

# 目標檢測模型

## Object Detection

國立東華大學電機工程學系 楊哲旻

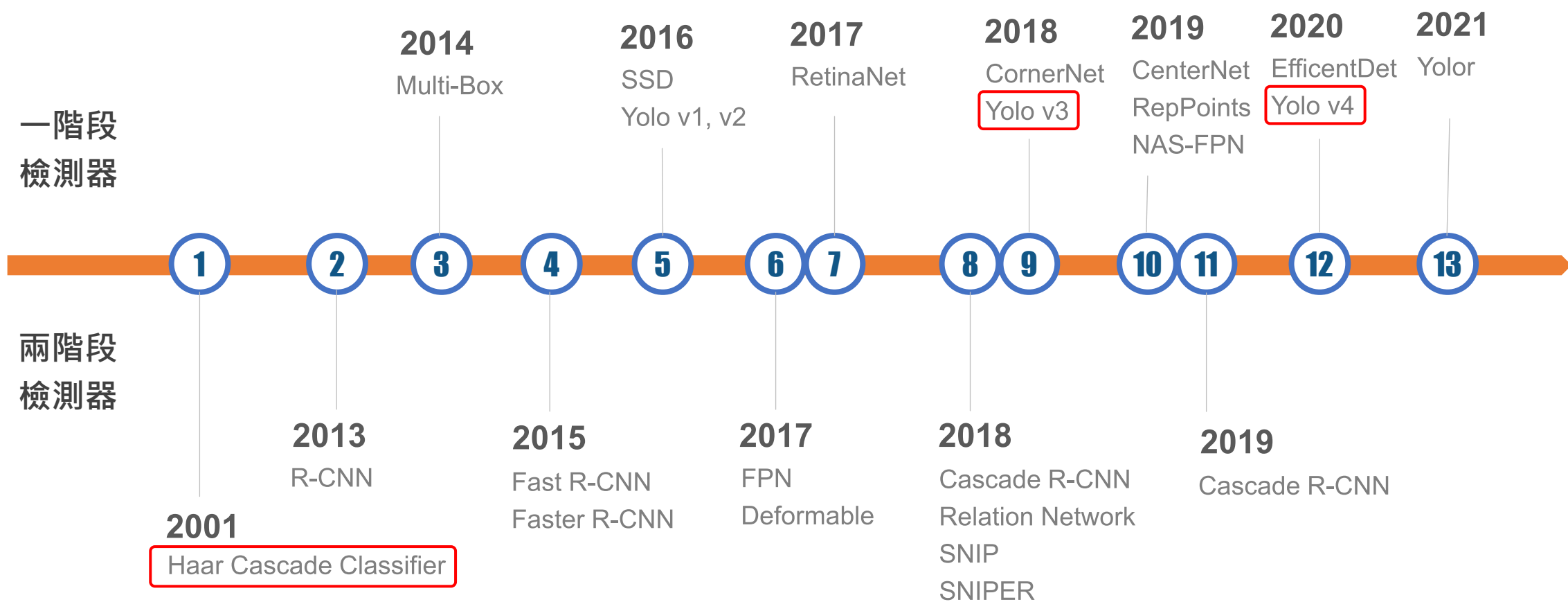
# Outline



- 1 技術發展
- 2 基於哈爾特徵級聯分類器
- 3 非最大抑制與並交比
- 4 平均精度
- 5 平均精度均值

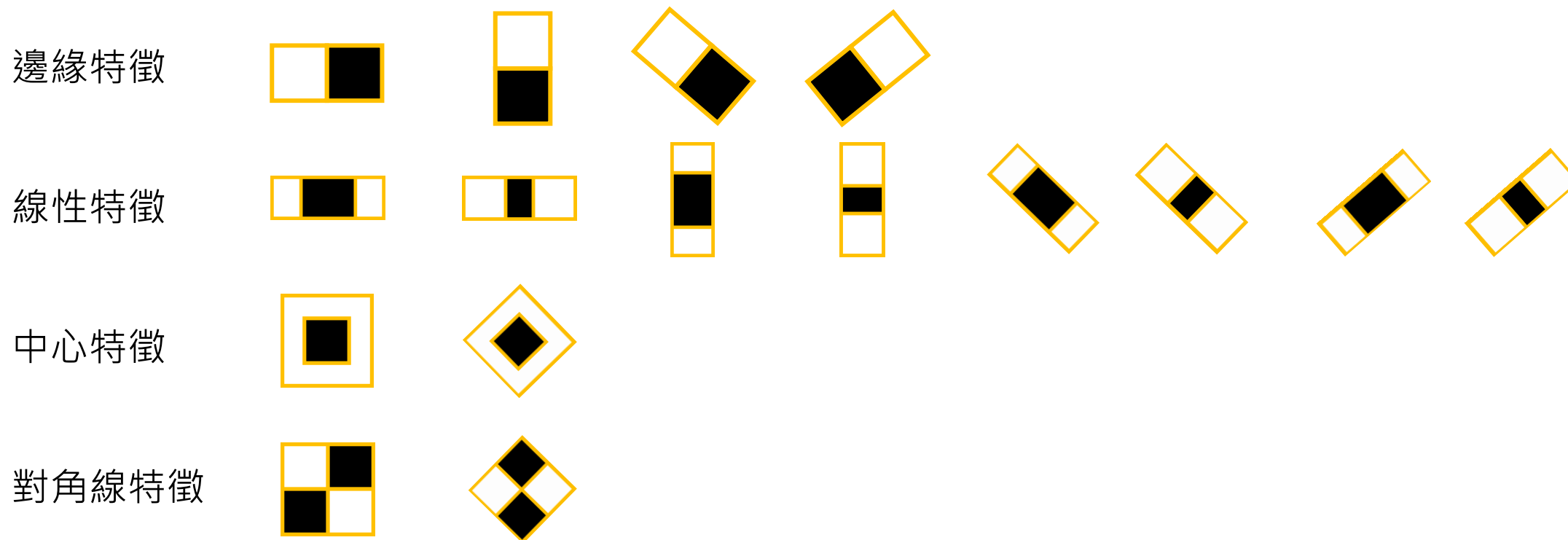


### 兩階段與一階段目標檢測模型 (Two and One Stage Object Detection Model)



### 哈爾特徵(Haar-like features)

哈爾特徵使用檢測視窗中的矩形，利用每特徵與矩形其差值計算積分圖，來對影像的子區域進行分類。哈爾特徵分為四類：**邊緣特徵**、**線性特徵**、**中心特徵**和**對角線特徵**，組合成特徵模板。



