Temperature and intensity-based analysis of infrared images for the state assessment of rolling bearings using Deep learning methods

Anonymous CVPR 2021 submission

Paper ID ****

Abstract

Infrared thermography (IRT) has arrose as a promising technology for the state assessment of rotating machinery. In this paper we evaluate the performance thermal based (TBA) and intensity based analysis (IBA) of IRT for the classification of the severity of outer race defect of rolling bearings. We propose EfficientNet as our baseline to evaluate both approaches, TBA and IBA, in the classification of the severity of the defect. Moreover, we propose a hybrid methodology that combines both approaches, and evaluate two ways to implement it: by concatenating thermal matrices to thermal images and using this 4-channel data is input to EfficientNet (4-Hy) and by using thermal images as input to the convolutional backbone of the network while concatenating to the linear classification layers a feature vector extracted from thermal matrices (fv-Hy).

We perform experimentation on different ways to implement both hybrid approaches, trained and tested on IRT data recorded at different speeds and angles over bearings with three levels of severity of outer race defect. Among all approaches, fv-Hy model had the best performance, surpassing our baseline with highest ACA, accurary and f1-score metrics.

1. Introduction

Rotating machinery can be found in almost any moving mechanical system, from machinery in a production line to the turbines in an airplane. In most cases they provide thrust or power in order to assure a proper and safe performance of a system. Their continuous operation is crucial in many supply chains given all the instances that depend on their output. Therefore, companies need to operate rotating machines continuously or reduce as much as possible the down-time due to the repair costs and the penalization for loss of operation. Hence, the failure prevention of rotating machinery has become an important focus regarding maintenance decisions [12].

Accordingly, the proper maintenance of machinery has become a critical activity in many industries, for identifying conditions under which they operate and control the different sources of failure [5]. Maintenance strategies have evolved during the last century, introducing more efficient and complex techniques as the effects of down-time become more dramatic.

We can these strategies into four categories [5]. The first is Run-to-Failure (R2F), where no repairing or restoring actions are taken until the occurrence of a failure. The second is referred to as Preventive Maintenance (PvM), where actions are carried out on time or usage intervals, regardless of the health state of the system in a planned schedule. Here, sometimes unnecessary maintenance is performed and that is why the last two categories are introduced: Condition-based maintenance (CBM) and Predictive Maintenance (PdM). In CBM the system is under continuous monitoring and actions are taken after the verification of one or more indicators of failure or possible failure of the equipment, whereas in PdM, actions are only taken when required, but prediction tools are implemented in order to identify when those actions are likely to be required, allowing for planning and scheduling.

The last two strategies are more complex but better in reducing down-time and increasing service life. CBM and PdM are usually implemented by continuously monitoring physical variables using sensor arrays, whose signals allow to perform analysis of vibrations, acoustics, alignment, lubrication, temperature, etc., which provide means to identify the health state of the system. Nevertheless, most of these techniques require invasive actions and expensive setups of sensors and data acquisition systems. As an alternative, infrared thermography (IRT) has arisen as a non-contact, non-intrusive temperature measuring method, currently used in the inspection of transformers and electrical installations, but still emerging in the state assessment and maintenance of rotating machinery [4].

One of the most omnipresent elements in rotating machinery are rolling bearings which also are susceptible to failure as a consequence of fatigue or aging. Hence, bearings are critical components in the state assessment of rotating machinery based on CBM and PdM. Given the above, in this paper we introduce a deep learning approach to assess the state of rolling bearings using infrared thermography as a promising technology in CBM of rotating machinery. These essential components may suffer from faults such as inner race defect, outer race defect, ball defect, cage defect, low lubrication defect, etc., however, in this paper we use IRT data to evaluate deep learning methods accuracy in classifying the severity of one type of defect on rolling bearings, outer race defect (ORD). We perform this analysis by two different approaches based on IRT output data: temperature based analysis using thermal pixel matrix data and intensity based analysis using thermal images.

