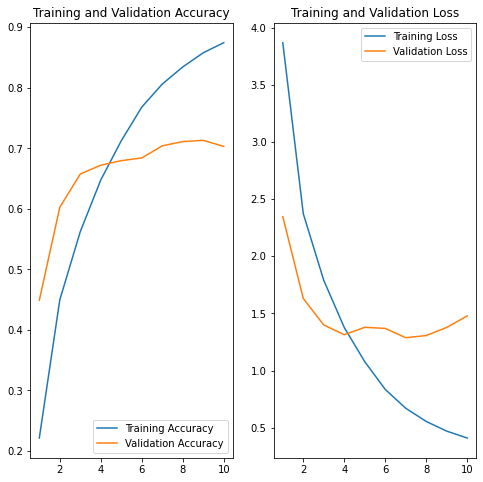


機器學習期末專題

BIRDS 400 - SPECIES IMAGE CLASSIFICATION



資三甲 ACS108128 楊斌翔

1. 前言  
   因筆電效能問題無法有效地在自身筆電訓練model故此專題為透過Google的產品Colab來訓練，有部分程式碼使用到Google Colab API，可能會導致無法透過本機執行，請多加注意。  
   此外，因為Colab有資源使用限制，所以無法使用太複雜的model訓練，在此有做點取捨，使用簡單一點的model，犧牲準確度，以在Colab正常運行。
2. 資料集介紹  
   題目：BIRDS 400 - SPECIES IMAGE CLASSIFICATION  
   連結：[BIRDS 400 - SPECIES IMAGE CLASSIFICATION | Kaggle](https://www.kaggle.com/datasets/gpiosenka/100-bird-species)  
   目標：訓練出一個分類器可區分出400種鳥類  
   資料庫說明：  
   資料集為400種鳥類種類的圖片，用於訓練出鳥類分類器。每張鳥類的圖片皆為224 x 224 x 3的RGB圖片，圖片中的鳥類佔總圖片超過50 %。每個種類在訓練集中的照片數量並不一致，但都超過120張，不會有資料太少的問題。資料唯一不平衡的地方為鳥的性別，雄雌性比例大約85:15，可能會無法判斷雄雌性，但這並非本專題的目標，故不影響。
3. 資料說明  
   種類：400種  
   圖片大小：224 x 224 x 3  
   Training dataset：總共58338張，400個資料夾區分種類，每個種類數量不等，但都至少超過120張。  
   Validation dataset：總共2000張，400個資料夾區分種類，每個種類各五張。  
   Testing dataset：總共2000張，400個資料夾區分種類，每個種類各五張。
4. 訓練方法與步驟  
   資料讀取：  
   因資料已先分類並存放在不同資料夾中，所以在讀取資料時使用keras.utils.image\_dataset\_from\_directory即可得到data/label的dataset。  
   資料預處理：   
   圖片的處理有嘗試使用過rgb\_to\_grayscale但準確度比沒處理來的低，所以最後僅將圖片做Rescaling，Rescaling會在model中執行。  
     
     
   model結構：  
   Input -> Rescaling -> [Convolution -> MaxPooling -> Normalization] \* 5 -> Flatten -> Dropout -> [Normalization -> Dense] \* 3 -> SoftMax  
   model 說明：  
   Input進來為 (224 x 224 x 3)，首先會先做Rescaling讓model能更好的分辨圖片，之後做CNN五次，過程為Convolution -> MaxPooling -> Normalization，Convolution會使用Relu解決梯度消失問題，Normalization用於降低梯度消失與加速收斂，五次Convolution的filter為 (256, 128, 64, 32, 16)，之後用Flatten將圖片展開並且Dropout，此時的大小為 (1600)，Dropout用於降低Overfitting的問題，最後做DNN，層數為 (1600, 800, 400)，前兩層會做Relu，最後一層為Output，會做SoftMax來將預測的機率分布標準化。  
   model設定：  
   loss function採用categorical\_crossentropy，用於多個class的model，其output會是one-hot category encoding。  
   optimizer使用Adam來做梯度下降，它能根據過去的梯度調整learning rate，能有效地調整參數。  
   訓練：  
   Batch Size設定32，Epoch設定10，每次會訓練1825張照片。
5. 結果  
   訓練過程：  
     
   Testing Data準確度：  
     
   分析：  
   圖中可看到在第5次後Validation Accuracy有趨緩，且Validation Loss有升高的跡象，判斷可能為Overfitting的問題，如果能做資料的平移，旋轉可能可以增加準確率，另外經實測彩圖會比灰階精確度來的高，原因可能是400種鳥類的分類其顏色可能會是一種特徵，最後Layer的層數也有可能會受到影響，如果能增加更多層的Layer有機會增加準確度會跟著上升。
6. 程式執行注意事項  
   執行環境：Python 3.7.13 (需在Colab執行)  
   Package：kaggle, numpy, matplotlib, tensorflow 2.8.2  
   資料下載：透過Kaggle API下載，程式碼會要求上傳Authentication，所以要先到Kaggle網頁取得，[參考連結請點這](https://www.kaggle.com/docs/api)。  
   訓練所需時間：將近一個半小時。