Differential Protection of Power Transformers Using Discrete Wavelet Transform and Convolutional Neural Network

Conference Paper · February 2025

DOI: 10.1109/TPEC63981.2025.10906891

5 authors, including:

Daksh Vyawhare
Institute of Infrastructure Technology Research and Management

Chiragkumar Parekh
RK University

Differential Protection of Power Transformers using Discrete Wavelet Transform and Convolutional Neural Network

Daksh Vyawhare, Jetal Chaudhari, Chirag Parekh, Manish Chaturvedi, Krupa Shah, Senior Member IEEE

***Abstract*—This research paper presents a hybrid approach for differential protection of transformer using discrete wavelet transform (DWT) and convolutional neural network (CNN). A simulation of a three-phase transformer is conducted to capture various current conditions such as normal, fault, and inrush currents. A statistical tool DWT is employed to decompose the input signal according to the mother wavelet Daubechies level. Further, it is used to extract required features such as mean, norm, standard deviation. The preprocessed data are feed into CNN, to discriminate between normal, fault, and inrush conditions accurately. In the training process, CNN demonstrated very good performance by achieving accuracy of 98%. To validate the effectiveness of the proposed method, a detailed analysis of the correct classification with actual true labels is carried out by employing performance metrics. The results indicate the potential of this hybrid DWT-CNN framework, to provide reliable and efficient differential protection of transformers that significantly improves reliability of the system.**

***Index Terms*—Convolution Neural Network (CNN), Differential Protection, Discrete Wavelet Transform (DWT), Transformer**

I. Introduction

Differential protection scheme is used as a primary scheme for the protection of the power transformers, generators, and motors wherein differential relays are utilised to protect the electrical apparatus from the internal fault. The currents flowing into and out of the apparatus under protection are monitored where a current transformer (CT) is being utilised to convert the current from the higher value to the proportionate current of a lower value. Under normal conditions, the current entering the primary winding should match with that of current leaving the secondary winding. This indicates the balanced condition. During the internal fault condition, the balance in the currents gets disturbed and the differential current starts flowing through the operating coil. This necessitates the need to issue a trip signal. The relay circuit will trip the breaker on both the sides of the transformer. When the fault is developed outside the transformer protection zone, it is referred to as an external fault. The differential relay scheme is built in such a way that only if a fault occurs within protection zone, the trip signal is issued. Transformers also experience magnetizing inrush current during energisation. The current could be higher than that of normal operating current and may lead to saturation and in turn phase shift. The CT core might be subjected to a symmetrical or asymmetrical saturation. The relay at times interprets such imbalance conditions as a fault, and issues a false trip signal. Therefore, discrimination of such conditions from the actual internal faults is essential to avoid mal-operation of relay [1], [2].

Differential protection of power transformer is a critical aspect in terms of power system reliability and safety. Traditional approaches for protection schemes at times suffer from limitations such as sensitivity for disturbances and susceptibility to false alarms. To challenge these limitations, innovative approaches imposing advanced wavelet transform and machine learning are considered. Wavelet transform has ability to decompose signal into multiple frequency bands and is suitable for extracting relevant features from the faulty signals, by capturing both time and frequency domain information. This can enhance the accuracy and sensitivity of fault detection [3]–[7]. On the other hand, CNN signify the performance for various pattern recognition tasks or classification tasks; that is done by learning about features from raw data; which can effectively classify fault type currents and differentiate from normal operating conditions. The Section II inspects the application/domain of wavelet transform and convolution neural networks (CNNs) for differential protection of transformers.

II. Literature Survey

Well-developed research has been carried out on percent differential relays for transformer protection over the years. As defined in [2], an innovative approach of computer-based flux-restraint current differential is proposed. Traditional approaches such as harmonic restraint and voltage restraint for current differential protection have shown certain limitations in identifying internal- and external-faults. To overcome the observed limitations, flux-restrained method exhibiting relationship between transformer flux and current is proposed that provides more accurate and reliable restrain function. Moreover, it requires fewer computational resources for calculating flux as compared to that of traditional harmonic current analysis method. The prime advantage of such flux-restrain approach is having direct correlation with the transformer saturation. However, it depends on voltage measurements, which then results in limitation when voltage data is unavailable. In such scenario, there is a requirement of utilising some alternate method such as harmonic restraint.

