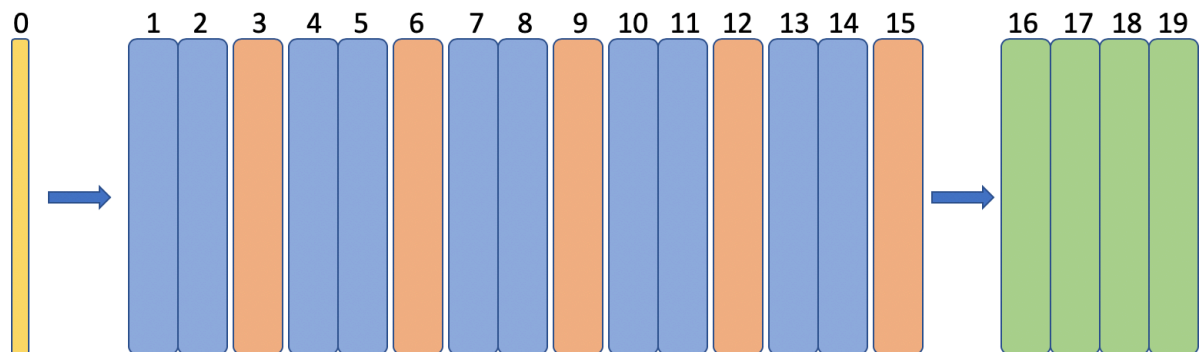


Machine Learning Spring 2020 - HW3 Report

學號:B07902064 系級:資工二 姓名:蔡銘軒

1. 請說明你實作的 CNN 模型(best model)，其模型架構、訓練參數量和準確率為何？(1%)

模型架構：



0: input (192x192x3) / 1: Conv2d (kernel 3x3, channel 64) / 2: Conv2d (kernel 3x3, channel 128)

3: Maxpool (2x2) / 4: Conv2d (kernel 3x3, channel 256) / 5: Conv2d (kernel 3x3, channel 256)

6: Maxpool (2x2) / 7: Conv2d (kernel 3x3, channel 256) / 8: Conv2d (kernel 3x3, channel 256)

9: Maxpool (2x2) / 10: Conv2d (kernel 3x3, channel 512) / 11: Conv2d(kernel 3x3, channel 512)

12: Maxpool (2x2) / 13: Conv2d (kernel 3x3, channel 512) / 14: Conv2d(kernel 3x3, channel 512)

15: Maxpool (2x2) / 16: Linear (512x6x6, 2048) / 17: Linear (2048, 2048)

18: Linear (2048, 1024) / 19: Linear (1024, 11)

每層Conv2d後面都接著LeakyReLU以及BatchNorm2d

每層Maxpool後面都有一層Dropout($p = 0.2$)

除了最後一層Linear (11) 為直接輸出，每層Linear後面都接著ReLU、BatchNorm1d、Dropout($p = 0.5$)

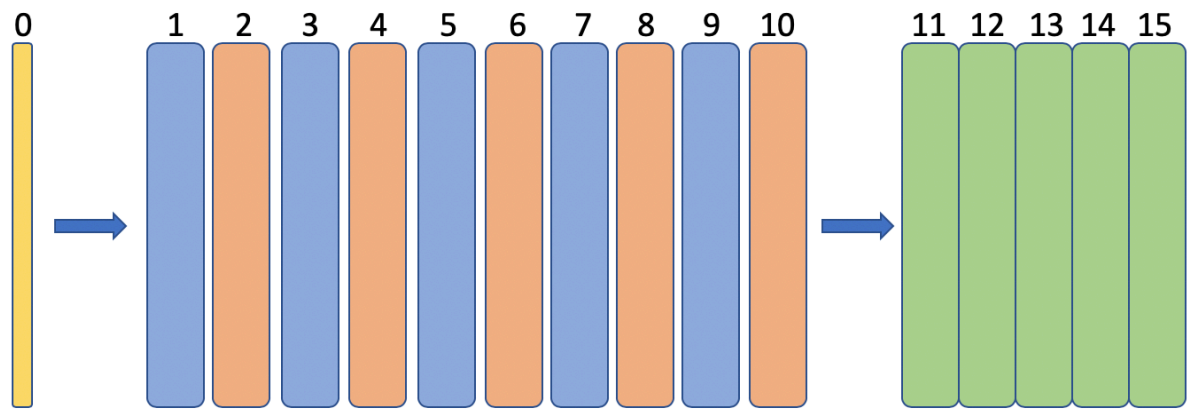
訓練參數量：54473995

準確率（觀察至收斂為止）：

	Accuracy	Loss
Training set	0.9115190365190365	0.008224935081976277
Validation set	0.836	0.01877435168127219

