TẠP CHÍ KHOA HỌC VÀ CÔNG NGHỆ NĂNG LƯỢNG - TRƯỜNG ĐẠI HỌC ĐIỆN LỰC
(ISSN: 1859 - 4557)

A MACHINE LEARNING APPROACH FOR UAV SIGNAL DETECTION BASED ON STFT SPECTROGRAMS AND 2D-DCT FEATURES

PHƯƠNG PHÁP HỌC MÁY CHO PHÁT HIỆN TÍN HIỆU UAV DỰA TRÊN PHỔ STFT VÀ ĐẶC TRƯNG 2D-DCT

Trong Thanh Nguyen*, Le Cuong Nguyen
Electric Power University
Received: March 1, 2026 Accepted: April 20, 2026

Abstract:

With the rapid growth of civilian and military UAV applications, reliable UAV signal detection has become increasingly important for spectrum monitoring and security. This research proposes a machine-learning method for UAV signal detection from raw IQ data. The IQ stream is converted into a time-frequency spectrogram using the Short-Time Fourier Transform (STFT), then transformed by two-dimensional Discrete Cosine Transform (2D-DCT) to exploit energy compaction and obtain a compact representation dominated by low-order coefficients. The DCT matrix is flattened and segmented in three ways: 15 segments over 300,000 coefficients, 15 segments over the first 20,000 coefficients, and 15 segments over the first 1,333 coefficients. The study also evaluates a coefficient selection strategy, narrowing from the initial 4,096 coefficients down to 640, and finally to 5 salient groups. . Thirteen statistical features are surveyed per segment, emphasizing RMS, peak value, shape factor, and harmonic mean. Feature vectors are classified by XGBoost, KNN, Decision Tree, and SVM, and evaluated by Accuracy, Precision, Recall, and F1-score. On an independent test set of 12,354 spectrogram images, XGBoost achieves 98.8% accuracy, showing that low-order DCT coefficients provide strong discriminative information for binary drone-versus-background detection at low computational cost.

Keywords: UAV; Raw IQ data; STFT; 2D-DCT; Statistical; Machine learning.

Tóm tắt:

Cùng với sự gia tăng nhanh của các ứng dụng UAV trong dân sự và quân sự, phát hiện tín hiệu UAV một cách tin cậy ngày càng cần thiết đối với giám sát phổ tần và an ninh. Trong nghiên cứu này, nghiên cứu đề xuất một phương pháp học máy để phát hiện tín hiệu UAV từ dữ liệu IQ thô. Trước hết, chuỗi IQ được chuyển thành ảnh phổ thời gian - tần số bằng biến đổi Fourier thời gian ngắn (STFT). Tiếp theo, biến đổi cosine rời rạc hai chiều (2D-DCT) được áp dụng nhằm khai thác tính dồn năng lượng, tạo ra biểu diễn gọn hơn với phần lớn thông tin tập trung ở các hệ số bậc thấp. Ma trận hệ số DCT được trải phẳng và chia thành ba cách: 15 đoạn trên 300.000 hệ số, 15 đoạn trên 20.000 hệ số đầu, và 15 đoạn trên 1.333 hệ số đầu. Nghiên cứu cũng phân tích chiến lược chọn hệ số từ 4.096 hệ số đầu, thu hẹp còn 640 hệ số đầu, rồi tiếp tục chọn 5 nhóm nổi bật. Trên mỗi đoạn, 13 đặc trưng thống kê được khảo sát, trong đó nhấn mạnh RMS, giá trị đỉnh, shape factor và harmonic mean. Các vector đặc trưng được phân loại bằng XGBoost, KNN, cây quyết định và SVM, được đánh giá theo Accuracy, Precision, Recall và F1-score. Trên tập kiểm tra độc lập gồm 12.354 ảnh phổ, XGBoost đạt Accuracy 98.8%, cho thấy các hệ số DCT bậc thấp mang thông tin phân biệt mạnh cho bài toán phát hiện nhị phân drone - background với chi phí tính toán thấp.

Từ khóa: Máy bay không người lái; Tín hiệu IQ thô; STFT; 2D-DCT; Đặc trưng thống kê; XGBoost.

