IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 8 August 2025, accepted 3 September 2025, date of publication 9 September 2025, date of current version 17 September 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3607751

RESEARCH ARTICLE

Hybrid Acoustic Fault Diagnosis in UAVs Using Wavelet Scattering Transform and Deep Learning

TEMEL SONMEZOCAK[1], (Member, IEEE), AND MERİH YILDIZ[2]
[1]Department of Electrical and Electronics Engineering, Yeni Yuzyil University, Istanbul 34010, Türkiye
[2]Department of Electrical and Electronics Engineering, Doğuş University, Istanbul 34775, Türkiye

Corresponding author: Temel Sonmezocak (temel.sonmezocak@yeniyuzyil.edu.tr)

This work was supported by Inci Automation Ltd.

ABSTRACT Unmanned aerial vehicles (UAVs) are increasingly employed in defense, agriculture, and logistics, where ensuring operational safety is critical. Propeller damages and rotor screw looseness represent two major fault types that can compromise flight reliability. Previous studies have typically addressed these faults separately, often relying on vibration analysis or conventional acoustic features. This paper introduces a hybrid model that combines Wavelet Scattering Transform (WST) and Long Short-Term Memory (LSTM) to simultaneously detect both fault types using microphone-recorded acoustic signals. Unlike traditional Fourier-based approaches or deep learning models utilizing Mel-Frequency Cepstral Coefficients (MFCCs) or spectrograms, the proposed framework leverages WST to extract deformation-stable, multi-resolution features, which are then modeled through an LSTM network. The model was trained and tested on a dataset comprising 750 one-second acoustic segments, approximately balanced between problem-free and faulty classes using stratified sampling and cross-validation. By integrating time–frequency-based multi-layered features into a time-sequential deep learning framework, the proposed model achieves reliable classification of both propeller faults and rotor screw looseness, reaching a high accuracy of 98.93%. These results highlight the potential of the WST-LSTM framework as a robust and innovative solution for UAV fault diagnosis, particularly in acoustic-based monitoring scenarios

INDEX TERMS Acoustical signal processing, deep learning, propellers, fault diagnosis, spectral analysis, unmanned aerial vehicles.

I. INTRODUCTION

Nowadays, unmanned aerial vehicles (UAVs) are extensively utilized across a wide range of applications, including agriculture, defense, mapping, and aerial photography [1], [2], [3]. However, ensuring the reliability and safety of these systems is of paramount importance. In particular, faults in propellers can lead to thrust loss, increased power consumption, and structural issues caused by airframe resonance. Furthermore, the loosening of rotor fastening components poses serious safety risks, such as the potential detachment of propellers during flight, which can ultimately result in UAV crashes. Such mechanical failures significantly threaten flight safety. In the literature, propeller faults in UAVs have mostly been investigated using vibration-based fault diagnosis techniques. Since these failures typically occur within the operational frequency range below 250 Hz, the interaction between the damaged propellers and the system's natural and operational frequencies can be effectively analyzed through vibration signals. In particular, the root mean square (RMS) values observed in each frequency band provide meaningful insights into fault conditions [4], [5].

To effectively interpret vibration data, MEMS sensors mounted on UAV airframes are typically used to capture oscillations along both horizontal and vertical axes.

The associate editor coordinating the review of this manuscript and approving it for publication was Mehrdad Saif.

While this approach introduces computational complexity, classification algorithms such as decision tree (DT) and support vector machine (SVM) have demonstrated high efficiency in terms of computational performance.

These models are particularly suitable for detecting low-frequency propeller faults operating in the 80–100 Hz range.

For instance, the use of 14 amplitude–frequency features with the DT algorithm has achieved a classification accuracy of 93.37%, whereas utilizing 64 features derived from broader frequency bands with the SVM algorithm has reached up to 98.21% accuracy [6]. Similarly, in our previous study, a cost-effective vibration-based fault diagnosis model for UAVs was developed using MEMS accelerometers and gyroscopes.

The proposed approach relied exclusively on vibration signals and employed DT & SVM based classification algorithms [7].

In recent years, deep learning models that utilize vibration sensor data across the UAV's X, Y, and Z axes have also gained popularity [8], [9]. Particularly for propeller faults, hybrid systems combining wavelet transform [10], [11] and deep neural networks using vibration signals within the operational band around 168 Hz have achieved classification accuracy of approximately 91% [9]. However, when detecting only propeller faults via vibration signals, frequency variations analyzed through wavelet scattering and long short-term memory (LSTM) networks have shown the potential to reach up to 99% accuracy [12].

In addition, the literature indicates that sub-200 Hz vibrations in UAVs can lead to structural deformations in the arms and frame, as well as loosened joints and weld defects [13], [14]. In this context, laboratory-based studies have shown that screw loosening conditions can be detected using Fuzzy Logic and Neural Network-based analyses [15]. However, under outdoor conditions, vibration amplitudes in UAV arms and frames vary with different speeds and maneuvers, which may lead to inaccurate diagnoses.

The field of UAV health monitoring is not limited to fault detection in propellers and structural arms but also encompasses overall system reliability and operational continuity. In [16], a weight-based landing platform was proposed to perform pre-flight safety inspections. This platform allows for the analysis of parameters such as load imbalance and deviations in the center of gravity, facilitating the identification of potential anomalies. However, the proposed approach is limited to pre-takeoff assessments. Even if the center of gravity remains balanced during flight, localized thrust losses caused by rotor or propeller faults can still affect the UAV's orientation. Such effects can be effectively detected using machine learning-based approaches [17].

In recent years, significant progress has been made in fault diagnosis studies on UAVs, particularly through the use of acoustic signals and conventional machine learning methods. In this context, algorithms such as DT, SVM and k-nearest neighbors (KNN) have been able to classify faults with high accuracy using statistical features such as mean, standard deviation, variance, correlation, kurtosis, and skewness. For example, classification accuracy exceeding 99% has been achieved on motors rated at 2200 KV (revolutions per volt). However, for motors with ratings of 1400 KV or 2700 KV, the accuracy drops to the range of 80–90% [18]. On the other hand, acoustic signals obtained from specially positioned microphones for detecting propeller fracture faults have achieved classification accuracies above 99% using mel-frequency cepstral coefficients (MFCC) and deep learning-based models [19].

As observed in the literature, most studies focusing on mechanical fault detection in UAVs rely on vibration-based conventional machine learning algorithms. Although studies based on acoustic signals are relatively limited, high classification accuracies have still been achieved using MFCC features in combination with deep learning approaches or conventional classifiers such as DT and SVM. However, most of these systems are constrained by narrow frequency bands and focus primarily on propeller faults occurring under fixed rotor speeds. In contrast, the present study aims to overcome these limitations by proposing a hybrid deep learning model that utilizes acoustic signals recorded via a microphone and integrates wavelet scattering transform (WST) with LSTM architecture. The proposed model is capable of simultaneously detecting both propeller damage and loosened rotor screws with high accuracy.

