ISSN: 2633-4828 Vol. 5 No.4, December, 2023

International Journal of Applied Engineering & Technology

1D-CONVOLUTION NEURAL NETWORK FOR TRANSFORMER CORE FAULT IDENTIFICATION

Priyanka Tiwari [1], Shweta Singh[2], Naresh B[3*], Shirish Jain[4] and Anand Goswami[5]
[1,2] Department of Electrical Engineering, Maharishi University of Information Technology, Lucknow, UP, India
[3] Department of Electrical Engineering, DU, Dinjan, Assam
[4,5] Department of Electrical Engineering, GGITS, Jabalpur, MP
arnatiwaripriyanka@gmail.com, shwetasingh580@gmail.com, nareshbangari@gmail.com and shirishjain@ggits.org

ABSTRACT
The design and implementation of one-dimensional convolutional neural networks (CNNs) for Transformer core fault detection via vibrations on the surfaces of the tank wall is described in this paper. 1D-CNNs are mostly developed for applications analysing sensor data in time-series domain. A Convolutional Neural Network employs a local receptive field, weight sharing, and spatial subsampling. Proposed 1D-CNN model facilitates a reduction in the computational burden of the network. The main focus of this study is to analyse vibration signals of transformer core to diagnosis fault condition, where data is measured in time domain. The constructed CNN model for transformer core fault detection, demonstrated an outstanding performance with a classification accuracy of 99.78%. Subsequently, the obtained results were compared with previous studies and thoroughly discussed.

Keywords: 1D- CNN, classification, power transformer, vibration analysis.

1.INTRODUCTION
The Power transformers play a crucial role in ensuring reliable power delivery within the power system and a key component responsible for voltage conversion, power distribution, and transmission. In the event of transformer failure, there is a high likelihood of large-scale blackouts, leading to significant direct and indirect economic losses. To minimize the incidence of unexpected failures, various methods and tools are employed for predicting faults in power transformers. In the case of Frequency Response Analysis (FRA) method, the transformer winding impedance or admittance is measured to know the condition of the core as presented in [1,2]. A high precision mathematical model to find the core health by using short circuit impedance (SCI) method is discussed in [3]. In [4] researcher presented a new approach, by using Ultra-wideband signal to monitor the condition of the power Transformer. Thermal imaging monitoring, as demonstrated in [5], can effectively detect abnormal thermal faults in transformers. Additionally, an advanced big data processing algorithm is implemented for the comprehensive analysis of power transformer operating status [6]. Despite the utility of these methods, some are characterized by high costs or insufficient accuracy.

The utilization of vibration signals for assessing transformer health is a relatively recent technique compared to other methods of monitoring transformer conditions, and ongoing research in this field is still in its early phases [7]. As a non-intrusive online approach, methods based on vibration prove to be well-suited for evaluating the condition of power transformers. The vibration analysis technique involves employing a vibration sensor to measure the vibration signal of the core. Time-frequency domain characteristics are then extracted to enable online condition monitoring of the transformer [8]. Despite these advantages, significant limitations persist in monitoring faults in winding and iron core components due to the lack of comprehensive research on the vibration characteristics of these elements and a shortage of experience in this domain.

In the past few years, convolutional neural networks (CNNs) have demonstrated significant success in the realm of pattern recognition, known for their capability to autonomously extract features from signals or images. Typically, these networks are built upon two-dimensional convolutional neural networks (2D CNNs) [9]. Recent studies indicate that the expertise dependence issue can be addressed through the application of Deep Learning (DL) methods. Key DL models, such as Deep Belief Networks (DBN), Recurrent Neural Networks (RNN) [10], Long Short-Term Memory (LSTM) Networks, and Convolutional Neural Networks (CNN), have been prominent in this context. In [11], a Neural model is specifically crafted for predicting interturn faults in transformers.

