Thermal Imaging and Vibration Based Multi-Sensor Fault Detection for Rotating Machinery

Olivier Janssens, Mia Loccufier, and Sofie Van Hoecke

Abstract—In order to minimize operation and maintenance costs and extend the lifetime of rotating machinery, damaging conditions and faults should be detected early on and automatically. To enable this, sensor streams should continuously be monitored, processed and interpreted. In recent years, infrared thermal imaging has gained attention for said purpose. However, the detection capabilities of a system that uses infrared thermal imaging is limited by the modality captured by this single sensor, as is any single sensor-based system. Hence, within this article a multi-sensor system is proposed which not only uses infrared thermal imaging data, but also vibration measurements for automatic condition and fault detection in rotating machinery. It is shown that by combining these two types of sensor data, several conditions/faults and combinations can be detected more accurately than when considering the sensor streams individually.

Index Terms—Machine learning, fault detection, preventive maintenance, feature extraction

I. INTRODUCTION

IN order to prevent machine failure, prolong the lifetime of machines, reduce operational costs and enhance operational uptime, condition monitoring and fault detection are required. This is done by monitoring a set of physical properties of a machine using different types of sensors. In recent years, infrared thermal imaging based approaches for online fault detection and condition monitoring have started to appear as is discussed in the recent review by Touret et al. [1]. Infrared thermal imaging enables non-contact, non-intrusive and fine-grained temperature measurements, which is ideal for automatic fault detection and condition monitoring. In order to let a system autonomously detect conditions and faults, infrared thermal imaging is mostly used in combination with image processing and machine learning [2]–[11]. Such a solution often consists of five steps (see also Fig. 1):

1) **Image extraction:** from an infrared thermal video, single frames are extracted.
2) **Region of interest extraction**: per frame, regions from which features have to be extracted are segmented. In this step irrelevant parts in the image are also removed.
3) **Feature extraction**: features are extracted from the segmented areas. These features are designed so that they contain information related to conditions that should be detectable.

O. Janssens and S. Van Hoecke are with imec - IDLab, Department of Electronics and Information Systems at Ghent University, e-mail: (odjansse.janssens@ugent.be)

Mia Loccufier is with the DySC Research Group, Ghent University, Technologiepark 914, 9052 Zwijnaarde, Belgium

4) **Feature selection/fusion**: possibly not all features are relevant, or features share information. Hence, they can be eliminated or fused together.
5) **Classification**: to let the system automatically decide if a fault or condition is present, a classification (machine learning) algorithm is required. This algorithm determines the condition using the extracted features.
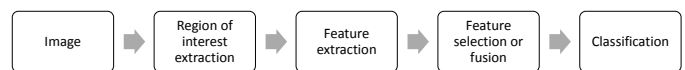

Fig. 1: General steps in an image processing/machine learning pipeline.

For automated infrared thermal imaging-based fault detection and condition monitoring, the focus of the field has been on data-driven approaches [10]. Such approaches exploit phenomena in the data that can be related to the conditions and faults. This is necessary because pre-existing model-based knowledge that can be extracted from the infrared thermal imaging data is currently not available.

In related literature, it has been demonstrated that several types of faults and conditions in rotating machinery are detectable to a certain extent using infrared thermal imaging, such as rotor imbalance, misalignment, coupling looseness, lubricant inadequacy and rolling element bearing damages [2]–[13]. For the detection of certain conditions, it is easy to see why infrared thermal imaging is advantageous, such as for lubricant inadequacy. One of the main purposes of lubricant is friction control. If there is too little lubricant in the bearing housing, excessive friction will occur, which results in heat that is observable by an infrared thermal camera. Additionally, if there is too much lubrication in the bearing, churning will occur, also resulting in additional heat that can be observed. However, for certain conditions, infrared thermal imaging-based approaches provide sub-optimal results, such as for the detection of damages to a rolling element bearing. For example, as demonstrated in [10], the outer-raceway damages can only be detect 45 % of the time using an infrared thermal imaging-based system.

Fortunately, other commonly used techniques such as acoustic emission [14], motor current signature analysis (MCSA) [15] and vibration analysis [16] have proven their merits w.r.t. damage detection of rolling element bearings. Therefore, within this article, additional to infrared thermal imaging we focus on vibration analysis as it is a well established, mature –and arguably the most frequently used– technique for bearing fault detection [17], [18]. Vibration analysis, which uses accelerometer data, is often done based

on the physics of the faults and the dynamics of the system. Therefore, several properties in the frequency spectrum of the vibration signal have been related to specific conditions and damages [16]. This knowledge allows to extract specific model-based features with well established techniques. Once processed, faults can be detected and machine's conditions can be assessed. For example, raceway damages will cause a peak at the bearings' fault frequencies for which the amplitude can be used as an indicative feature [16], [19].

As infrared thermal imaging-based approaches suffer from downsides that can be compensated using features extracted from established vibration analysis techniques, a combination of both approaches in a multi-sensor system will improve accuracy. Up until now, very little research has been done on combining vibration data with temperature data. One example where a multisensor system is developed is [20]. Their proposed system is designed to detect a cracked rotor, rotor rub or coupling misalignment in a rotating machine, using both vibration data and temperature (thermocouple) measurements. It should be noted that no thermal camera is used. A second relevant example is provided in [21], where MCSA is combined together with infrared thermography for fault detection. Their focus lies on faults such as bearing damage, misalignment and imbalance, but not on the combination of faults. It should also be noted that in contrast to vibration analysis, when using MCSA only rotating machinery using electrical actuators can be considered.

In this paper we propose a system that uses both infrared thermal imaging data and vibration data, which is the first of its kind to the best of our knowledge. The system extracts features from both sensor streams and fuses them together. The fused features are subsequently provided to a classification algorithm. We show that not only can the system detect a multitude of conditions/faults but also combinations of faults and conditions.

For this research, two data sets were constructed both consisting of vibration measurements and infrared thermal imaging measurements. In Section II, the set-up to create these data sets and the data sets themselves are discussed. Next, in Section III, an overview is provided of the preprocessing, feature extraction and classification steps of our system. Subsequently, in Section IV, the results of the proposed system applied on our data sets are discussed. Finally, in Section V, a conclusion is provided.

II. DATA

Two data sets were created –both containing vibration measurements and thermal imaging data– of conditions and faults we induced in a rotating set-up. The conditions present in each data set are listed in TABLE I and TABLE II respectively. Note that a condition in the data sets always exists out of a combination of two faults/conditions, e.g.: condition 4 in the first data set consists of (1) a bearing with an outer-raceway fault while (2) the machine is imbalanced.

A. Set-up

The set-up used to create the two data sets can be seen in Fig. 2. A detailed list of the relevant properties of the set-up

TABLE I: Summary of the 8 conditions in data set one. The weights in the column headers indicate the weight of the bolt added to the rotor.

	No imbalance	Imbalance: 13 g or 17.3 N
Healthy bearing (HB)	Condition 1	Condition 2
Outer-raceway fault (ORF)	Condition 3	Condition 4
Mildly inadequately lubricated bearing	Condition 5	Condition 6
Extremely inadequately lubricated bearing	Condition 7	Condition 8

are listed in TABLE III and for the infrared thermal camera in TABLE IV. The rolling element bearing in the housing at the right-hand side in the set-up was changed in-between test runs, hence this is the housing that was monitored by the thermal camera. Additional to the infrared thermal camera, a thermocouple was placed in the room to measure the ambient temperature. Furthermore, two accelerometers were mounted on this bearing housing to measure the accelerations in the x-direction and y-direction.
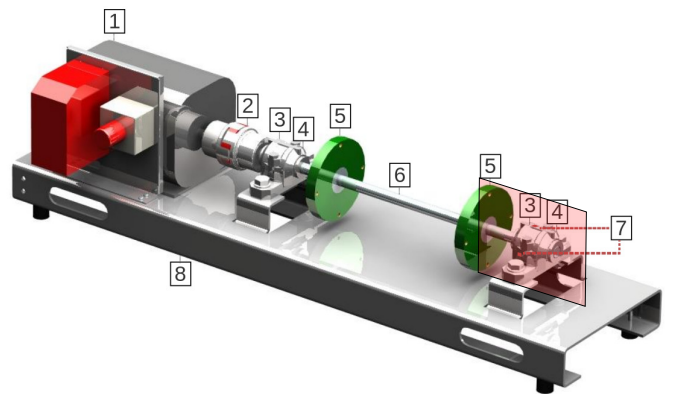

Fig. 2: 3D image of the set-up. The labels are: 1. servo-motor with a speed controller; 2. coupling; 3. bearing housing; 4. bearing; 5. disk; 6. shaft; 7. accelerometers; 8. metal plate. The red square indicates what the infrared thermal camera captures

To imitate outer-raceway faults (ORF) in the bearing, three small shallow grooves were added mechanically on the bearings' outer-raceway by a field expert (see Fig. 3 for an example of such a groove). For the respective data sets, two locations for the ORF faults were chosen. For data set one, the ORF is located at the 10 o'clock position in the housing. This is close to the top of the housing facing the infrared thermal camera as to better see the thermal effects of the raceway fault. For data set two, the ORF fault is located at the 6 o'clock position, which is the loaded zone, as to maximize the impact of the raceway fault. A cross-sectional view containing the locations of these ORFs can be seen in Fig. 4.

