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

Airborne objects appear very small on cameras. YOLOv7 is the state of the art real-time object detector optimized for general object detections. Thus, to detect airborne objects with YOLOv7, modifications are needed to be applied. The purpose of this research is to find a modification solution for YOLOv7 to optimize its small object detection capability especially for airborne objects. Modifications to be applied on YOLOv7 consists of architecture modifications and bag-of-freebies modifications. Architecture modifications consist of neck modification and head layer addition. Bag-of-freebies modifications consist of mosaic data augmentation dan active anchor recalculation. These modifications will be combined one with another and have their performance tested. Modifications that produces model with the highest mAP score on airborne objects dataset will be chosen as the optimization solution of this research.

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

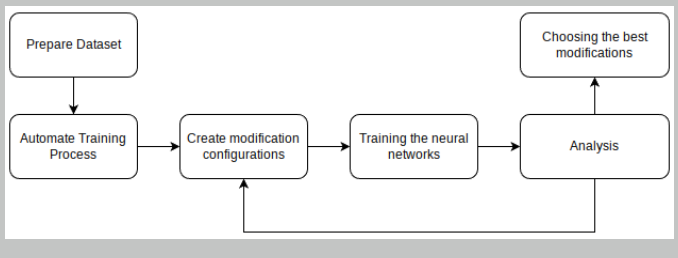
One of the greatest challenge in autonomous flight is about the problem of sensing and avoiding (SAA) airborne objects like bird, airplane, helicopter, and other. Camera is a popular choice of sensor for this task due to its cheaper price and small payload. One problem however, airborne objects appear very small on cameras. The size is about 4-1000 pixels in a 20 million pixels camera. Detecting such small objects is very challenging.
In this study, 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 its paper was published (July 2022). YOLOv7 can be scaled down so that it can run on edge computing devices such as Jetson TX2. 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.

Related Works

There are some similar attempts of modifying YOLO architecture to improve its object detection capability.

- YOLOZ
  - Based on YOLOv5.
  - Changed YOLOv5 backbone to DenseNet.
  - Changed the neck to biFPN.
  - Optimized to detect autonomous racing cones.
- exYOLO
  - Based on YOLOv3.
  - Added Receptive Field Block prior to feature fusion in the neck.

Methodology: Flow of Work



Methodology: Step by Step

To conduct this study, first the dataset must be prepared to conform with format understandable by YOLOv7. Then, to aid and accelerate the process of this study, we will automate the training process. Modification configurations will be made and then inputted to the trainer. These modifications might or might not include a combinations of modifications candidates. The trainer will built neural network according to the modification configuration and train them. The traned neural networks will be analyzed for their performance and to look for potential modifications that can be made to optimize the neural network more. If such modification was found, we will create a configuration for it and then feed it to the trainer. Finally, among all of the modifications, the best model will be chosen. The best model will be decided according to their mAP score.

Modifications Candidates

- Bag-of-Freebies Modifications
  - On-training Anchor Box Recalculation A layer will be added in the neural network to learn the optimal size for anchor box. In most YOLO architectures, the anchor box are constant and calculated prior to training.
  - Mosaic Augmentation Mosaic Augmentations has been proven to increase the object detection accuracy in YOLOv4 and YOLOv5.
- Architecture Modifications
  - Neck Modifications Some layer of the neck will be modified to take input from a more shallow layer of the backbone of YOLOv7.
  - Head layer addition Another head layer will be added so that YOLOv7 can detect on a higher scale.

Sampled Dataset Distribution

Division	Total image	Percentage				
		Airplane	Helicopter	Bird	*Other*	Negative
Train	54k	23,75%	23,75%	23,75%	23,75%	5%
Valid	3k	20%	20%	20%	20%	20%
Test	3k	20%	20%	20%	20%	20%

Contact Information

- Implementation: https://github.com/messier12/final-project-yolo-optimization.git
- Email: solang.dion@gmail.com
- Phone: +62 813 2752 2023