

Hw3 Report

0756545 楊秉澄

Github link: https://github.com/BingChengYang/cv_hw3

Reference: <https://github.com/facebookresearch/detectron2>

Introduction:

這次作業的 task 為 tiny Pascal Voc 的 Instance Segmentation，這個 dataset 總共 1349 筆訓練資料，100 筆測試資料，20 個分類項目，baseline 為 0.247mAP，為了達成這個目標，我使用 detectron2 中的 mask rcnn 來達成。

Methodology:

Environment:

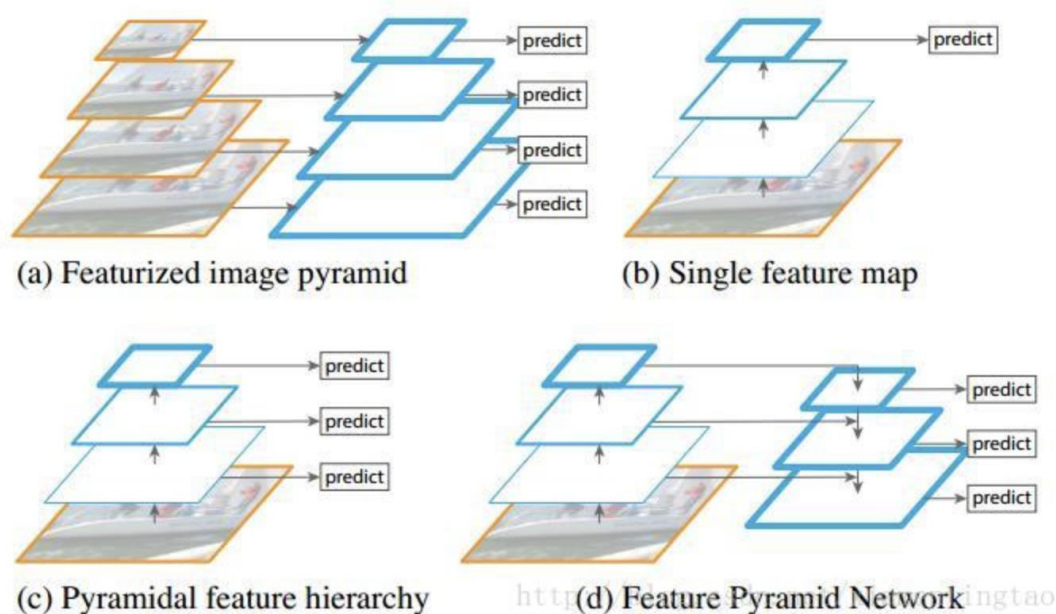
- CPU: AMD R5-2600
- RAM: 16G
- GPU: Nvidia RTX-2070 8G

Data Preprocessing:

由於我這次採用 detectron2 來實作 instances segmentation，只有對圖片使用 random flip 來作 data augmentation，來避免 overfitting 的發生。

Model Architecture:

Model 我使用 detectron2 中的 mask rcnn，並且使用 resnet101 當作我的 backbone 可以得到比 resnet50 有更好的效果，因為這次無法使用 coco/voc 的 pretrain model，所以我用 resnet101 的 ImageNet pretrain model 可以讓 model 更快的收斂且有比較好的結果，採用 feature pyramid network (fpn)是為了避免小偵測目標在最後一層的時候特徵都已經遺失了，所以採用 feature pyramid network 可以有效解決多尺度的檢測問題，可以將前幾層的 feature 也能被使用。



Hyperparameters:

參數分別有 learning rate(lr)、mini_batch_size、Max_iter 需要設置，lr 我試過 0.0001, 0.00005 兩種其中 0.00005 最終的 total loss 比較小，但是在 testing 的表現只有 0.36mAP，0.0001 的 loss 比較大但是最後的 testing 可以達到 0.387mAP，0.00005 有一點 overfitting 的現象發生。

Mini_batch_size 嘗試了 32、64、128 其中 3 者訓練起來效果差不多，但 32 訓練出來的 mAP 稍微好一點點，因此用 32 來訓練我最終的模型。

Max_iter 設定為 10000，也有試過 20000，但是有點 overfitting，所以最終以 10000 為我的設定。

Summary

這次的 task 訓練資料比較少，因此要注意 overfitting 的問題，resnet101 比起 resnet50 有較多的參數可以去避免 overfitting 以達到更好的，同樣的 learning rate 和 epoch 也要去注意避免太過於去擬合，整體而言只要使用 mask rcnn 就可以很輕鬆的擊敗 baseline。