2. Related Work

Research on IRT applications of CBM and PdM of rotating machinery has grown given its success over preventive and corrective maintenance techniques among a wide range of industries. Since a higher level of automation is required in CBM and PdM maintenance and modern industries wish to rely the least on expert knowledge. In consequence, machine learning techniques have been applied extensively for fault classification of rotating machinery. Variations and innovations have been proposed in the feature extraction processes, the data recording, and the utilized algorithms. For instance, Younus and Bo-Suk [3] proposed a feature extraction process for classifying fault type (not severity) of rotating machinery using IRT images of a test bench. Here, 2D Dirichlet Wavelet Transform is applied for image decomposition and then Mahalanobis distance and relief algorithm are used for feature selection of the wavelet coefficients. They pass computed features to a Support Vector Machine (SVM) algorithm for classification. Jassens [12] also proposed a classifier for identifying the fault type of rotating machinery, including bearing fault and shaft imbalance conditions, using localized IRT on the bearing's housing. They implement a random forest and an SVM classifier with some novel features as: the Gini coefficient, – commonly used in economics – moment of light (M20) and standard deviation. Authors extracted from IRT video recorded over the bearing case zone.

Duan et al. [7] propose segmentation methods for identifying faulty zones in rotating machinery using IRT on the whole machine instead of localized data. Once faulty zones are detected, first order statistical indicators like mean, standard deviation, kurtosis, skewness, etc., are used to define the representation space of each IRT as vectors passed to an SVM for classification. Deep learning methods have also been applied in the assessment of rotating machinery. Li and Du [16] applied LeNet-5 CNN in the classification of the fault type in a test bed using non-localized IRT data; faults included bearing's race defects and shaft unbalanced conditions. The proposed method has superior performance than traditional vibration approach and other DL models with 98.59% accuracy.

Furthermore, since rolling bearings are identified as critical components in rotating machinery, more research has been done on the suitability of IRT for the state assessment of these components. Jia and Liu [17] made a comparison between the performance of Bag of Words features and CNN features in the classification of rolling bearing fault condition type. Their approach resulted in better performance than the classical vibration data approach using CNNs. On the other hand, Janssens [11] proposed the use of transfer learning with a pretrained CNN from the Visual Geometry Group of Oxford in the classification of fault type of rotating machinery using infrared video, resulting in over 90% accuracy. In addition, Choudhary [4] made the analysis of the suitability of an SVM and Complex Decision Trees classifiers for the identification of faulty type in rolling bearings. Classifiers are trained with feature vectors made of first-order statistical indicators from the 2D Dirichlet Wavelet Transform of IRTs. That analysis was made using thermal pixel matrix data instead of intensity data, resulting in over 95% accuracy. Also, Choudhary [2] compared the performance of a LeNet-5 CNN architecture and Adversarial Neural Network in the classification of six different fault type using thermal images, resulting in 99.80% accuracy of the CNN, outperforming the ANN. All of these papers seek to classify the type of failure, but not the severity of the failure which can be useful in assessing the remaining lifespan of rolling bearings. In IRT applications to rolling bearings state assessment no public data-set or methodology with open source code have been found.

On the other hand, IRT applications different from maintenance have shown great performance in classification tasks. For example, Infrared Breast Thermography (IBM) has arisen as a promising method for early detection of breast temperature anomalies related to breast cancer [9]. In [15], Gogoi and Majumdar compared a thermal based analysis based using thermal pixel matrices (TBA), intensity based analysis using thermal images (IBA) an Tumor Location Matching of IBM in the classification of thermograms as benign, malign or healthy. They trained an SVM with an rbf kernel using features extracted from IBA and TBA and evaluated their performance. They obtained the best results by combining features from TBA and IBA, with 83.22% accuracy. In this paper, we give a similar approach to CBM and PdM application to the state assessment of rolling bearings by comparing TBA and IBA in the classification of the severity of ORD in rolling bearings using IRT data and deep learning methods following the procedures described in the next sections.

3. Data set

In order to assess the severity of ORD on rolling bearings, we define three health states based on the procedure followed by the Case Western University Bearing Data Center [10] for collecting acceleration data for bearings state assessment. Each health state is related to a severity level of the studied fault, defined by the size of single point defects induced over test bearings using electro-discharge machining:

Table 1. Severity levels specifications

Health state	**Fault specifications**
Healthy (S0)	No defect
First severity level (S1)	0.007 inches point fault
Second severity level (S2)	0.014 inches point fault
Third severity level (S3)	0.021 inches point fault

We test the bearings under three different speeds; 600, 1200 and 1800 RPM. We perform this experimental procedure on the test bench shown in Figure 1. In order to cover the most steady state data as possible, reached after 35 minutes on average, we run each bearing state during 150 minutes. We modelled the experiment as a second order system with and average time constant (τ) of 35 minutes, therefore, we recorded IRT every 5 minutes in order to account 5 times the average τ.

