#### 前言

主要有四個程式檔案，batch1\_eda、preprocessing、model、predict，各檔案最前面的cell都有說明該檔案的輸出與主要做了甚麼，也有貼在下面。另外說明一些注意事項，像是程式碼那些儲存格要怎麼處理；還有若要改善模型，有哪些可能的方向。

#### 程式執行流程

* Preprocessing 🡪 model 🡪 predict
* batch1\_eda可獨立執行
* **Preprocessing**
  + 檔案產出：從DB抓取資料，切分成train、valid、test三個資料夾，裡面各有12個pattern資料夾。裡面圖片都是切割過的。
  + 主要做的事情：資料前處理、切割圖片
* **Model**
  + 檔案產出：efficientnet.h5檔案(會在models資料夾下面)，即訓練完成的efficientnet模型
  + 主要做的事情：oversampling、data augmentation、model、hyperparameter tuning
* **Predict**
  + 檔案產出：輸入圖片(可為網址或檔案位置)，回傳預測pattern結果
  + 主要做的事情：將輸入圖片經過同樣的前處理後，輸入模型，預測出結果
* **batch1\_eda**
  + 檔案產出：請見程式輸出
  + 主要做的事情：從DB抓取資料，隨機印出所有類別的16張圖片

#### 注意事項

* **Preprocessing**

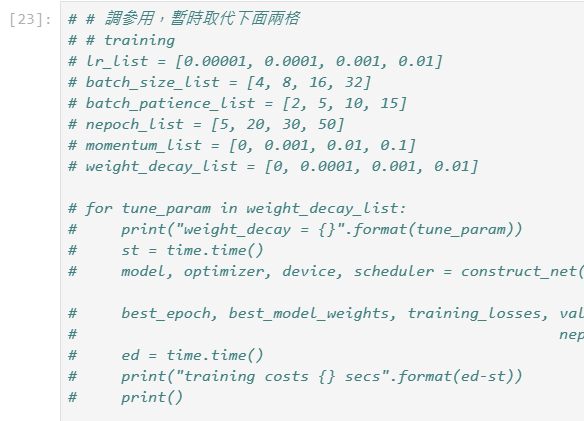


這格在切照片，已經將照片切成三個資料夾了，不須重作，所以註解掉。



這兩格就是把圖片分成三部分(三資料夾)，需要重跑的話按照上面的註解作即可。

* **Model**



這格是調參的時候用的，正常來說只需要這格下面的兩格就好。

#### 未來展望

若要改善模型，以下幾個方向提供參考：

* 改善data augmentation：



從這格去改，改datagen裡面就好，datagen\_valid\_test無須改動。

* 選擇其他模型：

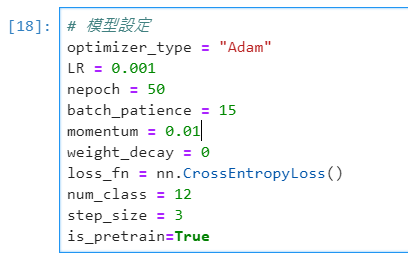


在這格把model變數換成其他想要的模型即可。

<https://blog.csdn.net/u014380165/article/details/90812249>

<https://arxiv.org/abs/1905.11946>

* 調整超參數：



這格是需要手動調整的超參數，其中LR代表學習率(learning rate)，是調整模型學習快慢的參數，動這個會讓模型表現變化較大。