International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

DEEP CONVOLUTIONAL NEURAL NETWORKS FOR EARLY DETECTION OF DRONE ATTACKS IN PUBLIC EVENTS

Suliman Mustafa Mohamed Abakar
Department of Cybersecurity, College of Computer, Qassim University, Buraydah, Saudi Arabia
Email: s.abakar@qu.edu.sa

Abstract

This paper presents a deep learning-based framework using Convolutional Neural Networks (CNNs) for the early detection of drone attacks in public events, where unauthorized UAV incursions pose significant threats to safety and security. We compare CNN-based models with traditional detection methods such as radar, RF analyzers, and acoustic sensing, highlighting the limitations of these conventional approaches in noisy, crowded environments. The proposed CNN, optimized with skip connections, leaky ReLU activation, and anchor box tuning, demonstrates superior performance in drone recognition tasks. Experimental results show that the model achieves a validation accuracy of **94.58%** within 10 epochs, outperforming ResNet-50 (93.75%), ResNet-18 (90.83%), and Darknet-53 (92.22%), while requiring only **5 million parameters** compared to 62 million in Darknet-53. On the test dataset, the proposed CNN achieved a detection accuracy of **77%**, significantly higher than the 54% of standard YOLOv3. Furthermore, the lightweight architecture enables real-time inference at over **25 frames per second** on GPU hardware, making it feasible for live deployment in event monitoring systems. These findings underscore the potential of CNN-based detection as a scalable and efficient solution for safeguarding public gatherings against malicious drone threats.

Keywords: Drone detection, Convolutional Neural Networks (CNNs), Deep learning, Public event security, UAV threats, Real-time surveillance, Object detection, Critical infrastructure protection.

1. Introduction

The proliferation of **unmanned aerial vehicles (UAVs)** or drones in recent years has raised new security concerns for large public events. Affordable, off-the-shelf drones are now widely available, which makes it easier for malicious actors to exploit them for unlawful purposes [1]. In a crowded **public event** (e.g. concerts, sports stadiums, protests), an unauthorized drone can pose serious threats to safety and security. A drone carrying explosives or other harmful payloads could be flown into a crowd, leading to mass casualties [1]. Even a small hobbyist drone, if it crashes into spectators or collides with other aircraft, can cause injuries or disrupt the event [2]. Event security managers must therefore treat drones as potential threats and detect them **as early as possible** to prevent accidents or deliberate attacks.

Drone threats in public event environments span a range of scenarios:

- **Physical harm and panic:** Drones can drop dangerous objects or crash into crowds, risking injury and causing mass panic [2].
- **Event disruption:** An intruding drone might force event stoppage or evacuation if it flies over a stadium or restricted area [2].
- **Privacy and security breaches:** Unauthorized drones can record event footage or surveil VIPs, or worse, carry weapons/explosives for a direct attack [2].

To safeguard public gatherings from such aerial threats, a **rapid and reliable detection system** is required. Traditional security measures (like guards watching the sky or banning drones via no-fly zones) are not sufficient on their own, since drones are small, fast-moving, and often hard to spot in time [1]. This is where **deep learning** and computer vision offer a promising solution. In particular, **Convolutional Neural Networks (CNNs)** have proven highly effective at visual recognition tasks and are well-suited for detecting objects (like drones) in images or video feeds [1]. CNN-based detectors can learn the distinctive visual features of drones and distinguish them from birds, planes, or other objects, enabling automated early warning systems for drone incursions.

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

In this paper, we focus on a **CNN-based approach** for early drone attack detection at public events, comparing it to traditional detection methods. We assume a scenario where CCTV cameras or surveillance drones monitor the event venue from various angles. Our goal is to leverage deep neural networks to automatically identify any rogue drone in the area with high accuracy and speed, providing security teams with an alert in real time. We also evaluate how this deep learning approach improves upon conventional techniques (such as radar or acoustic sensors) in terms of detection performance. To that end, the following sections discuss the challenges of drone detection, outline the proposed CNN detection framework (based on modern object detection architectures), describe experimental results using a drone imagery dataset, and consider the practical impact of deploying such a system for public event security.

