Detecting Bearing Defects under High Noise Levels: A Classifier Fusion Approach

Luana Batista, Bechir Badri, Robert Sabourin, Marc Thomas
lbatista@livia.etsmtl.ca, bechirbadri@yahoo.fr, {robert.sabourin, marc.thomas}@etsmtl.ca
École de technologie supérieure
1100, rue Notre-Dame Ouest, Montreál, QC, H3C 1K3, Canada

***Abstract*—Automatic bearing fault diagnosis may be approached as a pattern recognition problem that allows for a significant reduction in the maintenance costs of rotating machines, as well as the early detection of potentially disastrous faults. When these systems employ real vibration data obtained from bearings artificially damaged, they have to cope with a very limited number of training samples. Moreover, an important issue that has been little investigated in the literature is the presence of noise, which disturbs the vibration signals, and how this affects the identification of bearing defects. In this paper, a new strategy based on the fusion of different Support Vector Machines (SVM) is proposed in order to reduce noise effect in bearing fault diagnosis systems. Each SVM classifier is designed to deal with a specific noise configuration and, when combined together – by means of the Iterative Boolean Combination (IBC) technique – they provide high robustness to different noise-to-signal ratio. In order to produce a high amount of vibration signals, considering different defect dimensions and noise levels, the BEAring Toolbox (BEAT) is employed in this work. Experiments indicate that the proposed strategy can significantly reduin the presence of, even in the presence of very noisy signals.**

I. Introduction

Machine condition monitoring (MCM) systems have been gaining importance in the manufacturing industry, since they allow for a significant reduction in the machinery maintenance costs [1], and, most importantly, the early detection of potentially disastrous faults [2]. Mass unbalance, rotor rub, shaft misalignment, gear failures and bearing defects are exemples of faults that may lead to the machine's breakdown [3].

Besides the detection of the early occurence and seriousness of a fault, MCM systems may also be designed to identify the components that are deteriorating, and to estimate the time interval during which the monitored equipment can still operate before failure [4]. These systems continuously measure and interpret signals (e.g., vibration, acoustic emission, infrared thermography, etc.), that provide useful information for identifying the presence of faulty symptoms.

The focus of this work is aimed at detecting early defects on bearings of rotating machinery. Since they are the place where the basic dynamic loads and forces are applied, bearings represent a critical component. A defective bearing causes malfunction and may even lead to catastrophic failure of the machinery [5]. Vibration analysis has been the most employed methodology for detecting bearings defects [6]. Each time a rolling element passes over a defect, an impulse of vibration is generated. On the other hand, if the machine is operating properly, vibration is small and constant [7].

Automatic bearing fault diagnosis can be viewed as a pattern recognition problem, and several systems have been designed using well-known classification techniques, such as Artificial Neural Networks (ANNs) and Support Vector Machines (SVM). Since these systems employ real vibration data obtained from bearings artificially damaged, they have to cope with a very limited amount of samples. With exception of a few works [2], [8] – which consider a validation set, besides the training and test sets –, the choice of the system's parameters, including the feature selection step, has been done by using the same datasets employed to train/test the classifiers. This may lead to biased classifiers that will hardly be able to generalize on new data. Another important aspect that has been little investigated in the literature is the presence of noise in the signals and how this affects the identification of bearing defects [4].

In this paper, a new strategy based on the fusion of different SVMs is proposed in order to reduce noise effect in bearing fault diagnosis systems. Each SVM classifier is designed to deal with a specific noise configuration and, when combined together – by using the Iterative Boolean Combination (IBC) technique [9] – they provide high robustness to different noise-to-signal ratio.

In order to produce a high amount of bearing vibration signals, considering different defect dimensions and noise levels, the BEAring Toolbox (BEAT) is employed in this work. BEAT is dedicated to the simulation of the dynamic behaviour of rotating ball bearings in the presence of localized defects, and it was shown to provide realistic results, similar to those produced by a sensor during experimental measurements [10].

This paper is organized as follows. Section II briefly presents the state-of-the-art on automatic bearing fault diagnosis. Section III describes the experimental methodology, including datasets, measures used to evaluate the system performance, and the IBC technique. Finally, the experiments are presented and discussed in Section IV.

II. Automatic bearing fault diagnosis

As previously mentioned, the interaction of defects in rolling element bearings produces impulses of vibration. As these shocks excite the natural frequencies of the bearing elements, the analysis of the vibration signal in the frequency-domain, by means of the Fast Fourrier Transform (FFT),

has been an effective method for predicting the condition of bearings [5]. Generally, each defective bearing component produces a specific frequency, which allows for localizing different defects occuring simultaneously. BPFO (Ball Pass Frequency on an Outer race defect), BPFI (Ball Pass Frequency on an Inner race defect), FTF (Fundamental Train Frequency) and BSF (Ball Spin Frequency) – as well as their harmonics, modulating frequencies, and envelopes – are examples of frequency-domain indicators, calculated from kinematic considerations [10].

