

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%)

我的CNN 架構是由7層的convolution layer 最後再接上2048 X 1024 X 512 X 1 1的fully connected feedforward layer。每層convolution layer 完都會讓他再經過batch normalization 跟 PReLU function，唯一不同的是第一層我maxpooling 從(2,2,0)改成(2,1,1)，因此圖片的維度從128變成127。每層 fully connected feedforward layer 之後都會經過PReLU。Data augmentation的部分，我讓圖片在(-45度,45度)隨機旋轉、隨機翻轉以及隨機做仿射變換。訓練的參數量約為892,0852，epoch數量定為100，optimizer 是Adam，lr為0.001，而準確率約為0.83263。

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%)

此減半層數的模型架構與原本的CNN模型相似，僅拿掉了前三層的convolution layer，並在第二層和第三層maxpooling從(2,2,0)改為(4,4,0)，訓練參數量為747,0161，而準確率約為0.77824。

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%)

我使用第一題原本CNN的模型，然後將convolution layer全數拿掉，而為了使參數量相近，將DNN架構寫成

```
nn.Linear(3*128*128, 128)
nn.PReLU()
nn.Linear(128, 512),
nn.ReLU()
nn.Linear(512, 1024)
nn.PReLU()
```

```
nn.Linear(1024, 1024),  
nn.PReLU()  
nn.Linear(1024, 1024)  
nn.PReLU()  
nn.Linear(1024, 11)  
參數量為899,3424，而準確率為0.34130
```

4. 請說明由 1 ~ 3 題的實驗中你觀察到了什麼？(1%)

可以看到CNN層數多寡，會明顯的影響準確率的表現。我們拿掉一半的convolution layer就會造成準確率往下掉0.1多。當我們把CNN layer全部拿掉之後，單用DNN來訓練，準確率只有0.3初頭。雖然DNN看似表現不太好，但是還是比random盲猜的準確率還要高(0.09)

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%)

實作：

Data Normalization: transforms.normalization將圖片數據轉換成0~1

Data Augmentation:

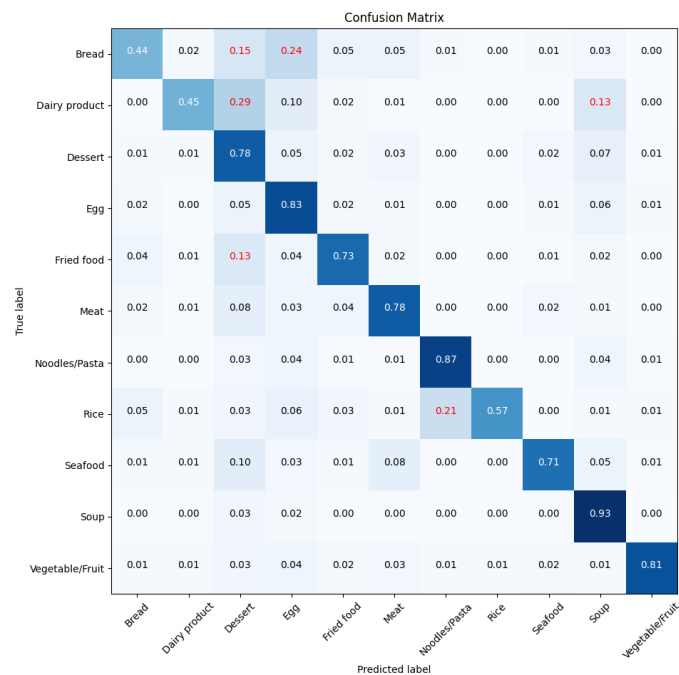
```
transforms.RandomHorizontalFlip(),
```

```
transforms.RandomRotation(45),
```

```
transforms.RandomAffine( degrees = 0, translate = (0.05, 0.05), scale = (0.9,  
1.1), shear = 5, resample = False, fillcolor = 0),
```

經過實驗後發現，不管是做data normalization 或是 data augmentation都會使得準確率提高(0.592 和 0.614)。當兩個同時實作時，會有更明顯的提(0.683)。

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%)



Bread 和 Dessert

Bread和Egg

Dairy Product 和 Dessert

Dairy Product 和 Soup

Fried Food 和 Dessert

Rice 和Noodles/Pasta

這幾類比較容易搞混，從圖片中可以看到我的model比較不會判斷Dessert，很容易把其他類的食物誤認成Dessert