Deep Learning Methods Apply in the Power Transformer Differential Protection

Elias Junio Muller Coelho [a]

[a] Academic Department of Electrical, Federal University of Technology – Parana (UTFPR), Cornélio Procópio, Brazil, eliasjunio@alunos.utfpr.edu.br

Abstract. Transformers play a key role in the transmission and distribution of power systems. Diagnostic faults of the power transformer is very important to ensure safe and stable operation of the power system. The objective of this article is to show methods for deep learning already applied in transformer differential protection and to share summary results of these methods. The methods addressed are Accelerated Convolutional Neural Network, Signal Localised Convolutional Neural Network, Fast GRNN and Dynamic Differential Current in real time with CNN. In the analysis of the articles, many analyses were carried out in different cases, with accuracy well above 95%, where in some cases it reached more than 99.5%. Therefore, the deep learning methods presented are effective and accurate, enabling for possible more advanced studies.

Keywords. Deep learning, transformer differential protection, electrical power system.

1. Introduction

The transformer plays a key role in the transmission and distribution of the power system, so the safe operation of the transformer directly affects the stability of the entire power system. Therefore, fault diagnosis of the power transformer is very important to ensure safe and stable operation of the power system [1]. A differential system can effectively protect a transformer because of the inherent reliability of the relays, which are highly efficient in operation, and the fact that equivalent ampere turns are developed in the primary and secondary windings of the transformer. The CTs on the primary and secondary sides of the transformer are connected in such a way that they form a circulating current system. Faults on the terminals or in the windings are within the transformer protection zone and should be cleared as quickly as possible. Although differential protection is very reliable for protecting power transformers, windings are not always fully protected, especially in the case of single-phase faults [2].

With the increasing complexity in time and memory of power system applications, the need for advanced statistical pattern recognition tools has led to the use of deep learning methodologies Clique ou toque aqui para inserir o texto. deep learning application can be thought of to solve problems related to transformer differential protection.

The objective of this article is to show methods for deep learning already applied in the transformer differential protection and to try to apply the methods addressed to solve the problem pertinent in transformer delta/star in which the star winding has been earthed via resistor. It is a well-known fact that overall phase differential protection cannot protect a large portion of an impedance-ground delta-star transformer [4].

2. Methodology

2.1 Theoretical Development

The book [2] deals with the problem described in the Introduction, follows the deduction of the problem below second the book:

Consider the case of a delta/star transformer in which the star winding has been earthed via a resistor. Assume that an internal earth fault occurs at point F at a distance X from the neutral point, involving X% turns, and that the resistor has been set so that normal current I_{nom} will flow for a fault on the terminals (with full line-to-neutral voltage applied between phase and earth). The numbers of primary and secondary turns are N_p and N_s, respectively. The secondary current for a fault at F is produced by X% of the line-to-neutral voltage. Therefore, by direct ratio, the current will be XI_{nom}. In addition, the number of turns involved in the fault is XN_s. The distribution of current in the delta side for an earth fault on the star side results in a line current I_L' equal to the phase current [2]. Therefore,

$$I_L' = XI_{nom} \text{ x } \left(\frac{XN_s}{N_p}\right) = X^2 I_{nom} \left(\frac{N_s}{N_p}\right) \quad 2.1$$

Under normal conditions, the line current in the delta side, I_L, is

$$I_L = \sqrt{3} I_{nom} \text{ x } \left(\frac{N_s}{N_p}\right) \quad 2.2$$

If the differential relay is set to operate for 20% of the nominal line current, then, for operation of the relay, the following should apply:

$$I_L' \geq 0.2 \text{ x } I_L \quad 2.3$$

That is,

$$X^2 I_{nom} \left(\frac{N_s}{N_p}\right) \geq 0.2 \text{ x } \sqrt{3} \text{ x } I_{nom} \text{ x } \left(\frac{N_s}{N_p}\right) \quad 2.4$$

$X^2 \geq 0.2\sqrt{3}$, i.e. $X \geq 59\%$

Therefore, 59% of the secondary winding will remain unprotected. It should be noted that protecting 80% of the winding ($X \geq 0.2$) would require an effective relay setting of 2.3% of the nominal primary current. This level of configuration can be very difficult to achieve with certain types of differential relays [2].

