Data-Driven Detection of False Data Injection Attacks in Smart Grids using Self-Attention Enabled Conditional GAN with Gradient Penalty

S. Murugesan, *Student Member, IEEE,* C. H. Ram Jethmalani, *Senior Member, IEEE,* K. Navin Sam, *Senior Member, IEEE,* A. Venkadesan, *Senior Member, IEEE,* and S. J. Sadheesh Kumar, *Senior Member, IEEE*

***Abstract*—The increasing integration of advanced information and communication technologies in power systems has increased their vulnerability to cyberattacks. False Data Injection Attacks (FDIAs) are a concerning and widely encountered cyberattack. FDIAs pose a critical threat to the security and reliability of modern power systems by manipulating measurement data. Traditional state estimation techniques often fail to detect stealthy FDIAs, particularly in large-scale grids. This paper proposes a Self-Attention-enabled Conditional Generative Adversarial Network with Gradient Penalty (SACGAN-GP) for effective FDIA detection. The framework leverages a self-attention mechanism to capture long-range dependencies among state variables, enhancing feature representation and improving detection performance. The gradient penalty term ensures training stability and mitigates mode collapse, a common issue in standard GANs. SACGAN-GP can effectively utilize limited labelled attack data and learn robust representations of normal and anomalous patterns, without relying heavily on large, labelled datasets as required by supervised models. Experimental validation on IEEE 14-bus and 118-bus test systems demonstrates that the proposed model performs better than other methods. The proposed method achieves an accuracy of 97.06%, F1-score of 98.28%, and an AUC of 100% on the IEEE 14-bus system, while attaining 95.59% accuracy, 97.39% F1-score, and 100% AUC on the IEEE 118-bus system. Detection times remain under 0.52 seconds, confirming the method's applicability for real-time scenarios. Furthermore, attention heatmaps generated by the model provide interpretable insights into the localized impacts of FDIAs, offering a promising direction for intelligent, secure power grid monitoring.**

Link to graphical and video abstracts, and to code: https://latamt.ieeer9.org/index.php/transactions/article/view/10238

***Index Terms*—Conditional generative adversarial network, Deep learning, False data injection attack, Gradient penalty, Intrusion detection, Self-attention mechanism, Smart grid security.**

The associate editor coordinating the review of this manuscript and approving it for publication was Carlos Thomaz (*Corresponding author: S. Murugesan*).

S. Murugesan, C. H. R. Jethmalani, K. N. Sam, A. Venkadesan, and S. J. S. Kumar are with the Department of Electrical and Electronics Engineering, National Institute of Technology Puducherry, Karikal 609609, India (e-mails: ee22d1004@nitpy.ac.in, ramjethmalani@nitpy.ac.in, navinsam.k@nitpy.ac.in, venkadesan@nitpy.ac.in, and sadheeshkumar@nitpy.ac.in).

I. Introduction

THE United Nations 2030 agenda underscores affordable, clean energy (SDG-7) and resilient, innovative infrastructure (SDG-9). Securing smart grids supports SDG-7 by ensuring reliable electricity, while SDG-9 is advanced through artificial intelligence that strengthens infrastructure resilience and fosters innovation.[1]. The increasing digitization of modern power systems has made them vulnerable to various cyber threats, particularly False Data Injection Attacks (FDIAs)[2]. These attacks manipulate measurement data from sensors and phasor measurement units (PMUs), deceiving control centers, potentially causing blackouts, equipment failures, and economic losses[3]. Ensuring cybersecurity in power grids is critical for maintaining grid stability, reliability, and operational integrity [4]. Modern power grids function as complex cyber-physical systems, integrating intelligent sensors, real-time communication, and advanced control mechanisms to manage bidirectional power and information flows [5]. However, their openness and heterogeneity expose them to significant risks, enabling attackers to infiltrate core systems remotely through external devices and networks. Consequently, it is imperative to detect cyber intrusions early by identifying abnormal patterns in grid measurement data [6]. Traditional model-based FDIA detection techniques rely on state estimation residuals and hypothesis testing [7]. For example, the k-smallest residual test and similar approaches compare estimated and measured values to flag inconsistencies [8]. However, these methods face two fundamental challenges: (i) the high computational cost of state estimation makes them unsuitable for real-time detection in large-scale grids, and (ii) unobservable attacks can evade detection by subtly altering measurements to bypass threshold-based methods.

The proliferation of smart devices and advanced metering infrastructure (AMI) has significantly increased the volume and speed of measurement data in power systems, necessitating efficient and intelligent data management solutions [9]. This data abundance, combined with advancements in artificial intelligence (AI), has motivated the use of machine learning (ML) and deep learning (DL) for FDIA detection. A hybrid intrusion detection system integrating particle swarm optimization and grey wolf optimization with convolutional

neural networks (CNN) and long short-term memory (LSTM) classifiers is used for FDIA detection in [10]. This method lacks deployment in large-scale power systems. A graph autoencoder model offering location-aware FDIA detection enhancing resilience against unknown attacks is proposed in [11]. However, the model's heavy computational requirements and dependence on diverse topological data make it difficult to deploy. An attention-based temporal graph variational autoencoder (AT-GVAE)–based locational detector that reconstructs power system states is proposed in [12]. However, this method's unsupervised design and lack of spatial dependency modelling may reduce its detection accuracy in complex power networks. A diffusion-based synthetic FDIA sample generator combining diffusion model with attention-based generative adversarial networks (GANs) is proposed in [13]. The model's manual tuning process adds complexity to deployment in practical smart grid environments. An ensemble of tree-based ML models (Extra Tree, Random Forest, XGBoost) are used in [14]. The models' dependence on idealized datasets restricts their effectiveness in unpredictable, data-scarce, and adversarial real-world grid conditions. Further a probabilistic deep auto encoder model is presented in [15]. Moreover, it is designed for datasets with discrete or limited consecutive outliers, which restricts its adaptability to continuous or large-scale data corruption. mixture Gaussian distribution learning method is proposed in [16]. However, its effectiveness is reduced when training data is limited or when the system operates under high variability. Additionally, the method assumes a fixed threshold and stable feature distributions, which may reduce adaptability in dynamic grid conditions. Support vector machines (SVMs) are proposed for FDIA detection in [17]. Moreover, their sensitivity to high-dimensional and imbalanced datasets restricts their scalability and performance. Deep Neural Networks (DNNs) are presented in [18] have shown promise in identifying FDIAs. however, their reliance on high-resolution temporal data makes them less effective under low sampling rates. CNN based approaches have been employed to capture spatial dependencies in grid data [19]. However, these methods typically require extensive offline training and careful parameter tuning, limiting their adaptability. Graph Neural Networks (GNNs) have also been used for topology-aware FDIA detection [20], but they often struggle to efficiently adapt to significant changes in grid topology, necessitating frequent retraining with updated data. Recent advancements have introduced federated DL frameworks for decentralized FDIA detection [21]. While these methods enhance data privacy, they demand considerable computational resources for encryption and distributed model training, making them less practical for large-scale deployment. A federated semi-supervised class-rebalanced (Fed-SCR) approach is proposed in [22] enables collaborative learning across substations without sharing raw data. Although it helps protect privacy, it requires substantial local computation. Similarly, a federated learning framework designed for edge-based FDIA detection is proposed in [23]. It preserves user privacy but can suffer reduced accuracy when confronted with highly dynamic or previously unseen attack patterns.

