AI-Driven Distributed Multi-Sensor Fusion for Real-Time Drone Detection in Urban Airspace

Neno Ruseno, Enrique Puertas, and Aurilla Aurelie Arntzen Bechina

Abstract— **The increasing density of unmanned aerial systems (UAS) in urban low-altitude airspace introduces significant safety and security challenges, particularly for detecting non-cooperative drones in environments where RADAR (Radio Detection and Ranging) deployment is impractical. This paper presents a distributed, artificial intelligence (AI)-enabled multi-sensor surveillance framework integrating visual, acoustic, and radio frequency (RF) sensing through weighted decision-level fusion.**

Each sensing modality is processed using dedicated deep learning models, while a decision-level fusion mechanism combines predictions based on confidence scores and reliability weights. The modular architecture enables asynchronous communication through a publisher–subscriber paradigm, supporting distributed deployment and resilience under partial sensor degradation.

The system is evaluated through both laboratory experiments and simulated urban environments with varying complexity. Results demonstrate that the proposed framework achieves over 90% detection precision, maintains false positive rates below 10%, and supports real-time processing exceeding 50 Hz. Furthermore, the fusion strategy effectively mitigates performance degradation in individual sensing modalities, particularly under noisy or obstructed conditions.

These results highlight the potential of AI-based distributed sensor fusion systems as a scalable and cost-effective solution for real-time drone surveillance in smart urban airspace, contributing to resilient monitoring within emerging U-space ecosystems.

I. Introduction

The rapid expansion of unmanned aerial systems (UAS) operations in low-altitude airspace is transforming the surveillance and security requirements of air traffic management (ATM). Forecasts associated with European U-space implementation anticipate a substantial increase in drone traffic density in urban and peri-urban environments over the coming decade, driven by logistics, inspection, infrastructure monitoring, and emergency response services [1, 2]. As operational density increases, so does the probability of non-conforming, malfunctioning, or malicious drone activity, particularly in proximity to restricted areas such as airport perimeters, critical infrastructure, governmental facilities, and densely populated urban zones.

Unlike conventional cooperative aircraft operating within certified Communications, Navigation, and Surveillance (CNS) frameworks, many small UAS operate without certified transponders and may intentionally disable broadcast identification mechanisms. Consequently, preventing unauthorized drone incursions into restricted airspace requires surveillance capabilities capable of detecting non-cooperative targets in real time. This security dimension introduces a significant operational challenge: ensuring resilient monitoring of sensitive airspace volumes where conventional RADAR (Radio Detection and Ranging) infrastructure is economically prohibitive, operationally impractical, or geographically constrained.

Recent advances in artificial intelligence (AI) and machine learning have enabled significant improvements in drone detection across vision-based, acoustic, radio-frequency (RF), and RADAR technologies [3]. RADAR-based systems provide robustness and long-range detection but involve high cost, spectrum management constraints, and deployment complexity, limiting their scalability in dense urban settings. Vision and acoustic systems offer affordable alternatives but degrade under adverse illumination, occlusions, and environmental noise. RF-based detection enables identification of control links but may fail in congested spectral environments or when drones operate autonomously without active communication.

These limitations highlight a critical research gap in the design of distributed AI-based surveillance systems that can integrate heterogeneous sensing modalities while maintaining robustness under partial sensor degradation and dynamic environmental conditions. In particular, there is a need for architectures that combine multi-sensor data fusion, real-time processing, and scalable communication mechanisms suitable for deployment in smart urban airspace systems.

To address these challenges, this paper proposes a distributed multi-sensor drone detection framework based on artificial intelligence and decision-level fusion. The system integrates visual, acoustic, and RF sensing modalities, where each sensor is processed by dedicated deep learning models. A meta-level fusion module combines the outputs using a weighted aggregation mechanism based on prediction confidence and sensor reliability. The architecture adopts a publisher–subscriber communication paradigm, enabling asynchronous data exchange and scalable deployment across distributed sensing nodes.

