
CIGRE Condition Monitoring, Diagnosis and Maintenance Conference
CMDM 2025 (8th edition)
October 28th–30th, 2025, Bucharest, Romania

**1011**

# Advanced Gas Turbine Condition Monitoring Using Acoustic Emission and Hybrid Deep Learning Approaches

**Babak Danesh**
**International Science and Technology University (ISTU)**
**babak.danesh@istu.edu.pl**
**Matin Hasanli**
**Middle East Technical University, Northern Cyprus Campus (METU NCC)**
**matin.hasanli_01@metu.edu.tr**
**Turkey**

## SUMMARY

This study develops and evaluates machine learning approaches for automated fault diagnosis in gas turbines using acoustic emission signals collected from a laboratory-scale P.9000 turbine under four conditions: unloaded/loaded normal operation and two/four-blade damage scenarios. Acoustic signals were transformed using continuous wavelet transform with Morlet wavelet and analyzed through traditional machine learning, custom CNN, hybrid models, and fine-tuned pretrained architectures. Results demonstrate exceptional performance of Random Forest (1.000±0.000 accuracy across most binary classifications), CNN-SVM (99.54% accuracy for four-blade damage), and ResNet50 (90.09% validation accuracy, 0.9846 ROC AUC). Despite some misclassification between loaded and two-blade damage conditions, this research establishes acoustic emission monitoring with advanced classification techniques as a highly effective non-invasive approach for gas turbine fault diagnosis, with significant commercialization potential in the growing turbine monitoring market.

## KEYWORDS

Acoustic emission, Gas turbine, Fault diagnosis, Deep learning, Transfer learning, Random For- est, Continuous wavelet transform

## 1. INTRODUCTION

Gas turbines play a vital role in power generation and aviation industries due to their efficiency and reliability. These systems comprise three essential components: a

compressor, combustion chamber, and turbine, where rotating blades extract energy from the hot gas flow. Blade faults represent a major cause of turbine failures, often resulting in costly operational interruptions and safety concerns [18]. Operating under harsh conditions including high temperatures, intense mechanical stress, and corrosive environments, turbine blades are susceptible to fatigue, creep, erosion, corrosion, and foreign object damage [11].

Conventional inspection techniques, including visual assessments and vibration-based analysis, often lack sensitivity for detecting initial defects. In contrast, acoustic emission (AE) monitoring has emerged as an effective real-time approach for detecting micro-cracks, fatigue, and impact-related damage in turbine blades [13]. Recent developments in deep learning, particularly convolutional neural networks (CNNs), have significantly improved AE signal analysis precision, facilitating early fault detection and classification [7].

Recent research has demonstrated promising results in AE-based turbine fault detection. Nashed et al. achieved 82.6% accuracy using ResNet50 for wear and rubbing detection [9], while Zhang et al. demonstrated over 98% accuracy by combining sound-pressure signals with LSTM networks [17]. Deep learning approaches have shown particular effectiveness, with CNNs achieving up to 96% accuracy in gas-turbine diagnostics when integrated with physics-driven monitoring systems [5].

This study evaluates various approaches for gas-turbine fault diagnosis using acoustic emission signals, including custom CNN architectures, hybrid CNN–traditional-classifier models, and fine-tuned deep learning models (ResNet50, VGG16, Xception) applied to time–frequency representations. The contributions include applying continuous wavelet transform (CWT) for time–frequency analysis, integrating CNNs with traditional classifiers, and a comparative evaluation of deep and hybrid models for enhanced early-stage fault detection in industrial applications.

## 2. METHODOLOGY

### 2.1. Data collection

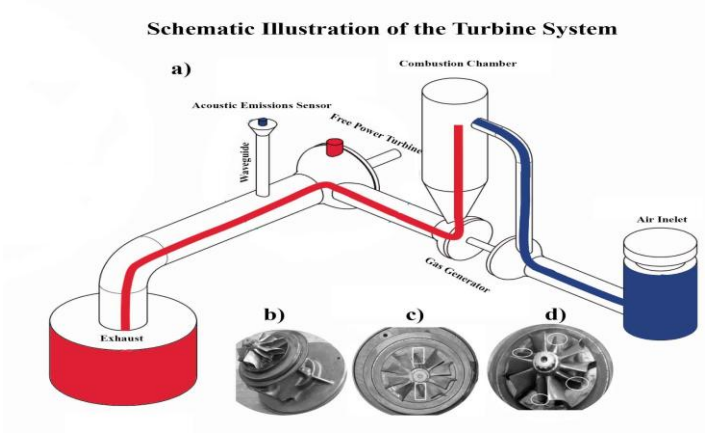


**Figure 2.1.** Turbine System and Acoustic Emissions Capture: (a) Turbine System, (b) Free Power Turbine, (c) Two Damaged Blades, (d) Four Damaged Blades

