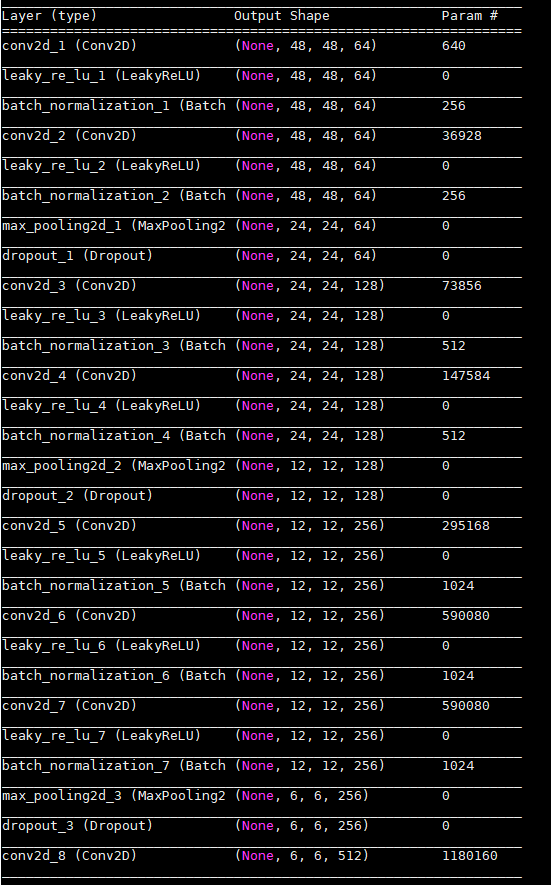
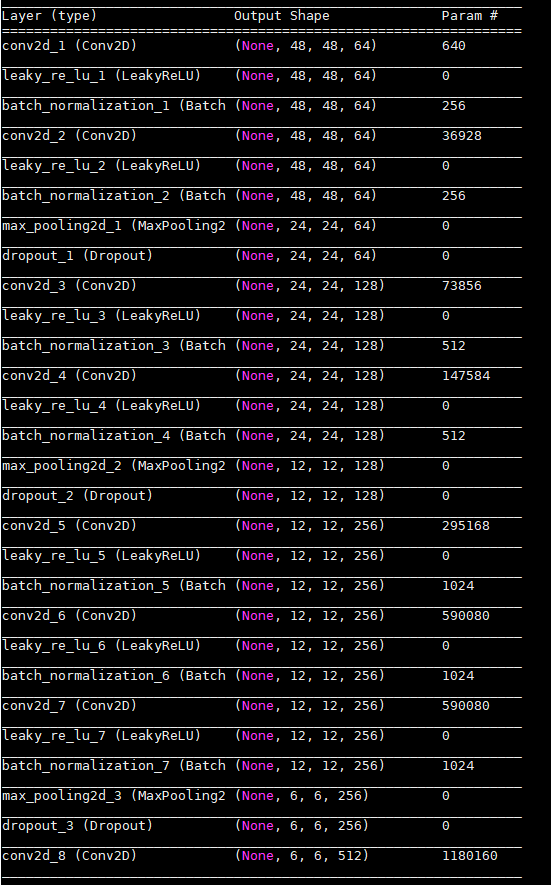
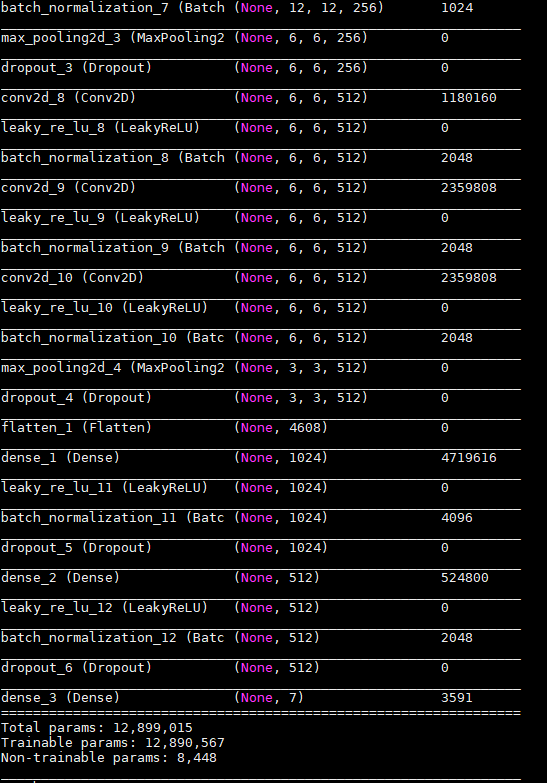