This research is conducted within the SESAR 3 AI4HyDrop project [4], which investigates AI-enabled services for safe and resilient drone operations in restricted and urban airspace. Within this context, the present study focuses specifically on the surveillance-layer challenges associated

Neno Ruseno (corresponding author) and Aurilla Aurelie Artnzen Bechina are with the Department of Science and Industry systems, University of South-Eastern Norway, Hasbergsvei 36, Krona, 3616 Kongsberg, Norway, e-mail: neno.ruseno@usn.no, aurilla.aurelie.arntzen@usn.no

Enrique Puertas is with the Department of Science and Computing, Universidad Europea de Madrid, C. Tajo, s/n, 28670 Villaviciosa de Odón, Madrid, Spain, e-mail: enrique.puertas@universidadeuropea.es

with detecting non-cooperative UAS and preventing unauthorized incursions into protected zones.

The main contributions of this paper are a distributed AI-based multi-sensor fusion architecture for drone detection in urban airspace, a decision-level fusion strategy that improves robustness under partial sensor degradation, a unified deep learning-based processing approach across heterogeneous sensing modalities (visual, acoustic, RF), and a real-time and scalable communication framework based on a publisher–subscriber model;

The remainder of this paper is organized as follows. Section II reviews related work on cooperative and non-cooperative drone surveillance approaches. Section III presents the proposed distributed multi-sensor architecture. Section IV describes the methodological framework and evaluation setup. Section V discusses performance results and resilience implications, and Section VI concludes the paper with recommendations for future surveillance architecture.

II. Related Works

A. *AI-Based Drone Detection Methods*

Recent research has increasingly adopted AI techniques to address the limitations of traditional sensing approaches. Existing studies can be categorized based on sensing modality including vision, acoustic, RF, and RADAR.

Vision-based detection approaches commonly formulate drone detection as a small-object detection problem. The YOLO (You Only Look Once) family of models has been widely adopted due to its balance between accuracy and real-time performance. Comparative studies demonstrate improved detection performance across YOLO variants, particularly in challenging scenarios with scale variation and motion [5, 6]. Other works explore lightweight architectures or traditional descriptors to reduce computational cost or improve robustness under low-resolution conditions [7, 8]. However, vision-based methods remain sensitive to illumination changes, occlusions, and adversarial perturbations [9].

Acoustic-based detection leverages the characteristic sound signatures of drone motors and propellers. Most approaches rely on time–frequency representations such as Mel-Frequency Cepstral Coefficients (MFCC) combined with machine learning models [10]. While deep learning methods improve performance under controlled conditions, their effectiveness is strongly influenced by environmental noise. Alternative approaches using handcrafted features and classical classifiers offer computational efficiency but may require careful tuning [11, 12].

RF-based detection exploits communication signals between drones and their controllers, offering long-range and non-line-of-sight capabilities. Traditional methods rely on handcrafted spectral features and classical classifiers [13, 14], while recent studies employ deep or hybrid architectures to improve robustness under interference [15, 16]. RF fingerprinting techniques further enable drone identification and classification [17, 18]. Nevertheless, RF-based methods are limited when drones operate autonomously or when signals are weak or absent.

RADAR-based detection provides strong performance in adverse conditions and cluttered environments. Recent work integrates RADAR sensing with deep learning models such as CNNs, LSTM, and Transformer architectures to enhance classification accuracy and robustness [19]. Advances in MIMO RADAR and feature optimization further improve detection capabilities [20, 21]. Hybrid approaches combining RADAR signal processing with machine learning have also demonstrated effectiveness in complex environments [22]. However, RADAR-based solutions require costly infrastructure and complex deployment, limiting their scalability for urban applications.

B. *Multi-Sensor Fusion Approaches and Research Gaps*

To overcome the limitations of individual sensing modalities, recent research has explored multi-sensor fusion techniques. Fusion approaches combining RADAR and vision [23], RF and vision [24], or RF and RADAR [25] demonstrate improved detection accuracy and reduced false alarms. More comprehensive frameworks integrating multiple modalities, including acoustic sensing, further enhance robustness under low signal-to-noise conditions [26].

