DOI: 10.61782/fa.2025.0400

FORUM ACUSTICUM EURONOISE 2025

FAULT DETECTION IN C-EPS BEARINGS BASED ON UNSUPERVISED LEARNING

Hyeon-Cheol Jo[1*] **Jun-Seo Park**[1] **In-je Cho**[1] **Jae-Yong Seo**[1] **Seong-Sik Yoo**[1] **Sangkyeong Lee**[1]
Seung-Je Bang[2] **Jeong-Lyul Oh**[2] **Chol-Hwan Jung**[2] **Sung-Zoon Cho**[2]
[1] Hyundai Mobis, 17-2, Mabuk-ro 240beon-gil, Giheung-gu, Yongin-si, Republic of Korea
[2] Big Data AI Center, Seoul National University, Republic of Korea

ABSTRACT

The Column Type Electric Power Steering (C-EPS) system consists of a motor, a reduction gear, and bearings that ensure structural stability and minimize friction. Noise in rotating systems mainly arises from motor noise, component defects, and frictional sounds caused by rotational dynamics.

Operational noise defects are typically managed within regulatory thresholds. Recently, machine learning-based anomaly detection models have gained popularity, often relying on labeled datasets for training. However, this process demands substantial human and time resources for labeling, and distinguishing between noise types remains a significant challenge.

In response to these challenges, this study introduces a method that preprocesses noise input data using Short-Time Fourier Transform (STFT), utilizes unsupervised learning for data encoding, and applies clustering to generate labels. The effectiveness of the proposed approach is demonstrated through a validation process.

Commonly Mel-spectrogram and MFCC transformations are used for AI noise input, but bearing noise often exhibits distinct high-frequency features. STFT was chosen to preserve high-frequency characteristic without attenuation. Various unsupervised learning techniques were utilized to encode the noise data effectively. As the goal was clustering rather than generation, and C-EPS noise shows limited temporal variation, experiments identified Convolutional Autoencoder as the effective unsupervised learning method for mapping noise.

Keywords: *Convolutional Autoencoder, Column Type Electric Power Steering, Self-Labeling, Dimensionality Reduction, Anomaly Detection*

1. INTRODUCTION

As shown Fig. 1, the steering system required to steer in a vehicle is divided into three main parts: the steering gear system that changes the rotational force into linear force, the Column Type Electric Power Steering system (C-EPS) [1] that helps the driver to steer with the power of motor [2] and the universal joint that connects the steering gear system and C-EPS. Among them, as shown Fig. 2, the C-EPS which consist of motor and the reducing gear system causes a lot of noise problems. In general, these noises cause emotional quality degradation to customers who drive the vehicle, which incurs steady field claim costs to companies. [3] One of the consistently problematic noise issues in C-EPS originates from the bearings in the worm shaft system. This bearing-induced noise can be broadly classified into frictional noise and rotational noise. However, there are limitations in detecting it using conventional quantitative noise indicators such as overall level and order noise level. Recently, supervised learning-based machine learning models have been introduced for anomaly detection. However, these methods still rely on subjective human evaluations, and there are ambiguous cases where classification based on subjective assessment is challenging. As a result, the accuracy of labeling inherently suffers, and a significant human resource is required. In addition, conventional quantification methods such as overall level

[*]***Corresponding author***: *clfclfdlsp@mobis.com.*

and order-based filtering are limited in effectively analyzing bearing noise in C-EPS systems.

Consequently, this study aims to identify an optimized model for detecting bearing noise defects in C-EPS using an unsupervised learning approach. In addition, the labeling of bearing noise based on unsupervised learning can be further utilized in combination with supervised learning and explainable AI (XAI) to establish new criteria. In other words, it enables the development of an end-to-end model capable of generating noise specifications even in the absence of prior knowledge about the noise characteristics.