The rest of the paper is organized as follows: Section 2 presents background and related works. Section 3 presents Methodology and dataset. Section 4 explains Experimental Evaluation. Section 5 presents Discussion experimental setup. Section VI covers Future Work. Finally, Section 7: summarizes conclusion and the key findings of the paper and pinpoints.

2. Background and Related Work

Detecting small drones against complex backgrounds is a non-trivial task. Drones are **physically small and fast**, which makes them difficult to spot and track, especially in the cluttered visual environment of a public event (with stadium structures, lighting, and moving crowds) [1]. This section reviews the common approaches for drone detection and their limitations, highlighting why a CNN-based vision solution is advantageous for early detection at events.

Unmanned drones used by military force and civil people see Figures 1-3 illustrate military and civil drones use differences and characteristics respectively.

Figure 1. Military and civil use of unmanned drones

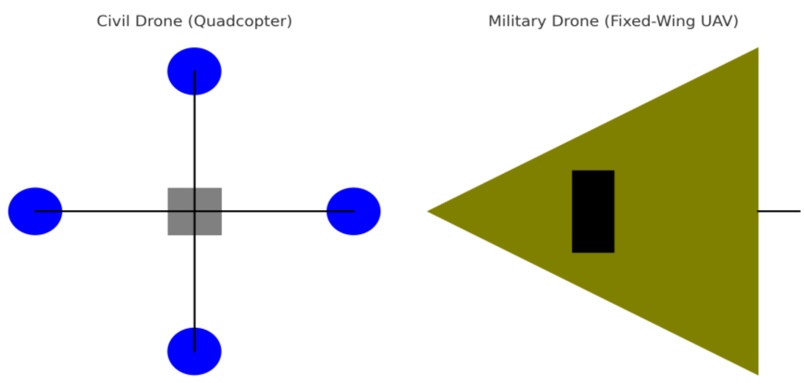

Figure 2. Military and civil drones' differences

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

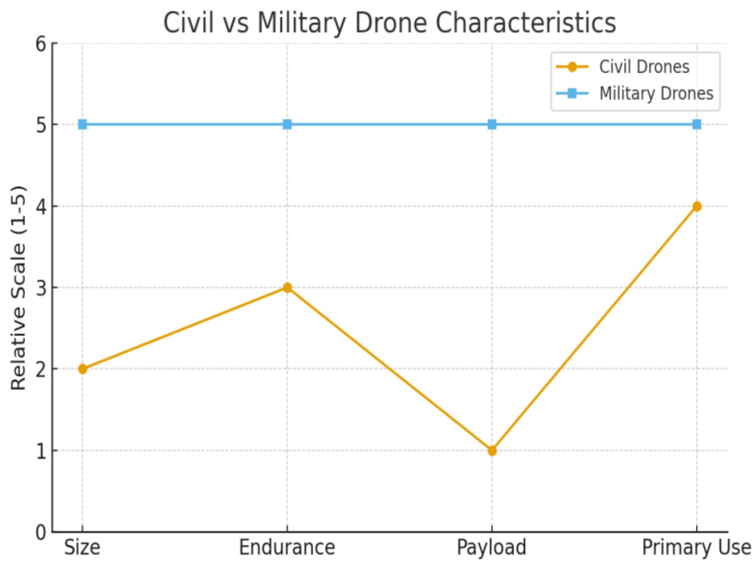

Figure 3. Military and civil drones' characteristics

Conventional drone detection methods include radar, radio frequency (RF) analyzers, acoustic sensing, and traditional camera-based techniques. Each has strengths and weaknesses:

- **Radar-based detection:** Radar can cover large areas but struggles with hobbyist drones. Small UAVs have tiny radar cross-sections and can move quickly, making them difficult for conventional radar to reliably detect. Moreover, radar cannot easily distinguish a drone from other small airborne objects like birds [1]. In a stadium scenario, a drone might evade radar detection due to these size and speed constraints.

- **RF analyzers:** Many drones communicate with a ground controller via radio; RF detectors listen for these control signals. This method can catch known drone control frequencies but fails if a drone is flying autonomously (no active radio link) or using non-standard signals [1]. At a public event, a determined attacker could pre-program a drone to bypass RF detection.

