Convolutional Neural Network-Assisted Fault Detection and Location Using Few PMUs

Tuna Yildiz
Department of Electrical and Computer Engineering
Northeastern University
Boston, U.S.A.
yildiz.tu@northeastern.edu

Ali Abur
Department of Electrical and Computer Engineering
Northeastern University
Boston, U.S.A.
a.abur@northeastern.edu

***Abstract*—Detection and identification of faults in large distribution systems with limited metering continue to pose significant challenges for system operators. The prevalence of installed phasor measurement units (PMUs) in power systems, provides an invaluable resource for fault detection and location. Therefore, there is a rich literature on methods that leverage the benefits of these PMUs for the purpose of event detection. However, a notable proportion of the proposed solutions assume the presence of one PMU at every bus, or at a large percentage of buses that will make the system observable. Despite concerted efforts to substantially increase the number of installed PMUs, majority of utilities have yet to attain comprehensive PMU coverage at the terminals of every transmission line. Thus, many of the proposed methods cannot be practically applied to existing power systems with limited number of installed PMUs. Consequently, the objective of this paper is to develop a method, aided by the application of Convolutional Neural Networks, to identify and precisely locate faults within a system, even when only a minimal number of Phasor Measurement Units (PMUs) are present.**

***Index Terms*—Fault Detection, Fault Location, CNN aid Fault Detection, Strategical PMU Placement, Ordinary Least Square (OLS)**

I. Introduction

Over the recent years, there has been a steady increase in the number of installations of Phasor Measurement Units (PMUs) in power systems. This surge in PMU installations can be attributed to the numerous applications that were developed to leverage the increasing numbers of these devices. These methods effectively address a variety of issues prevalent within power systems, including fault, line outage and other anomaly detection. Nevertheless, overall PMU numbers in most power grids still remain limited due to costs associated with their installation and maintenance. Consequently, a number of studies are motivated by addressing the challenges related to lack of sufficient PMUs for various applications.

Among the proposed methods, emergence of data-driven approaches has become particularly prominent, gaining substantial popularity in the detection and localization of faults as outlined in [1], [2]. This can be largely attributed to the advancements in technology that have fostered their development. Such methodologies propose various distinctive strategies, utilizing the unique structures of Artificial Neural Networks (ANN). The most commonly recognized models within this domain include the Graph Convolutional Neural Network (GCNN) [3] and the widely adopted Multi-Layer Perceptron (MLP) [4]–[6]. These models exemplify the diverse structures of ANNs that are designed for fault detection and localization in power systems. However, majority of these approaches suggest solutions predicated on the assumption that a PMU is either present at each bus, or is placed at a large proportion of buses, a condition which would yield the system observable. Thus, many of the proposed methods cannot be practically applied to existing power systems with limited number of installed PMUs.

In addition to the data-driven approaches, recently, an alternative formulation which is based on the so called "sparse estimation" method is shown to solve the fault location problem with very few PMUs [7], [8]. For an n-bus system with p PMUs, the approach makes use of a set of under-determined set of equations:

$$[Z_p] \cdot [\Delta I] = [\Delta V_p] \tag{1}$$

where,
$[Z_p]$ $(p \times n)$ contains p rows of the bus impedance matrix,
$[\Delta V_p]$ $(p \times 1)$ contains the changes in bus voltage phasors from pre-fault to post-fault conditions measured at p buses with PMU measurements,
$[\Delta I]$ $(n \times 1)$ contains the change in bus current injections due to the changes in bus voltages. This vector is expected to be sparse having only non-zeros corresponding to the terminal buses of the unknown faulted branch.

Applying the method of Least Absolute Shrinkage and Selection Operator (LASSO) [9] to the solution of (1), ΔI can be solved as shown in [7], [8] provided that there are sufficient number of strategically placed PMUs. The required number of PMUs may be large for certain systems and the performance of LASSO may deteriorate if PMU numbers are reduced.

Additionally, the location of PMUs is important for system operators in the execution of a range of applications, including but not limited to, State Estimation (SE), Line Outage

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office Award Number DE-EE0009356.

