



Real-Time YOLOv4 Object Detection with Web Camera in Google Colab

Mengyu Zhang
Yufeng Shi
Xuhua Sun
Zijian Wu

Introduction

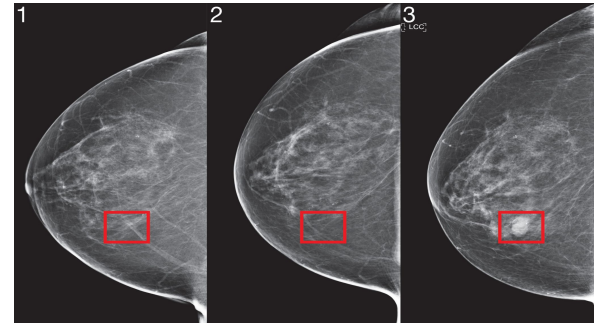
Why object detection matters?



E-commerce^[1]



Autonomous Driving^[2]



AI for Medicine^[3]

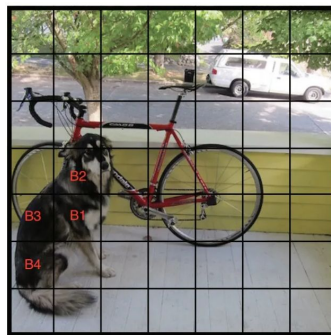
[1] <https://artificialintelligence.oodles.io/blogs/computer-vision-applications-for-ecommerce/>

[2] <https://blog.waymo.com/2020/07/opendataset-challenge-winners.html>

[3] <https://news.mit.edu/2021/robust-artificial-intelligence-tools-predict-future-cancer-0128>

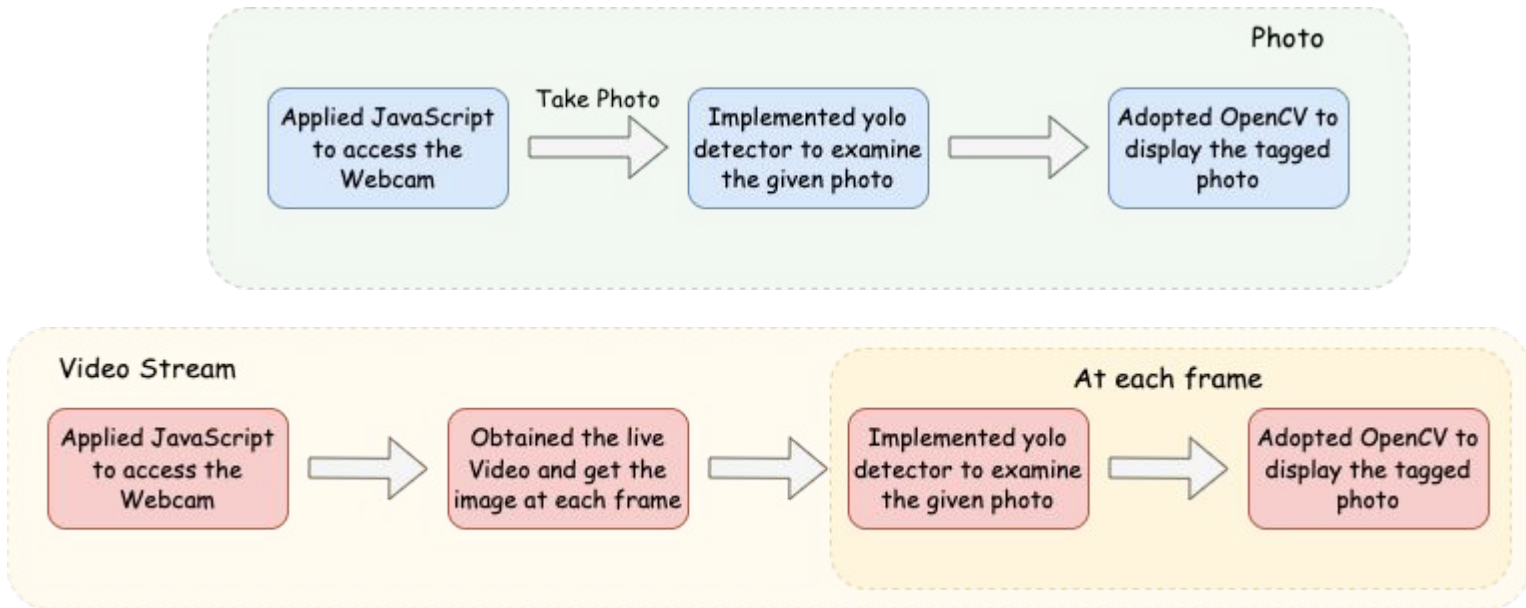
YOLO - You Only Look Once

- What is YOLO?
 - Another Object Detection Model
- Why choose YOLO?
 - Region-free (without traversing the image)
 - Higher Efficiency (compared to RCNN)
- Different Versions of YOLO
 - From v1 to v7, different backbone, loss function, input head

