CIRP Annals - Manufacturing Technology 74 (2025) 49–53

Contents lists available at ScienceDirect

CIRP Annals - Manufacturing Technology

journal homepage: https://www.editorialmanager.com/CIRP/default.aspx

Random wavelet kernels for interpretable fault diagnosis in industrial systems

Haoxuan Deng, Samir Khan, John Ahmet Erkoyuncu (2)*

Cranfield University, Faculty of Engineering and Applied Sciences, United Kingdom

ARTICLE INFO

Article history:
Available online 13 May 2025

Keywords:
Failure
Machine learning
Maintenance

ABSTRACT

Deep learning is a powerful method for fault diagnosis, but its "black-box" nature raises concerns in critical applications. This paper presents an interpretable, lightweight method combining random convolution kernel transformation (ROCKET) with wavelet kernels, which offer systematic time-frequency analysis and intuitive insights. Principal component analysis (PCA) is used to extract relevant patterns, forming a health indicator that guides maintenance decisions. A case study on linear actuator fault diagnosis demonstrates the method's balance of interpretability and computational efficiency, making it a valuable tool for reliable asset health monitoring in resource-limited settings.
© 2025 The Authors. Published by Elsevier Ltd on behalf of CIRP. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

1. Introduction

Fault diagnosis is essential for ensuring the reliability and safety of industrial systems [1]. While deep learning models are highly used due to their ability to extract complex patterns from large datasets, their lack of transparency creates challenges. The opaque decision-making process of these "black-box" models limits trust and usability in many critical applications [2]. This lack of interpretability raises concerns regarding the reliability, safety, and trustworthiness of these methods [3]. To address these challenges, interpretable machine learning methods are gaining traction. It offers diagnostic clarity by uncovering hidden patterns in degradation processes. Unlike deep models, they support actionable insights for maintenance while reducing reliance on opaque architectures. This capability has significant implications for maintenance strategies, allowing more effective interventions and extending the lifespan of critical industrial systems. Replacing deep models with transparent alternatives preserves performance while improving interpretability.

1.1. Limitations of current methods

Recent research has made significant strides in enhancing the interpretability of deep learning models, particularly in the domain of fault diagnosis. Techniques such as saliency maps, Layer-wise Relevance Propagation (LRP), and SHapley Additive exPlanations (SHAP) values have been proposed to provide explanations for model predictions by highlighting important features or regions of input data that influence the output [3,4]. These approaches provide valuable explanations but rely on inherently complex architectures, making them computationally expensive to train and deploy. For instance, interpretability tools like saliency maps and SHAP explain predictions but remain tied to complex deep models, limiting usability in real-time or resource-limited scenarios [3–5].

1.2. Proposed method: wavelet kernels for fault diagnosis

To overcome these challenges, this paper introduces a novel fault diagnosis framework based on wavelet kernels. This approach captures data patterns by combining temporal and frequency features, making it well-suited for analysing the signals commonly encountered in industrial systems [6,7]. Wavelet kernels provide a more structured and interpretable analysis compared to traditional methods, offering intuitive insights for operators.

To enhance interpretability, principal component analysis (PCA) is applied to identify the most relevant wavelet kernel features. These features are used to derive a health indicator that guides maintenance decisions, providing a clear and actionable metric for assessing asset conditions. The proposed method is validated through a case study on linear actuator fault diagnosis. Results demonstrate its effectiveness in monitoring asset conditions and planning maintenance activities, showcasing its potential for real-world applications.

Unlike deep learning models, which often require extensive computational resources and large datasets, the proposed method is lightweight and efficient. It avoids the need for complex training processes, making it ideal for resource-constrained industrial environments. Additionally, the model is designed in a shallow architecture with much fewer parameters, leading to less computation budgets and real-time diagnostic capabilities align with the demands of modern industrial applications, where quick and reliable fault detection is critical. Overall, this work advances the ongoing effort to develop

* Corresponding Author.
E-mail address: j.a.erkoyuncu@cranfield.ac.uk (J.A. Erkoyuncu).

https://doi.org/10.1016/j.cirp.2025.04.083
0007-8506/© 2025 The Authors. Published by Elsevier Ltd on behalf of CIRP. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

more interpretable and accessible machine learning models for practical industrial uses.