Not only frequency- but also time-domain indicators have been widely employed as input features to train a bearing fault diagnosis classifier. Time-domain indicators allow for representing the vibration signal through a single scalar value. For instance, *peak* is the maximum amplitude value of the vibration signal, RMS (Root Mean Square) represents the effective value (magnitude) of the vibration signal and *kurtosis* describes the impulsive shape of the vibration signal [6].

A bearing fault diagnosis system may be designed to provide different levels of information about the defect(s). The first and simpler issue investigated in the literature is the detection of the presence or absence of a defect [3], [8], [11]. The second issue is the determination of the defect location, which may occur in different components of a bearing, i.e., inner race (IR), outer race (OR), rolling element (RE) and cage [2], [4], [12], [13], [14], [15], [16], [17]. Finally, the severity of a bearing defect is the last and perhaps the most difficult information to be predicted. Through this information, it may be possible to estimate the time interval during which the equipment can still operate safely. In the literature, this issue has been partially investigated, by associating a different class to each defect dimension [13], [15], [18].

III. Methodology

The objective of this work is to detect the presence or absence of bearing defects by taking into account six levels of (white) noise, i.e., signal-to-noise ratio (SNR) ranging from 40 to 5 db. Noise robustness is achieved through the incorporation of noisy data during the training phase, along with the fusion of different SVMs, each one is designed to deal with a specific noise configuration.

The BEAT simulator is employed to generate vibration signals coming from the operation of a ball bearing type SKF 1210 ETK9. The rotational speed is 1800 RPM, subjected to a non-rotating load of 3000 N. From the simulated data, the following time-domain indicators are calculated: RMS, peak, kurtosis, crest factor, impulse factor and shape factor. As frequency-domain indicators, BPFO, BPFI, 2BSF, as well as their first two hamonics are calculated. It is worth noting that the frequency-domain indicators employed in this work are normalized with respect to the rotational speed. As for the noise levels, a SNR of about 15db corresponds to the typical response produced by BEAT for a defect of 1mm. By changing the simulation parameters, such as lubrication conditions, more or less noise can be added to the original signal. For more details on BEAT's implementation please refer to [10].

The rest of this section describes the datasets and the performance evaluation methods employed in the experiments, as well as the Iterative Boolean Combination technique.

A. Datasets

Six noise configurations ($nc = 1, 2, 3, 4, 5, 6$) are considered in this paper, as indicated in Table I. For each noise configuration, there is a specific database, that is, $DB_{(nc)}$. Each sample in the databases is composed of a set of frequency and temporal indicators, plus the defect diameter, d_{def}, related to each bearing component, i.e., $d_{def(OR)}$, $d_{def(IR)}$ and $d_{def(RE)}$. Eight classes of defects are defined in Table II. The flag = 1 indicates that there is a defect in the corresponding component, while flag = 0 indicates the absence of defect. For instance, class 6 corresponds to two different defects occuring simultaneously: one in the outer race, and another in the ball. For the non-defective components, d_{def} goes from 0mm to 0.016mm. Regarding the defective components, d_{def} goes from 0.017mm to 2.8mm.

TABLE I
NOISE CONFIGURATIONS (nc).

nc	training/validation	test
1	40 db	40,30,20,15,10,5 db
2	40+30 db	40,30,20,15,10,5 db
3	40+30+20 db	40,30,20,15,10,5 db
4	40+30+20+15 db	40,30,20,15,10,5 db
5	40+30+20+15+10 db	40,30,20,15,10,5 db
6	40+30+20+15+10+5 db	40,30,20,15,10,5 db

TABLE II
CLASSES OF DEFECTS.

	OR	IR	RE
class 0	0	0	0
class 1	1	0	0
class 2	0	1	0
class 3	0	0	1
class 4	1	1	0
class 5	1	0	1
class 6	0	1	1
class 7	1	1	1

Since the objective of this work is to indicate the presence or absence of a bearing defect, regardless its location, only two classes are considered, i.e, *faultless* and *faulty*. The *faultless* class corresponds to the class 0 (see Table II) and, in order to have two balanced classes, the *faulty* class contains subsets of samples from classes 1 to 7. Table III presents the way the samples are partitioned.

