Transformer Fault Prognosis using Vibration Signals and Forecasting Techniques

Askat Kural, BEng in Electrical and Electronic Engineering

Submitted in fulfilment of the requirements
for the degree of Master of Science
in Electrical and Computer Engineering

School of Engineering and Digital Sciences
Department of Electrical and Computer Engineering
Nazarbayev University

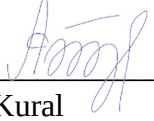
53 Kabanbay Batyr Avenue,
Nur-Sultan, Kazakhstan, 010000

Supervisor: Mehdi Bagheri
Co-supervisor: Amin Zollanvari

April 2022

DECLARATION

I hereby, declare that this manuscript, entitled “Transformer Fault Prognosis using Vibration Signals and Forecasting Techniques”, is the result of my own work except for quotations and citations which have been duly acknowledged. I also declare that, to the best of my knowledge and belief, it has not been previously or concurrently submitted, in whole or in part, for any other degree or diploma at Nazarbayev University or any other national or international institution.



Name: Askat Kural
Date: 03.05.2022

Abstract

Vibration signature analysis is considered as an advanced and economical methods to evaluate transformer operating condition and mechanical integrity. Transformer condition monitoring and fault prognosis have been investigated and discussed from the second decade of this century, while the modern and innovative approaches such Artificial Intelligence (AI), are very quickly under development in different applications and they have employed recently in this field.

In this thesis, we first discuss the advantages and disadvantages of the conventional techniques along with mathematical/practical approaches that are capable to monitor transformer working condition effectively. Afterwards, analytical approach to model the transformer vibration is conducted and vibrational model of transformer is provided. Then, for fault prediction deep neural networks, namely convolutional neural network (CNN) architectures, were employed. Different experimental works to emulate various crucial faults in transformer operational condition is conducted, and recorded vibrational data will be analyzed using CNN models. The experimental results are also compared with mathematical modeling by validating the recorded data approach in this study.

In this regard, two case studies were examined in the experimental laboratory. Firstly, transformer voltage excitation test was conducted to collect transformer vibration data using experimental transformer with different loads. Secondly, the transformer vibration waveforms were recorded by emulating transformer inter-turn short circuit using variable resistor. The focus of the thesis is observing the possibility of applying deep neural networks, in particular, CNNs, for time-series vibration signals to predict the transformer excitation and turn-to-turn short circuit fault. Therefore, 1D-CNN architecture was constructed by selecting the best predictive model from a prespecified space of hyperparameters.

The constructed CNN model for transformer excitation voltage exhibited a remarkable performance with RRSE of 4.49% and RAE of 2.49%. At the same time, the model constructed for the inter-turn short circuit fault classification achieved a remarkable accuracy of 99.86%. Finally, the achieved results were compared with previous studies and discussed in detail.

Acknowledgements

I would like to express my immense gratitude towards my supervisor, Prof. Mehdi Bagheri, and my co-supervisor, Prof. Amin Zollanvari, who supported me and made this thesis possible, for their supervision throughout the whole academic year. I appreciate all the time they spent for planning and discussion hours of works done under this thesis, their professionalism and patience shown during this master program.

I would like to thank Nazarbayev University for giving such a phenomenal opportunity to gain professional research and engineering skills.

Finally, I would like to appreciate my family, friends and colleges in the laboratory for sharing their guidance, knowledge and help.

Contents

List of Abbreviations & Symbols

AI	Artificial Intelligence
ANN	Artificial Neural Network
Bi-GRU	Bidirectional Gated Recurrent Unit
Bi-LSTM	Bidirectional Long Short-Term Memory
CNN	Convolutional Neural Network
DL	Deep Learning
FRA	Frequency Response Analysis
GRU	Gated Recurrent Unit
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
ML	Machine Learning
MLP	Multiplayer Perceptron
MSE	Mean Square Error
NA	Not Available
NN	Neural Network
RAE	Relative Absolute Error
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
RRSE	Root Relative Squared Error
SCI	Short-circuit Impedance
1D-CNN	One-Dimensional Convolutional Neural Network
2D-CNN	Two-Dimensional Convolutional Neural Network

List of Tables

List of Figures

Chapter 1 – Introduction

In the new century, asset management is one of the main and important tasks for industrial companies [1]. Instead of re-establishment of assets after catastrophic collapses and failures, intelligent predictive system is becoming beneficial strategy for risk management and fault prognosis systems that are able to distinguish the faults using monitoring systems with various sensors. The installation and maintenance of predictive assessment can cost far less than the restoration of expensive assets such as power transformers [2], [3].

In electric power systems, power transformer is one of the most crucial and costly equipment. Moreover, it is continuously in service under various mechanical and electrical stresses as well as in a variety of climate conditions that cause frequent and serious faults for power transformer. Beside on this, winding deformation of transformers is main consequence of the various stresses and faults in the electrical power system [4]. For instance, winding deformation provoked by short-circuits causes transformer out of service, which tends to 15 percent of all transformer failure [5].

1.1 Background Information

Real-time monitoring of transformer's internal mechanical stability is essential to prevent the electrical and mechanical parts from disbalance of their steady state working condition [6]. Based on this fact, there are different types of online and offline monitoring methods which have been implemented for observing and evaluating mechanical integrity of transformers [7]–[11]. One of the crucial methods for winding analysis is the Frequency Response Analysis (FRA) and it is still introduced as offline transformer condition monitoring method [7]. From one perspective, FRA can be performed in online basis, but it needs to be implemented perfectly [1]. From the other

perspective, the transformer mechanical integrity can be analyzed and evaluated using different methods such as online transformer sound analysis [12], short circuit impedance measurement [13], deformation coefficient [14], communication-based techniques and a locus diagram based methods for winding deformation detection of transformers [15]-[17]. Additionally, the transformer fault prognosis can be accomplished by vibration analysis that evaluates the repetitive movement of transformer inner parts. This approach does not require any complicated and expensive setup. It is possible to monitor mechanical integrity of transformer through analyzing the vibration signature of windings in real-time [1], [2], [6].

Modern technological revolutions are upgrading traditional control and classification methods into smart systems in the era of Industry 4.0. Nowadays, the simulation of human intelligence by machines are dramatically developing and becoming part of human's life. The integration of artificial intelligence (AI) to the industrial technologies, especially, in the power engineering sectors is making profound impacts on prolonging lifetime of costly equipment by prognosis failure in power systems [18], [19]. With the deployment of machine learning (ML) techniques, fault diagnosis process can be accomplished within milliseconds by preventing the equipment from significant damages [6]. There are numerous analysis mathematical tools to be able to analyze data for predicting a fault in power equipment, particularly, in power transformers such as regression methods [1], big data analysis and classification tools [6],[20].

The fault detection methods mainly have three large groups based on the applied techniques, precisely, model-based, signal-based and knowledge-based approach. Model-based approach will create a physical model involving analytical formulas for normal operating conditions and degradation process of a specified equipment [21]. The model developed by this approach is straightforward and very efficient since the model is used to monitor the consistency of measured parameters of the in-service equipment with predicted outputs. However, the full and

thorough understanding of specific equipment is essential. The signal-based method contains four main classes such as time-domain, frequency-domain, enhanced frequency and time-frequency analysis that can be employed as the signal processing techniques [22]. The computational cost of the method can be increased by complexity of tools, which tends to enhance in fault detection capability. The knowledge-based approach does not rely on domain experts and can be implemented based on symbolic and machine learning intelligence using collected parameters and big data. In other words, artificial intelligence (AI) introduces the knowledge-based approach including fault tree, diagraphs, unsupervised learning systems and supervised learning systems. Integration of signal processing methods with AI application is suitable for complex fault detection and prognosis problems, whereas the performance of this technique depends on the training data of algorithms and the quality of the selected features [22].

1.2. Aims and Objectives

This thesis aims to study and develop a technique for transformer condition and fault prognosis using vibration signature, the ML algorithms and forecasting techniques to predict the transformer fault based on the vibration patterns are conducted. This research has specifically focused on transformer faults such as voltage excitations (overvoltage and undervoltage) and inter-turn short-circuit.

The thesis reviews in detail the existing various studies on transformer mechanical assessment and winding conditions within analytical and theoretical approach. Moreover, research on transformer fault prognosis methods is divided into two main topics and reviewed. Firstly, the previous and recent literature about the fault modelling of the power transformer is circumstantially

discussed. Then the prognosis method using ML techniques is studied to apply suitable forecasting method for the prognosis of faults such as inter-turn short-circuit and voltage excitations.

Hence, the objectives of this research work are:

- to provide in detail an analytical approach for transformer vibration modeling;
- to perform the transformer under- and over-voltage excitation conditions along with the inter-turn fault detection, and evaluate them with the prognosis techniques using experimental setup vibration signature data in different operational and faulty states;
- to develop an algorithm for predicting the fault using vibration signal data;
- to develop a prognosis and predictive model for transformer faults using Convolutional Neural Network (CNN).

1.3. Literature Review

In this section, all the techniques of fault modelling are analyzed and advantages and drawbacks of the specific model are introduced.

1.3.1. Vibration Analysis

Repetitive movement of transformer inner part, or the movement of active part in transformer, can be considered as the transformer vibration, which has the reference position and moves around starting state [1],[2]. Transformer vibration can be illustrated by mechanical parameters such as acceleration, winding displacement and velocity. All the methods take reference position when transformer reaches once out of service moment for interpreting vibration level. According to [2], the acceleration is common choice as a parameter of interest for vibration analysis to make a proper decision as the mechanical parameter for analyzing and modelling vibration signature of transformer. Moreover, this technique is one of the economical methods for real-time

monitoring of transformer mechanical integrity as it can be integrated to the transformer inner part without any complicated setup while the transformer is in-service. However, it has its own drawbacks during data collection and monitoring process because the vibration signals can be combined with other vibrational noises dependent on various measurement environment.

Transformer vibration signal is the combination of mixed vibrations caused by environmental noises, background and noises of equipment near to transformer or accelerometer. Bagheri *et al*. [1] believes that the combination of vibration signal with different noise signals such as transformer active vibration parts with oil pump, cooling system and tap-changer, may cause problem for the development of high-quality vibration monitoring and interpretation system. Unpredictable behavior of the different signals will make the vibration signature analysis complex and challenging [2],[5].

To interpret vibration signal of transformer, the analytical model of each inner part has to be derived and discussed in detail to understand the vibration signature and its behavior. According to literature [1], [2], [15], [23]-[25], mainly three main parts of transformer are necessarily required to be modelled using mathematical derivations. They are core [1], [2], [15], winding [1], [2], [15], and the tank along with connected parts [23]-[25].

1.3.2. Frequency Response Analysis

Frequency response analysis (FRA) is considered as one of the advanced approaches in power engineering field. Effectiveness of this method is detecting transformer winding deformation and core displacement with the high accuracy, quickness, economy and indestructibility [9]. This method is based on swept frequency signal attenuation through the transformer winding. The technique helps to interpret faults such as turn-to-turn short-circuit

discussed in [5],[26], voltage excitation tested in [13],[14] and winding deformation from different faults [7]-[11], [26].

Zhao *et al.* [10] believe that the frequency response analysis is effective and powerful tool for mechanical deformation diagnosis of power transformer, however, still its interpretation has problems and challenges in its application. The cause of that is some frequencies can be produced by external disturbances and can be taken and recognized as a fault, incorrectly. Thus, this study proposes an improved frequency response analysis-based application that can be performed in binary morphology and extreme variations.

According to Bagheri *et al*. [11] frequency response analysis (FRA) has been used as the effective and economical diagnosis method in electrical machineries. Particularly, it can provide more accurate information about winding deformation of power transformers than short-circuit impedance (SCI) method. In [11], the off-line SCI and FRA measurements were conducted for a failed 400-MVA step-up transformer with the failed B phase winding. Author concludes that FRA measurement is faster and more economical, and it can provide more detailed information as compared to SCI which is recommended by IEC Standard 60076–5, Ed. 3.0, 20.

Another study by Bagheri *et al*. in [5] described a mathematical modeling for the turn-to-turn short circuit fault that can be explored in frequency response data and vibration signal. Authors mentioned that this fault is one of the common faults in transformers and it needs to be prevented at early stages before the fault is fully propagated. Otherwise, the maintenance of transformer is very challenging task and uneconomical to perform. Authors in [5] purpose to recognize transformer fault before it can fully take place, in very initial stages, therefore they have recommended to use a correlation between FRA and vibration analysis. The experimental work was conducted over a 2 kVA three-phase transformer as a routine distribution transformer. Consequently, both methods were sensitive to detect and identify turn-to-turn fault in transformer.

However, authors concluded that vibration signal detection in oil-filled transformers can be challenging task due to mentioned earlier reasons [1], [2]. In addition, according to [11], winding vibration signature is more sensitive to turn-to-turn fault than core vibration.

1.3.3. Lissajous Figure

According to Liao *et al*. [17], the voltage and windings' current values of transformer can support to monitor transformer winding deformation using Lissajous figure. It is possible to shape an ellipse using voltage and current values. The aim of the study was to identify effect of load on Lissajous figure. Liao *et al*. [17] have shown in results that Lissajous figure is proportional to load and the ellipse can be magnified or decreased by the load variation. However, power factor of load cannot affect the ellipse. As it can be seen from the Figure 1.1, after applying concepts of the following method, the faulty and healthy windings of experimental transformer have shown different diagrams in the Lissajous figure.

1.3.4. Locus Diagram

FRA is considered as one of the popular tools to detect mechanical deformation of the power transformer active parts. However, transformer cannot be tested while it is in service or connected to the system, which can cause the interruption to the power grid [16].

Abu-Siada *et al*. in [16] provided a new online technique to detect the internal faults of power transformers just using voltage and current values. This method does not need any complicated setup or equipment, it uses the metering devices of transformers. The method can be conducted using voltage-current locus diagram. The proposed technique was examined for different types of faults and it shows possible fault types correctly.

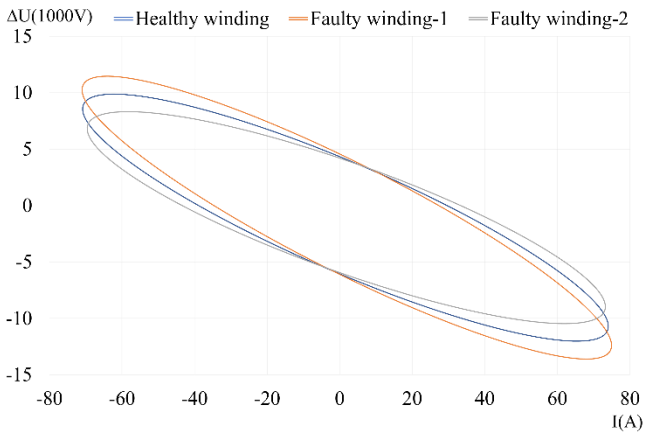


Figure 1.1: Lissajous figures of healthy and faulty winding [17].

1.3.5. Forecasting Techniques

Integration of modern technologies such as internet of things (IoT) and cloud computing into power engineering equipment is the next level of upgrading electrical equipment with high accuracy forecasting techniques, which tends to prolong in service lifetime of equipment, especially, power transformers [1], [2], [15]. In order to perform this, the integrated technologies should predict the fault progress signature in early stages [18], [27]. Depending on complexity and accuracy of the fault prediction methods using ML, previous studies were divided into two main categories: conventional methods and ANN [6]. The reason of naming it as “conventional” is simplicity of model construction and lower accuracy than artificial NN. These methods are the model tree [28], [29] and regression [1], [2].

Nowadays, parallel computation hardware has improved ANN technologies into desirable levels using a series of algorithmic progresses and achieved breakthrough results in a wide range of areas such as classification [6], [22], image recognition, optimization [30], condition monitoring [19] and natural language processing [20]. Early fault detection of essential equipment in power system is the main goal of industrial companies. For this purpose, ANN are widely in use to increase the quickness of fault prediction and accuracy specific fault detection [22]. Tian in [19]

claims that ANN methods does not assume any analytical model to perform equipment health condition prediction and it aims prediction modelling based on the collected condition monitoring data and neural networks. The adaptability, non-linearity and arbitrary function approximation ability of the ANN help to increase the efficiency of fault prediction and make it one of the promising tools for equipment health condition monitoring.

In several studies [31]-[34], different approaches of ANN technique are proposed. In [31], ANN-based approach with 13 training data and nine validation data patterns was applied to diagnose stator faults in induction machines. The accuracy of the predicted fault condition was more than 97.6%. Li *et al*. [32] developed motor bearing fault diagnosis using a NN methods with time and frequency-based features. The average accuracy rate of fault detection model with different hyperparameters (hidden layers and neurons) was between 88.75% and 96.25%. In [33], classification accuracy of healthy and damaged bearings was 85% applying two types of neural detectors such as feed-forward multilayer perceptron (MLP) and self-organized Kohonen's network. Lastly, Tung *et al*. [34] developed a classifier based on adaptive neuro-fuzzy inference system with 180 training and 90 test samples for six different fault diagnosis of induction motors. The predicted parameters were vibration and current signals, and the classification accuracy was 91.11% and 76.67%, respectively.

In addition, CNN has the inherent property that can unite the feature extraction and classification methods into one adaptive structure to monitor transformer condition accurately and detect the fault early and very fast [22]. This method can be conducted using vibration signal data that have time series vibration signature [21].

1.3.6. Convolutional Neural Networks

In recent years, DL has grown in popularity and many achievements have been made in a short period of time. One of the DL approaches is Convolutional Neural Network (CNN). CNN mostly used for image recognition, but nowadays there is the framework for time series classification and segmentation. In [22], authors have integrated adaptive 1-D CNN model and implemented a novel motor condition monitoring system. The advantage of the approach is combining the two main important steps of traditional and comprehensive fault detection methods such as feature extraction and classification into a single block of learning system. Ince *et al*. [22] believe that proposed monitoring system can be employed using any motor data and have the ability to train the model by extracting optimal features to get high classification accuracy. In this approach, convolutional layers of 1-D CNN with the back propagation training are used to extract optimized features from equipment condition data and perform fault classification with MLP layers. The effectiveness of the system was tested using motor current data and real-time condition monitoring potential of motor was demonstrated. The fault detection accuracy of the proposed method was more than 97%, which shows the high efficiency of the developed CNN model. The discussed method was tested for time series data, whereas CNN mainly employs to image data that has different characteristics than time series data. Liu *et al*. [21] explained that the main and important component of time series data is time, whereas image data generally involve RGB three channels and do not have time information. Therefore, time series data was converted to three-dimensional tensors. This success in results inspires the industry and it is the innovative opening that gives opportunity apply CNN to time series data [21].

Chapter 2 – Transformer Vibration Modeling

2.1. Transformer Core Vibration Modelling

Magnetostriction forces are applied to the ferromagnetic material causing the length alteration and vibration when the magnetic field is present in the core of the transformer. The several steady-state vibration oscillations with different frequencies can be created as the harmonic components and the fundamental component can be increased by the loosened sheets in the core, which also tends to appear high frequency components. The vibration signal pass through the transformer oil (if transformer is oil-immersed) to reach the transformer tank and tank walls start oscillation. The cause of the vibration is receiving orientation of magnetic domains in transformer core material. Under application of magnetic field magnetic domains are heading in one direction by themselves [1].

The relationship of the applied voltage and magnetic field in the core can be represented by equation (1),

$$U_0 \sin \omega t = -N_w \frac{d\phi}{dt} = -N_w A_c \frac{dB}{dt}, \tag{2.1}$$

where,

U_0 amplitude of the voltage-driving source;

ω angular frequency;

N_w number of winding turns;

B magnetic induction;

A_c cross-sectional area of single core arm.

