Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 12
ESAC 2026

A Survey on Deep Learning Methods for UAV Frequency-Hopping Signal Detection

Yanxi Liu

Detroit Green Technology Institute, Hubei University of Technology, Wuhan China

Abstract. To address the security threats posed by unauthorized UAV operations, commonly known as "black flights," traditional frequency-hopping signal detection methods often exhibit low accuracy and poor adaptability in environments characterized by low signal-to-noise ratios and complex electromagnetic conditions. This paper provides a systematic review of deep learning-based methods for detecting UAV frequency-hopping signals, followed by a detailed comparative analysis of three representative approaches. First, we introduce a detection method that synergizes compressed sensing with multi-level deep learning, which achieves the highest reported detection accuracy (≥99.3%) in controlled settings. Next, we describe an improved YOLOX-tiny model that incorporates an attention mechanism and a lightweight architecture, achieving a real-time processing rate of 57 FPS in mobile scenarios with a 40.6% reduction in parameters. Finally, a two-stage anti-interference framework named YOLOv3-CNN, which is designed for complex electromagnetic environments and demonstrates remarkable robustness (≥96% accuracy) under non-line-of-sight and strong interference, is analyzed. The results confirm that deep learning effectively breaks through the limitations of traditional methods in complex settings. Multi-scenario generalization, low-cost hardware adaptation, and integrated multi-task system development emerge as three key directions for future research to pave the way for the technology's transition from lab to field.

Keywords: UAV Frequency-Hopping Signal; Deep Learning; Signal Detection; Anti-Interference; Real-Time Processing.

1. Introduction

Unmanned Aerial Vehicles (UAVs) are aircraft that operate without human pilots, controlled via radio remote control or autonomous programming. Their applications have expanded from high-risk military missions to diverse civilian scenarios [1]. To date, the UAV industry has become a significant driver of socioeconomic growth, with a market scale exceeding 100 billion yuan, demonstrating substantial economic value and development potential [2]. However, the absence of unified industry standards and regulations has resulted in a large number of UAVs operating illegally (“black flights”). The emergence of abuses or illegal intrusions poses serious threats to national security, potentially endangering personal safety and causing property damage, creating multidimensional security risks [3].

Current traditional target detection technologies primarily include infrared detection, radar detection, acoustic detection, and radio frequency (RF) signal-based detection [4]. Existing frequency-hopping signal detection methods can be categorized into three types based on signal processing paradigms: Power spectrum analysis-based methods leverage the time-varying characteristics of frequency-hopping signal power spectra, distinguishing between frequency-hopping and fixed-frequency signals by calculating power spectral density (PSD) and power spectral cancellation [5]. This method is computationally simple but suffers from steep performance degradation in low signal-to-noise ratio (SNR) environments. In urban settings, it is susceptible to interference from Wi-Fi and Bluetooth signals, leading to reduced detection probability and increased false alarm rates. Time-frequency analysis methods employ tools like Short-Time Fourier Transform (STFT), Smoothed Pseudo-Wigner-Ville Distribution (SPWVD), and Extended B-Distribution Modulation (EMBD) to generate time-frequency plots. These plots visually represent frequency variations over time, enabling signal recognition and parameter estimation through image processing and clustering analysis [6]. However, reliance on manually designed image processing workflows results in poor adaptability to complex electromagnetic environments, and parameter estimation accuracy is susceptible to noise interference. Feature extraction and machine learning-based recognition methods extract manual features such as instantaneous characteristics, higher-order cumulants, and spectral features from frequency-hopping signals. These are then classified using algorithms like Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Random Forests (RF). While this approach offers improvements over pure signal processing methods, its performance heavily relies on the quality of manually selected features. It exhibits limited feature discrimination capabilities under low signal-to-noise ratios or complex electromagnetic environments. Furthermore, the feature engineering process is cumbersome, and the methods suffer from insufficient generalization capabilities, making them ill-suited for scenarios involving multiple coexisting signals.

