

人工智慧導論及實作 LAB 2 結報

姓名：虎冠廷

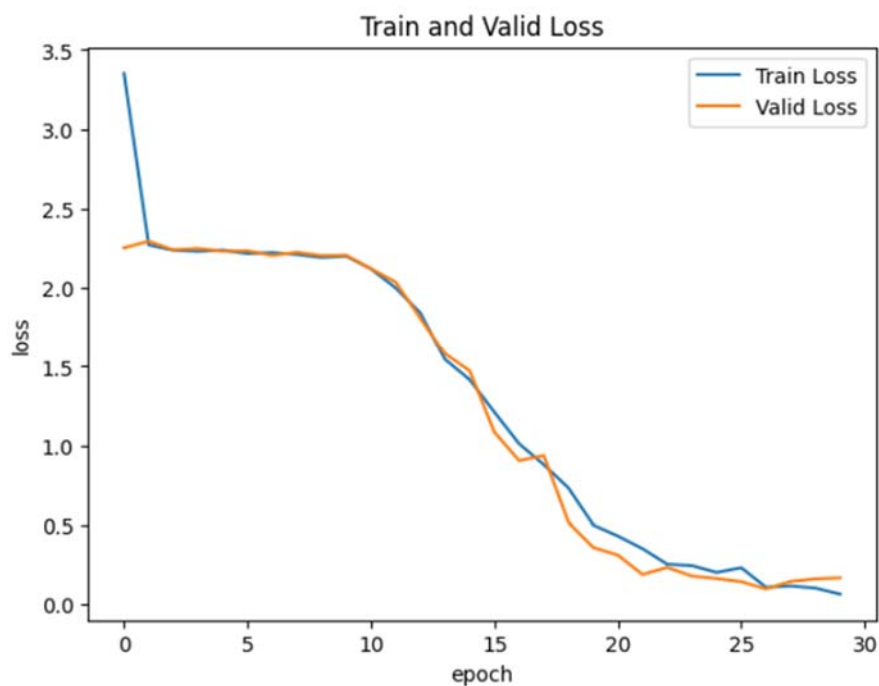
學號：F94096071

(5% for each question, except last one)

1. 貼上你的 Training Validation Accuracy 曲線。



2. 貼上你的 Training Validation Loss 曲線。



3. 請解釋你使用的 augmentation method 其優點與適用情況。

1. 隨機水平翻轉圖像：對於不受方向限制的物體識別任務效果好。
2. 隨機旋轉圖像(15 度內)：適用旋轉較為顯著的圖像識別任務（如文字識別）。

4. 請解釋 Residual block 和 Inception module 之概念與好處。

Residual block 是一種深度學習網絡的模塊，用來解決梯度消失或梯度爆炸的問題，Residual block 是除了原本模型的主架構外，另外還有一個捷徑連結，稱為 skip connection，而它的功能是將前一層的輸出直接加到後一層的輸入中，使信息更容易地傳遞到後面的層中，優點是就算模型網絡很深，能避免梯度消失或爆炸的問題，從而提高訓練效率和模型性能。

Inception module 包含多個分支，在對同一個輸入同時進行不同種運算後，輸出與原輸入截面積同樣大小的尺寸，然後把各個分支的輸出特徵圖拼接在一起作為總輸出，這樣可以提高模型對於不同尺寸和不同特徵的數據的處理能力，並且可以減少模型的參數量，Inception module 在 GoogleLeNet 中被廣泛應用，且取得了良好的效果。

5. 請解釋 Gradient vanishing 發生之原因與解決方法。

梯度消失會出現在反向傳播訓練深度神經網絡的時候，當反向傳播算法將訊息傳遞回輸入層時，梯度逐漸縮小且接近於零。當梯度變得非常小或接近於零時，模型的權重更新變得非常緩慢或停滯在每次訓練的迭代中。

梯度消失發生的原因通常和神經網絡的架構和激活函數有關，當神經網絡的深度增加時，梯度必須通過多個層傳遞回輸入層，神經網絡權重的更新值與誤差函數的偏導數成比例，然而在某些情況下，梯度值會幾乎消失，使得權重無法得到有效更新。像是一個典型的 activation function 如 tanh 函數，其梯度值在 $(-1, 1)$ 範圍內，反向傳播以鏈式法則來計算梯度，這樣相當於在 n 層網絡中，將 n 個這些小數字相乘來計算前端層的梯度，這就使梯度隨著 n 呈指數遞減，層數越深，前端層的訓練越緩慢。

解決辦法：使用多級層次結構、改用殘差網絡（Residual Networks，ResNets）、使用其他的激活函數（如 LeakyReLU 和 ELU）

6. 請解釋 Gradient exploding 發生之原因與解決方法。

梯度爆炸會出現在反向傳播訓練深度神經網路的時候，當反向傳播算法將訊息傳遞回輸入層時，梯度變得非常大，導致權重更新過大，從而破壞模型的穩定性和收斂性，導致模型無法收斂。

梯度爆炸發生的原因：如果模型非常深或非常寬，它可能會產生太多的變量，從而使梯度值急劇增加，或是權重的初始值設置得太大，也可能導致梯度爆炸問題。

解決方法：使用梯度剪切：對梯度設定閾值、weights regularization、使用其他的激活函數（導數在正數部分恆等於 1 如 RELU）、改用殘差網路（Residual Networks, ResNets）

7. 本次 LAB 有哪裡需要改進或可以做得更好的，LAB 難易度是否可接受？

無，LAB 難易度是可接受