TABLE III
DATA PARTITIONING FOR EACH $DB_{(nc)}$ ($1 \leq nc \leq 6$).

	positive class	negative class
trn	3500	3500
vld	1750	1750
tst (per noise level)	1750	1750

B. Performance Evaluation Methods

The ROC (Receiving Operating Characteristics) curve – where the true positive rates (TPR) are plotted as function of the false positive rates (FPR) – is a powerful tool for evaluating, comparing and combining pattern recognition systems [9]. Several interesting properties can be observed from ROC curves. First, the AUC (Area Under Curve) is equivalent to the probability that the classifier will rank a randomly chosen positive sample higher than a randomly chosen negative sample. This measure is useful to characterize the system performance through a single scalar value. In addition, the optimal threshold for a given class distribution lies on the ROC convex hull, which is defined as being the smallest convex set containing the points of the ROC curve. Finally, by taking into account several operating points, the ROC curve allows for analyzing these systems under different classification costs [19]. A similar way to evaluate systems is through a DET (Detection Error Trade-off) curve, in which the false negative rates (FNR) are plotted as function of the false positive rates, generally, on a logarithmic scale.

In this work, ROC and DET curves are computed from the output probabilities provided by the classifiers. The validation set, vld, is used for this task. In order to test a given classifier, its corresponding ROC operating points (thresholds) are applied to the set, tst. Results on test are shown as well in terms of equal error rate (EER), which is obtained when the threshold is set to have the false negative rate approximately equal to the false positive rate.

C. Iterative Boolean Combination (IBC)

Ensembles of classifiers (EoCs) have been used to reduce error rates of many challenging pattern recognition problems. The motivation of using EoCs stems from the fact that different classifiers usually make different errors on different samples. When the response of a set of $\mathcal{C}$ classifiers is averaged, the variance contribution in the bias-variance decomposition decreases by $\frac{1}{\mathcal{C}}$, resulting in a smaller classification error.

It has been recently shown that the Iterative Boolean Combination (IBC) [9] is an efficient technique for combining systems in the ROC space. IBC iteratively combines the ROC curves produced by different classifiers using all Boolean functions (i.e., $a \vee b$, $\neg a \vee b$, $a \vee \neg b$, $\neg(a \vee b)$, $a \wedge b$, $\neg a \wedge b$, $a \wedge \neg b$,$\neg(a \wedge b)$, $a \oplus b$, and $a \equiv b$), and does not require prior assumption that the classifiers are statistically independent. At each iteration, IBC selects the combinations that improve the Maximum Realizable ROC (MRROC) curve – i.e., the convex hull obtained from all individual ROC curves – and recombines them with the original ROC curves until the MRROC ceases to improve. Algorithm 1 explains how to combine a pair of ROC curves, R_a and R_b, considering a single IBC iteration. For more details about this technique, please refer to Algorithms 1 to 3 in [9].

Algorithm 1 : Boolean combination of two ROC curves.

Inputs: Thresholds of ROC curves, T_a and T_b, and Labels

let $m \leftarrow$ number of distinct thresholds in T_a
let $n \leftarrow$ number of distinct thresholds in T_b
Allocate F an array of size: $[2, m \times n]$ */*holds temporary results of fusions*/*
let BooleanFunctions $\leftarrow \{a \vee b,\ \neg a \vee b,\ a \vee \neg b,\ \neg(a \vee b),\ a \wedge b,\ \neg a \wedge b,\ a \wedge \neg b, \neg(a \wedge b),\ a \oplus b,\ a \equiv b\}$
Compute MRROC$_{old}$ of the original curves
for each bf in BooleanFunctions **do**
  for i=1,...,m **do**
    */*converting threshold of 1st ROC to responses*/*
    $R_a \leftarrow T_a \geq T_{a_i}$
    for j=1,...,n **do**
      */*converting threshold of 2nd ROC to responses*/*
      $R_b \leftarrow T_b \geq T_{b_j}$
      */*combined responses with bf*/*
      $R_c \leftarrow bf(R_a, R_b)$
      Compute (FPR,TPR) using R_c and Labels
      Push (FPR,TPR) onto F
    end for
  end for
  Compute MRROC$_{new}$ of F and store thresholds and corresponding Boolean functions that exeeded the MRROC$_{old}$
  MRROC$_{old} \leftarrow$ MRROC$_{new}$ */*Update ROCCH*/*
end for
return MRROC$_{new}$

IV. Simulation Results and Discussions

Two main experiments are performed. In the first experiment, each database $DB_{(nc)}$ $(1 \leq nc \leq 6)$ is employed in the generation of a baseline system $S_{(nc)}$. For each $DB_{(nc)}$:

- trn is used to train n different classifiers c_i, $1 \leq i \leq n$, by employing different SVM parameters;
- vld is used to validate each individual classifier c_i, by means of ROC curves, and select that one with the highest AUC. The select classifier is called $S_{(nc)}$;
- tst is used to test the performance of $S_{(nc)}$.

In the second experiment, the IBC technique [9] is used to combine the best classifier of each noise configuration.

A. Experiment 1

The goal of the first experiment was to obtain the best baseline system for each one of the noise configurations defined in Table I. For each database $DB_{(nc)}$ $(1 \leq nc \leq 6)$, several SVMs were trained using the grid search technique [20], so that the SVM providing the highest AUC is selected. To train the SVMs with RBF kernel, the following values were employed: $\gamma = \{2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^{0}\}$ and C = $\{2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^{0}, 2^{1}, 2^{2}, 2^{3}, 2^{4}, 2^{5}\}$, where γ is the RBF kernel parameter, and C is the SVM cost parameter.

