

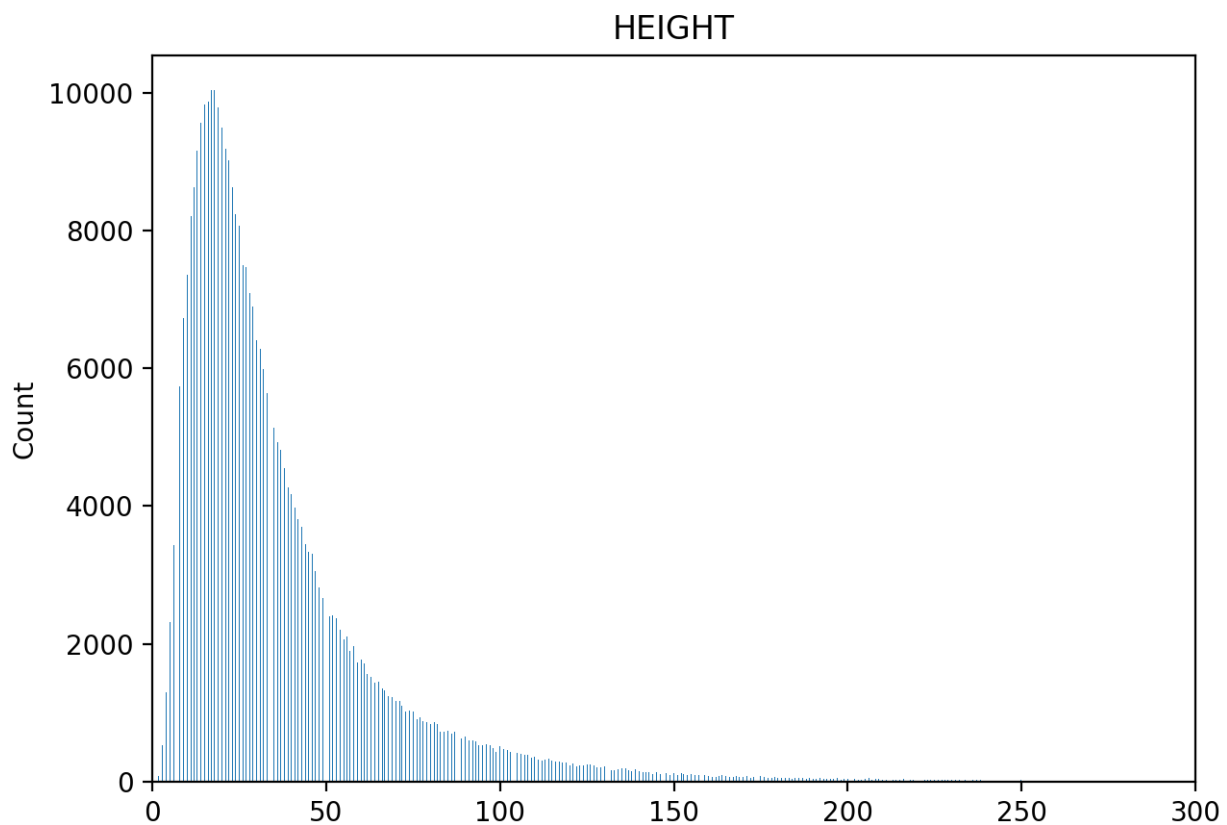
- Data exploration for VisDrone

- Class Statistics (**Training Data + Validation Data**)