Most existing supervised learning approaches for detecting FDIAs rely on large volumes of labelled data that span a wide range of attack scenarios. However, collecting such comprehensive labelled datasets is impractical in large-scale, dynamic power systems. To address this limitation, semi-supervised methods such as GANs have been explored. For instance, [24] proposed a semi-supervised GAN that leverages both labelled and unlabelled data. Despite these advantages, such models often assume balanced class distributions, which do not accurately reflect real-world FDIA occurrences. Furthermore, traditional GANs face challenges like training instability and mode collapse, resulting in inconsistent performance on actual power grid data [25]. Recent advancements, such as the self-attention-based GAN in [26], have demonstrated the benefits of incorporating attention mechanisms for FDIA detection in smart grids. However, these approaches often lack conditional data generation and gradient penalty, which are essential for learning structured dependencies and ensuring training stability. To overcome these issues, this paper introduces a Self-Attention enabled Conditional Generative Adversarial Network with Gradient Penalty (SACGAN-GP). The main contributions of this paper are summarized as follows:

1. To enhance the accuracy of FDIA detection in the power grid, propose a novel FDIA detection framework based on Conditional GANs enhanced with a self-attention mechanism and gradient penalty by integrating the topology and data of the power grid.
2. The self-attention module captures long-range dependencies in measurement data, facilitating the detection of stealthy and unobservable attacks.
3. The gradient penalty enforces Lipschitz continuity in the discriminator, enhancing training stability and convergence.
4. The proposed detection method is comprehensively evaluated on the standard IEEE 14-bus and 118-bus systems, and its performance is compared with state-of-the-art GAN-based approaches.

The remainder of this paper is structured as follows: Section II provides the background on power system state estimation, FDIAs, and formalizes the detection problem. Section III presents the proposed SACGAN-GP model architecture and training methodology. Section IV discusses experimental results and performance comparisons. Finally, Section V concludes the paper and outlines future research directions.

II. Background

FDIAs exploit power system communication infrastructure vulnerabilities to tamper with measurement data. By carefully crafting malicious inputs, attackers can alter state estimation results without triggering traditional False Data Detection (FDD) mechanisms. This section presents an overview of power system state estimation, the principle behind FDIAs, and the formal problem formulation for FDIA detection.

A. Power System State Estimation

Power system state estimation is used to determine the most probable real-time operating state of the grid. It estimates the voltage magnitudes and phase angles at various buses based on a set of measurements (e.g., power flows, injections, and voltage magnitudes.

The measurement model equation (1) is expressed as

$$z = Hx + v \quad (1)$$

where $z \in \Re^m$ is the measurement data and m is the number of measurement data. $x \in \Re^n$represents the system state variables (voltage, angles) and n represents the number of state variables, $H \in \Re^{mXn}$ is the Jacobian matrix determined by system topology, and v denotes the measurement noise vector, typically assumed to follow a Gaussian distribution with zero mean. The estimated state $\tilde{x}$ (2) is obtained by equation (2)

$$\tilde{x} = (H^T RH)^{-1} H^T Rz \quad (2)$$

Where R is the weight matrix, after power system state estimation, the FDD mechanism is used to detect, identify, and eliminate false data.

B. False Data Injection Attack

In FDIA detection mechanisms, the attack is generally expressed as $Z_a = z + a$, in practice the attack vector a originates from different sources of intrusion. Typical vectors include man-in-the-middle manipulation of Modbus/SCADA traffic, compromised RTUs or PMUs, and malicious firmware modifying outgoing measurements. These falsified values are constructed to remain consistent with the state-estimation model, producing unobservable attacks that bypass residual-based bad-data detection and mislead control decisions [27]. The proposed SACGAN-GP model addresses this risk by learning spatial dependencies among bus measurements; deviations caused by coordinated falsification result in elevated anomaly scores and improved detection reliability. An adversary injects a malicious vector $a \in \Re^m$into the original measurements (z) to produce a compromised vector z_a as in (3)

$$Z_a = z + a \quad (3)$$

If the attack vector is constructed as $a = H_c$, where $c \in \Re^n$ is the intended change in the estimated state, the resulting estimated state $\hat{x}_a$ results in (4)

$$\hat{x}_a = [H^T R^{-1} H]^{-1} H^T R^{-1} Z_a = \hat{x} + c \quad (4)$$

This manipulation enables the attacker to bias the estimated system state toward a malicious target while preserving consistency with the power system model, evading detection by traditional FDD mechanisms. These unobservable attacks are particularly dangerous, as they are crafted to mimic normal operating conditions, making them difficult to identify.

C. Problem Formulation

FDIA detection aims to build a robust model that identifies malicious measurement patterns, even when they evade conventional FDD techniques. Let the training dataset be defined as $T = \{X_1, X_2, \ldots X_\delta\}$ containing predominantly normal samples, where $\delta >> \beta$. A smaller labelled dataset $S = \{(X_1^S, Y_1), (X_2^S, Y_2), \ldots, (X_\beta^S, Y_\beta)\}$ includes a limited number of attack samples. Here,$Y_i \in \{0,1\}$ indicates whether a sample is normal or attacked.

The FDIA detection model is trained to assign an anomaly score A(X) to a given sample. A sample is flagged as malicious if the score exceeds a predefined threshold λ:

The detection performance is evaluated based on how accurately the model distinguishes attacked samples from normal samples under this decision rule.

Fig. 1. illustrates the workflow of the proposed False Data Injection Attack (FDIA) detection methodology. Power system measurements and transmission line parameters are first processed through a state estimation block. Observable FDIAs are identified using traditional residual-based detection methods. However, unobservable FDIAs bypass this stage and are further analyzed using a SACGAN-GP model. Historically

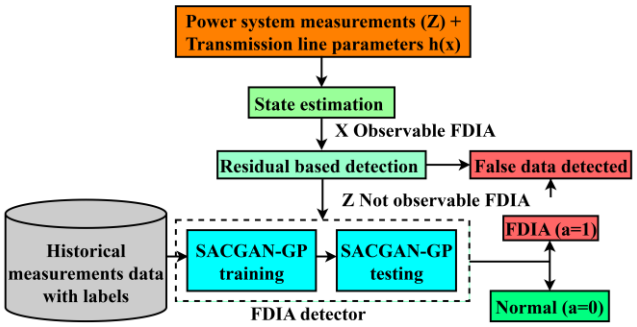

Fig. 1. FDIA detection framework integrating residual based state estimation and SACGAN-GP for observable and stealthy attacks.

labelled measurement data are used to train the SACGAN-GP-based FDIA detector, which classifies test instances as normal (label = 0) or FDIA (label = 1). This framework effectively identifies both observable and stealthy (unobservable) attacks, enhancing detection robustness and grid security.

III. Proposed SACGAN-GP for FDIA

The proposed SACGAN-GP framework is shown in Fig. 2. It consists of four main components: Generator, Discriminator, Self-Attention Mechanism and Gradient penalty. The model follows an adversarial learning approach, where the Generator creates synthetic attack samples, and the Discriminator distinguishes between real and fake data. The self-attention mechanism enhances FDIA detection in power systems by capturing long-range dependencies and complex feature interactions within measurement data. Gradient Penalty ensures training stability and sharp decision boundaries, contributing to improved generalization on unseen attack scenarios.

This architecture presents a SACGAN-GP for stabilized training, incorporating self-attention mechanisms and Gradient Penalty for improved feature learning. In this architecture a generator that synthesizes samples conditioned on binary labels (normal/anomaly) through concatenated noise and label embeddings, processed via dense layers and self-attention with residual connections, and a discriminator that evaluates samples using identical attention mechanisms while enforcing Lipschitz continuity through gradient penalties. During training, the

system employs a 5:1 update ratio between the discriminator and generator, utilizing Wasserstein loss for meaningful gradient signals. The discriminator's output scores are thresholded at the 75th percentile of normal samples for anomaly detection, enabling classification without additional networks. Integrating self-attention allows the capture of long-range dependencies in the data, while the Gradient Penalty framework ensures training stability. This unified approach demonstrates advantages in synthetic sample quality, training convergence, and anomaly detection performance through standard metrics, particularly benefiting from the discriminator's dual role in both adversarial training and anomaly scoring. The architecture represents an advancement over traditional GANs by combining conditional generation, attention mechanisms, and stabilized training for improved anomaly detection capabilities.

