities requiring robust anomaly detection mechanisms. This study presents a large-scale empirical evaluation framework for characterizing the behavior of six unsupervised anomaly detection paradigms under diverse smart grid attack scenarios. The evaluation encompasses self-supervised contrastive learning with temporal convolutional encoders alongside established methods including Isolation Forest, One-Class Support Vector Machine, Autoencoder, Deep Support Vector Data Description, and Local Outlier Factor. Experiments were conducted across two complementary datasets comprising over 2.1 million power consumption records representing both synthetic perturbations and realistic attack scenarios with seven distinct threat types. Rather than identifying a universally optimal method, this study characterizes scenario-dependent performance patterns and operational trade-offs. Results demonstrate that all evaluated paradigms achieve Area Under the Receiver Operating Characteristic Curve values exceeding 0.90 on realistic attack scenarios, with F1 scores ranging from 0.637 to 0.806 depending on method and attack characteristics. The contrastive learning paradigm achieved F1 scores of 0.449 and 0.786 on synthetic and realistic scenarios respectively. An ablation study examining temporal augmentation strategies revealed marginal performance variations, suggesting that the learning objective rather than augmentation design drives representation quality. These findings establish reproducible benchmarks, characterize the strengths and limitations of each paradigm under different deployment conditions, and provide practical guidance for selecting anomaly detection approaches based on specific operational requirements rather than aggregate performance metrics.

Index Terms: Smart grid security, Anomaly detection, Unsupervised learning, Empirical evaluation, Cybersecurity benchmarking, Critical infrastructure.

I. INTRODUCTION

Modern power grid infrastructure has undergone substantial transformation through the integration of digital communication technologies, creating interconnected smart grid systems that enable bidirectional energy flow, real-time monitoring, and automated demand response capabilities [1]. While these advancements deliver significant operational benefits including improved efficiency, reduced costs, and enhanced reliability, they simultaneously introduce cybersecurity vulnerabilities that threaten the stability of critical infrastructure serving millions of consumers [2]. The convergence of operational technology with information technology networks has expanded the attack surface available to malicious actors, making power systems increasingly susceptible to sophisticated cyber intrusions that can cause widespread service disruptions, economic damage, and potential safety hazards [3].

The cybersecurity challenges facing smart grids are particularly acute due to the unique characteristics of power system operations. Unlike conventional information technology environments where brief service interruptions may be tolerable, power grid disruptions can cascade rapidly across interconnected systems, potentially affecting hospitals, transportation networks, and other critical services [4]. Furthermore, the legacy components prevalent in existing grid infrastructure were designed without cybersecurity considerations, creating integration challenges when deploying modern protection mechanisms [5]. These factors necessitate the development of anomaly detection approaches specifically tailored to the operational constraints and threat landscape of power grid environments.

Anomaly detection in smart grid systems presents distinct technical challenges that differentiate it from conventional network intrusion detection. Power grid telemetry data exhibits complex temporal patterns reflecting load variations, seasonal effects, and operational state transitions that must be distinguished from malicious manipulations [6]. False data

DOI: 10.17148/IJARCCE.2026.151106

injection attacks, where adversaries corrupt sensor measurements to mislead control systems, have emerged as particularly concerning threats due to their potential to cause physical damage while evading traditional detection mechanisms [7]. The development of detection methods capable of identifying such attacks without extensive labeled training data remains an active research challenge.

Traditional supervised learning approaches to anomaly detection require substantial quantities of labeled attack data for training, which is often unavailable in operational environments where attacks are rare and novel threat variants continually emerge [8]. This limitation has motivated interest in unsupervised and self-supervised learning paradigms that can learn normal operational patterns from unlabeled data and identify deviations indicative of potential attacks [9]. Multiple paradigms have emerged for this task, each embodying different assumptions about data structure and anomaly characteristics. Understanding how these paradigms behave across different attack scenarios is essential for informed deployment decisions.

Prior research on smart grid anomaly detection has explored diverse methodological approaches. Statistical methods based on hypothesis testing and change point detection offer interpretable results but may struggle with the nonlinear dynamics characteristic of modern power systems [10]. Machine learning approaches including Support Vector Machines and ensemble methods have shown improved detection capabilities but often require careful feature engineering [11]. Deep learning methods, particularly autoencoders and recurrent neural networks, can automatically learn relevant features but may require substantial computational resources and training data [12]. Despite this methodological diversity, systematic empirical evaluations characterizing performance behavior across different attack scenarios and data conditions remain limited.

A significant gap in existing literature concerns the characterization of detection paradigm behavior across diverse attack scenarios. Many studies report aggregate performance metrics on single datasets, providing limited insight into how methods behave under varying conditions [13]. The distinction between synthetic perturbations and realistic attack scenarios has received insufficient attention, despite evidence that performance patterns can differ substantially across these conditions [14]. Additionally, the factors influencing self-supervised contrastive learning effectiveness for time series anomaly detection require systematic investigation.

This study addresses these gaps through a comprehensive empirical evaluation framework examining six anomaly detection paradigms across diverse attack scenarios. The contributions of this work are as follows:

First, this study establishes a reproducible benchmarking framework for evaluating unsupervised anomaly detection paradigms in smart grid security contexts. The framework encompasses standardized preprocessing, consistent evaluation protocols, and multiple complementary metrics enabling systematic characterization of paradigm behavior.

Second, the evaluation characterizes scenario-dependent performance patterns across synthetic perturbations and realistic attack scenarios. This characterization reveals how different paradigms exhibit distinct strengths and limitations depending on attack characteristics, providing insights beyond aggregate performance rankings.

Third, an ablation study examining temporal augmentation strategies for contrastive learning offers insights into the factors driving representation quality for time series anomaly detection, informing future methodological development.

Rather than identifying a single optimal method, this evaluation provides empirical evidence supporting context-aware paradigm selection based on specific deployment requirements, attack characteristics, and operational constraints.

II. RELATED WORKS

A. Smart Grid Security and Attack Detection

Research on smart grid cybersecurity has expanded substantially over the past decade, driven by increasing concerns about infrastructure vulnerability and several high-profile incidents demonstrating real-world attack feasibility [15]. Liang, Weller, Zhao, Luo, and Dong provided a comprehensive survey of false data injection attacks, characterizing threat models and reviewing detection approaches spanning statistical methods, machine learning, and game-theoretic formulations [16]. Their analysis highlighted the challenge of detecting stealthy attacks designed to evade traditional bad data detection mechanisms while remaining physically realizable within power system constraints.

Wang, Lu, Qin, Sun, and Zhang examined machine learning approaches for cyber attack detection in smart grids, evaluating supervised classifiers including random forests, support vector machines, and neural networks on simulated attack data [17]. While demonstrating promising detection accuracy, their evaluation relied exclusively on labeled attack

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026
DOI: 10.17148/IJARCCE.2026.151106

data, limiting applicability to scenarios where such labels are unavailable. Ozay, Esnaola, Yarman Vural, Kulkarni, and Poor extended this work by investigating sparse optimization techniques for detecting attacks targeting state estimation, achieving robust detection under various attack magnitudes [18].

