

CS655000 Computer Vision Homework 4

108062586 楊子儀

Q1. Two-Class Classification:

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

Training complete in 7m 25s Best val Acc: 0.790000	Training complete in 3m 15s Best val Acc: 0.797500	Training complete in 3m 16s Best val Acc: 0.810000
---	---	---

Average: 0.7991

Q1-2. Please report the validation accuracy of a pretrained Alexnet after it is fine tuned in the two class classification problem.

Training complete in 9m 39s Best val Acc: 0.885000	Training complete in 6m 52s Best val Acc: 0.890000	Training complete in 3m 57s Best val Acc: 0.885000
---	---	---

Average: 0.8866

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

Training complete in 3m 58s Best val Acc: 0.665000	Training complete in 3m 57s Best val Acc: 0.672500	Training complete in 4m 2s Best val Acc: 0.687500
---	---	--

Average: 0.675

Q1-4. Please discuss the results of Q1-1, Q1-2, & Q1-3.

從結果進行討論，重新 train 一個 model，其 validation 的準確率是最低的。而加入 fine tuned 的策略，可以得知除了作為分類的全連接層需要重 train，使得便是結果能符合自己的 dataset 外，若能適當地 fine tuned 一些 convolution layer，可以再提升辨識率。

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 pretrained Alexnet as a feature extractor in the two class classification problem.

Training complete in 3m 41s Best val Acc: 0.802500	Training complete in 3m 41s Best val Acc: 0.800000	Training complete in 3m 42s Best val Acc: 0.817500
---	---	---

Average: 0.8066

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 pretrained Alexnet after it is fine tuned in the two class classification problem.

Training complete in 4m 24s Best val Acc: 0.897500	Training complete in 4m 24s Best val Acc: 0.857500	Training complete in 4m 23s Best val Acc: 0.892500
---	---	---

Average: 0.8825

Q1-7. Please discuss the results of Q1-5 & Q1-6.

將 data augmentation，resize 後從影像中心位置做 crop，減少了切到人臉的機率，對於辨識率會有稍微的提升，但是否有使用 pretrained 的 weight 進行訓練，差別

會比較多，所以有 fine tuned 到 convolution 層的 model 效果仍會較好一些。

```
'train': transforms.Compose([
    # transforms.RandomResizedCrop(224),
    transforms.Resize(256),
    transforms.CenterCrop(224),
    transforms.RandomHorizontalFlip(),
```

Q1-8. Please try to achieve validation accuracy higher than 89.5% using a CNN other than Alexnet & ResNet 18 in the fine tuning case.

Training complete in 12m 35s Best val Acc: 0.912500	Training complete in 12m 36s Best val Acc: 0.912500	Training complete in 12m 36s Best val Acc: 0.902500
--	--	--

Net: VGG16

Average: 0.9091

Q1-9. Please discuss the results of Q1-9.

使用 data augmentation、pretrained model 進行 fine tuned 的策略，可以比重新訓練一個 model，獲得更高的 performance。而更複雜的架構，例如 VGG16，在辨識是否化妝的 data 上，也獲得更好的辨識結果。

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.

FCN8s	The highest mIOU is 0.46198667860324516 and is achieved at epoch-19 The highest pixel accuracy is 0.8583961486816406 and is achieved at epoch-20
	The highest mIOU is 0.45145309624861246 and is achieved at epoch-20 The highest pixel accuracy is 0.8493354797363282 and is achieved at epoch-20
	The highest mIOU is 0.46098138627104573 and is achieved at epoch-20 The highest pixel accuracy is 0.8605741882324218 and is achieved at epoch-19
	Average mIOU: 0.4580, Average pixel accuracy: 0.8560
FCN32s	The highest mIOU is 0.4064029535280595 and is achieved at epoch-20 The highest pixel accuracy is 0.8326734924316406 and is achieved at epoch-20
	The highest mIOU is 0.4031892052571331 and is achieved at epoch-20 The highest pixel accuracy is 0.8343037414550781 and is achieved at epoch-16
	The highest mIOU is 0.43002829503952084 and is achieved at epoch-19 The highest pixel accuracy is 0.8376283264160156 and is achieved at epoch-19
	Average mIOU: 0.4131, Average pixel accuracy: 0.8348

Q2-2. Please discuss the results of Q2-1.

FCN8s 相較於 FCN32s，可以切出更細部的細節。並且在 mIOU 和 pixel accuracy，FCN8s 也比較好。



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

The highest mIOU is 0.8250787798431164 and is achieved at epoch-19 The highest pixel accuracy is 0.9493478393554687 and is achieved at epoch-19
The highest mIOU is 0.8184026686166018 and is achieved at epoch-19 The highest pixel accuracy is 0.9473394775390624 and is achieved at epoch-15
The highest mIOU is 0.8170541870934019 and is achieved at epoch-19 The highest pixel accuracy is 0.9456251525878906 and is achieved at epoch-16

Average mIOU: 0.8201, Average pixel accuracy: 0.9474

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

mIOU 有顯著地提升。相較於原本 11 類，將車輛、行人、騎單車者合為一類，會使得 segmentation 後 cover 的面積變大，同時也降低辨識的複雜度。原本需再辨識街道、馬路等，會造成些微誤差，讓 mIOU 的值會比較低。

