

# Improved Bearing Fault Diagnosis by Feature Extraction Based on GLCM, Fusion of Selection Methods, and Multiclass-Naïve Bayes Classification

Mireille Pouyap<sup>1</sup>, Laurent Bitjoka<sup>1</sup>, Etienne Mfoumou<sup>2</sup>, Denis Toko<sup>3</sup>

<sup>1</sup>Energy, Signal, Imaging and Automation Laboratory (LESIA), University of Ngaoundere, Yaounde, Cameroon

<sup>2</sup>Division of Applied Research & Innovation, Nova Scotia Community College, Dartmouth, Canada

<sup>3</sup>Department of Mechanical Engineering, College of Technology, University of Buea, Buea, Cameroon

Email: pouyapmireille@gmail.com

**How to cite this paper:** Pouyap, M., Bitjoka, L., Mfoumou, E. and Toko, D. (2021) Improved Bearing Fault Diagnosis by Feature Extraction Based on GLCM, Fusion of Selection Methods, and Multiclass-Naïve Bayes Classification. *Journal of Signal and Information Processing*, 12, 71-85.

<https://doi.org/10.4236/jsip.2021.124004>

**Received:** August 24, 2021

**Accepted:** October 15, 2021

**Published:** October 18, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

The presence of bearing faults reduces the efficiency of rotating machines and thus increases energy consumption or even the total stoppage of the machine. It becomes essential to correctly diagnose the fault caused by the bearing. Hence the importance of determining an effective features extraction method that best describes the fault. The vision of this paper is to merge the features selection methods in order to define the most relevant features in the texture of the vibration signal images. In this study, the Gray Level Co-occurrence Matrix (GLCM) in texture analysis is applied on the vibration signal represented in images. Features selection based on the merge of PCA (Principal component Analysis) method and SFE (Sequential Features Extraction) method is done to obtain the most relevant features. The multiclass-Naïve Bayes classifier is used to test the proposed approach. The success rate of this classification is 98.27%. The relevant features obtained give promising results and are more efficient than the methods observed in the literature.

## Keywords

GLCM, PCA, SFE, Naïve Bayes, Relevant Features

---

## 1. Introduction

In industrial automation systems of recent years, machine movement is usually provided by rotational force. Bearings are a commonly used mechanical component in motor systems that perform this rotational motion and are used to

reduce friction. Early detection and diagnosis of rotating machinery, deteriorating condition, low efficiency and prevention of unexpected failures are becoming increasingly important in these systems. The main reasons for rotating machine failure are usually due to bearing faults. For example, metal bearing failures in asynchronous motors constitute 40% - 50% of system faults [1]. Therefore, several techniques have been developed for monitoring the condition of bearings to avoid such failures at an early stage. Apart from these techniques, fault analysis based on vibration signals has proved to be more advantageous in revealing bearing failure. In addition, it is impossible to avoid wear due to constant friction of mechanical components [2]. For this reason, condition monitoring based on bearing diagnostics should be applied to rotating machines in automation systems [3]. When the current literature is reviewed, methods based on vibration analysis and current analysis can be considered as the most applied fault monitoring methods. The data obtained in these studies are analyzed by methods such as time [4], frequency [5], and time-frequency [6] analysis and then supported by methods such as artificial intelligence techniques [7] [8].

Fault type identification and recognition uses the detection events as the start of the fault classification process in the monitored system. The vibration signal analysis method is widely used in the fault diagnosis of rotating machines, as an abnormal condition occurs when the vibration of the signal changes [9]. Vibration signal analysis requires attribute extraction to obtain an accurate diagnosis [10]. Several studies have been done on attribute extraction based on signal decomposition [11]. The vibration signal of defective bearings is usually very random, with strong interference and obvious irregularity. Thus, in practical engineering applications, it is not easy to classify the time-frequency images using conventional image recognition methods, such as syntactic recognition, two-dimensional linear discrimination and geometric transformation method, etc. The feature extraction step is performed by a computer programmer. The features extraction step is the most crucial part of the bearing fault diagnosis. In order to correctly diagnose the defect caused by the bearing, it is necessary to determine an efficient features extraction method that best describes the defect. Several features extraction methods are used in the literature among which the scale invariant feature transform (SIFT) is mainly used for its good robustness and high accuracy [12]. On the other hand, it has high time complexity and computation time requirements. Recently, features from gray level cooccurrence matrix (GLCM) have proven to be effective in a wide range of applications such as tumor classification in medical image analysis [13], texture analysis of bearing defect images [14] [15]. However, the feature extraction and selection step are known to be the most critical and difficult.

The originality of this work lies in the selection of the relevant features by merge of the PCA (Principal component Analysis) method and the SFE (Sequential Features Extraction) method to obtain the most relevant features. The most important advantages of the proposed methodology are: the application of GLCM on images representation obtained directly from the temporal vibration

signal, the selection of the relevant features by merge of the selection methods and finally the validation of the method by classification of the different classes of bearing faults.

The sections are organized as follows: first, an introduction giving bibliographical information on the subject of the study and general information on the classification of bearing vibration signals is given. Secondly, a description of the dataset used and the attribute extraction and selection approach are presented. In the third section, the obtained results are detailed and discussed. Finally, the fourth section concludes the work.