In this study, acoustic signals were collected under real laboratory conditions by flying the UAV at different speeds and maneuvering scenarios. Nevertheless, due to the indoor environment, wind effects were excluded from the analysis. The signals were segmented into 1-second intervals. The multilayered time-frequency features of each signal were then extracted using the Wavelet Scattering Transform (WST). These extracted features were then integrated into an LSTM network, enabling accurate classification of complex fault patterns encountered in UAV systems. As a result, not only was a high classification accuracy of up to 98.93% achieved, but two different fault types (propeller damage and rotor screw looseness) which had not been jointly addressed in previous studies were successfully evaluated within a single framework. Furthermore, the diversity of the data acquired at various rotor and maneuvering speeds under real environmental conditions enhanced the model's generalization capability. In this regard, the study presents a meaningful innovation in the literature by contributing both methodological diversity and practical applicability to real-world UAV fault diagnosis.

The key contributions of this study can be summarized as follows:

- A novel hybrid framework combining Wavelet Scattering Transform (WST) and Long Short-Term Memory (LSTM) is proposed for acoustic-based UAV fault diagnosis,

- The model enables simultaneous detection of two distinct fault types, propeller damage and rotor screw looseness, based solely on microphone-recorded signals,
- Unlike previous studies that focus on either vibration signals or a single fault type, this study demonstrates a generalized solution over diverse flight scenarios with high accuracy and robustness.

II. METHODOLOGY

A. EXPERIMENTAL SETUP AND DATA ACQUISITION

In this study, a quadrotor multicopter equipped with the DJI A2 flight control system is used to collect data under real flight conditions. The total weight of the multicopter, including 6S LiPo batteries, is approximately 6 kg, with a wingspan of 1.5 meters including the propellers. The propulsion system consists of T-Motor MN3520 KV400 motors, each driven by a T-Motor T60A ESC (Electronic Speed Controller). The UAV is powered by a JETFIRE 40C LiPo battery (6 cells, 7000 mAh), and the propellers are made of carbon fiber with a size of 16 inches. During the experiments, recordings are collected at various rotor speed levels, which is critical for identifying the corresponding operational frequencies. Therefore, the rotational speeds of each rotor are measured using a UNI-T brand non-contact tachometer, with a measurement uncertainty of ± 4.4 RPM. The tachometer is placed at a 1 cm distance from the rotating propeller to determine both the minimum and maximum RPM values.

In the UAV system, data were collected both during flight and on the ground. During flight, the UAV was operated at various speeds up to a maximum height of 10 meters within an indoor laboratory environment, simulating real flight conditions. Wind load effects were neglected. For performance analysis and fault detection involving propeller damage and rotor screw looseness, a total of 750 one-second audio recordings were obtained. Each of these recordings represents a single fault condition and was acquired independently at different times, without further segmentation i.e., each recording inherently corresponds to a 1-second segment. Consequently, the dataset consists of 283 recordings representing healthy condition (class 0), 307 representing propeller damage (class 1), and 160 representing rotor screw looseness (class 2). Although the classes are not strictly balanced, the data were divided into 75% training, 15% validation, and 10% testing sets. A 5-fold cross-validation (k-fold=5) was also applied within the training data to mitigate class imbalance and improve generalization. Furthermore, each one-second.wav recording was assigned exclusively to a single subset (training, validation, or test), ensuring that no recordings from the same flight condition appear in multiple sets. Since data acquisition time for rotor looseness was limited due to flight safety concerns, this class is represented with fewer samples. All recordings were captured at a sampling rate of 48 kHz, which is a commonly used technical standard in audio processing, aligned with the human auditory range ($\sim$20 kHz) and the Nyquist sampling theorem. This allows for detailed analysis of both low-frequency rotor vibrations and high-frequency propeller-related anomalies.

These recordings are stored as.wav format files on an external PC equipped with an Intel(R) Core(TM) i5-8265U CPU @ 1.60 GHz, 1800 MHz. During data acquisition, rotor speeds range from 2300 to 8500 RPM on the ground, 9000 to 10000 RPM during hover flight, and 10500 to 12000 RPM during free maneuvering in the air. These different speed levels are applied across three distinct test combinations to ensure comprehensive fault coverage. First, the baseline condition includes all propellers being balanced and undamaged, with no looseness in any of the rotor connections, representing a fully healthy state. Second, a fault condition is created by loosening the rotor screws of only one propeller, introducing a 1 mm physical gap to simulate rotor screw looseness. Third, a damaged-propeller scenario is applied in which a single propeller is intentionally damaged. Propeller damage is considered as a single fault class, encompassing light, moderate, and severe damage levels. Moreover, similar conditions are tested individually for the other propellers as well.

Figure 1 illustrates the experimental setup used in this study, along with the propellers and their corresponding fault configurations. As shown in the figure, the data are categorized into three distinct classes: Class 0 represents the healthy condition, where no propeller damage is present and all rotor connections are intact. Class 1 corresponds to propeller damage, and includes all recordings obtained with light, moderate, or severe propeller faults, which are grouped into a single fault class. Finally, Class 2 represents the faulty condition in which only the rotor screws are loosened, without any physical damage to the propeller itself.

B. WAVELET SCATTERING FEATURE EXTRACTION

The Wavelet Scattering Transform (WST) is a layered and fixed-size transformation method developed to represent a signal in a multiscale and time–frequency domain. This technique is both translation-invariant over time and capable of extracting detailed frequency features without significant information loss, making it a powerful tool for signal analysis. WST provides a representation that is robust to small variations in the signal, statistically meaningful, and capable of revealing nonlinear relationships between frequency components. As such, it is considered an effective feature extraction technique for analyzing complex structural degradations, such as those encountered in fault diagnosis.

The WST is implemented by applying a sequence of wavelet transforms followed by modulus (absolute value) operations in a hierarchical (layered) manner. The resulting coefficients carry information related to the signal across multiple scales and orientations. These coefficients form a representation that is both sparse and resistant to distortions [12], [20], [21], [22].Wavelet Scattering Transform (WST) was preferred in this study due to its ability to

