



You only look once v3

國立東華大學電機工程學系 楊哲旻 2020/05/27

Outline

1. Object Detection
2. Two stage Model : **Region-based Convolutional Neural Networks (RCNN)**
 - Region Proposal : Selective Search
3. One stage Model : **You only look once v3 (Yolo v3)**
 - Intersection over Union (IOU)
 - Non-Maximum Suppression (NMS)
4. Yolo v3 Demo :
 - Mask Detection
 - Face Detection

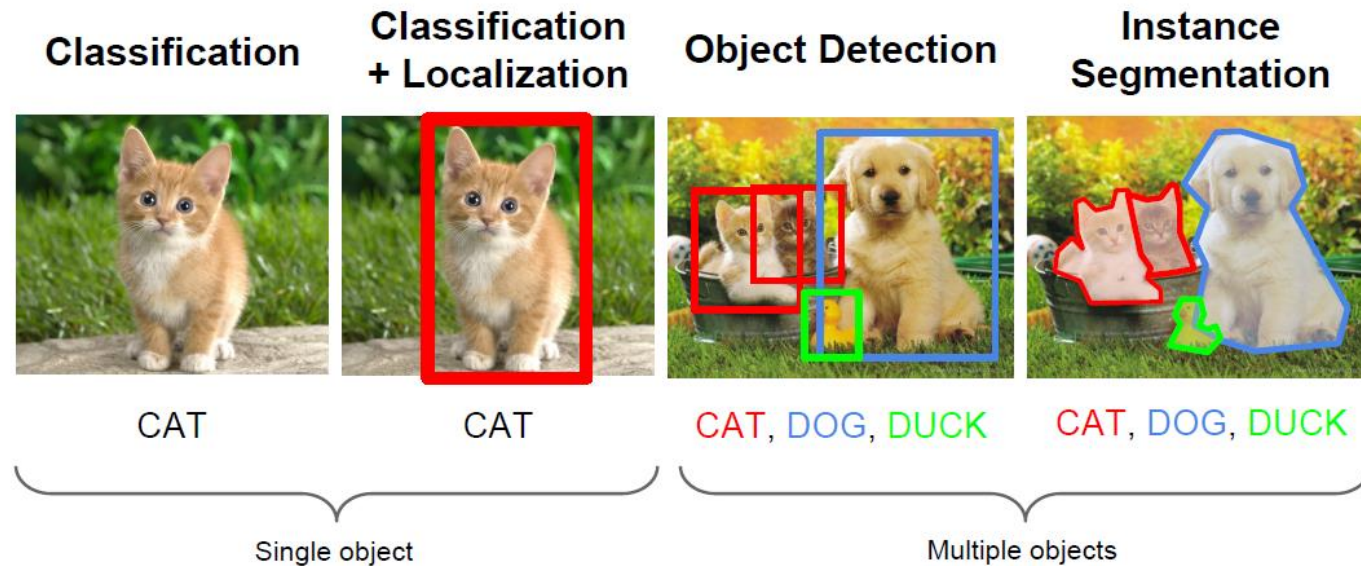
1. Object Detection

■ Two stage :

RCNN 、 Fes-RCNN 、 Faster RCNN

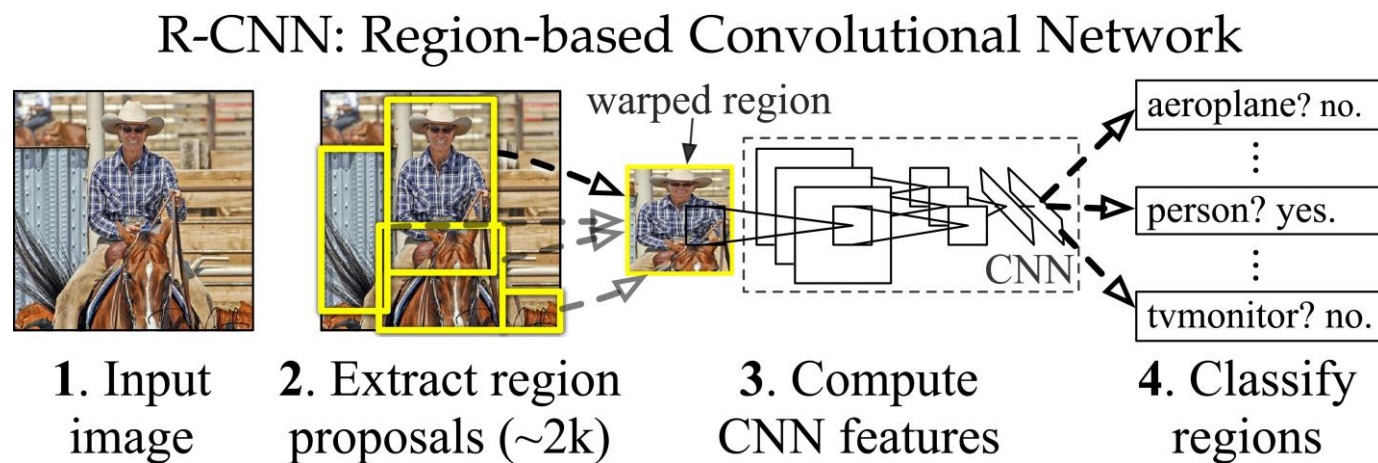
■ One stage (End-to-end training) :

