SNS COLLEGE OF TECHNOLOGY

(An Autonomous Institution)

COIMBATORE- 641 035

Department of Computer Science and Engineering

19CST302-Neural Networks and Deep learning

Object Detection

Object detection is a foundational computer vision technique that plays a crucial role in identifying and labeling objects within various types of visual data, including images, videos, and live footage. To enable object detection, models are trained using large datasets containing annotated visuals, which provide the necessary information for the model to recognize and localize objects accurately in new data.

The process of object detection is streamlined and straightforward, involving the input of visuals into the model, which then produces fully marked-up output visuals. This marked-up output typically includes bounding boxes that precisely outline the detected objects. Bounding boxes are clear-cut quadrilaterals, usually in the form of squares or rectangles, which encapsulate the spatial extent of the identified objects within the visual data.

An essential aspect of object detection is the association of each bounding box with a label that describes the object it encloses. These labels provide semantic information about the detected objects, such as "person," "car," or "dog," enabling users to understand the content of the visual data at a glance.

Bounding boxes are often accompanied by confidence scores, indicating the model's certainty in its predictions. Additionally, bounding boxes

can overlap to represent scenarios where multiple objects are present within a single shot or frame. However, this requires the model to have prior knowledge of the types of objects it is expected to detect, ensuring accurate localization and labeling of all relevant objects in the visual data.

In essence, object detection leverages bounding boxes and associated labels to enable computers to perceive and understand the contents of visual data, laying the groundwork for a wide range of applications in fields such as surveillance, autonomous driving, object tracking, and more.

Tasks in object detection

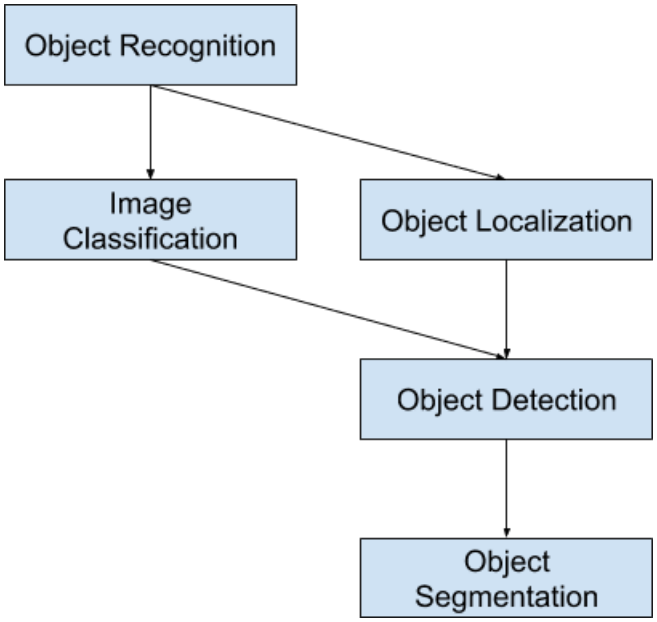

Image Classification:

Description: Image classification predicts the class or category of a single object within an image.

Task: Given an image, the model assigns it to one of several predefined classes (e.g., “cat,” “dog,” “car”).

Example: Identifying whether an image contains a specific animal or object.

Object Localization:

Description: Object localization involves identifying the location of one or more objects within an image.

Task: The model not only classifies the object but also draws bounding boxes around it.

Example: Detecting and localizing pedestrians in a street scene.

Object Detection:

Description: Object detection combines classification and localization to identify and classify multiple objects in an image.

Task: The model predicts bounding boxes and assigns class labels to each detected object.

Example: Detecting cars, pedestrians, and traffic signs in a traffic surveillance video.

Most Popular Object Detection Algorithms

YOLO (You Only Look Once): A real-time object detection system that divides images into a grid and predicts bounding boxes and class probabilities for each grid cell.

SSD (Single Shot MultiBox Detector): An algorithm that uses a single deep neural network to detect objects in images at different scales and aspect ratios.

R-CNN (Region-Based Convolutional Neural Networks): A family of algorithms that apply deep models to object detection, starting with region proposals and then classifying them.

Fast R-CNN: An improvement over R-CNN, Fast R-CNN uses a singleneural network to process the whole image and includes a Region of Interest (ROI) pooling layer.