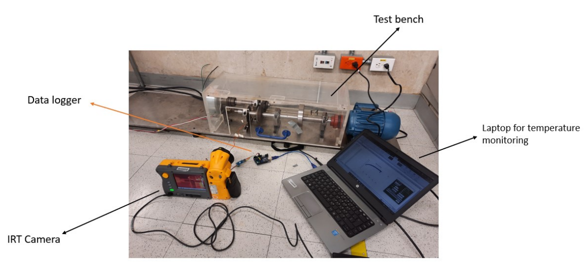

Figure 1. Test stand

For IRT data collection, we used Fluke Ti45 and Fluke Ti35 cameras, with emissivity of 0.9 for diagonal IRTs and 0.8 for lateral IRT at 30 centimeters form the target. To enhance generalization in the classification task, we collect IRT data using different palettes and at different angles ranging from 0 to 90 relative to the housing, and from below and above, as shown in figure 2.

We took images from distances between 15cm and 30cm from the target, according to the camera's data-sheet. Under these parameters we consolidated a data set of 2299 infrared thermographies each consisting of thermal image and thermal pixel matrix. The data set has four human annotated classes for classifying the health state of the bearings, corresponding to each severity of the failure: S0, S1, S2, and S3.

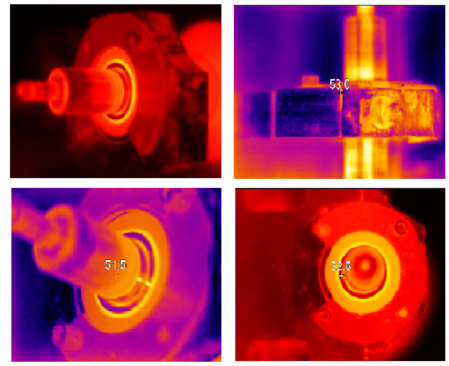

Figure 2. IRT angles shown at 90 degrees and 45 degrees from the target.

4. Aproach

4.1. Baseline

We selected EfficientNet scaling strategy as our baseline [14], a methodology that exhibited increased accuracy results on ImageNet classification with significant reduction in the number of parameters involved.

Tan et al. [14] study model scaling and propose a methodology to uniformly scale width (number of channels), depth (number of layers), and resolution (input image size) of convolutional neural networks in order to achieve better performance. Using neural achitecture, they desgin a new network and scale it up obtaining a family of networks: from EfficientNet-b0 to EfficientNet-b7, namend in increasing accuracy.

For dimension scaling, they propose three coefficients for each dimension; α, β and γ, and a compound coefficient ϕ that scales the network as follows:

$$\begin{aligned} depth: d &= \alpha^{\phi} \\ width: w &= \beta^{\phi} \\ resolution: r &= \gamma^{\phi} \\ s.t. \quad \alpha \cdot \beta^2 \cdot \gamma^2 &\approx= 2 \end{aligned}$$

The dimension parameters are determined by grid search, while ϕ is dependent on computational resources available for model scaling. The compound scaling method is applied in two steps:

1. Fix $\phi = 1$ and do a grid search on α, β and γ constrained to the above.

2. Fix α, β and γ and scale baseline network iterating over ϕ for optimal performance in accordance to available resources.

Authors start their experimentation with EfficientNet-b0 and changed the parameters until reaching the highest classification accuracy of 84.3% on ImageNet with 66 million parameters using EfficientNet-b7.

In addition, as part of EfficientNet backbone, Tan and Le apply two main methodologies that increased performance. The first one, is the use of an inverted residual structure from MobileNetV2 [13] as building block of the network. The second, in order to assess channel prioritization and consider their inter dependencies, they implement squeeze and excitation units as part of inverted residual blocks [1].

Therefore, the main purpose of using EfficientNet as our baseline is to assess the capacity of a high performance architecture on infrared data from rolling bearings. While related works have used older architectures such as VGG or LeNet, we propose EfficientNet with the potential of better results with fewer parameters. We start experimentation using EfficientNet-b0 over TBA and IBA with the following parameters [14].

