

Proceeding Paper

Intelligent Drone Patrolling with Real-Time Object Detection and GPS-Based Path Adaptation [†]

Gurugubelli V. S. Narayana *, Shiba Prasad Swain , Debabrata Pattnayak, Manas Ranjan Pradhan and P. Ankit Krishna

Department of Computer Science and Engineering, School of Engineering and Technology, GIET University, Gunupur 765022, Odisha, India; 22cse233.shibaprasadswain@giet.edu (S.P.S.); 22cse144.debabratapattnayak@giet.edu (D.P.); 22cse066.manasranjanpradhan@giet.edu (M.R.P.); 21cseaiml028.pankitkrishna@giet.edu (P.A.K.)

* Correspondence: gvs.narayana@giet.edu

[†] Presented at the 6th International Electronic Conference on Applied Sciences, 9–11 December 2025; Available online: <https://sciforum.net/event/ASEC2025>.

Abstract

Background: The need for autonomous aerial surveillance originates from weaknesses in manual monitoring, such as late response, low scalability and rigid patrol plans. AI and GPS-driven smart aerial monitoring present an attractive solution for continuous adaptive wide-area surveillance. **Objective:** In this paper, we aim at designing and validating experimentally a low-cost drone-based unmanned autonomous mission patrolling system with waypoint navigation, real-time video backhauling, AI-based human/object detection and GPS path re-planning when an event occurs to ensure the safety of patrol missions under battery constraints. **Methods:** The proposed architecture combines autonomous navigation and embedded flight-control with online analog video streaming and ground-station-based computer vision processing. Object detection based on deep learning for live aerial video is used, and the proposed system's performance is tested at different altitudes, lighting states and GPS patrol plans. **Results:** Experimental results show that the proposed method can obtain stable waypoint tracking with a clear real-time video downlink in patrol missions. The system is able to adaptively modify paths as a reaction to detected events and commence safe return-to-home functionality during low-battery conditions. The proposed detection model obtains a mean average precision of 87.4%, with an F1-score of 0.89 and real-time inference latency (20–25 ms per frame) that enables fast service without any interruption in practice during surveillance deployment. **Conclusions:** Experimental results show that the proposed method can obtain stable waypoint tracking with a clear real-time video downlink in patrol missions. The system can adaptively modify paths as a reaction to detected events and commence safe return-to-home functionality during low-battery conditions. The proposed detection model obtains a mean average precision of 87.4%, with an F1-score of 0.89 and real-time inference latency (20–25 ms per frame) that enables fast service without any interruption in practice during surveillance deployment.



Academic Editor: Eugenio Vocaturo

Published: 18 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

Keywords: autonomous drone; real-time object detection; MobileNet-SSD; GPS navigation; DroneKit; STM32 microcontroller; computer vision; AI surveillance

1. Introduction

UAVs have become more and more popular in different applications, including agriculture [1], urban traffic surveillance [2] and emergency search- rescue missions [3]. The swift construction of urban infrastructure, public facilities and large-scale social environment has further bolstered the need for a highly efficient and reliable surveillance system. Based on manual patrolling and fixed-camera monitoring, traditional monitoring methods have disadvantages such as limited coverage, untimely response, and inefficiency in the face of dynamic scenes. These applications are noticeably lacking in the area of perimeter security, disaster assessment, and remote-area monitoring, where long-term and intelligent situational awareness is needed. The autonomous aerial platforms have now become an efficient choice for wide-area surveillance, as they are mobile and robust to deploy at many different locations whilst having high sensing proficiency. Early airborne observation systems were generally confined to predetermined flight patterns and manual oversight, restricting their effectiveness in dynamic environments. New developments in GPS-guided mission planning and autonomous flight have allowed carnivore-based UAVs to undertake waypoint-driven patrols with limited human intervention. Nonetheless, being able to independently navigate a given environment does not ensure intelligent surveillance, since there is a necessity of real-time perception and context awareness in order to make relevant decisions. Aerial monitoring becomes much more powerful with AI and computer vision that is now able to automatically spot humans and objects of interest in live video streams. However, performing AI inference directly onboard low-cost aerial systems is still difficult due to limitations in onboard computing power, energy supply, and heat dissipation. Hence, most current solutions rely on powerful hardware or have low real-time performance, which makes them impractical and not scalable.

