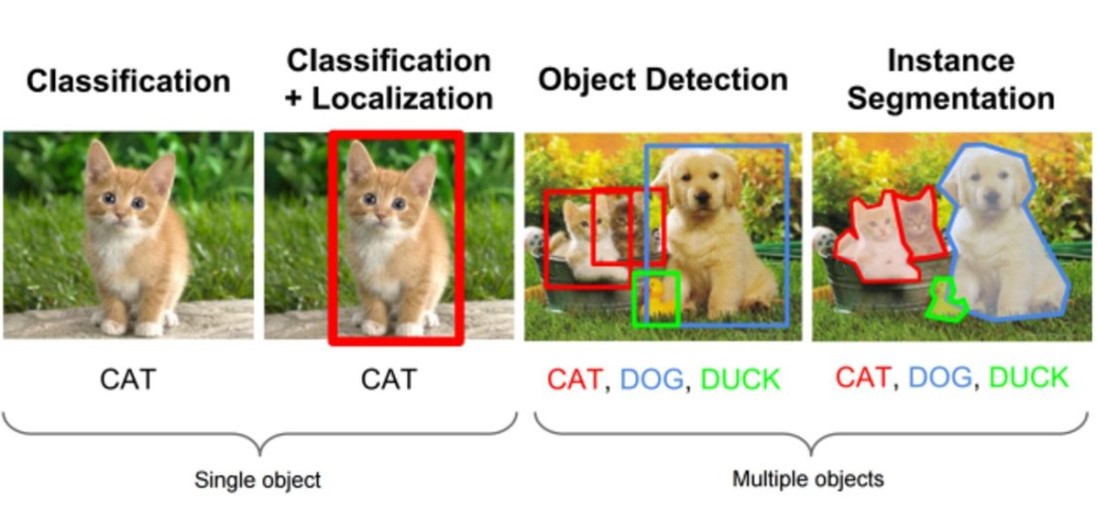
Final Project Report

Introduction:

我們這組做的是對ships in satellite imagery做Instance Segmentation, Instance Segmentation跟object detection主要差別是他會找到一個segmentation mask 來精確描繪物體的輪廓，如下圖所示。



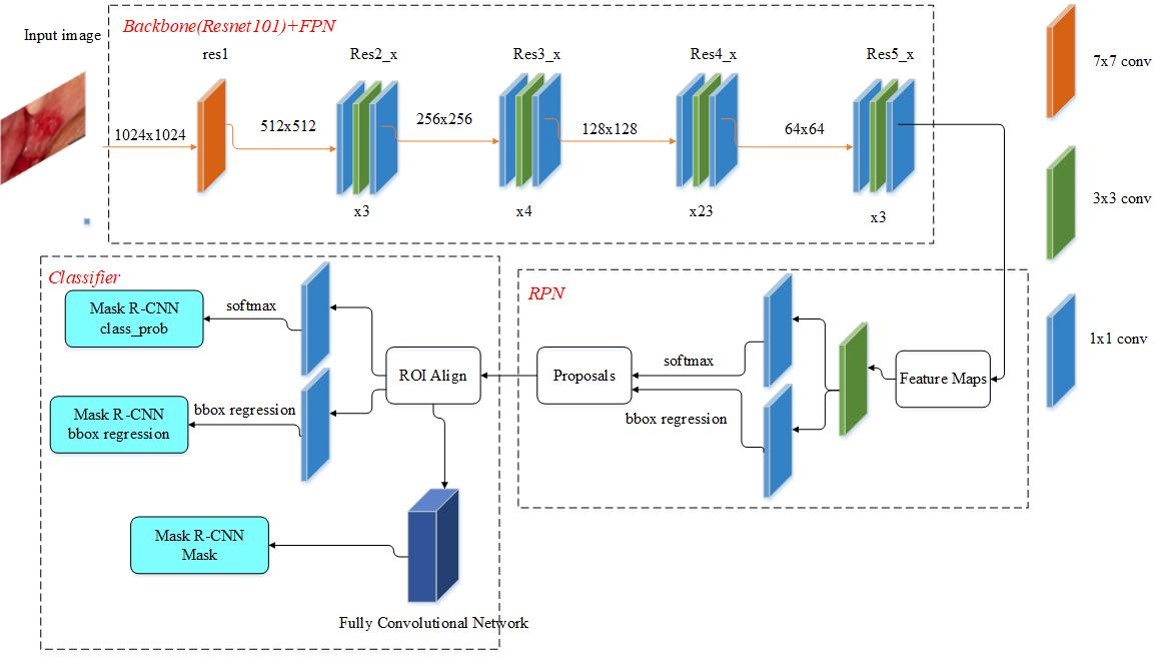
我們使用的data是ships in satellite imagery, training data 為每張80x80RGB的船隻，如下圖。