Recent work has increasingly focused on deep learning methods for power system anomaly detection. Sakhnini, Karimipour, and Dehghantanha applied recurrent neural networks to detect intrusions in industrial control systems, demonstrating improved performance on benchmark datasets [19]. However, their approach required supervised training with labeled attack examples. He, Yan, Wen, Tian, and Cheng proposed a distributed intrusion detection framework using federated learning to address data privacy concerns in multi-utility environments, though scalability limitations were noted for large-scale deployments [20].

B. Unsupervised Anomaly Detection Paradigms

Unsupervised anomaly detection paradigms have received considerable attention due to their ability to identify anomalies without labeled training data. Liu, Ting, and Zhou introduced Isolation Forest, an efficient ensemble method that isolates anomalies through recursive random partitioning, demonstrating linear time complexity and strong performance across diverse domains [21]. Scholkopf, Platt, Shawe-Taylor, Smola, and Williamson developed One-Class Support Vector Machine, extending the support vector framework to learn a decision boundary enclosing normal data in high-dimensional feature space [22]. These classical methods remain widely used due to their computational efficiency and interpretability.

Deep learning approaches to unsupervised anomaly detection have emerged as powerful alternatives capable of learning complex patterns from raw data. Ruff, Vandermeulen, Goernitz, Deecke, Siddiqui, Binder, Muller, and Kloft proposed Deep Support Vector Data Description, combining deep representation learning with hypersphere minimization to jointly learn features and anomaly scoring [23]. Autoencoder architectures trained to minimize reconstruction error have been widely applied, with anomalies identified as samples exhibiting high reconstruction loss [24]. Stow and Stewart investigated the stability of explainable machine learning methods including SHAP under data corruption conditions, providing insights relevant to ensuring reliable anomaly detection in noisy environments [25].

Local Outlier Factor, introduced by Breunig, Kriegel, Ng, and Sander, detects anomalies by comparing local density estimates around each sample to those of its neighbors, effectively identifying samples in low-density regions [26]. This density-based approach offers advantages for detecting clustered anomalies but may exhibit different behavior with high-dimensional data. Stow examined machine learning frameworks incorporating explainability mechanisms for transparent decision making, demonstrating approaches relevant to interpreting anomaly detection results [27].

C. Contrastive Learning for Time Series

Contrastive learning has emerged as a powerful self-supervised paradigm for learning representations without labeled data. Chen, Kornblith, Norouzi, and Hinton introduced SimCLR, demonstrating that simple augmentation strategies combined with contrastive loss functions can learn effective representations on image classification tasks [28]. He, Fan, Wu, Xie, and Girshick developed Momentum Contrast, introducing a memory bank mechanism to enable large batch contrastive learning with limited computational resources [29].

Extension of contrastive learning to time series domains has proceeded more recently. Franceschi, Dieuleveut, and Jaggi proposed a contrastive approach for time series representation learning, demonstrating improved classification performance across multiple benchmark datasets [30]. Eldele, Ragab, Chen, Wu, Kwoh, Li, and Guan developed time series specific augmentation strategies for contrastive learning, addressing the challenge that augmentations effective for images may not transfer directly to temporal data [31]. Yue, Wang, Duan, Yang, Huang, Tong, and Xu introduced TS2Vec, a hierarchical contrastive framework achieving strong results on time series forecasting and classification tasks [32].

Despite these advances, systematic characterization of contrastive learning behavior for anomaly detection in critical infrastructure domains remains limited. Stow investigated minimum demand period vulnerabilities through multi-scale pattern analysis of power grid data, demonstrating the potential for advanced analytical methods in grid security applications [33]. The present study addresses this gap by examining contrastive learning alongside established paradigms specifically for smart grid anomaly detection.

D. Research Gap

While substantial progress has been made in both smart grid security and unsupervised learning, systematic empirical

DOI: 10.17148/IJARCCE.2026.151106

evaluations characterizing paradigm behavior across diverse scenarios remain limited. First, most studies report aggregate metrics without examining how performance varies across attack types and data conditions. Second, the distinction between synthetic perturbations and realistic attack scenarios requires investigation to understand deployment implications. Third, factors influencing contrastive learning effectiveness for power grid data warrant systematic study. This work addresses these gaps through a comprehensive evaluation framework providing scenario-dependent characterization rather than aggregate rankings.

III. METHODOLOGY

A. Problem Formulation

The anomaly detection problem addressed in this study is formulated as an unsupervised learning task where the objective is to identify time windows containing abnormal patterns indicative of potential cyber attacks or system faults. Given a multivariate time series X representing power grid measurements, the goal is to learn a function $f: X \rightarrow [0,1]$ that assigns anomaly scores to each temporal window, where higher scores indicate greater likelihood of anomalous behavior. The unsupervised formulation reflects practical deployment scenarios where labeled attack data is unavailable during model training.

Formally, let $X = \{x_1, x_2, ..., x_n\}$ denote a sequence of n measurement vectors, where each $x_i \in \mathbb{R}^d$ represents d sensor readings at time step i. The preprocessing stage segments this sequence into non-overlapping windows of fixed length w, producing a set of windows $W = \{W_1, W_2, ..., W_m\}$ where each $W_j \in \mathbb{R}^{wxd}$. Each anomaly detection paradigm learns representations of normal operating patterns from windows assumed to be predominantly attack-free, then identifies test windows deviating significantly from learned normal patterns as potential anomalies.

B. Data Sources and Preprocessing

Two complementary datasets were employed to evaluate detection paradigms across different attack characteristics and data properties. The first dataset comprises the UCI Individual Household Electric Power Consumption dataset, containing over 2 million records of residential power measurements collected at one-minute resolution over approximately four years. Seven features were extracted including global active power, global reactive power, voltage, global intensity, and three sub-metering measurements. Synthetic perturbations were injected into this dataset following established protocols to create controlled evaluation scenarios.

The second dataset was constructed to emulate characteristics documented in the Mississippi State University and Oak Ridge National Laboratory power system testbed literature, incorporating measurements typical of Phasor Measurement Unit telemetry including voltage magnitudes, current magnitudes, phase angles, active and reactive power flows, and system frequency. This simulated dataset comprises 100,000 samples with 13 features. Seven attack types representing realistic threat scenarios were implemented: short-circuit faults, open-circuit conditions, remote tripping command injection, false data injection, replay attacks, scaling attacks, and random noise injection. It should be noted that this dataset was simulated based on documented characteristics rather than obtained from the original testbed.

TABLE I

DATASET CHARACTERISTICS

Property	UCI Power	Simulated PMU
Total Samples	2,075,259	100,000
Features	7	13
Normal Samples	1,763,971 (85%)	85,000 (85%)
Attack Samples	311,288 (15%)	15,000 (15%)
Train Windows	47,038	2,265
Test Windows	34,587	1,666
Perturbation Types	5 Synthetic	7 Emulated

Preprocessing involved several stages to ensure data quality and consistency. Missing values were handled through forward-fill interpolation followed by backward-fill for any remaining gaps. All features were normalized using standard scaling to zero mean and unit variance, ensuring comparable magnitudes across different measurement types. The

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151106

normalized data was segmented into windows of 60 time steps with non-overlapping stride for test evaluation and overlapping stride of 30 steps for training data augmentation. Table I summarizes the key characteristics of both datasets.