Therefore, the magnetic induction can be formulated as

$$B = \frac{-U_0}{N_w A_c} \int sin\, \omega t dt \; = \frac{U_0}{N_w A_c \omega} cos\, \omega t \; = B_0\, cos\, \omega t, \tag{2.2}$$

where B_0 is the magnitude of magnetic induction and it is equal or less than saturation level of magnetic induction (B_s) [1]. The magnetic intensity (*H)* is related to magnetic induction with magnetic permeability (μ). Once the magnetic intensity reaches the highest value (H_c), the magnetic induction can be calculated by

$$B = \frac{B_s}{H_c} H. \tag{2.3}$$

By replacing (2) in (3),

$$H = \frac{H_c B_0}{B_s} cos\, \omega t, \tag{2.4}$$

the magnetic field intensity (*H*) can be calculated and the variation of *H* alternates the length of the core laminate [1]. The maximum movement of the core laminate can be obtained by,

$$\begin{aligned} x_{core} = \frac{dL}{L} = \frac{\lambda_s}{H_c^2}\int_{-H}^{H} H dH = \frac{2\lambda_s}{H_c^2}\int_{0}^{H} |H| dH = \frac{\lambda_s H^2}{H_c^2} = \frac{\lambda_s}{H_c^2}\frac{H_c^2}{B_s^2} B_0^2\, cos^2\, \omega t \\ = \frac{\lambda_s}{H_c^2}\frac{H_c^2}{B_s^2}\left(\frac{U_0}{N_w A_c \omega}\right)^2 cos^2 \omega t = \frac{\lambda_s U_0^2}{B_s^2 N_w^2 A_c^2 \omega^2} cos^2 \omega t \end{aligned} \tag{2.5}$$

where, λ_s is the value of maximum magnetostriction. Lastly, the acceleration of core laminate can be calculated by

$$\ddot{x}_{core} = \frac{d^2 x_{core}}{dt^2} = -\frac{2\lambda_s L U_0^2}{B_s^2 N_w^2 A_c^2} cos\, 2\omega t. \tag{2.6}$$

According to (2.6), it can be seen that the fundamental frequency of core laminate vibration is proportional to fundamental frequency of the system. Therefore, if the system has the 50 or 60 Hz fundamental frequency, the core vibration frequency will be 100 or 120 Hz, respectively. However, there can be harmonic orders with more frequency value due to the loosened core laminates and they can be random numbers [1]. Also, the core vibration magnitude is proportional to the applied voltage square and can be increased if there is the excitation of applied voltage.

2.2. Transformer Winding Vibration Modeling

The applied electromagnetic force to the transformer winding part is induced by leakage magnetic flux, which flows through winding current, and main cause of the vibration in transformer winding [1]. Mechanical force direction in winding can be changed by changing electromagnetic force direction. The direction of the electromagnetic force depends on direction of leakage flux. According to [1], transformer winding is mechanically spring and can be derived using spring force model, which helps to model transformer winding vibration.

2.2.1. Free Vibration Without Damping Factor

Transformer winding can have simple string form without mechanical constraints and force applied to the string can be represented by,

$$F = -kx, \tag{2.7}$$

where, *k* is the spring factor and x is displacement, *F* is applied mechanical force. Thus, there is the spring weight and single impulse. Applying the Newton's second law,

$$\begin{aligned} ma = F' = W - F = W - (W + kx) \\ = \frac{W}{g}a = \frac{W}{g}\ddot{x} = W - (W + kx), \end{aligned} \tag{2.8}$$

where, *W* is the spring (winding) weight, *a* is acceleration factor, *g* is the gravity acceleration. Thus, natural motion of winding can be obtained by,

$$\frac{W}{g}\ddot{x} + kx = 0, \;\; \ddot{x} + \frac{g}{W}kx = 0. \tag{2.9}$$

Simplifying it,

$$\frac{gk}{W} = \alpha^2, \;\; \ddot{x} + \alpha^2 x = 0. \tag{2.10}$$

The solution is

$$x_1 = C_1 cos\alpha t + C_2 sin\alpha t, \tag{2.11}$$

where, C_1 and C_2 are initial condition constants and α is

$$\alpha = \frac{2\pi}{\tau_n}, \qquad \tau_n = 2\pi\sqrt{\frac{\delta_{st}}{g}}, \qquad f_n = \frac{1}{2\pi}\sqrt{\frac{g}{\delta_{st}}}, \tag{2.12}$$

where, f_n is natural oscillation frequency and damping factor is ignored [1].

2.2.2. Free Vibration with Damping Factor

The instantaneous force is the main cause of considering the damping factor for avoiding natural oscillation of windings and motion equation with damping factor (*c*) is

$$\frac{W}{g}\ddot{x} + c\dot{x} + kx = 0. \tag{2.13}$$

Simplification of (13) gives [1],

$$\ddot{x} + 2\beta\dot{x} + \alpha^2 x = 0, \qquad \frac{gc}{W} = 2\beta. \tag{2.14}$$

This linear equation (14) was solved by two ways [1]. First one is

$$\begin{gathered} x = e^{\mu t}, \\ \mu_1 = -\beta + j\alpha_1, \qquad \mu_2 = -\beta - j\alpha_1 \end{gathered} \tag{2.15}$$

where, *t* is the time.

The second solution has the winding displacement model below,

$$x = e^{-\beta t}(C_1 cos\alpha_1 t + C_2 sin\alpha_1 t)\,. \tag{2.16}$$

Chapter 3 – Methodology

In [1],[2],[6] several classification and prediction techniques were applied to forecast the fault using regression, classification methods, feature selection types for increasing the accuracy, Gated Recurrent Unit (GRU) and Long-Short Term Memory (LSTM) based Recurrent Neural Network (RNN). All the applied methods for predicting transformer faults were studied to reach the higher forecasting efficiency. In previous study [6], GRU and LSTM based RNN method is imposed to model and predict the transformer fault. In this work, it is planned to implement CNN along with a wide range of hyperparameters over new data obtained from a transformer in laboratory scale. In addition, the forecasting model with the best performance is going to be chosen according to accuracy of the fault prognosis.

3.1. Convolutional Neural Network

Various deep learning techniques have been previously implemented for vibration signal prognosis and transformer fault modelling [1, 2]. In [1], authors have estimated the essential features from unknown dataset and used them for different conventional regression techniques, such as support model regression, model tree, multilayer perceptron, and linear regression. All the features were extracted from vibration signals with a sum of sinusoids. However, in another study [6], prognosis models were constructed depending on the behavior of collected time series vibration signal dataset, with ignoring mathematical and analytical assumptions. In addition, the important patterns of the vibration signals were estimated and extracted over the model construction process. The predictive model was trained using LSTM unit of RNN architecture, specifically bi-directional LSTM.

In the present study, one of the popular approaches of DL, CNN architecture has been used to predict the transformer under and over excitations, and inter-turn short-circuit fault. CNN is a feedforward and constrained class of ANN that has similar ability based on cells in the human visual cortex [22], which helps it to be applied on data with grid-like topology such as time-series and images. Moreover, CNN has demonstrated state-of-the-art performances with dramatical achievements in pattern recognition [35].

CNNs have typically different layers that are responsible for specific tasks and targeted goals: convolutional layer, pooling layer and fully connected layer (Figure 3.1). The fundamental and essential parts of CNN, which distinguishes it from other ANNs, are convolutional layers [36]. In CNN, the kernels are used to extract features from input data by backpropagation algorithm using multidimensional input array with parameters (multidimensional array). The main operation of the convolutional layer is based on a mathematical linear operation, namely, convolution. In our study, the collected data is time-series vibration signals, and the kernel is slid in one-dimension in multivariate time-series input data, therefore, 1D CNN is used.

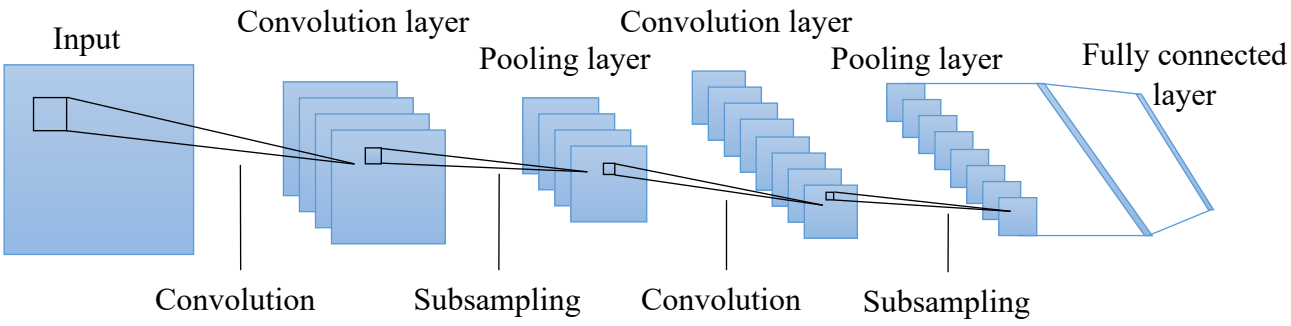


Figure 3.1. General structure of 2D-CNN.

3.1.1. 1D Convolutional Neural Network (1D-CNN)

1D-CNNs are mostly developed for applications analyzing sensor data in time-series domain. The main focus of this study is to analyze vibration signals of transformer under various faults and conditions, where data is measured in time domain.

The useful property of CNN is its feature extraction ability of convolutional and pooling layers. As it was discussed before, the input data of vibration signals with combined features (different frequencies and noises) is sent to convolutional kernel to extract and construct the feature vector. The pooling layer reduces the dimensions of the vector. The size of one-dimensional feature vector depends on convolutional layer parameters. The general architecture of 1D-CNN is illustrated in Figure 3.2.

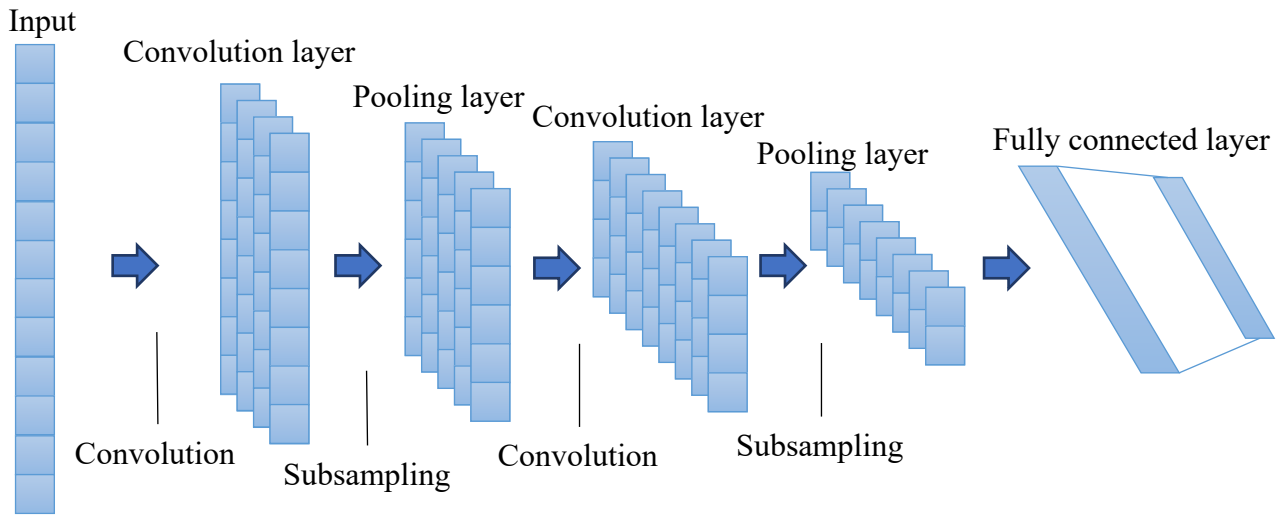


Figure 3.2. General structure of 1D-CNN.

3.2. Data Preparation

3.2.1. Transformer Under and Over Excitation

The experimental transformer, described in Section 4.1, was supplied by power source from 80 percent to 110 percent of rated nominal voltage. The core and winding vibration signals of

transformer under no-load, 50 percent load and full load were collected using three-dimensional Kistler sensor. However, vibration signals of only one dimension were chosen, because the core and winding vibrations accelerate mainly in one single dimension, i.e., the acceleration of one dimension is quite more dominant rather than other dimensions. This was also obvious in mathematical modeling of transformer core and winding. Hence, the input of the 1D-CNN model is the core and winding vibrations signals. The output of the model is the voltage rate supplied to the transformer.

Then the collected data was segmented by 50 samples per segment and there are 60 individual observations or segments in one second that is used for training 1D-CNN model as input raw data. After segmentation process, observations were randomly shuffled and 25 percent of the train dataset was decided to use as validation data. Rest of the training part was set aside as training data. Moreover, both training and validation datasets are normalized by dividing the maximum value of training data. In this way, all observations are scaled between -1 and 1. The model selection process is described in the next subsections in detail.

3.2.2. Transformer Inter-turn Short Circuit

Transformer inter-turn short circuit current measurement was discussed and described in Section 4.2.1. In this part, only the winding vibration signal was employed as an input dataset, because short circuit fault occurs in between turns (turn-to-turn short circuit) of the transformer. The output of the dataset was declared as “0” (non-fault) and “1” (fault). As it was discussed in Section 4.2.1, the turn-to-turn short-circuit current was initiated from 11 A to 15 A. The sampling rate of the winding vibration signals was remained unchanged for under and over excitation conditions, 3000 Hz. 50 samples of data per 1 segment and 60 segments in a second. Training and

validation sets are 75% and 25% of the total training data, respectively. All observations were randomly shuffled and all the values were normalized by dividing the maximum value of winding vibration. The input and out range was between -1 and 1. The training 1D-CNN model was selected by the assumptions below.

3.3. Training Model Selection

The space of CNN hyperparameters was defined based on the following assumptions:

- The batch sizes: 32 and 64;
- The epoch numbers: from 1 to 150 (without early stopping);
- The learning rates: 0.001 and 0.0001 with Adam optimizer;
- The number of convolutional layers (L): from 1 to 6;
- The number of filters in each layer: 2^{2+l} (increasing) and 2^{l+3-l} (decreasing) for l = [1:L];
- Kernel sizes: from 2 to 7;
- Activation function of each convolution layer: a ReLU function.

Each convolution layer has the max-pooling operation (except the last convolutional layer). Based on the aforementioneed assumptions, the cardinality of the hyperparameter space is 43,200: 2 (batch size) × 150 (epoch number) × 2 (learning rate) × 12 (convolutional layers or filters) × 6 (kernel size). For each element in this space of hyperparameters, a 1D-CNN model was trained and the model with best performance (lowest MSE) on the validation set was identified. Nonetheless, for each constructed models we also record other performance metrics such as RRSE, RAE and MAE. . All these metrics will be explained in the next subsections. Both case studies were trained with the same hyperparameter space, but with their own dataset.

3.4. Error Rate Calculation Methods

3.4.1. Mean Squared Error (MSE)

MSE is an absolute calculation of error and it is very efficient for the training the regression models. The loss function of the training model was chosen the MSE and it is defined as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(pv_i - av_i)^2, \quad (3.1)$$

where, *n* is the total number of the observations or segments in this study. MSE is the average of the squared deviation value of predicted value and actual value.

3.4.2. Mean Absolute Error (MAE)

MAE is the average of the absolute difference of the predicted and actual values and its formulation is obtained as,

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|pv_i - av_i|, \quad (3.2)$$

where *n* is the number of the observations.

3.4.3. Relative Absolute Error (RAE)

The next model evaluation type is RAE. This is one of the solutions to check the performance level of the predictive model. It takes the sum of absolute errors and normalizes it by dividing total absolute error.

$$RAE = \frac{\sum_{i=1}^{n}|pv_i - av_i|}{\sum_{i=1}^{n}|pv_i - \overline{av_i}|}, \quad (3.3)$$

where, n is the number of observations, $pv_i,\ av_i\ and\ \overline{av_i}$ are predicted, actual and mean of actual values, respectively.

3.4.3. Root Relative Squared Error (RRSE)

RRSE calculates the total squared error and normalizes it by dividing by the total square error.

$$RRSE = \sqrt{\frac{\sum_{i=1}^{n} |pv_i - av_i|^2}{\sum_{i=1}^{n} |pv_i - \overline{av_i}|^2}}, \quad (3.4)$$

where, n is the number of observations, pv_i, av_i and $\overline{av_i}$ are predicted, actual and mean of actual values, respectively.

Chapter 4 – Experimental Study

A set of vibration data is needed to construct the fault prognosis model. In this chapter, assembling of the experimental setup for data collection process is explained. As it was discussed before, the most common two types of the transformer faults, namely, transformer voltage excitations and inter-turn short circuit fault were chosen. For each fault, data was collected from separate experimental setups in the experimental laboratory.

4.1. Case Study 1: Transformer Under and Over Excitations

A three-phase transformer with three-phase resistive and reactive load connected to three-phase power supply is built as a practical setup and emulator of the small electrical grid system (Figure 4.1). The transformer characteristics is: 1200 VA, 42V/230V and 50-60 Hz (Figure 4.2). The accelerometer with the device DEWE 43 A is integrated to the transformer. DEWE 43 A is the 8 channel USB data acquisition system with the integrated DEWESOFTX software, which helps to convert analog vibration signals from the accelerometer to digital signals and visualize is using software device. The sampling rate of recording vibration data is 20 kHz. The sensor of the digital accelerometer is mounted to the transformer core and winding parts. Data acquisition part in this case was planned to collect the transformer vibration signal for under and over excitations with the injected voltage range between 80-110 percent of rated voltage with steps of 5 percent for three types of loads: no load, 50 percent load and full load, see Table 4.1. Vibration signals of core and winding of three-phase transformer were recorded in time domain using accelerometer (Kistler sensor).

The schematic diagram of the experimental setup is shown in Figure 4.3. Three-phase transformer was connected to the power supply using circuit breakers and CEM-U/EV equipment

was connected in parallel to the transformer input turns to record the injected voltage values. Variable resistor and motor were connected to the transformer via circuit breaker. The vibration sensors were mounted on top of core and winding parts of the transformer. During data acquisition process both data recorder equipment (DEWE 43 A and CEM-U/EV) recorded the vibration and input voltage data in parallel. Data collection process was conducted as following:

1. Setting the injected voltage to 80% of transformer rated input
2. Recording all data for 30 seconds for undervoltage and 10 seconds for overvoltage
3. Step up the injected voltage to 5% of rated input and repeat steps 2 and 3 until the injected voltage value reaches 110% of rated voltage

After data collection process data was collected into one file. Data preparation process is discussed in Chapter 5.

Table 4.1. Under and over excitation voltages of experimental transformer.

Excitation Voltage, %	80	85	90	95	100	105	110
Injected Voltage, V	320	340	360	380	400	420	440

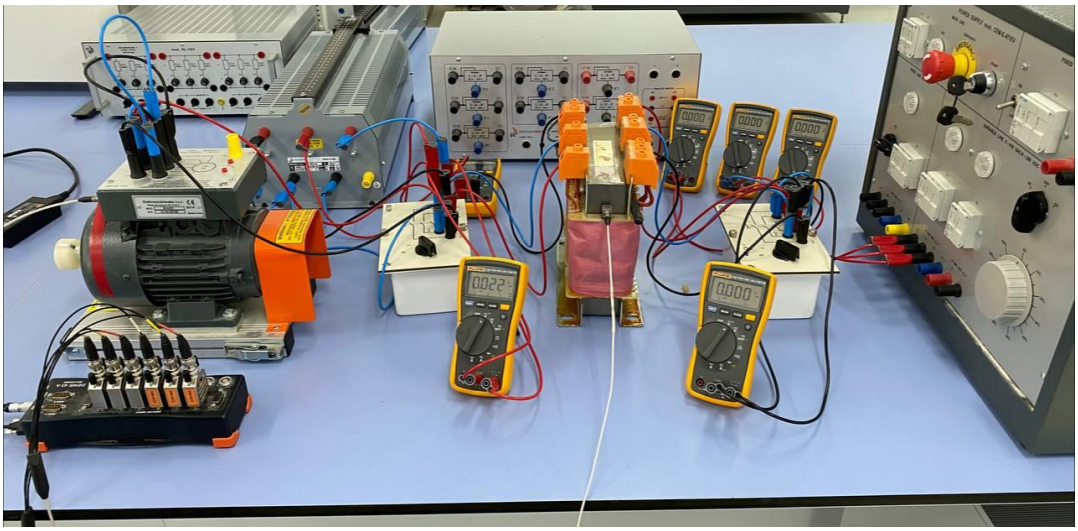


Figure 4.1. Experimental test set-up with three-phase transformer with active and reactive loads.

