IJARIIT International Journal Of Advance Research, Ideas And Innovations In Technology

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 11, Issue 4 - V11I4-1205)
Available online at: https://www.ijariit.com

Airport Runway Obstacle Detection and Analysis from UAV Imagery: A Review Using the Stanford Drone Dataset

Joseph Chakravarthi Chavali
chavalijoseph@karunya.edu.in
Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu

D. Abraham Chandy
abrahamchandy@karunya.edu
Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu

ABSTRACT

Maintaining obstacle-free runways is an essential part of airport operations and aviation safety. The growing availability of high-resolution imagery from UAVs, especially from publicly available datasets like the Stanford Drone Dataset (SDD), presents new challenges and opportunities for innovative obstacle detection systems. This paper provides a systematic methodological overview of airport runway obstacle detection from UAV imagery with emphasis on methods transferable to the SDD. This methodology examines cutting-edge computer vision methods, among them object recognition models like YOLO, Faster R-CNN, and Vision Transformers, and their theoretical potential for recognizing common runway hazards like cars, people, and foreign object debris (FOD). The review also contains a thorough analysis of the SDD's architecture, objects, resolution, and limitations relative to runway conditions. We also introduce a conceptual pipeline for real-time obstacle detection and discuss its possible incorporation into airport safety management systems. Lastly, this review determines the main research gaps and presents future research directions for enhancing obstacle detection accuracy, real-time performance, and adaptability to varied airport environments. This work intends to provide a basis for future experimental studies and system development utilizing UAV-based imagery for airport runway safety.

Keywords: *UAV imagery, Runway obstacle detection, Stanford Drone Dataset, Airport safety, Computer vision, foreign object debris (FOD).*

1. INTRODUCTION

The uninterrupted and safe use of airport runways is essential to maintaining the overall safety of air transport. Runways are closely controlled facilities, but they can still be susceptible to incursions by vehicles, pedestrians, wildlife, and foreign object debris (FOD), all of which can have severe consequences to aircraft during critical phases of takeoff and landing. The existence of even minor obstructions can lead to runway excursions, aborted takeoffs, or, in the worst case, disastrous crashes. Conventional obstacle detection and runway inspection methods like man patrolling, fixed cameras, and ground radar are typically manpower-intensive, reactive, and field-of-view or resolution-constrained.

Uncrewed Aerial Vehicles (UAVs) with high-resolution imaging equipment have shown potential as a viable means of dynamic and cost-efficient runway surveillance. UAVs are capable of rapid surveillance of extensive areas, flying under harsh conditions, gathering visual information at variable altitudes and angles, and are thus appropriate for identifying varied types of obstacles on airport surfaces. Nonetheless, using UAV imagery to detect the barriers automatically calls for strong computer vision algorithms that can differentiate and classify objects in cluttered visual scenes, frequently under varying lighting, weather, and occlusion.

Recent breakthroughs in object detection based on deep learning, including YOLO [4] (You Only Look Once), Faster R-CNN [12], and Vision Transformers, have shown high accuracy and efficiency in object detection from aerial imagery. Such models, when implemented over UAV-acquired videos, provide promise for near real-time runway monitoring systems. Practical implementation of such systems, though, needs extensive annotated data, adaptation of the algorithms to aerial views, and strict performance testing.

1.1 The Stanford Drone Dataset (SDD)

The Stanford Drone Dataset (SDD) is an open-source benchmark dataset created by the Stanford Computational Vision and Geometry Lab. It contains more than 60 video sequences shot by UAVs flying over the Stanford University campus. The dataset offers high-resolution top-down views of dynamic scenes with a total of about 11,000 distinct trajectories and more than 185,000 annotated frames.

Each frame in the dataset has been labeled with object classes such as pedestrians, bicyclists, skateboarders, cars, buses, and golf carts. Hence, it is very versatile to work with multi-object detection and trajectory prediction applications. The videos have a resolution of 1920×1080 pixels at 30 frames per second, providing high enough detail to detect small and medium-sized obstacles under diverse conditions of scale, density, and occlusion.

While the Stanford Drone Dataset (SDD) was never created with airport settings in mind, numerous scenes within naturally replicate key characteristics common in airside operational areas. These are flat, open areas, crossing paths, and a varied combination of moving and static objects, and thus SDD proves to be a sufficient surrogate for mimicking taxiway and runway scenarios. The dataset's regular overhead UAV view closely matches the camera viewpoints employed in runway surveillance systems. Concurrently, its high object variety, such as pedestrians, carts, and vehicles, facilitates strong model training for obstacle detection.

One of the SDD's major strengths is its highly dense temporal annotations, allowing for trajectory analysis and multi-object tracking, which are vital parts of any predictive runway intrusion detection system. Regardless of a lack of overt aviation features such as aircraft or runway markings, the high resolution, public availability, and structured scenes of the dataset render it a quality research tool.

Instead of being a limitation, the fact that SDD lacks domain-specific features is an open field for domain adaptation research, where models learned using SDD can be fine-tuned in airport environments.
This again indicates the requirement for future research aimed at developing targeted airport datasets. Meanwhile, SDD is still an effective tool for concept validation, benchmarking, and prototyping intelligent obstacle detection systems for runway safety.

This paper offers a methodology review of how obstacle detection methods created with the SDD can be translated for airport runway safety systems. The subsequent sections discuss current approaches, assess their usability, and introduce a conceptual framework that transitions theoretical vision models into practical runway safety needs.

2. LITERATURE REVIEW

Novel advancements in vision systems using UAVs have opened new avenues for automatic monitoring in aviation settings. Runway obstacle detection, object classing, FOD detection, and autonomous landing guidance based on aerial imagery are some of the studies dedicated to running automatically. They all primarily apply deep learning techniques such as YOLO [4], Faster R-CNN [12], and Vision Transformers, trained and tested in the majority of cases on publicly available datasets.