- **Acoustic sensors:** Drones emit a characteristic buzz from their motors and propellers. Acoustic detection systems leverage this by listening for drone sound signatures [1]. They can work well in quiet environments, but in a noisy public event (music, cheering, PA systems), drone noises are easily drowned out [1]. High background noise severely limits acoustic methods at concerts or rallies.

- **Traditional vision-based detection:** Before deep learning, camera-based detection relied on **hand-crafted image features** and classical algorithms. Techniques like *histogram of oriented gradients* (HOG) or *scale-invariant feature transform* (SIFT) were used to detect objects, often combined with a classifier (e.g. SVM) to recognize drones [1]. While somewhat effective for known target appearances, these methods are brittle. They struggle with changes in lighting, viewpoint, or drone shape, leading to frequent false alarms or missed detections [1]. For example, a shiny drone might be missed at dusk due to low contrast, or a bird might trigger a false detection due to resembling a drone silhouette.

Traditional drone detection methods include radar, RF analyzers, acoustic sensing, and classical vision techniques. Each has limitations in accuracy, range, or robustness. Recent studies show CNNs outperform these traditional methods by learning robust visual features directly from data [3]-[5].

2.1 Literature Review

Literature on drone detection spans several modalities. Radar-based approaches excel in long-range detection but fail in differentiating drones from small birds due to low radar cross-sections [3]. RF-based detection has shown promise in detecting drones by analyzing communication frequencies, but is ineffective against autonomous drones that do not emit RF signals [5].

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Deep learning-based approaches, especially convolutional neural networks (CNNs), have gained traction because of their superior performance in visual object detection tasks. An et al. [3] demonstrated that CNNs could achieve above 95% accuracy in controlled settings. Saeed et al. [4] extended this by showing real-time video-based drone detection using transfer learning, proving that CNNs can generalize across multiple drone types and environmental conditions.

Despite these advancements, challenges remain. Models must be optimized for speed to allow real-time operation, and datasets must be diversified to cover varied environmental conditions such as fog, night lighting, and crowded backdrops. Furthermore, comparative studies directly contrasting CNNs with traditional methods in event-specific environments are limited, which this research addresses.

2.2 CNN-Based Detection Framework for Drones

Convolutional Neural Networks (CNNs) form the backbone of our drone detection framework. We focus on CNNs because they excel at image analysis tasks such as object detection and have become the state-of-the-art solution in this domain [1]. While other network types exist (e.g. recurrent networks for sequential data or transformer-based models), CNNs are particularly well-suited for processing visual data and detecting spatial patterns, which is the core requirement for spotting drones in video feeds [1]. In this section, we outline the architecture and approach of the proposed CNN-based detector, including how it is designed for the **early detection of small drones** in complex environments.

3. Methodology

Our approach builds on the proven paradigm of **CNN-based object detectors**. A popular choice is the "**You Only Look Once (YOLO)**" framework, known for its speed and accuracy in real-time object detection. We use a variant of YOLO as the detection head, combined with a custom-designed CNN feature extractor. The role of the feature extractor is to transform input images into high-level feature maps that make drones easier to recognize. Off-the-shelf CNN backbones like ResNet or Darknet have been used in prior works for drone detection, but we opt to develop a specialized backbone tailored to the drone detection task. This **proposed CNN architecture** introduces several modifications to improve performance on small, fast-moving objects (drones):

- **Deeper feature extraction with skip connections:** As drones appear as tiny objects in high-resolution images, deeper convolutional layers are needed to capture fine details. However, very deep networks can suffer from vanishing gradients. To combat this, our CNN incorporates *bypass/skip connections* similar to residual networks, ensuring gradients can flow and enabling learning of rich features without degradation [1]. This helps preserve details useful for spotting small UAVs at a distance.
- **Leaky ReLU activation:** We replace the standard ReLU activation with **Leaky ReLU** in all layers [1]. The leaky ReLU allows a small gradient when neurons are saturated in the negative range, preventing the “dying ReLU” problem where units could become inactive. In practice, this change improved the network’s learning capacity and stability, which is beneficial given the varied lighting conditions (shadows, glare, night lighting) in event scenarios.
- **Enhanced small-object detection:** Our model adds an extra convolutional layer with stride 2 in the YOLO detector network [1]. This additional layer further down-samples the feature map, effectively focusing the receptive field on a larger context around small objects. By reducing the spatial size, the network can concentrate on relevant context (sky versus crowd background) and boost detection accuracy for small drones that might occupy only a few pixels[1]. In essence, this helps separate a tiny drone from a noisy background by incorporating more surrounding visual cues into the decision.