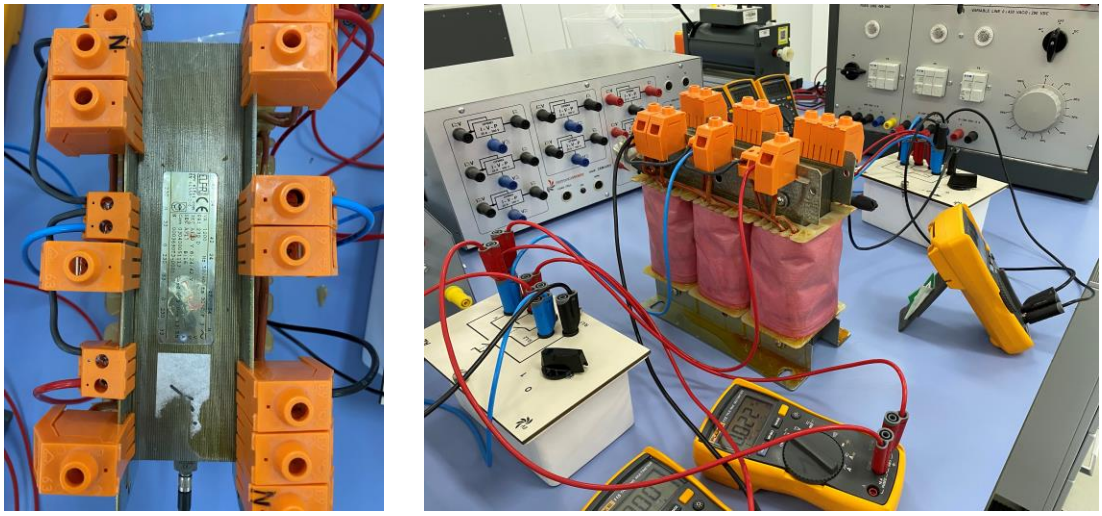


Figure 4.2. Three-phase experimental transformer.

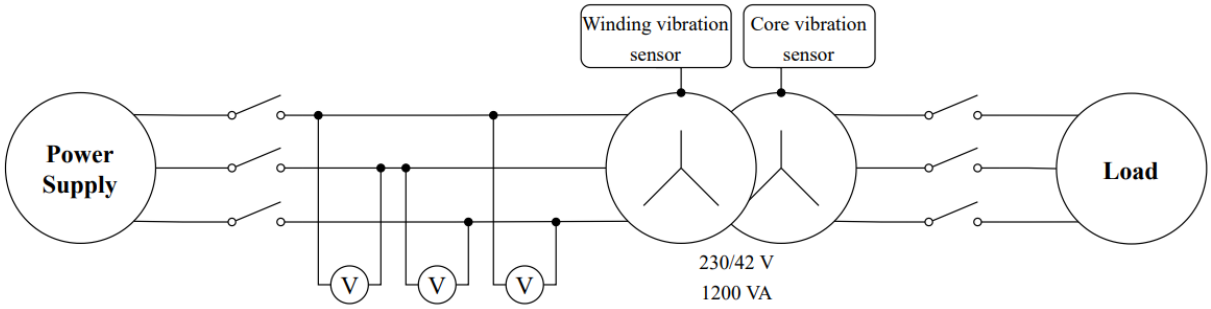


Figure 4.3. Schematic diagram of experimental setup.

4.2. Case Study 2: Inter-Turn Short-Circuit Fault

In the second case study inter-turn short-circuit fault was studied through the recording and analyzing the vibration data of transformer winding. The aim of this part study was to recognize the transformer inter-turn fault using the transformer windings' vibration signals. In this practical study a transformer was energized and loaded with different resistive loads to collect the data for training a new model. The data collection process was conducted using an open wounded 240/30/30 V, 252 VA single-phase transformer.

The rated nominal current for the transformer secondary side was 4.2 A. The primary side of the transformer was supplied by 240 V from the voltage sources and the secondary side of the transformer was connected to a variable resistive load. To increase the performance of the predictive model, the load value was increased step by step to obtain more different vibration signal and acceleration values as shown in Table 4.2. The two vibration sensors were mounted on top of the core and transformer windings. However, as the inter-turn short-circuit fault occurs over the transformer windings; only the winding vibration data was utilized for training the predictive model.

The experimental transformer has third winding and it was used to emulate turn-to-turn short circuit practically. The terminals of third winding were connected to a variable resistor to control the short-circuit current (Figure 4.4). Finally, the short circuit current was taken from 11 A to 15 A for the predictive model once the lower current values show the normal operational condition of transformer, and higher values should a real short-circuit in the secondary winding. Data preparation process for model construction is discussed in Chapter 5.

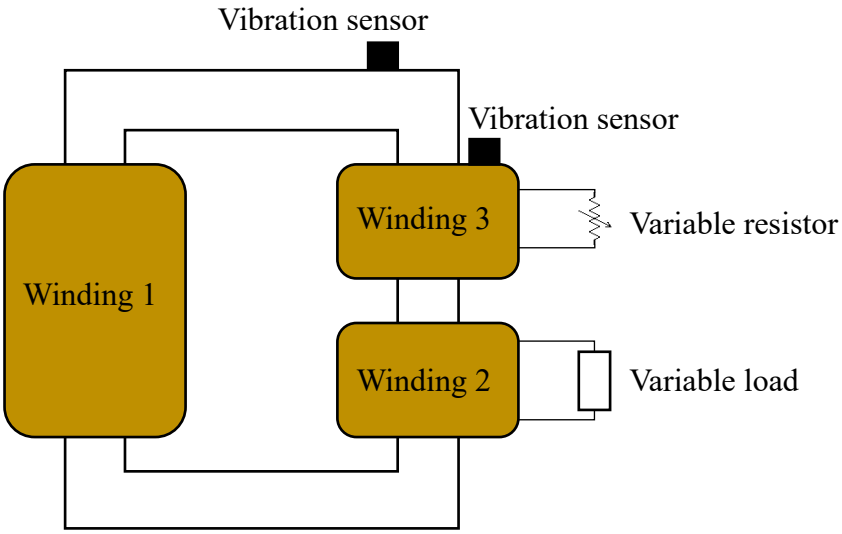


Figure 4.4. Single-phase transformer fault emulation scheme

Chapter 5 – Results & Discussion

In this thesis, two different datasets collected by experimental setup and equipment at the High Voltage Laboratory of Nazarbayev University. The data was collected from two different transformers for two different fault prognosis scenarios: the first one includes vibration signals under transformer voltage excitation and injected voltage values; the second dataset includes vibration signals under inter-turn short circuit fault and load values. In further subsections, the results of each prognosis model for each fault types are illustrated, analyzed and technically discussed.

5.1. Transformer Voltage Excitation Prognosis

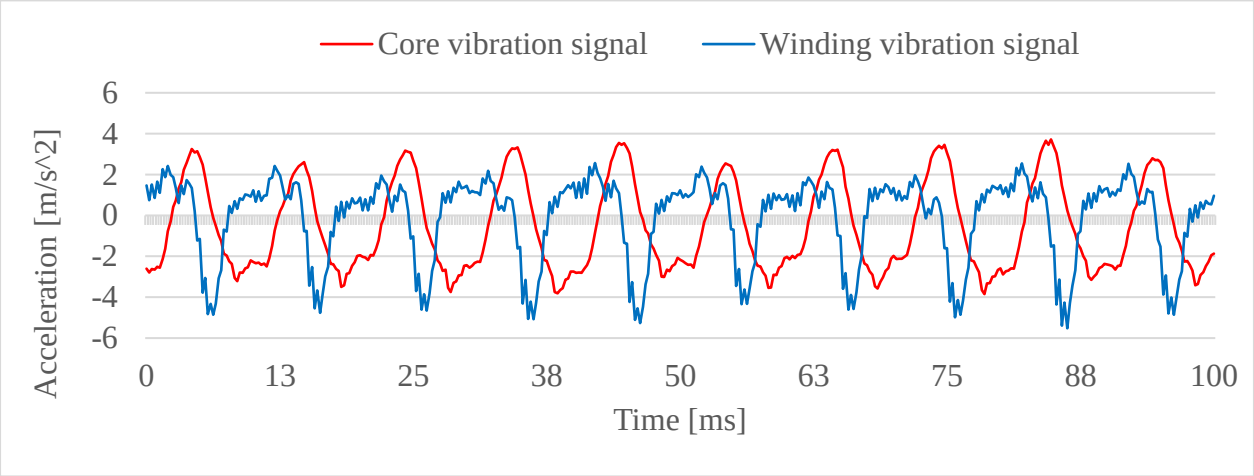
Using the experimental setup, discussed in Section 4.1, overvoltage and undervoltage data together with transformer vibration signals were measured and collected into one dataset table. It has two input features and one single output. The list of the features is shown below:

- Transformer winding vibration signals
- Transformer core vibration signals
- Transformer injected voltage (under, rated and over voltage values)

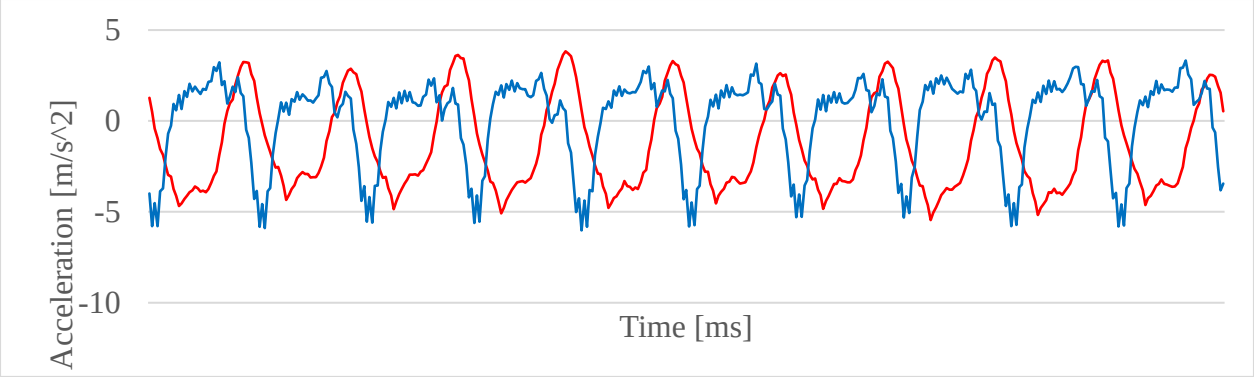
5.1.1. Data Collection and Preparation

Data was recorded by two different equipment at the laboratory: DEWE 43 A with Kistler sensor was used for measuring vibration signals of transformer core and winding under different load and injected voltage, and injected voltage values were recorded by CEM-U/EV equipment. The sampling rate of the vibration signals measured by DEWE 43 A data acquisition system (DAQ) using Kistler sensor (accelerometer) was 20 kHz. The injected voltage was in the range of 80 to 110% of rated voltage with steps of 5%. In each step of input voltage, the vibration data was

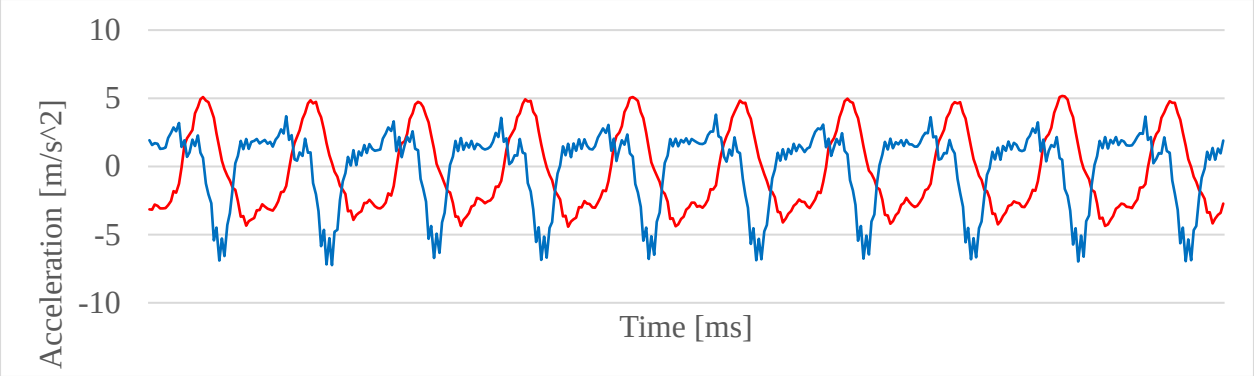
collected for 30 seconds until including rated voltage (100%). From 105% of the nominal voltage, time period for collecting data was decreased to 10 seconds to avoid any risks and hazards. Using DEWESOFTX software, the vibration sampling rate was decreased to 3000 Hz, which means 3000 data samples per second were recorded. The first 100 milliseconds of collected core and winding vibration signals for each step of 80% to 110% under full load condition was demonstrated in figures *(a)* to *(g)* in Figure 5.1, respectively.



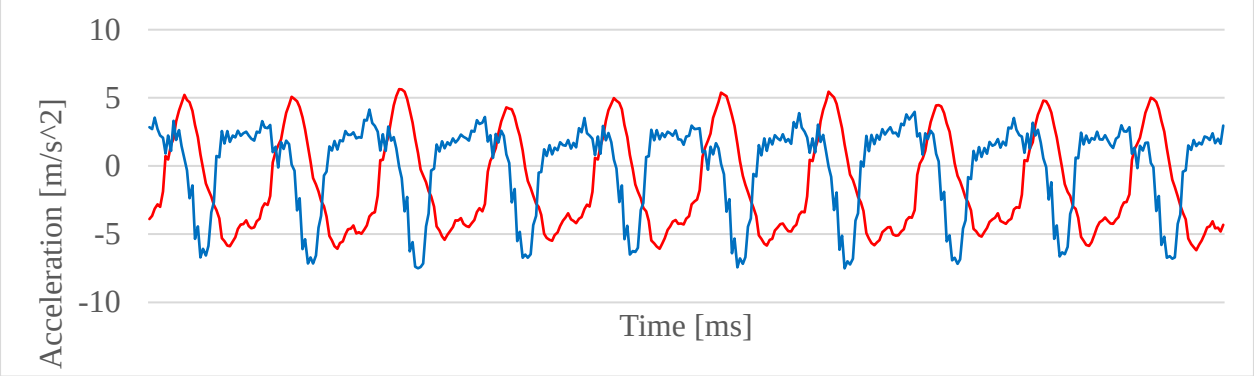
(a)



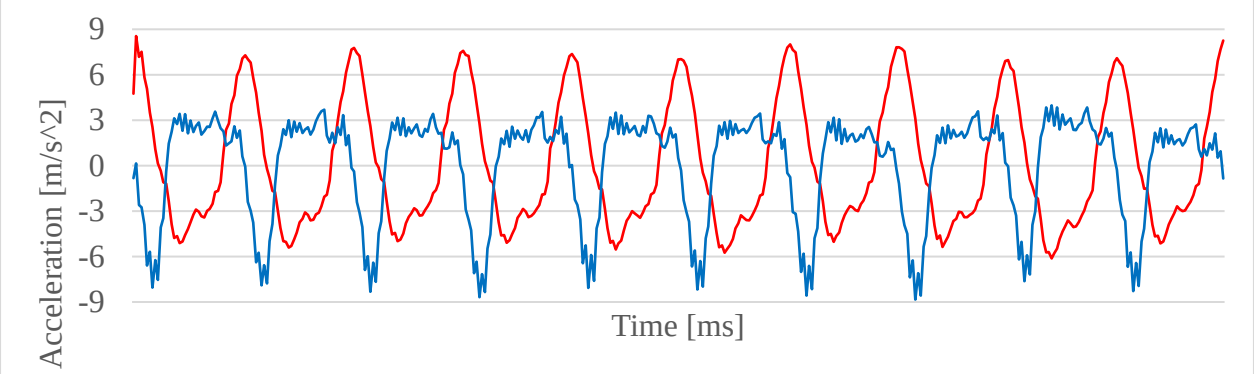
(b)


10
5
0
-5
-10
Acceleration [m/s^2]
Time [ms]

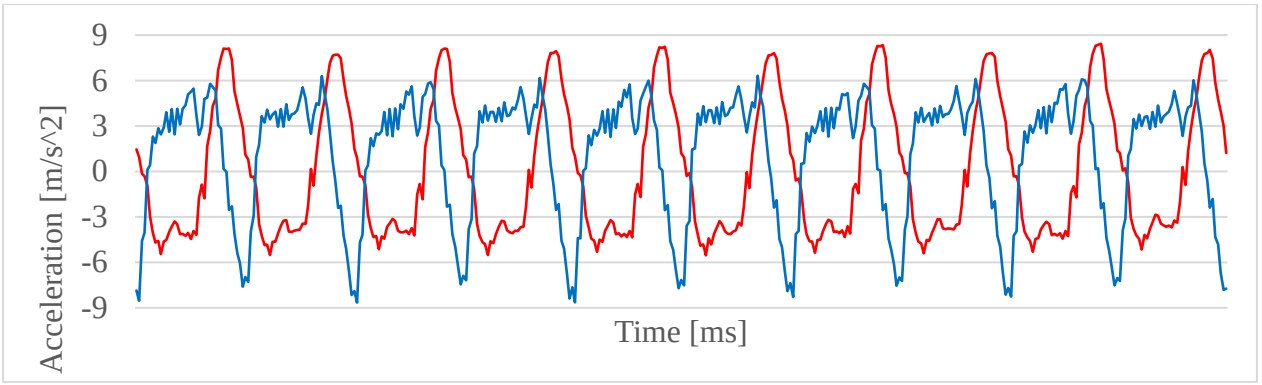
(c)


10
5
0
-5
-10
Acceleration [m/s^2]
Time [ms]

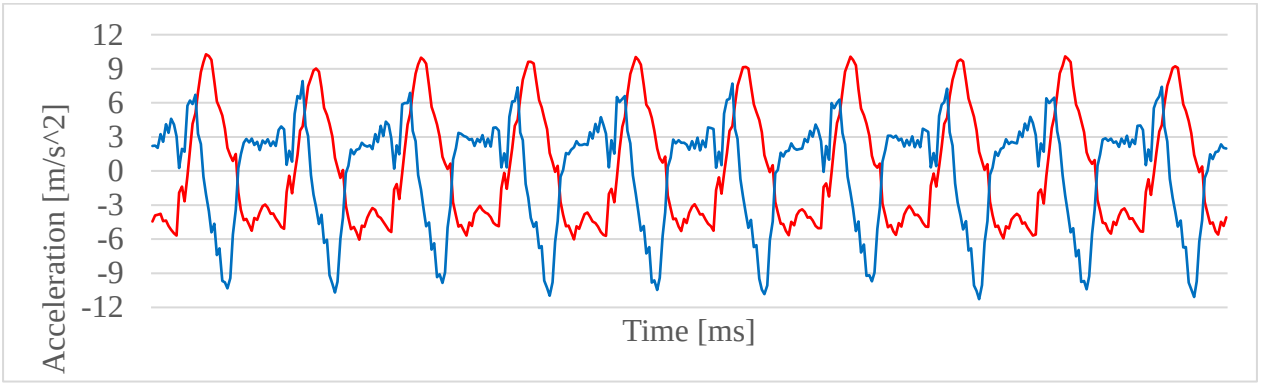
(d)


9
6
3
0
-3
-6
-9
Acceleration [m/s^2]
Time [ms]

(e)

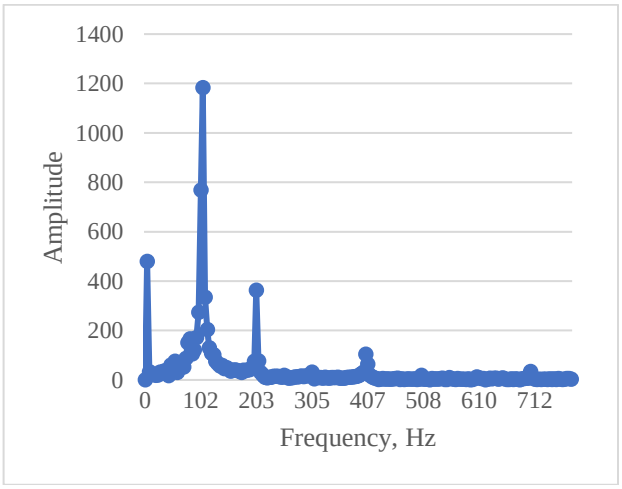


(f)

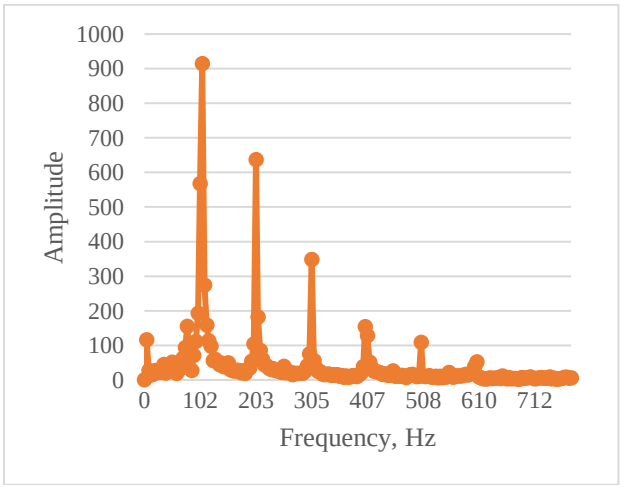


(g)

Figure 5.1. Transformer core and winding vibration signals under full load: (a) 80%, (b) 85%, (c) 90%, (d) 95%, (e) 100%, (f) 105%, (a) 110% of rated voltage injected to transformer.



