

DISCLAIMER: this YOLO implementation is still under development. No support will be provided during the development phase.

YOLO Object Detectors, You Only Look Once

arXiv Paper arXiv.1804.02767 arXiv Paper arXiv.2004.10934

This repository is the unofficial implementation of the following papers. However, we spent painstaking hours ensuring that every aspect that we constructed was the exact same as the original paper and the original repository.

- YOLOv3: An Incremental Improvement: [YOLOv3: An Incremental Improvement](#)
- YOLOv4: Optimal Speed and Accuracy of Object Detection: [YOLOv4: Optimal Speed and Accuracy of Object Detection](#)

Description

YOLO v1 the original implementation was released in 2015 providing a ground breaking algorithm that would quickly process images and locate objects in a single pass through the detector. The original implementation used a backbone derived from state of the art object classifiers of the time, like [GoogLeNet](#) and [VGG](#). More attention was given to the novel YOLO Detection head that allowed for Object Detection with a single pass of an image. Though limited, the network could predict up to 90 bounding boxes per image, and was tested for about 80 classes per box. Also, the model can only make predictions at one scale. These attributes caused YOLO v1 to be more limited and less versatile, so as the year passed, the Developers continued to update and develop this model.

YOLO v3 and v4 serve as the most up to date and capable versions of the YOLO network group. This model uses a custom backbone called Darknet53 that uses knowledge gained from the ResNet paper to improve its predictions. The new backbone also allows for objects to be detected at multiple scales. As for the new detection head, the model now predicts the bounding boxes using a set of anchor box priors (Anchor Boxes) as suggestions. Multiscale predictions in combination with Anchor boxes allow for the network to make up to 1000 object predictions on a single image. Finally, the new loss function forces the network to make better predictions by using Intersection Over Union (IOU) to inform the model's confidence rather than relying on the mean squared error for the entire output.

Authors

- Vishnu Samardh Banna ([@GitHub vishnubanna](#))
- Anirudh Vegesana ([@GitHub anivegesana](#))
- Akhil Chinnakotla ([@GitHub The-Indian-Chinna](#))
- Tristan Yan ([@GitHub Tyan3001](#))
- Naveen Vivek ([@GitHub naveen-vivek](#))

Table of Contents

- [Our Goal](#)
- [Models in the library](#)
- [References](#)

Our Goal

Our goal with this model conversion is to provide implementation of the Backbone and YOLO Head. We have built the model in such a way that the YOLO head could be connected to a new, more powerful backbone if a person chose to.

Models in the library

Object Detectors	Classifiers
Yolo-v3	Darknet53
Yolo-v3 tiny	CSPDarknet53
Yolo-v3 spp	
Yolo-v4	
Yolo-v4 tiny	
Yolo-v4 csp	
Yolo-v4 large	

Models Zoo

Requirements



DISCLAIMER: this YOLO implementation is still under development. No support will be provided during the development phase.