The dataset comprises acoustic-emission (AE) signals collected from a laboratory-scale gas turbine system under varying operational conditions [9]. The setup included a gas generator and a free power turbine operating at 170–600 RPS, generating up to 4 kW of power. AE signals were recorded using a Micro-80D sensor at a 5 MHz sampling rate with 12-bit resolution over 0.03 s windows. Four operational conditions were evaluated: normal operation without load (683 recordings), normal operation with load (545 recordings), two-blade damage D2 (621 recordings), and four-blade damage D4 (729 recordings), totaling 2,578 AE recordings. Fig- ure 2.1 provides a schematic overview of the turbine system and illustrates various levels of blade damage, while Figure 2.2 shows representative AE signals that reflect distinct temporal patterns under each condition.

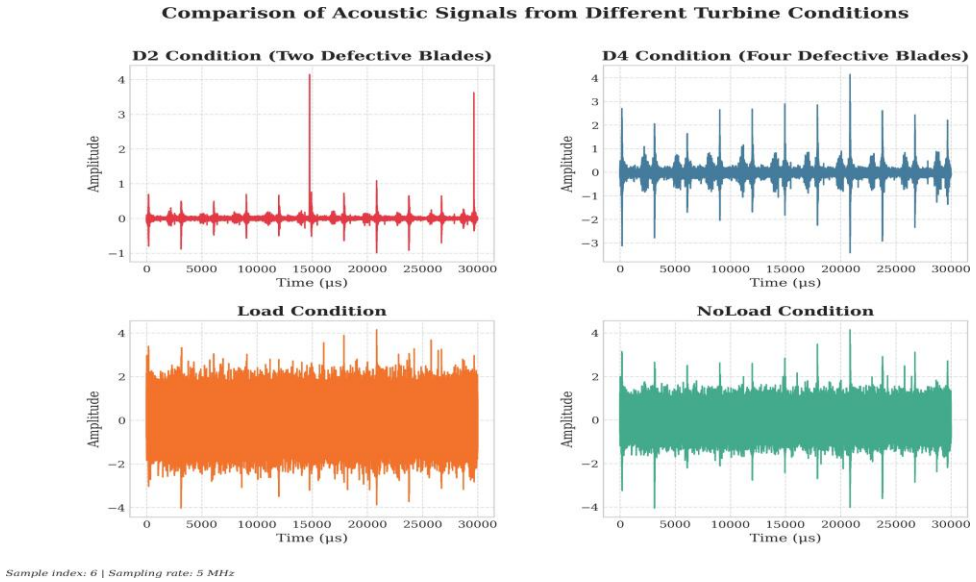


**Figure 2.2.** Time-domain acoustic signal examples for four different gas turbine conditions

## 2.2. Preprocessing of acoustic emission signals

*Low-Pass Filtering*

A fourth-order Butterworth low-pass filter with a 1 MHz cutoff frequency was applied to eliminate high-frequency noise [15]:

$$H(f) = \frac{1}{1 + \left(\frac{f}{f_c}\right)}, \tag{2.1}$$

where $f_c$ denotes the cutoff frequency and $n$ the filter order.

*Statistical and Nonlinear Feature Extraction*

Statistical and nonlinear features were extracted from time-domain, frequency-domain, and nonlinear-dynamical analyses. From an initial set of 31 features, 15 key discriminative features were selected based on ANOVA $F$-values, including mean, spectral bandwidth, skewness, energy, wavelet energy (levels 1–5), crest factor, energy entropy, kurtosis, dominant frequency, Shannon entropy, and peak amplitude.

*Continuous Wavelet Transform (CWT)*

Time–frequency analysis was performed using CWT with the Morlet wavelet [8]:

$$C(a, b) = \int_{-\infty}^{\infty} x(t)\, \psi^*_{a,b}(t)\, dt, \tag{2.2}$$

CWT coefficients were transformed into 224 × 224-pixel images and normalized for deeplearning compatibility. Figure 2.4 demonstrates the preprocessing pipeline, showing original acoustic signals, filtered signals, CWT representations, and resized images prepared for neural- network input across different turbine conditions.

