ML HW3 Report

學號：R08942025　系級：　電信碩一　姓名：徐瑋辰

1. (1%) 請說明這次使用的 model 架構，包含各層維度及連接方式。

沒有使用 Pretrained Model
採用 4 層 convolution layers 及 2 層 connected layers
Convolution layer: 增加 channels, leakyReLu, BatchNorm, MaxPool
Connected layer: 3*3*256 -> 256 -> 7

完整架構如下：

Convolution: 共 4 層

```
self.conv1 = nn.Sequential(
        nn.Conv2d(1, 32, kernel_size=5, padding=2),
        nn.LeakyReLU(negative_slope=0.05),
        nn.BatchNorm2d(32),
        nn.MaxPool2d(2),
)
self.conv2 = nn.Sequential(
        nn.Conv2d(32, 128, kernel_size=5, padding=2),
        nn.LeakyReLU(negative_slope=0.05),
        nn.BatchNorm2d(128),
        nn.MaxPool2d(2),
)
self.conv3 = nn.Sequential(
        nn.Conv2d(128, 256, kernel_size=3,padding=1),
        nn.LeakyReLU(negative_slope=0.05),
        nn.BatchNorm2d(256),
        nn.MaxPool2d(2),
)
self.conv4 = nn.Sequential(
        nn.Conv2d(256, 256, kernel_size=3,padding=1),
        nn.LeakyReLU(negative_slope=0.05),
        nn.BatchNorm2d(256),
        nn.MaxPool2d(2),
)
```

Connect: 2 層

```
self.fc = nn.Sequential(
        nn.Linear(3*3*256, 256),
        nn.ReLU(),
        nn.BatchNorm1d(256),
        nn.Linear(256, 7)
)
```

2. **(1%) 請附上 model 的 training/validation history (loss and accuracy)。**

Epoch: 1, train Loss: 1.5984, train Acc: 0.3964, time: 7.4132
Epoch: 1, valid Loss: 1.3875, valid Acc: 0.4666, time: 0.8540
Epoch: 2, train Loss: 1.2333, train Acc: 0.5371, time: 7.2585
Epoch: 2, valid Loss: 1.2376, valid Acc: 0.5365, time: 0.8645
Epoch: 3, train Loss: 1.0538, train Acc: 0.6070, time: 7.2537
Epoch: 3, valid Loss: 1.2193, valid Acc: 0.5430, time: 0.8485
Epoch: 4, train Loss: 0.9102, train Acc: 0.6641, time: 7.2434
Epoch: 4, valid Loss: 1.2850, valid Acc: 0.5458, time: 0.8627
Epoch: 5, train Loss: 0.7247, train Acc: 0.7380, time: 7.2971
Epoch: 5, valid Loss: 1.2796, valid Acc: 0.5723, time: 0.8565
Epoch: 6, train Loss: 0.5270, train Acc: 0.8124, time: 7.2789
Epoch: 6, valid Loss: 1.4342, valid Acc: 0.5620, time: 0.8625
Epoch: 7, train Loss: 0.3662, train Acc: 0.8738, time: 7.2699
Epoch: 7, valid Loss: 1.6603, valid Acc: 0.5687, time: 0.8470
Epoch: 8, train Loss: 0.2276, train Acc: 0.9271, time: 7.1904
Epoch: 8, valid Loss: 1.8062, valid Acc: 0.5716, time: 0.8469
Epoch: 9, train Loss: 0.1429, train Acc: 0.9558, time: 7.2065
Epoch: 9, valid Loss: 1.8379, valid Acc: 0.5749, time: 0.8488
Epoch: 10, train Loss: 0.1093, train Acc: 0.9676, time: 7.2145
Epoch: 10, valid Loss: 2.0112, valid Acc: 0.5748, time: 0.8678
Epoch: 11, train Loss: 0.0765, train Acc: 0.9799, time: 7.2608
Epoch: 11, valid Loss: 2.0336, valid Acc: 0.5803, time: 0.8496
Epoch: 12, train Loss: 0.0550, train Acc: 0.9863, time: 7.1905
Epoch: 12, valid Loss: 2.1993, valid Acc: 0.5785, time: 0.8469
Epoch: 13, train Loss: 0.0390, train Acc: 0.9898, time: 7.2256
Epoch: 13, valid Loss: 2.1998, valid Acc: 0.5839, time: 0.8590
Epoch: 14, train Loss: 0.0302, train Acc: 0.9937, time: 7.2346
Epoch: 14, valid Loss: 2.1729, valid Acc: 0.5849, time: 0.8568
Epoch: 15, train Loss: 0.0267, train Acc: 0.9941, time: 7.3055
Epoch: 15, valid Loss: 2.3041, valid Acc: 0.5876, time: 0.8665
Epoch: 16, train Loss: 0.0156, train Acc: 0.9962, time: 7.3128
Epoch: 16, valid Loss: 2.3134, valid Acc: 0.5934, time: 0.8678
Epoch: 17, train Loss: 0.0116, train Acc: 0.9969, time: 7.2546
Epoch: 17, valid Loss: 2.2637, valid Acc: 0.5973, time: 0.8496
Epoch: 18, train Loss: 0.0090, train Acc: 0.9969, time: 7.2907
Epoch: 18, valid Loss: 2.2851, valid Acc: 0.5948, time: 0.8649
Epoch: 19, train Loss: 0.0083, train Acc: 0.9973, time: 7.3089
Epoch: 19, valid Loss: 2.2603, valid Acc: 0.5962, time: 0.8645
Epoch: 20, train Loss: 0.0068, train Acc: 0.9970, time: 7.3096
Epoch: 20, valid Loss: 2.3338, valid Acc: 0.5943, time: 0.8651
Epoch: 21, train Loss: 0.0063, train Acc: 0.9975, time: 7.3086
Epoch: 21, valid Loss: 2.3154, valid Acc: 0.5998, time: 0.8578
Epoch: 22, train Loss: 0.0071, train Acc: 0.9974, time: 7.2333
Epoch: 22, valid Loss: 2.3576, valid Acc: 0.5941, time: 0.8548
Epoch: 23, train Loss: 0.0058, train Acc: 0.9975, time: 7.2429
Epoch: 23, valid Loss: 2.3689, valid Acc: 0.5947, time: 0.8537