In [8], the potential of second harmonic restraint differential protection of transformer (during internal faults) is studied with the help of three different algorithms such as LCA (Least-square Curve fitting Algorithm), FAA (Fourier Analysis Approach) and RTA (Rectangular Transform Approach). A

979-8-3315-4112-5/25/$31.00 ©2025 IEEE

three-phase transformer is used to simulate faults and inrush condition and several fault scenarios are tested. The results of FAA is extended for further analysis. Three different schemes are discussed such as instantaneous tripping of differential relay, voltage-restrained differential relay and, control of the second harmonic restrain with the phase voltage. It is observed from the study that the approach based on second harmonic can be applied as a primary differential protection scheme and long time delay can be avoided if phase voltage is utilised as a control signal. Transformer protection by performing high-frequency analysis of current transient signals is demonstrated in [9] for the purpose of accurately detecting inrush currents. Simulation study on EMTP software is performed to demonstrate the ability of precisely identifying various fault conditions and internal faults within the transformer protection zone, by ensuring its reliable protection and minimizing downtime.

In [4]–[7], an approach for identifying various types of the currents in power transformers, by employing two-level wavelet packet transform (WPT) is shown. The coefficients of the second-level high frequency sub-band are analysed and current types such as fault current and non-fault disturbances like magnetizing inrush currents are effectively identified. The optimal wavelet and resolution levels are considered using the minimum description length (MDL) criteria. In this approach, simulation study is performed for the test purposes, where WPT performed accurately and has rapidly identified different currents under various conditions. It is found that the WPT is suitable for analysing the non-stationary signals in the power systems. In [4]–[7], the study is also extended to hardware by building WPT using band-pass (BP) filters. It has offered a cost-effective and efficient solution that reduces the need for the complex circuits. The mentioned BP-filter WPT-based approach is utilized to build a relay for differential protection that detected faults successfully and restrained the relay from unnecessary operation. In [3], wavelet transform is utilized for accurate performance and detection of the internal and external faults for achieving an improved performance and reducing the complexity as compared to that of traditional methods; while in [10], current and voltage ratios are utilized to differentiate between inrush and fault currents. Each approach has shown promising results for enhancing power system protection in terms of reliability and security.

Since 1990, several authors have explored computation-based approaches utilising Artificial Neural Network (ANN) models as an alternative to traditional approaches to discriminate between magnetizing inrush and fault currents [11]–[14]. The use of ANN model for power transformer protection might offer advantages in terms of accurate discrimination of the faulty currents with faster response times and improved performance in terms of reliability under various fault scenarios. In [15], [16], machine learning algorithms such as Support Vector Machine (SVM), k-Nearest Neighbor (kNN), Decision Tree (DT), Random Tree (RT), Random Forest (RF) are utilised for the differential protection. It is observed that the algorithms are capable in differentiating the fault currents from the normal currents. Another approach is presented in [17]–[20] in which WT based technique is integrated with the machine learning algorithms such as ANN, PNN (Probabilistic Neural Network), DT, Ensemble Method based Decision Trees (EMD). It is found that the approaches offer promising improvements in terms of accuracy and reliability of protection scheme. Due to relevant feature extraction from the current signals, internal and inrush currents are accurately identified and false tripping rates are reduced. Moreover, power transformer protection approaches are also influenced by deep learning models [21]–[26]. In research, various models such as SVM, kNN, PNN, Neural Network (NN) and Convolutional Neural Networks (CNNs) are explored to classify the currents with complex patterns in current and voltage waveforms.

