Artificial Intelligence for Real-Time Topology Identification in Power Distribution Systems

Yifu Li, Shiyuan Wang, Li Li, and Payman Dehghanian
Department of Electrical and Computer Engineering
The George Washington University
800 22nd St NW, Washington, Suite 5900, DC 20052, USA.
{liyifu, shiyuan1225, lili1986, payman}@gwu.edu

***Abstract*—With the increase in demand for electricity and the number of end-use consumers, the operation and control of power grids have become more and more complex and challenging. Ensuring acceptable reliability and quality of the electricity supply has become particularly important to every aspect of our electrified economy. Due to the growing deployment of Micro-Phasor Measurement Units (μPMUs) in power distribution grids, an abundance of high-resolution measurements is available that can be harnessed for smarter operation and fault analyses in power distribution networks. Traditional models have revealed limitations on the network topology identification which may occupy manpower and material resources with no guaranty to effectively restore power in a short time period when facing faults and other disruptions. This paper suggests and implements a machine learning framework that uses the μPMU measurements as inputs and provides a full observation of the network topology in real-time. Specifically, the proposed framework employs a Convolutional Neural Network (CNN) to identify the physical state of the power network at all times. We evaluated the framework on the IEEE 34-node test feeder, where the experiments show that the proposed CNN can achieve a promising performance with high accuracy even when the μPMU measurements contain noises and missing entries.**

***Index Terms*—Micro-Phasor Measurement Unit (μPMU); Convolutional Neural Network (CNN); power distribution network; topology identification.**

I. Introduction

In order to generate and dispatch the electrical power efficiently and operate the power grid stably and safely, system operators need to be informed of the electrical network topology and demand profile at all times. The observability and controllability of the electrical network are essential to ensure its safe and economic operation. With the growing complexity in the power grid structure growingly reinforced with heterogeneous resources and the increasing demand for electricity needed for an electrified economy, Phasor Measurement Units (PMUs) have been introduced and widely deployed to observe the dynamic performance of the power grid with synchronized measurements. Compared to the traditional event detection schemes and infrastructure such as Feeder Terminal Unit (FTU), distribution Transformer supervisory Terminal Unit (TTU), and Remote Terminal Unit (RTU), Micro-PMUs (μPMUs) in power distribution grids offer yet-untapped potential for online situational awareness, i.e., event detection, classification, and high-fidelity high-resolution measurements.

Access to high-fidelity measurements in power distribution systems is particularly critical, and at the same time challenging due to the following reasons [1]:

(i) The length of the power distribution lines is usually between 5 to 10 kilometers, resulting in the phase angle difference between the two ends of the line to be commonly small (sometimes even lower than $0.1°$).
(ii) The proliferation and rushing arrival of renewables have increased the complexity in the grid structure and the way electricity flows in the network. Three-phase unbalance architectures are commonly seen in power distribution systems, which could result in more than 30% inter-harmonics and under 60dB noise conditions.
(iii) The fast switching characteristics of power electronic devices lead to more electrical transients, further mandating the higher efficiency and dynamic tracking capability of the event detectors in the power distribution sector.

While PMU measurements can be shared over communication networks in real-time and collected at a centralized platform, called Phasor Data Concentrators (PDC) [2] for further processing, the underlying network models are mostly unavailable or incomplete. This makes it challenging to get the most out of the synchronized measurements for fault detection and localization applications when the real-time network topology is unknown or not accurate. The problem of estimating the state of the power grid is often divided into two interrelated phases: the first is the state estimation in which the estimated value is the voltage at all buses across the network, and the second is topology processing and topology error detection, in which the breaker status is used to track the current topology of the grid, and to detect and correct the errors in the calculated topology. These two stages iterate, and the combined process is known as generalized state estimation [3]. With the measurements received from the μPMUs and when judiciously integrated with the topological processing of the visual image, the abnormal conditions in the distribution network (e.g., fault location, fault detection, etc.) can be better handled, further improving the network reliability, reducing the economic losses, and mitigating the electrical safety concerns.