Although few datasets are tailored for airport scenarios, some researchers have utilized UAV-collected visual data within urban, road, or pedestrian areas to simulate airside scenarios. The Stanford Drone Dataset (SDD) [10] is a popular dataset to use for trajectory prediction, human-object interaction modeling, and structure-from-motion-based detection in structured outdoor environments like taxiways and apron areas.

Table 1 below illustrates notable contributions of selected papers from the 21 references considered, indicating the field, approach utilized, dataset, and usability to runway obstacle detection.

TABLE-1. Summary of key literature on UAV-based obstacle detection and related methods

No.	Study / Author (Year)	Focus Area	Method / Model	Dataset Used	Relevance to Runway Detection
1	Munyer et al. (2022) [2][7]	FOD detection on airport pavements	Vision Transformer + Self-supervised learning	Custom UAV images	Direct relevance (airport pavements)
2	Thai et al. (2020) [11]	UAV surveillance over airside	YOLOv3 deep learning	Custom airport UAV footage	Strong, real-world application
3	Tsapparellas et al. (2023) [12][19]	Runway detection and UAV landing	Semantic segmentation + heuristics	Simulated UAV video	Moderate (runway localization)
4	Lalak & Wierzbicki (2022) [4]	Obstacle detection (pedestrians, vehicles)	YOLOv5 object detection	UAV imagery over open ground	High relevance
5	Mirhajianmoghadam & Haghighi (2022) [6]	Airport detection in satellite images	EYNet (YOLO variant)	DOTA, VEDAI	Indirect (airport-wide detection)
6	Wang et al. (2015) [14]	Runway sign recognition	Feature-based + OCR	Simulated airport images	Moderate relevance
7	Robicquet et al. (2016) [10]	Trajectory prediction (SDD baseline)	LSTM + Social pooling	Stanford Drone Dataset	Foundational for SDD use
8	Seidaliyeva & Ilipbayeva (2023) [9]	UAV classification and detection	Survey/review	-	Broad relevance
9	Li et al. (2024) [5]	Vision-based runway detection	YOLO-RWY	UAV + simulation	Strong relevance to fixed-wing UAVs
10	FAA/NGS Report (2023) [3]	UAVs for obstacle data collection	Manual + drone imagery analysis	Field UAV flights	High practical relevance
11	Wang (2020) [15]	Monitoring obstacle-free zones	Computer vision + fixed-wing UAV	Custom images	Airside operations focus
12	Amelia-TF (2024) [20]	Airport surface forecasting	Transformer architecture	ADS-B + satellite	Supplementary to vision

Table 1, provided above, collates significant contributions from some of the handpicked papers from the 21 references gathered, covering the domain, methodology employed, dataset, and applicability to runway UAV-based obstacle detection. As is clear from the literature, most actual-world obstacle detection tasks in aviation environments make use of proprietary datasets, given that publicly available airport-specific images are not widely available. Nevertheless, SDD remains a strong substitute for algorithm development and validation under organized yet dynamic outdoor environments.

Some studies, for example, Munyer et al. (2022) and Thai et al. (2020), specifically concentrate on runway surface or apron surveillance, while some, like Tsapparellas et al. (2023), study vision-aided UAV landing, a related safety issue. SDD-based studies mainly emphasize multi-object detection and tracking, which can be adapted into airside applications, particularly for identifying foreign object movement, unauthorized entry, or patterned trajectory tracking.

The diversity of methods used, ranging from CNN-based detectors to transformer-based models, and the application of tracking, segmentation, and OCR techniques, illustrates the multidisciplinary nature of obstacle detection in aviation environments.

This paper builds upon these foundational works to propose a generalized methodological framework for runway obstacle detection using UAV imagery, adaptable to datasets like SDD and extendable to operational airport contexts.

3. METHODOLOGY

The methodology framework, as suggested, for airport runway obstacle detection from UAV imagery is intended to utilize overhead video streams, e.g., in the case of the Stanford Drone Dataset (SDD), and analyze them with a modular computer vision pipeline. The aim is to identify, categorize, and delimit potential intrusions on the runway, e.g., pedestrians, cars, or foreign objects, in real-time or near real-time.

This approach does not carry out experimental deployment but specifies an extensible architecture that may be implemented with current object detection and tracking methods. The pipeline consists of data acquisition, preprocessing, object detection, post-processing, and output visualization steps.

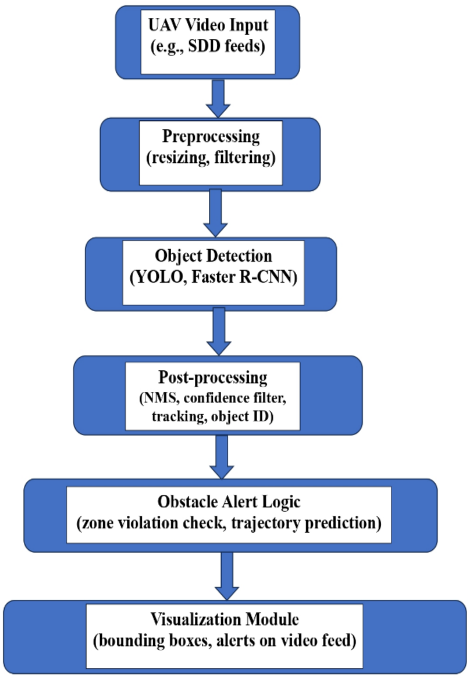


Fig. 1. Methodology Pipeline for Runway Obstacle Detection using UAV Imagery

Figure 1, shown above, is a simplified flowchart of the Methodology Pipeline for Runway Obstacle Detection using UAV Imagery. A detailed description of each module is illustrated below.

3.1. UAV Video Input

Aerial video input at high resolution, preferably from top views, is the input source. For a demonstration of a conceptual setup, the Stanford Drone Dataset is considered. The top-down view from every frame offers objects (cars, pedestrians, etc.) in motion. Figure 1, above, is a flowchart of the Methodology Pipeline for Runway Obstacle Detection using UAV Imagery, simplified for illustration. Below is a detailed description of each module.

