A Comprehensive Review on Intelligent Drone Detection and Tracking Systems: Deep Learning and Evolutionary Approaches

Rahul Kishore Jha*, Sudhanshu Vashistha**, Dr. Hukum Chand Saini***

*M.Tech Student, Department of Computer Science and Engineering, Jagan Nath University, Jaipur, Rajasthan, India

** Assistant Professor, Department of Computer Science and Engineering, Jagan Nath University, Jaipur, Rajasthan, India

** Associate Professor, Department of Computer Science and Engineering, Jagan Nath University, Jaipur, Rajasthan, India

Abstract:

The increasing proliferation of drones, or Unmanned Aerial Vehicles (UAVs), has introduced both opportunities and challenges in civilian and defense domains. While UAVs support diverse applications such as surveillance, delivery, and environmental monitoring, their misuse poses significant security and privacy threats. Intelligent drone detection and tracking systems are therefore essential for ensuring controlled and safe airspace. This paper presents a comprehensive review of recent advancements in drone detection and tracking, emphasizing the role of Deep Learning (DL) models—particularly Convolutional Neural Networks (CNNs) and Evolutionary Algorithms (EAs) such as Genetic Programming (GP) for optimization. The study explores the evolution of detection frameworks, multimodal sensing technologies, and hybrid learning models integrating DL and evolutionary optimization. Furthermore, key challenges, datasets, evaluation metrics, and future research directions are discussed to guide ongoing developments in intelligent UAV surveillance systems.

Keywords: UAVs, CNN, Genetic Programming, Machine Learning, Deep Learning, Drone Detection.

1. Introduction

The rapid advancement of Unmanned Aerial Vehicles (UAVs), commonly known as drones, has revolutionized various sectors including agriculture, logistics, disaster management, surveillance, environmental monitoring, and defense operations. These autonomous or semi-autonomous aerial systems offer unparalleled flexibility, mobility, and cost efficiency in comparison to traditional manned aircraft [1]. As a result, UAVs have become indispensable tools for real-time data collection, delivery services, and tactical applications across both civilian and military domains [2], [3].

However, the increasing accessibility and commercialization of UAV technology have also introduced several security and privacy challenges. Unauthorized or malicious drone activities such as illegal surveillance, smuggling, airspace violations, and potential terrorist exploitation pose significant risks to public safety and critical infrastructure [4]. These growing concerns have led to a heightened focus on the development of intelligent drone detection and tracking systems to ensure safe and regulated airspace management [5].

Traditional detection approaches, primarily based on radar, radio frequency (RF), acoustic, and optical sensors, often encounter

limitations when identifying small-sized, low-altitude, and fast-moving drones in complex or cluttered environments. Factors such as reduced radar cross-section (RCS), environmental noise interference, and limited visibility can significantly degrade detection performance [6]. To overcome these challenges, recent research has increasingly adopted artificial intelligence (AI)-driven methodologies, leveraging the power of Deep Learning (DL) and Computer Vision (CV) techniques [7], [8].

Among these, Convolutional Neural Networks (CNNs) have emerged as the most promising deep learning framework due to their exceptional ability to automatically extract spatial and contextual features from visual or multispectral input data. CNN-based systems such as YOLO (You Only Look Once), Faster R-CNN, and EfficientDet have demonstrated remarkable success in real-time UAV detection and classification tasks, offering improved accuracy and robustness under varying environmental conditions [9], [10].

In parallel, evolutionary computing techniques including Genetic Algorithms (GAs), Genetic Programming (GP), Particle Swarm Optimization (PSO), and Differential Evolution (DE) have been employed to optimize deep learning architectures and hyperparameters. These bio-inspired algorithms simulate natural evolutionary processes to achieve automatic network tuning, enhance model convergence, and improve generalization performance without exhaustive manual trial-and-error [11], [12].

By integrating deep learning frameworks with evolutionary optimization strategies, researchers have developed hybrid intelligent systems capable of adaptive learning, enhanced detection precision, and real-time performance. Furthermore, sensor fusion techniques, which combine multiple data sources such as radar, acoustic, and optical inputs, have further improved detection reliability in dynamic environments.

This review paper presents a comprehensive analysis of recent advancements in intelligent drone detection and tracking systems. It systematically explores the role of deep learning architectures, evolutionary optimization techniques, and multimodal sensor fusion approaches that collectively contribute to enhanced system accuracy, robustness, and adaptability. The discussion also highlights key research challenges, publicly available UAV datasets, performance evaluation metrics, and potential future directions for developing next-generation UAV detection and tracking frameworks.