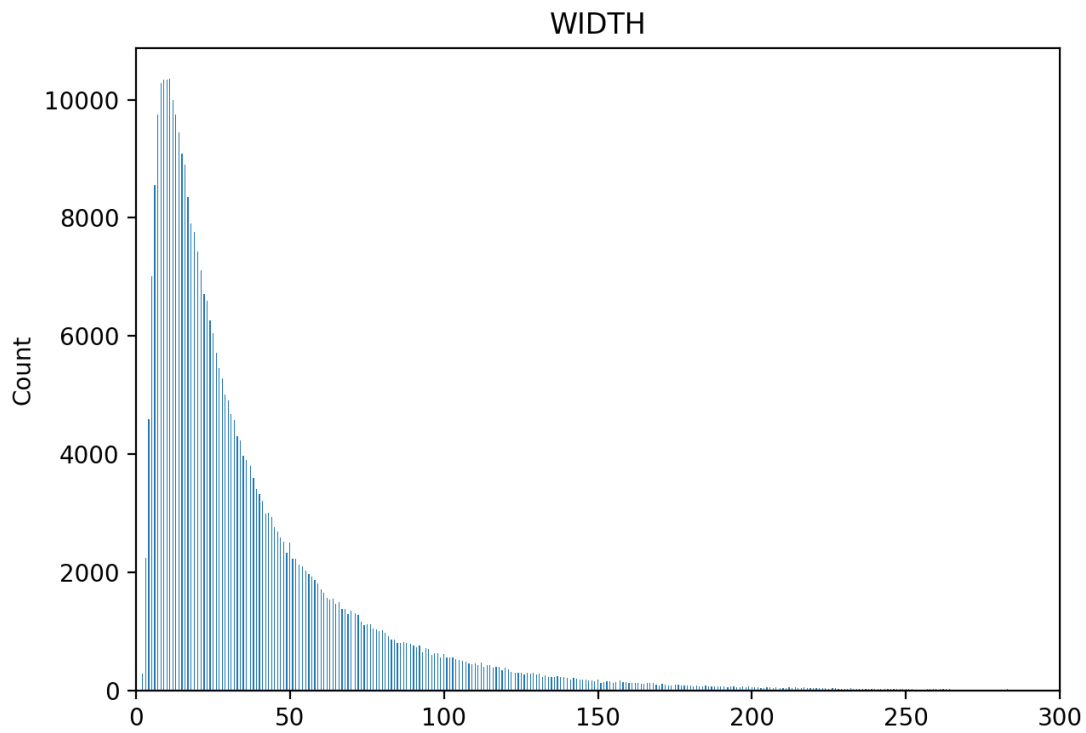
- I. 計算整個資料集 bounding box 的長、寬、面積統計資訊 : Max, Min, Average, Histogram(共  $3 \times 4 = 12$  個答案)