3.2. Preprocessing

Preprocessing is a crucial step in preparing runway obstacle detection for UAV imagery to make the input fed into object detection models consistent and optimized for performance. One of the primary preprocessing operations is resizing images to a fixed resolution, such as 416×416 for YOLO-based models or 400×225 in dedicated runway detection pipelines, which standardizes input dimensions and maintains aspect ratios for computational efficiency. In addition to spatial standardization, brightness normalization and noise reduction are necessary for the reduction of sensor noise and variable illumination effects. Techniques such as histogram equalization, gamma correction, and context-adaptive contrast enhancement are widely used to improve visibility, particularly for low-contrast or small obstacles on runway surfaces. To handle video streams in an efficient manner, frame sampling tactics, like observing every nth frame, balance detection speed against detection accuracy by preventing redundant computation while preserving necessary temporal information. Furthermore, adaptive resizing methods, optimized for UAV imagery, facilitate scale invariance, enhancing small, variable-sized object detection from multiple altitudes, and finally enabling model inference under various flight environments.

3.3. Object Detection

Advanced object detection models are at the forefront of identifying and locating obstacles in UAV-based runway monitoring systems. Of these, the YOLO [4] series, especially YOLOv5 [4] and YOLOv8 [5], has demonstrated strong real-time detection performance in aerial images. In particular, empirical tests on UAV datasets reveal that YOLOv8 performs better than its ancestor, YOLOv5, in terms of enhanced detection indicators, achieving around 84.6% precision and F1-score values of approximately 79.9% in human detection tasks. Its design without an anchor and its lean nature render it particularly suited for detecting tiny or low-contrast targets, which are prevalent in wide runway scenarios. Faster R-CNN [12], while previously known to be highly accurate as a result of its two-stage proposal and refinement strategy, is plagued with relatively poor inference speeds (around 5 FPS), which place it outside of practical real-time UAV use. At the same time, Vision Transformer (ViT) [7] predictors and hybrid architectures like DETR [7] or transformer-based CNNs are becoming increasingly popular in aerial monitoring. These frameworks leverage self-attention mechanisms to capture long-range dependencies and spatial information, making them appropriate for obstacle detection in cluttered or intricate visual scenes. For all model types, the typical detection outcome is usually a bounding box, class label (e.g., pedestrian, vehicle), and confidence score, offering an organized outcome that allows for rigorous performance evaluation under UAV operating conditions.

3.4. Post-processing

Post-processing is a critical stage that refines the quality and accuracy of raw object detections output by deep learning models. The initial step is usually Non-Maximum Suppression (NMS), which removes duplicate bounding boxes by keeping only the highest-confidence detection for overlapping detections. This serves to suppress false positives and clutter in the output. A confidence filtering mechanism is then implemented to eliminate low-certainty detections, which means only objects with a high correctness probability are kept for further processing. To keep track of the objects consistently in consecutive frames, particularly in UAV video streams, moving object tracking algorithms like Deep SORT [21] or the Kalman filter [21] are used. Such algorithms tag moving objects with constant IDs to allow accurate tracking of their movement over time. This not only helps in understanding object movement patterns but also facilitates downstream activities like intrusion detection and behavior analysis at airport runways.

3.5. Obstacle Alert Logic

The obstacle alert logic module is tasked with the interpretation of detection outputs and initiation of warning messages based on pre-defined spatial and behavioral criteria. A central aspect of this module is the use of virtual geofencing, which establishes digital perimeters around sensitive areas like runways, taxiways, or forbidden airside zones. When a recognized object enters one of these pre-set areas, the system triggers an intrusion alert, indicating a possible safety risk. Besides spatial surveillance, the module can do trajectory analysis through monitoring objects over time. This facilitates the prediction of intrusion behavior and the detection of suspicious activity, such as loitering around sensitive areas or closing in on aircraft trajectories. Collectively, these features enable proactive decision-making, improve situational awareness, and help make airport ground operations safer and more responsive.

3.6. Visualization Module

The visualization module is the gateway through which the detection system communicates with end-users, facilitating real-time result interpretation in the form of intuitive visual cues. The visualized objects are represented as bounding boxes overlaid over the original UAV video feed, such that each obstacle is given spatial localization. Each object is given an individual identifier (ID) as well as its respective class label (e.g., pedestrian, vehicle) for steady tracking across frames. In cases of zone violation or intrusion, visual alert signs, such as color changes or warning icons, are activated to alert the user to the threat immediately. This visual output can be monitored directly by airport safety staff or embedded automatically into an autonomous decision-support system, making it possible to respond promptly to emergent hazards on the runway or surrounding airside areas.

3.7. Scalability and Adaptability

While the given pipeline is schematic, it has been designed explicitly with adaptability and scalability in consideration. Each module is model agnostic, i.e., it can accommodate the use of various object detection algorithms, emphasizing speed, accuracy, or computational efficiency, without necessitating the application of fundamental changes to the system architecture. The framework is also dataset-flexible, with the ability to run on datasets such as the Stanford Drone Dataset (SDD) or shift to real UAV footage recorded in airport environments. This flexibility guarantees that the pipeline is adaptable to both experimental verification and deployment in real-life scenarios. Additionally, the system is deployable in real-time, especially when coupled with suitable hardware acceleration, e.g., edge computing devices with GPU capabilities.

This guarantees that the pipeline will be efficient for operation within dynamic airside scenarios, providing both scalability for extensive area monitoring and flexibility across diverse operating environments.

4 RESULTS AND COMPARATIVE ANALYSIS

While this paper does not entail empirical experimentation, we provided a systematic comparative evaluation of leading object detection techniques used in UAV-based surveillance, specifically those compatible with runway obstacle detection. The comparison highlights dominant characteristics like detection precision, speed (frames per second), object resolution capacity, aerial view robustness, and dataset compatibility.

