**YOLOX: Exceeding YOLO Series in 2021**

Most prominent improvements to YoloX are:

1. Switch the YOLO detector to an anchor-free manner
2. A decoupled head and the leading label assignment strategy SimOTA

Anchor-free model:

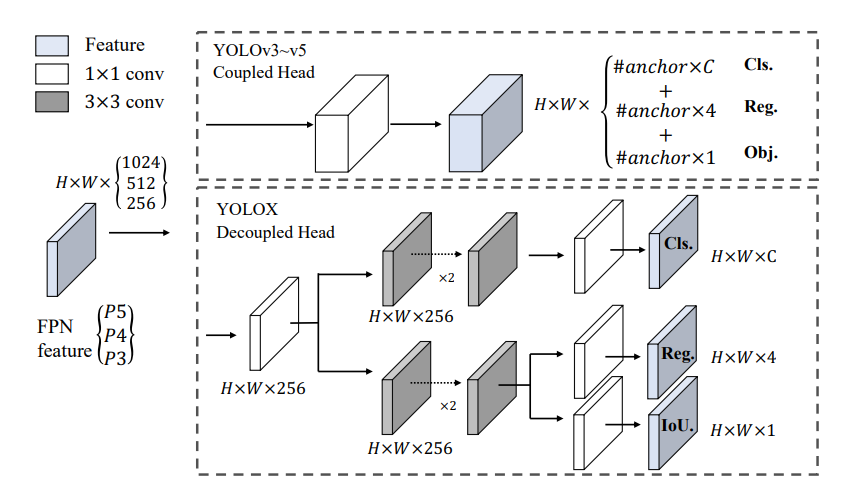
First we will talk about the disadvantages of anchor based models, that made us move towards an anchor-free model in YoloX

The anchor mechanism has many known problems. First, to achieve optimal detection performance, one needs to conduct clustering analysis to determine a set of optimal anchors before training. Those clustered anchors are domain-specific and less generalized.

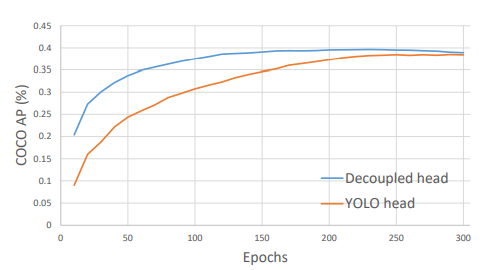
Second, anchor mechanism increases the complexity of detection heads, as well as the number of predictions for each image.

So in short moving to an anchor-free model significantly reduces the number of design parameters

Decoupled head:

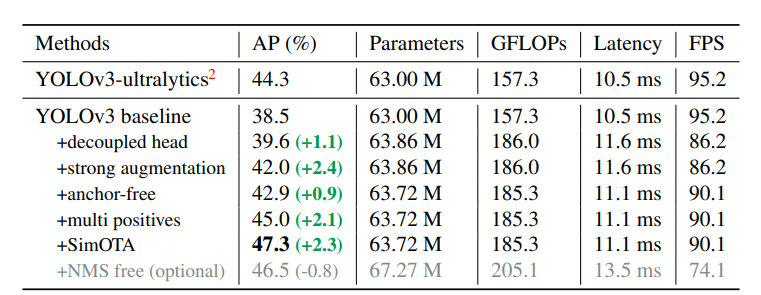


Replacing YOLO’s head with a decoupled one greatly improves the converging speed, as shown below:



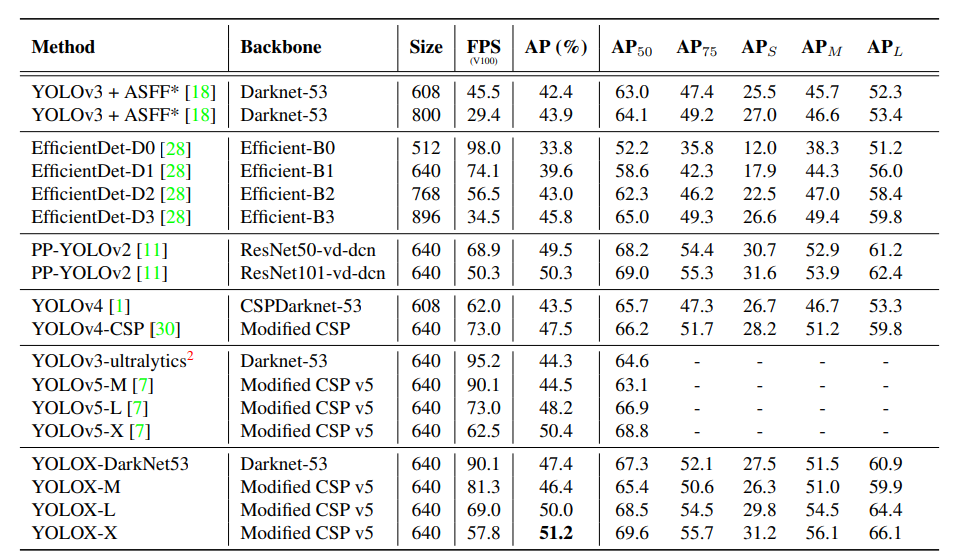
YOLOv3 Baseline

Since YOLOv4 and YOLOv5 are a little over-optimized for the anchor-based pipeline. The paper used YOLOv3 as a start point or a baseline model on which to build all of the ideas, and new techniques on. And in the end achieved the following results:



Results:

And here are the comparisons



Areas of improvement:

1. With everyone moving towards transformers, and the great results they have achieved in the last few years. An area for improvement is to have a version of YOLOx with a transformers backbone.

So far there has been 1 YOLO model that has used transformers as it’s backbone (ViT YOLO) and it hasn’t achieved the best results.

And I believe that the authors of this paper have said that they are working on a version that uses swine transformers in its backbone.

1. A larger model.
2. More features