Both the healthy bearings (HB) and those with an ORF are placed in a housing that contains a grease reservoir enclosing 20 g of grease [22]. Additionally, at the start 2.5 g of grease was added to these bearings as recommended by the manufacturer [9]. For the bearings with reduced lubricant in data set one, i.e. mildly inadequately lubricated bearings (MILB) and extremely inadequately lubricated bearings (EILB), no grease reservoir is present. For the MILBs, the grease on each individ-

TABLE II: Summary of the 12 conditions in data set two. The weights in the column headers indicate the weight of the bolt added to the rotor.

	No imbalance	Imbalance: 4.1 g or 5.5 N	Imbalance: 9.3 g or 12.4 N	Imbalance: 13 g or 17.3 N
Healthy bearing (HB)	Condition 1	Condition 2	Condition 3	Condition 4
Outer-raceway fault (ORF)	Condition 5	Condition 6	Condition 7	Condition 8
Hard particle contamination (HP)	Condition 9	Condition 10	Condition 11	Condition 12

TABLE III: Test set-up details

Property	Value
Accelerometer type	CCDL
Brand	Brüel & Kjaer
Accelerometer product type	4533-B
Bearing code	FAG 22205-E1-K
Bearing type:	Spherical roller bearing with tapered bore and adapter sleeve
Housing code	SNV052-F-L
Housing type	Closed plummer block
Grease	Molykote BR 2 plus
Chosen rotation speed	25 Hz which is 1500 rpm
Eigenfrequency of the set-up determined using an impact test	17.8 Hz
Accelerometer sample frequency	51,200 Hz
Motor type	Single phase asynchronous induction motor
Motor power	1.1 kW
Motor brand	SEW Eurodrive
Motor product ID	DRE90M2/FI/LN
Phase	three phase
Number of poles	2
Voltage supply	220 – 242 V (star) and 380 – 420 V (delta)
Motor speed controller	MC LTE-B0015-2B1-1-4

Fig. 3: Three shallow grooves in the outer-raceway of a bearing simulating an outer-raceway fault.

TABLE IV: Infrared thermal camera details

Property	Value
Thermal camera	FLIR SC655
Capture speed	6.25 frames per second (fps)
Resolution	640 x 480 pixels
Distance: camera - housing	38 cm
Emissivity	0.9
Lens	Macro lens
Spectral range	7.5 - 13 μm
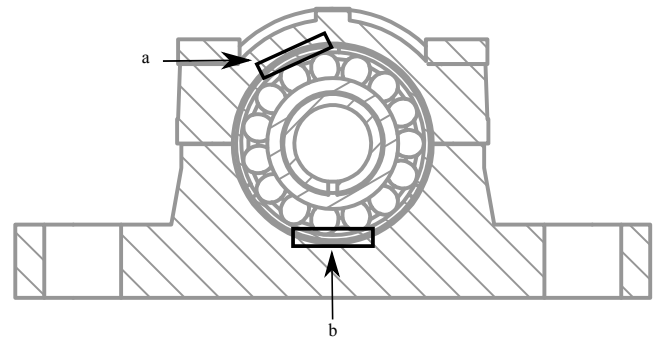
Fig. 4: Section view of where the outer raceway faults for data set one (a) and two (b).

ual bearing is superficially reduced (1.5 g reduction). Similarly, for the EILBs the grease in the bearings is decreased more (0.75 g reduction). For the hard particle faults (i.e. lubricant contamination) in data set two, 0.02 g of iron particles are mixed in the lubricant of the bearings. Experiments showed that the chosen lubricant conditions support the objectives of the research well.

To complete the data sets, all the different bearing conditions are also tested during imbalance, this is done by adding bolts to the rotor at a radius of 5.4 cm. The weight of the bolts can be found in TABLE I and II. The weight of the bolts results in imbalance that is below the G1 tolerance determined according to the ISO 1940. This is a strict tolerance, but ensures that the proposed approach will also work for scenarios where a less strict tolerance is suitable.
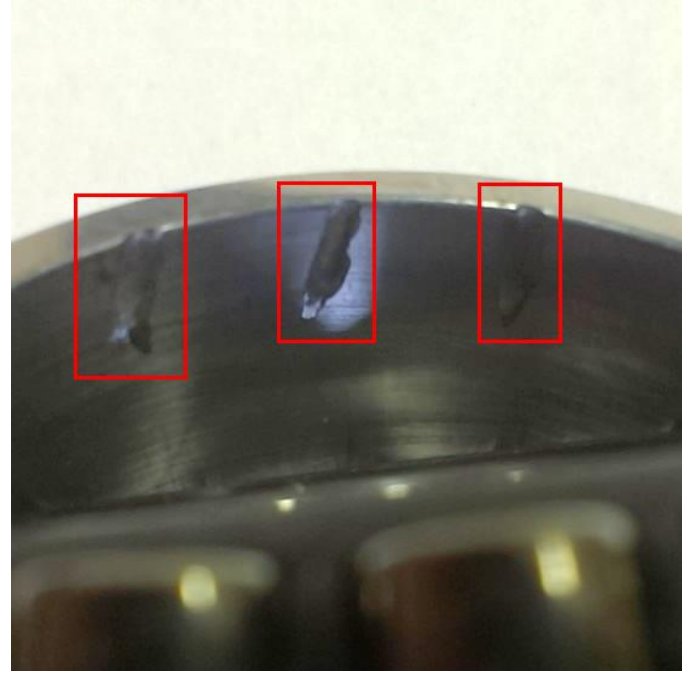
B. Data set

Every condition in both data sets is created for five different bearings. By using multiple bearings, variability is introduced in the data set due to manufacturing, mounting and grease distribution. Each bearing is run for one hour, and for the last 10 minutes —when steady-state is reached— infrared thermal video is captured together with accelerometer measurements. For data set one, in total, 5 bearings × 8 conditions = 40 recordings are made. For data set two, 5 bearings × 12

conditions = 60 recordings are made. An example of an infrared thermal image of the bearing housing can be seen in Fig. 5.
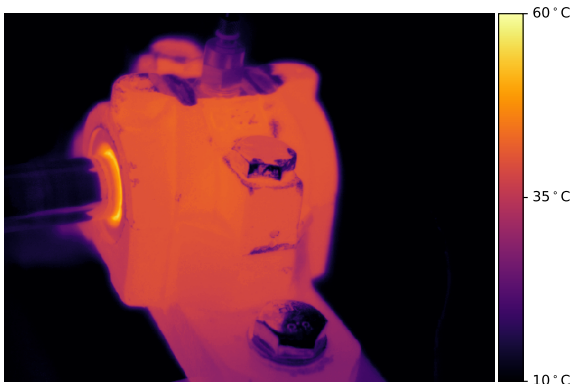

Fig. 5: Example of an infrared thermal image of the bearing housing.

III. METHODOLOGY

In our data sets, each vibration and infrared thermal recording is assigned two labels, i.e. one for the machine condition and one for the bearing condition. Hence, we regard the fault/condition detection task as a multi-label classification problem and propose a system with two classification pipelines. The first pipeline identifies the gradation of imbalance in the rotating machine and the second pipeline identifies the bearing's condition. In the end, the results of both pipelines are combined to identify the present condition of the rotating machine. This approach reduces the number of classes to be identified by a single classifier. For example, in data set two, there are 12 unique classes. When split into a multi-label problem, 3 unique classes remain for the first classifier and 4 for the second classifier.