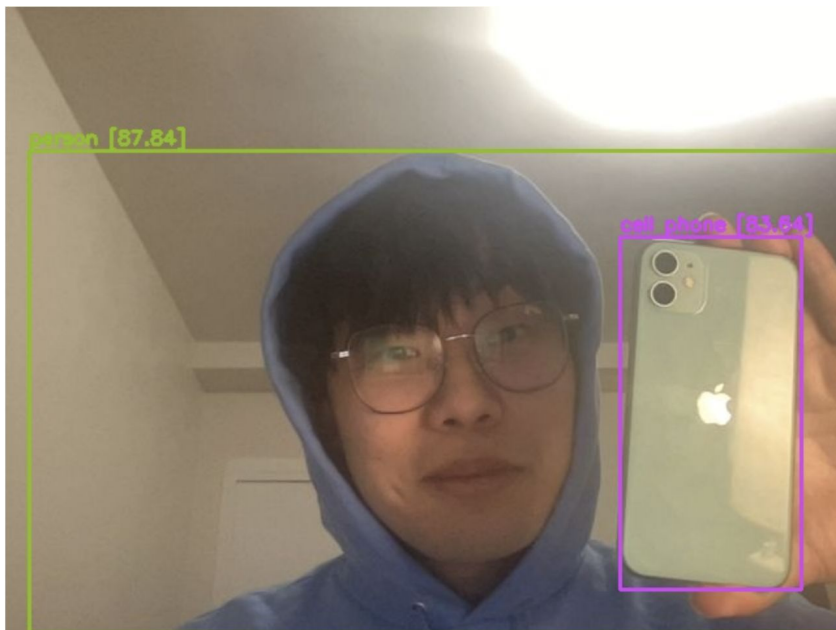


$$\begin{aligned} loss = & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\ & + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} \left[(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \right] \\ & + \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 \\ & + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 \\ & + \sum_{i=0}^{S^2} 1_i^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2 \end{aligned}$$

Implementing Process



Photo



Video Stream



TOY Application

Music Player

based on

Target Object Detection



colab

Target Object Detection

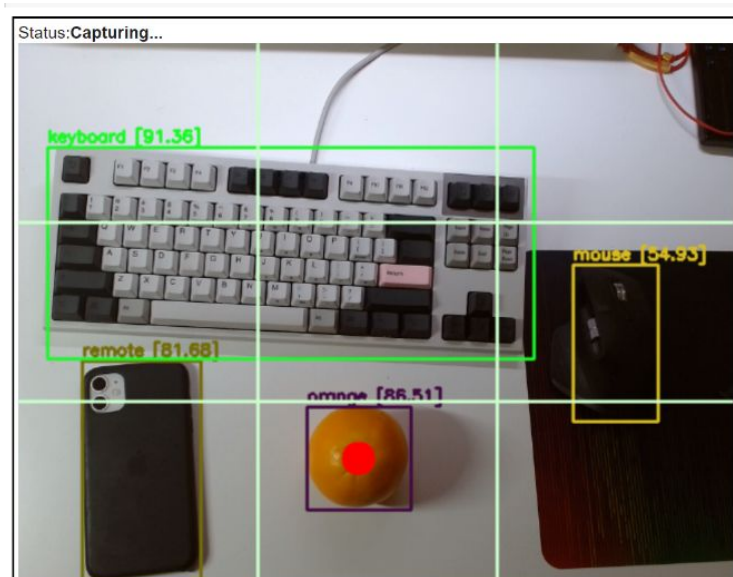


E

Sound Player



Orange Player





Live Example...

<https://colab.research.google.com/drive/1vl2rlRP6cNYxhE2eT5KVxIARk3LY3Ktq?usp=sharing>



Future Improvement

- YOLOv7 released
- Train with customized data
- Enable GPU with CUDA



Thanks For Watching~



Q & A