A variety of research methods have been proposed to identify the power system topology from synchrophasor measurements, and several methods of external network modeling were discussed to implement online security analyses [4]. In [5], a

978-1-7281-8192-9/21/$31.00©2021 IEEE

Jacobian-based equivalent approach is used for detecting the electrical network topology changes in the external system. An approach which is a hybrid of power flow and state estimation is discussed in [6]. A method to capture the network topology changes based on an extended Ward equivalent is discussed in [7]. H. Singh, et al. [8] introduced a technique that estimates the status of suspect lines as part of the state estimation process. Focused on transmission systems with inaccurate parameters, an offline REI (radial, equivalent, independent) equivalent [9] is suggested to be built from a base-case condition and to be updated using online data [10]. In order to improve the accuracy of the topology detection process, the problem of using telemetry data to correct and adjust the transmission parameters are considered in [11]. The network topology estimation accuracy could vary greatly depending on the information injections at the non-PMU buses. If the injections at the non-PMU buses are zero, the estimates will be the true equivalent at the PMU buses. In [12], a least-square "model-free" approach is proposed to estimate the equivalent power system topology by calculating the load variations with limited observation at each bus. In [13], a method for visualizing PMU data by reducing the system to an equivalent model at the PMU buses is discussed, which assumes the electrical network topology is known; an equivalent procedure is performed to reduce the network to a Ward type equivalent at the PMU buses. Another useful visualization technique has been done by biplots introduced in [14]. S. V. Wiel, et al. [3] developed a greedy search algorithm to estimate the current topology of a power grid from phasor measurements. It studies the PMU placement at strategic points in a distribution system [15] to achieve a promising sensitivity to single-line outages. G. Cavraro, et al. [16] proposed a novel method for topology detection in distribution networks called the Time-Series Signature Verification for Topology Detection (TSV-Top). This approach relies on measurement time series from PMUs and performs the projection of actual voltage phasor patterns onto a library of signals associated with possible topology transitions of a given distribution network [17].

The above literature review revealed that most of the power grid topology identification and estimation tools are based on mathematical models, the majority of them assuming an electrical network topology first and then measure the collected data to compare the features and determine the accuracy of the previously assumed network topology. Such strategies are time-consuming, less accurate, and with practical limitations. On the contrary, there are more recent strategies leveraging machine learning advancements. Instead of accurately modeling the system, recent works have focused on training artificial neural networks to automatically recognize the electrical network topology and solve complex problems. In [18] and [19], two learning algorithms based on nodal voltage graphical models are introduced which can estimate the network topology under varying topological restrictions. D. Deka, et al. [20] developed a learning framework to reconstruct the radial operational structure of the distribution grid from synchronized voltage measurements across the network subject to the exogenous fluctuations in nodal power consumption. For the economic purposes, P. K. Ghosh, et al. [21] proposed a novel approach for complete system and fault observability using a minimum number of strategically-placed PMUs.

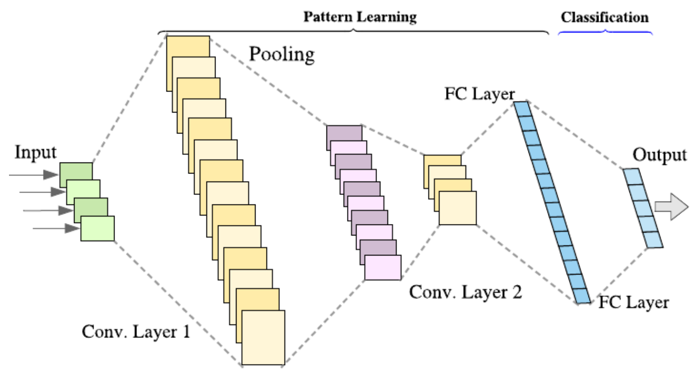

Fig. 1. A two-convolution-layer CNN structure.