(a)

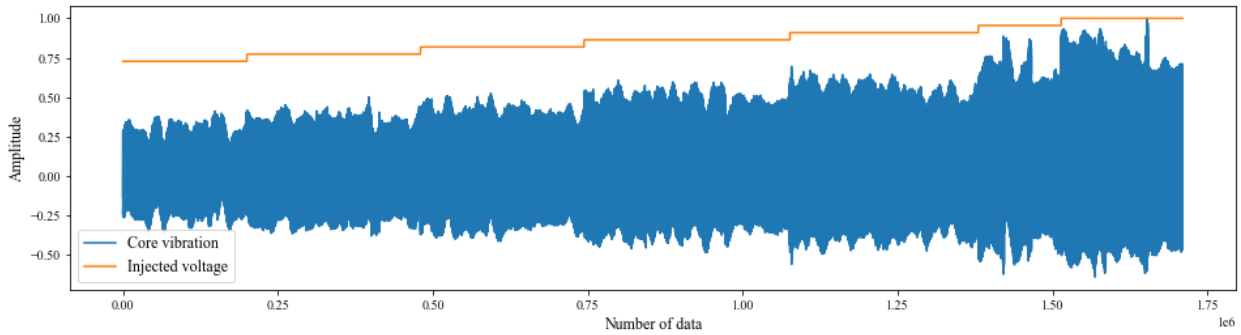


(b)

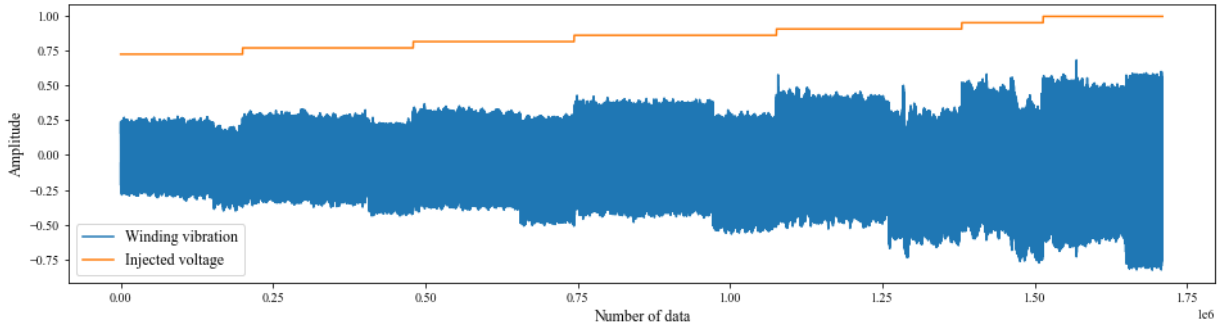
5.2. Transformer (a) core and (b) winding vibration signals in frequency domain.

To check the fundamental frequency of the core and winding vibration signals, Fourier analysis was applied to the collected transformer vibration signals. As it was discussed in Chapter 2, according to equation (2.6) the fundamental frequency of the core vibration should be proportional to system frequency or double of the system frequency. In the experimental part, the frequency of the power supplier was 50 Hz, which means the frequency of the system is 50 Hz. Therefore, the fundamental frequency of the vibration signals should be 100 Hz. According to Figure 5.2, both transformer core and winding frequency signals have dominant frequency of 100 Hz. However, there are some harmonic frequency orders because of the loosened core laminates and harmonics created by loads.

Firstly, data was not ready to predict voltage excitation, because some values of the features were missing or not available (NA). Also, the measured values were collected in different time units and it had different sampling rates. The reason of that is the equipment frequency for reading data from sensors. For instance, vibration signals were collected by 20 kHz and injected voltage of transformer was measured each 500 milliseconds. During data preparation process extra range of data values were dropped, such as the values of vibration and voltage under transition period of injected voltage, specifically, when injected voltage value is increased or decreased by power supplier. Also, all input and output features were normalized between -1 to 1. Normalized core and winding vibration signals with normalized injected voltage before segmentation process is shown in Figure 5.3. As it was discussed in Chapter 2, the amplitude and peak-to-peak values of the transformer vibration is proportional to square of input voltage, which means that in each step the vibration magnitude will be increased. Accordingly, Figure 5.3 illustrates the proportionality of vibration signal magnitudes on injected voltage value.



(a)



(b)

Figure 5.3. Normalized (a) core and (b) winding vibration signals with normalized injected voltage.

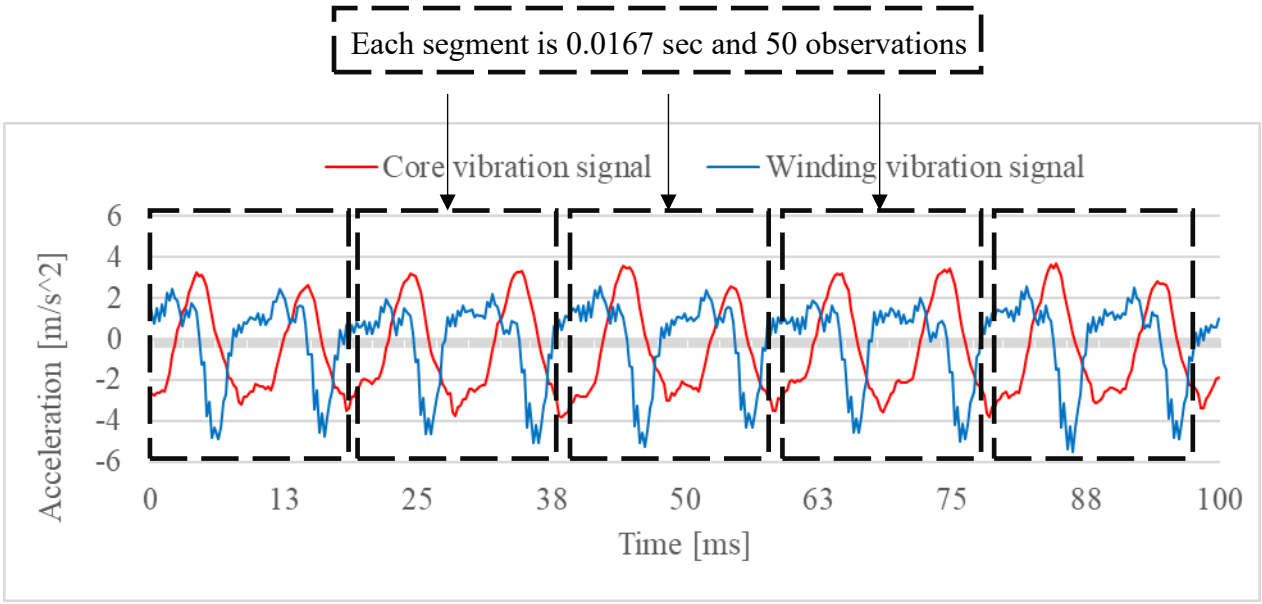


Figure 5.4. Segmentation process for a portion of collected vibration waveforms for 80% excitation.

After normalization of dataset, vibration signal data was chosen as the main input value of CNN model. Before starting to train the model, dataset was divided into segments. The reason of segmentation of the dataset is clear from the Figure 5.4. Actually, the sampling frequency of the data is 3 kHz, which means 3000 sample data in one second. 1 sample of the data does not give any information about vibration signal such as amplitude, peak-to-peak value, behavior or frequency of one cycle. Therefore, 50 samples or 0.0167 seconds of data were chosen as 1 segment of data, which contains approximately 2 cycle signals of the vibration as shown in Figure 5.4. Overall, 50 observations in one segment and 60 segments in one second data. All the segments were shuffled to increase the performance of the model. Lastly, testing data consists of 25 percent of whole dataset, 75 percent of remaining dataset was used as training data and 25 percent was left for validation part.

5.1.2. Model Construction, Selection and Validation Results

The listed assumptions in model selection section (Section 3.3) were constructed to search the best and high-performance models from possible space of established hyperparameters. 43200 models were trained from the combination of those choices of all possible parameters and only one epoch with the lowest validation loss was chosen from each 150 epochs. The remaining 288 CNN models with MSE, RRSE, and RAE values are listed in Table A.1 in Appendix. The best 30 models were given with all hyperparameters such as layer size, filter number in each layer, Adam optimizer, batch and kernel size in Table 5.1 and sorted by validation MSE value. Also, RRSE and RAE values were calculated for each constructed model and added in Table 5.1. According to Table 5.1, the best 10 models have at least 4 convolution layers, minimum kernel size 5 and Adam optimizer 0.001, which means 1D-CNN architecture constructed with more than 4 layers, and with

higher kernel size has better accuracy and performance for transformer voltage excitation prognosis.

Table 5.1. The selected 1D-CNN architecture for transformer excitation based on lowest MSE on the validation set (top 30 models).

Layer size, *N*	**Filter number**, *L*	**Batch size**	**Kernel size**	**Adam optimizer**	**MSE**	**RRSE**	**RAE**
5	[256, 128, 64, 32, 16]	32	6	0.001	7.7364E-06	4.4938	2.4863
4	[256, 128, 64, 32]	32	7	0.001	7.8655E-06	4.1612	2.7553
5	[256, 128, 64, 32, 16]	32	5	0.001	8.2211E-06	3.7985	1.4734
5	[256, 128, 64, 32, 16]	64	6	0.001	8.6697E-06	4.0723	2.0035
6	[256, 128, 64, 32, 16, 8]	32	5	0.001	9.2483E-06	3.7092	2.2312
5	[256, 128, 64, 32, 16]	64	7	0.001	9.3912E-06	3.9514	2.6350
4	[256, 128, 64, 32]	32	6	0.001	9.4796E-06	5.3838	3.1683
6	[256, 128, 64, 32, 16, 8]	64	7	0.001	9.4955E-06	4.6614	3.1618
6	[256, 128, 64, 32, 16, 8]	32	6	0.001	1.0342E-05	3.8244	2.0618
5	[256, 128, 64, 32, 16]	32	4	0.001	1.0545E-05	5.1230	3.1113
5	[256, 128, 64, 32, 16]	64	5	0.001	1.0814E-05	4.3217	2.4508
4	[256, 128, 64, 32]	64	5	0.001	1.0972E-05	4.4159	2.6754
4	[256, 128, 64, 32]	64	6	0.001	1.2035E-05	6.8311	5.3368
5	[8, 16, 32, 64, 128]	32	4	0.001	1.2840E-05	6.1684	4.5713
4	[256, 128, 64, 32]	32	4	0.001	1.3091E-05	4.7992	2.6838
5	[8, 16, 32, 64, 128]	32	7	0.001	1.3106E-05	5.1883	2.2773
6	[256, 128, 64, 32, 16, 8]	64	4	0.001	1.3316E-05	4.6016	3.0866
5	[8, 16, 32, 64, 128]	32	5	0.001	1.3667E-05	8.7269	8.9118

5	[256, 128, 64, 32, 16]	64	4	0.001	1.3844E-05	5.3139	3.8888
3	[256, 128, 64]	32	5	0.001	1.4450E-05	5.2235	4.1084
3	[256, 128, 64]	64	7	0.001	1.4935E-05	5.0575	3.8788
3	[256, 128, 64]	64	6	0.001	1.5077E-05	4.8116	3.6929
5	[256, 128, 64, 32, 16]	32	3	0.001	1.5170E-05	7.1981	5.3865
5	[8, 16, 32, 64, 128]	32	6	0.001	1.5598E-05	6.4959	3.5408
6	[8, 16, 32, 64, 128, 256]	32	7	0.001	1.5641E-05	5.3445	3.1393
4	[256, 128, 64, 32]	32	3	0.001	1.5798E-05	6.4545	4.8613
6	[8, 16, 32, 64, 128, 256]	64	7	0.001	1.5874E-05	7.0629	6.3522
3	[256, 128, 64]	32	6	0.001	1.5880E-05	4.9369	3.7870
4	[256, 128, 64, 32]	64	4	0.001	1.6624E-05	5.4340	4.2352
3	[256, 128, 64]	32	7	0.001	1.6744E-05	5.3343	4.0709

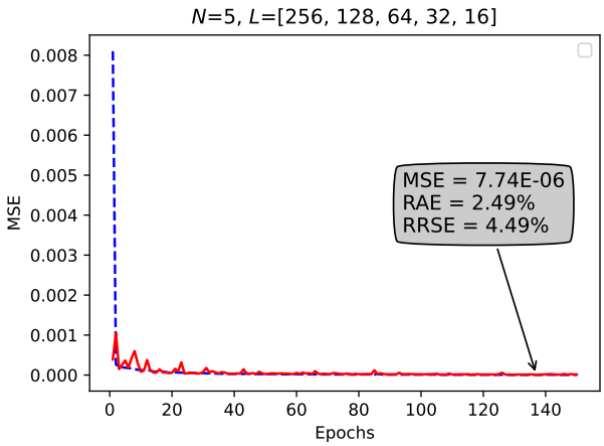
Table 5.2. The CNN model for transformer voltage excitation prognosis: batch size 32, convolution layers [256, 128, 64, 32, 16], kernel size 6.

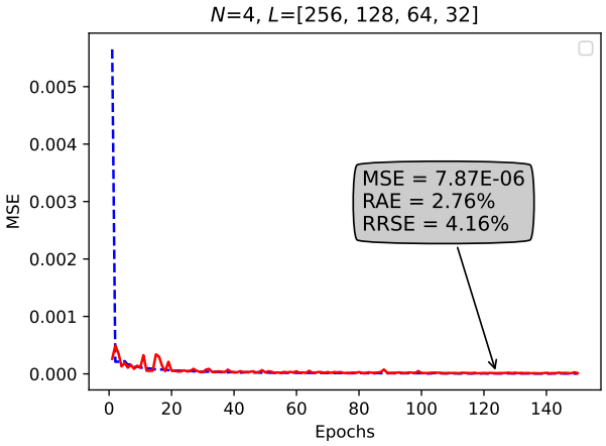
Layer (type)	Output Shape
conv1d	(45, 256)
max_pooling1d	(22, 256)
conv1d	(22, 128)
max_pooling1d	(11, 128)
conv1d	(11, 64)
max_pooling1d	(5, 64)
conv1d	(5, 32)
max_pooling1d	(2, 32)
conv1d	(2, 16)
flatten	(32)
dense	(50)
dense	(1)

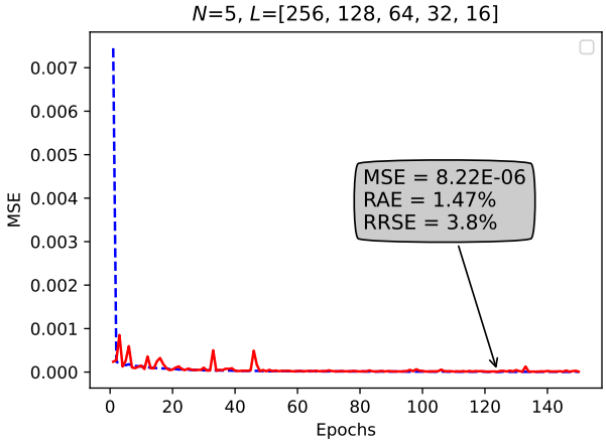
All CNN models were trained with all the assumed parameter combinations on training dataset by evaluating on both the training and validation sets. Training and validation processes were monitored by evaluating the MSE metric at each epoch, then the trained model with a lower validation loss is stored as the “best” model. MSE is chosen as the main evaluator of training and validation loss. After 150 epochs, the constructed model with the lowest validation loss is saved as the final model with specific hyperparameters.

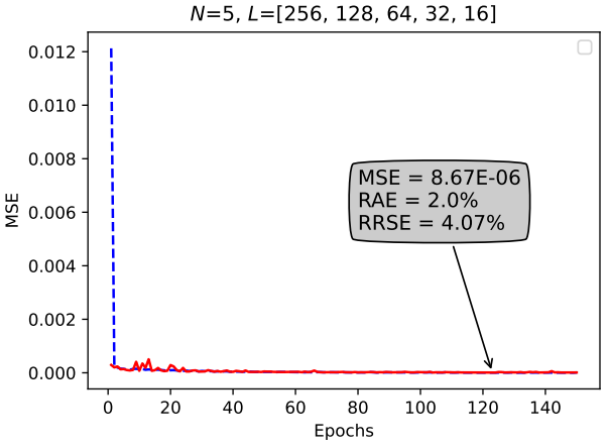
Figure 5.5 depicts the training (blue dash line) and validation loss (red line) for the best 12 constructed models. For each model, the RRSE and RAE values the best epoch (the epoch that led to lowest MSE) is shown in each plot. After analyzing the graphs, it can be clearly seen that all models have the best epoch between 100 and 150 epochs. Moreover, training and validation losses of the models are very close to each other, which indicates that how well constructed model fits the training and new data. The structure of the best constructed CNN architecture is summarized in Table 5.2.

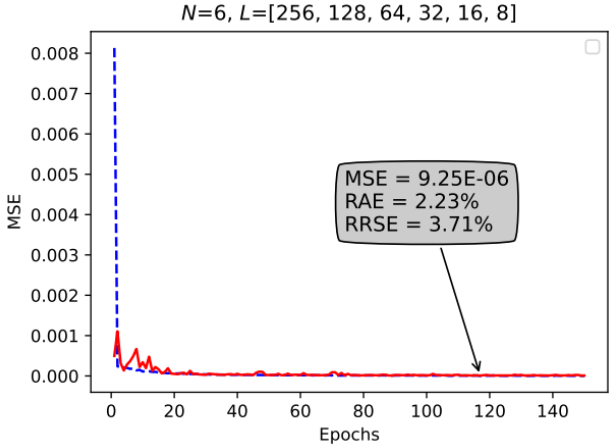
The comparison between performances of the constructed 1D-CNN models tabulated in Table 5.1 can be simplified using graphical visualization as shown in Figure 5.6. In this figure, the performance of the best 10 models was plotted using $-\log_{10}(\text{MSE})$ value. According to given graph, the best model in terms of the lowest MSE on the voltage excitation dataset is 1D-CNN model with kernel size is 6, batch size is 32, layer size is 5, filter number in each pair is [256, 128, 64,32,16] and epoch number is 139. In addition, RRSE and RAE values of this model are 4.49% and 2.49%, respectively.