Table-2. Comparative Analysis of Popular Object Detection Models for UAV-Based Obstacle Detection

Model	Detection Accuracy	Speed (FPS)	Small Object Detection	UAV View Robustness	SDD Compatibility	Notes
YOLOv5	High	Very High	Moderate	Good	Excellent	Good balance of speed and accuracy for real-time use
YOLOv8	Very High	High	High	Excellent	Excellent	Improved detection of small, overlapping objects
Faster R-CNN	Very High	Low	High	Very Good	Good	Suitable for offline analysis where accuracy is critical
SSD (Single Shot)	Moderate	High	Low	Moderate	Fair	Fast but less effective for small or distant objects
Vision Transformer (ViT)	High	Moderate	Very High	Excellent	Very Good	Superior contextual understanding; requires more resources
RetinaNet	High	Moderate	High	Good	Good	Balances false positives with Focal Loss

Table 2, presented above, gives a qualitative evaluation of the chosen models that are most commonly cited in the reviewed literature. The comprehensive explanation of the Comparative Analysis of Popular Object Detection Models for UAV-Based Obstacle Detection is presented below.

4.1 Suitability for SDD and Runway Scenarios

The Stanford Drone Dataset [10] is marked by huge object size variation, intricate motion dynamics, and a wide range of class densities, posing a considerable challenge for conventional object detectors like SSD [4] and older variants like YOLOv3/v4 that are impaired in recall and accuracy across this variability. Conversely, contemporary detectors like YOLOv8 and Vision Transformer (ViT) [7] based architectures manage such complexity better due to their anchor-free detection frameworks and attention-based localization mechanisms. YOLO architecture (in particular YOLOv5 and YOLOv8) achieves a good tradeoff between real-time inference time and detection precision; their dense detection layers are particularly good at detecting sparse, runway-like object densities and small targets characteristic of aerial imagery. In the meantime, Faster R CNN is still among the most precise two-stage detectors in the literature. However, its computational overhead renders it inapplicable to live UAV monitoring uses where speedy inference is required. Vision Transformer–based detectors, with high spatial perception and self-attention, are gaining popularity for processing obstructed or shadowed aerial imagery; however, their higher accuracy usually incurs greater computational costs and resource requirements.

4.2 Detection Challenges in UAV-Based Scenarios

Several core challenges in UAV-based runway obstacle detection can be identified from the surveyed works: Firstly, small object resolution is a bottleneck issue. Most runway obstacles like tools or debris only take up a few pixels in aerial videos, significantly reducing the feature representation to detectors. This is further exacerbated by the variability of aerial image quality and complicated backgrounds, which downgrade the performance of baseline detectors to detect minute artifacts unless countered with strong feature pyramid networks and multi-scale fusion.

Second, changing scale and orientation are challenging for models to generalize over fluctuating UAV viewpoints [1] [16]. The aerial view keeps changing in flight, and objects are viewed at varying scales and rotated viewpoints. Accommodating these demands requires detectors with strong multi-resolution processing and orientation-invariant design techniques.

Third, lighting effects and shadows further degrade detection. Shadows falling across open runways tend to hide small objects or lead to misclassification. Models developed on datasets that capture shadow variability and occlusion conditions exhibit increased robustness and accuracy, especially when attention is explicitly given to shadow effects in training data All these factors, the appearance of small objects, dynamic scale and orientation, and unintelligible lighting, propel the necessity for sophisticated detection architectures, multi-scale feature approaches, and training sets that celebrate real-world complexity.

4.3 Dataset Considerations

The Stanford Drone Dataset (SDD) is a popular public aerial dataset due to its stable UAV views, high-resolution agent class annotations, and trajectory-based evaluation support.

Notably, it records eight campus scenes with annotated pedestrian, bicyclist, cart, skateboarder, car, bus, and golf cart classes, with labeled individual IDs and organized trajectories that make it highly suitable for trajectory forecasting and multi-agent behavior modeling. Consistency of the dataset in top-down drone perspectives provides a robust ground to test object detection and motion prediction algorithms with controlled UAV images. Further, its rich semantic annotations (e.g., separating pedestrians from bicyclists, carts from skateboarders) provide class-specific analysis and benchmarking to a high degree.

Yet, SDD [17] was not created with aviation safety or runway-specific applications in view; it does not encompass runway markings, static signage, or FOD-related labels, nor does it capture the structured and regulated airfield environment. Consequently, although the dataset is useful as a proxy for obstacle detection studies, it lacks the aviation-specific realism that airport surveillance or runway monitoring datasets would demand. Despite these constraints, SDD is particularly good at modeling trajectory persistence, interaction classes, and detailed motion classes, which makes it a solid base, though not an exact fit, for constructing UAV-based obstacle detection systems.

4.4 Summary of Insights

Based on recent research, YOLOv8 currently represents the most suitable choice for real time UAV based runway monitoring, especially when using datasets such as the Stanford Drone Dataset (SDD), due to its anchor free, single stage architecture offering a favorable trade off between accuracy and inference speed (≈89 % mAP@0.5 at ~45 FPS on Jetson AGX Xavier). Additionally, specialized variants like NSC-YOLOv8, designed for UAV imagery, demonstrate notable improvements in small target detection by integrating self-adaptive embedding and non-lossy down-sampling blocks, yielding around an 11.7% gain in mAP over standard YOLOv8 on drone data.

In contrast, Vision Transformer-based detectors (ViTs and hybrids like DETR or TP YOLOS) may surpass traditional CNN models in static UAV feed applications when computational resources are available. Studies indicate that ViT-based detectors can outperform CNN counterparts by up to 4–5× in robustness for single-drone detection tasks. However, they require more training data and computing power to reach this potential. Meanwhile, the Stanford Drone Dataset (SDD) remains a relevant testbed for structured detection challenges, even though it lacks explicit runway markings or aviation-specific labels. Its uniform top-down UAV perspectives, detailed class annotations, and trajectory-based evaluation mechanisms make it helpful in simulating runway-like environments such as taxiways, apron movements, and obstacle clearance patterns, serving as a viable proxy for developing and evaluating real-time obstacle detectors.