## 2. Tools Used

### 2.1. GLCM

The GLCM quantify the spatial relation of neighboring pixels in an image. It's a comprehensive information of the image grayscale with regard to: the direction, the neighboring interval and the rangeability [16]. In simple terms, each element  $X_{d,\theta}(i, j)$  of the co-occurrence matrix represent the probability of occurrence to have the grayscale  $j$  and the grayscale  $i$ , at  $d$ -spatial distance and  $\theta$ -orientation. Usually, orientation is chosen among four directions namely, horizontal, left diagonal, vertical, and right diagonal, respectively:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$ . The spatial distance  $d$  belongs to the set of positive number and is usually one. Thus, for  $(d, \theta)$  fixed, GLCM is a matrix whose elements  $X(i, j)$  are obtained for  $(i, j) \in (N_G, N_G)$ , where  $N_G$  is the number of grayscale of image [17]. **Table 1** present the 14 features of GLCM defined by Haralick in 1973, the 5 features of GLCM defined by Soh in 1999 and the feature of GLCM defined by Clausi in 2002.

**Table 1.** Features computed from GLCM.

Haralick features [16]			
No.	Feature name	Notation used	Formulation
1	Contrast	CONTRA	$\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} (i-j)^2 \cdot P(i, j)$
2	Correlation	CORRE	$\frac{\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} (i-\mu_x)(j-\mu_y) \cdot P(i, j)}{\sigma_x \sigma_y}$
3	Energy	ENERG	$\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} [P(i, j)]^2$
4	Homogeneity	HOMOG	$\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} \frac{P(i, j)}{1+(i-j)^2}$
5	Sum of square: variance	SUMOF	$\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} (i-\mu)^2 \cdot P(i, j)$
6	Entropy	ENTRO	$-\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} P(i, j) \cdot \log[P(i, j)]$

Continued

7	Sum average	SUMAV	$\sum_{k=2}^{2N_G} k \cdot P_{x+y}(k)$
8	Sum entropy	SUMEN	$-\sum_{k=2}^{2N_G} P_{x+y}(k) \cdot \log[P_{x+y}(k)]$
9	Sum variance	SUMVA	$\sum_{k=2}^{2N_G} (k - \mu_{x+y})^2 \cdot P_{x+y}(k)$
10	Difference variance	DIFFVA	$\sum_{k=0}^{N_G-1} (k - \mu_{x-y})^2 \cdot P_{x-y}(k)$
11	Difference entropy	DIFFEN	$-\sum_{k=0}^{N_G-1} P_{x-y}(k) \cdot \log[P_{x-y}(k)]$
12	Information measure of correlation 1	INFO1	$HXY - HXY1/Max(HX, HY)$
13	Information measure of correlation 2	INFO2	$[1 - \exp(-2 \cdot HXY2 + 2 \cdot HXY)]^{1/2}$
14	Maximum correlation	MAXCOR	[Second largest eigenvalue of $Q$ ] <sup>1/2</sup>
Sohfeatures [18]			
15	Autocorrelation	AUTO	$\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} i \cdot j \cdot P(i, j)$
16	Dissimilarity	DISSI	$\sum_{i=1}^{N_G} \sum_{j=1}^{N_G}  i - j  \cdot P(i, j)$
17	Maximum probability	MAXIP	$Max(P(i, j)) \forall (i, j) \in (N_G, N_G)$
18	Cluster shade	CLUSHA	$\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} (i + j - \mu_x - \mu_y)^3 \cdot P(i, j)$
19	Cluster prominence	CLUSPRO	$\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} (i + j - \mu_x - \mu_y)^4 \cdot P(i, j)$
Clausi features [19]			
20	Inverse difference	INVDIF	$\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} \frac{P(i, j)}{1 +  i - j }$

### 2.2. PCA

The PCA is used for the purpose of dimensionality reduction of the high-dimensional feature vector including the extracted texture features, due to the fact that the high-dimensional feature vector can degrade classification performance [20]. PCA algorithm select only the relevant principal components (linear transformation of the original features which are uncorrelated). Among these principal components, only the most significant are used to find GLCM features which are more correlated. Those features are named, relevant GLCM features and utilized as inputs to the classifier.

### 2.3. SFE

Sequential feature selection [21]. This method has two components. Firstly, an objective function, called the criterion, which the method seeks to minimize over all feasible feature subsets. The common criteria for classification models are misclassification rate. Secondly, a sequential search algorithm adds or removes features from a candidate subset while evaluating the criterion. Since an exhaustive comparison of the criterion value at all  $2^n$  subsets of an  $n$ -features data set is typically infeasible (depending on the size of  $n$  and the cost of objective calls), sequential searches move only in one direction, always growing or always shrinking the candidate set. During the process, SFE selects the best features among all features data which are able to discriminate each class from others.

### 2.4. Classification Using Naïve Bayes

Naive Bayesian is one of the classification methods using the similarity of the characteristics of an object. This method is classified as a fairly simple method, but is widely used in the fields of medicine, biometrics, text classification, and many more. Naive Bayesian uses the Gaussian distribution by considering (2) important parameters, namely the average  $\mu$  and the variance  $\sigma$  [22]. In Naive Bayesian, Gaussian uses equation:

$$P(X_i = x_i / Y = y_i) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left[-\frac{1}{2}\left(\frac{x_i - \mu_{ij}}{\sigma_{ij}}\right)^2\right] \quad (1)$$

where:

$P$  = Probability of attribute  $x_i$ .

$x_i$  = Attribute sought.

$i$  = Index for the value of the attribute

$j$  = class index.

$Y$  = Represent the class sought

$\mu$  = The average value represented.

For variance ( $\sigma$ ), find the Equation (2)

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2 \quad (2)$$

To classify using the Naive Bayesian method, it is necessary to calculate the average and standard deviation of each class for each characteristic. For the final stage, the test data is entered into each class to determine the opportunities that exist in each class so that it can be determined in which class the image has the greatest opportunity [22].