Epoch: 24, train Loss: 0.0063, train Acc: 0.9973, time: 7.2210
Epoch: 24, valid Loss: 2.3302, valid Acc: 0.5902, time: 0.8565
Epoch: 25, train Loss: 0.0073, train Acc: 0.9973, time: 7.2786
Epoch: 25, valid Loss: 2.3530, valid Acc: 0.5960, time: 0.8539
Epoch: 26, train Loss: 0.0058, train Acc: 0.9976, time: 7.2915
Epoch: 26, valid Loss: 2.3955, valid Acc: 0.5813, time: 0.8641
Epoch: 27, train Loss: 0.0070, train Acc: 0.9974, time: 7.2112
Epoch: 27, valid Loss: 2.3125, valid Acc: 0.5891, time: 0.8485
Epoch: 28, train Loss: 0.0083, train Acc: 0.9971, time: 7.2249
Epoch: 28, valid Loss: 2.5378, valid Acc: 0.5959, time: 0.8671
Epoch: 29, train Loss: 0.0117, train Acc: 0.9972, time: 7.2652
Epoch: 29, valid Loss: 2.5065, valid Acc: 0.5831, time: 0.8571
Epoch: 30, train Loss: 0.0088, train Acc: 0.9969, time: 7.2287
Epoch: 30, valid Loss: 2.5097, valid Acc: 0.5919, time: 0.8541

判斷條件為：train acc > 0.99 的模型中，取 valid loss 較小者

3. (1%) 畫出 confusion matrix 分析哪些類別的圖片容易使 model 搞混，並簡單說明。

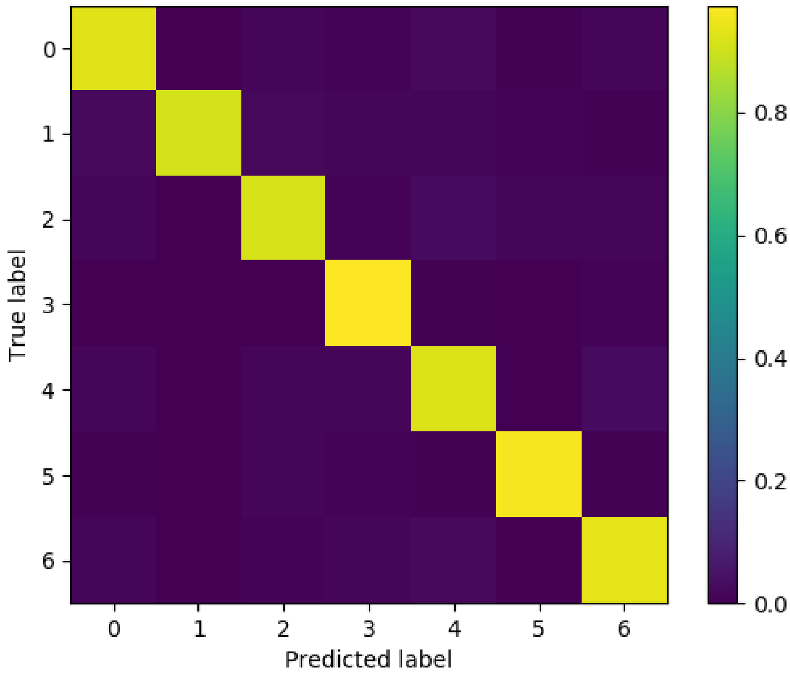
橫軸為 predict label, 縱軸為 true label
Label: (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral)

