STRIVEWORKS

AUTOMATED DEFECT DETECTION

Case Study: Aircraft General Visual Inspection Using Chariot

3520 Executive Ctr Dr, Ste 300
Austin, TX 78731-1634

STRIVEWORKS.COM

Executive Summary

Routine aircraft inspections are a large undertaking for a major Fortune 500 logistics company. The inspection process has traditionally been handled manually, with aircraft maintenance technicians (AMTs) using cherry pickers to investigate the aircraft and document defects with images. The company worked with Striveworks to automate this inspection process by surveying aircraft with drones and creating an MLOps model to detect defects from 4k video.

This inspection method shift contributes to cost savings by reducing fleet downtime and labor costs.

Automating elements of the aircraft inspection process reduces delays following severe weather events, reduces employee risk exposure during visual inspections, and improves inspection consistency, accuracy, and auditability.

Striveworks developed a framework to perform these inspections via MLOps.

- Processed **49x more data** per aircraft vs manual inspection
- Reduced data processing time by **50%**
- Accurately identified **25% more dents**

A defect detection model outperforms human inspectors in speed, accuracy, and scalability

Background

General visual inspection (GVI) is an inspection procedure for finding defects on an aircraft. Traditionally, these inspections are performed by mechanics who physically inspect each aircraft for approximately four hours.

In the event of catastrophic weather events, all aircraft are grounded and must be inspected. This can cause significant delays for customers and potentially hazardous conditions for the employees who must inspect the aircraft. An automated system can both decrease aircraft downtime and simplify logistics by parallelizing inspections and minimizing the need for manual effort.

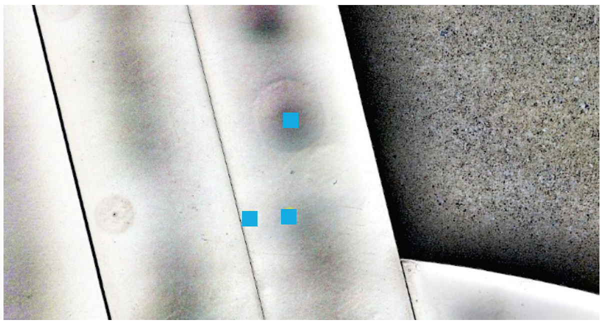
The major Fortune 500 logistics company worked with Striveworks to build a framework to identify the full contour of defects on aircraft from drone-captured video footage of the exterior of the aircraft.

Striveworks approached this endeavor via image segmentation, a well-established machine learning methodology. Segmentations provide the exact regions of an image that contain the defect. In this case, the segmentation focused on a particular type of defect, dents. The general idea is to create a pixel-wise classification (in this case ‘no damage,’ ‘dent,’ etc.) for every pixel in an image.

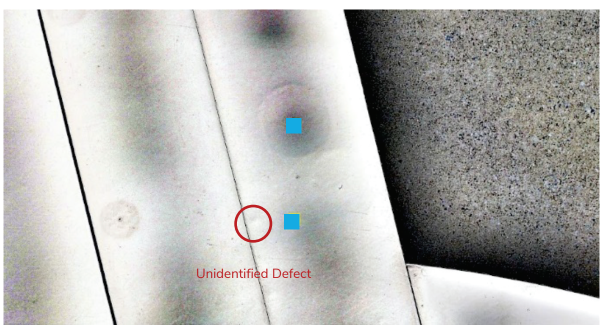
The company sought a solution it could deploy at the edge, so it did not need to coordinate physical or compute resources to a central location for review. This is something Chariot is well positioned to handle.

Technical Approach

Figure 1. Model Identifies Defect That Human Annotator Missed



Top: Predicted by Model



Bottom: Human annotated ground truth (missing defect)

The company used a semantic image segmentation model built in Chariot to operate on 4k video from the company's drones. The company's subject matter expert annotated a selection of video frames to identify dents for training data.

The model used lightweight MobileNetV3 architecture trained on the Chariot platform using the Chariot Training Runs service. Inference for this model was conducted by breaking the 4k video frames into 512 pixel x 512 pixel chips and stitching the results together. Training data was annotated in Chariot and new versions of the dataset were automatically tracked.

