Received December 17, 2020, accepted January 12, 2021, date of publication January 18, 2021, date of current version January 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3052144

Intelligent Classifiers in Distinguishing Transformer Faults Using Frequency Response Analysis

MEHDI BIGDELI[1], PIERLUIGI SIANO[2], (Senior Member, IEEE), AND HASSAN HAES ALHELOU[3], (Senior Member, IEEE)

[1]Department of Electrical Engineering, Zanjan Branch, Islamic Azad University, Zanjan 45156, Iran
[2]Department of Management and Innovation Systems, University of Salerno, 84121 Salerno, Italy
[3]Department of Electrical Power Engineering, Tishreen University, Lattakia 2230, Syria

Corresponding author: Hassan Haes Alhelou (alhelou@tishreen.edu.sy)

ABSTRACT With the expansion of the use of frequency response analysis (FRA) as a reliable tool for fault detection in transformers, more capabilities of this method are discovered every day. So that today the number of transformer faults that can be identified by FRA method has also increased. One of the most critical steps in fault detection with FRA is to distinguish faults and classify them in different classes. In this paper, well-known intelligent classifiers (probabilistic neural network, decision tree, support vector machine, and k-nearest neighbors) are used to classify transformer faults. For this purpose, the necessary measurements are performed on the model transformers under the healthy condition and under different fault conditions (axial displacement, radial deformation, disc space variation, short-circuits, and core deformation). Then, by dividing the frequency ranges of the measured transfer functions of the transformer, a new feature based on numerical and statistical indices for training and validation of classifiers is proposed. After completing the training process, the performance of the classifiers is evaluated and compared by applying the data obtained from real transformers.

INDEX TERMS Transformer, fault type detection, frequency response analysis (FRA), intelligent classifiers, measurement, numerical indices.

NOMENCLATURE

AD	Axial Displacement
ANN	Artificial Neural Networks
ASLE	Absolute Sum of Logarithmic Error
CC	Correlation Coefficient
CCF	Cross-Correlation Function
CSD	Comparative Standard Deviation
DABS	Absolute Difference
DCS	Deformation of Core Sheets
DSV	Disc Space Variation
DT	Decision Tree
ED	Euclidean Distance
FDA	Fisher Discriminant Analysis
FRA	Frequency Response Analysis
IA	Integral of Absolute difference
ID	Integral of Difference
IMGM	Interval Maximum to Global Maximum
k-NN	k-Nearest Neighbors
MAX	Maximum of difference
MM	Minimum-Maximum ratio
MMR	Maximum-Minimum Ratio
PNN	Probabilistic Neural Network
RD	Radial Deformation
RMSE	Root Mean Square Error
SC	Short Circuit
SD	Standard Deviation
SDA	Standardized Difference Area
SSE	Sum Squared Error
SSMMRE	Sum Squared Max-Min Ratio Error
SSRE	Sum Squared Ratio Error
SVM	Support Vector Machine
TF	Transfer Function
$E[\Delta]$	Expectation
σ_e	Standard deviation
σ	Spectrum deviation
σ_s	Stochastic spectrum deviation

The associate editor coordinating the review of this manuscript and approving it for publication was Fabio Massaro.

ρ Normalized correlation coefficient
R_{XY} Correlation factor

I. INTRODUCTION

Nowadays, intelligent monitoring of equipment in power systems is very important. Power transformers are one of the most critical and expensive equipment in a power generation and transmission network. Their failures will impose high costs and reduce the reliability of the power grid. Therefore, the care and protection of transformers during operation is necessary. One of the methods used in recent years to monitor the status of transformers is the FRA method. The FRA method, also known as the TF approach, is a comparative method [1]. In this method, the measurements are made on a typical transformer in the healthy condition are kept as the reference measurement with the customer or the manufacturer. After gathering the reference (healthy) TF in a typical transformer, the TFs results can be measured (from similar terminals) and compared with the reference at any desired time. By comparing the reference TF with new measurements, the type and severity of the transformer's fault can be determined. Unfortunately, the existing standards in the field of FRA [2], [3] have focused on measurement requirements and test circuits, and a precise standard for interpreting FRA measurement results has not yet been developed. Therefore, in recent years, many studies have been conducted on the interpretation of FRA measurement's results to obtain information about transformer faults [4]–[8]. Given that the first step is to identify the type of fault, the focus of current research is to classify faults.

The most important faults that occur in the transformer and can be identified by the FRA method are:

- Axial displacement (AD)
- Radial deformation (RD)
- Disc space variation (DSV)
- Short circuit (SC)

Besides, the deformation of core sheets (DCS) is also detectable by the FRA, which has not been well studied in the literature.

Various methods have been proposed in the literature to identify the type of fault and classify them, which can be divided into two main categories. The first category includes methods according to which faults' classification is solely based on the rate of variations in numerical indices (statistical and mathematical indices) in specific frequency ranges [9]–[18]. The second category includes methods that use intelligent classifiers to distinguish faults [19]–[25]. In these methods, the necessary features of frequency response (mainly the statistical and numerical indices) are extracted, and these features are used for training and testing classifiers.

In [9], the most important numerical indices have been introduced. By calculating them in different fault conditions and comparing them to the healthy condition, the indices that have a better ability to detect defects have been identified. In [10], the transformer's TFs in AD, RD and also simultaneous AD-RD fault conditions are obtained with the help of the finite element method. Then, using statistical indices such as ASLE and DABS, the transformer faults were detected. In [11], with the introduction of MMR based on the ED index; the type of transformer winding fault is identified. A new method using the wavelet technique and characteristic impedance for classifying defects has been proposed in [12]. Using antenna installation on the transformer and by analyzing the electromagnetic waves received from the antenna based on the ED index, a distinction is made between the faults in [13]. In [14], the measured TFs from the transformer are mapped to a two-dimensional space, and the fault type is detected based on a vector-based approach. In [15], the transformer's TFs are first estimated, and then a distinction is made between the faults by plotting the Nyquist diagram. Nonetheless, only two faults, AD and RD, have been considered in these studies. A new characteristic called α is introduced based on the ρ index to distinguish transformer faults in [16]. However, three faults, AD, RD, and DSV, have been investigated, while SC and DCS faults are not considered. In [17], CCF index is introduced, and various faults of the transformer are clustered. However, three faults AD, RD, and SC are clustered, and DSV and DCS are not considered in the fault detection process. In [18], a new index called IMGM is proposed for fault type detection and all five defects have been considered. The IMGM distinguishes faults with probability and it is unable to certainly detect the type of fault. Also, its conclusion emphasizes that intelligent classifiers should be used to detect transformer faults in the future.

