Estimating Breaker Status with Electrical State Images and Convolutional Neural Networks

Vladimiro Miranda
INESC TEC and
FEUP, Faculty of Engineering of the
University of Porto
Porto, Portugal
vmiranda@inesctec.pt

Luís Teixeira
INESC TEC and
FEUP, Faculty of Engineering of the
University of Porto
Porto, Portugal
up201406320@fe.up.pt

Jorge Pereira
INESC TEC and
FEP, Faculty of Economy of the
University of Porto
Porto, Portugal
jpereira@inesctec.pt

Abstract — This paper presents a method to identify the status (open or closed) of breakers in network branches, in the absence of status signal or electric measurements on the branch including the breaker. Indirect power measurements from the SCADA are combined to form a 2D image array, which is fed into a Convolutional Neural Network. The image construction is based on ranking measurements with the Cauchy-Schwarz divergence between two signal distributions (for breaker open and closed). The success rate obtained with this technique is close to 100% in the IEEE testbed adopted.

Keywords — Convolution neural networks, image recognition, topology estimation.

I. Introduction

Topology recognition and construction has always been a concern for a correct output of the State Estimation procedure running in a EMS/DMS environment. The process, however and as a rule, in practical environments, is still conducted very much based on a priori information and empirical rules, mostly based on sensor information on power flow through the devices.

Automatic topology identification is becoming a vital function in power networks, because of the growing requirement of self-healing behavior. This is mostly noted in distribution networks: switching is a major operation function and is essential is network reconfiguration following outages.

Two approaches have been proposed in literature, to deal with breaker status identification: 1. Making it a process within the SE procedure (also called co-estimation); 2. Making it an independent pattern recognition process. This paper presents a methodology belonging to the second approach. While the former is an approach that may be suitable for Energy Management Systems/Distribution Management Systems (EMS/DMS) environments, in Control Centers, the latter may be adopted in distributed environments, including as core in independent intelligent agents performing tasks in the network. This is why, in a way, these approaches are not competing but may rather be complementing each other.

The basic hypothesis behind this latter approach is to admit that a set of electric signals, surrounding the switching device of unknown status, forms distinct clusters or patterns when the device is open or closed, for many different power flow scenarios. The tool of choice to separate these clusters has been the Artificial Neural Network (ANN). Approaches have been distinct mostly in the type of ANN selected. Early attempts [1]-[5] used multi-layer perceptrons, with outputs generally of the binary type indicating status of open/ close for a device modeled, i.e. projecting the data onto a line where two results are discriminated. Recent proposals [6] adopted auto-associative neural networks (autoencoders), with a very distinct concept – learning the supporting manifold of the training vector data and then suggesting, via input-output error, if a new vector belongs to the learned manifold. This technique demanded the training of two autoencoders, learning from measurement data vectors, one for the case with breaker open and the other for breaker closed; then, in a competitive arrangement, the autoencoder displaying smaller input-output error would indicate to which set (open or closed) would a new vector belong.

The performance of the competitive autoencoder model was very accurate, when tested in a modified IEEE 24-bus network with added breakers. However, for each breaker, two autoencoders were required; plus, training (or re-training) an autoencoder is not at all an easy task, usually demanding specialized algorithmic approaches, because simple backpropagation iterations have difficulty in converging correctly (due to the bottleneck in the middle layer of the autoencoders). Thus, new attempts were made to develop a model with a single tapered neural network, instead of two, avoiding the autoencoder approach, but now resorting to a deep neural network [7]. A limited test suggested that this approach would surpass the former attempts.

This paper reports yet another approach, this time adopting Convolutional Neural Networks (CNN), with extremely good results. There are several elements of novelty in the new approach, which deserve to be underlined:

- the choice of a special type of neural network with specific architecture, emulating the connection in the human brain of the visual cortex to the hippocampus;
- the replacement of an input in the form of a vector (common to all previous approaches) to an input in the form of a 2D array, similar to an imaged represented by pixels;
- the selection of measurements and their arrangement within the 2D array, to favor an image recognition process by the CNN.

The paper compares the results obtained by the tree above-mentioned approaches: competitive autoencoders, deep neural network and the new CNN model. The testbed for comparisons will be the IEEE RTS 24-bus test system.