The performance limitations of traditional methods in low SNR and complex electromagnetic environments, coupled with their heavy reliance on manual features or processing workflows, have become bottlenecks for further performance enhancement. Consequently, deep learning-based analysis methods have emerged. Leveraging its strengths in automatic feature extraction and complex data modeling, DL offers new pathways for frequency-hopping signal detection [7]. Unlike traditional methods reliant on manual features, DL models automatically learn multi-level features from raw signals through nonlinear transformations in deep networks-low layers capture local time-frequency characteristics, middle layers integrate global patterns, and high layers abstract semantic features, eliminating dependence on expert knowledge [8]. Currently, deep learning-based frequency-hopping signal detection converges on two core approaches: First, low-sampling-rate schemes combined with compressed sensing (CS) overcome Nyquist sampling constraints, reducing data pressure while maintaining accuracy. Second, lightweight model design balances accuracy and real-time performance through network optimization, enabling embedded deployment. This paper focuses on reviewing these two categories, analyzing their principles and effectiveness.

The primary innovations and contributions of this work are as follows:

- This paper proposes a "scenario-technology-metric" triadic framework. It guides a systematic review by categorizing and contrasting the performance and suitability of methods across controlled, mobile, and strong-interference scenarios, to yield a clear decision map for selection.
- A paradigm shift from singular accuracy to balanced efficiency, robustness, and deployability is traced through three deconstructed solutions: the compressed sensing synergy, lightweight YOLOX-tiny, and two-stage YOLOv3-CNN.
- This work advances the field by framing an engineering-evolution from algorithm-centric to deployment-centric design—directly tackling bottlenecks in computation, cost, and multi-modal fusion to map the route to integrated systems.

2. Related Work

2.1. Convolutional Neural Networks (CNN)

As a pivotal technology category in deep learning, convolutional neural networks excel at automatically identifying relevant features without human supervision. They integrate feature extraction and classification tasks within a unified model framework, achieving the convergence of these two critical functions. CNNs are particularly adept at extracting spatially local correlation features from time-frequency maps [9]. Structurally, this network comprises multiple layers of neurons, including convolutional layers, pooling layers, fully connected layers, activation functions, and an output layer [10].

2.2. Drone Detection Network (DNN)

DNN is a directed acyclic graph composed of multiple computational layers. Its core application scenario is no-fly zones, primarily performing the task of detecting the "presence or absence" of drones—essentially a binary classification problem [11]. It preserves unique and critical information across layers by extracting higher-level abstractions or feature maps from input data [12]. Leveraging its robust nonlinear mapping capability, it directly learns high-dimensional nonlinear features from raw time-domain samples, providing more robust discrimination between frequency-hopping signals and fixed-frequency jammers/noise.

2.3. Multi-Channel Random Demodulation (MCRD) Sampling Structure Based on Compressed Sensing (CS)

Compressed sensing technology transcends traditional sampling theorems, liberating sampling rates from bandwidth constraints and instead making them dependent on signal information content. It exploits signal sparsity to acquire compressed measurements at rates below the Nyquist sampling rate, simultaneously performing sampling and compression. The original signal is then reconstructed through reconstruction algorithms. Drawing inspiration from Random Demodulation (RD) and Modulated Wideband Converter (MWC) approaches, a multi-channel parallel sampling architecture is constructed where each channel functions as an independent RD system. Unlike MWC's periodic mixed signals, MCRD employs random sequences as mixed signals. This simplifies system design by requiring only the generation of random sequences modulated onto different carrier frequencies.

2.4. YOLO Object Detection Algorithm

As a representative end-to-end object detection algorithm, YOLO (You Only Look Once) excels by simultaneously performing object classification and localization within a single forward pass of a deep neural network, achieving compact architecture and high computational efficiency [13]. Its core logic involves partitioning the input image into grid cells, with each cell predicting bounding boxes containing confidence information to detect and localize objects. The algorithm excels in balancing efficiency and accuracy, enabling simultaneous recognition and localization of multiple objects within an image while maintaining real-time detection capabilities.

