YOLOv7 Small Object Detection Optimization to Detect Airborne Objects

Author
Dion Andreas Solang
NRP 07211940000039

Advisors
Reza Fuad Rachmadi, S.T., M.T., Ph.D
Dr. I Ketut Eddy Purnama S.T., M.T.

Department of Computer Engineering
Institut Teknologi Sepuluh Nopember

ABSTRACT

In this research, we present an attempt to improve the detection capability of YOLOv7 for airborne objects. Airborne objects appear considerably small in camera images when they are located at a considerable distance from the camera. However, due to their high speed of movement, it is crucial to detect them while they are still far away. Therefore, to effectively detect these objects, YOLOv7 needs to be optimized for small objects. To address this challenge, we proposed several modifications that include changes in the architecture (adding an extra detection head, modifying the feature-map source, and replacing the detection head with a detached anchor-free head), application of bag-of-freebies techniques (anchor recalculation and mosaic augmentation), and change in the inference process (partitioning the image and performing inference on each partition). Through comprehensive experimentation, we have discovered that the combination of replacing the detection head with a detached anchor-free head, and performing inference on partitions yields the most promising results, with a significant increase in mean average precision (mAP) of 46.18% while still maintaining real-time inference speed (greater than 10 FPS). This improvement is notably higher compared to the unmodified plain YOLOv7, which achieved a mAP score of 0%.

Objective

The objective of this research is to find modifications that can be made to YOLOv7 such that it could detect airborne objects better.

Introduction

One of the greatest challenge in autonomous flight is about the problem of sensing and avoiding (SAA) airborne objects like bird, airplane, helicopter, and other. Camera is a popular choice of sensor for this task due to its cheaper price and small payload. One problem however, airborne objects appear very small on cameras. In Airborne Object Tracking Dataset 2023, the size is about 4-1000 pixels in a 20 million pixels camera. Detecting such small objects is very challenging.
In this research, we will try to optimize YOLOv7 to solve this challenge. YOLOv7 is an general real-time object detection architecture with the highest accuracy at the time this research was proposed (September 2022). YOLOv7 can be scaled down so that it can run on consumer GPU or edge computing devices. Its high accuracy and low computational cost were the reason why YOLOv7 was chosen for this study.
To optimize YOLOv7, we will carry out modifications to the bag-of-freebies and architecture of YOLOv7. These modifications however must not cause YOLOv7 lose its ability to detect objects in real-time.

Methodology: Modification Candidates

- Mosaic Augmentation
  We add mosaic augmentation to the training dataset to improve small object detection.
- Anchor Recalculation
  We performed anchor recalculation to have anchors that better fit the airborne object dataset.
- EIoU localization loss
  We replace the original CIoU loss of YOLOv7 to EIoU.
- Reroute Neck Feature-map Source
  We reroute the source of feature map for the head to an earlier scale of prediction to reduce the length of data propagation to the neural network, which would lead to fewer information loss.
- Additional Detection Layer
  We add a head detection layer to the neural network, increasing from the original 3 to 4.
- Replacing Detection Layer to Decoupled Anchor-Free Head
  We replace the head layer from anchor head to decoupled anchor-free head of YOLOv6.
- Partitioning Image for Inference
  We do detection by partitioning the image into 4.

Result

- Mosaic Augmentation and Anchor Recalculation

No	Model	mAP@50
0	`YOLOv7-plain`	0%
1	`YOLOv7-M`	0%
2	`YOLOv7-AR`	0%
3	`YOLOv7-MAR`	**11.2%**
	Improvement	**+11.2%**

- EIoU Localization Loss

No	Modification	mAP@50
0	`YOLOv7-MAR +CIoU` (original)	**11.2%**
1	`YOLOv7-MAR + EIoU`	0%
2	`YOLOv7-MAR + EIoU + Convexication`	4.92%
	Improvement	-6.28%

- Reroute Neck Feature-map Source

No	Modification	mAP@50
0	`YOLOv7-MAR`	11.2%
1	`YOLOv7-MAR + rerouting`	**14.09%**
	Improvement	**+2.98%**

- Additional Detection Layer

No	Modification	mAP@50
0	`YOLOv7-MAR`	**11.2%**
1	`YOLOv7-MAR + more head`	5.19%
	Improvement	-6%

- Replacing Detection Layer to Decoupled Anchor-Free Head

No	Model	mAP@50
0	`YOLOv7-MAR (anchor head)`	11.2%
1	`YOLOv7-MAR + anchor-free head`	0%
1	`YOLOv7-MAR + anchor-free head` *numerically stabilized	**32.98%**
	Improvement	**+21.78%**

- Partitioning Image for Inference

No	Model	Input Size	Partition4 mAP@50
0	`YOLOv7-AR`	960	2.9%
1	`YOLOv7-MAR`	640	18.46%
3	`YOLOv7-MAR`	960	37.69%
4	`YOLOv7-MAR + rerouting`	960	30.04%
5	`YOLOv7-MAR + more head`	960	10.53%
6	`YOLOv7-M + anchor-free`	640	37.57%
7	`YOLOv7-M + anchor-free`	960	**46.18%**

From the result, we can conclude that using anchor-free head with mosaic augmentation and partitioning, we can produce the greatest mAP score of 46.18%.

solang.dion@gmail.com