2. Random convolutional kernel transformation with wavelet kernel

In industrial scenarios, time series data is one of the most prevalent forms of information collected, particularly from sensors monitoring system operations. Fault diagnosis can often be framed as a time series classification problem, where the goal is to assign system states to predefined labels based on historical data. This process enables engineers to determine the system's condition and identify faults effectively.

2.1. Random convolution kernel transformation (ROCKET)

The Random Convolutional Kernel Transformation (ROCKET) [8] is a computationally efficient approach to time series classification. It employs a large set of convolution kernels with random weights for feature extraction, bypassing the need for parameter optimization. The extracted features are then fed into simple classifiers, such as logistic regression or ridge regression, which achieve a performance comparable to deep learning models like LSTM (Long-Short Term Memory), LSTM-FCN (LSTM Fully Convolutional Networks) and InceptionTime, ROCKET's efficiency stems from eliminating the extensive training required by deep learning models, making it particularly well-suited for time-sensitive or resource-constrained applications. The entire process can be divided into the following five parts:

1) Random kernel generation: ROCKET uses N convolution kernels with random weight denoted as $\{\omega_i | i = 1, \ldots, N\}$. These kernels are applied to the time series $\{x_i | i = 1, \ldots, T\}$, where T is the series length.

2) Convolution operation: Each kernel operates with a random dilation d and bias b, generating a feature map for each kernel. The convolution operation is mathematically expressed as:

$$Z_i(t) = \sum_{\tau=0}^{l_{ke}-1} x_{t+d\cdot\tau} \cdot \omega_i(\tau) + b \tag{1}$$

Where l_{ke} is the length of the convolution kernel and $t \in \{0, 1, \ldots, T - d(l_{ke} - 1)\}$.

3) Feature map creation: The convolution outputs form a feature map:

$$Z_i = [Z_i(0), Z_i(1), \ldots, Z(T - d(l_{ke} - 1))] \tag{2}$$

4) Down-sampling operations: Two down-sampling operations are performed on the feature map to reduce dimensionality::

- Maximum values (MAX): Extracts the maximum value from the feature map, akin to max-pooling in conventional convolutional neural networks.
- Portion of positive values (PPVs): Calculates the ratio of positive values in the feature map:

$$PPVs(Z_i) := \frac{1}{T - d(l_{ke} - 1)} \sum_{t=0}^{T-d(l_{ke}-1)} [Z_i(t) \geq 0] \tag{3}$$

5) Feature vector construction: After scanning the time series with N kernels, the results of the MAX and PPV operations are concatenated into a 2N-dimensional feature vector, which serves as input for classification. ROCKET's core idea is to measure the similarity between each kernel and time series segments via convolution, interpreted as a dot product. The resulting feature map captures these similarity scores, with the MAX highlighting the most aligned segment and the PPV reflecting the proportion of positive responses.

Building on this framework, we replace ROCKET's random convolutional kernels with wavelet kernels. These provide structured time-frequency analysis, capturing both temporal and spectral characteristics crucial for fault interpretation. This substitution enhances model interpretability while maintaining diagnostic performance, offering deeper insight into signal patterns for more effective condition monitoring and maintenance planning.

2.2. Replace random convolution kernel with wavelet kernel

The convolution operation can be viewed as a similarity measure between a time series and a kernel. Based on this interpretation, replacing ROCKET's random kernels with structured ones exhibiting regular patterns does not significantly compromise performance. Using kernels with well-defined, interpretable properties—such as symmetry or periodicity—enhances the model's transparency by revealing how specific time series segments align with meaningful signal features. Rather than manually designing each kernel, diverse patterns can be efficiently generated through parameterized functions. In this context, wavelet kernels—particularly Morlet wavelets—are well-suited due to their frequency selectivity and ability to capture multi-scale temporal structures.