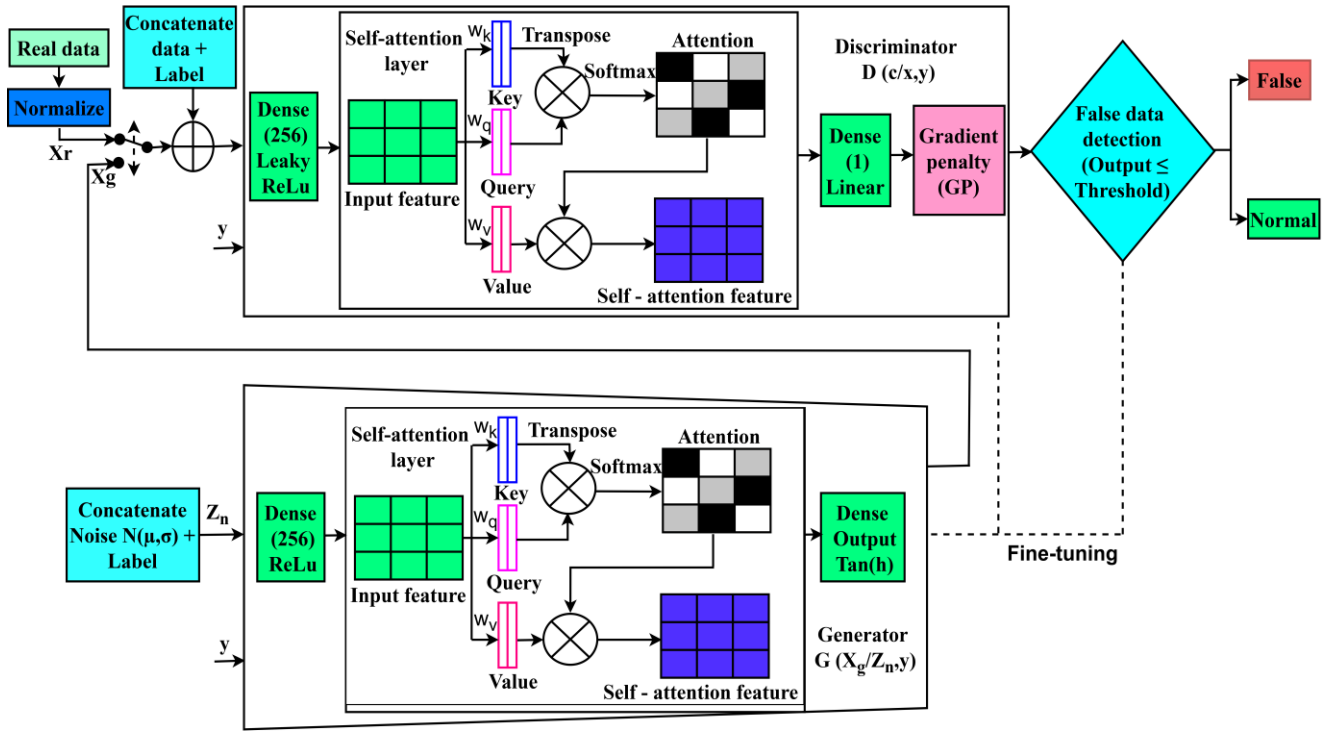

Fig. 2. The architecture of the proposed SACGAN-GP model for FDIA detection.

A. Generator Model (G^{θ}):

The generator model plays a central role in simulating realistic false data injection patterns to evaluate and enhance the resilience of power system detection mechanisms. The Generator aims to approximate the true data distribution $P_{data}(x)$ by mapping a random noise vector $z{\sim}p_z(z)$ into the data space. Additionally, the conditional nature of the model ensures that the generated data aligns with a given label y. $G^{\theta}(z,y)$ outputs a synthetic attack sample that mimics real measurements as given by equation (5)

$$G^{\theta}(z,y) = \Re^{d_z} \times \Re^{d_y} \rightarrow \Re^{d_x} \tag{5}$$

Where, $z{\sim}p_z(z)$ is sampled from a Gaussian noise distribution. $y \in \Re^{d_y}$ represents the class labels (i.e., normal or attack) and $z \in \Re^{d_z}$ is a latent vector sampled from a Gaussian distribution,

The loss function for the generator is given by equation (6)

$$L^{G} = -E_{z\sim p_z(z),y\sim p(y)}[D^{\varphi}G^{\theta}(z,y),y)] \tag{6}$$

Where D^{φ} Discriminator function parameterized by φ, outputs real/fake score. This promotes the generator to produce samples that fool the discriminator.

B. Discriminator Model (D^{φ}):

The discriminator is trained concurrently with the generator during the training process to enhance adversarial learning and improve detection performance, The Discriminator D^{φ} act as a binary classifier that distinguishes between non-attack and attack samples. It takes input x and the corresponding label y and outputs a score representing the likelihood that the sample is real, as expressed in the equation (7)

$$\mathrm{G}^{\varphi}(\mathrm{x},\mathrm{y}) = \Re^{\mathrm{d_x}} \times \Re^{\mathrm{d_y}} \rightarrow \Re \tag{7}$$

The Wasserstein loss with Gradient Penalty (WGAN-GP) is employed to enhance Training stability and enforce Lipschitz continuity. The discriminator's objective function is defined as in (8)

$$L^{D} = -E_{x\sim pdata}[D^{\varphi}(x,y)] - E_{\hat{x}}D^{\varphi}(\hat{x},y)] + \lambda E_{\hat{x}\sim p\hat{x}}\left[\left(\left\|\nabla_{\hat{x}}D^{\varphi}\left(\hat{x},y\right)\right\|_2 - 1\right)^2\right] \tag{8}$$

Where $x{\sim}pdata$ denotes real data samples, $\tilde{x} = G^{\theta}(z,y) \sim P_G$ are represented by generated fake samples, $\hat{x}{\sim}\ p_{\hat{x}}$ is sampled uniformly along straight lines between pairs of real and fake samples, and λ is the gradient penalty coefficient. The gradient

penalty ensures Lipschitz continuity, stabilizing training and avoiding mode collapse.

C. Self-Attention Mechanism in CGAN-GP:

Self-attention mechanisms enable the model to weigh the relevance of each input feature in relation to others, improving the detection of subtle and spatially distributed anomalies in power system data. In standard GANs, CNNs struggle to capture long-range feature dependencies in high-dimensional data. The self-attention mechanism resolves this by allowing each feature vector to attend to all others in the input space. For an input feature map $X \in \Re^{NXd}$ where N is the number of features and d is the feature dimension, the self-attention output is defined as in equation (9)

$$Attention(Q,K,V) = soft\,max\left(\frac{QK^T}{\sqrt{d_K}}\right)V \quad (9)$$

where $Q = XW_Q$, $K = XW_K$, and $V = XW_V$ are the query, key, and value matrices derived from learned projections, and d_K is the dimensionality of the key vectors. This formulation lets the model selectively focus on the most relevant features when generating or evaluating data samples. The self-attended features are then concatenated with CGAN features before being passed to the next layer. This enhances feature representation and improves FDIA detection accuracy**.** Integrates self-attention and a gradient penalty with CGAN, as discussed in the next section.