1. INTRODUCTION

Unmanned Aerial Vehicles (UAVs), commonly known as drones, have experienced a significant increase in deployment across a wide range of civil and commercial applications, including logistics, precision agriculture, infrastructure inspection, and disaster management [1- 4]. These aerial platforms provide transformative capabilities and cost-effective solutions for various practical tasks [1-3]. However, the widespread availability of low-cost UAV technology has also raised serious security and privacy concerns [1, 3, 5]. The misuse of UAVs for unauthorized surveillance, espionage, or disruption near sensitive areas such as airports and power plants poses substantial risks to public safety and critical infrastructure [1, 6, 7]. As a result, there is an urgent and growing need for robust monitoring and detection systems capable of identifying non-cooperative UAVs operating within a given airspace [1, 3, 8, 9].

Among different sensing modalities, including radar, vision-based, and acoustic systems, radio-frequency (RF)-based detection has emerged as a promising approach for UAV monitoring [1, 3, 4, 10, 11]. RF-based systems exploit the inherent characteristics of communication signals transmitted between a drone and its ground control station [1, 3, 4, 12]. These methods offer several advantages, such as long-range operation and the ability to work under non-line-of-sight (NLOS) conditions, independent of weather and visibility conditions [1, 4, 9, 11]. Nevertheless, RF-based detection still faces significant challenges, especially due to interference in the heavily congested 2.4 GHz Industrial, Scientific, and Medical (ISM) band, where common wireless technologies such as WiFi and Bluetooth also operate [1, 2, 4, 11, 13]. Distinguishing UAV-specific signal characteristics from background noise and other wireless emitters remains difficult, particularly in low signal-to-noise ratio (SNR) environments [2, 11, 14].

A major challenge in applying machine learning to RF signal analysis lies in the representation of the input data. Although raw in-phase and quadrature (IQ) signals preserve complete information about the captured transmission, directly using this one-dimensional time-domain data for classification is often impractical. Raw IQ data usually contains a large number of samples, resulting in high computational and storage demands. In addition, time-domain signals are sensitive to segmentation and can be strongly affected by noise or outliers, which may degrade classification performance [15]. To address these limitations, the Short-Time Fourier Transform (STFT) is widely used to represent signals in the time-frequency domain [1, 3, 16]. STFT is well suited for non-stationary signals such as UAV transmissions because it can capture dynamic spectral characteristics and reveal important patterns, including frequency hopping, that are useful for reliable signal identification [1, 3, 5, 16].

To improve feature extraction efficiency and reduce computational complexity, this study investigates the use of the two-dimensional Discrete Cosine Transform (2D-DCT) on

STFT spectrograms [17]. DCT is a well-established signal processing tool known for its energy compaction property, which concentrates the most informative signal content into a small number of low-order coefficients. For highly correlated data such as spectrogram images, 2D-DCT compresses most of the signal energy and structural information into the low-frequency region, typically located in the upper-left corner of the coefficient matrix. By focusing on these low-order coefficients, the method can emphasize the most relevant spectral structures of drone signals while suppressing small fluctuations caused by noise. This compact representation enables more efficient feature extraction and makes the overall framework more suitable for deployment on resource-constrained platforms.

Compared with STFT+CNN approaches, the proposed framework is computationally lighter and more interpretable because it does not rely on end-to-end deep feature learning. Unlike RF fingerprinting methods that often depend on learned deep representations, this study uses compact statistical features extracted from segmented low-order DCT coefficients. Thus, the contribution of this work lies in a lightweight and explainable framework for binary drone-versus-background detection.

In this paper, we propose a machine learning pipeline for binary UAV signal detection, specifically to distinguish drone signals from background activity. The proposed approach combines STFT-based time-frequency representation, 2D-DCT for energy compaction and dimensionality reduction, and statistical feature extraction from segmented DCT coefficients for final classification. The main contributions of this article are as follows:

• First, a lightweight and interpretable pipeline for binary drone-versus-background detection is proposed, combining STFT spectrogram representation, 2D-DCT-based energy compaction, segmented low-order DCT analysis, and statistical feature extraction.

• Second, multiple DCT coefficient segmentation strategies are investigated to identify the most discriminative regions for signal separation.

• Third, experimental evaluations are conducted using several machine learning models, including XGBoost, K-Nearest Neighbors (KNN), and Support Vector Machines (SVM), to verify the effectiveness and robustness of the proposed method for drone-versus-background detection.

The remainder of this paper is organized as follows: Section 2 presents the theoretical background and the proposed method; Section 3 reports and discusses the experimental results; and Section 4 concludes the paper and outlines directions for future research.

