Preprints.org

Article

Advances in Aircraft Skin Defect Detection Using Computer Vision: A Survey and Comparison of YOLOv9 and RT-DETR Performance

Nutchanon Suvittawat , Christian Kurniawan , Jetanat Datephanyawat , Jordan Tay , ZhiHao Liu , De Wen Soh , Nuno Antunes Ribeiro [*]

Article

Advances in Aircraft Skin Defect Detection Using Computer Vision: A Survey and Comparison of YOLOv9 and RT-DETR Performance

Nutchanon Suvittawat [1,†], Christian Kurniawan [1,†], Jetanat Datephanyawat [2,†], Jordan Tay [3], ZhiHao Liu [2], De Wen Soh [1,†] and Nuno Antunes Ribeiro [2,3,*]

[1] Information Systems Technology and Design, Singapore University of Technology and Design, Singapore

[2] Engineering Systems and Design, Singapore University of Technology and Design, Singapore

[3] Aviation Studies Institute, Singapore University of Technology and Design, Singapore

[*] Correspondence: nuno_ribeiro@sutd.edu.sg

[†] These authors contributed equally to this work.

Abstract: Aircraft skin surface defect detection is critical for aviation safety but is currently mostly reliant on manual or visual inspections. Recent advancements in computer vision offer opportunities for automation. This paper reviews the current state of computer vision algorithms and their application in aircraft defect detection, synthesizing insights from academic research (21 publications) and industry projects (18 initiatives). Beyond a detailed review, we experimentally evaluate the accuracy and feasibility of existing low-cost, easily deployable hardware (drone) and software solutions (computer vision algorithms). Specifically, real-world data were collected from an abandoned aircraft with visible defects using a drone to capture video footage, which was then processed with state-of-the-art computer vision models — YOLOv9 and RT-DETR. Both models achieved mAP50 scores of 0.70–0.75, with YOLOv9 demonstrating slightly better accuracy and inference speed, while RT-DETR exhibited faster training convergence. Additionally, a comparison between YOLOv5 and YOLOv9 revealed a 10% improvement in mAP50, highlighting the rapid advancements in computer vision in recent years. Lastly, we identify and discuss various alternative hardware solutions for data collection — in addition to drones, these include robotic platforms, climbing robots, and smart hangars — and discuss key challenges for their deployment, such as regulatory constraints, human-robot integration, and weather resilience. The fundamental contribution of this paper is to underscore the potential of computer vision for aircraft skin defect detection while emphasizing that further research is still required to address existing limitations.

Keywords: aircraft defect detection; computer vision; maintenance repair and overhaul (MRO); YOLO; RT-DETR

1. Introduction

Commercial airplanes undergo regular inspections and maintenance checks to meet operational safety standards and regulatory requirements before flying. During General Visual Inspection (GVI) tasks, Maintenance, Repair, and Overhaul (MRO) operators carefully examine airplanes for defects that could compromise safety and performance [1,2]. These inspections focus on critical aircraft components, including the fuselage, structural elements, landing gear, tires, and control surfaces. Identifying defects in the aircraft's exterior (or skin), such as cracks, corrosion, dents, or missing rivets, is particularly relevant for maintaining aircraft's structural integrity, preventing further deterioration, and mitigating risks to underlying systems.

The degree of each inspection varies based on the aircraft's operational phase. Pre-flight and post-flight inspections, conducted before and after every flight during the aircraft's turnaround, are typically brief and targeted, lasting approximately 10 to 30 minutes depending on the aircraft type

and available turnaround time. In addition to these rapid checks, more extensive routine inspections, categorized from A to D checks [1], are mandated by aviation regulations and occur at varying intervals [3]. Type A checks are performed every 1 to 2 months and take approximately 6 to 12 hours, while D checks, the most comprehensive, occur every 6 to 10 years and can span 3 to 7 days [2]. Type B and C checks fall somehow in between [3].

Currently, aircraft surface inspections are performed manually by teams of licensed MRO operators and engineers who must undergo extensive training and certification before being qualified for such task. While these inspections require a high level of precision, they are also physically demanding, especially when accessing challenging areas such as the upper surfaces of airplanes. Additional equipment, such as scaffolds or aerial platforms, are often necessary, adding to the cost, time, and labor-intensive nature of the process [4,5]. Moreover, manual inspections are prone to human error and inconsistencies, which can affect the overall reliability of defect detection [6].

To address these limitations, industry and academia are increasingly focused on integrating computer vision technologies into defect detection processes [7,8]. This push has given rise to a variety of innovative solution such as: (i) Unmanned Aerial Vehicles (UAVs) for capturing real-time images of hard-to-reach defects; (ii) mobile or climbing robots for navigating aircraft surfaces, or (iii) smart hangars equipped with advanced camera systems to capture images from multiple angles.

Academic research in this field has seen rapid growth, particularly in the development and application of object detection and image segmentation algorithms for identifying surface defects on aircraft. A review of existing literature, summarized in Table 1 (Section 2), identified 21 research papers focused on such algorithms. Notably, about half of these studies were published in 2024 (the year this paper is being written), highlighting the growing urgency and relevance of this topic.

In parallel, the aviation industry has been investing in research and development to operationalize computer vision-based inspection systems. Projects span both hardware, such as UAVs and robotic systems, and software, including advanced computer vision algorithms. Our review, summarized in Table 2 (Section 2), uncovered 18 publicly documented industry projects across different companies and regions. This likely represents only a fraction of ongoing efforts, as many initiatives are not disclosed publicly.

The surge in interest in automated aircraft defect skin inspection can be attributed to recent advancements in computer vision technologies, which have significantly improved accuracy and efficiency within a short period of time — for a detailed benchmark analysis of various object detection algorithms since 2015, refer to [9]. Moreover, the post-COVID-19 period has introduced labor challenges, with many licensed inspectors retiring or transitioning to other roles, further driving the need for automated solutions in aircraft inspection.

The aim of this research is to explore the current state-of-the-art in hardware and software solutions for aircraft skin defect detection. Beyond conducting a detailed review of academic literature and industry innovations, we also experimentally assess the accuracy and feasibility of existing low-cost (both in terms of financial cost but also ease of deployment) hardware and software solutions. To that end, real-world data was collected from an abandoned aircraft with visible defects using a drone to capture video footage. The footage was then processed using widely adopted computer vision algorithms, specifically YOLOv9 [10] and RT-DETR [11], and their accuracy was evaluated.

In summary, the main contributions of this paper are as follows:

---

[1] Additional inspections may be necessary if reports from pilots or sensor data indicate potential defects or anomalies during operation, or following severe events.

[2] These intervals are general estimates and may vary based on factors such as pilot- or crew-reported issues, aircraft usage, age, and type.

[3] B checks are typically performed every 800 to 1,500 flight hours (approximately 6 to 8 months), spanning 1 to 3 days of inspection. C checks occur every 3,000 to 6,000 flight hours (approximately 18 months to 2 years.)

1. *Thorough review of existing hardware and software solutions for aircraft defect detection:* The investigation highlights significant growth in academic research on computer vision for aircraft defect detection. Notably, all academic references are dated from 2019 onward, with over half of the publications appearing in 2024. On the industry side, numerous projects have been publicly disclosed since 2014, with half launched during or after the COVID-19 pandemic. These projects span various regions and industry sectors, including airlines (e.g., EasyJet, Air France, KLM, Singapore Airlines, Delta Airlines, Lufthansa), aircraft manufacturers (e.g., Airbus, Boeing), and organizations in the field of aerospace, robotics and artificial intelligence (e.g., Rolls-Royce, MainBlades, Donecle, SR Technics, ST Engineering, Jet Aviation, HACARUS) as well as research institutes (A*STAR).

2. *Collection of a new dataset of aircraft defects using a drone from an abandoned aircraft:* Using an affordable and commercially available Parrot Bebop Quadcopter Drone (14-megapixel full HD 1080p, 180° field of view), we demonstrate that current readily available technology can achieve accurate aircraft defect detection. We validate this hypothesis by investigating a preserved Boeing 747 with defects located in Korat, Nakhon Ratchasima, Thailand. The collected dataset has been made publicly available (see here [12]) to support testing and benchmarking of future algorithms for aircraft defect detection. To the best of our knowledge, although several other studies have collected defect data from drone imagery [13–18], their datasets remain inaccessible due to confidentiality. Existing publicly available datasets primarily focus on defect-specific images (e.g., [19]), often emphasizing close-up views of defects rather than encompassing the entire aircraft (with and without defects) or accounting for variations in image quality caused by distance and angle. The lack of publicly available data is a frequently cited challenge in research papers on this topic [18,20].

3. *Evaluation and Comparison of Leading Computer Vision Algorithms for Aircraft Defect Detection:* This work tests and compares YOLOv9 and RT-DETR, two of the most accurate computer vision models (refer to [9]). Our implementation of these algorithms for aircraft defect detection contributes in two major ways: (i) it offers a benchmark analysis highlighting the relative strengths of these models. The results show comparable overall performance, with RT-DETR slightly outperforming YOLOv9 prior to hyperparameter tuning, whereas YOLOv9 achieves superior results post-tuning; (ii) we demonstrate the practicality of leveraging modern computer vision algorithms for real-world aircraft defect detection, achieving moderate accuracy even without extensive tuning: mAP50 scores range from 0.699 for RT-DETR to 0.685 for YOLOv9. Additionally, a comparison with the YOLOv5 Extra-Large (YOLOv5x) model [21], developed in 2020, reveals a 10% improvement in mAP50 with YOLOv9, emphasizing the rapid advancements in computer vision.

