

Deep Learning HW1_2 Report

0853411 劉書維

1. Please describe in details how to preprocess images because of the different resolution images and various bounding boxes region in Medical Masks dataset and explain why.

我是先將 csv 檔讀入，再依照檔案中的 filename 開啟相關圖片，使用的工具是 cv2，來協助讀取灰階圖片，並擷取 ymin、ymax、xmin、xmax 的 bound，後調整大小成 28*28，之後 reshape 成一維 784 的 uint8 資料，最後轉換成 float32 的資料型態，就完成了圖片處理的部分。

而 label 的處理就較簡單，將 good 資料設置成 0，bad 設置成 1，none 設置成 2，最後再轉成 array 形式，就完成 label 資料轉換了。

這些資料轉換都是為了使用之後的 tensorflow 的 CNN 架構的資料前處理。

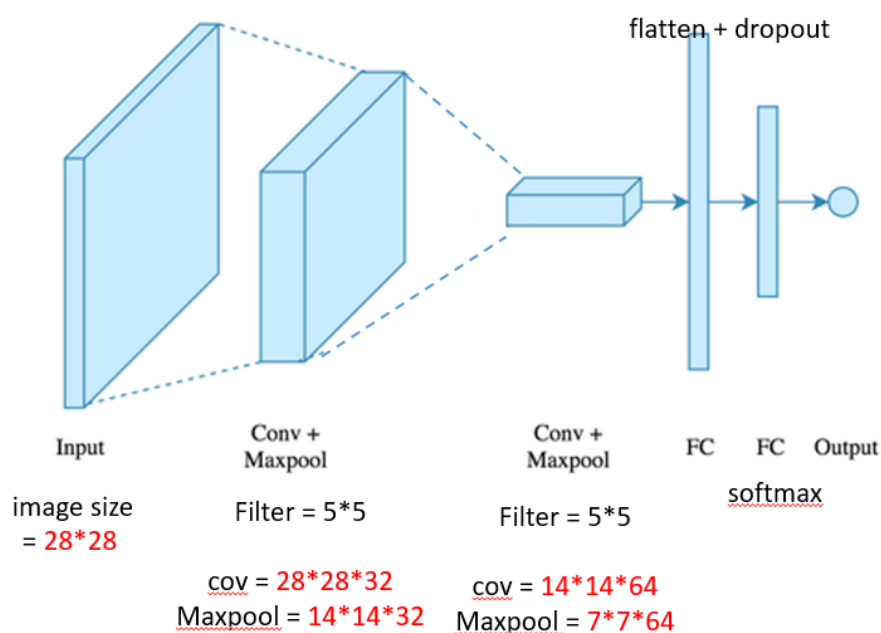
2. Please implement a CNN for image recognition by using Medical Masks dataset. You need to design the network architecture, describe your network architecture and analyze the effect of different settings including stride size and filter size.

Model : CNN (Tensorflow)

我的 network 是使用兩層的 Convolution 層和 maxpooling 來增加特徵值

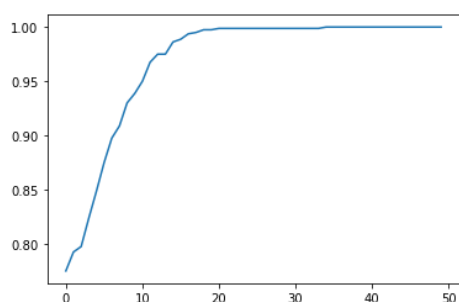
與降維，之後再加上 flatten 和 dropout，最後 softmax 來完成資料分類。

結構說明如下圖：

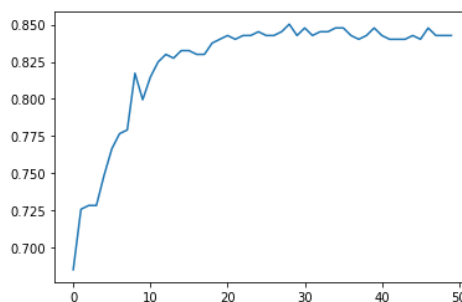


所以在這個架構下，參數設置為 1000 次的 iterations，每 20 次 output 一個 train accuracy 與 test accuracy，結果分別是在 1 與 0.85 收斂，似乎有點 overfitting。