2. PROPOSED METHOD

2.1. Overall Framework

The proposed framework for UAV signal detection is organized as a lightweight machine learning pipeline that transforms raw complex in-phase and quadrature (IQ) data

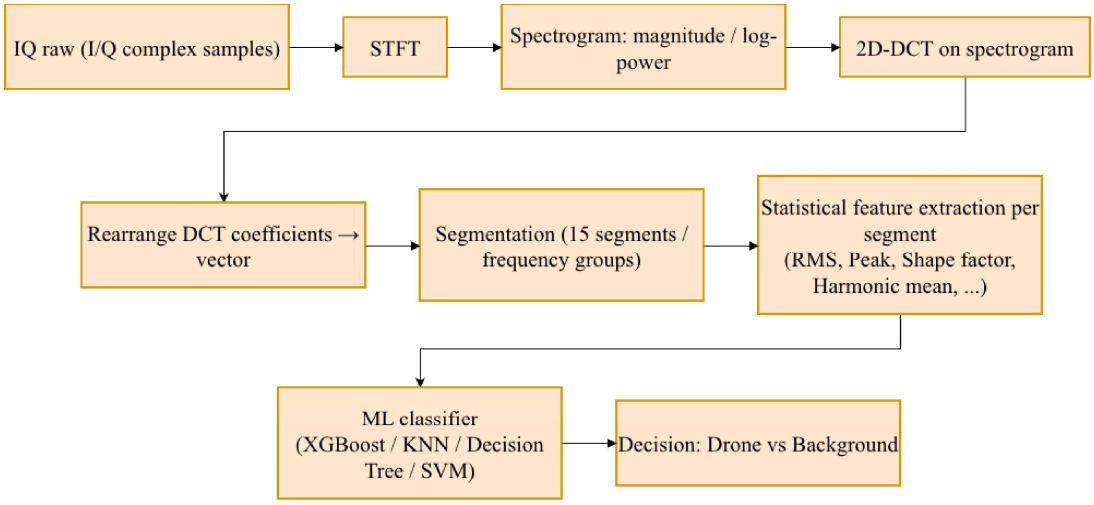

Figure 1. Overall framework of the proposed UAV signal detection method

into a binary detection label, namely drone or background. The objective of this framework is to efficiently process high-dimensional radio-frequency signals by transforming raw time-domain data into a compact and discriminative feature representation. The overall framework of the proposed method is illustrated in Figure 1.

The pipeline consists of several sequential stages. First, raw IQ samples are converted into time-frequency spectrograms using the Short-Time Fourier Transform (STFT) to capture the non-stationary characteristics of the signal. Next, a two-dimensional Discrete Cosine Transform (2D-DCT) is applied to the spectrogram images to exploit its energy compaction property. This transformation produces a more compact representation by concentrating most of the signal energy and structural information into the low-order coefficients. The resulting DCT coefficient matrix is then flattened into a vector and subjected to coefficient segmentation. This segmentation strategy enables a targeted investigation of different coefficient regions to identify those carrying the most discriminative information for signal separation. From these segmented coefficients, thirteen statistical features, including Root Mean Square (RMS), peak value, shape factor, and harmonic mean, are extracted to construct the final feature vector. These feature vectors are then used as inputs to machine learning classifiers, such as XGBoost, KNN), and Support Vector Machines (SVM), to perform binary drone-versus-background detection. Each stage of the framework is described in the following subsections.

2.2. STFT-Based time frequency representation

The raw input data of the proposed system consist of complex in-phase (I) and quadrature (Q) signals. Directly extracting discriminative patterns from one-dimensional time-domain IQ data is often challenging for machine learning models because of the high computational burden and the sensitivity of

the signal to segmentation. In addition, UAV transmissions are inherently non-stationary, meaning that their spectral characteristics, such as frequency hopping and signal variations, change over time. To address these issues, the Short-Time Fourier Transform (STFT) is employed to provide a suitable time-frequency representation. This transformation reveals the distribution of signal spectral components over time, thereby facilitating the capture of dynamic characteristics that are essential for reliable detection.

In the STFT process, the complex signal sequence is divided into short segments, and each segment is multiplied by a sliding window function to extract local spectral information. The discrete STFT of a complex signal sequence at time index n and frequency is mathematically defined as follows [18]:

$$X(n,\omega_k) = \sum_{m=-\infty}^{\infty} x[m]\omega[n - m]e^{-j\omega_k m} \quad (1)$$

where $\omega[n]$ denotes the window function. This transformation highlights the dynamic features and explicit spectral patterns of UAV transmissions, such as frequency hopping, which are difficult to identify directly in the raw time domain. The transform output is typically converted into a magnitude or log-power representation to form a two-dimensional spectrogram. The resulting spectrogram is then used as the input representation for the subsequent 2D-DCT analysis. In the experiments, the STFT was computed using a Hanning window, an overlap of samples 30%, and an FFT size of 1024. An example of the resulting STFT spectrograms for background and drone signals is illustrated in Figure. 2.