SSD(Single Shot MultiBox Detector) 、 Yolo(You only look once) v1-4

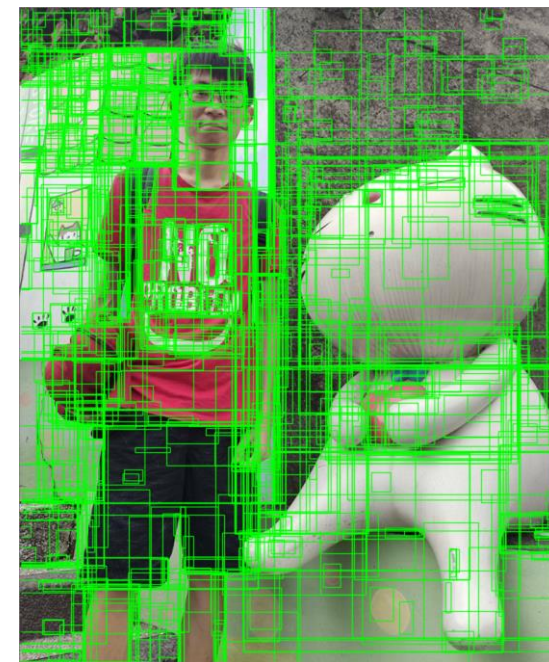
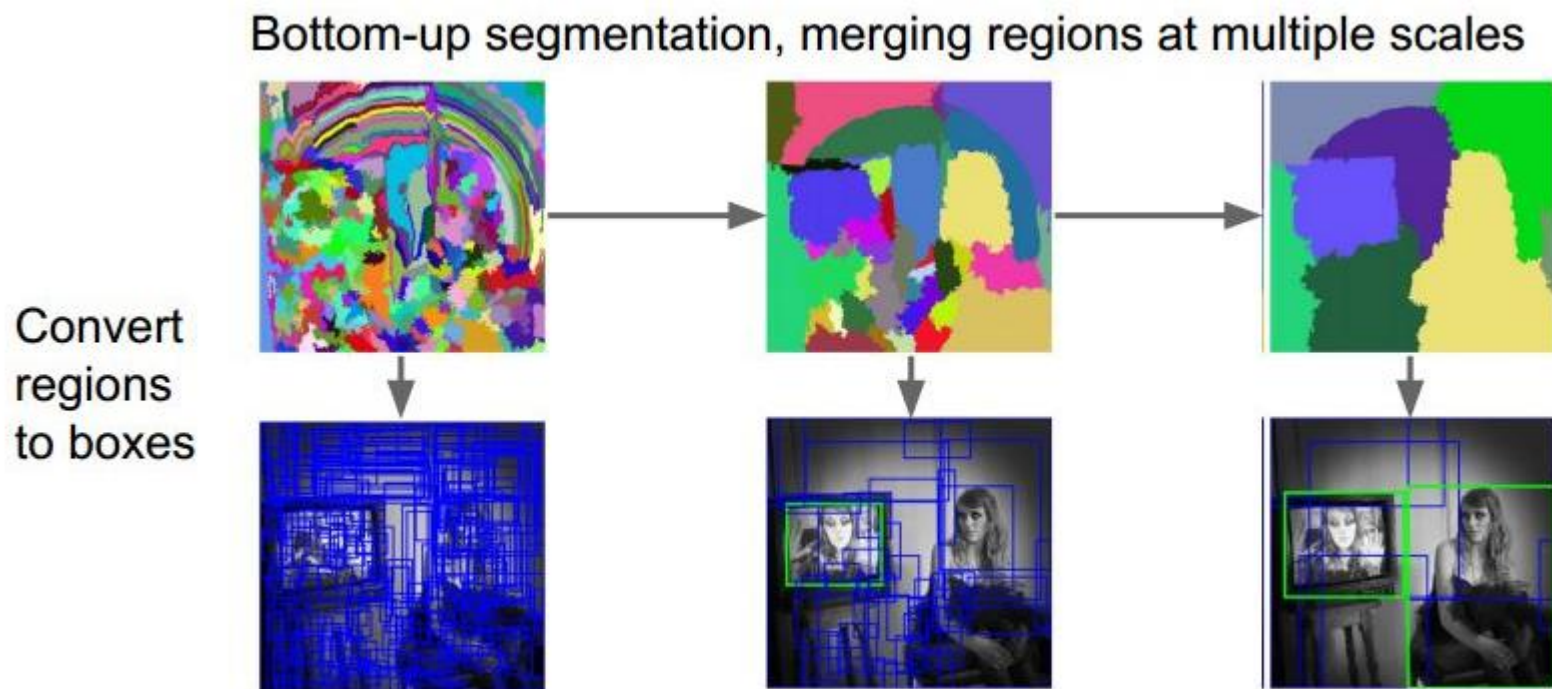


2. RCNN (Region-based Convolutional Neural Networks)

- 一、提取候選區域(Region Proposal)：選擇性搜索(Selective Search)
- 二、特徵提取(Feature Extraction)：卷積神經網路(Convolutional Neural Network, CNN)
- 三、分類(Classification)：SVM(Support Vector Machine)
- 四、邊框回歸(Bounding-box Regression)