Bounding Box	長	寬	面積
Max	790	983	328640
Min	1	2	3
Average	37.104	38.223	2394.567

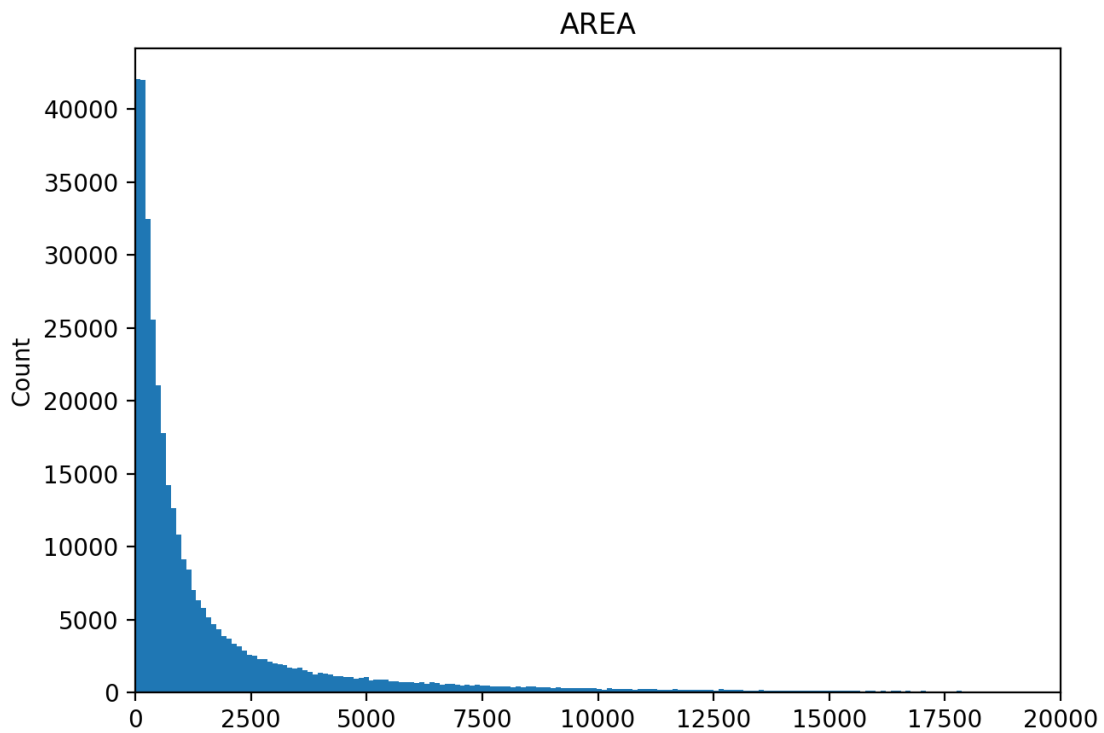
- 【Histogram】長**



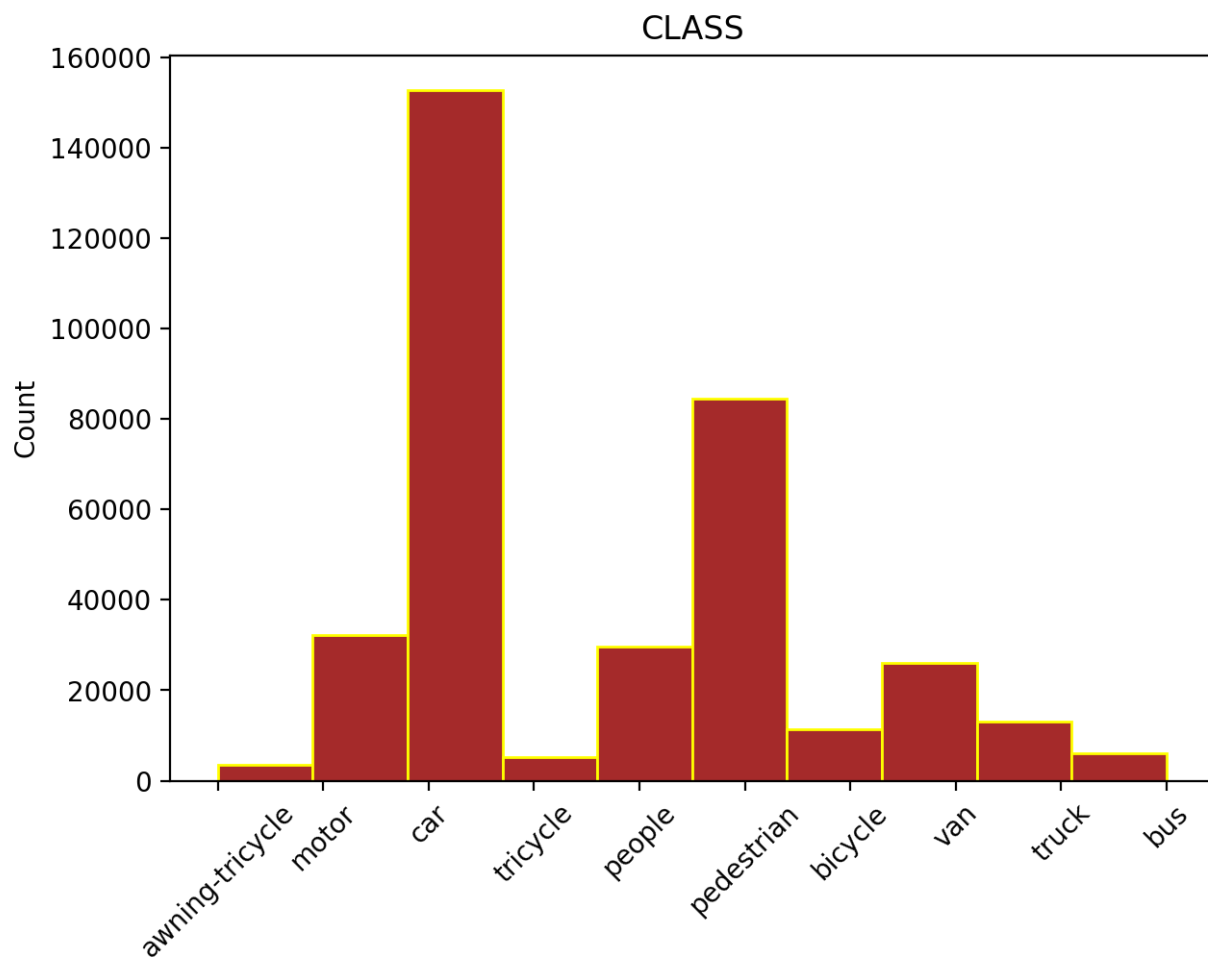
