Deep Learning Based Multi-Label Attack Detection for Distributed Control of AC Microgrids

Sheik M. Mohiuddin and Junjian Qi
Stevens Institute of Technology, Hoboken, New Jersey, USA
{smohiudd, jqi8}@stevens.edu

Sasha Fung, Yu Huang, and Yufei Tang
Florida Atlantic University, Boca Raton, Florida, USA
{sfung2017, yhwang2018, tangy}@fau.edu

***Abstract*—This paper presents a deep learning based multi-label attack detection approach for the distributed control in AC microgrids. The secondary control of AC microgrids is formulated as a constrained optimization problem with voltage and frequency as control variables which is then solved using a distributed primal-dual gradient algorithm. The normally distributed false data injection (FDI) attacks against the proposed distributed control are then designed for the distributed generator's output voltage and active/reactive power measurements. In order to detect the presence of false measurements, a deep learning based attack detection strategy is further developed. The proposed attack detection is formulated as a multi-label classification problem to capture the inconsistency and co-occurrence dependencies in the power flow measurements due to the presence of FDI attacks. With this multi-label classification scheme, a single model is able to identify the presence of different attacks and load change simultaneously. Two different deep learning techniques are compared to design the attack detector, and the performance of the proposed distributed control and the attack detector is demonstrated through simulations on the modified IEEE 34-bus distribution test system.**

I. Introduction

Microgrids are formed when distributed generators (DGs), energy storage systems, and loads are clustered as a single controllable entity to operate either independently or in conjunction with the main grid [1]–[3]. Microgrids can provide strong support to the grid by alleviating stresses, reducing feeder losses, and improving reliability, efficiency, and scalability [1]. For microgrid control, either centralized or distributed approach can be adopted [4]. In centralized control, a high-bandwidth, point-to-point communication is required between the central controller and local DG control units [1], which increases the communication and computational costs [1], [5]. The central controller also suffers from the risk of single point of failure [5]. By contrast, the distributed control that utilizes a sparse communication network provides a promising solution [1], [2], in which each DG only has access to the information of itself and its neighboring DGs [6], reducing computational complexity and the requirements on communication network, and improving scalability, reliability, and resiliency to faults and unknown system parameters [1].

For the distributed control of AC microgrids, the droop-free distributed control has recently been proposed [1], [2] which successfully achieves the objectives of average voltage regulation and active-reactive power sharing among the DGs. However, the control formulation in [1], [2] relies on extensive use of PI controllers and may not always theoretically guarantee convergence. Thus a generalized distributed control framework based on a formally formulated optimization problem is required for optimally coordinating the voltage regulation and power sharing objectives in AC microgrid control.

Furthermore, despite the advantages of distributed control, cyber-physical security has become a major concern [7], [8]. Due to a lack of central authority and relatively low security levels, distributed controllers are more susceptible to cyber attacks than their centralized counterparts. A malicious entity may inject false measurements to the exchanged data by attacking the nodes or the communication links [7]. Due to the collaborative nature of state update in the distributed control, a simple cyber attack such as false data injection (FDI) attack on an agent may make the controller deviate from the optimal solution or even make the system unstable [9].

The authors in [10] present FDI attacks on both the nodes and communication links in a droop-controlled microgrid and a mitigation approach based on the convergence of dual variables is developed. To mitigate the impact of FDI attack on frequency synchronization of AC microgrids, a distributed observer based attack detection strategy has been studied in [5]. The authors in [11] consider the application of a time-varying communication graph for FDI attack detection and mitigation purpose. In [12], Kullback Leibler (KL) divergence criteria is proposed for detecting FDI attacks. In [8], a distributed robust state estimation approach is integrated with the distributed control to enhance the attack resilience.

Experiments have demonstrated that data-driven deep learning algorithms can identify abnormal activities that cannot be detected by conventional bad data detection [13]. Various deep learning algorithms have been proposed to identify FDI attacks [14]–[16]. Detecting these attacks in real-time usually involves preprocessing historical sensor data for offline training and then using the trained model for new data in an online manner using multivariate data streams [13], [17]. This multivariate time series data scheme presents a challenge for classification [18]. Although different attack detection strategies are proposed for microgrids, application of multi-label deep learning based detection approach is relatively new [19].

Our major contributions can be summarized as follows.

This work was supported in part by the National Science Foundation under Grant Nos. ECCS-2103426 and OAC-2017597.