4. *Hyperparameter tuning using Bayesian optimization:* To avoid a myopic, single implementation of these algorithms, we applied hyperparameter tuning and analyzed its impact on performance metrics, particularly precision, recall, and mAP50. Our results show that tuning improves accuracy by 1.6% for RT-DETR and 6.6% for YOLOv9 across two datasets analysed — our drone-collected dataset [12] and a publicly available dataset [19]. We demonstrate that performance can be tailored to specific priorities by emphasizing different metrics during tuning. To encourage further research and facilitate refinement of the presented models, we have made our codes publicly available in [22].

5. *Discussion of limitations and challenges in implementing computer vision for aircraft defect detection:* Our findings highlight the potential of computer vision to enhance defect detection, particularly given the rapid evolution of computer vision algorithms. However, discussions with industry experts reveal significant challenges that remain. These include: (i) regulatory constraints that vary across countries, hindering the development of standardized solutions; (ii) the need for seamless human-robot collaboration, as automation is intended to complement rather than replace human inspectors; (iii) the necessity of achieving exceptionally high accuracy levels to meet stringent safety standards. Current mAP50 levels of 0.70–0.75 may still be insufficient due

to safety reasons. These findings emphasize the promise of this technology while motivating continued advancements to overcome these challenges.

The remainder of this paper is organized as follows: we begin with a review of existing work on computer vision for aircraft skin defect detection, highlighting publicly available projects from various industries in Section 2. Next, we describe the data collection process and provide details about the datasets used in this study in Section 3. The computer vision models employed for defect detection are introduced in Section 4, along with the metrics for evaluating algorithm performance in Section 4.3 and the hyperparameter tuning procedure in Section 4.4. In Section 5, we compare and analyze the performance of the object detection models, including the impact of hyperparameter tuning. We then delve into the challenges and important research directions in this field in Section 6. Finally, we conclude the paper in Section 7.

2. Research and Industry Survey

Our review identified a total of 21 research papers, with 14 published after 2020, illustrating the rapid growth in the field. These studies employ various algorithms, ranging from direct implementations of various versions of YOLO (though version 9, used in this paper, was not implemented) [16,17,20,23–28] to more advanced neural network models. For example, some researchers have enhanced existing architectures with additional layers, such as mask-scoring Region-based Convolutional Neural Network (R-CNN) [15] and U-Net [29]. A summary of these papers is presented in Table 1. The papers also address a wide range of defect types, including stains, lightning burns, dents, rust, cracks, corrosion, scratches, damaged and missing rivets or screws, among others.

Note that these papers focus on defects visible on the aircraft's surface or fuselage, aligning with the scope of this review. More detailed inspections, which require Non-Destructive Testing (NDT) methods such as ultrasonic or radiographic testing, are typically conducted during B to D checks to assess internal structures, engines, or overlapping panels [30–33]. Additionally, other studies have explored object detection algorithms for identifying manufacturing defects [34–37], detecting ice formation on critical components [38], aircraft ducts defects [39], and verifying the condition of aircraft components, such as ensuring vents are closed, valves are properly positioned, and static ports are uncovered [40,41]. While these studies are not included in Table 1, which focuses on defects in visible aircraft skin, we acknowledge their relevance and importance as part of the broader inspection process.

Table 1. Related academic research for advanced aircraft defects inspection.

References (Year)	Description
[42] (2019)	**Scope:** Utilizing a reconfigurable climbing robot for aircraft surface defect detection. **Data Source:** RGB images taken by climbing robot. **CV Algorithm:** Enhanced SSD MobileNet. **Focus:** Detecting between aircraft stain and defect.
[13] (2019)	**Scope:** Investigating defect classification on aircraft fuselage from UAV images. **Data Source:** Images taken using a drone. **CV Algorithm:** Combination of CNN and few-shot learning methods. **Focus:** Detecting paint defects, lightning burns, screw rashes, rivet rashes.
[24] (2019)	**Scope:** Adapting YOLO3 for faster crack detection in aircraft structures. **Data Source:** Images from aviation company and industrial equipment. **CV Algorithm:** YOLOv3-Lite. **Focus:** Aircraft cracks.
[14] (2020)	**Scope:** Developing a deep neural network (DNN) to detect aircraft defects. **Data Source:** Images taken using drone. **CV Algorithm:** AlexNet and VGG-F networks with a SURF feature extractor.

	Focus: Detecting fuselage defects (binary classification).
[43]	**Scope:** Developing a Mask R-CNN to detect aircraft dents.
	Data Source: Images from publicly available sources and within the hangar.
(2020)	**CV Algorithm:** Mask R-CNN.
	Focus: Aircraft dents.
[44]	**Scope:** Improving MASK R-CNN with augmentation techniques.
	Data Source: Images from publicly available sources and within the hangar.
(2020)	**CV Algorithm:** Pre-classifier in combination with MASK R-CNN.
	Focus: Wing images with and without aircraft dents.
[23]	**Scope:** Applying and Comparing YOLOv3 with Faster-RCNN.
	Data Source: Self-collected dataset of aircraft defect images.
(2020)	**CV Algorithm:** YOLO Neural Network and Faster-RCNN.
	Focus: Detecting: skin crack, thread corrosion, skin peeling, skin deformation, and skin tear.
[15]	**Scope:** Developing an improved mask scoring R-CNN.
	Data Source: Images taken using a drone.
(2022)	**CV Algorithm:** Mask Scoring R-CNN for instance segmentation with attention and feature fusion.
	Focus: Detecting paint falling and scratch.
[25]	**Scope:** Examining the use of YOLOv5 to detect defects on aircraft surface.
	Data Source: Self-collected dataset of aircraft defect images.
(2023)	**CV Algorithm:** YOLOv5.
	Focus: Detecting: cracks, dent, missing screws, peeling, and corrosion.
[45]	**Scope:** Exploring the use of CNNs for aircraft defect detection.
	Data Source: Images taken by drones and blimps.
(2024)	**CV Algorithm:** CNN.
	Focus: Detecting: missing rivet, corroded rivet.
[16]	**Scope:** Exploring the use of YOLOv8 for aircraft defect detection.
	Data Source: Images taken using drone, phone, and camera.
(2024)	**CV Algorithm:** YOLOv8.
	Focus: Detecting: panel missing, rivets/screws damaged or missing.
[26]	**Scope:** Inspecting hyperparameter tuning of YOLOv8 for aircraft defect detection.
	Data Source: Scanning electron microscopy (SEM) images of surface parts.
(2024)	**CV Algorithm:** YOLOv8 with MobileNet backbone and Bayesian optimization.
	Focus: Detecting: gap, microbridge, line collapse, bridge, rust, scratch, and dent.
[17]	**Scope:** Proposing an improved YOLOv5-based model for aircraft defect detection.
	Data Source: Publicly available images and images captured using a drone.
(2024)	**CV Algorithm:** YOLO model with deformable convolution, attention mechanisms, and contextual enhancement.
	Focus: Detecting: crack, dent, missing rivet, peeled paint, scratch, missing caps, and lost tools.
[29]	**Scope:** Proposing a bio-inspired CNN model for low-illumination aircraft skin defect detection.
	Data Source: Images taken by an airline.
(2024)	**CV Algorithm:** U-Net with residual blocks and attention mechanisms.
	Focus: Detecting image pixels with defects.
[46]	**Scope:** Applying YOLOv8 to aircraft defect detection.
	Data Source: Images taken using a drone.
(2024)	**CV Algorithm:** YOLOv8.
	Focus: Detecting: rust, scratches, and missing rivets.
[18]	**Scope:** Testing various deep learning architectures for aircraft defect detection.
	Data Source: Drone and handheld cameras used inside the hangar.
(2024)	**CV Algorithm:** Various CNN-based models with custom size-estimation.

	Focus: Detecting: dents, missing paint, screws, and scratches.
[20]	**Scope:** Improving YOLOv7 for aircraft defect detection.
	Data Source: Images collected using a mobile platform camera and publicly available
(2024)	datasets.
	CV Algorithm: FC-YOLO with feature fusion strategies.
	Focus: Detecting: paint peel, cracks, deformation, and rivet damage.
[27]	**Scope:** Improving YOLOv8n for detecting small aircraft defects.
	Data Source: Images collected using DJI OM 4 SE stabilizers paired with mobile
(2024)	cameras.
	CV Algorithm: Improved YOLOv8n with Shuffle Attention and BiFPN.
	Focus: Detecting: cracks, corrosion, and missing labels.
[47]	**Scope:** Data augmentation using Fourier GAN to improve defect detection accuracy.
	Data Source: Drone-collected and synthetic images.
(2024)	**CV Algorithm:** Fourier GAN for data augmentation.
	Focus: Detecting: loose components, corrosion, and skin damage.
[28]	**Scope:** Improving YOLOv8 for robust defect detection.
	Data Source: Publicly available images.
(2024)	**CV Algorithm:** YOLOv8 with CoTAttention and SPD-Conv modules.
	Focus: Detecting: cracks, dents, and rust.
[48]	**Scope:** Developing DyDET, a lightweight semi-supervised aircraft defect detection
	framework.
(2024)	**Data Source:** Images collected using camera.
	CV Algorithm: DyDET with dynamic attention and adaptive pseudolabeling.
	Focus: Detecting: scratches, paint-peeling, rivet damage, and rust.