In [19], [20], based on the TF estimation with the help of vector fitting, the necessary features of the measured TFs for four faults AD, RD, DSV, and SC were extracted and these faults are classified by using PNN [19] and SVM [20]. A distinction has been made between electrical and mechanical faults (AD and RD only) and the inrush current through ANN and the DT in [21]. Classification of RD, DSV, and SC faults in three different classes has been done in [22], [23]. In [22], the measured frequency response from the transformer is divided into three frequency ranges and using binary imaging and SVM, a distinction is made between the faults. In [23], statistical and numerical indices were used for SVM training and testing; nonetheless, the frequency division was not performed. A new windowing-based approach to determine the type of fault has been proposed in [24]. After windowing and calculating the indices, the FDA method was used to separate the faults. In [25], CCF was used to extract the characteristic, and these features were used for ANN training and testing to separate the three faults AD, RD, and SC.

As noted, several methods have been proposed to classify transformer faults, but none of these studies has considered DCS fault, and in many of these studies, two or eventually three faults have been investigated. The faults classification can be more difficult and complicated with the increase in the number of defects that can be identified by the FRA method. Also, one or two classifiers have been used in

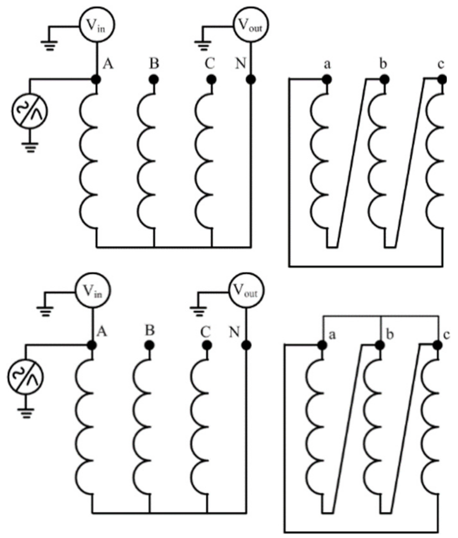


FIGURE 1. TFs measurement circuits.

previous researches, and the number of features extracted for classifiers training has been limited. To address these shortcomings, this article examines all five faults. For this purpose, the required TFs are extracted by performing the necessary measurements on different transformers in healthy condition and different fault conditions (AD, RD, DSV, SC, and DCS). Four intelligent classifiers, PNN, DT, SVM, and k-NN, are used to classify faults into five different classes. A new feature is proposed for the training and validation of intelligent classifiers by calculating the well-known statistical and numerical indices. Data obtained from real transformers are applied to classifiers to evaluate the performance of them. Analysis of the results shows that by dividing the frequency ranges of the measured TFs into 10 equal intervals and using SVM and k-NN classifiers with MAX, CSD, and R_{XY} numerical indices as the feature, the best performance is obtained.

The innovative contributions of the proposed approach if compared to previous studies are as follows:

1- Proposing a new feature based on numerical indices for training and testing of intelligent classifiers,
2- Comparison of the performance of four well-known intelligent classifiers (PNN, DT, SVM, k-NN) in distinguishing transformer faults and proposing the most suitable classifier,
3- Apply the proposed feature to real transformers which are faulty during operation,
4- Compare the performance of different numerical indices and determine the reliable method.

II. CASE STUDIES AND MEASUREMENTS

In the measurements carried out in this study, the circuit of Figure 1 is used [3]. It is important to note that in this figure, instead of the output voltage (V_{out}), the output current can be measured so that the TF will be of the admittance type.

To evaluate intelligent classifiers' performance, it is necessary to establish a database of transformers in healthy and faulty conditions (with different intensities of the fault). For this purpose, two sets of transformers are tested.

A. THE FIRST GROUP OF TRANSFORMERS

The first group of transformers is model transformers that the desired faults are intentionally created on them. A model transformer is a transformer whose structure is exactly the same as an actual transformer, but its voltage and power level may not be real. Therefore, it is used only for laboratory studies. In addition, different connections are available from its windings and it is possible to intentionally apply various defects on the transformer in order to obtain a more complete database. This group of transformers is tested with almost similar structures, and one of the studied faults (AD, RD, DSV, SC, and DCS) is applied to each of them. The description of these test objects is as follows:

Case 1: A transformer is consisting of a HV winding made of 31 pairs of 6-turns discs, and a LV winding made of 4 layers of 99 turns is used to study the AD of the windings relative to each other. The nominal power and voltage of this transformer are 1.3 MVA and 10 kV, respectively. In experiments performed, the internal layer winding is moved axially in 8 steps (1 cm in each step) respect to the outer disc winding.

Case 2: The transformer tested for RD study has a HV winding consisting of 30 pairs of 11-turns discs and a LV winding made of one layer of 23-turns. The rated power and voltage of this transformer are 1.3 MVA and 10 kV, respectively. In this test object, the disc winding is deformed in 4 steps in a radial direction [16].

Case 3: The transformer studied in this section is the same as Case 2. Another healthy winding has been selected to study the effect of DSV on frequency response. The distance between the discs in the healthy condition is 5 mm. This distance has been changed to 7.5, 10, 15, 20 and 25 mm to consider the effect of the DSV on the TFs. The distance between the discs was changed in 3 locations (discs 2, 4, and 16) to study the effect of the fault better.

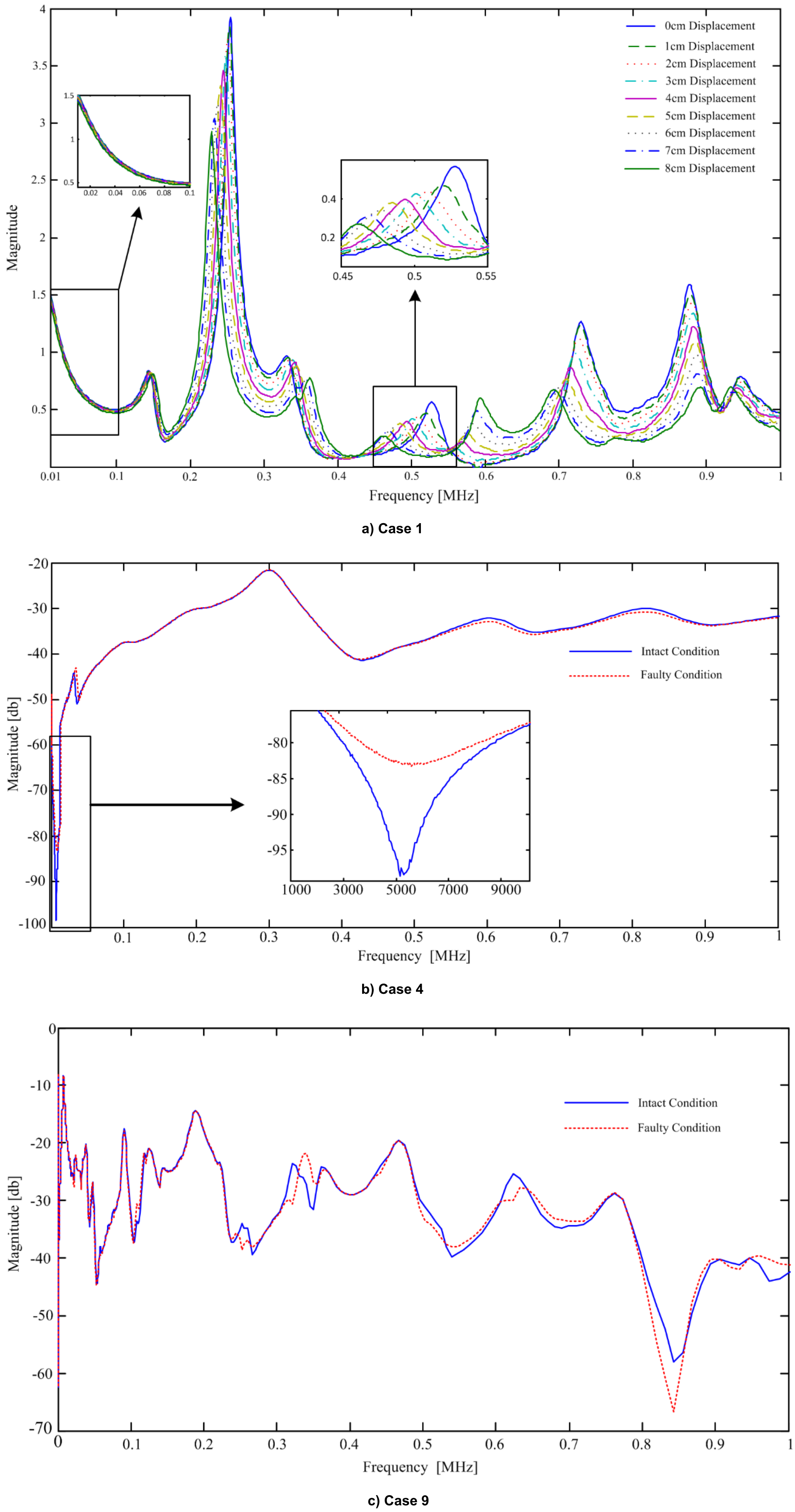
Case 4: A sample transformer is consisting of a HV winding made of 30 discs of 11-turns with 1 MVA rated power and 20 kV rated voltage is used to investigate the effect of DCS on the frequency response. The winding is mounted on the core, and the TF of this winding is measured in two conditions. First, in the healthy condition and then by deformation of the core, the TF is measured again.

