Q1: Transfer Learning

• Q1-1. Please report the validation accuracy of a pretrained Alexnet used as a feature extractor in the two-class classification problem. (5 pts)

在已經預訓練好的 Alexnet 之下，我們將除了最後一層(FC layer: 4096x2)之外的所有網路凍結. 在只訓練 FC layer 之下，我們得到模型的驗證準確度(validation accuracy)為 **0.80**.

```
Epoch 24/24
----------
train Loss: 0.8155 Acc: 0.7700
val Loss: 0.7581 Acc: 0.7900

Training complete in 8m 51s
Best val Acc: 0.800000
```

• Q1-2 Please report the validation accuracy of a pretrained Alexnet after it is finetuned in the two class classification problem. (5 pts)

在調整每一層參數之後，驗證準確度(validation accuracy) 為 **0.875**.

```
Epoch 24/24
----------
train Loss: 0.2984 Acc: 0.8590
val Loss: 0.2907 Acc: 0.8725

Training complete in 136m 9s
Best val Acc: 0.875000
```

• Q1-3 Please report the validation accuracy of a non-pretrained Alexnet after it is trained in the two-class classification problem. (5 pts)

在 non-pretrained Alexnet model 從新開始訓練之後，驗證準確度(validation accuracy) 為 **0.67**.

```
Epoch 24/24
----------
train Loss: 0.6518 Acc: 0.6210
val Loss: 0.6189 Acc: 0.6600

Training complete in 4m 1s
Best val Acc: 0.670000
```

- Q1-4 Please discuss the results of Q1-1, Q1-2, & Q1-3. (5 pts)

在各項 model 中, Q1-3 的 non-pretrained Alexnet model 起始的準確度最低(0.57), 且在同樣 Epoch 25 次之後, non-pretrained Alexnet model 的最佳準確度也是最低(0.67). 這說明了 transfer learning 可以有效幫助模型站在之前的經驗上進行訓練, 而不是從 0 開始.

此外 Q1-2(調整所有參數的模型), 比起 Q1-1(只調整輸出層), 雖然在起始(Epoch=1)的準確度比較低(0.67<0.775), 但是在最終訓練完 Epoch 25 次之後, 準確度是最高的(0.875>0.800). 這說明了訓練所有參數, 包含最終的分類網路(FC layer)與之前的圖型特徵萃取(feature extraction)網路, 會比只單存訓練 FC layer 的網路模型, 具有更優越的分類效果.

Table-1: Validation accuracy benchmark of Q1-1~Q1-3

Question	Situation	Initial Val Acc (Epoch=1)	Best Val Acc	Epoch of Best Val Acc
Q1-1	pretrained Alexnet used as a feature extractor	0.775	0.800	4
Q1-2	pretrained Alexnet after it is finetuned	0.670	**0.875**	17
Q1-3	non-pretrained Alexnet after it is trained	0.570	**0.670**	4

- Q1-5. Please try to correct the data augmentation strategy in order to let the entire face of each image be seen and report the validation accuracy of a pre-trained Alexnet as a feature extractor in the two-class classification problem. (5 pts)

在觀察過所有 training data set 之後, 我發現所有圖片中的 face 都位於圖片中心點附近. 因此我更改 data transform: RandomResizeCrop(隨機長寬比裁剪) 改成 transforms.CenterCrop(中心裁剪). 修改後的的驗證準確度達到 **0.83**. 比修改前準確度 0.80 來的高.