Since the obtained ROC curves reached AUC close to 1, as indicated in Table IV, DET curves on a log-log scale are presented instead (see Figure 1). Note that the curve representing system $S_{(nc=1)}$ does not appear in the graphic because a complete separation of both classes was obtained.

TABLE IV
ROC AUC ON VALIDATION DATA.

System	AUC
$S_{(nc=1)}$	1
$S_{(nc=2)}$	0.9999
$S_{(nc=3)}$	0.9999
$S_{(nc=4)}$	0.9996
$S_{(nc=5)}$	0.9992
$S_{(nc=6)}$	0.9989

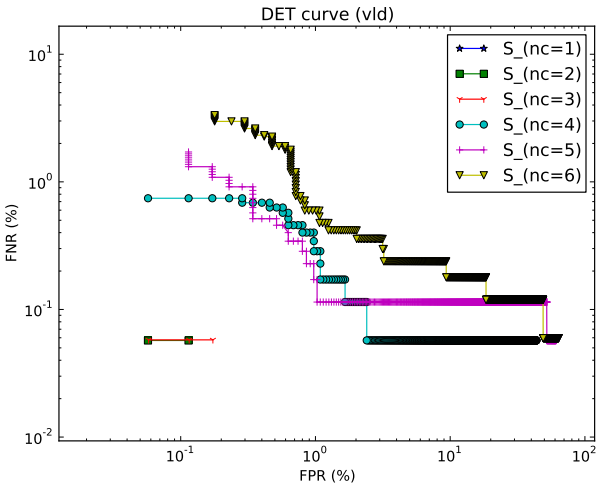


Fig. 1. DET curves of the selected systems $S_{(nc)}$, $1 \leq nc \leq 6$, using their respective validation sets (vld).

Figure 2 shows the DET curves obtained on test data (tst) using the validation operating points. Observe that DET curves plotted in a same graphic are the results of a same system on different test data. Therefore, these curves are useful in order to analyse the robustness of each system regarding individual noise levels. It is worth noting that system $S_{(nc=1)}$ provided a complete class separation for 40 db (that's why the corresponding DET curve does not appear in the graphic), and, in a similar way, $S_{(nc=2)}$ and $S_{(nc=3)}$ provided a complete class separation for 40 db and 30 db.

Table V presents the average EER obtained for each noise level during test, over 10 trials. The symbol 'κ' indicates that the system has a random (or worse than random) behaviour for a given test set. A similar situation was observed in the work of Lazzerini and Volpi [4], where classification accuracies of 50% or less were obtained for high levels of noise. As expected, the systems become more robust to higher noise levels as they are gradually incorporated to the training phase.

B. Experiment 2

In the second experiment, IBC was used to combine the best classifier of each noise configuration, found in the first experiment. For all classifiers, a same validation set containing all noise levels (i.e., 40+30+20+15+10+5 db) was employed. Since a high number of combinations is performed, the number thresholds per curve was limited to 500 (in the previous experiment, all validation scores were employed as thresholds).

Figure 3 shows the DET curve obtained with IBC, along with the DET curves of the six systems employed during the combination process. Note that IBC improved the Maximum Realizable DET (MRDET) curve of the individual systems.

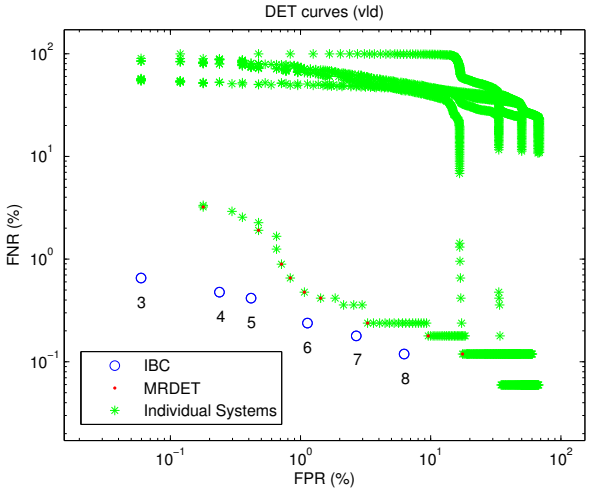


Fig. 3. DET curve obtained with IBC using a validation set containing all noise levels. The DET curves of the 6 individual systems and the Maximum Realizable DET curve (MRDET) are shown as well.