2. Overview of Drone Detection and Tracking Systems

Intelligent drone detection and tracking systems play a critical role in safeguarding airspace by identifying, localizing, and monitoring the movement of Unmanned Aerial Vehicles (UAVs). The general workflow of such systems consists of three main stages: data acquisition, feature extraction, and classification/tracking. Each stage contributes uniquely to the accurate and reliable identification of UAVs under varying operational and environmental conditions.

2.1 Detection Modalities

Drone detection relies on diverse sensing modalities, each offering unique advantages and limitations depending on environmental context, range, and object characteristics. The most common detection modalities include radar, acoustic, radio frequency (RF), and vision-based systems. Recently, sensor fusion approaches have emerged as a promising direction to overcome individual modality constraints.

- **Radar-Based Detection Systems**

Radar-based systems utilize electromagnetic wave reflections to detect and estimate the position, velocity, and size of aerial objects. They are particularly effective for long-range detection and all-weather operation, making them widely used in military and air traffic management applications. However, small commercial drones typically have a low radar cross-section (RCS), which reduces radar reflectivity and makes them difficult to distinguish from birds or background clutter. To address this issue, micro-Doppler signatures and machine learning-based radar

signal analysis have been employed to enhance target classification accuracy.

- **Acoustic-Based Detection Systems**

Acoustic detection systems identify drones by analyzing the acoustic signatures generated by their propellers and motors. These signatures can be captured using arrays of microphones and processed through algorithms such as Fast Fourier Transform (FFT) or Mel-frequency cepstral coefficients (MFCCs) to extract distinguishing frequency patterns. While this technique is cost-effective and suitable for short-range detection, it is highly susceptible to environmental noise, wind interference, and urban sound pollution. Additionally, the effective detection range is typically limited to a few hundred meters.

- **Radio Frequency (RF)-Based Detection Systems**

RF-based detection techniques exploit the wireless communication signals transmitted between drones and their controllers. By analyzing signal parameters such as frequency, modulation, and transmission patterns, RF-based systems can identify and classify specific UAV models. These systems are energy-efficient and capable of real-time operation; however, their performance degrades when dealing with autonomous drones operating without active communication links or under signal-jamming conditions. Advanced signal intelligence (SIGINT) approaches and deep learning-based RF fingerprinting are being explored to overcome these limitations.

- **Vision-Based Detection Systems**

Vision-based systems employ optical and infrared cameras to visually detect and track drones. With the integration of artificial intelligence (AI) and computer vision (CV), these systems can accurately recognize UAVs based on shape, motion, and texture features. Deep learning architectures such as You Only Look Once (YOLO), Faster R-CNN, and Single Shot MultiBox Detector (SSD) have demonstrated superior performance in real-time UAV detection tasks. Moreover, infrared (IR) and thermal imaging enhance visibility under low-light or nighttime conditions.

Despite these advantages, vision-based systems may struggle under adverse weather, occlusion, or illumination variability, necessitating the use of hybrid or fusion-based approaches.

- **Sensor Fusion Approaches**

Sensor fusion combines data from multiple sensing modalities such as radar, acoustic, RF, and visual sources to achieve robust and reliable detection under diverse conditions. By integrating complementary information, fusion systems mitigate the limitations of individual sensors. For instance, radar offers long-range detection, while optical sensors provide fine-grained object recognition. Advanced fusion frameworks employ Bayesian inference, Kalman filters, and deep multimodal learning models to aggregate heterogeneous data streams effectively. Sensor fusion is increasingly regarded as a cornerstone of intelligent UAV surveillance, particularly in dynamic and cluttered environments.

2.2 Tracking Approaches

Once a drone is detected, tracking aims to continuously estimate its position, trajectory, and motion parameters across successive time frames. Accurate tracking is crucial for threat assessment, interception, and flight path prediction.

- **Traditional Tracking Algorithms**

Classical tracking algorithms such as Kalman Filters (KF) and Particle Filters (PF) have been extensively used for UAV motion estimation.

- The Kalman Filter assumes linear motion and Gaussian noise, making it suitable for tracking drones moving at predictable velocities.
- The Particle Filter, on the other hand, handles nonlinear and non-Gaussian motion dynamics, offering greater flexibility in complex flight scenarios.

Other methods, including optical flow analysis, estimate pixel motion between video frames to determine drone velocity and direction.

- **Deep Learning-Based Tracking**

Recent advances in deep learning have significantly enhanced object tracking

performance. Models such as DeepSORT (Deep Simple Online and Real-Time Tracking) combine convolutional neural networks with motion estimation algorithms to achieve real-time, multi-object tracking. Similarly, YOLOv8-Track and SiamRPN++ (Siamese Region Proposal Network) leverage feature correlation and similarity matching between consecutive frames, enabling robust tracking even under partial occlusion or rapid drone maneuvers.

