



# WPT-Base Selection for Bearing Fault Feature Extraction: A Node-Specific Approach Study

Andrei Maliuk<sup>1</sup>  and Jong-Myon Kim<sup>1,2</sup>  

<sup>1</sup> Department of Electrical, Electronics and Computer Engineering, University of Ulsan, Ulsan 44610, Republic of Korea

[jmkim07@ulsan.ac.kr](mailto:jmkim07@ulsan.ac.kr)

<sup>2</sup> PD Technologies Co. Ltd., Ulsan 44610, Republic of Korea

**Abstract.** Wavelet packet transform (WPT) has found extensive use in bearing fault diagnosis for its ability to provide more accurate frequency and time-frequency representations of non-stationary signals. Traditional quantitative methods prioritize unequal node-energy distribution at the desired decomposition level as a criterion for WPT base selection. Decomposition results obtained with WPT-base selected using this approach can be characterized as having one WPT-node with high signal energy which is automatically considered as a component of interest containing information about bearing fault. However, prioritizing one WPT-node at this early stage of fault diagnosis process might not be optimal for all nodes in the WPT-tree decomposition level and might exclude components in other nodes, which may contain features potentially important for fault diagnosis. In this paper, we propose a node-specific approach for WPT-base selection to improve the quality of feature extraction. The new criterion evaluates WPT-bases upon their ability to generate a signal with the highest ratio of energy and entropy of the signal spectrum for a specific node. Using this criterion, the final WPT signal decomposition is constructed using the WPT-nodes produced by the bases with the highest criterion score. This approach ensures the preservation of all meaningful components in each node and their distinction from the noisy background, resulting in a higher quality feature extraction. To evaluate the effectiveness of the proposed method for bearing fault diagnosis, a comparative analysis was conducted using two sets of Paderborn University bearing fault experimental vibration data and the bearing vibration data from the Case Western University benchmark dataset. As a result, the proposed method showed better average performance across three datasets.

**Keywords:** Bearing Fault Diagnosis · Feature Extraction · Wavelet Packet Transform · Wavelet Packet Base Selection

## 1 Introduction

Bearings are crucial mechanical components found in various applications, and their faults can cause significant damage, loss of production, and even human casualties [1, 2]. With the availability of high-quality vibration sensors and Machine Learning

(ML) algorithms, data-based fault diagnosis methods are becoming more prevalent. These methods usually involve processing of the signal, extraction of the signal features, feature selection, and ML classification. Deep Learning (DL) can also be used for fault diagnosis, but the issue of explainability remains a challenge. Despite the advantages of DL, traditional ML methods remain a powerful alternative, particularly when data is limited [3].

In machine learning-based fault diagnosis algorithms, the first step is signal processing. Traditionally, Fast Fourier Transform (FFT) has been used, but it has limitations such as poor resolution, inability to capture transient signals, and spectral leakage [4–6]. Short-time Fourier Transform (STFT) solves these issues by sliding a window along the signal and performing FFT on each window to obtain a time-frequency representation. However, STFT has limitations regarding window length selection, which results in a tradeoff between frequency resolution and time resolution [7].

EMD is a time-frequency method that is used for decomposition of the signals into intrinsic mode functions (IMFs) without using a base function. It is particularly useful for non-stationary signals. However, the mode mixing phenomenon can occur when IMFs generated by EMD overlap, making it difficult to interpret and analyze them individually. To overcome this, techniques like EEMD or CEMD have been developed, but they require significant computing time and balancing the number of attempts and decomposition quality [8, 9].

The Wavelet Packet Transform (WPT) offers high time-frequency resolution and sensitivity to transient components by decomposing a signal into various sub-bands with different frequencies using wavelets as decomposition bases. WPT is less adaptive and flexible than EMD, but it is computationally less expensive. The selection of either method depends on the signal and application. Although there are many wavelet base functions available, selecting one remains a vulnerable part of WPT that lacks a general state-of-the-art solution method, creating a need for new solutions to be found.