Despite these advances, several limitations remain from a civilian U-space perspective. First, most studies emphasize detection accuracy without explicitly addressing operational constraints, such as detection latency and real-time responsiveness required for restricted-area protection. Second, many fusion architectures remain RADAR-centric or infrastructure-heavy, limiting scalability for distributed urban deployment. Third, limited attention has been given to resilience under partial sensor degradation, which is critical in dynamic urban environments where sensing conditions vary significantly.

These gaps highlight the need for distributed, non-RADAR multi-sensor architectures that prioritize scalability, resilience, and operational applicability. The present study addresses these challenges by evaluating a modular sensor fusion framework integrating visual, acoustic, and RF sensing for urban airspace protection.

III. Sensor Fusion Framework

To address the identified gaps in Section II, a multi-sensor drone detection framework is developed based on parallel machine learning models and a meta-model that fuses their outputs into a unified decision. The architecture follows a modular and decentralized design, where heterogeneous sensing modalities such as optical, acoustic, and RF, operate as independent agents as shown in Figure 1. Each agent processes sensor-specific data and produces detection outputs concurrently, enabling flexible deployment and scalability across different environments.

The system adopts a decision-level (late) fusion strategy, where each sensing modality independently generates predictions and associated confidence scores. These outputs are then aggregated by a meta-model to produce the final detection result. This approach allows the system to maintain operational capability even when one or more sensors degrade or become unavailable, thereby enhancing robustness in heterogeneous and dynamic urban environments.

Compared to feature-level fusion, the proposed late fusion approach reduces computational complexity and simplifies integration, as individual models can be developed, updated, or replaced independently. The modular design also supports the addition of new sensing modalities without requiring significant changes to the overall architecture.

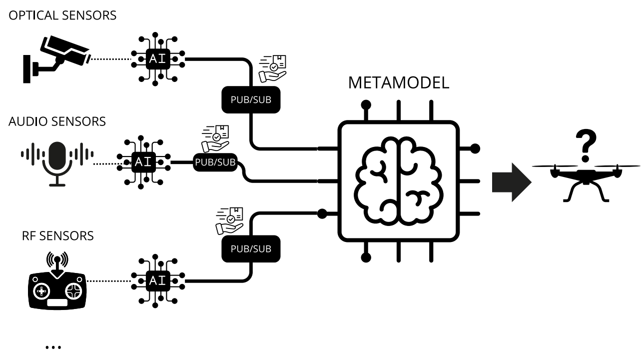


Figure 1. Sensor Fusion Framework

The final prediction P_{final} is obtained through a weighted aggregation mechanism, combining the prediction outputs, confidence scores, and predefined reliability weights of each sensor:

$$P_{final} = \frac{\sum_{i=1}^{n} w_i \cdot c_i \cdot p_i}{\sum_{i=1}^{n} w_i \cdot c_i} \tag{1}$$

where n represents the number of active individual models, p_i is the prediction of model i, c_i is its confidence score, and w_i is its assigned weight which can be dynamically adjusted based on contextual conditions and user-defined operational preferences. For instance, visual sensing may be assigned to higher weights under favorable lighting conditions, while acoustic and RF modalities can be prioritized in low-visibility or nighttime scenarios.

Communication between sensor agents and the fusion module is implemented using a publisher–subscriber (pub/sub) architecture, enabling asynchronous and scalable data exchange. Each sensor publishes its detection results to dedicated topics, while the meta-model subscribes to these streams and aggregates the incoming data at predefined intervals (e.g., one second).

Unlike existing multi-sensor fusion approaches that rely on centralized processing or infrastructure-heavy sensing such as radar, the proposed framework emphasizes distributed AI-based decision fusion using cost-effective sensing modalities. The integration of confidence-aware weighting and asynchronous communication enables the system to maintain detection performance under dynamic and partially degraded sensing conditions, which are common in urban environments.

IV. Methodology

The proposed multi-sensor drone detection system is developed using a unified deep learning framework, where all sensing modalities including optical, acoustic, and RF, are processed using convolutional neural networks (CNNs). To enable this, non-visual signals are transformed into two-dimensional representations, allowing consistent model architecture and training workflows across heterogeneous data sources. This unified approach simplifies integration and supports modular activation of sensing components based on operational needs.

