MVA2023 - Small Object Detection Challenge for Spotting Birds

Fayaz Moqueem Mohammed Gowtham Senthilnayaki
Boston University
fayaz@bu.edu gs@bu.edu

February 24, 2023

Abstract

The goal of this challenge is to detect small birds in images using minimal pixel features, which falls under the category of Small Object Detection (SOD) problems. Unmanned aerial vehicles (UAVs) such as drones must be able to detect distant birds in order to prevent bird attacks and deter harmful birds from damaging crops and rice fields. As a result, this challenge not only presents academic challenges in Computer Vision, but it also encourages the development of practical technologies that can be used in real-world scenarios. The dataset contains images captured by drones and introduces new challenges such as varying bird types, parallax, postures, and different degrees of motion blur.

1 Problem Definition

The MVA2023 Challenge on Small Bird Detection problem definition is to detect small birds in images using minimal pixel features. This issue is classified as a Small Object Detection (SOD) problem. This challenge's primary goal is to develop practical technologies that can be used in real-world scenarios where drones must detect distant birds to prevent bird attacks and deter harmful birds from damaging crops and rice fields.

As a result, the challenge not only offers computer vision academic problems, but also promotes the creation of useful technologies that can be used in real-world situations to advance the drone and agricultural industries.

2 Datasets

The challenge dataset has been divided into training, public test, and private test sets. We can access the images and annotations of the training data, as well as the images of the public test data, as participants.

The MVA2023 - Small Object Detection Challenge for Spotting Birds dataset is divided into four subsets: Train1, Train2, Public Test, and Private Test.

Train1, a modified version of a previously published dataset[4] contains 47,260 images with 60,971 annotated bird instances. The newly released Train2 subset for this competition contains 9,759 images with 29,037 annotated bird instances. The Public Test subset has 9,699 images, while the Private Test subset has around 10,000.

3 Algorithm and Goal

We will be making use of the MMDetection[2] algorithm to fit the needs of this project. A cutting-edge open-source object identification framework called MMDetection was created by OpenMMLab

.

and can be used to solve the problem of finding tiny birds in aerial photographs. A variety of pre-trained models and backbones are provided by this framework, which may be utilized as a jumping off point to create unique detection pipelines for the problem.

In order to increase the precision of the detection models[1], the MMDetection framework also incorporates a variety of data augmentation techniques, such as flipping, scaling, and rotating. Moreover, MMDetection offers connection with the widely known COCO dataset and API for object detection tasks.

By experimenting with different MMDetection framework modifications, such as adjusting hyperparameters, experimenting with various detectors and backbones, modifying data augmentation techniques, and fine-tuning pre-trained models, our team is committed to achieving better accuracy than the baseline model. By experimenting with these changes, we want to enhance the functionality of our model. Our goal is to place as high on the leaderboard as we can and maybe even win the challenge.

4 Evaluation Criteria

The MVA2023 Challenge on Small Bird Detection competition's evaluation metrics are based on the COCO mAP, specifically the AP0.5. The COCO API[3] is used to evaluate the detection results. While other metrics such as AP0.5, AP0.75, and AP small are reported, we believe that using the official COCO mAP is more beneficial for method development.

5 Previous Work

Small object detection has received a lot of attention in the computer vision community, and several methods have been proposed to solve it.

Lin et al.[6] proposed feature pyramid networks (FPN) as one of the widely used methods for small object detection. To generate feature maps at multiple scales, this approach employs a top-down architecture with lateral connections. After that, the feature maps are fused to generate a final set of object proposals.

The anchor-free method, which eliminates the need for pre-defined anchor boxes, is another option. CornerNet[5] is one such method that predicts object locations as the corners of bounding boxes using a single stage detector. On several object detection benchmarks, this method produced cutting-edge results.

Fujii et al.[4] proposed a deep learning-based method for small bird detection that uses a Siamese network for feature extraction and a region proposal network (RPN) for object detection. The proposed method was tested on a dataset of distant birds captured by drones, and the results were promising.

References

[1] Hank Chen. Awesome tiny object detection, n.d.

[2] Kai Chen, Jiangmiao Pang, Jiaqi Wang, Yu Xiong, Xiaoxiao Li, and Shuyang Sun. MMDetection: Open mmlab detection toolbox and benchmark. *arXiv preprint arXiv:1906.07155*, 2019.

[3] Piotr Dollar and Tsung-Yi Lin. COCO API, 2017.

[4] Sanae Fujii, Kazutoshi Akita, and Norimichi Ukita. Distant bird detection for safe drone flight and its dataset. In *2021 17th International Conference on Machine Vision and Applications (MVA)*. IEEE, 2021.

[5] He Law and Jia Deng. Cornernet: Detecting objects as paired keypoints. In *Proceedings of the European Conference on Computer Vision*, pages 734–750. Springer, 2018.

[6] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.