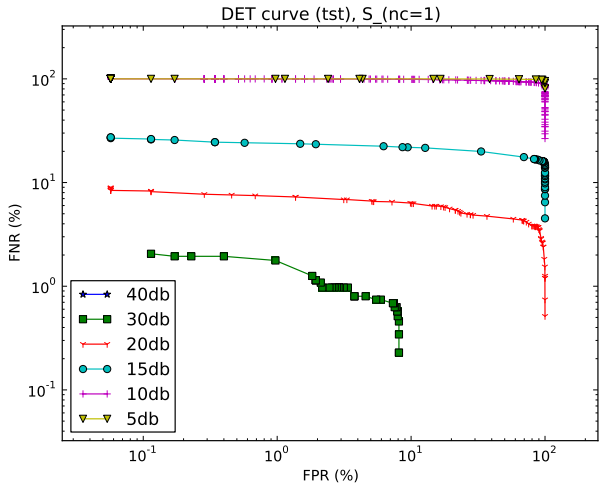
TABLE VI
OPERATING POINTS OF IBC DET CURVE.

operating point	FNR (%)	FPR (%)
1	100.00	0.00
2	0.89	0.00
3	0.65	0.06
4	0.48	0.24
5	0.42	0.42
6	0.24	1.13
7	0.18	2.68
8	0.12	6.25
9	0.00	15.65
10	0.00	100.00

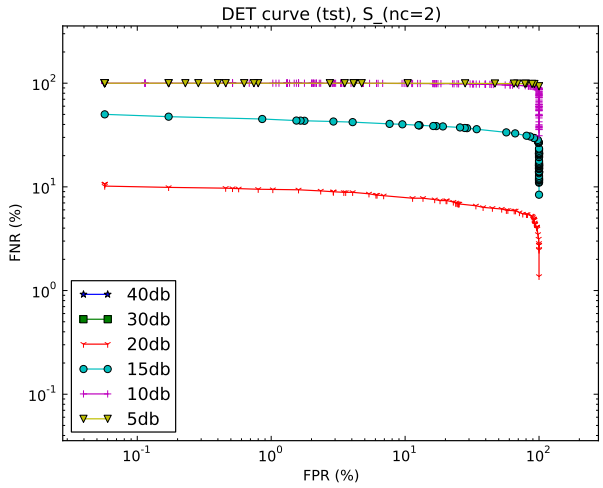
The operating points falling on the IBC curve are presented on Table VI. Each point is the result of a Boolean combination of different individual classifiers. For instance, the operating point 5, which gives the EER, corresponds to a boolean combination (BC) of all 6 classifiers (c_j, $1 \leq j \leq 6$), that is, $BC_{\{EER\}} = (c_1 \wedge c_2 \wedge c_3 \wedge c_4 \wedge c_5 \wedge c_6)$, using the decision thresholds indicated in Table VII.

TABLE VII
DECISION THRESHOLDS ASSOCIATED TO THE EER OPERATING POINT.

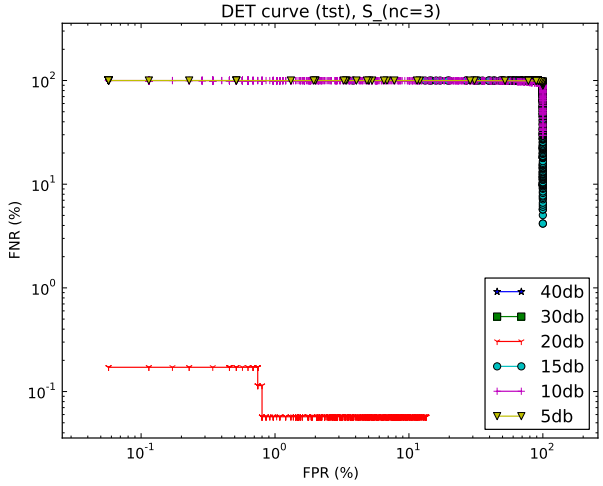
classifier	threshold
c_1	0.9919
c_2	0.9816
c_3	0.9916
c_4	1.5587e-004
c_5	0.0095
c_6	0.0452



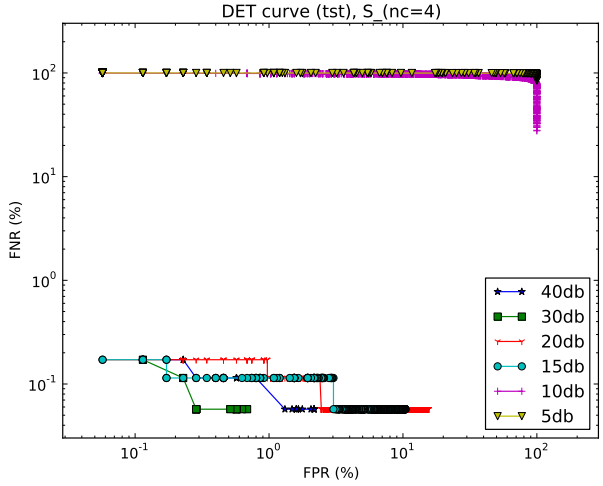
(a)