On a larger scale micro grid systems are focusing on the protection of transformers by utilizing the DWT and deep learning models and proven effective in the application of micro grid system [27]. However, DWT focusing on extracting relevant features by selecting mother wavelet that enables the improvement in performance of the model for classifying different fault types and in turn improve the system reliability. However, there still exists challenges pertaining to data quality available for the training of ANN model, computational complexity, model interpretation and integration of experimental data with actual sampling frequency. Unlike traditional methods, use of the DWT-transformed signals by considering the standard mother wavelet Daubechies level helps to extract relevant features and patterns. Moreover, choosing the appropriate sampling frequency during the wavelet analysis will decide the performance of the considered algorithm in terms of classification accuracy and convergence time. Therefore, this research focuses on a building a framework to fault classification for power transformers, using DWT and CNN. The primary focus of utilizing CNNs is to efficiently process the input data. It is crucial for feature learning and designing scalable architectures that is capable of handling large datasets. Hence, it is worthwhile to perform this investigation by considering wavelet based algorithm with the sampling frequency same as that used by the industry/utility to demonstrate the potential of neural network-based differential protection of power transformers.

The objective of this work is to combine the power of wavelet transform and CNN to develop, efficient differential protection scheme for power transformers. The proposed approach involves wavelet-based feature extraction and CNN based classification at utility permitted sampling frequency for enhancing the performance of differential protection approach. The details are explained in section III.

III. Proposed Method

In this work, a methodology is proposed in line with classifying the internal fault in power transformer using discrete wavelet transform (DWT) and convolution neural network (CNN). Wavelet Transform is a mathematical tool that enables the time-frequency representation to extract transient features in the time domain whereas, DWT provides efficient computation for the implementation of WT, by decomposing a

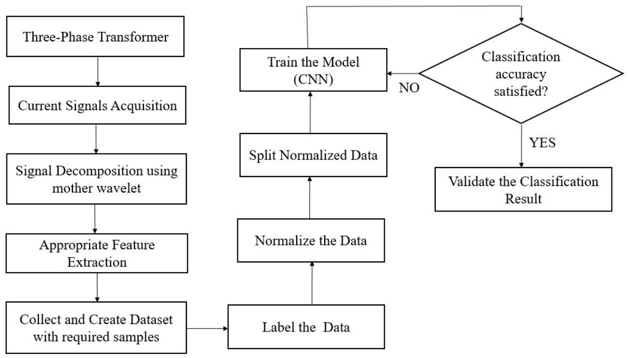

Fig. 1: Flow diagram for proposed methodology

signal into series of wavelet coefficients on different scale and location such as frequencies and time. With this, it is possible to extract suitable features. Analyzing the statistical properties such as standard deviation, mean etc. at different decomposition levels enables the extracted feature to be supplied to neural networks for the purpose of accurate classification of faults. The necessary steps of the method are enumerated below.

a) Data Collection: This phase involves generation of various current signals pertaining to fault-free conditions in the power system, including normal as well as inrush scenarios. Signals are generated for various switching instances to capture variations in the magnitude of the inrush current. Moreover, signals are also generated for various fault conditions such as line to ground (LG) fault, double line to ground (LLG) fault, triple line to ground (LLLG) fault, external fault considering various switching angles. Preprocessing of the data including noise cancellation can be performed for experimental datasets.

b) Feature Extraction Using Wavelet Transform: The second phase includes decomposition of the acquired differential current signals into multiple frequency bands using DWT and selection of an appropriate mother wavelet to obtain detailed and approximate coefficient effectively for each signal. The choice of mother wavelet such as Daubechies wavelet is crucial as it significantly impacts the feature extraction. Further, statistical indices such as standard deviation, mean, are analysed at various decomposition levels to fulfill the objective of fault-classification.

c) Data Processing: The third phase involves feature engineering for the feature extracted data, such as labelling, normalization, dividing the data into training, testing and validation sets. The labelling is carried out as it helps models to identify the appropriate labels and therefore, advantageous for the purpose of classifying the fault currents. Normalization is used to scale the data in certain ranges using methods such as standard scaler. Subsequently, each feature is centered by subtracting its mean and then scaled by dividing it by its standard deviation. Later, normalized data is divided into certain ratios such as 80:20 or 70:30 for the effective training/validation of the CNN.

