

Article

# Transformer Winding Condition Assessment Using Feedforward Artificial Neural Network and Frequency Response Measurements

Mehran Tahir \*  and Stefan Tenbohlen 

Institute of Power Transmission and High Voltage Technology (IEH), Stuttgart University, Pfaffenwaldring 47, 70569 Stuttgart, Germany; stefan.tenbohlen@ieh.uni-stuttgart.de

\* Correspondence: mehran.tahir@ieh.uni-stuttgart.de; Tel.: +49-711-685-67855

**Abstract:** Frequency response analysis (FRA) is a well-known method to assess the mechanical integrity of the active parts of the power transformer. The measurement procedures of FRA are standardized as described in the IEEE and IEC standards. However, the interpretation of FRA results is far from reaching an accepted and definitive methodology as there is no reliable code available in the standard. As a contribution to this necessity, this paper presents an intelligent fault detection and classification algorithm using FRA results. The algorithm is based on a multilayer, feedforward, backpropagation artificial neural network (ANN). First, the adaptive frequency division algorithm is developed and various numerical indicators are used to quantify the differences between FRA traces and obtain feature sets for ANN. Finally, the classification model of ANN is developed to detect and classify different transformer conditions, i.e., healthy windings, healthy windings with saturated core, mechanical deformations, electrical faults, and reproducibility issues due to different test conditions. The database used in this study consists of FRA measurements from 80 power transformers of different designs, ratings, and different manufacturers. The results obtained give evidence of the effectiveness of the proposed classification model for power transformer fault diagnosis using FRA.

**Keywords:** artificial neural network (ANN); condition assessment; feature generation; frequency response analysis (FRA); numerical indices; power transformer



**Citation:** Tahir, M.; Tenbohlen, S. Transformer Winding Condition Assessment Using Feedforward Artificial Neural Network and Frequency Response Measurements. *Energies* **2021**, *14*, 3227. <https://doi.org/10.3390/en14113227>

Academic Editor: Sérgio Cruz

Received: 27 April 2021

Accepted: 28 May 2021

Published: 31 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Power transformers are one of the most vital components in today's transmission and distribution infrastructure. Growing demand for electricity requires power transformers to operate at higher loading levels. Operating at higher demands can cause deterioration of transformer integrity due to mechanical, thermal, and electrical stresses. According to the transformer's reliability survey based on 964 major failures, winding failure is the dominant failure location in power transformers (38%) [1]. Therefore, it is necessary to assess the integrity of the transformer windings.

Frequency response analysis (FRA) has drawn attention as a powerful diagnostic method for detecting faults in the active part of power transformers [2–4]. FRA is a comparative diagnostic method in which a reference FRA signature is compared to the present FRA signature and based on the deviations between the two signatures the condition of the transformer is evaluated. At present, the assessment of FRA results demands skill-full personnel as there is no reliable standard code available in IEC and IEEE standards [5,6]. Thus, interpretation of FRA results is a significant challenge towards its practical application [7].

Many recent studies have been focused on FRA interpretation to detect the extent and the type of mechanical fault. Different algorithms have been proposed for this purpose that can be categorized into three main groups: simulation models (circuit models/FEM models) [8–10], numerical indices [11–13], and artificial intelligence (AI) techniques [14,15]. However, these methods have some drawbacks and limitations. Simulation techniques are

useful to study and investigate the effect of various winding faults on FRA traces. However, the diagnosis of transformer faults using these simulation techniques is limited by the high accuracy requirements to reproduce the FRA fingerprints [16]. Numerical indices are the most widely used method for transformer frequency response assessment that quantifies the differences in a pair of FRA traces. The final target is a decision about the condition of a transformer based on the quantified values of the indices. Some drawbacks are also associated with numerical indices. Firstly, the results are only numbers, and further interpretations are not possible. Secondly, the definition of the threshold has not been established yet [11,12]. Moreover, these numerical indices are calculated in fixed frequency sub-bands, whereas standards [5,6] declare that the ranges of frequency sub-bands depend on transformer ratings and cannot be fixed. Artificial intelligence (AI) and machine learning (ML) methods have been of great interest lately for diagnosing transformer faults using FRA results and are gaining popularity due to their effectiveness and high accuracy. Using AI methods, it is possible to classify the changes in the FRA traces and connect them to different fault types and even fault locations [14,15]. These methods are based on techniques such as genetic algorithms, fuzzy logic, support vector machines, and neural networks. The major dilemma of using AI methods is the need for a database from real power transformers, and such a database usually is not available. Some studies propose to use the outputs of model experiments and circuit models as a database [14,15]. However, the applicability of these models for large-scale power transformers is itself an issue.

This paper presents an intelligent fault detection and classification algorithm using FRA results. The algorithm is based on a multilayer, feedforward, backpropagation artificial neural network (ANN). Five numerical indices are employed using an adaptive frequency band division algorithm. Indices are then complemented with ANNs to detect and classify transformer winding faults.

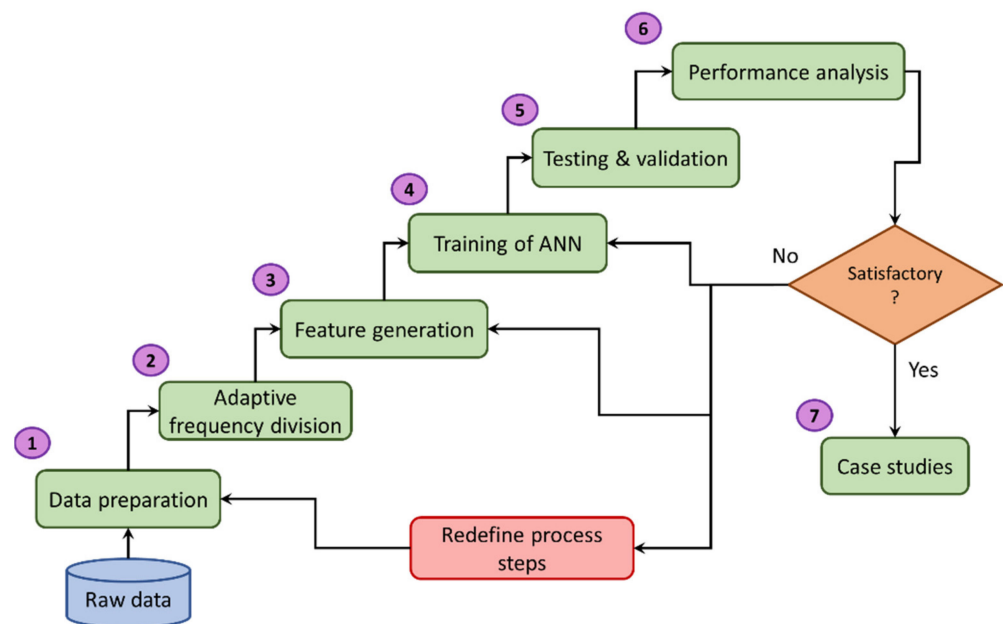
## 2. Application of Machine Learning in FRA

In literature, many efforts have been made where ML methods are implemented for the detection and identification of transformers winding faults using FRA. Zhijian et al. (2000) [17] used the correlation coefficient (CC) calculated in three fixed frequency sub-bands that serve as features for ANN. However, only 26 cases were used in total for training and testing, and only two classes, healthy and deformed, were identified. Bigdeli et al. (2012) [18] used the absolute ratio in the deformed and healthy state FRA to extract features. A support vector machine (SVM) was used to identify four types of winding faults. In this work, only two groups of setups are used. Gandhi et al. (2014) [15] used 9 indices and 90 cases for three-layer ANN. The major drawback of this work is that all faults are applied to a single transformer model. Ghanizadeh et al. (2014) [14] employed an ANN classifier to detect electrical and mechanical faults of the transformer winding. The limitation of this work is that all faults are implemented on a single 1.2 MVA transformer circuit model. Liu et al. (2019) [19] employed an SVM classifier to classify different mechanical faults. In this work, only a single transformer model is employed. Mao et al. (2019) [20] employed SVM to identify the winding type using FRA. In this study, a group of 400/275 kV transformers are tested. The main objective of this study was to provide an automatic approach for transformer asset management. The results of these reports show the potential of ML algorithms for fault diagnosis and classification. However, in all these studies, the used data sets are very small. Secondly, faults are applied to a small number of transformers, which reduces the diversity of fault patterns. Hence, a diverse dataset of different faults from the field is required to settle the criteria of using ML methods.

In contrast, the database used in this study consists of 139 FRA measurements from 80 real power transformers of different designs, ratings, and manufacturers. Moreover, six conditions of the transformer windings are identified, i.e., healthy windings, healthy windings with saturated core, mechanical deformations, shorted turn faults, open circuit faults, and reproducibility issues due to different test conditions.

### 3. Methodology

The necessary steps for building the intelligent fault detection algorithm (IFDA) for a given data set are described in Figure 1. The process includes seven major steps, i.e., data preparation, adaptive frequency band division, feature generation, training of ANN, testing and validation with unknown data, performance analysis, and validation with case studies. If the performance of the model is not satisfactory, it can be attributed to erroneous data preparation, poor feature generation, or simply due to less diversity among deviation patterns of different classes. In all of these cases, the IFDA recommends improving the performance by redefining the process steps.



**Figure 1.** Flowchart for intelligent fault detection algorithm.

#### 3.1. Database

The database used in this study consists of 139 FRA results from 80 power transformers of different designs, ratings, and different manufacturers. The database comprises different types of transformers: generator step-up unit, transmission, distribution, shunt reactor, GIS connector, dry type, etc. In the database, each FRA measurement belongs to a predefined state of the transformer. Mainly six conditions of the transformer are considered in this research: healthy windings, healthy windings with saturated core, mechanical deformations, shorted turn faults, open circuit faults, and reproducibility issues due to different test conditions. These conditions are also called ‘labels’ or ‘classes’ for IFDA. The distribution of classes in the database is shown in Figure 2.

#### 3.2. Data Preparation

Data preparation involves two main tasks: data labeling and noise removal. In data labeling, each FRA measurement is assigned a class (A–F). The FRA measurements performed in substation environments comprise a variety of noise which can influence the FRA signatures. Mainly, there are two types of sources: narrowband noise and broadband noise. Narrowband noise is due to the power frequency noise and its harmonics, whereas the broadband noise is caused by the noise bed of the connected FRA measurement equipment. The dynamic range defines the noise bed of an FRA instrument. IEC 60076-18 standard defines the minimum dynamic range for a device:  $-90$  dB to  $+10$  dB [5]. For denoising the FRA measurements, the moving average with Gaussian function is used in this work. The process of data de-noising is shown in the Figure 3. The identification of the noise frequency range and the number of main anti-resonances in the low-frequency

range should be determined first. Typically, the effect of both types of noise appears at low frequency (before the first resonance point in the end-to-end open circuit FRA signature), causing several small peaks and valleys as shown in Figure 4. The number of anti-resonances depends upon the type and configuration of the transformer. The Gaussian filter is applied with a moving window of variable size to remove small peaks and valleys due to noise until the number of valleys in this range is equal to the defined anti-resonance points. After this process, data are considered clean.

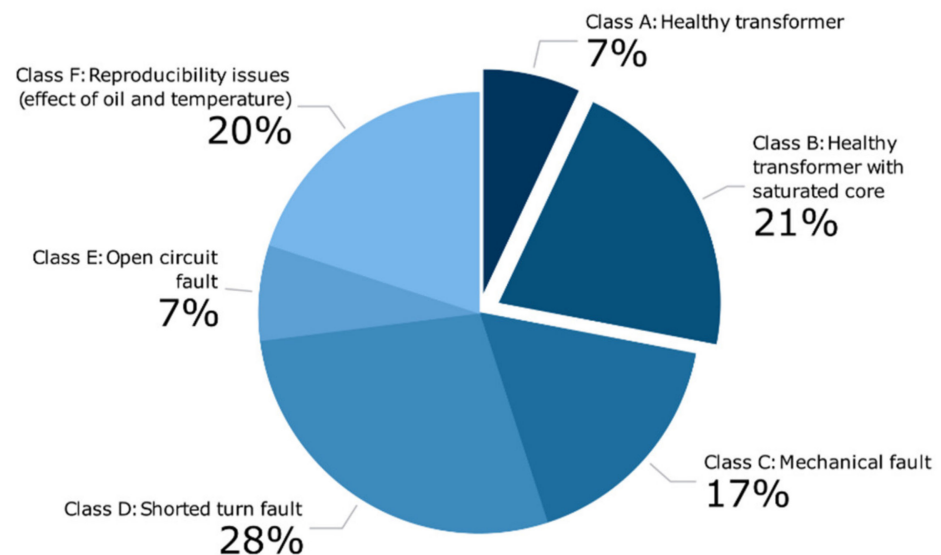


Figure 2. Breakdown of the entire database into six transformer conditions.

