Securing Wide-area Damping Controller Against Cyber Attacks Using Semi-supervised Generative Adversarial Network and Support Vector Machine-based Synthetic Minority Oversampling Technique

Abhishek Saini, *Student Member*, *IEEE*, and Pratyasa Bhui, *Member*, *IEEE*

***Abstract*—Wide-area measurement systems enable the transmission of measurement and control signals for wide-area damping controllers (WADCs) in smart grids. However, the vulnerability of the communication network makes the WADC susceptible to malicious cyber attacks, such as false data injection (FDI) attack and denial of service (DoS) attack. Researchers develope numerous supervised machine-learning and model-based solutions for attack detection. However, the partially labeled attack data, skewed class distributions, and the need for precise mathematical models present significant challenges for real-world attack detection. This paper introduces the cyber attack-resilient wide-area damping controller (CyResWadc) system framework to address these challenges. The proposed framework leverages semi-supervised generative adversarial network (SSGAN) model to handle partially labeled attack data. It utilizes the support vector machine-based synthetic minority oversampling technique (SVM-SMOT) for data oversampling to manage skewed class distributions. Furthermore, probing signals are used to stimulate the power system, facilitating the generation of synthetic attack scenarios under different operational conditions. If any attack is detected, an alternate pair of measurement and control signals is used for attack mitigation. The performance is validated on a developed hardware-in-the-loop (HIL) cyber-physical testbed built using the open parallel architecture laboratory-real time (OPAL-RT) simulator, industry-grade hardware, Network Simulator 3 (NS-3), and open platform for data collection (OpenPDC).**

***Index Terms*—Cyber attack, cyber security, false data injection (FDI) attack, attack detection, semi-supervised generative adversarial network (SSGAN), wide-area damping controller (WADC), support vector machine (SVM).**

I. Introduction

In the past, low frequency oscillations (LFOs) have been responsible for various blackouts and unscheduled tripping of generators, such as the blackout in the Western Interconnection (WSCC) in 1996, the blackouts in Denmark-Sweden, USA-Canada, and Italy in 2003, and the blackout in Chile in 2011 [1]. Unwanted generator tripping happens in the Indian grid due to sustained LFO [2]. Over the past two decades, various wide-area damping controllers (WADCs) have been proposed for inter-area oscillation modes (0.2-0.8 Hz) and implemented in several smart grids, e. g., WSCC and China Southern Power Grid [3]. However, the wide-area signals for WADC are transmitted through multiple devices, ranging from phasor measurement units (PMUs) to local phasor data concentrators (PDCs) and from local PDCs to super PDCs. The integration of cyber and physical layers makes it vulnerable to attacks that can manipulate control operations and destabilize the system. Notable incidents, such as the Stuxnet worm in 2010 [4] and the Venezuela cyber attack in 2019 [5], have demonstrated how cyber vulnerabilities in smart grids can be exploited.

Existing studies on methods of FDI attack detection in smart grids can be broadly classified into data-driven and model-based methods. Model-based methods are further divided into static state estimation (SSE) and dynamic state estimation (DSE) methods. The SSE methods are implemented by researchers. The SSE method proposed in [6] utilizes a non-linear filtering approach based on cyber-physical information derived from Kirchhoff's laws to detect FDI attacks on state estimation (SE). In [6], data from adjacent nodes are needed and can be implemented locally or distributively. The SSE method proposed in [7], which is called the reactance perturbation strategy, is used for FDI attack detection on power system SE for improving the security of SE without increasing the operational cost. However, the SSE methods in [6] and [7] are valid for the steady state, whereas WADC is designed to operate in a dynamic state as well. To address this limitation, researchers explore the DSE methods. For instance, a joint attack detection and compensation

Manuscript received: June 30, 2024; revised: November 8, 2024; accepted: December 17, 2024. Date of CrossCheck: January 26, 2025. Date of online publication: June 19, 2025.

This work was partially supported by Science and Engineering Research Board (No. CRG/2021/003827/EEC).

A. Saini and P. Bhui (corresponding author) are with the Department of Electrical, Electronics, and Communication Engineering, Indian Institute of Technology Dharwad, Dharwad, India (e-mail: abhishek.saini@iitdh.ac.in; pbhui@iitdh.ac.in).

DOI: 10.35833/MPCE.2024.000697

method is proposed by using the Kalman filter for automatic generation control (AGC) system [8]. Reference [9] proposes a robust parallel DSE method, which utilizes an extended Kalman filter and graphical processing unit for FDI attack detection by using a trusted set of PMU measurements [9]. However, the DSE methods in [8] and [9] require the following information: ① system models; ② system parameters; and ③ threshold selection. However, the DSE methods may suffer from divergence and scalability issues.

In contrast to the model-based methods, data-driven methods offer an alternative approach for FDI attack detection that does not require knowledge of the system parameters or system models in real time. Various supervised learning (SL)-based methods are implemented by researchers for this purpose. For example, the invertible automatic encoder (IEA) combined with a long short-term memory (LSTM) -based classifier to detect FDI attacks is proposed in [10]. A distributed support vector machine (SVM) model, which uses data preprocessed by principal component analysis (PCA), is trained for FDI attack detection in power systems [11]. A framework of bad data detection and convolutional neural network (BDD-CNN) is proposed in [12] to detect FDI attacks and identify the exact location of injected data in power systems. Unfortunately, the above-mentioned methods rely on a large feature dataset with complete labels, which presents a major challenge due to the scarcity of labeled data, especially in cases of cyber attacks. The task of labeling unlabeled data is exceedingly difficult, requiring expert domain knowledge of the system. As a result, SL-based methods may not be effective for attack detection in real power system [13]. However, semi-supervised learning (SSL) -based methods offer a better alternative in such cases. For example, the unobservable FDI attacks are detected in the distribution network using an adversarial auto-encoder (AAE)-based algorithm [14]. In addition, a graph-based semi-supervised learning (GBSSL) model is proposed for attack detection and classification in photovoltaic arrays in [15].

To enhance the cyber security of WADC, researchers propose various strategies, including the design of attack-resilient controllers and the development of data-driven attack detection techniques that can operate without modifying the existing controller architecture [16], [17]. A resilient adaptive WADC framework is proposed in [18] for FDI attack detection and correction on the measurement side, utilizing linear state estimation (LSE) as a data preprocessor. LSE can estimate the bus voltage and line current phasors and detect any bad data in the measurement. However, LSE relies on accurate system parameters for SE, which limits its adaptability for large and deregulated power systems with multiple system operators. Furthermore, a secure network predictive control (SNPC) -based resilient WADC is proposed to defend against the deception attacks in [19]. However, the method proposed in [19] is designed to detect attacks only during the communication process between the sensors and the control center. The data encryption standard (DES) algorithm is considered unsafe due to its short key length of 64 bits. Reference [20] shows that DES algorithm can be compromised through brute force attacks. A wide-area robust sliding mode controller (WARSMC) is proposed to defend against FDI attacks and mitigate inter-area oscillations in [21]. However, WARSMC depends on the redundancy of the measurement system, and the time delay is not considered for controller design. The time delay deteriorates the control performance and can destabilize the system [22]. A defense strategy based on simultaneous input and SE is proposed for the WADC system against modal resonance-oriented cyber-attack (MR-OCA) in [23]. A multiple-controller switching-based resilient wide-area damping controller (MCS-RWADC) is proposed for detecting strong cyber attacks in [24]. However, MCS-RWADC requires the deployment of multiple modules, which significantly increases the complexity of communication channels and escalates the implementation costs.

Researchers also implement the game theory framework to defend WADC against cyber attacks. An optimal cyber-layer defense strategy is developed using the Markov game to defend WADC against cyber attacks in [25]. However, the game theory framework assumes that the defender has perfect knowledge of the attacker strategy, which causes serious resource waste due to the high over-defense rate. Furthermore, due to the challenges of obtaining accurate system models in real time, data-driven approaches of PMU are proposed in studies for cyber attack detection [16], [26]. The *K*-nearest neighbor and decision tree-based methods are proposed for cyber attack detection in the WADC system in [16]. A wide-area measurement system (WAMS)-based high-voltage direct current (HVDC) damping framework is proposed in [26], which uses attack shuffle convolutional neural network and continuous wavelet transform (CWT) for data integrity attack detection. However, both [16] and [26] rely on large labeled datasets for model training. An online robust principal component analysis (RPCA) method is proposed in [17] for a malicious corruption-resilient WADC system. However, the performance of RPCA method highly depends on the accuracy of offline-generated subspace library.

Current research on the cyber security of WADC predominantly concentrates on cyber attack detection through attack-resilient WADC, as demonstrated by the methods proposed in [18], [23], and [24]. However, these methods are constrained either by their reliance on the accuracy of the system model and parameters or by the need for redundant measurement signals. Additionally, grid operators may be unwilling to modify the WADC due to operational and economic constraints. Additionally, SL-based methods, as proposed in [16] and [26], enable effective cyber attack detection by using labeled data without altering the controller architecture. Despite these advancements, SL-based methods are limited by the difficulty in obtaining accurate labels for attack scenarios. There is a noticeable lack of research exploring the potential of utilizing SSL-based methods with limited labeled data. The SSL-based methods operate without requiring changes to the existing controller architecture or real-time knowledge of system parameters and models. The SSL-based method has the potential to improve the accuracy of cyber attack detection by learning from both labeled and unlabeled data, and this study aims to fill this gap.