3755	3	55	35	100	19	74
11	408	9	6	7	4	2
64	1	3801	46	120	58	55
26	0	25	7099	50	27	57
78	3	76	60	4442	15	140
14	1	39	29	16	3050	23
61	1	40	64	122	17	4680

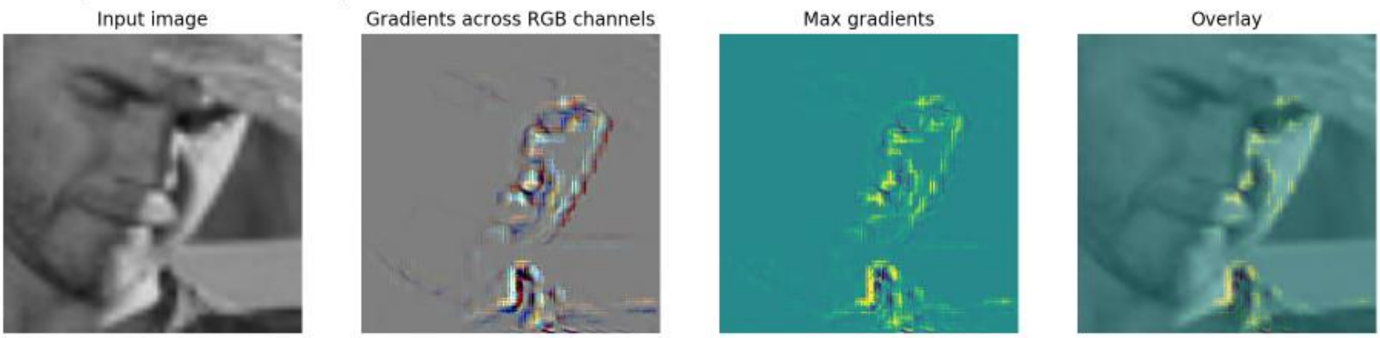
0.9292	0.0007	0.0136	0.0087	0.0247	0.0047	0.0183
0.0246	0.9128	0.0201	0.0134	0.0157	0.0089	0.0045
0.0154	0.0002	0.917	0.0111	0.029	0.014	0.0133
0.0036	0	0.0034	0.9746	0.0069	0.0037	0.0078
0.0162	0.0006	0.0158	0.0125	0.9227	0.0031	0.0291
0.0044	0.0003	0.0123	0.0091	0.005	0.9615	0.0073
0.0122	0.0002	0.008	0.0128	0.0245	0.0034	0.9388

綜合上表，可顯示出易被搞混的類別如下：
Angry -> 被判成 Sad
Disgust -> 被判成 Angry, Fear
Sad -> 被判成 Neutral
Neutral -> 被判成 Sad

Confusion Matrix 繪製出的圖如下

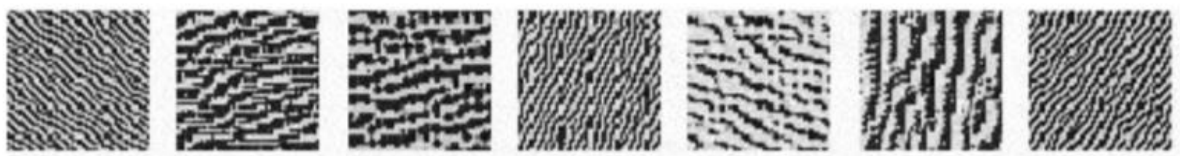


4. (1%) 畫出 CNN model 的 saliency map，並簡單討論其現象。



在 saliency map 中，人的面部表情特徵被提取出來，方便做進一步的分類。而這也表示圖中的輪廓部分對於分類的重要性較高。

5. (1%) 畫出最後一層的 filters 最容易被哪些 feature activate。



上排為不同 filter，下排為同一張圖片經 filter 的 output 結果

6. **(3%)Refer to math problem**

Convolution

input shape : (B. W. H input_ channels)