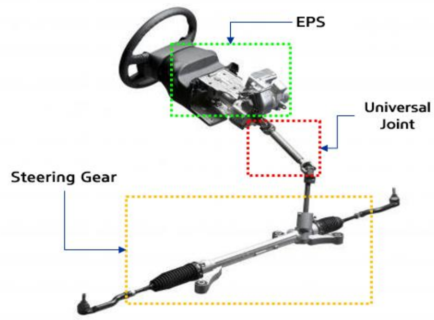


Figure 1. Composition of steering system of vehicle

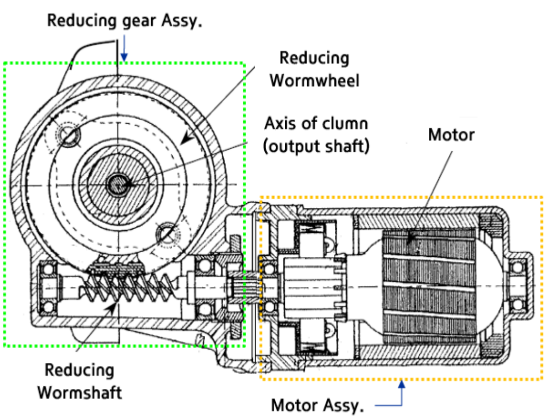


Figure 2. C-EPS Structure Diagram

2. EVALUATION OF NOISE AND VIBRATION

2.1 Evaluation Method

To measure the operational noise of the C-EPS, tests are conducted under constant velocity or constant acceleration conditions using equipment that controls the torque while considering the real vehicle conditions of the input and output, as shown in Fig. 3. As shown in the STFT colormap of the C-EPS operational noise in Fig.4, the advantage of constant acceleration evaluation is that it enables the identification of frequency characteristics across a range of rotational speeds in a single test. Reflecting this advantage, constant acceleration evaluation is commonly selected for assessing abnormal noise in rotating systems.

For the evaluation, both a microphone and multiple accelerometers were used to capture noise and vibration signals. The accelerometers were attached to the motor center, the motor lower area near the defective bearing, and the worm shaft. The reason for including accelerometer measurements was to ensure the applicability of the model not only during development but also in the production stage, where microphone-based measurements are often not feasible in End-of-Line environments.

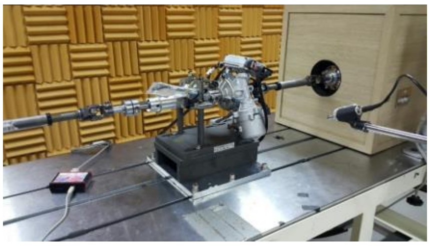


Figure 3. Evaluation setup for operational noise

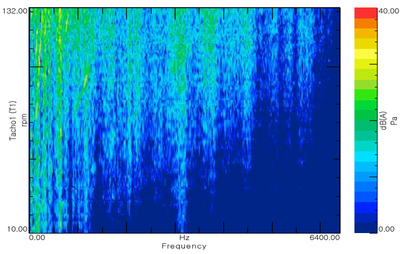


Figure 4. Colormap of the STFT Spectrogram

2.2 Description and analysis of noise

The bearing located at the upper end of the worm shaft, which is in contact with the motor, is subjected to a high load. Therefore, even minor defects can cause noise issues, which are generally classified into friction noise and rotational noise. Friction noise typically occurs when foreign substances such as aluminum particles or grease enter the bearing, whereas rotational noise is generated when there is damage to the raceway surface. Friction noise is characterized by broad-band frequency components that are independent of rotational speed. In contrast, rotational noise exhibits prominent order components that vary with speed.

In general, noise regulation is based on problematic order components and overall level, and the operational noise of C-EPS systems is also quantified using these two criteria. As shown in Table 1, the detection performance based on the overall level indicates that the microphone sensor failed to capture more than 50% of the bearing noise cases. The accelerometer attached to the motor lower section detected rotational noise with an accuracy of 61%, but was less effective for frictional noise, detecting only around 34%. Notably, the accelerometer placed at the motor center detected less than 15%, which is likely due to strong reflections of motor-specific characteristics. Theses results suggest that proper sensor placement is critical for effectively detecting bearing noise. Additionally, detection rates based solely on order components were extremely low and therefore not included in the table.