This work is integrated in the research effort supported by the ERDF – European Regional Development Fund (COMPETE 2020 Programme) and FCT- Fundação para a Ciência e Tecnologia (project POCI-01-0145-FEDER-016731 INFUSE), Portugal.

978-1-7281-3192-4/19/$31.00 ©2019 IEEE

II. Autoencoders

Auto-associative neural networks or autoencoders are feedforward networks that an input vector of the same size as the output and are trained to display an output equal to the input. Usually, they have a middle layer of smaller dimension than input-output layers (see example in Fig. 1).

A binary competitive arrangement of autoencoders is depicted in Fig. 2. After having been trained independently with data from two clusters, a new vector is presented to both networks. If the vector belongs to one of the sets, the input-output error of the corresponding autoencoder will be small, while in the other one it will be large. This allows the identification of the set the input vector is associated with.

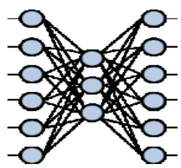

Fig. 1. Architecture of a 6-3-6 autoencoder with a single hidden layer

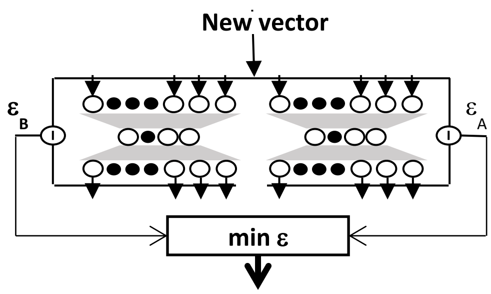

Fig. 2. Two autoencoders in parallel compete to produce the smallest error, thus identifying which cluster the new input vector belongs to.

III. Deep networks

A tapered neural network has an architecture with a stack of layers with progressively reduced size, such as in Fig. 3. The input is successfully projected in a smaller dimension space until the output. Training is mostly in unsupervised fashion, layer by layer, using some cost criterion that tends to maximize information content, even after dimension reduction from layer to layer. For instance, one may use Maximum Quadratic Entropy [7], where the entropy in the output layer is maximized - as entropy is a measure of information content of a probability density function, the maximum entropy in the output can only be obtained if all the information at the input is passed through to the output. The output layer is trained under supervised mode, using a logistic regression function, to make the output become associated with the labels defined a priori.

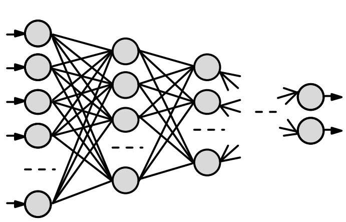

Fig. 3. Illustration of a tapered deep neural network, with several layers and ending in a 2-neuron layer.

IV. Convolutional neural networks

Convolutional Neural Networks (CNNs) are deep, feed-forward artificial neural network inspired in the biological vision process, with a connection pattern mimicking a visual cortex. The image processing and recognition features of CNNs have a solid mathematical background [8]-[10].

Layers in a CNN are not vectors, but arrays of neurons. The input layer can therefore receive an image and pass information to the following layers. The layers of the network are segmented in cells, in such a way that a block of neurons in a middle layer only receives input from selected blocks in the previous layer. This arrangement allows the network to recognize features and condense the extracted features in successive layers, taking advantage of the strong correlations between neighboring pixels that exist in images. The output may be matched to pre-defined labels, in a supervised process using some type of logistic regression function.

Such process is illustrated in Fig. 4.

There are not many reported applications to power system of CNNs – some examples are on insulator inspection from aerial images [11], on recognition of wind power patterns [12], or on recognition of network events in PMU frequency measurements [13]. The former model is based on actual images, while the two latter models are based on the construction of artificial images.

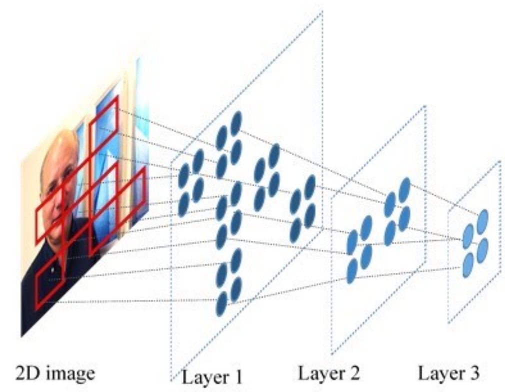

Fig. 4. Illustration of the action of a CNN over an image: segmentation and progressive merger.