2-1 Region Proposal : Selective Search

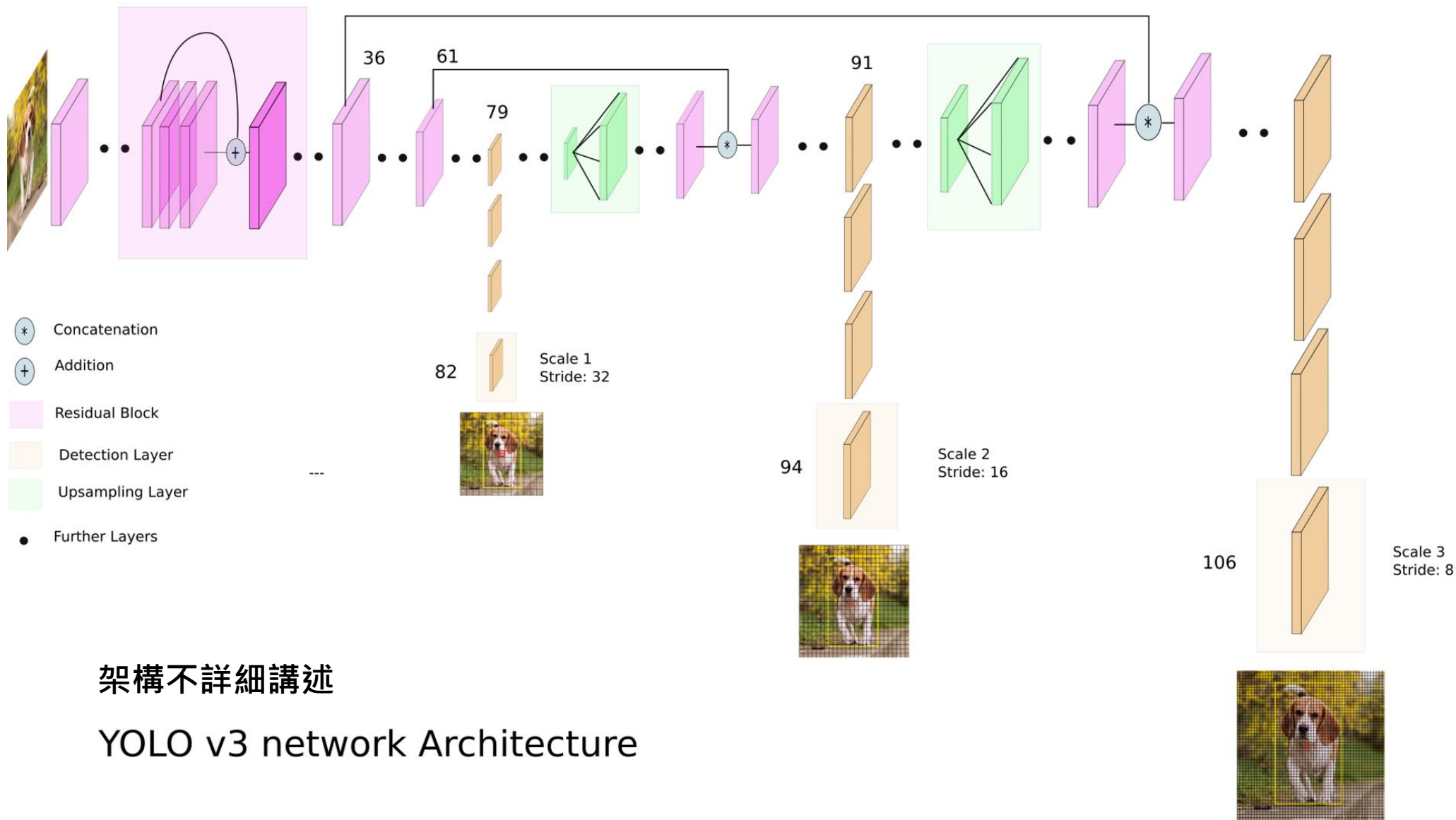


N=2000

選擇性搜索提取約2000個候選區域，再進行CNN特徵提取與SVM分類

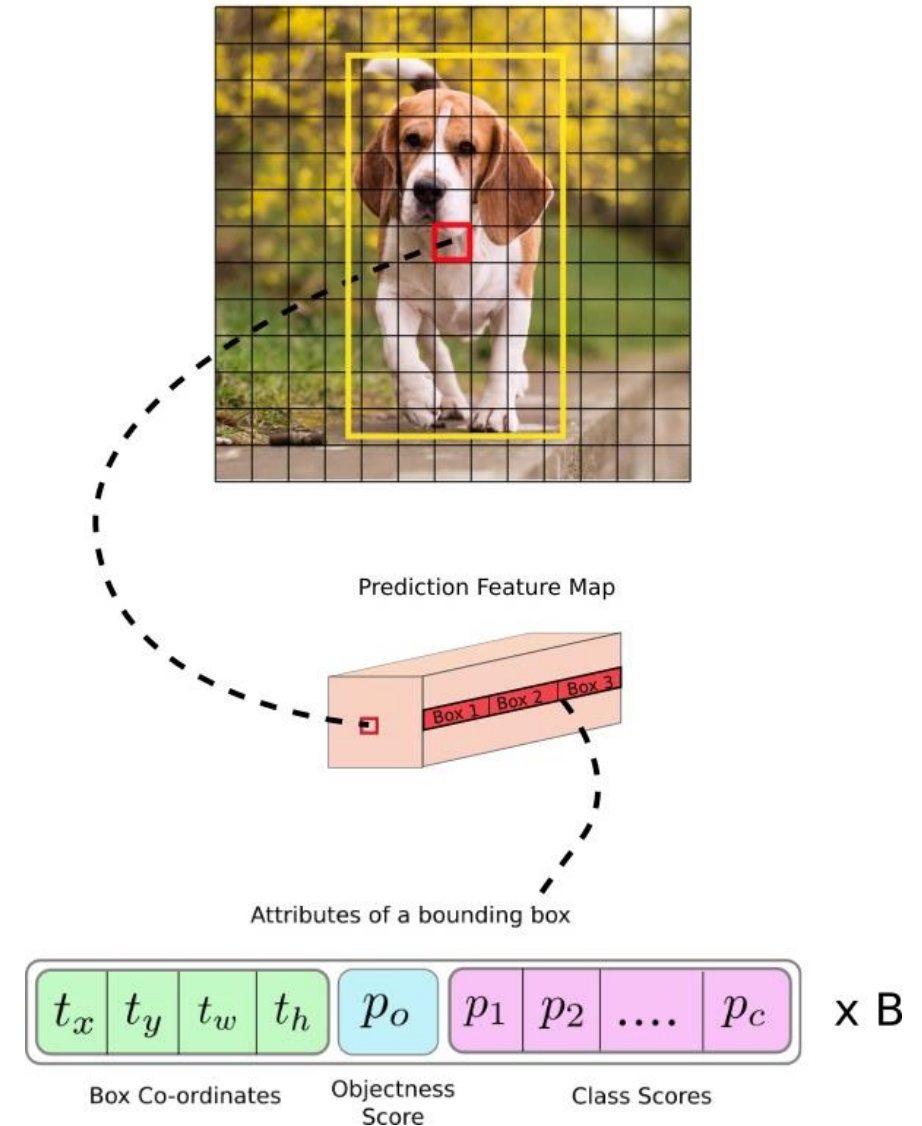
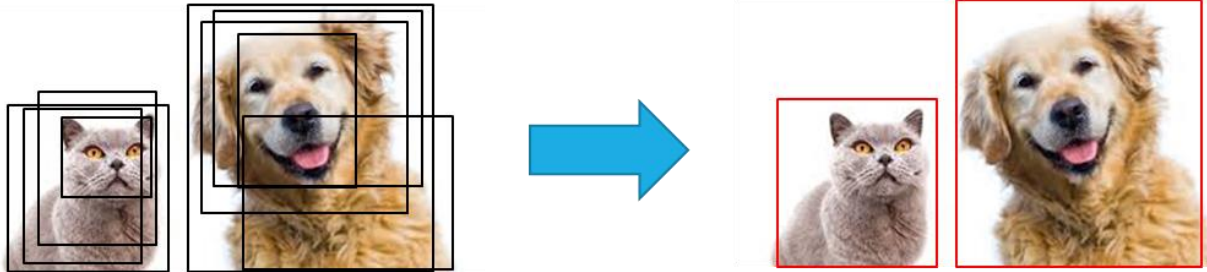
RCNN缺點：一張當前影像進行預測時，至少跑約2000 次的 CNN與SVM，非常耗時