In order to combine vibration data with infrared thermal data, feature fusion is employed. Feature fusion entails that features are first extracted from the respective data streams and are subsequently combined in feature vectors which are provided to the classification algorithm. Therefore, in this section, first, the features are discussed per pipeline. Features are extracted from the vibration data as well as the infrared thermal imaging data. Next, the classification step is discussed.

A. Features: Pipeline one

The goal of pipeline one is to detect the gradation of imbalance regardless of the bearing condition. First, the features extracted from the accelerometer measurements are discussed and subsequently, the features from the infrared thermal imaging data.

1) Vibration measurements: Using expert knowledge and signal processing, it is possible to detect imbalance by observing an unusually high amplitude at the rotation frequency of the machine. Hence, the amplitude at the rotation frequency is extracted as feature[1]. The first step to extract this feature is windowing. The goal of windowing is to minimize the influence of noise and outliers in the feature extraction step. A window contains one minute of vibration data, and overlaps by 50 % with its neighbouring window. The size of a window was determined empirically and with the purpose of being able to detect the condition of the machine every 30 seconds. Next, the Discrete Fourier Transform is calculated using the windowed signals. A single spectrum covers 25.6 kHz and the resulting frequency resolution is 0.0167 Hz. Finally, from this frequency spectrum, the amplitude at the rotation frequency, i.e. 25 Hz, is extracted as feature. A plot showing the peak at 25 Hz is shown in Fig. 6. This feature is extracted from the vibration signals of both accelerometers. As this is done per window, in the end there are 19 samples per 10 minute recording, each containing two features. The number of windows is determined using Eq. 1, where s is the length of the signal –which is 600 seconds–, w the window length –which is 60 seconds– and o the amount of overlap between the windows expressed as a percentage –which is 50 %–.

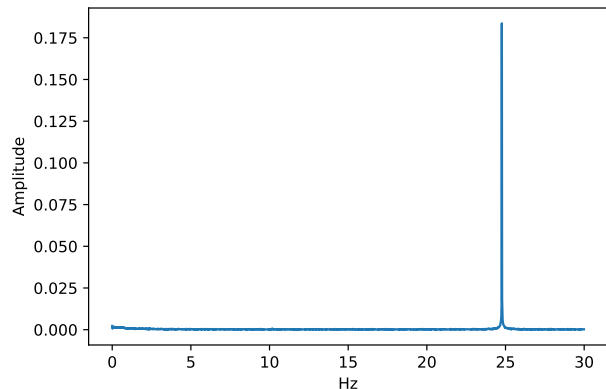$$\text{number of windows} = \frac{s - w}{w * o} + 1 \qquad (1)$$

Fig. 6: Example of the peak at the rotation frequency. Note that no spectral leakage occurs.

2) Infrared thermal imaging: In order to extract features from the infrared thermal imaging data to detect imbalance in the set-up, first windowing is applied similar to the vibration data processing. A window contains one minute of infrared thermal video and 30 seconds overlap. In the second step, consecutive frames are differenced, as such an infrared thermal video will reveal imbalance introduced oscillations. To detect

[1]It should be noted that a high amplitude at the rotation frequency can also be attributed to for example misalignment. However, we solely consider imbalance in this article and designed the experiment as such. If other faults, that cause a rise in the amplitude at the rotation frequency, should be detected, additional features should be engineered (i.e. amplitude at three times the rotation frequency for misalignment)

the degree of imbalance, the difference between frames needs to be quantified resulting in features. In the third step, for each differenced frame, along each column of pixels, the sum of the pixels is taken. Similarly, along each row of pixels, the sum of the pixels is taken. The operation on the rows can be expressed as $\mathbf{y} = \mathbf{Xj}$ and the operation on the columns as $\mathbf{x} = \mathbf{X}^T\mathbf{i}$; where $\mathbf{X} \in \mathbb{R}^{m \times n}$ represents a frame with m rows and n columns; $\mathbf{i} = [1, 1, ..., 1]^T \in \mathbb{R}^m$; $\mathbf{j} = [1, 1, ..., 1]^T \in \mathbb{R}^n$ and thus $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$. Examples of the output vectors displayed in line graphs along the image's axis are shown in Fig. 7. As can be seen, the width (i.e. range) of these graphs indicate the amount of movement in the differenced frame. Generally, if there is imbalance, the width will be larger. Hence, the standard deviation, calculated using these output vectors, is chosen as a feature. In the end, the extracted features are averaged per window to remove the effect of noise and possible outliers.

B. Features: pipeline two

The goal of pipeline two is to identify the specific condition of the bearing regardless of the imbalance gradation in the machine. Again, first the features extracted from the accelerometer measurements are discussed and subsequently, the features from the infrared thermal imaging data.

1) Vibration measurements: When a rolling element hits an aberration in the outer raceway, the natural frequency of the raceway is excited. This action results in a high frequency burst of energy which decays rapidly. Afterwards, the natural frequency is excited again as the next rolling element hits the fault. This high frequency impulse is superimposed, that is, amplitude modulated, on a carrier signal which originates from the rotating machine. To identify a fault, it is necessary to identify the frequency of occurrence of these high energy bursts. Therefore, an established technique, i.e. envelope detection, is employed [23]. First, a high pass filter is applied. All frequencies below 1 kHz (determined empirically), such as the carrier frequency, are removed. Next, the envelope signal is determined which will have a frequency equal to the frequency of occurrence of the high energy bursts. The envelope is constructed by taking the magnitude of the analytical signal, which is computed by the Hilbert Huang transform. When there is an outer-raceway fault, the frequency of the envelope signal will manifest itself at the ball pass frequency of the outer raceway (BPFO).

The BPFO is calculated using Eq. (2), where n is the number of rolling elements in the rolling element bearing, f the rotation frequency, d the diameter of the rolling elements, D the diameter of the rolling-element cage and α the contact angle. This results in a BPFO at 150.41 Hz for the chosen bearings. This envelope detection method is applied on the vibration signal coming from the accelerometer on top of the housing. From the resulting frequency spectrum of the envelope signal, the amplitude at the BPFO is chosen as feature.

Additionally, three statistical features are also extracted, i.e.: kurtosis, crest factor and the root-mean-square (RMS). Both the kurtosis and the crest factor are widely used statics in bearing fault detection [23]. Additionally, RMS has been shown to be indicative of the amount of separation between the rolling elements and the raceways due to the lubricant in a linear bearing [24]. Note that all the features, i.e. the amplitude at the BPFO, kurtosis, crest factor and the RMS are again calculated on windows extracted from the signal similar to as what is done for pipeline one.

2) Infrared thermal imaging: First, in order to increase robustness against environmental temperature changes, the ambient temperature, measured by the thermocouples, is subtracted from every pixel's temperature value as to end up with relative temperatures in the infrared thermal videos. Second, the relevant components of the set-up are segmented from the image using a threshold determined by the Otsu algorithm [25], a commonly used background segmentation algorithm.

Third, similar to pipeline one, the infrared thermal video is subdivided into overlapping windows. Next, from the preprocessed infrared thermal imaging data, data-driven features are extracted using image processing. These features are extracted from the individual frames, but averaged over the extracted windows. As a first feature, the standard deviation (σ) of the pixel values is chosen. As the temperature distribution varies for the different bearing conditions, the standard deviation is a suitable descriptor, e.g. when a grease reservoir is present the standard deviation will be higher. The standard deviation of the pixel values is not sufficient to discriminate between the bearing conditions. Therefore additional features are used.

The second feature is a measurement of concentration related to the spatial temperature distribution. The chosen feature is called the Moment of Light or the second-order moment of the pixels collectively containing 20 % of the brightest pixels (M_{20}) [26]. The M_{20} is influenced by the presence of a grease reservoir. Hence, it is very useful to indicate if a grease reservoir is present or not. Furthermore, if the M_{20} and the standard deviation are combined, a clear distinction can be made between the samples created when a grease reservoir is present, i.e. healthy bearings (HB) and the ones with an outer-raceway fault (ORF), and the samples without a grease reservoir, i.e. extremely and mildly inadequately lubricated bearings (EILB and MILB).