The industry's adoption of computer vision for defect detection is evident from the 18 publicly documented projects identified across different sectors. These initiatives span airlines (e.g., EasyJet, Air France, KLM, Singapore Airlines, Delta Airlines, Lufthansa), aircraft manufacturers (e.g., Airbus, Boeing), and organizations specializing in aerospace, robotics, and artificial intelligence (e.g., Rolls-Royce, MainBlades, Donecle, SR Technics, ST Engineering, Jet Aviation, HACARUS). Research institutes such as A*STAR are also actively contributing to advancements in this area. These projects not only focus on developing more robust object detection algorithms but also explore innovative hardware solutions for image and video collection. Emerging technologies include UAVs, drone swarms, mobile robotic platforms, climbing robots, and smart hangars equipped with high-quality cameras. Table 2 summarize these projects, detailing the hardware solutions employed and their respective applications.

Table 2. Existing Related Solutions and Projects.

References (Year)	Description
[49]	**Scope:** A French R&D project developing an autonomous wheeled robot capable of inspecting aircraft during maintenance operations.
(2013)	**Solution Type:** Robot Platform
	Offered Benefits: Enhances inspection reliability and repeatability; reduces inspection time and maintenance costs.
[50]	**Scope:** EasyJet integrates drones for aircraft inspection and explores 3D printing to reduce maintenance costs and downtime.
(2015)	**Solution Type:** Drones
	Offered Benefits: Speeds up inspection; reduces dependency on external suppliers for parts.
[51]	**Scope:** Automated visual inspections using autonomous drones with AI-powered software.

(2015)	**Solution Type:** swarm of drones **Offered Benefits:** Performs GVI, identifies lightning strikes, and provides close-up inspections more efficiently than manual processes.
[52] (2017)	**Scope:** Mainblades offers automated drone inspections for aircraft maintenance. **Solution Type:** Drones **Offered Benefits:** Detects defects such as dents, cracks, and lightning strikes; reduces inspection time and improves consistency.
[53] (2018)	**Scope:** Airbus uses autonomous drones equipped with high-resolution cameras to conduct visual inspections of aircraft surfaces. **Solution Type:** Drones **Offered Benefits:** Speeds up inspection (3 hours), enhances safety by reducing human exposure.
[54] (2018)	**Scope:** Rolls-Royce demonstrated multiple robotic solutions, including SWARM robots for visual inspections, FLARE for repairs, and remote boreblending robots for maintenance tasks. **Solution Type:** Swarm of drones **Offered Benefits:** Enables inspections and repairs without removing the engine; reduces downtime and maintenance costs.
[55] (2018)	**Scope:** Developed a robot to move on aircraft for inspection and deliver tools to difficult-to-access areas. **Solution Type:** Climbing Robot **Offered Benefits:** Tools can be delivered to engineers for improved maintenance efficiency.
[56] (2018)	**Scope:** A collaboration between Ubisense and MRO Drone to create an automated aircraft inspection and tool management system. **Solution Type:** Smart Hangar **Offered Benefits:** Decreases Aircraft on Ground (AOG) time by up to 90%; enhances efficiency and productivity in maintenance operations.
[57] (2019)	**Scope:** Developed automated drone inspection for exterior aircraft components on surfaces. **Solution Type:** Drones **Offered Benefits:** Reduce time and cost of aircraft maintenance. Can inspect both inside and outside of hangar.
[58] (2020)	**Scope:** Received CAAS authorization to use drones for aircraft inspections, ensuring faster and safer processes. **Solution Type:** Drones **Offered Benefits:** Reduces inspection time; authorized for regulatory compliance.
[59] (2022)	**Scope:** Korean Air's development of a drone swarm system to inspect aircraft exteriors. **Solution Type:** Swarm of drones **Offered Benefits:** Reduces inspection time by up to 60%; improves safety and accuracy; minimizes aircraft downtime.
[60] (2023)	**Scope:** GE Aerospace's soft robotic inchworm designed for on-wing jet engine inspection and repair. **Solution Type:** Climbing Robot **Offered Benefits:** Enables minimally invasive inspections; reduces maintenance burden; provides access to hard-to-reach engine areas.
[61] (2023)	**Scope:** A*STAR's Smart Automated Aircraft Visual Inspection System (SAAVIS) uses robots and hybrid AI for defect detection on airframes. **Solution Type:** Smart Hangar **Offered Benefits:** Reduces inspection time and human errors; detects over 30 defect types.

[62]	**Scope:** Employs robotic arms and AI to inspect aircraft engines and seat tracks, improving efficiency.
(2023)	**Solution Type:** Robot platform
	Offered Benefits: Saves over 780 man-hours annually; reduces staff fatigue and increases inspection accuracy.
[63]	**Scope:** AI-powered drone inspections for a range of aircraft maintenance needs.
	Solution Type: Drones
(2023)	**Offered Benefits:** Conducts inspections for aircraft structure, paint quality, and pre-purchase evaluations.
[64]	**Scope:** Delta is the first U.S. airline to use FAA-approved drones for maintenance inspections, improving speed and safety.
(2024)	**Solution Type:** Drones
	Offered Benefits: Speeds up inspection by 82%; reduces technician risk during inspections.
[65]	**Scope:** Boeing uses small drones with AI-based damage detection software to automate exterior aircraft inspections.
(2024)	**Solution Type:** Drones
	Offered Benefits: Cuts inspection time by half, improves accuracy, and reduces costs.
[66]	**Scope:** Tests drone-based inspections, 3D scanning, and exoskeletons for aircraft overhauls, aiming to streamline maintenance.
(2024)	**Solution Type:** Drones
	Offered Benefits: Speeds up maintenance, improves accuracy, and reduces costs.

3. Data Collection

For this study, we captured video images of an old Boeing 747 parked in the airplane park near Suranaree University of Technology (SUT), Thailand, shown in Figure 1a. We captured the video images with a Parrot Bebop 1 drone, shown in Figure 1b, with 1080p resolution using its 14-megapixel camera. Utilizing the image preprocessing tools available in Roboflow [67], we extracted 983 images from the 13 videos we took and manually annotated the visual defects into one of three defect categories, specifically, (1) missing-rivet, (2) rust, and (3) scratch. An example image with the three different defect types from this dataset can be seen in Figure 2. The annotated images were then randomly divided into 688 training, 197 validation, and 98 test images. We applied the Auto-Orient preprocessing option and resized the images into 640×640 pixels image size. The dataset is made publicly available and can be downloaded from [12]. For simplicity, we will call this dataset as Dataset I throughout this paper.

(a)   (b)

Figure 1. (a) An image of the old Boeing 747 parked in the airplane park near Suranaree University of Technology (SUT), Thailand, used in this study. (b) An image of the Parrot Bebop 1 drone employed in this study to capture the video footage of the old Boeing 747.
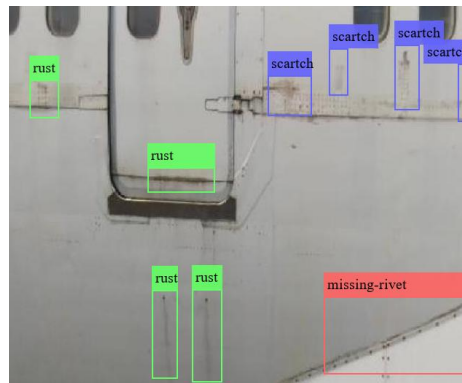
Figure 2. An example image from Dataset I with the three different defect types annotated manually.