3. You only look once v3 : Architecture




3. You only look once v3 : Output

1. 邊界框 (Bounding box, BBox) : x, y, w, h
2. 置信度 (Confidence Score)
3. 分類機率 (Classification Probability)



3-1. IOU (Intersection over Union)

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$


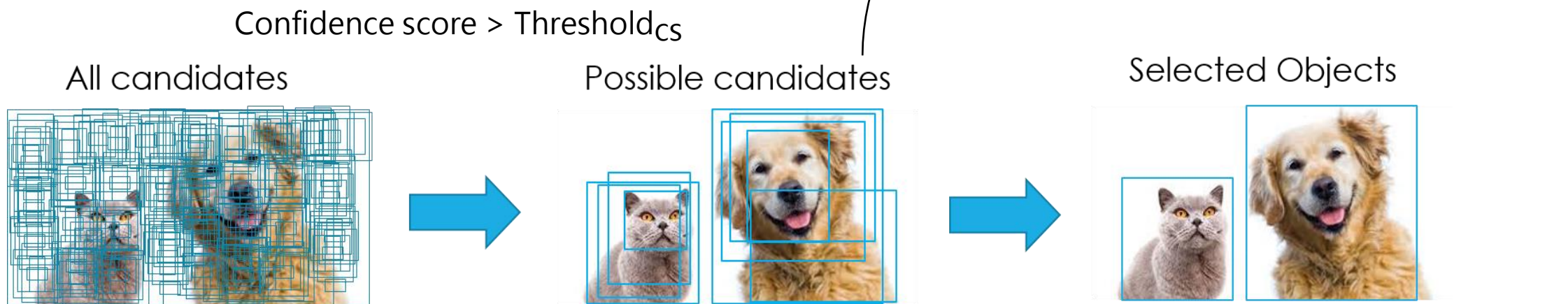


- 實際框 Ground Truth BBoxes
- 預測框 Anchor BBoxes (Detection Result)

3-2. NMS (Non-Maximum Suppression)

非最大抑制 (Non-Maximum Suppression, NMS) 方法
可以將同時包圍著單個物體的多個高置信度的邊界框，
消除到只剩下一個，即最佳邊界框。此方法需先設置置
信度閾值 (Threshold_{cs}) 與IoU閾值 (Threshold_{IoU})

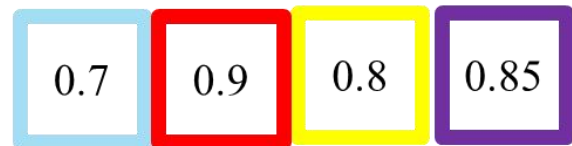
這裡IoU為實際框間的計算，並與實際框無關。



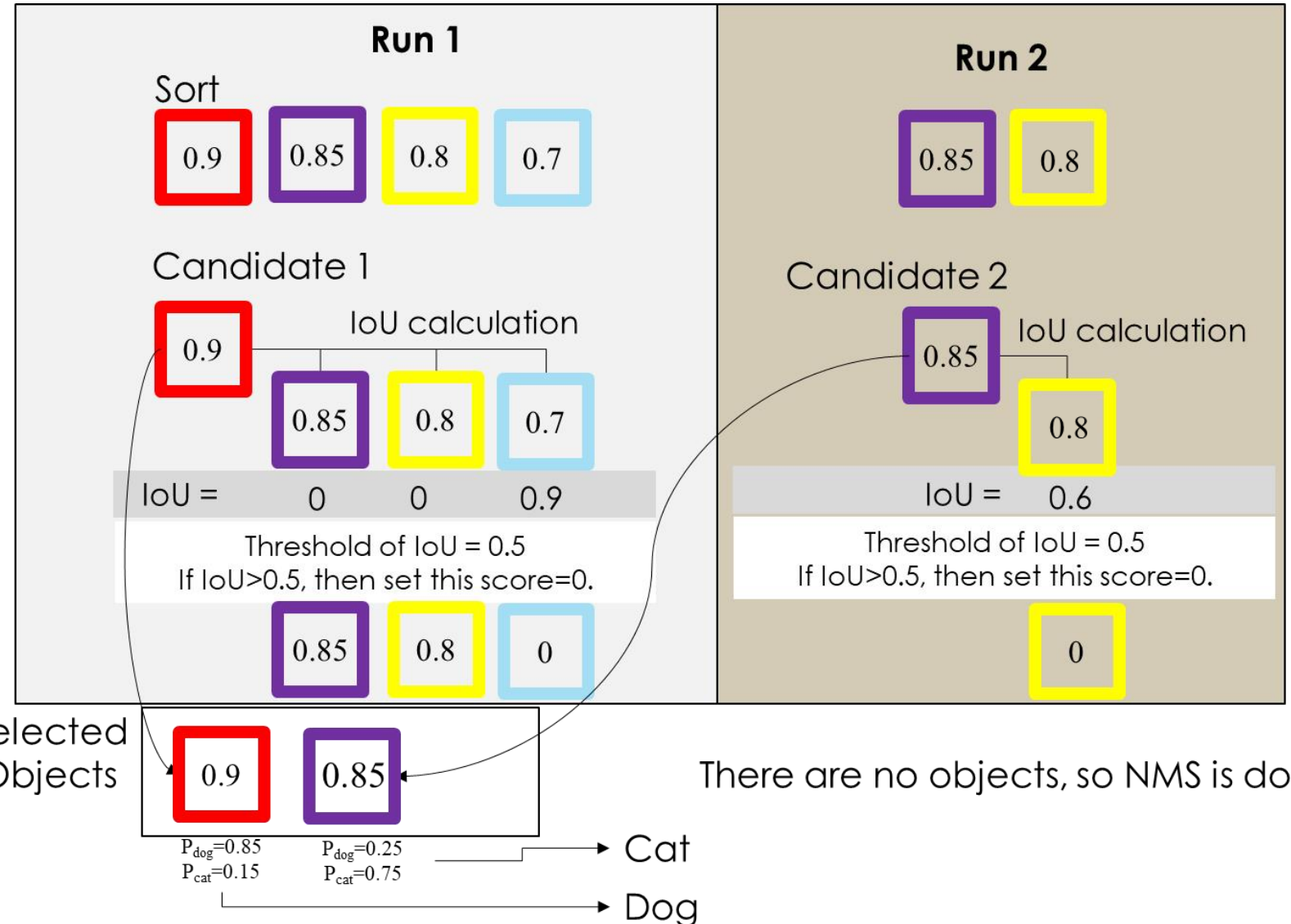
3-2. NMS (Non-Maximum Suppression)