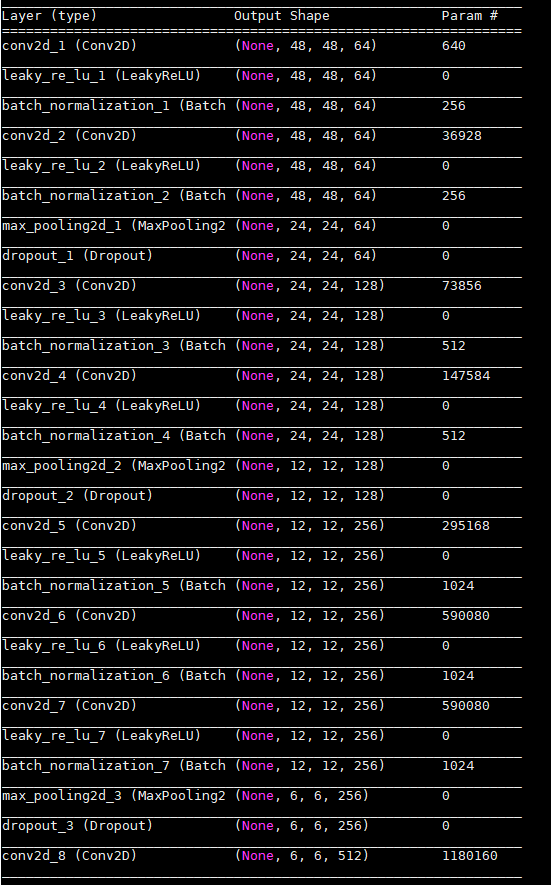
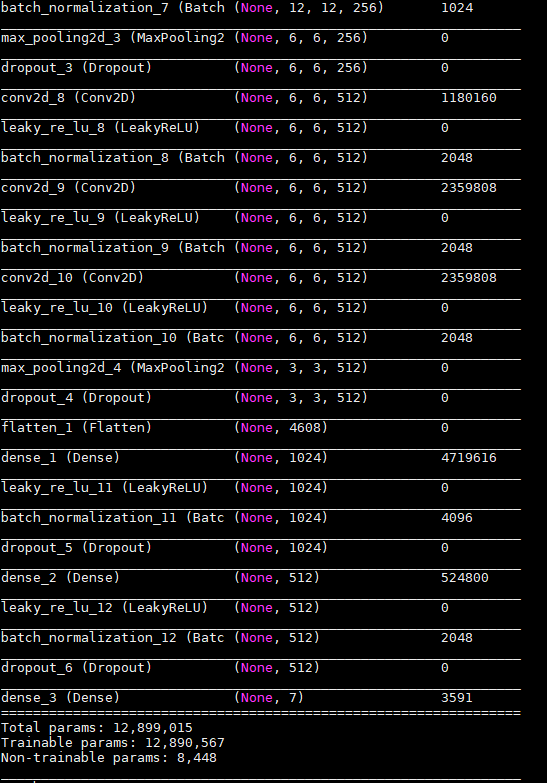
**Homework 3 Report – Image Sentiment Classification**