d) Convolutional Neural Network (CNN) Design: CNN is similar to feed-forward neural network and is specifically designed for the image processing related task in which features are extracted from the images utilising 2D CNN architecture. Similarly, CNN is also utilized for extracting statistical information with 1D architecture. CNN architecture consists of key layers that are convolutional layers, pooling layers and dense layers (also known as fully connected layers). In 1D architecture, the input layer accepts the features derived from the wavelet coefficients and, convolutional layers employ the filter called stride that moves with the defined size throughout the layer size and maps each feature of simulation data to generate the activation map. A ReLU activation function is applied with convolutional layer to get the data in the certain range. A pooling layer reduces the dimensions of the activation maps and prevents overfitting. It also reduces the computational cost. Lastly, the dense layer, also known as fully connected layers, resembles traditional neural networks, that are responsible to classify the features (that is, internal fault, external fault, inrush condition). To balance the computation speed, network depth, and accuracy, CNN model requires optimization functions such as Adam optimizer. The loss function is also very crucial part that measures the performance of a model by assessing how well its predictions align with the actual target values. For the multi-label classification problems with integer labels, the sparse categorical cross-entropy loss function is utilised.

As per the objective of the work, DWT is utilized to decompose current signals, and feed them to the CNN model whereas, CNN model is employed to classify the currents in various categories such as normal, inrush, internal fault or external fault as shown in Fig.1. After training the model, performance of the model is evaluated by measuring the accuracy and loss of the train data and validation data. Further, performance metrics including precision, recall, F1 Score are utilized for analysis purpose. If the results are not appropriate then the process of training the model is repeated till satisfactory results are obtained. Later, classification results are evaluated with the help of confusion matrix.

IV. Experimental Setup

This section represents the result acquired from the proposed methodology, by accomplishing circuit simulation, wavelet decomposition and convolution neural network implementation as described in section III.

a) Data Collection: An extensive dataset is generated using simulation of a 250 MVA, 138/13.8 kV, Y-Y, three-phase power transformer in MATLAB Simulink. The simulation circuit is shown in Fig.2, that employed to simulate various conditions including normal current, various fault current and inrush currents such as LG, LLG, LLLG. The total of 996 current signals samples collected from the simulation condition.

b) Feature Extraction using DWT: To extract meaningful features from this acquired raw current signals DWT approach is utilized. The collected samples consist of current signals then supplied to DWT, for that wavelet decomposition with

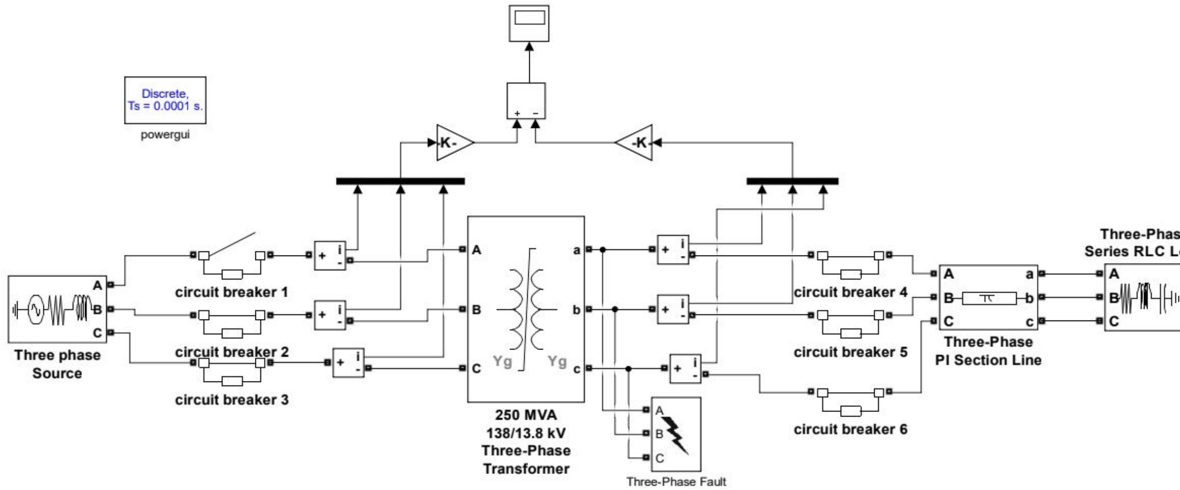

Fig. 2: Simulation circuit diagram

Daubechies wavelet family up to level 4 applied. The decomposition results into each phase with 12 wavelet coefficients. The statistical features were extracted for each coefficient such as mean, norm, and standard deviation, that result into a new dataset with 36 columns for statistical feature of each sample.