Hybrid systems combining traditional filtering techniques with deep neural networks have also been developed to achieve high precision with reduced computational complexity. These frameworks allow for adaptive learning, occlusion recovery, and long-term UAV tracking in dynamic outdoor environments.

3. Deep Learning Techniques in Drone Detection

The advent of deep learning (DL) has fundamentally transformed the field of UAV detection and tracking. Unlike traditional machine learning methods that rely on handcrafted features, deep learning models are capable of automatically learning hierarchical representations from raw input data such as images, videos, or sensor signals. These models excel at identifying subtle patterns, spatial dependencies, and motion dynamics that characterize UAV behavior in complex environments. In particular, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformer architectures have demonstrated state-of-the-art performance in drone detection and tracking applications.

A. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) have emerged as the foundational architecture for visual-based UAV detection due to their exceptional capability to extract and process spatial and contextual features from high-dimensional image data. By performing convolutional operations across image pixels, CNNs can identify edges, textures, and shapes that distinguish UAVs from other aerial objects such as birds, planes, or background clutter.

Several CNN-based architectures have achieved notable success in real-time drone detection tasks:

- YOLO (You Only Look Once) family (YOLOv5, YOLOv8) offers high detection accuracy with superior speed, making it suitable for real-time surveillance and counter-drone applications.
- Faster R-CNN employs a region proposal network (RPN) to generate candidate bounding boxes and refine them through classification, achieving superior precision in multi-scale drone detection.
- EfficientDet, based on compound scaling and efficient feature pyramids, delivers a balance between accuracy and computational efficiency, suitable for embedded or edge devices.

Furthermore, transfer learning using pre-trained CNNs (e.g., ResNet, VGG16, or MobileNet) has enabled the adaptation of large-scale vision models to drone datasets with limited samples. These networks, when fine-tuned on drone-specific imagery, have significantly improved the detection robustness under varying lighting and weather conditions.

B. Recurrent Neural Networks (RNNs) and LSTMs

While CNNs are effective for extracting spatial information from individual frames, they do not inherently capture temporal dependencies across video sequences. To address this limitation, Recurrent Neural Networks (RNNs) and their advanced variants, such as Long Short-Term Memory (LSTM) networks, are employed for temporal modeling and motion prediction in UAV tracking.

LSTMs maintain a memory of past frames, enabling the model to learn temporal patterns in drone trajectories, such as speed variations, direction changes, and hovering behavior. This temporal awareness enhances the accuracy of multi-frame drone tracking and motion-based classification, especially in dynamic environments where frame-to-frame variations are subtle. In hybrid frameworks, CNNs are often used to extract spatial features, which are then fed into RNN/LSTM networks for

sequential learning, forming CNN–LSTM architectures widely used in UAV behavior recognition and trajectory forecasting.

C. Transformer Models

In recent years, Transformer-based architectures have redefined the landscape of computer vision tasks, including UAV detection and tracking. Unlike CNNs that rely on local convolutional filters, Vision Transformers (ViTs) employ self-attention mechanisms to model global dependencies between image regions. This allows them to effectively capture contextual information and long-range relationships, which are crucial for distinguishing UAVs from visually similar background objects.

Variants such as the Swin Transformer and DETR (DEtection TRansformer) integrate hierarchical feature representations and attention-based localization, offering improved performance in cluttered or occluded scenes. Transformers also exhibit superior adaptability in multi-object tracking (MOT), where attention mechanisms dynamically focus on the UAV's region of interest, even in cases of partial occlusion or background noise. Moreover, transformer-based hybrid models that fuse CNN feature extractors with attention layers have achieved state-of-the-art detection accuracy with reduced false alarms.

D. Multimodal Deep Learning

Modern UAV detection frameworks increasingly rely on multimodal deep learning, which integrates multiple data sources such as visual, acoustic, radar, and RF signals to achieve robust and context-aware detection. Each modality contributes unique discriminative information:

- Visual data provide detailed appearance and motion cues.
- Acoustic signals capture drone-specific sound patterns.
- RF data offer communication signal characteristics.
- Radar provides distance and velocity measurements.

Multimodal fusion networks employ deep architectures that learn joint representations from these heterogeneous inputs. Early fusion methods combine data at the input level, while late fusion integrates features or decision outputs from multiple modalities. Recent works adopt attention-based fusion models that dynamically weigh the contribution of each modality based on environmental conditions. This strategy significantly enhances detection reliability in real-world scenarios affected by occlusion, background clutter, or signal loss.