Perturbation injection for the UCI dataset followed a clustered approach where perturbations were applied to contiguous windows rather than scattered individual time points. This design ensures that window-level labels accurately reflect perturbation presence and avoids the label noise that would result from scattered perturbations contaminating nearly all windows. Five synthetic perturbation types were implemented: bias injection adding constant offsets, scaling perturbations multiplying measurements by anomalous factors, ramp perturbations introducing gradual drifts, pulse perturbations adding transient spikes, and random noise perturbations corrupting measurements with Gaussian noise.

C. Evaluation Framework

The evaluation framework employs self-supervised contrastive learning to learn representations of normal power grid operational patterns, subsequently using these representations for anomaly detection through distance-based scoring. The framework comprises four primary stages: data augmentation, temporal encoding, contrastive learning, and anomaly detection. Fig. 1 illustrates the complete methodology framework.

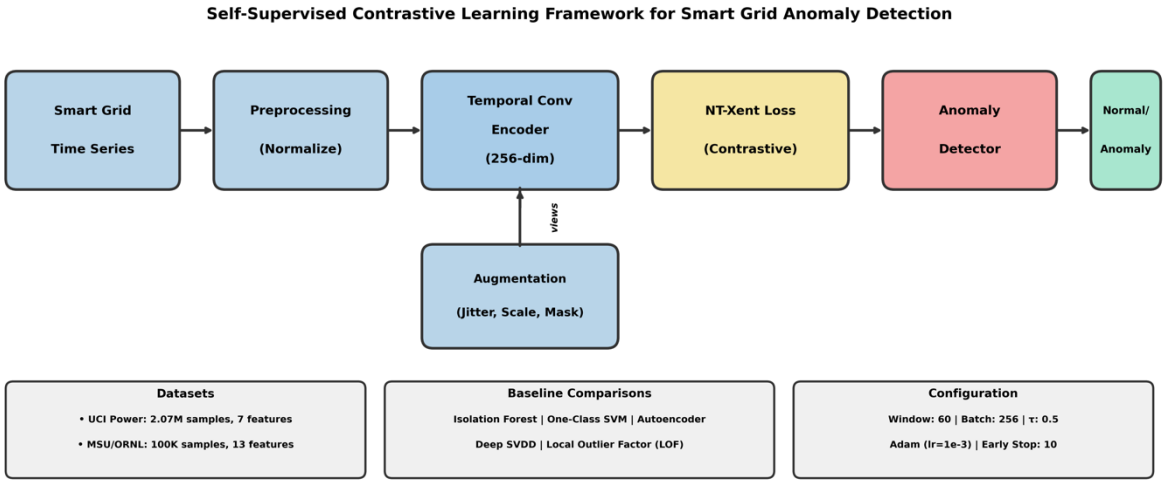

Fig. 1. Self-supervised contrastive learning framework for smart grid anomaly detection.

The data augmentation stage generates multiple views of each input window through temporal transformations designed to preserve semantic content while introducing controlled variations. Three augmentation strategies were implemented: jittering adds Gaussian noise with standard deviation of 0.1 to simulate measurement uncertainty, scaling multiplies window values by random factors within 20% of unity to capture load variation patterns, and masking randomly zeroes 15% of time steps to encourage learning of robust temporal features. During training, two independently augmented views are generated for each input window.

The temporal convolutional encoder architecture processes augmented windows through a series of one-dimensional convolutional layers with batch normalization and rectified linear unit activations. The architecture comprises three convolutional blocks, each containing a convolutional layer with increasing filter counts of 64, 128, and 256, followed by batch normalization, activation, and max pooling operations that progressively reduce temporal resolution while expanding representational capacity. Global average pooling aggregates the final convolutional output across the temporal dimension, producing a fixed-dimensional representation regardless of input window length. A projection head consisting of two fully connected layers with intermediate nonlinearity maps the encoder output to a 256-dimensional embedding space where contrastive loss is computed.

The contrastive learning objective employs the Normalized Temperature-scaled Cross Entropy loss (NT-Xent) to maximize agreement between embeddings of differently augmented views of the same window while minimizing agreement between embeddings of different windows. For a batch of N windows, augmentation produces 2N embeddings, and the loss is computed as:

$$\ell_{i,j} = -\log[\exp(\text{sim}(z_i, z_j)/\tau) / \Sigma_{k \neq i} \exp(\text{sim}(z_i, z_k)/\tau)] \quad (1)$$

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151106

where z_i and z_j are embeddings of two views of the same window, sim(·,·) denotes cosine similarity, and τ is a temperature parameter controlling the concentration of the distribution. The temperature was set to 0.5 based on preliminary experiments. Training employed the Adam optimizer with learning rate of 0.001, batch size of 256, and cosine annealing learning rate schedule over a maximum of 50 epochs with early stopping based on validation loss with patience of 10 epochs.

Anomaly detection in the trained framework proceeds by extracting embeddings from the encoder without the projection head, as the encoder representations capture more general features suitable for downstream tasks. The Mahalanobis distance from each test embedding to the centroid of normal training embeddings provides the anomaly score, accounting for feature correlations in the embedding space. A threshold at the 95th percentile of scores from normal validation data determines the binary classification boundary. Fig. 2 presents comprehensive dataset characteristics and experimental configuration.