c) Feature Engineering: The feature extracted dataset consist of 996 rows that indicates the samples that were simulated through simulation and 36 columns that indicates the statistical feature of each current signals. Later, appropriate labels were applied to each current samples such as normal, fault and inrush. The prepared dataset is ready to feed into the machine learning algorithm such as CNN. To smoothen the process one-hot encoding is applied to these labels, then supplied to normalization process. The statistical features need to be normalized before feeding into the CNN algorithm for that standard scalar approach is utilized. These normalized data are then split into two-part train data and test data in the ratio of 80:20 for CNN training.

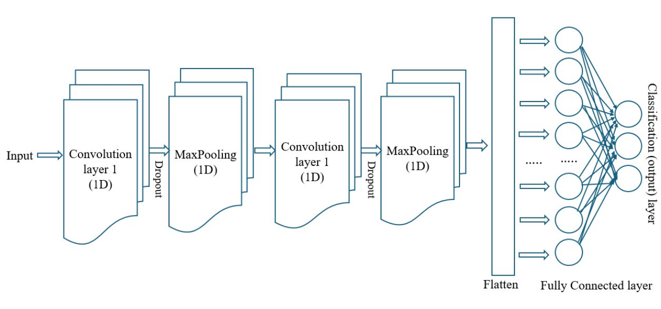

Fig. 3: CNN architecture for fault current classification

d) CNN: This study employs the 1D CNN architecture which consist of two convolutional layers, each followed by a dropout layer and max pooling layer as shown in Fig.3. To achieve the optimal configuration of 1D CNN architecture, the number and sizes for convolutional layers, dense layers, dropout rate, and the learning rate were varied. The number of convolutional layers (2-3), their filter sizes (32, 64, and 128), and the number of dense layers (1-3) before the softmax layer, dropout layers with dropout rates (0.2 to 0.5) and different learning rates (0.01, 0.001, and 0.0001). The Final Architecture consist of the kernel size of 3 employed on each convolution layer with an activation function ReLu. Each convolutional layer has their own filters 32 and 64 respectively. The dropout rate considered for this CNN architecture is 0.5 and 0.3 to mitigate overfitting. The max pooling layers with a pool size of 2 reduce the dimensions for next layers. After the convolution layers the resulting layer flattened and fed into two dense layer that is a fully connected layers for final classification result. The first fully connected layer has 64 units with ReLu activation function, and second fully connected layer has 3 units with softmax activation function utilized for multi-class classification. To overcome the overfitting issue with loss function l2 regularizer of 0.001 is applied to the fully connected layers. The model summary of 1D CNN is shown in Table I.

The CNN model trained with 80% of train dataset, with validation split of 10%. During the training CNN model was utilized with Adam optimizer with a learning rate of 0.001 to make learning performance more precisely, effectively and efficiently. This CNN model is iterated 150 times with batch size of 16. During the training the loss function is calculated using the sparse categorical cross-entropy loss function to quantify the difference between the predicted and true labels.

TABLE I: 1D CNN Model summary

Layer (type)	Output Shape	Parameter
conv1d (Conv1D)	(None, 34, 32)	128
dropout (Dropout)	(None, 34, 32)	0
max_pooling1d (MaxPooling1D)	(None, 17, 32)	0
conv1d_1 (Conv1D)	(None, 15, 64)	6208
dropout_1 (Dropout)	(None, 15, 64)	0
max_pooling1d_1 (MaxPooling1D)	(None, 7, 64)	0
flatten (Flatten)	(None,448)	0
dense (Dense)	(None, 64)	28,736
dense_1 (Dense)	(None, 3)	195

TABLE II: 1D CNN Performance

Data	Actual True Labels	Predicted True Labels	Accuracy
Train	775	769	98.42%
Test	194	179	92%
Validation	27	25	92.23%

To measure the performance of the model, the accuracy of the model is noted during training, validation and test process as shown in Table II.