To overcome the limitations of the traditional mathematical models, this paper proposes a machine learning framework for online identification of the distribution network topology. The neural network is trained using μPMU measurements across the network—voltage, current magnitudes, and their phase angles—and achieves the real-time network topology with high accuracy even under noise and missing entries in PMU measurements. The measured data are rearranged into a 2-D matrix (heatmap), where the suggested Convolutional Neural Network (CNN) [22] takes them as the input. The performance of the proposed algorithm is tested and verified in a radial three-phase unbalanced distribution network. The rest of the paper is structured as follows: Section II introduces the technical background on the design of the proposed convolutional neural network framework. Section III implements the suggested approach on the IEEE 34-node test feeder in MATLAB/Simulink platform and generates μPMU heatmaps under different system operating scenarios. Section IV presents the numerical studies and the network topology identification results in power distribution systems with noisy and missing measurements. Finally, we conclude the paper in Section V.

II. Technical Background

Within the family of neural networks, and to train the model with a grid-like topology such as images, deep CNN has been one of the greatest breakthroughs [23]–[25]. As it shows in Figure 1, the CNN structure consists of a convolutional layer, a pooling layer, and fully connected layers (FCs). When applied to single-label (multi-class) image classification, CNN can handle well-aligned images very well [26].

By definition, CNNs are simply neural networks that use convolution in place of the FC layer in that least one of their layers [27]. In general, the implementation of the convolution is actually the cross-correlation assessment and defined by

$$\boldsymbol{s}^p(m,n) = \sum_u \sum_v \sum_w \boldsymbol{I}^u(m+v, n+w)\boldsymbol{K}^p(v,w), \quad (1)$$

where $\boldsymbol{s}^p(m,n)$ is the output of the convolutional layer at position (m,n) and p-th channel, $\boldsymbol{I}^u$ is the u-th channel of the data volume, and $\boldsymbol{K}^p$ is the p-th convolutional kernel. A

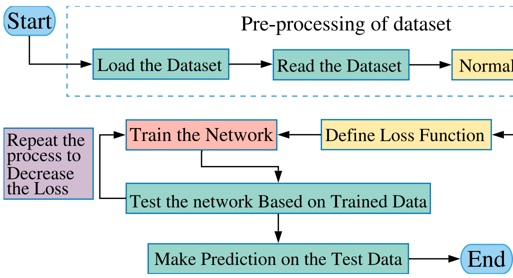

Fig. 2. CNN working flowchart [32].

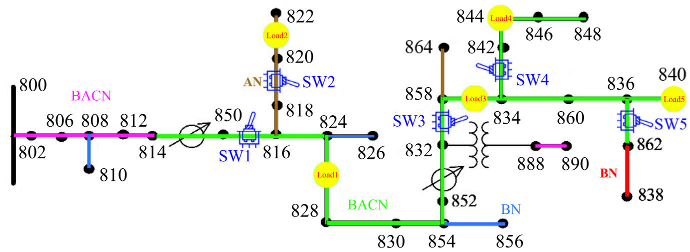

Fig. 3. IEEE 34-Node Test Feeder [33].

nonlinear activation function is applied to the output of the convolution output, and the final activations of neurons in a convolutional layer are

$$\boldsymbol{I}_l = \sigma(\boldsymbol{s}), \tag{2}$$

where $\boldsymbol{I}_l$ represents the output volume of the l-th layer, and $\sigma(\cdot)$ represents the non-linearity of the neurons. By stacking the convolutional layers, the abstraction capacity of the network increases [28].

The representations (outputs) of the last convolutional (Conv) layer are expanded to vectors and processed by the general fully-connected layers, which transform the representations with more nonlinearities and into spaces with different (higher or lower) dimensions. The final layer of a CNN usually reduces the dimensionality of the representations to the number of the classes; cross-entropy [29] is then employed to measure the "goodness" of the classification (Kullback-Leibler divergence between the predicted distribution and the target distribution). Finally, gradients of the cross-entropy loss function with respect to the parameters in the CNN are used to train the CNN by back-propagation.

In this paper, the proposed CNN for the heatmap classification in the IEEE 34-node test feeder has the following architecture: **Input(33×12)–Conv(32, 5×3)–Conv(32, 3×3)–FC(100)–FC(9)**. Note that the axes of the input heat-map are with different units; we, therefore, chose narrow kernel in the first Conv layer which could cover the 3-phase data in each group, and the stride of the convolution operation in the first layer is (3, 2)—other Conv layers' strides are (1, 1). This design processes the data in each group first, then combines the information of each group in the second Conv layer and the FC layer. Batch normalization [30] is used in each Conv layer. Dropout [31] is adopted in the last Conv layer and the FC layer to prevent over-fitting. Rectified Linear Unit (ReLU) were chosen as nonlinearities in the neural net.