Case 5: In this case, the HV winding of a 10 kV and 1.2 MVA transformer is tested. The winding is including 60 discs with 9 turns in each disc. All discs have an accessible branch that makes it possible to measure TFs at different locations along the winding.

B. THE SECOND GROUP OF TRANSFORMERS

The second group is real transformers that exposed to faults during operation. Initially, the type of fault is not known, and after opening its accessories, the type of fault is recognized. The specification of these transformers is as follows:

Case 6: This real case is a 20/0.4 kV and 0.4 MVA distribution transformer with a HV winding made of 40 discs, 17 turns in each disc, and a LV winding made of 2 layers,



a) Case 1

b) Case 4

c) Case 9

FIGURE 2. Some of the measured TFs of transformers.

9 turns in each layer. Due to improper transportation of the transformer from the production site to the installation site, the AD fault has occurred.

Case 7: This is a distribution transformer with the rated specification of 20/0.4 kV and 1 MVA with a HV winding made of 50 discs, 11 turns in each disc, and a LV winding made of 2 layers, 11 turns in each layer. This transformer has an RD fault as a result of a short circuit near the transformer terminals.

Case 8: In this transformer, a DSV fault occurred due to the accident of the vehicle carrying the transformer. The rated voltage and power are 63/20 kV and 30 MVA, respectively. The windings consist of a HV winding made of 80 discs, 15 turns in each disc, and a LV winding made of 5 layers, 64 turns in each layer.

Case 9: This real case is a 20/0.4 kV and 0.5 MVA distribution transformer with a HV winding made of 45 discs, 13 turns in each disc, and a LV winding made of 2 layers, 10 turns in each layer. The transformer was thrown to the ground while mounting on a vehicle, and a DSV fault has occurred.

Case 10: This is a transformer with the same specifications as the Case 7, which has a DCS fault.

The measurement results of some of these transformers are shown in Figure 2. The experimental results point out that each defect affects a specific frequency range. However, the type of faults cannot be determined by the appearance of the measured TFs, and a fault detection method should be used for this purpose, which will be discussed in the next sections.

III. INTELLIGENT CLASSIFIERS

In this paper, four methods of PNN, DT, SVM, and k-NN are used to detect the type of transformer fault. The theory of these methods has been extensively studied in the literature [26]–[29]. The ability of these methods to solve transformer faults classification problems has also been demonstrated in [19]–[25]. Therefore, in this section, these methods are briefly discussed.

A. PROBABILISTIC NEURAL NETWORK

The design of an ANN includes the selection of inputs, outputs, network structure, and neuron weight vectors [26]. The choice of network topology is made in an experimental process with trial and error to optimize the number of layers and neurons in the network. The PNN structure is a three-layer network. When an input vector is applied to the network, the first layer calculates the input vector distance from the training inputs, thus providing vector whose elements determine the distance between the new input and the training input. The second layer uses the output of the first layer to produce the vector of probabilities as the output of the network. Finally, the competitive transfer function in the third layer selects the maximum number of probabilities from the probability vector and produces 1 for that and 0 for the rest of the probabilities.

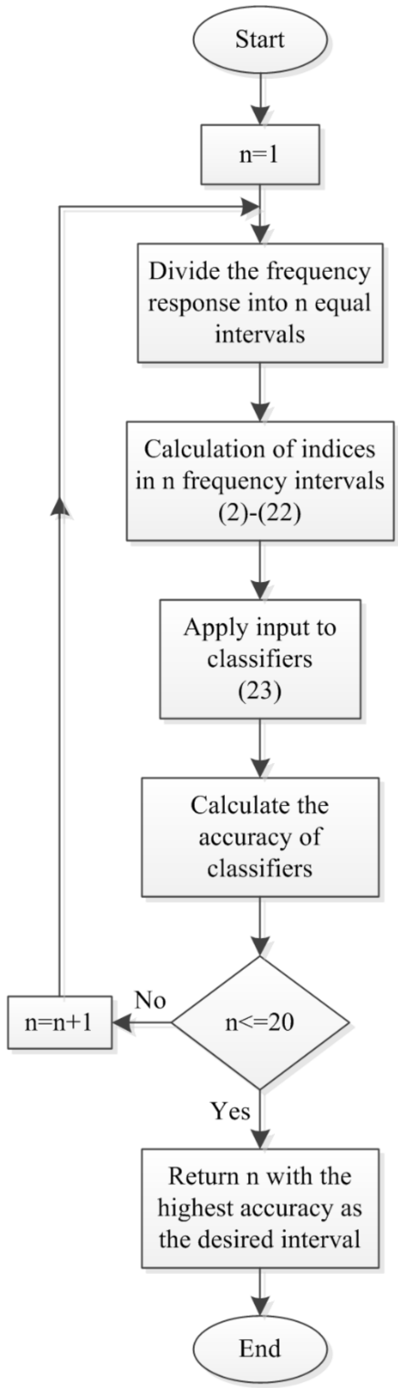


FIGURE 3. Flowchart of proposed method for finding the best interval for calculating indices.