- *ϕ:* 0
- *Resolution:* 224x224
- *Drop rate:* 0.2

4.1.1 Temperature Based Analysis (TBA)

Most-high level infrared thermography cameras output thermal pixel matrices with temperature information on each infrared pixel. By implementing temperature based analysis we seek to evaluate the utility of thermal patterns for the classification of the health state of rolling bearings.

First, we use TBA and a multi class SVM that could identify the severity of the failure on a rolling bearing, trained. We define a region of interest and first-order statistical indicators as mean, standard deviation, kurtosis, and energy were used as components of feature vectors that made the representation space of images. We apply principal component analysis to train the SVM, nevetheless, results were not promising, with an accuracy over the test set of 34% and f1-score of 51%.

Therefore, we held initial experimentation with EfficienNet-b0, training the network during 40 epoch of batch size 20, fixed learning rate of $1 * 10^4$ and Adam optimizer. Also, as Tan et al. [14], we use SiLu as activation function and apply stochastic depth. We use cross-entropy loss as a proper error function for classification tasks. An initial resizing was applied to thermal images in order to account for optimal 224x224 resolution of EfficientNet-b0. Results for our baseline on version b0 for TBA are shown in table 2.

4.1.2 Intensity Based Analysis

Thermal pixel matrix is not a universal output among thermal cameras, nevertheless, thermal images are and therefore, a pixel intensity analysis over the images containing pseudo color that represent surface temperature from IRT, is relevant. For that purpose, thermal images are resized to the 224 x 224 optimal resolution for EfficientNet and all RGB channels are fed into the network.

For IBA, we held initial experiments using EfficientNet-b0 by the same experimental parameters as for TBA: 40 epochs, batch size 20, learning rate of $1 * 10^4$ and Adam optimizer. Equally, we use SiLu and apply stochastic depth with cross entropy as the loss function. Table 2 show performance metrics

Table 2. Baseline results (14 039 680 trainable parameters for IBA and 14 039 104 for TBA).

Approach	ACA	Accuracy	F1-score	Loss_test
IBA	0.8908	0.8942	0.8898	0.3143
TBA	0.8531	0.8551	0.8447	0.4515

4.2. Proposed method

We consider an additional approach for IRT data, a hybrid model that combines both TBA and IBA approaches in a single pipeline in order to leverage the potential of both sources of information. Feeding a model with both types of data can improve feature learning and provide additional insights to enhance classification and pattern recognition among the data set. We propose two approaches to implement our hybrid model:

- *4th channel:* We first propose concatenating the thermal pixel matrix, with the temperature data, to the thermal image with pseudo colors. Then, this 4-channel matrix is fed to our EfficientNet baseline for classification.

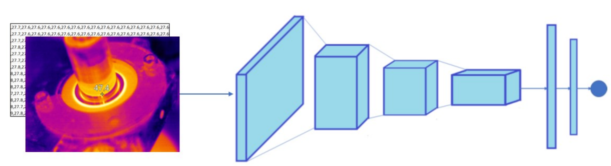

Figure 3. 4th channel hybrid model.

- *feature vector:* Another possible approach is to obtain a similar feature vector from the thermal pixel matrix used to train the SVM. Then, we feed thermal images to the convolutional backbone of the model and then, concatenate thermal feature vector to the flattened output of the convolutional layers as input to the linear layers for classification. The feature vector is made of first order statistical indicators: mean, median, standard deviation, minimum value and maximum value.

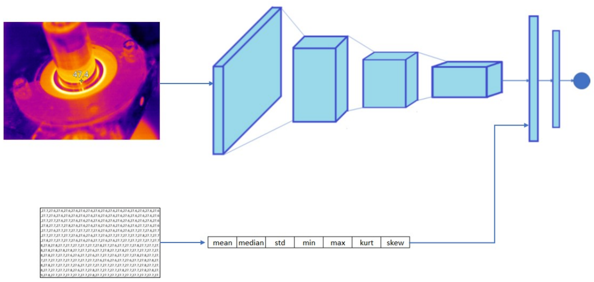

Figure 4. Feature vector hybrid model.