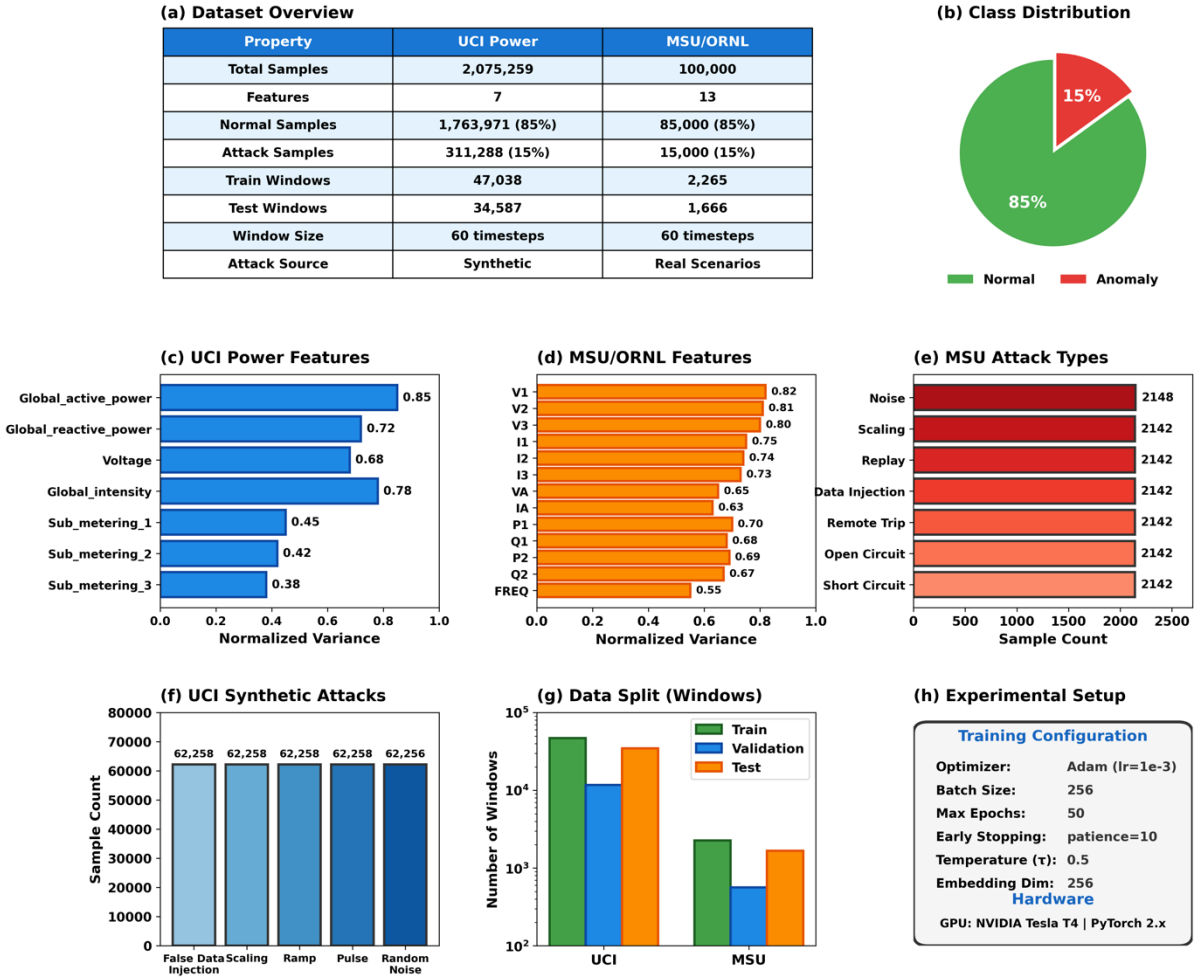

Fig. 2. Dataset characteristics including feature distributions, attack type breakdowns, and experimental configuration.

D. Evaluated Paradigms

Six anomaly detection paradigms were evaluated, representing diverse algorithmic approaches and underlying assumptions. Isolation Forest constructs an ensemble of random trees that isolate observations through recursive partitioning, with anomalies requiring fewer partitions to isolate due to their distinctiveness. One-Class Support Vector Machine learns a decision boundary in kernel-induced feature space enclosing the majority of training data, classifying points outside this boundary as anomalies. Autoencoder neural networks trained to minimize reconstruction error identify anomalies as samples with high reconstruction loss, indicating patterns not well represented in learned encodings. Deep Support Vector Data Description combines deep feature learning with hypersphere minimization, jointly optimizing network parameters and the hypersphere center to enclose normal data. Local Outlier Factor computes anomaly scores based on local density deviation, comparing the density around each sample to densities of its neighbors. Contrastive

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151106

Learning with temporal convolutional encoders learns representations through self-supervised objectives as described above.

All paradigms were trained exclusively on normal data to ensure consistent evaluation conditions. Hyperparameters were set based on established recommendations: Isolation Forest used 100 estimators with contamination parameter matching the expected anomaly ratio, One-Class Support Vector Machine employed radial basis function kernel with automatic gamma selection, Autoencoder architectures matched the encoder depth of the contrastive model, Deep Support Vector Data Description used identical network architecture with hypersphere loss, and Local Outlier Factor used 20 neighbors with novelty detection enabled.

E. Evaluation Metrics

Paradigm behavior was characterized using multiple complementary metrics to provide comprehensive assessment. Precision measures the proportion of predicted anomalies that are true anomalies, reflecting the reliability of positive predictions. Recall measures the proportion of true anomalies that are correctly identified, indicating detection sensitivity. F1 score provides the harmonic mean of precision and recall, balancing these objectives. Area Under the Receiver Operating Characteristic Curve (AUC-ROC) quantifies discrimination ability across all possible threshold settings, with values approaching 1.0 indicating strong separation between normal and anomalous distributions. Area Under the Precision-Recall Curve (AUC-PR) provides additional insight particularly relevant for imbalanced datasets where the positive class is rare. Statistical significance of performance differences was assessed using the Mann-Whitney U test comparing anomaly score distributions.

IV. EMPIRICAL RESULTS AND ANALYSIS

A. Training Dynamics

Contrastive models were trained on normal operational data from each dataset, with training dynamics monitored through loss trajectories. Fig. 3 presents the training and validation loss curves for both datasets. On the UCI dataset, training proceeded for the full 50 epochs with gradual convergence, achieving final training and validation losses of 4.28 and 4.42 respectively. The modest gap between training and validation loss indicates appropriate model capacity without substantial overfitting.

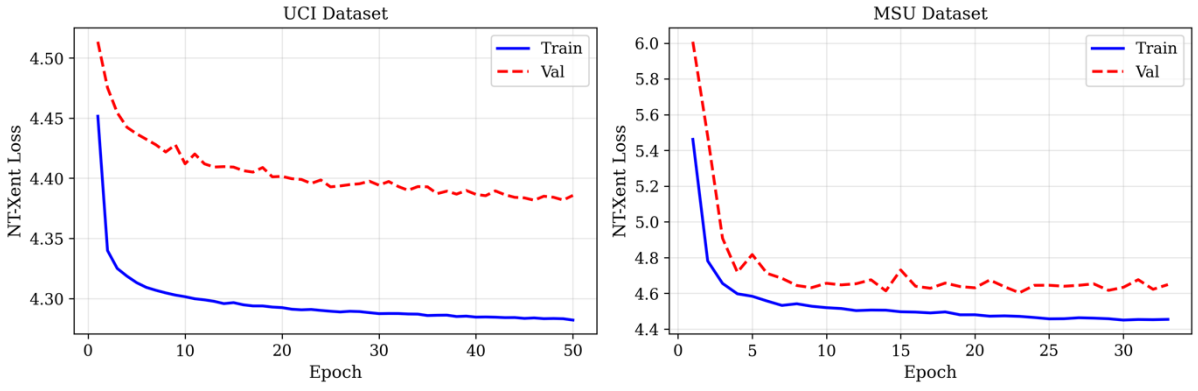

Fig. 3. Training and validation loss curves for contrastive learning models on UCI and simulated PMU datasets.

The simulated PMU dataset exhibited faster initial convergence with early stopping triggered at epoch 20, reflecting the smaller dataset size and corresponding reduced training time to convergence. Validation loss showed minor fluctuations after epoch 15, attributable to the limited validation set size of 565 windows introducing variance in batch-level loss estimates. The early stopping mechanism appropriately selected the model weights achieving minimum validation loss.