Hybrid surveillance architectures that break away from these limitations by realizing an offloading of AI processing to a ground control station and, at the same time, embedding autonomy in the aerial platform have been receiving considerable attention. This method can effectively process visual data in real time without putting too much load on onboard hardware. Nevertheless, existing systems tend to be detection-focused and lack consideration of system-level requirements such as adaptive GPS-based mission behavior or continuous patrol monitoring and battery-aware safety mechanisms needed for deployment in reality. This paper develops an intelligent, low-cost autonomous drone patrolling system comprising GPS-based waypoint navigation, real-time video communication handling and AI-powered event analysis enhanced with adaptive mission response. To meet the requirement, our framework allows the aerial platform to dynamically respond to what it perceives by performing “hover” or path-adjustment behaviors when some events are detected without affecting fundamental flight stability and its communication links, as well as safety-critical missions (e.g., return-to-autonomy home). Extensive experiments are carried out to demonstrate the effectiveness and robustness of the proposed system under different operating conditions.

Unlike many UAV surveillance studies that treat perception and navigation as independent functions, the proposed work presents a system-level autonomous surveillance framework in which perception directly affects flight behaviour. The system establishes a closed operational loop connecting real-time object detection, GPS-based navigation, and mission safety management on a resource-constrained platform. In this framework, detected events trigger adaptive waypoint modification and localized observation rather than passive video recording. The architecture further incorporates a battery-aware return-to-home mechanism to ensure operational safety during prolonged missions. By offloading computationally intensive vision processing to a ground station while maintaining stable onboard flight control, the platform achieves real-time responsiveness without requiring

high-performance embedded processors. Therefore, the novelty of this study does not lie in proposing a new detection algorithm, but in experimentally demonstrating autonomous surveillance behaviour emerging from the interaction between perception, communication, and flight control subsystems.

2. Literature Survey

Several works [4] have addressed UAV-enabled surveillance systems through the deployment of state-of-the-art sensor technology, autonomous path planning, and intelligent deployment, with which situational awareness can be improved. The combination of tiny and large AI models, as well as edge computing architectures for enhanced computation efficiency with lower response time, was also proposed. Yet current studies have consistently reported problems with limited energy supplies, delays in communication and issues regarding privacy. Performance trade-offs are observed in comparisons between the benchmark platforms and experimental prototypes, demonstrating a demand for practical, scalable, reliable, energy-efficient UAV surveillance systems.

Inspired by [5], the recent works on UAV-based object detection mainly concentrate on adapting one-stage models, in particular YOLO variants, for small target size, scale variance, occlusion and complex aerial background. As feature pyramid-based methods (FPN, PANet, AFPN) develop into one state-of-the-art single-scale and multi-scale object detection frameworks on hinge datasets such as VisDrone and UAVDT. Light-weight backbones, spatial pyramid pooling (SPP) and modules and adaptive feature fusion are proposed to compromise the need for accuracy and real-time nature. Evaluation often focuses on several criteria, such as mAP, Parameter count and FPS, showing the trade-off between efficiency and accuracy. But, precise detection of densely small objects with less model complexity is still a challenge in UAV imagery.

Ahmad et al. [6] present the increasing deployment of UAVs for ITS applications as influenced by technologies like IOT, artificial intelligence and edge computing. The literature focuses on the ability of UAVs to obtain high-resolution data from complex scenes, using real-time onboard processing. The existing works are to classify UAV autonomy and activeness levels, as well as present the core system architectures and key technologies. Applications vary from simple watching to completely autonomous weather control. However, some challenges are still remaining, such as system integration, computational efficiency and scalability with the creation of robust autonomous UAV frameworks.

Cordill et al. [7] discuss growing security and privacy concerns associated with the proliferation of UAVs in both the civilian sector and industry. These approaches take the whole system-level perspective, including hardware, software and communication attack surfaces. In addition to considerations over privacy (in a more traditional sense), ethical and regulatory aspects on a global level are highlighted. Lightweight cryptography, privacy-preserving machine learning, and blockchain-inspired security solutions are discussed. Nevertheless, issues about computation cost, regulatory variation and practical applications of computational designs are still intriguing.

Abdusalomov et al. [8] have studied deep learning-based fire outbreak detection to address the drawbacks related to traditional satellite and sensor-based surveillance in agricultural areas. Lighter-weight object-detection models designed for UAV deployment are particularly highlighted for real-time surveillance. The one-stage detectors with lightweight backbone networks can take advantage of effective detection and low computational cost. We demonstrate through comparative studies its superior accuracy and inference speed compared to the state-of-the-art methods. Nevertheless, open issues are present with regard to environmental variability, variety of firescale and robustness of edge-computing.

