Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 7-8, 2026

Intelligent Signal Classification for False Alarm Reduction via Multi-domain Feature Transformation

*DongHyeon Shin[a], Hee-Jae Lee[b], and Jonghyun Kim[b]**
[a]Korea University, Department of Electro-Mechanical Convergence Engineering, 2511, Sejong-ro, Sejong-si, Republic of Korea
[b]Korea Advanced Institute of Science and Technology, 291, Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea
**Corresponding author: jonghyun.kim@kaist.ac.kr*

***Keywords**∶* Loose Part Monitoring System, False Alarm Reduction, Signal Processing, Random Forest

1. Introduction

The Loose Part Monitoring System (LPMS) is a representative system designed to detect metallic loose parts or structural debris within a nuclear reactor. This system prevents secondary damage such as flow blockage in fuel channels, interference with control rod operation, or ruptures in steam generator tubes by providing early warnings. Despite its critical importance, many aging systems suffer from performance degradation and high false alarm rates [1]. Accordingly, there is a growing demand for the development of enhanced monitoring technologies that integrate precise signal processing and diagnostic tools. Regulatory guidelines, such as the U.S. Nuclear Regulatory Commission (NRC) Regulatory Guide 1.133, emphasize that the system must possess the signal discrimination capability to clearly distinguish actual impact signals from normal operational noise or electrical interference [2]. However, in actual nuclear power plant operations, false alarms occur due to non-impact signals such as cable vibration, which exhibits characteristics similar to those of impact signals [3]. As shown in Fig. 1, the conventional LPMS employs a threshold-based defect determination process for signals collected from accelerometers. This approach has a structural vulnerability in which the appropriate threshold must be adjusted according to the changing operating environment in the time domain. It requires continuous threshold adjustment due to changing operating conditions and rapid background noise variations during startup and shutdown. This often leads to cases where non-impact signals pass through the algorithm and trigger false alarms, serving as a major factor that degrades the diagnostic reliability of the system. The purpose of this study is to establish an intelligent classification schema to overcome these limitations. The proposed framework replaces existing methods with a multi-domain signal processing approach to achieve noise-robust diagnosis.

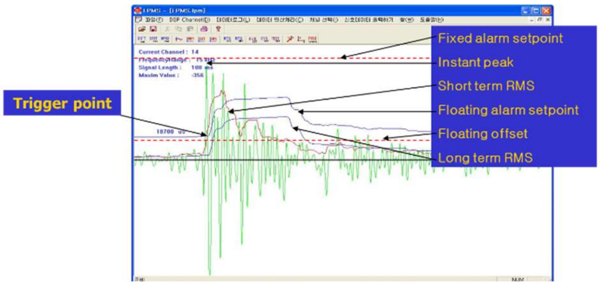
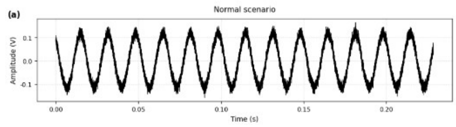


Fig. 1. Conventional LPMS Alarm Flow [3]

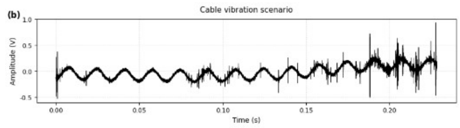
2. Data Collection

2.1 Data acquisition environment

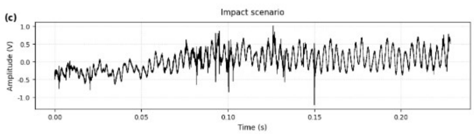
The experimental setup was configured as follows to reproduce signal patterns occurring in actual environments and to secure measurement data. B&K Type 5974 piezoelectric sensors attached to a stainless steel plate were placed at points approximately 5cm and 150cm away from the steel ball impact location, respectively, and data was collected for each 10-second scenario. The analysis dataset consisted of 80 normal states without external interference, 55 actual impacts using a steel ball, and 60 cable shaking noises.



