

Deep Learning Lab4 實驗討論

資工四乙 407262316 高挺恩

一、模型參數調整與準確率提升

這次實驗有著重在 convolutional 和 pooling layer 的嘗試，起初設計是各兩層的 convolutional 和 pooling layer，不過 filters 數量似乎給的有點小，導致訓練上準確率最好就只到 70% 左右，後來試著把 filters 的數量提高，讓整體的 parameters 數量大一點(最終約莫 30w)，如此一來訓練準確率也能在 20 代之內達到 98%。

不過，在驗證準確率上都會稍稍出現 overfitting 的情況，在 full-connected layer 下手調整，目前是有四層 fc layer，驗證準確率可以達 70%，測試準確率也能達到相同水準，就只是 loss 的部分會有稍稍的上下起伏。

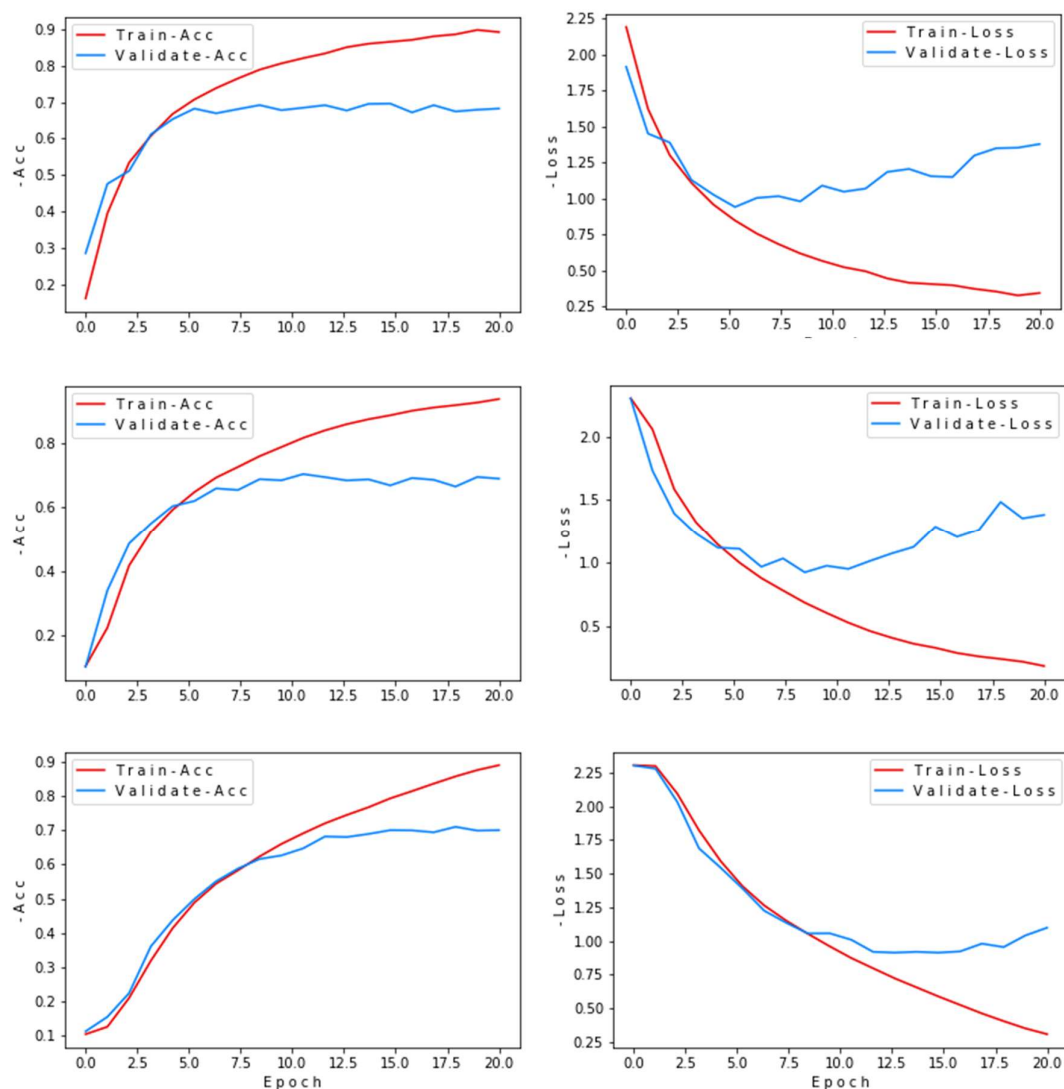
二、不同學習率的表現

由於目前的 batch size 設定為 20，照前一個 lab3 的分析討論，我同樣將其視為算是比較小的資料量去更新，所以學習率的部分同樣不設太大，一開始只保守給 0.1，可以得到 67%左右的測試準確率，經過幾次嘗試，發現學習率更低的話可以得到更好的結果，於是目前的學習率設定為 0.015，也成功讓測試準確率突破 70%大關啦!

話雖如此，但不能說比較大的學習率就比較不好，應該還是要搭配多少資料量更新這項條件去設定才好，而且學習率比較低收斂也要比較花時間，這部分還是一個很大的學問阿!

三、圖示分析

下圖分別是學習率 0.1、0.05、0.015 的情形，都是以 `cnn.py` 去執行產生的。



可以清楚的看到，在訓練部分的準確率都算是不錯的，但在較小的學習率上總是可以得到比較平滑的 loss 曲線，更能驗證說剛剛上述第二點的內容，是有一定的可信度的。

四、其他心得

這次 lab4 雖然一部分的分數只需要透過填空就可以得到，不過這也何嘗不是一個在短時間內學習的機會呢？尤其是剩下的 20%分數我認為才是一個挑戰的地方！包括自己寫寫看 dataset 的 class、自己找 dataloader 與 dataset 之間的關係等等，以及許多參數的測試，都真的很有做實驗的感覺，在優化器的部分我嘗試加入了 momentum 這個動量參數去動態輔助學習率，結果反而更差了，或許對於比較不複雜的資料集還用不到吧。

但同時也發覺到，電腦效能不足的無奈，老實說這份 lab4 我不是在自己電腦上跑，而是在 Colab 上執行，但它提供的 GPU 也是跑很久阿~ 要踏入這個領域，你不是很有錢、就是很有閒 XD