Improving Small Drone Detection Through Multi-Scale Processing and Data Augmentation

Rayson Laroca*,†, Marcelo dos Santos†, and David Menotti†

*Pontifical Catholic University of Paraná, Curitiba, Brazil
†Federal University of Paraná, Curitiba, Brazil

*rayson@ppgia.pucpr.br †{msantos, menotti}@inf.ufpr.br

***Abstract*–Detecting small drones, often indistinguishable from birds, is crucial for modern surveillance. This work introduces a drone detection methodology built upon the medium-sized YOLOv11 object detection model. To enhance its performance on small targets, we implemented a multi-scale approach in which the input image is processed both as a whole and in segmented parts, with subsequent prediction aggregation. We also utilized a copy-paste data augmentation technique to enrich the training dataset with diverse drone and bird examples. Finally, we implemented a post-processing technique that leverages frame-to-frame consistency to mitigate missed detections. The proposed approach attained a top-3 ranking in the 8th WOSDETC Drone-vs-Bird Detection Grand Challenge, held at the 2025 International Joint Conference on Neural Networks (IJCNN), showcasing its capability to detect drones in complex environments effectively.**

Fig. 1. The similarity in size and appearance often makes it challenging to differentiate a drone (blue box) from birds (red box), especially at a distance.

I. Introduction

Unmanned Aerial Vehicles (UAVs), commonly known as drones, have experienced a surge in popularity across diverse civil sectors [1], [2]. Their autonomy, flexibility, and affordability have driven their widespread adoption in applications such as search and rescue, package delivery, and remote sensing [3]. However, this rapid expansion presents significant security and privacy challenges [4], [5].

The versatility of drones also makes them susceptible to malicious uses, such as smuggling contraband, conducting intrusive surveillance, and executing physical attacks [6], [7]. Additionally, unauthorized drone operations can violate aviation safety regulations, posing direct threats to civilian aircraft and passengers while causing disruptions at airports, including flight delays [8], [9]. As a result, there is an escalating and urgent demand for advanced drone detection systems to address these growing security and privacy concerns [10], [11].

A major challenge in drone detection is accurately distinguishing drones from birds, given their similarities in size and appearance (see Fig. 1). Moreover, practical systems must be capable of identifying drones at long distances to allow for timely responses. This requires detecting very small objects and differentiating them from structured backgrounds and other complex visual elements. To spur innovation in this area, the International Workshop on Small-Drone Surveillance, Detection and Counteraction Techniques (WOSDETC) consortium has launched the 8th WOSDETC Drone-vs-Bird Detection Grand Challenge at the 2025 International Joint Conference on Neural Networks (IJCNN) [12]. The challenge seeks to attract research focused on developing novel signal-processing solutions for the problem of distinguishing birds from drones at long distances. To support participants, the consortium provides a video dataset, which is inherently difficult to acquire given the specific conditions and permissions required for drone flights. The dataset is progressively expanded with each challenge installment [13], [14] and subsequently released to the research community.

Considering the preceding discussion, this work presents a drone detection methodology that leverages the state-of-the-art YOLOv11 object detection model [15]. To enhance the detection of small-scale drones, we implemented a multi-scale approach that processes the input image both as a whole and in segmented components (simulating a zoom effect). Additionally, we employed extensive data augmentation, including a copy-paste technique, to increase the representation of drones and birds in the training images. A post-processing stage, utilizing adjacent frame predictions, was also employed to reduce missed detections. Our approach achieved a top-3 ranking (exact position not disclosed at the time of writing) in the 8th WOSDETC Drone-vs-Bird Detection Grand Challenge [12].

The rest of this paper is organized as follows: Section II briefly reviews related work. Section III describes the experiments conducted, including the datasets explored, the proposed method, and the experimental results. Finally, Section IV concludes the paper and suggests directions for future research.