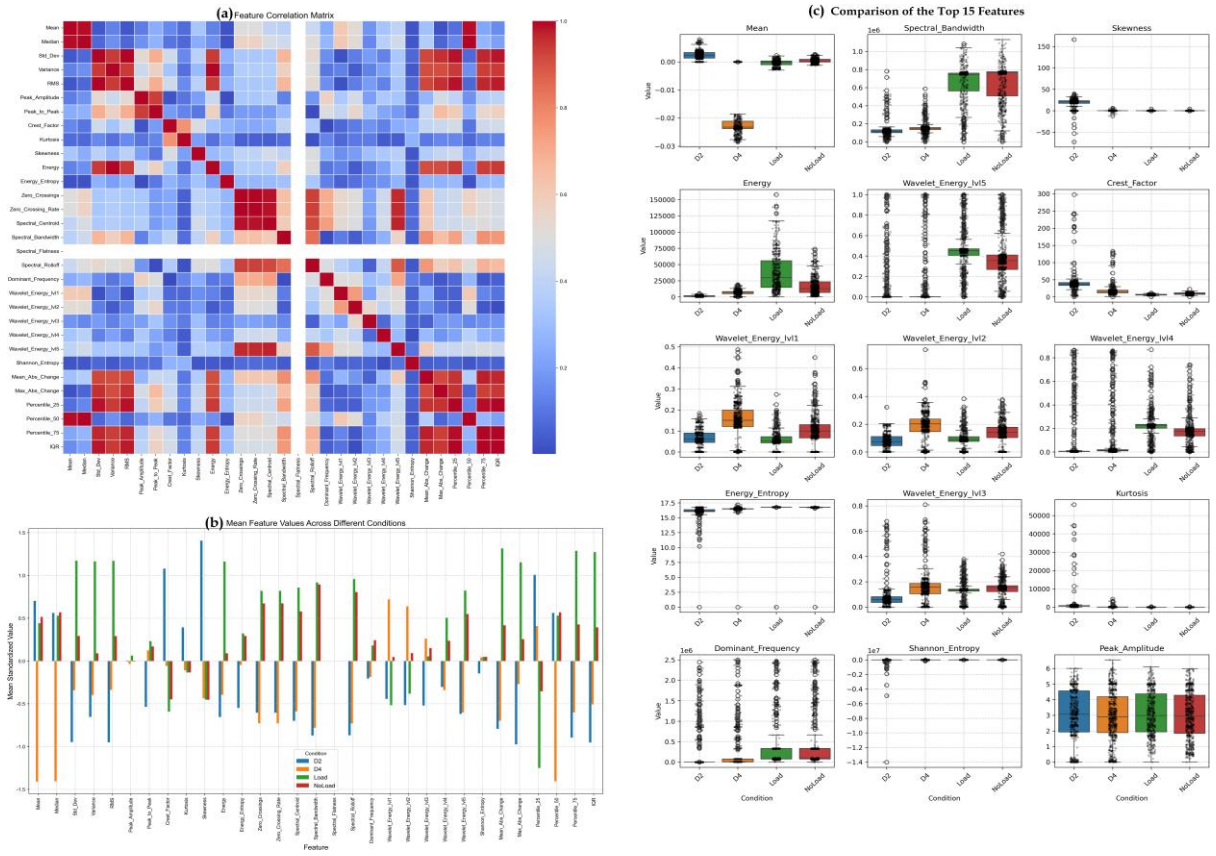


**Figure 2.3.** Visualization of extracted features: correlation analysis, mean feature values, and distribution of the top 15 discriminative features. Subfigure (a) shows the feature correlation matrix, highlighting relationships among all 31 features. Subfigure (b) depicts the mean feature values across four turbine conditions. Subfigure (c) compares the distributions of the top 15 features selected based on their discriminative power and statistical significance (ANOVA F-value). These visualizations provide insights into feature relevance and their potential for effective classification.