Fig-1: Image cropped before corrected the data augmentation strategy

Fig-2: Image cropped after corrected the data augmentation strategy

• Q1-6. Please try to correct the data augmentation strategy in order to let the entire face of each image be seen and report the validation accuracy of a pre-trained Alexnet after it is fine-tuned in the two-class classification problem. (5 pts)

承上一題，修改後的 pre-trained Alexnet after it is fine-tuned 驗證準確度達到 **0.8875**. 比修改前準確度 0.875 來的高.

```
Epoch 24/24
----------
train Loss: 0.0986 Acc: 0.9645
val Loss: 0.3527 Acc: 0.8850

Training complete in 4m 14s
Best val Acc: 0.887500
```

• Q1-7. Please discuss the results of Q1-5 & Q1-6. (5pts)

由於修改過 data augmentation 策略，在抓取圖像的時候，**避免了臉孔被截掉的現象**. 可以讓模型**專注於學習臉孔特徵，而非背景圖像**. 這位於臉孔分類器的訓練帶來效益.

Table-2: Validation accuracy benchmark of Q1-1~Q1-6

Question	Situation	Initial Val Acc (Epoch=1)	Best Val Acc	Epoch of Best Val Acc
Q1-1	pretrained Alexnet used as a feature extractor	0.7750	0.8000	4
Q1-2	pretrained Alexnet after it is finetuned	0.6700	0.8750	17
Q1-3	non-pretrained Alexnet after it is trained	0.5700	0.6700	4
Q1-5	pretrained Alexnet used as a feature extractor (correct the data augmentation strategy)	0.7725	0.8300	10
Q1-6	pretrained Alexnet after it is finetuned (correct the data augmentation strategy)	0.7670	**0.8875**	12

Q2: Semantic Segmentation

- Q2-1. Please try to “eliminate” the skip-connection so the output of convolution layers of FCN8s will be directly upsampled for 32x. Please report pixel accuracy and mIOU before and after. (10 pts)

在原有的 FCN8 模型上，經過訓練之後，我們得到最佳的 pixel accuracy 是 **0.846**，最佳的 mIOU 為 **0.436**. 這 2 項指標分別是在 epoch-17 與 20 所得到的表現..

在消除 skip-connection 之後，我們得到最佳的 pixel accuracy 為 **0.834**，最佳的 mIOU 為 **0.402**. 這 2 項指標分別是在 epoch-22 與 26 所得到的表現.

Fig-1: Training Result-FCN8s

```
The highest mIOU is 0.4362922363667073 and is achieved at epoch-20
The highest pixel accuracy  is 0.8460411071777344 and is achieved at epoch-17
```

Fig-2: Training Result-eliminate the skip-connection

```
The highest mIOU is 0.4027483643829762 and is achieved at epoch-26
The highest pixel accuracy  is 0.8344790649414062 and is achieved at epoch-22
```

- Q2-2. Please discuss the results of Q2-1. (10 pts)

原有的 FCN8s 模型藉由結合淺層學習所輸出的位置訊息，可以幫助模型學到更多的空間信息 (spatial location information). 但是在我們消除 skip-connection 之後，直接進行向上採樣(upsampled), 這反而會**丟失空間位置的資訊**. 進而降低模型學習的準確度 (mIOU: 0.436➔0.402).

- Q2-3. Please try to further reduce the number of classes from 11 to 3 and report the pixel accuracy & mIOU of FCN8s. (10 pts)

在減少類別項目(11 to 3)之後，模型訓練最佳的 mIOU 為 **0.478**, >0.436 (11 個類別的 mIOU). 而 pixel accuracy 為 0.324, < 0.8344.

Fig-3: Training Result-3 classes

```
The highest mIOU is 0.47844891725415795 and is achieved at epoch-14
The highest pixel accuracy  is 0.3241706848144531 and is achieved at epoch-20
```

- Q2-4. Please discuss the results of Q2-3. Was mIOU increased when the number of classes reduce? Please explain why! (10 pts)

在種類減少的情況下, mIOU 準確度會提高，主要是有 2 個原因: (1)由於種類減少，每個種類的影像數目提高，分類器得到的正確範例變多，這有助於提高分類器的學習效能. (2)種類越多，代表分類器要學習的特徵(feature)越多. 而每個種類的特徵可能彼此相關，這會造成 overfitting，降低分類器的準確度. 相反的種類越少，每個種類特徵的相關程度越低，這會提高分類器的準確度.