This paper proposes a method for a novel feature extraction approach by trying to resolve a fundamental drawback of the WPT in comparison with EMD—dependency on the wavelet base. To overcome this limitation and increase the quality of feature extraction using WPT, a new criterion for base wavelet selection is proposed along with the novel node-specific approach for constructing of the representation of the bearing vibration signal. The new criterion evaluates WPT bases based on their ability to generate a signal with the highest ratio of energy and entropy of the signal spectrum for a specific node. The final WPT signal decomposition is constructed using the WPT nodes produced by the bases with the highest criterion score. This approach aims to preserve all meaningful components in each node and distinguish them from the noisy background, resulting in higher-quality feature extraction.

Further this paper is organized as follows: Sect. 2 overviews the datasets used for the validation of proposed method, Sect. 3 provides some technical background on Wavelet Packet Transform, Sect. 4 describes the proposed criterion for WPT base selection and construction of the signal representation, Sect. 5 discusses the fault diagnosis framework used for performance evaluation and Sect. 6 concludes the manuscript.

## 2 Technical Background

### 2.1 Wavelet Packet Decomposition

The Wavelet Packet Decomposition or Transform (WPT) decomposes an input signal into a binary tree structure of wavelet packet nodes, which are indexed as  $(j, n)$ , with corresponding coefficients  $d_j^n$ . This allows for analysis of both low and high-frequency spectra, making it useful for characterizing non-stationary bearing fault signals. At the root of the WPT tree, the input signal is located in the node  $W(0, 0)$ , with  $W(1, 0)$  and  $W(1, 1)$  representing the low-pass and high-pass filtered branches, respectively, resulting in approximation and detail coefficients  $d_1^0$  and  $d_1^1$ . Further decomposition is performed in the same way at every level  $j$ . An schematic example of a WPT with  $j$  decomposition levels is shown in Fig. 1.

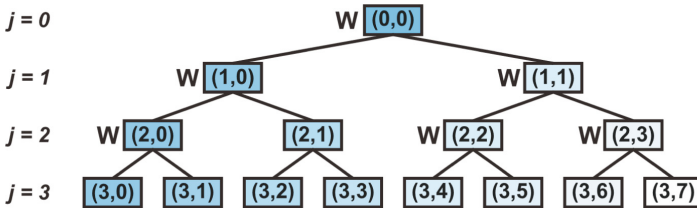


Fig. 1. Wavelet Packet Tree schematic.

The choice of the wavelet base in WPT decomposition is crucial. There are two main families of selection methods: qualitative and quantitative. Qualitative methods focus on properties like symmetry, compact support, orthogonality, regularity and vanishing moment to find the best fitting wavelet. However, relying solely on wavelet properties can be limiting due to the possibility of different wavelets sharing the same properties, making it hard to determine the most suitable one. To address this, researchers have explored shape matching, an alternative qualitative approach that analyzes the geometric shape of wavelets. It aims to find a wavelet base that resembles the shape of the target signal feature, improving signal component extraction. However, despite its benefits, manual shape matching is time-consuming and lacks automation.

Considerable research has been carried out to address the limitations of qualitative methods by exploring quantitative approaches. These approaches utilize various quantitative measures like signal energy, Shannon entropy, cross-correlation, Emlen’s modified entropy measure and distribution error criterion to determine the most appropriate wavelet base. In recent times, the criterion based on the ratio of the maximum energy to the Shannon entropy gained significant popularity as one of the leading quantitative methods for selecting the wavelet base. This criterion combines the widely used maximum energy metric and Shannon entropy metric, offering a reliable approach to wavelet base selection. The maximum energy method suggests that the most suitable wavelet base will enable the extraction of the highest possible energy from the analyzed discrete-time signal. The energy  $E_x$  of the signal  $x$  can be mathematically represented

as follows:

$$E_x = \sum_{n=1}^N |x_n|^2 \quad (1)$$

It's crucial to acknowledge that signals possessing identical energy levels can exhibit varying frequency distributions. For instance, one signal may have higher energy levels in frequency components important for fault diagnosis, while another signal may have a broad spectrum with uniform energy levels throughout the spectrum, which is not fruitful for fault feature extraction. To quantify the distribution of signal energy among nodes in a wavelet packet decomposition layer, Shannon Entropy  $H$  is employed as follows:

$$H = - \sum_{i=1}^N p_i \cdot \log_2 p_i \quad (2)$$

Here  $p_i$  denotes the probability distribution of the energy among the wavelet coefficients, is presented in the following manner:

$$p_i = \frac{|wt(s, i)|^2}{E_x(s)} \quad (3)$$

where  $wt(s, i)$  is a  $i$  th wavelet coefficient at the  $s$  level.