2. 請實作與第一題接近的參數量，但 CNN 深度（CNN 層數）減半的模型，並說明其模型架構、訓練參數量和準確率為何？(1%)

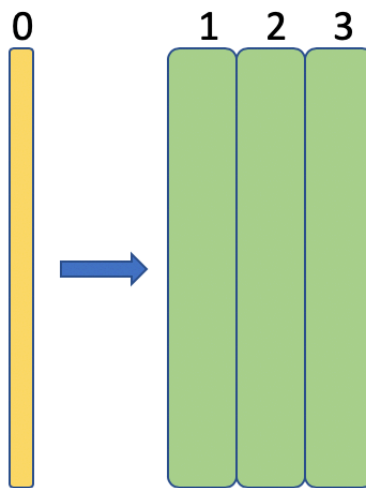
模型架構：



0: input (192x192x3) / 1: Conv2d (kernel 3x3, channel 64) / 2: Maxpool (2x2)
3: Conv2d (kernel 3x3, channel 256) / 4: Maxpool (2x2) / 5: Conv2d (kernel 3x3, channel 256)
6: Maxpool (2x2) / 7: Conv2d (kernel 3x3, channel 512) / 8: Maxpool (2x2)
9: Conv2d (kernel 3x3, channel 512) / 10: Maxpool (2x2) / 11: Linear (512x6x6, 2048)
12: Linear (2048, 2048) / 13: Linear (2048, 2048) / 14: Linear (2048, 1024) / 15: Linear (1024, 11)
每層Conv2d後面都接著LeakyReLU以及BatchNorm2d
每層Maxpool後面都有一層Dropout($p = 0.2$)
除了最後一層Linear (11) 為直接輸出，每層Linear後面都接著ReLU、BatchNorm1d、Dropout($p = 0.5$)
訓練參數量：52550027
準確率（觀察至收斂為止）：

	Accuracy	Loss
Training set	0.802933177933178	0.01819044312211832
Validation set	0.7856666666666666	0.021206362515687943

3. 請實作與第一題接近的參數量，簡單的 DNN 模型，同時也說明其模型架構、訓練參數和準確率為何？(1%)
- 模型架構：



0: input (192x192x3) / 1: Linear (192x192x3, 512) / 2: Linear (512, 32) / 3: Linear (32, 11)

除了最後一層Linear (11) 為直接輸出，每層Linear後面都接著ReLU、BatchNorm1d、Dropout ($p = 0.5$)

訓練參數量：56641483

準確率（觀察至收斂為止）：

	Accuracy	Loss
Training set	0.23251748251748253	0.06663037540149393
Validation set	0.273	0.09918265513579051

4. 請說明由 1 ~ 3 題的實驗中你觀察到了什麼？(1%)

1~3題逐漸地將Convolutional layer的數量漸少，並且用Fully connected layer代替。雖然參數的數量相近，但表現卻差了不少。

第三題在完全沒有Convolutional layer的情況下，準確率非常的低。可能的解釋是Convolutional layer比較能捕捉圖片裡相鄰像素之間的關聯，而Fully connected layer將每個pixel當成獨立的輸入，比較難捕捉到圖片裡面的各種資訊（例如邊、輪廓、形狀）等。

第二題減少了Convolutional layer的數量後準確率也稍微降低，可能是較少的Convolutional layer沒有捕捉到圖片裡比較多的細節。而Fully connected layer在夠寬的情況下，疊1~2層就可以有很好的能力，而我在best model裡已經疊了4層，在第二題繼續增加Fully connected layer的數量（用來彌補Convolutional layer裡少掉的參數量）可能就沒有太大的進步。推測重點還是在Convolutional layer捕捉圖片細節的能力，Fully connected layer夠用就好了。

5. 請嘗試 data normalization 及 data augmentation，說明實作方法並且說明實行前後對準確率有什麼樣的影響？(1%)

(Train Acc / Validation Acc)	with data normalization	without data normalization
with data augmentation	0.8144910644910645 / 0.7903333333333333	0.8111888111888111 / 0.8003333333333333
without data augmentation	0.8744483294483295 / 0.6266666666666666	0.8641219891219891 / 0.6046666666666667

