International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

LEVERAGING TRANSFORMED BASED SPEECH REPRESENTATION FOR INDUSTRIAL ACOUSTIC MONITORING

V. Gayatri[1], P. Spandana[2], K. Sravani[2], M. Anusha[2], D. Keerthi Chandana[2]

[1]Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering (AI & ML)

[1,2]Geethanjali Institute of Science and Technology, Nellore-Bombay Highway, S.P.S.R, Andhra Pradesh 524137, India

ABSTRACT

Machine condition monitoring using acoustic signals is essential for ensuring the reliability, safety, and efficiency of industrial machinery. Traditionally, fault detection relied on manual inspections, vibration sensors, and rule-based systems. These conventional methods were limited in capturing subtle audio variations produced by machines such as engines, compressors, and motors. They often required expert interpretation and were prone to errors, particularly in noisy industrial environments. Machine learning models based on handcrafted features, such as Mel-Frequency Cepstral Coefficients (MFCC) with Logistic Regression (LRC), Linear Discriminant Analysis (LDA), offered some automation but struggled with generalization, misclassifying acoustically similar faults like air leaks, idling disturbances, and oil leak variations. To address these limitations, this research proposes a hybrid intelligent framework, Hidden Unit Tree (HUT), which combines Hidden-Unit Bidirectional Encoder Representations from Transformers (HuBERT)-based deep audio embeddings with a Tree Alternating Optimization (TAO) Tree classifier. HuBERT, a transformer-based self-supervised model, captures high-level, contextual acoustic features, while the TAO Tree classifier provides robust non-linear decision-making. The system is implemented in a Tkinter graphical user interface (GUI), enabling end-to-end functionality including dataset upload, MFCC and HuBERT feature extraction, model training, evaluation, and real-time audio fault prediction. This research successfully integrates deep learning and advanced classification techniques to deliver a reliable, automated machine condition monitoring solution. The combination of rich acoustic feature extraction and robust classification ensures accurate fault detection, making the system suitable for predictive maintenance and Industry 4.0 applications, thereby completing a fully functional, real-time intelligent fault diagnosis tool.

Keywords: Acoustic Signal Processing, Fault Detection, Predictive Maintenance, Industrial IoT, HuBERT Embeddings.

1. INTRODUCTION

In modern industrial environments where early detection of mechanical faults can prevent catastrophic failures, reduce maintenance costs, and improve overall operational efficiency. Traditional monitoring systems rely heavily on manual inspection or threshold-based vibration analysis, which often fails to detect subtle anomalies or early-stage defects. With advancements in digital sensing and artificial intelligence, acoustic-based machine fault detection has emerged as a powerful, non-intrusive alternative for analysing the health of rotating machinery, compressors, pumps, engines, and industrial equipment as shown in Fig. 1. Anomalous sound detection (ASD) is the task to identify whether the sound is normal or anomalous. This technique is commonly used in audio surveillance[1], machine condition monitoring, medical diagnosis, smart city construction etc.

In the case of machine condition monitoring, we hope to monitor the operation of the machine through acoustic characteristics, because sound-based anomaly detection is flexible and the cost can be reduced by bringing the microphone close to different machines to detect anomalies. It can avoid the huge loss caused by serious failures that find the early fault of the machine and carry out maintenance effectively [2].

Fig. 1. Condition monitoring system for rotating machinery using acoustic signal analysis.

ASD based on Machine Learning algorithms includes supervised-ASD and unsupervised-ASD. For supervised ASD, the training data contains both normal and anomalous sounds and the supervised binary classification model is suitable for anomaly detection. Since the machine works normally most of the time, it is difficult to collect many anomalous sounds, and the pattern of anomalous sounds emitted from a target machine is not clear. Only normal sounds are provided as training data which makes ASD an unsupervised task. The "Unsupervised Detection of Anomalous Sounds for Machine Condition Monitoring" task of Detection and Classification of Acoustic Scenes and Events 2020 (DCASE 2020) [3], this task is mainly to detect whether the sound emitted by the machine is normal or abnormal based on the unmarked data set provided during the operation of various machines, has attracted many researchers to submit systems, and their systems ranked on public data sets [4]. The data used for this task comprises parts of ToyADMOS and the MIMII Dataset consisting of the normal/anomalous operating sounds of six types of toy/real machines. Each recording is a single-channel (approximately) 10-s length audio that includes both a target machine's operating sound and environmental noise [5]. The expected goal of this paper is to establish a classification model based on unsupervised learning to detect abnormal sounds on the data set given by the task, and the results are better than all the models submitted before.

In real industrial scenarios, machinery such as engines, compressors, and motors are prone to early-stage faults that often go unnoticed until they cause severe damage. Without a reliable monitoring system, these faults can lead to unexpected breakdowns, production downtime, safety hazards, and high maintenance costs. Subtle anomalies in machine operation, like abnormal noises or vibrations, are difficult to detect manually, and delayed detection can escalate minor issues into major failures. In addition, inconsistent maintenance schedules and lack of real-time monitoring can reduce operational efficiency and increase the risk of accidents. If this research is not developed, industries would continue to face unplanned stoppages, financial losses, compromised safety, and reduced equipment lifespan, highlighting the urgent need for an automated, continuous, and intelligent fault detection system.