V. Testbed

The results presented in the next sections all were obtained with the same testbed: a large set of randomly generated power flow scenarios, including sampling of breaker statuses. These scenarios were obtained on the IEEE RTS 24-bus network, with added breakers in 10 branches, as depicted in Fig. 5.

The dataset is composed of 20,000 power flow simulation results. To the calculated flows, Gaussian noise was superimposed, with $\sigma = 0.001$ p.u. in a 100 MVA power base. The resulting values were were standardized by converting the distribution of each one to mean 0 and variance 1. For each breaker, this set can be divided in to sub-sets, one for the breaker open and the other for breaker closed. The data set was divided in training and test sets, with approximately 70% of the data used in the training phase.

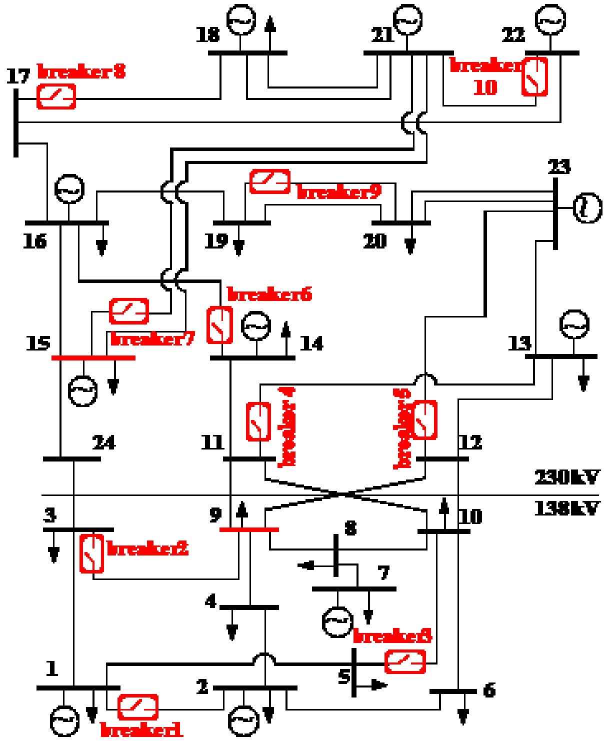

Fig. 5. Modified IEEE 24-bus RTS, with breakers added in 10 branches.

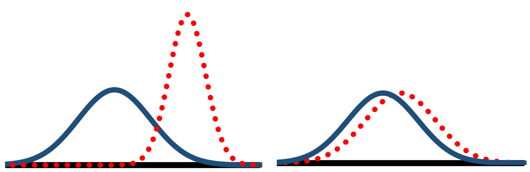

Fig. 6. Illustration of the distributions of the measurement values of two electric variables (A - left; B - right), for breaker open and closed. The variable associated to the left case is much more discriminative; the D_{CS} in the right case is closer to zero.

VI. Construction of Input Images

A. Discriminatory power of measurements

To serve as input to the CNN, a set of measurements must be selected. We define discriminatory power of a measurement, measured in the set of testbed scenarios, as the Cauchy-Schwarz divergence between two distributions of such variable: one for all the cases with breaker open, the other for all the cases of breaker closed. Given two discrete random variables P and Q, with finite distribution representations p and q, the Cauchy-Schwarz divergence D_{CS} is given by

$$D_{CS}(P\,\| \qquad \mathrm{g}\,\frac{\sum_{i=1}^{n} p_i q_i}{\sqrt{\sum_{i=1}^{n} p_i^{2} \sum_{i=1}^{n} q_i^{2}}} \tag{1}$$

Fig. 6 illustrates the case for two different measurement variables – their distributions in the test set for breaker open and closed have distinct overlaps. The D_{CS} values are different: in case A it is larger than in case B, showing that variable A is more helpful in discriminating the cases for breaker open or closed.

B. Building an image frame

An image frame is an array of measurement variables. An image is composed of actual values of measurements, for a specific case.