### 基於哈爾特徵級聯分類器(Haar Feature-based Cascade Classifiers)

基於哈爾特徵級聯分類器是由多種強分類器(Strong Classifier)所組成的級聯分類器(Cascade Classifier)，而每個強分類器由弱分類器(Weak classifier)，其分類器是透過多種特徵尋找將檢測圖像從左上至右下持續放大檢測。

基於哈爾特徵級聯分類器 (OpenCV)

常見可調的參數有：

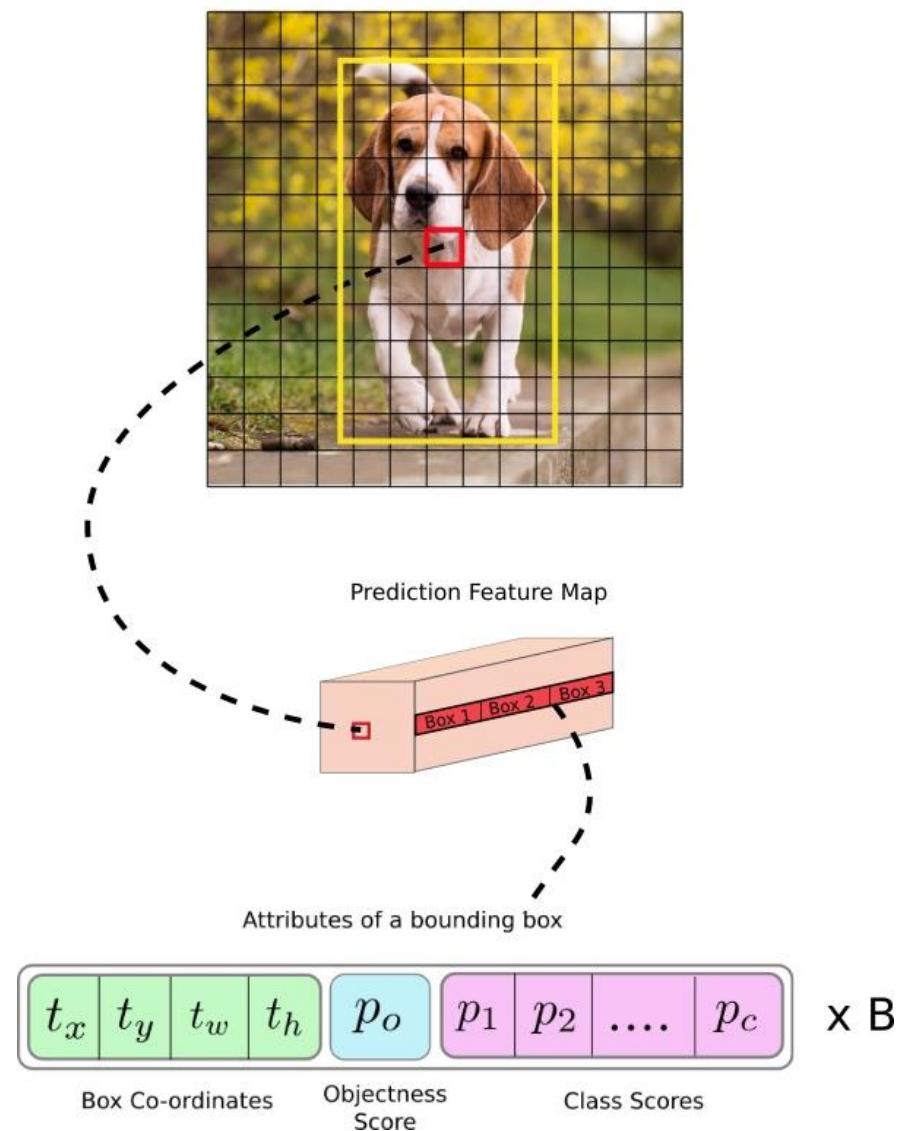
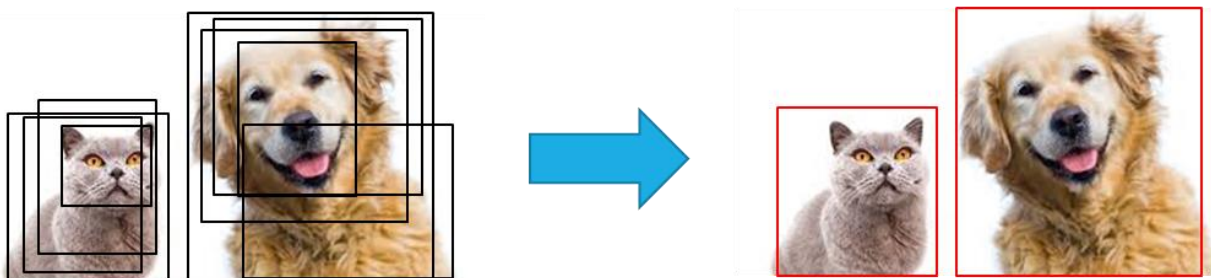
- minSize: 最小偵測區塊
- maxSize: 最大偵測區塊
- scaleFactor: 區塊大小放大倍率
- minNeighbors: 控制誤檢率參數




