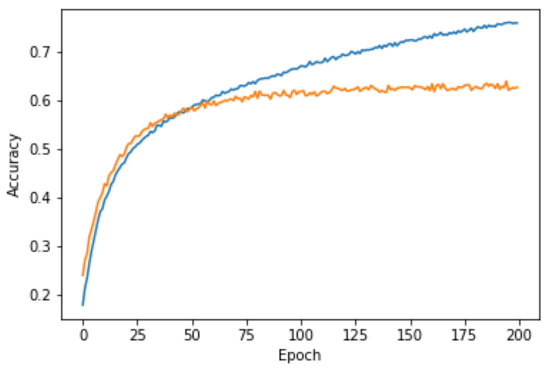
ML hw3 report

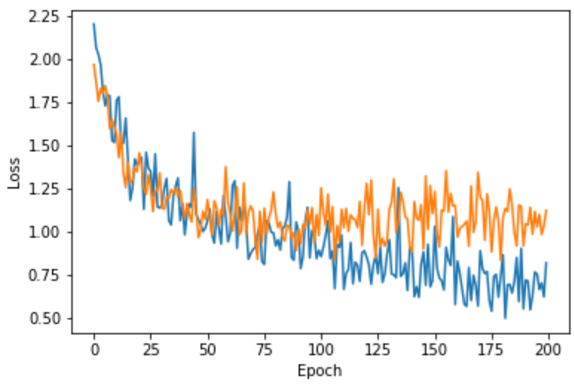
學號：B05705002 系級： 資管四 姓名：黃意芹

(1%) 請說明這次使用的model架構，包含各層維度及連接方式。

這次的model我是用四層CNN加上三層fully connected。第一層出去的channel是64層，之後經過Leaky ReLu, batch normalization, max pool(2,2)與dropout 25%。
第二層出去的channel是128層，之後經過Leaky ReLu, batch normalization, max pool(2,2)與dropout 30%。第三層出去的channel是512層，之後經過Leaky ReLu, batch normalization, max pool(2,2)與dropout 35%。第四層出去的channel是512層，之後經過Leaky ReLu, batch normalization, max pool(2,2)與dropout 40%。之後打參數攤平，連接到三層dense，第一層從512*3*3維到512維，之後經過ReLu, batch normalization, 與dropout 50%。第二層從512維到512維，之後經過ReLu, batch normalization, 與dropout 50%。第三層從512維到7維，最後給出分類結果。

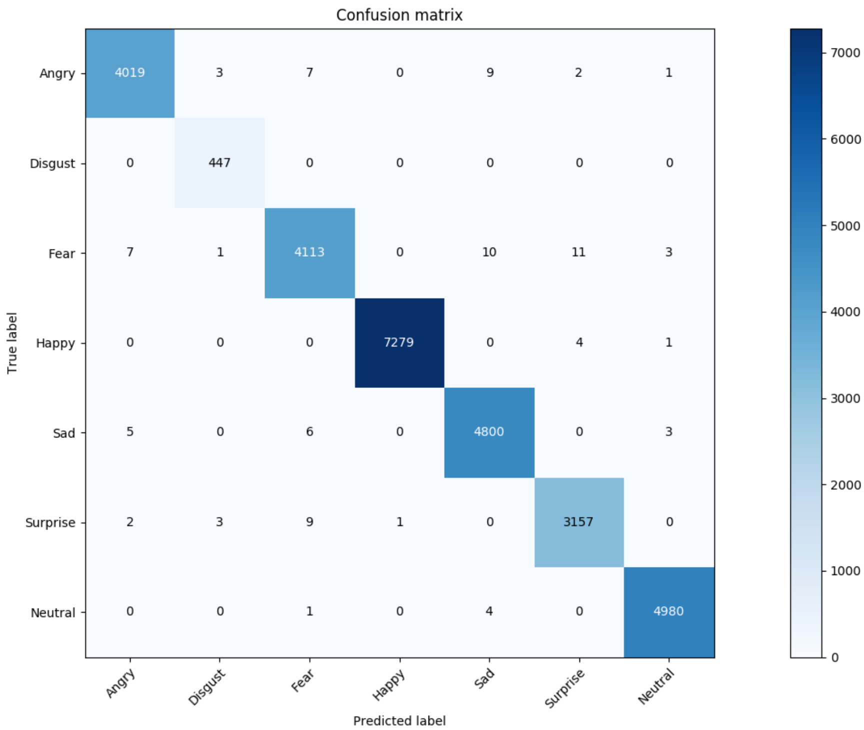
(1%) 請附上model的training/validation history (loss and accuracy)。





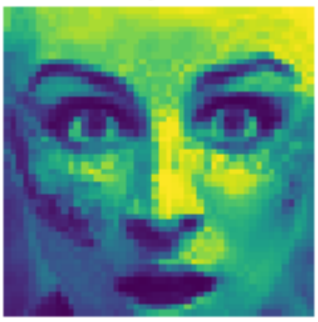
左圖藍線為training accuracy，橘色線為validation accuracy，可以觀察到training accuracy 持續上升，但是validation set從50個epoch後就上升很緩慢。
右圖藍線為training loss，橘色線為validation loss (loss為cross entropy) ，validation set從100個epoch後就下降的不太明顯。