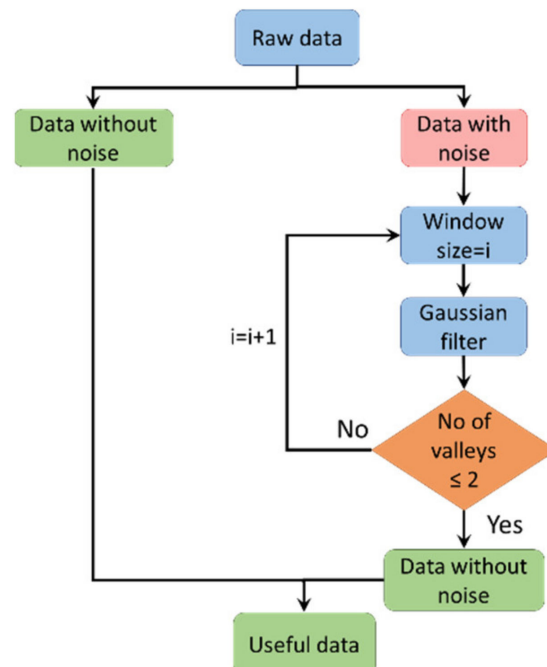
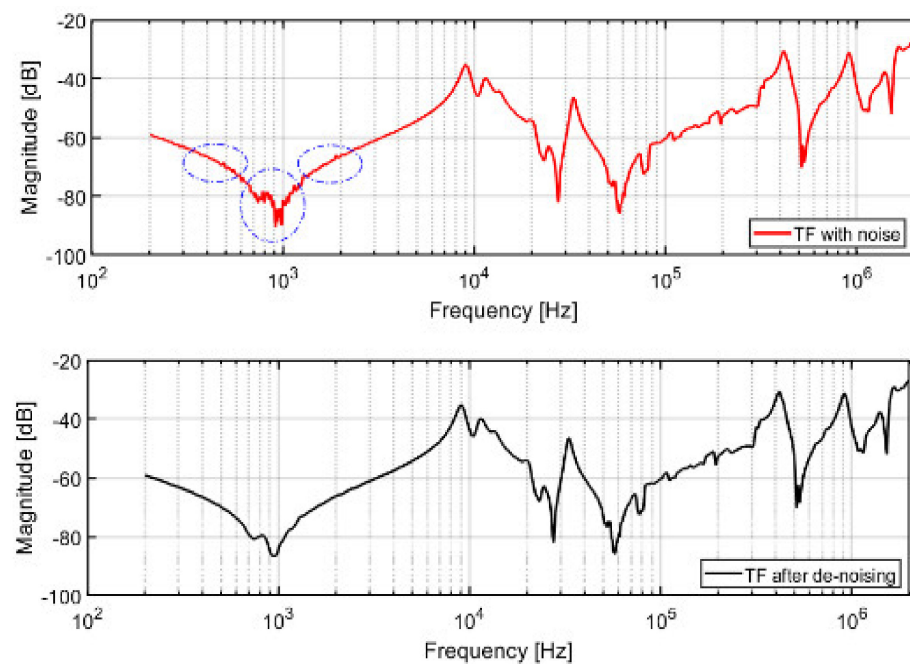


Figure 3. Workflow of data-denoising process.



**Figure 4.** Comparison of the transfer functions (TFs) before (up) and after noise removal (down).

### 3.3. Adaptive Frequency Division

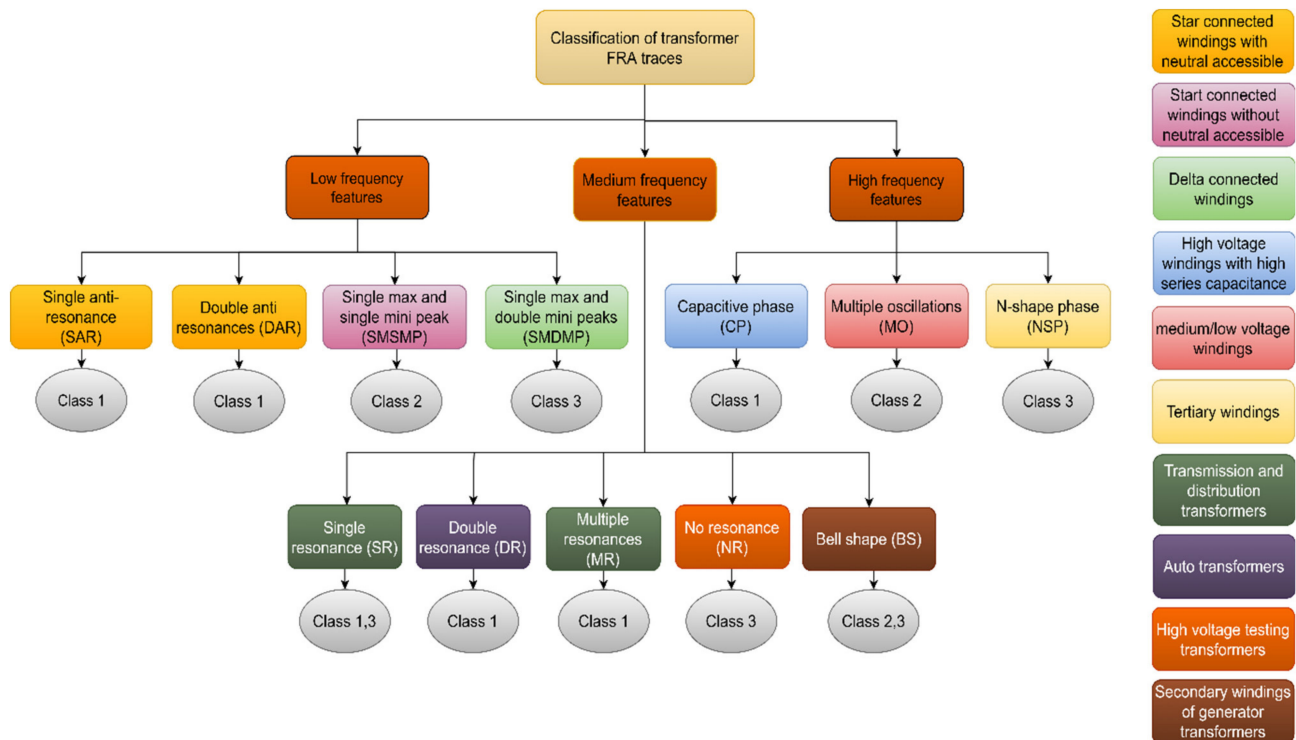
The frequency response of a transformer has a fundamental relationship with the physical parts of the transformer. These physical parts dominate the frequency response in different frequency regions. Hence, by identifying these regions, different faults in power transformers can be classified. However, the ranges of these frequency regions or sub-bands depend upon many factors such as rating, size, and core and winding structure, and a general range cannot be concluded [5,6].

In the literature, a few efforts have been made to divide the FRA spectrum into several sub-bands. Gonzales et al. [21] proposed a frequency-slicing algorithm based on phase zero-crossing points. However, this algorithm is not validated for FRA traces of different vector groups and windings. Hence, it fails to identify frequency limits for transformers having vector groups Dd0, Dyn0, and autotransformers. Similarly, identification of winding region in the frequency response of transformers having helical or ordinary disk-type winding structure is not feasible with this algorithm. Due to the low series capacitance of these windings, multiple resonances and anti-resonance appear in the high-frequency region; thus, the frequency region influenced by winding structure is hard to identify.

Lin et al. [22] proposed five frequency sub-bands structure using binary morphology. However, this method is based on deviation patterns between FRA pairs, and frequency ranges cannot be determined for healthy transformers. Likewise, this algorithm can misinterpret frequency regions for different faults on the same transformer because different faults lead to different deviation patterns. Velásquez et al. [23] also introduced an algorithm to divide the frequency spectrum into five sub-bands based on locations of poles and zeros. In this method, the low-frequency region is identified using the linear regression function. However, the characteristics of this linear regression function can vary for different frequency resolutions in the low-frequency region. Additionally, the identification of frequency ranges is not consistent for transformers having vector groups Dd0 and Dyn0. Hence, these automatic frequency division algorithms demand revision for proper fault classification. The adaptive frequency division algorithm presented in this work is based on different features appearing in FRA, which are characterized by the locations of resonances, anti-resonances points, and phase zero-crossings to ensure robustness for different FRA patterns.

The first approach for the development of an automatic frequency division algorithm is the identification and classification of different FRA patterns. In this work, the FRA

data from 80 power transformers of different sizes, ratings, and winding types are studied. Mainly, twelve features of FRA traces are recognized in different frequency sub-bands based on different transformer vector groups, winding structures, ratings of the winding, etc. The classification of FRA traces based on different features in different frequency sub-bands is illustrated in Figure 5.



**Figure 5.** Classification of transformer FRA traces based on different features in the low, medium, and high-frequency range.

Although there are at least 12 features of FRA traces, this does not imply that for identification of each feature independent algorithms have to be developed. Instead, after analyzing all possible features it was concluded that it is possible to group them into fewer classes, thus, the 12 features are further classified into three classes. The type of the transformer vector group and winding structure, etc., which are mainly responsible for the generation of different features, are also mentioned in the vertical column of Figure 5. Class 1 consists of FRA patterns belonging to star-connected windings with accessible neutral, autotransformers, and high-voltage windings with high series capacitance disk windings. The FRA patterns of star-connected winding without neutral accessible transformers, medium- and low-voltage windings, and secondary windings of generator transformers belong to class 2, whereas FRA patterns of delta connected primary and secondary windings, ordinary-disk or layer-type windings, and low-impedance windings are grouped in class 3.

Based on these features, an automatic frequency division algorithm is proposed that subdivides the entire frequency spectrum into four sub-bands, i.e., two low-frequency bands (LFB1 and LFB2), medium frequency (MFB), and high-frequency band (HFB). These frequency sub-bands are linked to different physical components of the transformer. For example, two low-frequency sub-bands (LFB1 and LFB2) are related to the core where magnetizing inductance ( $L_m$ ) and equivalent network capacitance ( $C_{net}$ ) dominate the response. Medium frequency sub-band (MFB) is dictated by the mutual inductances between windings ( $M_u$ ) and inter-winding capacitances ( $C_{iw}$ ), while the high-frequency region is controlled by the winding structure where a group of resonances caused by

the winding inductance and series and ground capacitances. The algorithms for the identification of LFB1, LFB2, MFB, and HFB are described in the following sections.

### 3.3.1. Determination of LFB1 and LFB2

The first low-frequency band (LFB1) starts at the start frequency of the sweep. To find the end of LFB1 and LFB2, two arbitrary frequencies  $f_1$  and  $f_2$  are used. The flow chart for the algorithm is shown in Figure 6. The algorithm starts with the search of frequency of the peak ( $f_{PhLC}$ ) at which the first transition of phase from inductive to capacitive takes place. If the peak at this frequency is a minimum peak, its frequency is assigned to  $f_1$  ( $f_1 = f_{PhLC}$ ). In this case, if the number of maximum peaks below  $f_1$  is less than 2, the algorithm corresponds to class 1, if not, it adapts to class 2.