放大倍率繼續從開始端檢測

### You only look once, Yolo

Yolo是一個經典的一階段檢測器(One-stage) 的目標檢測模型，用一個卷積神經網絡結構就可以從輸入圖像直接預測邊界框(Bounding Box, Box Co-ordinates)、類別概率(Classification Probability, Class Scores)與置信度(Confidence, Objectness Score)。由於預測的邊界框過多，通常需藉由**非最大抑制**方法來移除多餘的重疊框。



並交比(Intersection over Union, IoU)

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




 實際框 Ground Truth BBoxes

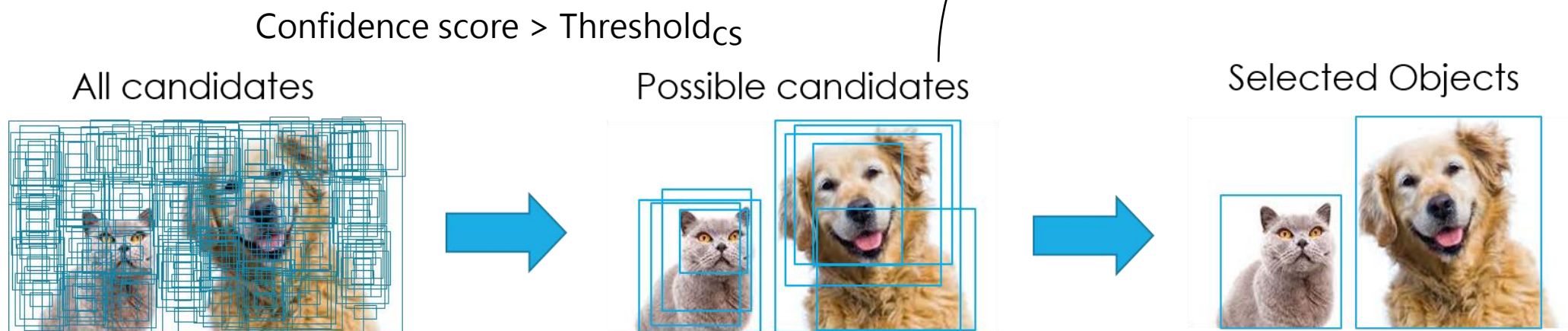
 預測框 Anchor BBoxes (Detection Result)



### 📄 非最大抑制 (Non-Maximum Suppression, NMS)

非最大抑制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 ( $\text{Threshold}_{cs}$ ) 與IoU閾值 ( $\text{Threshold}_{IoU}$ )

這裡IoU為實際框間的計算，並與實際框無關。



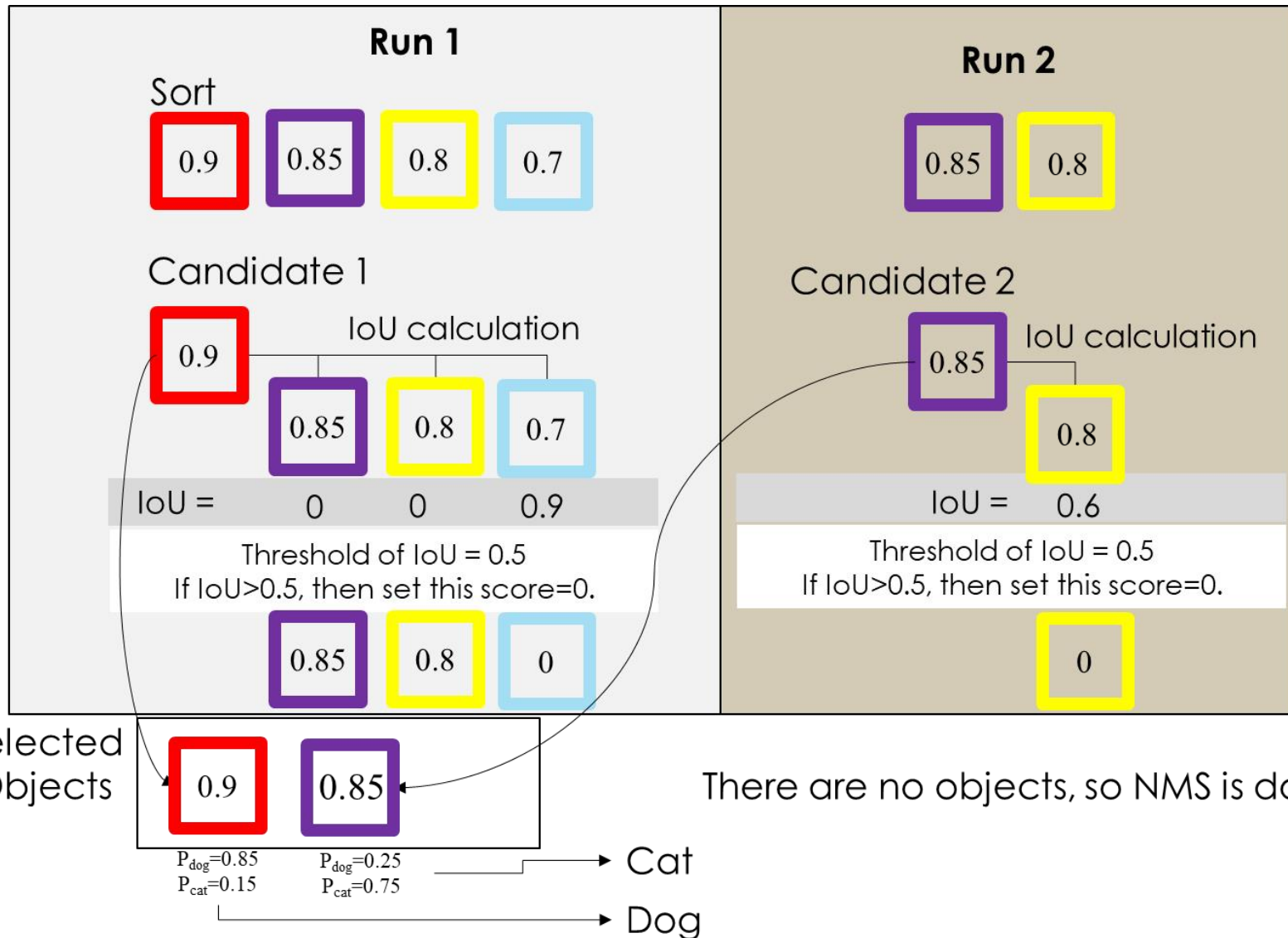




Score for 4 candidates



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



設置置信度閾值 ( $\text{Threshold}_{cs}$ ) 與IoU閾值 ( $\text{Threshold}_{IoU}$ ) 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