During training, the generator and discriminator are updated in an alternating manner using adversarial optimization, with conditional inputs guiding both generation and classification tasks. In the detection phase, the trained discriminator computes anomaly scores and classifies attack types based on conditional information, allowing for accurate and robust detection of false data injection attacks in complex power system environments. The Pseudocode for SACGAN-GP is shown in Algorithm 1.

D. Training Algorithm for SACGAN-GP:

Algorithm 1 shows the training process of the proposed SACGAN-GP based FDIA. The algorithm begins by loading the voltage and angle measurement dataset D, along with their corresponding FDIA labels. The feature matrix X and label vector y are extracted and normalized to a [0, 1] range using MinMaxScaler. The dataset is then divided into training and testing subsets.

The generator G^θ is designed by combining a random noise vector $z \in \mathbb{R}^{d_z}$ with conditional information attack type (label y) as in equation (5). This concatenated input is passed through dense layers and a self-attention block to capture spatial and temporal dependencies among buses, producing synthetic measurement data x_{attack}. The discriminator D^φ takes either normal or attacked samples concatenated with their respective labels and processes them through dense layers and a self-attention mechanism, ultimately producing a score that indicates whether the input is normal or an attack.

During training, the algorithm iteratively updates both the generator and discriminator through adversarial learning. For each epoch, the discriminator is first updated over several iterations by comparing its responses to normal samples (X_{normal}, y_{normal}) and attacked samples(x_{attack}, y_{attack}). The discriminator loss L^D is calculated using the Wasserstein distance formulation augmented with a gradient penalty term, which enforces the Lipschitz constraint to stabilize training. Following the discriminator update, the generator is optimized to fool the discriminator by generating data that it classifies as real. The generator loss L^G is computed as the negative expected output of the discriminator when evaluating fake samples. Both networks are optimized using the Adam optimizer. Once the training is completed after E epochs, the final trained generator and discriminator models are returned.

The entire structured sequence of steps involved in training the SACGAN-GP to learn a robust mapping from latent space to realistic power system measurements, conditioned attack labels, while effectively distinguishing between normal and manipulated data patterns for accurate FDIA detection.

IV. RESULTS AND DISCUSSION

All simulations were conducted on a system equipped with an Intel(R) Core (TM) i5-8400 CPU @ 2.80GHz, 8 GB RAM, and a 64-bit x64-based Windows operating system. The proposed SACGAN-GP was implemented using TensorFlow and Kera's libraries in Python.

A. Dataset Generation and Model Parameters

Since no publicly available dataset exists for False Data Injection Attacks (FDIAs), a synthetic dataset was generated using the Pandapower library. Power flow analyses were conducted using the Newton–Raphson method, and state estimation was carried out with the Weighted Least Squares (WLS) algorithm. The resulting system state variables were extracted from the measured data, including voltage magnitudes and phase angles.

Non-attacked data were produced by scaling the load demands at each bus to 25%, 50%, 75%, and 100% of their nominal values, with the corresponding voltage magnitudes and angles computed and stored as state variables. To simulate attacked data, FDIAs were introduced by modifying voltage magnitudes and angles by 5%–50% on selected buses (0, 1, and 5). Additionally, systematic angle deviations of up to 5° were applied to emulate coordinated attack strategies. These alterations bypass residual-based Bad Data Detection (BDD) mechanisms, thereby modifying the estimated states without triggering detection.

Each data sample represents the electrical state of the system under specific operating conditions, consisting of 34 features (voltage, angle, and load information). 336 labelled samples were generated, divided into 268 for training (80%) and 68 for testing (20%). Of the test set, 55 correspond to non-attacked states and 13 to attacked states, introducing class imbalance consistent with realistic scenarios. Attacked data points were labelled as "1," while normal samples were labelled as "0," forming a binary classification framework suitable for FDIA detection. To emphasize the severity and stealthiness of FDIAs, small deviations such as 5%–50% in voltage magnitudes and up to 5° in angles were sufficient to significantly distort state

estimation while remaining undetected by traditional BDD techniques. This study employs 336 samples for the IEEE 14-bus system and 2832 samples for the IEEE 118-bus system. Although the dataset size appears limited, the proposed SACGAN-GP framework is specifically designed to operate effectively in data-scarce environments. Its robustness under limited data conditions is primarily attributed to the semi-supervised learning paradigm of the conditional GAN and the stabilizing effect of the gradient penalty. In SACGAN-GP, the semi-supervised discriminator jointly models both normal and attacked operating states, enabling it to learn an implicit representation of the underlying data distribution even from a small number of samples.

Algorithm 1: Pseudocode for Training Procedure SACGAN-GP

Input: Dataset D (voltage/angle measurements and FDIA labels), total epochs E, batch size B, noise vector size dz, gradient penalty coefficient λ

Output: Trained generator G(θ) and discriminator D(ϕ)

Load dataset D
Extract feature matrix (X) and label vector (y)
Normalize (X) to range [0, 1] using MinMaxScaler
Split (X, y) into training and test sets: (X_{train}, y_{train}) and (X_{test}, y_{test})
Generator G(θ)
  Concatenate a random noise vector $z \in \mathbb{R}^{d_z}$ with a label embedding (y)
  Pass x through dense layers, the self-attention block, and the output layer to generate x_{attack}
Discriminator D(ϕ)
  Concatenate input (x) with label (y)
  Pass through dense layers, the self-attention block, and the output layer to produce a normal/attack score
For epoch e = 1 to E do
  For t = 1 to critic do
    Sample a batch of real samples (X_{normal}, y_{normal}) from the training data
    Sample noise $z \sim \mathcal{N}(0, I)$ and labels y_{attack}
    Generate attack samples: x_{attack} = G(z, y_{attack})
    Compute discriminator loss: $L^D = \mathbb{E}[D(x_{attack}, y_{attack})] - \mathbb{E}[D(x_{normal}, y_{normal})] + \lambda \cdot GP\ (x_{normal}, x_{attack})$(8)
    Update ϕ (discriminator parameters) using Adam optimizer
  End For
  Sample noise $z \sim \mathcal{N}(0, I)$ and labels (y)
  Generate attack samples: x_{attack}, = G (z, y)
  Compute generator loss: $L^G = -\mathbb{E}\ [D\ (x_{attack}, y)]$ (6)
  Update θ (generator parameters) using Adam optimizer
End For
Return trained generator G(θ) and discriminator D(ϕ)

Furthermore, the incorporation of a gradient penalty enforces Lipschitz continuity in the discriminator, which stabilizes adversarial training, mitigates mode collapse, and prevents overfitting. The joint modelling reduces reliance on large labelled datasets while preserving discriminative capability.

Model validation was done through performance comparison with baseline classifiers using standard evaluation metrics. In addition, ablation studies of architectural components and sensitivity analyses of key hyperparameters were performed. The training dynamics and convergence behaviour were also examined. Finally, the model's capability to distinguish between normal and attacked samples was visualized using t-distributed Stochastic Neighbor Embedding (t-SNE) for dimensionality reduction and clustering.

"The authors confirm that the complete source code and the synthetic dataset generation procedures used in this work will be made available publicly upon acceptance of the manuscript".

B. Quantitative Performance Evaluation

The detection performance was evaluated using standard classification metrics: Accuracy, Precision, Recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC) and Detection time. Fig. 3. and Fig. 4. shows the FDIA detection metrics of the six methods in the IEEE14-Bus and IEEE118-Bus system.