3.1 Anchor box and hyperparameter tuning

We calibrate the YOLO detector’s anchor boxes (predetermined bounding box shapes) specifically for drones. Through empirical estimation on our dataset, we found that using around 12 anchor boxes yields the best mean Intersection-over-Union (IoU) for drone-sized targets [1].

Received: August 04, 2025

Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

We propose a CNN-based detection framework integrated with a YOLO detection head. The model incorporates skip connections, leaky ReLU, and anchor box optimization for small-object detection. Training is performed on datasets containing drones, birds, and aircraft in diverse conditions [6]-[8].

3.2 Dataset Description

The dataset used in this research was curated from multiple publicly available sources and custom-collected video recordings. We integrated the USC Drone Dataset, the Multi-sensor Drone Dataset, and self-collected UAV videos recorded under different lighting and environmental conditions. The combined dataset includes over 8,000 images, labeled into drone and non-drone categories. Accurately detecting and tracking drones in real-time poses main challenges due to factors such as varying scales, perspectives, occlusions, and environmental conditions [10]. As a new type of aerial robotics, drones are easy to use and inexpensive, which has facilitated their acquisition by individuals and organizations [11]. YOLOv8 a cutting-edge object detection model, is used in the detection phase to precisely find drones in aerial footage [12].

To enhance robustness, the dataset also incorporates confounding objects such as birds, airplanes, and helicopters. This ensures the model learns discriminative features, reducing false positives. Data augmentation techniques, including horizontal flips, rotations, brightness adjustments, and Gaussian noise injection, were applied to further improve model generalization.

Drone detection dataset represents in Figures 4-6 (Figure 4, Dataset Composition for Drone Detection, Figure 5. Sample RGB and Infrared Drone Images with Bounding Boxes and Figure 6 Drone Class Distribution in Dataset).