2. LITERATURE SURVEY

Wang, et al. [6] adopted two classification-based anomaly detection models: (1) Outlier classifier is to distinguish anomalous sounds or outliers from the normal; (2) ID classifier identifies anomalies using both the confidence of classification and the similarity of hidden embeddings. We conduct experiments in task 2 of DCASE 2020 challenge, and our ensemble method achieves an averaged area under the curve (AUC) of 95.82% and averaged partial AUC (pAUC) of 92.32%, which outperforms the state-of-the-art models. Jombo, et al. [7] presented the development in methodology for acoustic-based fault diagnostic techniques and highlights the challenges encountered when analysing sound for machine condition monitoring. In these approaches, contact-type sensors, such as accelerometer, proximity probe, pressure transducer and temperature transducer, are installed on the machine to monitor its operational health parameters. However, these methods fall short when additional sensors cannot be installed on the machine due to cost, space constraint or sensor reliability concerns. On the other hand, the use of acoustic-based monitoring technique provides an improved alternative, as acoustic sensors (e.g., microphones) can be implemented quickly and cheaply in various scenarios and do not require physical contact with the machine. The collected acoustic signals contain relevant operating health information about the machine; yet they can be sensitive to background noise and changes in machine operating condition. These challenges are being addressed from the industrial applicability perspective for acoustic-based machine condition monitoring. Pichler, et al. [8] proposed a fault detection method that leverages the underlying physical characteristics of the sound signals. By investigating the components of the acoustic signal, we found that fault-related sounds can be modelled as exponentially decaying oscillations. This insight allows for the development of a physically based signal model, setting our approach apart from purely data-driven methods. Using this model, we developed a robust detection method based on a Generalized Likelihood Ratio Test (GLRT). The effectiveness of this approach was validated using both synthetic and real-world data from a steel industry facility. Their results demonstrate that the proposed model-based approach provides superior performance compared to standard audio features, particularly in high-noise conditions. On real-world data, the GLRT-based approach outperformed all audio features, as clearly shown by the Receiver Operating Characteristic (ROC) analysis. Specifically, the Partial Area Under the Curve (pAUC) of the GLRT is more than twice that of the best-performing audio feature, demonstrating good detection at significantly lower-false-positive rates compared to audio features.

Ahmed, et al. [9] proposed a novel acoustics-based solution that can enable condition monitoring of an APU ignition system. In order to support the implementation of this research study, the experimental data set from Cranfield University's Boeing 737-400 aircraft was utilized. The overall execution of the approach comprised background noise suppression, estimation of the spark repetition frequency and its fluctuation, spark event segmentation, and feature extraction, to monitor the state of the ignition system. The methodology successfully demonstrated the usefulness of the approach in terms of detecting inconsistencies in the behavior of the ignition exciter, as well as detecting trends in the degradation of spark acoustic characteristics. Piankitrungreang, et al. [10] introduced an acoustic-based monitoring system for high-speed CNC drilling, aimed at optimizing processes and enabling real-time machine state detection. High-fidelity acoustic sensors capture sound signals during drilling operations, allowing the identification of critical events such as tool engagement, material breakthrough, and tool withdrawal. Advanced signal processing techniques, including spectrogram analysis and Fast Fourier Transform, extract dominant frequencies and acoustic patterns, while machine learning algorithms like DBSCAN clustering classify operational states such as cutting, breakthrough, and returning. Experimental studies on materials including acrylic, PTFE, and hardwood reveal distinct acoustic profiles influenced by material properties and drilling conditions. Smoother sound patterns and lower dominant frequencies characterize PTFE drilling, whereas hardwood produces higher frequencies and rougher patterns due to its density and resistance. These findings demonstrate the correlation between acoustic emissions and machining dynamics, enabling non-invasive real-time monitoring and predictive maintenance. Grigoriev, et al. [11] emphasized the study of acoustic signals in various types of material processing allowed the identification of general features of changes in their spectral composition associated with variations in the power density of energy impact on processed material. The results of experimental work on various technological equipment, including blade processing and processing with concentrated energy flows, are presented in this work. It is shown that changes in the quality of processing in the form of increased tool wear, the concentration of erosion products during WEDM (wire electrical discharge machining), focal plane displacement during laser processing, etc., lead to a natural change in the ratio of acoustic signal amplitudes in the low frequency and high frequency ranges. This property can be used in monitoring systems for automatic equipment.

