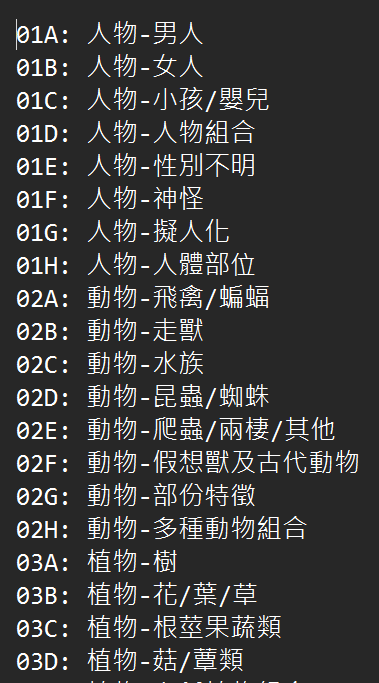
統計學習與深度學習Mini Project 2

B11705051 陳奕廷

專案內容：

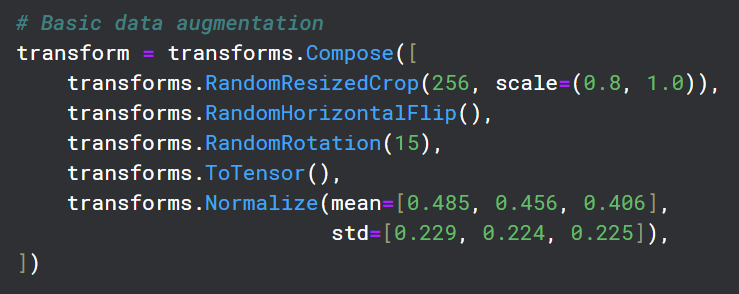
此份專案主要是針對 HW3的影像分類問題的延伸，Image資料集裡面，訓練資料集有 37866張照片，而要被預測的資料有8784張。

和作業不同的地方是，我們觀察到主要的難關是照片可以是多標籤的，因此需要使用另外的 loss 計算方式以及label 的方法，詳細後續會再補充。而標籤總共有 79 個，格式以下。有很多就不列舉了。



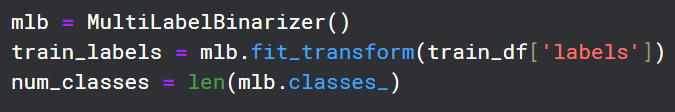
資料前處理：

由於我認為圖片已經相對非常清晰好看了，因此我也沒有多做什麼 transform 的變化，採用比較少變化的方法以免模型訓練時資料複雜度太高。

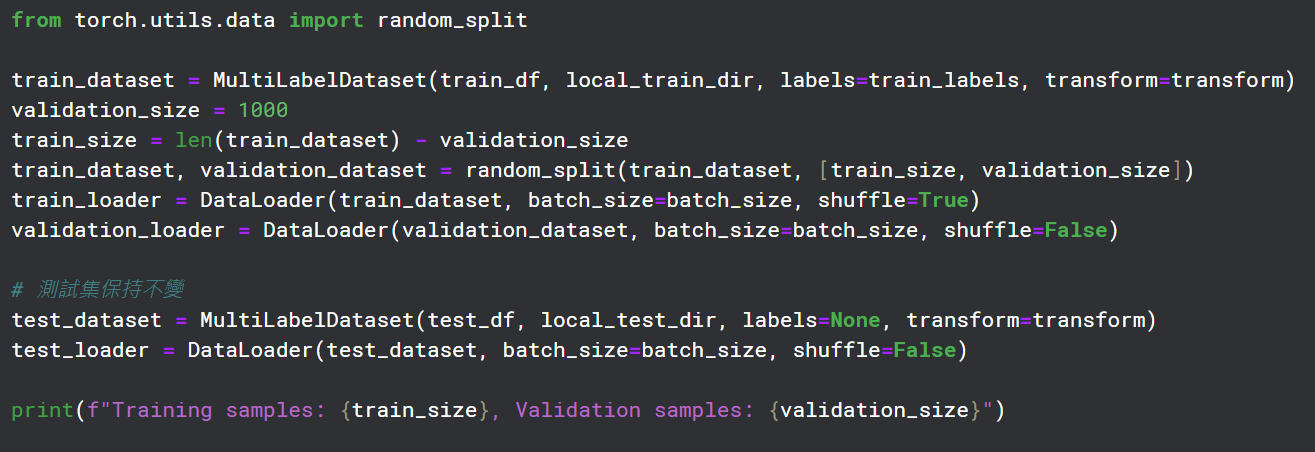


這裡僅採用隨機裁剪、翻轉、旋轉等操作增加數據多樣性，然後把 image vector正規化，好讓模型去做訓練。

接著是 label，只要訓練資料有的 label 會再 label vector上面改成 1 其餘不是的都是 0。偏向 One Hot 作法，只不過他現在是多標籤，因此有可能擁有多個 1。