Detection, and Fault Detection. Consequently, a breadth of methods designed to optimally place PMUs for various types of applications have been developed in the literature [10]–[12]. However, most of these methods are based on the assumption of an abundant availability of PMUs, which are then utilized to ensure the observability of the system. Furthermore, when considering fault detection and localization, instead of considering network observability, PMUs should be placed in a strategic manner to accurately capture faults system wide. Thus, this paper introduces an approach to PMU placement, explicitly designed to identify the best locations for these units.

Once all PMU locations are determined, an Artificial Neural Network (ANN) is used to assist and improve the sparse estimation problem solution when there are few installed PMUs. In order to accomplish this objective a Convolutional Neural Network (CNN) is employed as the chosen Artificial Neural Network (ANN) model. Furthermore, the approach is comparatively evaluated with respect to the commonly used method of the Multi-Layer Perceptron (MLP), to demonstrate the performance disparities between them. In both methodologies, the chosen input features consist of the changes in voltage phasors at the fault instant. Regarding the output of the model, one-hot encoding is employed for all transmission lines to indicate the fault location.

To generate a training dataset with the mentioned input and output structure, a 196-bus three-phase system is used. The pre-fault and post-fault network bus voltages are obtained in three-phase for a diverse range of faults that are introduced on each line in the system. Then, both models are trained using the created dataset and tested using unseen data to assess their performance in correctly identifying the faulted line. Once the faulted line is determined via ANN, the "Ordinary Least Squares (OLS)" method detailed in [13] is employed to accurately identify the fault location along the determined branch. Finally, the impact of the number of PMU measurements on the performance of the proposed approach is also investigated for different fault scenarios.

Further details of the processes mentioned above is given in following sections.

II. PMU Placement Strategy

Prior to the creation of the training and test set, it is important to determine PMU locations. In order to achieve this, it is observed that the positioning of PMUs around the leaf nodes significantly improves the accuracy in capturing the impacts of any fault occurring in the system. Consequently, to ensure a strategic deployment of the PMUs around the leaf nodes, the following strategic PMU placement scheme is employed.

Consider the following optimization problem for an n-bus power system with L branches:

$$\begin{aligned} \min \quad & x \\ \text{s.t.} \quad & Cx \geq 2 \\ & x \in \{0,1\} \end{aligned} \tag{2}$$

where:
x is an $n \times 1$ binary vector with $x_i = 1$ if a PMU is placed at bus i, 0 otherwise.
C is the coefficient matrix formed following the procedure described below.

Inverting the bus admittance matrix (Y_{bus}), bus impedance matrix (Z_{bus}) can be obtained. Columns of Z_{bus}, corresponding to two terminal buses k and m of branch $k-m$ are then extracted to form a $n \times 2$ sub-matrix c_k. The LU factorization is subsequently applied to the sub-matrix c_k, resulting in the formation of matrices $L(n \times 2), U(2 \times 2)$, and $P(n \times n)$. Following this procedure, one of the terminal buses of a branch is fixed. Then, linearly independent solution combinations are determined by examining the non-zero entries in the second column of the L matrix. Upon deriving all possible combinations, the respective columns of the "Buses" corresponding to these combinations will be set to "1" in the C matrix for the row of specified branch. This procedure is repeated for each branch $k = 1, \ldots, M$ and will be used to form C.

The constraints for the integer programming problem are encoded in C. Each row of C corresponds to a branch, and each column to a bus in the power system. Then the integer linear programming problem of (2) is solved to find the optimal placement of PMUs.

Algorithm 1 Strategical PMU Placement

```
Z_bus = inv(Y_bus)
C = zeros(num_branch,num_bus)
while counter ≤ num_branch do
    Z_l = Z(:,branch_list(counter,:))
    [L, U, P] = lu(Z_l)
    cl0 = find(P(1,:) == 1)
    row = find(L(:,2) > 0)
    cl1 = find(P(row,:) == 1)
    C(i,[cl0, cl1]) = 1
end while
ub = ones(num_bus,1);
lb = zeros(num_bus,1);
b = ones(num_branch,1) * 2;
intcon = num_bus;
intlinprog(ones(num_bus,1),intcon,-1*C,-1*b,[],[],lb,ub);
found_pmu_locations = find(x = 1);
```

where,

- P is defined as the permutation matrix, which represents the row interchanges executed during the factorization procedure ($A = P'LU$).