According to above three reasons, in our approach, we substitute ROCKET's random kernels with MMorlet wavelet kernels of random scales $\{s_i | i = 1, \ldots, M\}$, as shown in Fig. 1. The wavelet transformation for a segment of a time series, $f : \{1, \ldots, T\} \mapsto \{x_i\}$. is expressed as:

$$C(s_i, \tau) = \int_{-\infty}^{\infty} f(t)\psi^*\left(\frac{t-\tau}{s_i}\right)dt \tag{4}$$

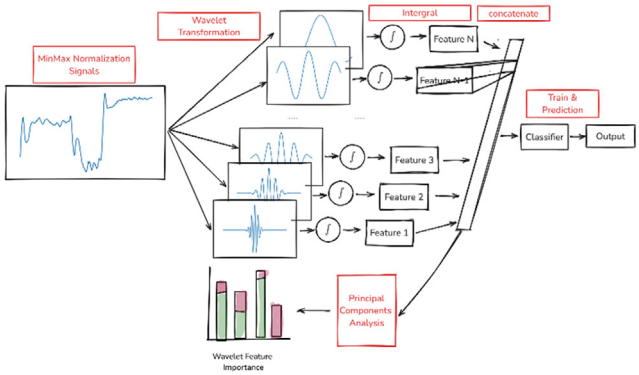

Fig. 1. The diagram of the proposed method.

Where τ is the translation parameter that shifts the wavelet in time, andψ^*(t)is the conjugate of the wavelet function. For our implementation, the Morlet wavelet is defined as:

$$\psi(t) = \exp^{-\frac{t^2}{2}} \cos(5t) \tag{5}$$

The resulting feature map from the wavelet transformation is:

$$C(s_i) = [C(s_i, \tau_0), C(s_i, \tau_1), \ldots, C(s_i, \tau_T)] \tag{6}$$

The feature map obtained from Eq. (6) exhibits a structure analogous to that of standard convolution operations, where each term reflects the similarity between a time series segment and a wavelet kernel at a specific scale. Unlike MAX or PPV down-sampling methods, we apply temporal integration to the feature map, aggregating similarity values over time into a single scalar. This approach captures the overall alignment between the wavelet kernel and the time series across the entire duration. This process is defined as:

$$IN(s_i) = \int C(s_i)d\tau \tag{7}$$

After applying wavelet transformations with Morlet kernels, the accumulated features form a feature vector, which is then used to train a classifier.

2.3. Feature importance analysis with PCA

Training a classifier is not the only part of the process. To gain deeper insights on interpretability and understand the classifier's reasoning, feature importance analysis is essential. By evaluating the contribution of each component in the feature vector to the

classifier's performance, we can gain insights into the underlying behaviour of the model.

Principal Component Analysis (PCA) is a powerful tool for this analysis. In this context, the feature vector consists of values derived from Eqs. (4)-(7), where the wavelet scale s_i is the only variable. Different scales produce distinct features, and PCA enables us to identify the most impactful scales contributing to classification. Key insights from PCA can be summarized as:

- PCA highlights the scales of the Morlet wavelet that are most relevant for classification.
- Visualizing these critical scales provides insights into the time-frequency patterns that influence fault diagnosis, improving transparency in model decision-making.

This approach enhances the interpretability of ROCKET, allowing us to align the extracted features with intuitive patterns in the time series and providing actionable insights for maintenance and fault diagnosis.

3. Case study: linear actuator fault diagnostics

3.1. Dataset description

The proposed method is validated using a publicly available time series dataset collected from a linear actuator system, as shown in Fig. 2 [9]. The implementation of our method is openly available and maintained as part of the RAVEN project on GitHub (https://github.com/BlueAsuka/RAVEN). This study directly reuses the dataset without re-running the experiments; detailed descriptions of the hardware and setup can be found in the original publication. The actuator was tested under a trapezoidal motion profile at constant speed across two load conditions (20 kg and 40 kg). Each scenario was repeated five times and replicated ten times, yielding 50 samples per load. Three types of mechanical faults were introduced at varying severity levels:

- Spalling (8 states), representing progressive wear.
- Lack of lubrication (2 states), simulating reduced system efficiency.
- Backlash (2 states), mimicking mechanical misalignment.

