# Self-adaptive Fault Diagnosis of Roller Bearings using Infrared Thermal Images

Zhiqiang Huo<sup>1,2</sup>, Yu Zhang<sup>1</sup>, Richard Sath<sup>3,1</sup>, Lei Shu<sup>2,1\*</sup>

<sup>1</sup>School of Engineering, University of Lincoln, Lincoln, UK

<sup>2</sup>Guangdong Provincial Key Laboratory on Petrochemical Equipment Fault Diagnosis,

Guangdong University of Petrochemical Technology, Maoming, China

<sup>3</sup>Mines Albi, University of Toulouse, Toulouse, France

Email: {zhuo, yzhang, rsath, lshu}@lincoln.ac.uk

Abstract-Fault diagnosis of roller bearings in rotating machinery is of great significance to identify latent abnormalities and failures in industrial plants. This paper presents a new selfadaptive fault diagnosis system for different conditions of roller bearings using InfraRed Thermography (IRT). In the first stage of the proposed system, 2-Dimensional Discrete Wavelet Transform (2D-DWT) and Shannon entropy are applied respectively to decompose images and seek for the desired decomposition level of the approximation coefficients. After that, the histograms of selected coefficients are used as an input of the feature space selection method by using Genetic Algorithm (GA) and Nearest Neighbor (NN), for the purpose of selecting two salient features that can achieve the highest classification accuracy. The results have demonstrated that the proposed scheme can be employed effectively as an intelligent system for bearing fault diagnosis in rotating machinery.

*Index Terms*—Fault Diagnosis, Roller bearing, Image processing, Infrared thermal image.

# I. INTRODUCTION

In recent decades, condition monitoring and fault diagnosis of mechanical faults in rotating machinery has received considerable attentions aiming at enhancing the confidence of the reliability and safety of rotating machinery. With the advent of smart sensors, a variety of powerful condition monitoring techniques have been developed for the purpose of providing intrinsic signatures measured from run-time machines, such as vibration monitoring, acoustic emission, and oil analysis [1], [2]. On the basis of useful intrinsic machine's signals increasing advances in fault diagnosis techniques have paved the way for minimizing performance degradation and avoiding dangerous situations in rotating machinery in real-time [3]. Nevertheless, vibration analysis or acoustic emission can be intrusive as sensors need to be mounted on, or within the machine. Oil analysis could detect debris in the lubrication system for defect diagnosis, but it is usually performed offline [4].

InfraRed Thermography (IRT), as an alternative technique to above mentioned methods, enables non-contact, non-intrusive, fine-grained, and single-sensor based temperature measurements [5], where the infrared radiation (wavelength lies between 0.75 and 1000 lm) emitted from the surface of the object

Corresponding author: Lei Shu (email: lshu@lincoln.ac.uk)

under inspection is recorded using an infrared camera. Those advantages of IRT make it suitable as an ideal solution to condition monitoring with the aim of autonomously diagnosing faults, which has attracted a great deal of attentions by detecting temperature anomaly caused by heat, friction, and excessive vibration, etc [6]. In order to monitor the conditions as well as diagnose the machine faults, signal processing and feature extraction are two of important stages in the field of fault diagnosis [7]–[9]. The 2-Dimensional Discrete Wavelet Transform (2D-DWT) enjoys the advantages of good time and frequency resolution, which has been widely applied in IRTbased fault diagnosis of rotating machinery [10], [11]. In addition, time and frequency domain features as well as non-linear features are popularly applied to discriminate fault information in thermal images acquired from machine [12]-[14]. Nevertheless, in most of those works, desired decomposition level, approximation coefficients, and features were selected by the experience of authors. Specifically speaking, the effectiveness and performance of those diagnosis methods developed by authors may be limited by the specific experiments, predetermined parameters adopted in 2D-DWT decomposition, and given features applied for discriminating fault information in different conditions of rotating components.