The objective of the placement algorithm is to identify strategic locations for PMUs around the leaf nodes within the system and enable capturing of the impact of any fault occurring in the system.

III. Training and Test Data Set Creation

In order to create the training and test data set, first the fault is simulated. Note that, when a fault occurs along a

transmission line, its impact will be observed at all system buses as a voltage transient. In the hypothetical (not realistic) scenario where none of the protective relays operate, the fault transients will diminish gradually, and bus voltages will settle at a new post fault steady state operating point. Therefore, all bus voltages will change with respect to their pre fault steady state conditions [7]. Faults can be visualized as shown in Fig. 1.

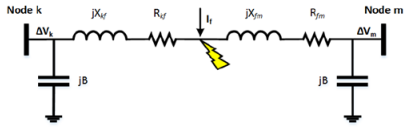

Fig. 1: Fault on branch k-m.

It can be shown that the fault current can be replaced by equivalent virtual current injections at the faulted line terminal buses as shown in Fig. 2.

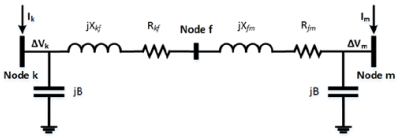

Fig. 2: Virtual Current Injections on adjacent buses for faulted branch k-m.

Moreover, by adjusting the magnitude and angular differences between these virtual current injections, it is possible to effectively modify the fault location and impedance. Consequently, the above-described modifications are employed to produce a comprehensive training data set.

This set includes a wide variety of cases, with a total of 420 distinct fault scenarios simulated for each line within the three-phase 196 Bus distribution system given in Fig. 3.

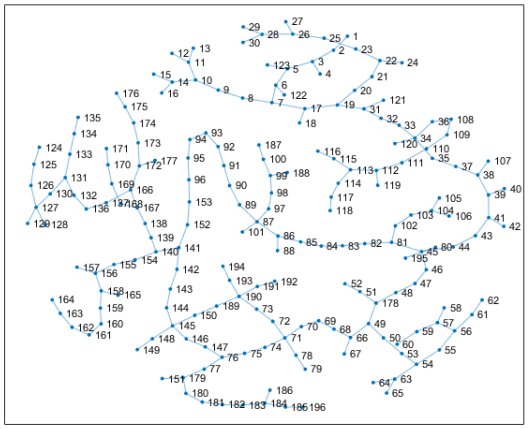

Fig. 3: Three-phase 196 Bus system.

To simulate the aforementioned fault scenarios, a three-phase power flow is employed by introducing virtual current injections as constant current sources at each terminal bus of a faulted branch. Then, for each branch within the system these scenarios are systematically repeated to obtain voltage magnitudes and voltage phase angles considering the given fault scenarios. Once all the pre and post fault voltage phasors are obtained, the voltage difference is calculated by using the $\Delta V = V_{post} - V_{pre}$ for these cases. Then, the strategic PMU placement algorithm is performed for the 196 Bus system.

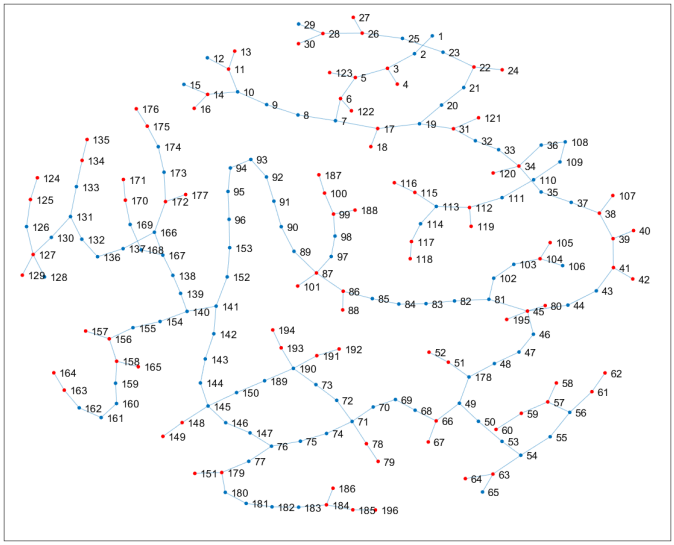

Fig. 4: Three-phase 196 Bus system with 50 PMU.