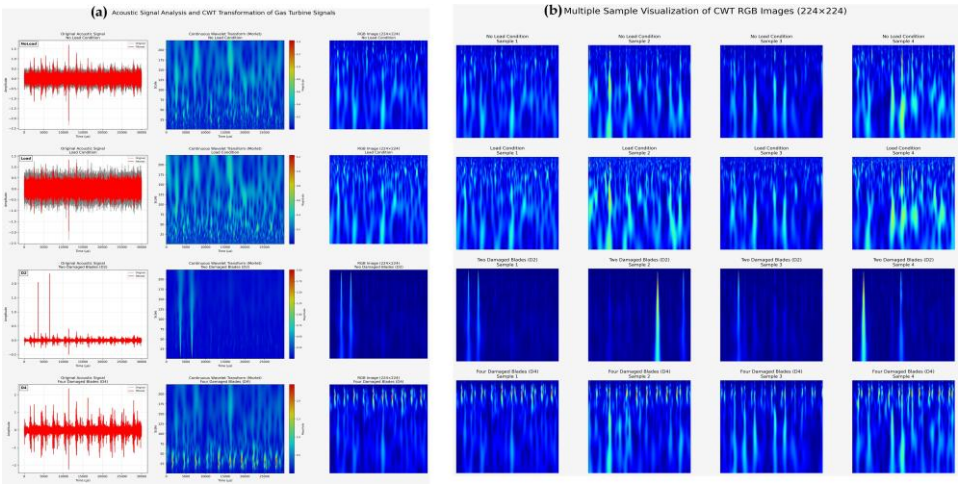


**Figure 2.4.** Preprocessing and visualization of gas turbine acoustic signals using continuous wavelet transform (CWT). Panel (a) shows the original acoustic signals, low-pass filtered signals, their continuous wavelet transform (CWT) representations, and resized CWT images (224×224 pixels) prepared for deep learning input. Panel (b) presents multiple sample CWT RGB images, illustrating differences between healthy and damaged turbine states under various operational conditions (No Load, Load, D2, and D4). These steps standardize and enhance the acoustic data, making it suitable for effective classification.

Model development

*DNN Structure*

A binary-classification DNN was implemented with three hidden layers (128, 64, and 32 neurons; ReLU activations) and a softmax output layer. Dropout and batch normalization were applied to improve generalization, and training used the Adam optimizer with sparse categorical cross-entropy (parameters in Table 2.1).

*Independent CNN Model*

A CNN was designed for classifying turbine states (no-load, load, two-blade damage, four-blade damage) from AE-based spectrograms. The architecture consisted of Conv2D and MaxPooling layers with ReLU, dropout, and dense layers, with softmax for multi-class output. Training used Adam, categorical cross-entropy, early stopping, and 10% validation.

*Hybrid CNN + Classical Models*

CNN-extracted features were integrated with SVM, RF, and DT to enhance classification. The CNN design matched the standalone model, with flattened features passed to each classifier.

- **SVM:** RBF kernel, hyperparameters tuned for optimal separation [3].
- **RF:** Ensemble of 100 decision trees to improve robustness [1].
- **DT:** Hierarchical partitioning of features for classification.

**Pre-trained Models (VGG16, ResNet50, Xception).** Three transfer learning models were fine- tuned using AE spectrograms (replicated to three channels). All models classified four operational states.

- **VGG16:** 16-layer CNN with uniform 3 × 3 filters and ReLU activations [14].
- **ResNet50:** 50-layer CNN with residual blocks to address vanishing gradients [6].
- **Xception:** Depthwise separable convolutions for efficient feature extraction [2].

Training settings and optimization parameters are summarized in Table 2.1. The proposed methodology is summarized in Figure 2.5.

**Table 2.1.** Summary of model architectures, learning parameters, and training configurations.

| Model | Architecture details | Learning parameters | Activation |
|---|---|---|---|
| ResNet50 | Input: (224,224,3); ImageNet weights; GlobalAveragePooling2D; Dense: 256; Frozen: all except last block | Adam; LR 1e-4; Binary cross-entropy; Early stopping patience 10; Batch 2; 50 epochs; 10-fold CV | ReLU (hidden), Sigmoid (output) |
| Xception | Input: (224,224,3); ImageNet weights; GlobalAveragePooling2D; Dense: 256; Frozen: all except last block | Adam; LR 1e-4; Binary cross-entropy; Early stopping patience 10; Batch 2; 50 epochs; 10-fold CV | ReLU (hidden), Sigmoid (output) |
| VGG16 | Input: (224,224,3); ImageNet weights; GlobalAveragePooling2D; Dense: 256; Frozen: all except last block | Adam; LR 1e-4; Binary cross-entropy; Early stopping patience 10; Batch 2; 50 epochs; 10-fold CV | ReLU (hidden), Sigmoid (output) |
| CNN (indep.) | (224,224,3) → Conv2D/MaxPool/Dropout → Dense (256) → Dropout → Dense (4) | Adam; Binary cross-entropy; Batch 2; 50 epochs; BatchNorm on | ReLU (hidden), Sigmoid (output) |
| DNN (features) | Input: $N$ features; hidden [128, 64, 32]; output: 1 | Adam; Binary cross-entropy; Batch 32; 50 epochs; Dropout 0.5; L1=0.005, L2=0.001; 10-fold CV | ReLU (hidden), Sigmoid (output) |
| SVM | RBF kernel | Standardization; SMO; 5-fold CV; $C$ = 1.0; $\gamma$ = scale; prob. estimates on; 10-fold CV | – |
| DT | Gini impurity | 5-fold CV; max depth: unlimited; min samples split: 2; 10-fold CV | – |
| RF | 100 trees (bootstrapped) | 10-fold CV; Gini; max features: auto; standardization applied | – |