Based on these two features, it is not possible to detect the level of lubricant degradation as the EILB(-IM) and MILB(-IM) samples overlap. Therefore, the Gini coefficient is added which proves to be useful to detect the different levels of lubrication inadequacy. However, none of these features clearly help to distinguish between a bearing with an ORF and a HB as is also observable when plotting the standard deviation of the pixels and the Gini coefficient in Fig. 8. The scatter plots show that there is some overlap between the ORF and HB samples[2]. This illustrates the weakness of the infrared thermal imaging-based approach, which can be solved by combining

$$BPFO = \frac{1}{2}nf(1 - \frac{d}{D}\cos\alpha) \qquad (2)$$

---

[2]For more information regarding these feature the reader is referred to [9].
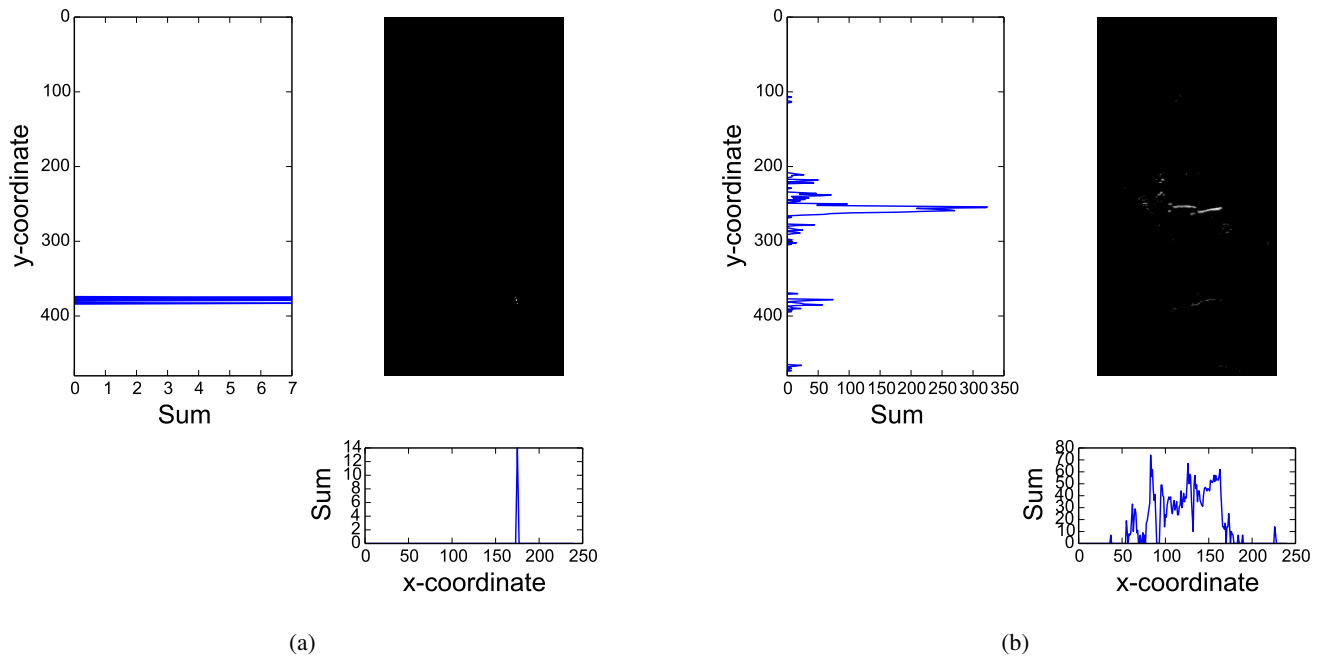
Fig. 7: Examples of the line graphs representing the sums along the x axis and y axis for a healthy bearing when the set-up is in balance (a) and imbalance (b). Note that in (a), the differenced frame is almost entirely black due to the fact that there is no movement in the infrared thermal video.

the infrared thermal features with the features extracted from the vibration measurements.

C. Classification

At this point, the features which are extracted from the vibration measurements and the infrared thermal imaging data in each pipeline have been presented and discussed. Next, per pipeline, these features are provided to classification algorithms. As mentioned at the start of this section, feature fusion is used to combine the two types of sensor data. Feature fusion is chosen to enable a single classifier to use information from both modalities. Therefore, both the features extracted from the vibration measurements and the features extracted from the infrared thermal imaging data are combined in feature vectors and provided to classification algorithms. To illustrate: one trial using the set-up results in 10 minutes of vibration measurements and 10 minutes of infrared thermal video. After applying windowing and feature extraction we end up with 19 samples, each containing two features for the 10 minutes vibration signal for pipeline one. For the 10 minutes infrared thermal video we end up with 19 samples each containing two features (also for pipeline one). By fusing (i.e. concatenating) these two sets of samples together, we end up with 19 samples containing four features. For dataset one and pipeline one there are in total 760 samples (i.e.: 19 samples × 8 conditions × 5 bearings) each containing 4 features.

In order to remove features that do not have a positive impact on the classification results, (exhaustive) feature selection was done. In TABLE V, the selected features per pipeline are listed. In this table, it can for example be seen that for imbalance detection the amplitude at the rotation frequency

(vibration feature) together with the standard deviation along the y-axis (infrared thermal imaging feature) are used.

TABLE V: Features used for the multi-sensor system for both data sets and pipelines. Note that pipeline one detects the different imbalance gradations and pipeline two the different bearing conditions.

Data set	Pipeline	Vibration features	Infrared thermal imaging features
One	One	Amplitude at the rotation frequency	Standard deviation along the y-axis
One	Two	Amplitude at the BPFO Amplitude at the rotation frequency	Gini coefficient Standard deviation M_{20}
Two	One	Amplitude at the rotation frequency	Standard deviation along the y-axis
Two	Two	Amplitude at the BPFO frequency RMS Kurtosis Crest Amplitude at the rotation frequency	Standard deviation

As classification algorithm, in this paper, the random forest classifier is used [27]. A random forest classifier is an ensemble approach that combines several decision trees. The decision trees are built, i.e. learned, using a subset of features and a subset of samples which are chosen at random with replacement. A decision tree learns to optimally split the training data into regions by minimizing the gini score using an optimal feature and feature value. A classification can subsequently be done by traversing the decision tree from top to bottom and assigning the sample to the class of the final node in the decision tree. It should be noted that these decision trees can be created in parallel, making the algorithm very efficient. Random forest classifiers have many additional advantages, such as they are not prone to overfit, require little preprocessing, can deal with both continuous
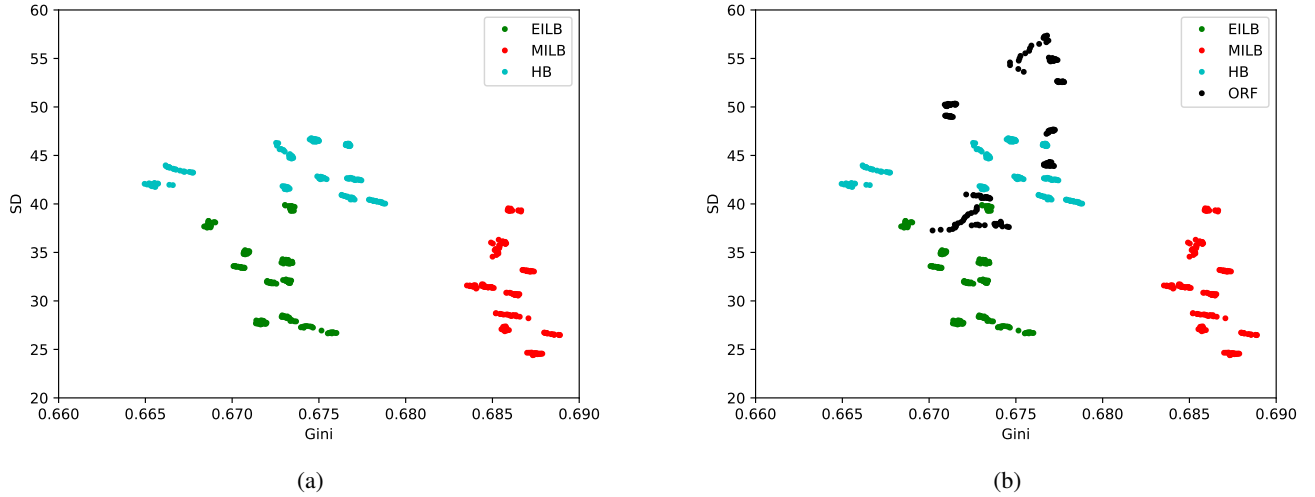
Fig. 8: 2D plots of the Gini coefficient with and without the ORF samples. (a) No overlap exist between the extremely and mildly inadequately lubricated bearing (EILB and MILB) samples. (b) Some overlap exists between the samples from the healthy bearings and the ones with an outer-raceway fault (HB and ORF).

and discrete data and require very few hyperparameters to be tuned. Furthermore, in two very recent comparisons of a plethora of machine learning algorithms, the random forest classifier always ended up as one of the best performing approaches currently available [28], [29].