2.5. Lightweight Network Optimization Techniques

To meet real-time detection demands on embedded platforms, lightweight network optimization techniques have emerged as a research focus in recent years. Core approaches include Coordinate Attention (CA), Deep Separable Convolution (DSC), and the Slim-Neck architecture. The CA mechanism captures long-range positional dependencies through spatial average pooling, enhancing the spatial localization capability of small-scale remote signals in the time-frequency domain and reducing boundary regression errors. DSC decomposes standard convolutions into depthwise and pointwise convolutions, substantially reducing model parameters and computational complexity while preserving feature extraction capabilities; The Slim-Neck architecture addresses feature extraction deficiencies caused by lightweighting through feature reshuffling and residual fusion, achieving a balance between detection accuracy and computational efficiency.

3. Deep Learning-Based Frequency-Hopping Signal Detection

3.1. UAV Detection and Classification Using RF CS and Multi-Level DL

This method proposes an integrated framework combining compressed sensing (CS) signal sampling with multi-stage deep learning (DL) detection and classification. It aims to address the issues of excessive bandwidth, heavy data processing demands, and insufficient accuracy in traditional radio frequency (RF) detection. The core approach utilizes the sparsity of RF signals exchanged between

UAVs and controllers to optimize sampling, while achieving precise identification of UAV presence, type, and flight mode through hierarchical deep learning.

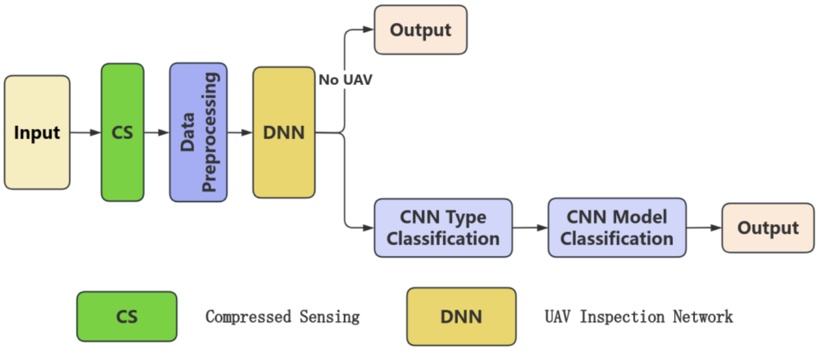


Figure 1. CS-Based Multi-Stage DL Algorithm Flowchart

3.1.1. Core Concept

This method relies on two breakthroughs: First, the RF signals between UAVs and controllers exhibit sparsity within the frequency band after filtering, meeting the conditions for compressed sensing sampling. Therefore, Multi-Channel Random Demodulation (MCRD) replaces Nyquist sampling to achieve simultaneous sampling and compression, reducing bandwidth and data volume. Second, it constructs a multi-stage model integrating DNN and CNN, eliminating manual feature extraction. Through deep learning, it automatically extracts deep signal features, sequentially detecting UAV presence, identifying types, and recognizing flight modes, thereby enhancing detection efficiency and accuracy.

3.1.2. Core Modules and Process

a. Data Sampling and Preprocessing Module

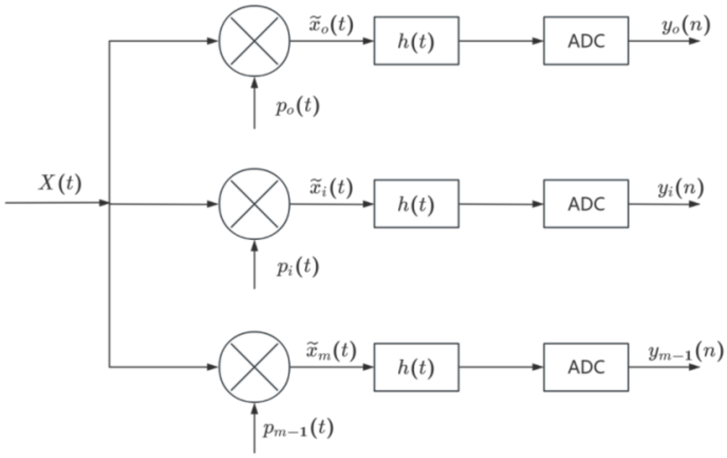


Figure 2. Multi-Channel Random Demodulation Sampling Architecture

Based on compressed sensing theory, the drone RF signal x(t) with bandwidth B is divided into m subbands, each corresponding to an independent random demodulation channel; Frequency shifting is achieved by multiplying the subband signals with random sequences in the time domain. After low-pass filtering, these signals are sampled at a reduced rate below 2B/m, ultimately generating a low-dimensional compressed measurement signal. The measurement matrix is a random Bernoulli matrix, balancing sampling efficiency with signal integrity.

