Vibration Analysis using Random Forest

Adwait Lokhande

Abstract**—** Vibration analysis for fault detection in rotating machinery is a crucial technique to identify early damage and prevent production disruptions. Machine learning methods, specifically Random Forest, were employed to streamline the analysis process and enhance diagnostic accuracy. We present a dataset designed for unbalance detection, featuring varying unbalance magnitudes applied to a rotating shaft using a 3D-printed holder. Vibrations were recorded by three sensors across a speed range of approximately 630 RPM to 2330 RPM at a sampling rate of 4096 values per second. Separate development and evaluation datasets are provided for each unbalanced strength. Utilizing this dataset, we tested Random Forest classification alongside other methods, including fully connected and convolutional neural networks, as well as Hidden Markov Models. The results demonstrated that the Random Forest classifier, using automatically extracted time series features, achieved the highest prediction accuracy of 98.6% on the evaluation dataset.

Keywords**—** Random Forest, Unbalance detection, Rotating machinery, Predictive maintenance, Vibration analysis

I. Introduction

The field of machine learning has seen remarkable progress in recent years, with notable achievements spanning a wide array of applications, including image recognition, natural language processing, and reinforcement learning. While these accomplishments have garnered significant attention in the media, it's essential to recognize that the potential of machine learning algorithms extends beyond these well-publicized domains. In particular, they offer a substantial opportunity for enhancing industrial processes and applications.

Within the realm of industrial applications, one area that has shown great promise is the analysis of vibrations in rotating machinery. This analysis serves a critical purpose, primarily revolving around two key objectives: the detection of unbalances and the identification of damage to roller bearings. In this context, we'll focus on the specific application of the Random Forest algorithm, a versatile and powerful machine learning technique.

The significance of detecting unbalances in rotating machinery cannot be understated. Unbalances in rotating shafts can lead to a decrease in the lifespan of crucial components, such as bearings and other machinery parts, resulting in elevated operational costs. Early detection of these unbalances is essential for minimizing maintenance expenses, preventing unwarranted production disruptions, and ultimately extending the overall service life of the machinery.

What makes algorithmic unbalance detection particularly appealing is that it involves minimal additional effort. Furthermore, it enables real-time analysis of streaming data, allowing for the swift identification and correction of unbalances, often before they cause any substantial damage to the machinery's drive train.

Despite the clear advantages of algorithmic unbalance detection, there exists a significant gap in the availability of publicly accessible condition monitoring (CM) datasets suitable for the rigorous testing and comparison of algorithms. While datasets for CM in various domains do exist, such as those pertaining to hydraulic systems and the detection of bearing damage, a conspicuous absence remains in the domain of unbalance detection, a known precursor to bearing damage. To address this gap, we have taken the initiative to introduce a dedicated dataset specifically designed for unbalance detection, centered around the analysis of vibration data. This dataset is made accessible through the Fraunhofer Fordatis database, thereby fostering research and innovation in this critical area.

In addition to providing the dataset, our study goes further by conducting comprehensive analyses. These analyses are geared towards determining which algorithms exhibit the highest accuracy in detecting unbalances and establishing the range of unbalanced strengths within which each algorithm performs reliably. Furthermore, to promote transparency and collaboration, we have open-sourced the Python code used in these analyses via a Github repository.

In summary, this study is firmly centered on the application of the Random Forest algorithm for the purpose of detecting unbalances in rotating machinery. Through this research, we aim to underscore the algorithm's potential to revolutionize maintenance practices, enhance the reliability of machinery, and contribute to the overall advancement of industrial processes.

II. Measurement Setup

The experimental setup for simulating defined unbalances and measuring resultant vibrations is orchestrated around an electronically commutated DC motor (WEG GmbH, type UE 511 T). This motor is impeccably regulated by a motor controller and securely affixed to an aluminum base plate using a galvanized steel bracket.