To optimize the hyperparameters (e.g.: number of trees and maximum depth of the individual trees) grid search was applied. This is done by defining a set of plausible hyperparameters and testing every combination of these hyperparameters. For brevity, only the results obtained using the optimal hyperparameters are reported next.

IV. RESULTS

Within this section, the results are presented and discussed of the proposed multi-sensor approach for both data sets.

A. Evaluation score

In order to evaluate how well the system performs, accuracy is chosen as evaluation score. This score specifies the ratio between the number of samples that are correctly classified and all the samples in total. By optimizing the accuracy, the false positives (i.e. flagging a sample as containing data from a faulty bearing, while no fault or damage is present) and false negatives (i.e. flagging a sample as containing data from a healthy bearing, while a fault or damage is present) are minimized.

B. Evaluation procedure

To objectively evaluate the performance of the systems, per-bearing cross validation is employed. This procedure is a type of cross validation wherein all the data, i.e. extracted features from the measurements, from all bearings but one are used to train the machine learning system and the data collected from trials using the remaining bearing is given to the trained

machine learning system for testing. This is done n times, where n equals the number of bearings in the data set (i.e. 5). To average out the fluctuations in the results due to randomness which are inherent to random forest classifiers, the five-fold cross-validation is executed 10 times.

In order to put the results of the multi-sensor system in perspective, we also include the results of two single-sensor systems that use either features extracted from vibration data or infrared thermal imaging data.

C. Results: data set one

The results for the two single sensor systems and the multi-sensor system can be seen in TABLE VI. A perfect imbalance detection is achieved with both the single-sensor systems and the multi-sensor system. However, detecting the specific bearing condition (EILB, MILB, ORF or HB) is more difficult. When observing the confusion matrix for the single-sensor infrared thermal imaging-based system (Fig. 9b) and the vibration based system (Fig. 9a), it can be concluded that the infrared thermal imaging-based system has difficulty detecting outer-raceway faults. Conversely, the vibration-based system has more difficulty detecting lubrication related conditions. Both single-sensor systems exhibit a specific weakness that can be compensated by the strengths of the other sensor. Hence, combining them in a multi-sensor solution results in a better overall system (Fig. 9c). Furthermore, there is an additional general improvement in all classes due to the interaction of features in the machine learning algorithms on top of eliminating the weaknesses of the respective single sensor solutions. This way, for example, outer-raceway faults can be classified perfectly (100 % accuracy), while at most 90 % accuracy could be achieved using vibration data only, respectively 74 % using infrared thermal imaging data only. Using a multi-sensor solution, the classification task becomes

easier as more information is available resulting in a perfect classification results.

TABLE VI: Classification results for data set one. σ is the standard deviation. IR = infrared based, VIB = vibration based. MILB = mildly inadequately lubricated bearing, EILB = extremely inadequately lubricated bearing, HB = healthy bearing, ORF = bearing with an outer-raceway fault.

IR	VIB	Conditions	Accuracy
IR		MILB, EILB, HB, ORF	88.25 % (σ = 8.07 %)
	VIB	MILB, EILB, HB, ORF	87.25 % (σ = 8.10 %)
IR	VIB	MILB, EILB, HB, ORF	100.00 % (σ = 0.00 %)
IR		balance and imbalance	100.0 % (σ = 0.00 %)
	VIB	balance and imbalance	100.0 % (σ = 0.00 %)
IR	VIB	balance and imbalance	100.0 % (σ = 0.00 %)
IR		All 8 conditions	88.25 % (σ = 8.07 %)
	VIB	All 8 conditions	87.25 % (σ = 8.10 %)
IR	VIB	All 8 conditions	100.0 % (σ = 0.00 %)

By using a random forest classifier it is possible to extract how important each feature is to the model [27], which is listed in TABLE VII for the multi-sensor bearing fault detection. The table illustrates that both the information coming from the infrared thermal imaging data and the vibration data is important for the decision making of the classification model. The effects of adding a certain feature are also directly noticeable in the results (TABLE VI). For example, the amplitude at the ball pass frequency of the outer raceway is very important to the model and adding this feature to the infrared thermal imaging features results in an accuracy gain of 9.25 %. In total, by using the strengths of both sensors, the accuracy rises by 11.75 % compared to an infrared thermal imaging-based system and 12.75 % compared to a vibration-based system.

TABLE VII: Feature importance in the final multi-sensor system according to the random forest classifier. The first three features are extracted from the infrared thermal imaging (IR) data and the last 2 features from the vibration (VIB) data. σ is the standard deviation.

Feature	Importance of the feature
Gini Coefficient (IR)	29.41 % (σ = 1.07 %)
Standard deviation (IR)	22.22 % (σ = 3.84 %)
M_{20} (IR)	20.29 % (σ = 3.26 %)
Amplitude at the BPFO (VIB)	25.63 % (σ = 1.82 %)
Amplitude at the rotation frequency (VIB)	2.45 % (σ = 0.40 %)

D. Results: data set two

The results of the two single-sensor system and the multi-sensor system can be seen in TABLE VIII. For pipeline two, a relatively low accuracy is achieved due to the overlap between the ORF and HB samples using infrared thermal imaging features. As the confusion matrix in Fig. 10b illustrates, the ORF class is hard to detect. This was also noticeable for data set one. Conversely, vibration-based bearing condition detection achieves a very high accuracy, and the classes do not get confused very often (Fig. 10a). When features from both systems are combined, the three bearing conditions can be detected perfectly (Fig. 10c).

TABLE VIII: Classification results for data set two. σ is the standard deviation. IR = infrared based, VIB = vibration based. HB = healthy bearing, ORF = bearing with an outer-raceway fault, HP = bearing with hard particles.

IR	VIB	Conditions	Accuracy
IR		HP, ORF, HB	65.00 % (σ = 16.16 %)
	VIB	HP, ORF, HB	91.67 % (σ = 12.91 %)
IR	VIB	HP, ORF, HB	100.00 % (σ = 0.00 %)
IR		Imbalance gradation	88.33 % (σ = 12.47 %)
	VIB	Imbalance gradation	75.00 % (σ = 9.13 %)
IR	VIB	Imbalance gradation	90.00 % (σ = 6.24 %)
IR		All 12 conditions	55.00 % (σ = 11.31 %)
	VIB	All 12 conditions	66.67 % (σ = 21.08 %)
IR	VIB	All 12 conditions	90.00 % (σ = 6.24 %)

The detection of the amount of imbalance is more difficult. Generally, the infrared thermal imaging-based system seems to outperform the vibration-based system. However, when both features from both modalities are used, the overall accuracy increases. The confusion matrices in Fig. 11 indicate that the classifier can confuse small imbalance gradation differences.

Pipeline one detects the different amounts of imbalance and pipeline two detects the specific bearing conditions. In total there are 12 conditions. When solely using infrared thermal imaging data the accuracy score is 55.00 % (σ = 11.31 %) and when solely using vibration data the accuracy score is 66.67 % (σ = 21.08 %). However, when both modalities are combined there is a major improvement as the multi-sensor system achieves an accuracy score of 90.00 % (σ = 6.24 %). When observing the overall confusion matrix (Fig. 12), it can be seen that when imbalance occurs together with contamination in the lubricant, it is more difficult to detect the conditions, hence there is some small interaction between these two types of faults.