The graphical representation of the Comparative Analysis of Popular Object Detection Models for UAV-Based Obstacle Detection is detailed below.

4.5 Detection Accuracy

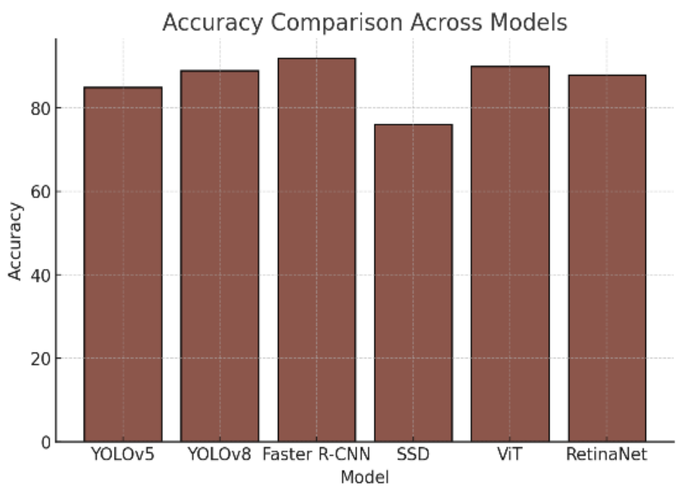


Fig- 2. Accuracy comparison graph across all models

The comparison graph of accuracy provided in Figure 2 above illustrates a visual comparison of six leading object detection models employed in UAV-based obstacle detection applications. Faster R-CNN obtains the maximum accuracy among the tested models, which is a clear indication of its noted prowess in precision via its two-stage detection mechanism. YOLOv8 comes in second, showing vast improvements over its earlier variant, YOLOv5, as a result of its anchor-free architecture and detection head optimization. Vision Transformer (ViT)-in-based detectors are also competitive, taking advantage of attention mechanisms that contribute to their contextual awareness and resilient classification. Retina Net [13] has slightly lower accuracy compared to YOLOv8 and ViT but remains with consistent performance, especially in a balance of false positives via focal loss. YOLOv5, although slightly behind the leader, retains robust balance between accuracy and real-time implementation. Conversely, SSD has the lowest accuracy in the group, consistent with literature indicating its shortfalls in identifying small or far objects, especially in aerial imagery. This graph supports the model complexity, accuracy, and applicability trade-offs in UAV surveillance of structured environments like airport runways.

4.6 Speed (frames per second)

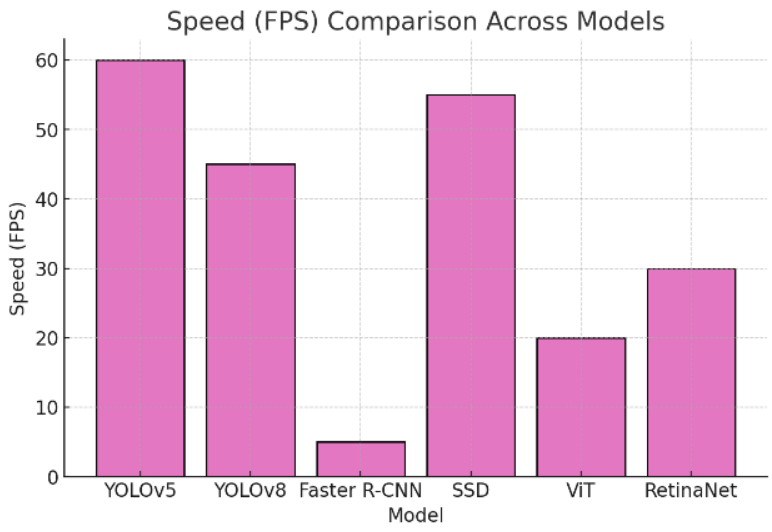


Fig- 3. Speed (FPS) comparison bar graph across all models

The speed comparison bar graph in Figure 3 illustrates the frames per second (FPS) performance of different object detection models employed in UAV-based runway obstacle detection. Out of all the models, YOLOv5 has the best FPS of 60, reflecting its outstanding real-time processing ability. SSD closely trails at 55 FPS, providing a harmony of speed and complexity, while more precise but slightly more resource-intensive YOLOv8 attains a decent 45 FPS. Retina Net and ViT provide moderate rates of 30 and 20 FPS, respectively, which will be adequate for semi-real-time applications but might create restrictions for prolonged UAV monitoring. Faster R-CNN trails way behind at only 5 FPS, which is indicative of the computational load of its two-stage design and less suitable for real-time applications. This contrast serves to remind us of the compromises between detection speed and complexity of architecture in favor of using models aligned with the operating constraints of UAV-based surveillance systems.

4.7. Object Resolution Capability

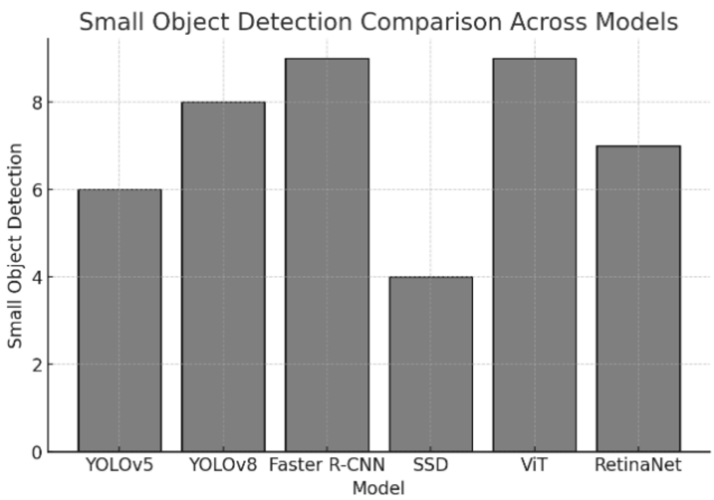


Fig- 4. Small Object Detection comparison bar graph across all models

The bar chart titled "Small Object Detection Comparison Across Models", shown in Fig. 4, presents a comparative evaluation of six object detection models based on their effectiveness in identifying small objects, which is essential in UAV-based runway monitoring where foreign object debris or distant intrusions often appear at low resolution. According to the chart, Faster R-CNN and Vision Transformer (ViT) exhibit the highest performance, each scoring nine, indicating their superior capacity to capture fine-grained visual features. YOLOv8 follows with a score of eight, reflecting advancements over YOLOv5, which scored six, due to improvements in detection architecture and feature representation. Retina Net achieves a balanced score of seven, demonstrating its ability to maintain detection performance through mechanisms like focal loss. SSD, on the other hand, shows the lowest performance with a score of four, suggesting limitations in resolving smaller targets within complex aerial scenes. This analysis highlights that transformer-based and two-stage detection models are more effective in small object recognition tasks and are better suited for ensuring runway safety through precise UAV-based obstacle detection.