V. Result Analysis

Comparative analysis of accuracy and loss between training and validation is carried out through corresponding graphs Fig. 4 and Fig. 5. As a result, minimum difference is noticed between training and validation performance. The accuracy of training is 98% and validation is 92.23%, and the difference between them is minimal that indicate the best fit of the model. Further, test process is conducted that result into 92% of accuracy.

To more precisely assess the result of the test, the performance evaluation metrics consist of precision, recall, F1 score along with the accuracy of each label shown in Table III. To verify the correct label classification and misclassification for all labels, the confusion matrix is employed as shown in Table IV. Lastly, Table II shows the number of actual true labels trained with predicted true labels during training and testing, and the observation indicates that 6 samples were misclassified during the training process and at the time of testing 15 samples were incorrectly classified. To look out, model misunderstanding with these misclassified labels, confusion matrix observation is employed; that indicates that the CNN model is misclassified the normal and inrush labels as fault label. To see significantly, the fault class is also misunderstood with normal label and inrush label that leads to further study to improvement of the model for actual true classification. To conclude this study,a comparison is carried out with various approaches,by considering various parameters such as configuration of simulated power transformers, the number of simulated cases/ cycles/ samples, the software platform used and their selected algorithms. The proposed method is compared with existing similar methods as shown in Table V, that shows the simulation parameters are low compared to other methods, and the considered parameters are similar as industry defined utility. Therefore, this method can be suitable for real-time application.For online monitoring purpose, this pre-trained CNN model needs to be deployed on a micro-controller like Raspberry PI. However, the detailed computational complexity analysis is required to make an

TABLE III: Model Performance Evaluation Metrics

Class	Accuracy	Precision	Recall	F1 Score
Normal	0.92268	0.950820	0.906250	0.928000
Fault	0.92268	0.870130	0.957143	0.911565
Inrush	0.92268	0.964286	0.900000	0.931034

TABLE IV: Confusion Matrix

True Label	Predicted Label		
	Normal	Fault	Inrush
Normal	58	5	1
Fault	2	67	1
Inrush	1	5	54