The above results in 13 fault categories and 650 samples per load condition. For evaluation, 30 % of the data was randomly assigned to the test set, while the remaining 70 % was used to train the diagnostic pipeline described in Section 2.

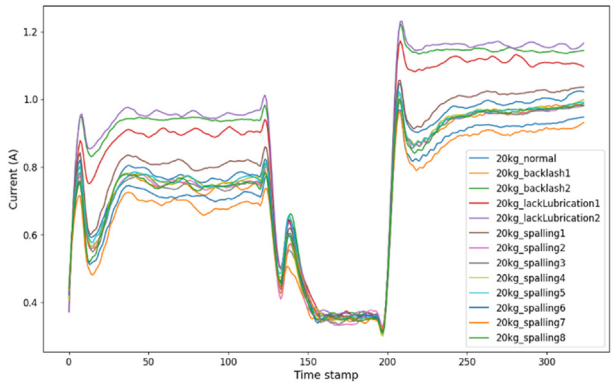

Fig. 2. All types of faults in the case study and corresponding time series signals under the load 20 kg.

3.2. The classification performance of the proposed method

To construct the pipeline, Python PyWavelet library was used to perform wavelet transformation on the time series data. We apply 300 randomly selected scales (a hyperparameter needed tuning to get a good performance), with the scales distributed uniformly between 0.1 and 500. The scale, defined as:

$$\frac{100}{central\ frequency\ of\ the\ wavelet} \tag{8}$$

allowed the wavelet trasformation to capture a wide frequency range from 0.2 Hz to 1000 Hz. This range ensures that the ability to extract both low-frequency and high-frequency patterns from the data, identifying subtle and complex features.

The wavelet-transformed data was used to extract features, which are subsequently input into two different classifiers for performance evaluation including Ridge Classifier (RC) and Linear Discriminant Analysis (LDA). Also, to benchmark the method, we compared it against five established algorithm's including LSTM, LSTM-FCN, InceptionTime, TST (Time Series Transformer), and ROCKET. There are two metrics for this comparison, i.e., the top-1 accuracy in the classification and the training time of the model.

Table 1
Performance comparison: accuracy vs. training time

Method	20 kg (%)	40 kg (%)	Training Time (s)
LSTM	37.95	53.58	10.57 (GPU)
LSTM-FCN	86.67	84.46	11.91 (GPU)
InceptionTime	87.08	84.56	29.15 (GPU)
TST	74.36	80.0	157.05 (GPU)
ROCKET	73.07	73.64	12.55 (CPU)
Proposed (RC)	72.31	70.98	10.41 (CPU)
Proposed (LDA)	85.13	83.42	10.57 (CPU)

GPU: NVIDIA RTX 4070, CPU Inter(R) Core i7−12700H.

The results are summarized in Table 1:

- InceptionTime delivers the top score among all models up to 87.08 % and 84.56 % under load 20 kg and 40 kg, respectively.
- LSTM-FCN follows closely with accuracies of 86.67 %/84.56 %.
- The proposed method, with the LDA delivered comparable results, achieving 85.13 %/83.42 % outperforming baseline methods like the original ROCKET.

In terms of computational efficiency, the proposed method was notably faster, completing training in just 10.57 s on a CPU, outperforming deep learning models such as InceptionTime and the LSTM-FCN, which require GPU acceleration.

3.3. PCA for the feature importance analysis

To better understand the classifier, the PCA was applied to analyse feature importance. This analysis highlights the contribution of different features, allowing insights into the classifier behaviour and the role of wavelet-transformed features. Fig. 3 shows the average time series for each fault type, along with the wavelet transformation coefficients at scales 5, 50, and 500. These scales represent different frequency bands:

- Scale 5: High-frequency components, capturing rapid changes in the signal.
- Scale 50: Mid-range frequencies, reflecting moderate oscillations.
- Scale 500: Low-frequency components, identifying slow trends in the data.