As shown in Fig. 4, a total of 50 PMUs ($50/196 \approx 25\%$) are strategically placed throughout the 196 Bus system. This arrangement is done based on the assumption that each PMU device possesses a single current channel, thereby enabling the calculation of an adjacent bus relative to the initial location of each PMU device. In Fig. 4 the PMU locations are marked with green color.

Then voltage phasors corresponding to those PMUs as a result of aforementioned scenarios are gathered from the power flow results and allocated to the data set. Subsequently, to train and evaluate the ANN during the training process, 80% of the data are reserved for training data set and 20% of the data are allocated to the validation data set.

IV. Convolution Neural Network Model

Convolutional Neural Networks (CNNs) offer a distinct advantage over traditional Multi-Layer Perceptrons (MLPs) when it comes to the classification problem, primarily due to their ability to effectively identify and exploit local patterns within the data [14]. This is facilitated through the use of convolutional layers that employ local filters or kernels, which slide across the input data, extracting valuable features from smaller sections. As a result, CNNs can effectively learn and recognize spatial hierarchies within the data, enabling them to detect patterns even in large datasets. Moreover, CNNs are well-suited for handling variable-length input data, since their convolutional layers can be applied to inputs of different lengths, automatically adapting to the size of the data. The

structure of the CNN is given in Fig. 5. The details of both MLP and CNN can be found in [15].

In contrast, MLPs process input data in a global manner, as each neuron in a given layer is fully connected to all neurons in the previous layer. This results in a considerable increase in the number of parameters to be trained, which can lead to longer training times and the risk of overfitting. Additionally, MLPs lack the inherent ability to recognize distinctive relationships within the data, making them less efficient at extracting and exploiting local patterns. Consequently, CNNs outperform MLPs in tasks where the identification and preservation of local patterns are crucial for accurate predictions. By leveraging their unique architecture and local feature extraction capabilities, CNNs provide a more efficient and powerful tool for 1D data analysis compared to traditional MLPs.

Therefore, building upon the findings, in this paper, it is aimed to improve the performance of the fault detection method by using the Convolutional Neural Network (CNN) as the Artificial Neural Network (ANN) architecture. This modification is expected to yield improved accuracy and effectiveness in detecting and pinpointing faults occurred within the power system.

The Convolutional Neural Network (CNN) model used in this study consists of multiple layers, including convolutional layers, activation layers, and fully connected layers. The architecture of the CNN model is summarised as follows:

1) Input layer: This layer accepts input data with dimensions $(X_{\text{train}}, 1)$.
   where,
   - Dimension of $(X_{\text{train}}, 1)$ is equal to the total number of PMUs within the system.
2) First convolutional block:
   - Conv1D layer with 64 filters,
   - Leaky ReLU activation function with a slope coefficient of α.
   - Conv1D layer with 64 filters.
   - Leaky ReLU activation function with a slope coefficient of α.
   - MaxPooling1D layer with a pool size of 2 and a stride of 1.
3) Second convolutional block:
   - Conv1D layer with 32 filters.
   - Leaky ReLU activation function with a slope coefficient of α.
   - Conv1D layer with 32 filters.
   - Leaky ReLU activation function with a slope coefficient of α.
   - MaxPooling1D layer with a pool size of 2 and a stride of 1.
4) Third convolutional block:
   - Conv1D layer with 16 filters.
   - Leaky ReLU activation function with a slope coefficient of α.
   - Conv1D layer with 16 filters.
   - Leaky ReLU activation function with a slope coefficient of α.
   - MaxPooling1D layer with a pool size of 2 and a stride of 1.
5) Flatten layer to convert the feature maps into a one-dimensional vector.
6) Dense layer with the same number of neurons as the dimensions of y_{train} and a Leaky ReLU activation function with a slope coefficient of α.
7) Output layer: Dense layer with the same number of neurons as the dimensions of y_{train} and a softmax activation function.
   where,
   - Dimension of y_{train} is equal to the number of branches where faulted branch is marked with "1" while others are kept as "0".

