**國立雲林科技大學資訊管理系**

**機器學習**

專案作業二

PREDICT MANGO QUALITY AND OBJECT RECOGNITION WITH CONVOLUTIONAL NEURAL NETWORK

以卷積神經網路預測芒果品質以及物件識別

M10723047 黃靖媛

M10923012 陳羿欣

M10923021 吳青芬

M10923023 張凱淇

指導教授：許中川

2021 年 05 月

# 摘要

本研究使用愛文芒果資料集和CIFAR-10資料集，利用卷積神經網路（Convolutional Neural Network，CNN）模型來建構圖像分類器，並使用ResNet50作為本實驗之預訓練模型。愛文芒果資料集主要進行預測芒果的品質等級，希望透過分類器作果篩自動化以提升篩選芒果的效率，且須達到一定的準確率；CIFAR-10 資料集主要進行10種的物件類別預測。在本研究中分類預測模型利用Precision、Recall及F1-score來每一類別之判別績效，經由評估指標計算顯示，以ResNet50作為預訓練模型所架設之分類器能在愛文芒果資料集和CIFAR-10資料集的分類任務以及物件辨識獲得良好的績效。

**關鍵字：**卷積神經網路、圖像分類、ResNet50、物件辨識

# 一、緒論

## 1.1動機

台灣芒果在近年來出口銷量逐年增加，成了我國外銷出口經濟最高的生鮮水果之一，愛文芒果採收後依品質篩選，分為上等、普通、瑕疵 三種等級，然而愛文芒果的品質若依靠人工篩選，除了農村人口流失導致人力短缺，篩選芒果（簡稱篩果）流程也因保鮮期壓縮地極短，導致篩果階段約有10％的誤差，其篩果處理有待改善的是採收後處理技術。影像辨識技術可利用於芒果的瑕疵檢測，因此本研究使用愛文芒果資料集，利用卷積神經網路（CNN）模型來進行芒果品質的分級。

在第二個資料集CIFAR-10中包含10個不同類別的圖片，分別代表飛機，汽車、鳥類、貓、鹿、狗、青蛙、馬、輪船和卡車，而每個類別皆有6,000張32×32像素的彩色圖像。藉由此資料集，本研究欲利用CNN模型架構來識別圖片中之物件。

## 1.2目的

本研究進行兩項類別預測任務，一為芒果品質分級預測，利用Kaiming He, Xiangyu Zhang, Shaoqing Ren & Jian Sun (2016)所提出的ResNet50架構來建構辨識模型，以提高篩果的效率及準確率，進而彌補人力短缺所造成的問題。二為CIFAR-10物件識別，也是利用ResNet50架構來建構辨識模型，將資料集中test data的所有圖片進行識別分類。兩項任務皆使用Precision、Recall及F1指標來評估模型的預測結果。

# 二、方法

## 2.1實作方法說明

本研究實驗資料集包括愛文芒果資料集、CIFAR-10 資料集，首先將影像進行前置處理，其包括影像正規化、類別編碼、資料增強以及資料分割，即為將訓練資料分成訓練資料（train data）以及驗證資料（validation data）。兩資料集皆使用ResNet50來進行影像特徵的擷取，再對影像辨識模型使用Precision、Recall及F1等指標對模型進行優劣評估，最後在預測測試資料（test data）之類別，得出識別準確率。

## 2.2程式執行方法說明

本研究利用Anaconda3的Jupyter notebook環境來進行開發，使用OpenCV和numpy等套件對愛文芒果資料集和CIFAR-10資料集進行影像的前處理，利用keras建立辨識模型，並對資料進行預測，再將模型建立時所記錄的每個epoch以及訓練後的loss和accuracy畫出檢視學習歷程，判斷是否正常後再使用Precision，Recall及F1等指標對模型進行優劣評估。

# 三、實驗

## 3.1資料集

本研究使用愛文芒果資料集以及CIFAR-10 資料集兩種資料集作為實驗資料，以下為兩種資料集前處理之前圖像規格與說明。

### 3.1.1 愛文芒果資料集

愛文芒果資料集中有3種類別，分別為A、B、C，如表1，A代表上等的芒果，B代表普通的芒果，C代表瑕疵的芒果，共有6130張訓練資料及250張測試資料，而每張圖片皆為大小不同的彩色影像（RGB）。

* 原始資料筆數：6130
* 測試資料筆數：250
* 類別數量：3
* 類別種類：A（上等）、B（普通）、C（瑕疵）

表 1

**愛文芒果資料集部分資料內容**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 影像 |  |  |  | D:\一下\ML\hw2\MANGO\Train_Image\00005.jpg |  |
| 檔名 | 00002.jpg | 00003.jpg | 00004.jpg | 00005.jpg | 00007.jpg |
| 類別 | C | C | B | A | A |

### 3.1.2 CIFAR-10資料集

CIFAR-10資料集中有10種類別，分別為Airplane、Automobile、Bird、Cat、Deer、Dog、Frog、Horse、Ship、Truck，如表2，而每種類別分為5000張訓練資料及1000張測試資料，每張圖片皆為32×32像素的彩色影像（RGB）。