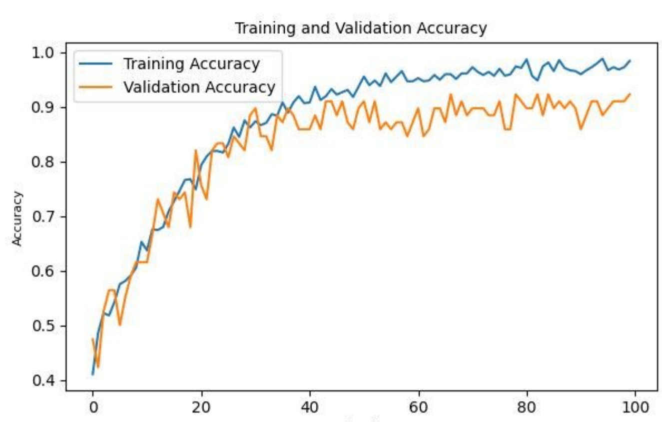

Fig. 4: Graph for accuracy v/s epoch for training and validation of 1D CNN

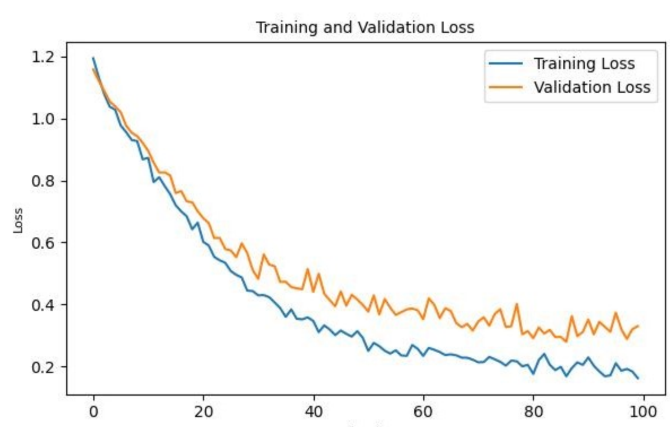

Fig. 5: Graph for loss v/s epoch for training and validation of 1D CNN

TABLE V: Comparison of Proposed Approach

Reference	Simulated Data Configuration	Algorithm
[17]	System 1: 750 MVA, 27/420 kV (150 cases); System 2: 35 MVA, 11/132 kV (50 cases); EMTP	ANN
[18]	500/230 kV - Star delta; 50 Hz sampling frequency; one cycle 800 sample for without load, with load and with load for secondary side; MATLAB-Simulink	PNN
[20]	315 MVA 400 kV/220 kV 50 Hz; normal, internal, external, over excitation, magnetizing inrush; train: 590 cases, test: 178 cases; PSCAD/EMTDC	EDT
Proposed Method	250 MVA 138/13.8 kV; normal, internal fault and inrush currents; MATLAB-Simulink	CNN

concrete claim about the feasibility of such deployment. This analysis is planned as a future work.

VI. Future Scope and Conclusion

This research presents a hybrid approach, combining DWT and CNN, to effectively classify the current data as normal, fault, and inrush conditions acquired from the simulation of a three-phase transformer. The proposed DWT-CNN framework achieves a remarkable training accuracy of 98% and a test accuracy of 92%. These results indicates the accurate performance in recognizing and categorizing the current data. A detailed analysis of the confusion matrix conclude that a few labels are misclassified. However, individual classification accuracy report is evident by the performance metrics, that demonstrate the ability of the model to accurately classify each label. This indicates that the DWT-CNN framework is capable of handling the complexities acquired in the dataset. To further enhance practical applicability of the proposed framework, future research direction focuses on the advancements on computing, enlarging the samples, and framework deployment on the industry acquired data for real time analysis.

References

[1] T. Sidhu and M. Sachdev, "Online identification of magnetizing inrush and internal faults in three-phase transformers," *IEEE Transactions on Power Delivery*, vol. 7, no. 4, pp. 1885–1891, 1992.

[2] A. G. Phadke and J. S. Thorp, "A new computer-based flux-restrained current-differential relay for power transformer protection," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-102, no. 11, pp. 3624–3629, 1983.

[3] R. P. Medeiros, F. B. Costa, and K. M. Silva, "Power transformer differential protection using the boundary discrete wavelet transform," *IEEE Transactions on Power Delivery*, vol. 31, pp. 2083–2095, Oct 2016.

[4] S. Saleh and M. Rahman, "Modeling and protection of a three-phase power transformer using wavelet packet transform," *IEEE Transactions on Power Delivery*, vol. 20, pp. 1273–1282, April 2005.

[5] S. Saleh and M. Rahman, "Real-time testing of a wpt-based protection algorithm for three-phase power transformers," *IEEE Transactions on Industry Applications*, vol. 41, pp. 1125–1132, July 2005.

[6] S. A. Saleh and M. A. Rahman, "Testing of a wavelet-packet-transform-based differential protection for resistance-grounded three-phase transformers," *IEEE Transactions on Industry Applications*, vol. 46, pp. 1109–1117, May 2010.

[7] S. A. Saleh, B. Scaplen, and M. A. Rahman, "A new implementation method of wavelet-packet-transform differential protection for power transformers," *IEEE Transactions on Industry Applications*, vol. 47, pp. 1003–1012, March 2011.

[8] P. Liu, O. Malik, D. Chen, G. Hope, and Y. Guo, "Improved operation of differential protection of power transformers for internal faults," *IEEE Transactions on Power Delivery*, vol. 7, no. 4, pp. 1912–1919, 1992.

