Object Detection — Anchor Box VS Bounding Box

Nikita Malviya · Follow

2 min read · Aug 6, 2023

▶ Listen　　⬆ Share　　••• More

Open in app ↗

Medium　🔍 Search　　🔔¹ 👤

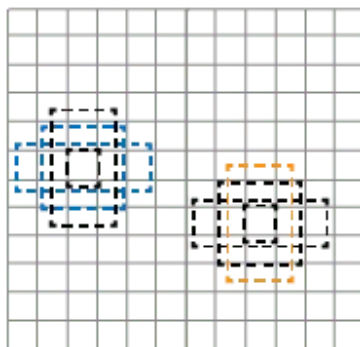**objectness scores. These predictions help refine the anchor boxes to match the ground-truth object locations.** Bounding boxes, on the other hand, **are the ground-truth annotations used to evaluate the model's accuracy in localizing objects** during training and testing.

Ground truth image and bounding boxes

Anchor boxes at each predefined location in each feature map

Predefined anchor box location
Predicted offset
Refined location of anchor box
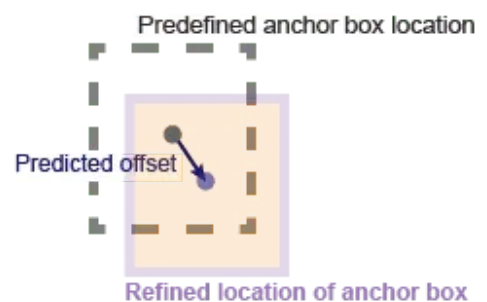
Bounding Boxes:

A bounding box is a rectangular region that is used to enclose an object or a specific region of interest within an image. It is represented by four coordinates: (x_min, y_min) representing the top-left corner, and (x_max, y_max) representing the bottom-right corner of the rectangle. Bounding boxes are **used to localize objects**

within an image and are commonly annotated in object detection datasets to provide ground-truth information about the object's location.

Anchor Boxes:

Anchor boxes, also known as **anchor priors or default boxes, are pre-defined bounding boxes with specific sizes, aspect ratios, and positions** that are **used as reference templates during object detection.** These anchor boxes **are placed at various positions across an image, often in a grid-like pattern, to capture objects of different scales and shapes.** During training and inference, anchor boxes are used to predict the locations and shapes of objects relative to these reference boxes.

Key Differences:

Definition and Purpose:

- Bounding boxes **represent the actual regions in an image** that enclose objects of interest.

- Anchor boxes are **reference bounding boxes used to predict object locations and shapes** during object detection.

Annotating vs. Predicting:

- Bounding boxes are **provided as annotations** in object detection datasets, **representing ground-truth** object locations.

- Anchor boxes are **used as templates for prediction** during model training and inference. The **model predicts adjustments to anchor boxes** to match objects' actual locations.

Static vs. Adaptive:

- Bounding boxes are **static** and **determined by annotators** based on the object's true boundaries.

- Anchor boxes **are adaptable** and **can be designed with different scales and aspect ratios** to handle various object sizes and shapes.

Localization vs. Reference:

- Bounding boxes are **used for object localization** and are compared to predicted boxes to measure accuracy.

- Anchor boxes **serve as reference points for predicting object locations** and shapes.

Object Detection   Bounding Box   Anchor Boxes   Yolo   Faster R Cnn

Follow

Written by Nikita Malviya

126 Followers   ·   15 Following

AI Engineer | Data Science | Computer Vision| Python | https://www.linkedin.com/in/nikita-malviya-/