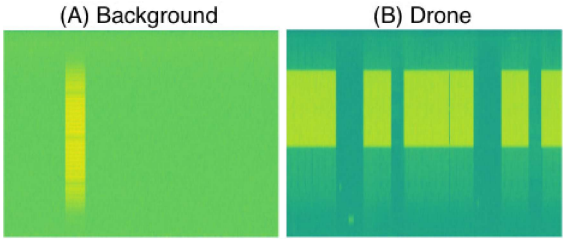

Figure 2. STFT spectrograms of (A) background and (B) drone signals.

2.3. Two-Dimensional Discrete Cosine Transform (2D-DCT)

The input to this stage is the STFT spectrogram matrix obtained from the preceding time–frequency representation process. The proposed method applies a two-dimensional 2D-DCT to these spectrogram images to exploit the well-known energy compaction property of the DCT. For highly correlated data such as spectrograms, 2D-DCT is an effective tool for dimensionality reduction, as it concentrates most of the signal energy and primary structural information into a small number of low-order coefficients, typically located in the top-left region of the transform matrix.

For each experimental setting, the selected coefficient range is divided into 15 consecutive segments of equal length. Specifically, the 300,000-coefficient setting yields 15 segments of 20,000 coefficients each, the first 20,000 coefficients are divided into 15 segments of approximately 1,333 coefficients each, and the first 1,333 coefficients are further divided into 15 smaller segments.

The 2D-DCT is commonly implemented by exploiting its separability property, which allows

the transform to be computed through successive one-dimensional operations along the rows and columns of the image. For an N×N matrix f(x,y), the 2D-DCT is mathematically defined as:

$$F(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1} f_x(x)\cos\left(\frac{\pi(2x+1)u}{2N}\right) \quad (2)$$

where u, v = 0, 1, …, N−1 are the discrete frequency indices.

By focusing on the low-order coefficients, the system can emphasize the essential spectral patterns of UAV transmissions while suppressing high-frequency variations that are often caused by noise.

DCT is selected in this work because it offers energy compaction, a fixed transform basis, and low implementation complexity for highly correlated spectrogram data. Compared with PCA and wavelet-based transforms, DCT is more convenient for lightweight coefficient segmentation and statistical feature extraction in the present framework.

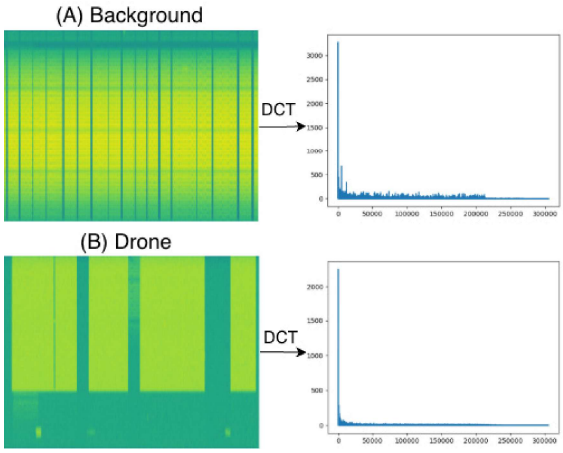

Figure 3. STFT spectrograms and corresponding DCT coefficient distributions for (A) background and (B) drone signals.

After the transformation, the DCT coefficient matrix is flattened into a one-dimensional vector and divided into segments to identify the most informative regions for drone-versus-background separation. This segmentation provides a compact representation and serves as the basis for subsequent statistical feature extraction. An example of the DCT-transformed STFT spectrogram is shown in Figure 3, where most of the signal energy is concentrated in the low-order coefficients, while higher-order coefficients have relatively low amplitudes.

2.4. Statistical Feature Extraction

The input to this stage consists of the partitioned DCT coefficients obtained from the previous segmentation process. For each segment, statistical features are computed to characterize the distribution of the coefficients, thereby mapping high-dimensional data into a more compact feature vector suitable for machine learning analysis. This step enables the extraction of representative spectral characteristics while reducing data dimensionality.