A. Optical Sensing Models

The optical detection pipeline follows a two-stage architecture designed to reduce false positives in complex urban environments as shown in Figure 2. The first stage employs a YOLO-based detector to identify candidate objects in real-time video streams. The model was trained on approximately 52,000 images of drones across diverse environments, using a resolution of 1600×900 pixels, 50 training epochs, and a batch size of 16.

Detected regions are then passed to a secondary CNN classifier for refinement, distinguishing drones from visually similar objects such as birds or helicopters. This classifier was trained on over 95,000 images at 256×256 resolution using a lightweight YOLOv11 Nano architecture, with data augmentation techniques including rotation, color adjustment, and mosaic composition.

To ensure operation under low-visibility conditions, an additional infrared-based model is incorporated. This model, trained on approximately 4,000 annotated thermal images, detects heat signatures from drone components. Despite lower resolution (640×640) and dataset size, the infrared model provides complementary capability in nighttime or degraded visual environments.

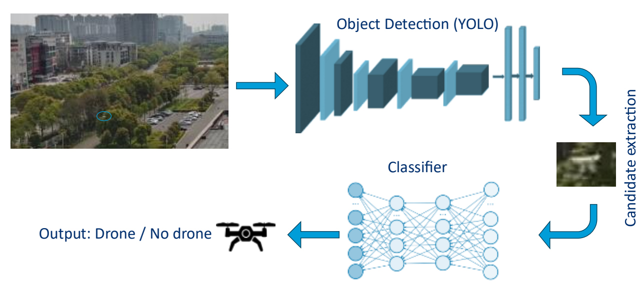


Figure 2. Two-stages processing pipeline for image-based models

B. Acoustic Sensing Models

The methodology for acoustic detection leverages the characteristic sounds produced by drone propellers and motors to identify their presence. Environmental audio is captured and segmented into fragments with a duration of one second to provide consistent temporal windows for analysis as shown in Figure 3. This segmentation process transforms a single continuous audio sample into multiple discrete fragments, which are then used as the basis for feature extraction.

Each one-second audio fragment is converted into a spectrogram by applying a Short-Time Fourier Transform (STFT). This process generates a visual representation of sound frequency over time, effectively translating an auditory signal into a graphic format suitable for neural network processing. This technical unification allows the system to identify specific acoustic patterns, such as the rhythmic noise of motors, as if they were visual features within an image.

The dataset utilized for training the acoustic model consists of 10,373 non-drone spectrograms and 1,333 drone spectrograms. For the training phase, 80% of the data was allocated to the training set while 20% was reserved for validation. The model was trained using the YOLO architecture for 152 epochs, focusing on differentiating drone signatures from ambient environmental noise and other extraneous sounds.

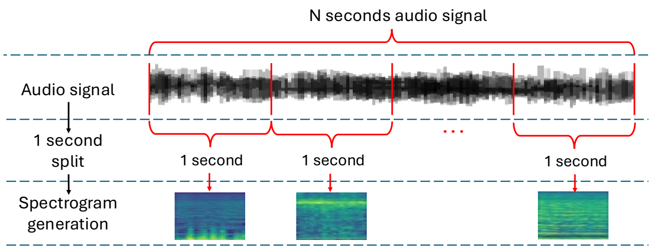


Figure 3. Split and Spectrogram generation of audio signals

A specific constraint applied during the training of the acoustic network was the deactivation of data augmentation options. While techniques such as rotations or flips are standard in image-based projects, their application to spectrograms would result in representations of non-existent audio signals. Introducing such variations would create noise and errors in the learning process, thus the training relied strictly on original spectrogram representations to maintain signal integrity.

C. *Radio Frequency Sensing Models*

Radio frequency-based detection focuses on the communication and control signals exchanged between a drone and its ground station. This sensing modality offers the advantage of long-range detection and the ability to identify drones that are not within the line-of-sight or are otherwise visually obscured. The model is designed to analyze the radio frequency spectrum for specific communication patterns that are characteristic of drone controllers and data transmission.