978-0-7381-3184-9/21/33.00$ ©2021 IEEE

1) We formulate the secondary control of AC microgrids as a constrained optimization problem with voltage and frequency as control variables. A distributed solving algorithm is developed based on the primal-dual gradient algorithm. Then FDI attacks against the proposed distributed control are designed for the DG output voltage, active power, and reactive power measurements.
2) We further formulate the FDI attack detection as a multi-label classification problem by considering the inconsistency and co-occurrence dependencies. The multivariate time series of power flow measurements are preprocessed and fed into deep learning models for feature extraction and attack detection. The advantage is that a single model is able to identify the presence of different attacks and load change simultaneously.

The remainder of this paper is organized as follows. Primal-dual gradient based distributed control and the FDI attack models are presented in Section II. The proposed deep learning based attack detection strategy is discussed in Section III. Validation through simulations is then presented in Section IV. Finally, the conclusions with our future research goal are presented in Section V.

II. Distributed Control and FDI Attack

A. Cyber-Physical Representation of AC Microgrids

Let the buses be those at the output of the LC filter of each DG. The remaining buses in the network are eliminated by Kron reduction. Denote the bus admittance matrix of the reduced network by $\mathbf{Y}$. The linearized approximation of the active and reactive power utilization ratios of DG i is [20]:

$$\lambda_{P_i} = \sum_{j=1}^{N} \big(G_{ij}v_j - B_{ij}\theta_j\big)/\overline{P}_i \tag{1}$$

$$\lambda_{Q_i} = -\sum_{j=1}^{N} \big(B_{ij}v_j + G_{ij}\theta_j\big)/\overline{Q}_i, \tag{2}$$

where N is the number of DGs, v_j and θ_j are the voltage magnitude and phase angle of bus j, G_{ij} and B_{ij} are the real and imaginary parts of the admittance matrix $\mathbf{Y}$, and $\overline{P}_i$ and $\overline{Q}_i$ are the active and reactive power limits of DG i.

For distributed control implementation we consider a sparse communication network that is modeled as a directed graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ in which nodes are represented as the agents and edges are the communication links connecting nodes. The communication network can be represented by an adjacency matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{N\times N}$ where $a_{ij} > 0$ if there is a connection between node i and j and $a_{ij} = 0$ otherwise. It is assumed that $\mathcal{G}$ has a spanning tree and a balanced Laplacian matrix [1], [2].

B. Secondary Control Problem Formulation

The design objectives of the secondary control of AC microgrids include: 1) regulating frequency back to nominal frequency, 2) regulating the average voltage of the output buses of all inverters to the rated voltage, and 3) achieving proportional active and reactive power sharing among all inverters. Thus the following optimization problem is defined for DG i:

$$\begin{aligned}\min_{\mathbf{v},\boldsymbol{\omega}} f_i = \frac{1}{2}\Bigg(&\sum_{j\in\mathcal{N}_i} a_{ij}\big(\lambda_{P_i} - \lambda_{P_j}\big)^2 \\ &+ \sum_{j\in\mathcal{N}_i} a_{ij}\big(\lambda_{Q_i} - \lambda_{Q_j}\big)^2 + \frac{1}{\tau}(\omega_i - \omega_0)^2\Bigg) \\ \text{s.t.}\quad & \mathbf{1}^\top\mathbf{v}/N - v^{\mathrm{r}} = 0,\end{aligned} \tag{3}$$

where v^{r} is the rated voltage in per unit and ω_0 is the rated frequency, $\tau > 0$ is a constant, $\mathbf{v} = [v_1, v_2, \cdots, v_N]^\top$ is the DG output voltage vector, $\boldsymbol{\omega} = [\omega_1, \omega_2, \cdots, \omega_N]^\top$ is the DG frequency vector, and $\mathcal{N}_i$ is the set of neighbors of DG i in $\mathcal{G}$. Note that the linearized approximation in (1)–(2) is adopted in order to get a convex optimization problem.

Let $h(\mathbf{v}) = \mathbf{1}^\top\mathbf{v}/N - v^{\mathrm{r}}$. The Lagrange function for the optimization problem in (3) can be defined as:

$$L_i(\mathbf{v}, \boldsymbol{\omega}) = f_i + \mu_i h, \tag{4}$$

where μ_i is the Lagrange multiplier for the equality constraint. In this Lagrange function we need the output voltage from all DGs to compute the global average which may not be locally available due to the distributed implementation of the control. To address this problem a distributed average voltage estimator is implemented in the paper. Specifically, the average voltage of all inverter output buses, $\mathbf{1}^\top\mathbf{v}[n]/N$, can be estimated by DG $i = 1, \ldots, N$ as $v_i^{\mathrm{av}}[n]$ using the following distributed observer based on dynamic consensus [1]:

$$v_i^{\mathrm{av}}[n] = v_i[n] + \sum_{t=0}^{n}\sum_{j\in\mathcal{N}_i} a_{ij}\big(v_j^{\mathrm{av}}[t] - v_i^{\mathrm{av}}[t]\big)\Delta t, \tag{5}$$

where Δt is the step size. It has been proven in [1] that for $\forall i = 1, 2, \cdots, N$, v_i^{av} converges to a consensus value which is the true global average voltage when the communication network has a spanning tree and a balanced Laplacian matrix.