Despite significant progress in UAV-based surveillance systems, existing studies primarily emphasize detection accuracy and model efficiency, with several works integrating waypoint navigation and offboard AI processing. However, limited attention has been given to practical system-level validation under real-world constraints. In particular, issues such as closed-loop interaction between perception and autonomous navigation, battery-aware mission continuity, latency stability under varying system loads, low-cost hardware feasibility, and communication reliability remain insufficiently addressed. In comparison to prior studies, most UAV surveillance research can be broadly categorized into two groups: detection-centric approaches and navigation-centric approaches. Detection-centric works primarily improve recognition accuracy and model efficiency but typically treat the UAV as a passive sensing platform. Navigation-centric works optimize waypoint tracking and autonomous flight but do not modify flight behaviour based on perception results. The proposed system differs by establishing a closed-loop interaction in which perception outcomes directly influence navigation decisions and mission safety behaviour. Rather than optimizing an individual algorithm, this work evaluates how perception latency, communication reliability, and flight control jointly affect real-time surveillance capability on a low-cost UAV platform. It is important to note that the proposed framework does not attempt to introduce new navigation hardware or a novel object detection algorithm. Instead, the research contribution lies in experimentally demonstrating how established components can be coordinated to produce autonomous surveillance behaviour. In many prior studies, these components operate independently, where detection results are used only for monitoring or post-analysis. In contrast, the present system integrates perception feedback into real-time flight decision-making, enabling adaptive patrol and safety response during mission execution. The cited references further highlight this gap. Survey and architectural studies such as [4,6] summarize UAV capabilities and applications but do not experimentally demonstrate adaptive mission behaviour during operation. Detection-oriented works including [5,8] primarily evaluate recognition accuracy and inference speed without modifying flight behaviour based on perception outcomes. Security and operational analyses in [6,7] discuss communication and safety concerns, yet they do not implement perception-driven navigation response. Consequently, although individual subsystems of UAV surveillance have been extensively studied, their real-time interaction and behavioural validation in an operational patrol scenario remain insufficiently demonstrated in prior literature. Therefore, there is a need for an experimentally validated framework that integrates perception, navigation, and safety mechanisms within a resource-constrained UAV platform. The proposed work specifically addresses these deployment-oriented challenges through system-level integration and real-time performance evaluation.

The proposed DEIMv2 builds upon the Dense O2O methodology and integrates DINOv3-pretrained backbones with a Spatial Tuning Adapter (STA) to generate enriched multi-scale features for detection. The model is evaluated on the COCO dataset using Average Precision (AP) as the primary metric, achieving 57.8 AP with 50.3 M parameters for DEIMv2-X and 50.9 AP with only 9.71 M parameters for DEIMv2-S. Ultra-lightweight variants such as DEIMv2-Pico (1.5 M parameters) also demonstrate competitive performance (38.5 AP), highlighting its superior performance–cost balance compared to existing models like YOLOv10-Nano. These results establish DEIMv2 as a scalable and efficient state-of-the-art detection framework [9].

Object detection has advanced significantly with single-stage detectors balancing speed and accuracy. On the COCO dataset, YOLOv4 integrates advanced methodologies within the Darknet framework, including Cross mini-Batch Normalization, Cross Stage Partial (CSP) connections, Self-Adversarial Training (SAT), Mosaic data augmentation, DropBlock regularization, and CIoU loss to improve regression and classification

performance. The model achieves 43.5% AP and 65.7% AP50 at 65 FPS on Tesla V100, demonstrating competitive real-time performance compared to prior YOLO versions and other state-of-the-art detectors. These results indicate that optimized architectural design and data augmentation strategies significantly enhance detection accuracy and efficiency in practical environments [10].

Recent advancements in UAV-based object detection [11] have largely relied on deep learning-driven one-stage detectors, particularly YOLO variants, due to their balance between speed and accuracy. Prior studies have highlighted challenges unique to aerial imagery, including small object scale, background clutter, and viewpoint variability, which often degrade detection performance in real-time scenarios. To mitigate these issues, researchers have explored multi-scale feature aggregation frameworks such as PANet, advanced data augmentation strategies like mosaic augmentation, and attention mechanisms including squeeze-and-excitation (SE) blocks to enhance feature representation. Despite these improvements, achieving robust small-object detection with faster convergence and computational efficiency remains an active research area. The proposed enhanced YOLOv5 framework builds upon these developments by integrating feature refinement and channel attention mechanisms to improve detection precision in UAV-view datasets.

Earlier approaches based on SIFT and optical flow were effective in controlled settings but incurred high computational overhead and showed reduced adaptability in rapidly changing environments. With the emergence of real-time object detectors such as YOLOv5, YOLOv7, and more recently YOLOv8 [12], significant improvements have been reported in detection accuracy (mAP), inference speed (FPS), and model efficiency for embedded UAV platforms. Several studies have focused on lightweight architectures and edge deployment; however, limited attention has been given to decision-aware weighting mechanisms that directly influence navigation behavior. Therefore, integrating a zone-based class weighting framework with YOLOv8 for real-time evasive maneuver optimization represents a meaningful advancement toward scalable and autonomous UAV obstacle avoidance systems.