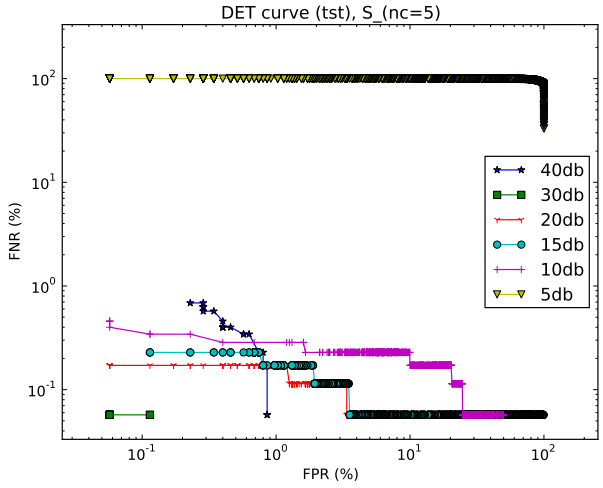
(b)



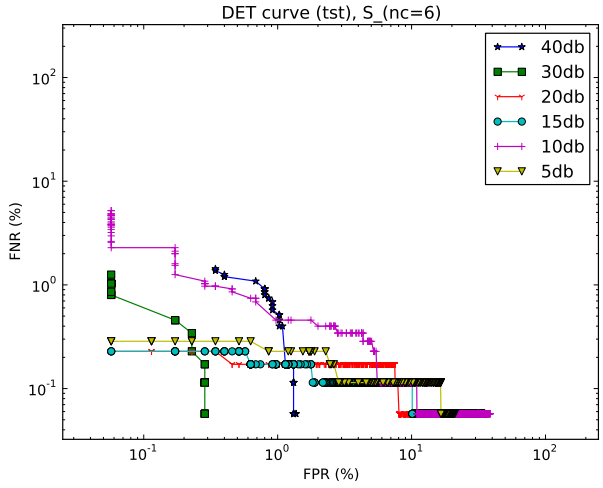
(c)



(d)



(e)



(f)

Fig. 2. DET curves of the selected systems $S_{(nc)}$, $1 \leq nc \leq 6$, using the test sets (tst).

TABLE V
AVERAGE EER (%) ON TEST DATA OVER 10 TRIALS.

tst	$S_{(nc=1)}$	$S_{(nc=2)}$	$S_{(nc=3)}$	$S_{(nc=4)}$	$S_{(nc=5)}$	$S_{(nc=6)}$
40 db	0.02	0.05	0.10	0.17	0.38	0.57
30 db	1.40	0.00	0.01	0.09	0.14	0.32
20 db	5.38	7.51	0.07	0.06	0.08	0.10
15 db	20.98	34.12	ʁ	0.11	0.16	0.15
10 db	ʁ	ʁ	ʁ	ʁ	0.27	0.68
5 db	ʁ	ʁ	ʁ	ʁ	ʁ	0.36

Figure 4 shows the DET curves obtained on test data, for 20, 15, 10 and 5 db noise levels, using the IBC points indicated in Figure 3; and Table VIII presents the average EER (over 10 trials) obtained with IBC, Majority vote and with the best single classifier. The Majority vote rule reached very low EER with respect to 40, 30, 20 and 15 db noise levels. On the other hand, a random behaviour was observed for 10 and 5 db noisy data. The reason is due to the fact that the majority of the individual classifers presents a random behaviour for high levels of noise. Observe that IBC provided an improvement for almost all test datasets with respect to the single best classifier obtained in the previous experiment.

TABLE VIII
AVERAGE EER (%) ON TEST DATA OVER 10 TRIALS.

tst	IBC technique	Majority vote	Single best ($S_{(nc=6)}$)
40 db	0.06	0.06	0.57
30 db	0.00	0.01	0.32
20 db	0.11	0.06	0.10
15 db	0.11	0.10	0.15
10 db	0.29	ʁ	0.68
5 db	0.33	ʁ	0.36

Finally, Table IX presents additional results of IBC on test data, when the threshold is set in order to reach FPR (%) = $\{1, 0.1, 0.01\ 0.001\}$. These intermediate points are obtained by using interpolation [21]. Note that the FPR decreases at the expense of an FNR increasing. In practice, the trade-off between FPR and FNR can be adjusted by the operators according to the current error costs.

V. CONCLUSION

In this paper, a new strategy based on the fusion of classifiers in the ROC space was proposed in order to reduce noise effect in bearing fault diagnosis systems. Noise robustness was achieved through the incorporation of noisy data (ranging from 40 to 5 db) during the training phase, along with the Iterative Boolean Combination of different SVMs, each one is designed to deal with a specific noise configuration.