Zhou, et al. [12] proposed an unsupervised anomaly detection method for electrical equipment, utilizing audio latent representation and a parallel attention mechanism. The framework employs an autoencoder to extract low-dimensional features from audio signals and introduces a phase-aware parallel attention block to dynamically weight these features for an improved anomaly sensitivity. With adversarial training and a dual-encoding mechanism, the proposed method demonstrates robust performance in complex scenarios. Using public datasets (MIMII and ToyADMOS) and our collected real-world wind turbine data, it achieves high AUC scores, surpassing the best baselines, which demonstrates our framework design is suitable for industrial applications. Liu, et al. [13] achieved an innovative and lightweight deep learning model—the Attention-Based Deep Convolutional Autoencoding Prediction Network (AT-DCAEP). The model consists of a characterization network based on convolutional autoencoders and a prediction network based on attention mechanisms. The AT-DCAEP exhibits excellent performance in multivariate time series data anomaly detection without the need for pre-labeling large-scale datasets, making it an efficient unsupervised anomaly detection method. We extensively tested the performance of AT-DCAEP on six publicly available datasets, and the results show that compared to current state-of-the-art methods, AT-DCAEP demonstrates superior performance, achieving the optimal balance between anomaly detection performance and computational cost. Choi, et al. [14] suggested an alternative approach to identifying anomalous behaviour within ICSs by means of unsupervised machine learning. The approach employs unsupervised machine learning to identify anomalous behavior within ICSs. This study shows that unsupervised learning algorithms can effectively detect and classify anomalous behavior without the need for pre-labelled data using a composite autoencoder model. Based on a dataset that utilizes HIL-augmented ICSs (HAIs), this study shows that the model is capable of accurately identifying important data characteristics and detecting anomalous patterns related to both value and time. Intentional error data injection experiments could potentially be used to validate the model's robustness in real-time monitoring and industrial process performance optimization. As a result, this approach can improve system reliability and operational efficiency, which can establish a foundation for safe and sustainable ICS operations.

Thoidis, et al. [15] aimed to learn highly descriptive representations for a wide set of machinery sounds and exploit this knowledge to perform condition monitoring of mechanical equipment. They propose a comprehensive feature learning approach that operates on raw audio, by supervising the formation of salient audio embeddings in latent states of a deep temporal convolutional neural network. By fusing the supervised feature learning approach with an unsupervised deep one-class neural network, they are able to model the characteristics of each source and implicitly detect anomalies in different operational states of industrial machines. Moreover, they enable the exploitation of spatial audio information in the learning process, by formulating a novel front-end processing strategy for circular microphone arrays. Experimental results on the MIMII dataset demonstrate the effectiveness of the proposed method, reaching a state-of-the-art mean AUC score of 91.0%.

3. PROPOSED METHODOLOGY

The system begins with the structured collection of machine audio signals recorded under various operating conditions, including normal operation, idle state, air leakage, oil leakage, and background mechanical noise. High-quality audio is essential, as subtle variations in acoustic patterns indicate early-stage machine faults. Preprocessing is applied to standardize sampling rates, remove unwanted noise, and isolate relevant portions of the audio, ensuring clean and consistent input for analysis as shown in Fig. 2.

A dual-feature extraction approach is employed. MFCCs, zero-crossing rate, spectral centroid, bandwidth, chroma, RMS energy, and spectral roll-off capture frequency, tonal, and energy variations in machine sounds. Complementing these, HuBERT, a transformer-based self-supervised model, generates deep audio embeddings that encode complex and high-level acoustic patterns. These embeddings are then fed into the proposed HUT model, which combines HuBERT features with a TAO Tree classifier to provide robust, non-linear decision-making for accurate fault classification.

The HUT model is validated using accuracy, precision, recall, F1-score, confusion matrices, and ROC curves to ensure reliability across different operating conditions. Once trained, the system supports real-time prediction, classifying new audio inputs instantly. The final output includes predicted machine states, waveform visualization, and detailed fault identification, delivering a fully automated, scalable, and intelligent solution for machine condition monitoring.