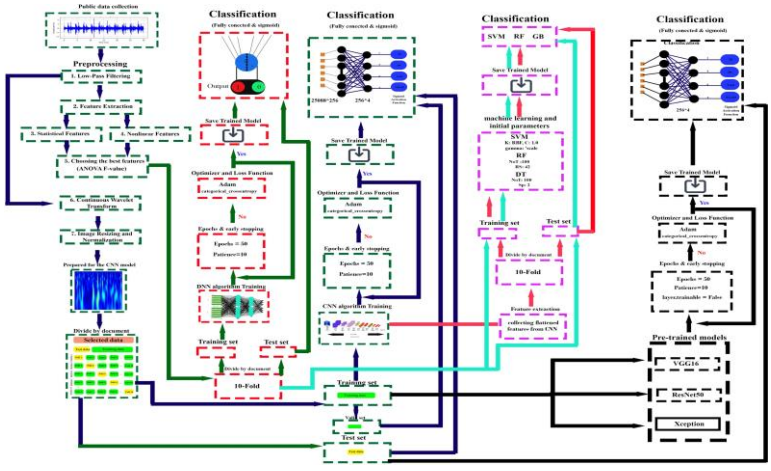


**Figure 2.5.** Proposed workflow for data processing and classification using DNN/CNN and pretrained models.

## 2.3. Model evaluation

*Performance Metrics*

Model evaluation employed standard metrics to address imbalanced and multi-class conditions, including accuracy, precision, recall, F1-score, ROC-AUC, and the confusion matrix [10] [16]. These metrics collectively provide a balanced view of predictive accuracy, class-wise sensitivity, and error distribution.

*Cross-Validation*

Robustness was assessed using k-fold cross-validation, where the dataset was partitioned into k folds, each serving once as validation. Final results were reported as the mean across folds, ensuring reliable performance estimation.

## 3. RESULTS AND DISCUSSION

This study analyzed 2,578 acoustic emission (AE) signals from a laboratory-scale gas turbine (Cussons P.9000) under four operational conditions: unloaded, loaded, and two fault states (D2, D4). Signals were captured with a Micro-80D sensor at 5 MHz, denoised using a 1 MHz low- pass filter, and transformed into spectrograms via CWT (Morlet wavelet). Two classification strategies were applied: (i) deep learning using fine-tuned ResNet50, Xception, VGG16, and custom CNN combined with SVM, DT, RF, and DNN classifiers, and (ii) traditional ML using statistical and nonlinear features. All models were evaluated using 10-fold cross-validation. Figure 3.1 summarizes results with the top 10 features (e.g., mean, spectral bandwidth, wavelet energy, entropy, kurtosis, dominant frequency).