This paper attempts to develop a self-adaptive system using IRT for bearing fault diagnosis with the purpose of providing desired decomposition parameters and salient fault features for fault classification. In the proposed method, thermal images measured from a test rig are first decomposed by using 2D-DWT, after which the histogram of approximation coefficients in five levels obtained from training images are applied to calculate mean values of Shannon entropy. Finally, the coefficients in given level having the largest mean entropy value is selected since it may contain the most prominent fault information; usually, the considered candidates have more dominant characteristics, thus convey useful information about the status of a system that can be applied to precisely identify fault patterns hidden in the machines conditions. Subsequently, Genetic Algorithm (GA) minimization method is adopted to automatically select two dominant features that can achieve best fault classification accuracy combining with Nearest Neighbor (NN) classifier, in which Mahalanobis distance is

	Younus et al. [10]	Taheri et al. [11]	Janssens et al. [12]	Duan et al. [13]	Glowacz et al. [14]
Year	2012	2015	2015	2016	2016
Objects considered	bearing fault, shaft misalignment, and unbalance	six types of radiator condition	bearing fault, rotor imbalance, and lubrication condition	rotor radial rubbing, shaft unbalance and misalignment, and base looseness	induction motor fault
Image preprocessing	2D-DWT (bior-3.5 and three levels)	2D-DWT ( $1^{st}$ level and four coefficients)	sub-sampling and windowing	image segmentation	binarization
Resolution of ROI (pixels)	158 × 25	N/A	640 × 480	N/A	320 × 240
Features used	M, SK, EN, KU, MA	M, SK, EN, STD, Smoothness, Energy	M of STD, M of Gini coefficient, M of $M_{20}$	M, SK, EN, KU, STD, Energy	Sum of pixel values
Classifiers applied	SVM and LDA	GA and ANN	Random decision forest classifier	SVM and Naïve bayes	NN and GMM

TABLE I: Comparison of related work for fault diagnosis based on IRT.

<sup>1</sup> ROI: Region of Interest, M: Mean, SK: Skewness, EN: Entropy, KU: Kurtosis, MA: Mean absolute deviation, STD: Standard deviation

applied for evaluating the similarity distance. Finally, 2 Dimensional (2D) feature space of thermal images is presented through applying this proposed self-adaptive fault diagnosis method. The main contributions in this paper are concluded as follows:

- The wavelet decomposition level and approximation coefficients are selected adaptively by evaluating the complexity and irregularity in the histograms of coefficients obtained from training images using 2D-DWT and Shannon entropy analysis. After that, the histograms of selected coefficients in given level having the largest mean value are finally selected as the desired input for feature extraction.
- Two salient features are selected by evaluating the classification success rate based on GA minimization technique along with NN classifier. Finally, the selected two features are used to discriminate fault characteristics in testing phrase and constructing 2D feature space.

The rest of this paper is organized as follows: Section II reviewed the recent works related to fault diagnosis based on IRT. Section III presents the theoretical framework of the proposed self-adaptive fault diagnosis method. Experimental validation and results are given in Section IV. Finally, conclusions are drawn in Section V.

# II. RELATED WORK

Advances in IRT have paved the way for the emergence of fruitful techniques which have been widely applied in the field of fault diagnosis. In the past decades, great efforts have been devoted to fault diagnosis of bearings in rotating machinery based on IRT. Table I gives comparative perspectives in the field of bearing fault diagnosis using IRT. As an example, Younus et al. [10] developed a rotary machine fault diagnosis method based on IRT, in which both of the coefficients obtained from 2D-DWT in three levels were applied to extract features. Time-domain and non-linear features were adopted for directly discriminating fault information from coefficients, before which the dimensionality and quality of features were determined by using statistical similarity analysis based on Mahalanobis distance and relief algorithm. Finally, the Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) were applied respectively for fault classification. Nevertheless, in their proposed method, the use of all coefficients in three levels applied can be very time-consuming. Moreover, features selected using relief algorithm based on the analysis of weights are greater than the thresholds, which probably could be affected by the selection of predetermined threshold in different experiments. Furthermore, Taheri et al. [11] proposed a diagnosis method by using 2D-DWT, GA and Artificial Neural Networks (ANNs), in which four coefficients in the first level after 2D-DWT were used, and six features were then applied to extract fault feature vectors from those coefficients. The use of GA and ANNs is for quantifying a subset of the relevant features which can achieve best classification accuracy. In addition, Janssens et al. [12] presented a novel automatic fault detection system using IRT where windowing was applied for the purpose of sub-sampling in time from a video, and three features were extracted from the preprocessed frames, e.g., characteristics in pixel values and temperature distribution. Additionally, for the purpose of reducing redundant information and minimizing the influence of noise in infrared images, Duan et al. [13] presented an image segmentation approach to enhance the feature extraction by selecting effective Region of Interest (RoI) which can discriminate fault representative regions from unrelated background information. Furthermore, Glowacz et al. [14] developed a motor fault diagnosis technique based on IRT, in which original infrared images were first transformed to preprocessed images with magenta color, the binary images of which were then calculated. After that, the sum of pixel values were used as feature vectors for fault recognition.

