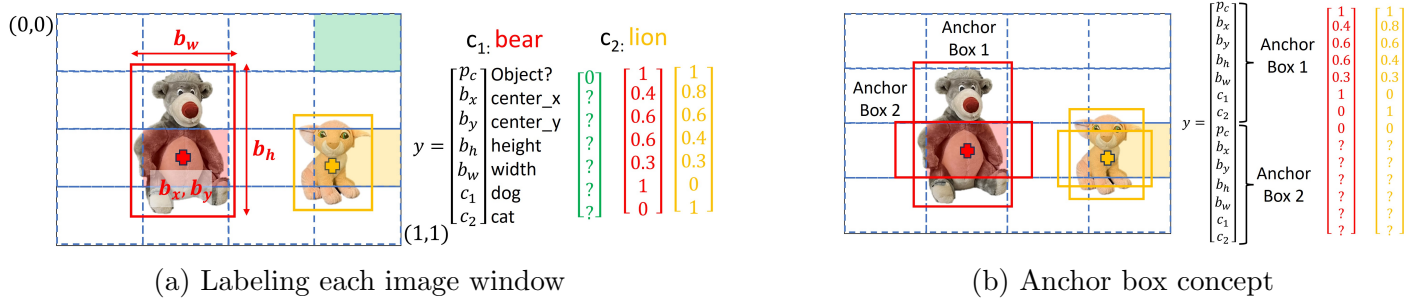


YOLO: You Only Look Once

YOLO (2015) is an object detection algorithm used in computer vision. It is fast and more efficient than traditional methods, making it widely used in applications like surveillance, autonomous driving, and object tracking.

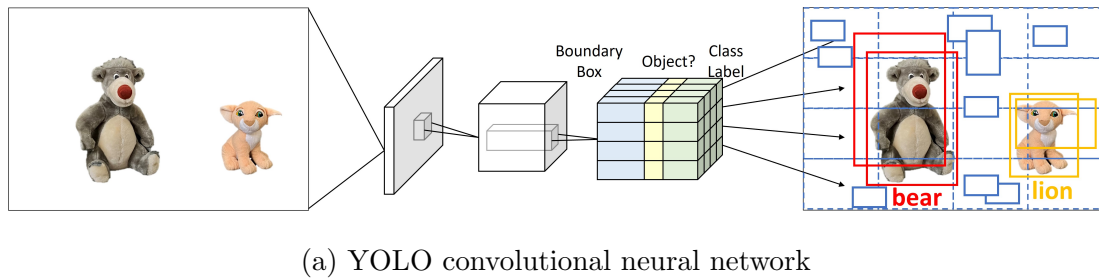
Step 1: Labeling each grid cell

Each grid cell has multiple anchor boxes which are used to detect multiple objects, overlapping objects, and objects of different scales. All grid cells with multiple anchors are labeled.



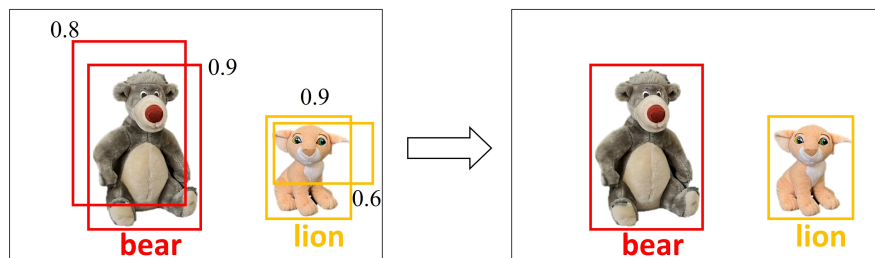
Step 2: Bounding boxes of each cell

During the training, the image is passed through YOLO CNN. To train the model, regression and classification losses are used for bounding boxes and class predictions respectively for each image grid cell.



Step 3: Select the right bounding box using Non-Maximum Suppression (NMS)

Each grid cell will predict anchor boxes during the testing. Discard low object probability boxes and for objects, apply NMS to obtain a single prediction per object. It includes picking the box with the highest probability output p_c as prediction and discarding any remaining box with $\text{IoU} \geq 0.5$ with box output in the previous step.



Summary:

Version	Year	New features
YOLOv2	2016	Batch normalization and anchor boxes
YOLOv3	2018	Efficient backbone network and spatial pyramid pooling
YOLOv4	2020	New loss function and anchor-free detection head.
YOLOv5	2020	Export formats, hyperparameter optimization, and integrated tracking.
YOLOv6	2022	Used for autonomous delivery robots
YOLOv7	2022	Pose estimation added.
YOLOv8	2023	Developed by Ultralytics and supports several tasks of computer vision.