Experiments performed with simulated vibration signals indicate that the proposed strategy can significantly reduce the error rates, even in the presence of high levels of noise. The results are comparable to those presented in [4] – with respect to noise robustness –, despite the use of different datasets, features and defect types. Future work consist of validating the proposed strategy with real vibration signals.

TABLE IX
ADDITIONAL ERROR RATES OBTAINED WITH IBC OVER 10 TRIALS.

Expected FPR = 1%			
tst	FNR	FPR	Average
40 db	0.02	10.70	5.36
30 db	0.01	8.61	4.31
20 db	0.13	5.34	2.73
15 db	0.16	6.48	3.32
10 db	0.18	8.94	4.56
5 db	0.09	21.87	10.98
Expected FPR = 0.1%			
tst	FNR	FPR	Average
40 db	0.03	0.37	0.40
30 db	0.01	0.15	0.08
20 db	0.18	0.01	0.09
15 db	0.23	0.09	0.16
10 db	0.36	1.35	0.85
5 db	0.15	5.74	2.94
Expected FPR = 0.01%			
tst	FNR	FPR	Average
40 db	0.05	0.10	0.07
30 db	0.03	0.03	0.03
20 db	0.50	0.00	0.02
15 db	0.56	0.02	0.29
10 db	1.60	0.14	0.87
5 db	0.28	0.95	0.61
Expected FPR = 0.001%			
tst	FNR	FPR	Average
40 db	0.09	0.11	0.10
30 db	0.07	0.03	0.05
20 db	0.63	0.00	0.31
15 db	0.94	0.01	0.47
10 db	3.38	0.06	1.72
5 db	0.48	0.51	0.49

REFERENCES

[1] S. Liang, R. Hecker, and R. Landers, "Machining process monitoring and control: The state-of-the-art," *Journal of Manufacturing Science and Engineering*, no. 2, pp. 297–310, 2004.

[2] H. Guo, L. Jack, and A. Nandi, "Feature generation using genetic programming with application to fault classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 35, no. 1, pp. 89–99, 2005.

[3] B. Samanta, K. Al-Balushi, and S. Al-Araimi, "Bearing fault detection using artificial neural networks and genetic algorithm," *EURASIP Journal on Applied Signal Processing*, pp. 366–377, 2004.

[4] B. Lazzerini and S. Volpi, "Classifier ensembles to improve the robustness to noise of bearing fault diagnosis," in *Pattern Analysis and Applications*, 2011, pp. 1–17.

[5] N. Tandon and A. Choudhury, "A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings," *Tribology International*, vol. 32, no. 8, pp. 469–480, 1999.

[6] M. Thomas, *Fiabilite, maintenance predictive et vibration des machines*. Ecole de technologie superieure, 2003.

[7] I. Alguindigue, A. Loskiewicz-Buczak, and R. Uhrig, "Monitoring and diagnosis of rolling element bearings using artificial neural networks," *IEEE Transactions on Industrial Electronics*, vol. 40, no. 2, pp. 209–217, 1993.

