

HW 4 report

Reproductivity of the results

使用的是kaggle notebook，沒有用到本地裝置的特別環境或Util，主要就是在epoch, lr, weight_decay的default部分進行調整。資料路徑是原始的絕對路徑，沒有進行更動。

FCN epoch = 70, lr = 1e-3, weight_decay = 1e-5, batch size = 4

UNet epoch = 70, lr = 1e-4, weight_decay = 1e-5, batch size = 2

Number of parameters

FCN : 14.72M

UNet : 13.39M

The difficulty during training

最難的部分是在自己的notebook上的Dice值很高，但是丟到競賽上的時候會將的比預期的低，花了一些時間做參數或模型的調整，再來就是選擇模型上的困難，不過丟個幾次就發現UNet的表現比FCN來得好，可能在FCN的模型上就需要做更多的調整來更好捕捉特徵。

再來就是在Wandb上的一些產出結果的理解，然後再回過頭去調整args裡的參數，比起之前的結果理解，這裡花了一些時間去理解根據給出的結果，如何去調整args裡的哪些參數，雖然最後真正動到的參數好像比我所想的少，可能是因為對於出來的結果的理解部分還不是很熟到可以根據他們去進行校正，可能之後多實作幾次後會更容易懂得如何去調整參數值。

另外在建UNet上時，直接看著圖去建比預期的要難一點，但參考一些資源後，對於在upsampling的input size與output size比較容易好理解，也比較好寫upsampling這裡，比起前面幾次作業，模型降維到最後只給出分類，這次再升維度部分耗費較多的時間去理解。

Model architecture

1. FCN

包含五個卷積塊，分別命名為 block1 到 block5。每個卷積塊包括多個卷積層和一個最大池化層。模型的前半部分通過這些卷積塊提取輸入圖像的高層次特徵。然後，模型通過反卷積層（或稱為轉置卷積層）將特徵圖上採樣，逐步恢復到原始輸入圖像的尺寸。最終，模型的輸出是一個與輸入圖像相同大小的 feature map，其中每個像素的值表示該像素屬於哪一個類別。

Encoder 部分（卷積塊 block1 到 block5）：

每個卷積塊包含多個卷積層，其中包括卷積核、激活函數（這裡未顯式指定，通常使用 ReLU）、和最大池化層。這些卷積層用於提取圖像的特徵。

block1 到 block5 依次提取不同層次的特徵，隨著層次的增加，感受野也增大。

Decoder 部分（反卷積層和預測層）：

transpose5、transpose4 和 transpose3 分別代表了 block5、block4 和 block3 的預測結果。這些預測結果透過反卷積層進行上採樣，以恢復到輸入圖像的尺寸。

最後，這些預測 feature map 通過相加，得到 addition_upsample。total_upsample 層再對 addition_upsample 進行進一步的上採樣，得到最終的分割結果 output。

2. UNet

由兩個主要部分組成：編碼器（Encoder）和解碼器（Decoder）。以下是 U-Net 模型的主要結構：

DoubleConv 模塊：

由兩個連續的卷積層組成，每個卷積層後面跟隨一個 ReLU 激活函數。用於構建網絡中的卷積層基本結構。

Down 模塊：

包含最大池化層和一個 DoubleConv 模塊。最大池化用於降低特徵圖的空間尺寸，而 DoubleConv 用於提取特徵。

Up 模塊：

包含上採樣層（使用 bilinear 插值進行上採樣）和一個 DoubleConv 模塊。上採樣增加特徵圖的空間尺寸，並且通過與來自編碼器的相應特徵圖的連接，實現跳躍式連接（skip connections）。

UNet 模型：

Encoder 部分：

包括三個 Down 模塊，每個 Down 模塊都將特徵圖的尺寸減半。

從64通道開始，透過 Down 模塊的連續應用，通道數逐漸增加。

在每個 Down 模塊之後，特徵圖的尺寸減半，通道數加倍。

Bottleneck（瓶頸）部分：

包括三個 DoubleConv 模塊，用於捕捉最深層的特徵。

Decoder 部分：

包括三個 Up 模塊，每個 Up 模塊都將特徵圖的尺寸加倍。

在每個 Up 模塊中，將來自 Encoder 的相應特徵圖與上採樣的特徵圖進行連接，實現跳躍式連接。

Output 部分：

包括一個 1x1 的卷積層，將最終的特徵圖映射為模型的輸出。這裡輸出通道數為 c_{out} ，通常對應於分割的類別數量。