UAVVaste: COCO-like dataset and effective waste detection in aerial images

Mateusz Piechocki, Bartosz Ptak, Marek Kraft
{ mateusz.piechocki, bartosz.ptak } @student.put.poznan.pl, marek.kraft@put.poznan.pl

Abstract

Detecting objects in aerial images is a much more difficult task than ordinary object detection which is more and more popular since deep learning methods was used. Nowadays applications which use aerial photos are very salable according to wide access to unmanned drones and flying vehicles. This difficulty due to significantly smaller objects' size, limitations in computing power related to using edge devices, energy consumption because of restricted power supply, Furthermore, conditions of application use also require efficiency and real time inference. Additionally, with increasing popularity of object detection tasks the world is struggling with the problem of growing garbage's amount. In many cities streets, parks and green places are strewn with litter and public services can't keep up with cleaning. Due to this issue we decided to introduce UAVVaste dataset which (presently) contain 772 images with 3716 hand-labeled annotations of rubbish in urban environment e.g. streets, parks and lawns. It is inspired by Trash Annotations in Context (TACO[3]) that however provides much larger objects from a near-earth perspective doesn't contain many examples similar to photos taken from the unmanned aerial vehicle. We aim to make the dataset public available in COCO format[2], including boxes annotations and segmentation masks. Moreover, using this data we develop waste object detectors based on the fast and effective YOLOv4[1] and EfficientDet[4] algorithms that allow inference on EDGE devices.

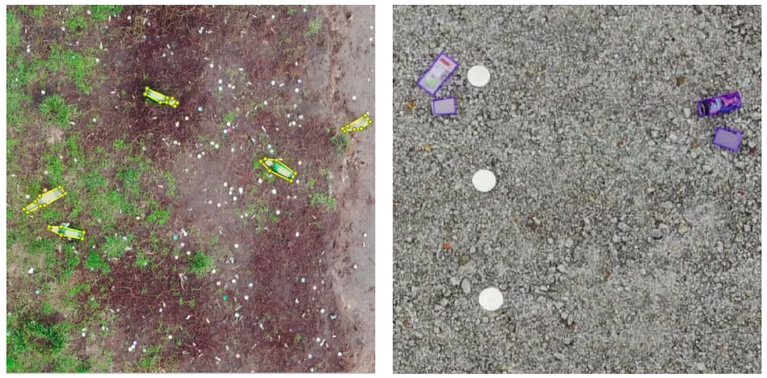
Dataset comparison

UAVVaste compared to TACO dataset contains less annotated images but the annotations count per image is one and a half times higher. Furthermore, introduced image set include much smaller objects than the base one.

Feature		TACO dataset	UAVVaste dataset
Images count		1500	772
Annotations count		4782	3718
Annotations per image		3.188	4.816
Image shape [px]	minimum	842x474	1280x960
	median	2448x3264	3840x2160
	maximum	6000x4000	4056x3040
Object shape [px]	minimum	20x4	5x10
	median	180x159	76x68
	maximum	2987x3680	1188x754

Statistical comparison of datasets with trash annotations

Efficient object detection models



Example of annotated image from UAVVaste dataset

UAVVaste dataset was prepared for using as a train data for unmanned aerial vehicle rubbish detection system. Due to resource constraints of UAV, efficient state of the art algorithms were used to detect objects. Models for waste images were fitted:

- YOLOv4 with an input image of 608 px,
- EfficientDet-d3 with an input image of 896 px.

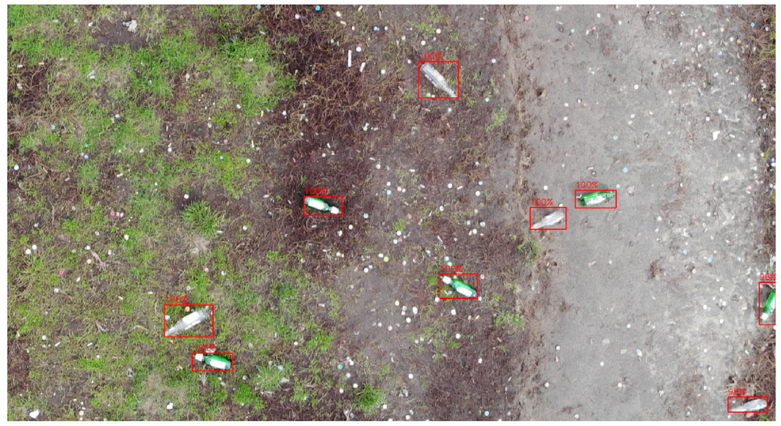
The obtained mean Average Precision (mAP) with an Intersection over Union (IoU) threshold rate equals 0.5 show that the created algorithms are able to learn to detect these objects regardless of the environment in which they are found. The performance gains with use of the algorithms' lighter versions (YOLOv4-tiny, EfficientDet-d1). Nevertheless, using of smaller detection neural networks decrease mAP and IoU metrics.

Model	Precision	mAP@0.5	Inference time [s]
YOLOv4-608	fp32	**0.785**	0.208
	fp16	0.784	0.073
YOLOv4-tiny-608	fp32	0.566	0.045
	fp16	0.566	**0.044**
EfficientDet-d3	fp32	0.751	0.406
	fp16	0.750	0.293
EfficientDet-d1	fp32	0.669	0.138
	fp16	0.669	0.110

Summary of UAVVaste results on different algorithms. Inference time measured on NVIDIA Jetson Xavier NX.

Applications

Quantization of model weights to half-precision floating-point allowed to obtain inference time on the embedded device, sufficient to build a real time application. Designing an unmanned aerial vehicle with the NVIDIA Jetson Xavier NX allows you to process data on the edge and reduce the amount of data transferred to the cloud. Possible use case of this applications for example is waste localization in a given area (e.g. squares, parks) or large area monitoring in order to search illegal landfills (e.g. fields, forests). **The goal is a modern and cleaner world!**



Example of inference using YOLOv4 algorithm

Acknowledgment & References

The project is implemented and financed under the project „System autonomicznego powietrznego i naziemnego monitoringu lotniska'' no. RO-076F/56/19/1114 with the number ERP 33/32/SIGR/0002.

[1] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection, 2020.

[2] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. Microsoft coco: Common objects in context, 2015.

[3] Pedro F Proença and Pedro Simões. Taco: Trash annotations in context for litter detection, 2020.

[4] Mingxing Tan, Ruoming Pang, and Quoc V. Le. Efficientdet: Scalable and efficient object detection, 2020.