To sum up, IRT has been widely applied for fault diagnosis by discriminating characteristics in infrared thermal images with the use of time and frequency domain features. In

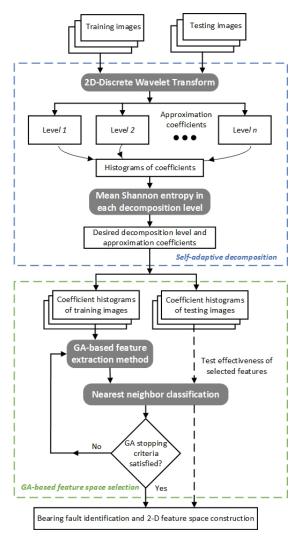


Fig. 1: Flow chart of the proposed approach.

practice, after the 2D-DWT, desired decomposition level and coefficients are generally selected by using the experience of authors. Furthermore, in most works, features were not selected adaptively and automatically by means of proving their effectiveness in classification success rate. Consequently, there is still strong need to provide effective approaches that can provide self-adaptive methods in both signal decomposition stage and feature extraction stage. For this purpose, this paper presents a self-adaptive technique which automatically select desired decomposition level and approximation coefficients in 2D-DWT, and two discriminatory features from 10 candidates in feature extraction phrase.

#### III. PROPOSED METHOD

This section briefly introduces the proposed method. Fig. 1 presents the procedure for bearing fault classification using the proposed approach based on 2D-DWT, Shannon entropy, GA, and NN techniques. The theories behind the proposed method are briefly reviewed as following:

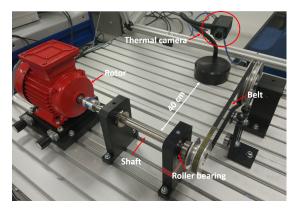


Fig. 2: PT 500 bearing fault kit and FILR thermal camera.

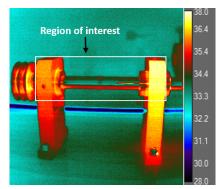
## A. Selection of wavelet coefficient

1) Discrete Wavelet Decomposition: The DWT plays a powerful mathematical role in the field of 2D signal processing in terms of image compression and denoising, which can provide multi-resolution analysis in fine-grained time and frequency domains. The 2D-DWT can be regarded as a chain of successive levels of decomposition [15], in which image is decomposed by applying 1D-DWT in both the horizontal and vertical directions of the image pixels of given level based on high pass filter and the cascade of sub-sampled low pass filter. The former gives the smoothing effect known as approximation coefficients, while the later gives the detailed coefficients. To achieve a better resolution, the process of decomposing the sequence into sub-sequences with half resolution is generally iterated on the lower band. The output by performing 2D-DWT decomposition, four types of wavelet coefficients, could be obtained from each infrared thermal image in each level, namely approximation coefficients, horizontal, vertical and diagonal detailed coefficients. For brevity, a more detailed explanation about 2D-DWT can be found in [15].

2) Shannon Entropy: Shannon entropy [16] provides a practical criteria for analyzing and estimating the similarity between distributions of probability within a time series. Given a time series, S with length N, Shannon entropy can be expressed below:

$$ShanEn = -\sum_{i=1}^{N} P_i \log_2 P_i \tag{1}$$

where *i* is the index of data points in a time series, *N* is the number of data points contained in a time series, and  $P_i$  is the distribution of the energy probability for each time series with  $\sum_{i=1}^{N} P_i = 1$ . Herein, the histograms of approximation coefficients were applied as input to feature extraction. Hence,  $P_i$  represents the frequency distribution probability in a histogram.  $P_i = E_i/E_{total}, i = 1, 2, \dots, N$ , in which  $E_i$  represents the frequency of a coefficient in the histogram, and  $E_{total}$  is the summation of frequencies for the overall *N* coefficients in the histogram. In this paper, in total 100 bins are applied in the calculation of the histogram of approximation coefficients. Hence, the mean Shannon entropy



(a) Original thermal image and region of interest

Fig. 3: Original thermal image and histogram for bearing A (healthy condition) after 30 minutes.

value can be obtained by averaging the summation of Shannon entropy values obtained from histograms of approximation coefficients in five levels respectively.