In this paper, a cyber attack-resilient wide-area damping

controller (CyResWadc) system framework is proposed for cyber attack detection and mitigation. The proposed framework is capable of dectecting cyber attack at both attack surfaces (measurement signal and control signal). To the best of the authors' knowledge, this paper innovatively employs an SSL-based method for detecting cyber attacks. Upon detecting a cyber attack, the mitigation module strategically switches the WADC input and output signals to minimize the impacts of cyber attacks. The main contributions of this paper can be summarized as follows.

1) We propose a CyResWadc system framework for cyber attack detection and mitigation in cyber-physical power system. The proposed framework incorporates a semi-supervised generative adversarial network (SSGAN) model to handle partially labeled data.

2) Support vector machine-based synthetic minority oversampling technique (SVM-SMOT) addresses the skewed class distributions for the WADC system. While limited attack data are expected to be available in the future, the lack of adequate labeled attack data at present can be solved by utilizing probing signals to generate synthetic attack scenarios under various operational conditions.

3) To capture the system dynamics, physics-aware features are utilized, such as the damping torque coefficient (DTC) and mode shape (MS).

4) Simulation test studies validate the proposed framework on 4-machine two-area system and IEEE 16-machine 68-bus systems, where the SSGAN+SVM-SMOT model is trained with unlabeled data and a small amount of labeled data.

5) A hardware-in-the-loop (HIL) cyber-physical testbed is developed utilizing the open parallel architecture laboratory-real time (OPAL-RT) simulator, industry-grade hardware, Network Simulator 3 (NS-3), and open platform for data collection (OpenPDC). The proposed framework is validated via the HIL testbed by replicating realistic cyber attack scenarios.

The rest of this paper is organized as follows. The system representation and problem formulation are briefly explained in Section II. Section III shows the proposed framework against FDI attack. The SSGAN model for attack detection is introduced in IV. The physics-aware features for FDI attack detection are defined in Section V, and the experimental setup and empirical evaluation are given in Section VI. Eventually, the conclusion is given in Section VII.

II. System Representation and Problem Formulation

A. System Respresentation

The wide-area power system exhibits inherent non-linearity, which is mathematically represented by complex non-linear differential-algebraic equations (DAEs). The non-linear wide-area power system under FDI attacks at the sensor and actuator locations is represented as:

$$\begin{cases} \boldsymbol{x}_t = \boldsymbol{f}(\boldsymbol{x}_{t-1}, \boldsymbol{u}_{t-1}, \boldsymbol{e}'_{t-1}, \boldsymbol{w}_{t-1}) \\ \boldsymbol{y}_t = \boldsymbol{h}(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{r}'_t) + \boldsymbol{v}_t \end{cases} \tag{1}$$

where subscript t is the time index; $\boldsymbol{x}_t$ is the n-dimensional state variable vector; $\boldsymbol{w}_{t-1}$ is the process noise vector; $\boldsymbol{y}_t$ and $\boldsymbol{v}_t$ are the l-dimensional vectors of measurement output signal and measurement noise, respectively; $\boldsymbol{u}_t$ is the q-dimensional vector of the wide-area control signal fed to the exciter; $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ are the functions that represent the state transition and measurement, respectively; $\boldsymbol{e}'_{t-1}$ is the malicious signal that is injected at the actuator; and $\boldsymbol{r}'_t$ is the attack signal that is employed at the measured data.

Reference [27] provides a detailed modeling of FDI attacks and their impacts on the WADC system, including a comparative analysis of pulse, sinusoidal, sawtooth, triangular, and random attack types. The magnitude of oscillation reaches its maximum for the sinusoidal attack when the frequency of the injected signal is similar to the inter-area mode frequency.

B. Problem Formulation

In the wide-area power system, an attacker can gain access to and compromise the sensors and actuator signals. To address this issue, this paper formulates the cyber attack detection problem as a semi-supervised classification problem. Specifically, the problem assumes the existence of a partially labeled data. We set S as the total data set, which includes a small amount of labeled attack/event data set $\mathcal{L}=\{(\boldsymbol{x}_l, y_l)$, $\boldsymbol{x}_l \sim P_d$, $y_l \in [1, C]$, $l \in [1, L]$ and a large amount of unlabeled data set $\acute{U}=\{\boldsymbol{x}_u\}$, $\boldsymbol{x}_u \sim P_d$, $u \in [1, U]$, where $\boldsymbol{x}_l$ is the labeled data instance; $\boldsymbol{x}_u$ is the unlabeled data instance; C is the number of classes; P_d is the actual distribution of data $\boldsymbol{x}_l$ and $\boldsymbol{x}_u$; y_l is the label of $\boldsymbol{x}_l$; L is the total number of labeled data instances; and U is the size of the unlabeled data instances. The primary objective is to obtain a robust semi-supervised classifier by using the limited labeled data. In addition to cyber attacks, this paper considers various power system events, such as various types of faults, line outages, and generator outages, which cause LFOs.

III. Proposed Framework Against FDI Attack

The schematic diagram in Fig. 1 illustrates the proposed framework for cyber attack detection and mitigation, where EXC is short for exciter; AVR is short for automatic voltage regulator; PSS is short for power system stabilizer; SVC is short for static var compensator; TCSC is short for thyristor controlled series capacitor; UPFC is short for unified power flow controller. Additionally, the control centre in Fig. 1 exhibits two switches denoted as T_1 and T_2. The proposed framework comprises two key components, which are cyber attack detection module and cyber attack mitigation module. The cyber attack detection module utilizes the deep learning (DL)-based model that is trained with physics-aware features to classify attacks and events. These physics-aware features are calculated using raw PMU data, thereby providing a comprehensive representation of the system dynamics. The raw PMU data vector $\boldsymbol{D}_{dataset}$ comprises frequency f, positive-sequence voltage magnitude V, angular speed ω, active power P, and reactive power Q.

$$\boldsymbol{D}_{dataset} = [f, V, \omega, P, Q] \tag{2}$$

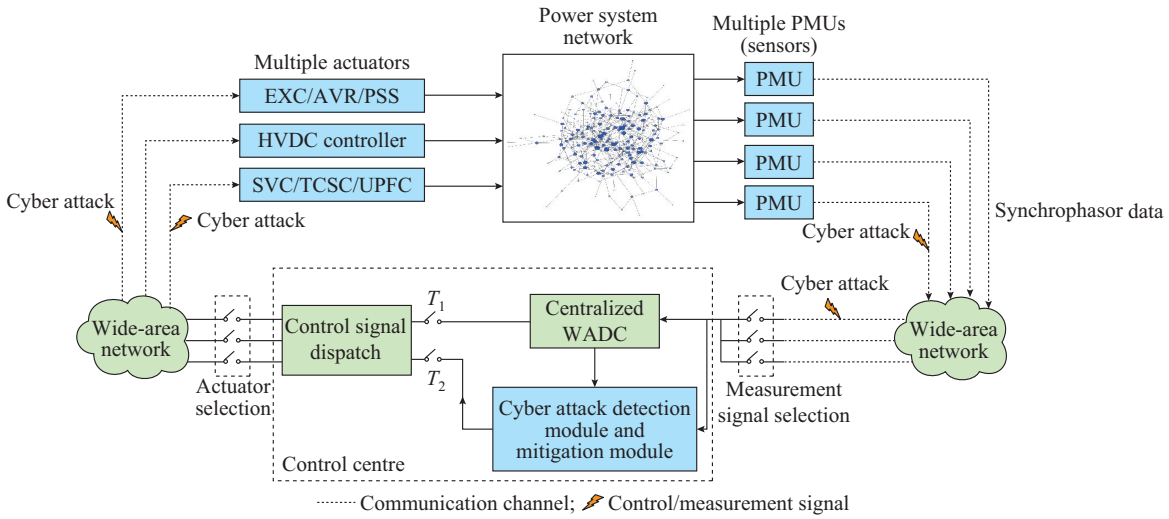

Fig. 1. Proposed framework for cyber attack detection and mitigation.

Figure 2 outlines sequential steps of training process of the proposed framework. The training process leverages physics-aware features to classify power system events and attacks. The physics-aware features are scaled for training process as illustrated in Fig. 2.

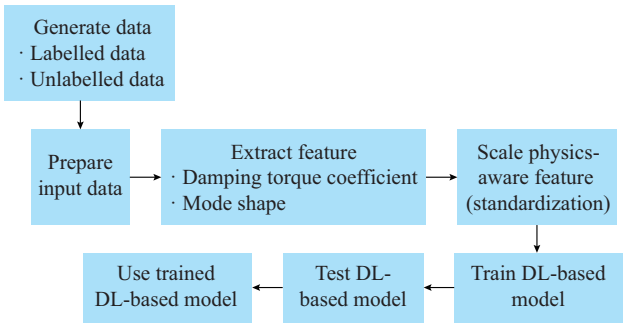

Fig. 2. Block diagram of sequential steps of training process of proposed framework.

During normal as well as event situations of the power system, T_1 remains connected. However, upon detecting a cyber attack by the proposed framework at the measurement/control signal, the switch configuration is altered from T_1 to T_2. Specifically, the cyber attack detection module is utilized to detect any cyber attacks on the measurement/control signal. Once a cyber attack is detected, the cyber attack mitigation module selects an alternative pair of measurement and control signals and adjusts the parameters of WADC system accordingly [28].