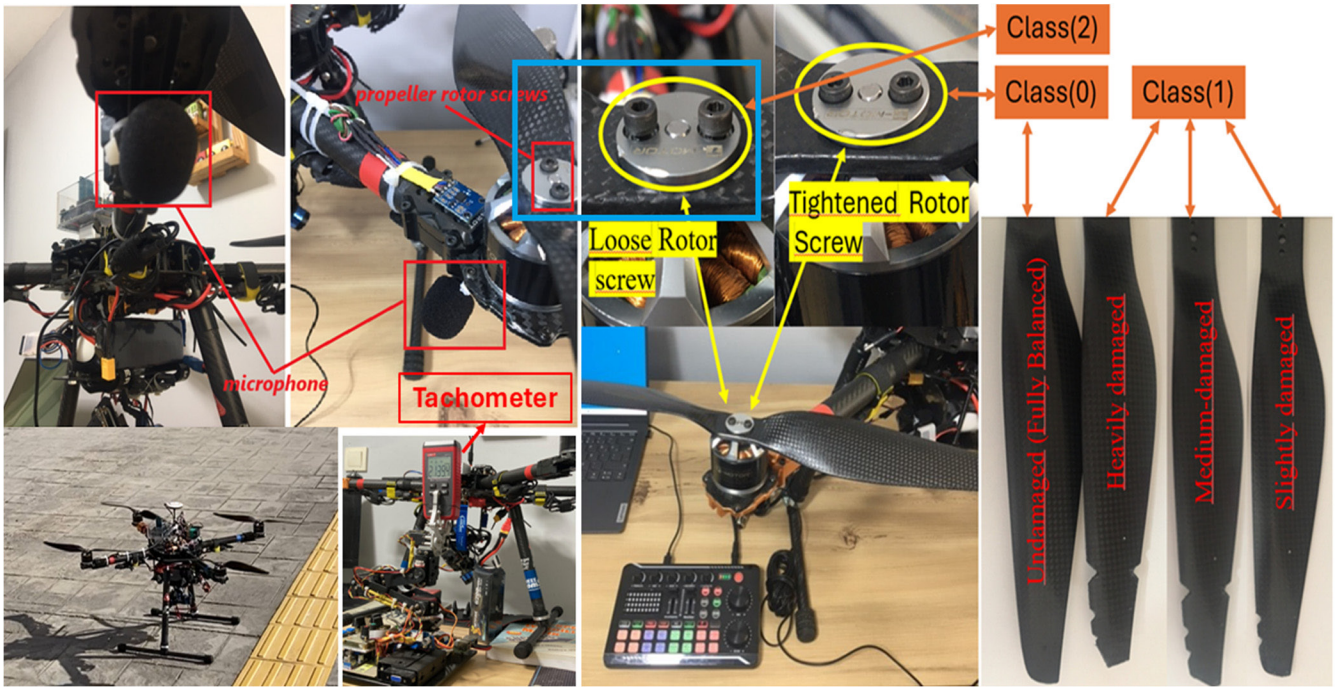


FIGURE 1. Experimental UAV System for Acoustic-Based signal acquisition, sensor placement, and implemented fault scenario.

perform high-resolution, multi-layered analysis compared to traditional time-frequency methods. Commonly used techniques in literature, such as MFCC and STFT, operate with fixed-length windows and therefore offer limited time-frequency resolution. In contrast, WST utilizes multi-scale wavelet filtering, allowing both low and high-frequency components to be preserved along the time axis, resulting in a more detailed representation. Furthermore, WST is more robust to noise, produces fixed-size outputs, and can be directly integrated with deep learning models. These properties make WST highly suitable for the acoustic fault detection tasks performed on UAV systems.

In this study, the initial step involves processing 750 seconds of acoustic signals, denoted as $x(t)$, which are acquired from the UAV and exist in the continuous-time domain. These signals are sampled at 48 kHz and segmented into 1-second intervals, resulting in a total of 750 segments. For each segment, the signal length is $N = 48kHz$ and it is represented in discrete time as $x_s[n] \in R^N$. Each segment is then convolved with a wavelet function $\psi\,[n]$centered at a frequency λ_1, as shown in (1).

$$U_1\,[n, \lambda_1] = \left|x_s\,[n] * \psi_{\lambda_1}\,[n]\right| \tag{1}$$

The wavelet function $\psi_{\lambda_1}\,[n]$ used in this study is defined as a Morlet wavelet. The Morlet wavelet is widely employed in multi-layered time–frequency analyses due to its favorable analytical properties, [21], [23]. It is a complex component modulated by a Gaussian envelope.

$$\psi_{\lambda_1}\,[n] = \pi^{-1/4}.e^{-j2\pi\lambda_1 n/N}.e^{-n^2/2\sigma^2} \tag{2}$$

The Morlet wavelet is defined as shown in (2) [23]. In this equation, the term $e^{-j2\pi\lambda_1 n/N}$ represents the complex carrier signal, while $e^{-n^2/2\sigma^2}$ serves as the Gaussian envelope for time scaling. As a result of the operation described in (1), the output $U_1\,[n, \lambda_1]$provides a time–frequency localized representation that captures the amplitude information of the signal within a specific frequency band. This representation is then convolved again with a second wavelet and followed by another modulus operation to obtain the second-layer coefficients, as shown in (3) [21].

$$U_2\,[n, \lambda_1, \lambda_2] = \left|U_1\,[n, \lambda_1] * \psi_{\lambda_2}\,[n]\right| \tag{3}$$

The second order scattering coefficients $U_2\,[n, \lambda_1, \lambda_2]$ are capable of accurately capturing modulations in the signal's frequency components, as well as transient acoustic changes caused by rotor or propeller-related disturbances. The coefficients obtained from each scattering layer, namely $U_1\,[n, \lambda_1]$ and $U_2\,[n, \lambda_1, \lambda_2]$ are then filtered using a Gaussian-based low-pass filter $\phi_j[n]$ with a scaling factor of2^j. Through this filtering process, the first- and second-order wavelet scattering coefficients, $S_1$and S_2, are obtained as defined by (4) and (5) [21], [22].

$$S_1\,[n, \lambda_1] = U_1\,[n, \lambda_1] * \phi_J\,[n] \tag{4}$$

$$S_2\,[n, \lambda_1, \lambda_2] = U_{12}\,[n, \lambda_1, \lambda_2] * \phi_J\,[n] \tag{5}$$

In this study, the feature coefficients obtained from Equations (4) and (5) are formatted into a feature matrix of size 12×412 to be used as input to the LSTM network. Here, 12 denotes the total number of scattering paths extracted from both first order and second-order components,

while 412 represents the number of time steps (frame samples) per segment. The wavelet scattering operations were implemented using MATLAB's Wavelet Toolbox and applied to each signal segment $x_s[n] \in R^N$, sampled at 48 kHz. Morlet wavelets ψ[n], known for their favorable time–frequency localization, were used in constructing the scattering network. The first layer utilized 8 band-pass filters per octave, and the second layer employed 1 wavelet, aligning with the definitions of $S_1\,[n, \lambda_1]$ and $S_2\,[n, \lambda_1, \lambda_2]$. A time invariance scale of approximately 0.5 seconds, defined by the low-pass filter $\phi_j[n]$, was used to obtain stable temporal representations across segments. The resulting time–frequency feature matrix captures both local and global variations in the acoustic signals and serves as sequential input to the LSTM classifier.