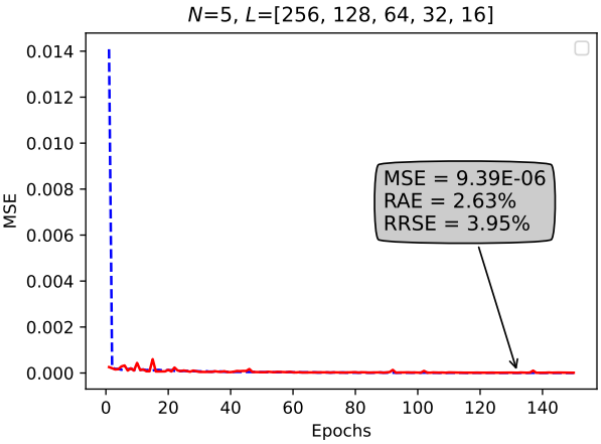

N=5, L=[256, 128, 64, 32, 16]
MSE = 7.74E-06
RAE = 2.49%
RRSE = 4.49%
MSE
Epochs


N=4, L=[256, 128, 64, 32]
MSE = 7.87E-06
RAE = 2.76%
RRSE = 4.16%
MSE
Epochs


N=5, L=[256, 128, 64, 32, 16]
MSE = 8.22E-06
RAE = 1.47%
RRSE = 3.8%
MSE
Epochs


N=5, L=[256, 128, 64, 32, 16]
MSE = 8.67E-06
RAE = 2.0%
RRSE = 4.07%
MSE
Epochs


N=6, L=[256, 128, 64, 32, 16, 8]
MSE = 9.25E-06
RAE = 2.23%
RRSE = 3.71%
MSE
Epochs


N=5, L=[256, 128, 64, 32, 16]
MSE = 9.39E-06
RAE = 2.63%
RRSE = 3.95%
MSE
Epochs

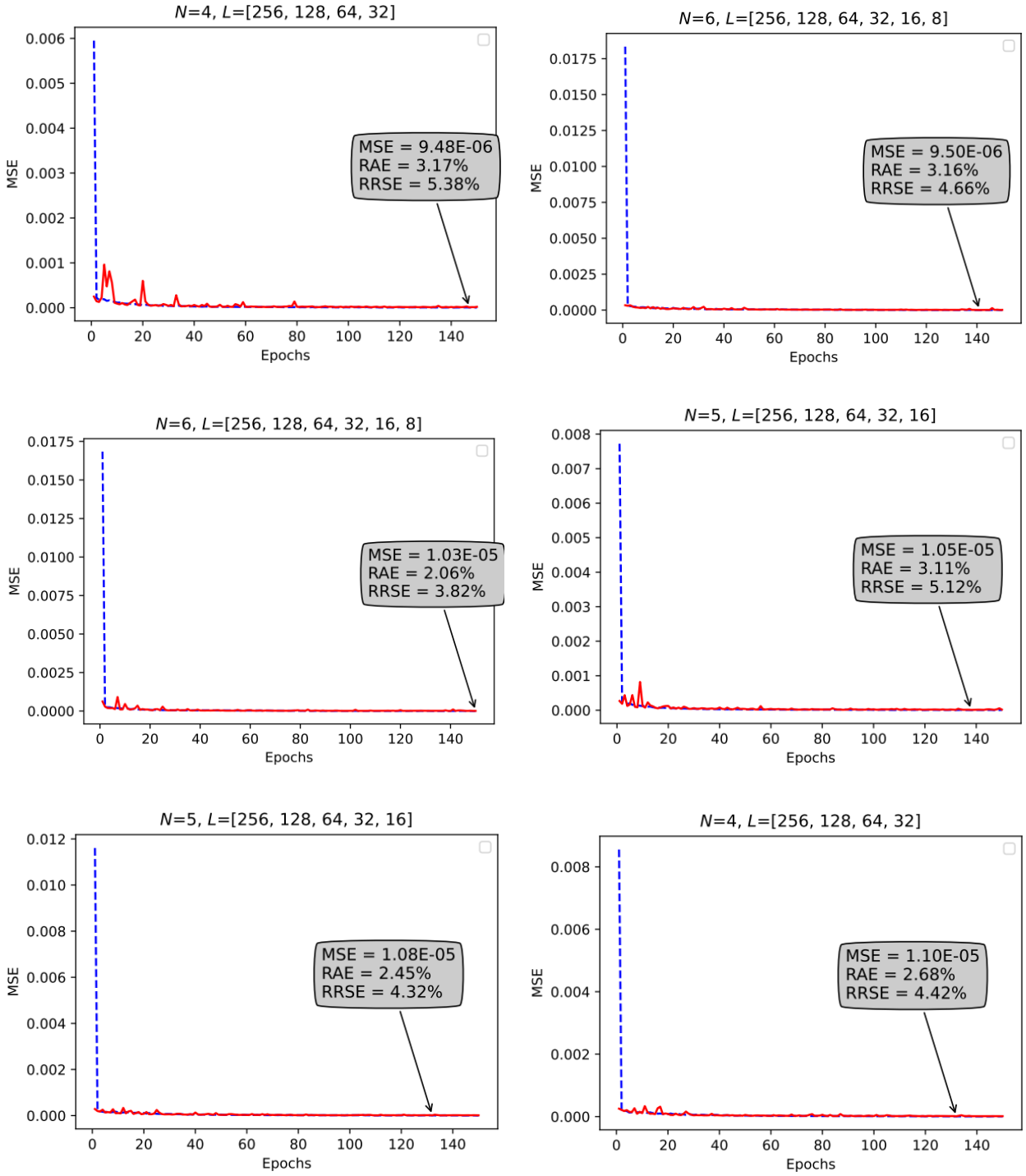


Figure 5.5. Performance of the best 12 1D-CNN models from Table 5.1 with arrows showing the best models.

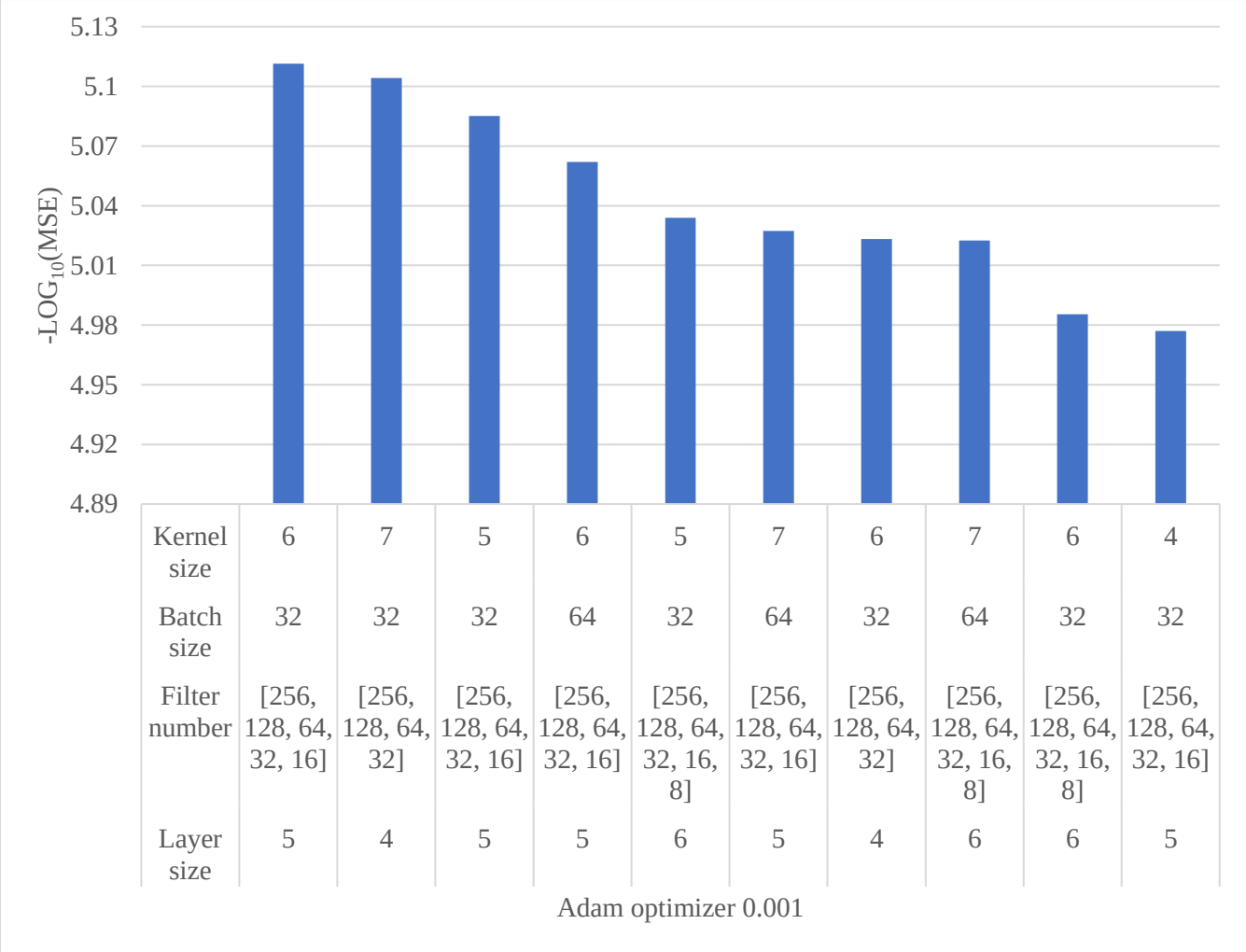


Figure 5.6. Performance of the top 10 constructed 1D-CNN models in terms of -log$_{10}$(MSE).

5.2. Transformer Inter-Turn Short-Circuit Fault

Transformer inter-turn short-circuit is very costly and can easily destroy the transformer. It can occur in all distribution transformers as well as in power transformers. The main cause of it is insulation failure between winding turns of the transformer. This fault was emulated in laboratory level to collect data from experimental transformer and the setup for data collection process was explained in detail in Section 4.2. But in this experiment only winding vibration is taken as the main attribute for collecting dataset by measuring the value of applied load of transformer.

Therefore, dataset has one input features and one output column. The list of the features is shown below:

- Transformer winding vibration signals;
- Transformer applied load.

5.2.1. Data Collection and Preparation

While emulating the turn-to-turn short circuit fault, the third turn of the experimental transformer was connected to variable resistor and its value was decreased to increase the load step by step. At each step winding vibration signals were measured and collected in parallel to load value. The loads applied to transformer is tabulated in Table 5.3. The short circuit fault starts from 11 A load. Thus, it was considered that from 11 A to 15 A is faulty condition and if load is less than 10 A it worked under non-fault state.

The data preparation process is similar to discussed process in Section 5.1.1 and normalized also in the same way between -1 and 1. The collected data, demonstrated in Figure 5.7, was segmented to 50 samples per segment or 0.0167 seconds in time duration same as in Figure 5.4. Thus, there are 60 segments per 1 second, because sampling rate is 3 kHz. The dataset was shuffled and divided into train, validation and testing set. Testing data was 25% of whole dataset. 75% of remaining dataset was used as training data and another 25% for validation set.

Table 5.3. Different loads and its values passing through experimental transformer including turn-to-turn short circuit current.

Load [p.u]	0.71	0.95	1	1.19	1.31	1.42	1.66	1.9
Load current [A]	3	4	4.2	5	5.5	6	7	8
Load [p.u]	2	2.14	2.26	2.38	2.61	2.85	3.09	3.57
Load current [A]	8.4	9	9.5	10	11	12	13	15

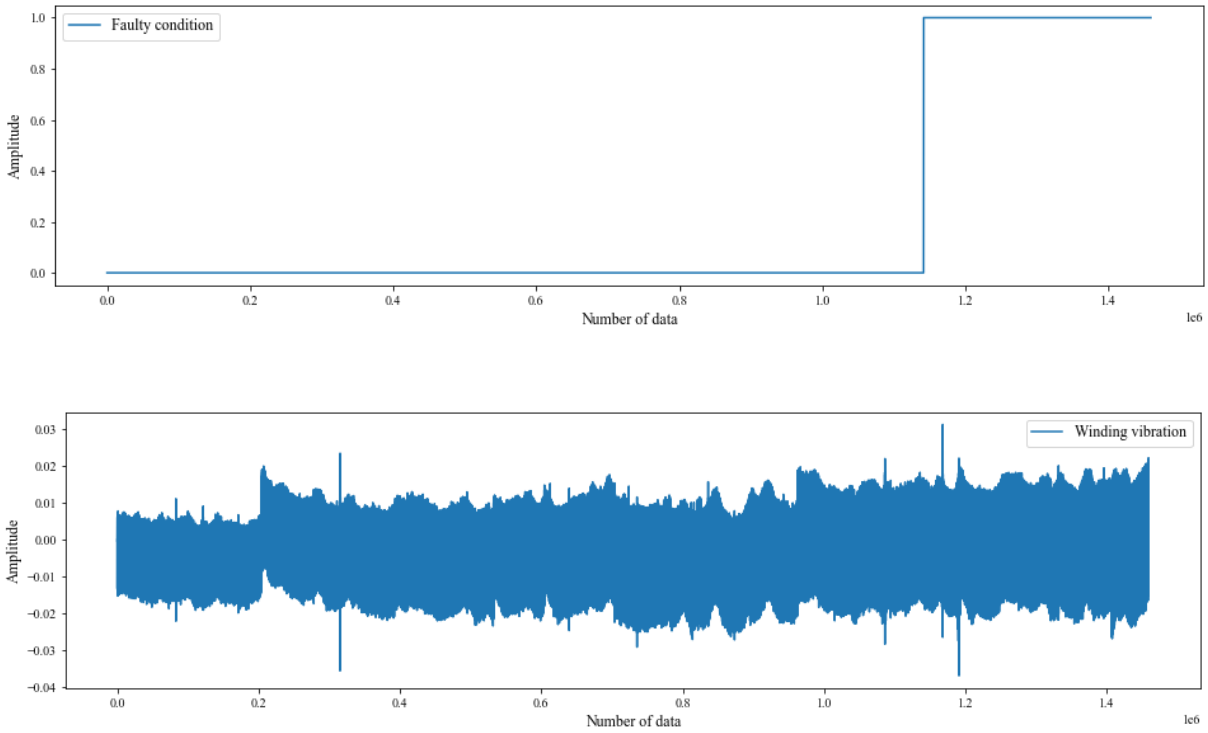


Figure 5.7. Normalized winding vibration and its value under non-faulty and faulty condition.

5.2.2. Model Construction, Selection and Validation Results

The CNN models were constructed with identical assumptions about hyperparameter space indicated in Section 3.3, such as optimizer type and learning rate, convolutional layer size and filter numbers in each layer, batch size, size and ReLu function. In previous case study, CNN regression model was constructed for voltage excitation prognosis. However, in this part, CNN model was constructed for classification. In other words, the model was developed to identify the faulty condition (fault or non-fault) of the transformer inter-turn windings. Model selection process relied on accuracy of the trained all possible 43200 models. From 150 epoch, only the best epoch was chosen and 288 models were trained and listed in Table A.2 in Appendix. The best 30 models from Table A.2 were illustrated in Table 5.4. Models accuracy was calculated by dividing all correct outputs to all outputs.

Table 5.4. The selected 1D-CNN architecture for transformer inter-turn fault based on highest accuracy on the validation set (top 30 models).

Layer size	Filter number, *L*	Batch size	Kernel size	Adam optimizer	Accuracy, %
1	[256]	64	3	0.001	99.8629
4	[256, 128, 64, 32]	32	3	0.0001	99.8458
3	[256, 128, 64]	64	5	0.001	99.8458
4	[8, 16, 32, 64]	32	6	0.001	99.8458
3	[256, 128, 64]	32	2	0.0001	99.8287
2	[256, 128]	32	2	0.0001	99.8287
1	[256]	64	7	0.001	99.8287
3	[8, 16, 32]	64	6	0.001	99.8287
2	[256, 128]	64	3	0.0001	99.8115
2	[256, 128]	32	4	0.0001	99.8115
1	[256]	64	5	0.001	99.7944
2	[256, 128]	32	5	0.0001	99.7944
1	[256]	64	4	0.001	99.7944
6	[256, 128, 64, 32, 16, 8]	64	7	0.001	99.7944
2	[256, 128]	64	5	0.001	99.7944
1	[256]	32	6	0.001	99.7944
1	[8]	32	4	0.001	99.7944
3	[256, 128, 64]	64	7	0.001	99.7944
2	[8, 16]	32	7	0.001	99.7944
5	[256, 128, 64, 32, 16]	32	4	0.0001	99.7944
2	[8, 16]	64	2	0.001	99.7944
4	[256, 128, 64, 32]	64	6	0.0001	99.7944
2	[8, 16]	64	3	0.001	99.7944
2	[256, 128]	64	4	0.0001	99.7773
4	[8, 16, 32, 64]	64	3	0.001	99.7773
2	[256, 128]	64	6	0.0001	99.7773
3	[256, 128, 64]	64	3	0.001	99.7773
2	[256, 128]	32	7	0.0001	99.7773
3	[256, 128, 64]	64	7	0.0001	99.7773
4	[8, 16, 32, 64]	64	6	0.001	99.7773

According to best models in Table 5.4, it is better to construct the 1D-CNN model with less than four convolutional layers and kernel size can be less than 5. Both learning rate of Adam optimizer shows same result, which means effect of learning rate is not important. Batch size of 32 and 64 are preferable to construct CNN model for inter-turn short-circuit fault prognosis. All the

models listed in Table 5.4 can be used to identify the faulty condition of transformer, because all of them have very high accuracy more than 99.77%. However, the model with one layer and 256 filters in layer, batch size of 64, kernel size of 3 and Adam optimizer with learning rate 0.001 was chosen to predict the short circuit fault in transformer inter-turn and its accuracy is 99.86%.

5.3. Discussion

In the previous research [1], authors considered only a mathematical sum of sinusoids with undetermined amplitudes, model order, frequencies, and phases in order to model the transformer vibration signals. Those parameters are then further calculated using the time-series data. This concept of applying the sum of sinusoids with unidentified parameters to model the small duration time-series segments is known to be the segmented Prony's approach for spectral line estimation. When the transformer is under operation, we can anticipate seeing a sinusoidal time series with a frequency of 2ω (where ω is the fundamental frequency) for vibration in the winding and core, as previously stated in [1] and [3]. Nevertheless, since the transformer vibration spectrum may face a higher harmonic order (>2ω), this sinusoidal model may not fully meet expectations to its analytical promises. Thus, to avoid such limitations, the segmented Prony's approach used - the sum of sinusoids on segmented observation windows. Following the features of these sinusoidal models were calculated on each segmented window, they had the option of making their prediction based on extracted features using a variety of machine learning techniques (multilayer perceptron, support vector regression, linear regression, and model trees). In study [3], authors have considered only data-driven techniques such as deep RNNs to thoroughly capture the hidden patterns of the vibration data with no intervention of any separate modelling of the winding vibration. In this regard, GRU, Bi-GRU, LSTM, and Bi-LSTM was taken into account. The results of [1] are

somewhat similar and comparable to the performances of these RNN techniques. However, it was not aimed at concluding about the GRU performance compared to LSTM in the scope of fault prognosis. On top of that, no machine learning community drew any conclusion about the outperformance of one of these techniques over the other. This can be clearly seen from the summary of the authors of [38]. Leaving aside, both cases of estimating the load current and excitation voltage based on the vibration samples depict that in the same batch size and learning rate, the bidirectional models are superior to the unidirectional ones. Furthermore, both cases were concerned in [3], a Bi-LSTM was the model of choice in the case of load current, whereas a Bi-GRU was selected model to predict the excitation voltage. Moreover, it is observed that the interturn fault prognosis is a way more challenging issue than over-/under-excitation in the framework of accomplishing a lower forecast error rate. It might have been caused by a random behaviour of interturn short-circuit fault in the transformers compared to over-/under-excitation fault.