The diagnostic performance of the classifier can be evaluated by average classification accuracy (Acc1) which is calculated using (3). It's the classifier success rate, where  $N_{Tp}$  is the number of images in class  $c$  that are correctly classified as class  $c$ ,  $N_{images}$  is the total number of images for all classes combined, and  $N_{classes}$  is the number of fault types or classes,  $N_{C\_images}$  is the total number of images for class  $c$  [23]. For our study, we used two types of accuracy, namely: Acc1 and

Acc2. Where, Acc2 is classifier identification rate between normal and faults vibration signal.

$$\text{Acc1} = \frac{\sum N_{\text{classes}} N_{\text{TP}}}{N_{\text{images}}} \cdot 100; \quad \text{Acc2} = \frac{N_{\text{TP}}}{N_{\text{C\_images}}} \cdot 100 \quad (3)$$

The aim of recognition is first of all to know how to find the positive examples “true positives”; it is also necessary to try to limit the number of false alarms “false positives”; these are objects that the system takes for normal but which are not.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (4)$$

**Table 2** helps to illustrate the confusion matrix and Equation (4) is the recognition rate of all signals.

### 3. Material and Method

#### 3.1. Material

##### 3.1.1. Data Description

The proposed approach is tested on failure test data collected and made publicly available by Case Western Reserve University [24] Center. The data was collected using a 2 HP motor with a torque transducer and dynamometer to apply different loads. The data of drives and tests, defects are diameter ranging from 0.007 to 0.021 in were tested located, the different types of defects are defined on ball, inner race and outer race. In this study, the data collected on the driving end bearings were included in the analysis. The bearings are SKF rigid ball bearings: 6205-2RS JEM and 6203-2RS JEM. **Table 3** gives information on the selected experimental data. The vibrations were measured using accelerometers placed at orthogonal, centered and opposite on the bearing housing. The data was collected using a 16-channel encoder at a sampling rate of 12,000 Hz. It should be noted that there is a variation in shaft speed in these data sets, from 1722 to 1796 rpm.

**Table 2.** Confusion matrix.

		Realit	
		Normal signal	Faulty signal
Prediction	Normal signal	True positive (TP)	False negative (FN)
	Faulty signal	False positive (FP)	True negative (TN)

**Table 3.** Data description.

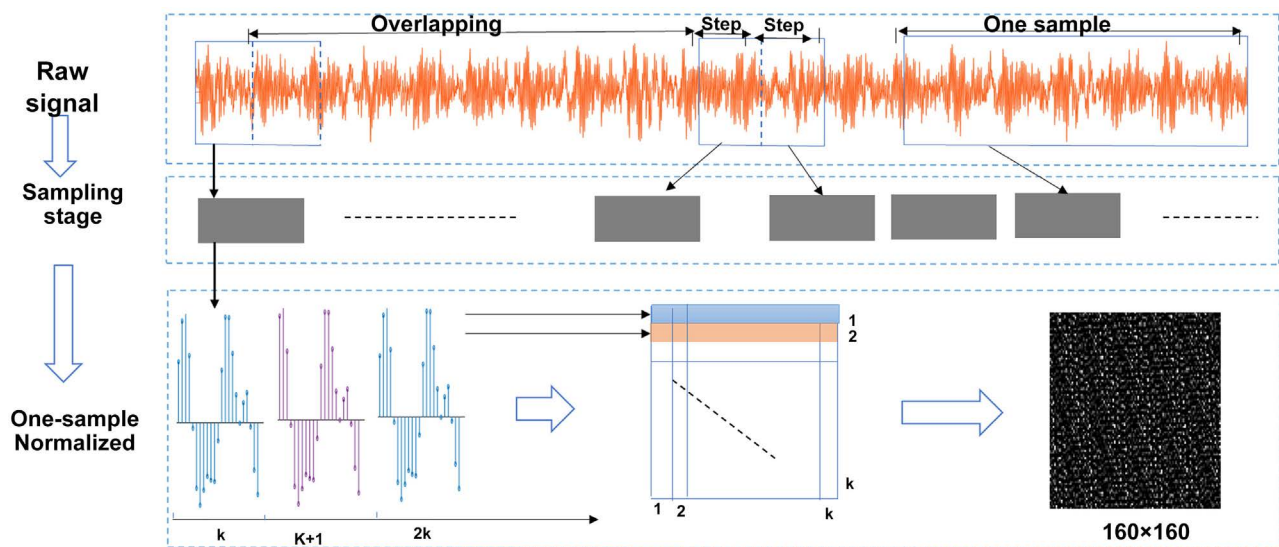
Class	Normal	0.014 ball	0.007 inner race	0.007 ball	0.007 outer race opposite	0.007 outer race centered	
Images number	60	60	60	60	60	60	
Class	0.007 outer race orthogonal	0.014 inner race	0.021 outer race opposite	0.021 outer race centered	0.021 outer race orthogonal		Total images
Images number	60	20	40	60	40		<b>580</b>

### 3.1.2. Converting the Signal into a Grayscale Image

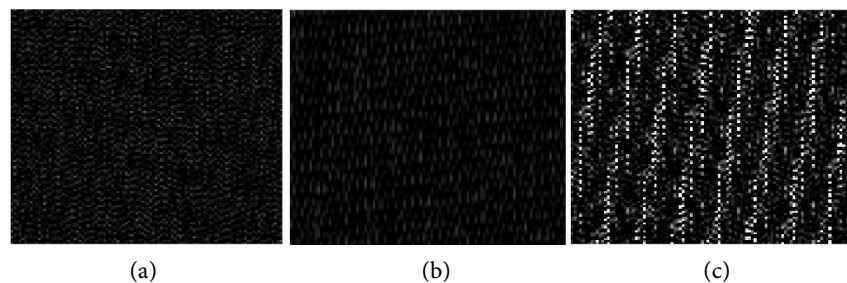
The process consists of converting the one-dimensional time signal into an image. This method allows us to explore the features in the two-dimensional domain of a signal. It should be noted that this method of data preprocessing can be archived without any predetermined parameters. **Figure 1** shows the process of converting the temporal signal into image. In this figure, the segmentation of the signal samples of size  $k^2$  is observed and arbitrarily extracted from the starting signal, the image obtained is of size  $K \times K$  by processing these samples. The intercepted signal segments are normalised from 0 to 255, which is the range of pixel intensity significant for a grayscale image. For this work, each data sample chosen for work has 25,600 points. The choice of  $160 \times 160$  in this paper is dependent on the volume of signal data.  $L(i) (i = 1, 2, \dots, K^2)$  denotes the value of the segment signal.  $P(j, k) (j = 1, 2, \dots, K; k = 1, 2, \dots, K)$  denotes the pixel strength of the image [25]. The process is described as:

$$P(j, k) = \frac{L(jK + k) - \text{Min } L}{\text{Max } L - \text{Min } L} \times 255 \quad (5)$$

**Figure 2** shows some images obtained after the conversion of 1D-vibration signal to grayscale image.



**Figure 1.** Signal-to-image conversion process [25].



**Figure 2.** Examples of 2D representation of vibration signals. (a) Normal; (b) 0.014 ball; (c) 0.021 outer race.

### 3.2. Proposed Method

Since the literature proposes a wide range of features that can be computed on the co-occurrence matrix, and if they are all used at the same time for classification, it is more likely that the recognition rates are not consistent due to the presence of some redundant features. Therefore, a selection step of the most relevant features is necessary. The PCA and the SFE are merged to define the most relevant features that will be used in the classification to improve the recognition rates. **Figure 3** represents the flowchart of the proposed methodology and the different steps of our work are organized as follows:

**Step 1:** Description of the vibration signal data of the bearings in normal and faulty conditions;

**Step 2:** The vibration signal can be split into random sub-samples, normalised and arranged in rows and columns to form a matrix; each matrix obtained is associated with a greyscale image of the vibration signal;

**Step 3:** The GLCM is calculated on each grey level image and its texture features are extracted on each GLCM;

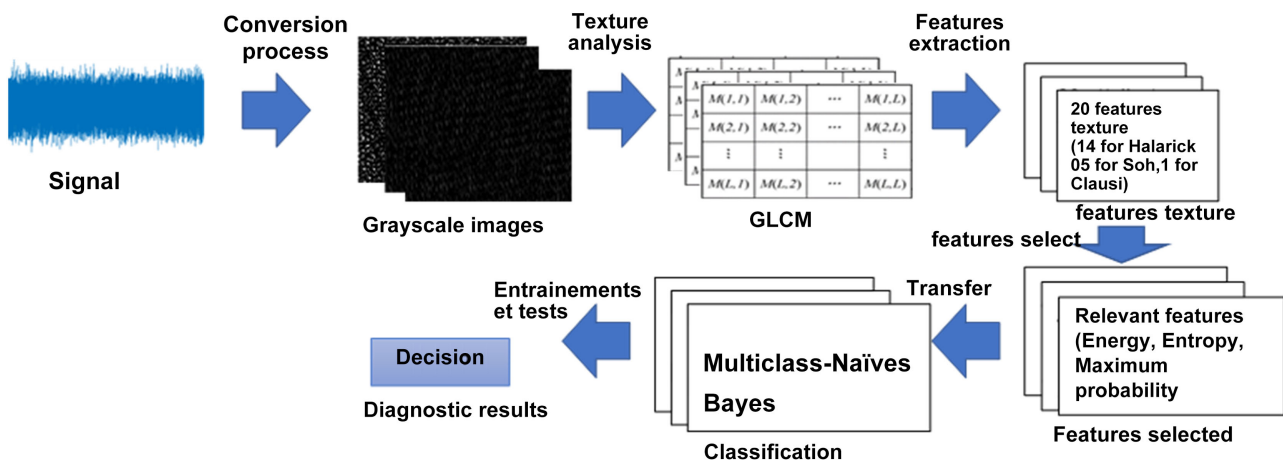
**Step 4:** features selection is done first by the PCA method to define the variables corresponding to the most significant features; then, by the SFE method; and finally, by the proposed PCA/SFE fusion method to obtain the most relevant features.

**Step 5:** The multiclass-Naïves Bayesis used to validate the relevance of these features by an optimum recognition rate of the bearing defect classes.

## 4. Experimental Results

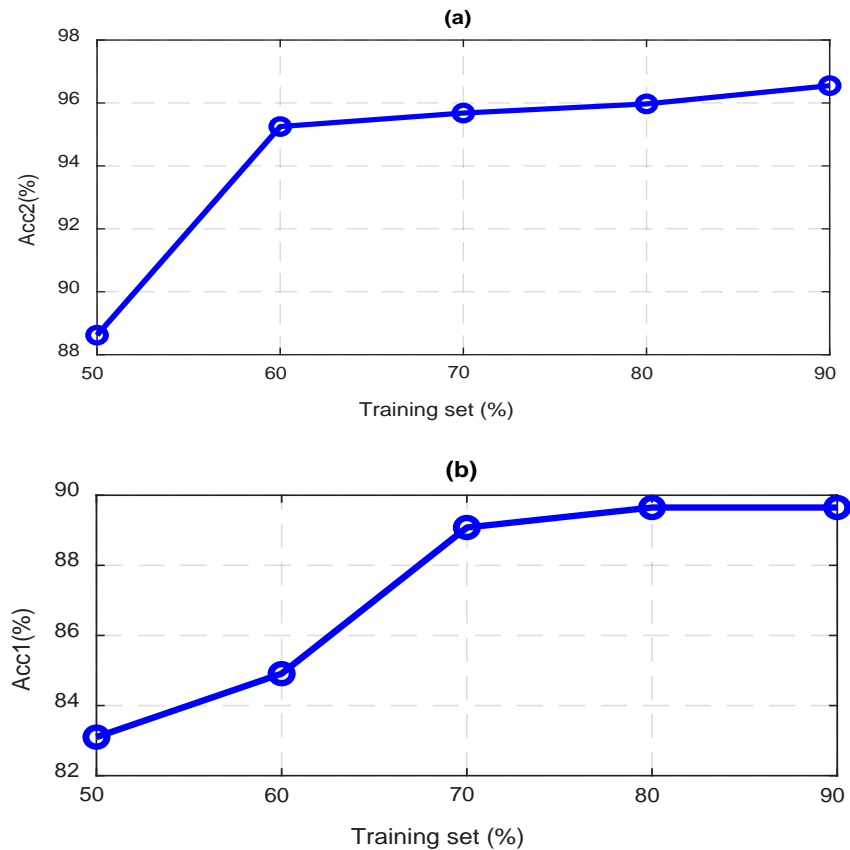
### 4.1. Bearing Fault Diagnosis with All Features