This is achieved by classifying indicators extracted from vibration signals using a Deep Recurrent Neural Network (RNN). Furthermore, the authors of [12] introduces a method where converted vibrating images are identified using CNN to distinguish between three working conditions of transformers. In contrast to DBN, RNN, and LSTM, a Convolutional Neural Network (CNN) employs a local receptive field, weight sharing, and subsampling within a spatial domain. Consequently, CNN helps mitigate the computational load of the network and the risk of overfitting, significantly enhancing the accuracy and efficiency of pattern recognition. Specifically, in the realm of 1D CNNs, they demonstrate effectiveness in extracting features from fixed-length segments of a dataset, with less emphasis on the precise location of these features within the segment. This capability is particularly advantageous when analysing time sequences from sensor data, such as that from a gyroscope or accelerometer, as well as other types of signal data over defined time intervals. The 1D CNNs offer a direct application to the raw signal, such as current, voltage, or vibration, eliminating the need for pre- or post-processing tasks like feature extraction, selection, dimension reduction, denoising, etc. In contrast, 2D deep CNNs typically require transforming 1D data into 2D and often demand datasets with substantial size. Additionally, the simplicity and compact design of adaptive 1D CNNs, performing linear 1D convolution involving scalar multiplications and additions, make them conducive to real-time and cost-effective hardware implementations.

This paper outlines an analytical methodology for constructing a 1D CNN model designed for predicting fault in transformer core. The model is based on analysing vibration modes that indicate the bending of limbs in the out-of-plane direction. Multiple experimental studies are under taken to simulate diverse critical faults in transformer operational conditions. The vibrational data recorded from these experiments with different load conditions are then analysed using CNN models.

2.THEORETICAL BACKGROUND

2.1Mechanism of Iron core vibration

The core vibration primarily originates from the interaction of two phenomena: (i) the vibration of windings induced because of interaction of leakage flux with electromagnetic force generated by transformer windings, and (ii) magnetostriction occurring in the silicon steel sheet and the Maxwell electromagnetic force. [13,14]. When a silicon steel sheet comes under influence of external magnetic field, it undergoes dimensional changes such as contraction or elongation. After the removal of the magnetic field, the sheet reverts back to its original dimensions. This phenomenon is known as the magneto strictive effect. The advancements in lamination techniques for manufacturing the iron core and the utilization of waterproof tape for binding have significantly improved the overall control over transformer core vibration, with the magnetostriction of the silicon steel sheet being the dominant factor. Magnetic materials have property of changing their shape under the external magnetic field is generally called as magnetostriction. This is given by equation (1)[15].

$$\delta = \frac{\Delta L}{L} \qquad (1)$$

Here, δ represents the axial magnetostriction ratio of the silicon steel sheet, ΔL denotes contraction of the silicon steel sheet, and L represents the original axial dimension of the silicon steel sheet. Assuming the supply voltage is V_1 = Vs sin ωt, then magnetic flux density in core is:

$$B = \frac{\varphi}{A} = \frac{V_s}{\omega NA}\cos\omega t = A_0 \cos\omega t \qquad (2)$$

Where φ magnetic flux in the magnetic material core, A is the area of cross section of the core, and $B_0 = (Vs/\omega NA)$ is the magnetic flux density. The magnetic flux density and magnetic field strength are having linear relationship and it is

$$H = \frac{B}{\mu} = B\frac{H_C}{B_s} = B_0 H_C \frac{\cos\omega t}{B_s} \qquad (3)$$

ISSN: 2633-4828 Vol. 5 No.4, December, 2023

Here, φ represents magnetic flux passing through core, A denotes cross-sectional area of the core, and B0 = (Vs/ωNA) is the magnitude of the magnetic flux density. In the unsaturated state of the iron core, there exists a linear relationship between the magnetic flux density and the magnetic field strength. The magnetic field strength within the iron core is

$$\frac{\nabla L}{L}\frac{1}{dH} = |H|\frac{2\varepsilon_s}{H_c^2} \quad (4)$$

Where ε_s is the saturation magnetostriction ratio of the silicon steel sheet. Based equation (4) the axial contraction of the silicon steel sheet due to external magnetic field is as follows:

$$\nabla L = L\int_0^H |H|\frac{2\varepsilon_s}{H_c^2}\,dH = L\frac{\varepsilon_s H^2}{H_c^2} = \frac{L\varepsilon_s^2 V_s^2}{(\omega N_{core} B_s\, s)^2}\cos^2\omega t \quad (5)$$

The period of magnetostriction variation is half that of the alternating electromagnetic field, whereas the core vibration induced by magnetostriction have twice the frequency of electromagnetic field variation, denoted as ω.