A total of thirteen statistical features are surveyed for each segment to describe the signal properties. Among them, four key features are emphasized in this study: RMS, peak value, shape factor, and harmonic mean. The RMS value represents the average energy level of the signal and is defined as:

$$x_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \quad (3)$$

where x_i denotes the i-th coefficient in a segment of length N. The peak value (x_{pv}) represents the maximum amplitude within the segment:

$$x_{pv} = \max_i x_i \quad (4)$$

To characterize the distribution shape of the coefficients, the shape factor (x_{sf}) is defined as the ratio between the RMS value and the mean of absolute values:

$$x_{sf} = \frac{x_{rms}}{\frac{1}{N}\sum_{i=1}^{N}|x_i|} \quad (5)$$

The harmonic mean (H_m) is used to reflect the influence of smaller coefficient values. To ensure numerical stability, especially when coefficients are close to zero, a small constant $\epsilon > 0$ is introduced to avoid division by zero:

$$H_m = \frac{N}{\sum_{i=1}^{N}\frac{1}{x_i + \epsilon}} \quad (6)$$

These features are computed independently for each DCT segment and then concatenated to form a multi-dimensional feature vector. This aggregated feature vector provides a compact representation of the signal characteristics and is subsequently used as the input to machine learning classifiers for binary drone-versus-background detection

2.5. Machine learning models

The feature vectors obtained from the statistical feature extraction stage are used as inputs to machine learning classifiers. The objective is to perform binary classification to distinguish between drone signals and background activity. To this end, four established models are considered: XGBoost, KNN, SVM, and Decision Tree.

XGBoost is employed as an efficient and scalable gradient boosting framework, well suited for modeling non-linear relationships in tabular feature data. KNN provides a distance-based classification approach and is used as a baseline model for comparison. SVM is adopted for its effectiveness in binary classification problems, particularly when the feature space exhibits separable structures. The Decision Tree model is included to provide classification based on a hierarchical decision structure.

All models are trained on the extracted feature vectors and evaluated using standard performance metrics, including Accuracy, Precision, Recall, and F1-score. This evaluation framework enables a consistent comparison of the classifiers in terms of their generalization capability and stability under varying signal conditions. The performance of these models is presented and discussed in the next section.

3. RESULTS AND DISCUSSION

3.1. Experimental Setup

The experimental evaluation is formulated as a binary classification task to distinguish between drone signals and background activity. The input signals are derived from IQ-baseband recordings of 10 consumer/prosumer UAV platforms, containing RF control and video transmission signals. The recordings were

collected in an isolated anechoic chamber, where the drone was placed on a turntable, the controller was located near the drone, and the measurement antenna was positioned on the opposite side of the chamber. The signals were recorded at center frequencies of 2.44 GHz and 5.8 GHz (5.7 GHz for Yuneec Typhoon H when applicable), with sampling rates of 120 MHz and 200 MHz, respectively. The recording duration was 1 s at 2.44 GHz and 0.5 s at 5.8 GHz. Explicit annotations for SNR range and class distribution are not provided in the source dataset; therefore, these factors are not further quantified in this study [19]. Each recording contains radio-frequency control and video transmission characteristics, including frequency hopping behavior. The raw IQ samples are transformed into time–frequency spectrograms for further processing.

The dataset is divided into training and validation sets for model development and hyperparameter tuning, while an independent test set of 12,354 spectrogram images is used to assess the generalization capability of the proposed method on unseen data. The statistical feature vectors extracted from segmented DCT coefficients are used as inputs to the machine learning models described in Section 2.5, including XGBoost, KNN, Decision Tree, and SVM. Model performance is evaluated using Accuracy, Precision, Recall, and F1-score, and the corresponding results are presented in the following subsection.

3.2. Performance Comparison of Classifiers

To evaluate the effectiveness of the proposed feature extraction approach, five models, namely XGBoost, CNN, SVM, Decision Tree, and KNN, are compared in terms of classification performance. The evaluation is conducted using standard metrics, including Accuracy, Precision, Recall, and F1-score.

Table 1. Performance comparison of classification models

Model	Acc (test)	Acc (valid)	Precision	Recall	F1-score
XGBoost	98.8	99.0	99	99	99
CNN	97.46	97.8	97	97.2	97.2
KNN	97	95.3	95.7	95	96
D - Tree	97.3	96.3	96	96	96
SVM	96.3	96.5	97	97	97

As shown in Table 1, XGBoost achieves the best overall performance, with a test accuracy of 98.8% and a validation accuracy of 99.0%. In addition, it consistently attains the highest precision, recall, and F1-score, all reaching 99%. These results indicate that XGBoost not only provides high classification accuracy but also maintains a strong balance between false positives and false negatives.