Then the ratio between energy and Shannon entropy can be expressed as.

$$R(s) = \frac{E_x(s)}{H(s)} \quad (4)$$

The given equation allows for the calculation of the  $R(s)$  ratio at the desired level of decomposition for each potential wavelet base. The candidate wavelet which exhibits the highest energy to Shannon entropy value is selected as the basis for decomposing the given signal or set of signals using the WPT method.

### 3 Proposed Methodology

#### 3.1 Envelope Analysis

The vibration signal from faulty bearings contains high-frequency components resulting from different mechanisms like impact, rubbing, or resonance. These high-frequency elements are frequently concealed by low-frequency components in the signal, which can arise from machine operation, background noise, or measurement noise. To extract these high-frequency components, the Hilbert Transform Envelope Extraction method is employed in the same way as in [6].

### 3.2 Wavelet Base Evaluation Criterion

The selection of a wavelet basis during the procedure of signal decomposition through Wavelet Packet Transform (WPT) significantly influences the spectral qualities of the resultant coefficients. Each wavelet base possesses distinct qualities that make them better suited for capturing particular types of spectral content or signal features, while others may be less effective in doing so. In the standard WPT procedure, outlined in the Sect. 3, a wavelet base is carefully selected from a poll containing  $W$  wavelet bases. Each wavelet base is then applied to decompose a representative subspace of the signal data. The resulting WPT coefficients are subsequently evaluated, and based on this assessment, a final decision is made regarding the most suitable wavelet base for the given signal data.

In contrast, the proposed method endeavors to portray a signal by employing WPT decomposition as a basis, yet it is not constrained to employing only one wavelet base.

In this approach, the signal data is initially decomposed into a specified level, denoted as  $j$ , using a poll of  $W$  wavelet bases. For each node within the decomposition, ranging from  $d_j^0$  to  $d_i^n$ , a score chart is created. This score chart has a length of  $W$ , with each element representing the evaluation results of reconstructed coefficients obtained from the WPT decomposition of this node using different wavelet bases. The assessment of these coefficients is based on the evaluation of their spectral content. This evaluation is performed by calculating the relation of the total power of the spectrum to the Shannon entropy of the signal power spectrum. Considering the Shannon entropy of the signal power spectrum  $H_{ps}$  is defined as:

$$H_{ps} = - \sum_{i=1}^N p_i \log_2(p_i) \tag{5}$$

where  $N$  stands for the number of frequency bins in the signal power spectrum and the probability of the signal power being in the  $i$ -th frequency bin  $p_i$  is defined as:

$$p_i = \frac{P_i}{\sum_{j=1}^N P_j} \tag{6}$$

where  $P_i$  is the power in the  $i$ -th frequency bin.

And the total power of the signal spectrum is computed as:

$$P_{ss} = \sum_{i=1}^N P_i \tag{7}$$

So, the final ratio is defined as:

$$R = \frac{P_{ss}}{H_{ps}} = \frac{\sum_{i=1}^N P_i}{-\sum_{i=1}^N \frac{P_i}{\sum_{j=1}^N P_j} \log_2 \left( \frac{P_i}{\sum_{j=1}^N P_j} \right)} \tag{8}$$

The proposed method utilizes the R-value criterion to assess the reconstructed coefficients and determine the most suitable mother wavelet for representing the signal.

This criterion enables the comparison of spectral content captured by each wavelet and identifies the one that offers the most effective representation. Specifically, the criterion measures the extent to which information in the signal power spectrum is concentrated within specific frequency bands rather than uniformly distributed across the entire spectrum. A preferred wavelet base is one that yields reconstructed coefficients with a higher ratio of signal spectrum total power to its Shannon entropy. This preference indicates a signal with a more predictable and structured spectral composition.