Train accuracy rate：



Test accuracy rate：

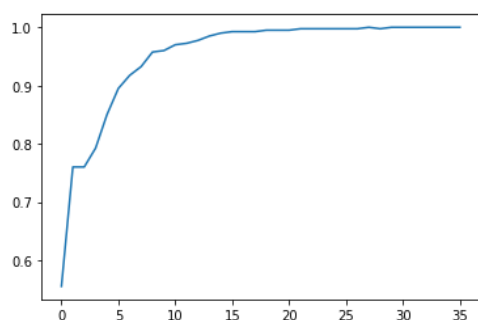


在 tensorflow，padding 有兩個選項「padding = VALID」和「padding = SAME」。padding = VALID 等於最一開始敘述的卷積計算，圖根據 filter 大小和 stride 大小而變小；padding = SAME，會用 zero-padding 的手法，讓輸入的圖不會受到 kernel map 的大小影響。所以可以協助我們進行以下討論。

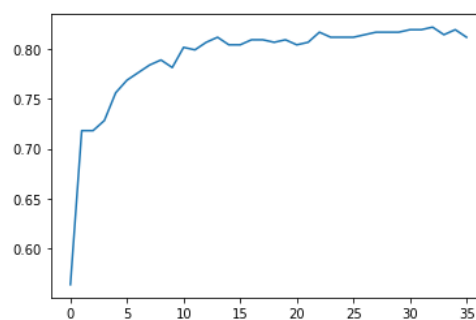
討論一：若 stride size 是 2*2

會使圖形變得更小，所以可以加速運算，但是精準度卻沒有比 1*1 來的高，但是也沒差很多。Train accuracy rate 一樣是在 1.0 收斂，也有提早收斂的趨勢（在 iter=20*20 收斂），但是 test accuracy rate 大致只有 0.82，不如 1*1 時的 0.85。

Train accuracy rate：

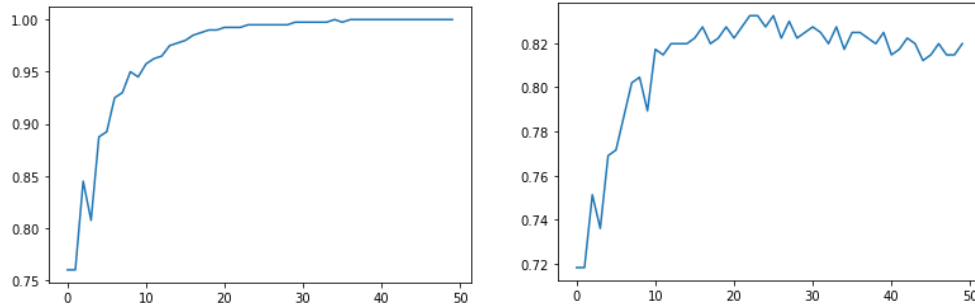


Test accuracy rate：



討論二：若 filter size 是 3*3

要是我們用 3*3 的 filter 的話，Train accuracy rate 一樣是在 1.0 收斂，但是 test accuracy rate 大致只有 0.83，不如 5*5 時的 0.85，而且 test accuracy 的波動也較大，不太穩定。

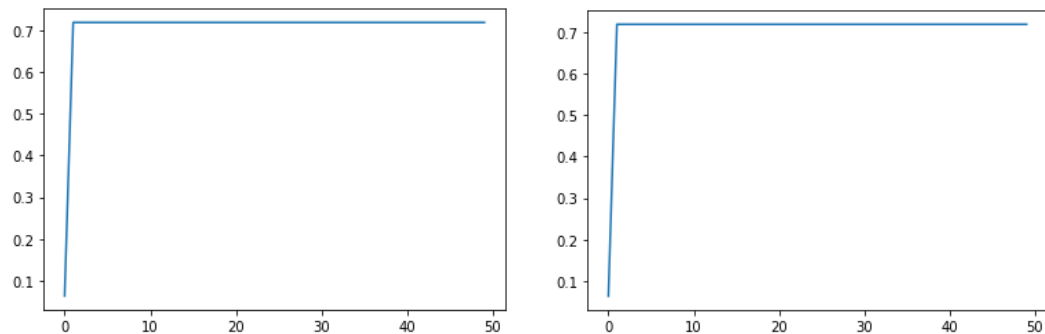


討論：若 filter size 是 7*7

若將 filter size 改成 7*7，則會有很快收斂，且兩者準確度都只有 0.7。

Train accuracy rate：

Test accuracy rate：



3. Show some examples of classification result, list your accuracy of each classes for both training and test data, and answer the following questions.

Training data 是完全分類正確，Test data 正確率與分類如下整理：

```
total acc: 0.8451776649746193
good label acc: 0.9646643109540636
bad label acc: 0.6629213483146067
none label acc: 0.045454545454545456
```

後面的比例為應標記該分類且標記正確的比例，可以發現 none label 的正

確率極低。可能的解決方法可以從資料下手，可以多增加一些額外 none 的

train 資料來增加該分類的準確率，另外也可以考慮多增加一層 conv + maxpooling 的 layer 來尋找更多特徵值（雖然會跑得更久）。