The CNN baseline also achieves strong performance, with a test accuracy of 97.46% and a validation accuracy of 97.8%, together with 97% precision, 97.2% recall, and 97.2% F1-score. These results confirm that spectrogram-based deep learning provides an effective benchmark for the binary UAV signal detection task. However, the CNN model remains slightly inferior to XGBoost across all evaluation metrics, indicating that the proposed feature-engineering framework can achieve competitive or better performance

without relying on end-to-end deep feature learning.

Although KNN achieves a comparable test accuracy of 97%, its validation accuracy decreases significantly to 95.3%, indicating limited generalization capability and high sensitivity to variations in the feature space. This behavior suggests that KNN is more influenced by local data distribution and may be less robust in the presence of noise or overlapping feature regions.

Decision Tree and SVM exhibit relatively stable performance, with validation accuracies of 96.3% and 96.5%, respectively. However, their overall performance remains slightly lower than that of XGBoost across all evaluation metrics.

To further examine the classification behavior of the best-performing model, the confusion matrix of XGBoost is presented in Figure 4. It can be observed that 99.20% of background samples are correctly classified, while only 0.08% are misclassified as drone signals. For drone samples, 98.10% are correctly detected, and 1.90% are incorrectly assigned to the background class. These results indicate that XGBoost provides strong discriminative capability for both classes, with very low false-alarm and missed-detection rates. In particular, the high correct classification rates for both background and drone signals confirm that the statistical features extracted from segmented DCT coefficients provide an effective representation for binary UAV signal detection.

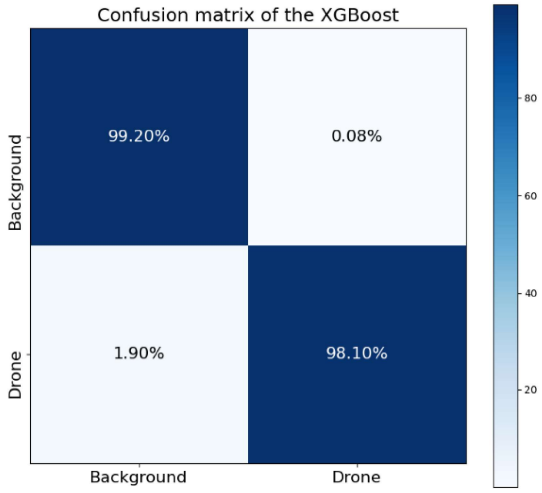

Figure 4. Confusion matrix of the XGBoost model for binary classification of drone and background signals

Overall, XGBoost demonstrates the most stable and consistent performance among the evaluated models. Its boosting mechanism enables it to capture complex relationships in the statistical feature space derived from segmented DCT coefficients, leading to stronger generalization capability for UAV signal detection.

3.3. Analysis of DCT Segmentation

To investigate the distribution of discriminative information in the DCT domain, several segmentation strategies are applied to the flattened DCT coefficient vector. As illustrated in Figure 5, after the 2D-DCT is performed on the STFT spectrogram, the resulting coefficients are rearranged into a one-dimensional vector and partitioned into 15 consecutive segments, denoted from to . In this study, three segmentation ranges are examined, including 15 segments over 300,000 coefficients, 15 segments over the first 20,000

coefficients, and 15 segments over the first 1,333 coefficients. By computing statistical features for each segment individually and evaluating the corresponding detection accuracy, the most informative coefficient regions for drone-versus-background separation can be identified.

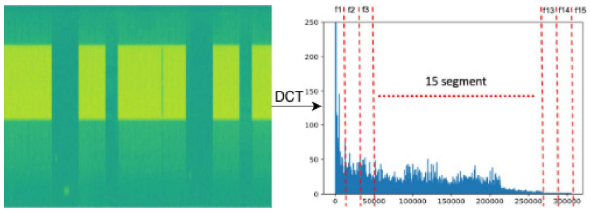

Figure 5. DCT coefficient segmentation into 15 segments

The accuracy comparison of the three segmentation strategies is presented in Figure 6. A clear trend can be observed across all cases. For the segmentation over 300,000 coefficients, only the first segment achieves high accuracy at 97.39%, whereas the remaining segments exhibit substantially lower performance, mostly below 15%. When the analysis is restricted to the first 20,000 coefficients, the first two segments remain more informative, with reaching 98.09% and achieving 76.61%, while the later segments show a clear reduction in discriminative capability. More notably, when the first 1,333 coefficients are divided into 15 segments, many early segments maintain high accuracy, with -all exceeding 94% and several subsequent segments remaining above 80%. This result indicates that the discriminative information is concentrated mainly in the initial portion of the DCT vector.