Where the mixed signal pi(t):

$$\mathrm{p}i(t) = 2\varepsilon n \cos(2\pi \frac{i \cdot f}{m} t), t \in [\frac{n}{W}, \frac{n+1}{W}), n = 0,1 \cdots W-1, i = 0,1 \cdots m-1. \tag{1}$$

Where εn takes ±1 with equal probability, and W is the Nyquist rate.

Compressed measurement value y_i(n):

$$\mathrm{yi(n)} = \int_0^{\mathrm{n \cdot Ts}} \mathrm{x}(\tau)\mathrm{d}\tau \mathrm{h}(\mathrm{n \cdot Ts} - \tau)\mathrm{d}\tau = \Phi\mathrm{i} \cdot \mathrm{x(n)} \tag{2}$$

$$\Phi\mathrm{i} = \int_0^{\mathrm{n} \cdot Ts} pi(\tau)h(n \cdot Ts - \tau)d\tau. \tag{3}$$

Where Ts denotes the sampling interval.

The compressed signal undergoes four-step processing: First, zero-centering to remove zero-frequency and offset components; second, calculating the power spectral density (PSD) to quantify signal energy distribution; third, spectral stitching using normalization factor c to eliminate spectral deviations between channels and ensure continuity; fourth, maximum normalization to map PSD values to the 0-1 range, generating model input data.

Calculate the power spectral density Pi(k) for each channel's yi(n):

$$Pi(k) = \frac{1}{N}\sum_{m=0}^{N-1}[yi(n)e^{j\frac{2\pi mk}{N}}]^2 \tag{4}$$

Where N denotes the length of sequence yi(n), and k ≤ n.

Normalization factor c:

$$\mathrm{c} = \frac{\sum_{q=0}^{Q} Pi(k)(N-q)}{\sum_{q=0}^{Q} Pi+1(k)(q)}, q = 0,1\ldots,Q-1 \tag{5}$$

Where N denotes the total number of samples in Pi(k).

b. Multi-stage Deep Learning Detection and Classification Module

Employing a hierarchical progressive logic, specialized networks are designed for different detection tasks. The workflow is as follows:

First, a DNN network performs UAV presence detection: For the binary classification task ‘UAV present/absent’, a 5-layer fully connected DNN is constructed using ReLU and Sigmoid activation functions with MSE as the loss function. Preprocessed RF background signals and UAV RF signals are labeled as 0 and 1, respectively, and 10-fold cross-validation is performed.

Then, a CNN network classifies UAV types and flight modes: For the multi-classification task, a 6-layer 1D convolutional CNN is designed with ReLU and Softmax activation functions and cross-entropy loss. Classifier 0 first identifies the UAV type, followed by Classifiers 1 and 2 identifying flight modes for Bebop and AR drones, respectively.

Final experiments achieved 100% detection accuracy and F1 score, with 99.3% accuracy for 10-class detection including background signals.

3.2. A Lightweight Signal Detection and Recognition Model Based on YOLOX-tiny

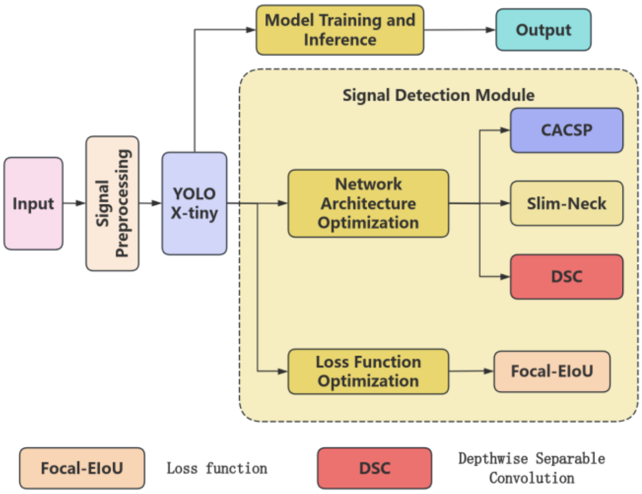


Figure 3. Algorithm Flowchart for Lightweight Detection Based on Improved YOLOX-tiny