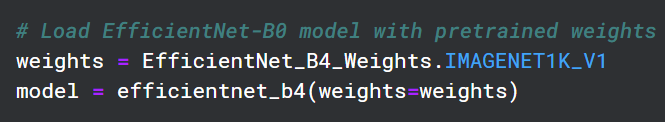
在訓練資料製作完成的同時，我把它split成訓練以及驗證資料 validation，驗證資料總共有 1000個。目的是要檢測 mAP 有沒有持續下降，若停滯那就 early stop，以免 epochs 過多導致的 overfit。





模型選擇：

這邊我使用的是 EfficientNet 家族的模型，EfficientNetB4。這邊會選擇使用 B4 的原因是，我事先跑過訓練更快的 B0，發現在 validation set上，early stop時，mAp 在 0.59左右，上傳Kaggle也大概落在0.52，因此我猜測使用更深更複雜一點的同家族模型，我的表現也會相當不錯。而訓練時第一個 epoch也史無前例的出現 0.4 (先前 B0 開頭是 0.25)。最後修改分類層 classifier，使其適配多標籤分類的輸出數量79個。



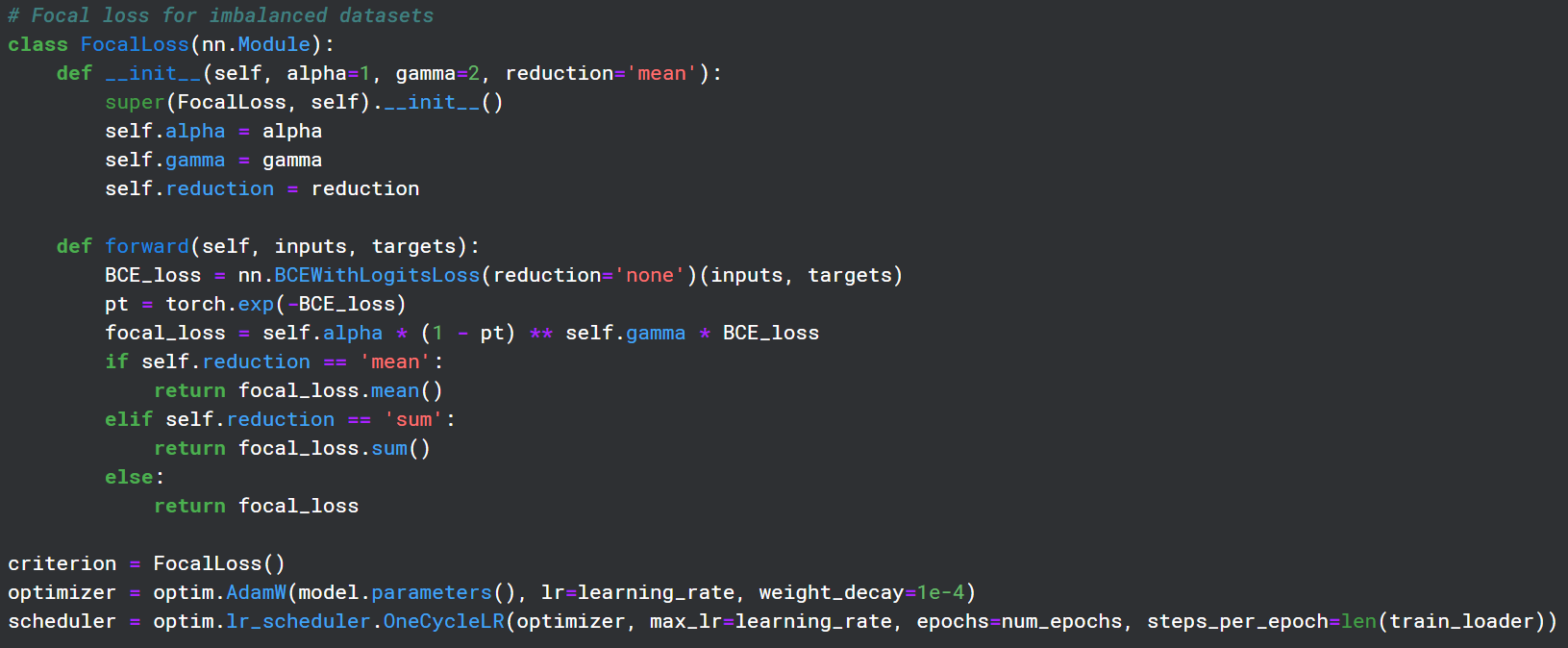
一些幫助自己的方法：

由於訓練時間較長，也可能常常遇到失敗，網路不穩，Kernel Dead 等等的情況，因此我在每一個epoch都把好的結果的model儲存下來，也寫了一套讀取 model的方法。





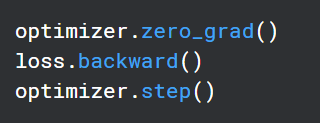
超參數的選擇：



這裡使用 **Binary Cross-Entropy（BCE），是因為它很適合多標籤分類的評估，並採用**Focal Loss 是 BCE 的變體，設計用來解決類別不平衡問題，特別是在檢測或分類問題中，負樣本往往遠多於正樣本。

