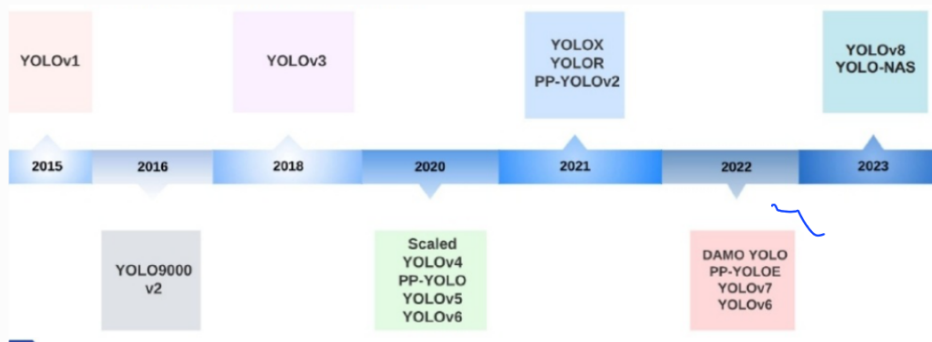


Yolo (You Only look once)

• เป็น Object Detection model (ดูทั้งภาพ, ทำภาพ)

Timeline



Process

- แบ่งรูปออกเป็น $s \times s$ grid โดยแต่ละ grid ก็จะมี Bounding Box
- Prediction Parameter ของ Bounding Box ได้แก่ $x, y, w, h, confidence$
- Yolo เป็น Regression model ประกอบด้วยตัวแปร x, y ซึ่ง

x เป็นภาพขนาด $w \times h \times RGB$ value $c \times 1$

y เป็นผลการทำนาย $5B + C$ โดย $5B$ คือ $x, y, w, h, confidence$ และ C คือ class predicted distribution

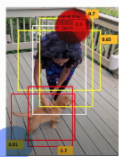


Non-maximal suppression

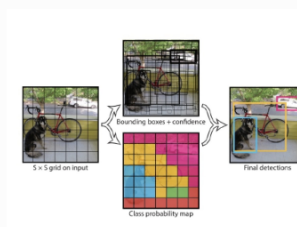
• ใช้ threshold เพื่อหา BB ที่มีค่าเป็นไปได้อาจจะมีซ้ำกันนั้นทิ้งออกไป

• สูตร: $\Pr(\text{Class}_i | \text{Object}) * \Pr(\text{Object}) * \text{IOU}_{\text{pred}}^{\text{truth}} = \Pr(\text{Class}_i) * \text{IOU}_{\text{pred}}^{\text{truth}}$

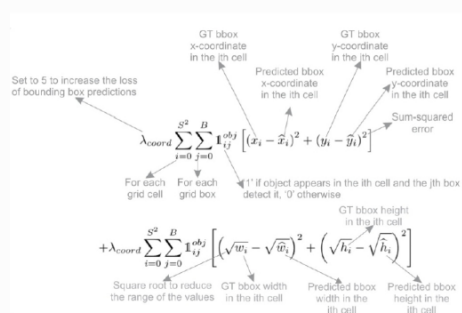
• มีค่าเป็นไปได้อาจจะเกิดซ้ำนี้



โดยมันจะเลือก BB ที่ค่า PC objects มากสุด



Localization Loss



Confidence loss

$$+ \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{obj} \left[(C_i - \hat{C}_i)^2 \right]$$

GT confidence score

Predicted confidence score

Confidence error when an object is detected in the i th cell

Set to 0.5 to decrease the loss for empty boxes

$$+ \lambda_{nobj} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{nobj} \left[(C_i - \hat{C}_i)^2 \right]$$

Confidence error when an object not detected in the i th cell

*1 if there is no object in the i th cell, 0 otherwise

Classification loss

$$+ \sum_{i=0}^{S^2} \mathbb{1}_i^{obj} \sum_{c \in \text{classes}} \left[(p_i(c) - \hat{p}_i(c))^2 \right]$$

Predicted conditional probability of an object of class c appearing in the i th cell

GT conditional probability of class c appearing in the i th cell

For each grid cell

For each class

YoloV8

- 95 Backbone instead of YoloV5 Tambahkan CSPLayer \rightarrow C2f module
- C2f module (cross-stage partial bottleneck with two convolutions): combine high-level feature with contextual information to improve accuracy
- YoloV8 architecture