Table 1. Detection performance according to Noise type based on overall RMS thresholding (O/A)

	detection rate [%]	
Sensor type and Location	Rotation	Friction
MIC O/A	54.3	33.1
Motor Center Acc O/A	20.4	9.2
Motor Lower Acc O/A	67.7	44.2
Worm Shaft Acc O/A	59.1	39.9

2.3 Data preprocessing

Although the STFT is commonly used for preprocessing bearing operational noise, various preprocessing techniques including Mel spectrogram, Mel Frequency Cepstral Coefficients (MFCC), and Continuous Wavelet Transform (CWT) were applied in this study to prepare the data as input for machine learning models.

Mel Spectrograms and MFCC are widely used feature representations in audio signal processing, particularly in machine learning-based approaches. Mel spectrograms provide a time-frequency representation based on Mel scale, which reflects the non-linear perception of pitch.[4] MFCC, in contrast, captures perceptually relevant characteristics by modelling the human auditory system and is effective for extracting timbral features.[5] Both techniques have been extensively applied in speech recognition, acoustic scene analysis, and environmental noise classification, serving as standard input features for various learning models.[6] As shown in Figure 5, the average colormaps of MFCC and Mel spectrogram for each noise type indicate that MFCC fails to reflect the timbral differences of bearing noise, and the Mel spectrogram similarly does not sufficiently highlight distinctive frequency characteristics compared to STFT. This is likely because both preprocessing methods apply scaling in the high-frequency range, resulting in dimensionality reduction that diminishes the high-frequency characteristics of bearing noise. Therefore, it was concluded that these methods offer no clear advantage over STFT in preserving the high-frequency characteristics of bearing noise.

And CWT is effective for preprocessing impulsive or transient signals due to its ability to capture localized time-frequency characteristics with high resolution. This makes it a suitable choice for fault detection tasks where short-duration, high-frequency components are critical.[7] As shown in the CWT colormaps for each noise type in Figure 5, the characteristic order components of rotational noise appear too be suppressed by the CWT representation. Considering these characteristics of CWT, it was concluded that STFT is more effective in capturing the features of both rotational and frictional noise.

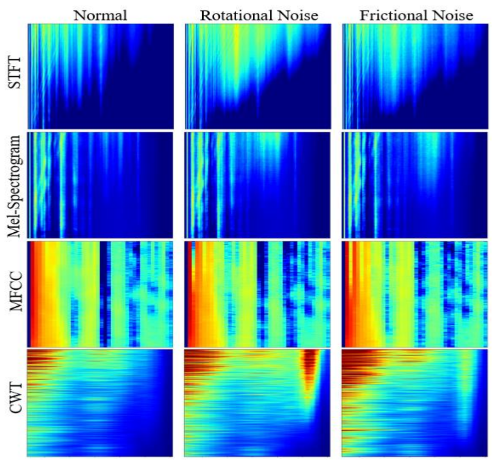


Figure 5. Average colormaps of each preprocessing method according to bearing noise defect types

3. UNSUPERVISED LEARNING

3.1 Unsupervised Learning

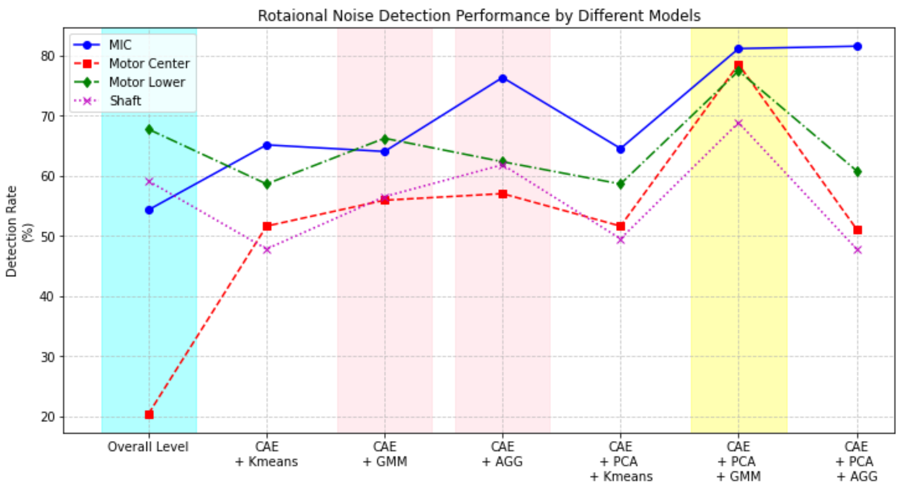
To embed the input data preprocessed using STFT through unsupervised learning, several models were employed, including Convolutional Autoencoder (CAE), Cluster GANs, Variational Autoencoder (VAE), ResNet-18 and LSTM-CNN Autoencoder. Since the primary objective of this study is clustering rather than generation, and the rotational noise signals exhibit minimal temporal variation, the experimental results suggest that the CAE is the most suitable architecture for unsupervised mapping of operational noise in C-EPS bearing.[8]