(a) Normal Scenario



(b) Cable Vibration Scenario



(c) Impact Scenario

Fig. 2. Field-Collected Data

2.2 Overview of the Signal Classification

The proposed signal classification algorithm is structured as a multi-stage pipeline. It consists of data preprocessing, multi-domain signal transformation, statistical feature extraction, and machine learning-based classification. First, the collected signals are segmented and normalized. Subsequently, seven signal processing techniques are applied to transform the signals. Five key statistical indicators are then extracted from each domain to form a fixed-length feature vector. Finally, Random Forest, an ensemble learning model, trains on this multi-dimensional feature set to distinguish between normal states, cable noise, and impact signals. This compensates for the limitations of a single analysis domain and enhances the classification precision and reliability of the diagnostic system.

2.3 Signal Processing methods

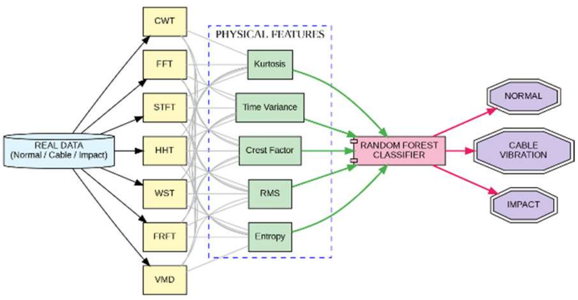


Fig. 3. Signal Classification Schema

In this section, signal characteristics were extracted through the following signal processing techniques

2.3.1 Wavelet Scattering Transform (WST)

WST is a technique that yields stable data even for minute changes in signals by combining the detailed feature extraction capability of wavelets with non-linear operations and low-pass filters [4]. It possesses translation invariance, where the output remains constant despite shifts in the time domain or local deformations. This allows for the generation of robust feature vectors regardless of location in this experimental environment, where distance changes between the sensor and the impact point are frequent. In particular, since feature extraction is possible through mathematical structure alone without separate training, high classification accuracy is secured without statistical bias even in environments with limited datasets.

2.3.2 Hilbert-Huang Transform (HHT)

HHT is a customized processing technique for analyzing irregular and highly variable signals, consisting of EMD (Empirical Mode Decomposition), which breaks down a signal into multiple components, and Hilbert spectral analysis, which calculates the instantaneous frequency of each component [5]. By adaptively deriving analysis criteria from the signal, HHT precisely captures sudden energy shifts and subtle frequency features during impact. However, a Mode Mixing phenomenon may occur during the signal decomposition process, where signals with different characteristics are mixed into a single mode. Consequently, if actual impact signals and noise such as cable shaking become mixed, the resolution for accurately separating the signals decreases in false alarm reduction situations where false signals must be filtered out.

2.3.3 Fractional Fourier Transform (FRFT)

FRFT is a domain transformation technique that rotates the time–frequency plane at an arbitrary angle. This enables exploration of intermediate regions between time and frequency to reveal latent signal characteristics. Optimized for analyzing signals with linear frequency modulation characteristics, it effectively detects overlapping impact signals that are difficult to distinguish in general domains and suppresses background noise [6]. However, when processing multiple scenario data, the optimal transformation order (α) that maximizes energy must be calculated through iterative operations, which slows down the overall processing speed when applied to a real-time diagnostic system.

2.3.4 Fast Fourier Transform (FFT)

FFT is a standard algorithm that significantly reduces the computational load of the Discrete Fourier Transform, quickly decomposing signals into frequency components. While it shows great strength in real-time monitoring due to its exceptional processing speed, it has a structural characteristic where time information is lost during the transformation process [7]. Because of this, it is impossible to know exactly at what point in time an impact occurred, making it difficult to precisely distinguish physical impact patterns that change over time from background noise.

2.3.5 Short-Time Fourier Transform (STFT)

STFT is a technique that observes frequency changes over time by dividing a signal into short segments and performing a Fourier transform on each segment. It enables intuitive analysis of frequency distribution at specific points in time by visualizing the time-frequency plane in a grid format [8]. However, STFT is limited by the reciprocal relationship between time and frequency resolutions. Consequently, the resolution decreases when accurately capturing metal impact sounds that occur in a very short instant or high-frequency components that disappear momentarily.