The flowchart of the proposed CNN is shown in Figure 2, where the first step is to collect μPMU data and normalize them into per-units. Such data with their corresponding topologies are then inputted into the neural network, and the trained network learns to identify distribution grid topology with μPMU measurements. The CNN used cross-entropy as the loss function. Finally, we used additional μPMU measurements beyond the training set to verify the model accuracy.

This CNN architecture will be used as a building block in the proposed framework that identifies the power distribution network topology in real-time. The relation between the μPMU measurements and the CNN module will be detailed in the next section.

III. Proposed Methodology

For the general loopy grids, there may be multiple paths between two nodes. If power distribution grids are featured with minimum cycle length greater than three, the nodal voltages are sufficient for efficient topology estimation without additional assumptions on the system parameters. In contrast, the detection of line failures or status change using nodal voltages does not require any structural assumption on the network [18], [19]. As for the case of the IEEE 34-node test feeder, the minimum cycle length in a radial graph is considered to be infinite as it has no cycles by definition. Hence, using nodal μPMU measurements, such as voltage and current phasors, real-time network topology estimation on the IEEE 34-node feeder system is effectively viable.

The problem of detecting the network topology change can recast as a classification problem based on the heatmaps which are obtained by the μPMUs measurements. The conventional classification approaches often involve manually designed features like thresholds and signatures in each scenario. However, these approaches require the human expertise and the type of topology it can detect would be limited. E.g., a threshold may be suitable for a certain topology, but if one node becomes offline in the electrical network, the threshold would not work anymore. We propose an artificial neural network platform that can learn the features (representations) of the data automatically. The proposed framework for online power system topology identification is illustrated in Figure 4. The μPMUs data is first used for offline training of the pre-built CNN model. The trained model is then used for online identification of the power distribution network topology.

To gain a full observation of the IEEE 34-node test feeder as shown in Figure 3, we set 33 μPMUs on each node expect node 800 (substation bus). As it shows in Table I, we also added 5 breakers in order to generate different electrical network topologies for the CNN training dataset. To generate more scenarios under one topology, we marked 5 loads, so the μPMU data could be varying under different realization of the load demand. Different colors are here used to mark the phasing status, e.g., we used pink to mark the line 800-812 as BACN, meaning that it is a three-phase four-wire segment in the distribution grid. The proposed network topology identification approach is applied on a three-phase unbalanced electrical distribution system.

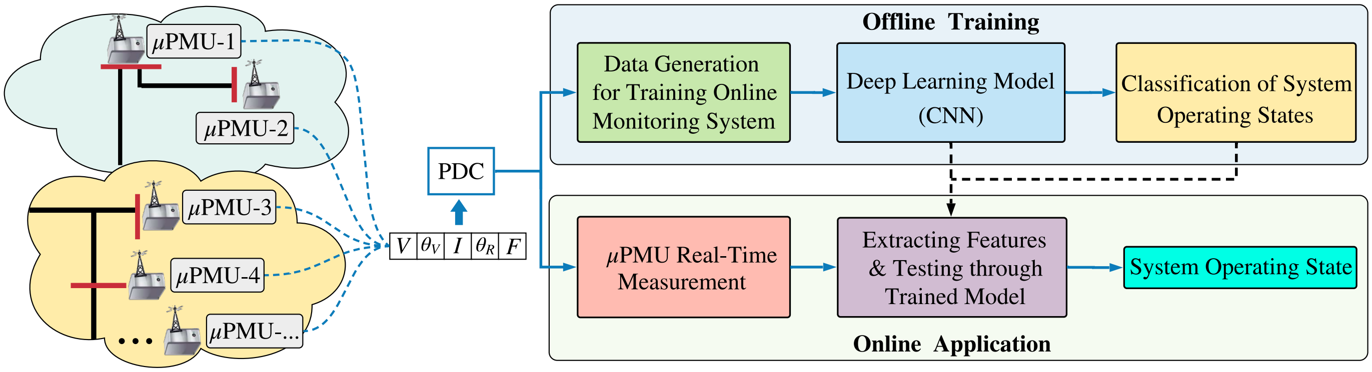

Fig. 4. The proposed framework for power system online topology identification. [34].

