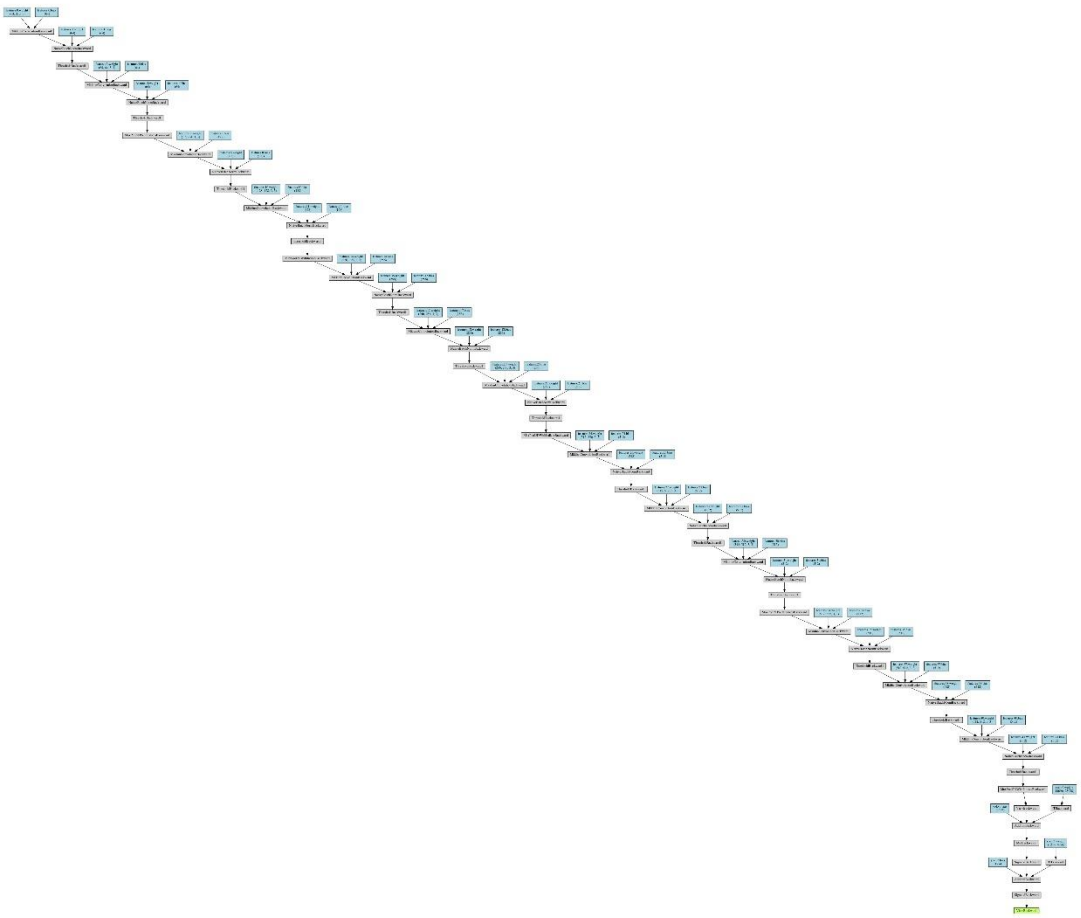
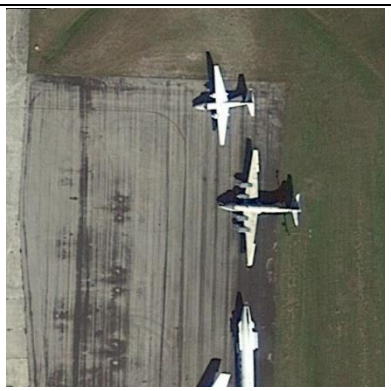
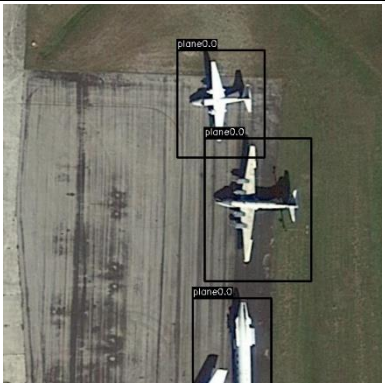
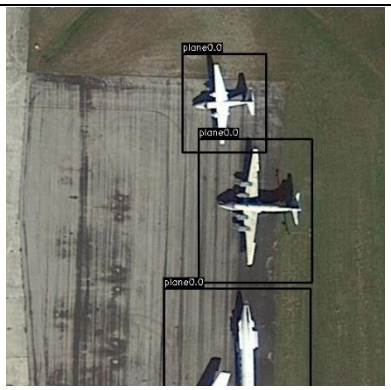
You can modify this report template, and upload your results in PDF format. Reports in other forms/formats will result in ZERO point. Reports written in either Chinese or English are both acceptable. The length of your report should NOT exceed 6 pages (**excluding bonus**).