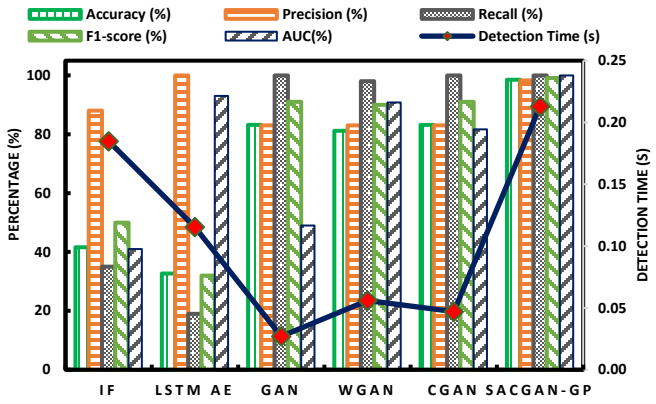

Fig. 3. Comparison of FDIA detection methods on the IEEE 14-bus system in terms of accuracy, precision, recall, F1-score, AUC and detection time.

Fig. 3. shows the comparative performance of various FDIA detection methods on the IEEE 14-bus system. The proposed SACGAN-GP model demonstrates superior performance across all key evaluation metrics, including accuracy, precision, recall, F1-score, and AUC. Notably, SACGAN-GP achieves near-perfect detection capability with an accuracy of 97.06%, recall of 100%, and an F1-score of 98.28%, outperforming traditional models such as Isolation Forest, LSTM Autoencoder, GAN, WGAN, and CGAN. Although the detection time slightly increases to approximately 0.21 seconds, it remains within acceptable limits for real-time applications. These results confirm the model's robustness, stability, and effectiveness in identifying stealthy FDIAs in power systems.

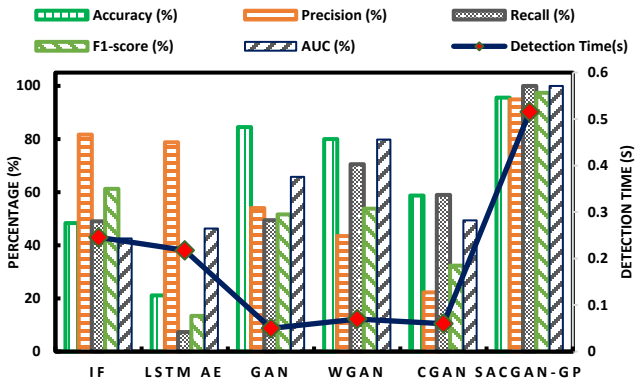

Fig. 4. Comparison of FDIA detection methods on IEEE 118-bus system demonstrating scalability and detection effectiveness of SACGAN-GP.

Fig. 4. illustrates the comparative performance of six FDIA detection methods evaluated on the IEEE 118-bus system. The proposed SACGAN-GP model consistently outperforms all baseline models Isolation Forest, LSTM Autoencoder, GAN, WGAN, and CGAN across all key metrics, including accuracy (95.59%), precision (94.91%), recall (100%), F1-score (97.39%), and AUC (100%). The performance and detection time are compared in Tables I and II for the IEEE 14-bus and 118-bus systems, respectively. As shown in Table I and II, SACGAN-GP consistently achieves improved detection metrics when compared to baseline methods. This enhanced detection performance, however, is accompanied by a modest increase in detection time. Specifically, the average detection time increases to 0.5158 seconds for the IEEE 118-bus system, compared to the significantly lower time 0.21 seconds observed in the IEEE 14-bus case. This increase is primarily attributed to the higher number of measurement features and the quadratic complexity of the self-attention operation with respect to system size. The performance improvement is beneficial as long as the detection time is within the control loop operation time of different layers of the power system.

Different layers of power system control function have varying temporal resolutions. For example, primary control actions, such as, immediate frequency regulation should respond within seconds; secondary control mechanisms, including Automatic Generation Control, typically act within a few seconds to several minutes; and tertiary control processes, including economic dispatch and state estimation, generally execute on a minute-scale. Therefore, a detection time below one second is acceptable [28-29]. The detection time of SACGAN-GP, below one second, allows it to identify malicious data injections before the subsequent state estimation cycle begins, enabling operators to intervene before corrupted measurements influence operational decisions. This response capability significantly reduces the risk of cascading errors that could compromise system stability, economic efficiency, or equipment safety. Consequently, the achieved detection time confirms that the proposed framework not only delivers high detection accuracy but is also computationally efficient enough to be integrated into real-time control center workflows, reinforcing its suitability for enhancing smart grid resilience, demonstrates its practical applicability for real-time power grid monitoring. The results confirm that integrating self-attention mechanisms and gradient penalty into the conditional GAN framework significantly enhances the model's capability to detect stealthy and complex FDIAs, making SACGAN-GP a reliable and scalable solution for securing large-scale power systems.

TABLE I
PERFORMANCE COMPARATION OF DETECTION METHODS IN THE IEEE14-BUS POWER SYSTEM

Structure	Accuracy (%)	Precision (%)	Recall (%)	F1score (%)	AUC (%)	Detection Time (s)
Isolation Forest	41.58	88.00	35.00	50.00	41.00	0.18
LSTM AE	32.67	100.00	19.00	32.00	93.00	0.12
GAN	83.17	83.00	95.00	91.00	49.00	0.03
WGAN	81.19	83.00	98.00	90.00	90.81	0.06
CGAN	83.17	83.00	99.00	91.00	81.65	0.05
SACGAN-GP	97.06	96.61	100.00	98.28	100.00	0.21

TABLE II
PERFORMANCE COMPARISON OF DETECTION METHODS IN THE IEEE118-BUS POWER SYSTEM

Structure	Accuracy (%)	Precision (%)	Recall (%)	F1score (%)	AUC (%)	Detection Time(s)
Isolation Forest	48.35	81.64	49.01	61.25	42.57	0.2456
LSTM AE	21.17	78.78	7.34	13.43	46.29	0.2177
GAN	84.47	54.02	49.47	51.64	65.76	0.05
WGAN	80	43.5	70.52	53.81	79.83	0.07
CGAN	58.73	22.31	58.94	32.36	49.37	0.06
SACGAN-GP	95.59	94.91	100	97.39	100	0.5158

The SACGAN-GP performed better than all baseline models, demonstrating its better generalization and robustness in detecting FDIAs. Notably, the higher precision and recall values indicate a low false positive and false negative rate, critical in real-world power systems where misclassification can lead to operational inefficiencies or vulnerabilities. It may be noted that the proposed SACGAN-GP has greater recall value for both the test systems.

C. Training Dynamics and Convergence Behaviour of SAGAN-GP

Fig. 5. illustrates the generator and discriminator loss trends across training epochs, along with the gradient penalty and conditional loss components. The SACGAN-GP exhibits smoother training dynamics, which help it to avoids mode collapse, a common challenge in adversarial learning.

D. Measurement Distribution and Dimensionality Reduction

To evaluate the model's ability to distinguish between real and generated data, t-distributed Stochastic Neighbor Embedding (t-SNE) was applied for dimensionality reduction and cluster visualization.

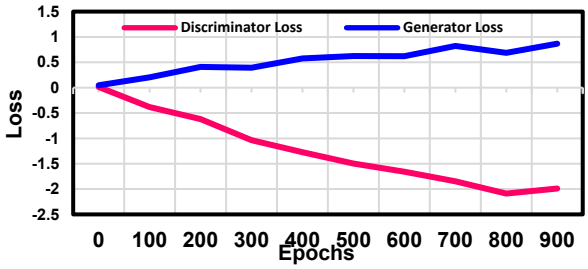

Fig. 5. Training convergence of SACGAN-GP showing generator and discriminator loss variation over epochs.