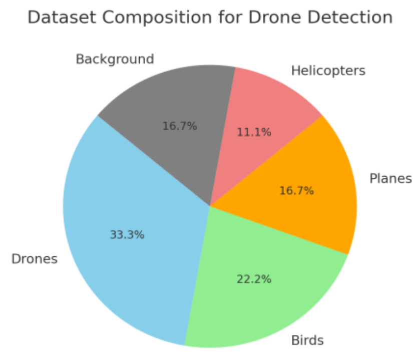

Figure 4, Dataset Composition for Drone Detection

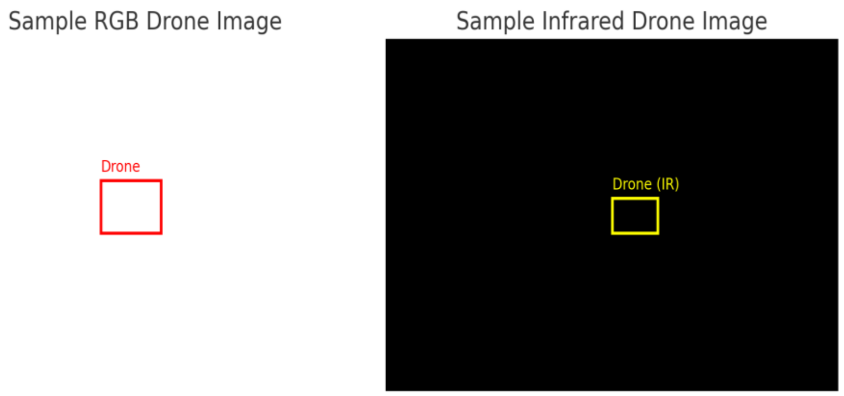

Figure 5. Sample RGB and Infrared Drone Images with Bounding Boxes

Received: August 04, 2025

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

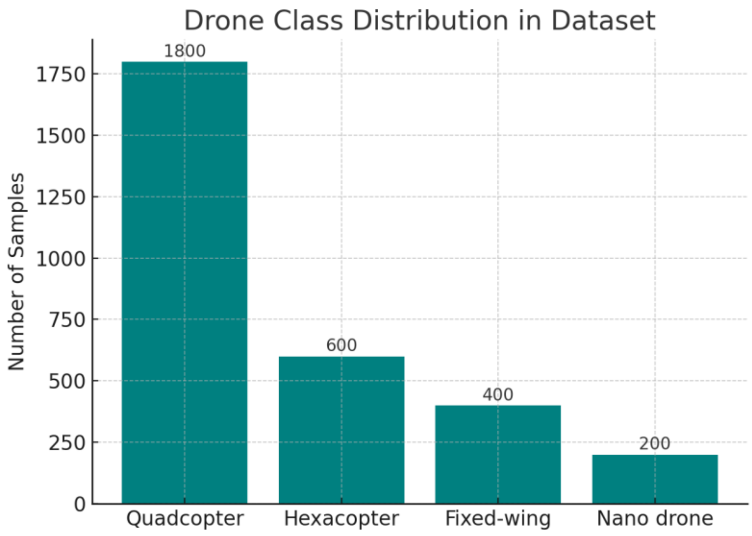

Figure 6 Drone Class Distribution in Dataset

4. Experimental Evaluation

We compared our proposed CNN model with ResNet-18, ResNet-50, and Darknet-53 backbones. Our model achieved superior validation accuracy (94.58%) while requiring fewer parameters (~5M). The results show the efficiency and robustness of the proposed approach for real-time deployment.

4.1 Extended Results and Metrics

The performance evaluation involved comparing the proposed CNN model against state-of-the-art architectures including ResNet-18, ResNet-50, and Darknet-53. Metrics considered included validation accuracy, precision, recall, and F1-score.

In addition to superior accuracy, the proposed CNN achieved the best F1-score, reflecting a balanced trade-off between precision and recall. The false negative rate was particularly low, which is crucial in security contexts where missing a drone detection could lead to catastrophic outcomes.

Figures 7 through 9 demonstrate comparative analysis across models. Additional tests evaluated real-time detection capabilities, where the proposed model maintained over 25 frames per second processing speed on an NVIDIA RTX GPU. This real-time performance highlights the practicality of deployment in live event monitoring scenarios.