## 【Histogram】寬



## 【Histogram】面積



II. 計算整個資料集中每一個類別占比: **Histogram**

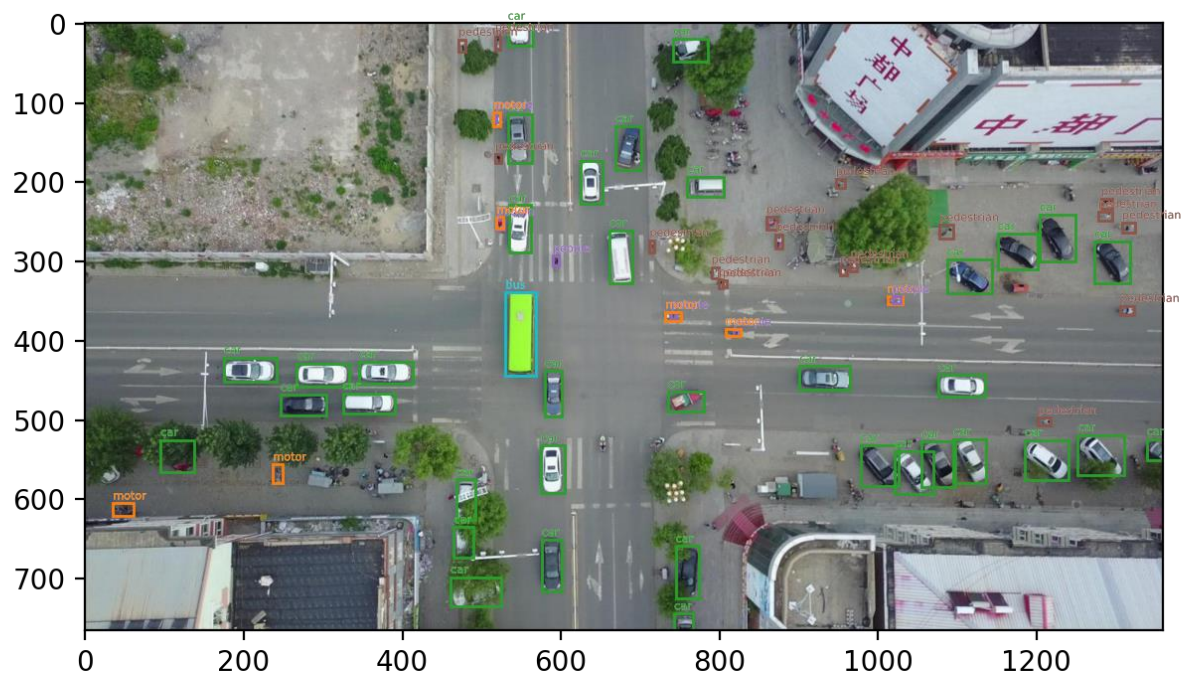


### III. Visualization

- 從資料集中隨機抽取影像(兩張)，將標準答案的框框畫上去，並標註類別名稱

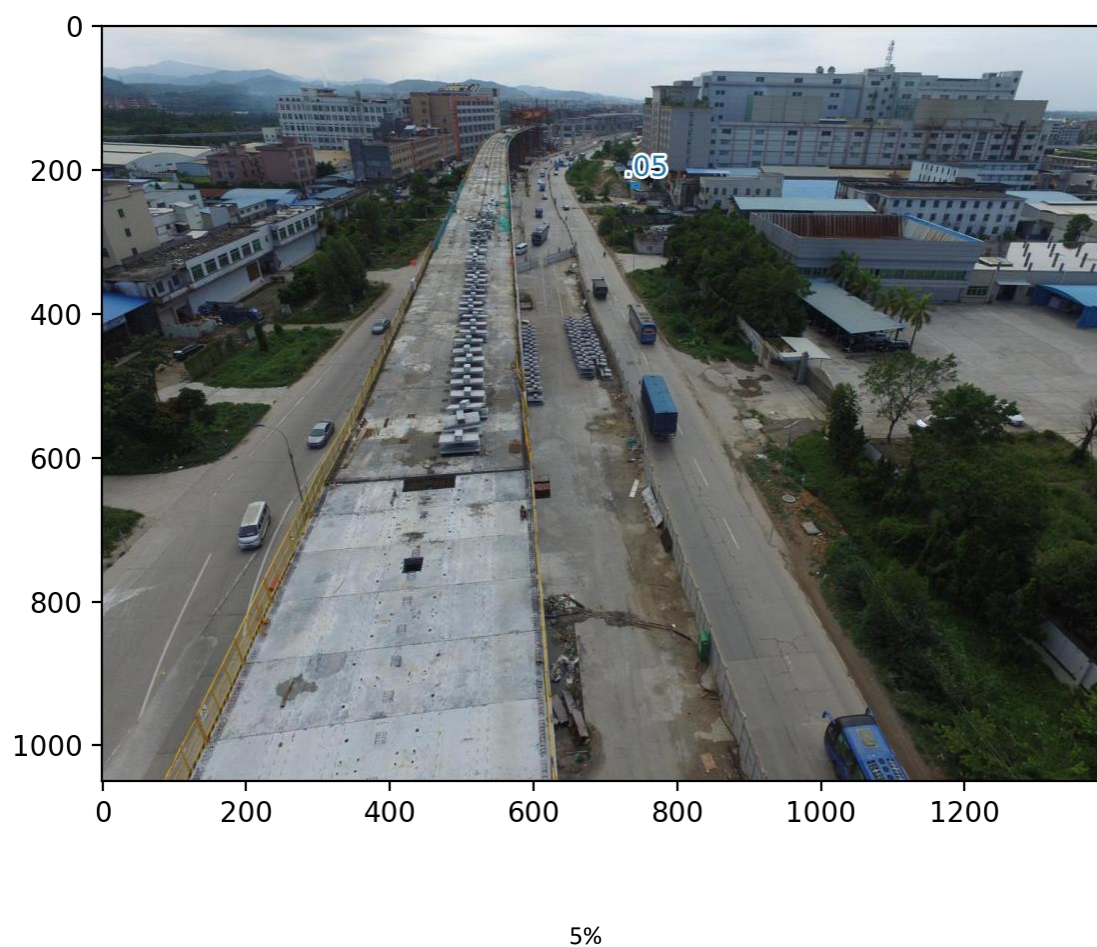