### 3.3 Signal Representation Using WPT with Node-Specific Bases

Aiming for more efficient feature extraction, the proposed method constructs a representation of the input signal using WPT as a fundament. The node-specific approach for the wavelet base selection uses a rating based on the R-value. Each of the  $n$  WPT nodes at decomposition level  $j$  has its unique rating of the WPT-bases, where the best base for a particular node has the highest R-value. It means that it is possible to obtain more useful spectral contents for efficient feature extraction from the reconstructed signal of each particular node if it was selected from a WPT tree decomposed using the wavelet base with the highest R-value. Thus, the final representation is obtained when each of the  $n$  nodes was selected. This process is illustrated in Fig. 2.

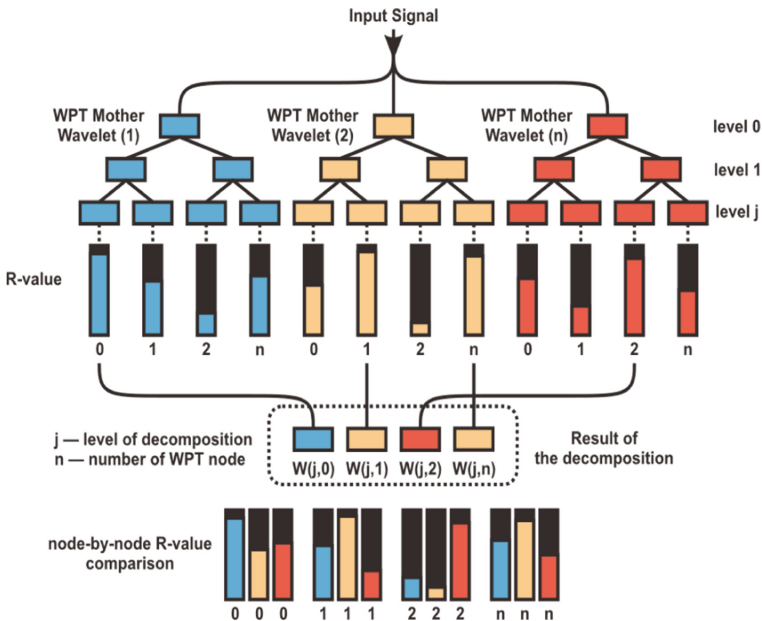


Fig. 2. Schematic for WPT node-specific base selection.

### 3.4 Statistical Feature Extraction

In order to represent the obtained data in fewer variables, 19 statistical features were extracted from the time and frequency domain of each of the eight reconstructed signals obtained in Sect. 4. Features from each node were concatenated and in the result one second sample of the vibration signal is now represented by 152 features. Table 1 contains the names and formulas for each of these 19 features.

**Table 1.** Statistical features and their formulas.

Feature	Equation	Feature	Equation
Peak	$X_p = \max_i  x_i $	5 <sup>th</sup> normalized moment	$HOMn5 = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^5}{\sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2}}$
Root-mean square	$X_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$	6 <sup>th</sup> normalized moment	$HOMn6 = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^6}{\sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2}}$
Kurtosis	$X_{kurtosis} = \frac{1}{N} \left( \frac{\sum_{i=1}^N (x_i - \mu)^4}{\sigma^4} \right)$	Skewness	$X_{skewness} = \frac{1}{N} \left( \frac{\sum_{i=1}^N (x_i - \mu)^3}{\sigma^3} \right)$
Crest factor	$C_f = \frac{X_p}{X_{RMS}}$	Shape factor RMS	$SF_{RMS} = \frac{X_{RMS}}{\mu}$
Clearance factor	$L = \frac{X_p}{\left( (1/N) \sum_{i=1}^N \sqrt{ x_i } \right)^2}$	Peak-to-peak value	$x_{pp} = \max  x  - \min  x $
Impulse factor	$L = \frac{\max_i \{ x_i \}}{\left( (1/N) \sum_{i=1}^N  x_i  \right)}$	Energy of signal	$e = \sum_{i=1}^N x_i^2$
Shape factor SMR	$SF_{SMR} = \frac{X_{SMR}}{\mu}$	Frequency center	$FC = \frac{\int_0^{\infty} f s(f) df}{\int_0^{\infty} s(f) df}$
Entropy	$H(x) = - \sum_{i=1}^N P(x_i) \cdot \log_2 P(x_i)$	RMS frequency	$RMSF = \frac{\sqrt{\int_0^{\infty} f^2 s(f) df}}{\sqrt{\int_0^{\infty} s(f) df}}$
Mean	$\mu = \frac{1}{N} \sum_{i=1}^N x_i$	Root variance frequency	$RVF = \sqrt{\frac{\int_0^{\infty} (f_i - FC)^2 s(f_i) df}{\int_0^{\infty} s(f_i) df}}$
Square mean root	$X_{SMR} = \left( \frac{\sum_{i=1}^N \sqrt{x_i}}{N} \right)^2$		

