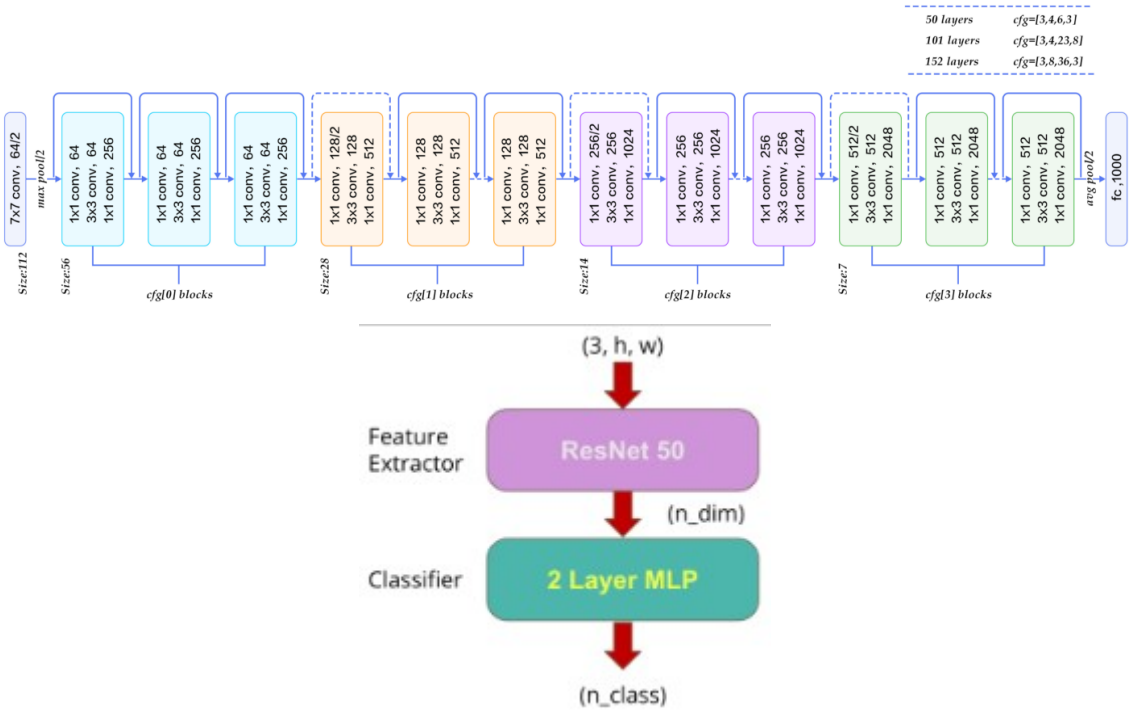
Report

Problem 1

1. Draw the network architecture of method A or B.
    - The graph should be brief and clear
    - It would be fine to straight copy the figure from the paper

    Model B



2. Report accuracy of your models (both A, B) on the validation set.

Model	Mine CNN (from scratch)	Resnet50 (pretrain)
Accuracy	0.37	0.864

3. Report your implementation details of model A.
    - Including but not limited to optimizer, loss function, cross validation method

      由於是圖片的緣故，第一個想法就是套用CNN，然後因為圖片原本很小，所以有用到ConvTranspose。

      optimizer使用的是AdamW，可以快速收斂

      loss function使用CrossEntropy，分類為題最常用的

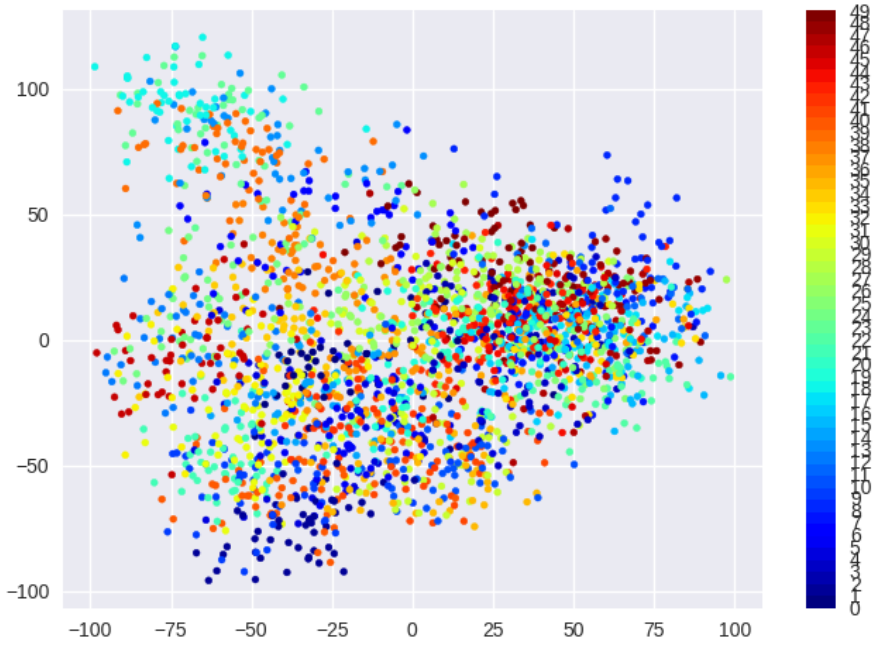
      cross validation用validation set來決定model的好壞而調整參數