4. Evolutionary Algorithms in Drone Detection

Evolutionary Algorithms (EAs) are bio-inspired optimization techniques that simulate the process of natural selection. They are particularly well-suited for addressing the complex, nonlinear, and high-dimensional optimization problems encountered in deep learning model design and training. In drone detection, EAs have been effectively utilized to optimize neural network architectures, feature selection, and hyperparameter tuning, thereby improving model accuracy, convergence rate, and generalization.

A. Genetic Algorithm (GA) Optimization

Genetic Algorithms (GAs) mimic the evolutionary processes of selection, crossover, and mutation to explore large solution spaces. In the context of UAV detection, GAs are commonly used to optimize hyperparameters such as learning rate, batch size, number of convolutional layers, and activation functions in CNNs. By iteratively evolving populations of potential solutions, GAs enable models to achieve superior performance without manual tuning. Furthermore, GA-based optimization can be applied to feature selection, ensuring that only the most discriminative features are retained, which enhances model interpretability and reduces computational overhead.

B. Genetic Programming (GP)

Genetic Programming (GP) extends the principles of GAs by evolving computer programs or mathematical expressions that represent model architectures. Unlike GAs that optimize predefined parameters, GP can

autonomously design novel CNN structures or decision rules for UAV detection. This automated architecture search reduces human bias and accelerates the discovery of optimal deep learning configurations. In drone tracking systems, GP has been employed to evolve adaptive feature extraction pipelines and lightweight classifiers, resulting in more efficient and interpretable detection frameworks suitable for deployment on embedded platforms.

C. Particle Swarm Optimization (PSO) and Differential Evolution (DE)

Beyond GA and GP, other evolutionary algorithms such as Particle Swarm Optimization (PSO) and Differential Evolution (DE) have also been integrated with deep learning models for UAV detection.

- PSO, inspired by the collective behavior of bird flocks, optimizes parameters by iteratively updating candidate solutions based on their own experience and that of neighboring particles. PSO has been utilized to tune network weights, enhance learning stability, and improve convergence in CNN-based detection systems.

- Differential Evolution (DE), on the other hand, relies on vector differences between population members to guide the search process, offering robustness in exploring non-convex and multimodal optimization landscapes. DE-based tuning has proven effective in fine-tuning drone detection thresholds and enhancing classification accuracy under uncertain environmental conditions.

D. Hybrid Deep–Evolutionary Frameworks

Recent research trends focus on hybrid frameworks that integrate deep learning with evolutionary optimization, combining the representation power of neural networks with the search efficiency of evolutionary algorithms. These deep–evolutionary models are capable of self-adaptive learning, wherein the network architecture and parameters evolve dynamically based on environmental feedback or performance metrics. For instance, CNNs integrated with GA or GP-based optimizers have been shown to improve detection accuracy, robustness to illumination changes, and real-time adaptability. Additionally, evolutionary meta-learning strategies are emerging, where EAs evolve the training strategies themselves, enabling models to adapt autonomously to new UAV types or operational scenarios.

5. Challenges and Research Gaps

Despite significant progress, several challenges persist:

- **Small-object detection:** UAVs appear tiny in high-altitude images, reducing detection confidence.
- **Environmental interference:** Weather, lighting, and occlusion degrade model performance.
- **Adversarial robustness:** Deep models can be fooled by camouflage or countermeasures.
- **Computational complexity:** Real-time operation requires lightweight architectures for edge devices.
- **Dataset diversity:** Public datasets are limited in representing varied UAV shapes, sizes, and backgrounds.

6. Conclusion

The evolution of intelligent drone detection and tracking systems marks a crucial advancement toward ensuring secure and regulated airspace amid rapid UAV proliferation. This review highlights how deep learning and evolutionary computing have significantly enhanced UAV surveillance. Deep learning models such as CNNs, RNNs, and Transformers effectively extract spatial-temporal and contextual features, enabling accurate detection and tracking even under challenging conditions like occlusion and variable illumination. Multimodal deep learning, which integrates visual, acoustic, radar, and RF data, further strengthens system robustness and adaptability. In parallel, evolutionary algorithms including GA, GP, PSO, and DE provide powerful optimization mechanisms for deep neural networks, facilitating automatic hyperparameter tuning

and adaptive model evolution. Hybrid deep–evolutionary frameworks that merge these techniques represent a promising direction for developing adaptive, efficient, and real-time UAV detection systems. However, challenges such as small-object detection, adverse weather performance, adversarial robustness, and limited dataset diversity persist. Future research should emphasize lightweight architectures for edge devices, advanced sensor fusion, self-supervised learning, and neuroevolutionary meta-learning for dynamic adaptation. Overall, the synergy between deep learning and evolutionary optimization paves the way for next-generation intelligent UAV surveillance systems that are more accurate, autonomous, and resilient against evolving aerial threats.