C. LONG SHORT-TERM MEMORY (LSTM)

LSTM is a neural network architecture built upon the structure of artificial neural networks (ANNs). This type of ANN belongs to the category of recurrent neural networks (RNNs), yet it offers a more powerful structure compared to conventional RNNs [24]. Owing to this capability, LSTM achieves high accuracy, particularly in processing time-dependent acoustic data and classifying fault types. Therefore, LSTM is widely used in various tasks involving speech and acoustic recognition algorithms [25], [26], [27]. It has also been extensively applied in data-driven diagnostics of rotating machinery [28], [29].

The LSTM cell primarily consists of four key components: the forget gate, input gate, cell state, and output gate. These components determine how much of the past information is retained and how much is updated at each step. Cell state can preserve information over long time intervals and enables the network to overcome the vanishing gradient problem. At each time step t, the LSTM network updates its cell state and output using the equations defined below [12]:

$$f_t = \sigma\left(W_f\,[h_{t-1}, x_t] + b_f\right) \tag{6}$$

$$i_t = \sigma\left(W_i\,[h_{t-1}, x_t] + b_i\right) \tag{7}$$

$$\tilde{C}_t = tanh\left(W_c\,[h_{t-1}, x_t] + b_c\right) \tag{8}$$

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \tag{9}$$

$$o_t = \sigma\left(W_o\,[h_{t-1}, x_t] + b_o\right) \tag{10}$$

$$h_t = o_t \tanh\left(C_t\right) \tag{11}$$

where x_t is the WST feature vector at time step t. $h_{t-1}$is the hidden state from the previous time step, and h_t is the current hidden state. Also W_f, W_i and W_o are the weight matrices for the forget gate f_t, the input gate i_t, and the output gate o_t,respectively. Similarly b_f, b_i and b_o are the corresponding bias parameters, C_{t-1}denotes the previous cell state,$\tilde{C}_t$ is the candidate cell state, and C_t is the current cell state. σ is the sigmoid function, calculated as shown in (12) [30].

$$\sigma = \frac{1}{1 + e^{-z}} \tag{12}$$

Here, the term z represents the weighted sum in the form of $W\,[h_{t-1}, x_t]\, +\, b$. As seen in the equations above, this weighted sum is processed through the sigmoid function at each gate.

Based on the LSTM cell structure, a custom LSTM network is designed in this study to process time series features extracted via WST as input. The WST coefficients obtained from each audio segment are fed into the LSTM layer as a time-step-based sequence, enabling the network to learn temporal patterns within the data. The LSTM layer in this architecture consists of 512 units, followed by a hidden layer with 256 neurons and ReLU activation. The 512-unit architecture selected for the LSTM model was determined as the optimal configuration based on tests conducted on different setups, providing the best balance between classification accuracy and computational load. The final output layer is a three-class classifier, corresponding to the following fault labels: Healthy (0), Propeller Damage (1), and Rotor Looseness (2).

Figure 2 illustrates the overall architectural structure of the proposed model. Additionally, Algorithm 1 presents the algorithmic structure of the WST-LSTM model for real-time applications. It explains how the trained model performs step-by-step real-time fault classification based on 1-second audio segments. The model processes each segment independently and continuously classifies incoming data in a streaming fashion.

The proposed model is trained using a supervised learning algorithm based on LSTM architecture. 75% of the data is allocated for training. During training, the Adam optimization algorithm is employed, which is highly efficient for large-scale problems involving extensive data or parameters [31]. The model weights are updated using the Backpropagation Through Time (BPTT) algorithm, based on the WST-based feature sequences. These updates aim to minimize the model's prediction error. The cross-entropy (CE) loss function is used throughout the training process [32].

$$L_{CE} = -\frac{1}{n}\sum\nolimits_{t=1}^{n} Y_t^T \log\left(\hat{Y}_t\right) \tag{13}$$

The CE equation is presented in (13). In this equation, n denotes the number of samples in the mini-batch, and $\hat{Y}_t \in R^3$represents the predicted output vector obtained from the final fully connected layer of the LSTM, expressed as $\sigma\,(Wh_t + b)$. The ground truth class vector is denoted by $Y_t \in R^3$.

Regarding the evaluation of the model's performance, accuracy, precision, and F1-score values obtained from the confusion matrix were utilized, along with the area under the curve (AUC) values derived from ROC (Receiver Operating Characteristic) curves. The ROC curves are used to assess the model's ability to distinguish between classes, with the vertical axis representing the True Positive Rate and the horizontal axis showing the False Positive Rate. These curves provide a visual representation of the classifier's discriminative performance. Ideally, the curve proceeds from point (0,0) to (0,1) and then to (1,1), where a trajectory closer to the top-left corner indicates higher accuracy with fewer

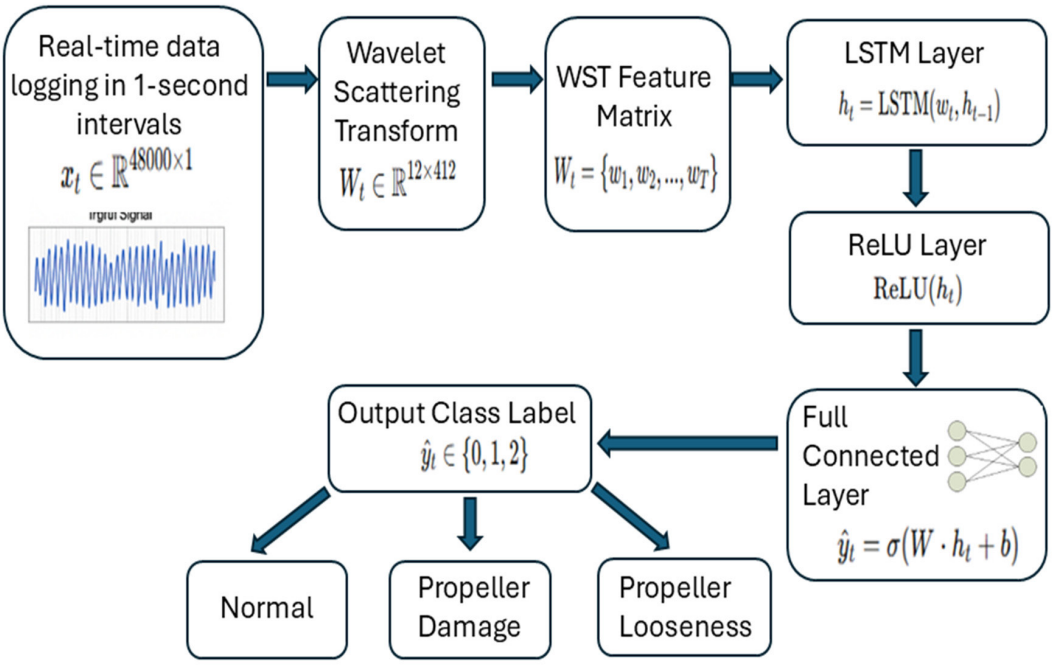


FIGURE 2. Block diagram of the proposed WST-LSTM model fault classification framework.