**Figure 4(a)** shows the classification accuracy Acc2, which is the detection rate of the successful classifier between the normal signal and the faults. **Figure 4(b)** shows the classification accuracy Acc1, for several training sets, when we use the twenty GLCM attributes from **Table 1** on the multiclass-Naïve Bayes. The training



**Figure 3.** Flow chart of the different diagnostic steps based on the selection of features.





**Figure 4.** Accuracy rate of the classification with the twenty GLCM attributes.

set is randomly selected between fifty and ninety percent of the database (580 images from **Table 1**). For each training set, the model classification is obtained and tested on the remaining images in the dataset.

**Figure 4(a)** shows that there is an accuracy rate on the detection of normal or faulty faults above 95% from 60% of the training set. Thus, with the set of twenty GLCM features, we have a good distinction between the normal and defect signal as shown in the confusion table, where we have only one case of false positive and four cases of false negative in the 116 test images. **Figure 4(b)** shows that there is an accuracy rate of over 88% from 60% of the training set in the case of identification.

**Table 4** shows the confusion matrix for the 20 attributes of the co-occurrence matrices in the case of defect detection for a training set greater than or equal to 70%.

#### 4.2. Relevant Features According to the PCA

The twenty GLCM attributes are realized on each of the five GLCMs obtained from five directions, namely  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ ; and the average of these four directions (Mo). Thus, with twenty GLCM features calculated on five GLCMs, a set of one hundred attributes is obtained. The first principal component obtained from the set of features represents 99.77% of the data. **Table 5** shows a

**Table 4.** Confusion matrix of the Naïve Bayes classifier.

		Reality	
		Normal signal	Faulty signal
Prediction	Normal signal	11	4
	Faulty signal	1	100

Accuracy = 95.68%.

**Table 5.** Correlation coefficient between the first principal component and the GLCM features.

<b>Features</b>	ENERG_0 ENERG_135 ENERG_Mo	ENERG_90 MAXIP_45	ENERG_45	MAXIP_Mo	MAXIP_90	MAXIP_135 MAXIP_0	SUMEN_90
<b>Correlation Coefficient</b>	0.9996	0.9990	0.9989	0.9985	0.9984	0.9977	0.9857
<b>Features</b>	SUMEN_Mo	SUMEN_135 SUMEN_0 ENTRO_90	SUMEN_45	DIFFEN_Mo	ENTRO_Mo	ENTRO_45	ENTRO_135 ENTRO_0
<b>Correlation Coefficient</b>	0.9838	0.9832	0.9824	0.9820	0.9818	0.9813	0.9810
<b>Features</b>	DIFFEN_45	DIFFEN_0 DIFFEN_135	DIFFEN_90	INFO2_135 INFO2_0	INFO2_Mo	INFO2_90	CORRE_Mo
<b>Correlation Coefficient</b>	0.9805	0.9791	0.9787	0.9630	0.9585	0.9427	0.9348
<b>Features</b>	CORRE_45	INFO2_45	CORRE_135 CORRE_0	CORRE_90	DIFFAV_90 DISSI_90	DIFFAV_Mo DISSI_Mo	DIFFAV_135 DIFFAV_0 DISSI_0 DISSI_135
<b>Correlation Coefficient</b>	0.9303	0.9293	0.8951	0.8951	0.8860	0.8747	0.8706

higher correlation coefficient between the first principal component and the different features. Each feature is presented with the notation (see **Table 1**) followed by the direction of the GLCM used. For example, “ENERG\_0” means the energy feature extracted from the GLCM obtained for the 0° direction. From these correlations, seven features of the GLCM most correlated with the first principal component are listed and represent the relevant features according to the PCA analysis.

**Figure 5** shows the performance of the classifiers when we use the relevant PCA features. These relevant GLCM features are used on the classifiers and we observe a certain stability of the recognition rate from 89.65% for the training and test dataset greater than or equal to 60%.