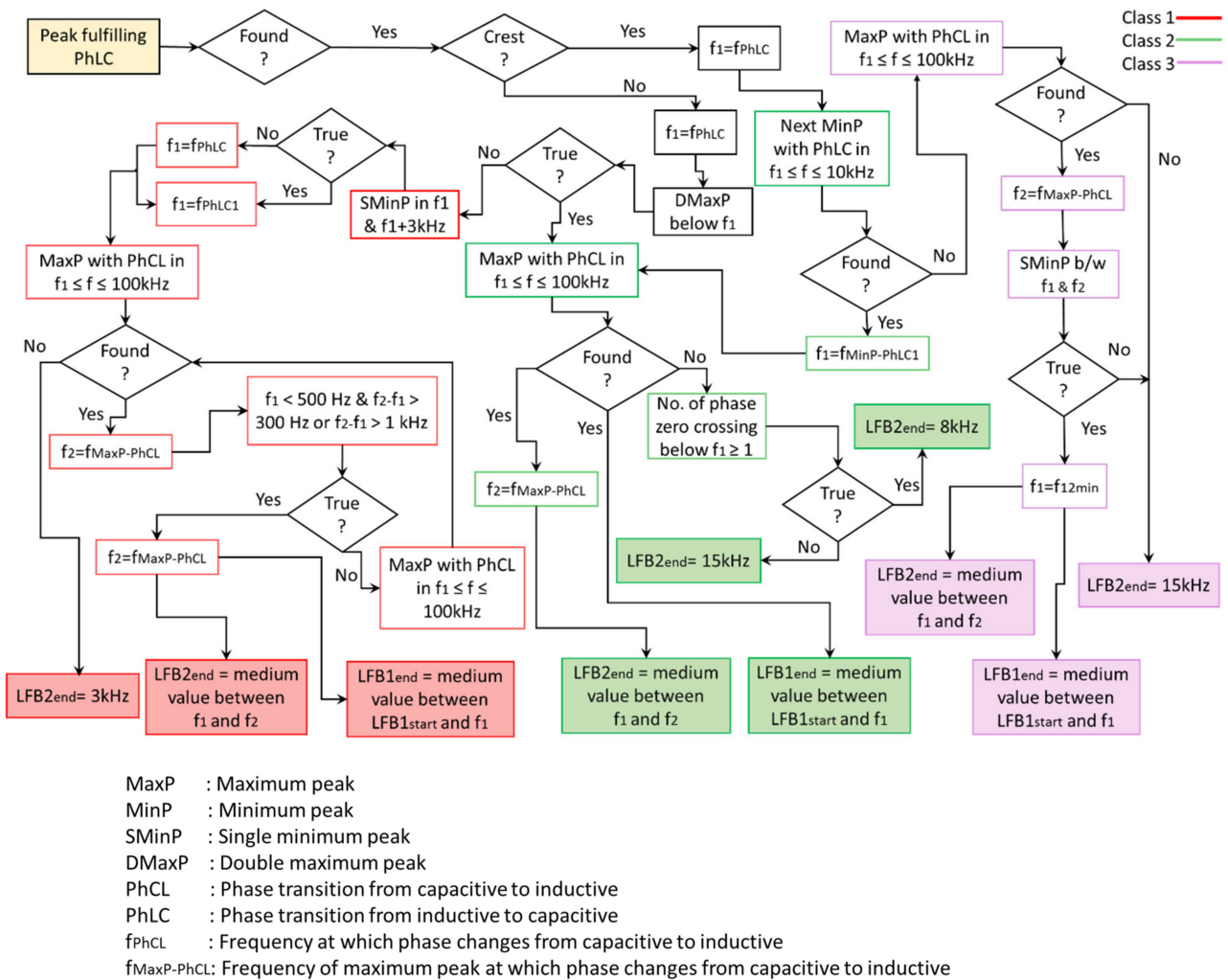


Figure 6. Workflow of the algorithms for identification of the low-frequency sub-bands.

The performance of this frequency division algorithm for class 1 is illustrated in Figure 7. In order to check if there is a single (for middle phase) or double minimum peak (for lateral phases) in the FRA trace, minimum peaks are searched between  $f_1$  and  $f_1 + 3$  kHz. If a minimum peak is found between  $f_1$  and  $f_1 + 3$  kHz, the frequency of this minimum peak is assigned to  $f_1$  ( $f_1 = f_{PhLC1}$ ), otherwise  $f_1$  remains unchanged. After fixing  $f_1$  for class 1, a maximum peak with phase transition from capacitive to inductive is searched in the range  $f_1$  and 100 kHz. If such a peak is found, the frequency of this crest is assigned to  $f_2$  ( $f_2 = f_{MaxP-PhCL}$ ). After determining  $f_2$ , if  $f_1$  is less than 500 Hz and the difference of  $f_2$  and  $f_1$  is greater than 300 Hz or the difference of  $f_2$  and  $f_1$  is greater

than 1 kHz, the algorithm for class 1 is complete. Hence, the  $LFB1_{end}$  is determined as a medium frequency between  $LFB1_{start}$  and  $f_1$ , whereas  $LFB2_{end}$  is determined as a medium frequency between  $f_1$  and  $f_2$ . Otherwise, a new maximum peak with phase transition from capacitive to inductive is searched in the range  $f_1$  and 100 kHz and the process repeats. If for any reason  $f_2$  cannot be determined,  $LFB2_{end}$  is set to 3 kHz.

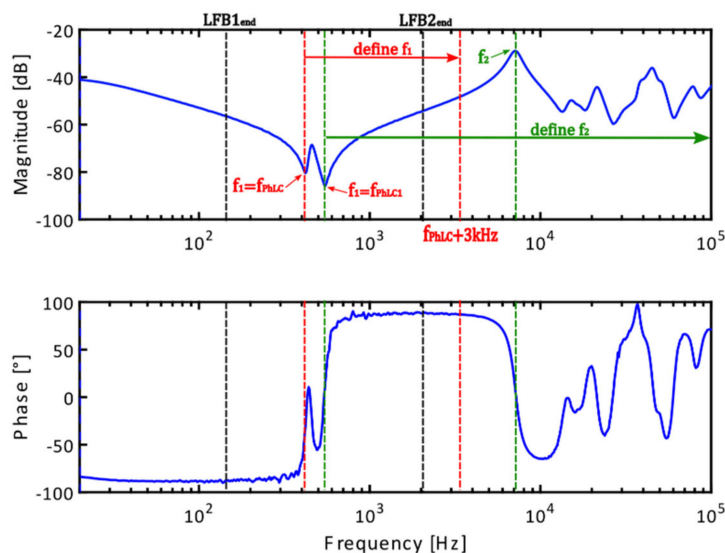


Figure 7. An example of class 1 FRA plot for identification of the  $LFB1_{end}$  and  $LFB2_{end}$ .

The algorithm corresponds to class 2 if  $fPhLC$  is a minimum peak and there are two maximum peaks below  $f_1$ . Another scenario for class 2 is that  $fPhLC$  is a crest. In this case, the algorithm searches for the next minimum peak in the range  $f_1$  and 10 kHz, for which phase crosses zero line from inductive to capacitive. If such a peak is found,  $f_1$  becomes the frequency of this point. After finding  $f_1$ , a maximum peak with phase transition from capacitive to inductive is searched between  $f_1$  and 100 kHz as illustrated in Figure 8. If this crest is found, the  $LFB1_{end}$  is determined as a medium frequency between  $LFB1_{start}$  and  $f_1$ , whereas  $LFB2_{end}$  is determined as a medium frequency between  $f_1$  and  $f_2$ . If  $f_2$  is not found, the algorithm checks the number of phase zero-crossings below  $f_1$ .  $LFB2_{end}$  is fixed at 8 kHz if phase zero-crossings below  $f_1$  are greater than or equal to 1; otherwise,  $LFB2_{end}$  becomes 15 kHz.

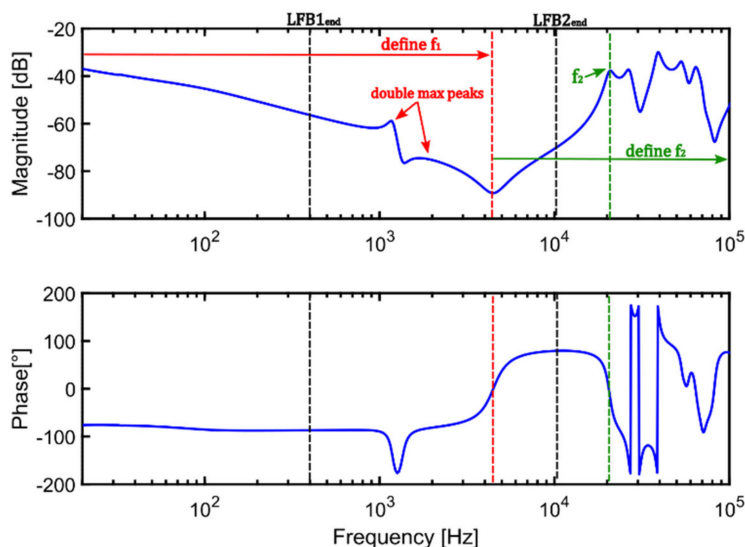
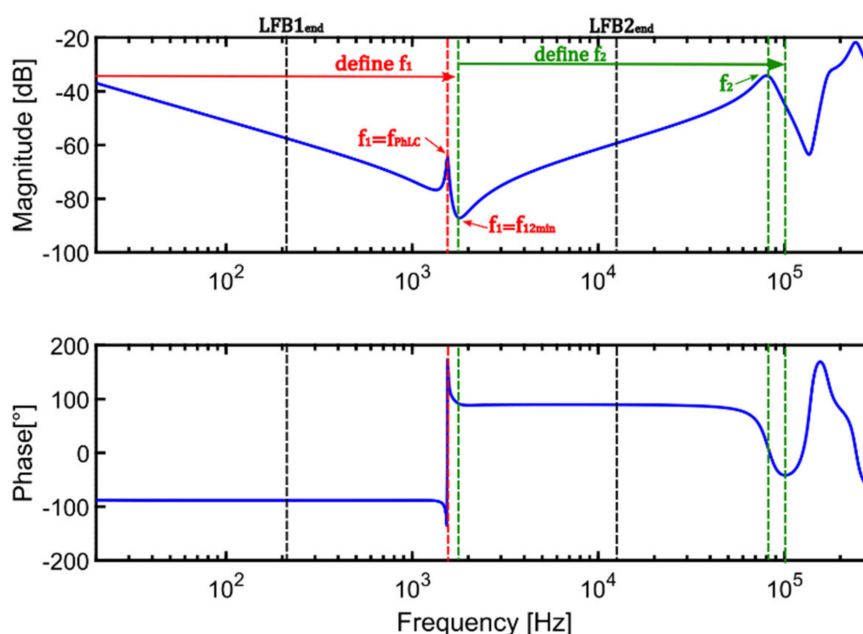


Figure 8. An example of class 2 FRA plot for identification of the  $LFB1_{end}$  and  $LFB2_{end}$ .



The algorithm adapts to class 3 if  $f_{PhLC}$  is a maximum peak and no minimum peak with phase transition from capacitive to inductive exist in range  $f_1$  to 10 kHz, as appreciated in Figure 9. In this case, the algorithm searches for a maximum peak with phase zero-crossing from positive to negative in range  $f_1$  and 100 kHz. If this peak is found,  $f_2$  becomes the frequency of this peak ( $f_2 = f_{maxP-PhCL}$ ). After finding  $f_1$  and  $f_2$ , the algorithm checks if there is a single minimum peak between  $f_1$  and  $f_2$ . If such a peak exists, the frequency of this minimum peak is assigned to  $f_1$  ( $f_1 = f_{12min}$ ), otherwise  $f_1$  remains unchanged ( $f_1 = f_{PhLC}$ ). Afterward, the  $LFB1_{end}$  is determined as a medium frequency between  $LFB1_{start}$  and  $f_1$ , whereas  $LFB2_{end}$  is determined as a medium frequency between  $f_1$  and  $f_2$ . If no such peaks are found,  $LFB2_{end}$  is set to 15 kHz.