C. Distributed Control Algorithm

In order to implement the proposed distributed control we need to evaluate the gradient of the Lagrange function L_i with respect to v_i and ω_i as follows:

$$\begin{aligned}\frac{\partial L_i}{\partial v_i} &= \frac{\partial(f_i + \mu_i h)}{\partial v_i} \\ &= \Bigg(\sum_{j\in\mathcal{N}_i} a_{ij}(\lambda_{P_i} - \lambda_{P_j})\Big(\frac{G_{ii}}{\overline{P}_i} - \frac{G_{ij}}{\overline{P}_j}\Big) \\ &\quad + \sum_{j\in\mathcal{N}_i} a_{ij}(\lambda_{Q_i} - \lambda_{Q_j})\Big(\frac{-B_{ii}}{\overline{Q}_i} + \frac{B_{ij}}{\overline{Q}_j}\Big)\Bigg) \\ &\quad + \mu_i\mathcal{D}_{v_i}|v_i^{\mathrm{av}} - v^{\mathrm{r}}|\end{aligned} \tag{6}$$

$$\begin{aligned}\frac{\partial L_i}{\partial \omega_i} &= \frac{\partial(f_i + \mu_i h)}{\partial \omega_i} \\ &= \Bigg(\sum_{j\in\mathcal{N}_i} a_{ij}(\lambda_{P_i} - \lambda_{P_j})\Big(\frac{-B_{ii}}{\overline{P}_i} + \frac{B_{ij}}{\overline{P}_j}\Big)\end{aligned}$$

Algorithm 1 Primal-dual gradient based distributed control algorithm for DG i

Initialization: Set $v_i[0] = v_i^{\text{meas}}$, $\omega_i[0] = \omega_i^{\text{meas}}$, $\theta_i[0] = \theta_i^{\text{meas}}$, $\mu_i[0] = 0$, and $v_i^{\text{av}}[0] = v_i[0]$. Set $n = 0$.
(S.1) Update $v_i[n+1]$ based on (9)
(S.2) Update $\theta_i[n+1]$ based on (10)
(S.3) Update $\omega_i[n+1]$ based on (11)
(S.4) Update $v_i^{\text{av}}[n+1]$ based on (5)
(S.5) Update $\mu_i[n+1]$ based on (12)
(S.6) increase n by 1 and go to **(S.1)**

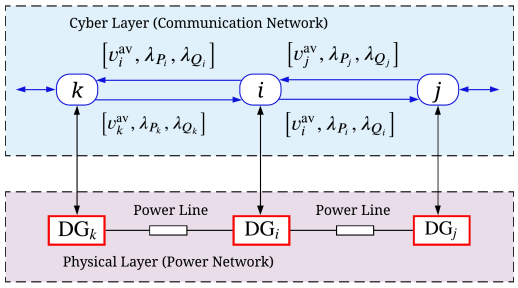

Fig. 1. Information flow of the proposed distributed control.

$$+\sum_{j\in\mathcal{N}_i} a_{ij}(\lambda_{Q_i}-\lambda_{Q_j})\Big(\frac{-G_{ii}}{\overline{Q}_i}+\frac{G_{ij}}{\overline{Q}_j}\Big)\Big)\frac{\partial\theta_i}{\partial\omega_i}$$
$$+\frac{1}{\tau}(\omega_i-\omega_0), \tag{7}$$

where $\mathcal{D}_{v_i}$ is the operator for subgradient with respect to v_i. Using Leibniz's rule for differentiation under the integral sign [21], there is:

$$\frac{\partial\theta_i}{\partial\omega_i}=\frac{\partial}{\partial\omega_i}\int_{t-\Delta t}^{t}\big(\omega_i(\tau)-\omega_0\big)d\tau=\Delta t. \tag{8}$$