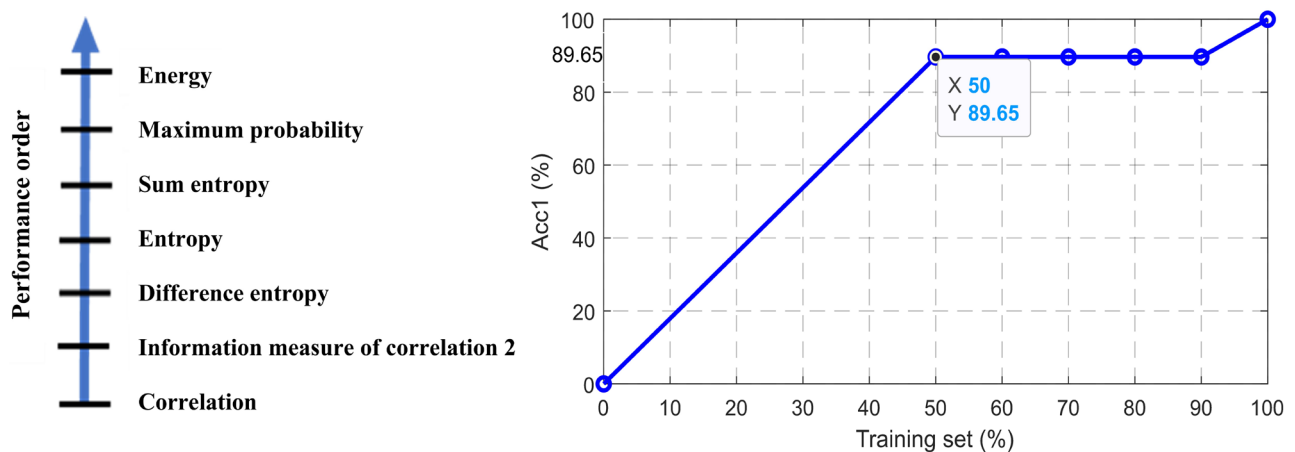
### 4.3. Relevant Features According to the SFE

**Table 6** shows the top seven features of each class according to the SFE. Since we

**Table 6.** The seven best features of the eleven classes.

Order	Classes					
	Normal	0.014 ball	0.021 outer race centered	0.007 inner race	0.007 ball	0.007 outer race opposite
1	DIFFEN_90	MAXIP_0	ENERG_45	MAXIP_45	MAXIP_45	MAXIP_0
2	CONTRA_0	ENERG_90	MAXIP_45	ENERG_0	MAXIP_90	MAXIP_90
3	CORRE_0	ENTRO_45	INFO1_0	MAXIP_0	CONTRA_0	CONTRA_0
4	HOMOG_0	CONTRA_0	CONTRA_0	CONTRA_0	CORRE_0	CORRE_0
5	SUMOF_0	CORRE_0	HOMOG_0	CORRE_0	ENERG_0	HOMOG_0
6	SUMAV_0	HOMOG_0	DIFFVA_0	HOMOG_0	ENTRO_45	SUMOF_0
7	SUMVA_0	SUMOF_0	INFO2_0	SUMOF_0	HOMOG_0	ENTRO_0

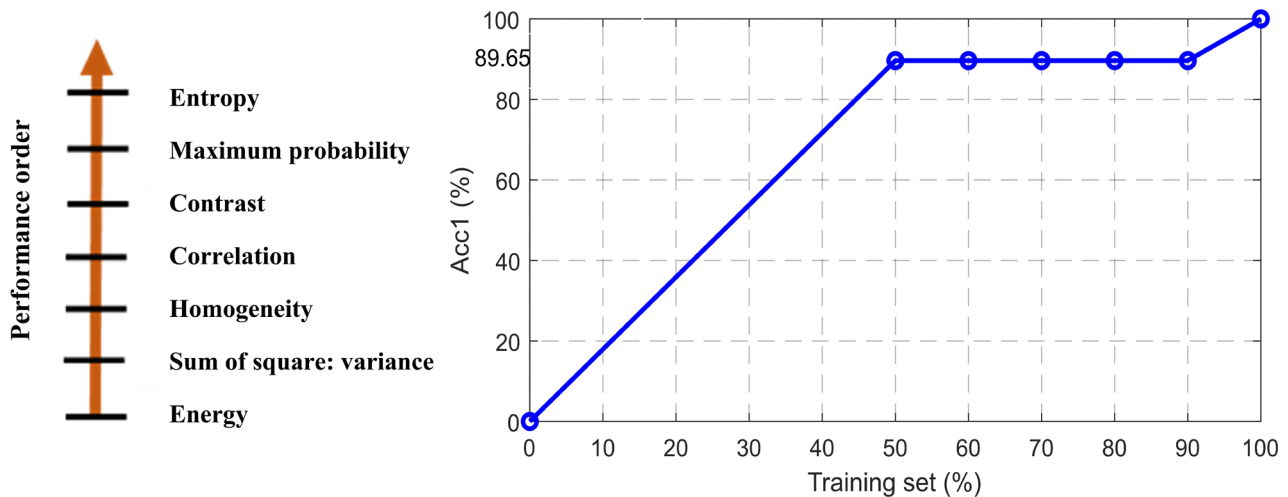
Order	CLASSES				
	0.007 outer race orthogonal	0.007outer race centered	0.014 inner race	0.021outer race opposite	0.021 outer race orthogonal
1	ENERG_90	ENERG_0	CONTRA_0	MAXIP_0	ENERG_90
2	ENERG_0	MAXIP_45	CORRE_0	MAXIP_90	ENERG_0
3	MAXIP_90	ENERG_90	ENERG_0	ENERG_90	ENERG_45
4	CONTRA_0	DIFFEN_0	HOMOG_0	MAXIP_135	MAXIP_0
5	CORRE_0	CONTRA_0	SUMOF_0	CONTRA_0	CONTRA_0
6	HOMOG_0	CORRE_0	ENTRO_0	CORRE_0	CORRE_0
7	SUMOF_0	HOMOG_0	SUMAV_0	HOMOG_0	HOMOG_0



**Figure 5.** Features performance order and classification accuracy for those relevant features of the PCA GLCM.

are looking for the best features, we calculated the number of occurrences of each feature for all classes. **Figure 6** shows the seven-high occurrence GLCM features that are the relevant features according to the SFE analysis.