Score for 4 candidates



$P_{\text{dog}}=0.9$ $P_{\text{dog}}=0.85$ $P_{\text{dog}}=0.01$ $P_{\text{dog}}=0.25$
 $P_{\text{cat}}=0.1$ $P_{\text{cat}}=0.15$ $P_{\text{cat}}=0.99$ $P_{\text{cat}}=0.75$



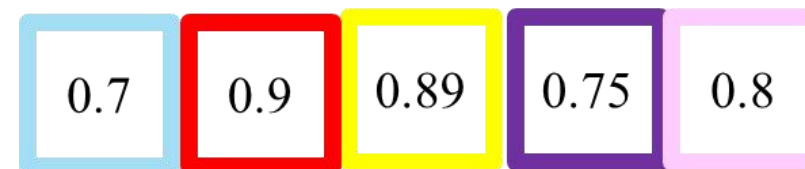
3-2. NMS (Non-Maximum Suppression)

設置置信度閾值 (Threshold_{cs}) 與IoU閾值 (Threshold_{IoU}) 會影響邊界框的數量：

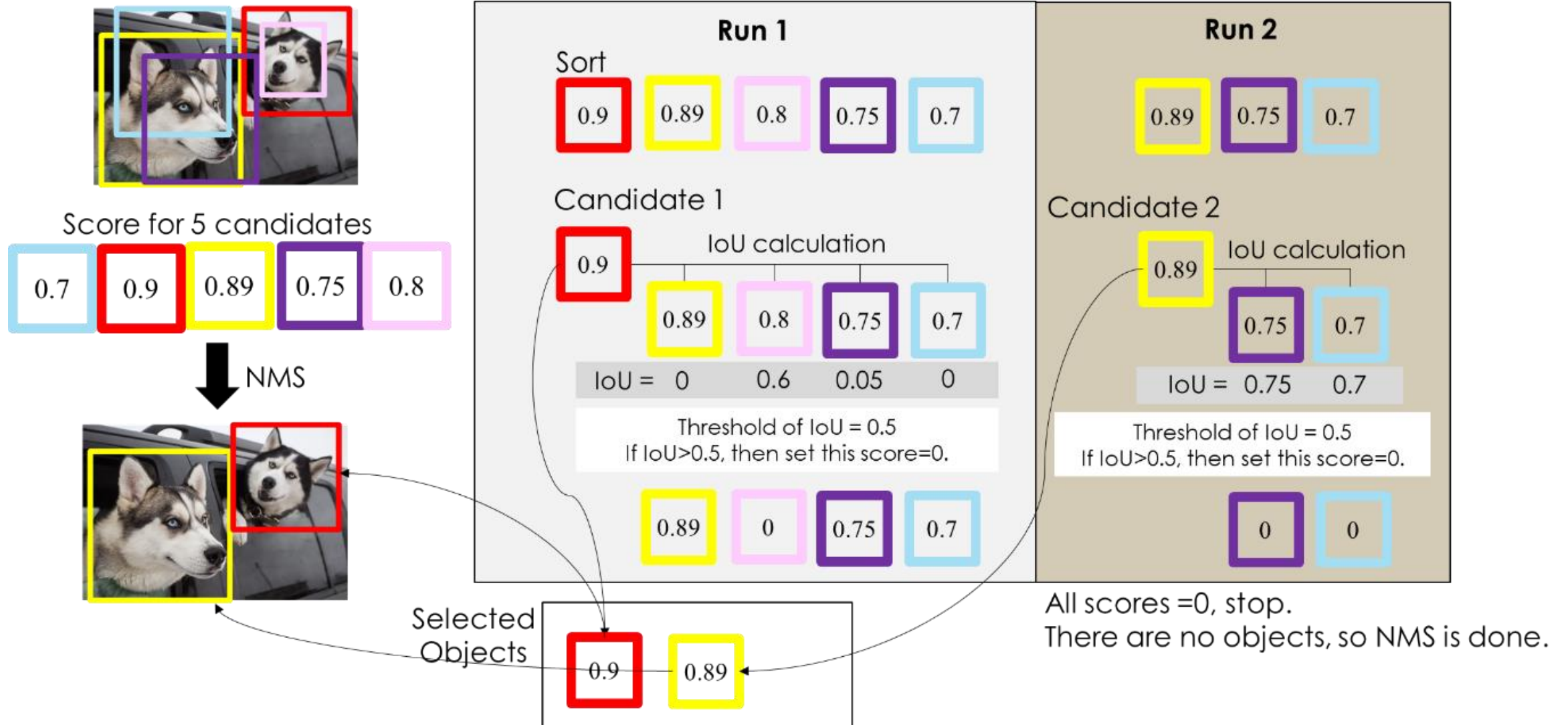
- 若需要嚴謹標準，以減少邊界框，則提高
置信度閾值或減少IoU閾值
- 若需要寬鬆標準，以增加邊界框，則減少
置信度閾值或增加IoU閾值