B. DECISION TREE

This method is a data mining tool for decision support that uses a tree-link model of decisions. It isn't easy to find the optimal tree in practice, but powerful algorithms such as ID3, C4.5, CART, and CHAID have been proposed for this purpose. One of the essential advantages of the C4.5 algorithm is the ability to convert trained trees into a series of if-then rules. On the other hand, no feature extraction technique is required in this method [27]. Therefore, this algorithm is used in this research.

C. SUPPORT VECTOR MACHINE

It is one of the most powerful classifiers that can perform both linear and nonlinear data separation operations. One of the best kernel functions used in SVM is the Gaussian function, which is defined as follows [28]:

$$K(X,Y) = \exp(-\frac{1}{2\sigma^2}\|X-Y\|^2) \tag{1}$$

where, X and Y are the input and output vectors, respectively, and is a constant value that is determined by the type of data. Also, SVM has another parameter called C, which controls the amount of over-fit or under-fit. Accurate determining the parameters of and C is very important and is usually determined by trial and error.

D. k-NEAREST NEIGHBORS

The k-NN method is an efficient tool for classifying data. The k-NN algorithm is straightforward, and it works by calculating the distance between data to put those that are similar to each other in a neighborhood. The number of neighbors (k) depends on the input data, but usually, an odd number is selected. When new data is applied to k-NN, the distance between that data and the k neighbors closest to it is calculated. Any class that has more neighbors will win the competition. The two most common methods for calculating the distance between data are Euclidean distance and Pearson methods [29]. In this paper, the Euclidean distance is used.

IV. THE PROPOSED METHOD FOR EXTRACTING THE FEATURES

One of the most critical parts of any pattern recognition system is the feature extraction. In detecting transformer faults, the feature extraction is based on the comparison of data of the measured TFs with the reference TF. One of the best ways to compare TFs with reference TF is to use numerical and statistical indices. Although numerical indices have been used in the diagnosis of transformer defects [9]–[12], [16]–[18], however, a comprehensive comparison of the performance of these indices for the purpose of training the intelligent classifiers is not seen in the literature. Therefore, in this paper, it is proposed that each of the indices be used separately as a feature. Then, the performance of each index in the transformer fault classification is evaluated to identify the most reliable indicator.

Various numerical indices have been presented to compare TFs, the most important and most commonly used are:

$$ED = \|Y-X\| = \sqrt{(Y-X)^T(Y-X)} \tag{2}$$

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

$$ID = \int (Y(f)-X(f))\,df \tag{4}$$

$$IA = \int |Y(f)-X(f)|\,df \tag{5}$$

$$SDA = \frac{\int |Y(f)-X(f)|\,df}{\int |X(f)|\,df} \tag{6}$$

$$ASLE = \frac{\sum_{i=1}^{N}\left|20\log_{10}Y(i)-20\log_{10}X(i)\right|}{N} \tag{7}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{|Y(i)|-|X(i)|}{\left(\frac{1}{N}\right)\sum_{i=1}^{N}|X(i)|}\right)^2} \tag{8}$$

$$E[\Delta] = \frac{1}{N}\sum_{i=1}^{N}\Delta(i);\ \Delta(i) = \frac{|Y(i)|-|X(i)|}{\left(\frac{1}{N}\right)\sum_{i=1}^{N}|X(i)|} \tag{9}$$

$$\sigma_e = \sqrt{var(\Delta)} = E[\Delta - E(\Delta))] \tag{10}$$

$$\sigma = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\left(\frac{X(i)-(X(i)+Y(i))/2}{(X(i)+Y(i))/2}\right)^2 + \left(\frac{Y(i)-(X(i)+Y(i))/2}{(X(i)+Y(i))/2}\right)^2} \tag{11}$$

$$\sigma_s = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{Y(i)-X(i)}{X(i)}\right| \tag{12}$$

$$MAX = \max(Y(i)-X(i)) \tag{13}$$

$$\rho = \frac{\sum_{i=1}^{N}(X(i)-\bar{X})(Y(i)-\bar{Y})}{\sqrt{\sum_{i=1}^{N}(X(i)-\bar{X})^2\sum_{i=1}^{N}(Y(i)-\bar{Y})^2}}$$

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N}X(i),\ \bar{Y} = \frac{1}{N}\sum_{i=1}^{N}Y(i) \tag{14}$$

$$CC = \frac{\sum_{i=1}^{N}X(i)Y(i)}{\sqrt{\sum_{i=1}^{N}(X(i))^2\sum_{i=1}^{N}(Y(i))^2}} \tag{15}$$

$$SSE = \frac{\sum_{i=1}^{N}(Y(i)-X(i))^2}{N} \tag{16}$$

$$SSRE = \frac{\sum_{i=1}^{N}(\frac{Y(i)}{X(i)}-1)^2}{N} \tag{17}$$

$$SSMMRE = \frac{\sum_{i=1}^{N}(\frac{\max(Y(i),X(i))}{\min(Y(i),X(i))}-1)^2}{N} \tag{18}$$

$$DABS = \frac{\sum_{i=1}^{N}|Y(i)-X(i)|}{N} \tag{19}$$

$$MM = \frac{\sum_{i=1}^{N}\min(X(i),Y(i))}{\sum_{i=1}^{N}\max(X(i),Y(i))} \tag{20}$$

$$CSD = \sqrt{\frac{\sum_{i=1}^{N}\left[(X(i)-\bar{X})-(Y(i)-\bar{Y})\right]^2}{N-1}} \tag{21}$$

$$R_{XY} = \begin{cases} 10 & if\ (1-\rho) \le 10^{-10} \\ -\log_{10}(1-\rho) & otherwise \end{cases} \tag{22}$$

where, X and Y are the magnitude vectors of the reference TF and the new TF, respectively, f is the measured frequency vector, and N is the number of samples in a vector.

Another shortcoming of past well-known studies [9]–[12] is that in order to extract the features, the indices have been calculated in the entire frequency interval (0-1MHz). Obviously, the result of this calculation will be just one numerical value, which is insufficient to train the classifier. In some studies [16], [23], the indices have been calculated in 3 frequency intervals. However, it is not clear in which frequency interval these indices are calculated, the best performance is obtained for the classifier. To solve this problem, in current research, the measured frequency response is divided into different intervals. This frequency division should be such that the number of data is not high because it makes the decision difficult for classifiers, and its speed will be reduced. Also, if the amount of data is lower than a certain level, classifiers will not be well trained and will produce incorrect output. Therefore, in this paper, different states of frequency division are examined. To achieve a suitable result, the number of intervals is changed from 1 to 20 and the output of the classifiers is evaluated in each interval. Figure 3 shows the proposed method for finding the best interval for calculating indices. The interval that produces the output with the highest accuracy for the classifier is selected as the most appropriate interval for calculating the numerical indices.