**Figure 6** shows the performance of the classifiers when we use the relevant attributes of the SFE. These GLCM relevant features are used on the classifiers



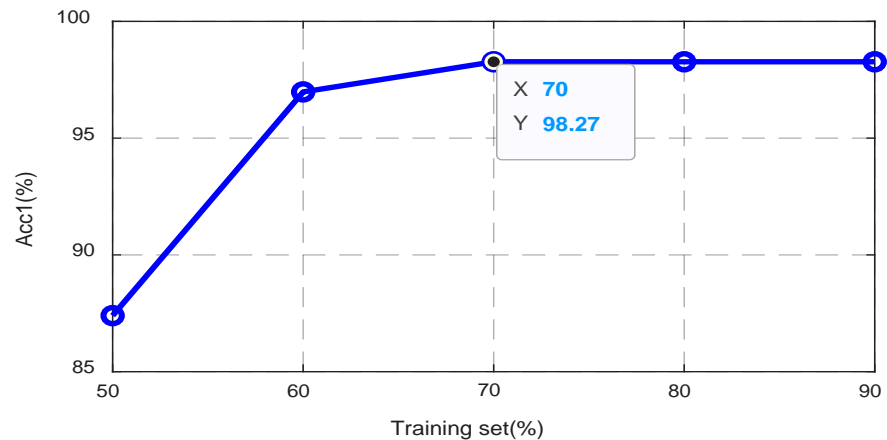
**Figure 6.** Features performance order and classification accuracy for these relevant features of the SFE GLCM.

and we observe a certain stability of the recognition rate from 89.65% for the training and test dataset greater than or equal to 60%.

#### 4.4. Relevant Features According to the PCA/SFE

To take advantage of both feature selection methods (PCA and SFE), we can select the features that appear in the best selection features of the PCA and SFE presented in **Figure 5** and **Figure 6** shows the performance of the classifier when we use the relevant features of PCA and SFE respectively. **Figure 7** show the performance of the classifiers when we use the relevant attributes of the fusion PCA/SFE. The relevant GLCM features of the fusion PCA/SFE show good classification performance with an accuracy above 99% for all training sets containing 60% or more data. Thus, the relevant GLCM features for bearing fault diagnosis among the twenty are the following four features: Energy, Entropy, Correlation and Maximum Probability.

The realisation of each classification system is based on the training and testing parameters. The classification system defined in this study is based on several training (50%; 60%; 70%; 80% and 90%) and testing (50%; 40%; 30%; 20% and 10%) samples. For each data item, an input vector is constructed by calculating the attributes of the GLCM. A study was first carried out on all 20 extracted features, then on the features by PCA and SFE and finally by merge PCA/SFE. The success rate of 89.65% was obtained on all training and test data sets of the 20 features (**Figure 4(b)**). This result already shows a feasibility in bearing diagnosis. Using the 04 relevant attributes obtained, the success rate is more than 98% on all data sets. We can observe the results of the classification of relevant features obtained by Naïve Bayes. We can observe the results based on the classification rate are listed here. It can be seen in **Figure 7** that the recognition rate is equal to 98.27% for a set of 70% of the training data samples and 30% of the test samples. This result is more accurate than when we take into account all the features computed without selection and even without fusion.



**Figure 7.** Classification accuracy for the relevant features of the PCA/SFE GLCM.

## 5. Conclusions

Studies have been made in the literature on bearing diagnosis by image processing. It should be noted that in none of the cases, texture analysis by GLCM was done on the images obtained by converting the temporal signal into a grayscale image. In this study, a new feature selection method based on the fusion of feature selection methods extracted from the GLCM of the vibration signal images was proposed. First, the vibration signals were converted into grayscale images and then the co-occurrence matrix was calculated on these images. Subsequently, PCA, SFE and PCA/SFE merge selection methods were applied to determine the most relevant features. The features of energy, entropy, correlation and maximum probability were obtained and used in the multiclass-Naïve Bayes classifier to validate the approach. The success rate of 89.65% was obtained for all training and test datasets on all 20 features of the GLCM. The classification of the relevant features obtained gave success rates above 96%. The present work addressed the automatic diagnosis of rolling defects by image processing. The impact of GLCM feature selection on the signal conversion images was presented on the classification results of rolling defects. The results showed that GLCM feature selection significantly increased the separability of the diagnostic results compared to those obtained without selection.

It should be noted that the diagnosis performed in this paper did not take into account the computation time. Therefore, an evaluation of the computation time of the method would be interesting for future work.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Ye, Z.M. and Wu, B. (2000) A Review on Induction Motor Online Fault Diagnosis. *Proceedings of the IPERC Third International Power Electronics and Motion Con-*