TABLE I
NETWORK MODEL SPECIFICATIONS AND VARIABLES

Breaker Name	**SW*1**	**SW2**	**SW3**	**SW4**	**SW5**
Breaker Location	850-816	818-820	832-858	834-842	836-862
Load Name	**Load1†**	**Load2†**	**Load3†**	**Load4**	**Load5**
Location	824-828	820-822	858-834	844	840

* SW: Breaker
† : Distributed load

Upon simulating each scenario in MATLAB Simulink environment, the resulting μPMU data will characterize a 33 by 12 heatmap matrix which contains three-phase voltage, three-phase voltage angle, three-phase current, and three-phase current angles. For the nodal measurements that contain single-phase or two-phase data, we let the remaining entries be zero. We line up the data in a heatmap format into four groups, process the data in each group individually, and finally integrate the information in each group together. A partially connected neural network is dedicated to the processing of these groups of heatmaps. In practice, a partially connected neural network is equivalent to a CNN. We design a CNN by carefully selecting the kernels in the first layer.

TABLE II
NETWORK TOPOLOGY REALIZATIONS WITH THE CORRESPONDING NUMBER OF GENERATED SCENARIOS.

Topology	**SW1**	**SW2**	**SW3**	**SW4**	**SW5**	**Number of Scenarios**
1	1	1	0	0	0	1600
2	1	0	1	0	0	2197
3	1	0	1	1	0	2401
4	1	0	1	1	1	2401
5	1	1	1	0	0	2401
6	1	1	1	0	1	2401
7	1	1	1	1	0	3125
8	1	1	1	1	1	3125

Table II shows the general configuration of all simulated scenarios. Take Topology 1 as an example, in which Breaker 1 and Breaker 2 are closed (status=1), Breaker 3 is open (status=0). For each network topology configuration, we generate different scenarios by adjusting the loads values. To keep the balance and avoid generating redundant training data for the CNN [35], keen considerations were taken into account in generating different load scenarios. We let the load change amplitude uniformly distributed between 95% to 105% of the rated demand at each load point. In Topology 1, only Load 1 and Load 2 are served through the connected distribution line and it is not necessary to adjust the remaining three load points for scenario generation. Assuming each load has 40 possible amplitudes in the constrained range above, the total number of scenarios is found 40^2 in this case, i.e., 1600. Under the network Topology 8, all five loads are being served in the distribution grid, and as each one is characterized with 5 possible amplitudes for the training process, there are 5^5, i.e., 3125, number of scenarios generated. The total number of generated scenarios that contribute to the training dataset is found 19651. As mentioned before, the simulation of each scenario results in a heatmap (see Figure 5 for a heatmap example); these heatmaps are used as the inputs to the CNN.

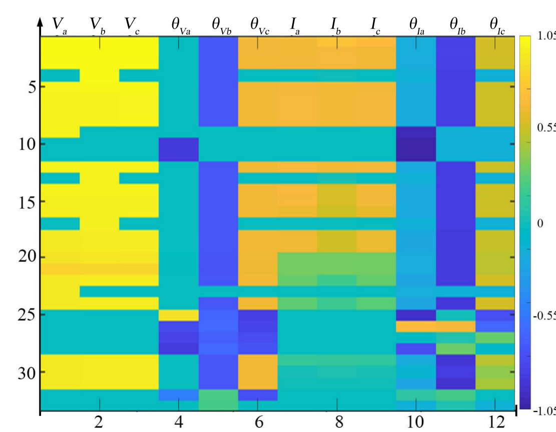

Fig. 5. A heatmap example of the generated μPMU measurement data sample.