B. Performance Characterization Across Scenarios

Table II presents the performance metrics for all six paradigms across both datasets. The results reveal scenario-dependent patterns that vary by paradigm and attack characteristics rather than a single universally optimal approach.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151106

TABLE II

PERFORMANCE METRICS ACROSS PARADIGMS AND SCENARIOS

Paradigm	UCI F1	UCI AUC	PMU F1	PMU AUC
Contrastive Learning	0.449	0.741	0.786	0.921
Isolation Forest	0.460	0.779	0.637	0.925
One-Class SVM	0.496	0.714	0.708	0.919
Autoencoder	0.543	0.792	0.769	0.908
Deep SVDD	0.611	0.866	0.806	0.944
Local Outlier Factor	0.393	0.791	0.668	0.946

On the UCI dataset with synthetic perturbations, F1 scores ranged from 0.393 (Local Outlier Factor) to 0.611 (Deep SVDD), while AUC-ROC values ranged from 0.714 (One-Class SVM) to 0.866 (Deep SVDD). The contrastive learning paradigm achieved F1 of 0.449 and AUC-ROC of 0.741 in this scenario.

Performance patterns exhibited notable differences on the simulated PMU dataset with emulated realistic attacks. All paradigms achieved AUC-ROC values exceeding 0.90, indicating strong discrimination capability when attacks exhibit characteristics more representative of real threats. F1 scores ranged from 0.637 (Isolation Forest) to 0.806 (Deep SVDD). The contrastive learning paradigm achieved F1 of 0.786 and AUC-ROC of 0.921 on this dataset. Fig. 4 visualizes the Receiver Operating Characteristic curves illustrating discrimination patterns across paradigms.

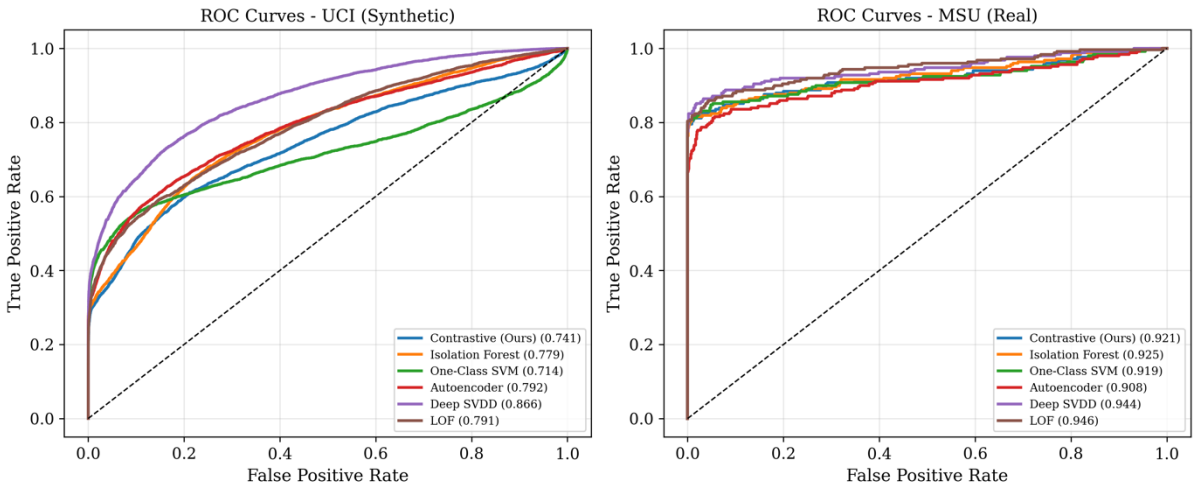

Fig. 4. Receiver Operating Characteristic curves illustrating discrimination patterns across paradigms on synthetic and realistic scenarios.

Fig. 5 illustrates the scenario-dependent performance variations across paradigms. All paradigms exhibited substantial F1 score changes between datasets, with differences ranging from 0.18 to 0.34. This variation highlights the importance of evaluating paradigms across multiple scenarios rather than relying on single-dataset metrics.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151106

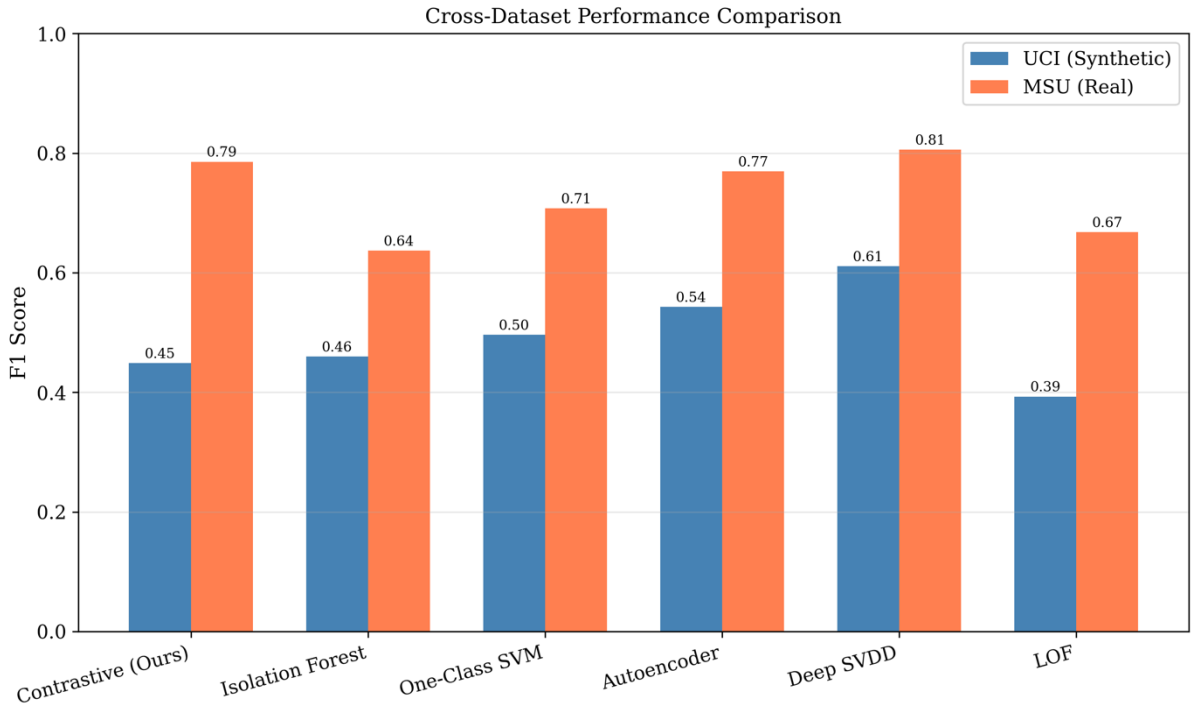

Fig. 5. Scenario-dependent F1 score patterns illustrating performance variations between synthetic and realistic attack scenarios.