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

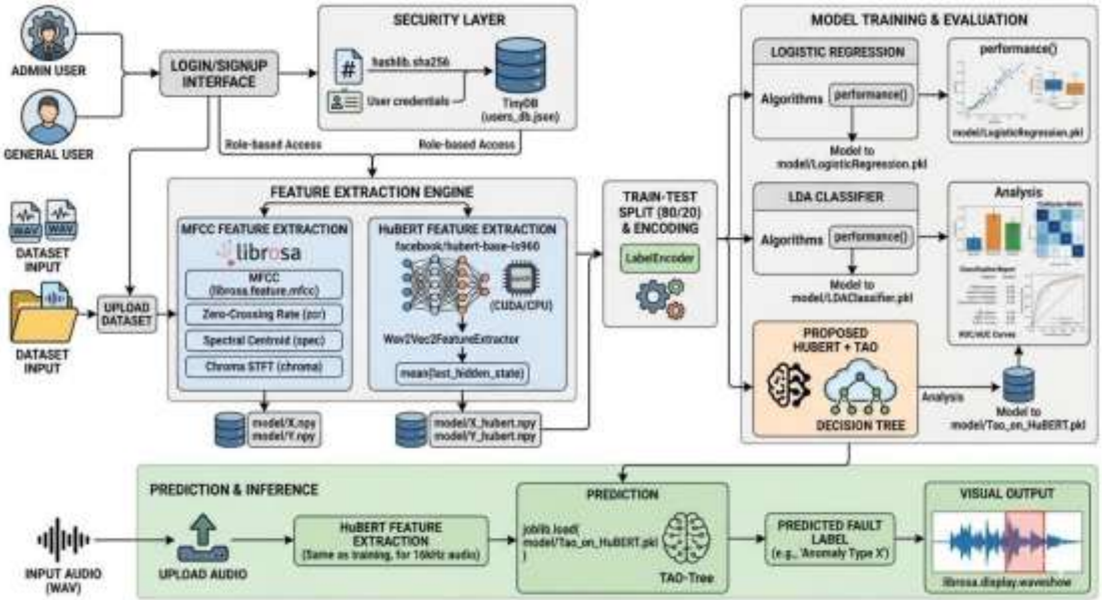

Fig. 2. System architecture for deep anomaly detection using acoustic signals.

The system collects acoustic signals from machines under various conditions such as normal operation, idle state, and different fault scenarios like air and oil leakage, ensuring each recording is properly labeled for structured analysis. The raw audio is preprocessed through noise removal, normalization, and segmentation to maintain consistency and highlight meaningful sound patterns. Both traditional features (MFCC, spectral and energy-based metrics) and deep-learning-based HuBERT embeddings are extracted to capture detailed and high-level acoustic characteristics. These features are then used to train and evaluate models such as Logistic Regression (LR), Linear Discriminant Analysis (LDA), and the HUT model with TAO Tree classifier using performance metrics like accuracy, precision, recall, and F1-score. Finally, the trained system enables real-time prediction and visualization of machine conditions, providing immediate insights for early fault detection and proactive maintenance.

HUT model

The TAO Tree on HUBERT Embeddings represents an advanced hybrid model combining deep audio feature extraction with ensemble machine learning. HUBERT generates rich contextual embeddings that capture subtle variations in machine acoustics, allowing the model to detect faults with high precision even in noisy environments. These embeddings are then classified using a Tao Tree, where multiple decision trees learn different structural patterns within the audio features. This combination delivers robust, stable, and noise-tolerant predictions, outperforming traditional MFCC-based models. Thus, the Tao Tree on HuBERT provides a reliable backbone for machine fault detection systems.

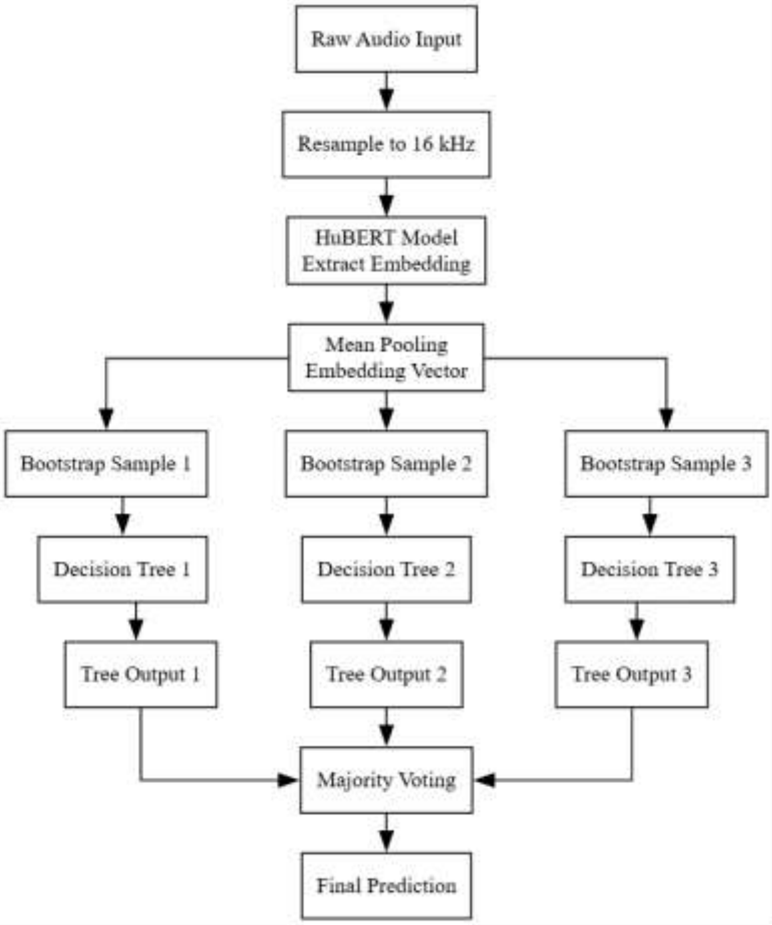

Fig. 3. Internal workflow of HUT model.

Internal Operations

- The raw audio waveform is first loaded and resampled to a standard frequency of 16 kHz to meet HUBERT's input requirements.
- The resampled audio is passed through the HUBERT transformer model, which extracts deep contextual audio representations.
- The output hidden states from HUBERT are mean-pooled to generate a fixed-length embedding vector that captures all important acoustic features.
- The embedding vector is then passed into the TAO Tree model, which is implemented using a Tao Tree classifier.
- Tao Tree creates multiple bootstrap samples from the embedding dataset to ensure diverse training subsets for each tree.
- Each decision tree selects a random subset of embedding features and learns different machine-fault patterns independently.