IV SSGAN Model for Attack Detection

A. Preliminary of GAN Model

The GAN model, which is a widely popular and efficient DL-based model, has been applied to solve diverse issues in various research fields such as video synthesis [29]. The GAN model consists of two competing neural networks: the generator and the discriminator. The generator is trained to produce novel synthetic samples from random noise input. In contrast, the discriminator classifies samples as either authentic (from the real data source) or synthetic (generated by the generator) [30]. The GAN model is trained through the zero-sum game theory between the generator and discriminator, where parameters are updated to reach Nash equilibrium.

To explain the training of GAN model, we set vector $\boldsymbol{x}$ and $\boldsymbol{k}$ to be the actual input data with distribution $p_{data}(\boldsymbol{x})$ and the latent/noise prior space with distribution $q(\boldsymbol{k})$, respectively. Furthermore, we set G and D to be differentiable functions, where G is the generator with input $\boldsymbol{k}$; and D is the discriminator with input $\boldsymbol{x}$. The output of D is mapped to interval $[0,1]$. The min-max optimization objective function $\mathcal{V}(D,G)$ is given as:

$$\min_G \max_D \mathcal{V}(D,G)=E_{\boldsymbol{x}\sim p_{data}(\boldsymbol{x})}[\log D(\boldsymbol{x})]+E_{\boldsymbol{k}\sim q(\boldsymbol{k})}[\log(1-D(G(\boldsymbol{k})))] \tag{3}$$

where $E[\cdot]$ is the expectation function.

The optimization objective function in (3) is solved using neural networks and the application of back-propagation via gradient ascent and gradient descent. Given a batch $\{\boldsymbol{x}_i,\boldsymbol{k}_i\}_{i=1}^n$ of training data and samples from the latent space, the optimization objective function in (3) can be reformulated into the optimization of two separate cost functions $\mathcal{C}(\cdot)$, i.e., one for the discriminator D and another for the generator G, which can be expressed as:

$$\begin{cases}\mathcal{C}(D)=-\dfrac{1}{2}\left(\displaystyle\sum_{i=1}^{n}\log D(\boldsymbol{x}_i)+\log(1-D(G(\boldsymbol{k}_i)))\right)\\ \mathcal{C}(G)=-\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\log D(G(\boldsymbol{k}_i))\end{cases} \tag{4}$$

where $\mathcal{C}(D)$ is the discriminator cost function, which aims to maximize the probability of correctly classifying real data as real and generated data as fake; and $\mathcal{C}(G)$ is the generator cost function, which aims to generate data that are classified as real by the discriminator.

B. Solution to Skewed Class Distribution

Unbalanced classification occurs when data are distributed unevenly among various classes, resulting in inconsistencies in the available data for each class. Training models with un-

balanced data, where one class has more examples than another (more event data than attack data), can lead to a biased model that performs poorly on the minority class (attack class) due to the skewed class distribution [31]. To address this issue, class balancing techniques are used before training the SSL-based method. The application of the class balancing techniques leads to an improved performance of the SSL-based method. The class balancing technique involves generating synthetic data for the minority class using oversampling techniques. Researchers propose numerous oversampling techniques, such as synthetic minority based on the probabilistic distribution (SyMProD) technique [31], synthetic minority oversampling technique (SMOT) [32], mahalanobis distance-based oversampling (MDO) technique [33], oversampling using propensity score (OUPS) technique [34], and SVM-based SMOT [35]. Moreover, a GAN variant, known as conditional GAN (CGAN), is also used to generate data for the attack class.

C. Integration of Oversampling Techniques with SSGAN Model

In the proposed framework, the issue of unbalanced data is resolved by integrating the oversampling techniques with SSGAN model, as shown in Fig. 3. Limited minority data points are used to generate new samples with a similar distribution. The synthetic data are then combined with real data set for SSGAN model training. The SSL-based method utilizes a combination of labeled and unlabeled data during the training data process. The SSGAN model consists of a generator and a discriminator, which is trained in supervised and unsupervised training modes. In unsupervised training mode, the discriminator distinguishes between real and fake samples. While the supervised training mode focuses on the ability of the discriminator to classify the samples into their respective class labels.

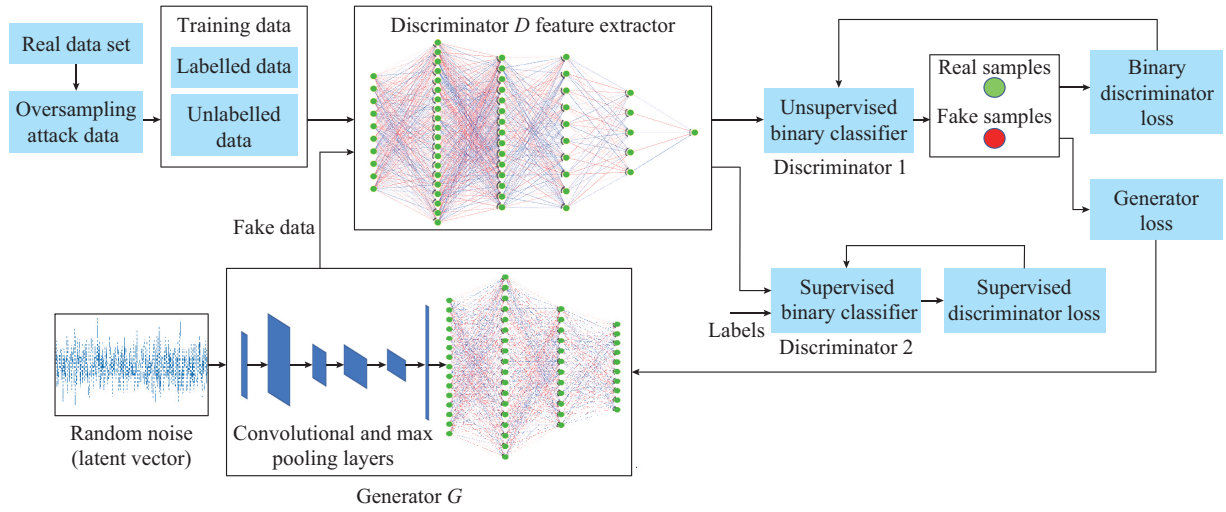

Fig. 3. Architecture of integrating oversampling techniques with SSGAN model.

As shown in Fig. 3, the discriminator in supervised and unsupervised training modes is arranged in a stacked manner. We denote the discriminator in supervised training mode as the supervised discriminator (discriminator 1) and the discriminator in unsupervised training mode as the unsupervised discriminator (discriminator 2). The supervised discriminator is trained to classify the samples into M classes by assigning the label to each sample $\boldsymbol{x}$ to give an M-dimensional vector of logits $[l_1, l_2, ..., l_M]$, which can be converted to class probabilities by applying the activation function. The unsupervised discriminator uses the same neural network as the supervised one, with the weights of the layers being reused. It is stacked on top of the output layer of the supervised discriminator just before the activation function. A normalized sum of exponential outputs is used as a custom activation function to predict the authenticity of real and fake samples, which can be given as:

$$\begin{cases} D(\boldsymbol{x}) = \dfrac{\mathcal{K}(\boldsymbol{x})}{\mathcal{K}(\boldsymbol{x})+1} \\ \mathcal{K}(\boldsymbol{x}) = \displaystyle\sum_{m=1}^{M} \exp\left(l_m(\boldsymbol{x})\right) \end{cases} \tag{5}$$

The SSGAN model involves augmenting the original dataset with extra samples generated by the generator, which increases the number of class labels from M to $M+1$, where $M+1$ is the newly added class label for the generated samples. This modification compels the discriminator to identify which of the $M+1$ classes the sample belongs to [36].

We set $p_{model}(y=(M+1)|\boldsymbol{x})$ to be the probability that the given input $\boldsymbol{x}$ is fake. The overall loss function of the stacked discriminator for training the classifier can be given as:

$$\begin{cases} J = -E_{\boldsymbol{x},y\sim p_{data}(\boldsymbol{x},y)}[\log p_{model}(y|\boldsymbol{x})] - \\ \quad E_{\boldsymbol{x}\sim G}[\log p_{model}(y=(M+1)|\boldsymbol{x})] = J_S + J_U \\ J_S = -E_{\boldsymbol{x},y\sim p_{data}(\boldsymbol{x},y)}[\log p_{model}(y|\boldsymbol{x}, y<M+1)] \\ J_U = -E_{\boldsymbol{x}\sim p_{data}(\boldsymbol{x})}[\log(1-p_{model}(y=(M+1)|\boldsymbol{x}))] + \\ \quad E_{\boldsymbol{x}\sim G}[\log p_{model}(y=(M+1)|\boldsymbol{x})] \end{cases} \tag{6}$$

where J is the overall loss; and J_S and J_U are the supervised loss and unsupervised loss, respectively.

In addition, J_S is the cross-entropy loss, which is incurred from the predicted distribution over M classes. Meanwhile, J_U consists of two separate terms. The first term in J_U corre-

sponds to the loss that arises from classifying real inputs, while the second term is associated with the loss incurred from classifying generated samples as fake. To achieve the optimal solution, it is necessary to minimize both J_S and J_U to acquire $\exp(l_m(\boldsymbol{x}))=h(\boldsymbol{x})p(y=i,\boldsymbol{x}), \forall i<M+1$ and $\exp(l_{M+1}(\boldsymbol{x}))=h(\boldsymbol{x})p_G(\boldsymbol{x})$ for any undetermined scaling function $h(\boldsymbol{x})$, where $p(y=i,\boldsymbol{x})$ is the joint probability that a sample x belongs to class i; and $p_G(\boldsymbol{x})$ is the probability that the same sample x is produced by the generators (the fake-data distribution).