We then implement a primal-dual gradient based distributed algorithm [22] to solve the optimization problem (3). For agent i the variable update equations are:

$$v_i[n+1]=v_i[n]-\alpha\partial L_i/\partial v_i \tag{9}$$
$$\theta_i[n+1]=\theta_i[n]+(\omega_i[n]-\omega_0)\Delta t \tag{10}$$
$$\omega_i[n+1]=\omega_i[n]-\tau\partial L_i/\partial\omega_i$$
$$=\omega_0-\tau\Bigg(\sum_{j\in\mathcal{N}_i}a_{ij}(\lambda_{P_i}[n]-\lambda_{P_j}[n])\Big(\frac{-B_{ii}}{\overline{P}_i}+\frac{B_{ij}}{\overline{P}_j}\Big)$$
$$+\sum_{j\in\mathcal{N}_i}a_{ij}(\lambda_{Q_i}[n]-\lambda_{Q_j}[n])\Big(\frac{-G_{ii}}{\overline{Q}_i}+\frac{G_{ij}}{\overline{Q}_j}\Big)\Bigg)\Delta t \tag{11}$$
$$\mu_i[n+1]=\mu_i[n]+\gamma|v_i^{\text{av}}[n+1]-v^{\text{r}}|. \tag{12}$$

Note that when the DGs achieve active and reactive power sharing in steady state the second term on the right-hand side of (11) becomes zero, which implies that the microgrid will achieve frequency synchronization in steady state.

The proposed algorithm is presented in Algorithm 1. Fig. 1 shows the information flow of the proposed control algorithm.

D. FDI Attack Against Distributed Control

In this paper we have considered the following three types of FDI attack models on the DG output voltage, active power, and reactive power measurements.

1) *Attacks on voltage measurements*: In this case, the DG output voltages are randomly changed by injecting an attack vector $\mathbf{u}_v^a$. The attack vector $\mathbf{u}_v^a$ is a normally distributed random vector with zero mean and standard deviation as 20% of the initial voltage. Due to the presence of attack, the output voltage of the DGs can be written as $\mathbf{v}^a=\mathbf{v}+\mathbf{u}_{\mathbf{v}}^a$, where $\mathbf{v}^a$ and $\mathbf{v}$ respectively represent the corrupted and actual measurements.
2) *Attacks on active and reactive power measurements*: For the attack on active power $\mathbf{P}$ and reactive power $\mathbf{Q}$ measurements, the corrupted measurements $\mathbf{P}^a$ and $\mathbf{Q}^a$ can be written as $\mathbf{P}^a=\mathbf{P}+\mathbf{u}_{\mathbf{P}}^a$ and $\mathbf{Q}^a=\mathbf{Q}+\mathbf{u}_{\mathbf{Q}}^a$, respectively, where $\mathbf{u}_{\mathbf{P}}^a$ and $\mathbf{u}_{\mathbf{Q}}^a$ are normally distributed random attack vectors with zero mean and standard deviations as 20% of the initial values.
3) *Attacks on voltage, active power, and reactive power measurements*: In this case we have considered the extreme scenarios in which attack vectors $\mathbf{u}_{\mathbf{v}}^a$, $\mathbf{u}_{\mathbf{P}}^a$, and $\mathbf{u}_{\mathbf{Q}}^a$ are injected to the DG output voltage, active power, and reactive power measurements.

III. Deep Learning Based Multi-Label Attack Detection Approach

We propose an effective FDI attacks detection mechanism using a multi-label classification scheme. In this section, we present the problem formulation and two different deep learning models for comparison.

A. Multi-Label Classification Problem Formation

Using one classifier to simultaneously evaluate multiple classes creates a substantial computational advantage over using multiple classifiers. For example, a single model can simultaneously detect anomaly events (i.e., change detection) and identify the anomaly types (i.e., anomaly diagnosis) after certain time steps using multivariate time series. Formally, we define the FDI attacks detection as a problem of multivariate time series classification with multi-label classes. Fig. 2 illustrates the problem setting and the notation. Given the historical data of n time series with length T, i.e., $\mathcal{X}=(x_1,\cdots,x_n)^\top\in\mathbb{R}^{n\times T}$, assume there is no anomaly before T. For the time segment, given a set of labels $\mathcal{Y}$, the task is to classify the data as that of regular behavior or anomaly. In our problem, the input multivariate time series $\mathcal{X}$ are voltage, active power, reactive power, and frequency, and the output labels $\mathcal{Y}\in\mathbb{R}^{1\times 4}$ are normal, load change, voltage attack, and power attack. Note that in the traditional transmission grid FDI attack detection problem the state estimate and bad data detection modules are usually applied before the attack detector. Most of the attack vectors will be filtered out before receiving by the deep learning based model. In this paper, the compromised measurements are directly incorporated into the proposed distributed control. The strong coupling will mix