# B. selection of features in 2-D feature space

1) Nearest Neighbor Classifier: Taking different conditions of roller bearings into consideration, the classification of this kind of fault is multi-class classification problem. NN algorithm is one of the most fundamental and commonly used methods for classification [17], which enables to consistently achieve high performance. The basic idea behind NN method is that a new sample can be classified by calculating the similarity distance between testing data sets and condition groups obtained from training data sets, after which the group label of testing samples can be determined using a given condition group having the nearest similarity distance between this undetermined sample. In this paper, Mahalanobis distance is applied as an alternative method to Euclidean and other similarity distances, since Mahalanobis is more unitless and scale-invariant because of taking the correlations of data sets into account [18].

2) Genetic Algorithm: Feature extraction in the field of fault diagnosis is of great importance to discriminate fault information hidden in complex signals with the aim at leading to higher learning accuracy for classification, lower computational cost, and better model interpretability [19], [20]. By means of applying time and frequency domains analysis or non-linear analysis, a large number of features can be used generally to extract features from time and frequency domains. Nevertheless, the use and application of features are needed to be selected with high efficiency rather than selected from the experience of authors. In this paper, for the purpose of drawing 2D feature space towards discriminating different conditions of bearings, a hybrid method is used for image pattern recognition based on GA and NN techniques by means of selecting the salient two features that can achieve the best classification accuracy. Herein, GA is used to select the two desired genes (features) and NN is applied to evaluate the classification effectiveness using given features and returned a minus success rate, which was used as a fitness function for selected subset

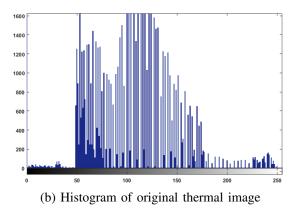


TABLE II: Details for experimental setup.

Equipment	Parameters			
	• Uncooled VOX microbolometer/7.5-13 $\mu$ m			
Thermal camera	$\bullet$ -40 °C to 550 °C temperature range			
(FLIR A35)	• 0.05 °C at °C thermal sensitivity			
	• 60 Hz image frequency			
	• 320 × 256 pixels			
	• Region of interest: $148 \times 40$ pixels			
	• Roller bearing faults kit			
PT500 test rig	• 2000 r.p.m rotating speed			
1 1200 test lig	• Roller bearing width: 14 mm			
	• Roller diameter: 7.5 mm			
	• Bearing types: healthy bearing (A), bearing with damage to outer ring (B), bearing with damage to inner ring (C), bearing with damage to rolling element (D), bearing with damage from B,C,D simultaneously (E), and bearing with severe wear (F)			

of genes. After several rounds of selection, recombination and mutation, the best individuals of a population are selected to reproduce and go into next generation. Finally, the two desired features achieving the highest classification accuracy can be obtained from the proposed self-adaptive system.

# IV. EXPERIMENTAL STUDY

In this section, a test rig using the PT 500 bearing fault kit and the FLIR thermal camera is first briefly introduced, and experimental results of the proposed method are then presented and discussed.

### A. Experiment Setup

In this study, PT 500 bearing fault kit is used as a supply of thermal images, and in total six different conditions of bearings are applied [21], as shown in Fig. 2. For each bearing condition, the speed of 2000 r.p.m is selected and maintained in the first 30 minutes until machine reaches its stable condition. After that, the process of image acquisition is taken

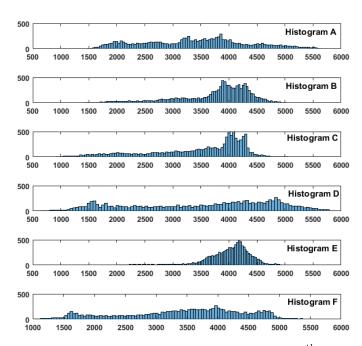


Fig. 4: Histogram of approximation coefficients in  $5^{th}$  level from bearing A to F.

by using FLIR camera. The detail parameters of experimental equipment are introduced in Table II. The experimental data sets contain six conditions of bearings. For each bearing, 20 thermal images are measured, and therefore the total number of data sets include 120 images. The entire data sets are split into two categories, namely training and testing data sets respectively. Fig. 3 illustrates one original thermal image of bearing A condition and histogram of this image.