To enhance the performance of the SSGAN model, various oversampling techniques are integrated into it. The modified models include SSGAN+SyMProD model, SSGAN+OUPS model, SSGAN+MDO model, SSGAN+SVM-SMOT model, and CGAN+SSGAN model. Additionally, a comprehensive comparison of the modified models with other state-of-the-art (SOTA) methods is conducted. Three SSL-based methods (namely semi-supervised boosting (SemiB) [37], label propagation (LP) [38], and semi-supervised SVM (S3VM) methods [39]) are considered in this study for comparison. Furthermore, to ensure a fair comparison, the data balancing methods are also integrated with the base learners, which are LP+SyMProD model, SemiB+SyMProD model, LP+OUPS model, and SemiB+OUPS model.

V Physics-aware Features for FDI Attack Detection

A. MS

The MS is a complex vector that describes the phase angle and magnitude of the dominant LFO mode observed in the PMU measurements. During system disturbances, MS provides information about the most dominant LFO mode and which generator is contributing the most to it. MS for a particular inter-area mode can be obtained conventionally from eigenvalue analysis based on the system model. It is unique irrespective of the type and duration of faults. However, it is affected to some extent by operational condition changes caused by line outages, generation outages, etc. Furthermore, when the cyber attack causes oscillation in the WADC system with a frequency different from the electromechanical mode, the MS may exhibit a peak magnitude. Additionally, at the bus where the WADC is installed, the phase angle will be leading. Therefore, the MS obtained for cyber attacks is different from those obtained for other power system events. Hence, this can be a useful feature for the cyber attack detections in WADC system. However, if the injected signal frequency is the same as the electromechanical mode frequency, resonance is likely to be induced. Thus, the MS may bear resemblance to the one derived from eigenvalue analysis. In such a case, the MS may not be useful for cyber attack detection, but this limitation is mitigated by utilizing DTC. The MS can be tracked in real time through a spectral analysis technique that employs time-synchronized PMU data as input, specifically generator speed measurements [40]. The dominant LFO mode for MS computation is filtered out by wavelet synchrosqueezing transform [41]. The phase angle and magnitude are computed using cross-power spectral density (CPSD) and power spectral density (PSD) techniques, respectively. The relationship defined by (7) can be utilized to extract phase angle information among the generators [40]. By simplifying (7), (8) is obtained to get phase angle information. The phase angle of all generators is calculated with respect to a high mode observability reference generator signal at the frequency of dominant mode ω_a. It is worth noting that $\lim_{T\to\infty}\frac{1}{T}E[|\boldsymbol{Z}_a(\omega_a)|^2]$ converges to a constant, where $\boldsymbol{Z}_a(\omega_a)$ is the finite Fourier transform (FT) of the signal $z_b(t)$ at frequency ω_a; and T is the total time. In addition, $\boldsymbol{\mathcal{U}}_{a,b}$ and $\boldsymbol{\mathcal{U}}_{a,c}$ provide information about the b^{th} and c^{th} signals in the a^{th} mode in (7), respectively.

$$\boldsymbol{\Psi}_{bc}(\omega_a)\cong\boldsymbol{\mathcal{U}}_{a,b}\boldsymbol{\mathcal{U}}_{a,c}\lim_{T\to\infty}\frac{1}{T}E[|\boldsymbol{Z}_a(\omega_a)|^2] \tag{7}$$

$$\boldsymbol{\Psi}_{bc}(\omega_a)\cong\angle\boldsymbol{\mathcal{U}}_{a,c}-\angle\boldsymbol{\mathcal{U}}_{a,b} \tag{8}$$

where $\boldsymbol{\Psi}_{bc}(\omega_a)$ is the CPSD between two generator signals.

Now, we consider the special case ($b=c$), which corresponds to the PSD. Substituting $b=c$ into (7) yields (9). The PSD is computed for each generator speed signal. The magnitude of the dominant mode presented in the generator speed signal is determined by (9). The PSD of each signal is scaled by the square of the magnitude $|\boldsymbol{\mathcal{U}}_{a,b}|$. Therefore, the PSD serves as a direct measure of the observability of the mode at the generator.

$$\boldsymbol{\Psi}_{bb}(\omega_a)\cong|\boldsymbol{\mathcal{U}}_{a,b}|^2\lim_{T\to\infty}\frac{1}{T}E[|\boldsymbol{Z}_a(\omega_a)|^2] \tag{9}$$

The normalized estimated $|\boldsymbol{\mathcal{U}}_{a,b}|$ is computed as:

$$|\boldsymbol{\mathcal{U}}_{a,b}|=\sqrt{\frac{\boldsymbol{\Psi}_{bb}}{\boldsymbol{\Psi}_{nn}}} \tag{10}$$

where $\boldsymbol{\Psi}_{bb}$ is the PSD for the generator under consideration; and $\boldsymbol{\Psi}_{nn}$ is the PSD of the reference generator at the mode frequency being estimated.

B. Damping Torque Coefficient (DTC)

The LFOs are associated with the dissipation and transmission of oscillation energy flow in the network. In the electromechanical transients, the energy concept is used for the analysis of LFOs. There is a strong correlation between the energy consumption/production of equipment during an oscillation period and its DTC [42]. By leveraging the wide-area measurement system (WAMS) data at the control center, the oscillation energy flow is initially calculated, followed by the subsequent computation of the DTC based on the derived energy flow. The oscillation energy flow W_{oe} between node o and node e is calculated by (11) and (12) [42]. We set W_{oe}^{O} to be the oscillation component. In addition, W_{oe}^{D} consists of the monotonic varying component pertaining to energy production/consumption, which can be computed as (13).

$$W_{oy}=\int(P_{oy}\mathrm{d}\Delta\theta_o+Q_{oy}\mathrm{d}(\Delta\ln V_o))+\int(2\pi P_{oy}\mathrm{d}\Delta f_o\mathrm{d}t+\Delta Q_{oy}\mathrm{d}(\Delta\ln V_o)) \tag{11}$$

$$W_{oe}=W_{oe}^{O}+W_{oe}^{D} \tag{12}$$

$$\begin{cases} W_{oe}^{D} = \int (\Delta P_{oe} \mathrm{d}\Delta\theta_o + \Delta Q_{oe} \mathrm{d}(\Delta \ln V_o)) = \\ \qquad \int (2\pi \Delta P_{oe} \Delta f_o \mathrm{d}t + \Delta Q_{oe} \mathrm{d}(\Delta \ln V_o)) \\ \Delta P_{oe} = P_{oe} - P_{oe,s} \\ \Delta \ln V_o = \ln V_o - \ln V_{o,s} \\ \Delta f_o = f_o - f_0 \\ \Delta Q_{oe} = Q_{oe} - Q_{oe,s} \end{cases} \tag{13}$$

where subscripts o and e denote node o and node e, respectively; ΔP is the active power variation; ΔQ is the reactive power variation; Δf is the frequency deviation; subscript s denotes the steady-state value; and $\Delta\theta_o$ is the phase angle variation.

For V, P, and Q, either per-unit or actual value can be selected. This selection may lead to the calculation results of W_{oe}^{D} having different units. Additionally, the frequency deviation Δf_o is selected in units of Hz.

Furthermore, W_{oe}^{D} is composed of potential energy W_{pot}, kinetic energy W_{ken}, and generation/consumption energy W_{gen}.

$$W_{oe}^{D} = W_{ken} + W_{pot} + W_{gen} \tag{14}$$

For monotonically varying W_{gen}, the difference of $W_{pot} + W_{gen}$ between two local minima l and m is:

$$\begin{aligned} &(W_{pot,m} + W_{gen,m}) - (W_{pot,l} + W_{gen,l}) = (W_{oe,m}^{D} - W_{ken,m}) - \\ &(W_{oe,l}^{D} - W_{ken,l}) = W_{gen,m} - W_{gen,l} = W_{gen,ml} \end{aligned} \tag{15}$$

where $W_{gen,ml}$ is the energy dissipation/production from point t_l to point t_m; and $W_{ken,l}$ is the kinetic energy. If $W_{gen,ml} > 0$, the generator is consuming energy resulting in a positive damping effect. Conversely, if $W_{gen,ml} < 0$, the generator is adding energy to the network and has a negative damping effect. By incorporating $W_{gen,ml}$ and angular speed ω, DTC K_{DTC} can be given as:

$$K_{DTC} = \frac{W_{gen,ml}}{\omega_0 \int_{t_l}^{t_m} \omega^2 \mathrm{d}t} = \frac{(W_{oe,m}^{D} - W_{ken,m}) - (W_{oe,l}^{D} - W_{ken,l})}{\omega_0 \int_{t_l}^{t_m} \omega^2 \mathrm{d}t} \tag{16}$$

where ω_0 is the rated speed deviation.

VI. Experiment Setup and Empirical Evaluation

A. Description of Test System

The two-area test system in Kundur consists of four generators, as depicted in Fig. 4. The PDC receives the PMU measurement signals from different buses, which are then transmitted to the control center via the wide-area network. The WADC system takes the line power deviation of lines 7 and 8 as input and transmits the generated control signal to generator G4. The WADC system may be placed at the control center (centralized control) or near the actuator (decentralized). In the case of a centralized WADC system, both measurement and control signals are subject to cyber attacks, while in the decentralized WADC system, only measurement signals are targeted.