The core of this method is to transform UAV signal detection and recognition into a computer vision problem within deep learning. By optimizing the structure of the YOLOX-tiny model and improving its loss function, it achieves efficient detection of UAV signals in complex environments while maintaining lightweight characteristics suitable for embedded deployment [10].

3.2.1. Key Concepts

Addressing challenges in UAV signal detection—such as multiple targets, co-channel interference, and weak signals—this approach prioritizes lightweight design. It maintains detection accuracy while reducing computational load and power consumption. By integrating attention mechanisms to enhance positional feature capture, adopting efficient convolutional structures to streamline model parameters, and optimizing the loss function to resolve sample imbalance, it achieves rapid and precise UAV signal detection suitable for embedded device deployment.

3.2.2. Frequency-Hopping Signal Detection Workflow

a. Signal Preprocessing:

Time-domain data of target signals is precisely extracted from acquired broadband signals via time-frequency filtering. A root raised cosine window is applied to frequency-domain signals, ensuring signal-to-background superposition aligns with real-world conditions.

b. Dataset Construction:

Expand samples using data augmentation techniques such as noise overlay, spectral shifting, and fading channel simulation. Acquire signal pixel coordinates via automated annotation tools to generate grayscale spatiotemporal map datasets for model training.

c. Model Inference:

The enhanced YOLOX-tiny model performs feature extraction and object detection on the spatiotemporal maps, outputting signal categories along with their temporal and frequency domain ranges.

d. Parameter Estimation:

Based on the normalized coordinates output by the model, key parameters are calculated for the drone video transmission signal (center frequency, bandwidth) and the remote control signal (frequency hopping bandwidth, period).

3.2.3. Core Modules

a. Enhanced Network Architecture:

The Coordinate Attention Mechanism (CACSP) module is introduced to strengthen signal location information utilization and reduce boundary regression errors. Deep separable convolutions replace standard convolutions, reducing parameters by 60%. The Slim-Neck structure optimizes the feature pyramid, balancing accuracy and computational efficiency.

b. Loss Function Optimization:

Employed the Focal-EIoU loss function, combining Intersection over Union (IoU) loss, center distance loss, and width-height loss. Weight allocation distinguishes high-quality from low-quality samples, addressing remote control signal detection and sample imbalance issues.

This approach enhances the utilization of spatial features in signal time-frequency plots. The improved remote control signal detection rate increased by 3.76%, while boundary regression error decreased by 1.69%. Model parameters are reduced by 40.6%, computational load decreases by 60%, and frame rate increases from 45 fps to 57 fps, balancing lightweight design with detection speed. While computational cost is reduced by 30%-40%, the feature pyramid enhances multi-scale signal fusion capability, boosting mAP from 83.74% to 87.50%.

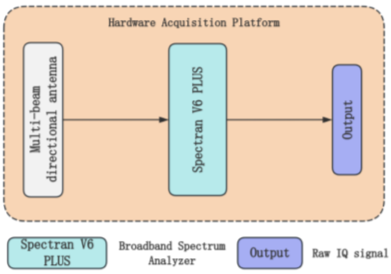
3.3. Two-Stage UAV Detection: RF-Assisted YOLOv3 in Complex Environments

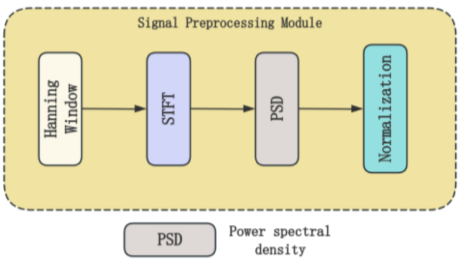
Addressing the challenges of interference susceptibility and performance degradation under non-line-of-sight (NLoS) conditions in complex electromagnetic environments, this study proposes a deep learning-based passive detection and recognition framework utilizing RF spectrum maps. Its core approach leverages the unique time-frequency structural features of frequency-hopping signals (FHS). Through modular design, functions such as signal acquisition, frequency hopping localization, and model recognition are integrated into a unified framework. This achieves synergistic optimization of interference suppression and high-precision recognition while balancing real-time performance and robustness [14].