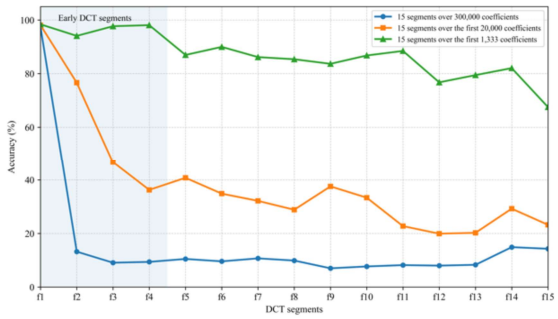

Figure 6. Accuracy of DCT segments under three segmentation strategies

This observation is consistent with the energy compaction property of the 2D-DCT, where most of the signal energy and major structural information are concentrated in the low-order coefficients. Since the early part of the flattened DCT vector corresponds to these low-order components, the first segments preserve the most relevant spectral characteristics for signal separation. Therefore, the segmentation analysis further supports the effectiveness of using low-order DCT coefficients as the basis for statistical feature extraction in lightweight UAV signal detection.

3.4. Computational Cost Analysis

Table 2 presents the runtime comparison between the proposed XGBoost-based framework and the CNN baseline. Since both methods use the same STFT-based spectrogram generation stage, they share an identical preprocessing cost of 80 ms per sample. The proposed method introduces an additional 2D-DCT stage with a small overhead of only 3 ms per sample. However, its model inference time is significantly lower, requiring only 10 ms per sample, compared with 400 ms for the CNN baseline. As a result, the total processing time of the proposed method is 93 ms (0.093 s)

per sample, whereas the CNN baseline requires 480 ms (0.48 s) per sample.

Table 2. Runtime comparison of the proposed method and CNN baseline

Stage	XGBoost (ms/sample)	CNN baseline (ms/sample)
STFT	80 ms	80 ms
2D-DCT	3 ms	nan
Model inference	10ms	400 ms
Total	93ms/0.093s	480ms/0.48s

These results show that, although the proposed framework includes an additional 2D-DCT transformation, its overall runtime remains substantially lower because the XGBoost classifier is much lighter than the CNN model. Therefore, the proposed method achieves a more favorable trade-off between accuracy and computational cost, making it more suitable for lightweight UAV signal detection systems.

4. CONCLUSION AND FUTURE WORK

This paper has presented a lightweight approach for UAV signal detection from raw complex IQ data. The proposed framework combines STFT-based time-frequency representation, 2D-DCT for energy compaction, and statistical feature extraction from segmented DCT coefficients. Experimental results demonstrate that the method is effective for binary classification between drone signals and background activity. Among the evaluated models, XGBoost provides the most reliable performance on both internal and independent test sets. In addition, the segmentation analysis shows that the initial segments of the flattened DCT vector contain the most discriminative information for signal separation, confirming the importance of low-order DCT coefficients.

The proposed method is computationally efficient and suitable for lightweight UAV detection systems. Future work will focus on extending the framework to multiclass UAV classification, supporting real-time deployment, and incorporating more advanced feature extraction or feature selection strategies to improve robustness under different operating conditions.

REFERENCES

[1]. M. H. Rahman, J.-I. Baik, M. A. Aziz, R. Tabassum, M. A. S. Sejan, and H.-K. Song, "1. Cascaded learning empowered classification of UAVs using radio frequency under wireless interference," Alex. Eng. J., vol. 121, pp. 201–212, May 2025, doi: 10.1016/j.aej.2025.02.031.

[2]. I. Alla, S. Yahia, and V. Loscri, "2. TRIDENT: Tri-modal Real-time Intrusion Detection Engine for New Targets," Comput. Secur., vol. 159, p. 104676, Dec. 2025, doi: 10.1016/j.cose.2025.104676.

[3]. L. Xu, R. Shi, and Y. Zhang, "3. A Radio Frequency Sensor-Based UAV Detection and Identification System Using Improved Vision Transformer-Based Model," IEEE Sens. J., vol. 25, no. 13, pp. 23437–23450, Jul. 2025, doi: 10.1109/JSEN.2025.3530937.