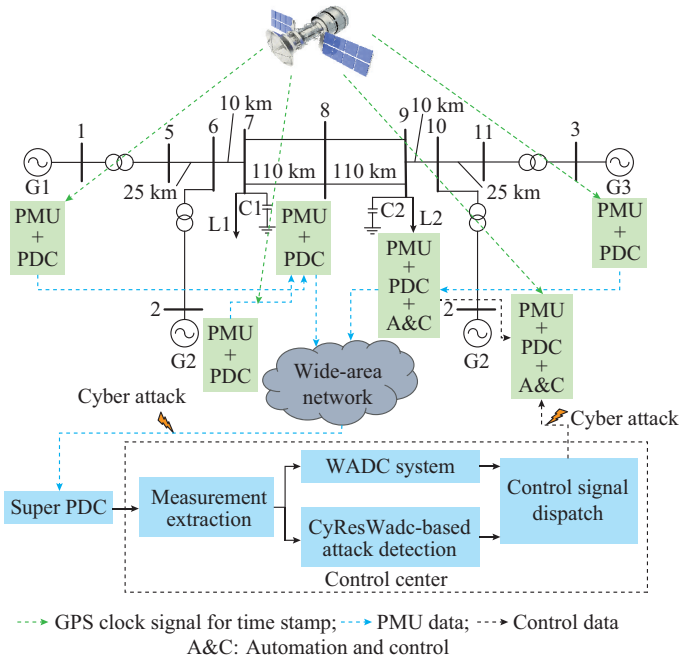

Fig. 4. Two-area test system.

B. WADC Design Based on Mixed H_2/H_∞

The non-linear power system is linearized around an operating point to obtain the state space model [43]. Then, the dominant inter-area mode is identified based on oscillation frequency and damping ratio. The WADC is designed using mixed H_2/H_∞ controller as it provides a balance between conflicting requirements such as improved transient response and frequency domain performance. The linear matrix inequality (LMI) method, implemented using MATLAB function *hinfmix*, provides a natural framework for formulating a multi-objective H_2/H_∞ output feedback control with regional pole placement [43].

C. Model Architecture

The architectures of modified generator and discriminator networks with learning parameters are presented in Tables I and II, respectively. The architecture describes the structure of the model, while the learning parameters provide information about the training process. The first layer in the generator network is the one-dimensional convolution (convention 1D) layer, which takes random noise as input. The LeakyReLU activation function and max pooling layers with stride are used subsequently. The flatten is used for changing data dimension into an array, then fully connected hidden layers are used for deeper architectures with ReLU and LeakyReLU activation functions in the generator, and the last layer uses sigmoid, which provides the required data format for the discriminator. In the case of the discriminator network, the last layer of the supervised discriminator uses a sigmoid activation function with Adam optimizer algorithm parameter set at learning rate $Lr = 0.002$ and $\beta = 0.5$. While unsupervised discriminator uses a custom activation function with optimizer setting $Lr = 0.0002$ and $\beta = 0.5$.

TABLE I
ARCHITECTURE OF GENERATOR NETWORK

Layer configuration details	Output dimension
Convolution1D (32 filter, size 3)+LeakyReLU ($\alpha=0.2$)	98, 32
Maxpooling1D (size 2, stride 2)	49, 32
Convolution1D (25 filter, size 2)+LeakyReLU ($\alpha=0.2$)	48, 25
MaxPooling1D (size 2, stride 2)	24, 25
Flatten	600
Dense+ReLU ($\alpha=0.2$)	15
Dense+ReLU ($\alpha=0.2$)	18
Dense+LeakyReLU ($\alpha=0.2$)	12
Dense+sigmoid	10

TABLE II
ARCHITECTURE OF DISCRIMINATOR NETWORK

Layer configuration details	Output dimension
Dense+LeakyReLU ($\alpha=0.2$)	20
Dense+LeakyReLU ($\alpha=0.2$)	15
Dense+LeakyReLU ($\alpha=0.2$)	10
Dense+ReLU ($\alpha=0.2$)	5
Dropout layer ($\alpha=0.2$)	5
Dense	1

D. Event and Attack Data

The proposed framework can handle partially available event and attack data. While the measured event data are generally available, adequate attack data may not be available. Therefore, probing signals in the WADC measurement or control signals can be used to generate limited cyber attack scenarios. However, since probing signals cannot be used too frequently, the imbalance in event and attack cases can be addressed using the SVM-SMOT [35].

1) Event Data

Various power system events, such as generator outage, line outage, single-line-to-ground fault, line-to-line fault, double-line-to-ground fault, and three-phase faults, are created for various operational conditions that consider generation-load changes randomly within a ±20% range [16]. The various faults are created at 2 s and cleared at different time ($T_{clear}=2.01, 2.05, 2.10, 2.15,$ and 2.20 s). Additionally, double-circuit lines are created to replace single lines, and line outage cases are generated. For generator outage cases, generators G2 and G3 are tripped for different operational conditions. The total number of the generated cases for events is 24735, out of which 5000 event cases are labeled and the rest are unlabelled.

2) Design and Execution of Probing Signal for Control/measurement Signals

Various types of probing signals, including sinusoidal, ramp, pulse, random, saw-tooth, and triangular attacks, are designed and injected into the control and measurement signals [16], [27]. The steps for designing the signals are as follows.

Step 1: for implementing covert attacks on control/measurement signals, the oscillation frequency of the probing signal is determined through Prony analysis of multiple measured event data, which fall within the inter-area frequency range of 0.50-0.55 Hz for the two-area test system. The frequency of the injected probing signal is selected to match the inter-area mode at 0.51 Hz, the local modes at 1.00 Hz, and half of the inter-area mode frequency at 0.25 Hz [23], [27].

Step 2: during various events, the minimum and maximum values of WADC signals are monitored , and the probing signal is kept within this range, making the modeled attacks difficult to detect. Additionally, to simulate stealthy replay attacks, recorded data from three-phase faults are used when there is no actual disturbance in the system.

Step 3: probing signals are injected under various operational conditions to generate a data set that simulates the attack scenarios.

The probing signal is injected at 2 s and ends at 30 s. The labeled data include only 100 attack scenarios, which is relatively low, and there are just 2000 unlabeled attack instances, which is significantly fewer than the event data.

E. Metrics for Evaluating Performance

The effectiveness of the proposed framework depends on its ability to accurately identify attacks while minimizing false alarms. To quantitatively evaluate the performance of all the base learners and modified models, several metrics are employed, including accuracy, F1-score [44], and Matthew correlation coefficient (MCC) [45]. A score approaching 1 for each metric is desirable, indicating that the model performs exceptionally well for attack detection. Four parameters are used to calculate these metrics: true positive *TP*, true negative *TN*, false positive *FP*, and false negative *FN*.

The MCC in [45] is used for quantifying the correlation between the predicted value and true value. The use of MCC showcases the reliability of the classifier. This metric yields additional insights into the performance of the base learners and modified models for unbalanced data beyond traditional measures such as accuracy and F1-score. The MCC is mathematically defined as:

$$MCC=\frac{TP\times TN-FP\times FN}{\sqrt{(TP+FP)(TN+FP)(TP+FN)(TN+FN)}} \quad (17)$$

F. Analysis and Comparison of Experimental Results

To evaluate the impact of training data composition, fifteen distinct training scenarios (S1 to S15) are generated using random sampling from the event and attack database. This allows us to investigate the effectiveness of the proposed framework under varying training scenario characteristics. However, it creates an uneven distribution of events and attacks in the training data, ultimately affecting the training of all the base learners and modified models. To ensure an accurate evaluation of the performance of all the base learners and modified models, the unseen samples of the test set remain consistent across all scenarios. These cases allow for the examination of the performance of the base learners and modified models under operational conditions of limited labeled data and an unbalanced distribution of attacks and events. The proportion of attack and event class data in the

unlabeled training data set varies between two cases: case 1 and case 2.

Case 1: the unlabeled training data set consists of 20% attack data and 80% event data.

Case 2: the unlabeled training data set consists of 30% attack data and 70% event data.

The resulting combination of labeled and unlabeled data sets is used to train all models to evaluate their performance.

1) Case 1

The objective of the test is to evaluate the performance of all base learners and modified models with physics-aware features and determine the impact of various oversampling techniques. Various models demonstrated promising performance across 15 scenarios. The accuracy comparison depicted in Fig. 5 demonstrates that the SSGAN+SVM-SMOT model outperforms the baseline classifiers in the maximum scenarios. In the heat map, yellow indicates higher performance of the model for a specific case, while green represents weaker performance. Specifically, the SemiB, LP, and LP-OUPS models achieved an average test accuracy of 0.7670, 0.7787, and 0.7827, respectively. The SSGAN, SSGAN+OUPS, and SSGAN+MDO models achieved an accuracy of 0.8660, 0.8997, and 0.8910, respectively, while the SSGAN+SVM-SMOT model provides an even higher accuracy of 0.9094. The heat map of F1-score among various models across fifteen scenarios (case 1) is shown in Fig. 6. SemiB model achieves an average F1-score of 0.7760, and SSGAN+OUPS model achieves that of 0.8676, while the SSGAN+SVM-SMOT model exhibites a superior F1-score of 0.8812. The MCC metric is employed to evaluate overall performance of the models concerning both attack and event classifications, and the results are shown in Fig. 7. It has been observed that the SSGAN+SVM-SMOT model shows the highest average MCC of 0.8177. These comparative analyses indicate the efficacy of SSGAN+SVM-SMOT model in detecting attacks across fifteen scenarios. The obtained results emphasize the critical importance of having a sufficient number of data points during the training process. The integration of oversampling techniques, such as SVM-SMOT, enables the generation of synthetic data points, effectively increasing the quantity and diversity of the training.