3.3.1. Overall Framework Design

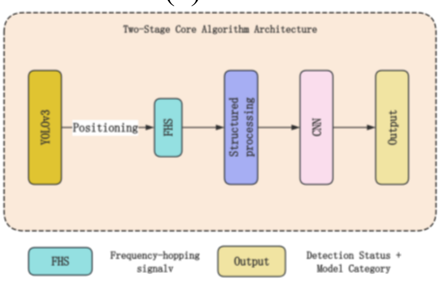
The framework integrates a hardware module for high-fidelity signal acquisition with an algorithmic core that executes detection and recognition via a two-stage deep learning process: specifically, the workflow entails RF signal acquisition, spectrum map generation and preprocessing, YOLO-based FHS regional localization, and CNN-based UAV model classification, forming a closed-loop system of perception, processing, and recognition.

3.3.2. Core Modules and Key Processes





(a) Hardware Acquisition Platform Flowchar (b) Mehod Flowchart of Signal Preprocessing Module



(c) Two-stage core algorithm architecture diagram

Figure 4. Core Framework Flowchart

(1) Hardware Support Module

The system employs a modular hardware architecture comprising a multi-beam directional antenna, a broadband spectrum analyzer (Spectran V6 PLUS), and a computing unit. The multi-beam antenna achieves full horizontal coverage and spatially selective reception, enhancing directional signal capture capabilities. The broadband spectrum analyzer supports an ultra-wide 9kHz-8GHz bandwidth, capturing frequency-hopping signals from UAVs in the 2.4GHz and 5.8GHz bands at a 1THz/s scan rate to ensure precise acquisition of short-duration burst signals. The computing unit, equipped with a GPU, handles real-time signal decoding, spectrum plot generation, and deep learning model operations.

(2) Core Algorithm Modules

First, I/Q signals are acquired and combined into a complex time-domain signal. Hann windowing is applied to suppress spectral leakage. Subsequently, a short-time Fourier transform (STFT) generates a time-frequency spectrum plot. Finally, the signal is converted into a power spectral density (PSD) in dBm units, undergoes normalization and enhancement, and provides high-quality input for subsequent models.

Utilizes a customized YOLOv3 architecture, with the core objective being the precise localization of FHS regions within complex spectrograms. This module takes preprocessed spectrum plots as input. It extracts time-frequency texture and frequency-hopping pattern features through the Darknet backbone network. By integrating multi-scale information via feature pyramids combined with upsampling, residual connections, and spatial pyramid pooling, it ultimately predicts FHS regions at different scales through three detection heads. It outputs bounding boxes and confidence scores, enabling effective separation of target signals from environmental interference and reducing computational load for subsequent recognition.

Addressing the limitation of traditional methods relying on global spectra—which are susceptible to interference from substantial non-target noise and signal overlap— this module takes YOLO-localized FHS regions as input. It first extracts the time-frequency coordinates of frequency-hopping points via block detection algorithms. Following normalization and structural encoding to preserve geometric features like hopping intervals and trajectory patterns, the data enters a five-layer convolutional neural network. Hierarchical structural features are extracted through ReLU activation and max pooling, culminating in multi-class UAV model classification via a Softmax layer [15].

This approach resolves interference issues in complex electromagnetic environments, overcomes traditional spectrum classification's reliance on global texture, and enhances robustness in non-line-of-sight (NLoS) scenarios. It ensures millisecond-level detection speed, employs a lightweight architecture to reduce computational overhead, and completes a single detection-recognition cycle in approximately 34 milliseconds.