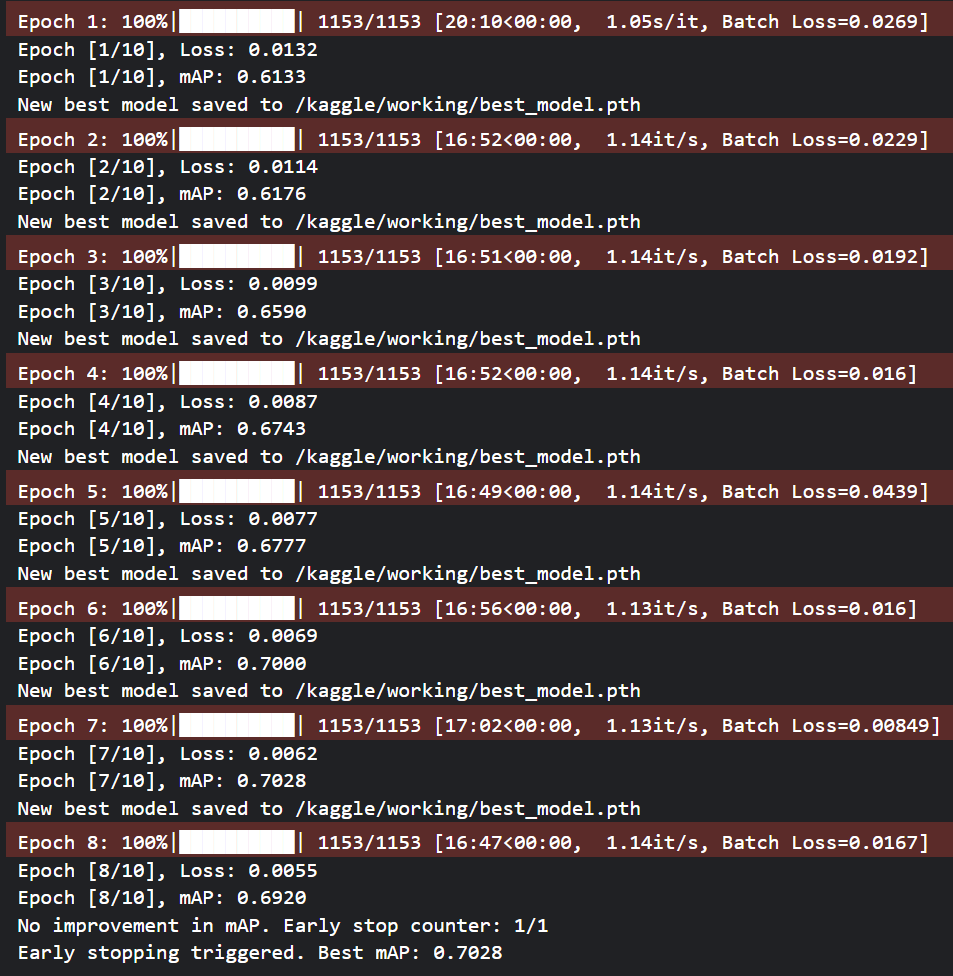
Optimizer使用AdamW，是改進版的 Adam 優化器，它能更好地處理權重衰減（weight decay），在深度學習模型中常用於穩定訓練並加快收斂。

並採用動態調整的學習率，使用已經生成的套件幫助模型挑選合適的學習率，並在每一個 epoch後調整。



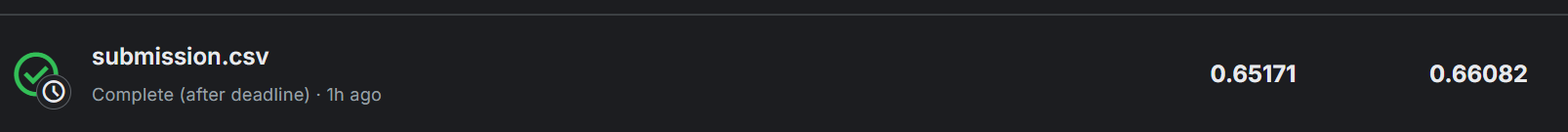
觀察現象：

這裡我發現，有切出 validation set並使用early stop 非常重要，我一開始認為直接訓練並輸出結果應該就有好分數了吧，但可以發現 loss 非常快就變很小，那對於使用者來說，loss 的變動這麼低，模型收斂了嗎，或者 overfit 了嗎。所以採用 validation & early stop 後，我發現丟上 Kaggle 的分數也好看很多，我同時也知道該在什麼時候停，即使每一個 epoch 的 loss 都非常非常小。下圖是讀取已經訓練三個 epoch 的模型繼續訓練，因為中間斷線。



成效：

從上圖可以看到，自己的 validation set的結果是 0.7028，而我丟上 Kaggle 後的分數是 0.65-0.66。差距不會說非常的巨大，是有估計價值的。



程式碼的使用：

在ipynb最後下面，有讀取 best model相關的程式碼，除此之外從上往下執行程式碼即可讀取資料並訓練，最後生成預測。