Video data allowed the company to exploit the temporal information and provided a wide range of perspectives. To leverage this video data, the company used Chariot to apply a tracking algorithm to the video. After an initial set of detections from the segmentation model were generated, the tracker aggregated individual image detections across frames to create robust dent detections. Robust dent detections were present in multiple frames, and the tracker exploited this to suppress many false positive detections without sacrificing recall.

Figure 1 shows an example of a single-frame segmentation mask for one segmentation model and the human annotator. Occasionally, the model found dents that the human annotator missed.

The analytics were packaged in Docker containers to create an application that featured a preprocessing stage, a detection stage, and a tracking stage before outputting videos with final frame-by-frame predictions.

Results

Evaluation Methodology

The team evaluated the performance of the MLOps solution by analyzing hold-out inspection videos; a subject matter expert validated each model detection as true positive or false positive. The subject matter expert also identified any dents the model had missed as false negatives.

Results

Overall, the model achieved a recall of 75%. Results vary on a video-by-video basis—a standard deviation of 20%. This is expected given the enormous amount of variability in lighting conditions, weather, time of day, and drone flight path. Precision is significantly improved with the deployment of the tracking algorithm over the entire video.

Table 1. Dent Recall on Evaluation Videos

Video	True Positive	False Positive	Recall
Video 1	65	27	0.71
Video 2	6	7	0.47
Video 3	26	23	0.53
Video 4	1	2	0.33
Video 5	5	4	0.56
Video 6	47	33	0.59
Video 7	26	73	0.26
Video 8	51	9	0.85
Video 9	49	5	0.91
Video 10	34	21	0.62
Total Recall			**0.75**

Impacts

Deployment of these computer vision models allowed 49 times more data to be processed relative to than human AMTs. This reduced inspection time by 50% and substantially improved the number of damage areas detected by 25%, which provides both cost savings and safety benefits.

Closing

This case study shows how using the Chariot platform can improve outcomes in aircraft inspections.

Performing physical aircraft inspections is a costly and time-consuming process. AMTs who perform physical inspections are highly trained and scarce; having them perform baseline inspections results in significant resource drain and opportunity cost when they cannot focus on “high confidence” areas of potential damage. Allocating these resources efficiently drives significantly higher fleet uptimes and improves safety for the general public.

Introducing computer vision models into the inspection process can create efficiencies in GVI without adding human resources. Some of the key challenges include:

- Fully representative datasets: Variations in lighting, weather, aircraft type, and drone flight path can materially affect model performance.
- False positive management: False positives take human time and effort to correct.
- False negative management: Due to the highly regulated nature of the business, regulatory agencies require an audited rate of false negatives in order to approve the analytic process.
- Incorporating domain-specific knowledge efficiently: As one example, most generic computer vision models struggle to address GVI because the rivets on the airframe can look like dents.
- Model management at scale: Once the custom models are built, maintaining and monitoring a diverse custom model array while maintaining auditability and regulatory compliance can be challenging.

With Striveworks, the customer can:

- Rapidly ingest, clean, and annotate their production data, using the Chariot annotation service.
- Quickly perform model experimentation, training, and validation using the Chariot training service.
- Reuse and modify well-performing models for additional operating environments and use cases.
- Build data processing workflows that incorporate deep learning-based computer vision models, heuristics built from domain expertise, and other algorithms to track potential dents.
- Build a culture of accessibility around data science using Chariot’s ‘process as code’ philosophy.
- Maintain audit and regulatory compliance around models.
- Improve cost and safety outcomes.

Striveworks Overview

Striveworks is a pioneer in operational data science for national security and other highly regulated spaces. Striveworks' flagship MLOps platform is Chariot, purpose-built to enable engineers and business professionals to transform their data into actionable insights. Chariot is an open-architecture data science platform, designed to enable the development, deployment, and use of AI/ML applications.

Founded in 2018, Striveworks was highlighted as an exemplar in the National Security Commission for AI 2020 Final Report.

Deploy Model in Hours, Not Months

Users can create and deploy solutions within a single business day

Retrain Model to Save Time & Money

Retrain existing model to answer new questions and handle new data

Achieve Fast ROI

Chariot is more cost-effective than ad-hoc model creation at just 5 models

STRIVEWORKS