4. Report your alternative model or method in B, and describe its difference from model A

   Resnet50 skip connection可以幫助model在深層時避免gradient vanishing，可以有效學習更多features

5. Visualize the learned visual representations of model A on the validation set by implementing PCA (Principal Component Analysis) on the output of the second last layer. Briefly explain your result of the PCA visualization.

PCA

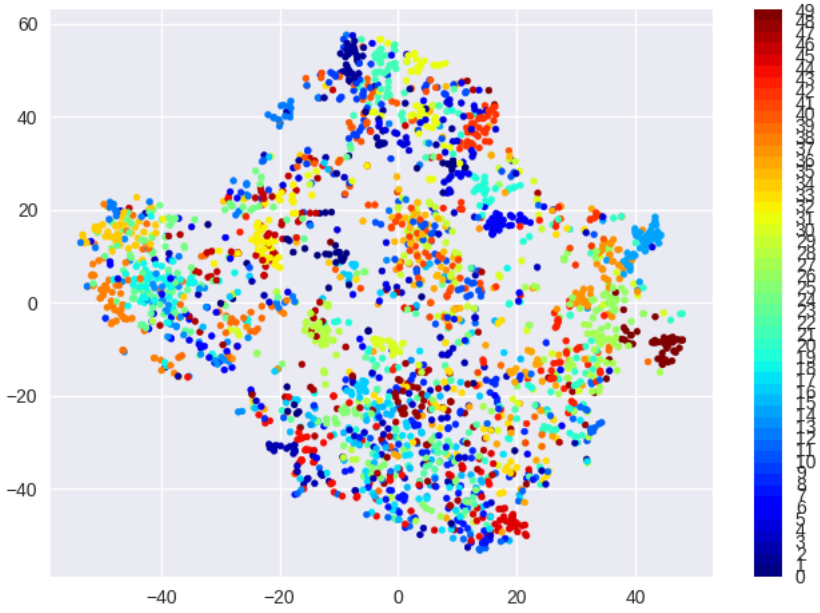


在PCA的情形下，由於NN是非線性的，但PCA是用線性的方式去降維，因此分析並不是到太符合我們要的資訊，他散落的程度還是不小。

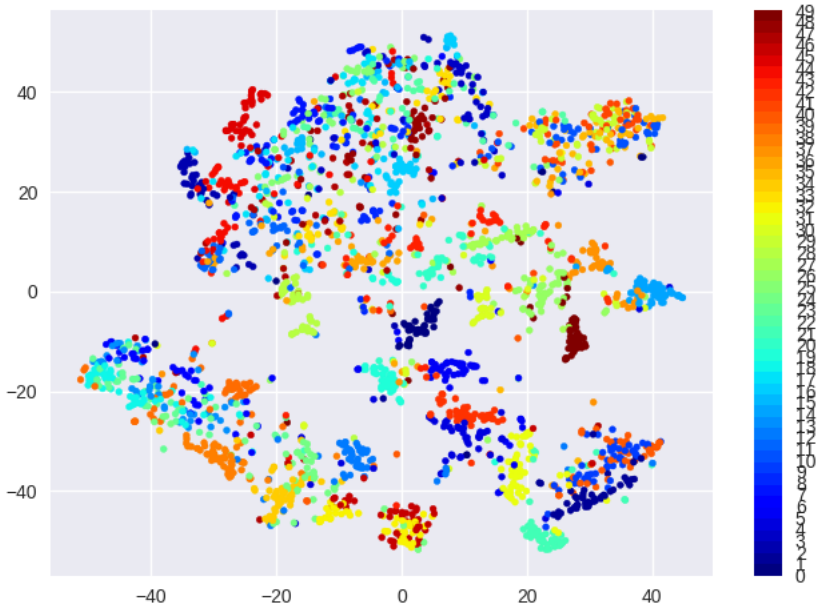
6. Visualize the learned visual representation of model A, again on the output of the second last layer, but using t-SNE (t-distributed Stochastic Neighbor Embedding) instead. Depict your visualization from three different epochs including the first one and the last one. Briefly explain the above results.