2.3.6 Continuous Wavelet Transform (CWT)

CWT is an analysis technique that measures the similarity with a signal by continuously adjusting the scale and shift parameters of a Mother Wavelet. It possesses multi-resolution analysis characteristics, making it excellent at capturing transient phenomena, including sudden impact components or rapid potential changes [9]. However, since performance depends heavily on the selected wavelet basis, an optimal function tailored to the experimental environment must be identified.

2.3.7 Variational Mode Decomposition (VMD)

VMD is a technique that analyzes an input signal by decomposing it into multiple components (intrinsic modes) centered around specific frequency bands. By minimizing the frequency range of each component, it blocks interference between signals and precisely isolates desired features [10]. However, since the number of modes (K) and the balancing parameter (α) suitable for the data characteristics must be set in advance, the algorithm complexity increases during the process of automating these settings.

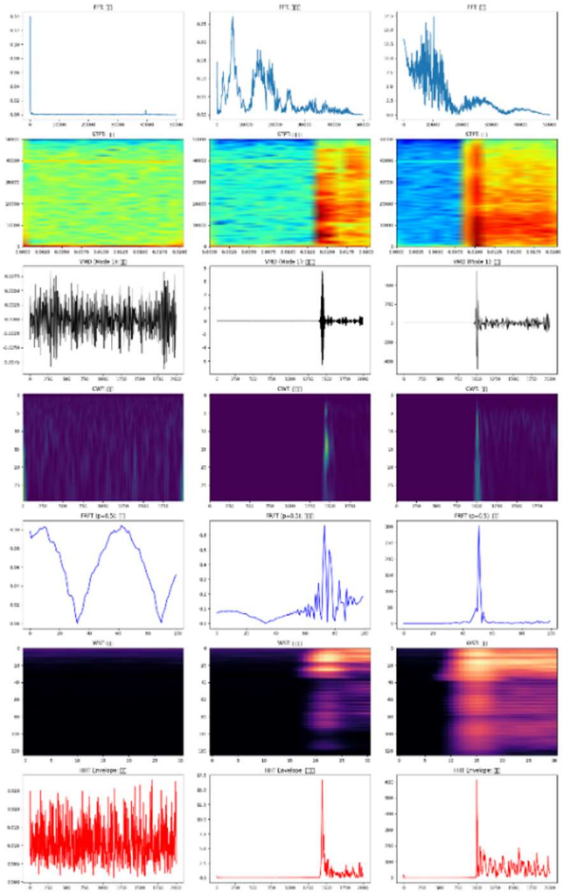


Fig. 4. Comparison of signal processing techniques across different states (Rows from top: FFT, STFT, VMD, CWT, FRFT, WST, HHT / Columns from left: Normal, Cable, Impact)

As confirmed through Fig. 4., the normal state maintains a constant baseline pattern with low energy distribution across all regions, whereas in the impact state, sharp vertical peaks or high-intensity energy bands where energy is highly concentrated are commonly observed across all techniques. Cable interference shows lower concentration than an impact but exhibits larger and more irregular energy variability than the normal state, forming a unique noise pattern. Looking at each technique, FFT and STFT represent the sudden energy rise of broadband frequencies occurring during an impact as density on a plane, while VMD and CWT reveal fine vibration components and transient phenomena included in the impact wave through precise amplitude changes and multi-resolution patterns. Additionally, FRFT clearly separates noise and signals through energy aggregation in the intermediate domain, WST shows the density differences in scattering coefficients uniquely formed for each state, and HHT intuitively extracts sharp peaks at the moment of impact.

3. Design of Intelligent Signal Classification Algorithm

3.1 Feature Selection

This section describes the design of the intelligent signal processing algorithm. The following physical features were used for signal classification in the algorithm development [11].

3.1.1 RMS (Root Mean Square)

RMS is the square root of the arithmetic mean of the squares of the values, representing the overall energy intensity of the signal. It is the most fundamental indicator for quantifying the magnitude of vibration or impact and is used to detect overall changes in the system's state.

3.1.2 Entropy

Entropy is an indicator that measures the complexity and irregularity of the signal distribution. The more disordered the signal is without a constant pattern, the higher the value, which is useful for determining the occurrence of unexpected abnormal signals in contrast to the normal state.