## 4 Experimental Test Bed and Data Collection

The proposed method was evaluated using four bearing-fault datasets to ensure validity and reliability. Three publicly available datasets were obtained from the KAt-DataCenter of Paderborn University (PU) in Germany and Case Western Reserve University (CWRU) [10, 11]. PU datasets are denoted as PUA and PUR for artificial and real bearing faults.

The PU dataset's vibration data was collected using a modular test rig comprising an electric motor (Hanning ElektroWerke GmbH & Co. KG), measuring shaft, a module for bearing installation, flywheel, and motor used for load simulation. The electric motor is a 425 W PMSM controlled by an industrial inverter (KEB Combivert 07F5E 1D-2B0A). Four different experimental conditions used for each bearing are shown in Table 2. Vibration data was obtained using a piezoelectric accelerometer installed on top of the bearing module and sampled at 64 kHz rate. For the current research the signals were downsampled with the factor of eight for faster computations. Each dataset sample represents one second of vibration signal. PU dataset has signals from six healthy bearings with a run-in period of one to 50 h. PUA dataset has 12 bearings with faults inflicted using electric discharge machine, electric engraving, and drilling. Faults have depths of 1–2 mm for EDM trenches and lengths of 1–4 mm for electric engraving. All faults are categorized as Healthy, Outer Ring Fault, and Inner Ring Fault, and their arrangement and codes are demonstrated in Table 3.

**Table 2.** PU test rig parameters of operation.

No.	Angular velocity [rad/sec]	Torque of the Load [Nm]	Radial Load [N]
0	157.08	0.7	$10^3$
1	94.25	0.7	$10^3$
2	157.08	0.1	$10^3$
3	157.08	0.7	$4 * 10^2$

**Table 3.** PUA dataset composition.

Bearing type and class	Bearing letter and code
Healthy	[K]: 001-006
Outer Ring Fault	[KA]: 01, 03, 05, 06, 07, 08, 09
Inner Ring Fault	[KI]: 01, 03, 05, 07, 08

The PUR dataset includes 14 bearings with accelerated lifetime faults inflicted by a specially developed machine, imitating natural fault infliction with the extensive radial load. Damages appear as pitting or plastic deformations caused by debris and were categorized into three levels based on the affected area's length on the ring surface.



Bearings were classified as having outer ring fault, inner ring fault, or both. Rolling elements remained intact. PUR dataset arrangement with bearing codes is presented in Table 4.

**Table 4.** PUR dataset composition.

Bearing type and class	Bearing letter and code
Healthy	[K]: 001-006
Outer Ring Fault	[KA]: 04, 15, 16, 22, 30
Inner Ring Fault	[KI]: 04, 14, 16, 17, 18, 21
Outer + Inner Ring Fault	[KB]: 23, 24, 27

The data obtained from Case Western Reserve University (CWRU) was collected from an experimental setup with a two-horsepower motor, accelerometers, one SKF6205 bearing positioned at the drive end and another installed at the fan end, and a device for measuring rotational force. Vibrations were recorded at 12000 and 48000 samples per second via a 16-channel DAT recorder. Faults were seeded on the following bearing parts: the inner race, the outer race, and the ball using spark erosion tool. The faults ranged in diameter from 0.007 inches to 0.040 inches and were positioned at three o'clock, six o'clock, and 12 o'clock. The data were digitized at a rate of 12000 samples per second and divided into one-second segments, resulting in a dataset with dimensions of  $1920 \times 12000$ . The arrangement of CWRU data is shown in Table 5.