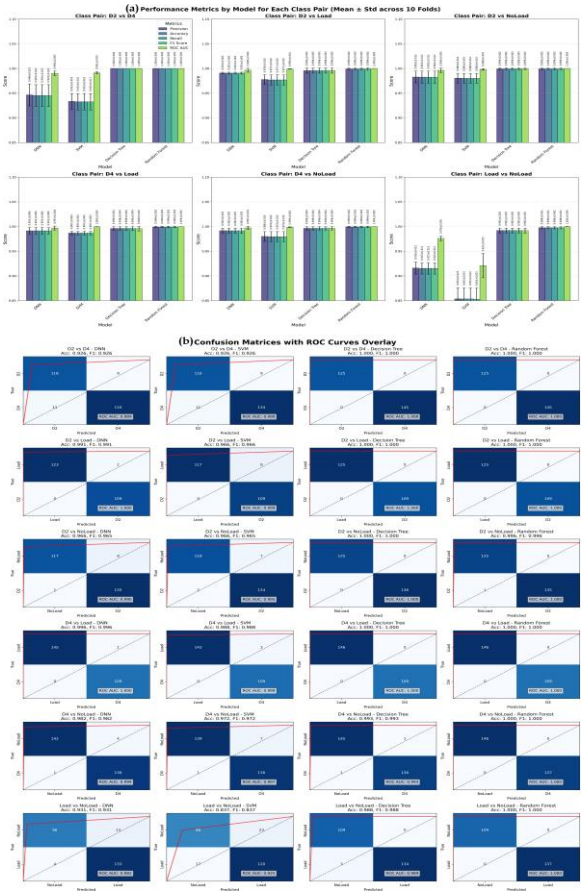


**Figure 3.1. –** Classification Results Based on the Top 10 Statistical and Nonlinear Features Extracted from Acoustic Emission Signals. (a) presents the performance metrics (Precision, Accuracy, Recall, F1 Score, and ROC AUC) calculated as the mean ± standard deviation across 10-fold cross-validation for each binary classification pair (D2 vs D4, D2 vs Load, D2 vs NoLoad, D4

vs Load, D4 vs NoLoad, Load vs NoLoad) using Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), and Deep Neural Network (DNN) models. (b) illustrates the confusion matrices overlaid with the corresponding ROC curves for each classification scenario.

Random Forest consistently achieved exceptional performance with near-perfect accuracy (≈1.0) across most binary classification tasks. Load vs. NoLoad classification proved comparatively challenging but remained highly accurate (0.997 ± 0.002 with RF). These results confirm that combining statistical and nonlinear features provides complementary information, with ensemble models effectively capturing subtle acoustic variations for reliable fault diagnosis.

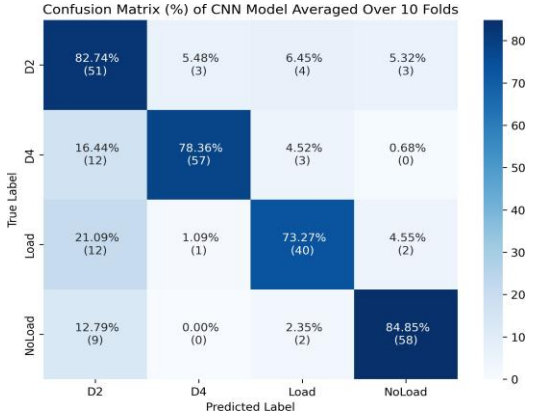


**Figure 3.2.** Confusion Matrix of the CNN Model Averaged Over 10 Folds for Gas Turbine Condition Classification

The classification outcomes presented in Figure 3.1 reveal the remarkable discriminative capability of combined statistical and nonlinear features in distinguishing between operational and fault states. The superior performance of Random Forest suggests ensemble learning effectively captures complex patterns in time-domain features such as kurtosis, entropy, and wavelet energy. From a mechanical perspective, blade damage states (D2, D4) introduce sharp, non- Gaussian signal artifacts due to structural impacts and vibrational irregularities, well-capturedby high-order statistical descriptors like skewness and kurtosis. Aerodynamically, loading conditions result in subtler acoustic profile modulations, reflecting smoother shifts in turbulent flow or combustion pressure, explaining the relatively lower classification performance (though still high, e.g., 0.997 ± 0.002 accuracy for Load vs NoLoad with RF). The exceptional performance of kurtosis and wavelet energy features suggests blade damage creates non-Gaussian signal artifacts characterized by intermittent high-amplitude events, differing significantly from stationary signals during normal operation. These findings align with previous studies demonstrating that structural defects produce unique acoustic signatures due to altered fluid-structure interactions. Figures 3.2 – 3.4 summarize the classification performance of different models for four turbine states (D2, D4, Load, NoLoad).