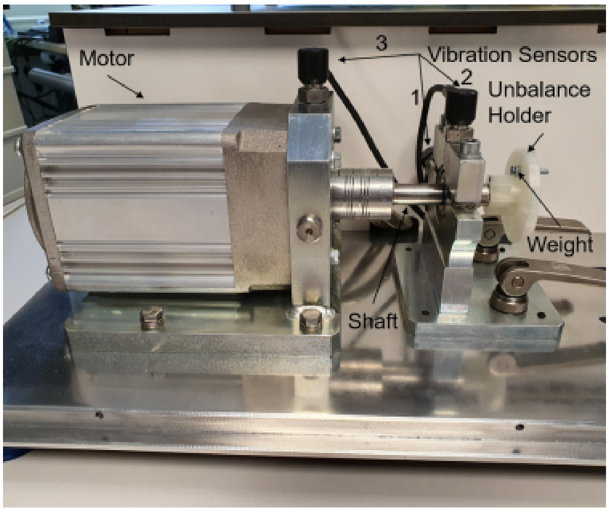
The motor controller grants precise control over the motor's rotation speed, offering a variable range spanning from approximately 300 to 2300 revolutions per minute (RPM). This speed can be seamlessly adjusted by manipulating the voltage supplied to the motor controller.

The motor itself drives a 12 mm diameter shaft, linked to another shaft of identical dimensions, measuring 75 mm in length. This coupling is facilitated by an Orbit Antriebstechnik GmbH component (type PCMR 29-12-12-A). The primary shaft passes through a roller bearing, firmly held in place by a roller bearing block constructed from galvanized steel. Directly behind the roller bearing, an unbalance holder is affixed. This holder, crafted using a 3D printer (Ultimaker 3, with nylon material), features a 52 mm diameter disc adorned with axially symmetric recesses. These recesses allow for the insertion of weights, effectively simulating unbalances.

To capture the vibrations generated during the experiments, vibration sensors (PCB Synotech GmbH, type PCB-M607A11 / M001AC) are strategically positioned on both the bearing block and the motor mounting. The data from these sensors is meticulously acquired using a 4-channel data acquisition system (PCB Synotech GmbH, type FRE-DT9837).

To keep a close eye on the motor's rotation speed, a frequency counter within the DT9837 is employed. This counter digitally records the periodicity of the rotor position signal emitted by the motor. For a visual representation of the setup, refer to Figure 1.

In essence, this experimental arrangement has been meticulously designed to simulate unbalances and capture the ensuing vibrations, providing a robust foundation for the research and analysis conducted in this study.



III. Results And Discussion

Results:

To assess the generalization ability of the employed algorithms and understand the potential impact of computational complexity on prediction accuracy, a classification experiment was carried out using a minimal set of features. This minimal feature set included the mean of the 'Measured RPM' values, as well as the standard deviation and kurtosis of the vibration values, calculated for each partitioned window of the datasets. Two variants of this feature calculation were performed. The first variant calculated the standard deviation and kurtosis solely for 'Vibration 1,' resulting in a total of 3 features, including the mean 'Measured RPM' values. The second variant employed data from all three vibration sensors, resulting in 7 features in total. Both of these variants are referred to as 'minimal features.' Subsequently, a Random Forest model was trained using these minimal features.

The classification training was executed twice: once with dataset pairs consisting of one unbalance strength and the unbalance-free case each, and another time with all existing unbalance strengths. The outcome of this experiment indicates that even with only 3 features, the highest unbalance can be detected nearly perfectly in both experiments. When using 7 features and training on the complete dataset, even the '3E' dataset can be classified close to 100% accuracy. However, with smaller unbalances, there is a notable decrease in prediction accuracy. In the second experiment, where all unbalanced strengths are considered, the unbalance-free case can only be detected to 82.2% accuracy (using 3 features) or 94.6% accuracy (using 7 features).

When evaluating results based on rotation speed, it is particularly evident that the high accuracy is maintained below 1200 RPM. In this range, the classification with the minimal feature set achieves significantly better results than in the previous approach. It's worth noting that the minimal features provide a robust basis for detecting substantial unbalances effectively.