To build an image frame, for a specific case of a breaker, first the measurement variables are ordered in descending D_{CS} value, calculated by (1). Then, the variable with the highest value is placed on the upper leftmost corner and the remainder are placed around this corner, in a way such that, the closest to the corner, the largest the D_{CS} value (distance is measured, in the array, in the Manhattan ou L_1 metric). The result of this process is depicted in Fig. 7 for an of array 11x11 measurements.

This array may be built based only on values from installed measuring devices, or on values obtained from power flow simulations. The suggestion is that the discriminatory power of measurements, based on equation (1), be used as priority criterion in selecting the variables to be measured (and measuring devices or sensors installed), so that the ropology recognition process becomes effective.

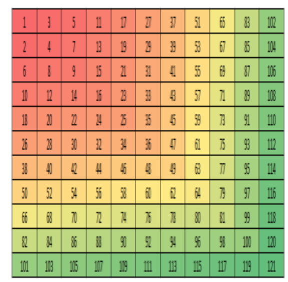

Fig. 7. Illustration, in a heat map, of the arrangement of measurement variables according to a decreasing Cauchy-Schwarz divergence value, from top left to bottom right. This map serves as indication on which measurements to collect, to obtain the desired result, the top left having priority.

C. Input images from indirect measurements

Over an image frame, a mask may be imposed to flatten out specific measurements that are unavailable. In this paper, we will assume that no direct measurements exist (neither active nor reactive power through the breaker), so classical heuristic rules cannot be used (such as: if power flow through the breaker is greater than zero, then the breaker is closed). For the unavailable measurements, a constant value is assumed, so such flattened measurement does not contribute to discriminate patterns. One example of mask is in Fig. 8, where only 13 measurements are available.

1	3	5	11	17	27
2	4	7	13	19	29
6	8	9	15	21	31
10	12	14	16	23	33
18	20	22	24	25	35
26	28	30	32	34	36

Fig. 8. Example of mask superimposed to the input matrix. The grey squares represent constant input values.

VII. Testbed results

The results obtained by the CNN models [14] are compared with other published results: competitive autoencoder model from [6], deep neural network from [7].

The results from [6] are available for all breakers in the test system, while from [7] only for breaker 2 are available.

In this work, we assumed that all possible power measurements were candidates. The process, therefore, is allowing the recognition of the most discriminatory measurements to be used; it is not using a measurement plan where only a limited set of measurements is fixed a priori.

A. Architectures used with autoencoders [6]

In the competitive autoencoder models, the sizes of the input/output and middle layers were the ones in the following table

Table I – Number of neurons in the input/output and middle layers of each autoencoder in the competitive arrangement

Breaker	1	2	3	4	5	6	7	8	9	10
I/O layers	14	18	16	16	16	14	16	12	12	16
Md. layer	9	13	11	11	11	9	11	7	7	11

B. Architecture of the deep neural network [7]

The deep neural network used in identifying the status of breaker 2 had 6 layers, with the sizes 18-14-10-6-4-2, with the two last layers providing patterns similar to [1,0] or [0,1], separating the cases of breaker closed and breaker open.

It therefore required 18 measurements to identify the breaker status, reproducing the same experiment as in the competitive autoencoder arrangement.

C. Architectures for CNNs

To conduct the tests with CNNs, two architectures were tested, one allowing an 11x11 input (model A) and the other a 6x6 input (model B). The objective was to use the smallest input possible (corresponding to the need of less measurements) to achieve the best recognition.

The CNN A had 9 layers, while the CNN B had 7 layers, including two last layers to smEsooth out the logistic regression approximating the targets [0,1] or [1,0].

D. Comparison of model accuracy

Table II shows, in percentage, the accuracy in correct identification of breaker status in a comparison of the three methods: in [6], [7] and the results now presented for CNNs, from [14]. It is evident that the Convolutional Neural Network models reached the highest accuracy – 100% in almost all cases, if we combine both models A and B. Such a result is obtained in 5 cases with only 13 measurements (model B).

Breaker 2 received more attention (possibly because it was the worst classified in the firstly published autoencoder model. In fact, in another publication [15], the autoencoder model trained with Maximum Quadratic Mutual Information for its first half reached the accuracy of 99.86%.

These results illustrate well how accurate a CNN model can be, in identifying the status of a breaker, only based on indirect local measurements.