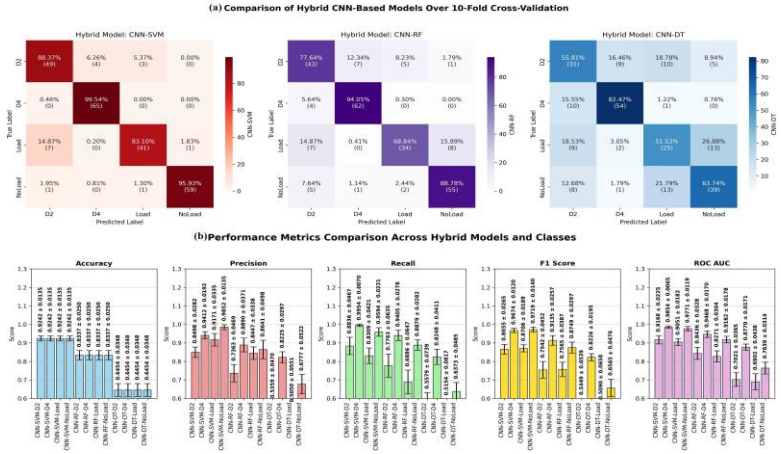


**Figure 3.3.** Performance Evaluation of Hybrid CNN-Based Models: (a) Confusion Matrices and (b) Performance Metrics for Gas Turbine Condition Classification.

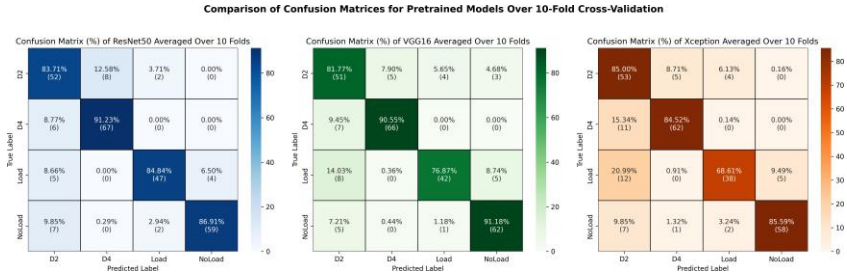


**Figure 3.4.** Confusion Matrices of Pretrained Models for Gas Turbine Condition Classification Averaged Over 10-Fold Cross-Validation

• **CNN:** The standalone CNN achieved good accuracy, especially for NoLoad (84.9%) and D2 (82.7%), but showed confusion between Load and D2, yielding an overall accuracy of 83.0% ± 0.01.

• **Hybrid CNN Models:** Combining CNN features with classical classifiers improved results. CNN-SVM outperformed others with 94.2% ± 1.4% accuracy and 98.5% ± 1.8% ROC AUC, while CNN-RF showed moderate performance and CNN-DT lagged behind, particularly under Load conditions.

• **Pre-trained Models:** Among ResNet50, VGG16, and Xception, ResNet50 achieved the best overall results (accuracy 86.2%, ROC AUC 0.98), particularly strong for D4 and NoLoad. VGG16 showed stable but slightly lower performance, while Xception struggled with Load classification and lower precision for D2.

Overall, integrating CNN with SVM and using ResNet50 as a pre-trained model provided the most reliable detection of turbine faults and operating conditions.

As shown in Figures 3.2-3.4, CNN models effectively learned distinctive AE patterns, with val- idation accuracy stabilizing at 87.50%. The hybrid models demonstrated varying classification effectiveness, with CNN-SVM achieving highest accuracy for severe damage (D4: 99.54%) and normal operation (NoLoad: 95.93%), while CNN-DT struggled with

overlapping conditions. Fine-tuned pre-trained models demonstrated that wavelet-transformed spectrograms effectively capture time-frequency characteristics, enabling differentiation of subtle mechanical variations. These results underscore the strong correlation between AE signal variations and underlying mechanical processes, where blade damage, combustion dynamics, and load-induced stress contribute to distinct signatures effectively leveraged by deep learning classification models for automated fault diagnosis.

Our results surpassed several recent studies: Random Forest achieved perfect classification compared to Zhang et al.'s 98% accuracy using LSTM networks, while our ResNet50 implementation exceeded previous benchmarks [17]. The CNN-SVM hybrid approach demonstrated effectiveness comparable to Reddy et al.'s 94.94% binary classification but distinguished between four distinct operational conditions [12]. However, unlike Bai et al., we did not address class-imbalance issues common in industrial settings [4].

Training times were evaluated using 10-fold cross-validation on an Intel Core i7 processor with NVIDIA RTX 3050 GPU. Models using CWT images required substantial computational resources (63-136 minutes), while hybrid CNN-classifier approaches reduced training time significantly (23-26 minutes).

Statistical feature-based models demonstrated the lowest comp

tational requirements, with Random Forest completing training in 6 seconds. This indicates a clear trade-off between computational complexity and efficiency, suggesting feature-based approaches offer practical solutions for resource-constrained applications without compromising performance.