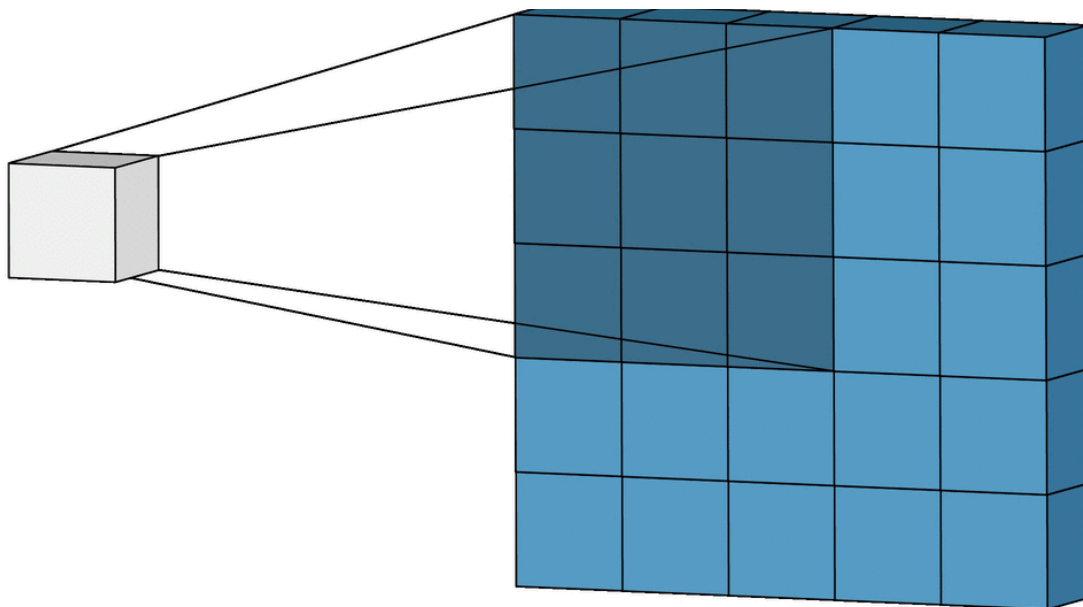# Responses (1)

What are your thoughts?

Respond

Manthan Manalwar
7 months ago

Explained in a simplified language , Reading this article an hour before exam . Helpful !
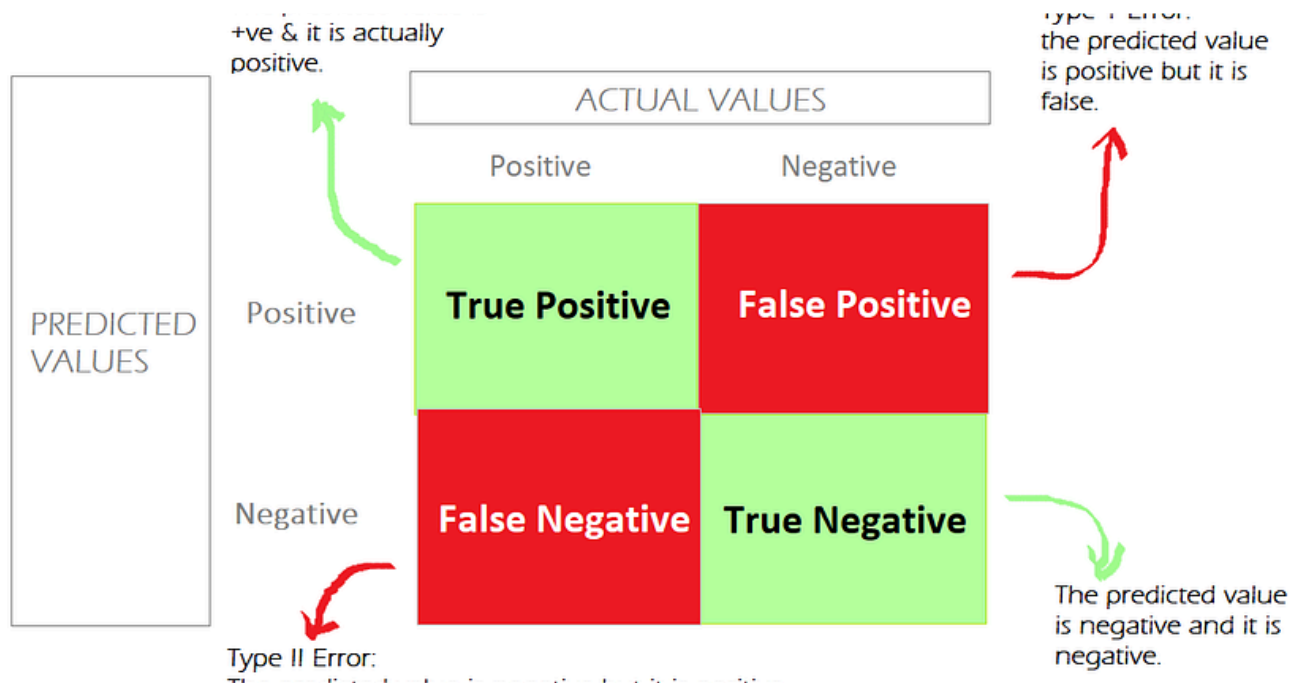
Reply