IV. NUMERICAL CASE STUDY

A. Data Generation and Preprocessing

For the case of IEEE 34-node test feeder, the obtained data are 33 by 12 heatmaps which stand by 33 μPMUs data points (rows) and are consisted of the voltage, current, and phase angle information (columns) from the μPM measurements.

These μPMU measurements are obtained under full observations of all scenarios in the power network as mentioned in Section III and generated in the MATLAB environment. We built the entire test feeder system in Simulink, and set all block parameters according to the data provided in [33]. In order to save the simulation time, we here used a technique of parallel simulation operation in Simulink [36], which could run simulations of multiple scenarios simultaneously. When the μPMU measurements are captured, all scenarios were classified into three folders: training dataset, testing dataset, and validation dataset, where each folder contains 8 topologies as presented in Table II.

In order to facilitate the CNN computation, we normalized the columns of the heatmaps (such as per-unit voltage and current values) through a zero-mean and unity variance distribution. We also randomly separate 80% of the total simulation outputs as the training dataset, 10% for testing dataset, and 10% for validation dataset.

B. Results Analysis

In order to test the performance of the proposed framework, we conduct experiments that are closer to the realistic situations. In the first group experiments, we applied different interferences in the dataset. The interferences include both the noise and the missing data. The accuracy of the proposed topology identification framework in the conducted experiments is shown in Table III. Wherein, one epoch of training means every sample in the training dataset is used in the training of the CNN once, and the SNR refers to the Signal-to-Noise Ratio. One can see that as the number of training epochs increases, the accuracy also increases. The "Validation Accuracy" in the Table III corresponds to the condition when the neural network was trained by the **training dataset**, and the accuracy of identifying the electrical network topology is assessed using the **validation dataset** which is included in the training dataset; and the "Prediction Accuracy" refers to the condition when the neural network was trained by the testing dataset, and the accuracy of the network topology identification is assessed using the **testing dataset** which is excluded in the training dataset.

As one can see, the accuracy of the proposed electrical network topology identification scheme was never found lower than 95%. This is because for one epoch training, all the testing data were included in the training dataset. For example, for full measurement containing 10dB SNR, and when the number of epochs is 20, the neural network was trained by the training dataset contained 10dB SNR, leading to an identification accuracy of 99.9%. Hence, the prediction accuracy cloud be achieved high as long as the neural network was trained well. In this situation, the prediction is actually called identification, because all scenarios are already known, taking advantage of a full observation. Additionally, when all measurements are available and the SNR is greater then 20dB, the proposed CNN can identify the system topology very accurately (the smaller the SNR, the greater the noise). Moreover, for well trained neural network, the more missing data in the training dataset, the greater the positive impact on the accuracy of the prediction engine.

TABLE III
THE IDENTIFICATION ACCURACY OF THE INTERFERED DATASET

Test Scenarios	Interference SNR (dB)	Number of Epochs	Best Validation Accuracy (%)	Best Prediction Accuracy (%)
Full Measurement	10	5	98.81	98.72
Full Measurement	10	10	98.81	98.81
Full Measurement	10	20	99.86	99.95
Full Measurement	20	20	100	100
Full Measurement	20-50*	20	100	100
Missing One Data	-	20	100	100
Missing Two Data	-	20	99.91	99.95
Missing One Data	20	20	100	99.95
Missing Two Data	20	20	100	99.95
Missing One Data	10	20	99.95	99.82
Missing Two Data	10	20	98.86	98.99
Missing Two Data	20-50*	20	100	100

*: the intensity of SNR is randomly selected in the range and applied on each data sample.

TABLE IV
THE IDENTIFICATION ACCURACY (%) BY TRAINING AND TESTING THE CNN IN DIFFERENT EXTENTS OF INTERFERENCES.