### 平均精度 (Average Precision, AP)

平均精度是目標檢測常用的指標之一，它由調整IoU閾值所繪製的PR曲線，並經由召回度的決斷值修正，其曲線下的面積即為平均精度

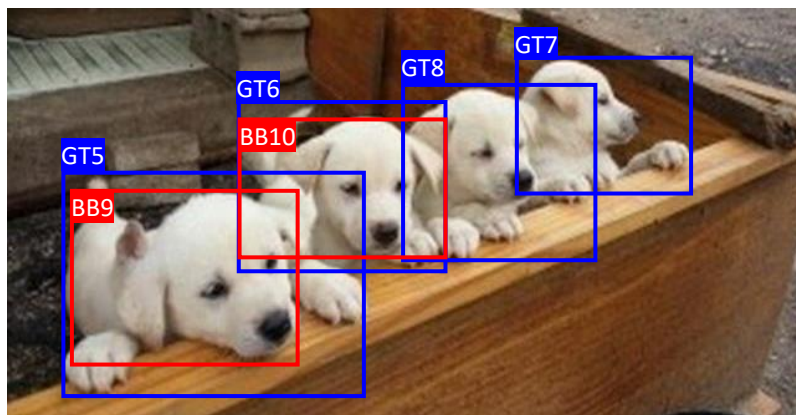
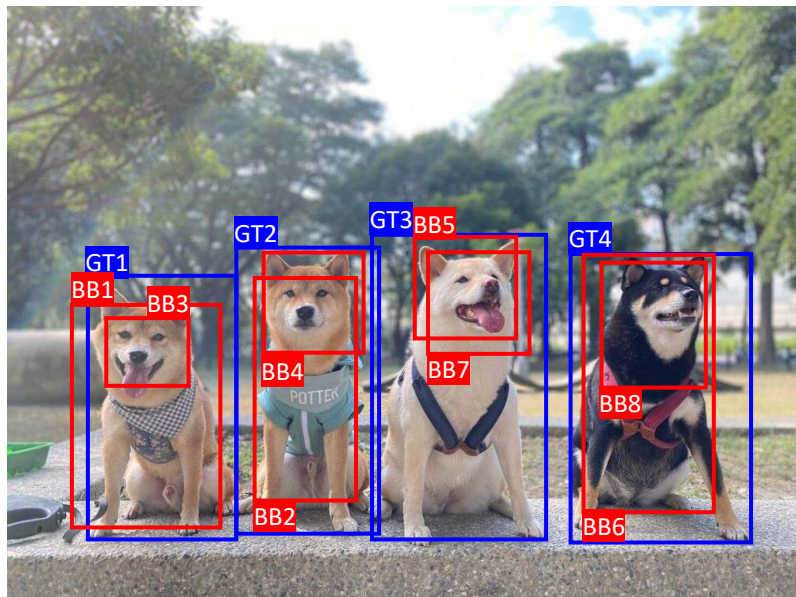
- 真陽性(True Positive, TP)：預測為有，實際上也有。
- 偽陽性(False Positive, FP)：預測為有，實際卻沒有。
- 真陰性(True Negative, TN)：預測為沒有，實際上也沒有。
- 偽陰性(False Negative, FN)：預測為沒有，實際卻有。

		Predicted Class	
		0	1
Actual Class	0	TN	FP
	1	FN	TP

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

1. 實際與預測皆為有的框，即IoU大於一個IoU閾值(通常是0.5) 的框之數量，則為TP
2. 由於TN是無限多個，即預測未被框到且實際也未被框到的數量，因此指標必須採用無TN的指標，所以目標檢測常見的指標除了IoU外，還有召回度(Recall)、精確度(Precision)以及這兩個指標所計算的平均精度



Ground Truth	Bounding Box	IoU	Confidence
GT1	BB1	0.8	0.9
GT2	BB2	0.7	0.9
GT1	BB3	0.1	0.2
GT2	BB4	0.2	0.8
GT3	BB5	0.1	0.2
GT4	BB6	0.7	0.7
GT3	BB7	0.1	0.3
GT4	BB8	0.2	0.4
GT5	BB9	0.7	0.7
GT6	BB10	0.8	0.9
GT7			
GT8			

■ 藍色為Ground Truth

■ 紅色為Bounding Box

有未預測到的，如GT7,8

或是框不完全導致IoU偏低，如BB3,4,5,7,8

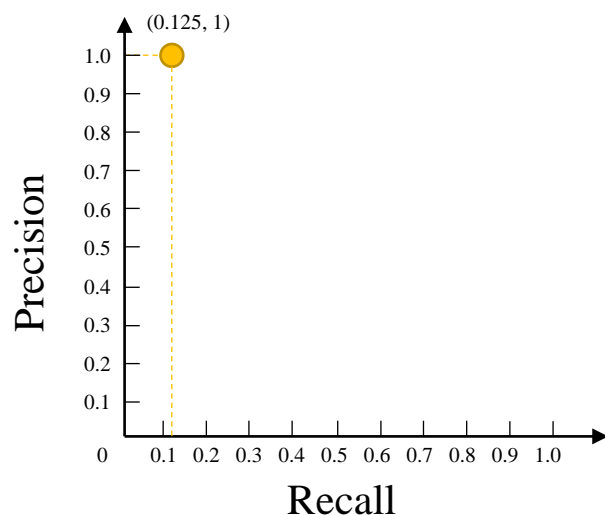
- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$  大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 1 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 0 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 7 ▲
- TP+FN = 全部共有八隻狗