For completeness, a late fusion approach was also created and tested for data set two. Late fusion entails the creation of a machine learning model for each sensor stream. Subsequently, the output of the machine learning models are combined to determine a final prediction. As the chosen algorithm, i.e. random forest classifier, provides probabilities of its predictions, we opted to combine the probabilities of the machine learning models created per sensor stream. The results of the experiments are listed in TABLE IX. As can be seen, the late fusion approach performs worse compared to the feature fusion approach. This can be contributed to the fact that late fusion does not allow feature interaction, which we determined to be beneficial.
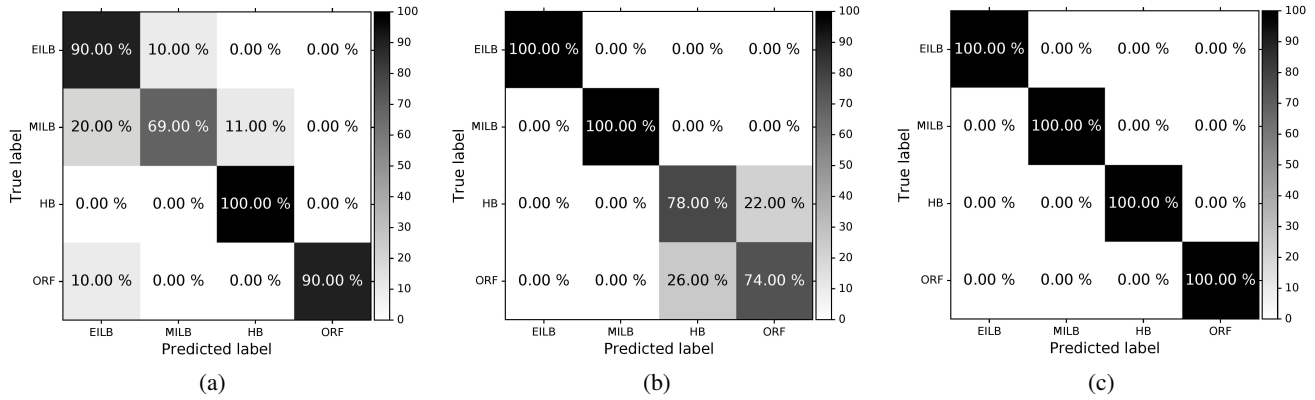
Fig. 9: Confusion matrices of the (a) vibration-based fault detection system, (b) the infrared thermal imaging-based fault detection system and the (c) multi-sensor fault detection system for data set one.
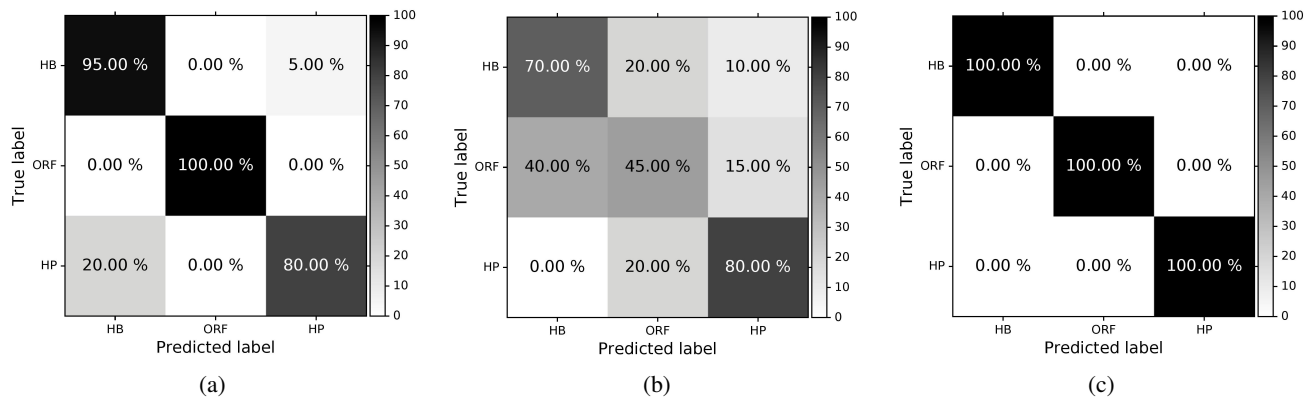

Fig. 10: Confusion matrices of the (a) vibration-based fault detection system, (b) the infrared thermal imaging-based fault detection system and the (c) multi-sensor fault detection system for data set two.
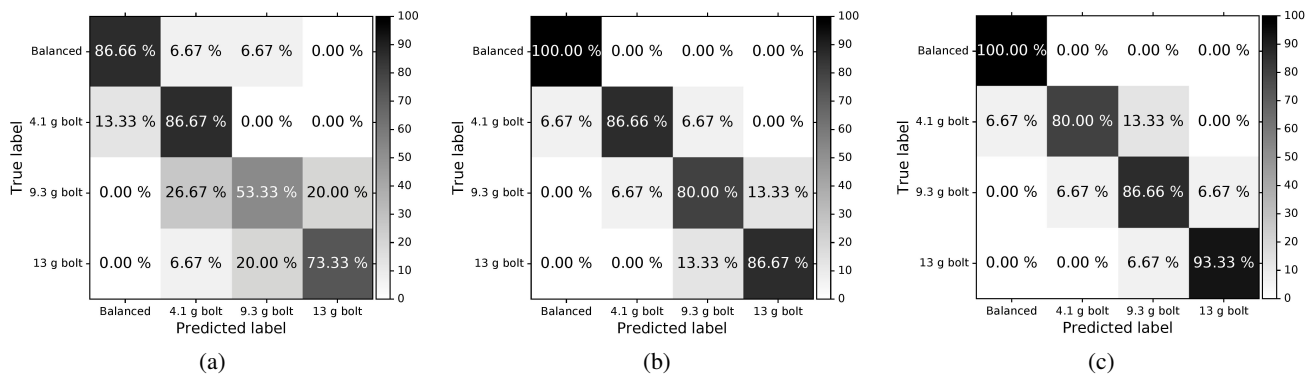

Fig. 11: Confusion matrices for (a) vibration-based imbalance detection, (b) the infrared thermal imaging-based imbalance detection and (c) multi-sensor based imbalance detection for data set two.
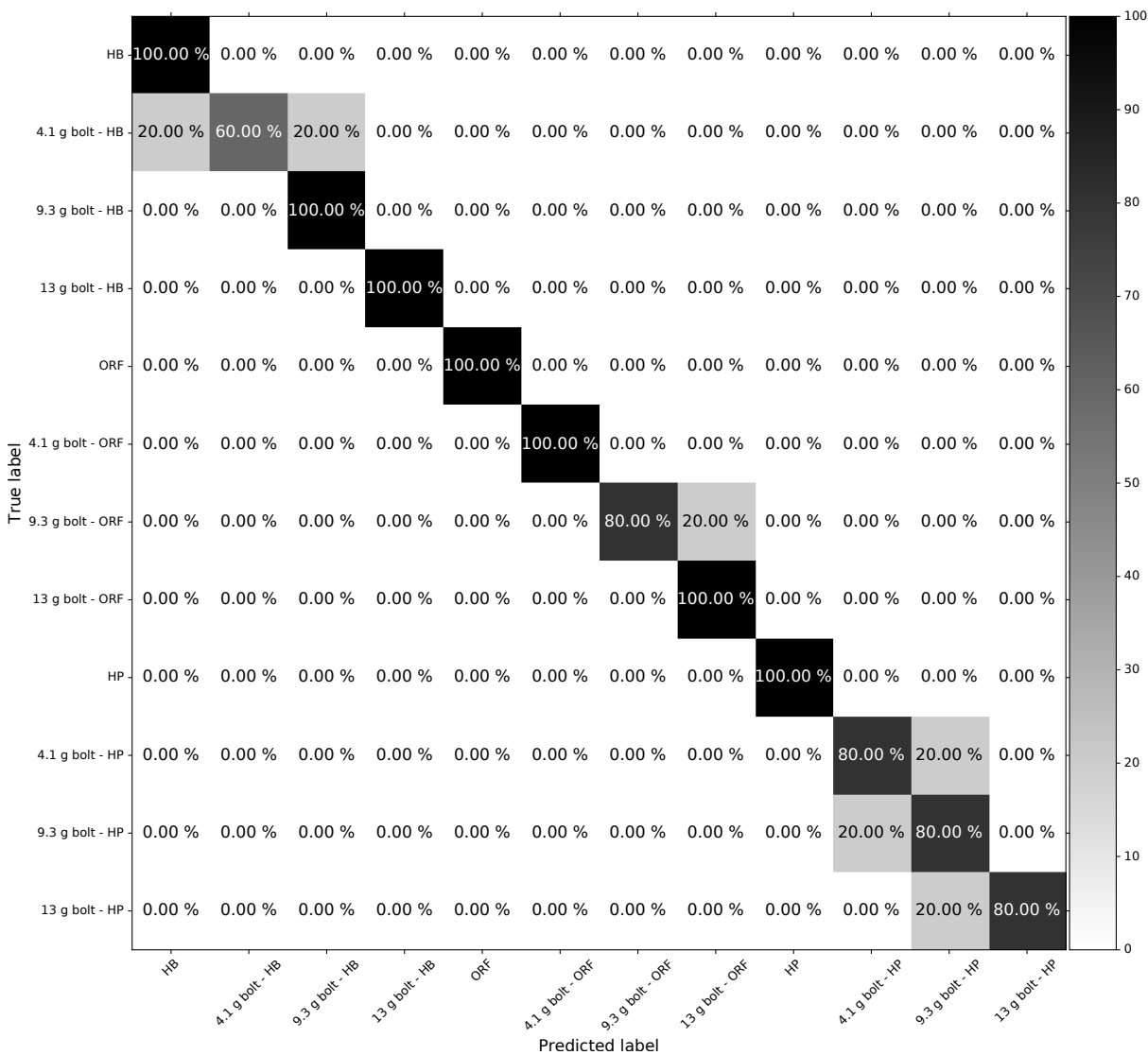
Fig. 12: Confusion matrix of the multi-sensor system for data set 2 for all 12 condition.