Recent progress in UAV-based image detection [13] has been largely driven by deep learning object detection models. These include two-stage approaches such as R-CNN, Fast R-CNN, and Faster R-CNN, as well as one-stage methods like YOLO and SSD. While two-stage detectors typically provide higher accuracy and more precise localization, one-stage models enable faster inference, making them better suited for real-time and latency-sensitive UAV applications. Multi-scale feature extraction techniques and backbone optimizations have further enhanced performance for small and densely distributed aerial targets. Metrics like mAP, precision, recall, and FPS are commonly used to evaluate how accurate and fast a UAV detection model performs. These measures help determine whether a model can provide reliable detection while operating efficiently within limited hardware resources.

3. Methodology

The intelligent drone patrolling infrastructure is conceptualised as a modular and autonomous airborne surveillance system which incorporates GPS-based navigation, wireless video broadcast service (VBS), and AI monitoring. The operational flowchart is depicted in Figure 1 and acts as a sequential pipeline to achieve such an effective, accurate patrol routine from both the computational and energy points of view. The purpose of this framework (Algorithm 1) is to experimentally evaluate how perception output influences autonomous navigation behaviour in a real UAV platform. Therefore, the methodology focuses on operational workflow and system interaction rather than development of a new navigation or detection algorithm.

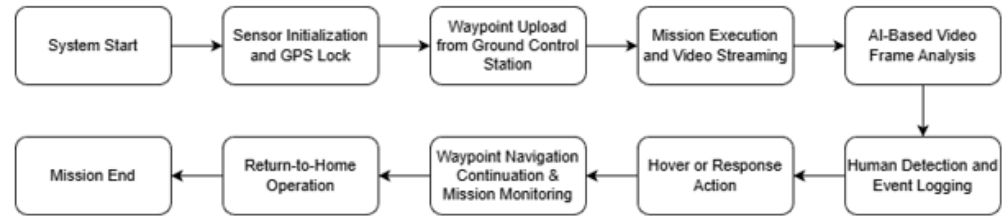


Figure 1. Process Flow diagram.

Algorithm 1 Intelligent Drone Patrol Framework

- 1: Initialize UAV system
 - 2: Load waypoint coordinates
 - 3: **while** mission is active **do**
 - 4: Capture frame
 - 5: Perform object detection
 - 6: **if** detection confidence > threshold **then**
 - 7: Trigger event
 - 8: Update patrol path
 - 9: **end if**
 - 10: Continue navigation
 - 11: **end while**
 - 12: End mission
-

3.1. Hardware Platform and Communication Setup

The developed system is mounted on a home-made flying platform built from standard off-the-shelf components to ensure cost-effectiveness, reconfigurability, and ease of replication. The platform integrates onboard flight control computation, navigation and motion sensing, communication interfaces, and real-time video streaming capability to support autonomous operation with visual feedback. The onboard hardware consists of an STM32-based flight controller, 40A electronic speed controllers, 920 KV brushless motors, an M8N GPS module for position estimation and waypoint navigation, a multirotor drone frame, and a 4200 mAh 3S Li-Po battery. A radio controller is provided for manual override and safety management. For visual monitoring, an analog camera payload captures aerial video, which is transmitted via an analog video transmitter and received at the ground control station (GCS) using an analog video receiver. The platform uses 9450 self-locking propellers to ensure stable flight. To maintain stable autonomous flight while addressing onboard resource limitations, computationally intensive perception tasks are executed at the ground control station. This offboard processing approach enables reliable data transmission and real-time visual analysis without requiring high-performance onboard processors. The selected hardware configuration was intentionally constrained to a low-cost platform to observe whether autonomous behaviour can be maintained under limited onboard computational resources. Hence, the platform serves as an experimental tested rather than a performance-optimized UAV design.