It should be noted that after calculating the mentioned indices in the custom frequency intervals of the TFs, these features are applied as inputs to the classifiers. Therefore, the input matrix for each indicator can be defined as Equation (23), as shown at the bottom of the page.

Classifier output can be numbered 1 to 5, which are assigned to AD, RD, DSV, DCS, and SC, respectively. Therefore, the output vector can be defined as Equation (24), as shown at the bottom of the page.

Where, i represents the intensity of AD from 1 to 8 cm, j is the degree of RD from 1 to 4, k is the intensity of DSV from 7.5 to 25 mm and for 3 locations 2, 4 and 16, m is the intensity of DCS and l is the location of SC. n also represents the number of frequency intervals, which in this article is equal to 10. In fact, with this formatting, the samples are placed on the matrix columns.

V. CLASSIFICATION RESULTS

In this section, by applying the features extracted from the TFs to intelligent classifiers, their performance is evaluated. In addition to test data, part of the model's transformers data is used to validate classifiers to prevent over-fitting. For this purpose, the K-Fold cross-validation method is used. In this method, the classification is done K times, and in each time, a fraction of 1/K of data is used for validation, and the rest is used for training. Then the mean of the errors is returned as the classification error. In this paper, the value of K is considered to be 5.

In the first step, the desired interval for calculating numerical indices must be determined. For this purpose, the proposed flowchart in Figure 3 is used. Figure 4 shows the accuracy of some indices for different frequency intervals of the measured TFs for the SVM Classifier. It is observed that the highest accuracy is related to the state where the measured frequency range is divided into 10 equal intervals. This has been done for all indices and classifiers, and a similar result has been obtained.

After calculating the numerical indices in 10 equal frequency intervals, the input matrix is created. Then, by applying the input matrix to the classifiers, the diagnosis error of each classifier is determined. Table 1 shows the diagnosis error in applying validation and test data to different classifiers for all indices.

Examining the results of Figure 4 and Table 1 shows that:

1- To extract the features, the indices can be calculated in different frequency intervals. The highest accuracy is obtained when the measured TF is divided into 10 equal frequency intervals. Therefore, it is proposed that the feature extraction be performed based on the calculation of numerical indices in 10 equal frequency intervals of the measured TF.

2- In PNN and DT classifiers, the diagnosis error of test data is not less than 40% for any indicator. Besides, the diagnosis error of validation data is above 10%. Therefore, in solving the classification problem mentioned in this article, the use of these classifiers is not recommended.

3- In all classifiers, the diagnosis error for ID, IA, RMSE, E, σ_e, σ, σ_s indices in test data is more than 50%. Also, the classification error in these indices is high in validation data. These indices fall into the first category.

4- In SVM and k-NN classifiers, the ED, SD, SDA, ASLE, ρ, CC, SSE, SSRE, SSMMRE, DABS, and MM indices have less error than the first category indices (validation error less than 15% and test error less than 50%). Nonetheless, the diagnosis error is still high. These indices fall into the second category.

5- The third category is indices that have less error than other indices. These indices are MAX, CSD, and R_{XY}.

The above analysis shows that for the extraction of the features, it is better to use the third category indices and use SVM and k-NN classifiers to classify the faults. Therefore, the results of these indices and classifiers will be examined in more detail.

Figure 5 shows the diagnosis accuracy of the validation data applied to the SVM and k-NN classifiers for the three MAX, CSD, and R_{XY} indices. As can be seen, the MAX index for SVM and k-NN classifiers has been misdiagnosed in two cases, which are related to AD, RD, and DCS faults. Misdiagnosis of DCS can significantly affect the classifier's

$$Input\ Matrix = \begin{bmatrix} Index_{1,ADi} & Index_{1,RDj} & Index_{1,DSVk} & Index_{1,DSVm} & Index_{1,SCl} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Index_{n,ADi} & Index_{n,RDj} & Index_{n,DSVk} & Index_{n,DSVm} & Index_{n,SCl} \end{bmatrix} \tag{23}$$

$$Output\ Vector = \begin{bmatrix} AD_1 \cdots AD_i \ RD_1 \cdots RD_j \ DSV_1 \cdots DSV_k \ DCS_1 \cdots DCS_m \ SC_1 \cdots SC_l \end{bmatrix} \tag{24}$$

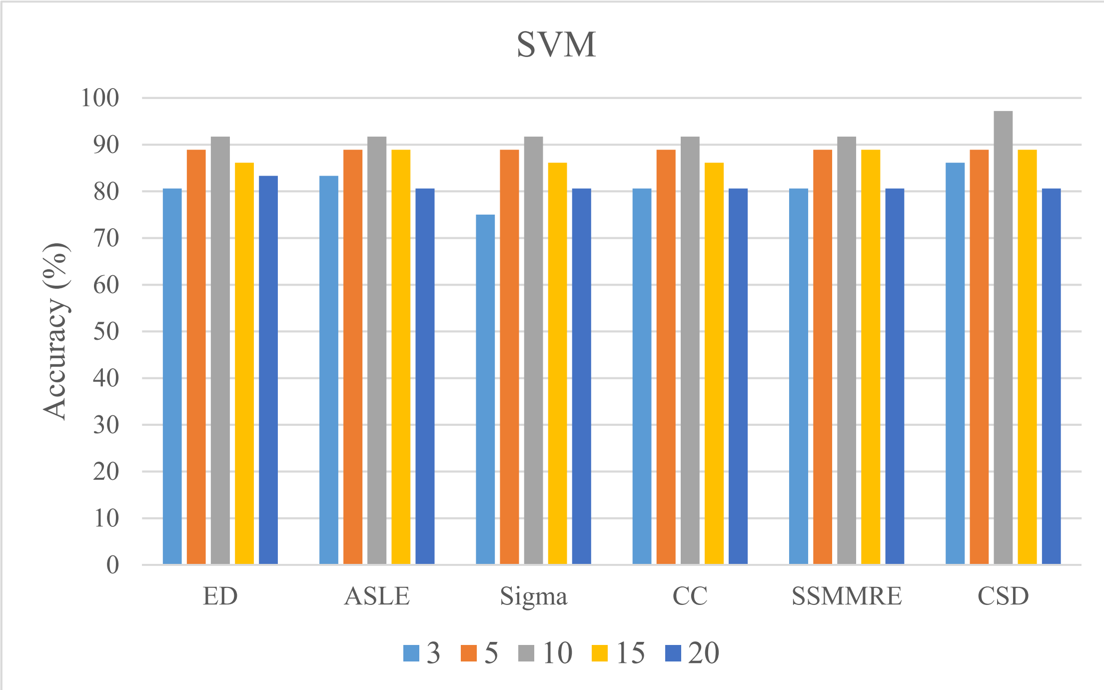


FIGURE 4. The accuracy of some indices by calculating them in different frequency intervals of measured TFs.

TABLE 1. Classifier diagnosis error for test and validation data (in percentage).