More from Nikita Malviya

Nikita Malviya

Convolution & Pooling

Convolution Layer:

Aug 3, 2023   👋 20   💬 1

Nikita Malviya

Confusion Matrix

A confusion matrix is a table that is used to describe the performance of a classification model on a set of data for which the true values...

Nikita Malviya

Bird's Eye View: A New Perspective

In computer vision, various viewpoints or perspectives can be utilized depending on the specific application and the sensor modalities...

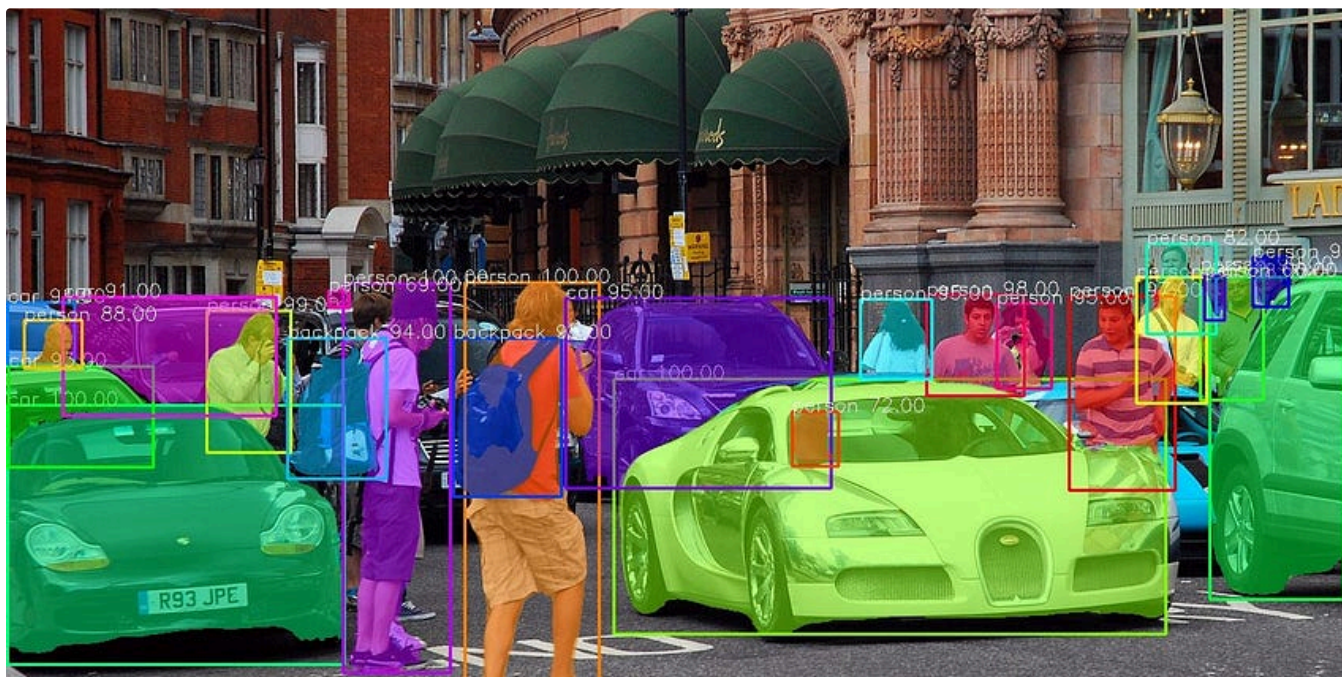Nikita Malviya

Managing Service Files in Linux

In the world of Linux system administration, managing services is a crucial aspect to ensure the proper functioning and availability of...

Aug 7, 2023   👏 5   🔖   •••

See all from Nikita Malviya

Recommended from Medium

🟢 Thoughts By Swarnali

Performance Comparison of Faster R-CNN and YOLO for Real-time Object Detection

Introduction

⭐ Jul 10   👏 61   🔖   •••
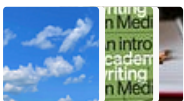
👤 Muhammad Rizwan Munawar

Object Counting using Ultralytics YOLOv8

This article provides a technical implementation overview of how to perform Object Counting using Ultralytics YOLOv8. You will learn the …
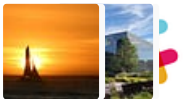
✦   Jun 30   👋 53   💬 1   🔖⁺   •••