3.2 Clustering

For clustering the latent features extracted by the convolutional autoencoder (CAE), we utilized Gaussian Mixture Models (GMM), K-means, and Agglomerative Clustering (AGG). Among the clustering methods applied to the CAE-extracted embeddings, as shown Fig.6, the GMM and AGG achieved notably better separation of frictional and rotational anomalies compared to K-means. The improved performance of GMM and AGG can be attributed to their flexibility in modeling non-spherical clusters, robustness to boundary ambiguities, and ability to capture complex structure in high-dimensional acoustic feature spaces.[9] Furthermore, it was observed that applying Principal Component Analysis (PCA) to the latent representations extracted by the CAE further improved the detection performance when combined with GMM clustering.[10] These results are considered to stem from PCA's ability to reduce noise and redundancy in the CAE-derived embeddings, thereby improving the separability of clusters in the GMM process. In conclusion, the results in Table 2 demonstrate that GMM clustering, when applied after sequential embedding via CAE and PCA, substantially improves the detection rate compared to the baseline approach based on overall level. The calculated value represents the proportion assigned to the two clusters classified as defective products through clustering, rather than as normal ones. Although the detection performance of unsupervised learning does not surpass that of supervised approaches, it offers greater potential in capturing previously unseen frictional or rotational noise. This is particularly important in real-world applications, where ambiguous acoustic anomalies often lack clear boundaries, and the initial labeling process is prone to subjectivity. Therefore, the level of detection performance achieved in this study is considered highly meaningful and practically valuable. Accordingly, in situations where no alternative indicators beyond overall level or order level are available, the proposed model can be recommended for detecting bearing-related noise.

Table 2. Detection performance by noise type based on overall RMS thresholding (O/A) versus CAE-PCA-GMM (C.P.M)

		Detection rate [%]	
Sensor	Method	Rotation	Friction
MIC	O/A	54.3	33.1
	C.P.M	81.1	76.1
M. Center (Acc)	O/A	20.4	9.2
	C.P.M	78.5	54.6
M. Lower (Acc)	O/A	67.7	44.2
	C.P.M	77.4	63.8
Worm Shaft (Acc)	O/A	59.1	39.9
	C.P.M	68.8	52.8