In addition to the dataset described above, we also chose to use a publicly available dataset provided by [19] in our study to further verify the models. We refer to this additional dataset as Dataset II in this paper. The dataset consists of 9651 training, 642 validation, and 429 test images of multiple defective airplanes. The defects in the images are annotated and classified into one of the five defect categories, i.e., (1) crack, (2) dent, (3) missing-head, (4) paint-off, and (5) scratch. Examples of images for these five defect types in Dataset II can be seen in Figure 3.
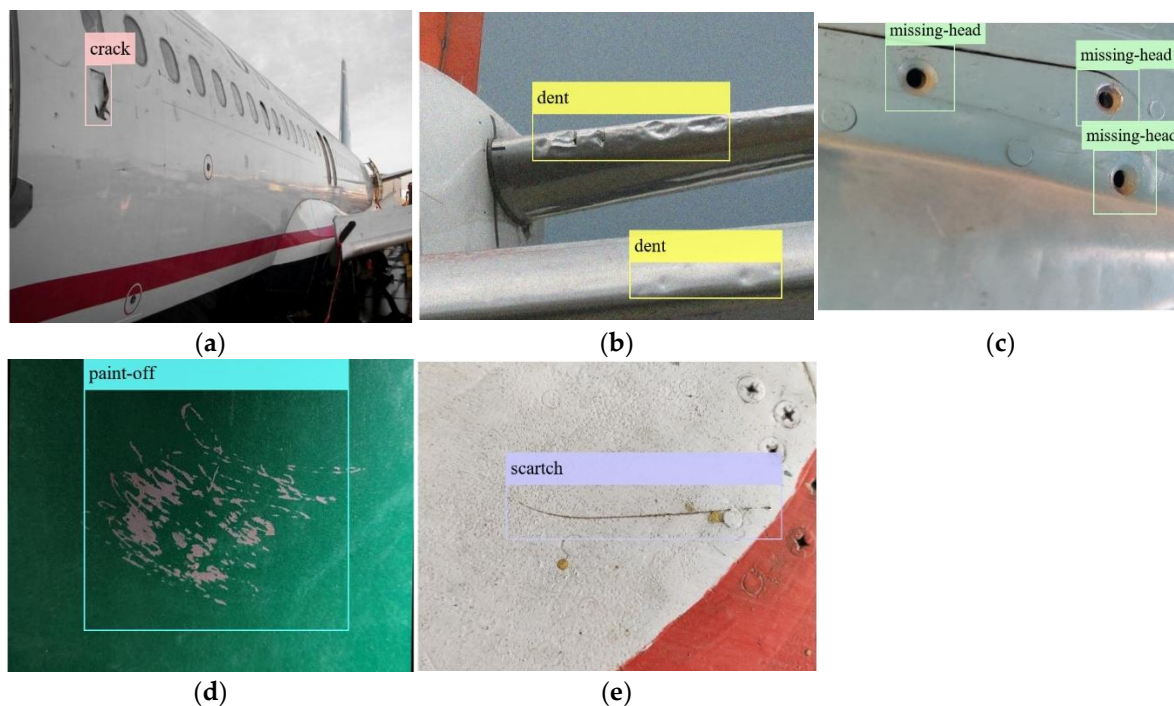
Figure 3. Example images from Dataset II with the five different defect types: (a) crack. (b) dent. (c) missing-head. (d) paint-off. (e) scratch.

4. Methodology

Object detection is a computer vision technique designed to identify and classify objects of interest within images. This capability aligns well with our goal of detecting and classifying defects on aircraft surfaces. With the development of various object detection models, numerous datasets have been proposed for benchmarking these models. The most widely used dataset for this purpose is the Microsoft Common Objects in Context (MS COCO) dataset, introduced by [68]. This dataset includes 180,000 training images, 5,000 validation images, and 41,000 test images, featuring millions of labeled objects across 80 categories. Performance comparisons of different object detection models

using the MS COCO dataset are detailed in [9]. For our study, we have chosen the two models well-known for their accuracy: RT-DETR and YOLOv9.

4.1. RT-DETR

DETR was first introduced by [69] in 2020, utilizing a transformer encoder-decoder architecture. Since its inception, DETR-based models have seen extensive development by various research groups (see [70] for a detailed review up to early 2023). To address the computational cost of DETR-based models, [11] proposed a new one-stage DETR variant called RT-DETR in late 2023, achieving real-time object detection. Their work introduced three key modifications to the original DETR model: (i) an efficient hybrid encoder was developed, combining Attention-based Intra-scale Feature Interaction (AIFI) and Convolutional Neural Network (CNN)-based Cross-scale Feature Fusion (CCFF) to alleviate computational bottlenecks; (ii) an uncertainty-minimal query selection scheme was implemented by integrating uncertainty into the loss function, reducing feature selection ambiguity; (iii) adjustments to the number of decoder layers were proposed, as the authors observed that changes in the number of decoder layers significantly impacted inference speed while maintaining accuracy. With these three modifications, RT-DETR reportedly outperformed earlier YOLO models and previous DETR series models in real-time object detection tasks on the MS COCO dataset. In this study, we implemented the RT-DETR Extra-Large (rtdetr-x) model available in [71].

4.2. YOLOv9

YOLO is one of the most well-known object detection models, recognized for its well-balanced tradeoff between efficiency and accuracy. The first version, YOLOv1 [72], introduced in 2015, was the first one-stage model to achieve true real-time object detection. Since then, YOLO has attracted widespread adoption. Numerous advancements and adaptations by different research groups have further improved its speed and accuracy while extending its capabilities to tasks such as image classification and segmentation. Readers interested in a detailed overview of the YOLO series models up to YOLOv8 are referred to [73]. In early 2024, [10] proposed YOLOv9, a significant advancement in the YOLO series [4]. YOLOv9 incorporates several key innovations: Programmable Gradient Information (PGI) for more reliable gradient updates, the Generalized Efficient Layer Aggregation Network (GELAN) to address data loss, and the integration of DETR and attention mechanisms to enhance detection capabilities. These improvements have reportedly resulted in superior performance on the MS COCO dataset compared to previous YOLO versions and RT-DETR. For our study, we utilized the YOLOv9 Large (yolov9e) model as implemented by [76].

4.3. Performance Metrics

The performance of object detection models can be assessed using several different metrics. In this study, we evaluated the accuracy of the considered algorithms based on their Precision (P), Recall (R), and mean Average Precision (mAP). We specifically used the metrics quantification method described in [77]. To quantify the three mentioned metrics, it is necessary to define the notion of True Positive (TP) calculated based on the Intersection over Union (IoU) of the bounding boxes with a **50**% threshold. Here, a bounding box is a rectangle added to tightly enclose a defect on the image. Given a predicted bounding box that intersects with a ground truth bounding box, the IoU is defined as the areas of the intersection divided by the union between the two bounding boxes, i.e.,

---

[4] Subsequently, [74] introduced YOLOv10, which reportedly achieves similar performance to YOLOv9 with 46% reduced latency and 25% fewer parameters. Additionally, in September 2024, [75] released YOLOv11 for public use, which improved the YOLO-series architecture and training methods for more versatile computer vision applications.

$$IoU = \frac{B_{pred} \cap B_{truth}}{B_{pred} \cup B_{truth}} \tag{1}$$

Here, B_{pred} and B_{truth} denote the area of the predicted and the ground truth bounding boxes, respectively, calculated in pixel units. The (TP) is then defined as the number of IoU instances that are higher than **50**% with the correct defect classifications.

The Precision (P) can then be computed as the ratio of the (TP) to the total number of predicted bounding boxes $\left(N_{B_{pred}}\right)$, i.e.,

$$P = \frac{TP}{N_{B_{pred}}} \tag{2}$$

Here, P quantifies the ability of the model to correctly classify the detected defects.

On the other hand, the Recall (R) is defined as the ratio between the (TP) and the actual number of defects present in the images $\left(N_{B_{true}}\right)$, i.e.,

$$R = \frac{TP}{N_{B_{truth}}} \tag{3}$$

Thus, R signifies the ability of the model to detect the defects present in the given images. To encapsulate the precision and recall of the model, we can then define the area under the precision-recall curve as Average Precision (AP). The average value of AP across all the defect types is defined as mean Average Precision (mAP). Because we calculated mAP at the **50**% threshold of the IoU, we will be using mAP50 to denote the mAP of the models in this study. Additionally, we also compared the inference speed of the models to detect and classify the defects, and the convergence speed of the models during training progress.

4.4. Hyperparameter Tuning

Training an NN model involves a number of hyperparameters that need to be configured beforehand, and the values we choose for these hyperparameters influence the performance of the algorithm [78], [79]. Given datasets for a specific application, one might resort to the default values of hyperparameters implemented by the NN packages or manually configure them based on available literature, experiences, and/or experiments.

Alternatively, a hyperparameter tuning strategy leveraging optimization techniques can also be implemented. Since this tuning strategy is generally computationally expensive, sequential model-based optimization methods, such as Bayesian optimization, are often employed [80–82]. In this approach, the optimization algorithm identifies the optimal point of an approximated model and iteratively updates the surrogate model, enabling the approximate optimum to be found with fewer evaluations of the computationally expensive original function.

In this study, we tuned the object detection models employing the Bayesian optimization approach by viewing the training processes as a black-box function. Specifically, we implemented the gp_minimize function available in [83]. We used the negative value of the weighted average of the accuracy metrics described in Section 4.3 as the objective function for the optimization scheme, i.e.,

$$\min_{\mathcal{H}}(-[w_P \quad w_R \quad w_{mAP50}][P \quad R \quad mAP50]^\top). \tag{4}$$

Where, $\mathcal{H}$ denotes the hyperparameters space we considered and $[w_P \quad w_R \quad w_{mAP50}]$ are the weight factors of the accuracy metrics. Specifically, the hyperparameter space $\mathcal{H}$ comprises 29 hyperparameters, as listed in Table 3. For the tuning process, we set a limit of 100 function calls, where each function call involves a complete training cycle of an object detection model using Dataset I. In our analysis, we assigned weight factors of $[\mathbf{0.3} \quad \mathbf{0.4} \quad \mathbf{0.3}]$ to maintain a balance across accuracy metrics. Additionally, we conducted a sensitivity analysis for other weight combinations (refer to Section 5.2).