Lists

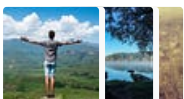**Staff picks**
781 stories · 1491 saves

Stories to Help You Level-Up at Work
19 stories · 891 saves

Self-Improvement 101
20 stories · 3117 saves

Productivity 101
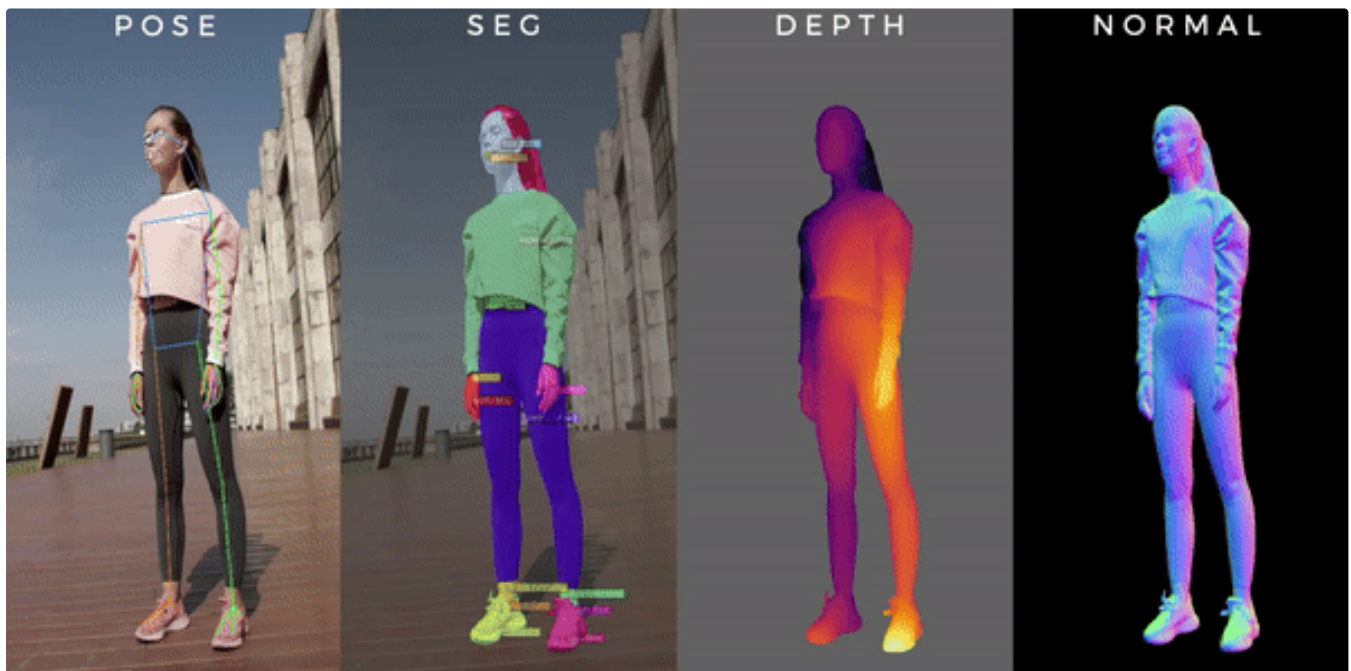20 stories · 2625 saves

👤 Smith

Yolo

∘ 1. Introduction ∘ 2. Understanding YOLO Architecture ∘ Key Components of YOLO ∘ 3. YOLOv3 Implementation ∘ 4. YOLO Evolution...

Aug 3 👋 1   🔖⁺   ⋯

🅰 AI Papers Academy

Sapiens by Meta AI: Foundation for Human Vision Models