4.8. Robustness to Aerial Perspective

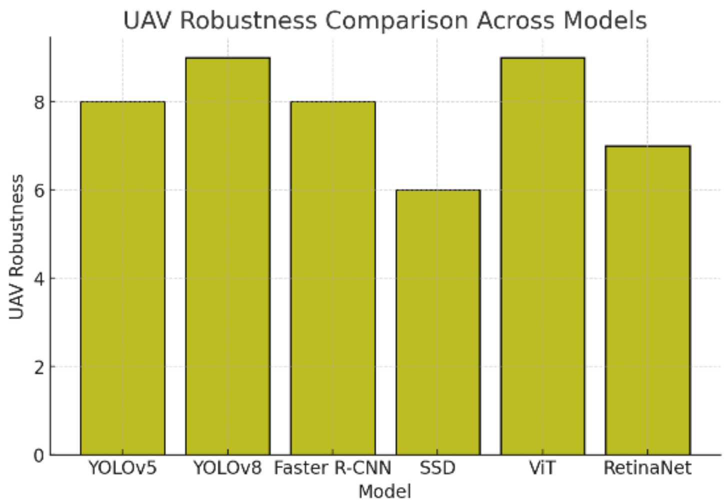


Fig. 5. UAV Robustness comparison bar graph across all models

The UAV robust comparison table, as presented in Figure 5, underscores the level at which various object detection models hold up to standard in the kind of variability seen in UAV-based imaging, including changing altitude, angles, motion blur, and outdoor lighting. Both YOLOv8 and ViT perform best, reflecting optimal resilience against aerial inconsistency and robust flexibility in dynamic drone-captured scenes. YOLOv5 and Faster R-CNN closely trail behind with slightly lower scores, proving consistent robustness but possibly more sensitivity to some dynamic conditions. Retina Net, with a medium score, works steadily across varied UAV clips, albeit not as responsively as transformer models. SSD is last, as expected, due to its issues with coping with the changing views and conditions of actual real-time UAV operations. These observations strengthen the benefit of anchor-free and transformer-based models for high-fidelity surveillance applications in airborne platforms.

4.9. Overall Model Comparison of UAV-Based Runway Obstacle Detection

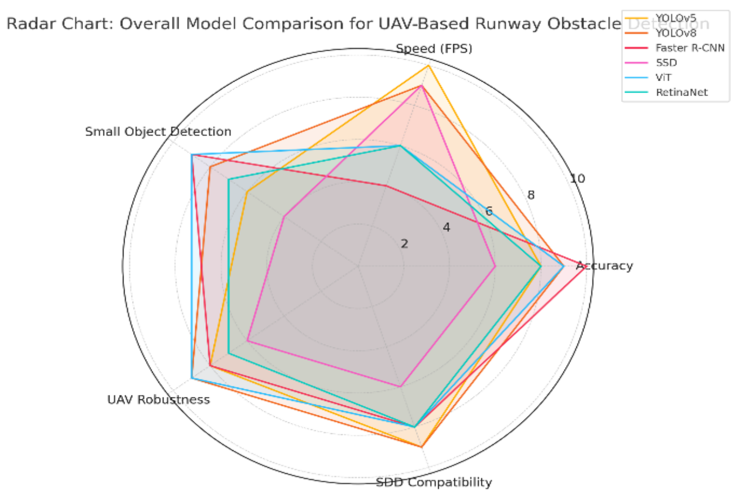


Fig. 6. Overall Model Comparison of UAV-Based Runway Obstacle Detection

The radar chart in Figure 6, Overall Model Comparison for UAV-Based Runway Obstacle Detection is a multi-criteria visual comparison of six widely used object detection models—YOLOv5, YOLOv8, Faster R-CNN, SSD, Vision Transformer (ViT), and Retina Net—on five performance aspects: detection accuracy, speed (frames per second), detection of small objects, robustness of UAV, and SDD compatibility.

Based on the chart, YOLOv8 shows one of the best-balanced and high-performing profiles. It achieves excellent scores in all five categories, particularly in detection precision, detection of small objects, and robustness against UAVs, with outstanding speed and compatibility with SDD. Its consistently large size on the radar plot indicates that it is one of the most generalizable options for real-time UAV-based runway monitoring.

Faster R-CNN has the maximum attainable score in detection accuracy and excels in small object detection and UAV robustness. The limitation lies in its speed (FPS), with a minimal range on that axis, confirming that it is not made for real-time applications despite its accuracy.

YOLOv5 is the fastest among all the models, leading on the FPS axis, which is vital for real-time observation. It also shows decent performance across various categories, especially in SDD compatibility and UAV robustness, thereby being a suitable trade-off model when accuracy as well as speed of processing are essential.

ViT models have great potential in small object detection and UAV resilience, primarily because of their self-attention, but lag in terms of speed, which means higher computational expenses. Their radar profile is almost symmetrical, showing robust contextual knowledge and, at the same time, resource richness.

Retina Net achieves a balanced performance with moderate to high marks in every aspect, particularly in small object detection and SDD compatibility. Nevertheless, it does not perform optimally in any one area, and the implication is that it would be helpful to use it more as a second- or benchmark-level model than as a first-line model to be deployed.

