Task: 08

YOLO

YOU ONLY LOOK ONCE -- V5

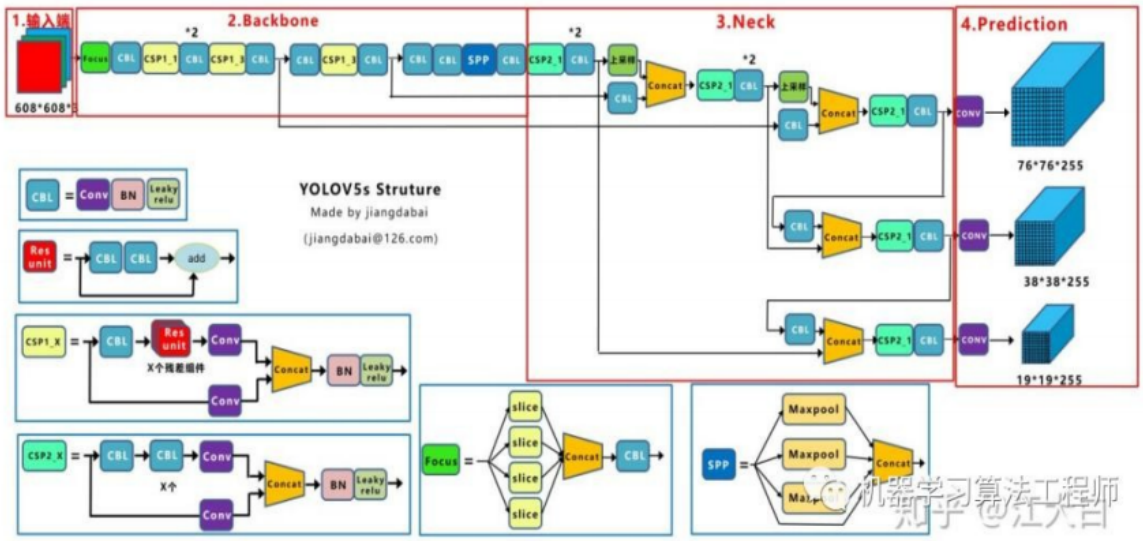
Yolo v5 is almost similar to yolo v4, but some modifications are done in architecture and optimization. Yolo v5 is implemented in the PyTorch framework.

The architecture of YOLOv4 and YOLOv5 is very similar and it makes many people dissatisfied with the name YOLOv5 (5th generation of YOLO) when it does not contain multiple outstanding improvements compared to the previous version YOLOv4.

Yolo v5 consists of:

Backbone: Focus structure, CSP network
Neck: SPP block, PANet
Head: YOLOv3 head using GIoU-loss

Architecture:



Backbone :

Cross Stage Partial Darknet(CSPDarknet) is used as Backbone for Yolo v5. Cross Stage Partial Net is the Modified version of DenseNet and DenseNet is the advanced version of ResNet.

Focus structure

1. Focus structure, there is no such structure in Yolov3 & Yolov4, and the key is the slicing operation.

2. For example, in the slice diagram on the right, the 4*4*3 image is sliced into a 3*3*12 feature map.

CSP structure

1. There are two CSP structures designed in Yolov5.
2. The difference between Yolov5 and Yolov4 is that only the backbone network in Yolov4 uses the CSP structure but in YOLOv5 both backbone and neck use CSP structure.

Neck:

1. When Yolov5 first came out, only the FPN structure was used, and the PAN structure was added later, and other parts of the network were also adjusted.
2. In the Neck structure of Yolov4, ordinary convolution operations are used. In the Neck structure of Yolov5, the CSP2 structure designed by CSPNet is adopted to strengthen the ability of network feature integration

Output :

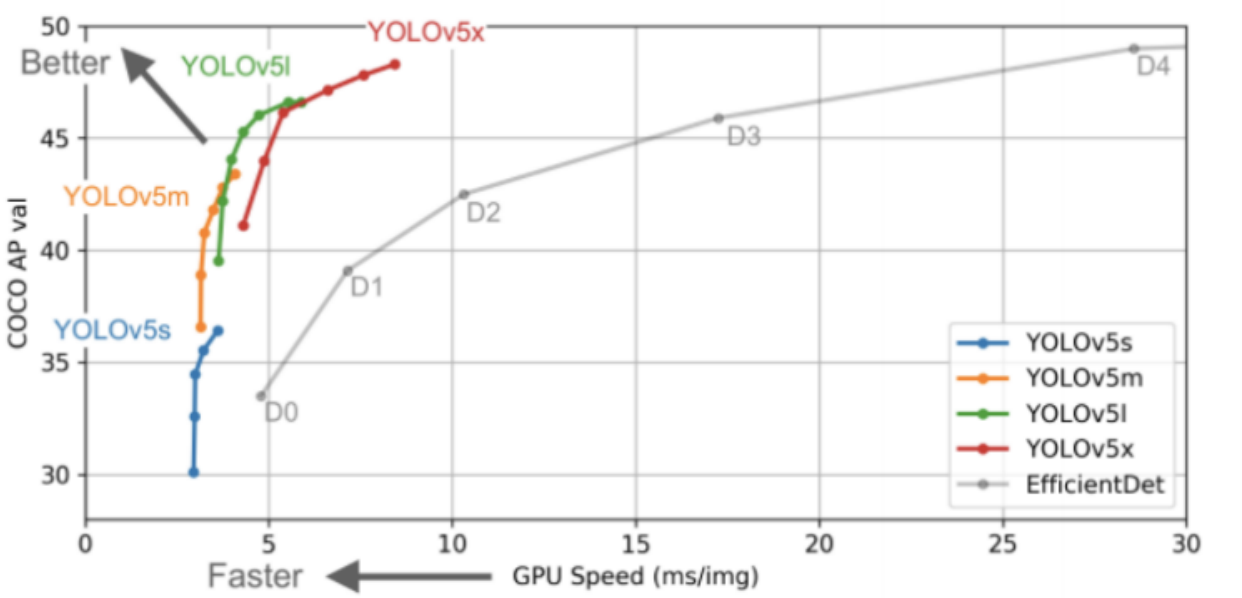
In the detection part all are the same as Yolo v4 , but the bounding box loss function is changed as Generalized Intersection of Union(GIoU).

1. Bounding box loss function:
    a. Yolov5 uses GIoU loss as the loss function of the Bounding box.
    b. In Yolov4, CIoU loss is used as the loss of the target Bounding box. If CIoU loss is used, there should be a faster and better convergence effect.

2. Non-maximum suppression:
    a. In the post-processing of target detection, the screening of many target frames usually requires nms operation.
    b. Yolov4 uses the DIoU nms method on the basis of DIoU loss, while the ordinary nms method is still used in Yolov5.

Optimization Techniques:

1. Mosaic data enhancement
2. Adaptive anchor frame calculation
3. Adaptive image scaling



Speed and Accuracy of Yolo v5 is higher than Yolo v4. Size is also very less. Yolo v5 is best suit for Embedded device applications