

DLP hw3

1. Overview of your lab 3 (10%)

用 Unet 和 resnet-unet 的混和模型預測圖片中屬於貓貓狗狗的部分

2. Implementation Details (30%)

A. Details of your training, evaluating, inferencing code

在 training 中，我使用 `batchsize=8`, `learning rate=0.001` 訓練，optimizer 使用 `adam`，沒有用 `scheduler`，使用 `train` 數據。

Training 後通常會自動銜接 evaluating，在 evaluating 中，我使用相同的 `batchsize` 來測試，使用 `validate` 數據。

在 inferencing 中，程式會自動銜接 evaluating，使用 `validate` 數據

B. Details of your model (UNet & ResNet34_UNet)

Unet 我完全按照論文中的架構，並且參考了一些網路資源來寫。

在 unet 中，資料先 `encode`，依序經過雙卷積層與池化層各四次，雙卷積層是依序兩次卷積層、標準化層、`relu` 層，因為經過四層池化層，這時圖片的邊長變成原本的 $1/16$ 。

接者資料再經過一層雙卷積，然後對資料 `decode`，讓含有特徵的圖片依序經過 `nn.ConvTranspose2d` 放大圖片、與

encode 時圖片邊長相同的資料 `concat`、雙卷積各四次，這時圖片的大小已經復原，最後再經過一層 `1x1` 大小的卷積，讓通道數量變成 1 就是預測結果。

Resnet 部分因為論文裡寫的不太清楚，有些地方大小對不上，我只能盡量按照論文裡的描述去還原，resnet34 encoder 的部分完全按照論文裡的架，因此 encode 之後圖片大小會變成原來的 $1/32$ ，還有四份資料在 skip connection 時傳遞，decoder 部分如果完全按照 unet decoder 的架構來寫，不僅圖片邊長和通道數不對，decode 完後的圖片大小也不正確，因此我在最後不使用 `1x1` 卷積，而是多進行一次 `nn.ConvTranspose2d` 放大圖片、再使用雙卷積輸出。

C. Anything more you want to mention

因為 `cpu` 跑的慢，`cuda` 的記憶體不足，在使用 `cuda` 時 `batchsize` 不能太大，因此只能選擇 `batchsize=8`。也因為模型複雜，我缺乏足夠的時間應證在 `resnet` 上的改動好不好，希望模型上的一點小瑕疵不會影響整體的表現。

3. Data Preprocessing (20%)

A. How you preprocessed your data?

因為 `template` 中 `transport` 的位置很奇怪，我更改了 `transport` 的位置，並將 `image` 資料 `/255-0.5`，把範圍從

0~255 移動到-0.5~0.5，作為一種標準化。

B. What makes your method unique?

我另外寫了一些方法展示資料集中的圖片與預測結果，在訓練與測試後展示。

C. Anything more you want to mention

Transport 原始的位置會讓改動後的資料在之後被轉為 nparray，如果不更改，就要在其他地方把資料轉成 tensor，很不方便

4. Analyze on the experiment results (20%)

A. What did you explore during the training process?

我發現訓練過程有點不穩定，表現不一定直線上升，可惜缺乏效能做更深入的比較，我不能確定確切原因。

B. Found any characteristics of the data?

都是可愛的貓貓和狗狗，浮水印等雜訊有點多，有些圖片就算用肉眼也不容易分辨

C. Anything more you want to mention

5. Execution command (0%)

A. The command and parameters for the training process

```
& C:/Users/darren/miniconda3/envs/plzcuda/python.exe  
c:/Users/darren/Documents/大學作業/深度  
/lab3/Lab3_Binary_Semantic_Segmentation/src/train.py -e 5 -b 8  
-lr 0.001
```

B. The command and parameters for the inference process

假設檔案名稱是"unet_10.pt"

```
& C:/Users/darren/miniconda3/envs/plzcuda/python.exe  
c:/Users/darren/Documents/大學作業/深度  
/lab3/Lab3_Binary_Semantic_Segmentation/src/inference.py --  
model "unet_10.pt" -b 8
```

6. Discussion (20%)

A. What architecture may bring better results?

我認為 unet 比較好，unet 的大小與通道數量在各層都能對應，在 decode 時也比較不會因為通道數量太少而丟失特徵

B. What are the potential research topics in this task?

Encoder-decoder 架構與 Information Bottleneck

C. Anything more you want to mention

Unet 的程式碼都是我寫的，但我有參考網路上的一支影片，影片裡的對 unet 程式的實現方式解釋的很詳細，很不錯，影片連結做為參考資料放在下方：

<https://www.youtube.com/watch?v=IHq1t7NxS8k>