In [3], the transformer excitation voltage and the load current can be estimated (predicted) with the trained RNN architectures by samples of the transformer vibration waveforms of in the previous 0.02 secs, which is a one-cycle 50 Hz fundamental frequency. It verifies that the electrical classical parameters indicating the transformer fault (i.e., load current and excitation voltage) are mapped to the measured vibration waveforms with 0.02 secs interval. It literally means that the vibration data enable to detect the fault in its early stages. In addition to this, transformer prognosis is some knowledge on how the transformer behaviour tends to turn out with accessing to the usual course of the fault. Going beyond a binary classification problem, where categorical variables are merely estimated (fault or non-fault conditions based on the vibration waveform), this study has dedicated to formulating a regression problem with continuous variables in terms of the excitation voltage and load current. Nevertheless, when it comes to practicality, there is a flexibility of

interpreting the faulty conditions based on the over-/under-excitation and/or short-circuit current rather than imposing the orthodox ‘faulty’ definition. In such circumstance, the detection of the faulty condition given the definition is directly correlated with how accurate the fault parameters were predicted such as excitation voltage and load current. Moreover, the duration of decision-making can be considered as another sophisticated point of our predictive models. It is impressive that authors employed the duration of 0.02 secs for vibration waveform for the RNNs to be trained for decision-making since previous study [1] was fed by the duration of as 5 times long as one-cycle vibration waveform. On the other hand, predictive RNN models are not perfectly accurate.

The CNN models in this thesis can predict the transformer inter-turn short circuit fault and voltage excitation by recorded vibration signals for 0.0167 seconds, which is less than fundamental frequency (50 Hz) of the electrical system. In other words, trained CNN models take 0.0167 seconds of recorded vibration signals to indicate the electric parameter of the transformer faulty conditions. Especially, the constructed models under this thesis are able to detect the faults named above in early stages or faulty transformer in initial phase, when it becomes suspicious, considering vibration waveforms. One of the main privileges of CNN is fastness in terms of the time complexity. In addition, CNN model constructed in this thesis for turn-to-turn short circuit fault detection in early stages has very good performance compared to results of previous two studies [1] and [3].

Chapter 6 – Conclusion and Future Work

In this thesis, various approaches to predict transformer voltage excitation and inter-turn short circuit fault using time-series vibration signals were discussed in detail. After reviewing previous studies, classical techniques that can be applied to detect faults using vibration data is following:

- Modelling the vibration signals using the mathematical approach (sum of sinusoids)
- Estimation of mathematical parameters in segmented short period of signals
- Constructing predictive mathematical models using feature extracting techniques.

The approaches that combine feature extraction methods with classical machine learning techniques are applied to time-series vibration data in the study [1]. They have compared the performance of different machine learning methods such as model trees, multilayer perceptron, support vector regression and linear regression. The results of constructed models to predict transformer underexcitation and overexcitaion voltage and turn-to-turn short circuit fault were measured by RAE value, 2.42% and 11.37%, respectively.

Another recent study [6] has demonstrated the potentials of recent developments in machine learning, particularly, in deep learning architectures. In [6], stacked unidirectional and bidirectional RNN models were constructed to extract important features from transformer vibration waveforms and train predictive models for under-excitation, over-excitation and inter-turn short circuit faults. Applying a brute-force model selection method with the four architectures (GRU, Bi-GRU, LSTM, and Bi-LSTM) discussed in Section 5.3, and 48 000 models were trained. The performance of the best predictive model was reported by RAE values of 0.56% and 17.58% for the voltage excitation and inter-turn short circuit faults, respectively.

In this research work, another high-performance approach of deep learning techniques was also conducted through different experimental works. One-dimensional CNNs (1D-CNN) were employed as they have ability to extract patterns of interest from input dataset, in this research it was transformer vibration waveforms, and construct predictive models. In other words, feature extraction and model construction are the combined process and not required further separate feature extraction process. After setting assumptions on hyperparameters for model selection, 1D-CNN architecture was considered to train 2 (batch size) × 150 (epoch number) × 2 (learning rate) × 12 (convolutional layers or filters) × 6 (kernel size) = 43 200 models for predicting transformer excitation voltage and inter-turn short circuit fault in early stages. The predictive performances of two faults were reported with different methods, as the predictive model for the case one was regression and for the second case it was classification. The best predictive model of the case 1 (transformer voltage excitation) revealed RRSE value of 4.49% and RAE value of 2.49%. The accuracy of the predictive model of case 2 was obtained also for 99.86%. The performance of the CNN models is comparable and shows that deep learning approaches have ability to capture all patterns of vibration waveforms in predicting the specific fault. Lastly, the future work over this research work are to improve the performance of the transformer voltage excitation prediction and employ further machine learning techniques. Developing a pipeline and system to be able to analyze the transformer vibration signal for fault prognosis is also suggested as the future work in this study.

Bibliography

[1] M. Bagheri, A. Zollanvari, and S. Nezhivenko, “Transformer fault condition prognosis using vibration signals over cloud environment,” IEEE Access, vol. 6, pp. 9862–9874, 2018.

[2] M. Bagheri, S. Nezhivenko, M. Naderi, and A. Zollanvari, “A new vibration analysis approach for transformer fault prognosis over cloud environment,” International Journal of Electrical Power and Energy Systems, vol. 100, pp. 104–116, Sept. 2018.

[3] S. Saponara, “Distributed measuring system for predictive diagnosis of uninterruptible power supplies in safety-critical applications,” Energies, vol. 9, no. 5, 2016.

[4] P. M. Joshi and S. V. Kulkarni, “Transformer winding diagnostics using deformation coefficient,” in 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, pp. 1–4, 2008.

[5] M. Bagheri and B. T. Phung, “Frequency response and vibration analysis in transformer winding turn-to-turn fault recognition,” in 2016 International Conference on Smart Green Technology in Electrical and Information Systems (ICSGTEIS), pp. 10–15, 2016.

[6] A. Zollanvari, K. Kunanbayev, S. Akhavan Bitaghsir, and M. Bagheri, “Transformer fault prognosis using deep recurrent neural network over vibration signals,” IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1–11, 2021.

[7] M. H. Samimi and S. Tenbohlen, “Fra interpretation using numerical indices: State-of-the-art,” International Journal of Electrical Power Energy Systems, vol. 89, pp. 115–125, 2017.

[8] M. F. M. Yousof, C. Ekanayake, and T. K. Saha, “Frequency response analysis to investigate deformation of transformer winding,” IEEE Transactions on Dielectrics and Electrical Insulation, vol. 22, no. 4, pp. 2359–2367, 2015.

[9] M. Bagheri, S. Nezhivenko, and B. T. Phung, “Loss of low-frequency data in on- line frequency response analysis of transformers,” IEEE Electrical Insulation Magazine, vol. 33, no. 5, pp. 32–39, 2017.

[10] Z. Zhao, C. Yao, C. Li, and S. Islam, “Detection of power transformer winding deformation using improved fra based on binary morphology and extreme point variation,” IEEE Transactions on Industrial Electronics, vol. 65, no. 4, pp. 3509–3519, 2018.

[11] M. Bagheri, M. S. Naderi, T. Blackburn, and T. Phung, “Frequency response analysis and short-circuit impedance measurement in detection of winding deformation within power transformers,” IEEE Electrical Insulation Magazine, vol. 29, no. 3, pp. 33–40, 2013.

[12] S. Naiqiu, Z. Can, H. Fang, L. Qisheng, and Z. Lingwei, “Study on ultrasonic measurement device for transformer winding deformation,” in Proceedings. International Conference on Power System Technology, vol. 3, pp. 1401–1404 vol.3, 2002.

[13] E. Arri, A. Carta, F. Mocci, and M. Tosi, “Diagnosis of the state of power transformer windings by on-line measurement of stray reactance,” IEEE Transactions on Instrumentation and Measurement, vol. 42, no. 2, pp. 372–378, 1993.

[14] M. A. Hejazi, G. B. Gharehpetian, G. Moradi, H. A. Alehosseini, and M. Mohammadi, “Online monitoring of transformer winding axial displacement and its extent using scattering parameters and k-nearest neighbour method,” IEE proceedings.Generation, transmission, and distribution., vol. 5, no. 8, pp. 824–832, 2011. Copyright - Copyright The Institution of Engineering Technology 2011.

[15] T.-T. He, J.-D. Wang, J. Guo, H. Huang, X.-X. Chen, and J. Pan, “A vibration based condition monitoring system for power transformers,” in 2009 Asia-Pacific Power and Energy Engineering Conference, pp. 1–4, 2009.

[16] A. Abu-Siada and S. Islam, “A novel online technique to detect power transformer winding faults,” IEEE Transactions on Power Delivery, vol. 27, no. 2, pp. 849–857, 2012.

[17] Y. X. Liao, T. Y. Zhu, Y. Q. Sun, J. Zhang, T. Cheng, and Y. Wang, "Load influence on lissajous figure for online transformer winding diagnosis," in 2016 IEEE International Conference on High Voltage Engineering and Application (ICHVE), pp. 1–4, 2016.

[18] Y. Eroglu and S. U. Sec¸kiner, "Early fault prediction of a wind turbine ˇ using a novel ann training algorithm based on ant colony optimization," Journal of Energy Systems, vol. 3, no. 4, pp. 139–147, 2019.

[19] Z. Tian, "An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring," Journal of Intelligent Manufacturing, vol. 23, no. 2, pp. 227–237, 2012.

[20] J. Benesty, J. Chen, and Y. Huang, "Automatic speech recognition: A deep learning approach," 2008.

[21] C.-L. Liu, W.-H. Hsaio, and Y.-C. Tu, "Time series classification with multivariate convolutional neural network," IEEE Transactions on Industrial Electronics, vol. 66, no. 6, pp. 4788–4797, 2019.

[22] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Realtime motor fault detection by 1-d convolutional neural networks," IEEE Transactions on Industrial Electronics, vol. 63, no. 11, pp. 7067–7075, 2016.

[23] B. Garcia, J. Burgos, and A. Alonso, "Transformer tank vibration modeling as a method of detecting winding deformations-part i: theoretical foundation," IEEE Transactions on Power Delivery, vol. 21, no. 1, pp. 157–163, 2006.

[24] J. Shengchang, L. Yongfen, and L. Yanming, "Research on extraction technique of transformer core fundamental frequency vibration based on olcm," IEEE Transactions on Power Delivery, vol. 21, no. 4, pp. 1981– 1988, 2006.

[25] S. Saponara, L. Fanucci, F. Bernardo, and A. Falciani, "Predictive diagnosis of high-power transformer faults by networking vibration measuring nodes with integrated signal processing," IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 8, pp. 1749–1760, 2016.

[26] R. Rajamani, M. Rajappa, K. Arunachalam, and B. Madanmohan, "Interturn short diagnosis in small transformers through impulse injection: on-line on-load self-impedance transfer function approach," Iet Science Measurement & Technology, vol. 11, pp. 961–966, 2017.

[27] B. Rao, P. S. Pai, and T. Nagabhushana, "Failure diagnosis and prognosis of rolling-element bearings using artificial neural networks: A critical overview," in Journal of Physics: Conference Series, vol. 364, p. 012023, IOP Publishing, 2012.

[28] Y. Wang and I. H. Witten, "Induction of model trees for predicting continuous classes," 1996.

[29] E. Frank, Y. Wang, S. Inglis, G. Holmes, and I. H. Witten, "Using model trees for classification," Machine learning, vol. 32, no. 1, pp. 63–76, 1998.

[30] Z. Boger and H. Guterman, "Knowledge extraction from artificial neural network models," in 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, vol. 4, pp. 3030–3035, IEEE, 1997.

[31] R. Di Stefano, S. Meo, and M. Scarano, "Induction motor faults diagnostic via artificial neural network (ann)," in Proceedings of 1994 IEEE International Symposium on Industrial Electronics (ISIE'94), pp. 220– 225, IEEE, 1994.

[32] B. Li, M.-Y. Chow, Y. Tipsuwan, and J. C. Hung, "Neural-network-based motor rolling bearing fault diagnosis," IEEE transactions on industrial electronics, vol. 47, no. 5, pp. 1060–1069, 2000.

[33] C. T. Kowalski and T. Orlowska-Kowalska, "Neural networks application for induction motor faults diagnosis," Mathematics and computers in simulation, vol. 63, no. 3-5, pp. 435–448, 2003.

[34] B.-S. Yang, M.-S. Oh, A. C. C. Tan, et al., "Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference," Expert Systems with Applications, vol. 36, no. 2, pp. 1840–1849, 2009.

[35] A. Krizhevsky, I. Sutskever, and G. Hinton, “Imagenet classification with deep convolutional neural networks,” in Proc. Adv. Neural Inf. Process. Syst. Conf., Lake Tahoe, Nevada, Dec. 2012, pp. 1097–1105.

[36] A. Serikbay, M. Bagheri, A. Zollanvari and B. T. Phung, "Accurate Surface Condition Classification of High Voltage Insulators based on Deep Convolutional Neural Networks," in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 28, no. 6, pp. 2126-2133, December 2021, doi: 10.1109/TDEI.2021.009648.

[37] Y. Li, L. Zou, L. Jiang and X. Zhou, "Fault Diagnosis of Rotating Machinery Based on Combination of Deep Belief Network and Onedimensional Convolutional Neural Network," in IEEE Access, vol. 7, pp. 165710-165723, 2019.

[38] J. Chung, Ç. Gülçehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” in Proc. NIPS, 2014, pp. 1–9.

Appendix

Table A.1. The selected 1D-CNN architecture for transformer excitation based on lowest MSE on the validation set.

Layer size	Filter number	Batch size	Kernel size	Adam optimizer	MSE	RRSE	RAE
5	[256, 128, 64, 32, 16]	32	6	0.001	7.7364E-06	4.4938	2.4863
4	[256, 128, 64, 32]	32	7	0.001	7.8655E-06	4.1612	2.7553
5	[256, 128, 64, 32, 16]	32	5	0.001	8.2211E-06	3.7985	1.4734
5	[256, 128, 64, 32, 16]	64	6	0.001	8.6697E-06	4.0723	2.0035
6	[256, 128, 64, 32, 16, 8]	32	5	0.001	9.2483E-06	3.7092	2.2312
5	[256, 128, 64, 32, 16]	64	7	0.001	9.3912E-06	3.9514	2.6350
4	[256, 128, 64, 32]	32	6	0.001	9.4796E-06	5.3838	3.1683
6	[256, 128, 64, 32, 16, 8]	64	7	0.001	9.4955E-06	4.6614	3.1618
6	[256, 128, 64, 32, 16, 8]	32	6	0.001	1.0342E-05	3.8244	2.0618
5	[256, 128, 64, 32, 16]	32	4	0.001	1.0545E-05	5.1230	3.1113
5	[256, 128, 64, 32, 16]	64	5	0.001	1.0814E-05	4.3217	2.4508
4	[256, 128, 64, 32]	64	5	0.001	1.0972E-05	4.4159	2.6754
4	[256, 128, 64, 32]	64	6	0.001	1.2035E-05	6.8311	5.3368
5	[8, 16, 32, 64, 128]	32	4	0.001	1.2840E-05	6.1684	4.5713
4	[256, 128, 64, 32]	32	4	0.001	1.3091E-05	4.7992	2.6838
5	[8, 16, 32, 64, 128]	32	7	0.001	1.3106E-05	5.1883	2.2773
6	[256, 128, 64, 32, 16, 8]	64	4	0.001	1.3316E-05	4.6016	3.0866
5	[8, 16, 32, 64, 128]	32	5	0.001	1.3667E-05	8.7269	8.9118
5	[256, 128, 64, 32, 16]	64	4	0.001	1.3844E-05	5.3139	3.8888
3	[256, 128, 64]	32	5	0.001	1.4450E-05	5.2235	4.1084
3	[256, 128, 64]	64	7	0.001	1.4935E-05	5.0575	3.8788
3	[256, 128, 64]	64	6	0.001	1.5077E-05	4.8116	3.6929
5	[256, 128, 64, 32, 16]	32	3	0.001	1.5170E-05	7.1981	5.3865
5	[8, 16, 32, 64, 128]	32	6	0.001	1.5598E-05	6.4959	3.5408
6	[8, 16, 32, 64, 128, 256]	32	7	0.001	1.5641E-05	5.3445	3.1393
4	[256, 128, 64, 32]	32	3	0.001	1.5798E-05	6.4545	4.8613
6	[8, 16, 32, 64, 128, 256]	64	7	0.001	1.5874E-05	7.0629	6.3522
3	[256, 128, 64]	32	6	0.001	1.5880E-05	4.9369	3.7870
4	[256, 128, 64, 32]	64	4	0.001	1.6624E-05	5.4340	4.2352
3	[256, 128, 64]	32	7	0.001	1.6744E-05	5.3343	4.0709
3	[256, 128, 64]	64	5	0.001	1.6899E-05	5.0201	4.0050
6	[256, 128, 64, 32, 16, 8]	64	5	0.001	1.6900E-05	5.0539	4.0316
4	[8, 16, 32, 64]	32	6	0.001	1.6983E-05	6.4453	4.0874
5	[256, 128, 64, 32, 16]	64	3	0.001	1.7280E-05	8.6135	8.2298
3	[256, 128, 64]	32	4	0.001	1.7645E-05	5.1679	3.7378
6	[8, 16, 32, 64, 128, 256]	32	6	0.001	1.8042E-05	6.1505	4.9852

5	[8, 16, 32, 64, 128]	64	7	0.001	1.8070E-05	5.5914	3.4748
3	[256, 128, 64]	64	4	0.001	1.8181E-05	6.4777	5.7205
6	[8, 16, 32, 64, 128, 256]	32	4	0.001	1.8611E-05	6.0565	2.3971
6	[8, 16, 32, 64, 128, 256]	64	6	0.001	1.8706E-05	6.3264	3.9414
2	[256, 128]	64	6	0.001	1.8885E-05	5.4415	4.4918
6	[8, 16, 32, 64, 128, 256]	32	5	0.001	1.9138E-05	6.5347	3.9383
2	[256, 128]	64	7	0.001	2.0061E-05	6.6063	6.0393
4	[8, 16, 32, 64]	32	7	0.001	2.0081E-05	6.6696	5.5829
4	[256, 128, 64, 32]	64	3	0.001	2.0178E-05	6.0558	4.4276
2	[256, 128]	32	7	0.001	2.0444E-05	5.9102	4.9733
5	[8, 16, 32, 64, 128]	64	6	0.001	2.0485E-05	5.6540	3.3889
3	[256, 128, 64]	64	3	0.001	2.0564E-05	5.5785	4.5500
6	[8, 16, 32, 64, 128, 256]	64	4	0.001	2.0836E-05	7.5616	4.3605
2	[256, 128]	32	7	0.0001	2.0849E-05	5.4300	4.6299
2	[256, 128]	64	4	0.001	2.1668E-05	5.7882	4.7279
2	[256, 128]	32	6	0.001	2.1736E-05	5.9335	4.6645
2	[256, 128]	32	6	0.0001	2.1921E-05	5.7647	4.9240
3	[256, 128, 64]	32	3	0.001	2.2143E-05	7.2858	6.4608
5	[8, 16, 32, 64, 128]	64	5	0.001	2.2255E-05	6.3297	4.9230
2	[256, 128]	64	5	0.001	2.2301E-05	7.7729	7.5242
6	[256, 128, 64, 32, 16, 8]	64	3	0.001	2.2310E-05	5.8031	4.4217
6	[8, 16, 32, 64, 128, 256]	32	3	0.001	2.2508E-05	5.8951	3.0691
4	[256, 128, 64, 32]	32	2	0.001	2.2779E-05	5.9170	3.9767
2	[256, 128]	32	5	0.0001	2.3789E-05	6.1164	5.4296
2	[256, 128]	32	5	0.001	2.3949E-05	6.0942	4.8935
3	[256, 128, 64]	32	5	0.0001	2.4388E-05	5.9157	5.0384
3	[256, 128, 64]	32	6	0.0001	2.4829E-05	6.5525	5.5470
2	[256, 128]	32	4	0.001	2.4860E-05	5.9293	4.7937
3	[256, 128, 64]	32	7	0.0001	2.5303E-05	6.3646	5.4387
4	[256, 128, 64, 32]	32	7	0.0001	2.5446E-05	13.3235	14.2956
4	[256, 128, 64, 32]	32	6	0.0001	2.6230E-05	7.6760	7.0897
5	[8, 16, 32, 64, 128]	64	4	0.001	2.6300E-05	7.1410	5.7360
2	[256, 128]	32	4	0.0001	2.6394E-05	7.3672	6.7771
3	[8, 16, 32]	32	6	0.001	2.6638E-05	9.8458	9.5144
2	[256, 128]	64	6	0.0001	2.7226E-05	6.2051	5.2735
2	[256, 128]	64	7	0.0001	2.7879E-05	7.4899	6.7220
3	[256, 128, 64]	32	4	0.0001	2.8174E-05	6.3122	5.2046
5	[256, 128, 64, 32, 16]	64	2	0.001	2.8249E-05	7.5891	6.4662
5	[256, 128, 64, 32, 16]	32	7	0.0001	2.8626E-05	6.4853	5.3724
3	[256, 128, 64]	64	6	0.0001	2.8860E-05	6.5586	5.4301
5	[8, 16, 32, 64, 128]	32	3	0.001	2.8901E-05	10.2785	9.8863
6	[8, 16, 32, 64, 128, 256]	32	2	0.001	2.9248E-05	6.6350	3.5006
4	[256, 128, 64, 32]	64	2	0.001	2.9762E-05	10.6401	10.9414