2.2 Deep Learning Methods

Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learnt [5].

This part is responsible for showing a brief summary to different methods for deep learning applied in the Power Transformer Differential Protection.

2.2.1 Accelerated Convolutional Neural Network (CNN)

In this cited article, an approach based on an accelerated convolutional neural network (CNN) is designed to discriminate between internal faults and inrush current [6].

The basic CNN is a type of deep learning network, which inherently extracts features based on the convolution operator. The basic CNN utilises conventional activation functions, such as the sigmoid activation function, which requires computing an exponent. This is a huge disadvantage when dealing with high non-linearity and can significantly increase training and evaluation times [6].

An accelerated CNN as a fast and compact version of the CNN is developed in the cited article to reduce both training and evaluation times, which performs on the basis of minimising the approximation error and decomposing the parameters of individual convolution layers and fully connected layers. This proposed algorithm also has the ability to merge feature extraction and fault detection blocks into a single deep neural network block, allowing the network to discover important features automatically, making the algorithm more efficient in terms of speed, hardware usage, and precision [6].

2.2.2 Signal Localised Convolution Neural Network (SLCNN)

This method approached in the cited article uses the distinct signal localisation, which is performed with the convolution process sequentially on the frequency and time coefficients which are obtained from the wavelet decomposition of the differential current signal [7].

Unlike a fully connected network, the network connection between the layers in the CNN has a localised region. The localisation on frequency and time information is made on the time-frequency spectrum as a spatial localisation, where the time domain signal is processed by the time-frequency transformation technique. It is essential to use a suitable dimension for the filter weight corresponding to each decomposition level, because the dimension of the time frequency is an irregular matrix, and therefore each level has a dissimilar number of coefficients. This results in an equal number of coefficients in each level of the second hidden layer. The convolution operation with the nonlinear activation function is performed between the time-frequency spectrum of the signal and shared filter weights. It means the extraction of features as performed on the respective frequency scale corresponding to time [7].

The results of convolution are obtained as the elements of the succeeding layer without a pooling operation. Generally, the pooling process is involved after the convolution process in each layer to reduce the size of the data. Here, the pooling process is not incorporated as the size of the matrix does not incur a substantial burden for signal processing applications. Hence, a set of feature mapping filters is used between any two layers to extract useful features of the signal [7].

The filter weights of each feature mapping in the second layer have a single column vector, which represents the feature extraction with an accumulation of all frequency features in the time scale. The convolution operation with the non-linear activation function is performed between the second layer elements and a shared filter weight vector, thereby it produces the output equal number of time scale in higher level decomposition (sixth level coefficients). This approach demonstrates the feature mapping on the time scale. Further, the output of each feature in the second layer is organized in a single vector, which is represented as elements of the third layer. Finally, the third layer is fully connected with the output layer as in the conventional neural network. Therefore, the output layer of the SLCNN has two neurones for the required classification of the trip and the restriction decision [7].

2.2.3 Fast GRNN

The differential protection in the power transformer should be able to perform based on raw data and fully learn the temporal features and changes in the transient signal, because the differential protection in the power transformer was always threatened by sending false trips subjected to external transient disturbances [8].

To propose a diagnosis scheme to implement a differential protection in real time, the cited article develops a GRU-based structure. GRU (gated recurrent unit) is the removal of one of the three gates in the LSTM (long short-term memory) structure, making it another faster and more accurate structure. It generally consists of two main ports, the update gate and the reset gate. To decrease the computational burden as well as improve the accuracy and reliability of the differential protection, the reset gate has been removed, thereby, the computational complexity has reduced by almost 42%, in the diagnosis process. This process is called FGRNN (fast gated recurrent neural network) [8].

The designed FGRNN, in the cited article, does not neglect the transient behaviour of internal faults, unlike the GRU, which neglects transient phenomena such as inrush current and internal faults in power transformers in the training process, by completely resetting the network. Learning sudden changes in the power transformers that can be beneficial in the discrimination process is fully fulfilled in the designed FGRNN network, as well as learning fully temporal features [8].