2.2Convolutional Neural Network (CNN)

The Convolutional Neural Network (CNN) represents an advanced iteration of artificial neural networks (ANN), primarily designed for feature extraction from matrix datasets with a grid-like structure. Comprising various layers such as the input layer, convolutional layer, pooling layer, and fully connected layers, CNN plays a crucial role in pattern recognition. The convolution layer is particularly significant, while the pooling layer downsizes input data, enhancing computational efficiency, reducing memory usage, and preventing overfitting. The fully connected layer is responsible for making the final predictions. As each layer feeds its output into the subsequent one, the extracted features progressively and hierarchically become more complex. Figure 1 illustrates the general structure of CNN.

The original 1D vibration signal, represented as a vector i[n], where n = 1, 2, . . . N. Feature maps are generated by the filters/kernels through convolution with the input signal. If we denote the convolution kernel with size (*z)* as k[n], the convolution output o[n] can be expressed as:

$$o[n] = i[n] * k[n] = \sum_{m=0}^{z-1} x[m].k[n-m] \quad (6)$$

In general, the convolved feature at the output of l th layer can be written as

$$o_i^l = \sigma\left(b_i^l + \sum_j o_j^{l-1} \times k_{ij}^l\right) \quad (7)$$

Where, o_i^l represents the ith feature in the lst layer; o_j^{l-1} denotes the j th feature in the (l − 1)th layer; k k_{ij}^l represents the kernel linked from ith to j th feature, b_i^l denotes the bias for this feature and σ represents the activation function.

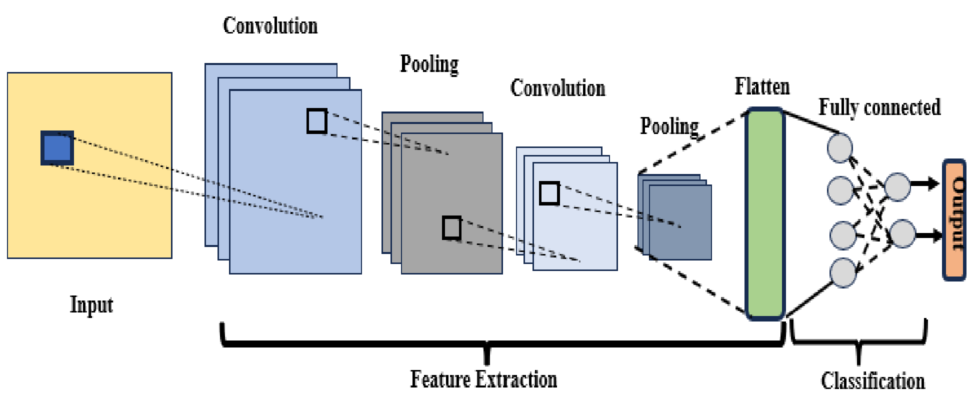

Figure 1 CNN general architecture diagram

The 1D -CNN model developed in this research work is to analyse the transformer core vibration signal. In the context of vibration signals, which are time-series data, the convolutional layers apply filters (kernels) to capture local patterns or features in the signal. These filters act as feature detectors and help identify important patterns at different scales and orientations. As the filters slide or convolve across the input signal, they create feature maps that highlight relevant information. In the convolution operation, pooling layers reduce the size of the feature map fed into the adjacent layer, while keeping the most significant information. This helps in creating a more compact and abstract representation of the input data, making it computationally efficient and less prone to overfitting.

In the CNNs, the hyperparameter space refers to the set of all possible combinations of hyperparameter values that can be used to configure and train the CNN model. Hyperparameters are external configuration settings that are not learned from the data but are set prior to the training process. Configuring the right set of hyperparameters is crucial for achieving good model performance. The hyperparameter space for a CNN is vast, as there are numerous possible combinations and values for these hyperparameters. Some assumed hyperparameters for CNN model are listed in Table 1. Exploring this space efficiently is the main aim of this paper. The goal is to find the set of hyperparameters that leads to the best performance on the specific task at hand.

Table 1 The CNN hyperparameters

S. No	Hyperparameters	Range
1	Batch size	32 and 64
2	Epoch numbers	1 to 100
3	Learning rates	0.001 and 0.0001
4	Convolution layers	1 to 6
5	Filter in each layer	2^3 to 2^8
6	Kernel size	2 to 7
7	Activation function	ReLu function

The size of the one-dimensional feature vector after the convolutional and pooling layers depends on several factor such as Number of filters, kernel size, stride and pooling type. By adjusting these parameters, you can control the size of the one-dimensional feature vector that is eventually fed into fully connected layers for classification or regression tasks. Experimenting with these hyperparameters allows you to fine-tune the network for optimal performance on your specific vibration signal processing task.