SHAP (SHapley Additive exPlanations) analysis revealed that wavelet energy at decomposition levels 1-3 and spectral kurtosis were consistently among the top five most influential features across all fault types. Feature importance patterns differed between loading conditions and blade damage scenarios, with entropy-based features proving critical for distinguishing loading states while kurtosis-based features excelled at identifying blade damage, suggesting distinct physical mechanisms generate acoustic signatures for different operational states.

Several limitations should be considered: the dataset consists solely of single-fault scenarios while real turbines often encounter multiple faults, experimental conditions were relatively stable compared to varying industrial environments, and current models focus on classification without prognostic capabilities. Future research should explore hybrid physics-data driven models, transfer learning for real-world applications, and integration of multiple sensing modalities to improve diagnostic reliability through complementary information fusion.

## 4. CONCLUSIONS

This study demonstrates the effectiveness of machine-learning approaches for gas-turbine fault diagnosis using acoustic-emission (AE) signals. Random Forest classifiers

achieved perfect classification accuracy (1.000 ± 0.000) across most binary tasks using statistical features, while ResNet50 emerged as the superior deep-learning model with 90.09% validation accuracy. The hybrid CNN–SVM model exhibited remarkable performance with 99.54% accuracy for four- blade damage detection. These findings highlight the significant potential of AE monitoring combined with advanced classification techniques for early fault detection, offering a non- invasive, highly accurate diagnostic solution. Despite limitations – Including single-fault scenarios and stable operating conditions – this research establishes a robust benchmark for future studies. Future work should focus on compound-fault cases, dynamic operating conditions, and prognostic capabilities to enhance practical applicability in industrial settings.

Data availability – The dataset used in this study is publicly available at the GitHub repository.

## 5. REFERENCES


[1] Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. doi: http://dx.doi.org/10.1023/A:1010933404324

[2] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

[3] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297. doi:https://doi.org/10.1007/BF00994018

[4] Fahmi, A.-T. W. K., Reza Kashyzadeh, K., & Ghorbani, S. (2024). Advancements in Gas Turbine Fault Detection: A Machine Learning Approach Based on the Temporal Convolutional Network–Autoencoder Model. *Applied Sciences*, 14(11), 4551.

[5] Fentaye, A. D., Zaccaria, V., & Kyprianidis, K. (2021). Aircraft engine performance monitoring and diagnostics based on deep convolutional neural networks. *Machines*, 9(12), 337.

[6] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

[7] Kim, H. J., Lee, J. H., Lee, S. Y., Lee, H. H., & Lee, S. H. (2023). Acoustic emission reflection signal classification of PVDF-type AE sensor using convolutional neural network–transfer learning. *Journal of Intelligent Manufacturing*, 1–20.

[8] Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Academic Press.

[9] Nashed, M., Renno, J., Mohamed, M. S., & Reuben, R. L. (2023). Gas turbine failure classification using acoustic emissions with wavelet analysis and deep learning. *Expert Systems with Applications*, 232, 120684.

[10] Powers, D. M. (2020). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*. doi:https://doi.org/10.48550/arXiv.2010.16061

[11] Rao, N., Kumar, N., Prasad, B., Madhulata, N., & Gurajarapu, N. (2014). Failure mechanisms in turbine blades of a gas turbine engine—An overview. *Int. J. Eng. Res. Dev*, 10(8), 48–57.

[12] Reddy, A., Indragandhi, V., Ravi, L., & Subramaniyaswamy, V. (2019). Detection of cracks and damage in wind turbine blades using artificial intelligence-based image analytics. *Measurement*, 147, 106823.

[13] Shahkar, S., & Khorasani, K. (2019). Gas turbine condition monitoring using acoustic emission sig- nals. *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems*, 2(3), 031005.

[14] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. doi:https://doi.org/10.48550/arXiv.1409.1556
[15] Smith, S. W. (1997). *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Pub.
[16] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. doi:https://doi.org/10.1016/j.ipm.2009.03.002
[17] Zhang, J., Chen, Y., Li, N., Zhai, J., Han, Q., & Hou, Z. (2023). A weak fault identification method of micro-turbine blade based on sound pressure signal with LSTM networks. *Aerospace Science and Technology*, 136, 108226.
[18] Zhang, Z., Yang, G., & Hu, K. (2018). Prediction of fatigue crack growth in gas turbine engine blades using acoustic emission.

* . *