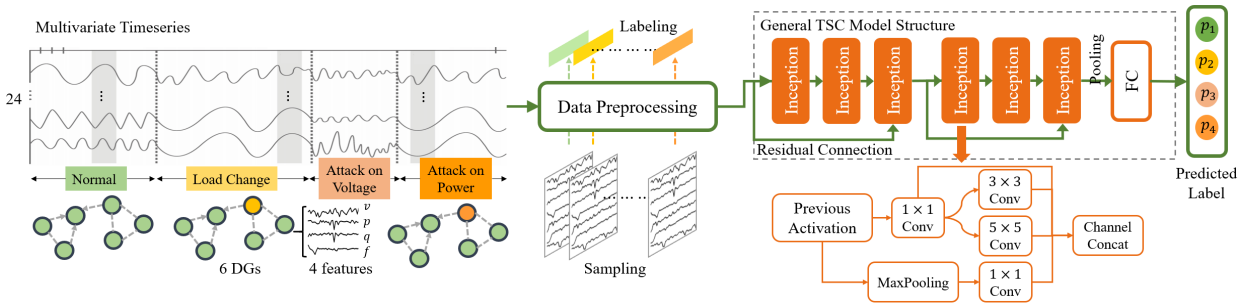

Fig. 2. Overall architecture of the proposed deep learning based FDI attack detection as a multi-label classification problem.

the attack patterns in the measurements and increase the attack detection and classification difficulty.

B. Data Preparation and Preprocessing

Features are extracted from the recorded multivariate time series, which are first preprocessed through a series of nonlinear transformations described as follows.

- Concatenation: There are four time series matrices that represent the four features for classification and six time series in each matrix that represent each of the six DGs. All of these time series are concatenated to create a multivariate time series feature matrix with 24 rows.
- Downsampling: Each 22-s simulation contains data that was sampled twice every 0.0001s. This data is down sampled to 100 samples per second.
- Window Slicing: The data matrix is segmented into windows of 24 by 500, which is (6 DGs $\times$ 4 features) by 500 time points. This is done using the window slicing technique in [23], where a window of the specified size slides across the time dimension of the data matrix with no overlap, effectively segmenting the time series.
- Labeling: The four labels are *normal*, *load change*, *attack on voltage*, *attack on power*. The samples from the first 30 s before the event are labelled as *normal*. The samples 15 s after the event is introduced are labelled according to the occurring events.

C. Deep Learning Models

Time series classification, especially that of multivariate nature, is an ongoing topic of research in machine learning. The classifiers have to be able to extract and effectively process the temporal qualities of the data. In this paper, the InceptionTime [24] which is considered to be the state-of-the-art for time series classification model is compared with the baseline ResNet model [25]. The basic structure of these models is depicted in Fig. 2. Both feature residual connections that are comprised of 1D convolutions and batch normalization. Inception has six residual blocks with max pooling and features a bottleneck layer while ResNet has three residual blocks with no pooling or bottleneck. The residual block layer connects to a global average pooling layer which then feeds to a dense fully connected network. The sigmoid function is employed at the output layer and the nonbinary values are transformed to binary using Matthew's correlation coefficient threshold calibration.

The data set consists of 1200 samples that are split with 33% reserved for testing. Glorot's uniform initialization is employed for all models. The training is done using mini-batch size of 15. The loss function is minimized for model optimization. The multi-label classification scheme minimizes the total loss—the sum of the binary cross-entropy loss function over each of the four labels in all training samples:

$$\mathcal{L} = \sum_{i=1}^{C} \Big(T_i \cdot \log_2 O_i + (1 - T_i) \cdot \log_2 (1 - O_i) \Big), \qquad (13)$$

where $C = 4$ is the class/label number, O_i is the predicted label of output node i, and T_i is the target label. The optimization is processed by a variant of Stochastic Gradient Descent (SGD), Adam, and the learning rate (with a minimum of 0.0001) is reduced by a factor of 0.5 each time the model's training loss has not improved for 10 consecutive epochs.

Swish Activation Function Traditionally, both the InceptionTime and ResNet models use the ReLu activation function between convolutional layers. This function was created in an attempt to rectify the vanishing gradient problem exhibited by logistic functions such as sigmoid and hyperbolic tangent. While rectifying the vanishing gradient, ReLu suffers from the "dying ReLu" problem due to the lack of gradient in the negative region, where the function is equal to 0. The swish activation function has been proposed as an alternate to the ReLU. The Swish creates a sigmoidal shape gradient in the negative region, where the sigmoidal shape can be tuned or trained using the parameter β in the swish function as $f(x) = x \cdot \text{sigmoid}(\beta x)$, where $\beta = 1$ is used in these experiments. This function has been shown to outperform the ReLu function in many situations [26]. The InceptionTime and ResNet models were modified to use the Swish function and the results are compared to the ReLu variants.