Table 3. List of hyperparameters considered in this study [5].

no.	argument	values	no.	argument	values
1	batch	$\{4,8,12,16,24,28,32\}$	15	hsv_h	$[0.0,1.0]$
2	lr0	$[10^{-10}, 10^{-1}]$	16	hsv_s	$[0.0,1.0]$
3	lrf	$[10^{-10}, 10^{-1}]$	17	hsv_v	$[0.0,1.0]$
4	momentum	$[0.0,1.0]$	18	degrees	$[-180,180]$
5	weight_decay	$[0.0,0.01]$	19	translate	$[0.0,1.0]$
6	warmup_epochs	$[1.0,10.0]$	20	scale	$[0.0,\infty)$
7	warmup_momentum	$[0.0,1.0]$	21	shear	$[-180,180]$
8	warmup_bias_lr	$[0.0,0.5]$	22	perspective	$[0.0,0.001]$
9	box	$[1.0,20.0]$	23	flipud	$[0.0,1.0]$
10	cls	$[0.0,1.0]$	24	fliplr	$[0.0,1.0]$
11	dfl	$[1.0,2.0]$	25	bgr	$[0.0,1.0]$
12	pose	$[10.0,15.0]$	26	mosaic	$[0.0,1.0]$
13	kobj	$[0.0,5.0]$	27	mixup	$[0.0,1.0]$
14	nbs	$\{4,8,12,16,24,28,32,36,$	28	copy_paste	$[0.0,1.0]$
		$40,44,48,52,56,60,64\}$	29	erasing	$[0.0,0.9]$

Additional preprocessing was applied to Dataset II to reduce training time during the tuning process, given its large size. The training and validation data was split into 20 random pools of 500 training and 200 validation images each. At the start of each function call during the tuning iteration, a pair of training and validation datasets was randomly selected from these pools. This approach ensures that each function call during tuning process utilized only 500 training and 200 validation images, while still leveraging the entire dataset of 9,651 training and 642 validation images over the complete tuning procedure, which involved 150 function calls.

5. Experimental Results

The experiments in this study were conducted on datasets described in Section 3. The tuning and training of the models were performed on the ASPIRE2A nodes provided by the National Supercomputer Center (NSCC) Singapore. Each training iteration was conducted utilizing one AMD EPYC Millan 7713 CPU and one NVIDIA A100-40G SXM GPU with 128 GB of DDR4 memory [84].

5.1. Performance Comparison

The performance results — including Precision (P), Recall (R), and mAP50 — for the two models, RT-DETR and YOLOv9, are presented in Tables 4 and 5 for Datasets I and II, respectively. These tables display results both before and after hyperparameter tuning, as described in Section 4.4. For the post-tuning results, we applied weight factors of $[\mathbf{0.3} \quad \mathbf{0.4} \quad \mathbf{0.3}]$ to maintain a balance across accuracy metrics, with slightly more emphasis on Recall (R) to minimize false negatives.

In addition to the RT-DETR and YOLOv9 results, we also include, for Dataset I, the performance of the YOLOv5 Extra-Large (YOLOv5x) model from [21] to examine the evolution of YOLO's accuracy since 2020. For Dataset II — which is an external dataset provided by [19] — we include the results of their AI model. These results are assumed to reflect post-tuning performance, although the specific tuning method used was not disclosed.

The results show improvement of newer object detection models, i.e., RT-DETR and YOLOv9, in most of the accuracy metrics, compared to the previously employed models, i.e., YOLOv5 and model used by [19]. The performance improvement of these newer models shows the benefits that can be gained from utilizing the updated state-of-the-art object detection models.

---

[5] Description of the hyperparameters can be found in Ultralytics (2023)..

Table 4. Accuracy comparison of the models to detect defects on Dataset I.

models	before tuning			after tuning		
	P	R	mAP50	P	R	mAP50
RT-DETR	0.731	0.706	0.699	0.777	0.701	0.710
YOLOv9	0.713	0.682	0.685	0.786	0.720	0.730
YOLOv5	0.682	0.651	0.626	0.723	0.726	0.698

Table 5. Accuracy comparison of the models to detect defects on Dataset II.

models	before tuning			after tuning		
	P	R	mAP50	P	R	mAP50
RT-DETR	0.749	0.693	0.716	0.794	0.664	0.716
YOLOv9	0.755	0.587	0.679	0.809	0.637	0.725
From [19]	−	−	−	0.795	0.652	0.696

Comparing the performance of RT-DETR and YOLOv9 in the current case study, the after-tuning results show a slight upper hand on the accuracy of YOLOv9 as seen in Tables 4 and 5, consistent with the results reported by [10]. Moreover, YOLOv9 also achieved a faster inference speed of approximately 66 frames per second, compared to the RT-DETR's inference speed of around 41 frames per second. However, comparing the training progress between the two models, we observed that RT-DETR has a faster convergence speed while maintaining a comparable accuracy as shown in Figure 4. In these plots, we show the average of the accuracy metrics from multiple training instances, at each epoch, as solid lines. The shaded areas in the plots display the one-standard deviation below and above the mean of the training instances. In this comparison, we considered the training progress of RT-DETR and YOLOv9 models for both Datasets I and II. Here, RT-DETR achieved comparable "P", "R", and "mAP50" before 100 epochs, while YOLOv9 needed more than 150 epochs.
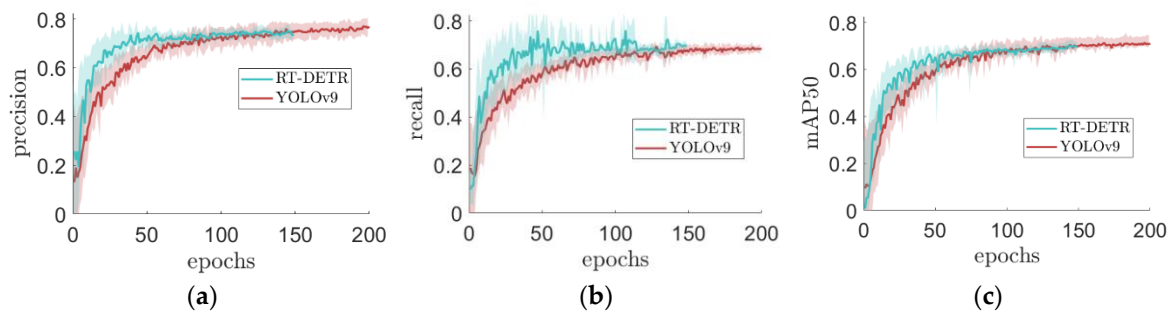
Figure 4. Training progress comparison of RT-DETR and YOLOv9 models based on the (a) precision, (b) recall, and (c) mAP50. The solid lines are the average of multiple training instances from both models using both Datasets I and II, while the shaded areas indicate the one-standard deviation below and above the mean of the training instances.

5.2. Accuracy Metrics Emphases During Tuning

As expected, from Tables 4 and 5, we can see that all algorithms show some performance improvement after tuning. Specifically, for YOLOv9 model, we can observe an average improvement around 8.6% in "P", 7% in "R", and 6.6% in "mAP50". However, it can be argued that the benefit of hyperparameter tuning is minimal for the RT-DETR model as it can be observed to have around 6% improvement in "P" but almost no improvement in "mAP50" and even an average of 3% decreased performance in "R". To further inspect this, we are interested in observing the effect of different emphases on the accuracy metrics during tuning procedure.

To compare the tuning effect of different emphases on the accuracy metrics, we varied the objective function (weights) of the optimization problem (4). Specifically, we considered three additional variations of weight factors [w_P w_R w_{mAP50}] that give different emphases to the

different accuracy metrics separately, i.e., $[\mathbf{1} \quad \mathbf{0} \quad \mathbf{0}]$, $[\mathbf{0} \quad \mathbf{1} \quad \mathbf{0}]$, and $[\mathbf{0} \quad \mathbf{0} \quad \mathbf{1}]$. We performed this tuning study on Dataset I. To ease the comparison, we plotted blue and red dotted lines to show the accuracy of RT-DETR and YOLOv9 models before tuning, respectively. Two different scales of the vertical axis are used to highlight the improvement in the accuracy of the models by zooming the upper part of the plots.

The comparison in Figure 5 demonstrates that a more balanced emphasis on accuracy metrics during tuning yields better overall performance in "P", "R", and "mAP50". Interestingly, although "mAP50" is intended to represent both precision and recall, focusing solely on this metric during the tuning process appears to lead to suboptimal overall accuracy. This is particularly evident in the case of the YOLOv9 model. For the RT-DETR model, although emphasizing "mAP50" during tuning results in higher recall improvements, the greatest overall accuracy improvement is achieved with a balanced weight.
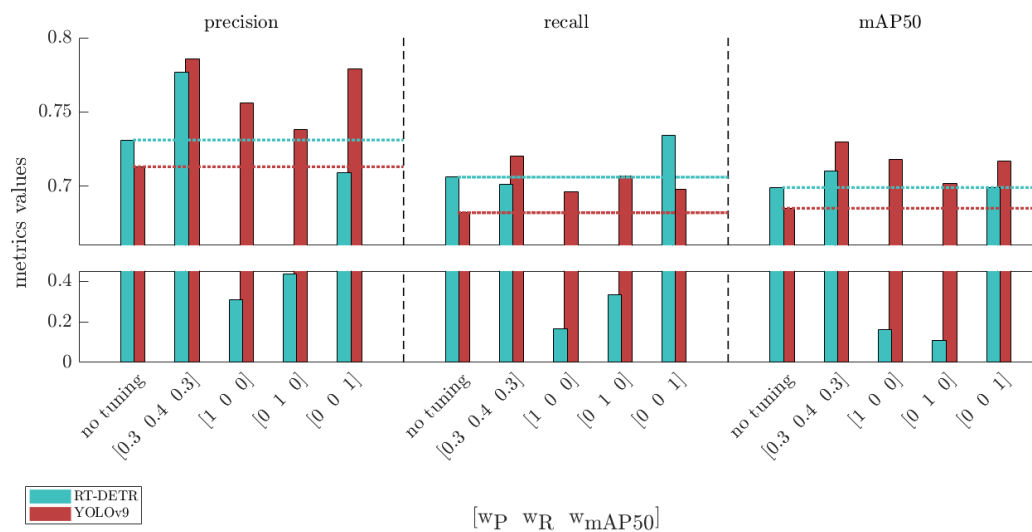
Figure 5. After-tuning comparison for Dataset I with different weight factors of the accuracy metrics. The blues and red dotted lines show the accuracy before tuning of RT-DETR and YOLOv9 models, respectively. Two different vertical axis scales are used to magnify the upper part of the plots, highlighting the accuracy improvements after tuning.

