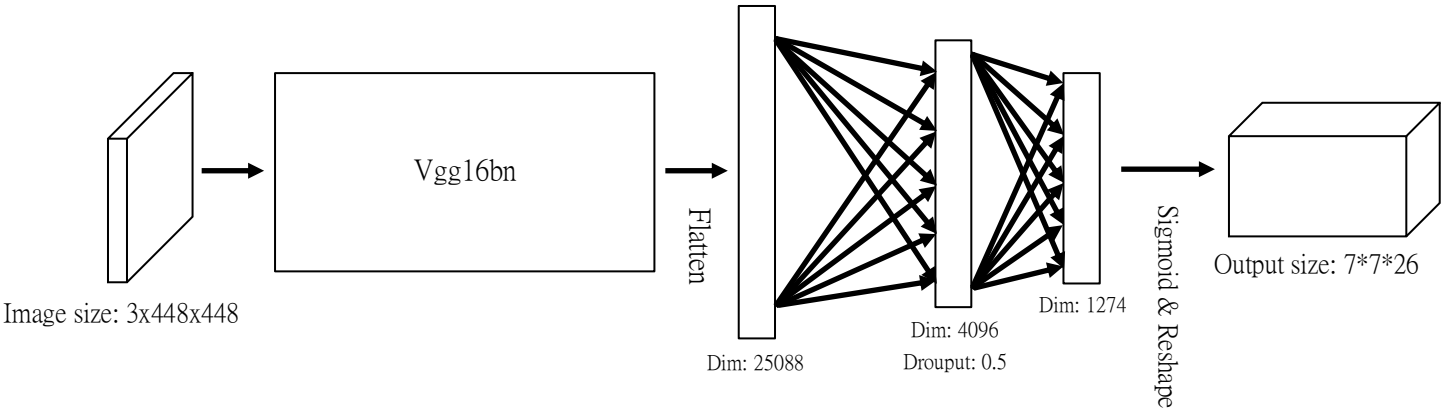
Name: 陳紘豪　Dep.:資工碩一　Student ID:R07922134

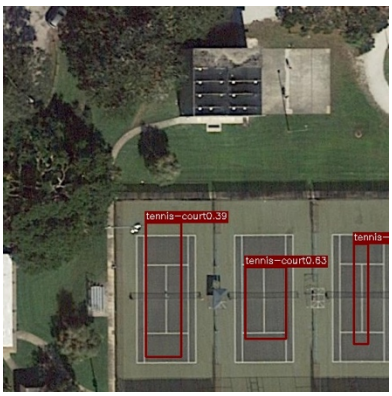
1. (5%) Print the network architecture of your YoloV1-vgg16bn model and describe your training config. (optimizer,batch size…and so on)