Algorithm 1 Real-Time Loop of the Trained WST-LSTM Model

Hyperparameters: Segment duation T = 1sec, sampling frequency fs=48kHz
Require: Trained LSTM model M
Require: Wavelet Scattering Transform configuration

1. Initialize system
2. while system is active do
3. Acquire 1-second audio segment $x_z \in \mathbb{R}^{4000\times^t}$
4. Normalize x_t if necessary
5. Apply WST on x_z to obtain $W_z \in \mathbb{R}^{12} \times 422$
6. Transpose W_t to match LSTM input: $W_t^\top \in \mathbb{R}^{412\times12}$
7. for t=1 to 412 do
8. Feed $W_t^T[t]$ into LSTM
9. Update hidden state h_t, cell state C_t
10. end for
11. Extract final hidden state $h_T \in \mathbb{R}^{512}$
12. Apply ReLU : a = ReLU $(h_T) \in \mathbb{R}^{236}$
13. Compute output: $\hat{y}_t$ = softmax(W · a + b) $\in \mathbb{R}^3$
14. Determine predicted class from $\hat{y}_t$
15. Output predicted class
16. end while

false positives. In contrast, a curve that lies near the diagonal suggests weaker class separability. Therefore, the AUC value serves as a critical metric that summarizes the overall classification performance of the model [33], [34].

III. EXPERIMENTAL RESULTS

A. FAULT CONDITION COMPARISON VIA WST FREQUENCY ANALYSES

In this study, the average WST-based feature distributions are analyzed for three different classes: healthy condition, propeller damage, and rotor screw looseness. These analyses provide important insights into the patterns left by each fault type in the time–frequency domain and feature space.

As shown in Figures 3(a) and 3(b), the frequency spectrogram corresponding to the healthy scenario reveals the presence of natural frequencies arising from the system's structural characteristics particularly prominent around 8 kHz.

The average WST-based feature distribution indicates that only a few low-index coefficients exhibit high amplitude values. This suggests that even in a fault-free condition, the system inherently carries information in a limited number of frequency bands, with minimal energy spread in the mid and high frequency ranges. In contrast, under propeller fault conditions as observed in Figures 3(c) and 3(d), significant changes appear in the frequency spectrogram.

While the system's natural frequency components remain visible, the energy is more broadly distributed across the frequency spectrum. This behavior reflects the presence of blade vibrations caused by rotational imbalance.

Similarly, in the average WST feature distribution, a slight reduction in energy is observed for low-index coefficients compared to the healthy case, whereas a noticeable increase is seen in mid-band features (approximately between indices 150–220). This indicates that propeller damage not only affects the fundamental modes of the system but also excites more complex sub-modes. Additionally, slight increases in some high-index coefficients are also observed.

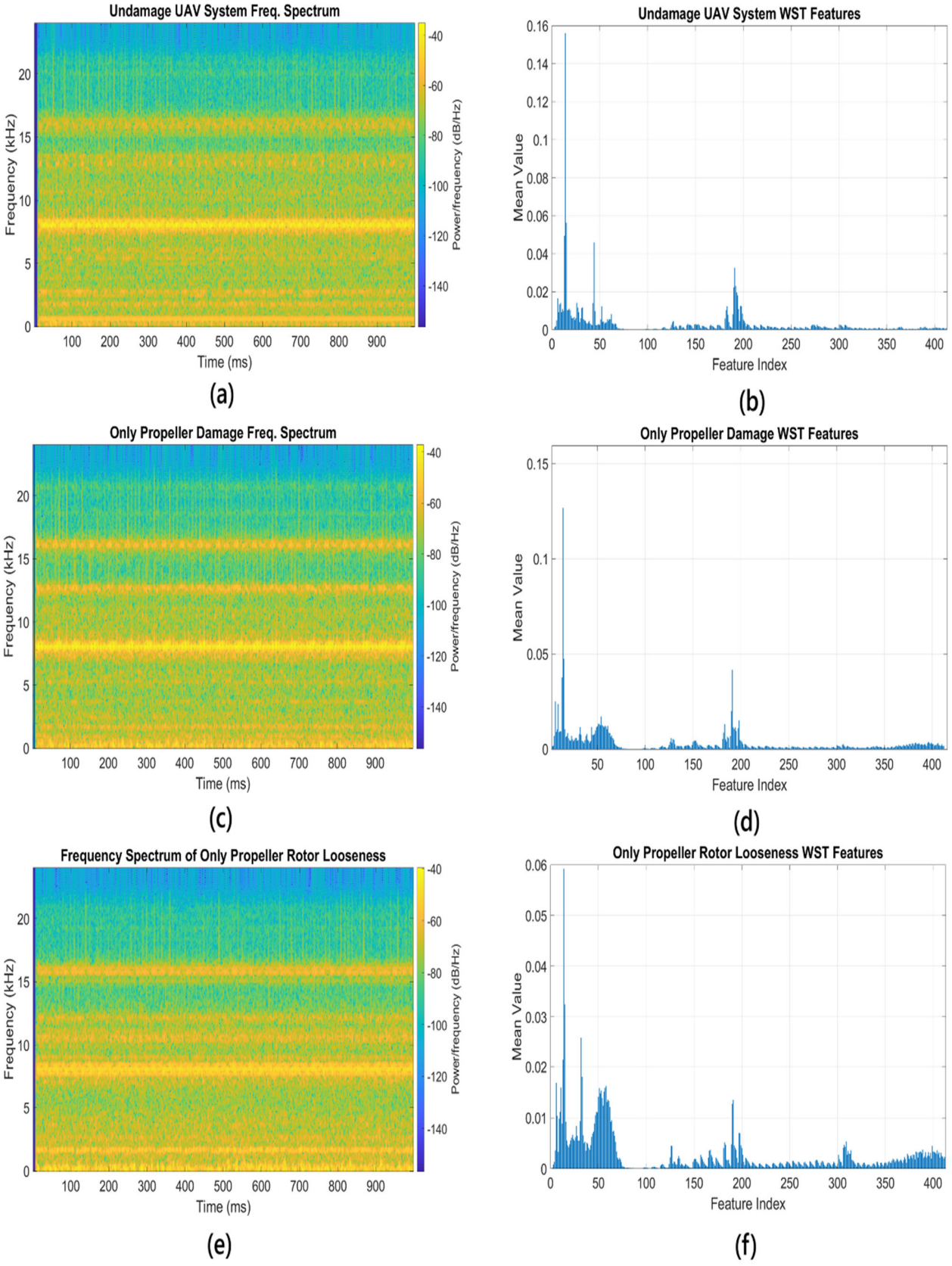


FIGURE 3. Time–frequency spectrograms (a, c, e) and corresponding average Wavelet Scattering Transform (WST) feature distributions (b, d, f) for undamaged UAV, propeller damage only, and rotor looseness only conditions.