2.2.4 Dynamic Differential Current-Based Transformer Protecton Using CNN

The dynamic differential current is presented in the cited article, with the objective of improving the generalisation performance and response speed of the multi-feature fusion-based transformer protection. The dynamic differential current merges the pre-disturbance and post-disturbance differential currents in real time, then develops a transformer protection based on dynamic differential current focussing on the characteristic changes of the differential current [9].

The image of differential current can comprehensively embody the feature changes resulting from any disturbance. In addition, a short window is sometimes sufficient to clearly reflect the internal fault because the differential current will change instantly when an internal fault occurs [9].

In order to identify the running states reliably in the shortest possible time, multiple images, which include the differential current from pre-disturbance one cycle to post-disturbance different time, are combined in order of time to define a dynamic differential current. The dynamic differential current serves as an input of CNN to identify the running states in real time, second the article [9].

3. Result and Discussion

The results presented below are a summary of each article cited, with their respective credits for the development of their work, simulations, and tests.

3.1 Accelerated CNN

Second, in the cited article, different cases with various external factors are simulated to calculate reliability indexes. Applied to a simulated 230-kV network and an experimental prototype. Five cases are discussed to address the advantages of the proposed differential protection scheme [6]. The cases are as follows:

1- simple benchmark;

2- CT saturation caused by fault or inrush current;

3- power transformer with series capacitor compensation;

4- Transformer with SFCL at neural point;

5- Experimental setup.

In Table 1 shows the minimal results in all the indexes classified by for each case cited in the article, where in certain cases the proposed method operates completely correctly with 100% performance [6].

Table 1 – Minimal results in all the indexes classified by for each case cited.

Cases	1	2	3	4	5
Minimal results (%)	98.46	99.17	98.46	97.70	98.46

In the article cited, an accelerated CNN was designed using the product quantisation technique to speed up the convolution and FCN layers. The accelerated CNN performs four times as fast as the basic CNN, not only without losing accuracy but also with an improvement of about 1% accuracy. The proposed machine learning-based protection method can be applied to different systems, regardless of system parameters once the CNN structure is set [6].

3.2 SLCNN

Second, in the cited article, three power transformer (frequency 50Hz) test systems are considered for evaluation of the proposed SLCNN where the training patterns of each transformer are generated for various operating conditions. This is shown in Table 2 the specifications of the power transformers [7].

Table 2 – Power transformers specifications

	Power (MVA)	Ratio (kV)	%Z	Conn
PT-1	40	132/11.5	13.56	Dyn11
PT-2	50	132/12	35.64	YNd1
PT-3	100	230/110/11	11.41	YNynd1

The performance of SLCNN based relay is evaluated in terms of accuracy, sensitivity, and specificity. It is shown in Table 3, also shown in more detail and compared with different methods in the article [7], The cases are:

1- Internal fault;

2- internal fault with inrush current;

3- Internal fault with CT saturation;

4- internal fault with inrush and CT saturation;

5- Inrush current;

6- Sympathetic inrush current;

7- CT saturation due to an external fault.

Table 3 – Performance of SLCNN.

	PT-1	PT-2	PT-3	PTs
Accuracy (%)	99.52	99.20	99.20	99.31
Sensitivity (%)	99.54	99.07	99.31	99.31
Specificity (%)	99.48	99.48	99.96	99.31

The SLCNN classifies almost all patterns correctly with an accuracy of 99.31% for all testing samples. However, the false classification is 0.69% which occurs due to the simultaneous occurrence of events of internal fault, inrush current, and CT saturation caused by an external fault. In a practical scenario, these simultaneous occurrences of events are very rare. Thus, it offers an effective method of classification for both multiple events and single events of internal fault, inrush current, sympathetic inrush current, and CT saturation due to an external fault [7].

3.3 Fast GRNN

Five case studies are discussed in the cited article, to investigate the superiority of the proposed differential protection technique. Case studies are briefly as follows:

1. A simple case without external factors;
2. Differential protection in the presence of CT saturation;
3. Series capacitor compensation;
4. transformer with SFCL on the neutral point;
5. Experimental prototype.