IV. Performance Evaluation

A. Test System and Control Performance

The distributed control based on the proposed Algorithm 1 is tested on the modified IEEE 34-bus distribution test system shown in Fig. 3. This system has 6 DGs and 9 loads. The line parameters are adopted from [27]. The simulations are

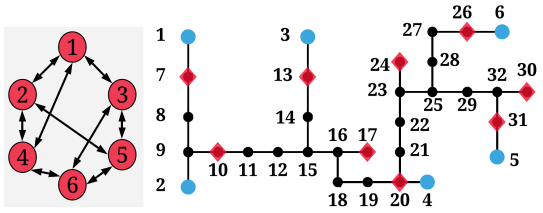

Fig. 3. Modified IEEE 34-bus distribution test system. Blue circles indicate DGs and red rectangles indicate loads. The communication network used in the proposed control is shown on the left-hand side.

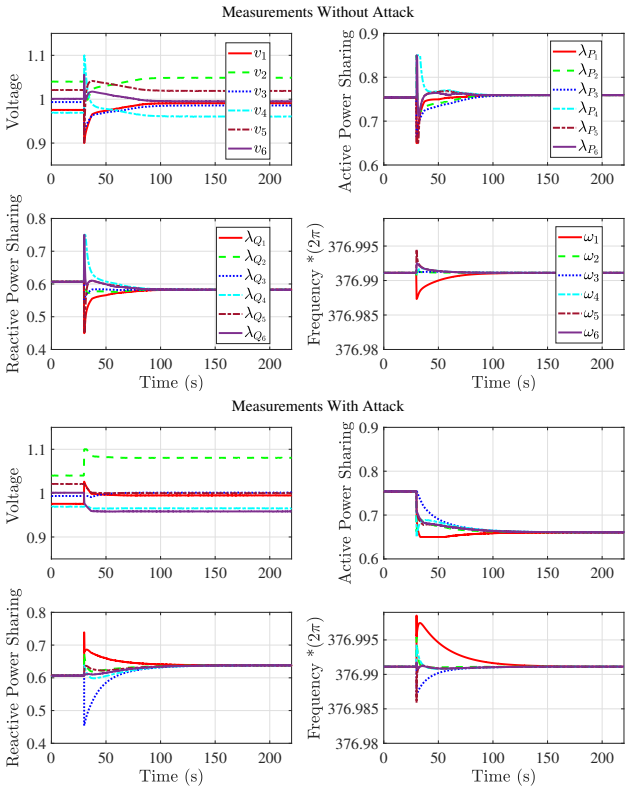

Fig. 4. DG output voltages, active/reactive power, and frequency: (a) Without FDI attack and only load change is applied at 30 s; (b) Both FDI attack on voltage, active power, and reactive power measurements and load change are applied at 30 s.

performed in Matlab without including the detailed zero-level control of the inverters. The communication network used in the proposed distributed control is shown on the left-hand side of Fig. 3. The α, τ, and γ in (9), (11), and (12) are respectively selected as 0.000075, 0.000075, and 0.0001.

For the data preparation we generated 110 test cases for each of the scenarios—load change, voltage attack under load change, active/reactive power attack under load change, and voltage, active-reactive power attack under load change. For applying load change we consider a normally distributed random change with zero mean and standard deviation as 20% of the initial loads.

Fig. 4 shows the DG output voltages, active/reactive power sharing, and output frequencies under the proposed distributed control. In Fig. 4a, we apply a load change at 30 s without any FDI attack. In Fig. 4b we apply both load change and FDI attack on voltage and active/reactive power measurements at 30 s. For Figs. 4a and 4b we consider the same load change but obtain different steady states due to the presence of FDI attack. The FDI attack can mislead the microgrid operator as normal load change events and is challenging to detect using existing approaches. This motivates us to investigate the deep-learning based multi-label attack detection approaches.