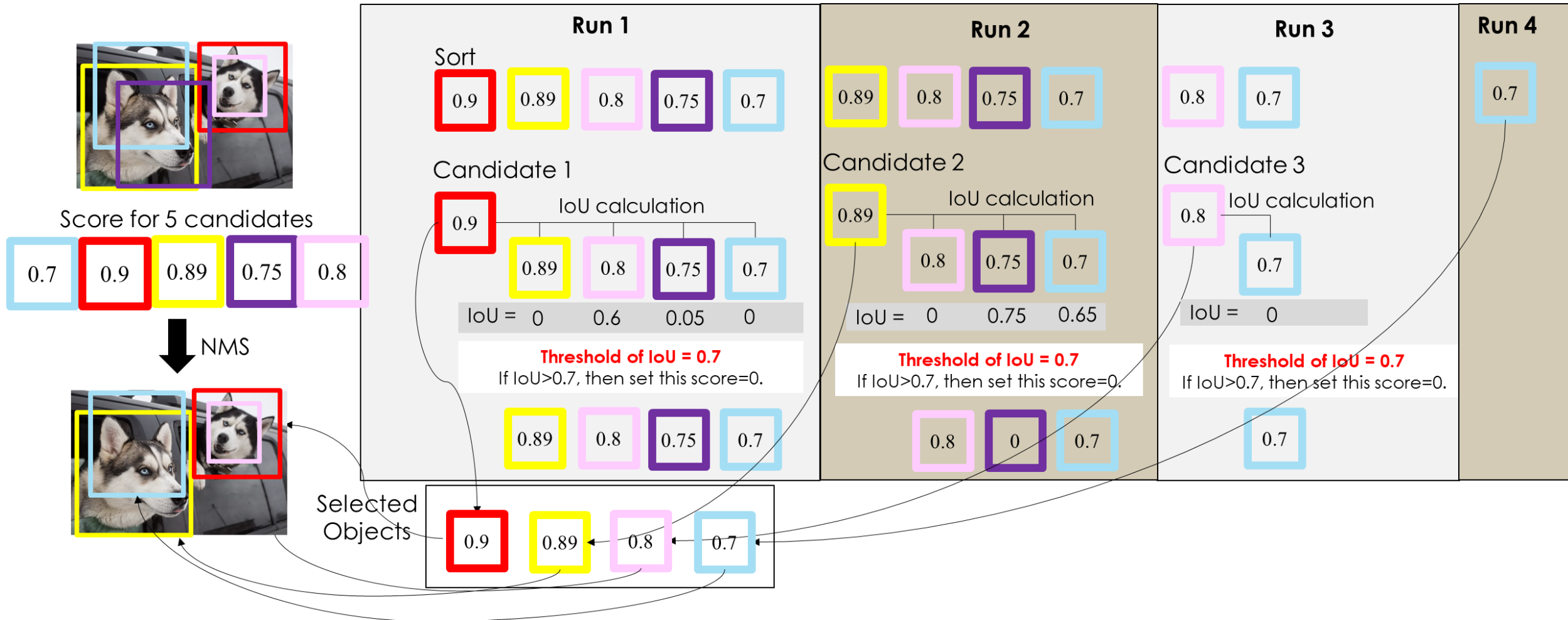
Score for 5 candidates



3-2. NMS (Non-Maximum Suppression)



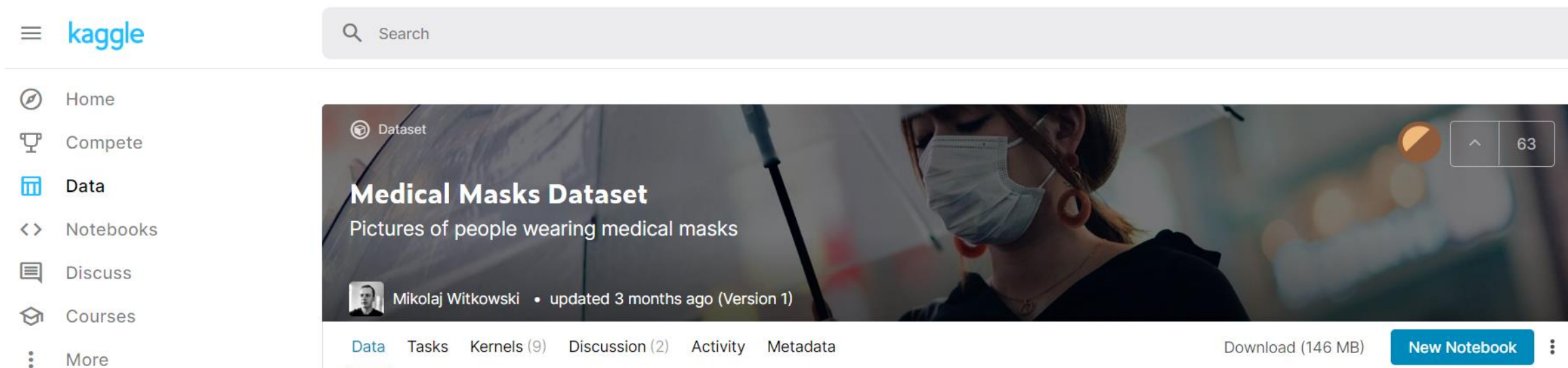
3-2. NMS (Non-Maximum Suppression)