Finally, when focusing specifically on faults caused by rotor screw looseness, Figures 3(e) and 3(f) reveal an increase in energy concentration at low-frequency components. This indicates that low-frequency vibrations resulting from structural looseness are reflected in the system's acoustic behavior. The average WST-based feature distribution supports these findings. In particular, a substantial number of coefficients within the feature index range of 0–60 attain significant values, suggesting that low-frequency components have become dominant in the system's dynamic behavior. Moreover, in the mid-band region especially between indices 180 and 220, a noticeable increase is observed in certain coefficients, implying that the looseness may gradually influence other system modes over time.

In summary, when these three conditions are evaluated collectively, the following general conclusions can be drawn:

- The healthy structure exhibits a stable and harmonic energy distribution in the frequency domain and is represented in the WST outputs by only a limited number of dominant components.
- Propeller damage causes the energy spectrum to spread across a wider frequency band and leads to noticeable increases in specific feature indices, particularly in the mid-frequency range. This behavior can be associated with rotational imbalance and aerodynamic disturbances.
- Rotor screw looseness predominantly affects the low frequency region of the spectrum and activates a large number of low-index components in the WST features. This is consistent with the fact that structural looseness generates dominant low-frequency vibrations within the system.

In conclusion, these findings demonstrate that WST-based feature extraction plays a complementary and explanatory role in distinguishing between different fault types. In other words, the discrete distribution of WST features and the transitions between frequency bands provide a highly discriminative foundation for classification algorithms.

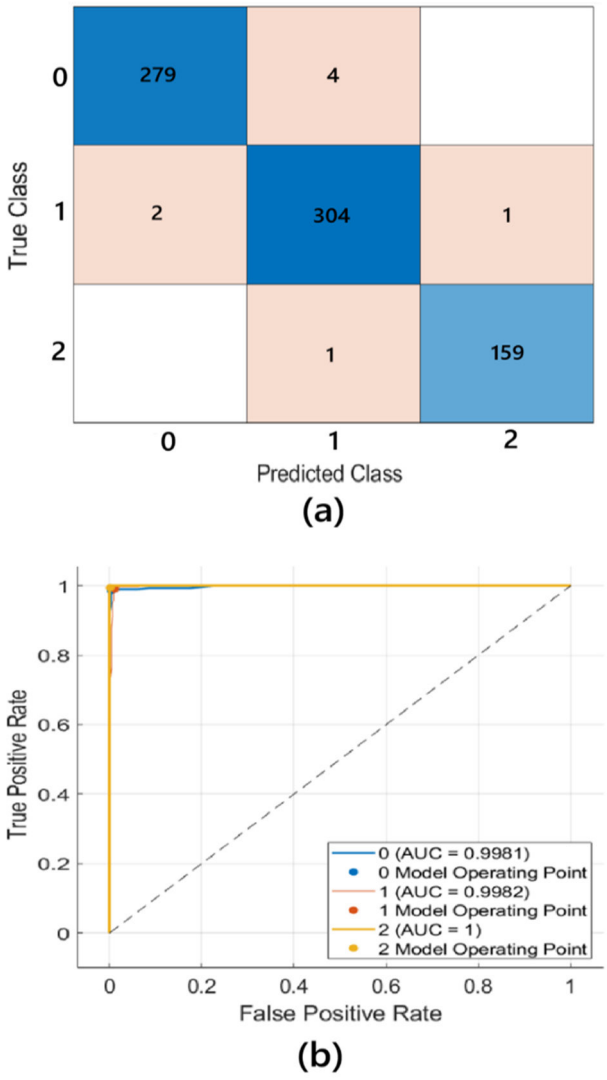


FIGURE 4. The Classification performance; (a) Confusion matrix results, (b) ROC curve AUC performance.

B. THE WST-LSTM MODEL PERFORMANCE

The classification performance of the proposed LSTM-based model is evaluated using the receiver operating characteristic (ROC) curve, which reflects AUC performance, and the confusion matrix. The model demonstrates a high discriminative capability across the three different conditions: 0 – Healthy, 1 – Propeller Damage, and 2 – Rotor Screw Looseness. All classification performance analyses are conducted using MATLAB R2024a, and the results are detailed in Figure 4 and Table 1.

As illustrated in the confusion matrix in Figure 4(a), the alignment between the model's predictions and the actual labels is clearly demonstrated. Out of a total of 750 samples obtained from both healthy and faulty/problematic UAV conditions, 742 were correctly classified using the proposed WST & LSTM-based model, yielding an overall classification accuracy of 98.93%. Specifically, for the health condition (Class 0), 279 samples were correctly classified, with only 4 misclassified as propeller damage. For the propeller damage condition (Class 1), 304 samples were accurately predicted, while only 3 samples were misclassified. In the rotor screw looseness condition (Class 2), nearly all the 160 samples are perfectly distinguished, with only 1 misclassification.

Similarly, based on the ROC curves shown in Figure 4(b), the AUC values were calculated as 0.9981 for class 0 (healthy state), 0.9982 for class 1 (propeller damage), and 1.000 for class 2 (rotor screw looseness). These high AUC scores indicate that the model is highly capable of accurately distinguishing between all three fault categories. The proximity of the curves to the upper-left corner of the plot suggests that the true positive rate is maximized while the false positive rate is minimized. In particular, the perfect AUC score of 1.000 for class 2 highlights the model's exceptional ability to identify rotor screw looseness cases with near-perfect precision.

TABLE 1. Classification Performance Results.

Output Class	Precision (%)	F1-Score (%)	AUC (%)
Class 0	98,58	98,58	98,81
Class 1	98,70	98,86	99,82
Class 2	99,38	99,38	100
Overall Accuracy Rate		98,93%	
Average AUC		99,88%	

Class 0: Undamaged system
Class 1: Only propeller damage
Class 2: Only rotor screw looseness

The performance results presented in Figure 4 are further detailed in Table 1. In summary, when the AUC values and confusion matrix results are evaluated together, the overall classification performance of the model is found to be exceptionally high. Both the ROC curves and the confusion matrix confirm that the model operates in a balanced and reliable manner. The time-stepped features extracted via WST, when processed by the LSTM network, provide powerful and discriminative information for acoustic-based fault detection. These high accuracy rates further support the practical applicability of the model in real-time UAV fault diagnosis scenarios.

IV. DISCUSSION AND CONCLUSION

Existing UAV fault diagnosis studies in the literature have typically been conducted under constant-speed and controlled laboratory conditions. This limits the generalizability of the proposed models, as they fail to adequately reflect the variability of real-world flight environments. Furthermore, many prior works have focused solely on either propeller damage or arm screw looseness, often neglecting critical mechanical faults such as rotor screw looseness [9], [10], [11], [12], [13], [14], [15]. For instance, Al-Haddad and Jaber [9] analyzed propeller faults using frequency components and achieved 91% accuracy with an AUC of 0.98, but their study was conducted only at a fixed speed. Similarly, Gazali and Rahiman [15] proposed a Fuzzy Logic and Neural Network-based approach for detecting arm screw looseness in laboratory conditions. However, these two fault types were not addressed concurrently in any single model. In contrast, our study introduces an innovative acoustic signal-based system capable of detecting both propeller damage and rotor screw looseness using a unified framework. Additionally, the experimental setup was designed to capture data across different throttle levels, i.e., varying rotational speeds. This approach helps mitigate overfitting in the datasets and enhances the model's generalizability.