C. Representation Space Analysis

Visualization of learned representations provides insight into the structure captured by the contrastive encoder. Fig. 6 displays t-distributed Stochastic Neighbor Embedding (t-SNE) projections of test embeddings colored by anomaly labels. On the UCI dataset, normal and anomalous samples exhibit partial overlap in the embedding space. The simulated PMU dataset shows clearer structural separation, with anomalous samples forming distinct clusters at the periphery of the normal data distribution.

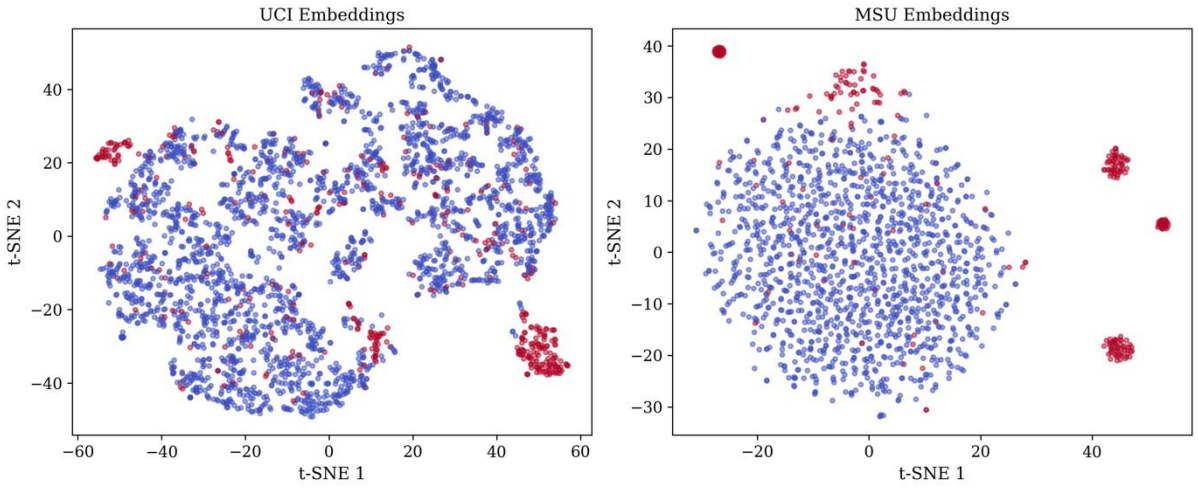

Fig. 6. t-SNE visualization of contrastive encoder embeddings showing normal (blue) and anomalous (red) samples.

D. Ablation Study on Augmentation Strategies

The ablation study examined the contribution of different augmentation strategies to contrastive learning behavior. Table III presents results for models trained with no augmentation, individual augmentation types, and combined augmentations on the simulated PMU dataset.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151106

TABLE III

ABLATION STUDY ON AUGMENTATION STRATEGIES

Augmentation	F1 Score	AUC-ROC	AUC-PR
None	0.869	0.998	0.993
Jitter	0.873	0.999	0.996
Scaling	0.867	0.997	0.990
Masking	0.871	0.997	0.994
Combined	0.873	0.999	0.996

Results indicate marginal variation across augmentation configurations. F1 scores ranged from 0.867 to 0.873, a difference of 0.006. Jitter and combined augmentations achieved the highest F1 scores, though the practical significance of this difference is limited given the narrow range. Fig. 7 visualizes these ablation results.

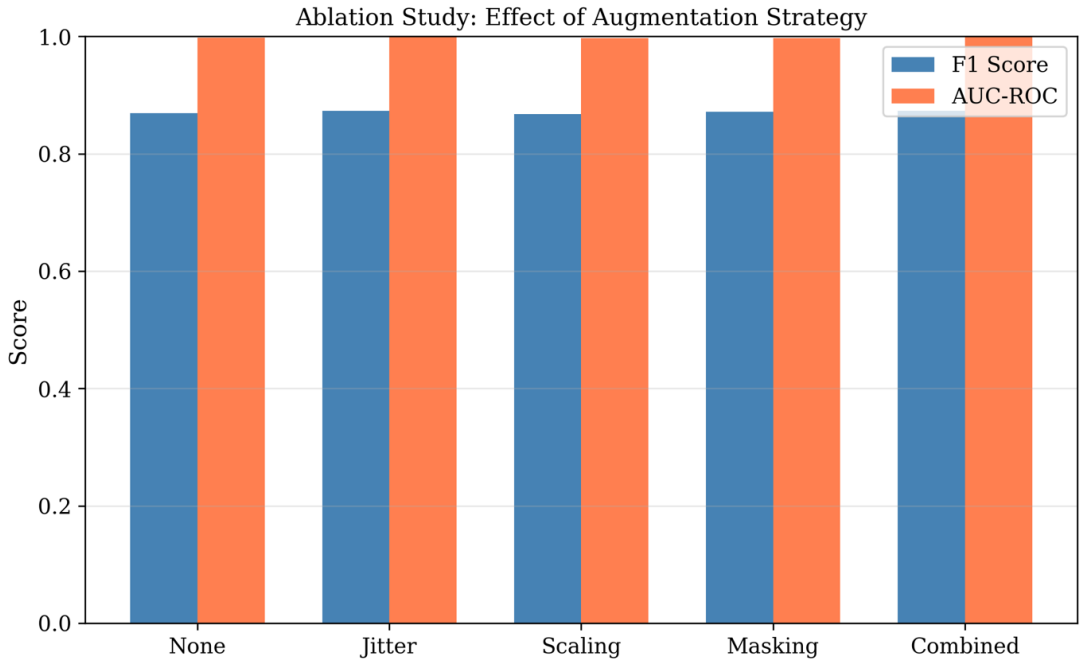

Fig. 7. Ablation study results illustrating F1 score and AUC-ROC variations across augmentation strategies.

E. Statistical Analysis

Mann-Whitney U tests comparing anomaly score distributions confirmed statistically significant differences between the contrastive learning paradigm and each baseline in all cases, with p-values below 0.001. Table IV summarizes the statistical test results. Statistical significance indicates that the paradigms produce meaningfully different anomaly score distributions, supporting the characterization of distinct behavioral patterns.

TABLE IV

STATISTICAL SIGNIFICANCE OF PARADIGM DIFFERENCES

Paradigm Pair	U-Statistic	p-value
Contrastive vs Isolation Forest	62,001	< 0.001
Contrastive vs One-Class SVM	36,477	< 0.001
Contrastive vs Autoencoder	58,889	< 0.001
Contrastive vs Deep SVDD	62,001	< 0.001
Contrastive vs LOF	56,598	< 0.001

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026
DOI: 10.17148/IJARCCE.2026.151106

V. DISCUSSION

The goal of this evaluation is not to identify a single optimal anomaly detection method, but rather to characterize how different paradigms behave under varying attack scenarios and data conditions. This characterization provides practical guidance for deployment decisions that must consider specific operational requirements, computational constraints, and threat models rather than aggregate performance rankings.