3.1.3 Kurtosis

Kurtosis is a measure of how sharp the probability distribution of a signal is. Since the value rises sharply as the signal contains sharp peaks or temporary impulsiveness, it plays a key role in capturing bearing defects or initial impact signals.

3.1.4 Time Variance(T-var)

T-var represents the variability of energy over time. By dividing the signal into several intervals and calculating the standard deviation of the RMS for each interval, it quantifies how constant or irregularly fluctuating the signal is over time.

3.1.5 Crest Factor (C.F)

C.F is the ratio of the signal's peak value to its RMS level. It quantifies the intensity of sudden impulses relative to the overall energy, effectively isolating instantaneous impacts from continuous vibrations.

3.2 Dataset Preparation

To establish the model's learning algorithm, prevent overfitting, and ensure objective performance verification, the entire dataset was divided into a 40:30:30 ratio (Train: Validation: Test). The Train Set (40%) is used to generate the decision tree structure of the Random Forest model and learn the weights for each feature, while the Validation Set (30%) serves as reference data to monitor for overfitting in real-time during the learning process and to verify generalization performance. Finally, the Test Set (30%) consists of an independent group of data that was not involved in the learning process at all, and it was used to objectively evaluate the final diagnostic reliability of the established intelligent algorithm.

3.3 Time Window Optimization

The time window was optimized for a 100 kHz sampling rate to ensure feature relevance and learning efficiency. Sufficient observation time is required to enable reliable pattern discrimination across the full impact event, making the window interval a critical factor in diagnostic performance. To this end, this study conducted repetitive experiments while variably adjusting the analysis range between 0.1 and 5 seconds. As a result, the 2 second interval was finalized as the unit of analysis, as it yielded the highest diagnostic accuracy by fully including the attenuation characteristics where energy gradually decreases after an impact while minimizing the loss of global patterns across the signal.

3.4 Classification Model

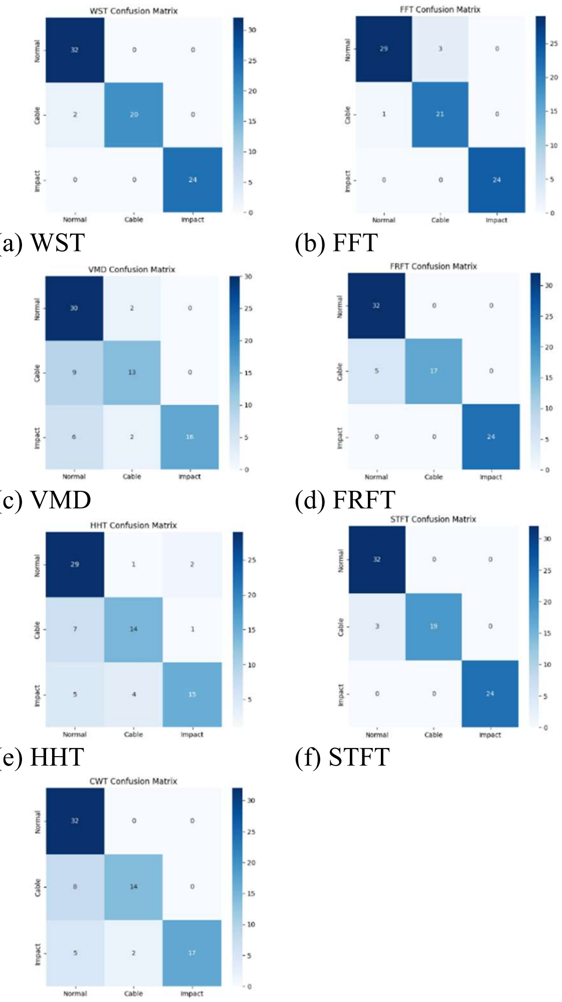
3.4.1 Random Forest:

Random Forest is an ensemble learning method consisting of a combination of tree predictors generated based on independently sampled random vectors. For classification, the final result is determined by the class that receives the most votes from each tree. As the number of trees increases, the error rate converges to a specific limit, making the model highly robust against overfitting problems and ensuring high learning reliability [12].