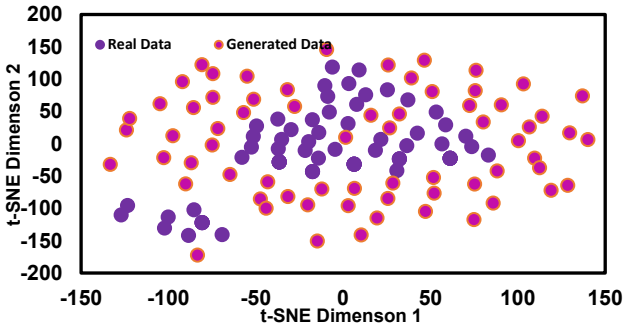

Fig. 6. t-SNE visualization illustrating cluster separation between normal and attacked measurement samples.

As shown in Fig. 6. the SACGAN-GP model successfully forms two well-separated clusters representing non-attacked and attacked states, confirming its ability to capture the underlying data distribution. The minimal overlap between real and synthetic samples suggests high generative accuracy, while the clear separation highlights the discriminator's effectiveness in distinguishing anomalies. A single generator was sufficient to represent the data distribution, as only two distinct classes normal and attacked were present in the dataset.

E. Anomaly Score Histogram for FDIA Discrimination

The anomaly score histogram is shown in Fig. 7. Demonstrates the model's ability to distinguish between normal (Non-Attacked) and attacked data. Normal samples show low anomaly scores clustered to the left, whereas attacked samples exhibit higher scores on the right. A clear threshold separates the two distributions, serving as an effective decision boundary. This separation confirms that the SACGAN-GP models accurately assign high scores to manipulated data and low scores to normal samples, minimizing false positives and negatives. the results demonstrate the models' effectiveness in robust and reliable FDIA detection.

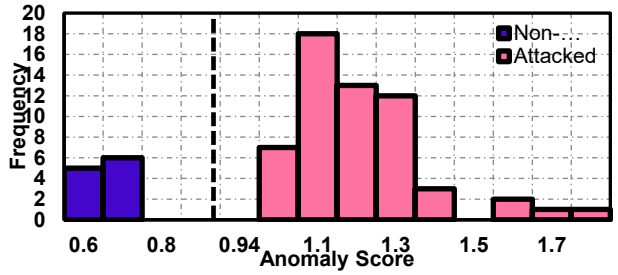

Fig. 7. Anomaly score distribution of normal and attacked samples with threshold-based separation.

F. Ablation Study

To evaluate the individual contributions of key components in the proposed SACGAN-GP architecture, an ablation study was performed. This involved systematically removing specific modules such as self-attention layer, gradient penalty, and conditional inputs, and analyzing their impact on detection performance. Tables III and IV present the performance comparison of ablated variants and baseline CGAN on the IEEE 14-bus and IEEE 118-bus systems, respectively. From Tables III and IV, it is inferred that SACGAN-GP model consistently performs better than its ablated versions in terms of Accuracy, Precision, Recall, F1-score, Area Under the Curve (AUC), and Detection Time.

TABLE III
PERFORMANCE COMPARISON OF ABLATED VARIANTS IN THE IEEE14-BUS POWER SYSTEM

Structure	Accuracy (%)	Precision (%)	Recall (%)	F1score (%)	AUC (%)	Detection Time(s)
SACGAN-GP	97.06	96.61	100.00	98.28	100.00	0.2130
CGAN-GP	82.35	83.58	98.25	90.32	26.48	0.3357
SACGAN	66.17	49.27	75.55	59.64	56.26	0.1329
CGAN	66.17	49.25	73.33	58.92	50.18	0.1062

TABLE IV
PERFORMANCE COMPARISON OF ABLATED VARIANTS IN THE IEEE118-BUS POWER SYSTEM

Structure	Accuracy (%)	Precision (%)	Recall (%)	F1score (%)	AUC (%)	Detection Time(s)
SACGAN-GP	95.59	94.91	100.00	97.39	100.00	0.5158
CGAN-GP	83.07	83.22	99.79	90.75	40.84	0.5192
SACGAN	66.66	50.00	74.60	59.87	66.89	0.8714
CGAN	65.34	47.43	73.62	57.69	64.71	0.3690

Particularly, in the IEEE 14-bus system, SACGAN-GP achieved an F1-score of 98.28% and an AUC of 100%, significantly performing better than CGAN-GP model. Also, in the larger and more complex IEEE 118-bus system, SACGAN-GP continued to deliver strong results, achieving an F1-score of 97.39% and an AUC of 100%. These results highlight the model's robustness in both small- and large-scale systems and confirm that the integration of self-attention, conditional generation, and gradient penalty contributes to superior detection accuracy and stable training dynamics. Removing the self-attention layer significantly drops both F1-score and AUC, confirming its role in enhancing spatial awareness and interpretability. Excluding the gradient penalty results in unstable training and poorer generalization, evident from reduced performance metrics.

G. Hyperparameter Sensitivity Analysis

The investigation of the sensitivity of the SACGAN-GP to key hyperparameters such as learning rate, noise dimension, batch size, gradient penalty coefficient (λ_{gp}), and attention head count for IEEE 14 bus system and IEEE 118 bus system are given in tables V and VI respectively. Here, the performance evaluation is carried out in terms of AUC range. The following inferences are drawn from table V and VI

1. Learning rate was tested at values of 1×10^{-4}, 1×10^{-5}, and 1×10^{-6}. The optimal value of 1×10^{-5} consistently resulted in higher AUC scores across both systems. Extremely low or high learning rates led to unstable or excessively slow convergence.
2. For noise dimension, values of 50, 100, and 200 were evaluated. A dimension of 100 was found optimal, achieving the highest AUC, and demonstrating a good balance between latent space complexity and data fidelity.
3. Batch size varied among 32, 64, and 128. A batch size of 64 produced the best results in terms of both convergence speed and generalization. Larger batch sizes led to underfitting, particularly in later epochs.
4. The gradient penalty coefficient was tested with values of 0.1, 1, and 10. A value of 1 yielded the most stable training and the highest AUC, confirming its importance in regularizing the discriminator and avoiding gradient explosion or vanishing.
5. The number of attention heads varied between 1, 2, and 4. Using 2 attention heads provided the best tradeoff between representational capacity and computational efficiency, resulting in the most favorable AUC range.

TABLE V
SACGAN-GP UNDER DIFFERENT HYPERPARAMETER SETTINGS: IEEE 14-BUS SYSTEM

Hyperparameter	Range Tested	Optimal Value	AUC Range
Learning Rate	1×10^{-4}, 1×10^{-5}, 1×10^{-6}	1×10^{-5}	0.925 – 0.973
Noise dimension	50, 100, 200	100	1.000 - 1.000
Batch Size	32, 64, 128	64	0.948 – 0.973
Gradient Penalty (λ_{gp})	0.1, 1, 10	1	0.951 – 0.973
Attention Heads	1, 2, 4	2	0.945 – 0.973

TABLE VI
SACGAN-GP UNDER DIFFERENT HYPERPARAMETER SETTINGS: IEEE 118-BUS SYSTEM

Hyperparameter	Range Tested	Optimal Value	AUC Range
Learning Rate	1×10^{-4}, 1×10^{-5}, 1×10^{-6}	1×10^{-5}	0.925 – 0.973
Noise dimension	50, 100, 200	100	0.999 - 1.000
Batch Size	32, 64, 128	64	0.948 – 0.973
Gradient Penalty (λ_{gp})	0.1, 1, 10	1	0.951 – 0.973
Attention Heads	1, 2, 4	2	0.945 – 0.973

Across both datasets, the model demonstrated strong robustness to moderate variations in hyperparameters. However, the analysis confirms that precise tuning, particularly of the learning rate, batch size, gradient penalty and attention configuration, can significantly enhance performance.

