Project Proposal - 13th April

Title:

"Scaled-YOLOv4: Scaling Cross Stage Partial Network ",
Chien-Yao Wang, Alexey Bochkovskiy, Hong-Yuan Mark Liao
https://doi.org/10.48550/arXiv.2011.08036

Name – Kishan Kumar Ravikumar
Course – CS 541 Artificial Intelligence
Mail Id – kravikumar@stevens.edu

Abstract

The authors show that the CSP-based YOLOv4 object identification neural network scales up and down and can be used to small and large networks while maintaining optimal speed and accuracy. The authors   suggest a network scaling method that alters not just the network's depth, width, and resolution, but also its structure. The YOLOv4-large model achieves state-of-the-art results, with 55.5 percent AP (73.4 percent AP50) for the MS COCO dataset at a speed of 16 FPS on Tesla V100, and 56.0 percent AP with test time augmentation (73.3 AP50). To the best of our knowledge, this is the most accurate study on the COCO dataset that has ever been released. On the RTX 2080Ti, the YOLOv4-tiny model scores 22.0 percent AP (42.0 percent AP50) at 443 FPS, while by using TensorRT, batch size = 4 and FP16-precision the YOLOv4-tiny achieves 1774 FPS.

Github link (in abstract): https://github.com/Kishanjr/CS-541-Artificial-intelligence.git

Data:

MS COCO dataset was released in its initial version in 2014. There are 164K photos in total, including 83K in training, 41K in validation, and 41K in testing. In 2015, a new 81K image test set was released, which included all of the previous test images as well as 40K new photographs.

The training/validation split was modified from 83K/41K to 118K/5K in 2017 in response to community feedback. The same images and notes appear in the new split. The 2017 test set is a subset of the 2015 test set, which consists of 41K photos. A new unannotated dataset of 123K photos is included in the 2017 release as well.