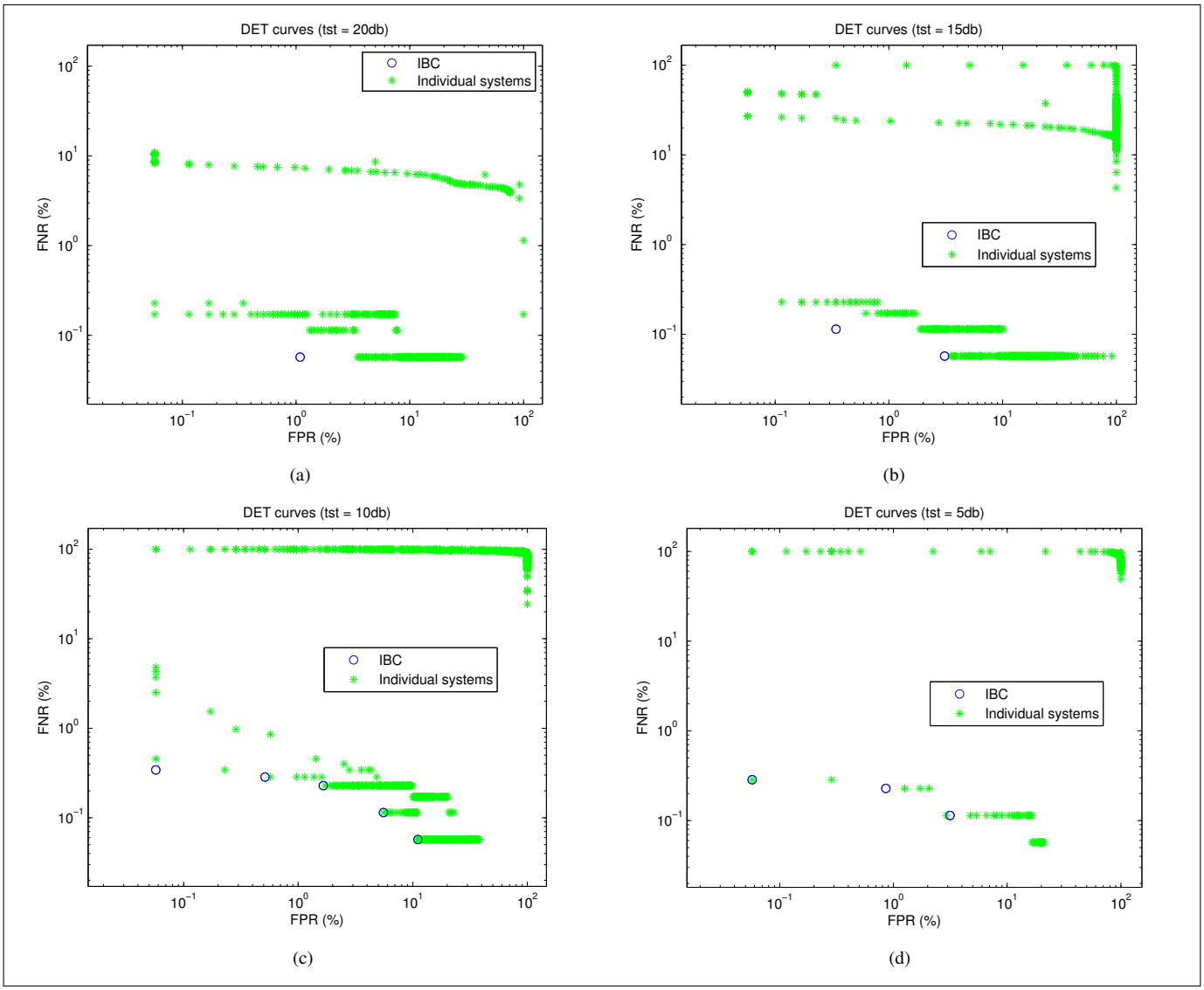


Fig. 4. DET curve obtained with IBC using the test sets (tst). The DET curves of the 6 individual systems are shown as well.

[8] L. Jack and A. Nandi, "Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms," *Mechanical Systems and Signal Processing*, vol. 16, no. 2-3, pp. 373–390, 2002.

[9] W. Khreich, E. Granger, A. Miri, and R. Sabourin, "Iterative boolean combination of classifiers in the roc space: An application to anomaly detection with hmms," *Pattern Recogn.*, vol. 43, pp. 2732–2752, August 2010.

[10] S. Sassi, B. Badri, and M. Thomas, "A numerical model to predict damaged bearing vibrations," *Journal of Vibration and Control*, no. 11, pp. 1603–1628, 2007.

[11] K. Teotrakool, M. Devaney, and L. Eren, "Bearing fault detection in adjustable speed drives via a support vector machine with feature selection using a genetic algorithm," in *IEEE Instrumentation and Measurement Technology Conference*, may 2008, pp. 1129 –1133.

[12] P. Kankar, S. Sharma, and S. Harsha, "Fault diagnosis of ball bearings using continuous wavelet transform," *Applied Soft Computing*, vol. 11, pp. 2300–2312, 2011.

[13] S. Volpi, M. Cococcioni, B. Lazzerini, and D. Stefanescu, "Rolling element bearing diagnosis using convex hull," in *IJCNN*, 2010, pp. 1–8.

[14] K. Bhavaraju, P. Kankar, S. Sharma, and S. Harsha, "A comparative study on bearings faults classification by artificial neural networks and self-organizing maps using wavelets," vol. 2, no. 5, pp. 1001–1008, 2010.

[15] A. Widodo, E. Kim, J. Son, B. Yang, A. Tan, D. Gu, B. Choi, and J. Mathew, "Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine," *Expert Systems with Applications*, vol. 36, no. 3, Part 2, pp. 7252–7261, 2009.

[16] B. Sreejith, A. Verma, and A. Srividya, "Fault diagnosis of rolling element bearing using time-domain features and neural networks," in *Third international Conference on Industrial and Information Systems*, 2008, pp. 1–6.

[17] A. Rojas and A. Nandi, "Practical scheme for fast detection and classification of rolling-element bearing faults using support vector machines," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1523–1536, 2006.

[18] M. Cococcioni, B. Lazzerini, and S. Volpi, "Automatic diagnosis of defects of rolling element bearings based on computational intelligence techniques," *International Conference on Intelligent Systems Design and Applications*, pp. 970–975, 2009.

[19] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, pp. 861–874, June 2006.

[20] C. Chang and C. Lin, "Libsvm: a library for support vector machines," in *http://www.csie.ntu.edu.tw/˜ cjlin/libsvm*, 2001.

[21] M. Scott, M. Niranjan, and R. Prager, "Realisable classifiers: improving operating performance on variable cost problems," 1998.