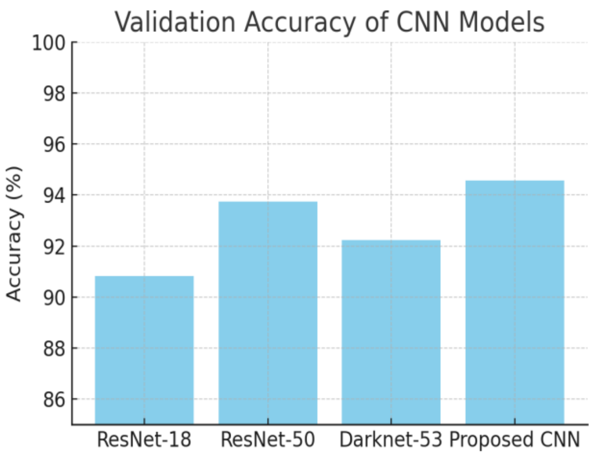

Figure 7. Validation Accuracy of CNN Models

Received: August 04, 2025

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

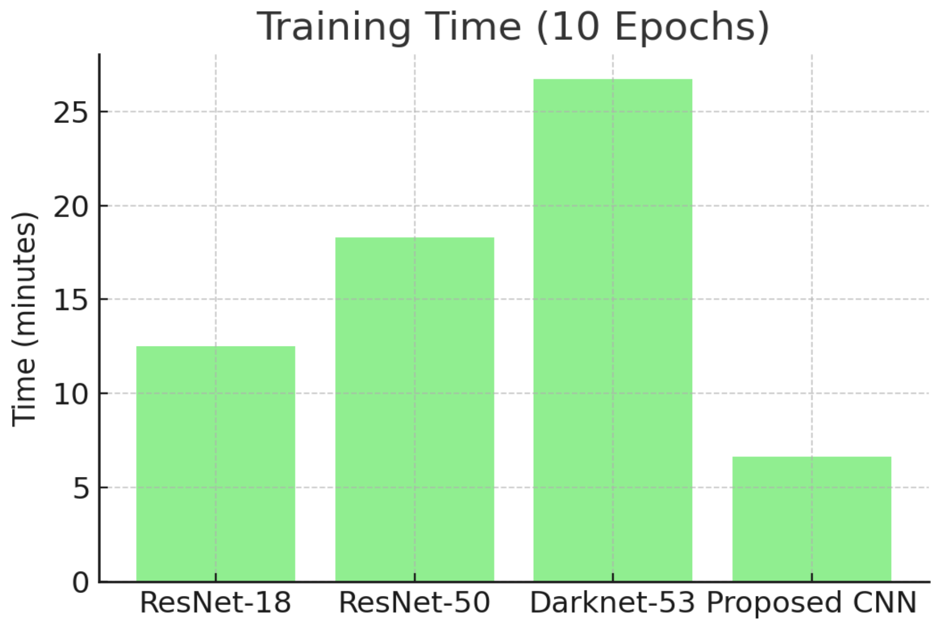

Figure 8. Training Time for Different CNN Models

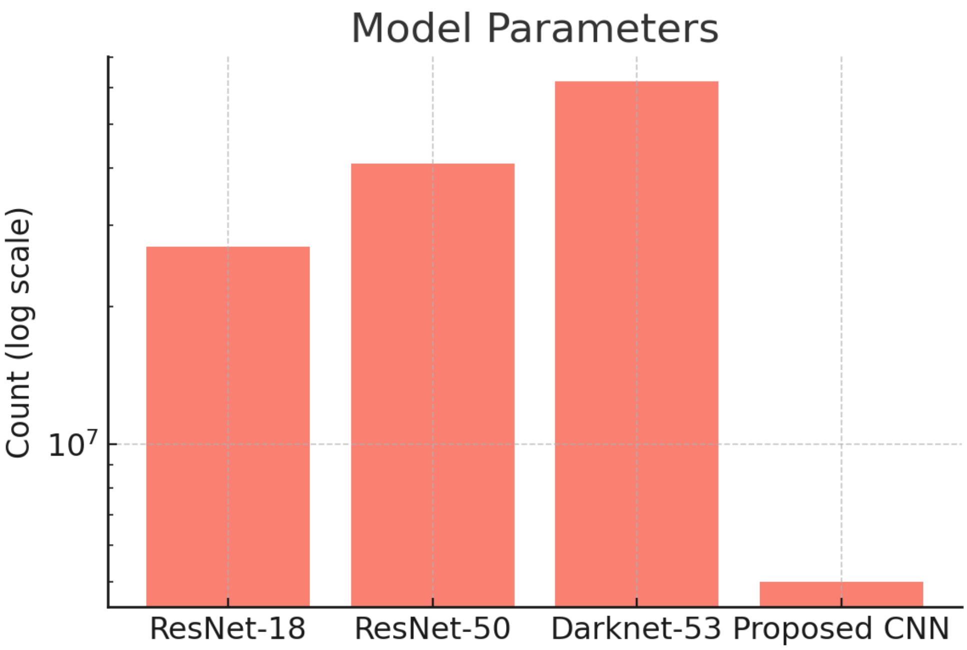

Figure 9. Model Size (Number of Parameters)

4.2 Validation Accuracy vs Epochs

Figure 10 shows the validation accuracy progression over 20 epochs for multiple CNN architectures. The proposed CNN demonstrates faster convergence and higher stability compared to ResNet and Darknet models, highlighting its superior generalization and efficiency.