3.Experimental setup

The experimental setup comprises a voltage regulator, an experimental transformer, and a load cabinet shown in Figure2. The voltage regulator has a rated capacity of 5 kVA. In the testing process, the voltage regulator outputs the rated primary voltage, which is then applied to the transformer under examination. The load cabinet consists of a resistance cabinet, The load cabinet facilitates the adjustment of load magnitude. The platform can undergo testing under conditions of no-load, half load, and full load.

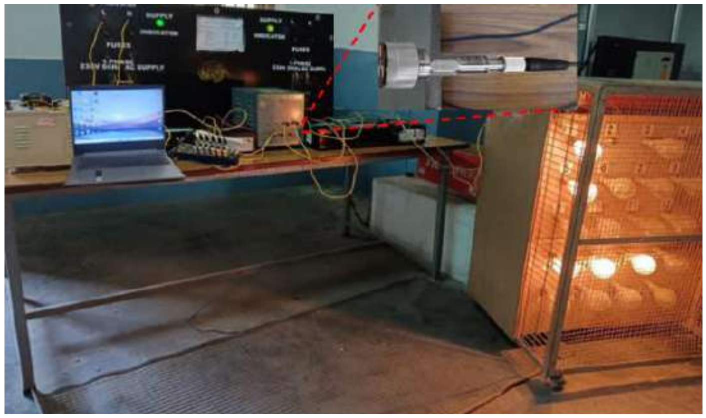

Figure 2 Experimental setup

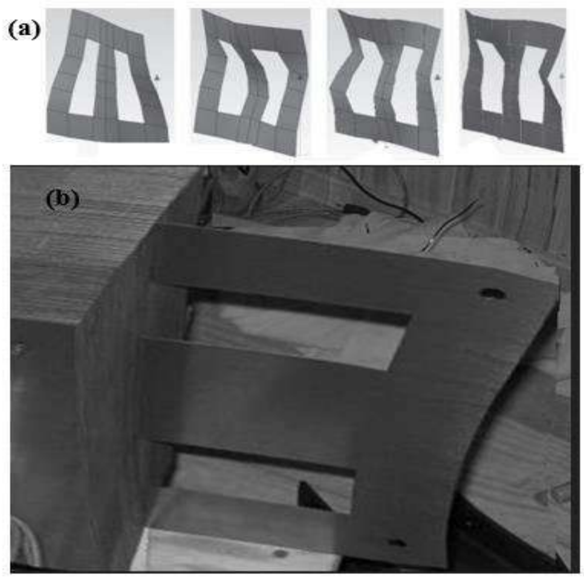

Figure 3 Bent core limb in out of plane direction

Table 2 2KVA Power Transformer Details

S. No	Parameter	Categories
1	Power rating	2KVA
2	High Voltage Rating	230V
3	Low Voltage Rating	115V
4	Operating Frequency	50Hz
5	Service conditions	Indoor

The accelerometer senther 310A is integrated to the transformer with flat magnet mounting method. Senther 310A sensor with Digital signal conditioner was used for measuring vibration signals from test transformer tank surface under different load conditions. The sensitivity of 100 mV/g and the sampling frequency used to collect vibration signals was 12KHz. To investigate the comprehensive vibration characteristics of the iron core, four distinct measuring points have been strategically positioned on the tank surface, two positions are illustrated in Figure

4(two more back side). These measurement locations have been carefully chosen to ensure proximity to the transformer core, which is securely fastened to the tank using bolts and nuts. This arrangement allows for a thorough analysis of the vibration behaviour and ensures that the measurements are taken in close proximity to the core-tank connection.