train2017/9999998\_00033\_d\_0000026.jpg

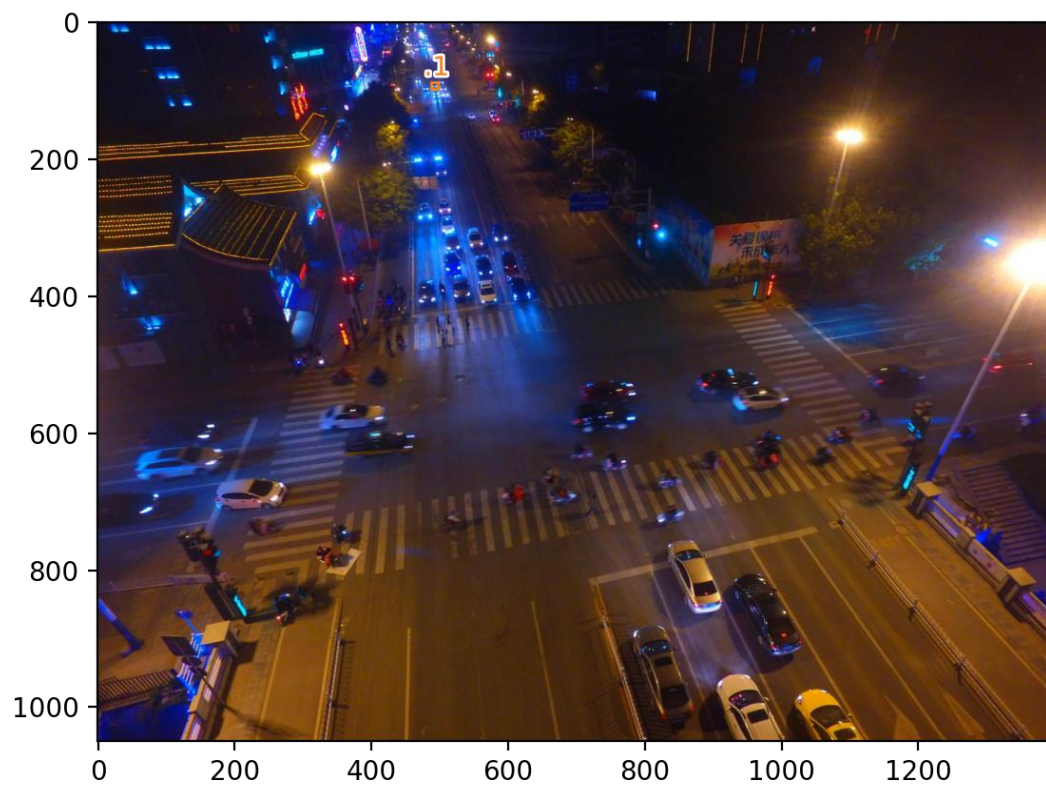


val2017/0000333\_00001\_d\_0000001.jpg

- 取出"car" 類別統計中面積統計資訊排序第 5% / 10% / 50% / 90% / 95% 的影像，除了將標準答案畫上去之外，也將該框框特別標註(用不同的顏色視覺化)