5. Experiments

5.1. Validation Experiments

We evaluate a alternative approach to the baseline on Table 2 in order to assess the information loss due to resizing. Therefore, we perform new experiments with no resolution change, namely 480×640 for thermal images and 240×321 for thermal matrices. Results for the baseline are shown in Table 3

Table 3. Baseline results without resizing

Approach	ACA	Accuracy	F1-score	Loss_test
IBA	0.8988	0.8942	0.8862	0.3370
TBA	0.8790	0.8783	0.8754	0.4784

Temperature data is more sensitive to information loss due resizing than thermal images, hence, in order to evaluate if parameter reduction by resizing compensates information loss, we test our hybrid models with and without resizing. We propose three ways to implement the 4th channel hybrid model and four ways to implement the feature vector hybrid model:

4th Channel hybrid model

- *4-HyRS*: Concatenating themal images and thermal pixel matrices after both being resized to optimal 224×224 b0 resolution.

- *4-HynoRS-0Pad*: Concatenating thermal images and thermal pixel matrices, minimizing resizing. Since both types of data have different sizes we propose zero-padding thermal matrices until 480×640size is reached. Then, we concatenate them to thermal images without modifying the initial temperature data acquired.

- *4-HynoRS-interpol*: Concatenating thermal images and thermal pixel matrices, by resizing thermal pixel matrices up to 480×640 by bicubic interpolation.

Feature vector hybrid model

- *fv-HynoRS*: Feeding original thermal images with no resizing to the convolutional backbone and feature vector obtained from the original thermal pixel matrices, with no modifications, fed into the linear layers for classification.

- *fv-HynoRS-0Pad*: Feeding original thermal images to the convolutional backbone and feature vector obtained from zero-padded thermal pixel matrix fed into the linear layers.

- *fv-HyRS*: Feeding resized thermal images to the convolutional backbone and feature vector obtained from resized thermal pixel matrix fed into the linear layers, both types of data with 224×224 resolution.

- *fv-HynoRS-interpol*: Hybrid model by feeding original thermal images to the convolutional backbone and feature vector obtained from the thermal pixel matrix resized to 480×640 resolution by bicubic interpolation, fed to the linear layers.

Tables 4 and 5 show the results for the above experimental framework. We verify that combining both data sources improve representational properties of the model since we achieve better performance compared to baseline results in Table 2. For the *4th channel* approach, the best result is for the 4-HynoRS-0Pad methodology which preserve the most amount of original information from data, by keeping the original resolution from thermal images and zero padding thermal pixel matrices in order to avoid modifying original recorded temperatures. On the other hand, the best results for the feature vector approach are obtained with the fv-HynoRS-interpol methodology with the highest metrics among hybrid models.

Table 4. Validation results for 4-channel hybrid approach (14039968 trainable parameters)

Approach	ACA	Acc	F1-score	Loss_test
4-HyRS	0.8964	0.9	0.8967	0.3840
4-HynoRS 0Pad	0.8984	0.9	0.8949	0.3808
4-HynoRS interpol	0.8973	0.9	0.8955	0.3883

5.2. Evaluation Experiments

Since we obtain the best performance over all proposed approaches with the *fv-HynoRS-interpol* methodology, as stated on Table 5, we performed additional experimentation over the test set with this methodology in the quest for better results. We started by implementing adaptive learning and increasing the number of epochs to 100, nevertheless,

Table 5. Validation results for feature vector hybrid approach (14039700 trainable parameters)

Approach	ACA	Acc	F1-score	Loss
fv-HynoRS	0.8587	0.8623	0.8519	0.5286
fv-HynoRS 0Pad	0.8587	0.8623	0.8519	0.5286
fv-HyRS	0.8587	0.8870	0.8807	0.4461
fv-HyRS noRS	0.8587	0.8623	0.8519	0.5286
fv-HynoRS interpol	**0.9079**	**0.9101**	**0.9062**	0.3407

we achieved performance results with ACA of 0.906, accuracy of 0.887 and f1-score of 0.883. Also, we evaluated a more efficient architecture, EfficientNet-b3, seeking for better performance with a small increase in the number of parameters, nevertheless we obtained similar results to version b0 with ACA of 0.9031, accuracy of 0.906 and f1-score of 0.903.

6. Discussion

Feature vector hybrid model has the best performance upon our baseline and the analyzed experimental framework. These increased metrics results for the feature vector hybrid model indicated that it is more suitable for identifying the health state of rolling bearings with reduced false positives and false negatives as shown in the confusion matrix in figure 5. We attribute these results to the following factors:

- Concatenating the feature vector to the classification layers' input, improves representational power of the network while limiting the increase in parameters due to additional information inputs, compared to the baseline.
- Channel-wise feature response recalibration performed by the squeeze and excitation blocks (SE) can be enhanced by channels with the same range as inputs. Since, thermal image pixel intensity range from 0 to 255 and thermal matrices range from minimum to maximum temperatures, e.g., from 21 to 45°C, concatenating these matrices might impact the performance of SE blocks, but future work is required in order to assess the impact of these range variations.