[9] Z. Bo, G. Weller, and T. Lomas, "A new technique for transformer protection based on transient detection," *IEEE Transactions on Power Delivery*, vol. 15, pp. 870–875, July 2000.

[10] E. Ali, A. Helal, H. Desouki, K. Shebl, S. Abdelkader, and O. Malik, "Power transformer differential protection using current and voltage ratios," *Electric Power Systems Research*, vol. 154, pp. 140–150, 2018.

[11] L. Perez, A. Flechsig, J. Meador, and Z. Obradovic, "Training an artificial neural network to discriminate between magnetizing inrush and internal faults," *IEEE Transactions on Power Delivery*, vol. 9, pp. 434–441, Jan 1994.

[12] J. Pihler, B. Grcar, and D. Dolinar, "Improved operation of power transformer protection using artificial neural network," *IEEE Transactions on Power Delivery*, vol. 12, pp. 1128–1136, July 1997.

[13] M. Zaman and M. Rahman, "Experimental testing of the artificial neural network based protection of power transformers," *IEEE Transactions on Power Delivery*, vol. 13, pp. 510–517, April 1998.

[14] A. Orille, N. Khalil, and J. Valencia, "A transformer differential protection based on finite impulse response artificial neural network," *Computers Industrial Engineering*, vol. 37, no. 1, pp. 399–402, 1999. Proceedings of the 24th international conference on computers and industrial engineering.

[15] A. M. Shah and B. R. Bhalja, "Discrimination between internal faults and other disturbances in transformer using the support vector machine-based protection scheme," *IEEE Transactions on Power Delivery*, vol. 28, pp. 1508–1515, July 2013.

[16] N. Shahbazi, S. Bagheri, and G. Gharehpetian, "Identification and classification of cross-country faults in transformers using k-nn and tree-based classifiers," *Electric Power Systems Research*, vol. 204, p. 107690, 2022.

[17] P. Mao and R. Aggarwal, "A novel approach to the classification of the transient phenomena in power transformers using combined wavelet transform and neural network," *IEEE Transactions on Power Delivery*, vol. 16, pp. 654–660, Oct 2001.

[18] S. Sendilkumar, B. Mathur, and J. Henry, "Differential protection for power transformer using wavelet transform and pnn," *International Journal of Electrical and Computer Engineering*, vol. 4, no. 3, pp. 564–570, 2010.

[19] S. Bagheri, Z. Moravej, and G. B. Gharehpetian, "Classification and discrimination among winding mechanical defects, internal and external electrical faults, and inrush current of transformer," *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 484–493, Feb 2018.

[20] N. Shrivastava, "Power transformer differential protection based on dwt and edt," *International journal of research in engineering, science and management*, vol. 5, Feb 2022.

[21] M. Tripathy, "Power transformer differential protection using neural network principal component analysis and radial basis function neural network," *Simulation Modelling Practice and Theory*, vol. 18, no. 5, pp. 600–611, 2010.

[22] S. Afrasiabi, M. Afrasiabi, B. Parang, and M. Mohammadi, "Integration of accelerated deep neural network into power transformer differential protection," *IEEE Transactions on Industrial Informatics*, vol. 16, pp. 865–876, Feb 2020.

[23] M. Afrasiabi, S. Afrasiabi, B. Parang, and M. Mohammadi, "Power transformers internal fault diagnosis based on deep convolutional neural networks," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 1, pp. 1165–1179, 2019.

[24] 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, pp. 1242–1251, April 2022.

[25] S. Key, S.-H. Kang, N.-H. Lee, and S.-R. Nam, "Bayesian deep neural network to compensate for current transformer saturation," *IEEE Access*, vol. 9, pp. 154731–154739, 2021.

[26] Z. Li, Z. Jiao, and A. He, "Knowledge-based convolutional neural networks for transformer protection," *CSEE Journal of Power and Energy Systems*, vol. 7, pp. 270–278, March 2021.

[27] P. Pan, R. K. Mandal, and M. M. Rahman Redoy Akanda, "Fault classification with convolutional neural networks for microgrid systems," *International Transactions on Electrical Energy Systems*, vol. 2022, p. 8431450, april.