Training Data / Test Data	10dB SNR	Missing One Data and 10dB SNR	Missing Two Data and 10dB SNR
40dB SNR	71.88	80.75	94.00
Missing One Data 40dB SNR	72.75	80.63	93.62
Missing Two Data 40dB SNR	72.13	82.13	94.50

In the second group of experiments, the training and testing data are interfered with at different levels. We trained three CNNs by applying the datasets (i) containing 10dB SNR, (ii) containing one missing data with 10dB SNR, and (iii) containing two missing data with 10dB SNR respectively. The models are individually tested each with 800 samples (i.e., we generated 100 μPMU data for each network topology) but with 40dB SNR. The network topology identification accuracy is shown in Table IV. Note that when generating the training dataset, the data are interfered by taking out the missing entries first, then adding noise. On the contrary, when testing the model, the data were added with noise first and then the missing entries were added. The test data were beyond the training dataset, which means the neural network should identify the system topology by estimating from the unknown input. For example, for μPMU data which contained 40dB SNR, using the network which was trained well by the dataset containing 10dB SNR to estimate the system topology will result in an overall identification accuracy of 71.88%. All these 9 cases achieve the accuracy greater than 70%. One can see that using the same testing data, if the training data has imperfections such as missing, and/or outlier values, interferences, but under a certain level, the trained neural network can provide more accurate results. This is because the training data's imperfections or complications can make the neural network become more versatile, and thus, the trained network would perform better. In all, the trained CNN under the greatest level of interference achieves a satisfactory topology identification accuracy, implying the proposed CNN has a very good capacity of generalizing the trained data to unseen inputs.

V. Conclusion

This paper presented a deep learning framework for online detection of power distribution system topology. The proposed framework can handle missing measurements under unbalanced operating states. The experiments show that the proposed CNN not only handles the data with the same level of interference (noise and missing measurements), but also has the capacity of estimating the interfered data which has different distributions from the training examples. Numerical experiments proved that our trained network can accurately identify the network topology corresponding to the observed data beyond the training dataset.

CNN was chosen as the machine learning engine in the proposed framework owing to its simplicity and efficiency for handling large-size three-phase data. Other machine learning models such as Autoencoder (AE) and Capsule Neural Network (CapsNet) could be potential candidates as well. Our future research will continue to explore utilizing more μPMU data in larger systems, such as the IEEE 123-bus distribution system, and different machine learning models for detecting the topology changes in power distribution systems.

References

[1] L. Hao, B. Tianshu, X. Quan, C. Shijie, S. Bonian, and X. Ancheng, "Scheme and prospect of high-precision synchrophasor measurement technology for distribution network," *Automation of Electric Power Systems*, [Online] Available at: http://doi.org/10.7500/AEPS20190814002, Accessed:2020.

[2] "openPDC," Phasor Data Concentrator, [Online] Available at: https://www.openpdc.codeplex.com.

[3] S. V. Wiel, R. Bent, E. Casleton, and E. Lawrence, "Identification of topology changes in power grids using phasor measurements," *Applied Stochastic Models in Business and Industry*, vol. 30, no. 6, pp. 740–752, 2014.

[4] F. F. Wu and A. J. Monticelli, "Critical review of external network modelling for online security analysis," *International Journal of Electrical Power and Energy Systems*, vol. 5, pp. 222–235, 1983.

[5] L. F. Alvarado, "Determination of external system topology errors," *Power Engineering Review, IEEE*, pp. 34–35, 1981.

[6] C. N. Lu, K. C. Liu, and S. Vemuri, "An external network modeling approach for online security analysis," *IEEE Transactions on Power Systems*, vol. 5, no. 2, pp. 565–573, 1990.

[7] K. L. Lo, L. J. Peng, J. F. Macqueen, A. O. Ekwue, and D. T. Y. Cheng, "An extended ward equivalent approach for power system security assessment," *Electric Power Systems Research*, vol. 42, pp. 181–188, 1997.

[8] H. Singh and F. L. Alvarado, "Network topology determination using least absolute value state estimation," *IEEE Transactions on Power Systems*, vol. 10, no. 3, pp. 1159–1165, 1995.

[9] P. Dimo, "Nodal analysis of power systems," 1975.

[10] G. Irisarri, A. M. Sasson, and J. F. Dopazo, "Real-time external system equivalent for on-line contingency analysis," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, no. 6, pp. 2153–2171, 1979.

[11] A. S. Debs, "Estimation of steady-state power system model parameters," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 5, pp. 1260–1268, 1974.