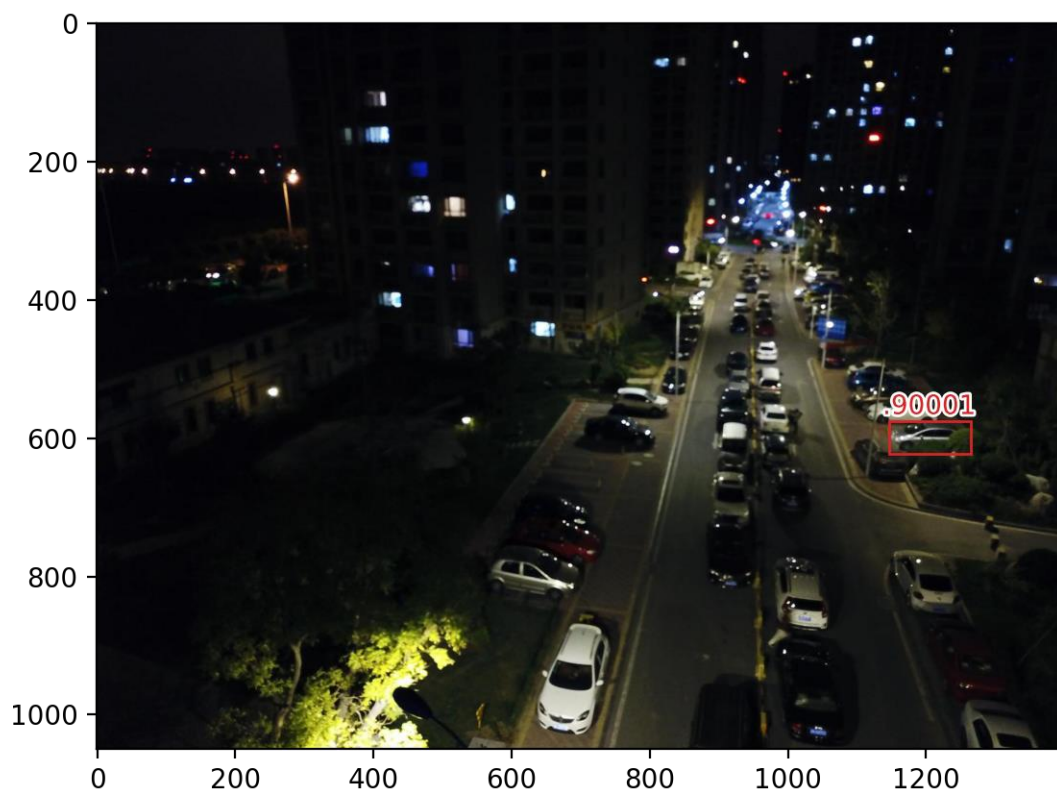




10%



50%



90% (90.001%)



95%



- 從上述的統計資訊中，請比較不同類別的統計資訊的差異(i)、類別不平衡的狀況(ii)，與你覺得經過 視覺化後，標註品質如何(iii)？
  - I. 長與寬總數最多的大約落在 20~25px，兩類型平均數差不多(37.1 vs 38.2)；面積的部分絕大多數都是小面積範圍的 Label。
  - II. Car 佔據了大多數，Tricycle 佔最少數(其中又分為兩類)。
  - III. 有些 Label 其實人眼也分辨不太出來，大多都是太細小的事物，但 Label 還是有標示出類別，圈選的範圍都相當的合理。

## ● Model

- 模型配置
  - I. Backbone: ResNet
  - II. neck: FPN
  - III. bbox\_head: RetinaHead
- 修改 'retinanet\_r50\_fpn\_1x.py'
  - I. Backbone 部分增加: norm\_eval=True
  - II. Neck 部分修改: add\_extra\_convs='on\_input'
  - III. bbox\_head 部分修改: loss\_bbox=dict(type='L1Loss', loss\_weight=1.0)
- 參數設定: 'config.py'
- 訓練時間: Epoch16 – 12hr