4. Training and Results

4.1 Training

In this study, the selected Random Forest model was used to train and comparatively analyze the performance of each signal processing technique on a batch basis, and the resulting Confusion Matrix for each methodology is shown in Fig. 6. To ensure consistency and reproducibility of the results, a fixed random seed (random_state=42) was applied to control the randomness of the training process. Furthermore, the maximum depth of the trees was limited to 8 (max_depth=8) to prevent the model from becoming overly optimized to local noise, which could degrade generalization performance during actual diagnosis. In addition, the weights for each class were automatically adjusted to increase the fairness of discrimination, ensuring that impact signals with a small amount of data were not marginalized during training. The final performance evaluation was conducted by inputting independent test data (30%) which had not been used for training into the model optimized during the validation phase, through which each technique successfully filtered out impact signals from complex noise.



(a) WST (b) FFT

(c) VMD (d) FRFT

(e) HHT (f) STFT

(g) CWT

Fig. 5. Confusions of signal processing techniques

4.2 Results

After testing the accuracy of each signal processing technique across various Time Window environments, WST (Wavelet Scattering Transform) recorded a relatively superior performance trend among the compared methods. According to the test results, WST maintained a steady level up to the 1.5 second mark, showed a significant leap in performance upon entering the 2 second interval, and consistently maintained high accuracy even as the analysis time increased thereafter.

While STFT showed a gradual upward trend as time resources increased, reaching its peak performance at the 4 second mark, FFT formed an inflection point at 1.5 seconds, with accuracy actually decreasing as the analysis time lengthened. This decline may be attributed to the absence of temporal localization in FFT, which makes transient features less distinctly represented in extended windows. Other techniques, such as CWT, HHT, and VMD, started in a lower performance band and showed a modest improvement. The classification accuracy for cable vibrations varied significantly depending on the signal processing technique applied. Wavelet Scattering Transform (WST) achieved the highest performance at 90.91%, whereas Variational Mode Decomposition (VMD) reached only 54.55%. This performance gap stems from how each technique transforms the raw signal into a feature space. Methods such as FFT, STFT, WST, and FRFT represent signals through global frequency characteristics within a consistent structure. In this process, the energy distribution characteristics of cable vibrations remain relatively stable, which is consistently reflected in statistical metrics like RMS and Entropy. Consequently, a feature space is formed where the distinction between steady-state and impact signals is clearly defined.

In contrast, adaptive decomposition techniques like CWT, VMD, and HHT offer the flexibility to decompose signals based on local variations. However, for irregular, low-energy signals like cable vibrations, the signal components tend to be dispersed across multiple decomposition levels. Even when using the same statistical measures, this dispersion causes inconsistent feature representations across different techniques, leading to blurred boundaries between classes. Ultimately, while the same Random Forest classifier was used, the inherent separability of the input feature space varied by technique, directly resulting in the observed differences in final classification performance.

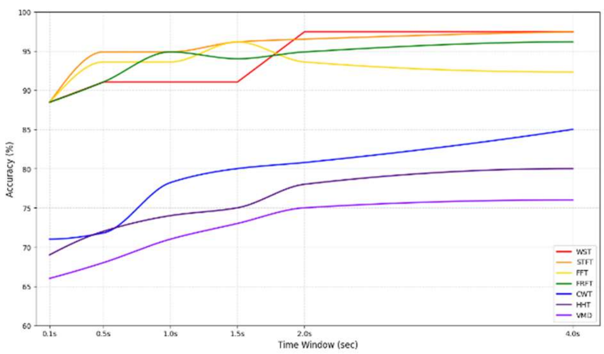


Fig. 6. Performance Analysis by Time Window

Table I. Classification Report

Signal Processing	Scenario	Accuracy	%
WST	Normal	32/32	100
	Cable Vibration	20/22	90.91
	Impact	24/24	100
	Total	76/78	97.44
FFT	Normal	32/32	100
	Cable Vibration	18/22	81.82
	Impact	24/24	100
	Total	74/78	94.87
STFT	Normal	32/32	100
	Cable Vibration	19/22	86.36
	Impact	24/24	100
	Total	75/78	96.15
CWT	Normal	32/32	100
	Cable Vibration	14/22	63.64
	Impact	17/24	70.83
	Total	76/78	80.77
VMD	Normal	32/32	100
	Cable Vibration	12/22	54.55
	Impact	13/24	54.17
	Total	76/78	73.08
FRFT	Normal	32/32	100
	Cable Vibration	17/22	77.27
	Impact	24/24	100
	Total	76/78	93.59
HHT	Normal	30/32	93.75
	Cable Vibration	13/22	59.09
	Impact	16/24	66.67
	Total	76/78	75.64