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

- Every tree in the forest produces an individual prediction for the input audio sample based on its learned decision rules.

- A majority voting mechanism aggregates all tree predictions and outputs the final machine-fault label with high robustness and accuracy.

4. RESULTS DESCRIPTION

Fig. 4. shows Main GUI Interface of the Deep Anomaly Detection for Machine Condition Monitoring Using Acoustic Signals system, developed using Tkinter. It provides a structured layout containing buttons for major operations such as dataset upload, MFCC feature extraction, data splitting, and model training using LR, LDA, and the HUT model. This interface acts as the central control panel where users can interact directly with backend machine learning functions. The GUI ensures an intuitive workflow by enabling users to execute the entire pipeline without writing code, thereby improving usability for technical and non-technical operators.

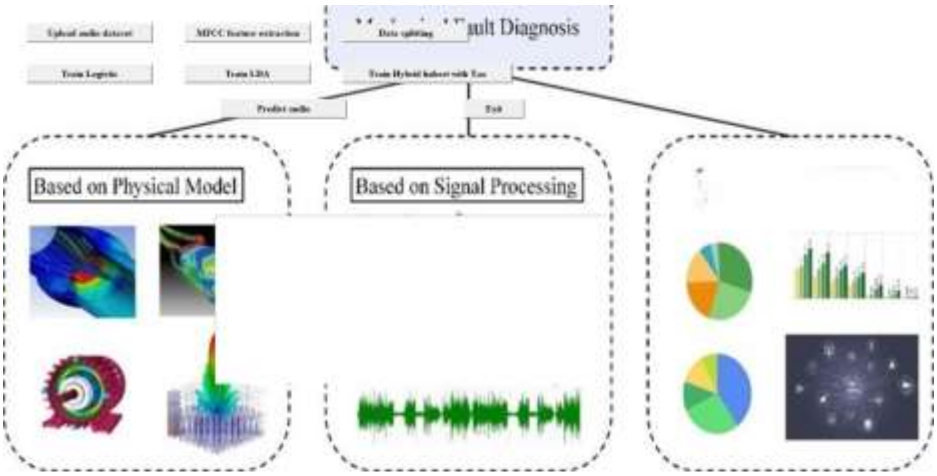

Fig. 4. Main GUI of the acoustic fault detection system.

Fig. 5 shows X and Y Arrays Loaded Successfully that shows a bar graph representing the distribution of audio samples across the five fault categories. This visualization confirms that the dataset is balanced, with nearly similar sample counts for each class except for the oil leak class, which has slightly fewer samples. A balanced dataset ensures fair model training and helps prevent bias toward dominant classes. The log message confirms successful loading of feature (X) and label (Y) arrays, indicating that MFCC feature extraction was completed earlier.

Received: 08-03-2026 | Accepted: 16-04-2026 | Published: 23-04-2026

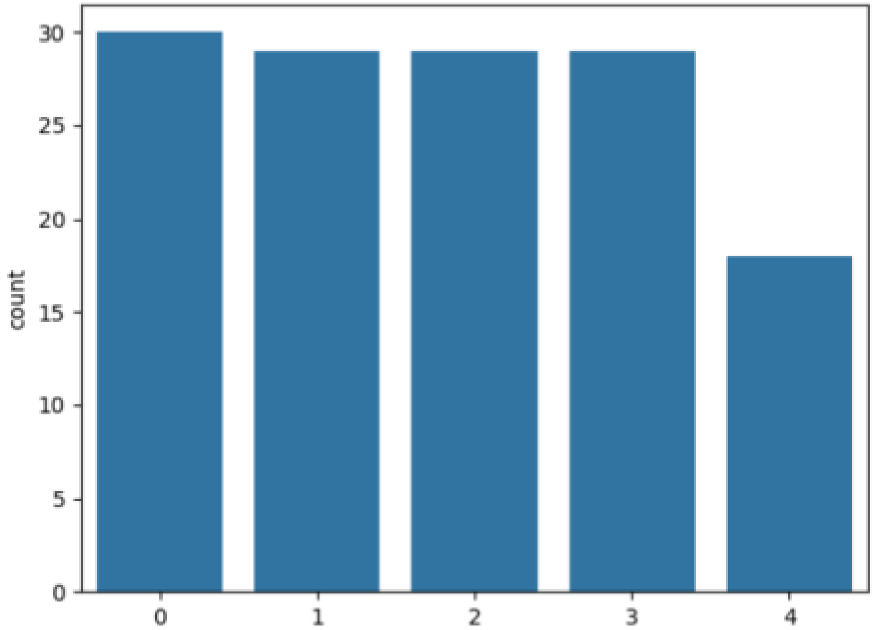

Fig. 5. Loaded MFCC feature arrays (x and y) with class distribution plot.