- trol Conference* (IEEE Cat. No.00EX435), Beijing, 15-18 August 2000, 1353-1358. <https://doi.org/10.1109/IPEMC.2000.883050>
- [2] Kaplan, K., Kaya, Y., Kuncan, M., Minaz, M. R. and Metin Ertunç, H. (2020) An Improved Feature Extraction Method Using Texture Analysis with LBP for Bearing Fault Diagnosis. *Applied Soft Computing*, **87**, Article ID: 106019. <https://doi.org/10.1016/j.asoc.2019.106019>
- [3] Ertunc, H. M., Ocak, H. and Aliustaoglu, C. (2013). ANN- and ANFIS-Based Multi-Staged Decision Algorithm for the Detection and Diagnosis of Bearing Faults. *Neural Computing and Applications*, **22**, 435-446. <https://doi.org/10.1007/s00521-012-0912-7>
- [4] Shen, C., Wang, D., Liu, Y., Kong, F. and Tse, P.W. (2014) Recognition of Rolling Bearing Fault Patterns and Sizes Based on Two-Layer Support Vector Regression Machines. *Smart Structures and Systems*, **13**, 453-471. <https://doi.org/10.12989/sss.2014.13.3.453>
- [5] Cai, J.H. (2014) Fault Diagnosis of Rolling Bearing Based on Empirical Mode Decomposition and Higher Order Statistics. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, **229**, 1630-1638. <https://doi.org/10.1177/0954406214545820>
- [6] Berredjem, T. and Benidir, M. (2018) Bearing Faults Diagnosis Using Fuzzy Expert System Relying on an Improved Range Overlaps and Similarity Method. *Expert Systems with Applications*, **108**, 134-142. <https://doi.org/10.1016/j.eswa.2018.04.025>
- [7] Zhang, S., Zhang, S., Wang, B. and Habetler, T.G. (2019) Deep Learning Algorithms for Bearing Fault Diagnostics—A Review. 2019 *IEEE 12th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, Toulouse, 27-30 August 2019, 257-263. <https://doi.org/10.1109/DEMPED.2019.8864915>
- [8] Kuncan, M., Kaplan, K., Minaz, M.R., Kaya, Y. and Metin Ertunç, H. (2020) A Novel Feature Extraction Method for Bearing Fault Classification with One Dimensional Ternary Patterns. *ISA Transactions*, **100**, 346-357. <https://doi.org/10.1016/j.isatra.2019.11.006>
- [9] Cempel, C. and Tabaszewski, M. (2007) Multidimensional Condition Monitoring of Machines in Non-Stationary Operation. *Mechanical Systems and Signal Processing*, **21**, 1233-1241. <https://doi.org/10.1016/j.ymssp.2006.04.001>
- [10] Li, W., Zhu, Z., Jiang, F., Zhou, G. and Chen, G. (2015) Fault Diagnosis of Rotating Machinery with a Novel Statistical Feature Extraction and Evaluation Method. *Mechanical Systems and Signal Processing*, **50-51**, 414-426. <https://doi.org/10.1016/j.ymssp.2014.05.034>
- [11] Yu, Y., Yu, D. and Cheng, J. (2006) A Roller Bearing Fault Diagnosis Method Based on EMD Energy Entropy and ANN. *Journal of Sound and Vibration*, **294**, 269-277. <https://doi.org/10.1016/j.jsv.2005.11.002>
- [12] Wang, Y., Ban, X., Chen, J., Hu, B. and Yang, X. (2015) License Plate Recognition Based on SIFT Feature. *Optik*, **126**, 2895-2901. <https://doi.org/10.1016/j.ijleo.2015.07.040>
- [13] Cho, H. and Park, H. (2017) Classification of Low-Grade and High-Grade Glioma Using Multi-Modal Image Radiomics Features. 2017 *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju, 11-15 July 2017, 3081-3084. <https://doi.org/10.1109/EMBC.2017.8037508>
- [14] Zhang, X.D. (2020) A Matrix Algebra Approach to Artificial Intelligence. Springer, Singapore. <https://doi.org/10.1007/978-981-15-2770-8>

- [15] Kuncan, M. (2020) An Intelligent Approach for Bearing Fault Diagnosis: Combination of 1D-LBP and GRA. *IEEE Access*, **8**, 137517-137529. <https://doi.org/10.1109/ACCESS.2020.3011980>
- [16] Haralick, R.M., Shanmugam, K. and Dinstein, I. (1973) Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, **3**, 610-621. <https://doi.org/10.1109/TSMC.1973.4309314>
- [17] Gong, S., Liu, C., Ji, Y., Zhong, B., Li, Y. and Dong, H. (2019) Advanced Image and Video Processing Using MATLAB. Vol. 12, Springer International Publishing, Heidelberg.
- [18] Soh, L.-K. and Tsatsoulis, C. (1999) Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-Occurrence Matrices. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, 780-795. <https://doi.org/10.1109/36.752194>
- [19] Clausi, D. (2002) An Analysis of Co-Occurrence Texture Statistics as a Function of Grey Level Quantization. *Canadian Journal of Remote Sensing*, **28**, 45-62. <https://doi.org/10.5589/m02-004>
- [20] Uddin, J., Kang, M., Nguyen, D.V. and Kim, J.-M. (2014) Reliable Fault Classification of Induction Motors Using Texture Feature Extraction and a Multiclass Support Vector Machine. *Mathematical Problems in Engineering*, **2014**, Article ID: 814593. <https://doi.org/10.1155/2014/814593>
- [21] Ruckstieß, T., Osendorfer, C. and van der Smagt, P. (2011) Sequential Feature Selection for Classification. *Australasian Joint Conference on Artificial Intelligence*, Perth, 5-8 December 2011, 132-141. [https://doi.org/10.1007/978-3-642-25832-9\\_14](https://doi.org/10.1007/978-3-642-25832-9_14)
- [22] Muntasa, A. (2015) Pengenalan Pola. Graha Ilmu, Yogyakarta.
- [23] Ali Khan, S. and Kim, J.-M. (2016) Automated Bearing Fault Diagnosis Using 2D Analysis of Vibration Acceleration Signals under Variable Speed Conditions. *Shock and Vibration*, **2016**, Article ID: 8729572. <https://doi.org/10.1155/2016/8729572>
- [24] Case Western Reserve University Bearing Data Center Website (n.d.). <http://csegroups.case.edu/bearingdatacenter/home>
- [25] Zhang, J.Q., Sun, Y., Guo, L., Gao, H., Hong, X. and Song, H.L. (2019) A New Bearing Fault Diagnosis Method Based on Modified Convolutional Neural Networks. *Chinese Journal of Aeronautics*, **33**, 439-447. <https://doi.org/10.1016/j.cja.2019.07.011>