B. Deep Learning Performance Metrics

We evaluate the results with precision, recall, and F1 score, which are typical multi-label classification performance metrics to evaluate deep learning based models. The precision is the ratio $tp/(tp + fp)$ where tp is the number of true positives and fp the number of false positives. The precision is intuitively the ability of the classifier not to label as positive a sample that is negative. The recall is the ratio $tp/(tp + fn)$ where tp is the number of true positives and fn the number of false negatives. The recall is intuitively the ability of the classifier to find all the positive samples. F1 score can be interpreted as a weighted average of the precision and recall:

$$\mathrm{F1} = 2 \cdot \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}, \tag{14}$$

where F1 micro is calculated globally by counting the total true positives, false negatives and false positives while F1 macro is calculated for each label, and find their unweighted mean.

The relative trade-off between tp rate and fp rate is depicted using AUC, the area under the ROC. This measure shows how good the model is at making correct predictions. An AUC closer to 1 signifies excellent performance. The micro-average and macro-average ROC curves over all labels are plotted to assess the multi-label performance. Further analysis is done on the performance of single labels.

C. FDI Attack Detection Results

The ResNet and Inception models are both tested for their abilities to distinguish between load change only and attack only scenarios. Both models are above 99% accurate in this case. The following results show how well the models can identify attacks when they are coordinated with load changes.

The three performance metrics are compared among the models with the results shown in Table I. As can be observed, the implementation of the Swish activation function significantly improves the overall accuracy of the models. This shows that more samples have all labels correctly classified when using Swish.

TABLE I
MODEL PERFORMANCE RESULTS

Model	Activation	Precision (%)	Recall (%)	F1 (%)
Inception	ReLU	91.03	95.31	87.13
ResNet	ReLU	95.97	95.12	96.84
Inception	Swish	97.06	96.27	97.87
ResNet	Swish	97.13	97.84	96.45

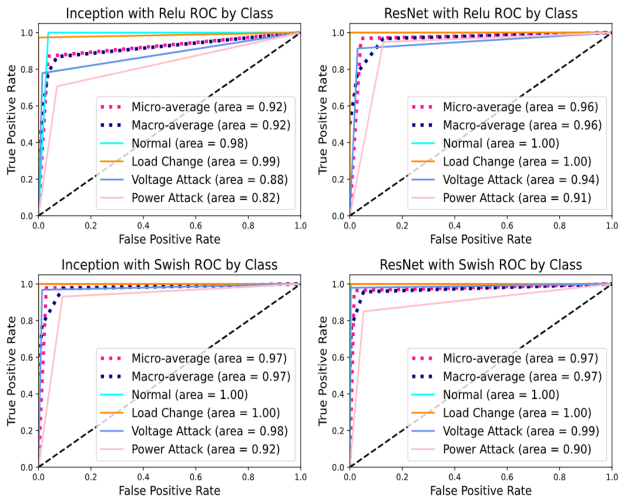

Fig. 5. ROC curves. Top: Inception and ResNet wish ReLU. Bottom: Inception and ResNet with Swish.

The ROC curves are shown in Fig. 5. All models have macro and micro average AUC greater than or equal to 0.92, showing very good performance over multiple labels. The individual labels show that the power attack detection was slightly worse than the rest of the labels. The performance of Inception and ResNet is comparable, but ResNet is shown to have better performance in both the ReLu and Swish models. ResNet with Swish is the best performing model with the highest overall accuracy at 91%. The ResNet model also trains twice as fast as the Inception due to less residual connections, making it a more desirable model[1].

V. Conclusion

In this paper, we propose a deep-learning based attack detection strategy for distributed control of AC microgrids. A distributed primal-dual gradient based algorithm is developed to control the distributed generators in the microgrid and random FDI attack vectors are injected to the DG output voltage and active/reactive power measurements. A deep learning based multi-label attack detection technique is developed to detect the presence of attacks which gives a high accuracy for attack detection. In our future work, we will incorporate the physical model and domain knowledge of the microgrid into the data-driven deep-learning model to improve accuracy.

References

[1] V. Nasirian, Q. Shafiee, J. M. Guerrero, F. L. Lewis, and A. Davoudi, "Droop-free distributed control for AC microgrids," *IEEE Trans. Power Electron.*, vol. 31, no. 2, pp. 1600–1617, Feb. 2016.

[2] S. M. Mohiuddin and J. Qi, "Droop-free distributed control for AC microgrids with precisely regulated voltage variance and admissible voltage profile guarantees," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 1956–1967, May 2020.

[3] D. E. Olivares, A. Mehrizi-Sani, A. H. Etemadi, C. A. Cañizares, R. Iravani, M. Kazerani, A. H. Hajimiragha, O. Gomis-Bellmunt, M. Saeedifard, R. Palma-Behnke *et al.*, "Trends in microgrid control," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1905–1919, Jul. 2014.