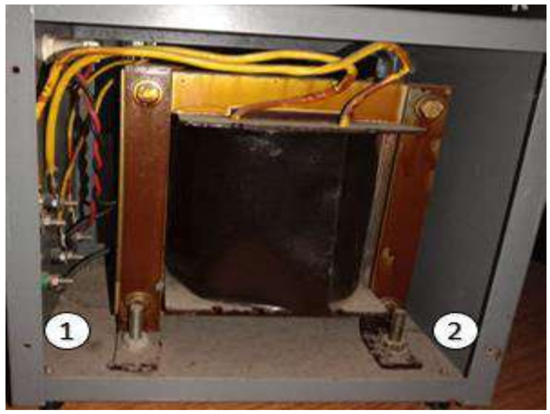

Figure 4 Vibration measuring points on Tank surface

To train the 1D-CNN, vibration signals are gathered for the transformer's no-load test, short-circuit test, and various load operation states. Specifically, vibration signals are collected during load operations with currents set at 0.25IF, 0.5IF, 0.75IF, and IF. The corresponding vibration signals are recorded for each load current step, with a data collection duration of 30 seconds per step.

4.RESULTS AND DISCUSSION

To train the 1D-CNN, vibration signals are gathered for the transformer's various load operation states. Specifically, vibration signals are collected during load operations with currents set at $0.25I_F$(I_F is full load current), $0.5I_F$, $0.75I_F$, and I_F. The corresponding vibration signals are recorded for each load current step, with a data collection duration of 30 seconds per step and the recorded data shown in below Figure 5.

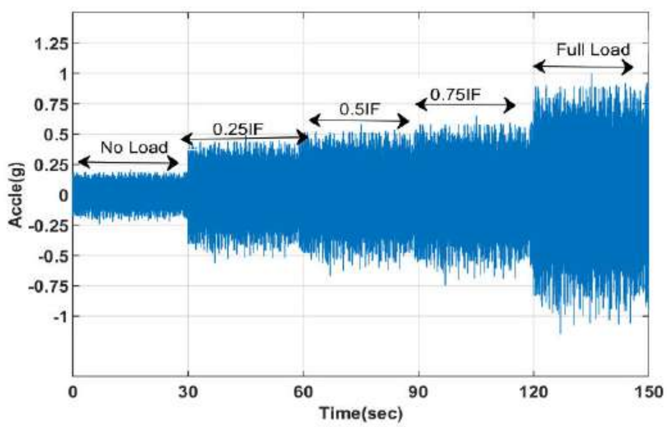

Figure 5 Recorded vibration signal under load test.

As discussed in the section 2.1 and following equation (2.5) the fundamental frequency of the vibration signal collected from the transformer tank is has to be proportional to twice the supply frequency. To determine this, Fourier analysis is carried out on vibration signal and the fundamental frequency of the vibration signals is 100 Hz. Figure 6 illustrates that the transformer core frequency signals prominently display a dominant frequency of 100 Hz.

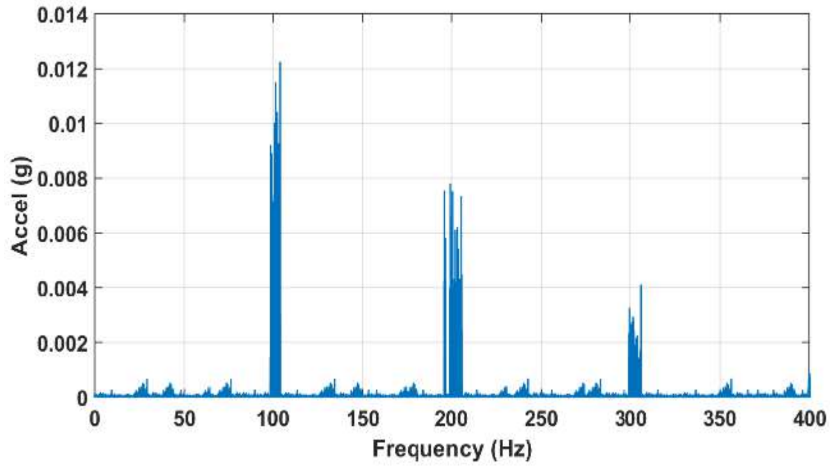

Figure 6 Core Vibration signal in frequency domain at full load current.