Name: 謝忱 Dep.:電機碩二 Student ID:R06921088

1. (5%) Print the network architecture of your YoloV1-vgg16bn model and describe your training config. (optimizer,batch size….and so on)
   optimizer=torch.optim.SGD(params,lr=learning_rate,momentum=0.9, weight_decay=5e-4)
   learning_rate = 0.002,num_epochs = 100,batch_size = 16
   10epoch:learning_rate = 0.0015
   20epoch:learning_rate = 0.0010
   50epoch:learning_rate = 0.0005
   60epoch:learning_rate = 0.0001
   80epoch:learning_rate = 0.00001
   VGG16:



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. (For example, results of 1st, 10th, 20th epoch)

	0076.jpg	0086.jpg	0907.jpg
Epoch0 (early)			
Epoch30 (middle)			
Epoch80 (final)			

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

learning_rate = 0.002,num_epochs = 30,batch_size = 16

10epoch:learning_rate = 0.0015

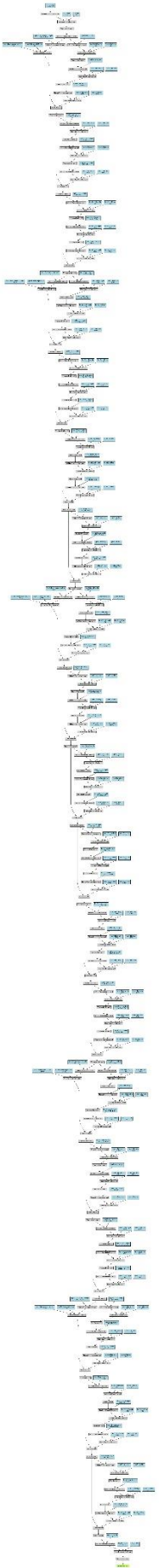
20epoch:learning_rate = 0.0010

50epoch:learning_rate = 0.0005

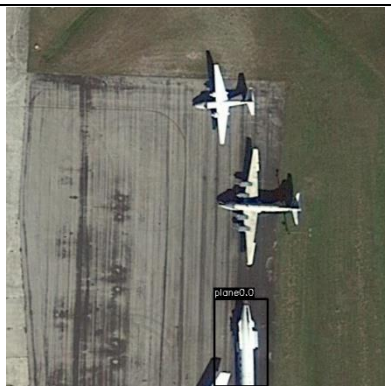
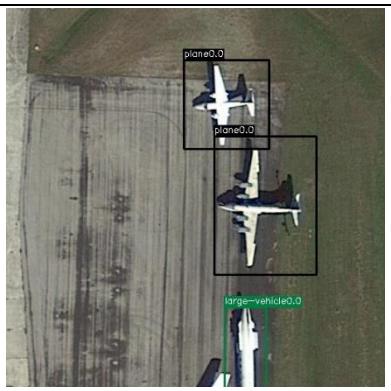
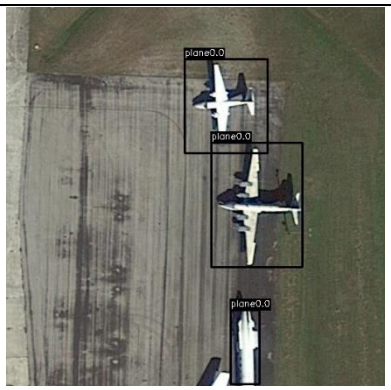
60epoch:learning_rate = 0.0001

80epoch:learning_rate = 0.00001

ResNet50:

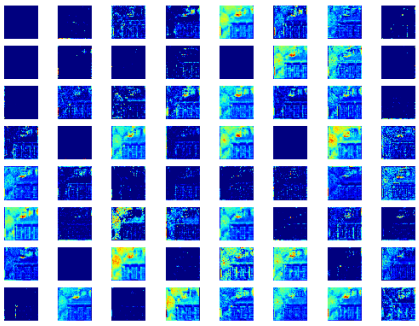
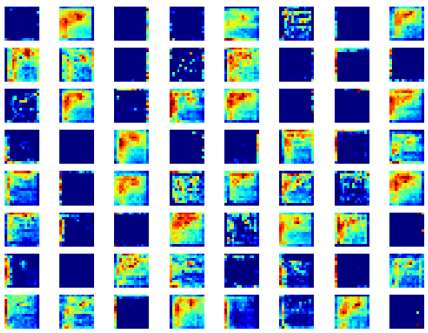


4. (10%) Show the predicted bbox image of “val1500/0076.jpg”, “val1500/0086.jpg”, “val1500/0907.jpg” during the early, middle, and the final stage during the training process of this improved model.