PCA revealed that specific wavelet scales contributed disproportionately to classification performance. Low-frequency scales were particularly effective in distinguishing different states. In this case, the wavelet with scale 495 (central frequency at 0.2 Hz) contributes the most to the final performance. This analysis underscores the versatility of wavelet transformations in capturing critical patterns relevant to fault diagnosis. By leveraging PCA, we identified the most important features for classification, enabling both improved interpretability and performance optimization.

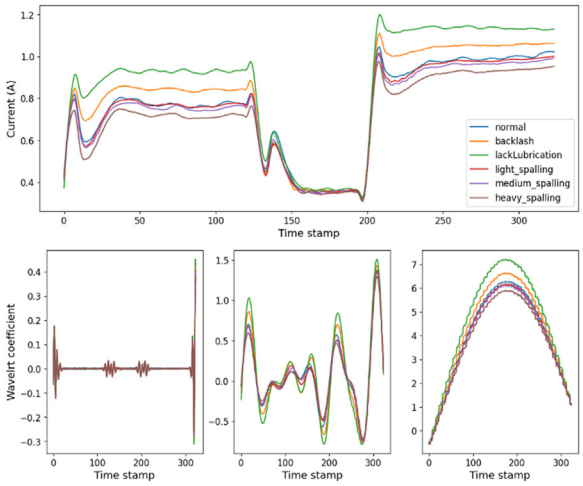

Fig. 3. The time series obtained by averaging each fault. The coefficient of the wavelet transformation with scale 5, 50, and 500, respectively.

4. Discussion

From Table 1, the proposed method demonstrates its capability in terms of lightweight architecture and training, while achieving comparable classification accuracy to state of the art algorithms. Beyond algorithmic performance, this approach offers insights into proactive maintenance in industrial systems.

The feature importance analysis highlights that wavelets with higher scales for transformation can reduce the overlap in the coefficient curves across different fault types. By applying Eq. (7) with the scale at 495, an indicator value can uniquely represent the original time series. This condensed value serves as a critical and efficient health indicator for the system, enabling fast and effective health monitoring.

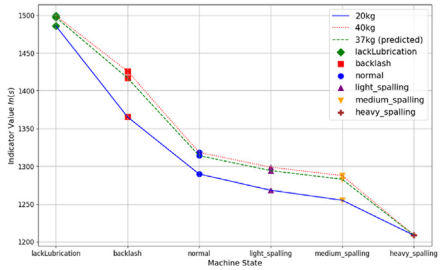

Fig. 4. The health indicator for each state derived by the Eq. (7).

As shown in Fig. 4, the health indicators for different states under various loads, denoted as $In(s)$, exhibit distinct patterns:

- Spalling events are associated with lower health indicator values.
- A decreasing health indicator value may signal the onset of spalling.
- Unlike black-box models, our method produces actionable health indicators, enabling proactive interventions

Additionally, the health indicator can be used to infer system behaviour in previously unseen scenarios. For instance, based on the health indicators in Fig. 4, as for a new load, say 37 kg, the system's state under this condition can be estimated by linear interpolating the values based on the 20 kg and the 40 kg curves. The estimated result is presented as the green dash line in Fig. 4. Although the data is not collected under this specific working condition, the algorithm can still perform inference and estimate the health state of the system by leveraging previously acquired knowledge. This capability enables generalization to situations not covered by historical data, further enhancing the method's applicability. It is important to note that linear interpolation serves as an initial and reasonable hypothesis for generalization; however, more advanced methods for extending to out-of-distribution cases should be explored in future research.

5. Limitations and future work

Ambiguities in health indicators: One limitation is that the health indicators for backlash and lack of lubrication are both higher than the normal value, potentially causing confusion for operators, as an increase in the indicator could signal either issue. To resolve this, we propose adding a second sensor to provide another measurement, enriching the health indicator and distinguishing the trends of different states.

Model simplicity: Another limitation is the simplicity of the health state indication model. While more complex models could better capture diverse features, patterns and regulations of the system, the key point lies not in increasing complexity, but in enabling a model update mechanism that can adapt to human requirements and newly collected data. This self-bootstrapping process would allow the model to reflect the machine's true state more accurately under varying operating conditions, ensuring continuous adaptation and improvement over time.