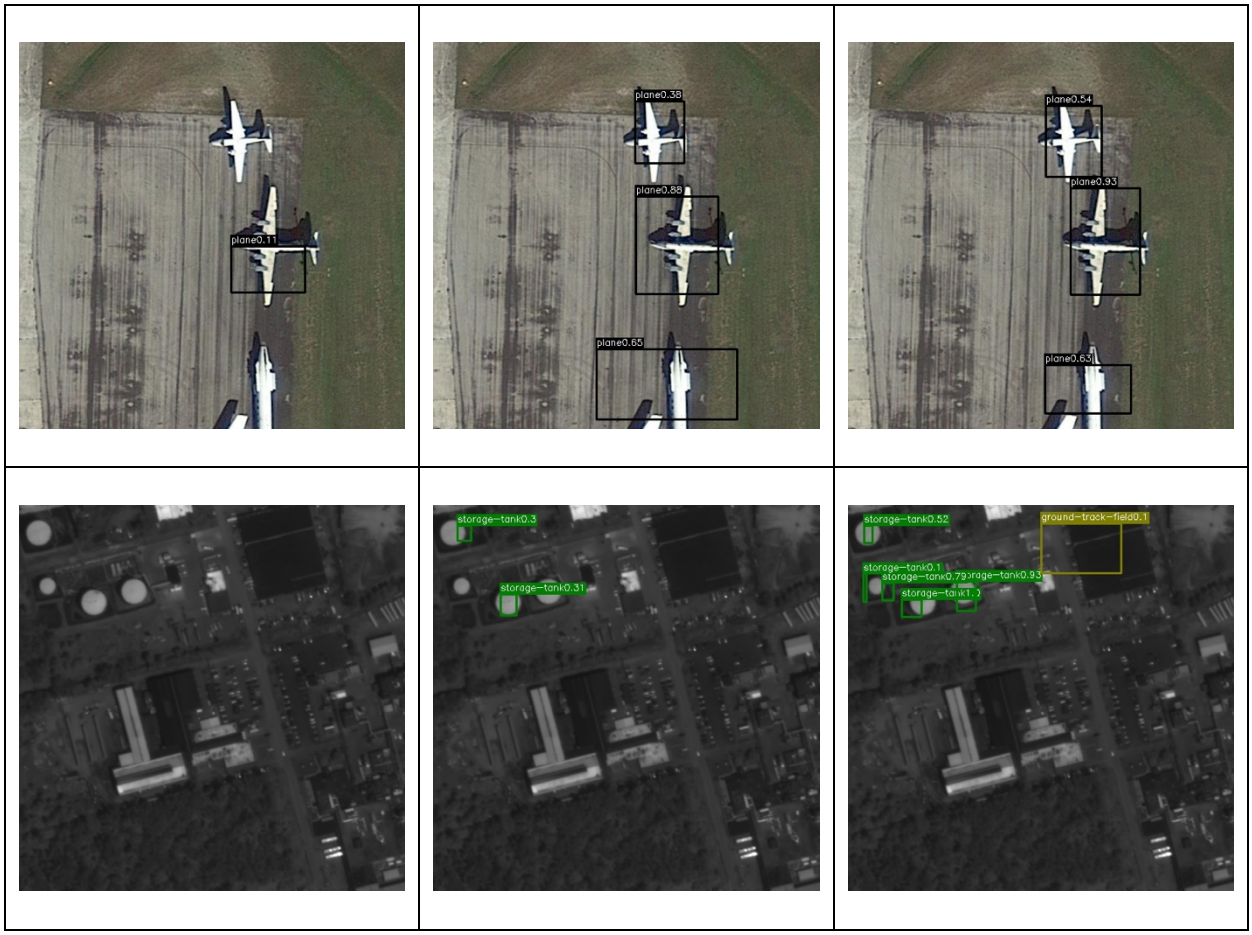
Training config

- Epoch：10
- Optimizer：Adam（除了 lr 外都採 default 值）
- lr：1e-4
- Batch size：16
- Score threshold：0.1
- NMS IoU threshold：0.4
- Loss lambda：5, 0.5（default）
- Input size：448 * 448
- Data augmentation：水平翻轉、垂直翻轉、水平+垂直翻轉
- 因為有 data augmentation，training set 變成了60000張圖片，因此 train 10 個 epoch 即可過 baseline 了。



2. (10%) Show the predicted bbox image of “val1500/0076.jpg”, “val1500/0086.jpg”, “val1500/0907.jpg” during the early, middle, and the final stage during the training stage.