2	[256, 128]	64	5	0.0001	2.9797E-05	7.2412	6.3731
2	[256, 128]	64	3	0.001	2.9913E-05	7.0252	6.0528
5	[256, 128, 64, 32, 16]	32	4	0.0001	3.0125E-05	9.7618	9.5835
2	[256, 128]	32	3	0.0001	3.0279E-05	6.6870	5.7165
4	[256, 128, 64, 32]	32	4	0.0001	3.0423E-05	7.5400	6.7123
5	[256, 128, 64, 32, 16]	32	6	0.0001	3.0466E-05	6.8638	5.7239
6	[256, 128, 64, 32, 16, 8]	32	7	0.0001	3.0647E-05	7.1240	6.0459
2	[256, 128]	32	3	0.001	3.0951E-05	6.6159	5.4601
6	[8, 16, 32, 64, 128, 256]	64	3	0.001	3.1204E-05	7.0508	4.8351
5	[256, 128, 64, 32, 16]	32	5	0.0001	3.1269E-05	6.8345	5.6361
4	[8, 16, 32, 64]	32	5	0.001	3.1772E-05	6.9430	5.5354
3	[256, 128, 64]	32	3	0.0001	3.1776E-05	7.1414	6.1583
5	[8, 16, 32, 64, 128]	64	3	0.001	3.2058E-05	6.8249	5.2569
4	[256, 128, 64, 32]	64	6	0.0001	3.2227E-05	6.8016	5.5634
4	[256, 128, 64, 32]	32	3	0.0001	3.2274E-05	6.7910	5.6922
5	[256, 128, 64, 32, 16]	64	6	0.0001	3.2434E-05	7.5764	6.4085
3	[256, 128, 64]	64	5	0.0001	3.2438E-05	6.7730	5.7008
2	[256, 128]	64	4	0.0001	3.2973E-05	6.9857	5.9514
4	[8, 16, 32, 64]	64	7	0.001	3.3032E-05	7.8163	6.9925
6	[256, 128, 64, 32, 16, 8]	32	6	0.0001	3.3057E-05	6.9676	5.7459
3	[256, 128, 64]	64	4	0.0001	3.3417E-05	8.0075	7.0559
5	[256, 128, 64, 32, 16]	64	7	0.0001	3.3667E-05	7.0628	5.7938
4	[8, 16, 32, 64]	64	5	0.001	3.4126E-05	7.1479	5.9726
6	[256, 128, 64, 32, 16, 8]	32	5	0.0001	3.4220E-05	7.1739	5.7749
4	[256, 128, 64, 32]	64	4	0.0001	3.4664E-05	7.4887	6.3650
5	[256, 128, 64, 32, 16]	64	5	0.0001	3.4893E-05	9.5522	9.1745
3	[256, 128, 64]	64	3	0.0001	3.5034E-05	8.1297	7.1875
6	[256, 128, 64, 32, 16, 8]	64	2	0.001	3.5392E-05	8.1108	7.0877
2	[256, 128]	64	3	0.0001	3.5418E-05	7.2606	6.1968
5	[256, 128, 64, 32, 16]	32	3	0.0001	3.5419E-05	7.2239	6.0559
4	[8, 16, 32, 64]	64	6	0.001	3.5442E-05	10.2144	9.9202
3	[256, 128, 64]	64	2	0.001	3.5517E-05	7.4901	6.3907
3	[8, 16, 32]	32	5	0.001	3.5822E-05	7.3482	6.1812
4	[256, 128, 64, 32]	64	3	0.0001	3.6209E-05	7.2035	5.9543
1	[256]	32	7	0.0001	3.6491E-05	7.4023	6.2946
3	[256, 128, 64]	32	2	0.001	3.6735E-05	7.5819	6.5463
4	[8, 16, 32, 64]	32	3	0.001	3.6941E-05	7.2782	6.1920
1	[256]	32	6	0.0001	3.6944E-05	7.6453	6.5724
2	[256, 128]	64	2	0.001	3.7366E-05	7.6457	6.5574
2	[256, 128]	32	2	0.0001	3.7747E-05	8.1517	7.2709
5	[256, 128, 64, 32, 16]	64	3	0.0001	3.8011E-05	8.6028	7.8046
4	[8, 16, 32, 64]	64	4	0.001	3.8734E-05	7.8477	6.8590
3	[8, 16, 32]	64	6	0.001	3.8745E-05	7.4407	6.2585

5	[256, 128, 64, 32, 16]	64	4	0.0001	3.8785E-05	7.4677	6.0580
1	[256]	32	5	0.0001	3.8928E-05	8.6191	7.8591
5	[8, 16, 32, 64, 128]	32	2	0.001	3.8964E-05	7.4232	5.8983
6	[256, 128, 64, 32, 16, 8]	32	4	0.0001	3.9530E-05	14.7633	15.8659
4	[256, 128, 64, 32]	32	2	0.0001	3.9633E-05	8.5976	7.5891
1	[256]	32	4	0.0001	3.9641E-05	10.9760	10.4847
3	[256, 128, 64]	32	2	0.0001	3.9812E-05	7.7569	6.6759
2	[8, 16]	64	5	0.001	3.9864E-05	8.7902	7.8959
2	[256, 128]	32	2	0.001	3.9949E-05	12.3220	12.4769
2	[8, 16]	32	6	0.001	3.9969E-05	7.8608	6.8063
2	[256, 128]	64	2	0.0001	4.0149E-05	7.6364	6.4608
5	[8, 16, 32, 64, 128]	32	7	0.0001	4.0386E-05	9.2353	8.4358
5	[256, 128, 64, 32, 16]	32	2	0.0001	4.0488E-05	7.6985	6.4446
3	[8, 16, 32]	64	7	0.001	4.0502E-05	8.1438	7.0603
3	[8, 16, 32]	32	4	0.001	4.0567E-05	9.3023	8.4586
3	[256, 128, 64]	64	2	0.0001	4.1058E-05	8.8311	8.0063
4	[8, 16, 32, 64]	64	3	0.001	4.1185E-05	10.8772	10.6928
1	[256]	64	5	0.001	4.1371E-05	8.2350	7.1423
4	[256, 128, 64, 32]	64	2	0.0001	4.1585E-05	7.6687	6.4220
1	[256]	64	7	0.0001	4.1657E-05	7.7315	6.4918
5	[8, 16, 32, 64, 128]	64	2	0.001	4.1731E-05	7.6822	5.9499
6	[256, 128, 64, 32, 16, 8]	32	3	0.0001	4.1806E-05	9.9034	9.4658
1	[256]	64	5	0.0001	4.2024E-05	7.9248	6.8086
6	[256, 128, 64, 32, 16, 8]	64	6	0.0001	4.2242E-05	7.7291	6.2509
3	[8, 16, 32]	64	5	0.001	4.2549E-05	8.8618	7.9430
2	[8, 16]	64	7	0.001	4.2832E-05	9.1999	8.3191
1	[256]	64	7	0.001	4.2959E-05	7.7944	6.6078
1	[256]	32	3	0.0001	4.3150E-05	8.2791	7.2473
1	[256]	64	3	0.001	4.3534E-05	14.3531	14.5583
1	[256]	64	6	0.0001	4.3726E-05	9.0153	8.1073
6	[8, 16, 32, 64, 128, 256]	32	6	0.0001	4.3766E-05	8.1686	6.9900
6	[8, 16, 32, 64, 128, 256]	32	4	0.0001	4.4823E-05	9.4509	8.6268
4	[8, 16, 32, 64]	32	6	0.0001	4.5011E-05	8.5422	7.2933
6	[256, 128, 64, 32, 16, 8]	32	2	0.0001	4.5054E-05	8.1255	6.8913
2	[8, 16]	32	5	0.001	4.5176E-05	8.9116	7.9973
1	[256]	64	4	0.001	4.5336E-05	8.0177	6.9415
1	[8]	32	6	0.001	4.5565E-05	8.3735	7.3077
3	[8, 16, 32]	64	3	0.001	4.5701E-05	8.0547	6.9460
4	[8, 16, 32, 64]	32	7	0.0001	4.5798E-05	10.9760	10.6451
3	[8, 16, 32]	64	4	0.001	4.5950E-05	8.1253	6.9967
3	[8, 16, 32]	32	3	0.001	4.6553E-05	9.6739	8.5948
2	[8, 16]	32	3	0.001	4.6618E-05	9.4216	8.5517
1	[256]	32	4	0.001	4.6704E-05	10.8536	10.3771

4	[8, 16, 32, 64]	32	5	0.0001	4.6747E-05	9.1209	8.1430
4	[8, 16, 32, 64]	32	2	0.001	4.6883E-05	8.1426	6.7637
6	[8, 16, 32, 64, 128, 256]	32	3	0.0001	4.6896E-05	8.1504	6.8386
1	[256]	64	4	0.0001	4.6956E-05	8.3176	7.0890
6	[8, 16, 32, 64, 128, 256]	32	7	0.0001	4.7151E-05	8.5824	7.4050
1	[256]	32	6	0.001	4.7219E-05	8.7486	7.6457
2	[8, 16]	64	6	0.001	4.7555E-05	13.1416	13.3979
1	[256]	32	3	0.001	4.7669E-05	8.4208	7.2520
1	[256]	64	6	0.001	4.7922E-05	8.3067	6.9167
5	[8, 16, 32, 64, 128]	64	7	0.0001	4.7953E-05	8.9456	8.0048
6	[256, 128, 64, 32, 16, 8]	64	3	0.0001	4.8188E-05	8.8828	7.6803
1	[256]	32	5	0.001	4.8535E-05	8.9707	7.8452
2	[8, 16]	64	4	0.001	4.8845E-05	9.4909	8.6136
2	[8, 16]	32	4	0.001	4.9019E-05	8.3564	7.1980
6	[8, 16, 32, 64, 128, 256]	32	2	0.0001	4.9800E-05	9.1201	7.9597
1	[8]	32	7	0.001	4.9924E-05	11.4599	10.9542
5	[8, 16, 32, 64, 128]	32	4	0.0001	5.0302E-05	9.0543	7.9154
5	[8, 16, 32, 64, 128]	64	4	0.0001	5.0593E-05	9.2717	8.2716
2	[8, 16]	32	2	0.001	5.0896E-05	8.7176	7.5590
3	[8, 16, 32]	32	6	0.0001	5.0945E-05	10.7654	10.1071
1	[8]	64	7	0.001	5.1190E-05	8.8011	7.7635
5	[8, 16, 32, 64, 128]	32	3	0.0001	5.1208E-05	11.7008	11.2963
1	[8]	64	5	0.001	5.1453E-05	8.6886	7.5686
3	[8, 16, 32]	32	7	0.0001	5.1684E-05	8.9837	7.8358
1	[256]	64	2	0.001	5.1994E-05	8.7453	7.6658
6	[8, 16, 32, 64, 128, 256]	64	6	0.0001	5.2152E-05	8.5880	7.3877
4	[8, 16, 32, 64]	32	3	0.0001	5.2423E-05	12.2273	12.0013
5	[8, 16, 32, 64, 128]	64	6	0.0001	5.2495E-05	8.6162	7.3447
3	[8, 16, 32]	32	2	0.001	5.2910E-05	8.8738	7.7278
1	[8]	32	5	0.001	5.3051E-05	12.7803	12.6718
1	[8]	32	3	0.001	5.3182E-05	12.5405	11.8025
6	[256, 128, 64, 32, 16, 8]	64	2	0.0001	5.3589E-05	10.5080	9.7921
6	[8, 16, 32, 64, 128, 256]	64	4	0.0001	5.3612E-05	8.7074	7.4108
6	[8, 16, 32, 64, 128, 256]	64	5	0.0001	5.3735E-05	8.9742	7.9781
1	[8]	64	6	0.001	5.3943E-05	8.9078	7.7393
4	[8, 16, 32, 64]	32	4	0.0001	5.4198E-05	8.9744	7.8350
5	[8, 16, 32, 64, 128]	32	2	0.0001	5.4721E-05	9.1934	8.1380
5	[8, 16, 32, 64, 128]	32	5	0.0001	5.4752E-05	9.7656	8.7175
4	[8, 16, 32, 64]	64	6	0.0001	5.5013E-05	8.9215	7.6279
3	[8, 16, 32]	32	5	0.0001	5.5345E-05	8.8501	7.5463
1	[256]	32	2	0.001	5.5910E-05	11.3640	10.8817
4	[8, 16, 32, 64]	64	7	0.0001	5.6399E-05	8.9309	7.7435
1	[8]	32	4	0.001	5.6590E-05	9.4211	8.2691

4	[8, 16, 32, 64]	64	2	0.001	5.6628E-05	9.2089	7.9730
6	[8, 16, 32, 64, 128, 256]	64	3	0.0001	5.6764E-05	9.0866	7.8791
2	[8, 16]	64	3	0.001	5.6970E-05	10.5624	9.6171
5	[8, 16, 32, 64, 128]	64	3	0.0001	5.7666E-05	9.0493	7.7355
2	[8, 16]	32	3	0.0001	5.7684E-05	9.1177	7.9368
3	[8, 16, 32]	32	4	0.0001	5.7787E-05	10.6995	9.6784
1	[8]	64	2	0.001	5.8764E-05	10.3674	9.5558
1	[8]	64	4	0.001	5.9191E-05	9.6411	8.6374
1	[8]	32	2	0.001	5.9698E-05	9.6269	8.3624
4	[8, 16, 32, 64]	64	5	0.0001	5.9912E-05	9.2770	8.1124
3	[8, 16, 32]	64	2	0.001	6.0674E-05	10.1039	9.0388
2	[8, 16]	32	4	0.0001	6.1486E-05	11.3913	10.7654
1	[8]	64	3	0.001	6.1858E-05	11.0286	10.1521
3	[8, 16, 32]	32	2	0.0001	6.2129E-05	10.2622	9.0622
3	[8, 16, 32]	64	6	0.0001	6.2130E-05	9.5066	8.3071
2	[8, 16]	64	7	0.0001	6.2380E-05	9.3924	8.1916
1	[256]	32	2	0.0001	6.2424E-05	10.9937	10.1535
4	[8, 16, 32, 64]	32	2	0.0001	6.2498E-05	9.4013	8.2449
4	[8, 16, 32, 64]	64	4	0.0001	6.3379E-05	9.7164	8.3831
2	[8, 16]	32	7	0.0001	6.3732E-05	9.4937	8.2356
4	[8, 16, 32, 64]	64	3	0.0001	6.3832E-05	10.0154	8.9515
3	[8, 16, 32]	64	7	0.0001	6.4387E-05	9.6171	8.4653
3	[8, 16, 32]	32	3	0.0001	6.4478E-05	10.0395	8.9891
2	[8, 16]	32	6	0.0001	6.4988E-05	9.7129	8.6144
1	[256]	64	3	0.0001	6.5383E-05	9.8162	8.6328
2	[8, 16]	32	2	0.0001	6.5486E-05	9.6234	8.4923
2	[8, 16]	64	2	0.001	6.5558E-05	10.4136	9.3131
2	[8, 16]	32	5	0.0001	6.6275E-05	10.1561	8.9875
2	[8, 16]	64	6	0.0001	6.6614E-05	9.8782	8.7862
2	[8, 16]	64	3	0.0001	6.7338E-05	9.9258	8.7272
2	[8, 16]	64	4	0.0001	6.8259E-05	9.9306	8.6978
3	[8, 16, 32]	64	3	0.0001	6.9511E-05	10.1211	8.9512
5	[8, 16, 32, 64, 128]	64	2	0.0001	6.9670E-05	10.0463	9.0034
3	[8, 16, 32]	64	5	0.0001	6.9893E-05	10.1104	8.9060
2	[8, 16]	64	2	0.0001	7.0501E-05	10.1131	9.1142
1	[8]	32	5	0.0001	7.1010E-05	10.2512	9.1622
1	[8]	32	6	0.0001	7.1874E-05	10.7316	9.6821
2	[8, 16]	64	5	0.0001	7.2323E-05	10.8969	9.8204
3	[8, 16, 32]	64	2	0.0001	7.2343E-05	10.1147	9.0176
4	[8, 16, 32, 64]	64	2	0.0001	7.2624E-05	10.2785	9.1403
1	[8]	32	4	0.0001	7.3905E-05	10.5495	9.3637
1	[256]	64	2	0.0001	7.5849E-05	10.8004	9.6905
1	[8]	32	3	0.0001	7.7711E-05	10.7055	9.5225

1	[8]	32	2	0.0001	7.8002E-05	10.7104	9.6599
1	[8]	64	3	0.0001	8.0652E-05	10.8689	9.8926
1	[8]	64	7	0.0001	8.2221E-05	10.7832	9.6394
1	[8]	64	6	0.0001	8.5979E-05	11.0268	10.0616
1	[8]	64	2	0.0001	8.8843E-05	11.2090	10.2801
1	[8]	64	5	0.0001	8.9671E-05	11.2611	10.2309
1	[8]	64	4	0.0001	9.3078E-05	11.6804	10.6817
6	[256, 128, 64, 32, 16, 8]	64	4	0.0001	7.4126E-01	1023.8622	1219.2432
4	[256, 128, 64, 32]	64	7	0.0001	7.4126E-01	1023.8622	1219.2432
3	[256, 128, 64]	64	7	0.0001	7.4126E-01	1023.8622	1219.2432
6	[8, 16, 32, 64, 128, 256]	64	7	0.0001	7.4126E-01	1023.8622	1219.2432
6	[8, 16, 32, 64, 128, 256]	64	2	0.0001	7.4126E-01	1023.8622	1219.2432
4	[256, 128, 64, 32]	64	7	0.001	7.4126E-01	1023.8622	1219.2432
6	[256, 128, 64, 32, 16, 8]	64	6	0.001	7.4126E-01	1023.8622	1219.2432
5	[8, 16, 32, 64, 128]	64	5	0.0001	7.4126E-01	1023.8622	1219.2432
6	[8, 16, 32, 64, 128, 256]	64	5	0.001	7.4126E-01	1023.8622	1219.2432
6	[256, 128, 64, 32, 16, 8]	64	7	0.0001	7.4126E-01	1023.8622	1219.2432
6	[8, 16, 32, 64, 128, 256]	64	2	0.001	7.4126E-01	1023.8622	1219.2432
3	[8, 16, 32]	64	4	0.0001	7.4126E-01	1023.8622	1219.2432
4	[256, 128, 64, 32]	64	5	0.0001	7.4126E-01	1023.8622	1219.2432
5	[256, 128, 64, 32, 16]	64	2	0.0001	7.4126E-01	1023.8622	1219.2432
6	[256, 128, 64, 32, 16, 8]	64	5	0.0001	7.4126E-01	1023.8622	1219.2432
2	[8, 16]	32	7	0.001	7.4126E-01	1023.8622	1219.2432
4	[256, 128, 64, 32]	32	5	0.001	7.4126E-01	1023.8622	1219.2432
6	[256, 128, 64, 32, 16, 8]	32	4	0.001	7.4126E-01	1023.8622	1219.2432
5	[8, 16, 32, 64, 128]	32	6	0.0001	7.4126E-01	1023.8622	1219.2432
1	[256]	32	7	0.001	7.4126E-01	1023.8622	1219.2431
5	[256, 128, 64, 32, 16]	32	7	0.001	7.4126E-01	1023.8622	1219.2432
4	[8, 16, 32, 64]	32	4	0.001	7.4126E-01	1023.8622	1219.2432
6	[256, 128, 64, 32, 16, 8]	32	3	0.001	7.4126E-01	1023.8622	1219.2432
6	[256, 128, 64, 32, 16, 8]	32	7	0.001	7.4126E-01	1023.8622	1219.2432
6	[256, 128, 64, 32, 16, 8]	32	2	0.001	7.4126E-01	1023.8622	1219.2432
5	[256, 128, 64, 32, 16]	32	2	0.001	7.4126E-01	1023.8622	1219.2432
4	[256, 128, 64, 32]	32	5	0.0001	7.4126E-01	1023.8622	1219.2432
6	[8, 16, 32, 64, 128, 256]	32	5	0.0001	7.4126E-01	1023.8622	1219.2432
1	[8]	32	7	0.0001	7.4126E-01	1023.8622	1219.2432
3	[8, 16, 32]	32	7	0.001	7.4126E-01	1023.8622	1219.2432

Table A.2. The selected 1D-CNN architecture for transformer inter-turn fault based on highest accuracy on the validation set.