.

References

[1]. S. S. Unde, V. K. Kurkute, S. S. Chavan, et al., "The expanding role of multirotor UAVs in precision agriculture with applications, AI integration, and future prospects," Discover Mechanical Engineering, vol. 4, p. 38, 2025.

[2]. G. Pandey, M. Manimaran, A. Mishra, and A. Sapra, "Unmanned aerial vehicles: An inevitable armamentarium in Armed Forces Medical Services," Medical Journal Armed Forces India, 2025.

[3]. V. Semenyuk, I. Kurmashev, A. Lupidi, D. Alyoshin, L. Kurmasheva, and A. Cantelli-Forti, "Advances in UAV detection: Integrating multi-sensor systems and AI for enhanced accuracy and efficiency," International Journal of Critical Infrastructure Protection, vol. 49, 2025.

[4]. L. Xu and Z. Luo, "Anti-UAV detection and identification technology: Fundamentals, methods and challenges," Physical Communication, vol. 71, 2025.

[5]. S. O. de Macedo, M. Caetano, and R. M. da Costa, "Drone detection in airport environments: A literature review," Array, vol. 28, 2025.

[6]. J. Zhu, J. Rong, W. Kou, Q. Zhou, and P. Suo, "Accurate recognition of UAVs on multi-scenario perception with YOLOv9-CAG," Scientific Reports, 2025.

[7]. P. Upadhyay, K. K. Sharma, R. Dwivedi, and P. Jha, "A statistical machine learning approach to optimize workload in cloud data centre," in Proc. 7th Int. Conf. on Computing Methodologies and Communication (ICCMC), 2023, pp. 276–280.

[8]. I. Yadav, V. Shekhawat, K. Gautam, G. K. Soni, and R. Yadav, "Artificial intelligence for cybersecurity: Emerging techniques, challenges, and future trends," in Proc. 3rd Int. Conf. on Sustainable Computing and Data Communication Systems (ICSCDS), 2025, pp. 1176–1180.

[9]. M. K. Jha, "Recent trends and emerging applications of the Internet of Things: Transforming the way we live and work," International Journal of Engineering Trends and Applications (IJETA), vol. 12, no. 4, pp. 239–244, 2025.

[10]. M. K. Jha, S. Agarwal, and V. Kabra, "Artificial intelligence at work: Transforming industries and redefining the workforce landscape," International Journal of Engineering Trends and Applications (IJETA), vol. 12, no. 4, pp. 416–424, 2025.

[11]. H. Mathur and R. Ajmera, "Optimizing service efficiency and safeguarding privacy in distributed computing environments via a MapReduce-powered framework," International Development Planning Review, vol. 47, 2023.

[12]. M. Jha, "A study of ISA server for providing fast Internet access with a single proxy," SGVU Journal of Engineering & Technology, vol. 1, no. 1, pp. 15–18, 2015.

[13]. G. Sharma, N. Hemrajani, S. Sharma, A. Upadhyay, Y. Bhardwaj, and A.

Kumar, "Data management framework for IoT edge-cloud architecture for resource-constrained IoT applications," Journal of Discrete Mathematical Sciences and Cryptography, vol. 25, no. 4, pp. 1093–1103, 2022.

[14]. A. Kumar and N. Hemrajani, "Congestion avoidance in TCP based on optimized random forest with improved random early detection algorithm," International Journal of Image and Graphics, 2024.

[15]. H. Bali and N. Hemrajani, "Attack analysis and designing of quality of service framework for optimized link state routing protocol in MANET," International Journal of Intelligent Engineering & Systems, vol. 11, no. 5, 2018.

[16]. A. Maheshwari and R. Ajmera, "A comprehensive guide to natural language processing in Sanskrit with named entity recognition," in Proc. ACM Int. Conf. on Information Management & Machine Intelligence, 2023.

[17]. J. Xiao, J. H. Chee and M. Feroskhan, "Real-Time Multi-Drone Detection and Tracking for Pursuit-Evasion With Parameter Search," in IEEE Transactions on Intelligent Vehicles, vol. 9, no. 10, pp. 6027-6037, Oct. 2024.

[18]. A. Anastasiou, R. Makrigiorgis, P. Kolios and C. Panayiotou, "Hyperion: A Robust Drone-based Target Tracking System," 2021 International Conference on Unmanned Aircraft Systems (ICUAS), pp. 927-933, 2021.