Different fault types, fault location, switching angles, transformer winding types, and source impedances are used to generate data sets for the case studies mentioned, simulating in a 230kV network [8].

In Table 4 shows the minimal results in all the indexes classified by for each case cited in the article, where in certain cases, the proposed method operates completely correctly with 100% performance [8].

Table 4 – Minimal results in all the indexes classified by for each case cited.

Cases	1	2	3	4	5
Minimal results (%)	98.47	99.17	98.51	96.97	97.16

The results of the robust FGRNN method show that the proposed method has an average computational time of less than 6.5ms [8].

3.4 Dynamic Differential Current

In the study to the article cited, the training and test samples were both collected in PSCAD. The PSCAD simulation has fully considered the influences of the transformer parameters and scenarios on the features of dynamic differential current. The transformer parameters of the training and test samples are completely different from each other to improve and verify the generalisability of the proposed protection method [9].

	Ratio (kV)	Sampling Frequency (kHz)	Conn
PT-1	230/11	20	Y/ Δ-11
PT-2	230/35	20	Y/ Δ-11
PT-TEST	35/500	4	Δ / Y-11

The dynamic model experiments verified that the accuracy of experimental scenarios reached 99.10% for all running states, with the average window length of 1ms and 10.14ms for internal fault and faulty transformer energization, respectively [9].

4. Conclusions

Power transformer fault diagnosis is very important to ensure the safe and stable operation of the power system, because safe operation of the transformer directly affects the stability of the entire power system, and a power transformer is a fundamental part of the energy transmission and distribution system. Although differential protection is reliable and widely used in transformer protection, in some cases the transformer windings are not fully protected, resulting in the need to use more advanced methods to solve these problems, in this case, deep learning.

In the articles analysed, many analyses were carried out in different cases, where it was observed that the application of deep learning in the differential protection of transformers has a very high accuracy, well above 95%, and in some cases reaching more than 99.5%.

Therefore, it is concluded that the deep learning methods presented are effective and accurate, enabling possible more advanced studies, especially in the case of the delta/star connection problem grounded with a resistor in a single-phase fault.

5. References

[1] M. Liu *et al.*, "Power Transformer Discharge Fault Identification Based on Deep Learning," in *Proceedings - 2022 4th International Conference on Applied Machine Learning, ICAML 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 339–344. doi: 10.1109/ICAML57167.2022.00071.

[2] J. M. Gers, E. J. Holmes, and Institution of Electrical Engineers., *Protection of electricity distribution networks*. Institution of Electrical Engineers, 2011.

[3] M. Khodayar, G. Liu, J. Wang, and M. E. Khodayar, "Deep learning in power systems research: A review," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 2. Institute of Electrical and Electronics Engineers Inc., pp. 209–220, Mar. 01, 2021. doi: 10.17775/CSEEJPES.2020.02700.

[4] J.-C. Tan, D. Tholomier, and H. Wei, "A New Restricted Earth Fault Protection."

[5] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, May 27, 2015. doi: 10.1038/nature14539.

[6] S. Afrasiabi, M. Afrasiabi, B. Parang, and M. Mohammadi, "Integration of Accelerated Deep Neural Network into Power Transformer Differential Protection," *IEEE Trans Industr Inform*, vol. 16, no. 2, pp. 865–876, Feb. 2020, doi: 10.1109/TII.2019.2929744.

[7] S. K. Murugan, S. P. Simon, and R. R. Eapen, "A Novel Signal Localized Convolution Neural Network for Power Transformer Differential Protection," *IEEE Transactions on Power Delivery*, vol. 37, no. 2, pp. 1242–1251, Apr. 2022, doi: 10.1109/TPWRD.2021.3080927.

[8] S. Afrasiabi, M. Afrasiabi, B. Parang, M. Mohammadi, H. Samet, and T. Dragicevic, "Fast GRNN-Based Method for Distinguishing Inrush Currents in Power Transformers," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 8, pp. 8501–8512, Aug. 2022, doi: 10.1109/TIE.2021.3109535.

[9] "Dynamic differential current-based transformer protection using convolutional neural network," *CSEE Journal of Power and Energy Systems*, 2022, doi: 10.17775/cseejpes.2021.02120.