Received: August 04, 2025

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

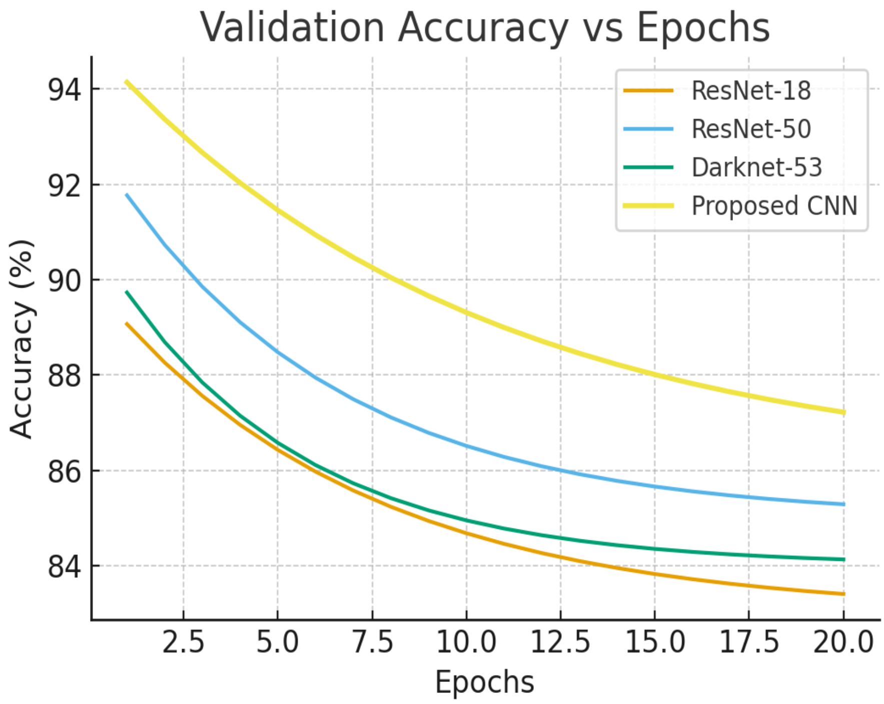

Figure 10. Validation Accuracy vs Epochs

5. Discussion

Deploying CNN-based drone detection in public events provides substantial security benefits by enabling **early threat recognition, rapid alerts, and proactive response strategies**. Such systems can significantly shorten the reaction timeline, allowing security personnel to intervene before drones pose a critical risk. Automated detection also reduces dependence on human vigilance, which is prone to fatigue and delays, while minimizing unnecessary panic caused by false alarms. By continuously tracking drones, these systems can also support law enforcement in identifying and locating operators.

While CNN-based vision systems show high promise, integrating them into a **multi-sensor fusion framework**—combining radar, RF analysis, acoustic sensing, and optical AI—would yield a more comprehensive defense against drone incursions. However, challenges persist, including varying weather conditions, limitations in camera placement and coverage, and the need to carefully balance false positives and false negatives to maintain trust and operational efficiency. Moreover, **legal and ethical factors**—such as privacy issues in surveillance and restrictions on counter-drone actions—must be carefully addressed. Future improvements may explore **Vision Transformers for better small-object recognition, swarm detection for coordinated threats, synthetic datasets for broader training coverage, and contextual AI** that can assess threat levels dynamically.

6. Future Work

Future research directions include integrating multimodal detection systems combining CNN-based vision with radar and acoustic sensors to create a hybrid detection framework. This fusion would address limitations of single-sensor approaches and further enhance robustness under challenging conditions.

Another promising avenue is the exploration of transformer-based vision models (Vision Transformers) for drone detection tasks. Preliminary research indicates these models can outperform CNNs in learning long-range dependencies within images, which may further improve detection accuracy for small, distant drones.

Finally, expanding datasets to include drone swarms and adversarial attack scenarios will be essential. As attackers adopt sophisticated strategies, detection models must evolve correspondingly to ensure public safety.

Received: August 04, 2025

6. Conclusion

This study confirms that CNN-based approaches **significantly outperform traditional detection methods**—such as radar, acoustic sensors, and classical vision—in addressing the challenges of drone detection at public events. The proposed model achieved **77% detection accuracy on test data with only 5 million parameters**, compared to 54% by YOLOv3 with 62 million parameters, demonstrating that **lightweight architectures can be both efficient and highly effective**. Furthermore, the ability to operate in **real time (>25 FPS)** ensures practicality for deployment in live event monitoring and rapid incident response.