The raw radio frequency signals are transformed into two-dimensional representations using the STFT similar to the audio signal. This transformation encodes the one-dimensional frequency spectrum series as images with a resolution of 640x640 pixels, capturing the unique spectral fingerprints of drone activity. This approach is particularly effective for identifying drone presence in congested electromagnetic environments where simple signal detection might fail.

The dataset for this model includes 21,142 images, comprising 13,187 drone signals and 7,955 non-drone signals. The data was partitioned into three sets, including a primary training partition of 92%, followed by test and validation partitions of 4% each. This distribution ensures a broad basis for the model to learn the distinct bands emitted during drone control commands, which are relatively easy to differentiate from other patterns.

The process spanned 50 epochs with a batch size of 64 and included specific parameters such as a dropout value of 0.3 to prevent overfitting. Additionally, early stopping with a patience of five epochs was implemented to ensure the model reached optimal convergence without unnecessary computational cycles.

D. *Evaluation and Performance Metrics*

The evaluation of the drone detection system was conducted through a dual-phase approach, involving both controlled laboratory tests of individual models and a validation exercise within a simulated urban environment.

For the individual evaluation of model in lab conditions, we used usual performance metrics for object detection and classification: for object detection, Intersection over Union (IoU) measures the spatial overlap between predicted boxes and ground truth. This is summarized by mean Average Precision (mAP), specifically mAP@50 and mAP@50-95, to evaluate the model across different spatial precision requirements. These scores indicate how well the models can locate and identify drones. Classification models are measured by precision, recall, and accuracy. Precision indicates how accurate the model is when it predicts a drone's presence, which helps minimize false alarms. Recall measures the model's ability to find all actual drones, reducing the chance of missing a target. Accuracy represents the proportion of total correct predictions made by the model. These metrics work together to provide a balanced view of model performance.

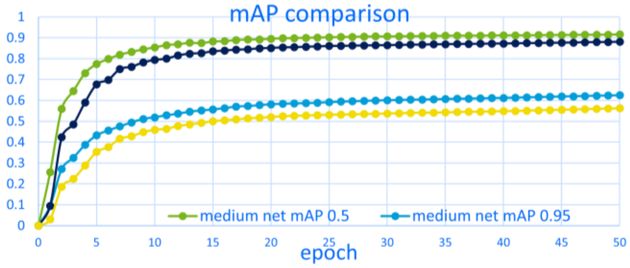
The second phase of evaluation involved a validation exercise, which utilized a highly realistic 3D simulated urban environment to assess the system's real-time performance. In this validation we evaluated the performance of the systems when not all the models were activated. This validation focused on the parallel execution of visible and audio models, but no radio frequency models were deployed during this exercise, to test a realistic scenario where not all sensors are available. Two distinct types of scenarios were analyzed to test the system under varying levels of environmental difficulty. The first scenario represented a complex urban layout with high-rise buildings, obstructions, and high traffic density, while the second one utilized a simpler layout with clearer detection conditions.

V. Results and Discussion

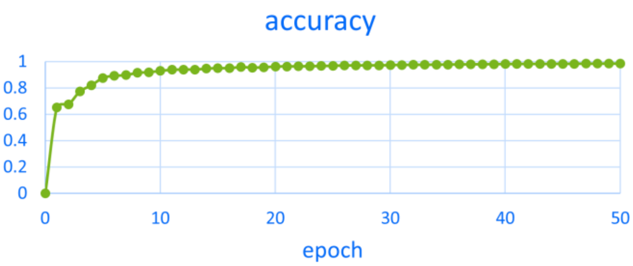
A. *Individual Model Performance in Laboratory Settings*

During the laboratory phase, each machine learning model was assessed to establish baseline performance metrics for detection and classification. The visible spectrum detection model, utilizing the YOLO architecture, demonstrated high accuracy with a mAP@50 of **0.931** and a mAP@50-95 of **0.646**. These metrics were complemented by a secondary classification model that achieved a Top-1 accuracy of **0.985**, showing strong capability in distinguishing drones from airplanes, birds, and helicopters with a 99% true positive rate for the drone category as shown in Figure 4.