Although fv-hybrid model has the overall best results, non resized IBA models also outperforms our baseline (table 3). This can be beneficial when data acquisition is performed with less complex cameras incapable of obtaining thermal pixel matrices and bearing state assessment is required only from thermal images. Most cameras capable of recording thermal images and thermal pixel matrices

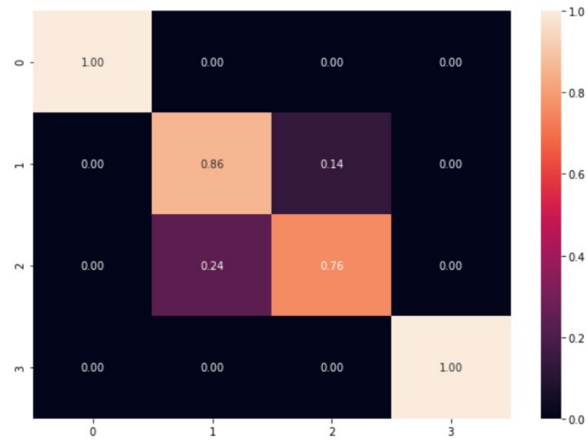

Figure 5. Confusion matrix for fv-HynoRS-interpol

are high-cost equipment manufactured by few companies. Hence, for users with low cost equipment, non-resized approach could be useful.

An analysis of the imaging patterns on which our hybrid model relies for classification might be important in order to identify the most informative patterns across the data set. Further work on the visualization of the saliency maps will be required.

Furthermore, our study had limitations. First, our data was acquired with two IRT cameras for bearings of the same manufacturer and reference. In consequence, in order to enhance the applicability and generalization of our study, models must be trained with data from more variate subjects and acquired by more than two cameras from different manufacturers. Second, our experimental framework is centered on EfficientNet-b0, so future work on the implementation and test of other EfficientNet networks is required in order to increase performance, in particular EfficientNet b5 or higher.

7. Ethical considerations

7.1. P-F curves

To analyze the ethical impact of the implementation of a thermal based system for the state assessment of rolling bearings, it is important to understand its possible failure consequences and how it relates to other methodologies.

Usually, the reliability of CBM or PdM methodologies is measured as its capacity to anticipate the failure of an asset. This is represented with a PF-curve which situates the methodology in the time span of an asset, that ends when it reaches a failure state. Companies are usually interested in systems that can anticipate failure and therefore, are located as back as possible from the failure state, inside the Potential Failure (PF) interval. As seen in figure 6, thermography is inside the PF interval, however, cannot anticipate a failure as back in time as vibration methodology, which is currently the most implemented approach in the industry for bearing state assessment. Therefore, in an eventual implementation of an IRT system, false negatives, i.e, the wrong classification of the health state indicating a bearing as healthy when it is in a critical state (S3 state), increases the probability of failure due to the location of this methodology in the PF-curve.

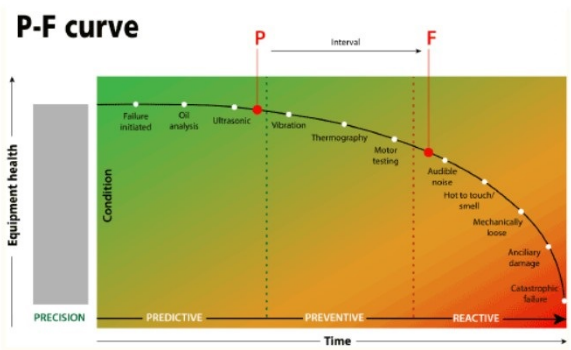

Figure 6. PF-curve [6]

If our proposed IRT methodology misclassifies the health state of a bearing in an asset and it fails, failure consequences can be classified into four categories [8]:

- **Operational consequences:** When the failure of an asset affects the production line and has an impact in the operations and processes of the whole company, involving time and money losses.
- **Safety consequences:** When the failure of an asset affects the safety of the workforce associated to the system. These consequences range from psychological impact to injuries and even death.
- **Environmental consequences:** When the failure of an asset has a direct impact on the environment due contaminating remnants released during the system breakdown.
- **Non-operational consequences:** These are consequences that are not directly related to the operation of the company.