4-1. Demo : Mask Detection by Yolo v3

Medical Masks Dataset

- 此數據集包含戴著醫用口罩者的JPG影像及詳細說明的XML文件。
- 有682張影像且超過3000張戴著口罩，以及大約700張錯誤或完全沒有佩戴口罩。XML文件包含其位置並標記為good、none或bad



4-1. Demo : Mask Detection by Yolo v3

Google Colab

NVIDIA Tesla P100 GPU

CUDA Version: 10.1

- yolov3.data
- yolov3.cfg
- yolov3.name

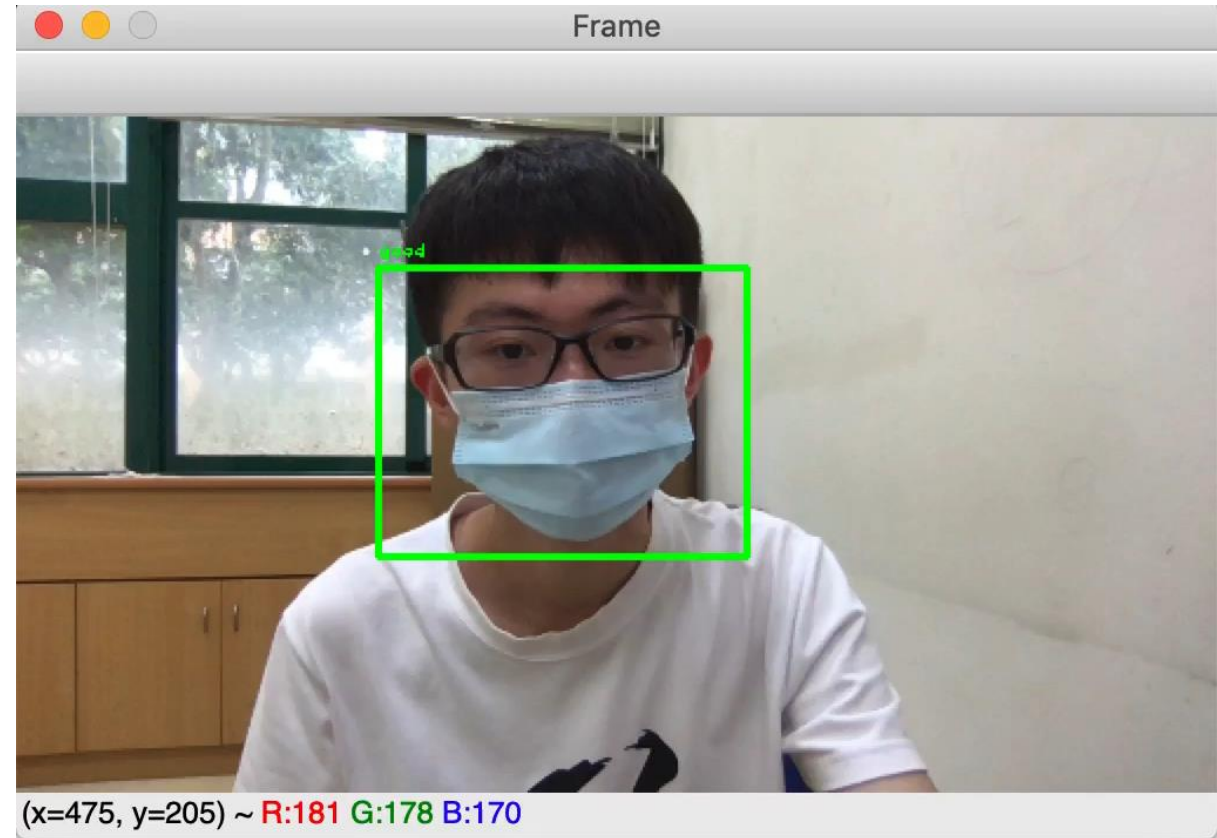
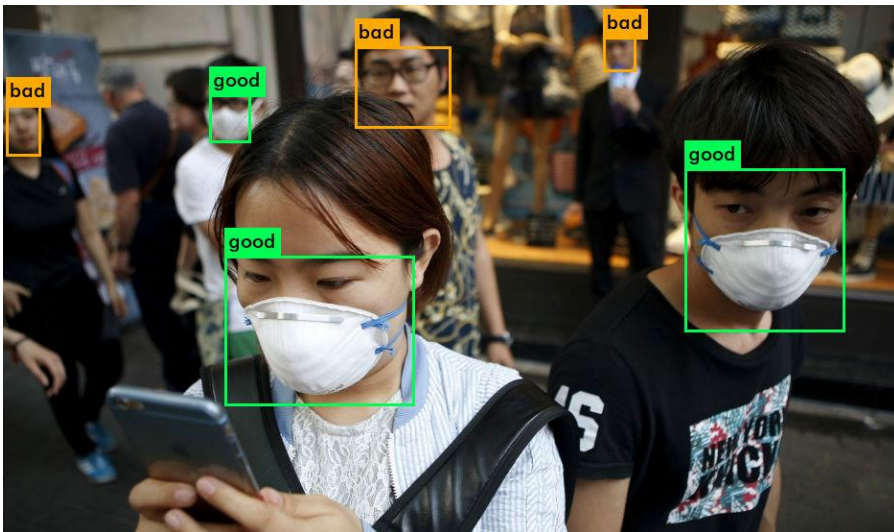
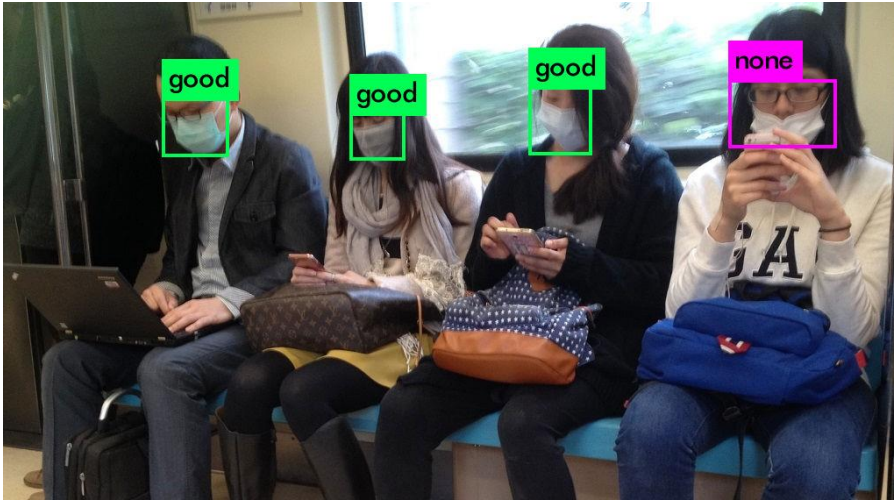
NMS