Index	Classifier							
	PNN		DT		SVM		k-NN	
	Validation	Test	Validation	Test	Validation	Test	Validation	Test
ED	13.9	60	13.9	40	8.3	40	2.8	40
SD	13.9	40	13.9	40	5.6	40	2.8	40
ID	21.2	80	19.4	80	8.3	80	8.3	60
IA	16.7	80	27.8	60	5.6	60	8.3	60
SDA	11.1	40	11.1	40	8.3	40	2.8	40
ASLE	13.2	40	11.2	40	8.3	40	5.6	40
RMSE	19.4	60	19.4	60	11.1	80	8.3	60
E	21.2	60	16.7	80	11.1	80	5.6	80
σ_e	25	60	35.1	80	16.7	60	11.1	80
σ	27.8	60	21.2	60	8.3	60	5.6	60
σ_s	19.4	80	21.2	80	8.3	60	5.6	60
MAX	11.1	40	13.9	40	5.6	20	5.6	20
ρ	21.2	40	19.4	60	8.3	40	8.3	40
CC	16.7	60	16.7	60	8.3	40	8.3	40
SSE	19.4	60	16.7	40	8.3	40	2.8	40
SSRE	21.2	40	25	40	11.1	40	2.8	40
SSMMRE	21.2	40	25	40	11.1	40	8.3	40
DABS	16.7	40	19.4	40	8.3	40	8.3	40
MM	25	40	41.7	40	8.3	40	8.3	40
CSD	11.1	40	11.1	40	2.8	0	2.8	20
R_{XY}	13.9	40	11.1	40	2.8	20	2.8	20

performance because only one measurement result for this fault is available in the model transformers and not recognizing it correctly means that the classifier has not been properly trained for this fault and produces a 100% error. In the CSD index, only one condition of AD and RD faults has been misdiagnosed for each of classifiers. In the R_{XY} index, the performance of SVM and k-NN classifiers is the opposite of their performance in the CSD index.

Despite the good results obtained in the training and validation phase of MAX, CSD, and R_{XY} indices, but to prove the capabilities of the proposed method and generalize it, various TFs of different transformers must be examined. For this purpose, the measured data from the second group of transformers are used to test and verify the proposed method. Figure 6 shows the results of applying test data to SVM and k-NN classifiers for three indices: MAX, CSD, and R_{XY}.

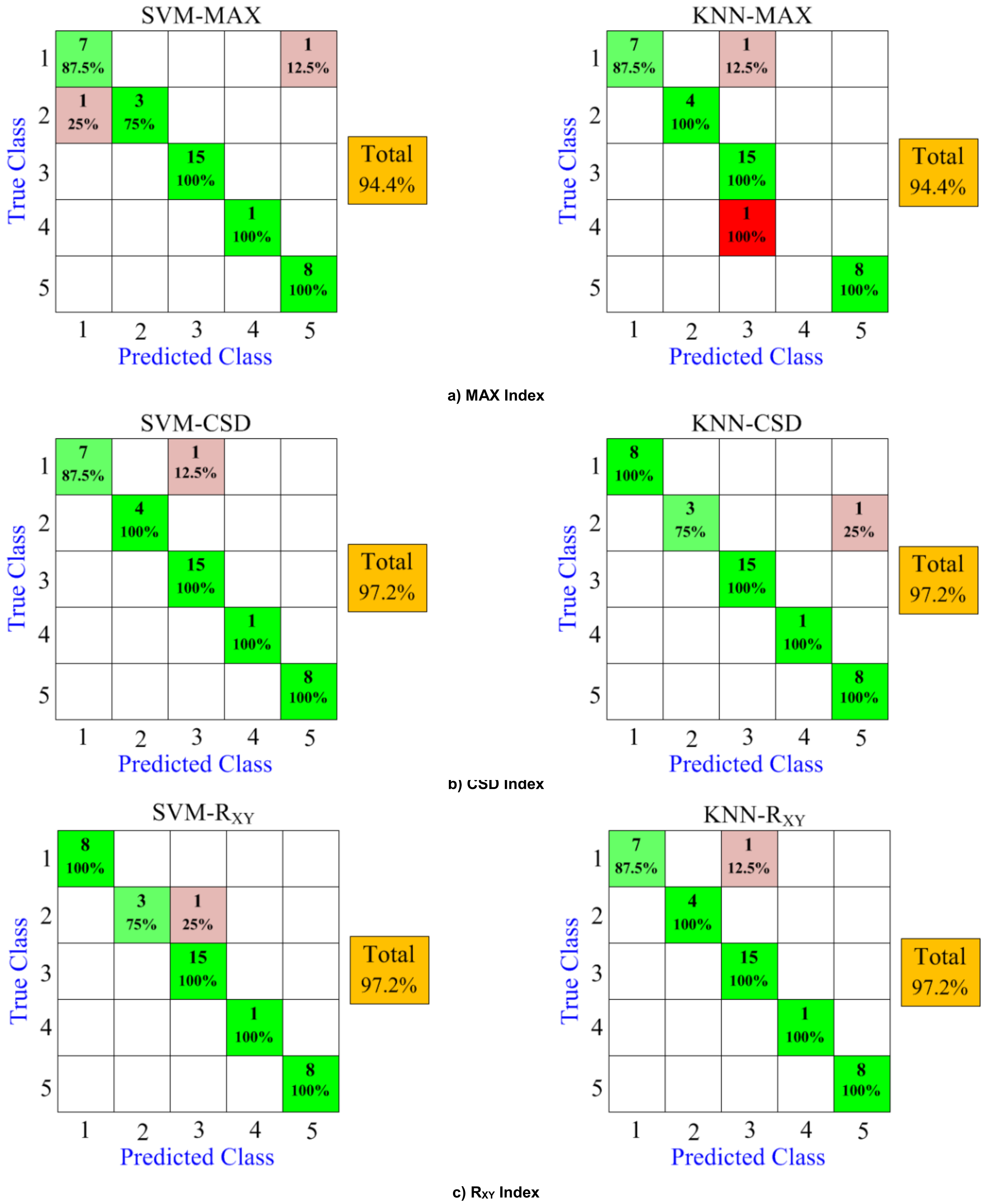
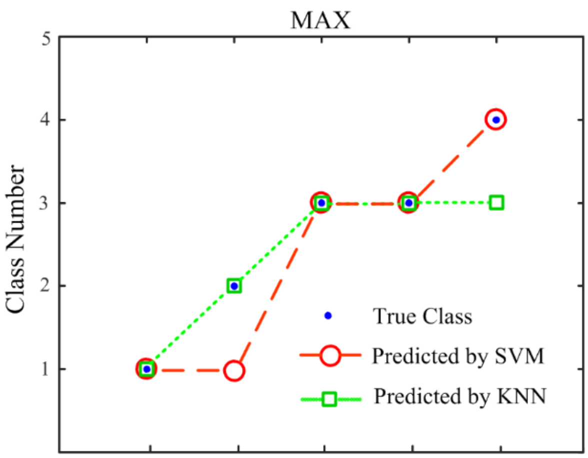


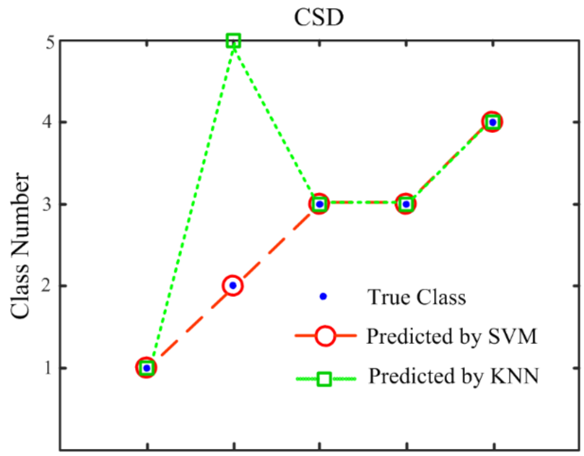
FIGURE 5. Performance of third category indices using K-Fold cross-validation.