✓ Precision =  $1/(1+0)=1$

✓ Recall =  $1/(7+1)=0.125$



GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	▲	▲	▲	▲	▲	▲	▲

Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence
GT1	BB1	0.8	1	1	0.9
GT2	BB2	0.7	2	1	0.9
GT6	BB10	0.8	3	1	0.9
GT2	BB4	0.2	4	0	0.8
GT4	BB6	0.7	5	1	0.7
GT5	BB9	0.7	6	1	0.7
GT4	BB8	0.2	7	0	0.4
GT3	BB7	0.1	8	0	0.3
GT1	BB3	0.1	9	0	0.2
GT3	BB5	0.1	10	0	0.2
GT7					
GT8					

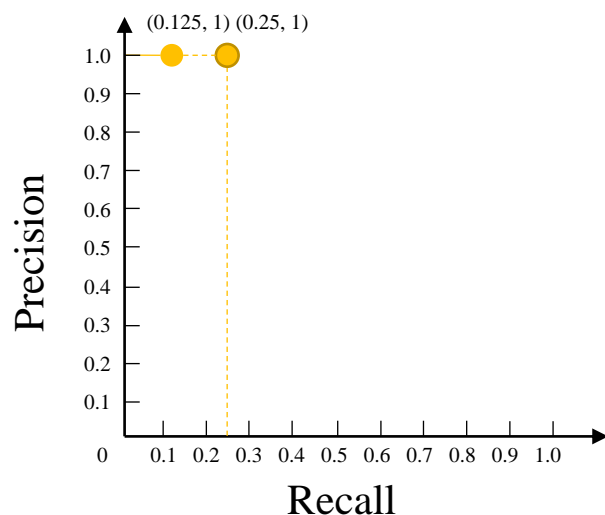
- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$  大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1-2:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 2 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 0 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 6 ▲
- TP+FN = 全部共有八隻狗

✓ Precision =  $2 / (2 + 0) = 1$

✓ Recall =  $2 / (6 + 2) = 0.25$



GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	✓	▲	▲	▲	▲	▲	▲

Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence
GT1	BB1	0.8	1	1	0.9
GT2	BB2	0.7	2	1	0.9
GT6	BB10	0.8	3	1	0.9
GT2	BB4	0.2	4	0	0.8
GT4	BB6	0.7	5	1	0.7
GT5	BB9	0.7	6	1	0.7
GT4	BB8	0.2	7	0	0.4
GT3	BB7	0.1	8	0	0.3
GT1	BB3	0.1	9	0	0.2
GT3	BB5	0.1	10	0	0.2
GT7					
GT8					

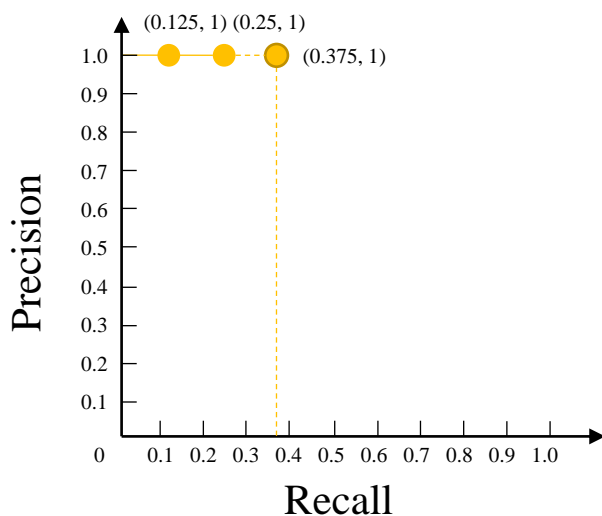


- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$  大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1-3:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 3 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 0 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 5 ▲
- TP+FN = 全部共有八隻狗

- ✓ Precision =  $3 / (3 + 0) = 1$
- ✓ Recall =  $3 / (5 + 3) = 0.375$



GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	✓	▲	▲	▲	✓	▲	▲

Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence
GT1	BB1	0.8	1	1	0.9
GT2	BB2	0.7	2	1	0.9
GT6	BB10	0.8	3	1	0.9
GT2	BB4	0.2	4	0	0.8
GT4	BB6	0.7	5	1	0.7
GT5	BB9	0.7	6	1	0.7
GT4	BB8	0.2	7	0	0.4
GT3	BB7	0.1	8	0	0.3
GT1	BB3	0.1	9	0	0.2
GT3	BB5	0.1	10	0	0.2
GT7					
GT8					



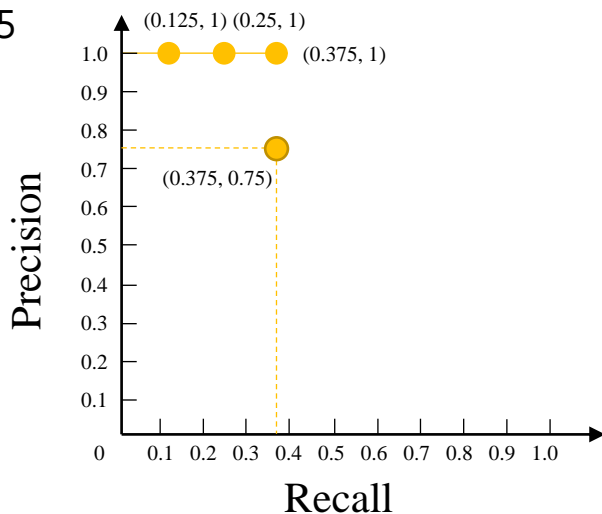
- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$  大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1-4:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 3 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 1 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 5 ▲
- TP+FN = 全部共有八隻狗

