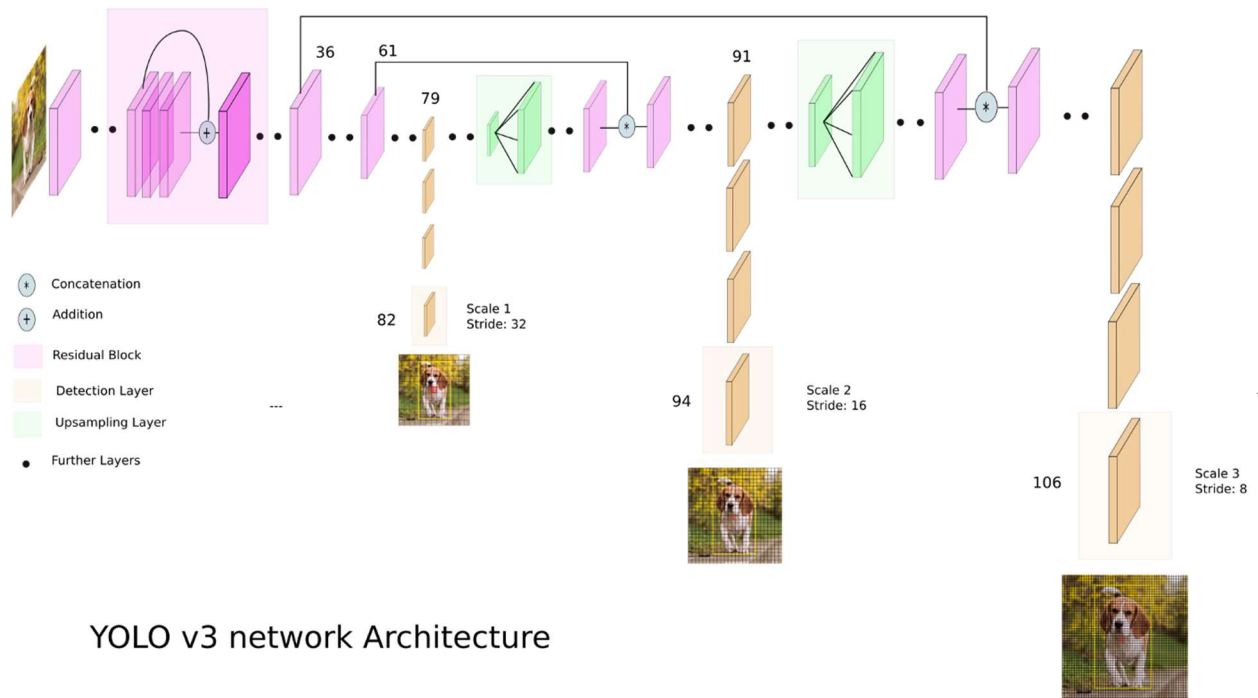


Here is my YOLOV3 implementation using Pytorch

Code is available on my github <https://github.com/jeremywinston/computervision>

YOLOV3 architecture is based on a darknet-53, the configuration file of darknet-53 is available on the /cfg/yolov3.cfg. Then darknet.py will read all of this text information and transform it into a neural network using torch.nn.module. We can import this python file later for the object detection program.



YOLO v3 network Architecture

YOLOV3 makes use of only convolutional layers. The prediction is also done by using a convolutional layer which uses 1×1 convolutions, so the output is a feature map and YOLOV3 has 3 of them at 3 different scale with stride 32, 16, 8. (for an 416×416 image the feature map is 13×13 , 26×26 , 52×52).

For each cell, the depth size is $(3 \times (5 + C))$. 3 represent 3 bounding boxes, where each bounding box have $5 + C$ attributes. 5 represent the center coordinates, the dimensions, the objectness score, and C represent the number of classes confidences for each bounding box.

In this project YOLOV3 is trained on coco data set which has 80 classes, so the total number of feature map is $1 \times (13 \times 13 + 26 \times 26 + 52 \times 52) \times (5 + 80) = 1 \times 10647 \times 85$ where 1 is the batch size.

The program from image detection is object_detection.py. For this project I not training the YOLO instead I'm using pre-trained model. The weights of the model located in yolov3.weight. First in the program we want to create a model by importing the darknet, afterward we load the pre trained weights to the model. Next, we will load all the images from COCO dataset available on the imgs folder and create a batch to speed up the process. This batch will be calculated in parallel fashion to speed up the inference process. After that we can predict the image and save the result in the det folder. To evaluate the model generalization, I download random pictures from the internet and tried to do object detection on that model and the result is reasonably good.

Result on COCO dataset



Result on picture from internet