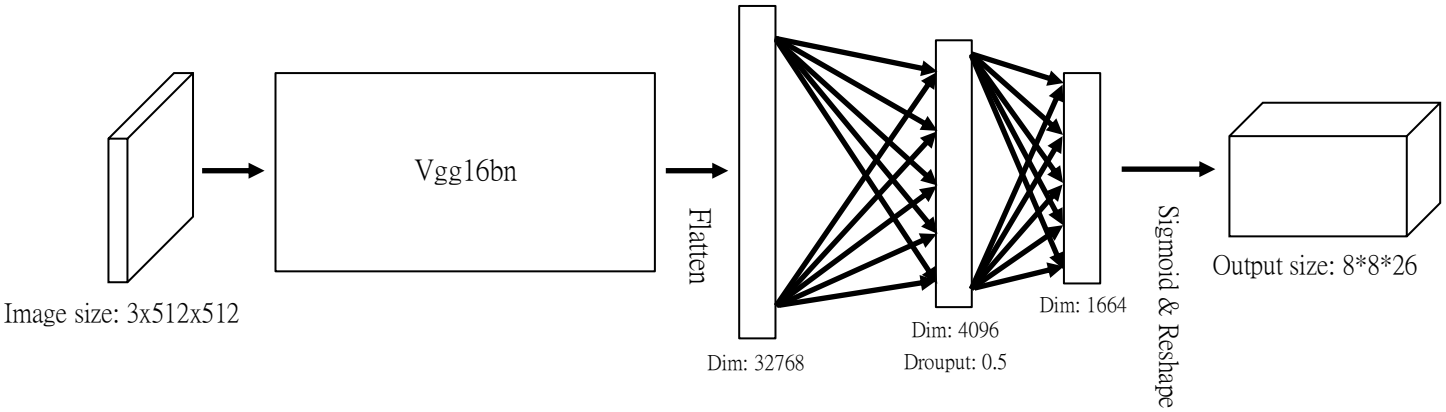
Epoch 1	Epoch 5	Epoch 10
		



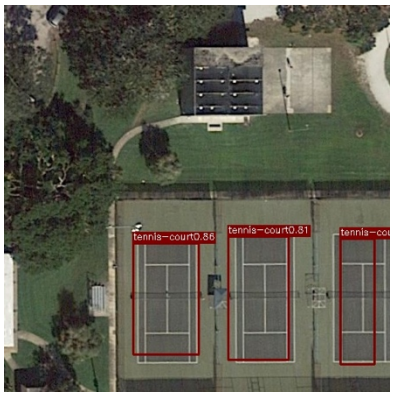
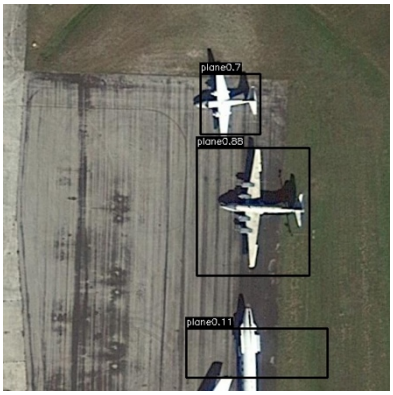
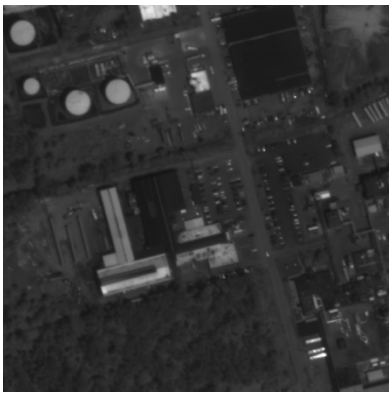
3. (10%) Implement an improved model which performs better than your baseline model. Print the network architecture of this model and describe it.

Training config

- Epoch：10
- Optimizer：Adam（除了 lr 外都採 default 值）
- lr：1e-4
- Batch size：16
- Score threshold：0.1
- NMS IoU threshold：0.4
- Loss lambda：5, 0.5（default）
- Input size：512 * 512、Grid number：8
- Data augmentation：水平翻轉、垂直翻轉、水平+垂直翻轉
- 我想看 Grid number 對 mAP 的影響，於是改用原圖大小以及 Grid number = 8 的設定去 train model，並觀察結果。



4. (10%) Show the predicted bbox image of “val1500/0076.jpg”, “val1500/0086.jpg”, “val1500/0907.jpg” during the early, middle, and the final stage during the training process of this improved model.

Epoch 1	Epoch 5	Epoch 10
		
		
		

5. (15%) Report mAP score of both models on the validation set. Discuss the reason why the improved model performs better than the baseline one. You may conduct some experiments and show some evidences to support your reasoning.

Baseline model: mAP = 0.09974 = 9.974%

```
(hh) titan@titan-desktop:~/Desktop/test$ python3 hw2_evaluation_task.py orig hw2_train_val/val1500/labelTxt_hbb/
classname: plane
ap:  0.16420940376844553
classname: baseball-diamond
ap:  0.23142482517482518
classname: bridge
ap:  0.015673981191222573
classname: ground-track-field
ap:  0.045454545454545456
classname: small-vehicle
ap:  0.0303030303030303
classname: large-vehicle
ap:  0.05303030303030303
classname: ship
ap:  0.09090909090909091
classname: tennis-court
ap:  0.37930617084836393
classname: basketball-court
ap:  0.08585858585858586
classname: storage-tank
ap:  0.0303030303030303
classname: soccer-ball-field
ap:  0.18958818958818957
classname: roundabout
ap:  0.11688311688311688
classname: harbor
ap:  0.08292011019283746
classname: swimming-pool
ap:  0.08
classname: helicopter
ap:  0.0
classname: container-crane
ap:  0.0
map: 0.09974152396909919
```

