What is YOLO algorithm

Ishani Dey · Follow
5 min read · Jul 12, 2023

👏 --   💬 1   🔖   ▶   ⬆

You only look once(YOLO) is a state-of-the-art, real-time object detection system. It is so fast, that it has become the standard way of detection of objects in the field of computer vision. The algorithm was introduced in 2015 by Joseph Redmon. And ever since it came out it has surpassed other algorithms such as sliding window object detection, R CNN, Fast R CNN, Faster R CNN, etc.

How it works

> *Prior detection systems use classifiers or localisers to perform detection. They apply the model to an image at multiple locations and scales. High scoring regions of the image are considered detections.*
>
> *We use a totally different approach. We apply a single neural network to the full image. This network divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities.*

Taken from https://pjreddie.com/darknet/yolo/

"Imagine you built a YOLO application that detects players and soccer balls from a given image.

But how can you explain this process to someone, especially non-initiated people?

→ That is the whole point of this section. You will understand the whole process of how YOLO performs object detection; how to get image (B) from image (A)"

The algorithm works based on the following four approaches:

- Residual blocks

- Bounding box regression

- Intersection Over Unions or IOU for short

- Non-Maximum Suppression.

1- Residual blocks

This first step starts by dividing the original image (A) into NxN grid cells of equal shape, where N in our case is 13 shown on the image on the left. Each cell in the grid is responsible for localizing and predicting the class of the object that it covers, along with the probability/confidence value.

Bounding box regression

Bounding box regression is a technique used in object detection tasks to predict the coordinates of a bounding box that tightly encloses an object of interest within an image.

YOLO determines the attributes of these bounding boxes using a single regression module in the following format, where Y is the final vector representation for each bounding box.

Y = [pc, bx, by, bh, bw, c1, c2]

This is especially important during the training phase of the model.

- **pc** corresponds to the probability score of the grid containing an object.

- **bx, by** are the x and y coordinates of the center of the bounding box with respect to the enveloping grid cell.

- bh, bw correspond to the height and the width of the bounding box with respect to the enveloping grid cell.

- c1 and c2 correspond to the two classes Player and Ball. We can have as many classes as your use case requires.

Intersection Over Union (IOU):

Intersection over Union, commonly referred to as IOU, is a metric used to evaluate the overlap between two bounding boxes or regions of interest. It quantifies the similarity or agreement between the predicted bounding box and the ground truth bounding box. IOU is calculated as the ratio of the intersection area to the union area of the two bounding boxes. It is often used as a criterion for evaluating the performance of object detection algorithms, where a higher IOU indicates better detection accuracy.

4- Non-Max Suppression or NMS

Setting a threshold for the IOU is not always enough because an object can have multiple boxes with IOU beyond the threshold, and leaving all those boxes might include noise. Here is where we can use NMS to keep only the boxes with the highest probability score of detection.

How does YOLO work?

Let's understand the architecture of YOLO better. The architecture of the CNN model that forms the backbone of YOLO is shown below.

The first 20 convolution layers of the model are pre-trained using ImageNet. Then, this pre-trained model is converted to perform detection.YOLO's final fully connected layer predicts both class probabilities and bounding box

coordinates. And the what happens next has already been discussed as the approaches before.

Comparisons with other detectors

YOLOv3 is extremely fast and accurate.

> *Moreover, you can easily tradeoff between speed and accuracy simply by changing the size of the model, no retraining required!*

Taken from : [*https://pjreddie.com/darknet/yolo/*](https://pjreddie.com/darknet/yolo/)

Conclusion

Yolo is used for object detection in images and videos as well. It can be implemented in many cases depending on the requirement. If you want to know how to detect objects yourself using a pre-trained model or train YOLO on different kinds of data, please have a look at this : https://pjreddie.com/darknet/yolo/

Hello👋 I am Ishani, final year engineering student. I love web development and I love programming and I love travelling and I love photography and I love gardening and I love painting and I love…Medium❤️
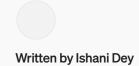
If you enjoyed reading this post, I would appreciate your support by giving it a clap. Additionally, I encourage you to check out my other posts for more content that you may find interesting. Thank you!

Knowledge gathered from:

> [*https://pjreddie.com/darknet/yolo/*](https://pjreddie.com/darknet/yolo/)
>
> [*https://www.v7labs.com/blog/yolo-object-detection*](https://www.v7labs.com/blog/yolo-object-detection)
>
> [*https://www.datacamp.com/blog/yolo-object-detection-explained*](https://www.datacamp.com/blog/yolo-object-detection-explained)

Yolo   Machine Learning   Science   Object Detection   Neural Networks

👏 --   💬 1   🔖   ⬆️

Written by Ishani Dey

6 Followers

Follow