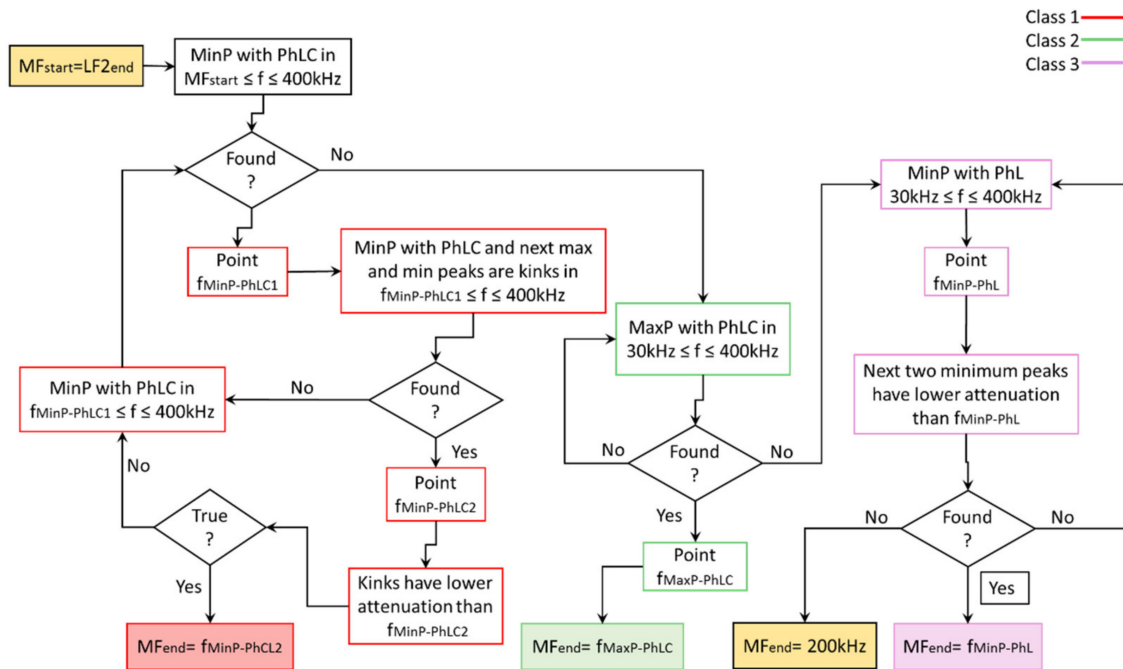


**Figure 9.** An example of class 3 FRA plot for identification of the  $LFB1_{end}$  and  $LFB2_{end}$ .

### 3.3.2. Determination of MFB

For the beginning of the medium frequency sub-band (MFB),  $MFB_{start}$  is set to the end of  $LFB2_{end}$ . The algorithm is adapted to three classes similar to LFB1 and LFB2. The workflow of the algorithms for the identification of the medium frequency sub-band ( $MFB_{end}$ ) is demonstrated in Figure 10. At the start, a minimum peak with phase zero-crossing from inductive to capacitive is searched between  $MFB_{start}$  and 400 kHz. If such a peak is found, the frequency of this peak is noted as  $f_{MinP-PhLC1}$  and the algorithm corresponds to class 1, as illustrated in Figure 11.

Next, another minimum peak with phase transition from inductive to capacitive is searched between  $f_{MinP-PhLC1}$  and 400 kHz, whose next maximum and minimum are kinks. If this peak is found, the frequency of this peak is assigned to  $f_{MinP-PhLC2}$ . After finding this point, it is confirmed that the  $MaxK$  and  $MinK$  have lower attenuation than  $MinP-PhLC2$ . If this condition is true, such that  $MagMaxK > MagMinP-PhLC2$  and  $MagMinK > MagMinP-PhLC2$ ,  $MFB_{end}$  becomes  $f_{MinP-PhLC2}$ . Otherwise, a new peak will be searched between  $f_{MinP-PhLC1}$  and 400 kHz, and the whole process repeats.



MaxK :Maximum peak which is a kink

MinK :Minimum peak which is a kink

Mag<sub>MinP</sub> :Magnitude of minimum peak

$f_{\text{MinP-PhLC}}$  :Frequency of minimum peak at which phase changes from inductive to capacitive

$f_{\text{MinP-PhLC1}}$  :Frequency of minimum peak at which 1st phase transition from inductive to capacitive occurs

$f_{\text{MinP-PhLC2}}$  :Frequency of minimum peak at which 2nd phase transition from inductive to capacitive occurs

Figure 10. Workflow of the algorithms for identification of the medium frequency sub-band ( $MFB_{\text{end}}$ ).

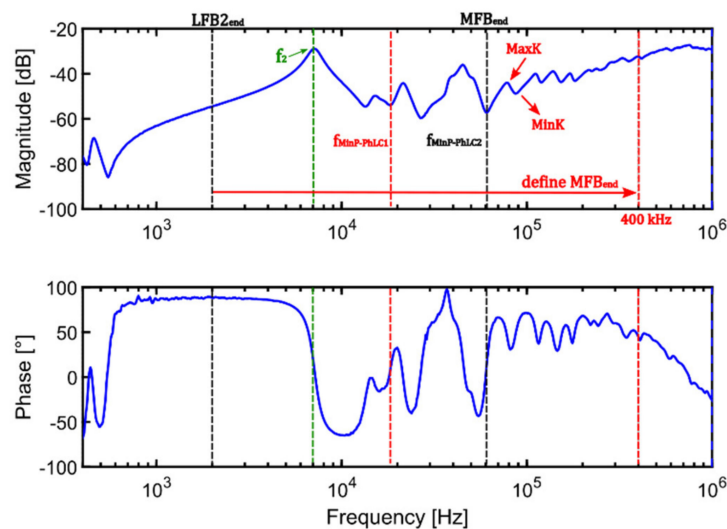


Figure 11. An example of class 1 FRA plot for identification of the  $MFB_{\text{end}}$ .

The algorithm adapts to class 2 if no minimum peak is found in the range  $MFB_{\text{start}}$  to 400 kHz for which phase changes from inductive to capacitive ( $\text{MinP-PhLC1}$ ). In this class, a maximum peak is searched between 30 kHz and 400 kHz, which fulfills the condition of phase zero-crossing from inductive to capacitive, as shown in Figure 12. If such a peak is found,  $MFB_{\text{end}}$  becomes the frequency of this point,  $MFB_{\text{end}} = f_{\text{MaxP-PhLC}}$ ; otherwise, a new maximum peak is searched between 30 kHz and 400 kHz, and process repeats.

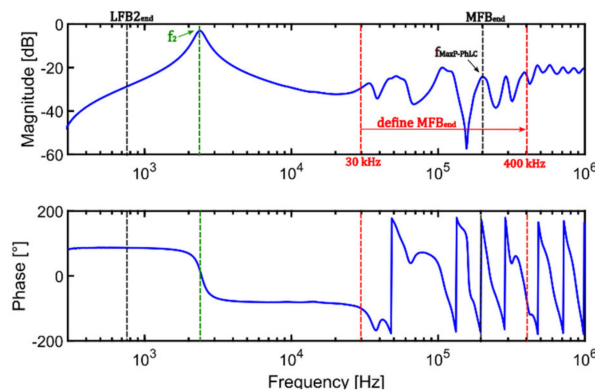


Figure 12. An example of class 2 FRA plot for identification of the  $MFB_{end}$ .

If no maximum or minimum peak with phase zero-crossing from inductive to capacitive is found, the algorithm corresponding to class 3 is to be followed. For this class, the algorithm searches for a minimum peak in the range of 30 kHz to 400 kHz. The phase of this minimum peak must be inductive ( $f_{MinP-PhL}$ ), as shown in Figure 13. If such a peak is found,  $MFB_{end}$  becomes the frequency of this point. If none peak fulfilling these conditions is found,  $MFB_{end}$  is fixed at 200 kHz.

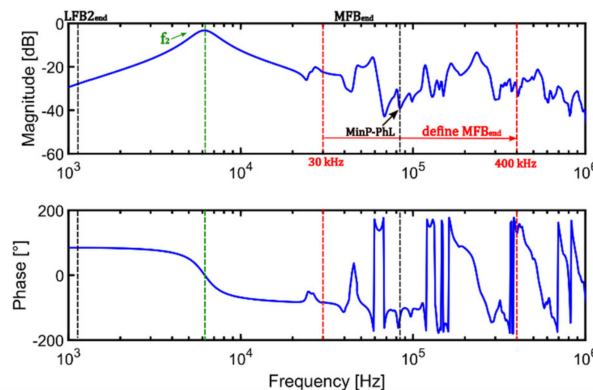


Figure 13. An example of class 3 FRA plot for identification of the  $MFB_{end}$ .

### 3.3.3. Determination of HFB

The HFB starts at the end of MFB. The determination of  $HFB_{end}$  depends upon the rating of the transformer winding. According to Ref [5], the FRA interpretation range is 2 MHz for windings with voltage less than 72.5 kV, whereas it is 1 MHz for windings with voltage greater than 72.5 kV. Hence,  $HFB_{end}$  is user-defined and depends upon the voltage rating of the winding. The decision algorithm for the identification of the high-frequency sub-band ( $HFB_{end}$ ) is shown in Figure 14.

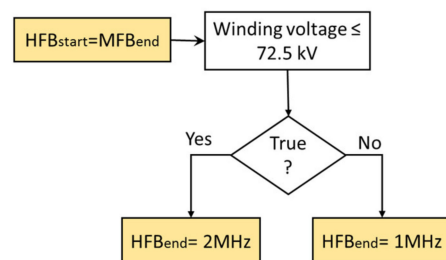


Figure 14. Decision workflow of the algorithm for identification of the high-frequency sub-band ( $HFB_{end}$ ).

### 3.4. Feature Generation

After dividing the FRA plot into four frequency sub-bands, five numerical indices (CCF, LCC, SD, CSD, and SE) are employed as given in Equations (1)–(5). These numerical indices are calculated in four sub-bands using reference and current TFs as shown in Figure 15. These indices quantify the deviation patterns for different fault types. This process is called feature generation. Each index gives four features from the FRA magnitude plot that serve as features for the classification models.