As expected, the k-NN method in the MAX index failed to identify the DCS fault correctly. Also, the SVM method in this index could not accurately identify the AD fault. For both classifiers, the R_{XY} index only misdiagnosed AD. In the CSD index, for the k-NN classifier, the AD fault has been misdiagnosed, but in the SVM method, all fault conditions have been correctly identified.

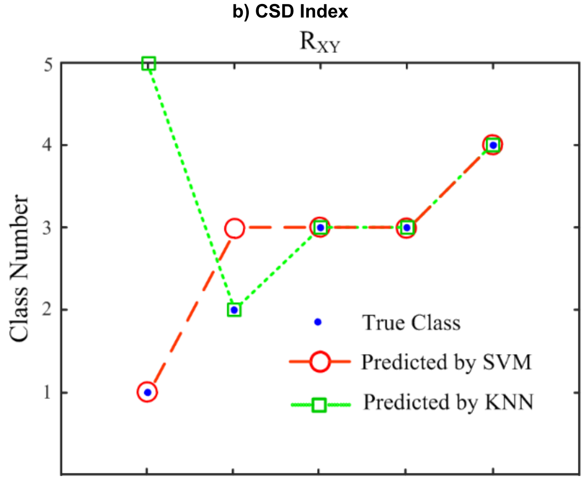
The results of the above analysis show that the classification error of a classifier may be somewhat high in the validation phase. Nonetheless, in the testing phase, good accuracy is obtained. The reason for this is that training and validation of classifiers are performed with different intensities of a fault, and maybe just for certain intensity, validation is not correct that particular intensity is not available in test data.



a) MAX Index



b) CSD Index



c) R_{XY} Index

FIGURE 6. Performance of third category indices in response to test data.

For example, in the MAX index and with the k-NN classifier, for the low intensity (displacement of 1 cm) in the AD fault in the validation phase, the fault type is recognized as DSV. However, in the test phase, the classifier's diagnosis was correct for the transformer that had an AD fault because the intensity of the fault in this transformer has been severe (displacement of more than 4 cm).

In general, for transformers studied in current research, the CSD index with SVM classifier has the best performance in classifying transformer winding faults. Also, to extract the features, it is better to calculate the numerical indices in 10 equal frequency intervals of the measured TF.

VI. CONCLUSION

Due to the increasing use of FRA in detecting transformer faults and increasing the number of faults that can be detected with FRA's help, it is necessary to provide a reliable method for classifying faults. Therefore, in this paper, the most important intelligent classifiers (PNN, DT, SVM, and k-NN) were used to classify faults. To train and test classifiers, a new feature based on statistical and numerical indices was proposed. In the proposed feature, statistical indices were calculated in 10 equal frequency intervals of the measured TFs. The required data for the extraction of the features were obtained by performing the necessary measurements on different transformers in the healthy condition and the different faults conditions (AD, RD, DSV, DCS, and SC). By calculating the indices and applying the extracted features to the intelligent classifiers, their performance was evaluated. The obtained results showed that the three features, MAX, CSD, and R_{XY} with SVM and k-NN classifiers, have fewer errors compared to other indices and classifiers. Among these, the SVM classifier with CSD feature has 97.2% accuracy and 100% accuracy in diagnosing validation and testing data, respectively. Therefore, it is recommended as a reliable method in the industry.

ACKNOWLEDGMENT

M. Bigdeli would like to thank Dr. E. Rahimpour for his valuable help in providing the experimental data of the case studies 1 to 3.

REFERENCES

[1] E. Rahimpour, J. Christian, K. Feser, and H. Mohseni, "Transfer function method to diagnose axial displacement and radial deformation of transformer winding," *IEEE Trans. Power Del.*, vol. 18, no. 2, pp. 493–505, Apr. 2003.

[2] *IEEE Guide for the Application and Interpretation of Frequency Response Analysis for Oil-Immersed Transformers*, IEEE Standard C57.149, 2012.

[3] *IEC Standard on Power Transformers, Part 18: Measurement of Frequency Response*, Standard IEC 60076-18, 2012.

[4] A. Abu-Siada, M. I. Mosaad, D. Kim, and M. F. El-Naggar, "Estimating power transformer high frequency model parameters using frequency response analysis," *IEEE Trans. Power Del.*, vol. 35, no. 3, pp. 1267–1277, Jun. 2020.

[5] M. H. Samimi, P. Hillenbrand, S. Tenbohlen, A. A. Shayegani Akmal, H. Mohseni, and J. Faiz, "Investigating the applicability of the finite integration technique for studying the frequency response of the transformer winding," *Int. J. Electr. Power Energy Syst.*, vol. 110, pp. 411–418, Sep. 2019.

[6] N. Jianqiang, Z. Zhongyong, T. Shan, C. Yu, Y. Chenguo, and T. Chao, "The actual measurement and analysis of transformer winding deformation fault degrees by FRA using mathematical indicators," *Electr. Power Syst. Res.*, vol. 184, pp. 1–11, Jul. 2020.

[7] X. Zhao, C. Yao, Z. Zhou, C. Li, X. Wang, T. Zhu, and A. Abu-Siada, "Experimental evaluation of transformer internal fault detection based on V–I characteristics," *IEEE Trans. Ind. Electron.*, vol. 67, no. 5, pp. 4108–4119, May 2020.