To further enhance the feature set, the Python package tsfresh was leveraged to extract a broader range of time series features. A total of 748 features belonging to the class EfficientFCParameters() were extracted for 'Vibration 1' and employed as input for a Random Forest algorithm. The prediction results on the evaluation dataset reveal significant improvements in the detection rate for smaller unbalances when compared to the minimal features. The overall prediction accuracy achieved with this extended feature set was 93.2% when trained on all unbalanced strengths and a mean prediction accuracy of 79.9% when trained with dataset pairs. This performance is comparable to that of Convolutional Neural Networks in the previous approach.

Discussion:

The results and discussion of this study highlight the efficacy of the Random Forest algorithm as a central tool for unbalance detection in rotating machinery. When using minimal features, Random Forest exhibited excellent performance, particularly in detecting substantial unbalances, demonstrating its robustness in practical applications. The algorithm excelled in scenarios with rotation speeds below 1200 RPM, surpassing other classification approaches.

Furthermore, the utilization of an extended feature set, extracted using tsfresh, significantly improved the detection rate for smaller unbalances. This enhanced feature set provided a higher level of accuracy when classifying unbalanced strengths. It's noteworthy that this performance was comparable to that of Convolutional Neural Networks, underscoring the power of Random Forest in solving this complex problem.

In conclusion, the Random Forest algorithm emerges as a strong candidate for unbalance detection in rotating machinery. Its flexibility, robustness, and ability to accommodate both minimal and extensive feature sets make it a valuable tool for predictive maintenance and fault detection in industrial applications.

IV. CONCLUSION

The study demonstrates that the Random Forest algorithm serves as a highly effective and versatile tool for unbalance detection in rotating machinery, showcasing its pivotal role in predictive maintenance and industrial fault detection.

Through the evaluation of minimal features, it became evident that Random Forest excels in detecting substantial unbalances, achieving near-perfect accuracy even with a limited set of features. The algorithm's robustness is particularly notable when applied in scenarios with rotation speeds below 1200 RPM, outperforming alternative classification approaches.

Furthermore, by expanding the feature set using the tsfresh package, the study enhanced the algorithm's capacity to detect smaller unbalances. This augmented feature set significantly improved detection accuracy, aligning it with the performance levels of Convolutional Neural Networks used in previous approaches.

In conclusion, Random Forest demonstrates its adaptability, robustness, and reliability in unbalanced detection across varying scenarios. Its capacity to work effectively with both minimal and extensive feature sets makes it a valuable asset for industrial applications, fostering the enhancement of predictive maintenance practices and the advancement of machinery reliability. This research reaffirms the pivotal role of Random Forest in the context of industrial machinery health monitoring and maintenance.

REFERENCES

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE, vol. 86, no.11, pp. 22782324, 1998.

[2] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, Nature, vol. 521, no. 7553, pp. 436444, 2015.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ImageNet classification with deep convolutional neural networks, Commun. ACM, vol. 60, no.6, pp. 8490, 2017.

[4] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, in AAAI Conference on Artificial Intelligence, 2017.

[5] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, Enriching Word Vectors with Subword Information, Transactions of the Association for Computational Linguistics, vol. 5, pp. 135146, 2017.

[6] A. Conneau, D. Kiela, H. Schwenk, L. Barrault, and A. Bordes, Supervised Learning of Universal Sentence Representations from Natural Language Inference Data, May. 2017. Accessed: Mar. 4 2020.

[7] T. Young, D. Hazarika, S. Poria, and E. Cambria, Recent Trends in Deep Learning Based Natural Language Processing [Review Article], IEEE Comput. Intell. Mag., vol. 13, no. 3, pp. 5575, 2018.

[8] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, XLNet: Generalized Autoregressive Pretraining for Language Understanding, Advances in neural information processing systems, 2019.