4. Result Analysis

4.1. Comparative Analysis of Results

Table 1. Comparison of Core Metrics

Comparision dimensions	Detection accuracy rate	Recognition accuracy	Processing speed
[3]	100%	99.6%	-
[7]	87.50%	87.50%	57 [FPS]
[14]	≥96%	≥96%	34 [ms /T]

Based on the comparison table of core metrics above, the results can be analyzed as follows:

4.1.1. Accuracy Comparison.

[3] Achieves the highest accuracy of ≥99.3% by combining compressed sensing with multi-stage deep learning, but relies on fixed datasets and does not address complex electromagnetic interference scenarios, making it more suitable for high-precision detection in controlled environments; [14] demonstrates outstanding performance in NLoS scenarios: accuracy improves by 15% over traditional methods, though overall accuracy slightly trails [3]; [7] exhibits lower average precision than the former two but is designed for multi-target and co-frequency interference scenarios, featuring low parameter estimation error: video transmission error <1%, better aligning with practical deployment needs for detection and parameter extraction. In summary: [3] offers optimal performance, [7] provides balanced capabilities, and [14] emphasizes interference resistance.

4.1.2. Real-time performance and efficiency comparison.

[7] demonstrates significant lightweight advantages with a 57FPS frame rate and 0.14-second processing time on Jetson Nano. Through deep separable convolutions and model acceleration, it meets lightweight application demands such as handheld devices; [14] achieves near-equivalent real-time performance at 34ms/cycle; [3] lacks explicit frame rate but reduces data transmission pressure through low sampling rate design, making it suitable for long-range networked detection. However, its staged training approach increases complexity and challenges embedded deployment. Thus, [7] excels in lightweight optimization, while [14] balances performance for fixed scenarios.

4.1.3. Interference Resistance and Environmental Adaptability.

[14] Designed for complex electromagnetic environments, it effectively filters WiFi and Bluetooth interference via FHS signal localization. In NLoS scenarios, it achieves 14.9%–20.1% higher accuracy than baseline models, significantly outperforming traditional methods. [7] employs time-frequency filtering and data augmentation to maintain effective detection at -10dB SNR, demonstrating strong resistance to co-channel interference and a 3.76% recognition rate improvement over baseline models. [3] did not focus on verifying anti-interference performance, as the dataset contained minimal interference. Therefore, it can be concluded that [14] is optimal in NLoS scenarios with the most significant performance improvement in special scenarios, while [7] performs best under co-channel interference and also enhances overall performance.

4.2. Application Scenario Analysis

Table 2. Adaptive Scenarios and Scenario Requirement Analysis

Literature	Suitable Scenarios	Scenarios Requirements
[3]	Minimal-Interference Controlled Environments: Enclosed Testing and Long-Range Industrial Monitoring.	High-precision, long-range detection across multiple tasks via low bandwidth to reduce costs in a non-interfering environment.
[7]	Lightweight mobile scenarios: handheld security detectors and urban vehicle patrol systems.	Rapid Multi-UAV Detection in Dense Interference: Balancing Anti-Jamming and Flexibility.
[14]	High-security fixed zones. Such as airports and military bases.	Specialized EMI/Shielding Mitigation for Critical Airspace Security.

5. Future Development Directions and Research Outlook

5.1. Deep Optimization of Multi-scenario Adaptability and Anti-interference Capabilities

Current technologies demonstrate excellent performance in specific scenarios but lack cross-scenario robustness. Future efforts should integrate the core strengths of existing technologies to optimize signal preprocessing for complex environments like urban canyons and adverse weather conditions. Design adaptive anti-interference algorithms while expanding concurrent detection capabilities for multiple UAVs to resolve signal overlap issues from co-frequency multiple targets. This will achieve a breakthrough from single-scenario adaptation to full-scenario coverage.

5.2. Engineering Implementation of Low-Cost, Lightweight Technologies

Existing technologies face hardware dependency or deployment barriers, often requiring specialized spectrum analyzers and posing challenges for embedded deployment. Future efforts should focus on developing low-cost hardware, optimizing power consumption and form factor for compatibility with mobile terminals like handheld devices and UAVs. Concurrently, advancing model compression techniques—through quantization, pruning, and other methods—will enhance real-time performance, reduce deployment costs, and drive the transition from lab validation to large-scale engineering applications in civilian security and campus management.