   TSNE

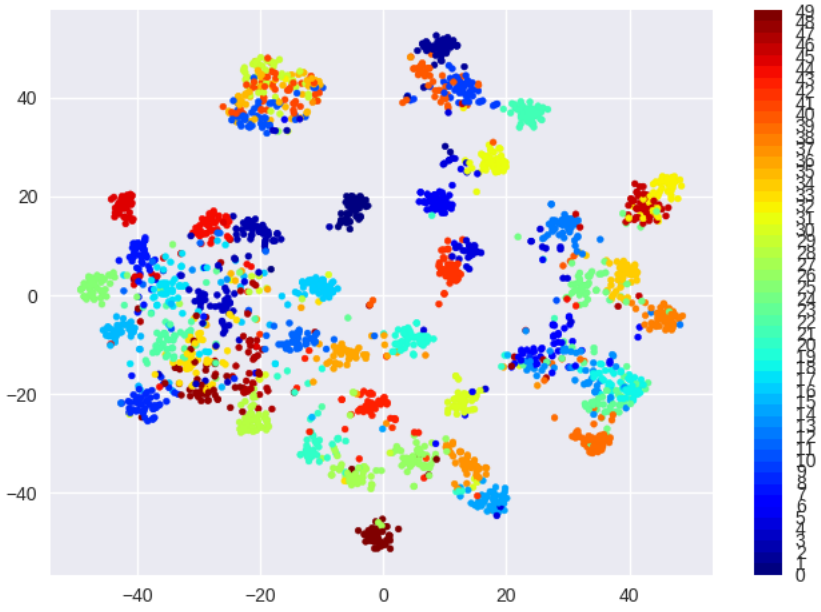
   - first



- middle



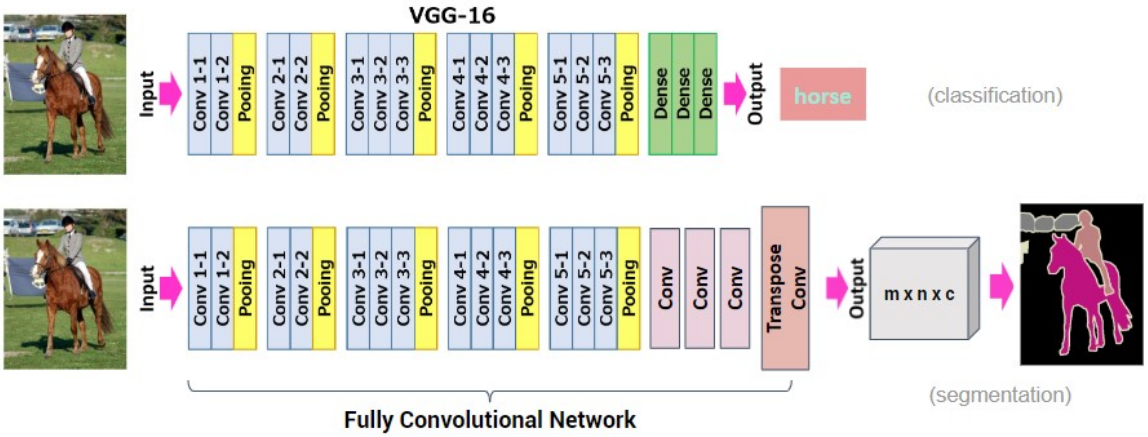
- last

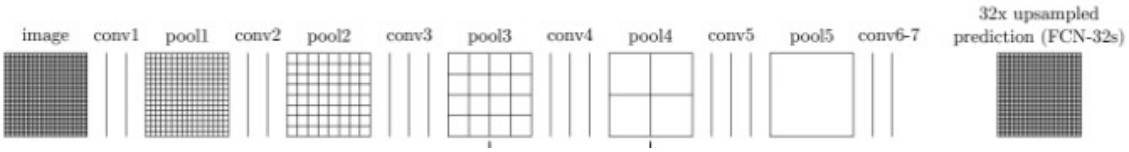


t-SNE看起來比PCA的分布還要群聚，也許是因為他本身是非線性的緣故比較貼合Model，可議看出他比較沒有同個顏色散落在各處的情形，因此認為model train的還不錯

Problem 2

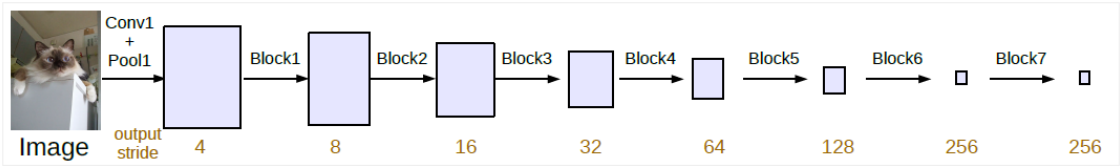
1. Draw the network architecture of your VGG16-FCN32s model (model A).