Conv2D (input_ channels, output_ channels, kernal_ size = (k_1, k_2),
stride = (s_1, s_2), padding = (p_1, p_2)):

(B. W, H. input_ channels) $\xrightarrow{\text{padding}}$ (B. $W+2p_1$, $H+2p_2$, input_channels)

$\xrightarrow{\text{kernel. stride}}$ (B, $\frac{W+2p_1-k_1}{s_1}+1$, $\frac{H+2p_2-k_2}{s_2}+1$, output_ channels) ✗

Batch Normalization

$$\frac{\partial l}{\partial \hat{x}_i} = \frac{\partial l}{\partial y_i} \cdot \gamma$$

$$\frac{\partial l}{\partial \sigma_B^2} = \sum_{i=1}^{m} \frac{\partial l}{\partial \hat{x}_i} (x_i - \mu_B) \cdot \frac{-1}{2} (\sigma_B^2 + \epsilon)^{-3/2}$$

$$\frac{\partial l}{\partial \mu_B} = \left(\sum_{i=1}^{m} \frac{\partial l}{\partial \hat{x}_i} \cdot \frac{-1}{\sqrt{\sigma_B^2 + \epsilon}} \right) + \frac{\partial l}{\partial \sigma_B^2} \frac{\sum_{i=1}^{m} -2(x_i - \mu_B)}{m}$$

$$\frac{\partial l}{\partial x_i} = \frac{\partial l}{\partial \hat{x}_i} \cdot \frac{1}{\sqrt{\sigma_B^2 + \epsilon}} + \frac{\partial l}{\partial \sigma_B^2} \frac{2(x_i - \mu_B)}{m} + \frac{\partial l}{\partial \mu_B} \cdot \frac{1}{m}$$

$$\frac{\partial l}{\partial \gamma} = \sum_{i=1}^{m} \frac{\partial l}{\partial y_i} \cdot \hat{x}_i$$

$$\frac{\partial l}{\partial \beta} = \sum_{i=1}^{m} \frac{\partial l}{\partial y_i}$$

Softmax and Cross Entropy

$$\mathrm{softmax}(z_t) = \frac{\exp(z_t)}{\sum_i \exp(z_i)} \qquad \hat{y}_t = \mathrm{softmax}(z_t)$$

cross entropy: $L(y, \hat{y}) = -\sum_i y_i \log \hat{y}_i$

$$\frac{\partial L}{\partial z_t} = \sum_i \frac{\partial L}{\partial \hat{y}_i} \frac{\partial \hat{y}_i}{\partial z_t} \longrightarrow \frac{\partial \hat{y}_i}{\partial z_t} = \frac{\partial\left(\exp(z_i) \cdot \sum_i^{-1} \exp(z_i)\right)}{\partial z_t}$$

$$= \sum_i \left(-y_i \cdot \frac{1}{\hat{y}_i}\right) \cdot \frac{\partial \hat{y}_i}{\partial z_t}$$

$$= -y_i + \sum_i y_t \hat{y}_i$$

$$= \hat{y}_t - y_i \quad \#$$

$$\Rightarrow \begin{cases} \text{if } i = t: \dfrac{\partial \hat{y}_i}{\partial z_t} = \dfrac{\partial\left(\exp(z_i) \sum^{-1} \exp(z_i)\right)}{\partial z_i} \\ \qquad = \dfrac{\exp(z_i) \sum_i \exp(z_i) - e^{z_i} e^{z_i}}{\sum_i^2 \exp(z_i)} \\ \qquad = \dfrac{\exp(z_i)}{\sum \exp(z_i)} \cdot \dfrac{\sum \exp(z_i) - \exp(z_i)}{\sum \exp(z_i)} \\ \qquad = \hat{y}_i (1 - \hat{y}_i) \\ \text{if } i \neq t: \dfrac{\partial y_i}{\partial z_t} = \dfrac{0 - e^{z_i} e^{z_t}}{\sum_i^2 \exp(z_t)} = -\hat{y}_i \hat{y}_t \end{cases}$$