5.3. Multi-Task Fusion and Intelligent Interoperability System Development

Current technologies predominantly focus on the isolated “detection-identification” phase, failing to form a closed-loop functional system. Future developments should integrate countermeasure systems to generate targeted jamming strategies, creating a closed-loop process from detection to interference. Additionally, multimodal data fusion should be incorporated, supplementing radar and visual signals to overcome the limitations of single RF signals in scenarios involving obstructions or weak signals. This will ultimately establish an intelligent UAV control system tailored to high-end security requirements such as airports and large-scale events.

6. Conclusion

This paper addresses the practical challenges of detecting UAV frequency-hopping signals in complex environments by systematically reviewing and comparing three mainstream deep-learning approaches. Our work confirms the superior efficacy of deep learning in overcoming low-SNR and complex interference while establishing a clear “scenario-driven” adaptation paradigm: optimal accuracy in controlled settings is achieved through compressed sensing and multi-level deep learning; an efficiency-accuracy balance in mobile scenarios is enabled by a lightweight YOLOX-tiny model; and exceptional robustness in high-interference environments is provided by a two-stage anti-interference architecture. Therefore, future efforts should prioritize overcoming the key bottlenecks of cross-scenario generalization, low-cost deployment, and multi-modal integration to transition this technology from discrete algorithm validation into real-time, reliable integrated systems—a critical step toward solidifying its role in intelligent airspace security.

References

[1] Cai C. Recognition and Intercept Method of Unmanned Aerial Vehicle (UAV) TT&C Signals [D]. Chengdu : Recognition and Intercept Method of Unmanned Aerial Vehicle (UAV) TT&C Signals, 2017.

[2] Cui, K. (2021). A new frequency hopping signal detection of civil UAV based on improved K-means clustering algorithm. IEE.

[3] Mo Y, Huang J, Qian G. Deep learning approach to UAV detection and classification by using compressively sensed RF signal [J]. Sensors, 2022, 22 (8): 3072.

[4] Qiao Q, Yin T, Zhang Q, et al. Research on UAV classification techniques based on time-frequency features [J]. Journal of Ordnance Equipment Engineering, 2025, 46 (7).

[5] Ming S T, Xu G M. A Frequency Hopping Communication Signal Detection Method for Drones [J].Wireless Communication Technology, 2025, 34 (01): 39-43.

[6] Yan X, Han B, Su Z, et al. SignalFormer: Hybrid Transformer for Automatic Drone Identification Based on Drone RF Signals [J]. Sensors, 2023, 23 (22): 9098.

[7] Ma R P.Research on Intelligent Detection and Identification System of UAV Signal Based on Deep Learning [D]. Zhengzhou University, 2023. DOI: 10.27466/d.cnki.gzzdu.2023.001123.

[8] Li J J. Research on UAVCommunication Signal Detection and Parameter Estimation [D]. Xian: Xidian University, 2019.

[9] Alzubaidi L, Zhang J, Humaidi A J, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions [J]. Journal of big Data, 2021, 8 (1): 53.

[10] Guo Q Q. Pedestrian Detection in UAV Scene Based on Convolutional Neural Network [D]. Dalian University of Technology, 2019.

[11] LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision [C] // Proceedings of 2010 IEEE international symposium on circuits and systems. IEEE, 2010: 253-256.

[12] Li G, Hari S K S, Sullivan M, et al. Understanding error propagation in deep learning neural network (DNN) accelerators and applications [C] // Proceedings of the international conference for high performance computing, networking, storage and analysis. 2017: 1-12.

[13] Jiang P, Ergu D, Liu F, et al. A Review of Yolo algorithm developments [J]. Procedia computer science, 2022, 199: 1066-1073.

[14] Zhu G, Briso C, Liu Y, et al. An Intelligent Passive System for UAV Detection and Identification in Complex Electromagnetic Environments via Deep Learning [J]. Drones, 2025, 9 (10): 702.

[15] Li M, Hao D, Wang J, et al. Intelligent identification and classification of small UAV remote control signals based on improved yolov5-7.0 [J]. IEEE Access, 2024, 12: 41688-41703.