Upon finalizing the structure of the model, hyperparameters, including the learning rate, batch size, among others, are adjusted in accordance with the observed performance of the model. Furthermore, a learning schedule strategy is implemented during the training process. For every 30 epochs in the training, the learning rate is reduced by a factor of $\frac{1}{10}$. This approach seeks to achieve more effective and accurate results. Then, utilizing the training dataset acquired as described in Section III, the CNN model is trained and validated using the test dataset. A comprehensive analysis of the training outcomes can be found in Section VI.

V. Ordinary Least Squares Estimation

Upon identification of the fault location utilizing the trained model, the subsequent step is the estimation of virtual currents via the Ordinary Least Squares Estimation (OLS) method. The reason behind employing the OLS estimation is to determine the virtual fault currents and pinpoint the fault locations with greater precision. To achieve this, initially, all variables associated with the phasor domain are transformed into the sequence domain yielding:

$$\begin{bmatrix} Y_{seq,ee} & Y_{seq,ei} \\ Y_{seq,ie} & Y_{seq,ii} \end{bmatrix} * \begin{bmatrix} \Delta V_{seq,e} \\ \Delta V_{seq,i} \end{bmatrix} = \begin{bmatrix} \Delta I_{seq,e} \\ \Delta I_{seq,i} \end{bmatrix} \tag{3}$$

- subscript i refers to buses with PMUs and,
- subscript e refers to buses without PMUs,
- ΔV is the voltage differences between post and pre outage event,
- ΔI is the virtual current injection vector,

Then QR transformation is applied to eliminate ill-conditioning caused by the collinearity:

$$R_{seq,ii} * \Delta V_{seq,ii} = [Q^T_{seq,ei} Q^T_{seq,ii}] * \begin{bmatrix} \Delta I_{seq,ee} \\ \Delta I_{seq,ii} \end{bmatrix} \tag{4}$$

where,

- Subscript seq refers to "zero, positive, and negative" sequences,

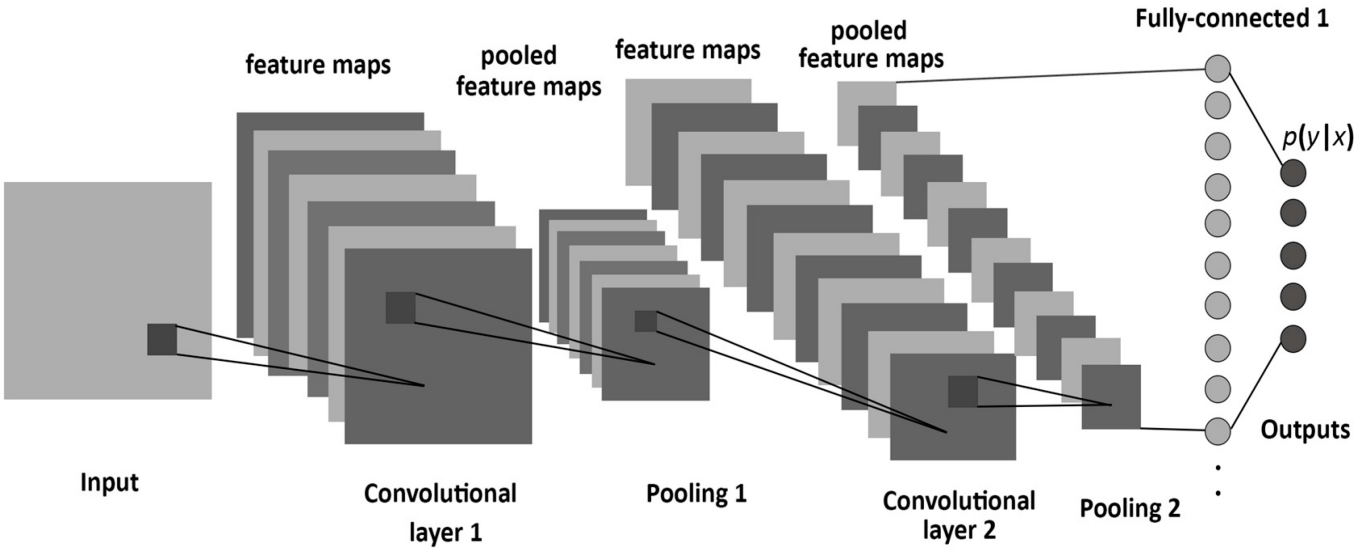

Fig. 5: The Architecture of CNN.