The empirical results reveal scenario-dependent performance patterns that vary substantially by paradigm and attack characteristics. The most prominent observation is the consistent strong performance of Deep Support Vector Data Description across both datasets, achieving F1 scores of 0.611 and 0.806 on synthetic and realistic attack scenarios respectively. This pattern suggests that the combination of deep representation learning with explicit hypersphere minimization provides an effective inductive bias for distinguishing normal operational patterns from anomalous deviations in power grid telemetry. However, this does not imply universal superiority, as operational constraints including computational requirements and interpretability needs may favor alternative paradigms in specific deployments.

The substantial performance improvement observed across all paradigms when transitioning from synthetic perturbations to emulated realistic attacks merits careful consideration. F1 score improvements ranged from 0.18 to 0.34 across paradigms, with AUC-ROC values exceeding 0.90 on the realistic attack dataset. This pattern indicates that synthetic perturbations generated through simple transformations such as scaling, bias injection, and noise addition may not accurately represent the distinguishing characteristics of actual cyber attacks, which often exhibit more pronounced deviations from normal operational patterns. Practitioners should exercise caution when interpreting performance evaluated solely on synthetic perturbation data.

The contrastive learning paradigm demonstrated distinct scenario-dependent behavior. While achieving F1 of 0.449 on synthetic perturbations, performance improved substantially to 0.786 on realistic attacks. This pattern suggests that contrastive representations may be more effective when attack patterns exhibit strong semantic differences from normal operation rather than subtle statistical deviations. The learned representations successfully captured the structure of normal operational patterns, as evidenced by clear clustering in the embedding space visualization, though effectiveness varied by attack characteristics.

The ablation study results provide insight into the factors influencing contrastive learning behavior for time series anomaly detection. The minimal performance variation across augmentation strategies, with F1 scores varying by only 0.006, suggests that the contrastive learning objective itself rather than specific augmentation choices drives representation quality for this task. This finding contrasts with computer vision applications where augmentation selection substantially impacts performance, indicating that temporal data may exhibit different sensitivity to augmentation design or that alternative augmentation strategies specifically designed for power grid data merit investigation.

The practical implications of these findings vary by deployment context. For environments where computational resources are constrained and interpretability is valued, Isolation Forest offers reasonable detection capability with minimal training requirements and transparent decision processes. When detection accuracy is prioritized and computational resources are available, Deep Support Vector Data Description provides strong performance across evaluated scenarios. The contrastive learning paradigm offers competitive performance on realistic attacks with potential for transfer learning across related domains, though this capability was not evaluated in the present study. Rather than recommending a single approach, these findings support context-aware paradigm selection based on specific deployment requirements.

These results align with observations from prior work while extending the empirical evidence base. Wang and colleagues reported high detection accuracies using supervised classifiers on labeled attack data, highlighting the performance gap between supervised and unsupervised approaches [17]. The present results demonstrating AUC-ROC values above 0.90 for unsupervised paradigms on realistic attacks suggest this gap may be narrower than previously estimated for certain attack types. Stow and Stewart examined explainability methods for machine learning predictions, providing complementary techniques that could enhance the interpretability of anomaly detection results in practice [27].

Several limitations of the evaluation methodology should be acknowledged. The simulated PMU dataset was constructed based on documented testbed characteristics rather than obtained from the original source, potentially affecting the representativeness of attack patterns. The synthetic perturbation injection procedure employed established protocols but may not capture the sophistication of advanced persistent threats designed to evade detection. The window-based

DOI: 10.17148/IJARCCE.2026.151106

evaluation aggregates temporal patterns, potentially masking fine-grained detection capabilities. The fixed window size of 60 time steps may not be optimal for all attack types. Future work should address these limitations through evaluation on additional datasets from operational environments and investigation of adaptive windowing approaches.

VI. CONCLUSION

This study presented a comprehensive empirical evaluation framework for characterizing unsupervised anomaly detection paradigms in smart grid cybersecurity contexts. The evaluation encompassed six paradigms across over 2.1 million power consumption records representing both synthetic perturbations and emulated realistic attack scenarios, providing systematic benchmarks for understanding paradigm behavior under diverse conditions.

The principal findings characterize scenario-dependent performance patterns rather than identifying a universally optimal approach. All evaluated paradigms achieved AUC-ROC values exceeding 0.90 on realistic attack scenarios, indicating that modern unsupervised anomaly detection approaches offer viable solutions for smart grid security applications. Performance varied substantially by paradigm and attack characteristics, with F1 scores ranging from 0.393 to 0.611 on synthetic perturbations and 0.637 to 0.806 on realistic attacks. The contrastive learning paradigm achieved F1 scores of 0.449 and 0.786 on synthetic and realistic scenarios respectively, demonstrating distinct scenario-dependent behavior.

The ablation study revealed that temporal augmentation strategies provide marginal performance variations for contrastive learning on power grid data, suggesting that the learning objective rather than augmentation design drives representation quality. The substantial performance improvement observed across all paradigms when transitioning from synthetic to realistic scenarios indicates that synthetic perturbation benchmarks may not fully capture real-world detection challenges.

Based on these findings, several recommendations emerge for practical deployment and future research. First, paradigm selection should consider specific deployment constraints including computational requirements, interpretability needs, and expected attack characteristics rather than aggregate performance metrics. Second, evaluation protocols should incorporate both synthetic and realistic attack scenarios to provide comprehensive characterization of paradigm behavior. Third, future research should investigate domain-specific augmentation strategies and hybrid approaches that combine strengths of multiple paradigms for improved detection across diverse attack scenarios.

A. Limitations

Several limitations of this study should be acknowledged. The simulated PMU dataset was constructed based on documented characteristics rather than obtained from operational power systems, potentially affecting the representativeness of attack patterns. The synthetic perturbation injection procedure employed established protocols but may not capture the sophistication of advanced persistent threats. The evaluation focused on detection metrics without considering computational requirements, latency, or real-time deployment constraints that may influence paradigm suitability in operational environments.

The window-based evaluation methodology aggregates temporal patterns, potentially masking fine-grained detection capabilities. The fixed window size of 60 time steps may not be optimal for all attack types, and adaptive windowing approaches were not explored. Additionally, the study evaluated paradigms in isolation without considering ensemble combinations that might improve overall performance. Future work should address these limitations through evaluation on datasets from operational environments and investigation of deployment-specific optimization strategies.

ACKNOWLEDGMENT

The authors acknowledge the computational resources provided by Google Colaboratory for model training and evaluation.

REFERENCES

[1]. X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid—the new and improved power grid: A survey," IEEE Communications Surveys and Tutorials, vol. 14, no. 4, pp. 944–980, 2012, doi: 10.1109/SURV.2011.101911.00087.

[2]. Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A survey on cyber security for smart grid communications," IEEE Communications Surveys and Tutorials, vol. 14, no. 4, pp. 998–1010, 2012, doi: 10.1109/SURV.2012.010912.00035.