3.2. Perception-Navigation Framework and Mission Operation

For visual processing, images from the COCO (Common Objects in Context) dataset are used to train the object detection model due to its diverse human and object annotations across varying lighting conditions, viewpoints, and background complexity. The trained MobileNet-SSD model is then employed for real-time inference using the live aerial video stream received at the ground control station. The selected hardware configuration was intentionally constrained to a low-cost platform to observe whether autonomous behaviour can be maintained under limited onboard computational resources. Hence, the platform

serves as an experimental tested rather than a performance-optimized UAV design. During operation, the system follows a closed-loop perception–action workflow. After take-off, the UAV autonomously follows predefined GPS waypoints to perform area coverage patrol. The onboard camera continuously captures aerial video and transmits it to the ground control station, where each frame is processed using the MobileNet-SSD detection model. When a target object is detected, the detection output, including object location and confidence score, triggers an event response. The UAV temporarily modifies its flight path to perform localized observation of the detected target. If no detection occurs, the UAV continues normal way point navigation. Consequently, the UAV behaviour depends on perception output, forming a closed control loop in which detection results act as a trigger for navigation decisions rather than passive monitoring information. This continuous interaction between visual perception and navigation control enables autonomous monitoring and real-time reaction during patrol missions. Hence, the system operates as an integrated perception–navigation framework in which detection results directly influence flight behaviour rather than functioning only as a passive monitoring system.

3.3. System Initialization and Mission Setup

Operation begins with platform startup, during which the flight controller, sensing units, communication interfaces, and power management components are activated. Sensor calibration and health verification are performed, followed by acquisition of a stable GPS fix to ensure accurate localization. After successful initialization, predefined patrol waypoints are uploaded from the ground control station, establishing the surveillance route and operational boundaries.

3.4. Autonomous Navigation and Video Transmission

Once mission parameters are set, the platform operates autonomously following GPS waypoints. A constant streaming video transmission sending a real-time video to the ground station is provided, not in control of the patrol operation via an analogue FPV link. This stream flows in low-latency, allowing you to see the footage on your mobile device without significant interruption or delay. The analog FPV link provides low-latency transmission necessary for real-time perception processing; however, the communication channel is susceptible to noise and signal degradation at longer distances.

3.5. AI-Based Event Detection and Adaptive Response

The input video stream to the ground station is analysed on a frame-by-frame level using computer vision AI methods to detect humans or interesting objects. When detected, the reported events are recorded with all relevant metadata such as geographic coordinates and timestamps. According to the detected event, predefined response actions are taken, for example, hovering, localized observation or dynamic path adaptation to improve awareness.

3.6. Patrol Continuation and Safety Management

After the event has been handled, the platform resumes its original patrol route or follows an adaptively modified trajectory according to mission needs. System health, communications and battery are monitored throughout the patrol. Auto Return Home When the energy level is below a certain minimum, or the trajectory is complete, an automatic recovery procedure guarantees safety. Importantly, the return-to-home function is independent of the perception module, ensuring that temporary perception or communication failure does not compromise flight stability or mission safety.

3.7. Mission Termination and Data Logging

When they come to the home base, making a safe landing means the mission is over. All critical operational data, such as flight logs, detection events and system status logs, are recorded for post-operation analysis and performance monitoring. In summary, our methodology integrates customized aerial platform design together with autonomous navigation and real-time visual sensing as well as AI-based event analyses into an integrated operational pipeline. A structured workflow guarantees that patrols are performed reliably, response to observed events is adapted, and missions can be safely terminated, this methodology allows observation of system-level behaviour, specifically the interaction between perception latency, communication reliability, and autonomous navigation stability during real flight operations.

4. Results and Discussion

Experimental tests were carried out on the testing platform of the intelligent drone patrolling framework to fully verify the stability of autonomous navigation, the effectiveness of real-time surveillance and alert detection, perception ability based on AI techniques, and adaptive patrol strategy under practical operation. Both quantitative indicators and qualitative images are illustrated by the associated figures to provide support for analysis and exhibit the validity, responsiveness, and practicality of the proposed system.

The experimental verification was performed on a purpose-made flyable platform for autonomous patrolling tasks. The platform was developed out of readily available parts, making the integration of customizable hardware like flight control, navigation sensors, communication interface and video transmission possible. Such modularised system design enables customisation of the entire system and guarantees stable flight operation and reliable real-world usage.

The autonomous flight experiments show that the platform behaves uniformly and accurately during takeoff, waypoint-based patrolling, and landing. Figure 2, Custom aerial platform for field test with fully integrated sensing, communication and flight control. Consistent real-time trajectory execution and GPS-based navigation were achieved, even though the tested scenarios, proving that the platform is well-suited to performing autonomous patrol. Real-time monitoring is provided through a live video stream to the ground control station. Example of detection results in live aerial video are shown in Figure 3, where multiple people are detected and localized on-the-fly. These observations show that the vision pipeline can be successfully used to detect human subjects in an aerial scenario, in outdoor environments. The quantitative perception results are shown in Figure 4, the detection accuracy with the SSD-MobileNetV2 model. The model achieves an mAP of 87.4% and an F1-score of 0.89, which indicates that the detection performance would be credible for different environments. This accuracy is good enough for practical application in real-time video surveillance. Although direct comparison with alternative UAV platforms or detection networks was not performed due to hardware and deployment constraints, MobileNet-SSD was selected as a practical real-time baseline because it is a widely adopted lightweight detector suitable for edge and embedded vision systems. The objective of this study is to validate system behaviour and closed-loop surveillance capability rather than to compete with state-of-the-art detection benchmarks. Accordingly, the reported accuracy and latency values should be interpreted as indicators of operational feasibility for autonomous patrol missions.