Improved model: mAP = 0.11350 = 11.35%

```
(hh) titan@titan-desktop:~/Desktop/test$ python3 hw2_evaluation_task.py improve hw2_train_val/val1500/labelTxt_hbb/
classname: plane
ap:  0.257634811046095
classname: baseball-diamond
ap:  0.15584415584415584
classname: bridge
ap:  0.1111111111111111
classname: ground-track-field
ap:  0.15151515151515152
classname: small-vehicle
ap:  0.022727272727272728
classname: large-vehicle
ap:  0.098294263388603
classname: ship
ap:  0.06818181818181818
classname: tennis-court
ap:  0.3853277458116738
classname: basketball-court
ap:  0.09090909090909091
classname: storage-tank
ap:  0.045454545454545456
classname: soccer-ball-field
ap:  0.18181818181818182
classname: roundabout
ap:  0.05916305916305916
classname: harbor
ap:  0.07652474108170311
classname: swimming-pool
ap:  0.1116427432216906
classname: helicopter
ap:  0.0
classname: container-crane
ap:  0.0
map: 0.11350929320463451
```

原因：

一樣是 train 10 epoch 的結果，但是 grid number = 8 的 performance 比 grid number = 7 的 performance 還要好，可以推測 mAP 的好壞跟 grid number 有一定程度上的關係。

grid number 比較大的話，grid size 會比較小，其中所包含的 object 數量自然也少，所以每個 grid 可以專心預測該 grid 中有的 object，而 grid number 比較少的話，例如取極端值 grid number = 1，整張圖片就是一個 grid，而我們又限制一個 grid 只能預測裡面的一個 object，這樣遇到裡面有一堆小物體的圖片，performance就會整個被拉下來。

但是 grid number 太大的話，也不一定會讓 performance 真的很好，grid number 太大可能會使得 grid size 很小，小到什麼資訊都裝不下，這樣也不行，所以要在取參數時小心別適得其反。

實驗：

為了驗證，我再去訓練一個 grid number = 1 的 model，也就是整張圖片就只有一個 grid，其表現如下：

```
(hh) titan@titan-desktop:~/Desktop/dlcv$ python3 hw2_evaluation_task.py grid_1/ hw2_train_val/val1500/labelTxt_hbb/
classname: plane
ap:  0.0
classname: baseball-diamond
ap:  0.0
classname: bridge
ap:  0.0
classname: ground-track-field
ap:  0.0
classname: small-vehicle
ap:  0.0004308487720809996
classname: large-vehicle
ap:  0.0
classname: ship
ap:  0.0
classname: tennis-court
ap:  0.0
classname: basketball-court
ap:  0.0
classname: storage-tank
ap:  0.0
classname: soccer-ball-field
ap:  0.0
classname: roundabout
ap:  0.0
classname: harbor
ap:  0.0
classname: swimming-pool
ap:  0.0
classname: helicopter
ap:  0.0
classname: container-crane
ap:  0.0
map: 2.6928048255062476e-05
```

mAP 只有 2.6*10^-5，非常的低，因此驗證了我所說的 grid number 真的會對 performance 有影響，而且 grid number 太少 performance 會很低。

6. bonus (5%) Which classes prediction perform worse than others? Why? You should describe and analyze it.

觀察到 container-crane 、helicopter 的 ap 不管在哪個 model 都是 0.0，把 training data 中所有出現的 class 統計出現次數後得到以下：

```
~/Documents/Courses/DLCV/dlcv_hw2/hw2-hhccode  master
$ python3 Count.py
Numbers of each class in the training data.
plane: 8723
ship: 34585
storage-tank: 5199
baseball-diamond: 515
tennis-court: 3279
basketball-court: 661
ground-track-field: 621
harbor: 7457
bridge: 2114
small-vehicle: 116228
large-vehicle: 23746
helicopter: 434
roundabout: 537
soccer-ball-field: 590
swimming-pool: 1977
container-crane: 136
```

可以看到 container-crane 出現的次數是所有 class 中最少的，而且 helicopter 的次數是第二少的，這種 imbalance class 的發生導致了 model 無法正確辨識。
若要增加 imbalance class 的 ap 的話，可以試著增加那些 class 的資料，讓 model 可以去學到那些 class 的特性，使得整體 performance 上升。