[12] K. M. Rogers, R. D. Spadoni, and T. J. Overbye, "Identification of power system topology from synchrophasor data," in *2011 IEEE/PES Power Systems Conference and Exposition*. IEEE, 2011, pp. 1–8.

[13] C. M. Davis, J. E. Tate, and T. J. Overbye, "Wide area phasor data visualization," in *2007 39th North American Power Symposium*. IEEE, 2007, pp. 246–252.

[14] K. R. Gabriel, "The biplot graphical display of matrices with application to principal component analysis," *Biometrika*, vol. 58, no. 3, p. 453–467, 1971.

[15] E. Lawrence, S. V. Wiel, and R. Bent, "Model bank state estimation for power grids using importance sampling," *Technometrics*, vol. 55, no. 4, pp. 426–435, 2013.

[16] G. Cavraro and R. Arghandeh, "Power distribution network topology detection with time-series signature verification method," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3500–3509, 2017.

[17] G. Cavraro, R. Arghandeh, K. Poolla, and A. Von Meier, "Data-driven approach for distribution network topology detection," in *2015 IEEE power & energy society general meeting*. IEEE, 2015, pp. 1–5.

[18] D. Deka, S. Talukdar, M. Chertkov, and M. Salapaka, "Graphical models in loopy distribution grids: Topology estimation, change detection and limitation," *arXiv*, pp. arXiv–1905, 2019.

[19] ——, "Graphical models in meshed distribution grids: Topology estimation, change detection & limitations," *IEEE Transactions on Smart Grid*, 2020.

[20] D. Deka, M. Chertkov, and S. Backhaus, "Topology estimation using graphical models in multi-phase power distribution grids," *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 1663–1673, 2019.

[21] P. K. Ghosh and A. Tahabilder, "Optimal pmu placement for complete system observability and fault observability using graph theory," in *2017 International Electrical Engineering Congress (iEECON)*. IEEE, 2017, pp. 1–4.

[22] S. Wang, P. Dehghanian, L. Li, and B. Wang, "A machine learning approach to detection of geomagnetically induced currents in power grids," *IEEE Transactions on Industry Applications*, vol. 56, no. 2, pp. 1098–1106, 2019.

[23] Y. LeCun, Y. Bengio *et al.*, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.

[24] S. Wang, P. Dehghanian, and L. Li, "Power grid online surveillance through pmu-embedded convolutional neural networks," *IEEE Transactions on Industry Applications*, vol. 56, no. 2, pp. 1146–1155, 2019.

[25] L. Li, M. Doroslovački, and M. H. Loew, "Discriminant analysis deep neural networks," in *2019 53rd annual conference on information sciences and systems (CISS)*. IEEE, 2019, pp. 1–6.

[26] ——, "Loss functions forcing cluster separations for multi-class classification using deep neural networks," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2019, pp. 2106–2110.

[27] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[28] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 8, pp. 1798–1828, 2013.

[29] J. S. Bridle, "Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition," in *Neurocomputing*. Springer, 1990, pp. 227–236.

[30] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015.

[31] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[32] A. Rao, "Convolutional neural network tutorial (cnn) – developing an image classifier in python using tensorflow," edureka.co, [Online] Available at: https://www.edureka.co/blog/convolutional-neural-network, Accessed:2020.

[33] "34-bus feeder case," IEEE PES, [Online] Available at: https://site.ieee.org/pes-testfeeders/resources, Accessed:2010.

[34] B. Shinde, S. Wang, P. Dehghanian, and M. Babakmehr, "Real-time detection of critical generators in power systems: A deep learning hcp approach," in *2020 IEEE Texas Power and Energy Conference (TPEC)*. IEEE, 2020, pp. 1–6.

[35] A. Santacruz, "Why it is important to work with a balanced classification dataset," amsantac.co, [Online] Available at: http://amsantac.co/blog/en/2016/09/20/balanced-image-classification-r.html, Accessed:2016.

[36] "Run parallel simulations," MathWorks, [Online] Available at: https://www.mathworks.com/help/simulink/ug/running-parallel-simulations.html, Accessed:2020.