[4]. M. Ezuma, F. Erden, C. Kumar Anjinappa, O. Ozdemir, and I. Guvenc, "4. Detection and Classification of UAVs Using RF Fingerprints in the Presence of Wi-Fi and Bluetooth Interference," IEEE Open J. Commun. Soc., vol. 1, pp. 60–76, 2020, doi: 10.1109/OJCOMS.2019.2955889.

[5]. C. J. Swinney and J. C. Woods, "5. Unmanned Aerial Vehicle Operating Mode Classification Using Deep Residual Learning Feature Extraction," Aerospace, vol. 8, no. 3, Mar. 2021, doi: 10.3390/aerospace8030079.

[6]. S. O. De Macedo, M. Caetano, and R. M. Da Costa, "6. Drone detection in airport environments: A literature review," Array, vol. 28, p. 100511, Dec. 2025, doi: 10.1016/j.array.2025.100511.

[7]. B. Taha and A. Shoufan, "7. Machine Learning-Based Drone Detection and Classification: State-of-the-Art in Research," IEEE Access, vol. 7, pp. 138669–138682, 2019, doi: 10.1109/ACCESS.2019.2942944.

[8]. S. Glüge, M. Nyfeler, A. Aghaebrahimian, N. Ramagnano, and C. Schüpbach, "8. Robust Low-Cost Drone Detection and Classification in Low SNR Environments," IEEE J. Radio Freq. Identif., vol. 8, pp. 821–830, 2024, doi: 10.1109/JRFID.2024.3487303.

[9]. Y. Awad and S. I. Al-Dharrab, "9. Malicious UAV Detection Over Rician Fading Channel: Performance Analysis," IEEE Access, vol. 12, pp. 34681–34690, 2024, doi: 10.1109/ACCESS.2024.3373309.

[10]. A. Özkaner and Y. Akça, "10. Mini/Micro UAV detection in the presence of ISM or spurious signals and an experimental application on an SDR," Eng. Sci. Technol. Int. J., vol. 49, p. 101591, Jan. 2024, doi: 10.1016/j.jestch.2023.101591.

[11]. Y. Vasant Ahirrao, R. P. Yadav, and S. Kumar, "11. RF-Based UAV Detection and Identification Enhanced by Machine Learning Approach," IEEE Access, vol. 12, pp. 177735–177745, 2024, doi: 10.1109/ACCESS.2024.3502754.

[12]. R. Nelega, R. V. F. Turcu, B. Belean, and E. Puschita, "12. Radio Frequency-Based Drone Detection and Classification using Deep Learning Algorithms," in 2023 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia: IEEE, Sep. 2023, pp. 1–6. doi: 10.23919/SoftCOM58365.2023.10271669.

[13]. M. Zuo, S. Xie, X. Zhang, and M. Yang, "13. Recognition of UAV Video Signal Using RF Fingerprints in the Presence of WiFi Interference," IEEE Access, vol. 9, pp. 88844–88851, 2021, doi: 10.1109/ACCESS.2021.3089590.

[14]. T. M. Ilyas and B. Mustapha, "14. RF-based UAV classification using CNN in noisy environments," Eng. Res. Express, vol. 7, no. 3, p. 035386, Sep. 2025, doi: 10.1088/2631-8695/ae0872.

[15]. S. Yang, Y. Luo, W. Miao, C. Ge, W. Sun, and C. Luo, "15. RF Signal-Based UAV Detection and Mode Classification: A Joint Feature Engineering Generator and Multi-Channel Deep Neural Network Approach," Entropy, vol. 23, no. 12, p. 1678, Dec. 2021, doi: 10.3390/e23121678.

[16]. T. Delleji and F. Slimeni, "16. RF-YOLO: a modified YOLO model for UAV detection and classification using RF spectrogram images," Telecommun. Syst., vol. 88, no. 1, p. 33, Mar. 2025, doi: 10.1007/s11235-025-01264-4.

[17]. "17. The Discrete Cosine Transform (DCT)," in Digital Image Processing, in Texts in Computer Science. , London: Springer London, 2008, pp. 367–373. doi: 10.1007/978-1-84628-968-2_15.

[18]. D.L.Jones, Digital signal processing: A user's guide (2014).

[19]. M. Vuorenmaa, J. Marin, M. Heino, M. Turunen, and T. Riihonen, "18. Radio-Frequency Control and Video Signal Recordings of Drones." Zenodo, Nov. 05, 2020. doi: 10.5281/ZENODO.4246392.