✓ Precision =  $3/(3+1)=0.75$

✓ Recall =  $3/(5+3)=0.375$



GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	✓	▲	▲	▲	✓	▲	▲

Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence	
GT1	BB1	0.8	1	1	0.9	✓
GT2	BB2	0.7	2	1	0.9	✓
GT6	BB10	0.8	3	1	0.9	✓
GT2	BB4	0.2	4	0	0.8	✗
GT4	BB6	0.7	5	1	0.7	
GT5	BB9	0.7	6	1	0.7	
GT4	BB8	0.2	7	0	0.4	
GT3	BB7	0.1	8	0	0.3	
GT1	BB3	0.1	9	0	0.2	
GT3	BB5	0.1	10	0	0.2	
GT7						
GT8						

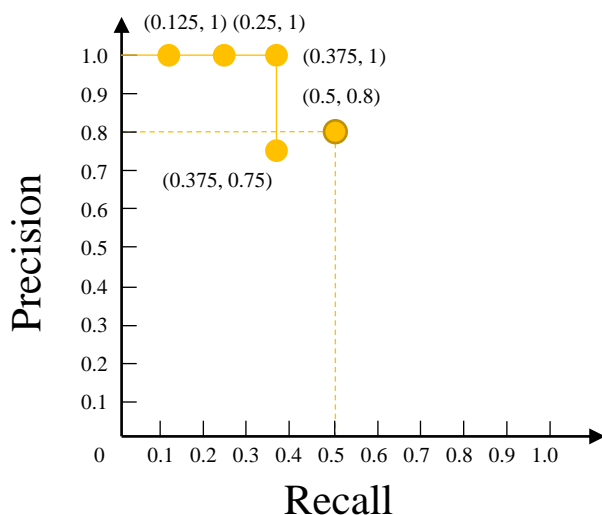
- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$  大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1-5:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 4 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 1 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 4 ▲
- TP+FN = 全部共有八隻狗

✓ Precision =  $4 / (4 + 1) = 0.8$

✓ Recall =  $4 / (4 + 4) = 0.5$



GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	✓	▲	✓	▲	✓	▲	▲

Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence	
GT1	BB1	0.8	1	1	0.9	✓
GT2	BB2	0.7	2	1	0.9	✓
GT6	BB10	0.8	3	1	0.9	✓
GT2	BB4	0.2	4	0	0.8	✗
GT4	BB6	0.7	5	1	0.7	✓
GT5	BB9	0.7	6	1	0.7	
GT4	BB8	0.2	7	0	0.4	
GT3	BB7	0.1	8	0	0.3	
GT1	BB3	0.1	9	0	0.2	
GT3	BB5	0.1	10	0	0.2	
GT7						
GT8						

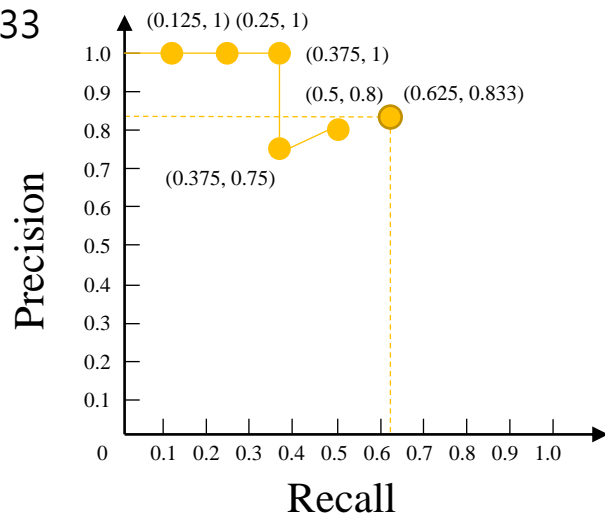
- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$  大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1-6:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 5 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 1 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 3 ▲
- TP+FN = 全部共有八隻狗

✓ Precision =  $5 / (5 + 1) = 0.833$

✓ Recall =  $5 / (5 + 3) = 0.625$



GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	✓	▲	✓	✓	✓	▲	▲

Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence	
GT1	BB1	0.8	1	1	0.9	✓
GT2	BB2	0.7	2	1	0.9	✓
GT6	BB10	0.8	3	1	0.9	✓
GT2	BB4	0.2	4	0	0.8	✗
GT4	BB6	0.7	5	1	0.7	✓
GT5	BB9	0.7	6	1	0.7	✓
GT4	BB8	0.2	7	0	0.4	
GT3	BB7	0.1	8	0	0.3	
GT1	BB3	0.1	9	0	0.2	
GT3	BB5	0.1	10	0	0.2	
GT7						
GT8						

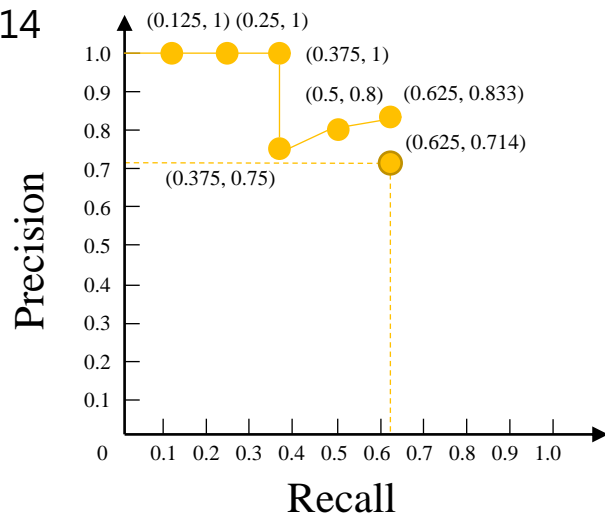
- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$  大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1-7:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 5 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 2 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 3 ▲
- TP+FN = 全部共有八隻狗