Limited use case: the proposed method has only been tested on a specific use case. To fully demonstrate its general applicability, the method must be extended to other industrial applications. This would allow for a broader evaluation of its effectiveness in different contexts and tasks, establishing its versatility in real-world industrial settings.

6. Conclusion

In this paper, we propose a novel method for time series fault diagnosis, inspired by the ROCKET algorithm with a key modification: replacing the random kernels with wavelet kernels with integral down-sampling operation.

The method was tested on a linear actuator dataset showing:

- Fast training with relatively low computational resources,
- Comparable top-1 accuracy to other deep learning methods.

The case study highlighted how the health indicator derived from wavelet kernels can accurately capture degradation trends, enabling timely interventions and improved maintenance outcomes. Feature importance analysis using wavelet-based health indicators revealed their potential for real-time system health monitoring. While some limitations remain, addressing them will pave the way for broader industrial applicability.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Haoxuan Deng: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation. **Samir Khan:** Writing – review & editing, Validation, Supervision, Software, Investigation. **John Ahmet Erkoyuncu:** Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition, Formal analysis.

Acknowledgment

The research was partially supported with EPSRC funding (EP/P027121/1). Data supporting this study are openly available from https://doi.org/10.17862/cranfield.rd.5097649.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cirp.2025.04.083.

References

[1] Gao RX, Krüger J, Merklein M, Möhring HC, Váncza J (2024) Artificial Intelligence in manufacturing: state of the art, perspectives, and future directions. *Annals of the* CIRP 73(2):723–749. https://doi.org/10.1016/J.CIRP.2024.04.101.

[2] Ribeiro MT, Singh S, Guestrin C (2016) Why should i trust you?' Explaining the predictions of any classifier. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778/SUPPL_FILE/KDD2016_RIBEIRO_ANY_CLASSIFIER_01-ACM.MP4..

[3] Bach S, Binder A, Montavon G, Klauschen F, Müller KR, Samek W (2015) On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS One* 10(7). https://doi.org/10.1371/JOURNAL.PONE.0130140.

[4] Ancona M, Ceolini E, Öztireli C, Gross M (Nov. 2017) Towards better understanding of gradient-based attribution methods for Deep Neural Networks. *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, . Accessed: Jan. 13, 2025. [Online]. Available: https://arxiv.org/abs/1711.06104v4.

[5] Bukkapatnam STS, Afrin K, Dave D, Kumara SRT (2019) Machine learning and AI for long-term fault prognosis in complex manufacturing systems. *Annals of the CIRP* 68 (1):459–462. https://doi.org/10.1016/J.CIRP.2019.04.104.

[6] Calaon M, Chen T, Tosello G (2024) Integration of multimodal data and explainable artificial intelligence for root cause analysis in manufacturing processes. *Annals of the CIRP* 73(1):365–368. https://doi.org/10.1016/J.CIRP.2024.04.014.

[7] Mallat S (2008) A wavelet tour of signal processing: the sparse way. *A Wavelet Tour of Signal Processing: The Sparse Way* : 1–805. https://doi.org/10.1016/B978-0-12-374370-1.X0001-8.

[8] Dempster A, Petitjean F, Webb GI (2019) ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels. Data *Min* Knowl Discov 34(5):1454–1495. https://doi.org/10.1007/s10618-020-00701-z.

[9] Ruiz-Carcel C, Starr A (2018) Data-based detection and diagnosis of faults in linear actuators. *IEEE Trans Instrum Meas* 67(9):2035–2047. https://doi.org/10.1109/TIM.2018.2814067.

https://dspace.org/

Random wavelet kernels for interpretable fault diagnosis in industrial systems

Deng, Haoxuan; Khan, Samir; Erkoyuncu, John Ahmet

2025
Elsevier
https://doi.org/10.1016/j.cirp.2025.04.083
Deng H, Khan S, Erkoyuncu JA. (2025) Random wavelet kernels for interpretable fault diagnosis in industrial systems. CIRP Annals, Volume 74, Issue 1, 2025, pp. 49-53

Downloaded from a DSpace repository. https://dspace.org/