TABLE IX: Classification results for data set two using late fusion. σ is the standard deviation. IR = infrared based, VIB = vibration based, HB = healthy bearing, ORF = bearing with an outer-raceway fault, HP = bearing with hard particles.

IR	VIB	Conditions	Accuracy
IR	VIB	HP, ORF, HB	90.00 % ($\sigma = 6.24$ %)
IR	VIB	Imbalance gradation	72.80 % ($\sigma = 12.47$ %)
IR	VIB	All 12 conditions	72.80 % ($\sigma = 12.47$ %)

V. CONCLUSION

In this article, a multi-sensor fault detection system for rotating machinery is proposed that uses both infrared thermal imaging and vibration data. In order to achieve this, feature fusion is used wherein model-driven features are extracted from the vibration measurements, and data-driven features from the infrared thermal imaging data. Subsequently, the extracted features are fused together and provided to random forest classifiers for the actual fault detection. We show that by using the multi-sensor approach, the shortcomings, inherent to the modality (e.g. vibrations or heat), are compensated by the other modality type. Therefore, the system is able to outperform single sensor based systems on our two data sets. We demonstrate that our proposed multi-sensor approach that includes infrared thermal imaging can provide a significant improvement in fault detection performance (e.g. 35 % absolute improvement for data set two). However, it should be noted that the financial costs increases as two sensors are required. Hence, a cost-benefit analysis is required when considering a multi-sensor approach.

REFERENCES

[1] T. Touret, C. Changenet, F. Ville, M. Lalmi, and S. Becquerelle, "On the use of temperature for online condition monitoring of geared systems – a review," *Mechanical Systems and Signal Processing*, vol. 101, no. Supplement C, pp. 197 – 210, 2018.

[2] A. Widodo, D. Satrijo, T. Prahasto, G.-M. Lim, and B.-K. Choi, "Confirmation of Thermal Images and Vibration Signals for Intelligent Machine Fault Diagnostics," *International Journal of Rotating Machinery*, vol. 2012, pp. 1–10, 2012.

[3] H. Fandino-Toro, O. Cardona-Morales, J. Garcia - Alvarez, and G. Castellanos-Dominguez, "Bearing Fault Identification using Watershed-Based Thresholding Method," ser. Lecture Notes in Mechanical Engineering, 2014, pp. 137–147.

[4] V. T. Tran, B.-S. Yang, F. Gu, and A. Ball, "Thermal image enhancement using bi-dimensional empirical mode decomposition in combination with relevance vector machine for rotating machinery fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 38, no. 2, pp. 601–614, 2013.

[5] G.-m. Lim, Y. Ali, and B.-s. Yang, "The Fault Diagnosis and Monitoring of Rotating Machines by Thermography," in *Engineering Asset Management and Infrastructure Sustainability*, J. Mathew, L. Ma, A. Tan, M. Weijnen, and J. Lee, Eds. Springer London, 2012, pp. 557–565.

[6] G.-M. Lim, D.-M. Bae, and J.-H. Kim, "Fault diagnosis of rotating machine by thermography method on suport vector machine," *Journal of Mechanical Science and Technology*, vol. 28, no. 8, pp. 2948–2952, 2014.

[7] A. Younus and B. Yang, "Wavelet co-efficient of thermal image analysis for machine fault diagnosis," in *International Conference on Mechanical Engineering*, 2009, pp. 1–7.

[8] A. M. Younus and B.-S. Yang, "Intelligent fault diagnosis of rotating machinery using infrared thermal image," *Expert Systems with Applications*, vol. 39, pp. 2082–2091, 2012.

[9] 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.

[10] O. Janssens, V. Slavkovikj, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. V. de Walle, and S. V. Hoecke, "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331 – 345, 2016.

[11] T. Bai, L. Zhang, L. Duan, and J. Wang, "Nsct-based infrared image enhancement method for rotating machinery fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2293–2301, 2016.

[12] O. Janssens, R. V. de Walle, M. Loccufier, and S. V. Hoecke, "Deep learning for infrared thermal image based machine health monitoring," *IEEE/ASME Transactions on Mechatronics*, no. 99, pp. 1–10, 2017.

[13] D. Lopez-Perez and J. Antonino-Daviu, "Application of infrared thermography to failure detection in industrial induction motors: case stories," *IEEE Transactions on Industry Applications*, 2017.

[14] A.-b. Ming, W. Zhang, Z.-y. Qin, and F.-l. Chu, "Dual-impulse response model for the acoustic emission produced by a spall and the size evaluation in rolling element bearings," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6606–6615, 2015.

[15] S. Singh and N. Kumar, "Detection of bearing faults in mechanical systems using stator current monitoring," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1341–1349, 2017.

[16] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the case western reserve university data: A benchmark study," *Mechanical Systems and Signal Processing*, vol. 64 – 65, pp. 100 – 131, December 2015.

[17] 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.

[18] H. Henao, G.-A. Capolino, M. Fernandez-Cabanas, F. Filippetti, C. Bruzzese, E. Strangas, R. Pusca, J. Estima, M. Riera-Guasp, and S. Hedayati-Kia, "Trends in fault diagnosis for electrical machines: A review of diagnostic techniques," *IEEE industrial electronics magazine*, vol. 8, no. 2, pp. 31–42, 2014.

[19] B. P. Graney and K. Starry, "Rolling element bearing analysis," *Materials Evaluation*, vol. 70, no. 1, pp. 78 – 85, 2012.

[20] A. D. Nembhard, J. K. Sinha, A. J. Pinkerton, and K. Elbhbah, "Combined vibration and thermal analysis for the condition monitoring of rotating machinery," *Structural Health Monitoring*, vol. 13, no. 3, pp. 281–295, 2014.

[21] A. Garcia-Ramirez, L. Morales-Hernandez, R. Osornio-Rios, A. Garcia-Perez, and R. Romero-Troncoso, "Thermographic technique as a complement for mcsa in induction motor fault detection," in *IEEE International Conference on Electrical Machines (ICEM)*, 2014, pp. 1940–1945.

[22] Schaeffler, "Fag split plummer block housings of series snv," https://goo.gl/HvzXpN, 2015, [Online; accessed 28-February-2018].

[23] E.-T. Idriss and J. Erkki, "A summary of fault modelling and predictive health monitoring of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 6061, pp. 252 – 272, 2015.