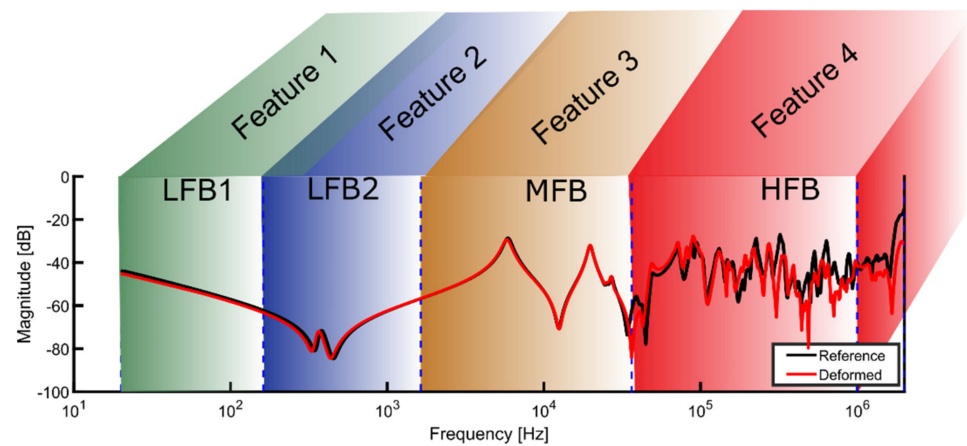
$$LCC = \frac{2S_{XY}}{(\bar{Y} - \bar{X})^2 + S_Y^2 + S_X^2} \quad (1)$$

$$CCF = \frac{\sum_{i=1}^N (X(i) - \bar{X})(Y(i) - \bar{Y})}{\sqrt{\sum_{i=1}^N [X(i) - \bar{X}]^2 \sum_{i=1}^N [Y(i) - \bar{Y}]^2}} \quad (2)$$

$$SD = \sqrt{\frac{\sum_{i=1}^N (Y(i) - X(i))^2}{N - 1}} \quad (3)$$

$$CSD = \sqrt{\frac{\sum_{i=1}^N [(Y(i) - \bar{Y}) - (X(i) - \bar{X})]^2}{N - 1}} \quad (4)$$

$$SE = \frac{\sum_{i=1}^N (Y(i) - X(i))}{N} \quad (5)$$



**Figure 15.** An example of feature generation from four frequency sub-bands of magnitude plot of FRA.

## 4. Training of ANN

### 4.1. Structure of ANN

To select the optimum structure of ANN, a sensitivity study is performed. For this purpose, overall accuracy is compared using a different number of hidden layers and five commonly used activation functions are tested in these hidden layers. Each hidden layer consists of eight neurons. Increasing the number of neurons in the hidden layer increases the power of the network but requires more computation and is more likely to produce overfitting. The neurons of the hidden layer are chosen through trial and error. Hidden neurons are increased if the training performance is poor, while they are decreased if training performance shows overfitting of the data sets. In this work, the hidden layer with eight neurons gives a better training performance. The results of the sensitivity study are presented in Table 1. Here, the overall accuracy of SE feature set is compared. It is clear that tansig and logsig outperformed the other activation functions. The logsig in

five hidden layers and tansig in one hidden layer show the highest overall accuracy. It is important to note that accuracy has minimal impact on increasing the number of hidden layers. However, the computation cost and complexity increase with the increase of hidden layers. This may lead to overfitting. To avoid these issues, a simple one hidden layer feedforward network with a tansig activation function for the hidden layer and a linear activation function for the output layer is used in this work, as shown in Figure 16. The figure demonstrates an associative ANN that consists of an input layer with ‘ $r$ ’ nodes, a hidden layer with ‘ $m$ ’ nodes, and an output layer with ‘ $n$ ’ nodes. The neurons of the input layer receive four features as inputs from different indicators ( $139 \times 4$  matrix). The output vector is a column matrix with six rows that correspond to six classes ( $6 \times 1$  matrix). The signal from the input layer is transmitted to the output layer through the hidden layer. The relationship between input vector  $F_i$ , hidden layer vector  $h_j$ , and the output vector  $y_k$  can be represented by Equations (6) and (7):

$$h_j = f\left(\sum_{i=1}^r W_{ij}F_i + b_j\right) \quad (6)$$

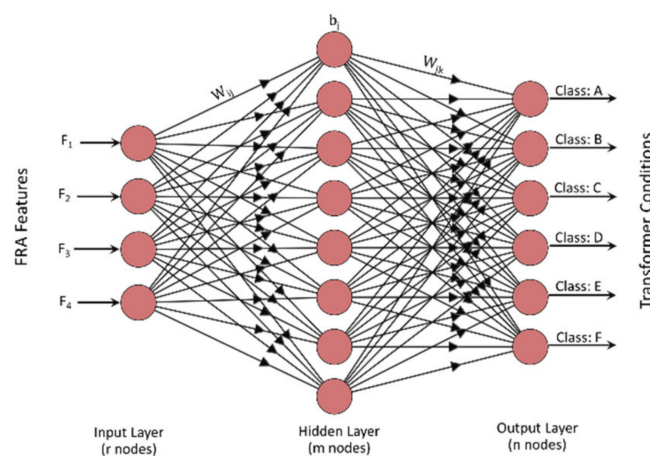
$$y_k = f\left(\sum_{j=1}^m W_{jk}h_j + b_k\right) \quad (7)$$

$$i = 1, 2, \dots, r, j = 1, 2, \dots, m, k = 1, 2, \dots, n$$

where  $F_i$  is the input vector,  $W_{ij}$  is the weight vector or connection strength between the input and hidden layer neurons,  $b_j$  is the bias of the hidden layer neuron,  $W_{jk}$  is the weight vector or connection strength of hidden and output layer neurons, and  $b_k$  is the bias of output layer neuron.

**Table 1.** Percentage overall accuracy for different activation functions and number of hidden layers.

Activation Function	Abbreviation	Number of Hidden Layers				
		1	2	3	4	5
Linear	purelin	60.1	60.9	60.9	60.9	60.9
Triangular basis	tribas	92	94.2	94.2	95.7	94.8
Radial basis	radbas	94.2	96.4	94.9	91.3	94.9
Hyperbolic tangent sigmoid	tansig	97.1	96.6	97.1	96.6	96.6
Log-sigmoid	logsig	95.3	95.8	96.3	94.8	97.6



**Figure 16.** Three-layer structure of feedforward backpropagation ANN.

## 4.2. Training Process

For a feed-forward network, the training algorithm Levenberg-Marquardt (LM) was applied in this work due to its fastest computing speed for a small network [16]. It updates the weights and biases in the direction in which the performance increases with a decreased gradient. The process for training the network has the following four steps.

### 4.2.1. Data Preprocessing

As the first step, the input data of the value for indices (CCF, LCC, SD, CSD, and SE) need to be normalized before training the network since the training process will be very slow if the input is very large. For instance, the sigmoid transfer function becomes essentially saturated when the net input is greater than three ( $\exp(-3) = 0.05$ ), thus, the gradients are very small which slow down the training process. The default function for preprocessing of the feed-forward network (mapminmax) was applied to normalize the data in the range  $[-1, 1]$ .

### 4.2.2. Data Division

Before training the network, the data should be divided into three data sets: training data set, validation data set, and test dataset. The training set is used for computing the gradient and updating the network weights and biases, while the validation set is used to validate that the network is generalizing and to stop training before overfitting. The test set is not used during training, but it is useful to test the generalization of the network. In this work, the data sets were divided into a training set of 70%, validation set of 10%, and test set of 20%, respectively. It is worth noting that the dividing ratio of each sub-set needs to be identical for every class due to a small amount of data set.

### 4.2.3. Training of Network

The process of training a neural network involves tuning the values of the weights and biases of the network to optimize network performance, as defined by the network performance function ( $F$ ). For training multilayer feedforward networks, the optimization algorithms (the gradient and the Jacobian) are calculated using a technique called the backpropagation algorithm, which involves performing computations backward through the network.

$$F = mse = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - a_i)^2 \quad (8)$$

There are two different ways in which training can be implemented: incremental mode and batch mode. For most problems, batch training is significantly faster and produces smaller errors than incremental training; thus, batch training mode is used in this work. In batch mode, all the inputs in the training set are applied to the network before the weights are updated. A basic structure of the ANN needs to be firstly chosen at the beginning of the training, which should be changed by analyzing the performance of the network. Then, some parameters of the network (such as learning rate, epochs, and cost function) are set to a fixed value to stop the training, after the weights and biases were initialized. The training of the network begins at epoch = 1. At each epoch, the weight adjustments are applied whose size can be determined by the learning rate, until the network is converged. If some conditions are fulfilled, the training of the network can stop. In this way, there are generally two methods to stop the training: early stopping and reaching the best performance. The condition (reaching the maximum error goal) can be used to early stop the training if the validation error increases for a predefined value of this parameter, but the weights and biases at the minimum of the validation error and minimum gradient are returned. If the maximum validation error is not reached, then the mean square error (mse) is able to validate whether the feedforward network has the best performance. If the best goal is reached, then the training will be stopped.

### 5. Analysis of Training and Validation Performance of ANN

The performance of training and validation is analyzed by several variables during the course of training, such as the value of the performance function and the magnitude of the gradient. The performance function used in this work is mean square error (MSE), and its threshold value is set to 0.01. After reaching a minimum if the MSE does not decrease for the next six consecutive iterations (epochs), the training will stop. Hence, the number of validation checks is set to 6.

#### 5.1. Performance of ANN with CSD Feature Set

The confusion matrices of ANN trained with CSD feature set are presented in Figure 17. The diagonal cells show the number of cases that are correctly classified, and the off-diagonal cells show the misclassified cases. The blue cell in the bottom right shows the total percentage of correctly classified cases (in green) and the total percentage of misclassified cases (in red). The training confusion matrix shows the training performance of the ANN, while validation and test confusion matrices show the performance of ANN with unseen cases, whereas all confusion matrices show the overall performance. It can be seen that the general training and validation performance is excellent, as 100% of the cases are correctly classified. It can be appreciated that ANN shows very good test performance, as 96.3% of the cases are correctly classified and only one case from class B is misclassified as class D. Hence, ANN has correctly classified cases with an overall accuracy of 99.3%.

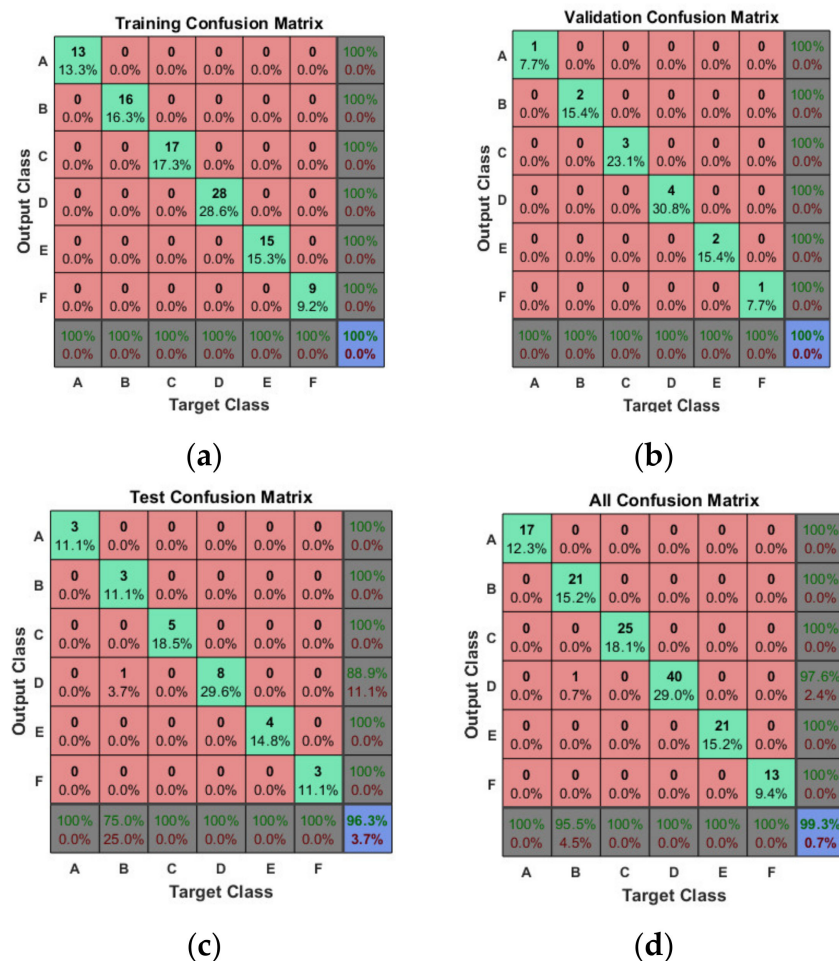
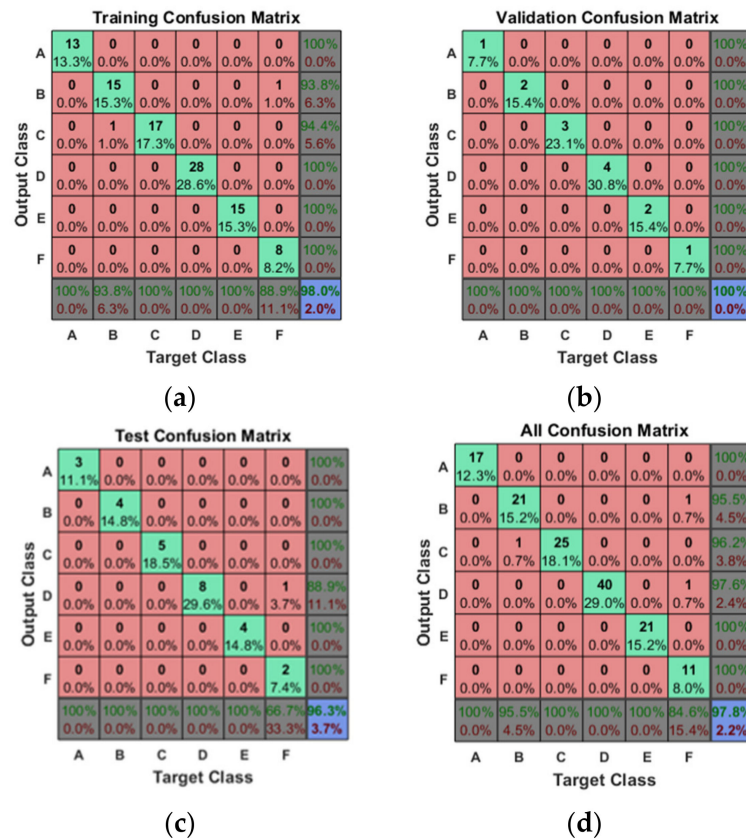


Figure 17. The confusion matrices of ANN trained with CSD feature set: (a) training confusion matrix, (b) validation confusion matrix, (c) testing confusion matrix, and (d) overall confusion matrix.