✓ Precision =  $5 / (5 + 2) = 0.714$

✓ Recall =  $5 / (5 + 3) = 0.625$



GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	✓	▲	✓	✓	✓	▲	▲

Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence	
GT1	BB1	0.8	1	1	0.9	✓
GT2	BB2	0.7	2	1	0.9	✓
GT6	BB10	0.8	3	1	0.9	✓
GT2	BB4	0.2	4	0	0.8	✗
GT4	BB6	0.7	5	1	0.7	✓
GT5	BB9	0.7	6	1	0.7	✓
GT4	BB8	0.2	7	0	0.4	✗
GT3	BB7	0.1	8	0	0.3	
GT1	BB3	0.1	9	0	0.2	
GT3	BB5	0.1	10	0	0.2	
GT7						
GT8						

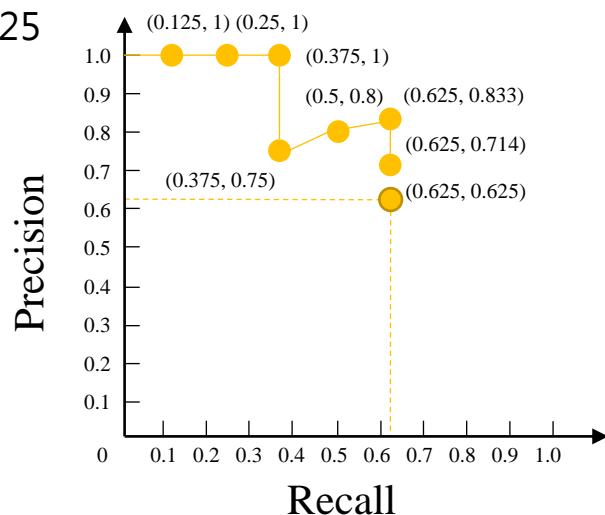
- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$ 大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1-8:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 5 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 3 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 3 ▲
- TP+FN = 全部共有八隻狗

✓ Precision =  $5 / (5 + 3) = 0.625$

✓ Recall =  $5 / (5 + 3) = 0.625$



GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	✓	▲	✓	✓	✓	▲	▲

Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence	
GT1	BB1	0.8	1	1	0.9	✓
GT2	BB2	0.7	2	1	0.9	✓
GT6	BB10	0.8	3	1	0.9	✓
GT2	BB4	0.2	4	0	0.8	✗
GT4	BB6	0.7	5	1	0.7	✓
GT5	BB9	0.7	6	1	0.7	✓
GT4	BB8	0.2	7	0	0.4	✗
GT3	BB7	0.1	8	0	0.3	✗
GT1	BB3	0.1	9	0	0.2	
GT3	BB5	0.1	10	0	0.2	
GT7						
GT8						

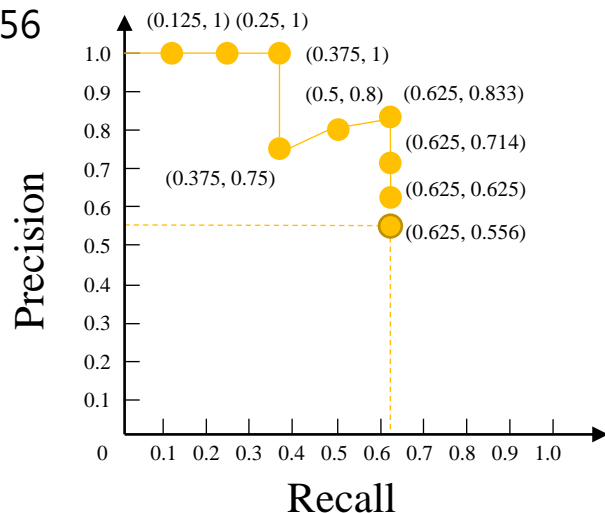
- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$ 大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1-9:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 5 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 4 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 3 ▲
- TP+FN = 全部共有八隻狗

✓ Precision =  $5 / (5 + 4) = 0.556$

✓ Recall =  $5 / (5 + 3) = 0.625$



GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	✓	▲	✓	✓	✓	▲	▲

Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence	
GT1	BB1	0.8	1	1	0.9	✓
GT2	BB2	0.7	2	1	0.9	✓
GT6	BB10	0.8	3	1	0.9	✓
GT2	BB4	0.2	4	0	0.8	✗
GT4	BB6	0.7	5	1	0.7	✓
GT5	BB9	0.7	6	1	0.7	✓
GT4	BB8	0.2	7	0	0.4	✗
GT3	BB7	0.1	8	0	0.3	✗
GT1	BB3	0.1	9	0	0.2	✗
GT3	BB5	0.1	10	0	0.2	
GT7						
GT8						

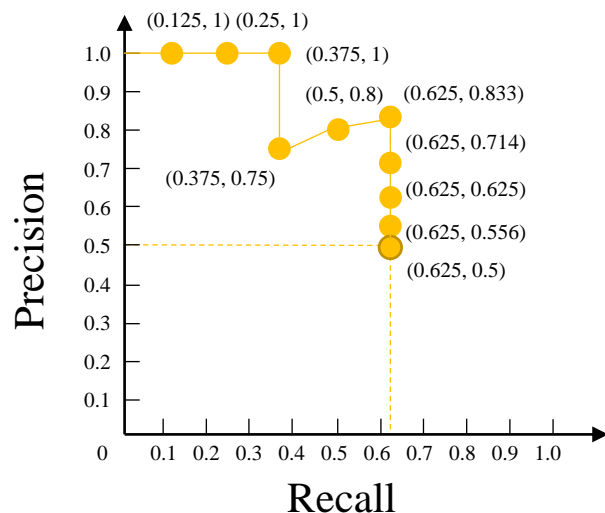
- IoU閾值( $Th_{IoU}$ )是0.5，新增 $IoU > 0.5$ 欄
- 預測框進行排序，置信度大至小，其次 $IoU > 0.5$  大至小，並標註為Rank 1至n
- 根據不同Rank累計計算Precision與Recall

Rank 1-9:

- TP = 實際有狗 被框出且被判斷( $IoU > 0.5$ )為狗的個數 = 5 ✓
- FP = 實際有狗 被框出但未判斷( $IoU < 0.5$ )為狗的個數 = 5 ✗
- FN = 實際有狗 卻 未被框出與有被框出未判斷為狗的個數 = 3 ▲
- TP+FN = 全部共有八隻狗

✓ Precision =  $5 / (5 + 5) = 0.5$

✓ Recall =  $5 / (5 + 3) = 0.625$

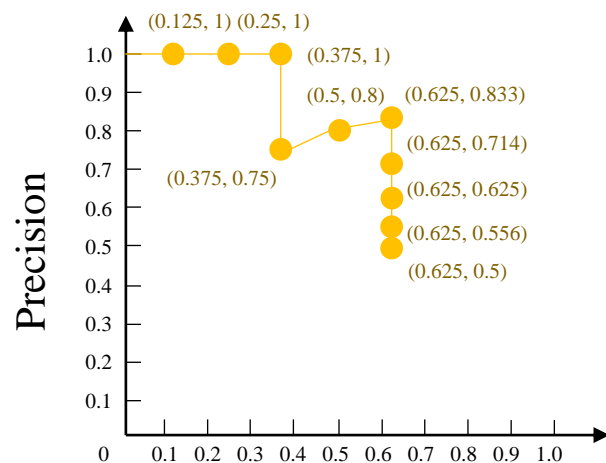


GT1	GT2	GT3	GT4	GT5	GT6	GT7	GT8
✓	✓	▲	✓	✓	✓	▲	▲

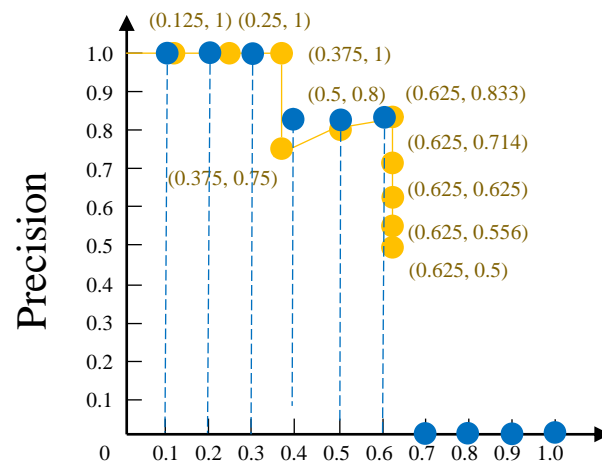
Ground Truth	Bounding Box	IoU	Rank	IoU>0.5	Confidence	
GT1	BB1	0.8	1	1	0.9	✓
GT2	BB2	0.7	2	1	0.9	✓
GT6	BB10	0.8	3	1	0.9	✓
GT2	BB4	0.2	4	0	0.8	✗
GT4	BB6	0.7	5	1	0.7	✓
GT5	BB9	0.7	6	1	0.7	✓
GT4	BB8	0.2	7	0	0.4	✗
GT3	BB7	0.1	8	0	0.3	✗
GT1	BB3	0.1	9	0	0.2	✗
GT3	BB5	0.1	10	0	0.2	✗
GT7						
GT8						



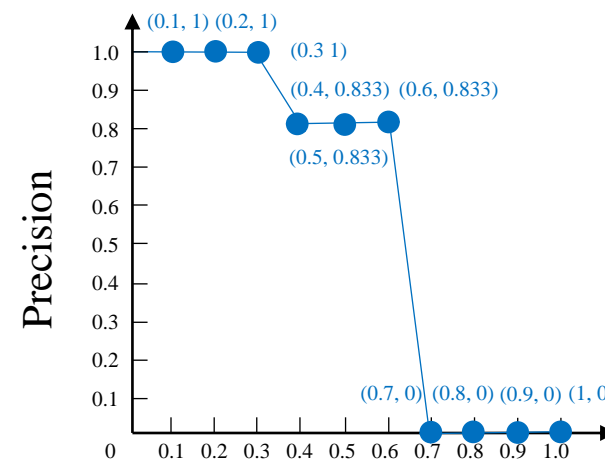
1. IoU閾值( $TH_{IoU}$ )是0.5，新增IoU>0.5欄
2. 預測框進行排序，置信度大至小，其次IoU>0.5 大至小，並標註為Rank 1至n
3. 根據不同Rank累計計算Precision與Recall
4. 召回度(Recall)的決斷值修正：選取Recall=0, 0.1, 0.2,..., 1，共11個地方取大於等於Recall中Precision最大值為新的Precision



修正前



修正後



修正後

5. 計算修正後的Precision-Recall curve下的面積，即為平均精度 $AP = \frac{1}{11} (1 + 1 + 1 + 1 + 0.833 + 0.833 + 0.833 + 0 + 0 + 0 + 0)$

### 平均精度 (Average Precision, AP)

- 通常**未詳細標註**的平均精度AP泛指為IoU閾值( $TH_{IoU}$ )是0.5所計算的，此時AP可以表示為 $AP(TH_{IoU}=0.5)$ 。
- 但在公開COCO 資料庫常比較的為更複雜的AP，例如 $AP[.50:.05:.95]$ 的意思為計算 $AP(TH_{IoU}=0.5)$ 、 $AP(TH_{IoU}=0.55)$ 、...、 $AP(TH_{IoU}=0.95)$ 平均後的AP值

$$AP[.50:.05:.95] = \frac{\sum_{i=0}^n AP(TH_{IoU} = 0.05 \times i + 0.5)}{n + 1}, n = \frac{0.95 - 0.5}{0.05} = 9$$

$$AP[x:y:z] = \frac{\sum_{i=0}^n AP(TH_{IoU} = y \times i + x)}{n + 1}, n = \frac{z - x}{y}$$

### 平均精度均值 (Mean Average Precision, mAP)

- AP值指的為某一目標類別的偵測結果，通常一個模型能判斷出多種目標類別的結果，因此計算每個目標類別的AP後的平均，則為平均精度均值 (Mean Average Precision, mAP)