Class-Incremental Object Detection for aerial surveillance systems

Aleksandr Setkov[1], Romaric Audigier[1], Hélion du Mas des Bourboux[2]

1 – *Université Paris-Saclay, CEA, List, 91120, Palaiseau, France*
2 – *Thales LAS France, 2, Av. Gay Lussac, 78990, Elancourt, France*

Identification of Use Cases

Automatic target detection and recognition (ATDR) solutions are being increasingly deployed on military aircraft. TALIOS AI, integrated on Rafale and Mirage platforms, has been undergoing DGA qualification flight testing and is expected to enter operational qualification in the coming years. This AI capability detects ground objects in TALIOS near-infrared imagery and localizes them on a map for the crew, with the objective of improving situational awareness and supporting faster mission adaptation.

At the same time, operational conditions in conflict zones are evolving rapidly. Recent theaters have highlighted the emergence of new types of equipment, including land, air, and maritime drones, as well as new ways of employing existing systems. More broadly, the multiplication of contested environments increases the need for ATDR systems that can be updated quickly to preserve or improve performance under changing operational conditions.

This creates a key operational requirement: the ability to integrate new target classes into an existing detection system without retraining the entire model from scratch and without degrading performance on previously learned classes. Such a capability is especially important for embedded airborne systems, where updates must remain time-efficient and compatible with certification processes, and limited access to large quantities of annotated data.

Although recent deep learning approaches, particularly foundation and generative models, have shown strong zero-shot and few-shot capabilities, they are generally not yet well suited to real-time embedded deployment on airborne platforms. Their performance can also be less convincing on low-resolution air-to-ground imagery, where targets may occupy only a few tens of pixels. In this context, class-incremental object detection appears to be a particularly relevant approach for aerial surveillance systems, as it addresses the need to learn novel classes while preserving previously acquired knowledge within operational constraints.

Proposed Technical Approach

Automatic Target Detection and Recognition (ATDR) systems require object detectors that can be quickly updated to recognize new target classes without compromising performance on previously learned targets, even when the original training data are unavailable. ***Class-Incremental Object Detection (CIOD)*** addresses this requirement by incrementally learning novel classes while preserving knowledge of previously learned classes. In this work, we propose **a novel class-incremental experimental protocol to investigate CIOD for aerial imagery**, including both remote sensing and UAV

viewpoints, representative of the operational scenarios encountered by Automatic Target Detection and Recognition (ATDR) systems.

Our experimental protocol aims to tackle **aerial-specific challenges** by using two **public datasets** including small objects and viewpoint variability: *DOTA* and *VisDrone*, made of satellite and drone images respectively, contain up to 15 classes including mobile objects of very small size (e.g., people and car) and static (e.g., infrastructure) objects. Images present diverse, dense, occluded, and cluttered scenes captured under day/night conditions.

We build our detection model from recent SOTA after evaluating which models (one-stage, two-stage vs transformers) are effective, not only on common life benchmarks like COCO, but also in aerial views and small objects. As the chosen aerial datasets were not evaluated in incremental learning, we follow the knowledge distillation paradigm which gives best results on academic benchmarks.

For each dataset, the **CIOD experiments** start with training a base model on a '*base*' subset of labeled images containing only base classes. Then, this base model is updated via incremental learning on a disjoint '*increment*' subset of images featuring one or more new (labeled) classes and possibly (unlabeled) base classes. Through metrics like the Average Precision and Average Recall, we assess the effectiveness of the *incremental update* by comparing it with the ideal *full supervision on* the *entire* labeled dataset.

As acquiring the labeled 'increment' subset is often drastically limited in real applications (due to cost and time constraints), we also study how the incremental model's performance varies with smaller numbers of new class training images. In addition, the experimental protocol includes varied choices for base or novel classes to investigate ***which factors impact the effectiveness of incremental learning***.

In particular, we address the following *research questions* about effective CIOD: What is the influence of (i) the **number of new class training instances**, (ii) the **visual appearance similarity** between classes (e.g., bus and truck), (iii) the **co-localization of object instances from different classes**, (iv) the **apparent object size**, (v) the **initial performance of the detector** on the base classes and the **quantity of base class objects** in the incremental training step, enabling the base knowledge distillation? Finally, beside the classification of each detected object, how do all these training factors impact the **quality of new class object localization**?

Based on the experimental results, several *conclusions* can be drawn, notably: The CIOD model can effectively address the catastrophic forgetting problem, maintaining strong performance on base classes (e.g., only a 2%-6% decrease in AP@50 for base classes on the tested datasets) while effectively learning new classes (e.g., the performance of new class can sometimes reach that of the non-incremental setting) when conditions are met, like providing sufficient training instances and other discussed factors. This study presents CIOD models and assesses factors impacting incremental learning effectiveness from experiments on publicly available civilian data, simulating conditions encountered by defense systems, and corresponding to a low TRL of 3-4. *Future work* includes: using more datasets, transposing to military scenarios and corresponding sets of fine-grained classes, extending the number of incremental steps and the number of new classes selected for each incremental update, optimizing algorithm and hardware for real-time embedded systems, addressing changing environments, in parallel of class-incremental paradigm, through domain-incremental learning and domain adaptation approaches.

All details about the CIOD methods, experimental results, and conclusions related to the formulated research questions will be presented at the conference upon acceptance of this extended abstract.