Threshold_{cs} = 0.5

Threshold_{IoU} = 0.3

```
1 batch = 64 # 批量數量
2 subdivisions = 2
3 width = 416 # 輸入寬度
4 height = 416 # 輸入高度
5 channels = 3 # 通道數(三維彩色)
6 momentum = 0.9 # 梯度下降法的動量超參數
7 decay = 0.0005 # 正則化超參數
8 angle = 0 # 數據增量：旋轉角度
9 saturation = 1.5 # 數據增量：飽和度
10 exposure = 1.5 # 數據增量：曝光量
11 hue=0.1 # 數據增量：色調
12 learning_rate = 0.001 # 梯度下降法的學習速率超參數
13 max_batches = 500200 # 訓練次數
```

4-1. Demo : Mask Detection by Yolo v3



4-2. Demo : Face Detection by Yolo v3

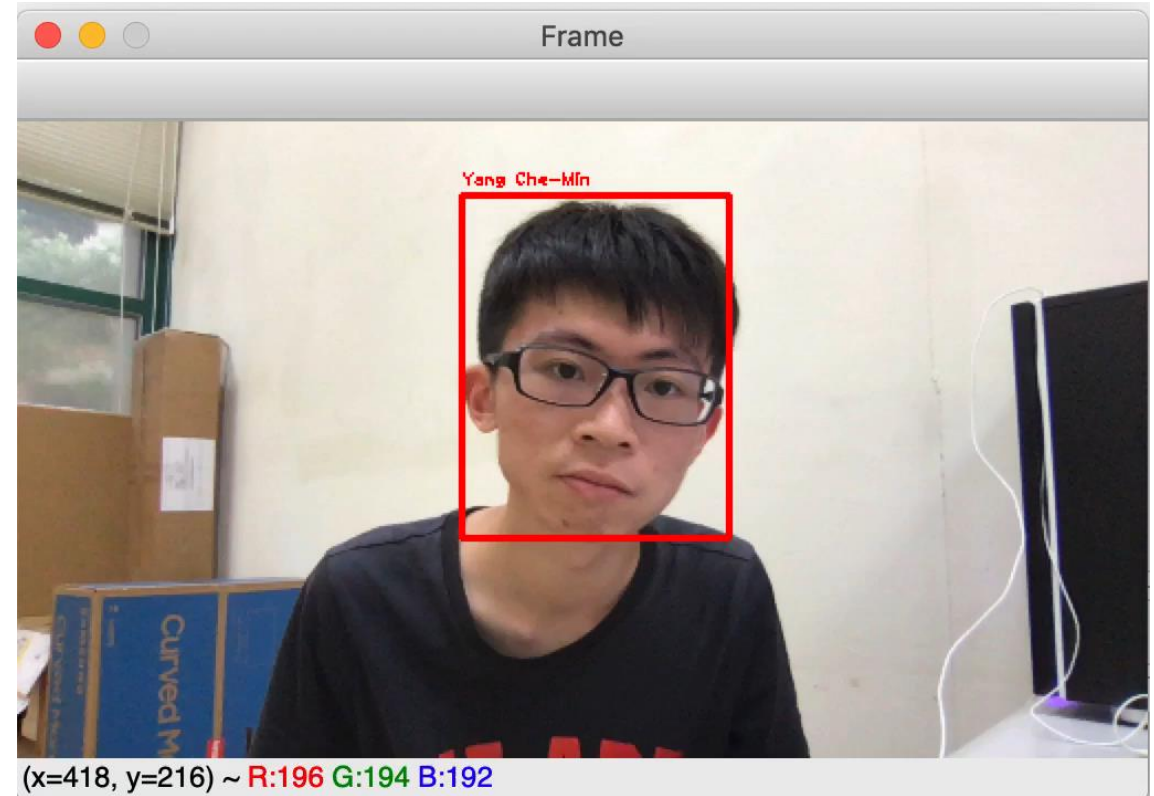
Face Dataset

- 自己蒐集的數據集包含人臉的JPG影像及詳細說明的XML文件。
- 有123張影像，人臉分別為楊哲旻(Yang Che-Min)、陳零九(Nine Chen)、邱鋒澤(Qiu Feng Ze)。XML文件包含其位置並標記為Yang Che-Min、Nine Chen或Qiu Feng Ze

NMS

$$\text{Threshold}_{cs} = 0.2$$

$$\text{Threshold}_{IoU} = 0.6$$



4-2. Demo : Face Detection by Yolo v3

LabelImg 影像標註工具

```
pip install labelImg
```

1. 開啟 LabelImg 之後，首先開啟想要進行標註的影像檔，開啟時可以選擇「Open」開啟單張影像檔，或是以「Open Dir」開啟整個目錄中所有的影像檔。「Change Save Dir」則可以改變xml與影像的存檔路徑。
2. 資料夾中labelImg-master\data\predefined_classes.txt 可更改類別標籤
3. 右鍵中點選「Create RectBox」可進行標註，以選擇預標籤的類別，並再點選「Save」儲存。可以點選「Next/Prev Image」進行上張或下張的影像作修改。