#### **B.** Experimental Results

Thermal images measured from rotating machinery were first decomposed by using 2D-DWT. Approximation coefficients were then evaluated based on Shannon entropy to select the desired coefficients as the input into feature extraction stage. In this study, the mother wavelet "dmey" and in total five decomposition levels were adopted in 2D-DWT method. Indeed, among the coefficients obtained from 2D-DWT (approximation and detail coefficients), approximation coefficients which passed through the low pass filter were considered since the low frequency signals may contain most prominent information in original images. Hence, there are five approximation coefficients can be finally obtained from one image; therefore, the histogram of each coefficient was obtained. After that, the mean entropy value of histogram of approximation coefficients obtained from all training images were calculated by using Shannon entropy. In this study, the  $5^{th}$  level was finally obtained since it obtained the highest average values of Shannon entropy in the total five levels. Typical histograms of approximation coefficients in  $5^{th}$  using "dmey" wavelet are illustrated in Fig. 4. As can be seen, the occurrence frequencies in the histogram of coefficient representing the condition of bearing A are uniformly distributed,

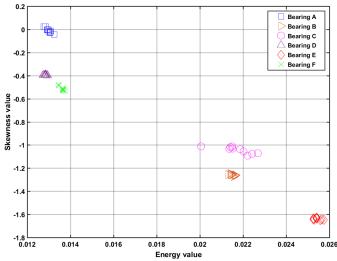


Fig. 5: 2D Feature space of six conditions of bearings based on skewness and energy features.

and the peak values are no more than 300. Furthermore, for bearing A, D and F, they roughly have the similar changing trend. That is, for those three bearings, they almost have the larger range along the X axis. For bearing B, C, and E, their peak values intensively locate among 4000 and 4500 in the X axis. This phenomenon can be found in the Fig. 5 where testing samples separately disperse in general two ranges, namely in the upper left corner and in the lower right corner. Interestingly, multiple faulty bearing E has the most narrow changing range in contrast with others' conditions.

In this work, after 2D-DWT decomposition, two salient features were selected by using GA and NN methods. With the purpose of selecting desired features, there are in total ten kinds of features including time-domain features and entropy features applied in the stage of feature selection. That is, mean value, standard deviation, kurtosis, skewness, root mean square, peak value, crest factor, clearance factor, impulse factor, and energy. In image processing, energy descriptor provides a measure of how the pixel values are distributed along the gray-level range. Usually, the image with few gray levels will have higher energy than the others with many gray levels. Herein, the energy represents the distribution of occurrence frequency of coefficient values in histogram of coefficient. By applying NN methods, the efficiency of two given features adopted for feature extraction for six conditions of bearings can be obtained. The use of GA minimization method helps to select two features which can achieve the highest classification accuracy by using iterative processes in which the best two features of a population are selected to reproduce and pass to the next generation. Therefore, in order to select two desired features and reduce dimensions, feature selection was performed by combining GA and NN techniques. The proposed procedure for feature space selection is presented in Fig. 1. Consequently, skewness and energy features finally outperformed the others, which can achieve

100% classification accuracy for classifying six conditions of bearings using IRT in this study. The 2D feature space is shown in Fig. 5. As can be seen, testing data sets are separately distributed in different ranges. The healthy conditions of bearings (represented by square) intensively locate in the upper left corner. Interestingly, the multi-fault conditions of bearings (represented by diamond) locate in the lower right corner. In addition, it can be notice that the outer race fault and inner race fault conditions of bearings locate closely within a short distance from each other. In conclusion, dominant and discriminatory fault characteristics in six conditions of bearings can be extracted from histograms of approximation coefficients and two salient features obtained by using the proposed system.

## V. CONCLUSIONS

This paper presents a new self-adaptive method for bearing fault diagnosis in rotating machinery using InfraRed Thermography (IRT). Six conditions of bearings were applied including healthy bearing, outer race damage, inner race damage, roller element damage, multi-fault damage, and worn damage to bearings. The purpose of the proposed system is to provide an automatic method that can select desired wavelet decomposition level and approximation coefficients using entropy analysis based on the use of histogram of approximation coefficients. In addition, two salient features were adaptively chosen by using Genetic Algorithm (GA) and Nearest Neighbor (NN) techniques from ten types of timedomain and non-linear features, which can achieve the highest classification accuracy. The proposed method was performed using a test rig, the results of which demonstrated that the classification success rate can achieve 100% by using the "dmey" as the mother wavelet, approximation in the  $5^{th}$ decomposition level, and skewness and image energy as fault features. Finally, the 2D feature space was constructed in which six conditions can be seen separately located in different regions.