Table II – Comparative accuracy of models in breaker status identification, in the modified IEEE RTS 24.bus test system.

Breaker	Autoencoder [6]	Deep NN [7]	CNN [14]	
			A	B
1	100.00%		100.00%	95.73%
2	98.69%	99,96%	100.00%	99.87%
3	100.00%		99.93%	100.00%
4	100.00%		100.00%	100.00%
5	100.00%		99.97%	100.00%
6	100.00%		100.00%	100.00%
7	99.89%		100.00%	93.20%
8	99.33%		99.93%	99,90%
9	99.97%		100.00%	99.67%
10	99.96%		100.00%	100.00%

It is also obvious that for some breakers, status identification is easier than for other breakers. This depends mostly on the network topology and on where the breaker is inserted. The case of breaker 8 is enlightening on these difficulties, because it was not possible to reach 100% accuracy with any of the models suggested. Therefore, yet another attempt was made, by enlarging the input array from 13 to 16 most discriminatory measurements (while still not including direct measurements). Applying model B CNN, we finally reached 100% with an image array of 6x6 with a mask flattening 20 measurements.

It must also be added that the strategy of enlarging the image size or number of significant pixels (measurement variables) is not a guarantee to improved performance. In some cases, adding more measurements led to poorer performance; this may be explained by the presence of more noise in the images, instead of more information content.

It is therefore possible to state that, using the convenient CNN model and the convenient set of inputs, one could always find a CNN that could 100% correctly identify a breaker status of open or closed, by discriminating a pattern formed by the values of measurements of active and reactive power flows. In other words, the new technique is the only one that allowed discovering a process of reaching 100% accuracy in all cases (by carefully selecting the model to be applied). The comparative case of the competitive autoencoders [6] is shown in the second column of Table II: its results were also reached with a best choice of autoencoder architecture in each case (each breaker) and the accuracy obtained, while remarkable, remained below the accuracy not reached with the virtual image processing CNN model.

VIII. Conclusions

Having a correct recognition of the topology of a power network is of paramount importance in an operation and control center. Most EDS/DMS functions depend on a correct identification of topology, starting with State Estimation, which is usually preceded by an algorithm denoted “configurator” that establishes network connections and build up the necessary grid admittance matrices. An error in setting up the correct network configuration puts in jeopardy all subsequent functions and leads to operator misjudgment or even incorrect triggering of automatic functions, with serious implications on operation security and safety.

It is therefore relevant that an intelligent function for breaker status definition may be added to SCADA-EMS/DMS environments. This function can be used in a pre-processing phase, in three ways:

1. Producing convenient input to the configurator;
2. Replacing missing breaker status signals;
3. Checking breaker status signals arriving at the SCADA, for inconsistencies, and signaling possible sensor malfunction.

The two latter processes are especially useful in distribution networks, because of the pervasive number of breakers and switches and the not so infrequent sensor failure or communication problems.

This intelligent function is totally independent of any other algorithm or pre-processing. Therefore, one can also foresee it embedded as core function in distributed intelligent agents, in charge of the automatic operation of switches and breakers, within the operational concept of network self-healing. The fact that such a function may establish redundancy with a specific local sensor increases the reliability of operation. Complex self-healing procedures demand some level of local information and communications – which are compatible with the concept.

The work reported has the value of confirming that it is indeed possible, though pattern recognition on electric measurements, to identify breaker status and recognize network topology. It also shows that this is not a trifle task, because it requires a sophisticated model to achieve maximum accuracy.

The art in the process is to have a way of building 2D images from a vector of measurements such that patterns more distinguishable may be identified by a CNN. The paper showed that the concept of discriminatory power of a measurement can be translated numerically by the Cauchy-Schwarz divergence between two distributions for the same variable, obtained for a random set of power flow scenarios with a breaker open and with the same breaker closed. It also showed that a convenient arrangement, in a 2D array, of measurement values is one that packs together the variables with greater discriminatory power.

However, theoretically, there is a zone of ambiguity in this classification process, i.e., the mapping of electric states onto the space of breaker statuses is not a strict 1-to-1 mapping. There are, in fact, power flow states that may have correspondence in both the "breaker open" and "breaker closed" states – when the voltages at both extremes of the branch including the breaker are equal. In this case, the power flow through the device is zero and the network state is indifferent to the breaker status. This can only be resolved with direct observation.