Then, each sequence component can be estimated as follows:

$$[Q^{fault}_{bus1,seq}\ Q^{fault}_{bus2,seq}] * \begin{bmatrix} \Delta I^{fault}_{bus1,seq} \\ \Delta I^{fault}_{bus2,seq} \end{bmatrix} \approx \Delta V'_{seq} \tag{5}$$

where,

- $Q^{fault}_{bus1,seq}$ and $Q^{fault}_{bus2,seq}$ denotes the reduced $[Q^T_{ei}\ Q^T_{ii}]$ matrix for the corresponding sequence, whose columns correspond to the terminal buses of the faulted branch and whose rows correspond to the buses with PMU measurements,
- $\Delta I^{fault}_{bus1,seq}$ and $\Delta I^{fault}_{bus2,seq}$, are the entries of the 2×1 unknown current vector for the corresponding sequence, at the terminal buses of the faulted branch,
- $\Delta V'_{seq} = R_{ii,seq} * \Delta V_{ii,seq}$,

An optimal solution of (5) can be found by a standard Ordinary Least Squares (OLS) algorithm [13]:

$$\hat{\beta} = (X^T X)^{-1} X^T y \tag{6}$$

where:

$$\beta = \begin{bmatrix} \Delta I^{fault}_{bus1,seq} \\ \Delta I^{fault}_{bus2,seq} \end{bmatrix},$$

$$y = \Delta V'_{seq},$$

$$X = [Q^{fault}_{bus1,seq}\ Q^{fault}_{bus2,seq}].$$

Then, the computed virtual sequence currents are transformed back to the phase domain and the fault distance from each terminal bus can be calculated as:

$$FaultLocation(m) = \frac{|I_{fk}|}{|I_{fk}| + |I_{fm}|} * 100\% \tag{7}$$

$$FaultLocation(k) = \frac{|I_{fm}|}{|I_{fk}| + |I_{fm}|} * 100\% \tag{8}$$

where,

- I_{fm} is the virtual fault current at bus m,
- I_{fk} is the virtual fault current at bus k.

VI. Test Results

Proposed methodology is comparatively evaluated with respect to the prevalent approach (MLP) by employing the Keras library in Python for model construction [16]. Subsequently, the training and test datasets, as obtained in Section III are used to train both the CNN and MLP networks under the assumption that the PMUs are placed as shown in Fig. 4.The outcomes of the training process are quantified in both loss and accuracy graphs as shown in Fig. 6 and 7, respectively.

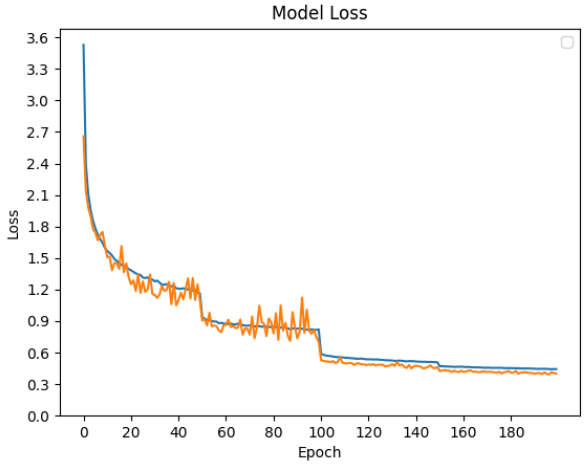

Fig. 6: Loss graph of MLP network with 50 PMU.