### 5.2. Performance of ANN with SD Feature Set

The confusion matrices for the SD feature set are shown in Figure 18. The training performance of ANN is very good since 98% of the cases are correctly classified. However, the training performance of the class B and F is slightly low, as one case of class B is misclassified as class C, while one case of class F is misclassified as class B. The validation performance is excellent, as 100% of the cases are correctly identified. The test performance is very good since 96% of the cases are correctly classified. However, the classification performance of class F is poor, as one case of class F is misclassified as class D. In summary, the ANN has correctly classified faults with an overall accuracy of 97.8%.



**Figure 18.** The confusion matrices of ANN trained with SD feature set: (a) training confusion matrix, (b) validation confusion matrix, (c) testing confusion matrix, and (d) overall confusion matrix.

### 5.3. Performance of ANN with SE feature Set

The confusion matrices of ANN for the SE feature set are shown in Figure 19. The general training performance of ANN is very good, as 96.9% of the cases are correctly classified. However, the training performance of class C is fairly low (88%) since two cases are misclassified as class A. While the training performance of class F is also low (88.9%), one case of class F is misclassified as class B. The general validation performance is excellent since 100% of the cases are correctly identified. The general test performance of ANN is very good since 96% of the cases are correctly classified; only one case of class F is misclassified as class D. In summary, the ANN has correctly classified faults with an overall accuracy of 97.1% with SE feature set.

### 5.4. Performance of ANN with LCC Feature Set

The performance matrices of ANN when supplied with Lin's concordance coefficient (LCC) as feature set are shown in Figure 20. In training, the ANN shows a very good performance (96.9% accuracy) where it has correctly identified all cases from classes A, B, D, and E with 100% accuracy, whereas two cases from class C and one case from class F



are misclassified as class A and B, respectively. In validation, all classes are identified with 100% accuracy except class C, for which one case is misclassified as class A, reducing the overall validation performance to 92.3%. When supplied with test data, the ANN correctly classify 92.6% of cases. Only one case from each class B and F is wrongly identified as class D and B, respectively. Thus, the overall accuracy of ANN is 95.7% as shown in all confusion matrices. Classes A, D, and E are always correctly identified in training, validation, and testing. Most errors are seen for class C, for which three cases are misclassified followed by class F (2 misclassified cases) and class B (1 misclassified case).

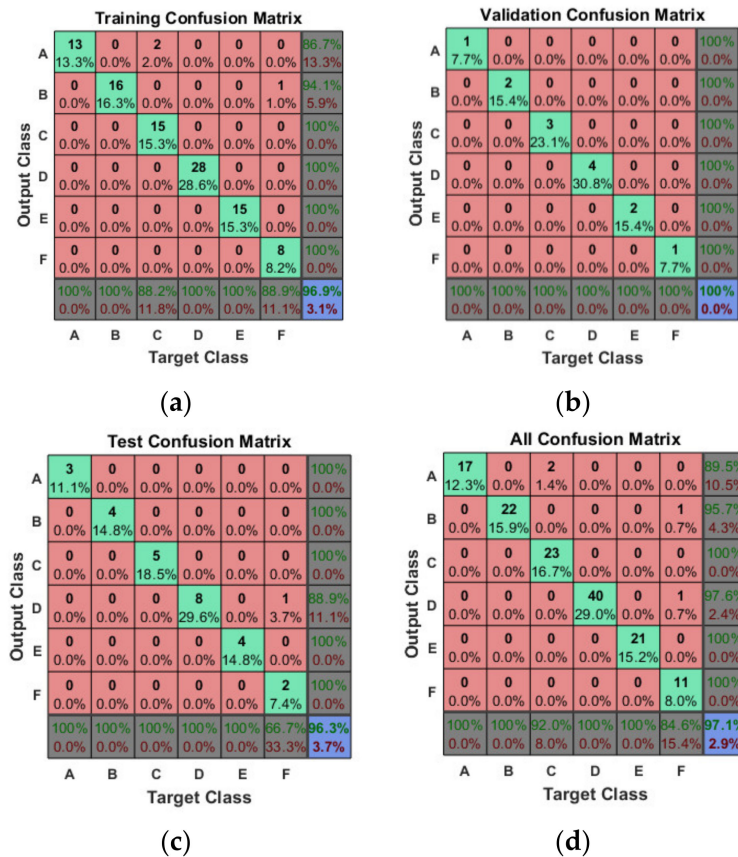


Figure 19. The confusion matrices of ANN trained with SE feature set: (a) training confusion matrix, (b) validation confusion matrix, (c) testing confusion matrix, and (d) overall confusion matrix.

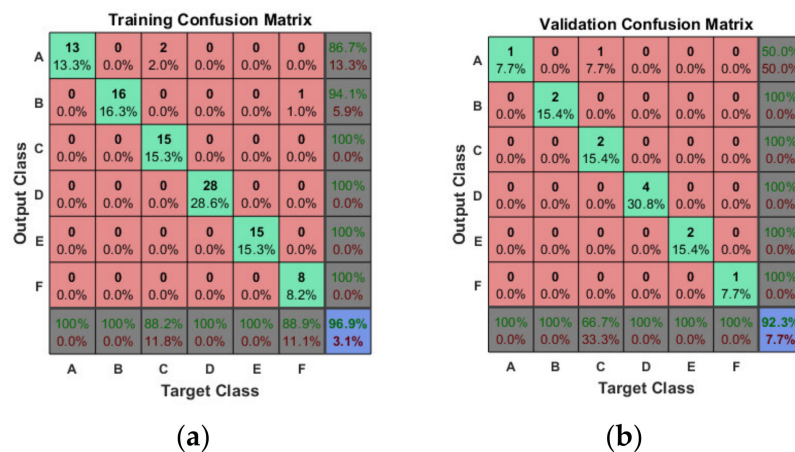
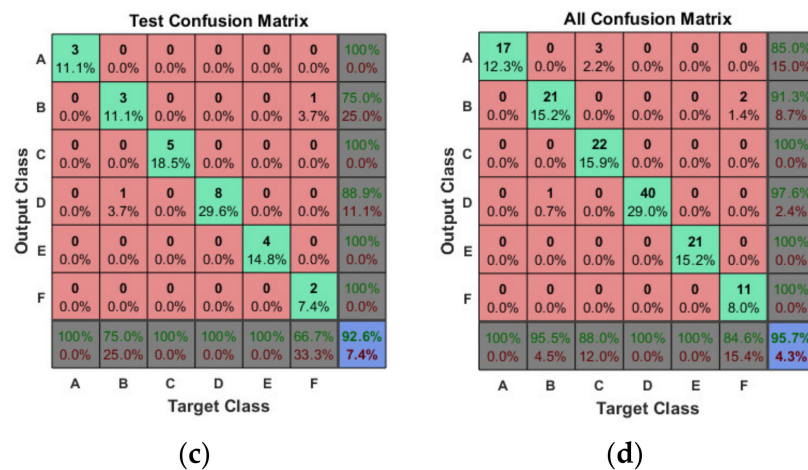


Figure 20. Cont.



**Figure 20.** The confusion matrices of ANN trained with LCC feature set: (a) training confusion matrix, (b) validation confusion matrix, (c) testing confusion matrix, and (d) overall confusion matrix.

### 5.5. Performance of ANN with CCF Feature Set

The classification performance of ANN with CCF features is shown in Figure 21. It can be seen that with CCF the general training accuracy of ANN is 92.3% as one case from class B and five cases from class C are misclassified as class A. Furthermore, one case from class F is classified as class B. In validation, the ANN shows very good performance where only one case from class C is misclassified as class F. It should be noted that the ANN has misclassified more cases with the CCF feature set than any other feature sets. The same trend can be observed in the test confusion matrix. When supplied with test data, the ANN gives an accuracy of 85.2%. In the test confusion matrix, two cases from class E are wrongly identified as class F, and two cases from class D and F (one case from each class) are misclassified as class D. Thus, the overall accuracy of ANN with CCF features is 91.3%. Only classes A and D are identified with 100% accuracy. In class B and E, 90% of cases are correctly classified, followed by class F, where 84.6% cases are correctly identified. For class C, many cases are misclassified which reduces the ANN accuracy to 76% for this class. The reason behind this is that the FRA traces of a slight mechanical change are similar to the FRA traces of a normal transformer.

Figure 22 shows the accuracy comparison of ANNs trained with different feature sets (CSD, SD, SE, LCC, and CCF). ANN trained with CSD feature set shows the best performance, as average overall accuracy is 99%. SD, SE, and LCC also show reasonable performance in classifying different faults since the average overall accuracy is above 95%. However, the performance of ANN trained with CCF is relatively low since only 90% of the cases are correctly classified. Moreover, it possesses the lowest accuracy for unseen data sets provided during the test. It should be noted that the ANN has misclassified more cases with the CCF feature set than any other feature sets. These results confirm the ability of different indices in detecting and classifying different transformer winding faults. In summary, provided with the feature set of CSD, all the classes are effectively learned by the network, indicating the best feature set for detection and classification of winding faults using FRA results.

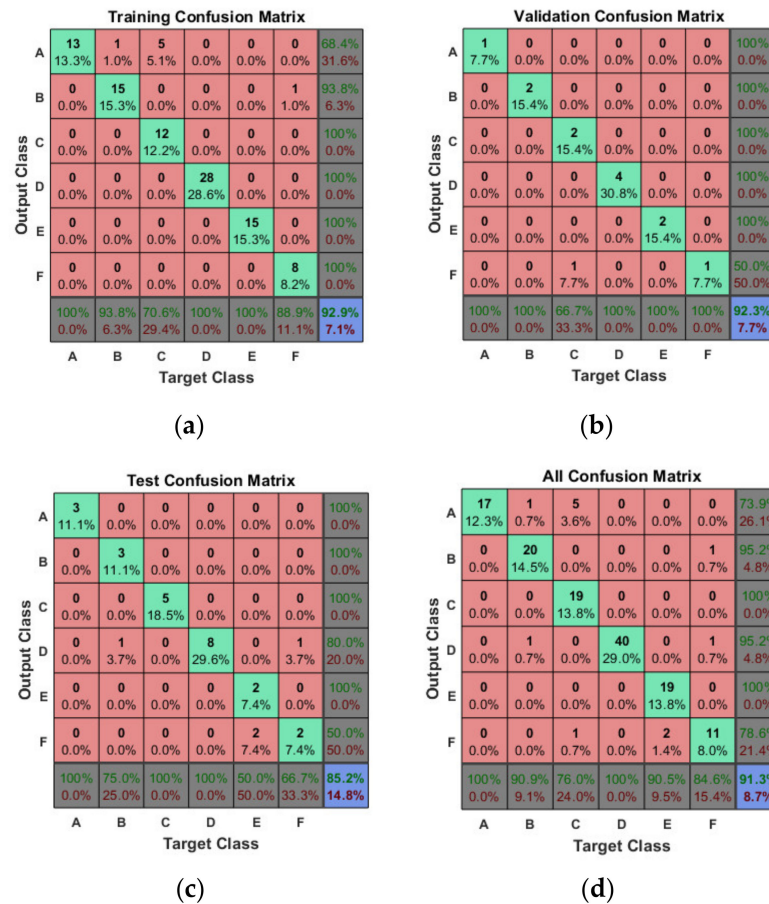


Figure 21. The confusion matrices of ANN trained with CCF feature set: (a) training confusion matrix, (b) validation confusion matrix, (c) testing confusion matrix, and (d) overall confusion matrix.