	0076.jpg	0086.jpg	0907.jpg
Epoch0 (early)			
Epoch10 (middle)			
Epoch30 (final)			

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.

VGG-16	ResNet50
xiec@xiec-System-Product-Name: ~/DLCV/hw2-xiechen0692 xiec@xiec-System-Product-Name:~/DLCV/hw2-xiechen0692$ python3 hw2 l1500/labelTxt_hbb/ classname: plane ap: 0.3484548483080413 classname: baseball-diamond ap: 0.0 classname: bridge ap: 0.16764247125925216 classname: ground-track-field ap: 0.0106951871657754 classname: small-vehicle ap: 0.03223378630014256 classname: large-vehicle ap: 0.15889464594127808 classname: ship ap: 0.07181425485961124 classname: tennis-court ap: 0.4537876339346927 classname: basketball-court ap: 0.0 classname: storage-tank ap: 0.10699226484110205 classname: soccer-ball-field ap: 0.0 classname: roundabout ap: 0.0 classname: harbor ap: 0.20524422121036945 classname: swimming-pool ap: 0.16414141414141414 classname: helicopter ap: 0.0 classname: container-crane ap: 0.0 map: 0.10749379549760495 xiec@xiec-System-Product-Name:~/DLCV/hw2-xiechen0692$	xiec@xiec-System-Product-Name: ~/DLCV/hw2-xiechen0692 xiec@xiec-System-Product-Name:~/DLCV/hw2-xiechen0692$ python3 t/ hw2_train_val/val1500/labelTxt_hbb/ classname: plane ap: 0.5432282368649283 classname: baseball-diamond ap: 0.28306356000321514 classname: bridge ap: 0.17420970295301846 classname: ground-track-field ap: 0.08054226475279107 classname: small-vehicle ap: 0.0895690548086013 classname: large-vehicle ap: 0.24611221536422917 classname: ship ap: 0.15311329140882654 classname: tennis-court ap: 0.7665448582140132 classname: basketball-court ap: 0.1194144838212635 classname: storage-tank ap: 0.21400097228974235 classname: soccer-ball-field ap: 0.11554621848739496 classname: roundabout ap: 0.11586452762923351 classname: harbor ap: 0.275713295543999 classname: swimming-pool ap: 0.37438905180840665 classname: helicopter ap: 0.0 classname: container-crane ap: 0.0 map: 0.22195698337185396 xiec@xiec-System-Product-Name:~/DLCV/hw2-xiechen0692$

VGG,使用 0076.jpg 從 VGG16 的第二層 conv 中提取 64 個特徵圖	
ResNet-50，使用 0076.jpg 從 ResNet-50 的第二層 conv 中提取 64 個特徵圖，同樣是 best model，可以看出 resnet 中提取出的特徵更加明顯。CNN 提取特徵的能力越强，對於圖片任務的效果就會越好。	

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

baseball-diamond， basketball-court， soccer-ball-field， roundabout， helicopter， container-crane 判斷的 ap 都是 0,ground-track-field 也接近 0,統計所有類別出現的次數如下，可以看出這些比其他預測的表現更差的類別，這些類別的樣本數量都排在倒數，並且他們的數目只有幾百個，相對於其他類別幾千或者幾萬的數量，會造成樣本的不均衡，導致這些類別的預測表現較差。

各類別出現的次數：

Counter({'small-vehicle': 116228, 'ship': 34585, 'large-vehicle': 23746, 'plane': 8723, 'harbor': 7457, 'storage-tank': 5199, 'tennis-court': 3279, 'bridge': 2114, 'swimming-pool': 1977, 'basketball-court': 661, 'ground-track-field': 621, 'soccer-ball-field': 590, 'roundabout': 537, 'baseball-diamond': 515, 'helicopter': 434, 'container-crane': 136})

Reference: https://github.com/xiongzihua/pytorch-YOLO-v1