Several machine learning algorithms, including gradient-based optimization methods, exhibit faster convergence when the input features share a similar scale. In cases where the features have significantly different scales, the algorithm might experience delayed convergence or even fail to converge. Thus, collected vibration data is normalized between -1 and 1. Before commencing the model training, the dataset was partitioned into segments. Given that the data has a sampling frequency of 12 kHz, implying 12000 data samples per second, a single sample alone lacks adequate information about the vibration signal, including aspects such as amplitude, peak-to-peak value, behaviour, or frequency of one cycle. To address this limitation, a decision was made to consider 3000 samples, equivalent to 0.25 seconds of data, as one segment. This choice was made to encompass approximately four cycles of the vibration signal within each segment. By doing so, the intention is to provide the model with a more comprehensive view of the vibration data, enabling it to capture essential characteristics and patterns associated with input signal. This segmentation strategy aims to enhance the model's ability to discern meaningful features and relationships within the vibration data during the training process.

The CNN models employed for fault classification underwent training based on hyperparameters. Performance assessment during the training and validation processes was conducted by evaluating the classification accuracy. The model with the highest accuracy was identified and saved as the final model, featuring specific hyperparameters. The top 10 models were presented in Table 3, sorted by validation accuracy value. The final 1D-CNN structure encompasses five convolution layers, four pooling layers, and two fully connected layers. Figure 7 illustrates the detailed network structure parameters.

Table 3 Designed 1D-CNN architectures for core fault based on accuracy. (after testing 30 models)

S. No	Batch Size	Kernel Size	Layers Numbers	Filter	Accu-racy (%)
1	32	4	5	[128,64,32,16,8]	99.78
2	32	7	6	[256,128,64,32,16,8]	99.74
3	32	5	5	[128.64,32,16,8]	99.67
4	64	6	5	[128.64,32,16,8]	99.62
5	32	5	4	[64,32,16,8]	

99.57				
6	64	5	5	[128.64,32,16,8]
99.57				
7	64	7	4	[64,32,16,8]
99.54				
8	32	6	5	[128,64,32,16,8]
99.51				
9	64	4	5	[128,64,32,16,8]
99.51				
10	32	7	4	[64,32,16,8]
99.51				

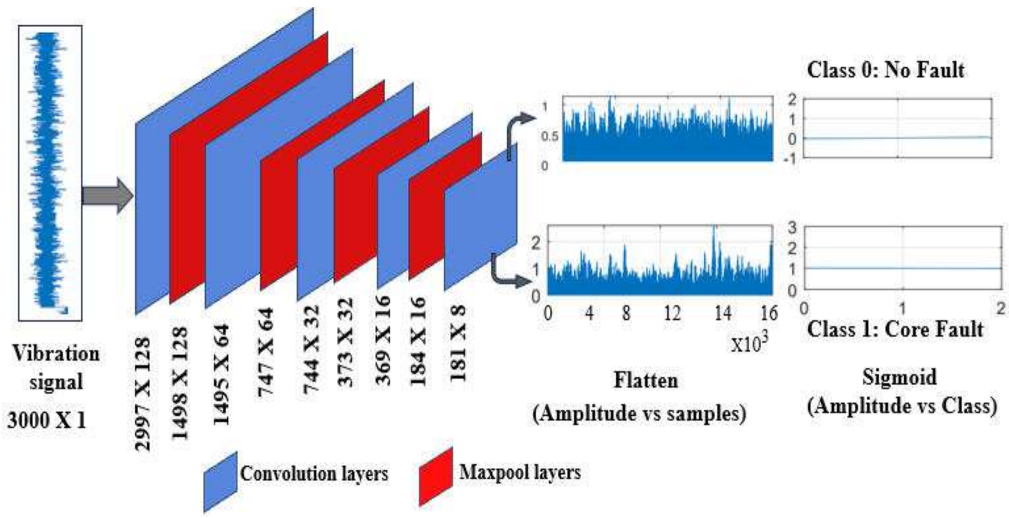

Figure 7 Proposed 1D-CNN network structure for transformer core fault diagnosis.

Within each convolutional layer (depicted in Figure 7), the suitable number and size of the convolution kernel conduct one-dimensional convolution operations. In this context, the rectified linear unit (ReLU) serves as the activation function for all five convolution layers. The stride is a hyperparameter that dictates the step size taken by the convolutional filter as it traverses the input volume. In convolution layers, the typical stride value is 1, while in max pooling layers, it is o set to 2.