Conversely, SSD has the most limited coverage on the radar plot. It is relatively good at speed but lacks performance in detecting small objects and UAV robustness. Its limited region indicates a lower overall quality for UAV-based obstacle detection tasks, particularly in challenging or dynamic settings such as airport runways.

As a whole, the radar plot accurately conveys the performance trade-offs between various object detection models. It emphasizes the need for selecting a model that is appropriate for the particular operational requirements—whether that is real-time speed, detection accuracy, or resilience to UAV imagery variability. The reasoning behind this analysis is that YOLOv8 and ViT are exceptional in terms of overall performance. However, Faster R-CNN is more suitable for offline accuracy-critical applications, and YOLOv5 has the most favorable real-time processing benefit.

5. DISCUSSION AND FUTURE WORK

5.1. Discussion

The conclusions derived from this methodological review highlight the potential and existing limitations of computer vision systems based on UAVs for detecting obstacles on airport runways. By highlighting the state of the art in deep learning techniques coupled with aerial video datasets like the Stanford Drone Dataset (SDD), it becomes clear that most of the building blocks for efficient runway monitoring are present in the public domain. A few technical and contextual loopholes remain to be addressed.

While YOLO-family models (especially YOLOv8) and Vision Transformers exhibit robust performance in object detection tasks with aerial perspectives, detecting small and stationary foreign object debris (FOD), which is often crucial for runway safety, is problematic for them. Such models must also be optimized for high-confidence detection under challenging lighting, weather fluctuation, and varied terrain texture, frequent conditions in real-life airport environments.

The Stanford Drone Dataset has been a valuable substitute dataset, particularly in training and testing object detection and tracking pipelines in organized open spaces. It lacks domain-specific aspects like Aircraft presence and proximity interactions, Runway markings, lighting, or signs, Typical FOD types (e.g., luggage components, wildlife, screws, tools).

Consequently, although SDD makes prototype development possible, deployment to operational airports demands domain adaptation, bespoke data augmentation, or additional real-world UAV recordings.

In addition, there isn't much benchmark data on obstacle detection for airports. Most of the current UAV image datasets (e.g., Vis Drone, UAVDT, DOTA) are designed for urban or surveillance settings. Thus, even state-of-the-art models, when trained on such datasets, might perform suboptimally in real airside scenarios because of domain shift.

Lastly, the review highlights a requirement for holistic systems that integrate detection, tracking, geofencing, and alert generation into a deployable real-time module. There is little literature presently that treats standalone detection algorithms or their integration with airport safety measures or ATC systems.

5.2. Future Work

To plug the gaps identified, we suggest the following future research directions:
1. Airport-Runway-Specific UAV Dataset Development
Development of a labeled dataset using real UAV video from runways, taxiways, and apron areas, along with classes of obstacles (e.g., wildlife, equipment, debris, personnel), will significantly increase detection precision and model resilience.

2. Multi-Sensor Fusion Systems
Integration of RGB UAV imagery with thermal, LiDAR, or radar inputs can facilitate enhanced obstacle detection, especially in low-visibility or nocturnal operations.

4. Model Optimization for Embedded Deployment
Lightweight models (such as YOLO-Nano, MobileNet variants) need to be tested for real-time use on UAV or edge GPU platforms for autonomous decision-making without cloud reliance.

5. Self-Supervised Pretraining for Domain Adaptation
Pretraining models with the unlabeled airside video and subsequent fine-tuning using smaller labeled airport datasets could enhance generalization in the limited data case.

5. Trajectory and Behavior Prediction for Proactive Safety
Merging object trajectory estimation (facilitated through platforms such as SDD) with geofenced areas may facilitate early warning systems for would-be incursions.

6. Interoperability with Air Traffic Control Systems
Future systems must be compatible with current ATC or airside management software, enabling dynamic alerts, hand-off of objects, and event logging.

6. CONCLUSION

The integrity of airport runways hinges critically on early and accurate detection of any obstacles like vehicles, pedestrians, wildlife, and foreign object debris (FOD). With the emergence of UAV-based surveillance and advances in computer vision, there is enormous potential for intelligent, real-time runway monitoring systems. This article offered a methodological overview of UAV imagery for runway obstacle detection with emphasis on the usability of the Stanford Drone Dataset (SDD) as a stand-in for airside surveillance settings.

We inspected and contrasted the latest object detection models, including YOLOv5/YOLOv8, Faster R-CNN, and Vision Transformers, based on detection accuracy, resilience, and appropriateness for overhead UAV views. Our comparative study demonstrated that although there are current models that can adequately address structured outdoor environments, modifications are necessary for efficient use within airport contexts, especially where small objects, changing light conditions, and operating realism are issues.

The Stanford Drone Dataset, while aviation-nonspecific, gives us a dense, annotated video corpus that can be used well to simulate runway scenarios and evaluate pipeline detection. Its organized environments, varied object classes, and regular overhead views make it an asset to use for prototyping.

We outlined a conceptual methodology and pipeline describing how UAV video input can be processed via detection, tracking, and alerting modules to aid runway safety operations. While there are numerous promising technologies, the necessity for airport-specific datasets, embedded deployment-optimized models, and integration with real-time airport systems remains urgent.

This survey is a starting point for researchers and stakeholders in airport safety to develop, test, and implement UAV-based obstacle detection systems for runways. Future research should focus on establishing domain-specific datasets, investigating data fusion from multiple sensors, and extending horizon-adaptive learning models to address the stringent requirements of real-world airport safety operations.

7. ACKNOWLEDGMENT

The authors would like to thank the Karunya Institute of Technology and Sciences for providing them with excellent opportunities and support to pursue this research work.