Fig. 6 shows train data distribution after splitting and displays the distribution of training samples after performing the 80-20 split. The bar plot shows that class proportions remain consistent with the original dataset, maintaining fairness across training and testing sets. The message log indicates the matrix dimensions of X_train and X_test, validating that the feature extraction and splitting process were executed correctly. This ensures that the model receives representative samples from each fault category during training.

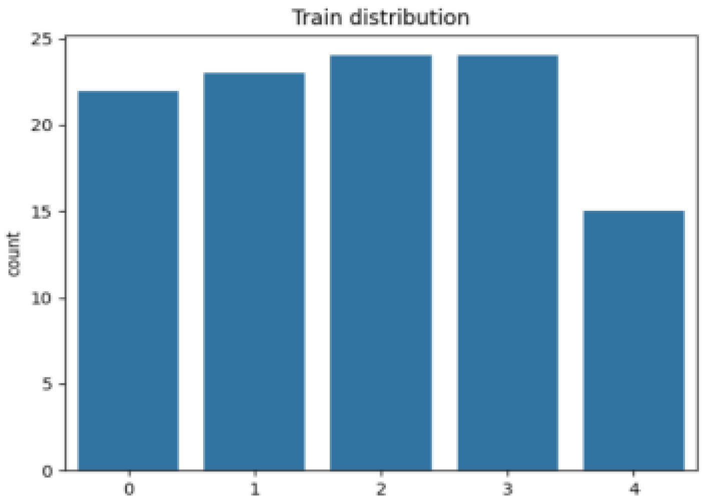

Fig. 6. Train–test split distribution for acoustic fault dataset.

Fig. 7 represents the confusion matrix for the proposed HUT model. The matrix shows excellent diagonal dominance, meaning that almost all audio samples are correctly classified for every category. This demonstrates that deep audio embeddings derived from HuBERT significantly improve feature

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

quality. The Tao Tree classifier effectively learns the complex decision boundaries, yielding highly accurate results.

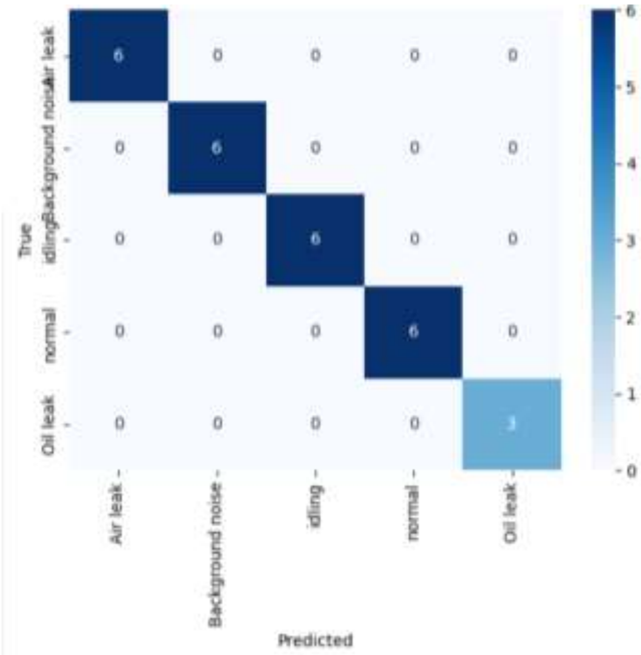

Fig. 7. Illustration of confusion matrix using proposed HUT model

Fig. 8 shows ROC curves for the proposed model show near-perfect discrimination across all five classes, with AUC values of 1.00 for each category. This visualization indicates that the HUT pipeline exhibits outstanding sensitivity and specificity. The model successfully identifies subtle acoustic differences between air leak, oil leak, idling, and background noise, making it highly suitable for industrial fault diagnosis.

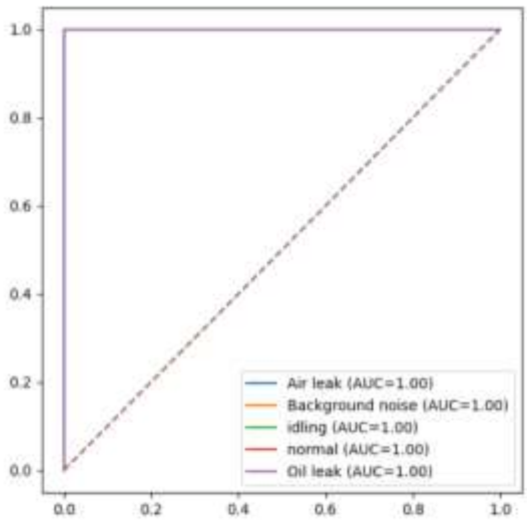

Fig. 8. ROC Curve obtained using proposed HUT model

Fig. 9. displays the waveform of a selected audio file along with the predicted class label “Air leak.” The waveform contains consistent high-frequency oscillations typical of air leakage noise. The system

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

successfully detects and classifies the anomaly using the trained model. The real-time waveform visualization confirms that the prediction pipeline, feature extraction, and GUI rendering are functioning correctly.