* 原始資料筆數：50000
* 測試資料筆數：10,000
* 類別數量：10
* 類別標籤：Airplane、Automobile、Bird、Cat、Deer、Dog、Frog、Horse、Ship、Truck。

表 2

**CIFAR-10資料集部分資料內容**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 影像 |  |  |  |  |  |
| 檔名 | aeroplane\_s\_000004.png | ambulance\_s\_000101.png | accentor\_s\_000014.png | alley\_cat\_s\_000009.png | alces\_alces\_s\_000001.png |
| 類別 | Airplane | Automobile | Bird | Cat | Deer |
| 影像 |  |  |  |  |  |
| 檔名 | blenheim\_spaniel\_s\_000006.png | alytes\_obstetricans\_s\_000006.png | american\_saddle\_horse\_s\_000001.png | abandoned\_ship\_s\_000004.png | aerial\_ladder\_truck\_s\_000001.png |
| 類別 | Dog | Frog | Horse | Ship | Truck |

## 3.2前置處理

### 3.2.1愛文芒果資料集

圖 1愛文芒果資料集-前置處理流程圖

* 影像正規化：將每張影像的像素值正規化至0到255的灰階值。
* 類別編碼：將每種類別編號。
* 資料增強：將影像重新縮放、旋轉、水平翻轉或垂直翻轉等動作來製作成新的影像，以彌補資料量的不足。
* 資料分割：將資料之類別與欲用來訓練之資料分割，以8：2比例分割成4905筆訓練資料以及1225筆驗證資料。

3.2.2 CIFAR-10資料集

圖 2 CIFAR-10資料集-前置處理流程圖

* 影像正規化：將每張影像的像素值正規化至0到255的灰階值。
* 類別編碼：將每種類別編號。
* 資料分割：將資料之類別與欲用來訓練之資料分割，以8：2比例分割成40000筆訓練資料以及10000筆驗證資料。

## 3.3實驗設計

本研究在實驗設計進行兩種模型實驗測試，一為淺層CNN，二為ResNet50；淺層CNN採用5層Convolution layers，目的是為了將圖像大特徵轉變成小特徵，並用Maxpooling保留重要特徵，每次進行卷積時，為了防止整體數值分布偏右或偏左，因此在卷積作業後使用Batch Normalization 以保持數值原本之分布，最後使用softmax輸出每一類的機率取最大之當作預測結果，在愛文芒果資料集之準確率為0.73，在CIFAR-10資料集之準確率為0.75，經實驗證實ResNet50相對於淺層CNN來說，較能獲得較高之準確率，因此本研究主要以ResNet50作為主要之實驗模型，以下為愛文芒果資料集與CIFAR-10資料集的實驗流程步驟：

圖 3實驗流程圖

### 3.3.1 愛文芒果資料集

在愛文芒果資料集中，預訓練模型使用ResNet50架構，在模型中使用Dropout，參數設置為0.3，使模型在訓練時去除不必要之特徵以及防止模型overfitting；Batch size之設置會影響模型優化程度以及收斂速度，由於本資料集之訓練資料筆數經由資料分割後為4905筆，因此Batch size參數設置為8；若模型訓練在10回合Loss尚未降低，則提前結束訓練（early stopping），在模型之Input shape設定為224×224×3，Weights為imagenet，而在ResNet後接上分類器之網路層的Activation function為Relu，最後輸出層使用softmax。

1. 模型建構：設定模型初始參數

* Epoch：1000
* Batch Size：8
* Activation: Relu, softmax
* Loss：categorical\_crossentropy
* Optimizer：Adam
* Metrics：accuracy

1. 訓練模型：將資料匯入模型中並進行訓練
2. 計算指標：使用Precision, Recall及F1指標做為模型評估指標
3. 輸出預測績效結果。

表 3

**愛文芒果資料集品質分類模型架構表**

|  |  |  |
| --- | --- | --- |
| Layer (type) | Output Shape | Param # |
| flatten\_2(Flatten) | (None,100352) | 0 |
| dense\_4(Dense) | (None,4096) | 411045888 |
| dense\_5(Dense) | (None,2048) | 8390656 |
| dropout\_2(Dropout) | (None,2048) | 0 |
| dense\_6(Dense) | (None,1024) | 2098176 |
| dense\_7(Dense) | (None,3) | 3075 |
| Total params: 445,125,507 | | |
| Trainable params: 445,072,387 | | |
| Non-trainable params: 53,120 | | |

### 3.3.2 CIFAR-10資料集

在CIFAR-10資料集中，預訓練模型使用ResNet50架構。 在模型設置Dropout，使模型在訓練時去除不必要之特徵以及防止模型overfitting；參數設置為0.3，Batch size之設置會影響模型優化程度以及收斂速度，由於本資料集之訓練資料筆數經由資料分割後為40000筆，因此Batch size參數設置為16；若模型訓練在10回合Loss尚未降低，則提前結束訓練（early stopping），在本實驗中模型之Input shape設定為224×224×3，Weights為imagenet，而在ResNet後接上之網路層的Activation function為Relu，最後輸出層使用softmax。