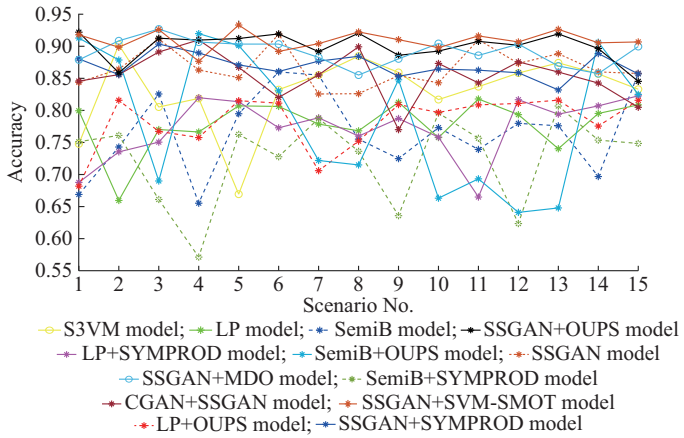

Fig. 5. Accuracy comparison among various models across fifteen scenarios (case 1).

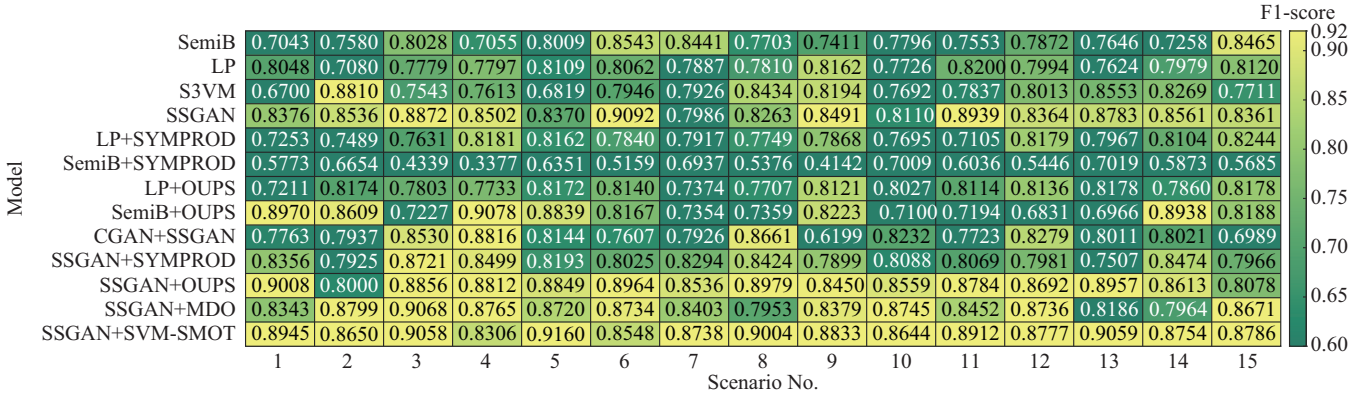

Fig. 6. Heat map of F1-score among various models across fifteen scenarios (case 1).

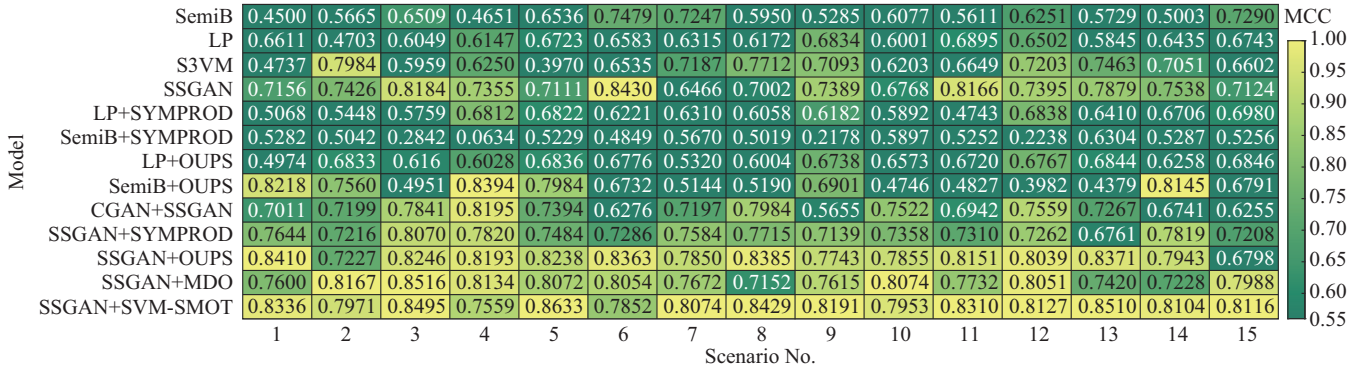

Fig. 7. Heat map of MCC among various models across fifteen scenarios (case 1).

2) Case 2

The objective of the test is to evaluate and compare the performance of various models as the proportion of actual attack data increased to 30% in the training data set. The accuracy in case 2 for various models is illustrated in Fig. 8.

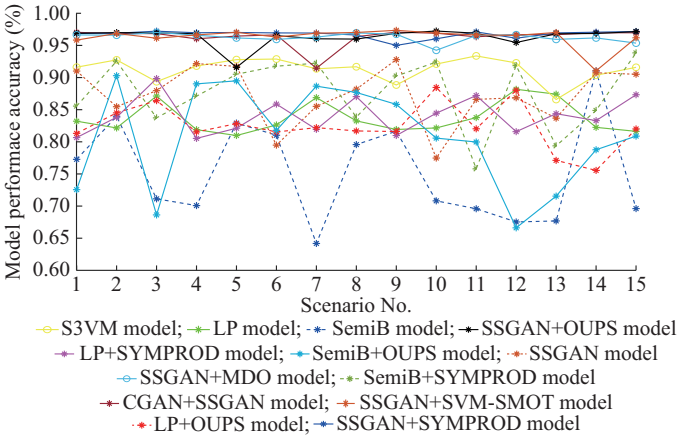

Fig. 8. Accuracy comparison among various models across fifteen scenarios (case 2).

It can be observed that the baseline classifier models, including SemiB model (0.7522), LP+SyMProD model (0.8409), SemiB+OUPS model (0.8084), SSGAN+OUPS model (0.9637), and SSGAN+MDO model (0.9629) exhibit low average performance while the SSGAN+SVM-SMOT model provides 0.9633. The corresponding heat map of F1-score among various models across fifteen scenarios (case 2) is depicted in Fig. 9. It can be observed that the SSGAN+SVM-SMOT model achieves an average F1-score of 0.9579. The heat map of MCC among various models across fifteen scenarios (case 2) is presented in Fig. 10.

Interestingly, the average scores of baseline models are SemiB+SyMProD model (0.7612), SSGAN+OUPS model (0.9280), and SSGAN+MDO model (0.9258) while the SSGAN+SVM-SMOT model achieves the highest average score of 0.9278. These results confirm that increasing the number of actual attack data points and subsequently generating artificial samples lead to an improvement in the performance of all base learners and modified models. Additionally, the F1-score and MCC increase for certain models like SSGAN-OUPS model due to the expanded training dataset with more attack samples. However, the SSGAN+SVM-SMOT model maintains consistent performance across both cases with low and high attack samples.

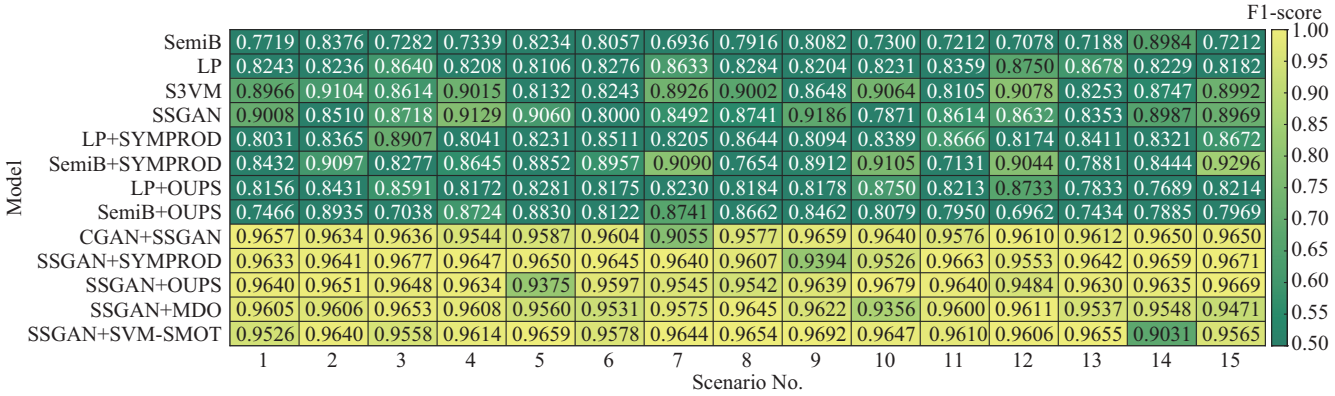

Fig. 9. Heat map of F1-score among various models across fifteen scenarios (case 2).