Professor Pei-Yuan Wu

EE5184 – Machine Learning

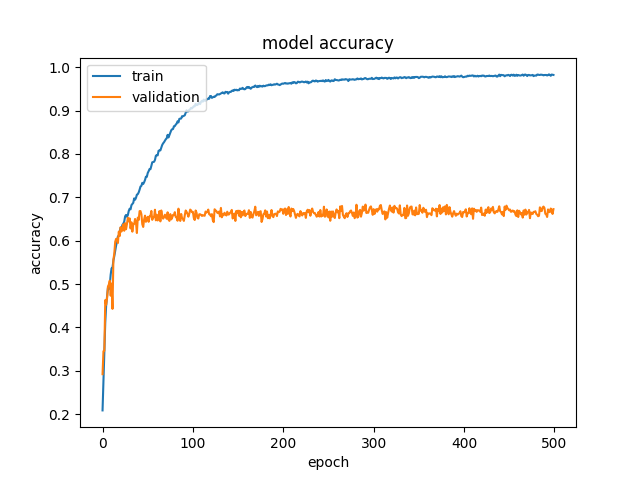
**學號：R07943023 系級：電子所碩一 姓名: 楊仲萱**

**1. (1%) 請說明你實作的CNN model，其模型架構、訓練過程和準確率為何?**

 右圖是我所實作的CNN model 架構，大約的架構是經過兩到三層的Convolution Layer之後，再經過Maxpooling2D Layer並且加上 Dropout Layer。最後加上Fully Connected Layer，把Matrix利用Flatten Layer後一樣加上Dense Layer，最後一層的Dense Layer剛好要output出七個class，對應到答案。

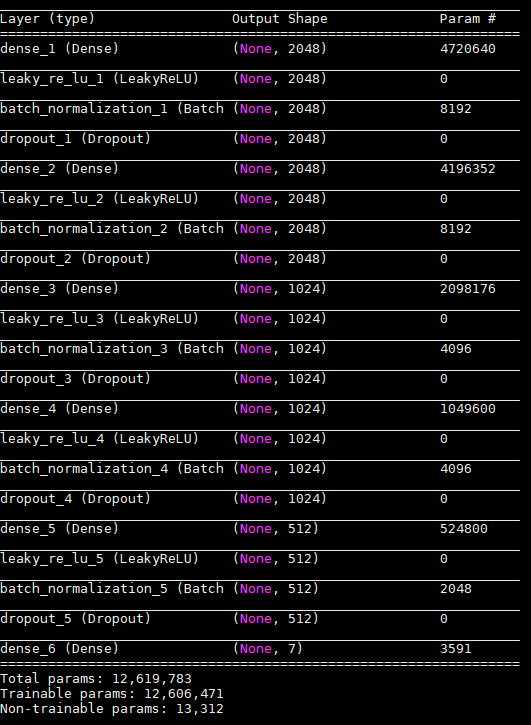
細節上來說，我在filter數為64和128的Conv Layer用了兩層，而在filter數為256和512的Conv Layer則是用了三層。這樣的設計概念是來自於VGG16的思維。每疊了一層Convolution Layer後就會加上LeakyReLU的Activation function，以及Batch Normalization。

除了這些架構我也再加上了Keras內建的ImageDataGenerator進行Data Augmentation。

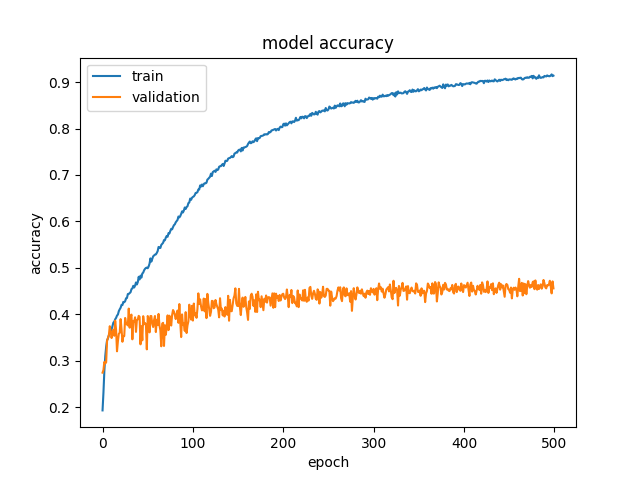


訓練過程的如右圖，當到500個epoch時，validation data的accuracy已經趨於緩和，就算training accuracy還是微微的在上升也是維持約**0.68**的正確率。而training accuracy約是**0.98**。在Kaggle上的分數為0.67818，過strong baseline。