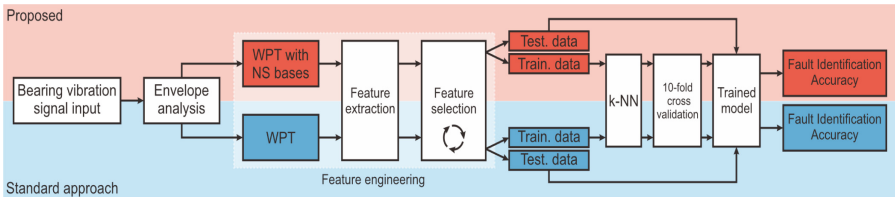
**Table 5.** CWRU dataset composition.

Class names	Bearing identification number
Healthy	97-100
Outer Ring Fault	130-133, 144-147, 156-160, 197-200, 234-237, 246-249, 258-261
Inner Ring Fault	056-059, 105-108, 169-172, 209-212
Ball Damage	048-051, 118-121, 185-188, 222-225

## 5 Performance Evaluation and Discussion

The proposed method constitutes only a portion of the bearing fault diagnosis framework. However, to evaluate its effectiveness for fault diagnosis through comparative analysis, two complete fault diagnosis frameworks were constructed. The structure of these frameworks is illustrated in Fig. 3. Signal processing and feature extraction stages were established based on the description in Sect. 3. The only discrepancies between the two frameworks lie in the feature engineering stage, specifically in the way the

vibration signal is processed after envelope analysis. The framework highlighted in red employs the proposed WPT with node-specific base selection, while the framework highlighted in blue uses the standard WPT. The white elements are identical for both frameworks. To ensure the validity of the comparison, three feature selection approaches were employed. The first approach utilizes the entire feature vector obtained after feature selection, the second approach utilizes Principal Component Analysis to reduce dimensionality by employing a linear combination of original features, and the third approach is a wrapper-based Boruta method that selects features based on their importance score for the Random Forest model. After feature selection, the data is randomly divided, leaving 80% of the data for training and 20% for testing. The 80% chunk is used for training of the k-NN model, and the performance of the trained model is validated using k-fold cross-validation. K-NN model is selected due to its low computational expensiveness and its instance-based nature. It is a non-parametric algorithm that refrains from making any presumptions about the inherent data distribution. Instead, its method of classification involves assessing new instances by contrasting them against the labeled instances present in the training data. Its performance heavily relies on the quality of the features. If the feature space is not well-defined or if irrelevant features are included, the algorithm may not perform well, which means that the quality of the feature set can be fairly assessed by the k-NN performance.



**Fig. 3.** Bearing fault diagnosis framework for performance comparison.

This bearing diagnosis framework underwent testing on three datasets, and Table 6 presents the accuracy of fault identification for each dataset individually, as well as the average performance across all three datasets. By testing the proposed feature extraction method combined with various feature selection methods, it can be concluded that the suggested approach of WPT node-specific base selection produces signal representations that enhance the performance of fault diagnosis frameworks employing traditional statistical feature extraction and selection techniques, as well as conventional Machine Learning models.

**Table 6.** Fault identification accuracy comparison.

WPT type	Feature Selection	PU (art) Acc. %	PU (real) Acc. %	CWRU Acc. %	Avg. Acc. % all datasets
Proposed	PCA	90.97	98.43	87.76	<b>92.39 (+1.09)</b>
	Boruta 20 features	98.00	99.84	97.14	<b>98.33 (+1.78)</b>
	Boruta 10 features	96.61	99.69	96.88	<b>97.73 (+3.17)</b>
	Boruta 5 features	96.70	99.68	95.05	<b>97.14 (+2.96)</b>
	None	93.40	98.75	94.53	<b>95.56 (+1.24)</b>
Standard	PCA	90.02	98.44	85.42	<b>91.29</b>
	Boruta 20 features	96.35	100.00	93.29	<b>96.55</b>
	Boruta 10 features	96.18	100.00	87.50	<b>94.56</b>
	Boruta 5 features	94.96	99.84	87.76	<b>94.19</b>
	None	91.84	99.45	91.67	<b>94.32</b>