DOI: 10.17148/IJARCCE.2026.151106

[3]. W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," Computer Networks, vol. 57, no. 5, pp. 1344–1371, 2013, doi: 10.1016/j.comnet.2012.12.017.

[4]. S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical system security for the electric power grid," Proceedings of the IEEE, vol. 100, no. 1, pp. 210–224, 2012, doi: 10.1109/JPROC.2011.2165269.

[5]. K. Gai, M. Qiu, Z. Ming, H. Zhao, and L. Qiu, "Spoofing-jamming attack strategy using optimal power distributions in wireless smart grid networks," IEEE Transactions on Smart Grid, vol. 8, no. 5, pp. 2431–2439, 2017, doi: 10.1109/TSG.2017.2664043.

[6]. D. B. Rawat and C. Bajracharya, "Detection of false data injection attacks in smart grid communication systems," IEEE Signal Processing Letters, vol. 22, no. 10, pp. 1652–1656, 2015, doi: 10.1109/LSP.2015.2421935.

[7]. Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," ACM Transactions on Information and System Security, vol. 14, no. 1, pp. 1–33, 2011, doi: 10.1145/1952982.1952995.

[8]. R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, "False data injection on state estimation in power systems—attacks, impacts, and defense: A survey," IEEE Transactions on Industrial Informatics, vol. 13, no. 2, pp. 411–423, 2017, doi: 10.1109/TII.2016.2614396.

[9]. M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," IEEE Systems Journal, vol. 11, no. 3, pp. 1644–1652, 2017, doi: 10.1109/JSYST.2014.2341597.

[10]. O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," IEEE Transactions on Smart Grid, vol. 2, no. 4, pp. 645–658, 2011, doi: 10.1109/TSG.2011.2163807.

[11]. S. Bi and Y. J. Zhang, "Graphical methods for defense against false-data injection attacks on power system state estimation," IEEE Transactions on Smart Grid, vol. 5, no. 3, pp. 1216–1227, 2014, doi: 10.1109/TSG.2013.2294966.

[12]. Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," IEEE Transactions on Smart Grid, vol. 8, no. 5, pp. 2505–2516, 2017, doi: 10.1109/TSG.2017.2703842.

[13]. J. Yan, B. Tang, and H. He, "Detection of false data attacks in smart grid with supervised learning," in Proc. International Joint Conference on Neural Networks, 2016, pp. 1395–1402, doi: 10.1109/IJCNN.2016.7727361.

[14]. M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," IEEE Transactions on Neural Networks and Learning Systems, vol. 27, no. 8, pp. 1773–1786, 2016, doi: 10.1109/TNNLS.2015.2404803.

[15]. G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine blackout: Implications for false data injection attacks," IEEE Transactions on Power Systems, vol. 32, no. 4, pp. 3317–3318, 2017, doi: 10.1109/TPWRS.2016.2631891.

[16]. G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," IEEE Transactions on Smart Grid, vol. 8, no. 4, pp. 1630–1638, 2017, doi: 10.1109/TSG.2015.2495133.

[17]. Y. Wang, Z. Lu, Y. Qin, X. Sun, and J. Zhang, "A machine learning approach for cyber attack detection in smart grid," in Proc. IEEE Power and Energy Society General Meeting, 2020, pp. 1–5, doi: 10.1109/PESGM41954.2020.9282058.

[18]. M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor, "Sparse attack construction and state estimation in the smart grid: Centralized and distributed models," IEEE Journal on Selected Areas in Communications, vol. 31, no. 7, pp. 1306–1318, 2013, doi: 10.1109/JSAC.2013.130714.

[19]. J. Sakhnini, H. Karimipour, and A. Dehghantanha, "Smart grid cyber attacks detection using supervised learning and heuristic feature selection," in Proc. IEEE 7th International Conference on Smart Energy Grid Engineering, 2019, pp. 108–112, doi: 10.1109/SEGE.2019.8859946.

[20]. Y. He, G. Yan, C. Wen, S. Tian, and J. Cheng, "Distributed intrusion detection based on federated learning for industrial internet of things," Security and Communication Networks, vol. 2022, pp. 1–12, 2022, doi: 10.1155/2022/3710871.

[21]. F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation forest," in Proc. IEEE International Conference on Data Mining, 2008, pp. 413–422, doi: 10.1109/ICDM.2008.17.

[22]. B. Scholkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," Neural Computation, vol. 13, no. 7, pp. 1443–1471, 2001, doi: 10.1162/089976601750264965.

[23]. L. Ruff, R. Vandermeulen, N. Goernitz, L. Deecke, S. A. Siddiqui, A. Binder, E. Muller, and M. Kloft, "Deep one-class classification," in Proc. International Conference on Machine Learning, 2018, pp. 4393–4402.

[24]. D. Bank, N. Koenigstein, and R. Giryes, "Autoencoders," Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook, pp. 353–374, 2023, doi: 10.1007/978-3-031-24628-9_16.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151106

[25]. M. Stow and A. A. Stewart, "Empirical analysis of SHAP stability under data corruption across datasets and model architectures," International Advanced Research Journal in Science, Engineering and Technology, vol. 12, no. 8, pp. 92–110, 2025, doi: 10.17148/IARJSET.2025.12810.

[26]. M. M. Breunig, H. P. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying density-based local outliers," in Proc. ACM SIGMOD International Conference on Management of Data, 2000, pp. 93–104, doi: 10.1145/342009.335388.

[27]. M. Stow and A. A. Stewart, "Interpreting machine learning predictions with SHAP and LIME for transparent decision making," International Journal of Computer Science and Mathematical Theory, vol. 11, no. 8, pp. 22–49, 2025, doi: 10.56201/ijcsmt.vol.11.no8.2025.pg22.49.

[28]. T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in Proc. International Conference on Machine Learning, 2020, pp. 1597–1607.

[29]. K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 9729–9738, doi: 10.1109/CVPR42600.2020.00975.

[30]. J. Franceschi, A. Dieuleveut, and M. Jaggi, "Unsupervised scalable representation learning for multivariate time series," in Proc. Advances in Neural Information Processing Systems, vol. 32, 2019, pp. 4650–4661.

[31]. E. Eldele, M. Ragab, Z. Chen, M. Wu, C. K. Kwoh, X. Li, and C. Guan, "Time-series representation learning via temporal and contextual contrasting," in Proc. International Joint Conference on Artificial Intelligence, 2021, pp. 2352–2359, doi: 10.24963/ijcai.2021/324.

[32]. Z. Yue, Y. Wang, J. Duan, T. Yang, C. Huang, Y. Tong, and B. Xu, "TS2Vec: Towards universal representation of time series," in Proc. AAAI Conference on Artificial Intelligence, vol. 36, no. 8, 2022, pp. 8980–8987, doi: 10.1609/aaai.v36i8.20881.

[33]. M. Stow, "Revealing minimum demand period vulnerabilities through multi scale pattern analysis of power grid data," International Journal of Scientific and Research Publications, vol. 15, no. 8, 2024.