#### ACKNOWLEDGEMENT

This work is partially supported by International and Hong Kong, Macao & Taiwan collaborative innovation platform and major international cooperation projects of colleges in Guangdong Province (No.2015KGJHZ026), The Natural Science Foundation of Guangdong Province (No.2016A030307029), and Maoming Engineering Research Center on Industrial Internet of Things (No.517018).

#### REFERENCES

- P. Tavner, "Review of condition monitoring of rotating electrical machines," *IET Electric Power Applications*, vol. 2, no. 4, pp. 215–247, 2008.
- [2] D. Rossetti, Y. Zhang, S. Squartini, and S. Collura, "Classification of bearing faults through time-frequency analysis and image processing," in *Mechatronics-Mechatronika (ME), 2016 17th International Conference* on. IEEE, 2016, pp. 1–7.
- [3] X. Dai and Z. Gao, "From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2226–2238, 2013.

- [4] I. El-Thalji and E. Jantunen, "A summary of fault modelling and predictive health monitoring of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 60, pp. 252–272, 2015.
- [5] S. Bagavathiappan, B. Lahiri, T. Saravanan, J. Philip, and T. Jayakumar, "Infrared thermography for condition monitoring-a review," *Infrared Physics & Technology*, vol. 60, pp. 35–55, 2013.
- [6] A. N. Huda and S. Taib, "Application of infrared thermography for predictive/preventive maintenance of thermal defect in electrical equipment," *Applied Thermal Engineering*, vol. 61, no. 2, pp. 220–227, 2013.
- [7] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," *Data Classification: Algorithms and Applications*, p. 37, 2014.
- [8] A. Rai and S. Upadhyay, "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings," *Tribology International*, vol. 96, pp. 289–306, 2016.
- [9] Z. Huo, Y. Zhang, P. Francq, L. Shu, and J. Huang, "Incipient fault diagnosis of roller bearing using optimized wavelet transform based multi-speed vibration signatures," *IEEE Access*, 2017.
  [10] A. M. Younus and B.-S. Yang, "Intelligent fault diagnosis of rotating
- [10] A. M. Younus and B.-S. Yang, "Intelligent fault diagnosis of rotating machinery using infrared thermal image," *Expert Systems with Applications*, vol. 39, no. 2, pp. 2082–2091, 2012.
- [11] A. Taheri-Garavand, H. Ahmadi, M. Omid, S. S. Mohtasebi, K. Mollazade, A. J. R. Smith, and G. M. Carlomagno, "An intelligent approach for cooling radiator fault diagnosis based on infrared thermal image processing technique," *Applied Thermal Engineering*, vol. 87, pp. 434– 443, 2015.
- [12] O. Janssens, R. Schulz, V. Slavkovikj, K. Stockman, M. Loccufier, R. Van de Walle, and S. Van Hoecke, "Thermal image based fault diagnosis for rotating machinery," *Infrared Physics & Technology*, vol. 73, pp. 78–87, 2015.
- [13] L. Duan, M. Yao, J. Wang, T. Bai, and L. Zhang, "Segmented infrared image analysis for rotating machinery fault diagnosis," *Infrared Physics* & *Technology*, vol. 77, pp. 267–276, 2016.
- [14] A. Glowacz and Z. Glowacz, "Diagnostics of stator faults of the singlephase induction motor using thermal images, moasos and selected classifiers," *Measurement*, vol. 93, pp. 86–93, 2016.
- [15] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
- [16] C. Shannon, "A mathematical theory of communication, bell system technical journal 27: 379-423 and 623–656," *Mathematical Reviews* (*MathSciNet*): MR10, 133e, 1948.
- [17] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [18] R. De Maesschalck, D. Jouan-Rimbaud, and D. L. Massart, "The mahalanobis distance," *Chemometrics and Intelligent Laboratory Systems*, vol. 50, no. 1, pp. 1–18, 2000.
- [19] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [20] Y. Zhang, C. Bingham, M. Martínez-García, and D. Cox, "Detection of emerging faults on industrial gas turbines using extended gaussian mixture models," *International Journal of Rotating Machinery*, vol. 2017, 2017.
- [21] "PT 500 machinery diagnostic system," (Date last accessed 31-July-2017). [Online]. Available: www.gunt.de/static/s3680\_1.php