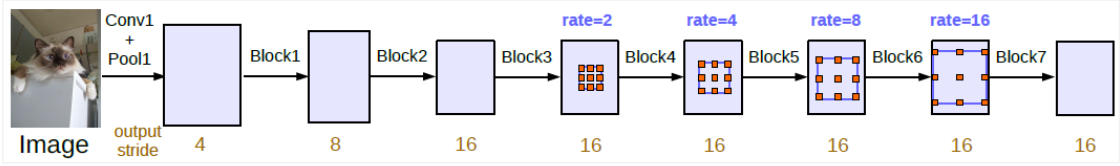


2. Draw the network architecture of the improved model (model B) and explain it differs from your VGG16-FCN32s model.

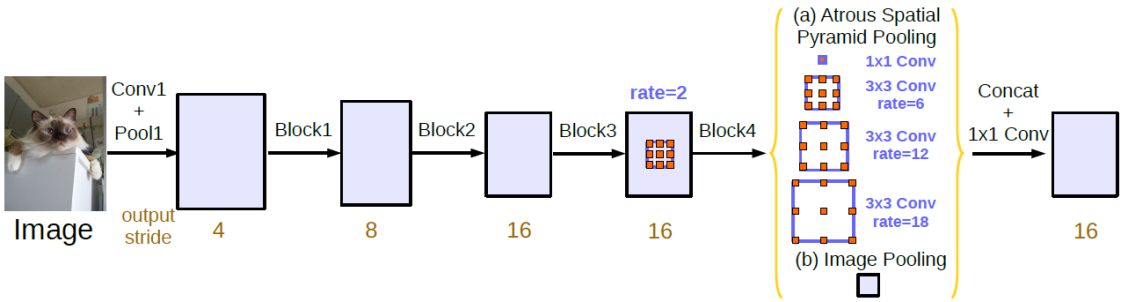
   DeepLab v3



(a) Going deeper without atrous convolution.



(b) Going deeper with atrous convolution. Atrous convolution with $rate > 1$ is applied after block3 when $output\_stride = 16$.



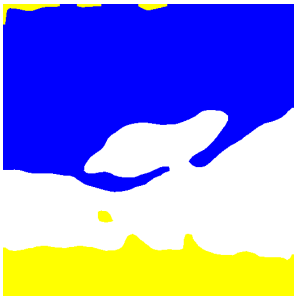
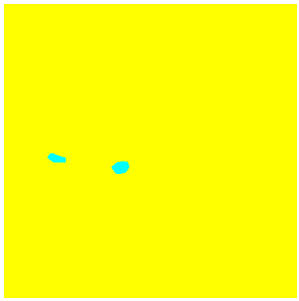
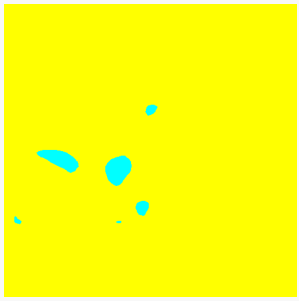
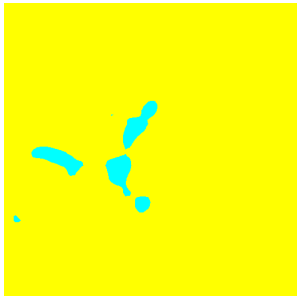
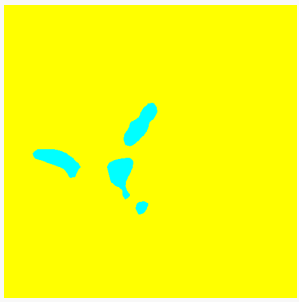
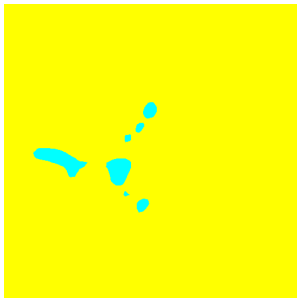
   1. backbone變強了(VGG16-Resnet-50
   2. ASPP能關注不同範圍的feature

3. Report mIoUs of two models on the validation set.

Epoch / Model	VGG16_FCN32s	DEEPLAB v3
5	0.61	0.71
10	0.68	0.72
15	0.69	0.73
20	0.70	0.73

Epoch / Model	VGG16_FCN32s	DEEPLAB v3
25	0.69	0.73

4. Show the predicted segmentation mask of “validation/0013_sat.jpg”, “validation/0062_sat.jpg”, “validation/0104_sat.jpg” during the early, middle, and the final stage during the training process of the improved model.
    - Tips: Given n epochs training, you could save the 1st, (n/2)-th, n-th epoch model, and draw the predicted mask by loading these saved models.

Epoch	0013_sat	0062_sat	0104_sat
5			
10			
15			
20			
25			

Epoch	0013_sat	0062_sat	0104_sat
Ground Truth			
Image