Other modalities provided diverse results based on their respective signal characteristics. The radio frequency model achieved nearly 100% accuracy due to the distinct spectral bands emitted by drone control commands, which are easily differentiated from background noise. In contrast, the infrared model recorded lower values, with a mAP@50 of **0.662** and a mAP@50-95 of **0.329**. This performance gap is attributed to the smaller training dataset and lower resolution of thermal imagery compared to visible spectrum data. The audio classification model also performed well in isolated tests, yielding a true positive rate of **0.974** and a true negative rate of **0.996** when processing characteristic motor sounds.



(a)



(b)

Figure 4. Performance metrics for detection and classification: mAP comparison (a) and accuracy (b)

B. Validation in Simulated Urban Environments

The system-level validation was conducted in a realistic 3D urban simulation environment to evaluate performance under operational conditions. The proposed framework maintained real-time processing speeds exceeding 50 frames per second, demonstrating its suitability for real-time applications on standard computing hardware

In the simple urban scenario, the system achieved a precision of approximately 94.12% with an incorrect detection rate of only 5.88%. In the complex urban scenario, precision remained high at **90.20%**, even with the increased presence of obstructions and erratic drone movements. These results indicate that the system maintains robust detection performance even under challenging conditions.

C. Comparison and Discussion of Findings

A comparison between laboratory tests and urban simulations reveals both the strengths and the challenges of the multi-sensor approach. The visible spectrum models translated effectively from isolated testing to the simulated city environment, maintaining an overall accuracy of over 90%. The use of a weighted fusion technique, which combines predictions from different models into a final consensus decision, was key in keeping false positive detections below the 10% threshold in both urban scenarios. This chaining of detection and classification significantly improves the reliability of the system compared to using single sensors.

However, the validation exercise identified a notable discrepancy in the performance of the audio detection model. While laboratory tests using real-world recordings showed high precision, the model's accuracy decreased when integrated into the simulated city environment. This reduction suggests that the dataset used to train the audio-based model did not represent the actual acoustic conditions that can be found in real environments and the wide variety of background sounds and audio in real urban environments. This finding emphasizes the importance of training models with a wider range of real-world data to ensure they remain functional when moved from controlled environments to varied operational settings.

The overall results indicate that while individual models provide a foundation for detection, integration through decision-level fusion is what enables the system to handle the complexities of urban airspace. The system's performance in the complex scenario (**90.20% precision**) relative to the simple scenario (**94.12% precision**) shows that environmental density and obstructions do impact accuracy, yet the results remain well above the established operational requirements.

The visual detection pipeline operated at over 50 frames per second during validation, indicating that frame-level object detection was performed in real time under the tested computational conditions. The decision-level fusion module aggregated predictions from active sensor agents at preset intervals of one second using a publisher–subscriber architecture. This fusion interval therefore represents the primary synchronization cycle governing final detection decisions.

As a result, while individual sensor-level detection occurs at sub-second time scales, the effective system-level detection update rate is bounded by the one-second aggregation interval. The present study does not evaluate the cumulative end-to-end delay from target appearance to alert generation.

In addition, the proposed multi-sensor fusion framework provides an inherent level of resilience against false data injection (FDI) by integrating heterogeneous sensing modalities and combining their outputs through confidence-weighted decision fusion. The framework could reduce the influence of corrupted or manipulated sensor inputs. While explicit detection of FDI is beyond the scope of this work, the results provide a promising foundation for developing more secure and resilient drone monitoring systems.

VI. CONCLUSION AND RECOMMENDATIONS

This paper presented the development and evaluation of a distributed AI-based multi-sensor fusion framework integrating visual, acoustic, and RF sensing for drone detection in urban airspace environments. The objective was to examine whether affordable and widely deployable sensing modalities, combined through weighted decision-level fusion, can improve detection robustness under heterogeneous urban conditions.