**2. (1%) 承上題，請用與上述CNN接近的參數量，實作簡單的DNN model，其模型架構、訓練過程和準確率為何? 試與上題結果作比較，並說明你觀察到了甚麼。**

原本上題CNN中所用的參數為12,899,015，所以為了讓實作的DNN Model可以有接近的參數量，我使用了兩層的2048個unit的Dense Layer，兩層的1024 units的Dense Layer以及一層的512個units的Dense Layer。也都有像CNN model加上LeakyReLU和Dropout Layer。

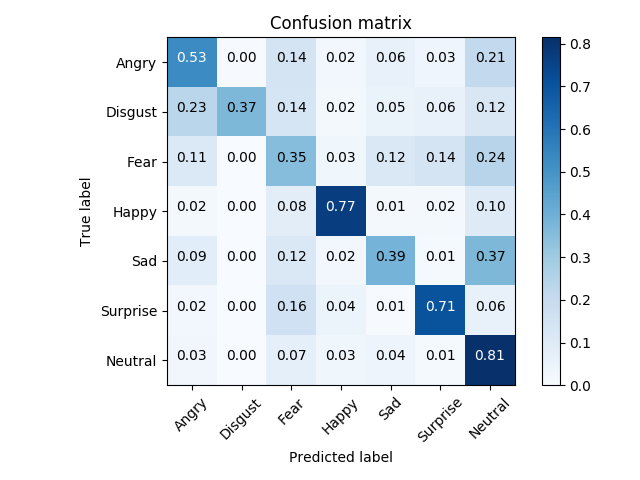
可以從左邊的架構圖中看到Total params為12,619,783是接近的參數量。



訓練過程的部分則是由左圖可以看到，雖然最終的training accuracy可以達到0.92左右的高正確率，但是不難發現在validation的部分，正確率卻一直都只有0.45左右。和上題比較後發現，雖然參數量差不多但是validation accuracy表現卻沒有CNN好，也不像使用CNN Model時可以很快達到收斂的正確率，震盪的幅度也大於CNN Model。在training時也可以發現DNN執行所要花的時間也少於CNN Model。

**3. (1%) 觀察答錯的圖片中，哪些class彼此間容易用混? (繪出 confusion matrix 分析)。**

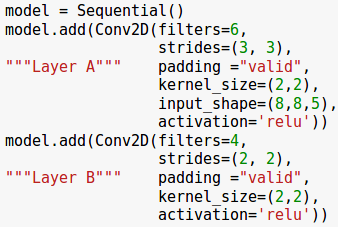
利用confusion matrix後，可以得到下面的圖。

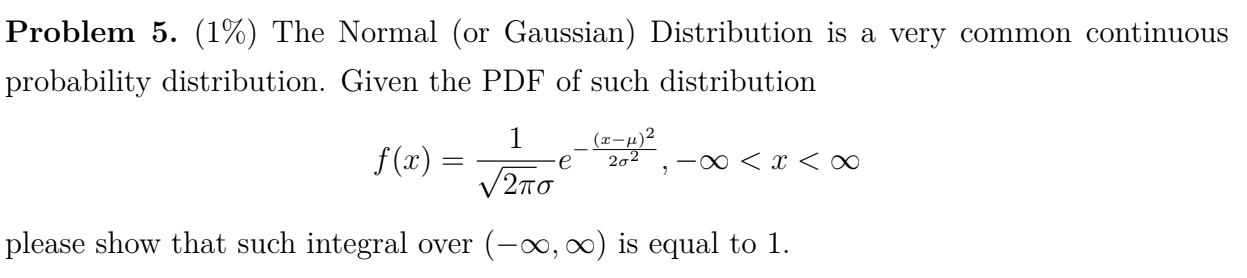


由圖可以發現，Angry容易被誤判成Fear以及Neutral，Disgust容易被誤判成Angry，Fear容易被誤判成Neutral，Sad容易被誤判成Neutral，Surprise容易被誤判成Fear，，Neutral和Happy則可以較好的被判別。

**4. (1.5%, each 0.5%) CNN time/space complexity:**

**For a. b. Given a CNN model as**



**5. (1%) (no collaborator)**

**6. (1%) (no collaborator)**