Moreover, in our previous study, vibration signals obtained from MEMS-based accelerometers and gyroscopes were successfully used to detect both propeller faults and mechanical looseness at rotor/arm joints. Two separate models were developed for low-cost embedded platforms using Decision Trees (DT) and Neural Networks (NN), respectively [7]. However, that earlier work relied solely on vibration data. Due to the nature of frequency-based features and the physical sensor placements, microstructural deformations in some UAV regions could limit fault detection performance. To address these limitations, the present second study proposes a hybrid WST-LSTM approach using microphone-based acoustic signals to identify both propeller faults and rotor screw looseness. This architecture enables the extraction of multi-layered time-frequency features, effectively distinguishing frequency transitions in each signal segment with a high classification accuracy of 98.93%. Notably, for structural failures such as rotor screw looseness, the model achieved an AUC value of up to 100%, underscoring its strength in both sensitivity and generalization. The exceptional performance of the proposed model stems from the WST's ability to extract deformation-stable, multi-layered time-frequency features, combined with the LSTM network's capacity to learn temporal dependencies between them. The joint evaluation of spectral amplitude information and temporal patterns provides significant advantages in capturing both low-frequency propeller faults and more irregular rotor looseness cases. In other words, the model successfully distinguishes between different fault types, including both propeller damage and rotor screw looseness.

While some vibration-based studies employing LSTM and WST architectures have achieved classification accuracies up to 99% [12], they typically focus only on propeller faults, disregarding rotor looseness and lacking acoustic signal analysis. Similarly, although a few acoustic-based fault diagnosis studies exist, they commonly rely on MFCC features [19] and remain limited to propeller faults, ignoring rotor-related issues. In our proposed WST-LSTM-based hybrid model, acoustic signals are analyzed in a layered structure that captures both mid-frequency and high-frequency components. The distinct time-frequency effects of each fault type are isolated, enabling unified identification of propeller faults and rotor screw looseness with high accuracy.

Another notable contribution of this study lies in the comprehensive testing of all fault scenarios under varying rotor speeds, with the UAV placed on the ground, hovering, and in motion. This improves the model's generalization capability. In this respect, the study distinguishes itself from prior work, both methodologically and in terms of data collected under diverse flight modes. Nonetheless, the proposed system also has certain limitations. The current model was trained under controlled testing conditions in a closed laboratory environment with a fixed microphone position. Therefore, the impact of background noise on model performance was not evaluated in this study. It is anticipated that environmental noise, wind effects, and variability in microphone placement during real-time flights may influence the system's accuracy. Furthermore, while this study focused

on post-acquisition signal processing, future research will aim to develop a real-time fault detection system that can not only acquire but also process the data instantaneously. Accordingly, the system is planned to be tested under more complex and noisy conditions and adapted for real-time diagnostics.

In conclusion, this study introduced robust system architecture that combined WST features with LSTM network to effectively detect both transient and continuous anomalies in acoustic signals. It successfully addressed a gap in the existing literature by demonstrating that both propeller damage and rotor screw looseness could be identified within a unified model using UAV flight data under various operating modes. This achievement significantly advanced the capabilities of acoustic-based fault diagnosis in UAVs, offering not only high classification accuracy but also strong potential for real-time implementation in practical UAV health monitoring systems.

In future work, it is planned to develop a deployable system that can simultaneously acquire and process acoustic data in real-world conditions, considering the impact of environmental noise, wind effects, and variations in microphone placement on model performance.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Ibrahim Okten, the owner of the company (Inci Automation Ltd.), for his support.

REFERENCES

[1] V. Puri, A. Nayyar, and L. Raja, "Agriculture drones: A modern breakthrough in precision agriculture," *J. Statist. Manage. Syst.*, vol. 20, no. 4, pp. 507–518, Jul. 2017.

[2] M. A. Ma'sum, M. K. Arrofi, G. Jati, F. Arifin, M. N. Kurniawan, P. Mursanto, and W. Jatmiko, "Simulation of intelligent unmanned aerial vehicle (UAV) for military surveillance," in *Proc. Int. Conf. Adv. Comput. Sci. Inf. Syst. (ICACSIS)*, Sanur Bali, Indonesia, Sep. 2013, pp. 161–166.

[3] M. H. M. Ghazali, K. Teoh, and W. Rahiman, "A systematic review of real-time deployments of UAV-based LoRA communication network," *IEEE Access*, vol. 9, pp. 124817–124830, 2021.

[4] A. Bondyra, P. Gąsior, S. Gardecki, and A. Kasiński, "Development of the sensory network for the vibration-based fault detection and isolation in the multirotor UAV propulsion system," in *Proc. 15th Int. Conf. Informat. Control, Autom. Robot.*, 2018, pp. 102–109.

[5] L. A. Al-Haddad and A. A. Jaber, "Influence of operationally consumed propellers on multirotor UAVs airworthiness: Finite element and experimental approach," *IEEE Sensors J.*, vol. 23, no. 11, pp. 11738–11745, Jun. 2023.

[6] A. Baldini, R. Felicetti, F. Ferracuti, A. Freddi, and S. Iarlori, "Real-time propeller fault detection for multirotor drones based on vibration data analysis," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, Art. no. 106343.

[7] T. Sonmezocak, "Intelligent UAV health monitoring: Detecting propeller and structural faults with MEMS-based vibration," *Eng. Sci. Technol., Int. J.*, vol. 69, Sep. 2025, Art. no. 102130.

[8] L. Al-Haddad, A. Jaber, P. Neranon, and S. Al-Haddad, "Investigation of frequency-domain-based vibration signal analysis for UAV unbalance fault classification," *Eng. Technol. J.*, vol. 41, no. 7, pp. 1–9, Feb. 2023.

[9] L. A. Al-Haddad and A. A. Jaber, "An intelligent fault diagnosis approach for multirotor UAVs based on deep neural network of multi-resolution transform features," *Drones*, vol. 7, no. 2, p. 82, Jan. 2023.

[10] L. A. Al-Haddad, W. Giernacki, A. A. Shandookh, A. A. Jaber, and R. Puchalski, "Vibration signal processing for multirotor UAVs fault diagnosis: Filtering or multiresolution analysis?" *Eksploatacja i Niezawodność Maintenance Rel.*, vol. 26, no. 1, Dec. 2023, Art. no. 176318.

[11] A. Bondyra, P. Gąsior, S. Gardecki, and A. Kasiński, "Fault diagnosis and condition monitoring of UAV rotor using signal processing," in *Proc. Signal Process., Algorithms, Archit., Arrangements, Appl.*, Poznań, Poland, 2017, pp. 233–238.