The experimental results demonstrate that multi-sensor integration enhances detection reliability compared to single-modality approaches. Visual detection models maintained over 90% detection accuracy when transitioning from laboratory validation to simulated urban environments. In the simulated validation scenarios, the system achieved 94.12% precision in the simple urban scenario and 90.20% precision in the complex urban scenario. Furthermore, the weighted fusion strategy effectively limited false positive detections to below 10% across both scenarios.

At the same time, validation revealed performance degradation in the acoustic-based model when applied to simulated urban conditions. While laboratory recordings yielded high precision, the reduced performance in complex environments indicates that the training dataset did not sufficiently represent real-world acoustic variability. This observation highlights the importance of environment-representative data when deploying machine learning models beyond controlled settings.

The framework maintained real-time processing performance exceeding 50 images per second during validation, demonstrating that the proposed architecture operates within practical computational constraints under the tested configuration. However, end-to-end operational latency was not experimentally measured in this study and should be addressed in future investigations.

Future work should prioritize expanded data collection across diverse urban environments, extended field validation, and systematic evaluation of operational constraints relevant to restricted-area monitoring and U-space contexts. Further investigation is also required to assess integration pathways with higher-level monitoring services while maintaining transparency, reliability, and interoperability.

Acknowledgment

This research is part of AI4HyDrop project and supported by the SESAR 3 Joint Undertaking and its founding members and co-funded by the EU's research and innovation program Horizon Europe under Grant Agreement no 101114805.

References

[1] S. J. Undertaking, "European ATM Master Plan: Roadmap for the safe integration of drones into all classes of airspace," 2022.

[2] S. J. Undertaking, "U-space Blueprint," 2017. [Online]. Available: https://www.sesarju.eu/u-space-blueprint

[3] H. Lee *et al.*, "CNN-Based UAV Detection and Classification Using Sensor Fusion," *IEEE Access,* vol. 11, pp. 68791–68808, 2023, doi: 10.1109/access.2023.3293124.

[4] AI4HyDrop. "An AI-based Holistic Dynamic Framework for a safe Drone's Operations in restricted and urban areas." https://ai4hydrop.eu/ (accessed 29 December, 2025).

[5] D. Kaur, N. Battish, A. Bhavsar, and S. Poddar, "Analysis of vision-based Air-to-Air UAV detection using YOLO based architectures," *Robotics and Autonomous Systems,* vol. 193, 2025, doi: 10.1016/j.robot.2025.105074.

[6] B. Aydin and S. Singha, "Drone Detection Using YOLOv5," *Eng,* vol. 4, no. 1, pp. 416–433, 2023, doi: 10.3390/eng4010025.

[7] M. Kassab, R. A. Zitar, F. Barbaresco, and A. E. F. Seghrouchni, "Drone Detection With Improved Precision in Traditional Machine Learning and Less Complexity in Single-Shot Detectors," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 60, no. 4, pp. 3847–3859, 2024, doi: 10.1109/taes.2024.3368991.

[8] I. Stancic, L. Veic, J. Music, and T. Grujic, "Classification of Low-Resolution Flying Objects in Videos Using the Machine Learning Approach," *Advances in Electrical and Computer Engineering,* vol. 22, no. 2, pp. 45–52, 2022, doi: 10.4316/aece.2022.02006.

[9] I. Ud Din, A. Almogren, and J. J. P. C. Rodrigues, "Adversarial Machine Learning for Robust and Secure UAV Detection in Consumer Applications," *IEEE Transactions on Consumer Electronics,* vol. 71, no. 2, pp. 3360–3367, 2025, doi: 10.1109/tce.2025.3572255.

[10] S. Kümmritz, "The Sound of Surveillance: Enhancing Machine Learning-Driven Drone Detection with Advanced Acoustic Augmentation," *Drones,* vol. 8, no. 3, 2024, doi: 10.3390/drones8030105.

[11] E. Akbal, A. Akbal, S. Dogan, and T. Tuncer, "An automated accurate sound-based amateur drone detection method based on skinny pattern," *Digital Signal Processing,* vol. 136, 2023, doi: 10.1016/j.dsp.2023.104012.