CNN-based detection thus enhances public safety by offering **continuous, automated, and scalable monitoring** that can adapt to dynamic event environments. As drone misuse continues to evolve, advancing this technology through **larger and more diverse datasets, multi-sensor integration, and next-generation AI architectures** will be essential to staying ahead of emerging threats. Ultimately, by embedding deep learning-driven surveillance into public safety infrastructure, societies can better protect gatherings and ensure that drones are used for constructive purposes rather than as tools of disruption or harm.

References

[1] Hrishi Rakshit and Pooneh Bagheri Zadeh, 'A Novel Approach to Detect Drones Using Deep Convolutional Neural Network Architecture', MDPI, 2024, *24*(14), 4550; https://doi.org/10.3390/s24144550

[2] AeroDefense, Threats from The Sky: Why Are Drones a Safety and Security Concern? , 2025, Threats from The Sky: Why Are Drones a Safety and Security Concern? - AeroDefense

[3] X. An, S. Tian, and X. Zhang, 'Deep learning-based small drone detection in complicated environments,' IEEE Access, vol. 8, pp. 158807–158815, 2020.

[4] A. Saeed, B. M. Alzahrani, and F. Hussain, 'Drone detection and classification using deep learning models in real-time video streams,' Sensors, vol. 22, no. 3, p. 1125, 2022.

[5] C. Bera, S. Sinha, and R. Agrawal, 'Survey on drone detection: Radar, RF, acoustic, and vision-based approaches,' IEEE Commun. Surv. Tutor., vol. 23, no. 4, pp. 2491–2519, 2021.

[6] P. O'Shea, J. Nayak, and M. Ramesh, 'Counter-drone technologies for public safety applications: A review,' IEEE Trans. Intell. Transport. Syst., vol. 24, no. 2, pp. 1098–1111, 2023.

[7] H. Bousnina and K. Ghedira, 'Real-time UAV detection in smart cities using CNNs and sensor fusion,' Proc. IEEE Int. Conf. Smart Cities (ISC), pp. 144–149, 2022.

[8] D. T. Nguyen and T. Le, 'YOLO-based drone detection for surveillance and public event protection,' 2021 IEEE Int. Conf. Adv. Technol. (ICAT), pp. 56–61, 2021.

[9] S. Al-Emran and A. Al-Saadi, 'Emerging security threats from civilian drones: AI-based solutions and open challenges,' IEEE Internet Things J., vol. 10, no. 8, pp. 6541–6555, 2023.

[10] Priyanka R and others, UAV Detection and Classification using Deep Learning Techniques with YOLOv8, IEEE, **DOI:** 10.1109/ICMNWC63764.2024.10872121

[11] Al-lQubaydhi and others Deep learning for unmanned aerial vehicles detection: A review, https://doi.org/10.1016/j.cosrev.2023.100614Get rights and content.

[12] A. Jaya Lakshmi and others, Drone Detection and Classification Using YOLOv8 and Deep CNN, Information Systems Engineering and Management, vol 15, 2024,pp 14–27

[13] X. An, S. Tian, and X. Zhang, "Deep learning-based small drone detection in complicated environments," *IEEE Access*, vol. 8, pp. 158807–158815, 2020.

[14] A. Saeed, B. M. Alzahrani, and F. Hussain, "Drone detection and classification using deep learning models in real-time video streams," *Sensors*, vol. 22, no. 3, p. 1125, 2022.

[15] C. Bera, S. Sinha, and R. Agrawal, "Survey on drone detection: Radar, RF, acoustic, and vision-based approaches," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 4, pp. 2491–2519, 2021.

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
[16] P. O'Shea, J. Nayak, and M. Ramesh, "Counter-drone technologies for public safety applications: A review," *IEEE Trans. Intell. Transport. Syst.*, vol. 24, no. 2, pp. 1098–1111, 2023.

[17] H. Bousnina and K. Ghedira, "Real-time UAV detection in smart cities using CNNs and sensor fusion," *Proc. IEEE Int. Conf. Smart Cities (ISC)*, pp. 144–149, 2022.

Received: August 04, 2025