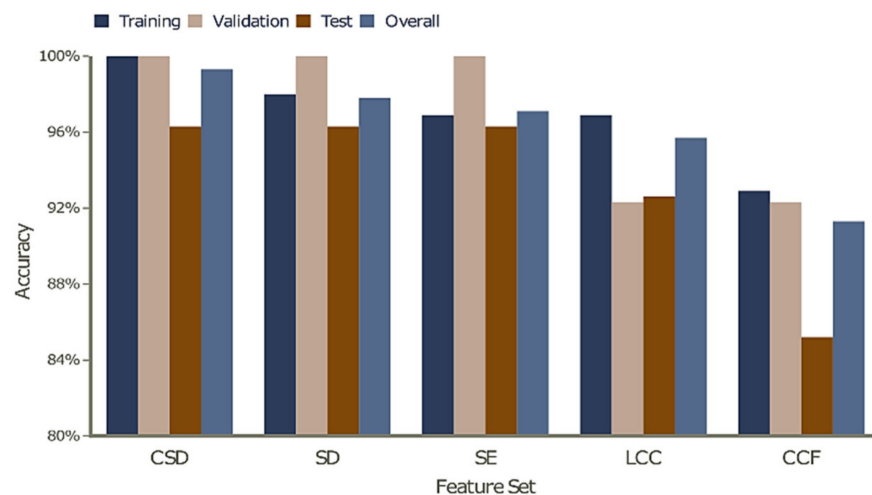


Figure 22. Performance comparison of ANN trained with different numerical indices (CSD, SD, SE, LCC, and CCF).

## 6. Case Studies

### 6.1. Case 1: Axial Collapse after Clamping Failure

In case 1, the unit is a three-phase, 240 MVA, 400/132 kV autotransformer. The unit was switched out of service for investigation after a Buchholz alarm. FRA measurements on common winding before and after the fault are shown in Figure 23 [24]. The visual analysis

of the FRA results shows some deviations and shifts of resonances in the high-frequency sub-bands, but these deviations yet need to be identified as normal or investigable? After strip-down, irreparable damage such as axial collapse and twisting to A-phase LV winding were found as shown in Figure 24. In order to verify the performance of the ANN in diagnosing this case, the case was further tested with the proposed ANNs. The ANNs trained with different feature sets (CSD, SD, CCF, LCC, and SE) were employed. The performance metrics are shown in Table 2. It can be seen that ANN has successfully diagnosed this case as a mechanical fault. Moreover, all the feature sets have diagnosed the same class.

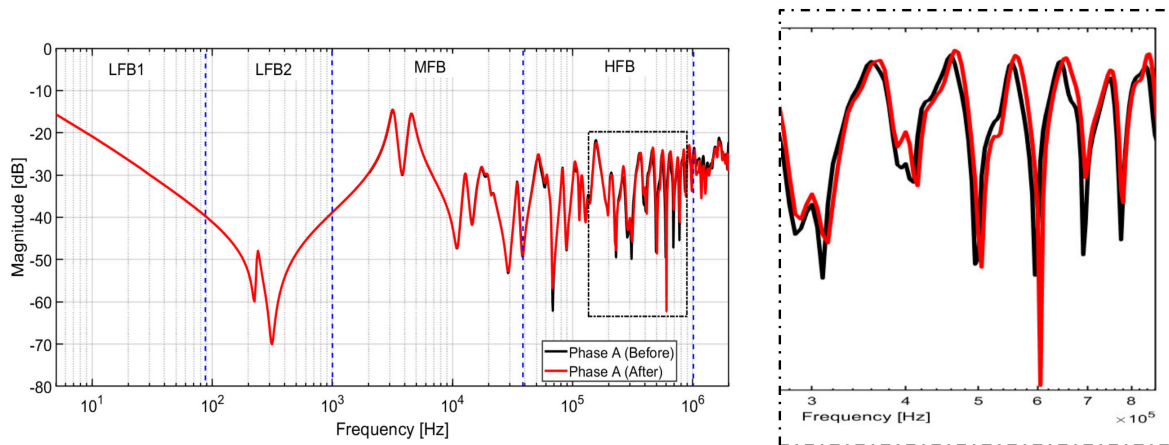


Figure 23. TF of LV winding before and after axial collapse [24].



Figure 24. Axial collapse of LV winding due to clamping failure [24].

Table 2. Performance matrices of ANN for different indicators in assessing case 1.

Indices	Actual Class	Predicted Class	Comment
SD	C	C	Pass
CSD	C	C	Pass
SE	C	C	Pass
CCF	C	C	Pass
LCC	C	C	Pass

SD-LFB1 = 0, SD-LFB2 = 0, SD-MFB = 0.3942, SD-HFB = 2.3219,  
 CSD-LFB1 = 0, CSD-LFB2 = 0, CSD-MFB = 0.3915, CSD-HFB = 2.3184,  
 SE-LFB1 = 0, SE-LFB2 = 0, SE-MFB = 0.1834, SE-HFB = 1.1839,  
 CCF-LFB1 = 1, CCF-LFB2 = 1, CCF-MFB = 0.9983, CCF-HFB = 0.9372,  
 LCC-LFB1 = 1, CC-LFB2 = 1, LCC-MFB = 0.998, LCC-HFB = 0.9370

### 6.2. Case 2: Shorted Turn Failure

In case 2, the unit is a three-phase, 60 MVA, 105/6.6/22 kV transformer. Figure 25 shows the HV open-circuit measurements. The response of the short-circuited winding is deviated in the low and medium frequency bands. After dismantling of the transformer, it was found phase-U turns shorted due to lightning surge. This case was further tested with the proposed ANNs. Again, the ANNs trained with different feature sets (CSD, SD, CCF, LCC, and SE) were employed for fault investigation. The performance metrics are shown in Table 3. It can be appreciated that ANN has diagnosed this case as shorted-turn failure (class D). Moreover, all the feature sets have diagnosed the same class.

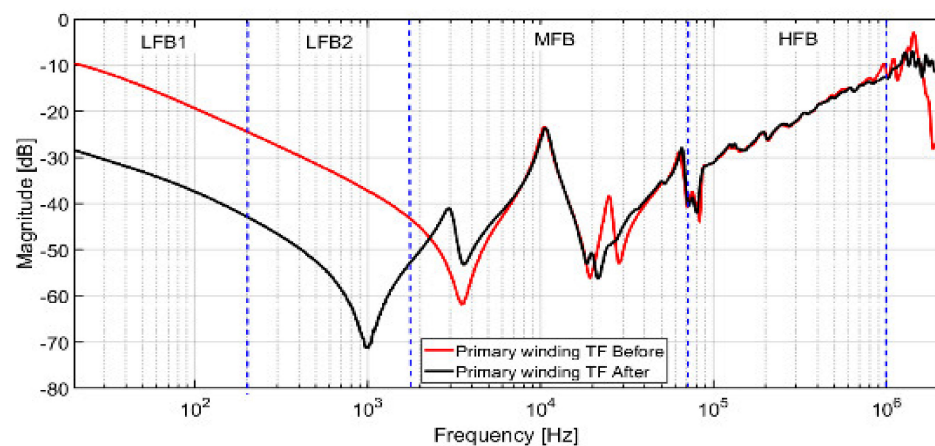


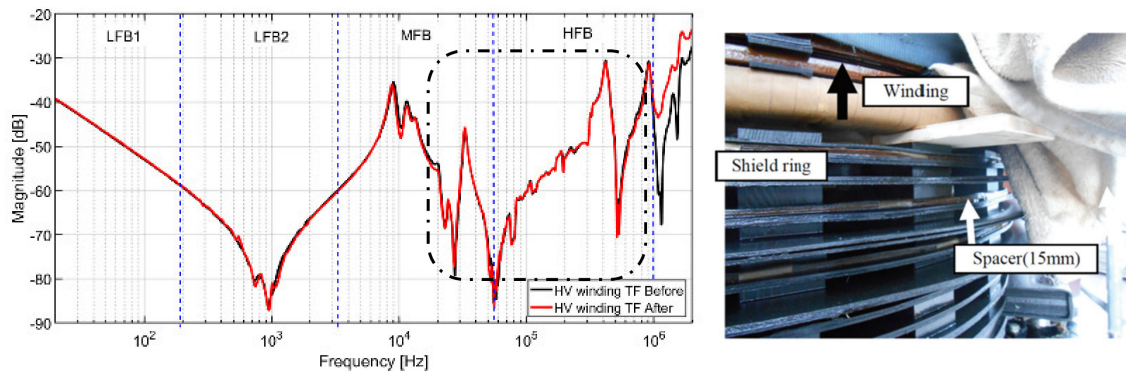
Figure 25. Frequency response of HV winding before and after the fault.

Table 3. Performance matrices of ANN for different indicators in assessing case 2.

Indices	Actual Class	Predicted Class	Comment
SD	D	D	Pass
CSD	D	D	Pass
SE	D	D	Pass
CCF	D	D	Pass
LCC	D	D	Pass
SD-LFB1 = 18.63, SD-LFB2 = 24.39, SD-MFB = 5.32, SD-HFB = 2.53 CSD-LFB1= 0.447, CSD-LFB2 = 5.160, CSD-MFB = 5.135, CSD-HFB = 2.535 SE-LFB1= 18.59, SE-LFB2 = 23.74, SE-MFB = 3.54, SE-HFB = 1.266 CCF-LFB1= 0.997, CCF-LFB2 = 0.55, CCF-MFB = 0.846, CCF-HFB = 0.972 LCC-LFB1= 0.170, CC-LFB2 = 0.039, LCC-MFB = 0.826, LCC-HFB = 0.972			
D: Shorted turn fault			

### 6.3. Case 3: Marginal Mechanical Faults (Axial Displacement) in HV Winding

In case 3, the FRA traces of a 27 MVA, 154 kV/10.5 kV transformer are employed. In this case, axial displacement fault is applied on U-phase HV winding as described in [25], where HV winding of U-phase is axially displaced up to 15 mm (0.6% of height) by inserting spacers as shown in Figure 26. The open-circuit FRA traces are measured before and after the fault are also shown in Figure 26. From the FRA comparison, it can be seen that marginal faults are difficult to interpret manually since the deviations between FRA curves are very small. However, with the application of the proposed ANNs, it is possible to detect such slight deviations between FRA curves. The performance metrics are shown in Table 4. It can be appreciated that ANN has diagnosed this case as a mechanical failure mode. However, ANN trained with CCF failed to detect the AD fault, indicating class A.



**Figure 26.** Frequency response of U-phase HV winding before and after the fault (left) and axial displacement of U-phase HV winding (right) [25].

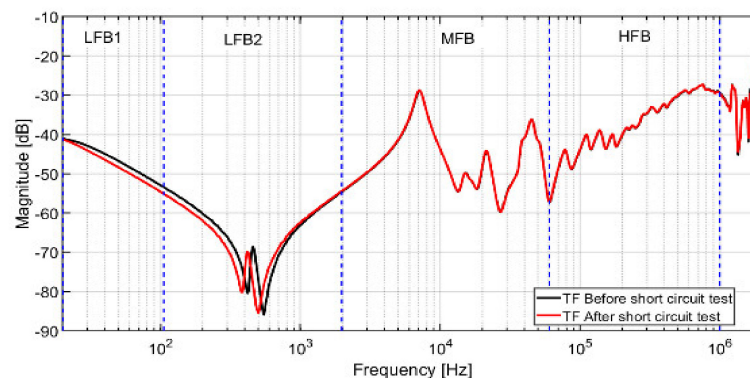
**Table 4.** Performance metrics of ANN for different indicators in assessing case 3.