[12] S. Lee, A. Abdulboriy Abdulkhay Ugli, and J. S. Shin, "More Realistic Audio-Based Drone Detection and Identification Approaches With Machine Learning," *IEEE Access,* vol. 13, pp. 170328–170350, 2025, doi: 10.1109/access.2025.3613683.

[13] T. Delleji, F. Slimeni, A. Ayadi, R. Salhi, and A. Siala, "Real-Time Detection and Identification of mini-UAVs Using RF Signal Features and a Customized MLP Model," *Journal of Electrical Engineering & Technology,* vol. 20, no. 5, pp. 3585–3600, 2025, doi: 10.1007/s42835-025-02217-2.

[14] R. Kılıç, N. Kumbasar, E. A. Oral, and I. Y. Ozbek, "Drone classification using RF signal based spectral features," *Engineering Science and Technology, an International Journal,* vol. 28, 2022, doi: 10.1016/j.jestch.2021.06.008.

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

[16] K. Raina, T. Alladi, V. Chamola, and F. R. Yu, "Detecting UAV Presence Using Convolution Feature Vectors in Light Gradient Boosting Machine," *IEEE Transactions on Vehicular Technology,* vol. 72, no. 4, pp. 4332–4341, 2023, doi: 10.1109/tvt.2022.3227192.

[17] Y. Vasant Ahirrao, R. P. Yadav, and S. Kumar, "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.

[18] P. Choudhary, V. Sihag, G. Choudhary, and N. Dragoni, "DrIfTeR: A Drone Identification Technique using RF signals," *Forensic Science International: Digital Investigation,* vol. 54, 2025, doi: 10.1016/j.fsidi.2025.301948.

[19] M. Larrat and C. Sales, "Classification of Flying Drones Using Millimeter-Wave Radar: Comparative Analysis of Algorithms Under Noisy Conditions," *Sensors (Basel),* vol. 25, no. 3, Jan 24 2025, doi: 10.3390/s25030721.

[20] W. Wang *et al.*, "HQNN-SFOP: Hybrid Quantum Neural Networks with Signal Feature Overlay Projection for Drone Detection Using Radar Return Signals—A Simulation," *Computers, Materials & Continua,* vol. 81, no. 1, pp. 1363–1390, 2024, doi: 10.32604/cmc.2024.054055.

[21] A. N. Sayed, H. Abedi, O. M. Ramahi, and G. Shaker, "Enhanced UAV Detection and Classification Using Machine Learning and MIMO Radars," *IEEE Transactions on Microwave Theory and Techniques,* vol. 72, no. 11, pp. 6716–6727, 2024, doi: 10.1109/tmtt.2024.3400889.

[22] S. Eiadkaew, A. Boonpoonga, K. Athikulwongse, K. Kaemarungsi, and D. Torrungrueng, "UAV Detection and Classification in Complex Environments Using Radar and Combined Machine-Learning Approaches," *IEEE Transactions on Microwave Theory and Techniques,* vol. 73, no. 8, pp. 5457–5470, 2025, doi: 10.1109/tmtt.2025.3551626.

[23] V. Mehta, F. Dadboud, M. Bolic, and I. Mantegh, "A Deep Learning Approach for Drone Detection and Classification Using Radar and Camera Sensor Fusion," presented at the 2023 IEEE Sensors Applications Symposium (SAS), 2023.

[24] M. Aledhari, R. Razzak, R. M. Parizi, and G. Srivastava, "Sensor Fusion for Drone Detection," presented at the 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), 2021.

[25] D. Lofù, P. D. Gennaro, P. Tedeschi, T. D. Noia, and E. D. Sciascio, "URANUS: Radio Frequency Tracking, Classification and Identification of Unmanned Aircraft Vehicles," *IEEE Open Journal of Vehicular Technology,* vol. 4, pp. 921–935, 2023, doi: 10.1109/ojvt.2023.3333676.

[26] A. Frid, Y. Ben-Shimol, E. Manor, and S. Greenberg, "Drones Detection Using a Fusion of RF and Acoustic Features and Deep Neural Networks," *Sensors (Basel),* vol. 24, no. 8, Apr 10 2024, doi: 10.3390/s24082427.