因為best model train一輪需要較長的時間，這題使用比較簡易的CNN model並train 50個epoch來比較差異。

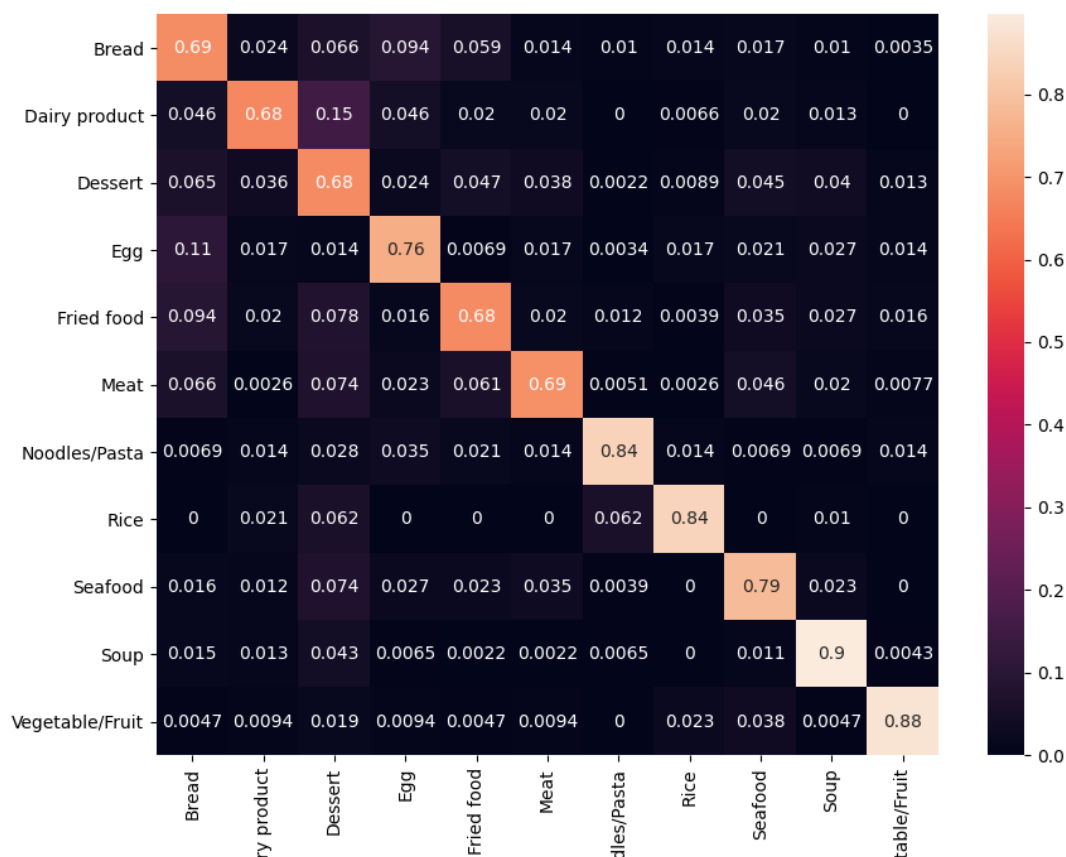
在這次的實驗中，沒有data augmentation的兩個model在50個epoch後，validation set上的表現已經收斂了。只有training set的Acc繼續提升，因此可以視為model已經train完了（Early stopping）。

加上data augmentation的model，Acc上升的速度比較慢，到50個epoch結束時，training set跟validation set的Acc都還有微幅上升的跡象，但已經開始出現震盪（Acc從80%掉到60%，再慢慢爬回80%等等現象），且Acc的表現已經足夠拿來比較，因此停在50個epoch。

從結果來看，data normalization在這次的模型中沒有太大的影響，Acc的差距並沒有非常顯著。在我的best model裡，加上data normalization甚至使Acc略為下降1%~2%，在這個較簡易的model裡也有類似的現象。

影響較大的是data augmentation。雖然加上data augmentation後，training set上的Acc上升較慢，在50個epoch結束時Acc跟沒有data augmentation的model比起來有一段差距，但validation set上的表現明顯提升。這個結果符合一開始選擇使用data augmentation的原因，可以讓training set似乎多了更多data，避免model overfit的問題。

6. 觀察答錯的圖片中，哪些 class 彼此間容易用混？繪出 confusion matrix 分析(1%)



從confusion matrix來看，比較明顯的錯誤是會把Dairy product分類成Dessert。另外比較沒那麼嚴重的是Rice, Seafood也比較容易分類成Dessert。整體而言Soup的分類正確率有90%為最高，Rice的正確率雖然較低，但錯誤大多集中在分類成Dessert或是Dairy product，比較少誤認成其他的項目。而最容易分類錯誤的是Fried Food，正確率68%，主要容易分類成Bread以及Dessert。