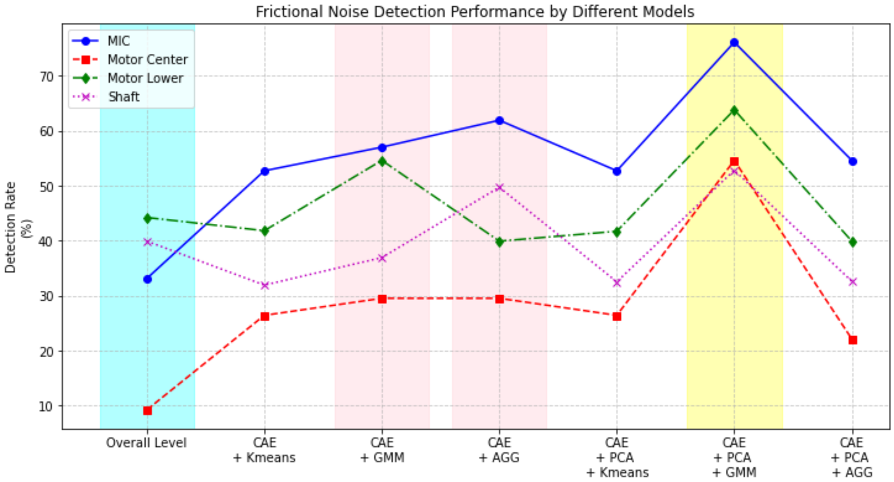


Figure 6. Detection rates for rotational and frictional noises using unsupervised learning with different clustering

3.3 End-to-End model for feature extraction

In addition to performing anomaly detection using unsupervised learning and clustering, the label sets obtained through these unsupervised approaches can further be utilized to enable supervised learning and XAI. This allows for the extraction of class-specific feature maps corresponding to different types of bearing noise. The overall framework can thus be the end-to-end model for extracting discriminative feature maps of bearing noise. As shown in Fig. 7, the process begins with STFT-based time frequency transformation, followed by dimensionality reduction through CAE and PCA. Unsupervised learning with GMM clustering yields label sets, which are then utilized to train a supervised classification model. Finally, XAI methods are applied to interpret the learned representations and extract feature maps associated with each noise type. Initially, both the ResNet18 and ResNet18 with Self-Attention models were applied for supervised learning. [11-12] Due to the lack of significant differences between the two, a classification model was developed using ResNet18. Subsequently, XAI tools such as Smooth Grad, Gradient-weighted Class Activation Mapping (Grad CAM), Score-Weighted Visual Explanations for Convolutional Neural Networks (Score CAM), and Local Interpretable Model-agnostic Explanations (LIME) were utilized to obtain the average heatmap of feature maps for each noise type. [13-15]

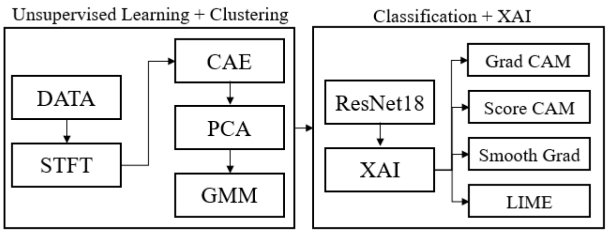


Figure 7. Diagram of end-to-end model

As shown Fig.8, the heatmaps generated by each XAI tool revealed that rotational noise primarily identifies the diagonal components of the order elements as key factors, while frictional noise is characterized by broad bands regardless of speed. In conclusion, an end-to-end process is established where labeling is performed through unsupervised learning and clustering. Subsequently, a classification model is developed based on the labeled data using unsupervised learning. Finally, XAI tools are employed to automatically extract features according to the characteristics of each noise type. If additional noise types are anticipated, unsupervised learning can be employed to add new classes. These feature maps can then be utilized as weighted filters for the specifications of noise and vibration. We believe that recalculating the noise using such feature maps and managing it through thresholding is significantly more effective than conventional overall-level-based approaches.

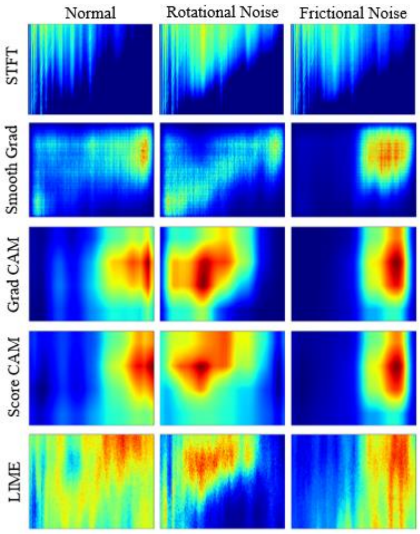


Figure 8. Average heatmaps of each XAI method according to bearing noise defect types

4. CONCLUSIONS

This study demonstrates the effectiveness of anomaly detection in diagnosing bearing defects in C-EPS systems by optimizing unsupervised learning and clustering methodologies. To preprocess the input data, STFT was employed to retain the high-frequency characteristics of bearing noise.