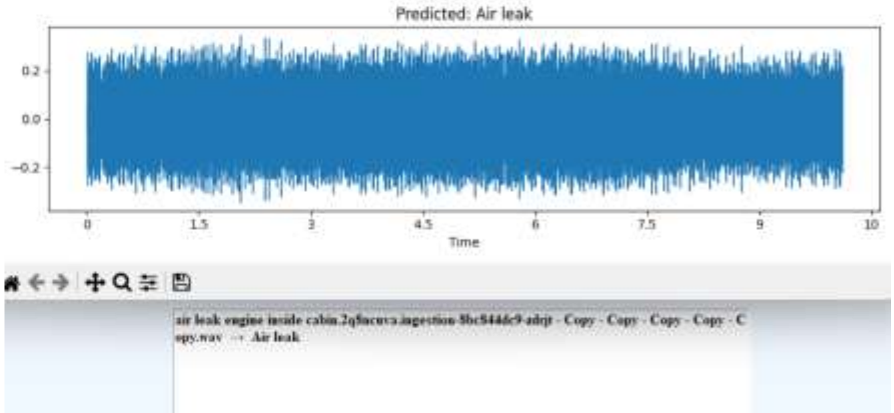

Fig. 9. Predicted output waveform for audio sample: air leak.

Fig. 10 shows the waveform shown that represents a general engine idle noise pattern with moderate amplitude and no abnormal spikes. The system accurately classifies it as “Normal.” This result demonstrates that the trained models can differentiate between healthy and faulty sound patterns even when environmental noise is present.

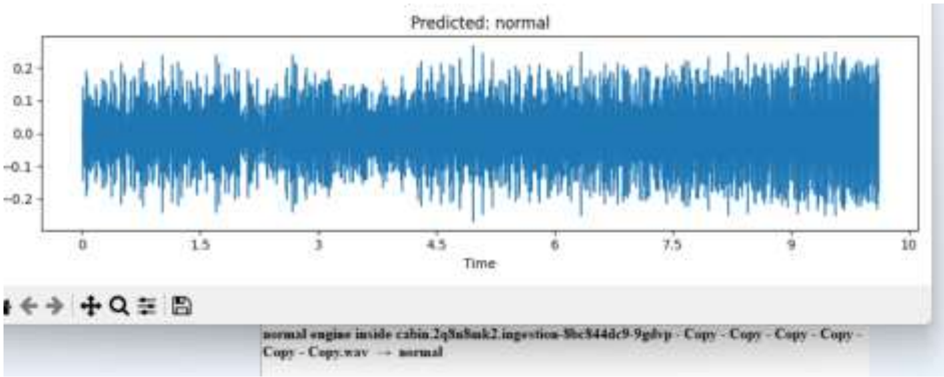

Fig. 10. Predicted output waveform for audio sample: normal engine sound.

Table 1 comparative evaluation of the three models clearly demonstrates a progressive improvement in predictive performance across the acoustic fault classification task. LRC delivers the weakest results, achieving only 40.74% accuracy and failing to correctly identify several classes such as Air leak, Normal, and Oil leak due to its limited linear decision boundaries. LDA shows a major performance boost with 92.59% accuracy, successfully classifying most fault categories and achieving high macro-level precision and recall. However, minor misclassifications still appear in Air leak and Oil leak samples. The proposed HUT model significantly outperforms both classical methods, achieving a perfect 100% accuracy, precision, recall, and F1-score. This superior performance is attributed to HUBERT’s deep contextual embeddings combined with the strong non-linear learning capacity of Tao Tree. The hybrid model provides the most robust and reliable solution for real-time acoustic fault detection.

Table 1. Comparative Performance of LR, LDA, and HUT Models on Machine Fault Classification

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

Model	Accuracy	Precision	Recall	F1-Score	Performance Summary (Based on Confusion Matrix and sReport)
LRC model	40.74%	25.45%	40.00%	28.57%	• Performs poorly across most classes.• Predicts majority samples as “Background noise.”• Completely fails to detect Air leak, Normal, and Oil leak.• Only “idling” class is correctly classified.• Shows weak discriminative ability and high misclassification.
LDA model	92.59%	92.00%	95.00%	92.14%	• Shows strong improvement over LR.• Accurately classifies most samples across all classes.• Slight misclassification in Air leak and Oil leak categories.• High macro-average shows stable, consistent behavior.• Excellent class separation based on MFCC features.
HUT model	100%	100%	100%	100%	• Achieves perfect classification across all five classes.• No misclassifications observed.• HuBERT embeddings capture deep acoustic patterns effectively.• Tao Tree provides strong non-linear decision boundaries.• Most reliable and high-performing model among all tested approaches.