The consequence of this is that a true 100% accuracy, using a pattern recognition approach, is not theoretically attainable. However, the probability of an ambiguity case happening, in a real network, must be considered as extremely low – and this would also affect any method of co-estimation, using mathematical models.

Therefore, having available a sharp tool for breaker status identification must be considered as a valuable add-on to network operation. This paper demonstrates that a CNN-based model, with input images constructed adequately, is for the moment the most reliable tool proposed.

References

[1] A. A. da Silva, V. Quintana, and G. K. H. Pang, "A pattern analysis approach for topology determination, bad data correction and missing measurement estimation in power systems", Proceedings of the Twenty-Second Annual North American, pp. 363 - 372, 1990.

[2] A. P. Alves da Silva, V. H. Quintana, and G. K. H. Pang, "Neural networks for topology determination of power systems," Proceedings of the First International Forum on Applications of Neural Networks to Power Systems, Seattle, pp. 297-301, 1991.

[3] A. Alves da Silva, V. Quintana, and G. Pang, "Solving data acquisition and processing problems in power systems using a pattern analysis approach", IEE Proceedings-C, vol. 138, no. 4, pp. 365–376, 1991.

[4] D. Vinod Kumar, S. Srivastava, S. Shah, and S. Mathur, "Topology processing and static state estimation using artificial neural networks," IEE Proceedings - Generation, Transmission and Distribution, vol. 143, no. 1, pp. 99–105, 1996

[5] J. Souza, A. Leite da Silva and A. P. Alves da Silva, "Online topology determination and bad data suppression in power system operation using artificial neural networks", IEEE Transactions on Power Systems, vol. 13, no. 3, pp. 3311-3318, 2013.

[6] J. Krstulovic, V. Miranda and J. Pereira, "Towards an auto-associative topology state estimator," IEEE Transactions on Power Systems, vol. 28, no. 3, pp. 3311-3318, 2013)

[7] R. Oliveira, R. Bessa and V. Miranda, "Identifying topology in power networks in the absence of breaker status sensor signals," 19th IEEE Mediterranean Electrotechnical Conference (MELECON), Marrakech, Morocco, 2-7 May 2018. DOI: 10.1109/MELCON.2018.8379086

[8] Y. LeCun, K. Kavukcuoglu, C. Farabet, "Convolutional networks and applications in vision," in Proc. of the 2010 IEEE International Symposium on Circuits and Systems (ISCAS), Paris, France, 30 May-2 June 2010

[9] J. Bruna and S. Mallat, "Invariant scattering convolution networks," IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 35, no. 8, pp. 1872–1886, 2013

[10] T. Wiatowski and H. Bolcskei, "A Mathematical Theory of Deep Convolutional Neural Networks for Feature Extraction," IEEE Transactions on Information Theory , Vol. 64, no. 3, March 2018

[11] Z. Zhao, G. Xu, Y. Qi, N. Liu, and T. Zhang, "Multi-patch deep features for power line insulator status classification from aerial images," in Proc. of the IEEE International Joint Conference on Neural Networks (IJCNN), pp 3187–3194, July 2016.

[12] H. Wang, G. Li, G. Wang, J. Peng, H. Jiang, and Y. Liu. "Deep learning based ensemble approach for probabilistic wind power forecasting". Applied Energy, 188:56–70, 2017.

[13] V. Miranda, P. A. Cardoso, R. J. Bessa and I. Decker, "Through the looking glass: Seeing events in power systems dynamics," International Journal of Electrical Power & Energy Systems, Vol. 106, pp. 411-419, March 2019. DOI: 10.1016/j.ijepes.2018.10.024

[14] L. Teixeira, "MISS SAIGON – Missing Signal Appraising in Globally optimized Networks," M.Sc. thesis, Faculty of Engineering of the University of Porto, July 26, 2019

[15] V. Miranda, J. Krstulovic, J. Hora, V. Palma and J. C. Príncipe, "Breaker status uncovered by autoencoders under unsupervised maximum mutual information training," ISAP 2013 – International Conference on Intelligent Systems Applications to Power Systems, Tokyo, Japan, Jul 1-4, 2013.