Figure 2. Custom-built autonomous aerial platform used in experimental evaluation.

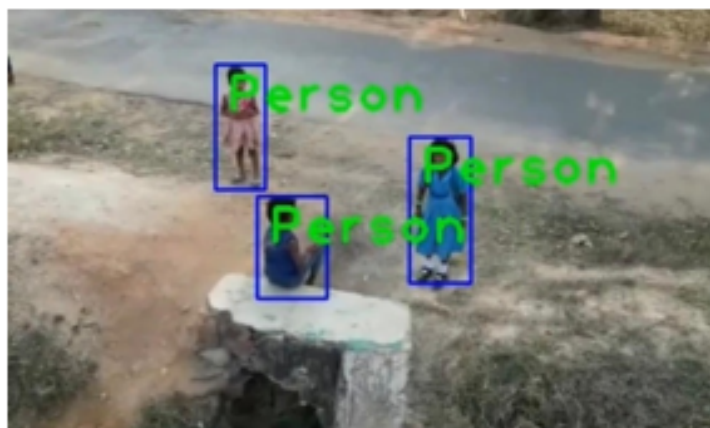


Figure 3. Real-time human detection results from live aerial video.

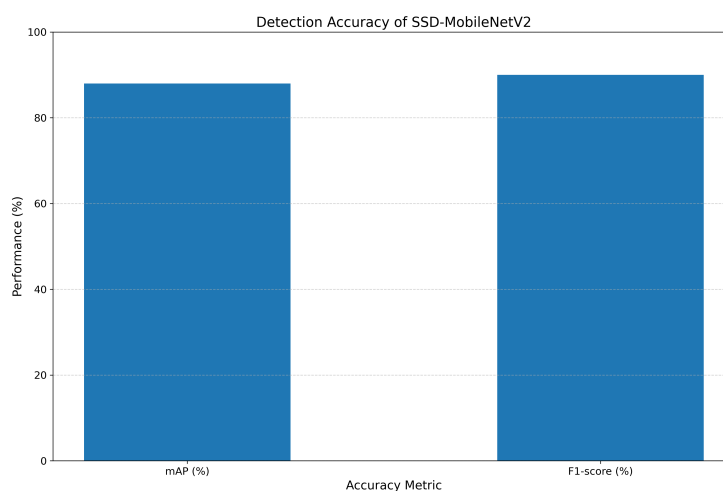


Figure 4. Detection accuracy of the SSD-MobileNetV2 model (mAP and F1-score).

Inference efficiency under varying system loads is analyzed in Figure 5, which presents the inference latency observed under idle, low-load, medium-load, and high-load conditions. As system load increases, a gradual rise in latency is observed; however, processing delay remains within 20–25 ms per frame across all scenarios. This consistency confirms that the perception pipeline maintains real-time detection capability even under increased computational demand. The reverse throughput behaviour is also shown in Figure 6, which shows the number of frames processed per second for different load levels. Despite this decrease in throughput with increasing system load, processing rates at the proposed approach range between 40 and 60 frames per second, which is enough for carrying out real-time surveillance in any case. Meanwhile, the trade-off between latency and throughput is

balanced, which guarantees steady and responsive perception performance in autonomous patrol tasks.

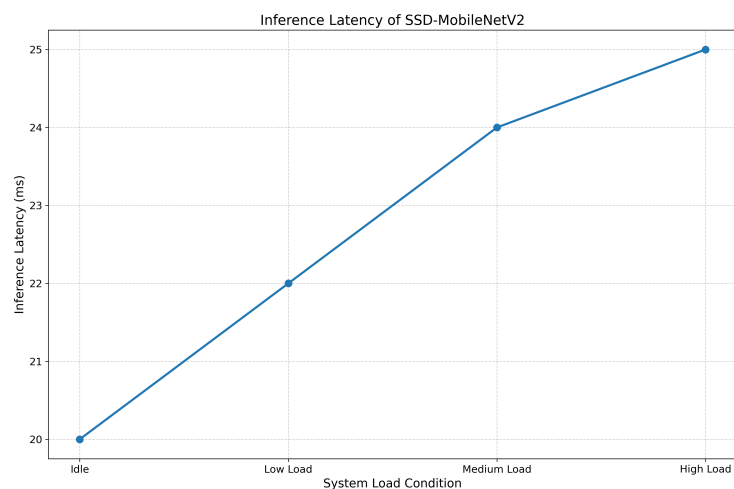


Figure 5. Inference latency under different system load conditions.