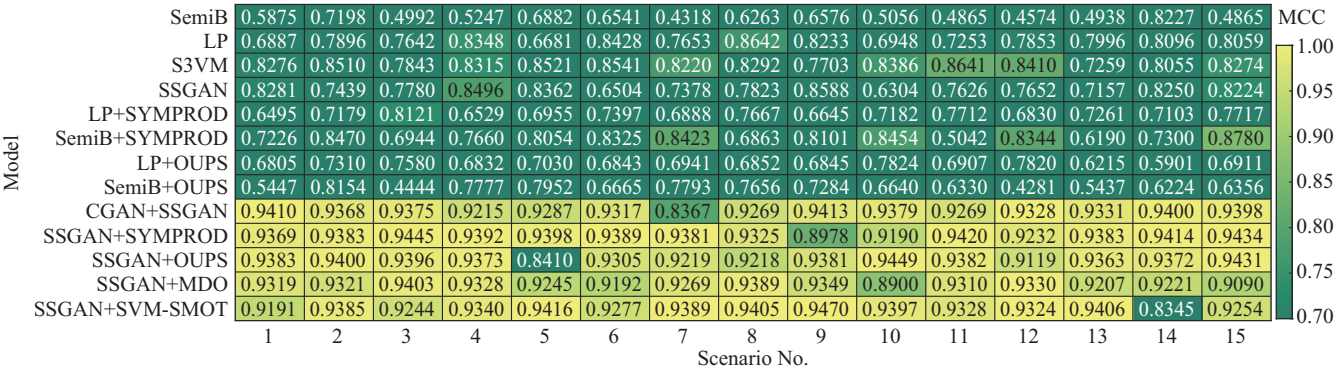

Fig. 10. Heat map of MCC among various models across fifteen scenarios (case 2).

G. Real-time HIL Test Analysis

The HIL cyber-physical testbed depicted in Fig. 11 is used to validate the proposed framework by using an OPAL-RT simulator (model OP5707XG) test bench. System parameters are identical to those used in the MATLAB/Simulink simulation. The experimental evaluation provides a real-time platform for analysis that incorporates delays and errors, which are absent in offline MATLAB simulations. The RT-Lab software is utilized for real-time interaction with the simulator. In this setup, the Schweitzer Engineering Laboratories (SEL-

2407) GPS clock is used to synchronize devices with coordinated universal time (UTC) by directly connecting to the Oregano syn1588-PCIe card in OPAL-RT simulator and the SEL-421 PMU using an inter-range instrumentation group time code-B (IRIG-B) signal. The SEL-421 PMU is used to transmit synchrophasor measurements to the local PDC. The SEL-3350 real-time automation controller (RTAC) acts as the local PDC and receives synchrophasor data according to the IEEE C37.118 standard. The NS-3 is utilized to emulate the communication infrastructure of a wide-area network. By utilizing the Linux kernel networking subsystem with the NS-3 tap bridge feature to connect NS-3 substation gateways (local PDC) with the control center application (super PDC). The measurement data are directed via NS-3 to the super PDC, in this case, OpenPDC. Similarly, NS-3 is also deployed to create wide-area network communication between the control center and the actuator. This arrangement allows for the introduction of time delays in the data transmission path, both on the measurement side between the local PDC and super PDC and on the control signal side between OpenPDC and the actuator. The time delays considered range from a minimum of 60 ms to a maximum of 600 ms [46]. The wireshark packet analyzer (Pcap) is used to verify the received data packets by OpenPDC following the IEEE C37.118 protocol. In real time, data are saved in a CSV file format, and a Python script is executed to calculate the WADC control signal for the proposed framework online. Subsequently, the calculated control signal is transmitted to the actuator simulated in the OPAL-RT simulator via NS-3.

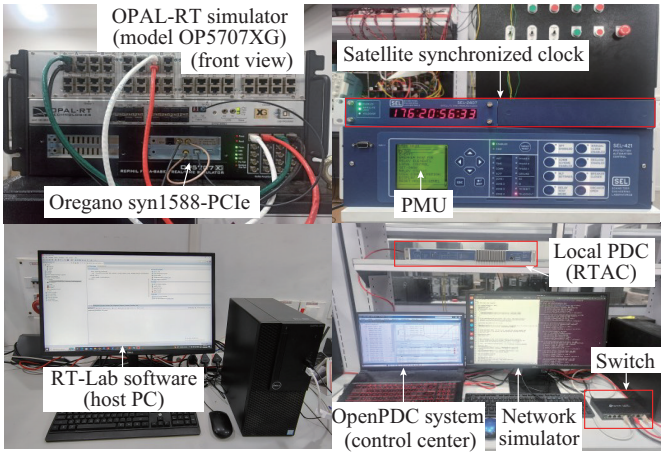

Fig. 11. Implementation of proposed framework in HIL cyber-physical testbed utilizing OPAL-RT simulator .

The trained SSGAN+SVM-SMOT model is deployed to the OpenPDC system (control center) to detect cyber attacks. To illustrate, a sinusoidal attack is executed at 1 s on the line power flow deviation of lines 7 and 8 (measurement signal), and the control signal is sent to G4. The test system shows sustained oscillations and moves toward instability. But when the proposed framework is implemented, the attack is detected, and a new set of measurement and control signals is selected. The input to the controller is changed to the power flow in lines 8 and 9, and the control signal is given to G3, which compensates for the adverse effect caused by the attack, as shown in Fig. 12.

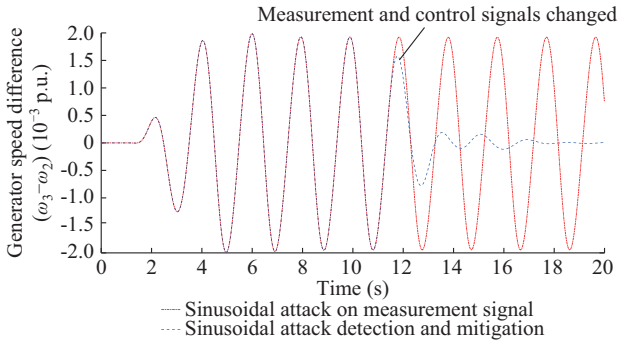

Fig. 12. Implementation of proposed framework in test system.

H. Performance on IEEE 16-machine 68-bus System

To further investigate the practicality of the proposed framework in a more complex power system, the tests are carried out on the IEEE 16-machine 68-bus test system [47]. There are three poorly damped inter-area modes, with frequencies of 0.514, 0.620, and 0.780 Hz, respectively. The generators that have the highest involvement in these modes are G13, G14, and G15. WADC is designed with active power flow measurements from lines 9-29, 13-17, 14-41, 16-18, and 54-1 as its inputs, and G3, G9, G13, G14, and G16 as the actuator plants. The SSGAN+MDO and SSGAN+SVM-SMOT models are implemented with physics-aware features to assess the efficacy of cyber attack detection. The results of unbalanced data cases are presented in Table III. From the table, it can be observed that SSGAN+SVM-SMOT model achieves the highest accuracy of 97.04%, with an F1-score of 96.54% and an MCC score of 94.04%.

TABLE III
RESULTS OF UNBALANCED DATA CASES

Model	Case	Accuracy	F1-score	MCC
SSGAN+MDO model	Case 1 (20%-80%)	0.8859	0.8449	0.7743
SSGAN+SVM-SMOT model	Case 1 (20%-80%)	0.8924	0.8580	0.7854
SSGAN+MDO model	Case 2 (30%-70%)	0.9632	0.9536	0.9258
SSGAN+SVM-SMOT model	Case 2 (30%-70%)	0.9704	0.9654	0.9404

VII. CONCLUSION

The conclusions of this paper are summarized as follows.

1) The CyResWadc system framework is proposed to detect and mitigate FDI attacks. The proposed framework uses an SSGAN model integrated with SVM-SMOT for the attack detection.

2) SVM-SMOT synthesizes new data instances by sampling along the decision boundary. It uses few nearest neighbors and applies interpolation or extrapolation depending on the density of the majority class data points around it. As a result, the SVM-SMOT generates realistic synthetic samples and ensures unbiased model training.

3) Probing signals are used to generate measurements that

resemble attack scenarios required for the SSL-based method training. It eliminates the need for actual attack data, which may not be possible to obtain at present.

4) The SSGAN+SVM-SMOT model is evaluated against existing models that incorporate oversampling techniques. The performance of the proposed framework has been validated in two-area and IEEE 16-machine 68-bus test systems. Real-time evaluation on a developed HIL cyber-physical testbed validates that the proposed framework effectively detects and mitigates attacks while maintaining system performance within acceptable limits.

REFERENCES

[1] D. Kosterev, C. Taylor, and W. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 967-979, Aug. 1999.

[2] S. R. P. Committee. (2017, Feb.). Record notes of deliberations of the 4th meeting on PSS tuning of generators held on 13th January 2017. [Online]. Available: https://www. srpc. kar. nic. in/website/2016/meetings/special/m4pssm13-01-17.pdf

[3] D. Trudnowski, B. Pierre, F. Wilches-Bernal *et al*., "Initial closed-loop testing results for the Pacific DC intertie wide area damping controller," in *Proceedings of 2017 IEEE General Meeting*, Chicago, USA, Jul. 2017, pp. 1-5.

[4] T. Chen and S. Abu-Nimeh, "Lessons from stuxnet," *Computer*, vol. 44, no. 4, pp. 91-93, Apr. 2011.