1. 設定模型初始

* Epochs：1000
* Batch size：16
* Activation : Relu, Softmax
* Loss：Categorical\_crossentropy
* Optimizer：Adam
* Metrics：Accuracy

1. 將資料匯入模型中並進行訓練
2. 使用Precision, Recall及F1指標做為模型評估指標。
3. 輸出預測績效結果

表 4

**CIFAR-10資料集物件識別模型架構表**

|  |  |  |
| --- | --- | --- |
| Layer(type) | Output Shape | Param # |
| flatten\_2(Flatten) | (None,100352) | 0 |
| dense\_4(Dense) | (None,4096) | 411045888 |
| dense\_5(Dense) | (None,2048) | 8390656 |
| dropout\_2(Dropout) | (None,2048) | 0 |
| dense\_6(Dense) | (None,1024) | 2098176 |
| dense\_7(Dense) | (None,10) | 10250 |
| Total params:445,132,682 | | |
| Trainable params:445,079,562 | | |
| Non-trainable params:53,120 | | |

## 3.4實驗結果

### 3.4.1 愛文芒果資料集

表 5

**分類模型預測績效表**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 訓練資料績效 | 驗證資料績效 | 測試資料績效 |
| **Accuracy** | 0.848 | 0.796 | 0.808 |
| **Loss** | 0.369 | 0.474 |  |

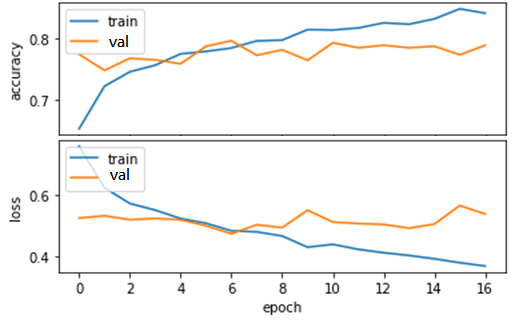


圖 4品質分類模型訓練績效

表 6

**分類模型評估表**

|  |  |  |  |
| --- | --- | --- | --- |
| 類別 | **Precision** | **Recall** | **F1-Score** |
| A | 0.84 | 0.82 | 0.83 |
| B | 0.78 | 0.74 | 0.76 |
| C | 0.81 | 0.89 | 0.84 |

### 3.4.2 CIFAR-10資料集

表 7

**分類模型預測績效表**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 訓練資料績效 | 驗證資料績效 | 測試資料績效 |
| **Accuracy** | 0.993 | 0.919 | 0.939 |
| **Loss** | 0.019 | 0.325 |  |

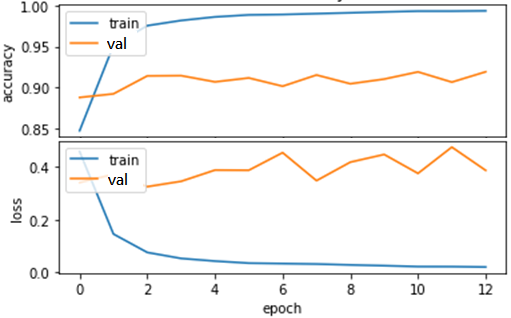


圖 5物件識別模型訓練績效

表 8

**分類模型評估表**

|  |  |  |  |
| --- | --- | --- | --- |
| **類別** | **Precision** | **Recall** | **F1-Score** |
| airplane | 0.94 | 0.94 | 0.94 |
| automobile | 0.97 | 0.97 | 0.97 |
| bird | 0.94 | 0.92 | 0.93 |
| cat | 0.89 | 0.85 | 0.87 |
| deer | 0.92 | 0.95 | 0.94 |
| dog | 0.91 | 0.86 | 0.88 |
| frog | 0.95 | 0.97 | 0.96 |
| horse | 0.95 | 0.97 | 0.96 |
| ship | 0.96 | 0.98 | 0.97 |
| truck | 0.96 | 0.96 | 0.96 |

# 四、 結論

在本研究中使用愛文芒果資料集與CIFAR-10資料集，作為分類預測之實驗資料集，並利用ResNet50作為預訓練模型，再分別建構分類器模型，以達到分類預測之任務。在愛文芒果資料集中，進行芒果分級的預測任務，希望能透過CNN模型判別出芒果的品質，該資料集包含了上等貨、普通、瑕疵三種不同品質的芒果圖像，在該實驗中預測正確之數量為202張，錯誤張數為48張，準確率高達81%，透過表5與表6綜合評估，該分類器可用於芒果果篩任務判別大部分的芒果品質。

在CIFAR-10資料集中，包括了10種不同類別之物件，在該物件識別任務中預測正確數量為9386張，錯誤張數為614張，準確率高達93.86%，透過表7與表8綜合評估，該分類器可預測出大部分圖像中之物件。

**參考文獻**

**英文文獻**

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp.770-778).