H. Impact of Measurement Noise

Presence of noise in real-world power systems is inevitable. Therefore, evaluating the robustness of the detection method under noisy conditions is essential. Hence, the proposed method is tested under different noise levels, namely low-level noise at 1% (PMU noise), medium-level noise at 5% (SCADA noise) and high-level noise at 10% (sensor noise). The performance of the proposed SACGAN-GP is presented in Tables VII and VIII.

TABLE VII
PERFORMANCE AT DIFFERENT NOISE LEVELS IN IEEE14-BUS SYSTEM

SACGAN-GP	Without Noise	1% Noise level ($\sigma = 0.01$)	5% Noise level ($\sigma = 0.05$)	10% Noise level ($\sigma = 0.10$)
Accuracy (%)	97.06	95.59	95.65	94.12
Precision (%)	96.61	95.00	94.83	94.92
Recall (%)	100.00	100.00	96.49	98.25
F1score (%)	98.28	97.44	95.65	96.55
ROC-AUC score (%)	100.00	100.00	100.00	98.41
PR-AUC score	1	1	1	0.99

TABLE VIII
PERFORMANCE AT DIFFERENT NOISE LEVELS IN IEEE118-BUS SYSTEM

SACGAN-GP	Without Noise	1% Noise level ($\sigma = 0.01$)	5% Noise level ($\sigma = 0.05$)	10% Noise level ($\sigma = 0.10$)
Accuracy (%)	95.59	95.41	94.89	95.41
Precision (%)	94.91	94.90	94.87	94.91
Recall (%)	100	99.79	99.14	99.79
F1score (%)	97.39	97.28	96.96	97.39
ROC-AUC score (%)	100	100	98.04	99.89
PR-AUC score	1	1	0.99	0.98

From tables VII and VIII, it can be inferred that there are slight variations in the performance metrics for different scenario of noise injection do not significantly affect the detection capability of the SACGAN-GP model, highlighting its robustness to measurement noise. Further, the stable values of the Area Under the Receiver Operating Characteristic curve (ROC -AUC) score characterizes the presence of a stable trade-off between true positive and false positive rates across decision thresholds. The Precision-Recall (PR-AUC) scores near unity show that the proposed SA-CGAN-GP model maintains stable detection performance even under increasing noise levels, demonstrating robustness to practical sensor inaccuracies and confirming its applicability in real smart-grid environments.

I. Limitations and Future Directions

While the attention mechanism captures long-range dependencies effectively, it also adds a significant computational burden as the system size increases. For very large power systems, the expected quadratic growth in computational complexity can be effectively addressed by adopting scalable attention variants, such as sparse and hierarchical attention mechanisms. These approaches can significantly reduce computational and memory overhead while retaining the ability to capture critical system-wide dependencies. The proposed SACGAN-GP framework use simulated data, along with the assumption of a centralized training environment. These limitations can be mitigated with

the availability of real-time operational datasets and by adopting scalable variants of the proposed framework. Future works can also be focused on extending the proposed framework to detect multiple coordinated attack types and classify them based on their locational impact. This can be achieved by hybridizing the SACGAN-GP with methods such as federated learning and edge computing paradigms.

V. Conclusion

To advance the cyber-resilience of modern power grids, this work investigates the FDIA detection problem by applying an attention-driven conditional GAN, which has exhibited state-of-the-art efficacy and adaptability in diverse research contexts. This research proposes a novel SACGAN-GP for detecting FDIAs in power systems. The proposed approach is entirely data-driven and does not require prior knowledge of the system model, making it adaptable and suitable for real-world deployment across different power grid configurations. The SACGAN-GP architecture captures complex spatial and contextual dependencies within measurement data by leveraging self-attention mechanisms. The model sustains high detection capability under adverse conditions such as unbalanced datasets. Experimental results on IEEE 14-bus and IEEE 118-bus test systems confirm that SACGAN-GP significantly performs better than other methods, achieving higher accuracy, precision, recall, F1-score, and ROC metrics. The higher ROC-AUC and PR-AUC values at different noise level shows the model stability even in presence of noise. These results confirm the model's ability to detect FDIAs effectively in the power grid. Furthermore, dimensionality reduction and visualization techniques such as t-SNE validate the model's ability to generate realistic and distinguishable attack patterns. The limitations of the proposed method, the possible mitigation strategies, and future direction are also discussed.

References

[1] G. Del Río Castro, M. C. González Fernández, and Á. Uruburu Colsa, "Unleashing the convergence amid digitalization and sustainability towards pursuing the Sustainable Development Goals (SDGs): A holistic review," Jan. 20, 2021, *Elsevier Ltd*. doi: 10.1016/j.jclepro.2020.122204.

[2] X. Wang, H. Zhu, X. Luo, and X. Guan, "Data-Driven-Based Detection and Localization Framework Against False Data Injection Attacks in DC Microgrids," *IEEE Internet Things J*, vol. 12, no. 17, pp. 36079–36093, 2025, doi: 10.1109/JIOT.2025.3579915.

[3] L. Xiao, H. Chen, S. Xu, Z. Lv, C. Wang, and Y. Xiao, "Reinforcement Learning-Based False Data Injection Attacks in Smart Grids," *IEEE Trans Industr Inform*, vol. 21, no. 4, pp. 3475–3484, 2025, doi: 10.1109/TII.2025.3528571.

[4] H. Feng, Y. Han, F. Si, and Q. Zhao, "Detection of False Data Injection Attacks in Cyber-Physical Power Systems: An Adaptive Adversarial Dual Autoencoder With Graph Representation Learning Approach," *IEEE Trans Instrum Meas*, vol. 73, pp. 1–11, 2024, doi: 10.1109/TIM.2023.3331398.

[5] J. Xie, A. Rahman, and W. Sun, "Bayesian GAN-Based False Data Injection Attack Detection in Active Distribution Grids with DERs," *IEEE Trans Smart Grid*, vol. 15, no. 3, pp. 3223–3234, 2024, doi: 10.1109/TSG.2023.3337340.

[6] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *ACM Transactions on Information and System Security*, May 2011. doi: 10.1145/1952982.1952995.

[7] G. Chaojun, P. Jirutitijaroen, and M. Motani, "Detecting False Data Injection Attacks in AC State Estimation," *IEEE Trans Smart Grid*, vol. 6, no. 5, pp. 2476–2483, 2015, doi: 10.1109/TSG.2015.2388545.

[8] G. Cheng, Y. Lin, J. Zhao, and J. Yan, "A Highly Discriminative Detector Against False Data Injection Attacks in AC State Estimation," *IEEE Trans Smart Grid*, vol. 13, no. 3, pp. 2318–2330, 2022, doi: 10.1109/TSG.2022.3141803.

[9] N. Li, J. Zhang, D. Ma, and J. Ding, "Enhancing Detection of False Data Injection Attacks in Smart Grid Using Spectral Graph Neural Network," *IEEE Trans Industr Inform*, vol. 21, no. 6, pp. 4543–4553, 2025, doi: 10.1109/TII.2025.3545044.

[10] S. H. Mohammed *et al.*, *Dual-hybrid intrusion detection system to detect False Data Injection in smart grids*, vol. 20, no. 1 January. 2025. doi: 10.1371/journal.pone.0316536.

[11] A. Takiddin, M. Ismail, R. Atat, K. R. Davis, and E. Serpedin, "Robust Graph Autoencoder-Based Detection of False Data Injection Attacks Against Data Poisoning in Smart Grids," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 3, pp. 1287–1301, 2024, doi: 10.1109/TAI.2023.3286831.