Indices	Actual Class	Predicted Class	Comment
SD	C	C	Pass
CSD	C	C	Pass
SE	C	C	Pass
CCF	C	A	Fail
LCC	C	C	Pass

SD-LFB1 = 0.171, SD-LFB2 = 0.672, SD-MFB = 1.398, SD-HFB = 0.857  
 CSD-LFB1 = 0.0378, CSD-LFB2 = 0.66, CSD-MFB = 1.366, CSD-HFB = 0.799  
 SE-LFB1 = 0.161, SE-LFB2 = 0.458, SE-MFB = 0.960, SE-HFB = 0.5711  
 CCF-LFB1 = 1, CCF-LFB2 = 0.996, CCF-MFB = 0.992, CCF-HFB = 0.997  
 LCC-LFB1 = 0.999, CC-LFB2 = 0.996, LCC-MFB = 0.991, LCC-HFB = 0.996

6.4. Case 4: Healthy Transformer

In case 4, the unit is a three-phase, 47 MVA, 120/26.4 kV transformer. This case belongs to the class of normal transformers which have successfully passed the short-circuit test. The FRA traces are measured before and after the short-circuit test event. After inspection, no winding deformations were found in this transformer. The FRA curves of A-phase HV winding are presented in Figure 27. It can be seen that the FRA curves perfectly lie on each other, indicating no deviation except in the low-frequency regions (LFB1 and LFB2) where the deviations are due to a different core magnetization. The case is further tested with the proposed ANNs. The results are presented in Table 5. It can be appreciated that ANN has diagnosed this case as healthy transformer with core saturation. Moreover, all the feature sets have diagnosed the same class.



**Figure 27.** Frequency response of A-phase HV winding before and after the short circuit test [24].

**Table 5.** Performance metrics of ANN for different indicators in assessing case 4.

Indices	Actual Class	Predicted Class	Comment
SD	B	B	Pass
CSD	B	B	Pass
SE	B	B	Pass
CCF	B	B	Pass
LCC	B	B	Pass
SD-LFB1 = 1.419, SD-LFB2 = 3.431, SD-MFB = 0.164, SD-HFB = 0.149 CSD-LFB1 = 0.401, CSD-LFB2 = 3.36, CSD-MFB = 0.150, CSD-HFB = 0.117 SE-LFB1 = 1.357, SE-LFB2 = 2.437, SE-MFB = 0.136, SE-HFB = 0.136 CCF-LFB1 = 0.998, CCF-LFB2 = 0.909, CCF-MFB = 0.999, CCF-HFB = 0.999 LCC-LFB1 = 0.938, LCC-LFB2 = 0.905, LCC-MFB = 0.999, LCC-HFB = 0.999			

## 7. Conclusions

In this work, an intelligent fault detection algorithm (IFDA) was proposed for automatic condition assessment of transformer windings. The algorithm was based on a multilayer, feedforward, backpropagation artificial neural network (ANN). Mainly, six conditions of transformers were identified, namely, A: healthy winding, B: healthy winding with saturated core, C: mechanically deformed winding, D: short-circuited winding, E: open-circuited winding, and F: healthy winding tested with different oil and temperature. For classification and feature generation, an adaptive frequency slicing algorithm was developed that could satisfactorily indicate low-, medium-, and high-frequency sub-bands in FRA signatures of different transformers. The implementation and test of the adaptive frequency slicing algorithm with real case studies confirmed their validity. The data from different sub-bands were transformed into features by using five main statistical indicators: CSD, SD, CCF, LCC, and SE. The classification abilities of all five indicators were studied by performance metrics of ANNs trained with different indicators. It was found that ANN trained with CSD feature sets showed the highest average accuracy since 99% of the cases were correctly classified in the given database, while ANNs trained with CSD, SD, and LCC showed acceptable performance since they could classify up to 95% of the unseen data sets. However, the performance of ANN trained with CCF was relatively low, as only 90% of the cases were correctly classified. Moreover, it possessed the lowest accuracy for unseen data sets provided during the test.

The performance of the algorithm was further illustrated from selected case studies of real power transformers. Four case studies were selected from the database. Case studies included cases with a variety of winding faults and one case where the transformer had no-fault. Results showed that the proposed algorithm can detect and classify the variety of windings faults. Results obtained provide evidence that the proposed machine learning algorithm can precisely assess the transformer winding condition and identify the fault type with good accuracy without much human intervention and thus solve the major challenge of the FRA for industrial application, namely, the reliable automatic assessment of FRA results.

**Author Contributions:** M.T. developed the intelligent fault detection algorithm (IFDA), conceptualized the adaptive frequency division methods, performed the classification task, and drafted the manuscript. S.T. acquired the funding, supervised the research, and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—380135324.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank Satoru Miyazaki from Central Research Institute of Electric Power Industry (CRIEPI) Japan and Mohammad Hamed Samimi from University of Tehran, Iran for their contribution in the database. The authors also appreciate the discussions within Cigré Working Group A2.53 “Advances in the interpretation of transformer Frequency Response Analysis (FRA)”.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. CIGRE WG A2.37. *Transformer Reliability Survey*; CIGRE Brochure 642; CIGRE: Paris, France, 2015; ISBN 978-2-85873-346-0.
2. Secue, J.R.; Mombello, E. Sweep frequency response analysis (SFRA) for the assessment of winding displacements and deformation in power transformers. *Electr. Power Syst. Res.* **2008**, *78*, 1119–1128. [\[CrossRef\]](#)
3. Tenbohlen, S.; Coenen, S.; Djamali, M.; Müller, A.; Samimi, M.; Siegel, M. Diagnostic Measurements for Power Transformers. *Energies* **2016**, *9*, 347. [\[CrossRef\]](#)
4. Rahimpour, E.; Christian, J.; Feser, K.; Mohseni, H. Transfer function method to diagnose axial displacement and radial deformation of transformer windings. *IEEE Trans. Power Deliv.* **2003**, *18*, 493–505. [\[CrossRef\]](#)
5. IEC60076-18. *Measurement of Frequency Response, ED. 1'*, IEC Std. March 2012; International Electrotechnical Commission: Geneva, Switzerland, 2012.
6. IEEE Std. C57.149-2012. *IEEE Guide for the Application and Interpretation of Frequency Response Analysis for Oil-Immersed Transformers*; IEEE Standard Association: New York, NY, USA, 2013.
7. Picher, P.; Tenbohlen, S.; Lachman, M.; Scardazzi, A.; Patel, P. Current state of transformer FRA interpretation. *Procedia Eng.* **2017**, *202*, 3–12. [\[CrossRef\]](#)
8. Miyazaki, S.; Tahir, M.; Tenbohlen, S. Detection and quantitative diagnosis of axial displacement of transformer winding by frequency response analysis. *IET Gener. Transm. Distrib.* **2019**, *13*, 3493–3500. [\[CrossRef\]](#)
9. Tahir, M.; Tenbohlen, S. A Comprehensive Analysis of Windings Electrical and Mechanical Faults Using a High-Frequency Model. *Energies* **2019**, *13*, 105. [\[CrossRef\]](#)
10. FRA lookup charts for the quantitative determination of winding axial displacement fault in power transformers. *IET Electr. Power Appl.* **2020**. [\[CrossRef\]](#)
11. Tahir, M.; Tenbohlen, S.; Miyazaki, S. Analysis of Statistical Methods for Assessment of Power Transformer Frequency Response Measurements. *IEEE Trans. Power Deliv.* **2020**. [\[CrossRef\]](#)
12. Tahir, M.; Tenbohlen, S.; Samimi, M.H. Evaluation of Numerical Indices for Objective Interpretation of Frequency Response to Detect Mechanical Faults in Power Transformers. In *Proceedings of the 21st International Symposium on High Voltage Engineering, Budapest, Hungary, 26–30 August 2019*; Németh, B., Ed.; Springer International Publishing: Cham, Switzerland, 2020; Volume 598, pp. 811–824. ISBN 978-3-030-31675-4.
13. Nirgude, P.M.; Ashokraju, D.; Rajkumar, A.D.; Singh, B.P. Application of numerical evaluation techniques for interpreting frequency response measurements in power transformers. *IET Sci. Meas. Technol.* **2008**, *2*, 275. [\[CrossRef\]](#)
14. Ghanizadeh, A.J.; Gharehpetian, G.B. ANN and cross-correlation based features for discrimination between electrical and mechanical defects and their localization in transformer winding. *IEEE Trans. Dielect. Electr. Insul.* **2014**, *21*, 2374–2382. [\[CrossRef\]](#)
15. Gandhi, K.R.; Badgajar, K.P. Artificial neural network based identification of deviation in frequency response of power transformer windings. In *Proceedings of the 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD)*, Kottayam, India, 24–26 July 2014; pp. 1–8.
16. Tenbohlen, S.; Tahir, M.; Rahimpour, E.; Poulin, B.; Miyazaki, S. *A New Approach for High Frequency Modelling of Disc Windings*; CIGRE: Paris, France, 2018; pp. A2–A214.
17. Jin, Z.; Li, J.; Zhu, Z. Diagnosis of transformer winding deformation on the basis of artificial neural network. In *Proceedings of the 6th International Conference on Properties and Applications of Dielectric Materials (Cat. No.00CH36347)*, Xi'an, China, 21–26 June 2000; Volume 1, pp. 173–176.
18. Bigdeli, M.; Vakilian, M.; Rahimpour, E. Transformer winding faults classification based on transfer function analysis by support vector machine. *IET Electr. Power Appl.* **2012**, *6*, 268. [\[CrossRef\]](#)
19. Liu, J.; Zhao, Z.; Tang, C.; Yao, C.; Li, C.; Islam, S. Classifying Transformer Winding Deformation Fault Types and Degrees Using FRA Based on Support Vector Machine. *IEEE Access* **2019**, *7*, 11. [\[CrossRef\]](#)
20. Mao, X.; Wang, Z.; Jarman, P.; Fieldsend-Roxborough, A. Winding Type Recognition through Supervised Machine Learning using Frequency Response Analysis (FRA) Data. In *Proceedings of the 2019 2nd International Conference on Electrical Materials and Power Equipment (ICEMPE)*, Guangzhou, China, 7–10 April 2019; pp. 588–591.
21. Gonzales, J.C.; Mombello, E.E. Automatic detection of frequency ranges of power transformer transfer functions for evaluation by mathematical indicators. In *Proceedings of the 2012 Sixth IEEE/PES Transmission and Distribution: Latin America Conference and Exposition (T&D-LA)*, Montevideo, Uruguay, 3–5 September 2012; pp. 1–8.
22. Lin, H.; Zhang, Z.; Tang, W.; Wu, Q.; Yan, J. Equivalent gradient area based fault interpretation for transformer winding using binary morphology. *IEEE Trans. Dielect. Electr. Insul.* **2017**, *24*, 1947–1956. [\[CrossRef\]](#)



23. Velasquez-Contreras, J.L.; Kolb, D.; Sanz-Bobi, M.A.; Koltunowicz, W. Identification of transformer-specific frequency sub-bands as basis for a reliable and automatic assessment of FRA results. In Proceedings of the Conference on Condition Monitoring and Diagnosis 2010 (CMD2010), Tokyo, Japan, 6–11 September 2010.
24. CIGRE WG A2.53. *Advances in the Interpretation of Transformer Frequency Response Analysis (FRA)*; CIGRE Brochure 812; CIGRE: Paris, France, 2020.
25. Miyazaki, S.; Mizutani, Y.; Taguchi, A.; Murakami, J.; Tsuji, N.; Takashima, M.; Kato, O. Diagnosis Criterion of Abnormality of Transformer Winding by Frequency Response Analysis (FRA). *Electr. Eng. Jpn.* **2017**, *201*, 25–34. [[CrossRef](#)]