6. Discussion

The results of this research demonstrate that a relatively inexpensive UAV system equipped with a 14-megapixel camera can achieve moderately accurate results (mAP50: 0.70–0.75) in identifying aircraft defects. This suggests that existing methods have potential as supportive tools for aircraft defect detection. However, feedback from MRO experts highlights that these accuracy levels remain insufficient due to safety concerns. This underscores the need for further advancements in computer vision algorithms specifically tailored to aircraft defect detection to surpass 90% accuracy levels. One major challenge faced by computer vision researchers is access to comprehensive datasets, which constitutes a significant bottleneck in advancing current algorithms. In this paper, we aim to address this issue by providing access to the recorded and collected data [12]. We also include a summary of links to publicly available defect image datasets in the Section titled Available Datasets of Aircraft Defects.

In addition to algorithmic improvements, discussions with MRO experts revealed uncertainties regarding the optimal hardware solutions for defect data collection in real-world context. We identified four main categories of hardware solutions — UAVs, smart hangars, mobile robot platforms, and climbing robots — each with distinct advantages and limitations. A summary of the key insights from these discussions is provided below, serving as a starting point for future investigations into hardware selection and operational integration.

Unmanned Aerial Vehicles (UAVs):

Advantages:

- Capable of reaching difficult areas, such as the upper fuselage and tail sections, without the need for scaffolding or lifts;
- Able to conduct inspections quickly, making them particularly useful for pre/post-flight inspections;
- Swarm of drones can make routine inspections inside hangars even faster;
- As suggested by our review of literature, drones have been already widely tested by both academia and industry;
- Relatively inexpensive. As shown through the experiment in this paper even inexpensive drones and cameras can already provide good levels of accuracy;
- Easily transferable between inspections. Can be deployed across multiple aprons/stands.

Limitations:

- UAV operations are subject to strict aviation regulations, which may limit their deployment in certain areas of the airport. Additionally, regulatory differences across airports pose challenges for developing standardized solutions;
- Drones may only be permitted for use inside hangars, reducing their utility for pre/post-flight inspections. Solutions like tethers attached to drones could address this, but may limit their range and effectiveness;
- May obstruct other maintenance activities, and therefore human-drone coordination requires planning;
- Battery life and payload capacity can restrict flight duration and the types of sensors that can be equipped;
- Planning 3D trajectories for drones around aircraft surfaces can be challenging, particularly when coordinating a swarm of drones.

Smart Hangars with High-Resolution Cameras:

Advantages:

- Less constrained by regulations as the system operates within a controlled environment and does not interfere with other operations;
- Easily integrates with human-led maintenance activities without obstructing ongoing tasks;
- Once installed, these systems provide reliable, consistent coverage without needing repositioning, and maintenance costs remain relatively low after the initial setup.

Limitations:

- Can only be used inside hangars limiting its use at pre/post flight inspections;
- Fixed cameras may struggle to capture certain parts of the aircraft, especially at specific angles. Additionally, the quality of images may be reduced compared to drones or robots that capture close-up visuals;
- The system is not transferable, requiring airports to dedicate specific hangars equipped with this technology and plan aircraft inspection schedules accordingly;
- Installation and integration of high-resolution camera systems involve substantial upfront investment.

Mobile Robot Platforms:

Advantages:

- Relatively inexpensive and easy to implement, as they can be adapted from technologies already used in other industries (e.g., infrastructure and industrial inspections, as reviewed in [85]);
- Easily transferable between inspections, can be deployed across multiple aprons/stands;
- Can accommodate a more diverse range of equipment (different cameras, sensors, etc) due to their greater payload capacity (not as constrained as drones or climbing robots).

Limitations:

- Cannot access elevated areas like the top of the fuselage or tail sections without additional equipment, such as extendable cameras;
- May obstruct other maintenance activities, requiring careful planning and coordination between human workers and the robotic platform, which can be more challenging compared to drones;
- Inspections using ground-based robots may take longer than drones or fixed cameras to cover the entire aircraft.

Climbing Robots:

Advantages:

- Capable of traversing complex geometries, these robots can inspect detailed areas of the aircraft, including the fuselage, wings, and tail sections, without the need for scaffolding or lifts;
- Since they adhere to the aircraft's surface rather than flying freely like drones or moving on the ground like traditional robot platforms, climbing robots may face fewer regulatory constraints during operation;
- Their operation is less likely to obstruct other maintenance activities, especially during routine inspections;
- Easily transferable between inspection. Can be deployed across multiple aprons/stands.

Limitations:

- While effective for detailed examinations, these robots may require more time to inspect the entire aircraft surface, making them more suitable for scheduled maintenance rather than quick pre/post-flight checks;
- Adhesion mechanisms and technologies tailored for aircraft inspection can make climbing robots more expensive compared to other inspection tools;
- Technology is still under development. Although existing models, such as those developed by SR Technics, and SUTD [42,55], demonstrate progress. Further research and development are required to address the advanced design and technology requirements of these robots.

In addition to advancing visual detection algorithms and identifying the most suitable hardware solutions for aircraft defect detection, several other critical considerations must be addressed:

- **Regulatory Framework:** The regulatory aspect is paramount, particularly given that aviation is a global industry. Automation of defect detection operations requires international coordination and recognition to ensure consistency across countries. Standardizing regulations would enable the development of solutions that are globally applicable rather than tailored to specific regional contexts. While the International Civil Aviation Organization (ICAO) provides overarching safety management frameworks, such as Standards and Recommended Practices (SARPs) [86], these currently lack detailed guidance and regulations specific to automated aircraft defect detection.
- **Human-Robot Interaction:** The interaction between humans and robotic systems in defect detection remains undefined and requires research. The role of humans in the defect detection

process, once these solutions are implemented, needs clarification. Should their role be passive, focused on analyzing feedback from these systems, or active, complementing the robots/drones detection search? Additionally, these systems may obstruct other critical operations, particularly during aircraft turnaround processes. Some work has already begun to explore this topic [87].

- **Path Planning:** Robust path planning mechanisms are needed to address two key objectives: minimizing interference with other operations and ensuring complete aircraft coverage with high-quality images within the limited time available. Research is already underway to tackle this challenge [88–90], but further advancements are crucial.

- **Weather Resilience:** Weather conditions may significantly impact automated defect detection, especially during pre- and post-flight inspections. Ensuring that automated solutions remain resilient and functional under adverse weather conditions is essential to prevent disruptions. This resilience must be investigated and integrated into future systems to ensure reliability across a wide range of environmental scenarios.

7. Conclusion

This paper reviews the current state of computer vision algorithms for supporting aircraft surface defect detection during inspection routines, whether pre/post-flight during turnarounds or at routine intervals (e.g., every few months). Specifically, we analyze advances from both academic (21 publications identified) and industry perspectives (18 projects identified).

Through this review and expert discussions, we identified existing object detection algorithms tailored for aircraft defect detection and explored hardware solutions for image and video collection, including drones (UAVs), robot platforms, climbing robots, and smart hangars equipped with cameras.

This study goes beyond literature review by applying these concepts to real-world scenarios. We collected real data on aircraft defects using a drone, making this dataset publicly available for further research. Additionally, we evaluated and compared the performance of two state-of-the-art object detection models — RT-DETR and YOLOv9. Results show that both models achieved mAP50 scores between 0.70 and 0.75. YOLOv9 demonstrated a slight edge in accuracy and inference speed, while RT-DETR achieved similar accuracy with faster training convergence. A comparison with YOLOv5 from 2020 highlights the rapid evolution of computer vision algorithms, with YOLOv9 achieving a 10% improvement in mAP50. Despite these advancements, we acknowledge that an mAP50 of 0.70–0.75 is insufficient for meeting the risk requirements of aviation operations. Nonetheless, the continuous and rapid progress in computer vision is encouraging.

We also identify and discuss important challenges that still need to be adressed such as regulatory constraints, human-robot integration, path planning optimization for defect image collection, and weather resilience.

In conclusion, this paper provides a comprehensive overview of the current state of computer vision in aircraft defect detection, serving as a foundation for future research in a field that has gained significant attention in recent years.