5. Conclusion

The results indicate that WST is the most stable and reliable technique among the evaluated methods. Its invariant feature extraction across time and frequency domains significantly improved Random Forest performance. For the total of 78 test samples, WST recorded an accuracy of 76/78 (97.44%), exhibiting the highest classification performance across all conditions: Normal 32/32 (100%), Cable Vibration 20/22 (90.91%), and Impact 24/24 (100%).

While the conventional FFT and STFT showed accuracies of 74/78 (94.87%) and 75/78 (96.15%) respectively, they demonstrated high volatility, with accuracy fluctuating irregularly depending on the time window settings. In contrast, WST exhibited robust characteristics against minute signal variations; after significantly improving at the 2-second interval, it maintained consistent performance without degradation even as the analysis time increased. Therefore, WST effectively enhances classification performance in complex signal environments while reducing variability.

The proposed framework will incorporate RNN and Transformer architectures to automate feature extraction and improve diagnostic performance. By applying multiple techniques in parallel and conducting multi-faceted comparative analyses, we aim to derive an optimized model and establish an intelligent diagnostic algorithm capable of flexibly responding to real-world environmental changes.

Acknowledgements

This work was supported by the Innovative Small Modular Reactor Development Agency grant funded by the Korea Government (MIST) (No. RS-2023 00258052) .

REFERENCES

[1] W. D. Rhodes and D. W. Langenberg, Export control guide: loose parts monitoring systems for nuclear power plants (ORNL/TM-2012/601), Oak Ridge National Laboratory, 2012.

[2] U.S. Nuclear Regulatory Commission, Loose-part detection program for the primary system of light-water-cooled reactors, Regulatory Guide 1.133, Rev. 1, 1981.

[3] D. Yoon, Personal communication, March 20th, 2024.

[4] R. N. Toma, Y. Gao, F. Piltan, K. Im, D. Shon, T. H. Yoon, D. S. Yoo, and J. M. Kim, Wavelet scattering transform, Encyclopedia, Vol. 2, No. 2, pp. 1161-1172, 2022.

[5] D. C. Bowman and J. M. Lees, The hilbert-huang transform: a high resolution spectral method for nonlinear and nonstationary time series, Seismological Research Letters, Vol. 84, No. 6, pp. 1074-1080, 2013.

[6] D. Mendlovic and H. M. Ozaktas, Fractional fourier transforms and their optical implementation: i, Journal of the Optical Society of America A, Vol. 10, No. 9, pp. 1875-1881, 1993.

[7] H. J. Nussbaumer, The fast fourier transform, Fast Fourier Transform and Convolution Algorithms, Springer Series in Information Sciences, Vol. 2, 1981.

[8] D. Goyal and B. S. Pabla, Condition based maintenance of machine tools-a review, CIRP Journal of Manufacturing Science and Technology, Vol. 10, pp. 24-35, 2015.

[9] M. Wolter, F. Blanke, J. Garcke, and C. T. Hoyt, ptwt - the pytorch wavelet toolbox, Journal of Machine Learning Research, Vol. 25, No. 80, pp. 1-7, 2024.

[10] K. Dragomiretskiy and D. Zosso, Variational mode decomposition, IEEE Transactions on signal processing, Vol. 62, No. 3, pp. 531-544, 2014.

[11] W. Caesarendra and T. Tjahjowidodo, A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing, Machines, Vol. 5, No. 4, p. 21, 2017.

[12] L. Breiman, Random forests, Machine Learning, Vol. 45, No. 1, pp. 5-32, 2001.