Primarily, the input data are the one-dimensional signal that has a length of 3000. First convolution layers use 128 convolution kernels of size 4 × 1, second layer use 64 of size 4 × 1, third layer use 32 of size 4 × 1 layer, fourth layer use 16 of size 4 × 1 and fifth layer has 8 of size 4 × 1. The output feature maps generated from convolution layers (layer 1 to 4) are fed into the max pooling layer as an input. which carries out a 2x2 max-pooling operation. Within the flatten layer, the features extracted from the five convolution layers are expanded into a one-dimensional vector. The output layer comprises a single neuron. For this study, a sigmoid activation function is employed, representing binary classification post-training.

The effectiveness of the proposed architecture for core fault classification is assessed through training and testing processes. For this evaluation, 60% of the data is allocated to training, 20% for testing, and another 20% for validation. Figures 8 illustrate the training curves, showcasing the progression of validation and training loss, as well as validation and training accuracy throughout the training epochs. The depicted curves indicate that the proposed architecture effectively learns from the provided data within a few epochs without overfitting.

Ultimately, the achieved average classification accuracy for the proposed architecture is an impressive 99.78%. The comparative outcomes between the proposed methodology and existing works are presented in Table 4.

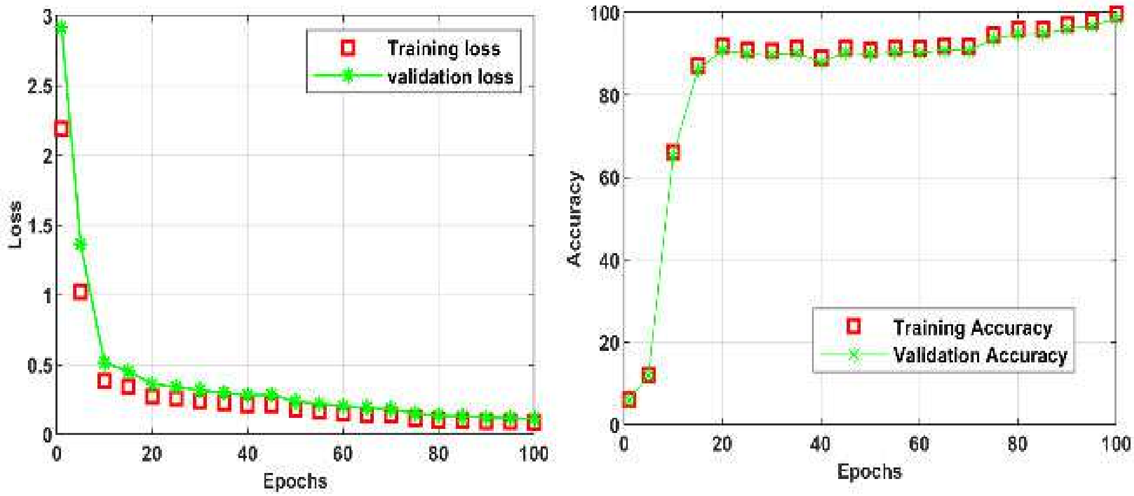

Figure 8 1D-CNN Training curves for core fault classification

Table 4 Comparison Table of Related Work

S.No	Reference	Year	Method	Accuracy%
1	[17]	2021	KNN	93.70
2	[18]	2019	SVM	94
3	[19]	2021	Fuzzy logic	93.85
4	[20]	2020	Fuzzy+RL	99.70
5	proposed	2024	CNN	99.78

5.CONCLUSION AND FUTURE WORK

In the proposed work, a single-channel vibration measurement system to capture and measure the vibration signals occurring on the dry tank surface of an operational transformer. When a mechanical breakdown occurs in the transformer core such as bending of limbs in out-of-plane direction were recognized through this vibration signal and a one-dimensional convolutional neural network (1D-CNN) architecture is proposed for core fault identification. The experimental vibration dataset of the transformer core serves as the basis for evaluating the effectiveness of the proposed CNN model. The important classification accuracy of 99.78% is justified by experimental results. The straightforward and space-efficient construction of adaptive 1D CNNs, which engage in linear 1D convolution through scalar multiplications and additions, renders them well-suited for real-time and economical hardware implementations.

REFERENCES

[1] Akhmetov, Y.; Nurmanova, V.; Bagheri, M.; Zollanvari, A.; Gharehpetian, G.B. A new diagnostic technique for reliable decisionmaking on transformer FRA data in interturn short-circuit condition. IEEE Trans. Ind. Inform. 2020, 17, 3020–3031.