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

References

1. FAA, "Advisory Circular: Visual Inspection for Aircraft." U.S. Department of Transportation, 1997.
2. IMRBPB, "Clarification of Glossary Definitions for General Visual (GVI), Detailed (DET), and Special Detailed (SDI) Inspections." 2011.
3. R. Saltoğlu, N. Humaira, and G. İnalhan, "Aircraft Scheduled Airframe Maintenance and Downtime Integrated Cost Model," *Advances in Operations Research*, vol. 2016, no. 1, p. 2576825, 2016.
4. Y. Liu, J. Dong, Y. Li, X. Gong, and J. Wang, "A UAV-Based Aircraft Surface Defect Inspection System via External Constraints and Deep Learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–15, 2022.
2. J. P. Sprong, X. Jiang, and H. Polinder, "Deployment of Prognostics to Optimize Aircraft Maintenance – A Literature Review," *Journal of International Business Research and Marketing*, vol. 5, no. 4, pp. 26–37, May 2020.
3. U. Papa and S. Ponte, "Preliminary Design of An Unmanned Aircraft System for Aircraft General Visual Inspection," *Electronics*, vol. 7, no. 12, 2018.
4. S. Le Clainche, E. Ferrer, S. Gibson, E. Cross, A. Parente, and R. Vinuesa, "Improving Aircraft Performance Using Machine Learning: A Review," *Aerospace Science and Technology*, vol. 138, p. 108354, 2023.
5. D. A. Rodríguez, C. Lozano Tafur, P. F. Melo Daza, J. A. Villalba Vidales, and J. C. Daza Rincón, "Inspection of Aircrafts and Airports Using UAS: A Review," *Results in Engineering*, vol. 22, p. 102330, 2024.
6. COCO, "Object Detection on COCO test-dev." https://paperswithcode.com/sota/object-detection-on-coco, 2023.
7. C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," *arXiv preprint arXiv:2402.13616*, 2024.
8. Y. Zhao et al., "DETRs Beat YOLOs on Real-time Object Detection," *arXiv preprint arXiv:2304.08069*, 2024.
9. SUTD, "Aircraft AI Dataset," *Roboflow Universe*. https://universe.roboflow.com/sutd-4mhea/aircraft-ai-dataset ; Roboflow, 2024.
10. J. Miranda, J. Veith, S. Larnier, A. Herbulot, and M. Devy, "Machine Learning Approaches for Defect Classification on Aircraft Fuselage Images Aquired by An UAV," in *International conference on quality control by artificial vision*, 2019, p. 1117208.
11. T. Malekzadeh, M. Abdollahzadeh, H. Nejati, and N.-M. Cheung, "Aircraft Fuselage Defect Detection Using Deep Neural Networks," *arXiv preprint arXiv:1712.09213*, 2020.
12. D. Meng, W. Boer, X. Juan, A. N. Kasule, and Z. Hongfu, "Visual Inspection of Aircraft Skin: Automated Pixel-Level Defect Detection by Instance Segmentation," *Chinese Journal of Aeronautics*, vol. 35, no. 10, pp. 254–264, 2022.
13. L. Connolly, J. Garland, D. O'Gorman, and E. F. Tobin, "Deep-Learning-Based Defect Detection for Light Aircraft With Unmanned Aircraft Systems," *IEEE Access*, vol. 12, pp. 83876–83886, 2024.

14. B. Huang, Y. Ding, G. Liu, G. Tian, and S. Wang, "ASD-YOLO: An Aircraft Surface Defects Detection Method Using Deformable Convolution and Attention Mechanism," *Measurement*, vol. 238, p. 115300, 2024.

15. A. Plastropoulos, K. Bardis, G. Yazigi, N. P. Avdelidis, and M. Droznika, "Aircraft Skin Machine Learning-Based Defect Detection and Size Estimation in Visual Inspections," *Technologies*, vol. 12, no. 9, p. 158, 2024.

16. I. Hangar, "Innovation Hangar v2 Dataset," *Roboflow Universe*. https://universe.roboflow.com/innovation-hangar/innovation-hangar-v2/dataset/1 ; Roboflow, 2023.

17. W. Zhang, J. Liu, Z. Yan, M. Zhao, X. Fu, and H. Zhu, "FC-YOLO: an aircraft skin defect detection algorithm based on multi-scale collaborative feature fusion ," *Measurement Science and Technology*, vol. 35, no. 11, p. 115405, 2024.

18. G. Jocher, "Ultralytics YOLOv5." https://github.com/ultralytics/yolov5, 2020.

19. C. Kurniawan, "Aircraft-Skin-Defect-Detection-YOLOv9-Vs.-RT-DETR." https://github.com/cparyoto/Aircraft-Skin-Defect-Detection-YOLOv9-Vs.-RT-DETR, 2024.

20. Z. Deyin, W. Penghui, T. Mingwei, C. Conghan, W. Li, and H. Wenxuan, "Investigation of Aircraft Surface Defects Detection Based on YOLO Neural Network," in *2020 7th international conference on information science and control engineering (ICISCE)*, IEEE, 2020, pp. 781–785.

21. Y. Li, Z. Han, H. Xu, L. Liu, X. Li, and K. Zhang, "YOLOv3-Lite: A Lightweight Crack Detection Network for Aircraft Structure Based on Depthwise Separable Convolutions," *Applied Sciences*, vol. 9, no. 18, p. 3781, 2019.

22. J. D. Maunder et al., "AI-based General Visual Inspection of Aircrafts Based on YOLOv5," in *2023 IEEE canadian conference on electrical and computer engineering (CCECE)*, 2023, pp. 55–59.

23. S. Pasupuleti, R. K, and V. M. Pattisapu, "Optimization of YOLOv8 for Defect Detection and Inspection in Aircraft Surface Maintenance using Enhanced Hyper Parameter Tuning," in *2024 international conference on electrical electronics and computing technologies (ICEECT)*, 2024, pp. 1–6.

24. H. Wang, L. Fu, and L. Wang, "Detection algorithm of aircraft skin defects based on improved YOLOv8n," *Signal, Image and Video Processing*, vol. 18, no. 4, pp. 3877–3891, 2024.

25. Q. Ren and D. Wang, "Aircraft Surface Defect Detection Based on Improved YOLOv8," in *2024 5th International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)*, IEEE, 2024, pp. 603–606.

26. J. Zhao et al., "Biological Visual Attention Convolutional Neural Network for Aircraft Skin Defect Detection," *Measurement: Sensors*, vol. 31, p. 100974, 2024.

27. L. Chen, L. Zou, C. Fan, and Y. Liu, "Feature weighting network for aircraft engine defect detection," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 18, no. 3, p. 2050012, 2020.

28. Y. Abdulrahman, M. M. Eltoum, A. Ayyad, B. Moyo, and Y. Zweiri, "Aero-engine blade defect detection: A systematic review of deep learning models," *IEEE Access*, vol. 11, pp. 53048–53061, 2023.

29. A. Upadhyay, J. Li, S. King, and S. Addepalli, "A deep-learning-based approach for aircraft engine defect detection," *Machines*, vol. 11, no. 2, p. 192, 2023.

30. Y. Qu, C. Wang, Y. Xiao, J. Yu, X. Chen, and Y. Kong, "Optimization algorithm for surface defect detection of aircraft engine components based on YOLOv5," *Applied Sciences*, vol. 13, no. 20, p. 11344, 2023.

31. A. Niccolai, D. Caputo, L. Chieco, F. Grimaccia, and M. Mussetta, "Machine Learning-Based Detection Technique for NDT in Industrial Manufacturing," *Mathematics*, vol. 9, no. 11, 2021.

32. I. Shafi et al., "Deep learning-based real time defect detection for optimization of aircraft manufacturing and control performance," *Drones*, vol. 7, no. 1, p. 31, 2023.

33. A. Rosell et al., "Machine Learning-Based System to Automate Visual Inspection in Aerospace Engine Manufacturing," in *2023 IEEE 28th international conference on emerging technologies and factory automation (ETFA)*, 2023, pp. 1–8.

34. N. Prakash, D. Nieberl, M. Mayer, and A. Schuster, "Learning Defects from Aircraft NDT Data," *NDT & E International*, vol. 138, p. 102885, 2023.

35. J. Zhuge, W. Zhang, X. Zhan, K. Wang, Y. Wang, and J. Wu, "Image matching method based on improved harris algorithm for aircraft residual ice detection," in *2017 4th international conference on information, cybernetics and computational social systems (ICCSS)*, IEEE, 2017, pp. 273–278.

36. R. Wen, Y. Yao, Z. Li, Q. Liu, Y. Wang, and Y. Chen, "LESM-YOLO: An improved aircraft ducts defect detection model," *Sensors (Basel, Switzerland)*, vol. 24, no. 13, p. 4331, 2024.

37. I. Jovančević, I. Viana, J.-J. Orteu, T. Sentenac, and S. Larnier, "Matching CAD model and image features for robot navigation and inspection of an aircraft," in *International conference on pattern recognition applications and methods*, SCITEPRESS, 2016, pp. 359–366.

38. J. R. Leiva, T. Villemot, G. Dangoumeau, M.-A. Bauda, and S. Larnier, "Automatic visual detection and verification of exterior aircraft elements," in *2017 IEEE international workshop of electronics, control, measurement, signals and their application to mechatronics (ECMSM)*, IEEE, 2017, pp. 1–5.

39. B. Ramalingam et al., "Visual Inspection of the Aircraft Surface Using a Teleoperated Reconfigurable Climbing Robot and Enhanced Deep Learning Technique," *International Journal of Aerospace Engineering*, vol. 2019, no. 1, p. 5137139, 2019.