Figure 6. Throughput performance across varying system loads.

The obtained latency range (20–25 ms) indicates that perception feedback is fast enough to influence navigation decisions in real time. Considering a typical UAV cruising speed of approximately 4–6 m/s, the system can react to detected events within a spatial distance of roughly 0.08–0.15 m between consecutive perception cycles, as shown in Table 1. This demonstrates that the perception module actively contributes to flight behaviour, confirming closed-loop response rather than passive video monitoring. Combined perception and navigation allow adaptive patrol. Once humans or other areas of interest are detected, the GPS-based waypoints can be re-created on the fly, providing focused observation of a desired area with stable flight. This also closed-loop interaction of detection and control validates the possibility of event-triggered autonomous patrolling integrating offboard intelligence.

While the general reliable system performance is largely preserved, some slight deprivation in detection confidence levels can be seen under very large angles of view, fast movement or analog noise. These side effects naturally fall out of the aerial imaging and analog transmission; however, they do not prevent uninterrupted patrol or the adaptation response. The experimental results as a whole prove the possibility of realising intelligent and adaptive aerial surveillance at an affordable price via reasonable system partitioning. Additional operational observations were made when communica-

tion quality degraded. Under increased distance and electromagnetic interference, the analog video link exhibited visible noise and temporary reduction in detection confidence. However, the flight controller continued autonomous waypoint tracking because navigation was independent of the perception module. In cases where visual detection became unreliable, the UAV maintained its patrol mission and successfully executed the return-to-home safety behaviour during low battery conditions. These observations indicate that communication degradation affects surveillance quality but does not compromise flight safety. Nevertheless, the operational coverage remains limited by the communication range, and larger-area deployment would require more robust digital communication or multi-UAV networking.

Table 1. Quantitative System Performance Evaluation.

Metric	Observed Value	Operational Interpretation
mAP	87.4%	Reliable human detection for patrol monitoring
F1-score	0.89	Balanced precision and recall in real scenes
Inference latency	20–25 ms	Real-time detection capability
Frame rate	40–60 FPS	Continuous surveillance coverage
Communication range	~200 m	Suitable for local perimeter patrol
Waypoint tracking	Stable	Accurate navigation without drift
Return-to-home trigger	Successful	Safety recovery during low battery

5. Conclusions

This paper proposed an intelligent low-cost autonomous drone covering system combining GPS-based waypoint flying, live video streaming and AI-based event detection, with adaptive mission response. Meanwhile, the suggested solution allows the computationally expensive vision tasks to be offloaded towards a ground station in order to provide responsive surveillance while significantly preserving stable autonomous flight and power-efficient operation on resource-limited embedded hardware. Experimental results show that our proposed framework can effectively and efficiently assist in real-time patrols, while also dynamically adjusting the robot navigation trajectory based on the observed events without introducing deviations in waypoint tracking or influencing system stability during flight. Persistent video streaming, together with latency-aware offboard inference, helps achieve useful situation awareness and battery-conscious trajectories for return-to-home operations, guaranteeing flight safety during long missions.

However, there are clear limitations to these strengths. First of all, depending on the transmission of analogue videos, it could introduce sensitivity to noise or poor visual quality in harsh conditions, as well as interference, which can influence detection performance. Second, offboard AI computation speeds up the aerial platform's response to situational changes; however, it requires a robust communication link between the airframe and the ground station, which restricts its applicability in very remote or low-bandwidth regions. Moreover, the existing implementation is limited in terms of detection classes and single-platform operation, which also confines its scalability and multi-target situational awareness. Although the experimental evaluation demonstrates the operational feasibility of adaptive aerial surveillance, the validation was conducted on a single prototype platform and within controlled field conditions. Therefore, the results should be interpreted as a proof-of-concept demonstration of closed-loop perception-guided navigation rather than a fully generalizable deployment solution. Large-scale operation, long-distance communication reliability, and multi-UAV coordination were not experimentally verified and remain subjects for future investigation.

In general, the presented framework can serve as a practical and economical approach toward air-deployed systems for independently flying surveillance tasks such as security patrols, disaster area monitoring and large-area observation. In future investigations, these limitations will be tackled by investigating the on-board AI acceleration and robust digital communication links; LoRa, as a case in point, coordinates cooperative multi-drone capabilities. Multi-class detection system will prolong the generalisation ability of our system, which tends towards improving the scalability and reliability of the system.