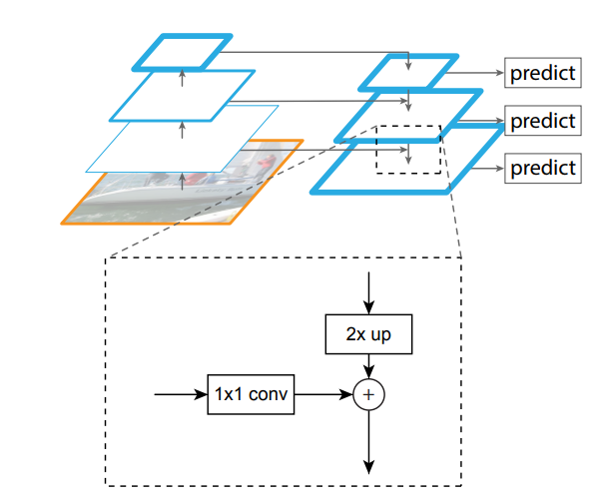
而testing data為較完整的衛星影像圖，每張長1xxx~2xxx，寬2xxx~3xxx不等，如下圖所示，



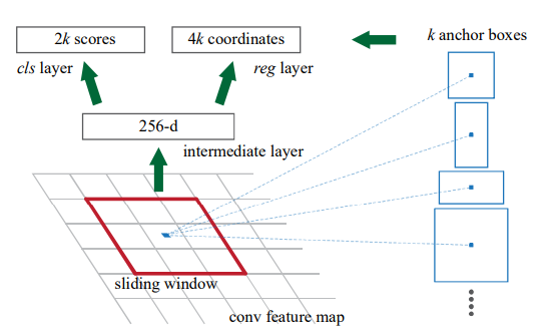
我們使用的model為Mask R-CNN，大致架構如下，Backbone為resnet101搭配feature pyramid network，用來產生特徵圖，regional proposal network(RPN)用來找出候選區域，最後在分類器進行物件的定位和分類。



feature pyramid network的動機是希望能最大限度的提取多尺度的特徵，在resnet101的是意圖如下，每個layer的output會進行up-sampling並與上一個layer做bit-wise addition.

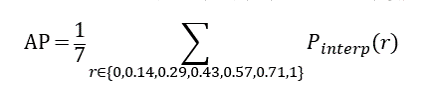


**Region Proposal Network** (RPN) 為透過sliding window的方式去找候選區域，示意圖如下:



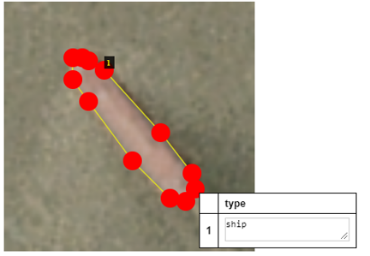
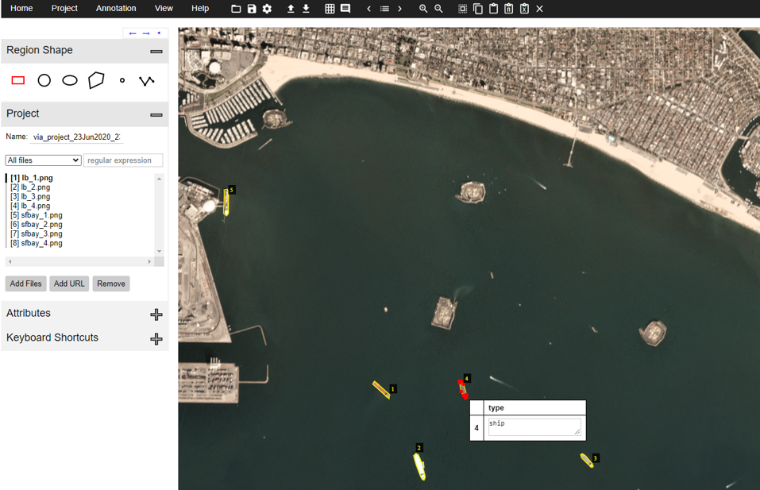
Evaluation Metric:

我們使用mAP作為我們的評估指標，mAP 為各類別AP的平均。



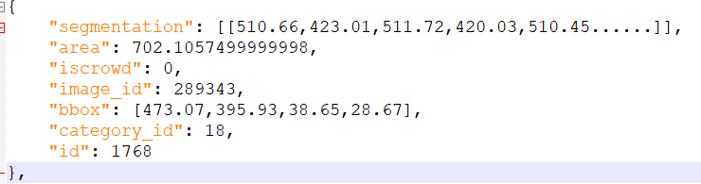
Data preprocessing:

這部分為我們主要實作的地方，我們首先使用vgg image annotator 對資料進行標註。

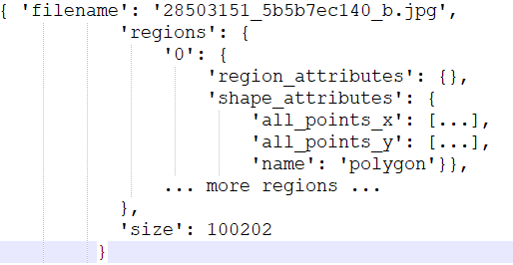


我們使用的Mask R-CNN版本讀取MS COCO的標註格式，我們需要將他改成讀取VIA的格式。

COCO:

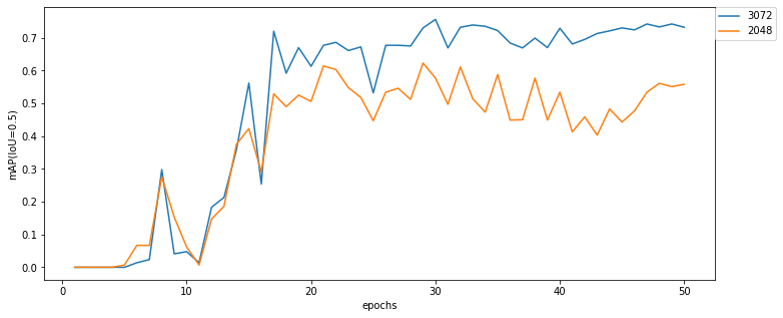


VIA:



我們更改的地方為ship.py當中load\_ship以及load\_mask的部分。

為了讓training data中一艘船的大小跟testing data中的一艘船差不多，我們盡量讓traing data的資料越小越好，而testing data越大越好，因此，traing的照片resize成256x256(模型能接受的最小數值)，而testing的部分比較2048x2048以及3072x3072。

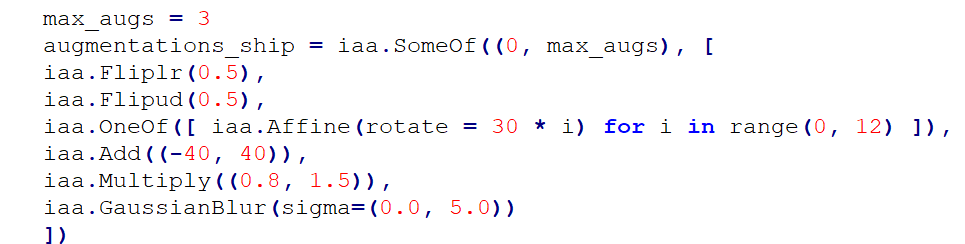


可是發現確實我們讓testing data越大，成效是越好的。

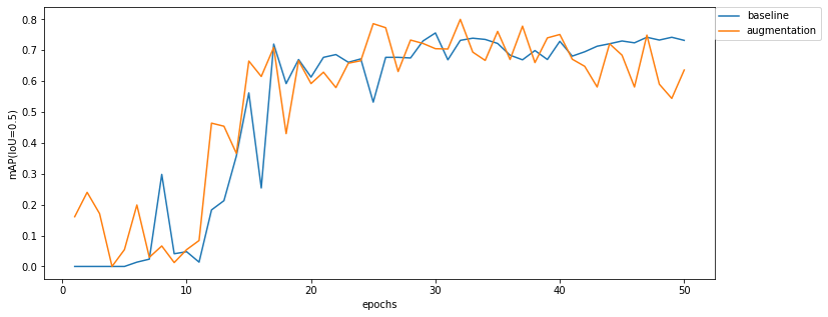
|  |  |  |
| --- | --- | --- |
| Image Size | mAP(%, IoU = 0.5) | Inference time(s) |
| 2048x2048 | 61.4 | **9.21** |
| 3072x3072 | **75.6** | 19.1 |

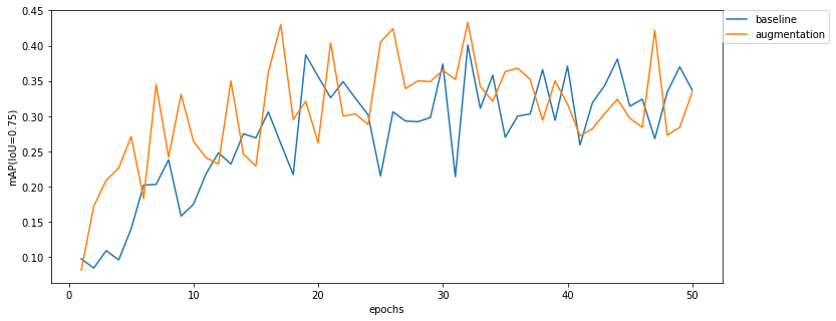
但相對來說Inference time變長。

Data augmentation的部分我們使用Flip/mirror, rotate, add, multiply, blur這幾種來確保資料多樣性。



mAP50及mAP75的結果如下:





|  |  |  |
| --- | --- | --- |
|  | mAP(%, IoU = 0.5) | mAP(%, IoU = 0.75) |
| baseline | 75.6 | 40.1 |
| Augmentation | **80.0** | **43.3** |

相對baseline來說有少許效能的提升。

最後是detection 的結果:



False Positive過高，因此提升DETECTION\_MIN\_CONFIDENCE