[2] Wu, Z.; Zhou, L.; Wang, D.; Zhou, M.; Jiang, F.; Yu, X.; Tang, H.; Zhao, H. Feature Analysis of Oscillating Wave Signal for Axial Displacement in Autotransformer. IEEE Trans. Instrum. Meas. 2021, 70, 1–13.

[3] Ye, Z.; Yu, W.; Gou, J.; Tan, K.; Zeng, W.; An, B.; Li, Y. A Calculation Method to Adjust the Short-Circuit Impedance of a Transformer. IEEE Access 2020, 8, 223848–223858.

[4] Alehosseini, A.; Hejazi, M.A.; Mokhtari, G.; Gharehpetian, G.B.; Mohammadi, M. Detection and classification of transformer winding mechanical faults using UWB sensors and Bayesian classifier. Int. J. Emerg. Electr. Power Syst. 2015, 16, 207–215.

[5] Mariprasath, T.; Kirubakaran, V. A real time study on condition monitoring of distribution transformer using thermal imager. Infrared Phys. Technol. 2018, 90, 78–86.

[6] Research on Power Transformer Test Plan based on Big Data Processing Algorithm Kaiping Li1, Dehong Liu1, Ye Li1 1Wenshan Power Supply Bureau of Yunnan Power Grid Co., Ltd, China, 663099.

[7] F. Zhang, S. Ji, Y. Shi, C. Zhan, and L. Zhu, "Investigation on vibration source and transmission characteristics in power transformers," *Applied Acoustics*, vol. 151, pp. 99–112, 2019.

[8] Y. C. Ma, J. Mo, G. Z. Wang et al., "Test and Experimental Study on vibration and noise characteristics of distribution transformer," IOP Conference Series: Earth and Environmental Science, vol. 526, Article ID 012136, 2020.

[9] Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9**, 611–629,2018.

[10] Liu, H.; Zhou, J.; Zheng, Y.; Jiang, W.; Zhang, Y. Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders. ISA Trans. 2018, 77, 167–178.

[11] A. Zollanvari, K. Kunanbayev, S. Akhavan Bitaghsir and M. Bagheri, "Transformer Fault Prognosis Using Deep Recurrent Neural Network Over Vibration Signals," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-11, 2021, Art no. 2502011, doi: 10.1109/TIM.2020.3026497.

[12] Hong, K.; Jin, M.; Huang, H. Transformer winding fault diagnosis using vibration image and deep learning. IEEE Trans. Power Deliv. 2021, 36, 676–685.

[13] Bicen, Propositional logic concept for fault diagnosis in complex systems. Engineering Science and Technology, an International Journal. Y. (2020).

[14] B.Pfützner, H. Relevance of magnetostriction and forces for the generation of audible noise of transformer cores. IEEE Trans. Magn. **2000**, 36, 3759–3777.

[15] Ramirez-Mendoza, Ricardo A.,Qiang, He,Jingkai, Nie,Songyang, Zhang,Weimin, Xiao, "Study of Transformer Core Vibration and Noise Generation Mechanism Induced by Magnetostriction of Grain-Oriented Silicon Steel Sheet" 1070-9622 , 2021.

[16] Eren, L.; Ince, T.; Kiranyaz, S. A Generic Intelligent Bearing Fault Diagnosis System Using Compact Adaptive 1D CNN Classifier. J. Signal Process. Syst. 2019, 91, 179–189.

[17] Kherif, O., Benmahamed, Y., Teguar, M., Boubakeur, A., and Ghoneim, S. S. M. (2021). Accuracy improvement of power transformer faults diagnostic using knn classifier with decision tree principle. IEEE Access 9, 81693–81701, 2021.

[18] Shang, H., Xu, J., Zheng, Z., Qi, B., and Zhang, L. (2019). A novel fault diagnosis method for power transformer based on dissolved gas analysis using hypersphere multiclass support vector machine and improved d–s evidence theory. Energies 12, 4017. 2019.

[19] Poonnoy, N., Suwanasri, C., and Suwanasri, T. (2021). Fuzzy logic approach to dissolved gas analysis for power transformer failure index and fault identification. Energies 14, 36. 2021.

[20] Malik, H., Sharma, R., and Mishra, S. Fuzzy reinforcement learning based intelligent classifier for power transformer faults. ISA Trans. 101, 390–398. 2020.