[12] E. C. Ozkat, "Vibration data-driven anomaly detection in UAVs: A deep learning approach," *Eng. Sci. Technol., Int. J.*, vol. 54, Jun. 2024, Art. no. 101702.

[13] K. Chen, W. Meng, J. Wang, K. Liu, and Z. Lu, "An investigation on the structural vibrations of multi-rotor passenger drones," *Int. J. Micro Air Vehicles*, vol. 15, pp. 1–10, Jan. 2023.

[14] J. Verbeke and S. Debruyne, "Vibration analysis of a UAV multirotor frame," in *Proc. IEEE Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Orlando, FL, USA, Sep. 2016, pp. 2329–2337.

[15] M. H. M. Ghazali and W. Rahiman, "Vibration-based fault detection in drone using artificial intelligence," *IEEE Sensors J.*, vol. 22, no. 9, pp. 8439–8448, May 2022.

[16] Z. Zhou and Y. Liu, "A smart landing platform with data-driven analytic procedures for UAV preflight safety diagnosis," *IEEE Access*, vol. 9, pp. 154876–154891, 2021.

[17] F. R. Lopez-Estrada, A. Mendez-Lopez, I. Santos-Ruiz, G. Valencia-Palomo, and E. Escobar-Gomez, "Fault detection in unmanned aerial vehicles via orientation signals and machine learning," *Rev. Iberoam. Autom. Inform. Ind.*, vol. 18, no. 3, pp. 254–264, 2021.

[18] A. Altinors, F. Yol, and O. Yaman, "A sound based method for fault detection with statistical feature extraction in UAV motors," *Appl. Acoust.*, vol. 183, Dec. 2021, Art. no. 108325.

[19] A. Bondyra, M. Kołodziejczak, R. Kulikowski, and W. Giernacki, "An acoustic fault detection and isolation system for multirotor UAV," *Energies*, vol. 15, no. 11, p. 3955, May 2022.

[20] S. Mallat, "Group invariant scattering," *Commun. Pure Appl. Math.*, vol. 65, no. 10, pp. 1331–1398, Oct. 2012.

[21] J. Andén and S. Mallat, "Deep scattering spectrum," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4114–4128, Aug. 2014.

[22] R. N. Toma, Y. Gao, F. Piltan, K. Im, D. Shon, T. H. Yoon, D.-S. Yoo, and J.-M. Kim, "Classification framework of the bearing faults of an induction motor using wavelet scattering transform-based features," *Sensors*, vol. 22, no. 22, p. 8958, Nov. 2022.

[23] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1872–1886, Aug. 2013.

[24] S. M. Mendis, G. M. K. B. Karunasena, and D. H. R. J. Wimalasiri, "Non-intrusive motorcycle engine fault diagnosis using LSTM network and spectrogram-based audio analysis," in *Proc. Int. Res. Conf. Smart Comput. Syst. Eng. (SCSE)*, Colombo, Sri Lanka, Apr. 2025, pp. 1–6.

[25] A. Seralieva, M. Ilyas, A. Gissa, and M. H. Ali, "Hybrid CNN-LSTM model for non-invasive fault detection in induction motors using acoustic data," in *Proc. IEEE 12th Conf. Syst., Process Control (ICSPC)*, Malaysia, Dec. 2024, pp. 136–141.

[26] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019.

[27] K. Smagulova and A. P. James, "A survey on LSTM memristive neural network architectures and applications," *Eur. Phys. J. Special Topics*, vol. 228, no. 10, pp. 2313–2324, Oct. 2019.

[28] B. Lindemann, B. Maschler, N. Sahlab, and M. Weyrich, "A survey on anomaly detection for technical systems using LSTM networks," *Comput. Ind.*, vol. 131, Oct. 2021, Art. no. 103498.

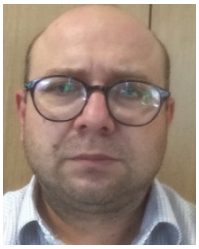
[29] M. Jalayer, C. Orsenigo, and C. Vercellis, "Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, fast Fourier and continuous wavelet transforms," *Comput. Ind.*, vol. 125, Feb. 2021, Art. no. 103378.

[30] N. Kyurkchiev and S. Markov, *Sigmoid Functions: Some Approximation and Modelling Aspects*. Saarbrücken, Germany: Lap Lambert Academic, 2015.

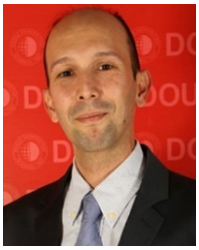
[31] R. N. Singarimbun, E. B. Nababan, and O. S. Sitompul, "Adaptive moment estimation to minimize square error in backpropagation algorithm," in *Proc. Int. Conf. Comput. Sci. Inf. Technol. (ICoSNIKOM)*, Nov. 2019, pp. 1–7.

[32] Y. Wang, X. Ma, Z. Chen, Y. Luo, J. Yi, and J. Bailey, "Symmetric cross entropy for robust learning with noisy labels," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Jan. 2019, pp. 322–330.

[33] A. N. Kamarudin, T. Cox, and R. Kolamunnage-Dona, "Time-dependent ROC curve analysis in medical research: Current methods and applications," *BMC Med. Res. Methodology*, vol. 17, no. 1, pp. 1–19, Dec. 2017.
[34] M. V. Arteaga, J. C. Castiblanco, I. F. Mondragon, J. D. Colorado, and C. Alvarado-Rojas, "EMG-driven hand model based on the classification of individual finger movements," *Biomed. Signal Process. Control*, vol. 58, Apr. 2020, Art. no. 101834.



TEMEL SONMEZOCAK (Member, IEEE) was born in Edirne, in 1978. He received the B.S. degree from the Electrical Education Department, Technical Education Faculty, University of Marmara, Istanbul, in 2001, the M.Sc. degree from the Department of Electrical and Electronics Engineering, Istanbul Aydin University, in 2015, and the Ph.D. degree in education from the Electronics Engineering Program, Yıldız Technical University, in 2021. From 2022 to 2025, he was an Assistant Professor (Dr.) with the Department of Electrical and Electronics Engineering, Doğuş University. He is currently with the Department of Electrical and Electronics Engineering, Yeni Yüzyıl University. His research interests include circuits, systems, biomedical, sensor networks, lighting systems, mechatronics, and instrumentation.



MERİH YILDIZ received the B.S. and M.Sc. degrees in electronics and communication engineering and the Ph.D. degree in electronics engineering from Istanbul Technical University, Istanbul, Türkiye, in 2000, 2003, and 2009, respectively. He was a Field Support Engineer with Nortel Networks-Netas, from 2000 to 2001. He is currently an Assistant Professor with the Department of Electrical and Electronics Engineering, Doğuş University, Istanbul. His current research interests include current-mode circuits, analog circuit design, and analog signal processing.