Since CBM and PvM are usually implemented in crucial and critical assets of companies, most of the times their failure has serious operational consequences that involve massive economical losses and the possibility of having safety and environmental consequences, involving human and environmental losses.

7.2. Workforce

Automation in a maintenance context usually involves a lessening of the resources needed in the asset monitoring process. CBM and PvM strategies require fewer resources once implemented and therefore, may imply a workforce reduction. As maintenance systems increase automation, less human intervention is needed, hence, some positions inside companies are being eliminated which in many contexts is associated with unemployment of technical staff.

7.3. Bias and pre-trainning

Our proposed infrared state assessment methodology, and the CNN are biased by the training data acquired in a 6 rolling bearings of a unique SKF reference, under 9 operating states, which do not represent industry operating conditions of rolling bearing. Therefore, our methodology can not be directly implemented for industry asset management, nonetheless, our pretrained model on IRT data can increase performance of other models trained on industry bearing data.

8. Conclusion

Simultaneous intensity and thermal based analysis of IRT data for the state assessment of rolling bearings outperforms independent IBA and TBA approaches. Our proposed hybrid model feeds thermal images to the convolutional backbone of a CNN model (EfficientNet-b0) while concatenating to the linear classification layers a feature vector conformed by first order statistical indicators from thermal pixel matrices. We show that this approach is implemented with a slight increment in computational requirements due to low increase in the number of trainable parameters. This model has the potential to replace traditional vibration CBM and PdM systems currently used on the industry for the state assessment of rolling bearings.

References

[1] 4
[2] T. Mian A. Choudhary and S. Fatima. Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images. *Measurement*, 2021. 2
[3] Bo-Suk Yang Ali MD Younus. Intelligent fault diagnosis of rotating machinery using infrared thermal image. *Expert Systems with Applications*, 2012. 2
[4] A. Choudhary and S. Sudha Letha. Infrared thermography-based fault diagnosis of induction motor bearings using machine learning. *IEEE SENSORS JOURNAL*, 2021. 1, 2
[5] A. Beghi G. Susto and C. De Luca. A predictive maintenance system for epitaxy processes based on filtering and prediction techniques. *IEEE SENSORS JOURNAL*, 2021. 1
[6] Clark Kevin. The p-f curve: One of the first, yet hardest, things to learn. 2020. 7
[7] Jinjiang Wang Tangbo Bai Laibin Zhang Lixiang Duan, Mingchao Yao. Segmented infrared image analysis for rotating machinery fault diagnosis. *Infrared Physics Technology*, 2016. 2

[8] J. Moubray. *Reliability-centered Maintenance*. Industrial Press, 2001. 7

[9] EYK Ng N Golestani, M EtehadTavakol. Level set method for segmentation of infrared breast thermograms. *EXCLI Journal Experimental and Clinical Sciences*, 2014. 2

[10] Case Western University School of Engineering. Case western reserve university bearing data center website. 3

[11] Mia Loccufier Sofie Van Hoecke Olivier Janssens, Rik Van de Walle. Deep learning for infrared thermal image based machine health monitoring. *IEEE/ASME TRANSACTIONS ON MECHATRONICS*, 2018. 2

[12] Viktor Slavkovikj Kurt Stockman Mia Loccufier Rik Van de Walle Sofie Van Hoecke Olivier Janssens, Raiko Schulz. Thermal image based fault diagnosis for rotating machinery. *Infrared Physics Technology*, 2015. 1, 2

[13] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. 2018. 4

[14] Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. 2019. 3, 4

[15] Mrinal Kanti Bhowmika Anjan Kumar Ghosha Usha Rani Gogoia, Gautam Majumdarb. Evaluating the efficiency of infrared breast thermography for early breast cancer risk prediction in asymptomatic population. *Infrared Physics Technology*, 2019. 2

[16] Fangyi WAN Xianzhi WANG Huangchao YU Yongbo LI, Xiaoqiang DU. Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging. *Chinese Journal of Aeronautics*, 2019. 2

[17] Chi-Man Vong Michael Pecht Zhen Jia, Zhenbao Liu. A rotating machinery fault diagnosis method based on feature learning of thermal images. *IEE access*, 2018. 2