REFERENCES

[1] Abu-Jbara, A., et al. (2014). *A robust vision-based runway detection and tracking algorithm for automatic UAV landing*. ResearchGate. https://www.researchgate.net/publication/282982367
[2] Alonso, P., Iñiguez de Gordoa, J. A., Ortega, J. D., García, S., Iriarte, F. J., & Nieto, M. (2024). *Automatic UAV-based airport pavement inspection using mixed real and virtual scenarios*. arXiv. https://arxiv.org/abs/2401.06019
[3] FAA/National Geodetic Survey. (2023). *Evaluation of unmanned aircraft systems for airport obstacle data collection*. FAA Technical Report. https://www.airporttech.tc.faa.gov
[4] Lalak, M., & Wierzbicki, D. (2022). Automated detection of atypical aviation obstacles from UAV images using a YOLO algorithm. *Sensors, 22*(17), 6611. https://doi.org/10.3390/s22176611
[5] Li, Y., Xia, Y., Zheng, G., Guo, X., & Li, Q. (2024). YOLO-RWY: A novel runway detection model for vision-based autonomous landing of fixed-wing uncrewed aerial vehicles. *Drones, 8*(10), 571. https://doi.org/10.3390/drones8100571
[6] Mirhajianmoghadam, H., & Bolourian Haghighi, B. (2022). EYNet: Extended YOLO for airport detection in remote sensing images. arXiv. https://arxiv.org/abs/2203.14007
[7] Munyer, T., Brinkman, D., Zhong, X., Huang, C., & Konstantzos, I. (2022). Foreign object debris detection for airport pavement images based on self-supervised localization and Vision Transformer. arXiv. https://arxiv.org/abs/2210.16901

[8] Robicquet, A., Sadeghian, A., Alahi, A., & Savarese, S. (2016). Learning social etiquette: Human trajectory prediction in crowded scenes. *European Conference on Computer Vision (ECCV)*. https://cvgl.stanford.edu/projects/uav_data/

[9] Seidaliyeva, U., & Ilipbayeva, L. (2023). Advances and challenges in drone detection and classification techniques: A state-of-the-art review. *Sensors, 23*(2), 456. https://doi.org/10.3390/s23020456

[10] Stanford Computational Vision and Geometry Lab. (2016). *Stanford Drone Dataset*. https://cvgl.stanford.edu/projects/uav_data/

[11] Thai, P., Alam, S., Lilith, N., Tran, P. N., & Thanh, B. N. (2020). Deep4Air: A novel deep learning framework for airport airside surveillance. arXiv. https://arxiv.org/abs/2010.00806

[12] Tsapparellas, K., Waters, J., Mihaylova, L. S., et al. (2023). Vision-based runway detection and landing for uncrewed aerial vehicles. *Proceedings of the International Conference on Mechatronics and Automation (ICMA)*. https://eprints.whiterose.ac.uk/200327/

[13] Tuzikov, A. (2024). *AirWall: Malicious drones detection using YOLO*. CS231n Project Report, Stanford University. https://cs231n.stanford.edu/2024/papers/airwall-malicious-drones-detection-using-yolo.pdf

[14] Wang, Q., & Wang, H. (2015). Airport signs and markings recognition for enhanced runway incursion avoidance. *EE368 Class Project, Stanford University*. https://web.stanford.edu/class/ee368/Project_Spring_1415/Reports/Wang_Wang.pdf

[15] Wang, L. (2020). Design of an airport obstacle-free zone monitoring UAV system based on computer vision. *Sensors, 20*(11), 3123. https://doi.org/10.3390/s20113123

[16] Yuan, L., Zhang, Y., & Guo, X. (2021). A review on uncrewed aerial vehicle remote sensing: Vision, processing, and applications. *Drones, 7*(6), 398. https://doi.org/10.3390/drones7060398

[17] Zhou, T., & Wu, Z. (2022). Stanford Drone Dataset object detection project. *Roboflow*. https://universe.roboflow.com/abp-e6yxw/stanford-drone-dataset-xhhtl

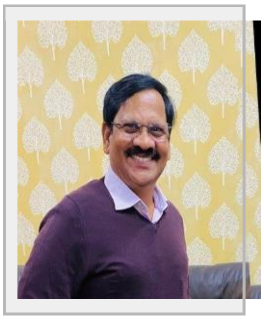
[18] Lalak, M., & Wierzbicki, D. (2022). Methodology of detection and classification of selected aviation obstacles based on UAV dense image matching. *ResearchGate*. https://www.researchgate.net/publication/358419403

[19] Joshua J Andle. et al. (2023). Stanford Drone Dataset is more complex than we think: An analysis of key characteristics. *ResearchGate*. https://www.researchgate.net/publication/359927030

[20] Didier Aldana Rodríguez, et al. (2024). Inspection of aircraft and airports using UAS: A review. *ScienceDirect*. https://www.sciencedirect.com/science/article/pii/S2590123024005851

[21] Wei, Z., Meng, Z., Lai, M., Wu, H., Han, J., & Feng, Z. (2021). Anti-collision technologies for uncrewed aerial vehicles: Recent advances and future trends. arXiv. https://arxiv.org/abs/2109.12832

BIOGRAPHIES

	**Joseph Chakravarthi Chavali** is an accomplished professional with over 38 years of extensive leadership experience, including three decades at Air India, specializing in aviation safety, operations training, and regulatory compliance. He holds advanced qualifications in Electronics Engineering, Digital Systems, and Human Resource Management, and is currently pursuing a Ph.D. in Image Processing at Karunya University, Coimbatore. As a published researcher in reputed Scopus and IEEE-indexed platforms, his expertise includes integrating Environmental, Social, and Governance (ESG) principles into strategic governance and corporate oversight.
	**Dr. D. Abraham Chandy** is an Associate Professor at Karunya University, Coimbatore, with expertise in Digital Image Processing, Artificial Intelligence, and Machine Learning. He has authored numerous high-impact publications in medical image analysis, edge computing, and optimization, with over 300 citations and an h-index of 9. His contributions include pioneering work in mammogram retrieval, retinal image processing, and deep learning-based ultrasound segmentation. Dr. Chandy is a verified researcher and active academic mentor with a strong presence in IEEE and other reputed platforms.