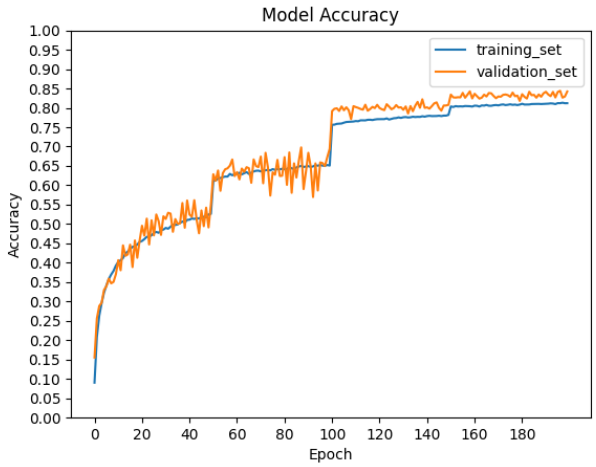

Fig. 7: Accuracy graph of MLP network with 50 PMU.

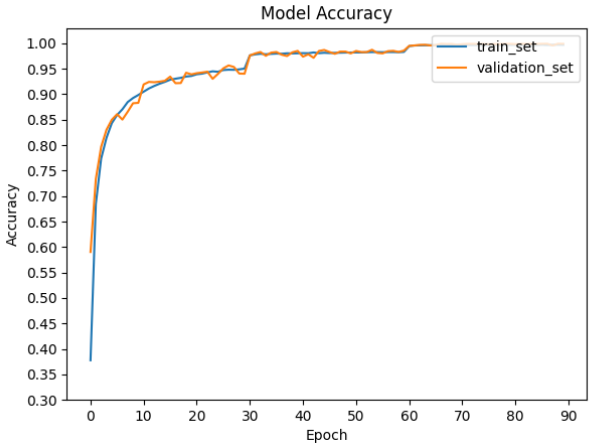

Fig. 9: Accuracy graph of CNN network with 50 PMU.

Upon the strategic placement of 50 Phasor Measurement Units (PMUs) throughout the system, the Multi-Layer Perceptron (MLP) structure yields an accuracy of approximately 85%. Moreover, the training process necessitates close to 200 epochs to achieve this accuracy level. In contrast, when the proposed method is trained with identical PMU locations, the training process yields the loss and accuracy charts given in Fig. 8 and 9 respectively.

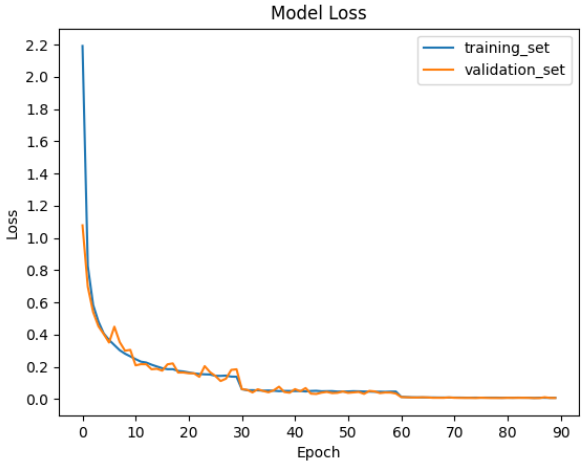

Fig. 8: Loss graph of CNN network with 50 PMU.

As evident from Fig. 8 and 9, use of the CNN network architecture significantly improves the accuracy from 85% to 99.9% for the identical data set by using only 50 PMUs. Furthermore, utilizing the CNN architecture substantially reduces the number of epochs, thereby increasing time efficiency of the training process.

Upon completion of the model training, in order to further evaluate its performance, new fault cases comprising unseen fault locations and magnitudes are produced. Subsequently, the trained model is performed to identify the locations of these faults, which were previously unseen by the trained model. The results of detecting fault locations using both proposed method and MLP are given in Table I.

After predicting the locations using the trained model, the OLS estimation method detailed in Section V is employed to determine both the fault type and its precise location on the line. The results of the OLS process, subsequent to identifying fault locations for previously unseen fault cases using the proposed CNN model, are presented in Table II.

TABLE I: The performance comparison of proposed method and MLP for unseen fault cases.

Total Fault Cases	**Utilized PMU Number**	**Used Method**	**Detected Fault Locations**	**Performance**
4116	50	CNN	4056	98.54%
4116	50	MLP	3085	74.95%

In Table I, results for a total of 4,112 created fault cases are reported. During the generation of these unseen fault cases, three distinct fault locations ”15%, 45%, and 70%,” of he line length are used; these locations were not utilized during the training process. Once the fault distances are determined, different types of faults labeled as ”ABC, AB, AC, A, B, C to ground” were systematically simulated based on the specified locations for each branch within the system. Then, both post-fault and pre-fault values were collected for corresponding PMU locations. Finally, trained models are used to determine the faulted branch within the system and the results are given in Table I. As evident from the performance results in Table I, the proposed method has much higher percentage of success compared to the MLP in terms of detecting the previously unseen cases by the trained model.