[24] H. Ohta, Y. Nakajima, S. Kato, and H. Tajimi, "Vibration and acoustic emission measurements evaluating the separation of the balls and raceways with lubricating film in a linear bearing under grease lubrication," *Journal of Tribology*, vol. 135, no. 4, 2013.

[25] O. Nobuyuki, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62– 66, Jan 1979.

[26] J. M. Lotz, J. Primack, and P. Madau, "A New Nonparametric Approach to Galaxy Morphological Classification," *Astronomical Journal*, vol. 128, pp. 16—182, Jul. 2004.

[27] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2001.

[28] R. S. Olson, W. L. Cava, Z. Mustahsan, A. Varik, and J. H. Moore, "Data-driven advice for applying machine learning to bioinformatics problems," in *Pacific Symposium on Biocomputing*, 2018, pp. 192–203.

[29] C. Zhang, C. Liu, X. Zhang, and G. Almpanidis, "An up-to-date comparison of state-of-the-art classification algorithms," *Expert Systems with Applications*, vol. 82, pp. 128–150, 2017.

REFERENCES

[1] T. Touret, C. Changenet, F. Ville, M. Lalmi, and S. Becquerelle, "On the use of temperature for online condition monitoring of geared systems – a review," *Mechanical Systems and Signal Processing*, vol. 101, no. Supplement C, pp. 197 – 210, 2018.

[2] A. Widodo, D. Satrijo, T. Prahasto, G.-M. Lim, and B.-K. Choi, "Confirmation of Thermal Images and Vibration Signals for Intelligent Machine Fault Diagnostics," *International Journal of Rotating Machinery*, vol. 2012, pp. 1–10, 2012.

[3] H. Fandino-Toro, O. Cardona-Morales, J. Garcia - Alvarez, and G. Castellanos-Dominguez, "Bearing Fault Identification using Watershed-Based Thresholding Method," ser. Lecture Notes in Mechanical Engineering, 2014, pp. 137–147.

[4] V. T. Tran, B.-S. Yang, F. Gu, and A. Ball, "Thermal image enhancement using bi-dimensional empirical mode decomposition in combination with relevance vector machine for rotating machinery fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 38, no. 2, pp. 601–614, 2013.

[5] G.-m. Lim, Y. Ali, and B.-s. Yang, "The Fault Diagnosis and Monitoring of Rotating Machines by Thermography," in *Engineering Asset Management and Infrastructure Sustainability*, J. Mathew, L. Ma, A. Tan, M. Weijnen, and J. Lee, Eds. Springer London, 2012, pp. 557–565.

[6] G.-M. Lim, D.-M. Bae, and J.-H. Kim, "Fault diagnosis of rotating machine by thermography method on suport vector machine," *Journal of Mechanical Science and Technology*, vol. 28, no. 8, pp. 2948–2952, 2014.

[7] A. Younus and B. Yang, "Wavelet co-efficient of thermal image analysis for machine fault diagnosis," in *International Conference on Mechanical Engineering*, 2009, pp. 1–7.

[8] A. M. Younus and B.-S. Yang, "Intelligent fault diagnosis of rotating machinery using infrared thermal image," *Expert Systems with Applications*, vol. 39, pp. 2082–2091, 2012.

[9] 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.

[10] O. Janssens, V. Slavkovikj, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. V. de Walle, and S. V. Hoecke, "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331 – 345, 2016.

[11] T. Bai, L. Zhang, L. Duan, and J. Wang, "Nsct-based infrared image enhancement method for rotating machinery fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2293–2301, 2016.

[12] O. Janssens, R. V. de Walle, M. Loccufier, and S. V. Hoecke, "Deep learning for infrared thermal image based machine health monitoring," *IEEE/ASME Transactions on Mechatronics*, no. 99, pp. 1–10, 2017.

[13] D. Lopez-Perez and J. Antonino-Daviu, "Application of infrared thermography to failure detection in industrial induction motors: case stories," *IEEE Transactions on Industry Applications*, 2017.

[14] A.-b. Ming, W. Zhang, Z.-y. Qin, and F.-l. Chu, "Dual-impulse response model for the acoustic emission produced by a spall and the size evaluation in rolling element bearings," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6606–6615, 2015.

[15] S. Singh and N. Kumar, "Detection of bearing faults in mechanical systems using stator current monitoring," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1341–1349, 2017.

[16] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the case western reserve university data: A benchmark study," *Mechanical Systems and Signal Processing*, vol. 64 – 65, pp. 100 – 131, December 2015.

[17] 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.

[18] H. Henao, G.-A. Capolino, M. Fernandez-Cabanas, F. Filippetti, C. Bruzzese, E. Strangas, R. Pusca, J. Estima, M. Riera-Guasp, and S. Hedayati-Kia, "Trends in fault diagnosis for electrical machines: A review of diagnostic techniques," *IEEE industrial electronics magazine*, vol. 8, no. 2, pp. 31–42, 2014.

[19] B. P. Graney and K. Starry, "Rolling element bearing analysis," *Materials Evaluation*, vol. 70, no. 1, pp. 78 – 85, 2012.

[20] A. D. Nembhard, J. K. Sinha, A. J. Pinkerton, and K. Elbhbah, "Combined vibration and thermal analysis for the condition monitoring of rotating machinery," *Structural Health Monitoring*, vol. 13, no. 3, pp. 281–295, 2014.

[21] A. Garcia-Ramirez, L. Morales-Hernandez, R. Osornio-Rios, A. Garcia-Perez, and R. Romero-Troncoso, "Thermographic technique as a complement for mcsa in induction motor fault detection," in *IEEE International Conference on Electrical Machines (ICEM)*, 2014, pp. 1940–1945.

[22] Schaeffler, "Fag split plummer block housings of series snv," https://goo.gl/HvzXpN, 2015, [Online; accessed 28-February-2018].

[23] E.-T. Idriss and J. Erkki, "A summary of fault modelling and predictive health monitoring of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 6061, pp. 252 – 272, 2015.

[24] H. Ohta, Y. Nakajima, S. Kato, and H. Tajimi, "Vibration and acoustic emission measurements evaluating the separation of the balls and raceways with lubricating film in a linear bearing under grease lubrication," *Journal of Tribology*, vol. 135, no. 4, 2013.

[25] O. Nobuyuki, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62– 66, Jan 1979.

[26] J. M. Lotz, J. Primack, and P. Madau, "A New Nonparametric Approach to Galaxy Morphological Classification," *Astronomical Journal*, vol. 128, pp. 16—182, Jul. 2004.

[27] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2001.

[28] R. S. Olson, W. L. Cava, Z. Mustahsan, A. Varik, and J. H. Moore, "Data-driven advice for applying machine learning to bioinformatics problems," in *Pacific Symposium on Biocomputing*, 2018, pp. 192–203.

[29] C. Zhang, C. Liu, X. Zhang, and G. Almpanidis, "An up-to-date comparison of state-of-the-art classification algorithms," *Expert Systems with Applications*, vol. 82, pp. 128–150, 2017.

Olivier Janssens received his Master degree in industrial engineering focussing on information and communication technology from the University College of West Flanders in 2012. Afterwards he obtained his PhD degree in engineering technology in 2017 at the Ghent University-imec's IDLab. His research focusses on data-driven multi-sensor condition monitoring.

Mia Loccufier received a M.S degree of electromechanical engineer, a M.S. degree of automatic control engineer and a PhD degree in electromechanical engineering from Ghent University. She is a professor at the DySC research group of the Department of Electrical Energy, Systems and Automation, Faculty of Engineering, Ghent University, Belgium. She is a lecturer in mechanical vibrations, structural dynamics and systems dynamics. The field of research is focussed on the dynamics of technical systems. The main research themes are passive control, especially nonlinear tuned mass dampers of mechanical systems and structures, dynamics of rotating machinery; stability and bifurcation analysis of nonlinear systems and structures, control of underactuated mechanical systems.

Sofie Van Hoecke received her MSc in Computer Science from Ghent University in 2003. Following up on her studies, she achieved a PhD in computer science engineering at the Department of Information Technology at the same university. Her research concentrates on the design of multi-sensor architectures, innovative ICT solutions for care, and multi-sensor condition monitoring. Currently, she is an associate professor at Ghent University and senior researcher at IDLab, Ghent University-imec.