Layer size	**Filter number**	**Batch size**	**Kernel size**	**Adam optimizer**	**Accuracy**, %
1	[256]	64	3	0.001	99.8629
4	[256, 128, 64, 32]	32	3	0.0001	99.8458
3	[256, 128, 64]	64	5	0.001	99.8458
4	[8, 16, 32, 64]	32	6	0.001	99.8458
3	[256, 128, 64]	32	2	0.0001	99.8287
2	[256, 128]	32	2	0.0001	99.8287
1	[256]	64	7	0.001	99.8287
3	[8, 16, 32]	64	6	0.001	99.8287
2	[256, 128]	64	3	0.0001	99.8115
2	[256, 128]	32	4	0.0001	99.8115
1	[256]	64	5	0.001	99.7944
2	[256, 128]	32	5	0.0001	99.7944
1	[256]	64	4	0.001	99.7944
6	[256, 128, 64, 32, 16, 8]	64	7	0.001	99.7944
2	[256, 128]	64	5	0.001	99.7944
1	[256]	32	6	0.001	99.7944
1	[8]	32	4	0.001	99.7944
3	[256, 128, 64]	64	7	0.001	99.7944
2	[8, 16]	32	7	0.001	99.7944
5	[256, 128, 64, 32, 16]	32	4	0.0001	99.7944
2	[8, 16]	64	2	0.001	99.7944
4	[256, 128, 64, 32]	64	6	0.0001	99.7944
2	[8, 16]	64	3	0.001	99.7944
2	[256, 128]	64	4	0.0001	99.7773
4	[8, 16, 32, 64]	64	3	0.001	99.7773
2	[256, 128]	64	6	0.0001	99.7773
3	[256, 128, 64]	64	3	0.001	99.7773
2	[256, 128]	32	7	0.0001	99.7773
3	[256, 128, 64]	64	7	0.0001	99.7773
4	[8, 16, 32, 64]	64	6	0.001	99.7773
3	[8, 16, 32]	64	5	0.001	99.7773
3	[256, 128, 64]	32	7	0.0001	99.7773
3	[8, 16, 32]	32	7	0.001	99.7773
2	[256, 128]	64	4	0.001	99.7773
3	[256, 128, 64]	32	6	0.001	99.7602
2	[256, 128]	32	6	0.001	99.7602
2	[256, 128]	32	7	0.001	99.7602
6	[256, 128, 64, 32, 16, 8]	32	6	0.001	99.7602
4	[8, 16, 32, 64]	32	7	0.001	99.7602

1	[256]	64	6	0.001	99.7602
4	[256, 128, 64, 32]	64	6	0.001	99.7602
3	[256, 128, 64]	32	3	0.001	99.7602
2	[256, 128]	64	5	0.0001	99.7602
3	[256, 128, 64]	64	5	0.0001	99.7602
1	[8]	32	5	0.001	99.7602
3	[256, 128, 64]	64	6	0.001	99.7602
5	[256, 128, 64, 32, 16]	32	3	0.0001	99.7430
1	[8]	32	7	0.001	99.7430
2	[256, 128]	64	6	0.001	99.7430
2	[8, 16]	32	2	0.001	99.7430
4	[256, 128, 64, 32]	64	7	0.0001	99.7430
6	[256, 128, 64, 32, 16, 8]	32	5	0.0001	99.7430
6	[8, 16, 32, 64, 128, 256]	64	6	0.001	99.7430
2	[256, 128]	64	2	0.0001	99.7430
6	[256, 128, 64, 32, 16, 8]	32	7	0.001	99.7430
1	[256]	32	2	0.001	99.7430
2	[256, 128]	32	2	0.001	99.7430
3	[256, 128, 64]	32	2	0.001	99.7430
4	[256, 128, 64, 32]	32	4	0.001	99.7430
3	[256, 128, 64]	32	5	0.0001	99.7430
4	[256, 128, 64, 32]	64	4	0.0001	99.7430
4	[256, 128, 64, 32]	32	6	0.0001	99.7430
3	[8, 16, 32]	32	3	0.001	99.7259
2	[256, 128]	64	7	0.001	99.7259
2	[8, 16]	64	7	0.001	99.7259
4	[256, 128, 64, 32]	32	4	0.0001	99.7259
5	[8, 16, 32, 64, 128]	64	7	0.001	99.7259
5	[8, 16, 32, 64, 128]	32	4	0.0001	99.7259
1	[256]	32	7	0.001	99.7259
4	[256, 128, 64, 32]	32	5	0.0001	99.7259
3	[256, 128, 64]	64	2	0.0001	99.7259
5	[256, 128, 64, 32, 16]	32	6	0.0001	99.7259
3	[256, 128, 64]	64	4	0.0001	99.7259
2	[256, 128]	64	7	0.0001	99.7259
5	[8, 16, 32, 64, 128]	32	7	0.0001	99.7088
2	[8, 16]	64	4	0.001	99.7088
1	[8]	64	3	0.001	99.7088
5	[256, 128, 64, 32, 16]	32	5	0.0001	99.7088
6	[256, 128, 64, 32, 16, 8]	32	7	0.0001	99.7088
4	[256, 128, 64, 32]	32	7	0.0001	99.7088
5	[256, 128, 64, 32, 16]	64	4	0.001	99.7088
3	[256, 128, 64]	32	4	0.0001	99.7088

6	[256, 128, 64, 32, 16, 8]	64	5	0.001	99.7088
2	[8, 16]	64	6	0.001	99.7088
2	[8, 16]	64	5	0.001	99.7088
6	[256, 128, 64, 32, 16, 8]	64	7	0.0001	99.7088
5	[256, 128, 64, 32, 16]	64	7	0.001	99.7088
5	[8, 16, 32, 64, 128]	32	4	0.001	99.7088
4	[256, 128, 64, 32]	64	5	0.0001	99.7088
2	[8, 16]	32	3	0.001	99.7088
4	[256, 128, 64, 32]	64	3	0.0001	99.7088
6	[256, 128, 64, 32, 16, 8]	32	5	0.001	99.7088
4	[256, 128, 64, 32]	32	2	0.0001	99.7088
2	[256, 128]	32	4	0.001	99.6916
5	[256, 128, 64, 32, 16]	64	4	0.0001	99.6916
6	[8, 16, 32, 64, 128, 256]	32	4	0.001	99.6916
4	[8, 16, 32, 64]	64	4	0.001	99.6916
5	[256, 128, 64, 32, 16]	32	4	0.001	99.6916
2	[8, 16]	32	5	0.001	99.6916
6	[8, 16, 32, 64, 128, 256]	64	5	0.0001	99.6916
5	[8, 16, 32, 64, 128]	64	5	0.0001	99.6916
3	[256, 128, 64]	64	6	0.0001	99.6916
5	[256, 128, 64, 32, 16]	32	5	0.001	99.6916
6	[256, 128, 64, 32, 16, 8]	32	6	0.0001	99.6916
1	[8]	32	6	0.001	99.6916
4	[8, 16, 32, 64]	32	3	0.001	99.6916
6	[256, 128, 64, 32, 16, 8]	32	2	0.001	99.6916
5	[256, 128, 64, 32, 16]	32	7	0.001	99.6916
5	[256, 128, 64, 32, 16]	32	2	0.001	99.6916
4	[8, 16, 32, 64]	64	2	0.001	99.6916
5	[256, 128, 64, 32, 16]	64	7	0.0001	99.6916
2	[256, 128]	32	3	0.001	99.6745
1	[256]	64	2	0.001	99.6745
6	[256, 128, 64, 32, 16, 8]	64	3	0.001	99.6745
1	[256]	32	7	0.0001	99.6745
1	[256]	32	4	0.001	99.6745
6	[256, 128, 64, 32, 16, 8]	64	6	0.001	99.6745
3	[8, 16, 32]	64	2	0.001	99.6745
1	[256]	32	5	0.0001	99.6745
6	[8, 16, 32, 64, 128, 256]	32	2	0.0001	99.6745
4	[256, 128, 64, 32]	32	6	0.001	99.6745
5	[256, 128, 64, 32, 16]	64	5	0.001	99.6745
6	[8, 16, 32, 64, 128, 256]	32	6	0.001	99.6745
3	[256, 128, 64]	32	5	0.001	99.6745
4	[8, 16, 32, 64]	32	2	0.001	99.6574

4	[256, 128, 64, 32]	32	2	0.001	99.6574
5	[8, 16, 32, 64, 128]	64	2	0.001	99.6574
2	[256, 128]	32	3	0.0001	99.6574
5	[8, 16, 32, 64, 128]	64	3	0.0001	99.6574
3	[256, 128, 64]	64	2	0.001	99.6574
5	[256, 128, 64, 32, 16]	64	2	0.001	99.6574
6	[8, 16, 32, 64, 128, 256]	64	7	0.0001	99.6574
3	[256, 128, 64]	32	4	0.001	99.6574
5	[8, 16, 32, 64, 128]	64	5	0.001	99.6574
4	[256, 128, 64, 32]	32	3	0.001	99.6574
3	[256, 128, 64]	32	6	0.0001	99.6574
3	[8, 16, 32]	64	3	0.001	99.6574
2	[8, 16]	32	4	0.001	99.6574
2	[256, 128]	64	3	0.001	99.6574
6	[8, 16, 32, 64, 128, 256]	64	6	0.0001	99.6574
3	[8, 16, 32]	32	6	0.0001	99.6574
3	[8, 16, 32]	32	5	0.001	99.6574
5	[8, 16, 32, 64, 128]	64	6	0.0001	99.6574
5	[8, 16, 32, 64, 128]	64	4	0.001	99.6574
6	[256, 128, 64, 32, 16, 8]	64	4	0.0001	99.6402
4	[256, 128, 64, 32]	64	2	0.0001	99.6402
5	[8, 16, 32, 64, 128]	64	7	0.0001	99.6402
1	[8]	64	2	0.001	99.6402
5	[8, 16, 32, 64, 128]	32	7	0.001	99.6402
6	[8, 16, 32, 64, 128, 256]	32	4	0.0001	99.6402
4	[8, 16, 32, 64]	32	4	0.0001	99.6402
5	[8, 16, 32, 64, 128]	32	3	0.0001	99.6402
6	[256, 128, 64, 32, 16, 8]	32	3	0.001	99.6402
5	[256, 128, 64, 32, 16]	32	2	0.0001	99.6402
2	[256, 128]	32	5	0.001	99.6402
1	[256]	32	6	0.0001	99.6402
4	[256, 128, 64, 32]	32	7	0.001	99.6402
6	[256, 128, 64, 32, 16, 8]	32	4	0.0001	99.6231
3	[8, 16, 32]	64	4	0.001	99.6231
1	[8]	64	4	0.001	99.6231
1	[8]	64	7	0.001	99.6231
6	[8, 16, 32, 64, 128, 256]	64	7	0.001	99.6231
1	[8]	32	2	0.001	99.6231
3	[8, 16, 32]	32	6	0.001	99.6231
6	[8, 16, 32, 64, 128, 256]	32	5	0.001	99.6231
4	[8, 16, 32, 64]	32	4	0.001	99.6231
1	[256]	32	3	0.001	99.6231
4	[8, 16, 32, 64]	32	7	0.0001	99.6231

2	[256, 128]	64	2	0.001	99.6231
6	[8, 16, 32, 64, 128, 256]	32	7	0.0001	99.6231
6	[8, 16, 32, 64, 128, 256]	64	5	0.001	99.6231
2	[8, 16]	32	6	0.001	99.6060
4	[8, 16, 32, 64]	64	6	0.0001	99.6060
4	[8, 16, 32, 64]	32	3	0.0001	99.6060
6	[8, 16, 32, 64, 128, 256]	32	3	0.0001	99.6060
4	[256, 128, 64, 32]	64	5	0.001	99.6060
5	[256, 128, 64, 32, 16]	32	7	0.0001	99.6060
6	[8, 16, 32, 64, 128, 256]	32	6	0.0001	99.6060
5	[8, 16, 32, 64, 128]	32	5	0.0001	99.6060
4	[8, 16, 32, 64]	32	5	0.0001	99.5888
5	[8, 16, 32, 64, 128]	32	2	0.001	99.5888
5	[8, 16, 32, 64, 128]	32	6	0.0001	99.5888
6	[256, 128, 64, 32, 16, 8]	64	6	0.0001	99.5888
5	[256, 128, 64, 32, 16]	64	6	0.0001	99.5888
6	[256, 128, 64, 32, 16, 8]	32	2	0.0001	99.5888
4	[256, 128, 64, 32]	32	5	0.001	99.5888
5	[256, 128, 64, 32, 16]	64	3	0.0001	99.5888
6	[8, 16, 32, 64, 128, 256]	64	3	0.001	99.5888
6	[8, 16, 32, 64, 128, 256]	64	4	0.001	99.5888
2	[256, 128]	32	6	0.0001	99.5888
6	[8, 16, 32, 64, 128, 256]	64	4	0.0001	99.5717
4	[8, 16, 32, 64]	64	4	0.0001	99.5717
1	[256]	32	5	0.001	99.5717
6	[256, 128, 64, 32, 16, 8]	64	3	0.0001	99.5546
6	[256, 128, 64, 32, 16, 8]	64	5	0.0001	99.5546
6	[8, 16, 32, 64, 128, 256]	32	5	0.0001	99.5546
3	[8, 16, 32]	32	4	0.001	99.5546
5	[256, 128, 64, 32, 16]	64	5	0.0001	99.5546
1	[8]	64	5	0.001	99.5546
3	[256, 128, 64]	32	3	0.0001	99.5374
4	[8, 16, 32, 64]	64	7	0.0001	99.5374
3	[256, 128, 64]	32	7	0.001	99.5374
5	[256, 128, 64, 32, 16]	64	2	0.0001	99.5374
3	[8, 16, 32]	32	2	0.001	99.5374
6	[8, 16, 32, 64, 128, 256]	64	2	0.001	99.5374
5	[256, 128, 64, 32, 16]	64	6	0.001	99.5374
4	[8, 16, 32, 64]	64	5	0.0001	99.5374
4	[8, 16, 32, 64]	64	5	0.001	99.5203
4	[8, 16, 32, 64]	64	3	0.0001	99.5203
3	[256, 128, 64]	64	3	0.0001	99.5203
4	[8, 16, 32, 64]	32	5	0.001	99.5032

4	[8, 16, 32, 64]	64	7	0.001	99.5032
4	[256, 128, 64, 32]	64	3	0.001	99.5032
4	[8, 16, 32, 64]	32	2	0.0001	99.5032
3	[8, 16, 32]	32	4	0.0001	99.5032
1	[256]	64	3	0.0001	99.5032
4	[256, 128, 64, 32]	64	2	0.001	99.4860
5	[8, 16, 32, 64, 128]	64	2	0.0001	99.4860
6	[256, 128, 64, 32, 16, 8]	32	3	0.0001	99.4860
2	[8, 16]	32	4	0.0001	99.4860
4	[256, 128, 64, 32]	64	4	0.001	99.4689
6	[256, 128, 64, 32, 16, 8]	64	4	0.001	99.4689
4	[256, 128, 64, 32]	64	7	0.001	99.4689
1	[256]	32	3	0.0001	99.4518
6	[8, 16, 32, 64, 128, 256]	32	2	0.001	99.4518
3	[8, 16, 32]	32	7	0.0001	99.4518
1	[8]	32	3	0.001	99.4518
5	[8, 16, 32, 64, 128]	64	3	0.001	99.4518
3	[8, 16, 32]	64	7	0.001	99.4346
4	[8, 16, 32, 64]	32	6	0.0001	99.4346
1	[256]	64	5	0.0001	99.4175
3	[256, 128, 64]	64	4	0.001	99.4175
5	[8, 16, 32, 64, 128]	32	3	0.001	99.4004
3	[8, 16, 32]	64	7	0.0001	99.4004
3	[8, 16, 32]	64	6	0.0001	99.3490
1	[256]	32	2	0.0001	99.3318
5	[8, 16, 32, 64, 128]	32	6	0.001	99.3318
6	[256, 128, 64, 32, 16, 8]	64	2	0.001	99.3318
5	[8, 16, 32, 64, 128]	64	4	0.0001	99.3147
2	[8, 16]	32	6	0.0001	99.2805
1	[256]	64	6	0.0001	99.2633
3	[8, 16, 32]	32	5	0.0001	99.2633
3	[8, 16, 32]	32	3	0.0001	99.2462
3	[8, 16, 32]	32	2	0.0001	99.2291
6	[8, 16, 32, 64, 128, 256]	64	2	0.0001	99.2291
6	[256, 128, 64, 32, 16, 8]	32	4	0.001	99.2119
5	[8, 16, 32, 64, 128]	32	2	0.0001	99.2119
5	[256, 128, 64, 32, 16]	32	6	0.001	99.1948
1	[256]	64	7	0.0001	99.1777
5	[8, 16, 32, 64, 128]	32	5	0.001	99.1605
6	[8, 16, 32, 64, 128, 256]	32	3	0.001	99.1605
3	[8, 16, 32]	64	5	0.0001	99.1263
6	[8, 16, 32, 64, 128, 256]	64	3	0.0001	99.1263
2	[8, 16]	32	3	0.0001	99.0920

2	[8, 16]	64	3	0.0001	99.0920
1	[256]	32	4	0.0001	99.0749
6	[256, 128, 64, 32, 16, 8]	64	2	0.0001	99.0749
5	[8, 16, 32, 64, 128]	64	6	0.001	99.0577
2	[8, 16]	32	5	0.0001	99.0577
1	[256]	64	2	0.0001	99.0577
5	[256, 128, 64, 32, 16]	32	3	0.001	99.0577
1	[8]	64	6	0.001	99.0406
2	[8, 16]	32	7	0.0001	99.0235
4	[8, 16, 32, 64]	64	2	0.0001	99.0063
6	[8, 16, 32, 64, 128, 256]	32	7	0.001	98.9721
5	[256, 128, 64, 32, 16]	64	3	0.001	98.9207
3	[8, 16, 32]	64	3	0.0001	98.9207
1	[256]	64	4	0.0001	98.8008
3	[8, 16, 32]	64	4	0.0001	98.8008
2	[8, 16]	64	5	0.0001	98.7836
2	[8, 16]	32	2	0.0001	98.7665
1	[8]	32	2	0.0001	98.7665
2	[8, 16]	64	7	0.0001	98.7151
2	[8, 16]	64	6	0.0001	98.5780
1	[8]	32	4	0.0001	98.5609
2	[8, 16]	64	2	0.0001	98.3553
3	[8, 16, 32]	64	2	0.0001	98.3553
1	[8]	32	5	0.0001	98.3382
1	[8]	64	6	0.0001	98.3211
1	[8]	32	6	0.0001	98.2183
1	[8]	32	3	0.0001	98.0983
1	[8]	64	7	0.0001	97.9270
2	[8, 16]	64	4	0.0001	97.9099
1	[8]	64	4	0.0001	97.7386
1	[8]	64	3	0.0001	97.7214
1	[8]	32	7	0.0001	97.6872
1	[8]	64	2	0.0001	97.4302
1	[8]	64	5	0.0001	97.2417

Back Cover