As evident from Table II, it is apparent that after identifying the locations through the trained ANN models, OLS is utilized to ascertain both the fault type and its exact location among

TABLE II: The performance of the OLS process for determining the fault types and distances.

Method	Detected Locations	Utilized PMU Number	Used Method	Detected Fault Type & Distance	Performance	Overall Performance
CNN	4056	50	OLS	3999	98.60%	97.15%
MLP	3085	50	OLS	3035	98.38%	73.73%

the detected faulted branches with considerable accuracy using the voltage differences gathered from the deployed PMUs.

As depicted from both Table I and II, using the proposed method provides satisfactory results for not only the training process but also the detection and identification of previously unseen fault cases in the 196-bus three phase distribution system utilizing the capabilities of CNN architecture and OLS.

VII. Conclusion

Accurate identification of types and locations of faults in power networks is needed to avoid or minimize service interruptions. This paper introduces an approach which is aided by artificial neural networks to achieve this goal with a high rate of success. One significant and distinguishing feature of the proposed approach is its reliance on a limited number of synchronized voltage measurements received from a set of strategically placed PMUs. Effectiveness of the proposed approach is illustrated by simulated scenarios on a typical three-phase distribution system.

References

[1] G. Andersson, "Modelling and analysis of electric power systems," *ETH Zurich*, pp. 5–6, 2008.

[2] N. Tleis, *Power systems modelling and fault analysis: theory and practice*. Elsevier, 2007.

[3] K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He, "Fault location in power distribution systems via deep graph convolutional networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 119–131, 2020.

[4] A. Aljohani, A. Aljurbua, M. Shafiullah, and M. A. Abido, "Smart fault detection and classification for distribution grid hybridizing st and mlp-nn," in *2018 15th International Multi-Conference on Systems, Signals & Devices (SSD)*, pp. 94–98, 2018.

[5] S. I. Ahmed, M. F. Rahman, S. Kundu, R. M. Chowdhury, A. O. Hussain, and M. Ferdoushi, "Deep neural network based fault classification and location detection in power transmission line," in *2022 12th International Conference on Electrical and Computer Engineering (ICECE)*, pp. 252–255, 2022.

[6] S. A. Aleem, N. Shahid, and I. H. Naqvi, "Methodologies in power systems fault detection and diagnosis," *Energy Systems*, vol. 6, pp. 85–108, 2015.

[7] A. Mouco and A. Abur, "Fault location using sparse l1 estimator and phasor measurement units," 06 2019.

[8] G. Feng and A. Abur, "Fault location using wide-area measurements and sparse estimation," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2938–2945, 2016.

[9] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[10] D. R. Shrivastava, S. A. Siddiqui, and K. Verma, "Optimal pmu placement for coordinated observability of power system under contingencies," in *2017 IEEE International Conference on Circuits and Systems (ICCS)*, pp. 334–339, 2017.

[11] S. Kumar, B. Tyagi, V. Kumar, and S. Chohan, "Multi-phase pmu placement considering reliability of power system network," in *2018 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, pp. 1–9, 2018.

[12] M. Göl and A. Abur, "Optimal pmu placement for state estimation robustness," in *IEEE PES Innovative Smart Grid Technologies, Europe*, pp. 1–6, 2014.

[13] C. Dismuke and R. Lindrooth, "Ordinary least squares," *Methods and Designs for Outcomes Research*, vol. 93, pp. 93–104, 2006.

[14] S. Ben Driss, M. Soua, R. Kachouri, and M. Akil, "A comparison study between mlp and convolutional neural network models for character recognition," in *SPIE Conference on Real-Time Image and Video Processing*, (Anaheim, CA, United States), Apr 2017. HAL Id: hal-01525504.

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

[16] F. Chollet *et al.*, "Keras," 2015. https://github.com/fchollet/keras.