40. B. Soufiane, D. Anıl, A. Ridwan, A. Reyhan, and S. Joselito, "Towards automated aircraft maintenance inspection. A use case of detecting aircraft dents using mask r-cnn," *A use case of detecting aircraft dents using Mask R-CNN*, 2020.

41. A. Doğru, S. Bouarfa, R. Arizar, and R. Aydoğan, "Using Convolutional Neural Networks to Automate Aircraft Maintenance Visual Inspection," *Aerospace*, vol. 7, no. 12, p. 171, 2020.

42. S. Merola, "Digital Optics and Machine Learning Algorithms for Aircraft Maintenance," *Materials research proceedings*, 2024.

43. N. Suvittawat and N. A. Ribeiro, "Aircraft Surface Defect Inspection System Using AI with UAVs," in *International conference on research in air transportation (ICRAT)*, 2024, pp. 1–4.

44. H. Li, C. Wang, and Y. Liu, "Aircraft Skin Defect Detection Based on Fourier GAN for Data Augmentation ," in *2024 international conference on advanced robotics and mechatronics (ICARM)*, IEEE, 2024, pp. 449–454.

45. X. Zhang, J. Zhang, J. Chen, R. Guo, and J. Wu, "A Semisupervised Aircraft Fuselage Defect Detection Network With Dynamic Attention and Class-Aware Adaptive Pseudolabel Assignment," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 7, 2024.

46. Air-Cobot Project, "Air-Cobot: Collaborative Mobile Robot for Aircraft Inspection." Collaborative Project by Akka Technologies, Airbus Group,; Partners; https://en.wikipedia.org/wiki/Air-Cobot, 2013.

47. P. Davies, "EasyJet Reveals Drone Inspection and 3D Printing Plans." https://travelweekly.co.uk/articles/54445/easyjet-reveals-drone-inspection-and-3d-printing-plans; Travel Weekly, 2015.

48. Donecle, "Automating Your Aircraft Inspections." https://www.donecle.com/, 2015.

49. MainBlades, "Aircraft Inspection Automation." https://www.mainblades.com/, 2017.

50. Airbus, "Airbus Launches Advanced Indoor Inspection Drone to Reduce Aircraft Inspection Times and Enhance Report Quality." https://www.airbus.com/en/newsroom/press-releases/2018-04-airbus-launches-advanced-indoor-inspection-drone-to-reduce-aircraft, 2018.

51. Rolls-Royce, "Rolls-Royce Demonstrates The Future of Engine Maintenance with Robots that can Crawl Inside Engines." https://www.rolls-royce.com/media/press-releases/2018/17-07-2018-rr-demonstrates-the-future-%20of-engine-maintenance-with-robots.aspx, 2018.

52. SR Technics, "Robots Driving Innovation at SR Technics." https://www.srtechnics.com/news/press-releases-blog-social-media/2018/february-2018/robots-driving-innovation-at-sr-technics/, 2018.

53. Ubisense, "Ubisense and MRO drone launch world's first 'smart hangar' solution." https://ubisense.com/ubisense-and-mro-drone-launch-worlds-first-smart-hangar-solution/, 2018.

54. Airfrance, "In Motion AWACS: Drone Inspection." https://www.afiklmem.com/en/robotics/drone-inspections, 2019.

55. ST Engineering, "ST Engineering Receives First-ever Authorisation from CAAS to Perform Aircraft Inspection Using Drones." https://www.stengg.com/en/newsroom/news-releases/st-engineering-receives-first-ever-authorisation-from-caas-to-perform-aircraft-inspection/, 2020.

56. E. Kulisch, "Korean air develops drone swarm technology to inspect aircraft." https://www.flyingmag.com/korean-air-develops-drone-swarm-technology-to-inspect-aircraft/, 2022.

57. G. Aerospace, "GE aerospace service operators: Meet your 'mini' robot inspector companions." https://www.geaerospace.com/news/press-releases/services/ge-aerospace-service-operators-meet-your-mini-robot-inspector-companions, 2023.

58. A*STAR I^2R, "Smart Automated Aircraft Visual Inspection System (SAAVIS)." https://www.a-star.edu.sg/i2r/research/I2RTechs/research/i2r-techs-solutions/SAAVIS, 2023.

59. E. Loi, "SIA Engineering Lifts Productivity as Latest Robots Help Inspect Aircraft Engines." https://www.straitstimes.com/singapore/sia-engineering-boosts-productivity-with-robots-inspecting-aircraft-engines-and-seat-tracks; The Straits times, May 2023.

60. Jet Aviation, "AI Drone Inspections." https://www.jetaviation.com/ai-drone-inspections/, 2023.

61. A. Koplin, "Industry First: FAA Accepts Delta's Plan to Use Drones for Maintenance Inspections." https://news.delta.com/industry-first-faa-accepts-deltas-plan-use-drones-maintenance-inspections, 2024.

62. C. Biesecker, "Boeing Expanding Effort to Autonomously Inspect Aircraft." https://www.aviationtoday.com/2024/07/11/boeing-expanding-effort-to-autonomously-inspect-aircraft-july-28/; Aviation Today, 2024.

63. Lufthansa Technik, "Aircraft Overhaul: Testing Future technologies." https://www.lufthansa-technik.com/en/innovation-bay, 2024.

64. B. Dwyer, J. Nelson, T. Hansen, et al., "Roboflow (Version 1.0)." https://roboflow.com, 2024.

65. T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," *arXiv preprint arXiv:1405.0312*, 2015.

66. N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-End Object Detection with Transformers," *arXiv preprint arXiv:2005.12872*, 2020.

67. A. Khan et al., "A Survey of The Vision Transformers and Their CNN-Transformer Based Variants," *Artificial Intelligence Review*, vol. 56, no. 3, pp. 2917–2970, 2023.

68. Ultralytics, "Baidu's RT-DETR: A Vision Transformer-Based Real-Time Object Detector." https://docs.ultralytics.com/models/rtdetr/, 2024.

69. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *arXiv preprint arXiv:1506.02640*, 2015.

70. A. Vijayakumar and S. Vairavasundaram, "YOLO-based Object Detection Models: A Review and its Applications," *Multimedia Tools and Applications*, 2024.

71. A. Wang et al., "YOLOv10: Real-Time End-to-End Object Detection," *arXiv preprint arXiv:2405.14458*, 2024.

72. G. Jocher and J. Qiu, "Ultralytics YOLO11." https://github.com/ultralytics/ultralytics/yolov11, 2024.

73. Ultralytics, "YOLOv9: A Leap Forward in Object Detection Technology." https://docs.ultralytics.com/models/yolov9/, 2024.

74. Ultralytics, "Performance Metrics Deep Dive." https://docs.ultralytics.com/guides/yolo-performance-metrics/, 2024.

75. P. Probst, A.-L. Boulesteix, and B. Bischl, "Tunability: Importance of Hyperparameters of Machine Learning Algorithms," *Journal of Machine Learning Research*, vol. 20, no. 53, pp. 1–32, 2019.

76. H. J. P. Weerts, A. Mueller, and J. Vanschoren, "Importance of Tuning Hyperparameters of Machine Learning Algorithms," *arXiv preprint arXiv:2402.13616*, 2020.

77. F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Sequential Model-Based Optimization for General Algorithm Configuration," in *Learning and Intelligent Optimization*, C. A. Coello, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 507–523.

78. J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian Optimization of Machine Learning Algorithms," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. Burges, L. Bottou, and K. Q. Weinberger, Eds., Curran Associates, Inc., 2012, pp. 1–9.

79. B. Bischl, J. Richter, J. Bossek, D. Horn, J. Thomas, and M. Lang, "mlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions," *arXiv preprint arXiv:1703.03373*, 2018.

80. T. Head, M. Kumar, H. Nahrstaedt, G. Louppe, and I. Shcherbatyi, "scikit-optimize/scikit-optimize." https://github.com/scikit-optimize/scikit-optimize, 2020.

81. National Supercomputer Center (NSCC) Singapore, "ASPIRE2A General QuickStart Guide." https://help.nscc.sg/wp-content/uploads/ASPIRE2A-General-Quickstart-Guide.pdf, 2024.

82. D. Lattanzi and G. Miller, "Review of robotic infrastructure inspection systems," *Journal of Infrastructure Systems*, vol. 23, no. 3, p. 04017004, 2017.

83. ICAO, "SARPs - standards and recommended practices." https://www.icao.int/safety/safetymanagement/pages/sarps.aspx, 2024.

84. F. Donadio, J. Frejaville, S. Larnier, and S. Vetault, "Human-robot collaboration to perform aircraft inspection in working environment," in *Proceedings of 5th international conference on machine control and guidance (MCG)*, 2016.

85. A. Bircher et al., "Structural inspection path planning via iterative viewpoint resampling with application to aerial robotics," in *2015 IEEE international conference on robotics and automation (ICRA)*, IEEE, 2015, pp. 6423–6430.

86. Y. Liu, J. Dong, Y. Li, X. Gong, and J. Wang, "A UAV-based aircraft surface defect inspection system via external constraints and deep learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–15, 2022.

87. Y. Sun and O. Ma, "Drone-based automated exterior inspection of an aircraft using reinforcement learning technique," in *AIAA SCITECH 2023 forum*, 2023, p. 0107.