[1]Source code and testing data for reproducing the results are available at https://github.com/IRES-FAU/Multi-Label-Attack-Detection-for-Microgrids.

[4] H. Sun, Q. Guo, J. Qi, V. Ajjarapu, R. Bravo, J. Chow, Z. Li, R. Moghe, E. Nasr-Azadani, U. Tamrakar, G. N. Taranto, R. Tonkoski, G. Valverde, Q. Wu, and G. Yang, "Review of challenges and research opportunities for voltage control in smart grids," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2790–2801, Jul. 2019.

[5] S. Abhinav, H. Modares, F. L. Lewis, F. Ferrese, and A. Davoudi, "Synchrony in networked microgrids under attacks," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6731–6741, Nov. 2018.

[6] S. M. Mohiuddin and J. Qi, "A unified droop-free distributed secondary control for grid-following and grid-forming inverters in AC microgrids," in *IEEE Power & Energy Society General Meeting*, 2020, pp. 1–5.

[7] J. Qi, A. Hahn, X. Lu, J. Wang, and C.-C. Liu, "Cybersecurity for distributed energy resources and smart inverters," *IET Cyber-Physical Syst.: Theory & Appl.*, vol. 1, no. 1, pp. 28–39, Dec. 2016.

[8] S. M. Mohiuddin and J. Qi, "Attack resilient distributed control for AC microgrids with distributed robust state estimation," in *2021 IEEE Texas Power and Energy Conference (TPEC)*, 2021, pp. 1–6.

[9] H. Zhang, W. Meng, J. Qi, X. Wang, and W. X. Zheng, "Distributed load sharing under false data injection attack in an inverter-based microgrid," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1543–1551, Feb. 2019.

[10] L. Lu, H. J. Liu, H. Zhu, and C. Chu, "Intrusion detection in distributed frequency control of isolated microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6502–6515, Nov. 2019.

[11] A. Bidram, B. Poudel, L. Damodaran, R. Fierro, and J. M. Guerrero, "Resilient and cybersecure distributed control of inverter-based islanded microgrids," *IEEE Trans. Ind. Informat.*, pp. 1–1, Sep. 2019.

[12] A. Mustafa, B. Poudel, A. Bidram, and H. Modares, "Detection and mitigation of data manipulation attacks in AC microgrids," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2588–2603, May 2020.

[13] X. Niu, J. Li, J. Sun, and K. Tomsovic, "Dynamic detection of false data injection attack in smart grid using deep learning," in *2019 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2019, pp. 1–6.

[14] J. Wei and G. J. Mendis, "A deep learning-based cyber-physical strategy to mitigate false data injection attack in smart grids," in *2016 Joint Workshop on Cyber- Physical Security and Resilience in Smart Grids (CPSR-SG)*, 2016, pp. 1–6.

[15] Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2505–2516, 2017.

[16] S. Basodi, S. Tan, W. Song, and Y. Pan, "Data integrity attack detection in smart grid: a deep learning approach," *International Journal of Security and Networks*, vol. 15, p. 15, 01 2020.

[17] 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*, pp. 1–1, 2020.

[18] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Deep learning for time series classification: a review," *Data Mining and Knowledge Discovery*, vol. 33, no. 4, pp. 917–963, 2019.

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

[20] J. Yang, N. Zhang, C. Kang, and Q. Xia, "A state-independent linear power flow model with accurate estimation of voltage magnitude," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3607–3617, Sept. 2017.

[21] Y. Xu, Z. Qu, and J. Qi, "State-constrained grid-forming inverter control for robust operation of AC microgrids," in *2020 European Control Conference (ECC)*, May 2020, pp. 471–474.

[22] Y. Nesterov, *Introductory lectures on convex optimization: A basic course*. Springer Science & Business Media, Dec. 2013, vol. 87.

[23] A. Le Guennec, S. Malinowski, and R. Tavenard, "Data augmentation for time series classification using convolutional neural networks," in *ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data*, 2016.

[24] L. B. F. G. e. a. Ismail Fawaz, H., "Inceptiontime: Finding alexnet for time series classification." *Data Min Knowl Disc.*, vol. 34, 2020.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[26] P. Ramachandran, B. Zoph, and Q. V. Le, "Swish: a self-gated activation function," *arXiv: Neural and Evolutionary Computing*, 2017.

[27] N. Mwakabuta and A. Sekar, "Comparative study of the IEEE 34 node test feeder under practical simplifications," in *Proc. 39th North Amer. Power Symp.*, 2007, pp. 484–491.