[8] A. Shamlou, M. R. Feyzi, and V. Behjat, "Interpretation of frequency response analysis of power transformer based on evidence theory," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 17, pp. 3879–3887, Sep. 2019.
[9] M. H. Samimi, S. Tenbohlen, A. A. S. Akmal, and H. Mohseni, "Evaluation of numerical indices for the assessment of transformer frequency response," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 1, pp. 218–227, Jan. 2017.
[10] M. Mahvi, V. Behjat, and H. Mohseni, "Analysis and interpretation of power auto-transformer winding axial displacement and radial deformation using frequency response analysis," *Eng. Failure Anal.*, vol. 113, Jul. 2020, Art. no. 104549.
[11] K. Pourhossein, G. B. Gharehpetian, E. Rahimpour, and B. N. Araabi, "A probabilistic feature to determine type and extent of winding mechanical defects in power transformers," *Electr. Power Syst. Res.*, vol. 82, no. 1, pp. 1–10, Jan. 2012.
[12] A. J. Ghanizadeh and G. B. Gharehpetian, "Application of characteristic impedance and wavelet coherence technique to discriminate mechanical defects of transformer winding," *Electr. Power Compon. Syst.*, vol. 41, no. 9, pp. 868–878, Jul. 2013.
[13] H. Rahbarimagham, S. Esmaeili, and G. B. Gharehpetian, "Discrimination between radial deformation and axial displacement in power transformers using analysis of electromagnetic waves," *IEEE Sensors J.*, vol. 17, no. 16, pp. 5324–5331, Aug. 2017.
[14] K. Pourhossein, G. B. Gharehpetian, E. Rahimpour, and B. N. Araabi, "A vector-based approach to discriminate radial deformation and axial displacement of transformer winding and determine defect extent," *Electr. Power Compon. Syst.*, vol. 40, no. 6, pp. 597–612, Mar. 2012.
[15] P. Karimifard, G. B. Gharehpetian, A. J. Ghanizadeh, and S. Tenbohlen, "Estimation of simulated transfer function to discriminate axial displacement and radial deformation of transformer winding," *COMPEL-Int. J. Comput. Math. Electr. Electron. Eng.*, vol. 31, no. 4, pp. 1277–1292, Jul. 2012.
[16] E. Rahimpour, M. Jabbari, and S. Tenbohlen, "Mathematical comparison methods to assess transfer functions of transformers to detect different types of mechanical faults," *IEEE Trans. Power Del.*, vol. 25, no. 4, pp. 2544–2555, Oct. 2010.
[17] A. R. Abbasi, M. R. Mahmoudi, and Z. Avazzadeh, "Diagnosis and clustering of power transformer winding fault types by cross-correlation and clustering analysis of FRA results," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 19, pp. 4301–4309, Oct. 2018.
[18] M. Bigdeli, D. Azizian, and G. B. Gharehpetian, "Detection of probability of occurrence, type and severity of faults in transformer using frequency response analysis based numerical indices," *Measurement*, vol. 168, Jan. 2021, Art. no. 108322.
[19] M. Bigdeli, M. Vakilian, and E. Rahimpour, "A probabilistic neural network classifier-based method for transformer winding fault identification through its transfer function measurement," *Int. Trans. Electr. Energy Syst.*, vol. 23, no. 3, pp. 392–404, Apr. 2013.
[20] M. Bigdeli, M. Vakilian, and E. Rahimpour, "Transformer winding faults classification based on transfer function analysis by support vector machine," *IET Electr. Power Appl.*, vol. 6, no. 5, p. 268, 2012.
[21] S. Bagheri, Z. Moravej, and G. B. Gharehpetian, "Classification and discrimination among winding mechanical defects, internal and external electrical faults, and inrush current of transformer," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 484–493, Feb. 2018.
[22] Z. Zhao, C. Yao, C. Tang, C. Li, F. Yan, and S. Islam, "Diagnosing transformer winding deformation faults based on the analysis of binary image obtained from FRA signature," *IEEE Access*, vol. 7, pp. 40463–40474, 2019.
[23] J. Liu, Z. Zhao, C. Tang, C. Yao, C. Li, and S. Islam, "Classifying transformer winding deformation fault types and degrees using FRA based on support vector machine," *IEEE Access*, vol. 7, pp. 112494–112504, 2019.
[24] H. Tarimoradi and G. B. Gharehpetian, "Novel calculation method of indices to improve classification of transformer winding fault type, location, and extent," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1531–1540, Aug. 2017.
[25] A. J. Ghanizadeh and G. B. Gharehpetian, "ANN and cross-correlation based features for discrimination between electrical and mechanical defects and their localization in transformer winding," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 21, no. 5, pp. 2374–2382, Oct. 2014.
[26] S. Haykin, *Neural Networks and Learning Machines*, 2rd ed. London, U.K.: Pearson, 2009.
[27] V. N. Vapnik, *Statistical Learning Theory*. Hoboken, NJ, USA: Wiley, 1998.
[28] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 660–674, May 1991.
[29] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, "Top 10 algorithms in data mining," *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, 2008.

MEHDI BIGDELI received the B.Sc. and M.Sc. (Hons.) degrees in electrical engineering from the Iran University of Science and Technology, and Zanjan University, in 2004 and 2006, respectively, and the Ph.D. (Hons.) degree in electrical engineering from the Islamic Azad University, Science and Research Branch, Tehran, Iran, in 2012.

From 2005 to 2006, he was affiliated as a M.Sc. scholar with the Iran Transformer Research Institute (ITRI). In 2006, he joined Islamic Azad University, Zanjan Branch, where he is currently an Assistant Professor. Since 2019, he has also been working with Zanjan Regional Electric Company and Iran Transfo Corporation as a Consultant and Project Manager. His main research interests include the fault detection, transient modeling and optimization of electrical systems especially transformers, renewable energy, and artificial intelligence.

PIERLUIGI SIANO (Senior Member, IEEE) received the M.Sc. degree in electronic engineering and the Ph.D. degree in information and electrical engineering from the University of Salerno, Salerno, Italy, in 2001 and 2006, respectively. He is a Professor and Scientific Director of the Smart Grids and Smart Cities Laboratory with the Department of Management and Innovation Systems, University of Salerno. His research activities are centered on demand response, on energy management, on the integration of distributed energy resources in smart grids, on electricity markets, and on planning and management of power systems. In these research fields he has coauthored more than 500 articles including more than 300 international journal papers that received in Scopus more than 9450 citations with an H-index equal to 47. He received the award as 2019 highly cited researcher by ISI Web of Science Group. He has been the Chair of the IES TC on Smart Grids. He is Editor for the Power and Energy Society Section of IEEE Access, IEEE Transactions on Industrial Informatics, IEEE Transactions on Industrial Electronics, Open Journal of the IEEE IES and of *IET Renewable Power Generation*.

HASSAN HAES ALHELOU (Senior Member, IEEE) is currently a faculty member with Tisheen University, Lattakia, Syria.

He is included in the 2018 and 2019 Publons list of the top 1% best reviewer and researchers in the field of engineering. He was a recipient of the Outstanding Reviewer Award from *Energy Conversion and Management Journal* in 2016, *ISA Transactions Journal* in 2018, *Applied Energy Journal* in 2019, and many other Awards. He was a recipient of the best young researcher in the Arab Student Forum Creative among 61 researchers from 16 countries at Alexandria University, Egypt, 2011. He has published more than 100 research papers in the high quality peer-reviewed journals and international conferences. He has also performed reviews for high prestigious journals including IEEE Transactions on Industrial Informatics, IEEE Transactions on Industrial Electronics, Energy Conversion and Management, Applied Energy, *International Journal of Electrical Power and Energy Systems*. He has participated in more than 15 industrial projects. His major research interests are power systems, power system dynamics, power system operation and control, dynamic state estimation, frequency control, smart grids, micro-grids, demand response, load shedding, and power system protection.