## 6 Conclusions

The current study proposes a node-specific approach for wavelet packet transform (WPT) base selection for feature selection in bearing fault diagnosis. The traditional approach focuses on a single WPT node with high signal energy, potentially excluding important features present in other nodes. To address this limitation, a criterion based on the energy-entropy ratio of the signal spectrum for each node was introduced. This criterion evaluates the ability of WPT bases to generate signals with concentrated energy in specific frequency bands. By selecting the WPT bases with the highest criterion score, our method ensures the preservation of meaningful components and their distinction from noise. Upon the evaluation using three bearing fault datasets, it was found that on average across three datasets, the proposed method outperforms the traditional approach in fault diagnosis performance.

Our approach provides several benefits, including a comprehensive representation of the signal, explicable diagnostic procedures, and low computational cost. Nevertheless, it's crucial to acknowledge that while our suggested technique has its roots in the Wavelet Packet Transform (WPT), it does not adhere to certain fundamental properties of wavelet decomposition, such as energy conservation and superposition. Thus, it's inappropriate to label it as an advanced WPT technique. Nonetheless, our proposed method remains a valuable tool for feature extraction, leveraging well-founded mathematical principles that underlie the applied signal manipulations, establishing a strong basis for its use. As a result, the extracted features offer valuable information about the signal, which can be applied in various applications including signal processing, classification, and pattern recognition. In conclusion, our node-specific approach improves the accuracy and reliability of bearing fault diagnosis by enhancing feature extraction capabilities. Future work can focus on optimizing the criterion and exploring its applicability to other signal analysis tasks as long as performing more tests with the Deep Learning models.

**Acknowledgements.** This work was supported by the Technology Innovation Program (20023566, Development and Demonstration of Industrial IoT and AI Based Process Facility Intelligence Support System in Small and Medium Manufacturing Sites) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea). This work was also supported by the Technology Infrastructure Program funded by the Ministry of SMEs and Startups (MSS, Korea).

## References

1. Bazurto, A.J., Quispe, E.C., Mendoza, R.C.: Causes and failures classification of industrial electric motor. In: 2016 IEEE ANDESCON, pp. 1–4. IEEE, Arequipa, Peru (2016)
2. Smith, W.A., Randall, R.B.: Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study. *Mech. Syst. Signal Process.* **64–65**, 100–131 (2015)
3. Chen, L., Xu, G., Tao, T., Wu, Q.: Deep Residual network for identifying bearing fault location and fault severity concurrently. *IEEE Access* **8**, 168026–168035 (2020)
4. Skora, M., Ewert, P., Kowalski, C.T.: Selected rolling bearing fault diagnostic methods in wheel embedded permanent magnet brushless direct current motors. *Energies* **12**(21), 4212 (2019)
5. Maliuk, A.S., Prosvirin, A.E., Ahmad, Z., Kim, C.H., Kim, J.M.: Novel bearing fault diagnosis using Gaussian mixture model-based fault band selection. *Sensors* **21**(19), 6579 (2021)
6. Maliuk, A.S., Ahmad, Z., Kim, J.M.: Hybrid feature selection framework for bearing fault diagnosis based on Wrapper-WPT. *Machines* **10**(12), 1204 (2022)
7. Pham, M.T., Kim, J.M., Kim, C.H.: Accurate bearing fault diagnosis under variable shaft speed using convolutional neural networks and vibration spectrogram. *Appl. Sci.* **10**(18), 6385 (2020)
8. Ke, Z., Di, C., Bao, X.: Adaptive suppression of mode mixing in CEEMD based on genetic algorithm for motor bearing fault diagnosis. *IEEE Trans. Magn.* **58**(2), 1–6 (2022)
9. Lei, Y., Lin, J., He, Z., Zuo, M.J.: A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.* **35**(1–2), 108–126 (2013)
10. Lessmeier, C., Kimotho, J.K., Zimmer, D., Sextro, W.: Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: a benchmark data set for data-driven classification, p. 17 (2016)
11. “Welcome to the Case Western Reserve University Bearing Data Center Website|Case School of Engineering|Case Western Reserve University,” Case School of Engineering, 10 August 2021. <https://engineering.case.edu/bearingdatacenter/welcome>. Accessed 20 Oct 2022