(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。



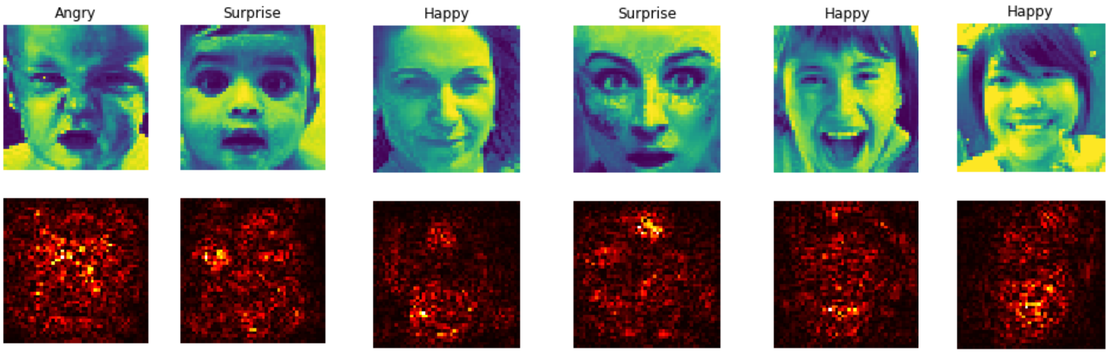
由上圖可以看出 Sad & Fear, Surprise & Fear, Sad & Angry, Fear & Angry 是幾個比較容易搞混的類型。以分錯最多的 Surprise & Fear為例，兩者都會瞪大眼睛和張嘴巴，連人看都不見得能區分，所以模型也容易搞混。

Surprise



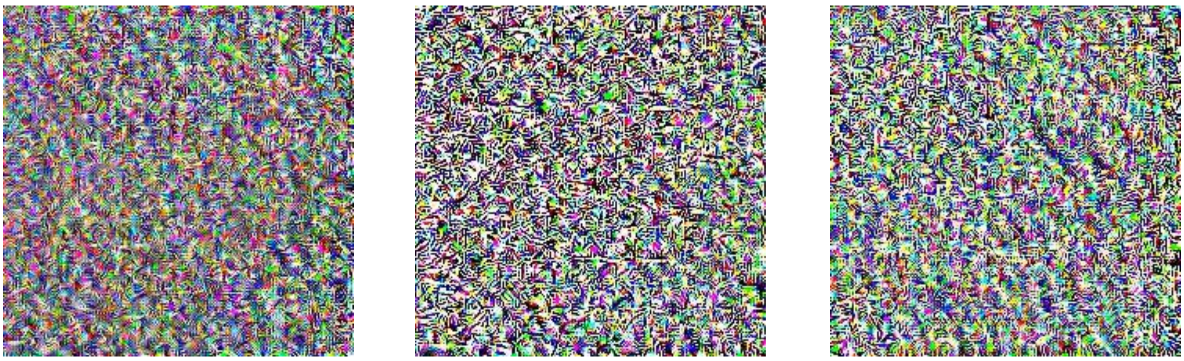
Fear

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



上圖中saliency map的亮處代表我的模型focus圖片的部分。可以觀察到，model比較會focus在臉部，尤其五官處更明顯，例如第一張憤怒的圖著重在皺起的眉頭與眼睛，第二、四張驚訝的圖模型捕捉到瞪大的眼睛，第三、五、六張快樂的圖則著重在笑的嘴吧。

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



這幾張圖是我用model最後一層其中三個filters畫出的圖。但可能因為model的效果不是很好，畫出來看不出特別的特徵，感覺有點像眼睛的紋路。

(3%)Refer to math problem

ML HW3 黃晏祈 B05705001

1、 $\left(\text{output\_channels}, \dfrac{W-k_1+p_1+S_1}{S_1}, \dfrac{H-k_2+P_2+S_2}{S_2}\right)$

2、

$$l=F_2(F_1(u,\theta_1),\theta_2)$$

where $\theta = \arg\min_{\theta} \frac{1}{N}\sum_{i=1}^{N} l(x_i,\theta)$

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

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

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

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

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

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

3、

$$\frac{\partial L_t}{\partial z_t}=\frac{-\partial y_t\log\hat{y}_t}{\partial z_t}=-y_t\frac{\partial\log\hat{y}_t}{\partial z_t}=-y_t\frac{1}{\hat{y}_t}\frac{\partial\hat{y}_t}{\partial z_t}$$

$$=-\frac{y_t}{\hat{y}_t}\frac{\partial\hat{y}_t}{\partial z_t}=-\frac{y_t}{\hat{y}_t}\hat{y}_t(1-\hat{y}_t)=-y_t+y_t\hat{y}_t=\hat{y}_t-y_t$$