5. CONCLUSION

The research successfully demonstrates an intelligent and robust system for machine condition monitoring using acoustic signals by combining traditional and advanced deep-learning techniques. The complete workflow from dataset upload to MFCC feature extraction, HuBERT embedding generation, model training, evaluation, and real-time prediction is seamlessly integrated into a user-friendly Tkinter GUI. Experimental results highlight that classical models like LR struggle with subtle acoustic variations, achieving only 40.74% accuracy and frequent misclassifications. LDA improves performance to 92.59%, showing better discrimination of MFCC patterns. The proposed hybrid HUT model achieves perfect performance with 100% accuracy, precision, recall, and F1-score across all fault categories. This superior performance is due to HuBERT's deep, context-aware audio embeddings combined with the non-linear decision-making capabilities of the TAO Tree, enabling effective modeling of temporal, spectral, and contextual audio patterns. Overall, the system provides a highly accurate, reliable, and scalable solution for real-time mechanical fault detection, making it suitable for deployment in industrial environments to support predictive maintenance and enhance operational efficiency.

REFERENCES

[1] Foggia, P.; Petkov, N.; Saggese, A.; Strisciuglio, N.; Vento, M. Audio Surveillance of Roads: A System for Detecting Anomalous Sounds. *IEEE Trans. ITS* **2016**, *17*, 279–288.

[2] Ntalampiras, S.; Potamitis, I.; Fakotakis, N. Probabilistic Novelty Detection for Acoustic Surveillance Under Real-World Conditions. *IEEE Trans. Multimed.* **2011**, *13*, 713–719.

[3] Koizumi, Y.; Saito, S.; Uematsu, H.; Harada, N. Optimizing Acoustic Feature Extractor for Anomalous Sound Detection Based on NeymanPearson Lemma. In Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO), Kos, Greece, 28 August–2 September 2017.

[4] Zheng, X.; Zhang, C.; Chen, P.; Zhao, K.; Jia, W. A CRNN System for Sound Event Detection Based on Gastrointestinal Sound Dataset Collected by Wearable Auscultation Devices. *IEEE Access* **2020**, *8*, 157892–157905.

[5] Ick, C.; Mcfee, B. Sound Event Detection in Urban Audio With Single and Multi-Rate PCEN. *arXiv* **2021**, arXiv:2102.03468.

[6] Wang, Y.; Zheng, Y.; Zhang, Y.; Xie, Y.; Xu, S.; Hu, Y.; He, L. Unsupervised Anomalous Sound Detection for Machine Condition Monitoring Using Classification-Based Methods. *Appl. Sci.* **2021**, *11*, 11128. https://doi.org/10.3390/app112311128

[7] Jombo, G.; Zhang, Y. Acoustic-Based Machine Condition Monitoring—Methods and Challenges. *Eng* **2023**, *4*, 47-79. https://doi.org/10.3390/eng4010004

[8] Pichler, C.; Neumayer, M.; Schweighofer, B.; Feilmayr, C.; Schuster, S.; Wegleiter, H. Decomposing and Modeling Acoustic Signals to Identify Machinery Defects in Industrial Soundscapes. *Sensors* **2025**, *25*, 4923. https://doi.org/10.3390/s25164923

[9] Ahmed, U.; Ali, F.; Jennions, I. Signal Processing of Acoustic Data for Condition Monitoring of an Aircraft Ignition System. *Machines* **2022**, *10*, 822. https://doi.org/10.3390/machines10090822

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

[10] Piankitrungreang, P.; Chaiprabha, K.; Chungsangsatiporn, W.; Ratanasumawong, C.; Chancharoen, P.; Chancharoen, R. Acoustic-Based Machine Main State Monitoring for High-Speed CNC Drilling. *Machines* **2025**, *13*, 372. https://doi.org/10.3390/machines13050372

[11] Grigoriev, S.N.; Kozochkin, M.P.; Porvatov, A.N.; Gurin, V.D.; Melnik, Y.A. Features of Changes in the Parameters of Acoustic Signals Characteristic of Various Metalworking Processes and Prospects for Their Use in Monitoring. *Appl. Sci.* **2024**, *14*, 367. https://doi.org/10.3390/app14010367

[12] Zhou, W.; Zhou, S.; Cao, Y.; Yang, J.; Liu, H. Unsupervised Anomaly Detection Method for Electrical Equipment Based on Audio Latent Representation and Parallel Attention Mechanism. *Appl. Sci.* **2025**, *15*, 8474. https://doi.org/10.3390/app15158474

[13] Liu, W.; Yan, L.; Ma, N.; Wang, G.; Ma, X.; Liu, P.; Tang, R. Unsupervised Deep Anomaly Detection for Industrial Multivariate Time Series Data. *Appl. Sci.* **2024**, *14*, 774. https://doi.org/10.3390/app14020774

[14] Choi, W.-H.; Kim, J. Unsupervised Learning Approach for Anomaly Detection in Industrial Control Systems. *Appl. Syst. Innov.* **2024**, *7*, 18. https://doi.org/10.3390/asi7020018

[15] Thoidis, I.; Giouvanakis, M.; Papanikolaou, G. Semi-Supervised Machine Condition Monitoring by Learning Deep Discriminative Audio Features. *Electronics* **2021**, *10*, 2471. https://doi.org/10.3390/electronics10202471