Author Contributions: Conceptualization, P.A.K. and G.V.S.N.; Methodology, P.A.K. and G.V.S.N.; Software, S.P.S. and M.R.P.; Validation, S.P.S. and D.P.; Formal Analysis, D.P.; Investigation, G.V.S.N. and D.P.; Resources, P.A.K. and M.R.P.; Data Curation, S.P.S. and G.V.S.N.; Writing—Original Draft Preparation, P.A.K. and G.V.S.N.; Writing—Review and Editing, M.R.P. and D.P.; Visualization, D.P. and M.R.P.; Supervision, G.V.S.N.; Project Administration, P.A.K. and M.R.P.; Funding Acquisition, G.V.S.N. and P.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yang, X.; Smith, A.M.; Bouchier, R.S.; Hodge, K.; Ostrander, D.; Houston, B. Mapping flowering leafy spurge infestations in a heterogeneous landscape using unmanned aerial vehicle Red-Green-Blue images and a hybrid classification method. *Int. J. Remote Sens.* **2021**, *42*, 8930–8951. [[CrossRef](#)]
2. Feng, J.; Yi, C. Lightweight detection network for arbitrary-oriented vehicles in UAV imagery via global attentive relation and multi-path fusion. *Drones* **2022**, *6*, 108. [[CrossRef](#)]
3. Alsamhi, S.H.; Shvetsov, A.V.; Kumar, S.; Shvetsova, S.V.; Alhartomi, M.A.; Hawbani, A.; Rajput, N.S.; Srivastava, S.; Saif, A.; Nyangaresi, V.O. UAV computing-assisted search and rescue mission framework for disaster and harsh environment mitigation. *Drones* **2022**, *6*, 154. [[CrossRef](#)]
4. Kumar, S.; Tiwari, A.; Ahirwar, Y.; Kumar, G.; Arafat, M.Y. The Rise of UAV-Based Smart Surveillance: A Systematic Review of Trends and Technologies. *IEEE Access* **2025**, *13*, 181553–181575. [[CrossRef](#)]
5. Wang, H.; Zhang, Y.; Zhu, C. DAFPN-YOLO: An Improved UAV-Based Object Detection Algorithm Based on YOLOv8s. *Comput. Mater. Contin.* **2025**, *83*, 1929. [[CrossRef](#)]
6. Ahmad, T.; Morel, A.; Cheng, N.; Palaniappan, K.; Calyam, P.; Sun, K.; Pan, J. Future UAV/drone systems for intelligent active surveillance and monitoring. *ACM Comput. Surv.* **2025**, *58*, 35. [[CrossRef](#)]
7. Cordill, B.; Fang, D.; Xu, S. A comprehensive survey of security and privacy in UAV systems. *IEEE Access* **2025**, *13*, 117843–117866. [[CrossRef](#)]
8. Abdusalomov, A.; Umirzakova, S.; Tashev, K.; Egamberdiev, N.; Belalova, G.; Meliboev, A.; Atadjanov, I.; Temirov, Z.; Cho, Y.I. AI-driven UAV surveillance for agricultural fire safety. *Fire* **2025**, *8*, 142. [[CrossRef](#)]
9. Huang, S.; Hou, Y.; Liu, L.; Yu, X.; Shen, X. Real-time object detection meets DINOv3. *arXiv* **2025**, arXiv:2509.20787. [[CrossRef](#)]
10. Geetha, A.S. YOLOv4: A Breakthrough in Real-Time Object Detection. *arXiv* **2025**, arXiv:2502.04161. [[CrossRef](#)]
11. Han, Y.; Zhang, K.; Yan, X.; Liu, X.; Shang, Y.; Qi, P. An improved YOLOv5 for real-time UAV detection. In *Proceedings of the IET Conference Proceedings CP889*; The Institution of Engineering and Technology: Stevenage, UK, 2024; Volume 2024.
12. Vohra, D.S.; Garg, P.K.; Ghosh, S.K. Boosting Unmanned Aerial Vehicle (UAV) Navigation Capabilities: Improving Obstacle Detection and Evasion using the You Only Look Once (YOLO) Algorithm. *J. Aerosp. Sci. Technol.* **2025**, *77*, 214–227. [[CrossRef](#)]
13. Cheng, X.; Lu, B.; Qu, S. Application of Deep Learning in Drone Image Detection. *Highlights Sci. Eng. Technol.* **2024**, *94*, 510–515. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.