[12] Y. Wang, Y. Zhou, J. Ma, and Q. jin, "A locational false data injection attack detection method in smart grid based on adversarial variational autoencoders [Formula presented]," *Appl Soft Comput*, vol. 151, no. July 2023, p. 111169, 2024, doi: 10.1016/j.asoc.2023.111169.

[13] K. Li, F. Li, B. Wang, and M. Shan, "False data injection attack sample generation using an adversarial attention-diffusion model in smart grids," *AIMS Energy*, vol. 12, no. 6, pp. 1271–1293, 2024, doi: 10.3934/ENERGY.2024058.

[14] A. Shees, M. Tariq, and A. I. Sarwat, "Cybersecurity in Smart Grids: Detecting False Data Injection Attacks Utilizing Supervised Machine Learning Techniques," *Energies (Basel)*, vol. 17, no. 23, 2024, doi: 10.3390/en17235870.

[15] Y. Lin and J. Wang, "Probabilistic Deep Autoencoder for Power System Measurement Outlier Detection and Reconstruction," *IEEE Trans Smart Grid*, vol. 11, no. 2, pp. 1796–1798, 2020, doi: 10.1109/TSG.2019.2937043.

[16] S. A. Foroutan and F. R. Salmasi, "Detection of false data injection attacks against state estimation in smart grids based on a mixture Gaussian distribution learning method," *IET Cyber-Physical Systems: Theory & Applications*, vol. 2, no. 4, pp. 161–171, 2017, doi: 10.1049/iet-cps.2017.0013.

[17] 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 Trans Neural Netw Learn Syst*, vol. 27, no. 8, pp. 1773–1786, 2016, doi: 10.1109/TNNLS.2015.2404803.

[18] J. J. Q. Yu, Y. Hou, and V. O. K. Li, "Online False Data Injection Attack Detection with Wavelet Transform and Deep Neural Networks," *IEEE Trans Industr Inform*, vol. 14, no. 7, pp. 3271–3280, 2018, doi: 10.1109/TII.2018.2825243.

[19] G. Zhang, J. Li, O. Bamisile, D. Cai, W. Hu, and Q. Huang, "Spatio-Temporal Correlation-Based False Data Injection Attack Detection Using Deep Convolutional Neural Network," *IEEE Trans Smart Grid*, vol. 13, no. 1, pp. 750–761, 2022, doi: 10.1109/TSG.2021.3109628.

[20] S. Liu, C. Wu, and H. Zhu, "Topology-Aware Graph Neural Networks for Learning Feasible and Adaptive AC-OPF Solutions," *IEEE Transactions on Power Systems*, vol. 38, no.

6, pp. 5660–5670, 2023, doi: 10.1109/TPWRS.2022.3230555.
[21] Y. Li, X. Wei, Y. Li, Z. Dong, and M. Shahidehpour, "Detection of False Data Injection Attacks in Smart Grid: A Secure Federated Deep Learning Approach," *IEEE Trans Smart Grid*, vol. 13, no. 6, pp. 4862–4872, 2022, doi: 10.1109/TSG.2022.3204796.
[22] M. Abdel-Basset, N. Moustafa, and H. Hawash, "Privacy-Preserved Generative Network for Trustworthy Anomaly Detection in Smart Grids: A Federated Semisupervised Approach," *IEEE Trans Industr Inform*, vol. 19, no. 1, pp. 995–1005, 2023, doi: 10.1109/TII.2022.3165869.
[23] M. R. Uddin, R. Rahman, and D. C. Nguyen, "False Data Injection Attack Detection in Edge-based Smart Metering Networks with Federated Learning," 2024, doi: 10.1109/CCNC54725.2025.10976228.
[24] Y. Zhang, J. Wang, and B. Chen, "Detecting False Data Injection Attacks in Smart Grids: A Semi-Supervised Deep Learning Approach," *IEEE Trans Smart Grid*, vol. 12, no. 1, pp. 623–634, 2021, doi: 10.1109/TSG.2020.3010510.
[25] N. Costilla-Enriquez and Y. Weng, "Attack Power System State Estimation by Implicitly Learning the Underlying Models," *IEEE Trans Smart Grid*, vol. 14, no. 1, pp. 649–662, 2023, doi: 10.1109/TSG.2022.3197770.
[26] S. Aldhaheri, "Aldhaheri, Sahar, and Abeer Alhuzali. 'SGAN-IDS: Self-attention-based generative adversarial network against intrusion detection systems.' Sensors 23.18 (2023): 7796.," 2023, doi: 10.3390/s23187796
[27] G. N. Ericsson, "Cyber security and power system communication-Essential parts of a smart grid infrastructure," *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1501-1507, Jul. 2010, doi:10.1109/TPWR D.2010.2046654.
[28] P. Kundur et al., "Definition and classification of power system stability," IEEE Transactions on Power Systems, vol. 19, no. 3, pp. 1387–1401, 2004, doi**:** 10.1109/TPWRS.2004.825981.
[29] T. Huang, D. Wu, and M. Ilić, "Cyber-resilient automatic generation control for systems of AC microgrids," IEEE Transactions on Smart Grid,vol.15, no.1, pp. 886-898, Jan.2023, doi:10.1109/TSG.2023.3272632.

Murugesan S. received the B.E. degree from Park College of Engineering and Technology, Coimbatore, India (Anna University) in 2007, and the M.E. degree in control systems from the P.S.G. College of Technology, Coimbatore, India (Anna University), in 2011. He has around ten years of teaching experience at various engineering institutions. He is working toward a PhD in power system security, including cyber-attack detection in the smart grid.

Ram Jethmalani C. H. received the B.E. degree in electrical and electronics engineering from PSNA College of Engineering and Technology, Dindigul, India, in 2008, the M.E. degree in power systems from Adhiyamaan College of Engineering, Hosur, India, in 2011, and the Ph.D. degree from the National Institute of Technology, Tiruchirappalli, in 2018. Since 2021, he has been an Assistant Professor with the Department of Electrical and Electronics Engineering, National Institute of Technology Puducherry, Karaikal, India. His technical interests include power system operation, economics and scheduling, and soft computing applications to power systems.

Navin Sam K. received the B.E. degree from the Dr. Sivanthi Aditanar College of Engineering, Tiruchendur, India (Anna University), in 2009, the M.E. degree in power electronics and drives from the A.C. Government College of Engineering and Technology, Karaikudi, India (Anna University), in 2011, and the Ph.D. degree from the National Institute of Technology, Tiruchirappalli, in 2017. Since 2018, he has been an Assistant Professor with the Department of Electrical and Electronics Engineering, National Institute of Technology Puducherry, Karaikal, India. His research focuses on renewable energy electric conversion systems.

Venkadesan Arunachalam received the B.Tech. degree in electrical and electronics engineering, the M.Tech. degree in electrical drives and control, and the Ph.D. degree in AI techniques applied to power electronics and drives from Pondicherry Engineering College, Pondicherry University, Puducherry, India, in 2007, 2009, and 2014, respectively. He is a Professor with the Department of Electrical and Electronics Engineering, National Institute of Technology Puducherry, Karaikal. He has around ten years of total teaching experience. His interests include electric drives control and artificial intelligence techniques.

Sadheesh Kumar S. J. received the B.Tech. degree in electrical and electronics engineering, the M.Tech. degree in Power Electronics from Pondicherry University, Puducherry, India, in 2009 and 2014, respectively. Pursuing the Ph.D. degree in AI techniques applied to Power system forecasting from National Institute of Technology Puducherry, Karaikal, Puducherry, India, He is currently a Senior Technical Assistant with the Department of Electrical and Electronics Engineering, National Institute of Technology Puducherry, Karaikal. He has around 15 years of expertise in electrical engineering. His area of interest includes Time series forecasting and renewable energy systems.