[5] R. Vaz, "Venezuela's power grid disabled by cyber attack," *Green Left Weekly*, no. 1213, p. 15, Mar. 2019.

[6] I. Lukicheva, D. Pozo, and A. Kulikov, "Cyberattack detection in intelligent grids using non-linear filtering," in *Proceedings of 2018 IEEE PES Innovative Smart Grid Technologies Conference Europe*, Sarajevo, Bosnia and Herzegovina, Oct. 2018, pp. 1-6.

[7] C. Liu, J. Wu, C. Long *et al*., "Reactance perturbation for detecting and identifying FDI attacks in power system state estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 763-776, Aug. 2018.

[8] M. Khalaf, A. Youssef, and E. El-Saadany, "Joint detection and mitigation of false data injection attacks in AGC systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 4985-4995, Sept. 2019.

[9] H. Karimipour and V. Dinavahi, "Robust massively parallel dynamic state estimation of power systems against cyber-attack," *IEEE Access*, vol. 6, pp. 2984-2995, Dec. 2017.

[10] Y. Li, W. Huo, R. Qiu *et al*., "Efficient detection of false data injection attack with invertible automatic encoder and long-short-term memory," *IET Cyber-Physical Systems*: *Theory & Applications*, vol. 5, no. 1, pp. 110-118, Mar. 2020.

[11] M. Esmalifalak, L. Liu, N. Nguyen *et al*., "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1644-1652, Sept. 2017.

[12] S. Wang, S. Bi, and Y. Zhang, "Locational detection of the false data injection attack in a smart grid: a multilabel classification approach," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8218-8227, Sept. 2020.

[13] T. Hu, Q. Guo, X. Shen *et al*., "Utilizing unlabeled data to detect electricity fraud in AMI: a semisupervised deep learning approach," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3287-3299, Jun. 2019.

[14] Y. Zhang, J. Wang, and B. Chen, "Detecting false data injection attacks in smart grids: a semi-supervised deep learning approach," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 623-634, Jan. 2021.

[15] Y. Zhao, R. Ball, J. Mosesian *et al*., "Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Transactions on Power Electronics*, vol. 30, no. 5, pp. 2848-2858, May 2015.

[16] R. Gelli and M. Govindarasu, "Anomaly detection and mitigation for wide-area damping control using machine learning," *IEEE Transactions on Smart Grid*, vol. 15, no. 6, pp. 5939-5951, Nov. 2024.

[17] K. Mahapatra, M. Ashour, N. Chaudhuri *et al*., "Malicious corruption resilience in PMU data and wide-area damping control," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 958-967, Mar. 2020.

[18] Y. Zhao, W. Yao, J. Nan *et al*., "Resilient adaptive wide-area damping control to mitigate false data injection attacks," *IEEE Systems Journal*, vol. 15, no. 4, pp. 4831-4842, Dec. 2021.

[19] W. Yao, J. Nan, Y. Zhao *et al*., "Resilient wide-area damping control for inter-area oscillations to tolerate deception attacks," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4238-4249, Sept. 2021.

[20] X. Xu and N. Tian, "The search and improvement of DES algorithm for data transmission security in SCADA," in *Proceedings of 2019 International Conference on Intelligent Computing*, *Automation and Systems*, Chongqing, China, Dec. 2019, pp. 275-279.

[21] M. Li and Y. Chen, "Wide-area robust sliding mode controller for power systems with false data injection attacks," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 922-930, Mar. 2020.

[22] L. Levaggi and E. Punta, "Analysis of a second-order sliding-mode algorithm in presence of input delays," *IEEE Transactions on Automatic Control*, vol. 51, no. 8, pp. 1325-1332, Aug. 2006.

[23] Z. Wang and S. Bu, "Design and defense of modal resonance-oriented cyber-attack against wide-area damping control," *IEEE Transactions on Smart Grid*, vol. 15, no. 2, pp. 2164-2178, Mar. 2024.

[24] Y. Zhao, W. Yao, C. Zhang *et al*., "Resilient wide-area damping control to mitigate strong cyber attack: a multiple-controller switching approach," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 2326-2337, May 2023.

[25] S. Liu, I. Zenelis, Y. Li *et al*., "Markov game for securing wide-area damping control against false data injection attacks," *IEEE Systems Journal*, vol. 15, no. 1, pp. 1356-1365, Mar. 2021.

[26] K. Sun, W. Qiu, Y. Dong *et al*., "WAMS-based HVDC damping control for cyber attack defense," *IEEE Transactions on Power Systems*, vol. 38, no. 1, pp. 702-713, Jan. 2023.

[27] A. Saini, P. Bhui, A. Singh *et al*., "Impact of false data injection attacks in wide area damping control," in *Proceedings of 2022 22nd National Power Systems Conference*, New Delhi, India, Dec. 2022, pp. 218-223.

[28] V. Pradhan, A. Kulkarni, and S. Khaparde, "A model-free approach for emergency damping control using wide area measurements," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4902-4912, Sept. 2018.

[29] T. Wang, M. Liu, J. Zhu *et al*. (2019, Oct.). Video-to-video synthesis. [Online]. Available: https://doi.org/10.48550/arXiv.1910.12713

[30] I. Goodfellow, J. Pouget-Abadie, M. Mirza *et al*., "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139-144, Oct. 2020.

[31] I. Kunakorntum, W. Hinthong, and P. Phunchongharn, "A synthetic minority based on probabilistic distribution (SyMProD) oversampling for imbalanced datasets," *IEEE Access*, vol. 8, pp. 114692-114704, Jun. 2020.

[32] N. Chawla, K. Bowyer, L. Hall *et al*., "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, Jun. 2002.

[33] L. Abdi and S. Hashemi, "To combat multi-class imbalanced problems by means of over-sampling techniques," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 1, pp. 238-251, Jan. 2016.

[34] W. Rivera and P. Xanthopoulos, "*A priori* synthetic over-sampling methods for increasing classification sensitivity in imbalanced data sets," *Expert Systems with Applications*, vol. 66, pp. 124-135, Dec. 2016.

[35] H. Nguyen, E. Cooper, and K. Kamei, "Borderline over-sampling for imbalanced data classification," *International Journal of Knowledge Engineering and Soft Data Paradigms*, vol. 3, no. 1, p. 4, Jan. 2011.

[36] T. Salimans, I. Goodfellow, W. Zaremba *et al*., "Improved techniques for training gans," *Advances in Neural Information Processing Systems*, vol. 29, pp. 2226-2234, Dec. 2016.

[37] P. K. Mallapragada, R. Jin, A. K. Jain *et al*., "SemiBoost: boosting for semi-supervised learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 2000-2014, Nov. 2009.

[38] X. Zhuĺ and Z. Ghahramaniĺн, "Learning from labeled and unlabeled data with label propagation," in *Proceedings of the 2002 International Joint Conference on Neural Networks*, Honolulu, USA, May 2012, pp. 1-12.

[39] F. Gieseke, A. Airola, T. Pahikkala *et al*., "Fast and simple gradient-based optimization for semi-supervised support vector machines," *Neurocomputing*, vol. 123, pp. 23-32, Jan. 2014.

[40] D. J. Trudnowski, "Estimating electromechanical mode shape from synchrophasor measurements," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1188-1195, Aug. 2008.

[41] I. Daubechies, J. Lu, and H. Wu, "Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 243-261, Mar. 2011.

[42] L. Chen, Y. Min, and W. Hu, "An energy-based method for location

of power system oscillation source," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 828-836, May 2013.
[43] Y. Zhang and A. Bose, "Design of wide-area damping controllers for interarea oscillations," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1136-1143, Aug. 2008.
[44] D. Powers. (2020，Oct.). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. [Online]. Available: https://arxiv.org/abs/2010.16061
[45] A. Al-abassi, A. N. Jahromi, H. Karimipour *et al*., "A self-tuning cyber-attacks' location identification approach for critical infrastructures," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 7, pp. 5018-5027, Jul. 2022.
[46] B. J. Pierre, F. Wilches-Bernal, D. Schoenwald *et al*., "Design of the Pacific DC intertie wide area damping controller," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3594-3604, Sept. 2019.
[47] M. Bento, "Design of a wide-area damping controller to tolerate permanent communication failure and time delay uncertainties," *Energy Systems*, vol. 13, no. 1, pp. 235-264, Jan. 2022.

Abhishek Saini received the M.Tech. degree in power systems from the National Institute of Technology Hamirpur (NITH), Hamirpur, India, in 2020. He is currently pursuing the Ph.D. degree with the Department of Electrical, Electronics, and Communication Engineering, Indian Institutes of Technology Dharwad, Dharwad, India. His research interests include cyber-security in smart grid, machine learning, deep learning, and dynamic state estimation application in modern power system.

Pratyasa Bhui received the M.Tech. degree in power and energy systems from the Indian Institute of Technology Kharagpur, Kharagpur, India, in 2013, and the Ph.D. degree in electrical engineering from Indian Institute of Technology Delhi, New Delhi, India, in 2017. He received the Power System Operation Corporation Award for one of the best doctoral theses on power systems in India. From 2017 to 2018, he was a Postdoctoral Fellow at Texas A&M University, Texas, USA. He is currently an Associate Professor in Indian Institute of Technology Dharwad, Dharwad, India. He is also worked as a Visiting Professor at West Virginia University, Morgantown, USA, in 2024. His research interests include power system dynamics, wide-area measurement system, and cybersecurity in smart grid.