In this post we dive into Sapiens, a new family of computer vision models by Meta AI that show remarkable advancement in human-centric...
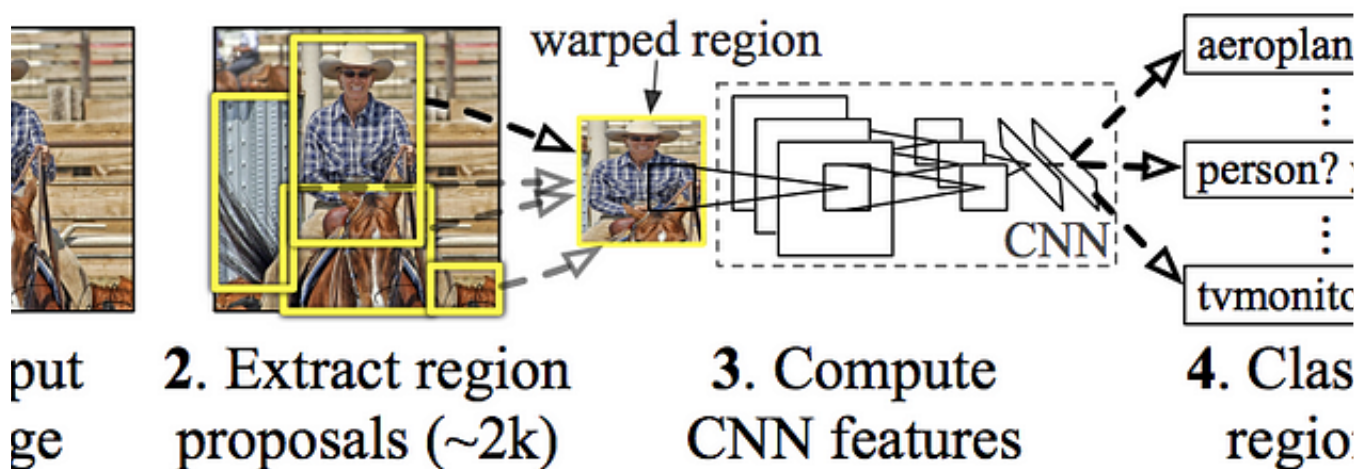
In The Deep Hub by Jorgecardete

Convolutional Neural Networks: A Comprehensive Guide

Exploring the power of CNNs in image analysis

J Jo Wang

R-CNN family explanation

The R-CNN (Regions with Convolutional Neural Networks) family includes several models that improve object detection by leveraging region...

Oct 20   👋 4

See more recommendations