Unsupervised learning was conducted using CAE, with PCA applied to improve performance. The embedded features were clustered using GMM, resulting in superior detection capabilities compared to the traditional overall level metric for noise and vibration. Although both noise and vibration were analyzed, vibration was prioritized due to limitations of noise measurement in EOL environments.

Additionally, a classification model was developed using supervised learning based on noise-type labels derived from optimized unsupervised learning and clustering. This model integrates XAI to extract feature maps for each noise type, enabling identification of the speed and frequency at which anomalies occur. While the end-to-end model does not guarantee perfect detection, it offers a fast and practical auxiliary tool for EOL and early-stage development.

5. ACKNOWLEDGMENTS

This research was supported by Hyundai Mobis.

6. REFERENCES

[1] Y. Kozaki, G. Sekiya, and Y. Miyaura: “Electric power steering (EPS)”, *Motion & Control*, *vol.6*, pp. 9-15, 1999.

[2] Tanaka, Toshinori: "Motors for electric power steering," *Motor Technologies for Industry and Daily Life Edition*, 2003.

[3] Freund, Mary Kate, and Oumar Rafiou Barry: "Noise and vibration suppression of electric power steering systems: a survey" *International Journal of Vehicle Noise and Vibration*, 15.1, pp.68-88, 2019.

[4] Hong, G. and Suh, D.: “Mel Spectrogram-based advanced deep temporal clustering model with unsupervised data for fault diagnosis” *Expert Systems with Applications*, *217*, p.119551, 2023.

[5] Gupta, S., Jaafar, J., Ahmad, W.W. and Bansal, A.: “Feature extraction using MFCC” *Signal & Image Processing: An International Journal*, *4*(4), pp.101-108, 2013.

[6] Rawat, P., Bajaj, M., Vats, S. and Sharma, V.: “A comprehensive study based on MFCC and spectrogram for audio classification” *Journal of Information and Optimization Sciences*, *44*(6), pp.1057-1074, 2023.

[7] Rioul, O. and Duhamel, P.: “Fast algorithms for discrete and continuous wavelet transforms” *IEEE transactions on information theory*, *38*(2), pp.569-586, 1992.

[8] Guo, Xifeng, Xinwang Liu, En Zhu, and Jianping Yin, “Deep clustering with convolutional autoencoders." In *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14-18, 2017, Proceedings, Part II 24*, pp. 373-382. Springer International Publishing, 2017.

[9] Reynolds, Douglas A.: “Gaussian mixture models”. *Encyclopedia of biometrics*, *741*(659-663), p.3, 2009.

[10] Abdi, Herve and Lynne J., Williams: “Principal component analysis” *Wiley interdisciplinary reviews: computational statistics*, *2*(4), pp.433-459, 2010.

[11] Targ, S., Almeida, D. and Lyman, K., “Resnet in resnet: Generalizing residual architectures” *arXiv preprint arXiv:1603.08029*, 2016.

[12] Stephen, A., Punitha, A. and Chandrasekar, A.: “Designing self-attention-based ResNet architecture

for rice leaf disease classification" *Neural Computing and Applications*, *35*(9), pp.6737-6751, 2023.

[13] Panati, Chandana, Simon Wagner, and Stefan Brüggenwirth, "Feature relevance evaluation using grad-CAM, LIME and SHAP for deep learning SAR data classification." In *2022 23rd International Radar Symposium (IRS)*, pp. 457-462. IEEE, 2022.

[14] Wang, Haofan, Zifan Wang, Mengnan Du, Fan Yang, Zijian Zhang, Sirui Ding, Piotr Mardziel, and Xia Hu, "Score-CAM: Score-weighted visual explanations for convolutional neural networks." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp. 24-25. 2020.

[15] Smilkov, Daniel, Nikhil Thorat, Been Kim, Fernanda Viégas, and Martin Wattenberg, "Smoothgrad: removing noise by adding noise." *arXiv preprint arXiv:1706.03825*, 2017.