中華大學資訊工程學系

111學年度專題製作報告

**基於深度學習鞋子影像辨識之購買系統**

指導老師： 張欽智 教授

組長:H10818015 陳振笙  
組員:B10802039 陳建安

組員:B10802005 葉士賢

**中華民國 112年 1月 10日**

# 目錄

[目錄](#_Toc45490468) 1

[一、 研究背景與動機](#_Toc45490470) 2

[二、 專題內容](#_Toc45490471) 2

[三、 研究方法](#_Toc45490472) 3

[四、 流程圖](#_Toc45490473) 6

[五、 成果展示](#_Toc45490474) 11

[六、 工作分配與進度甘梯圖](#_Toc45490475) 12

[七、 結論與未來展望](#_Toc45490478) 13

[八、 參考文獻](#_Toc45490472) 14

1. 研究背景與動機

一開始在討論要做哪種專題時，看到老師之前學生做的「蔬菜辨識」，有了想要嘗試的想法，之後透過實作後可以成功辨識出蔬菜，於是我們想做「便當辨識」、「麥當勞辨識」等等，但發現這些都有人做過了，後來我們想到鞋子，平時在路上看到很多潮鞋，但不知如何搜尋及購買通路，因此我們想做一個辨識鞋子的系統，可以透過辨識、品牌辨識系列等等細節，但難度太高，最後決定退一步做「鞋子種類辨識」。

因此我們製作了一個APP名為「SHOESEE」，此APP主要是以鞋子為主題來進行辨識，操作流程只須拍照或選擇照片後上傳後就可以找到鞋子資訊了，並推薦5雙相同種類的鞋子給使用者，也可以讓使用者自行透過APP瀏覽商品目錄，尋找到自己想要的鞋子並完成購買。

1. 專題內容

1.利用手機內部的相機進行拍照及選擇照片的功能上傳至伺服器進行鞋子辨識，辨識完回傳結果至App內告訴使用者是什麼類型的鞋子，並推薦相同5雙類型的鞋子給使用者選購。

2.使用者也可以直接使用APP來選購鞋子，在商城裡有依照鞋子的類型來進行分類，讓使用者可以自行依照喜歡的類別來購買。

3.建立網頁商城，可選購商品並且也綁定金流系統

1. 研究方法

在製作鞋子種類辨識時，我們是使用ResNet架構去撰寫訓練模型，ResNet主要有五種主要形式：ResNet18，ResNet34，ResNet50，ResNet101，ResNet152 ，整體架構都可以分為三大部分：

1.輸入部分: 使用一般的convolution，並且用大的stride降低解析度。

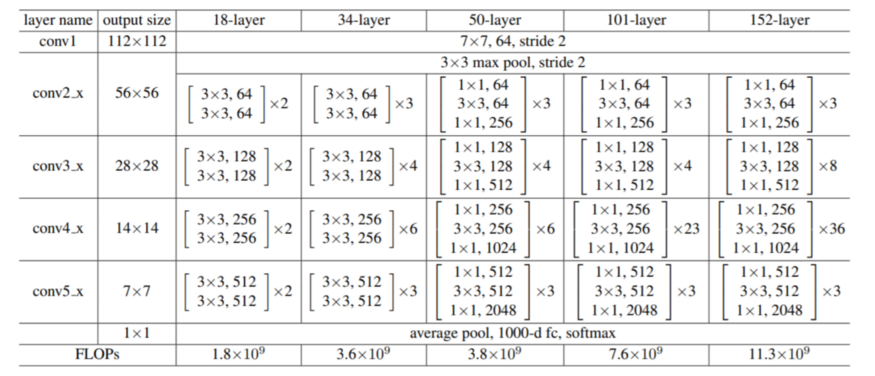
2.卷積部分: ResNet共有4個Stage block，每個stage block都是由數個building block堆疊而成。不論是用stride或是pooling，每個stage一般都會先降低解析度並加大寬度 (channel)，再做一連串的residual learning。

3.輸出部分：依照任務，設計不同的輸出。一般來說這邊會隨著任務轉變，所以通常不算在ResNet的backbone裡。

由於每個 stage第一個 building block的輸入與 residual path連接的地方解析度與網路寬度都不同，所以第一個 block會多一個 convolution，做解析度與寬度的調整。而不同深度的ResNet的差別有兩個：

1.ResNet18、ResNet34使用一般的residual block，而ResNet50、ResNet101、ResNet152使用了expansion為4的bottleneck block。

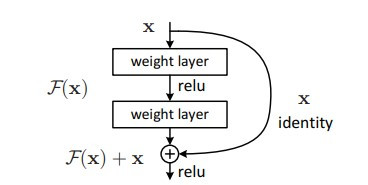
2.剩下的差異就在於每個stage堆疊的building block層數不同，詳細差異可以參考下圖表格。



**圖3-1 不同層數的ResNet**

ResNet特色

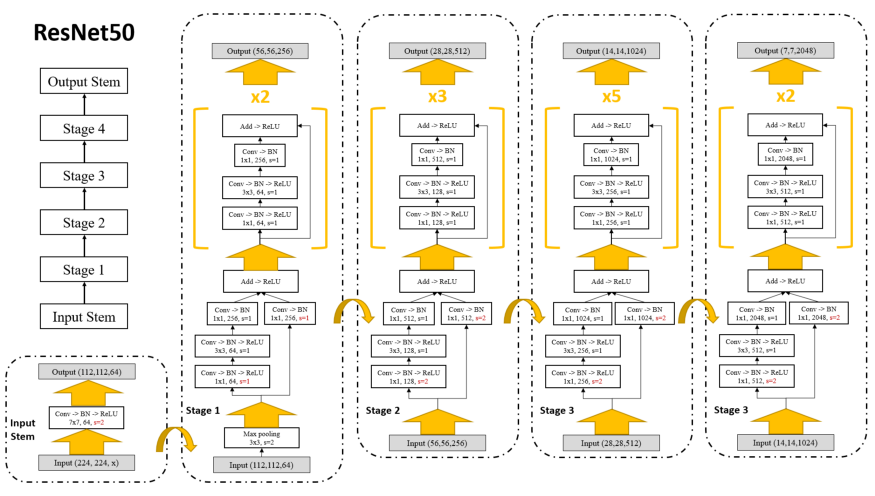
1.ResNet在VGG的基礎上，新增直連通路(skip connection)，不會新增任何參數，但是可以解決梯度消失的問題

2.直連通路與卷積層進行元素點(element-wise)相加，而非直接連接。  
這使某一層模型只需要學習輸入與輸出之間的殘差(residual)，大幅度降低了模型學習的難度。  


3.殘差塊使用bottleneck形式，將一個3x3卷積核拆分成一組(1x1,3x3,1x1)卷積核，  
可減少模型參數使用量。

4.綜合前面3點，ResNet保障了深度神經網路的可訓練性，  
並且在深度模型上也可以避免使用過多參數量。

鑒於上述ResNet的優點，我們決定使用ResNet50為基礎來製作本次專案。我們挑選高跟鞋、運動鞋、靴子、拖鞋這4種鞋子作為辨識的種類。以4種鞋子各50張圖片，總共200張作為test資料，1603張作為training資料，200張作為Validation資料，共訓練350代，將辨識完的分類結果輸出為圖文，並顯示每一代訓練的過程及損失函數。在最後儲存最好的一代。



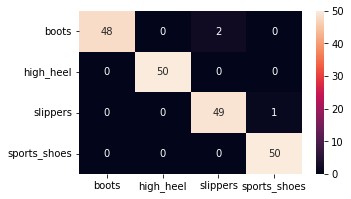
**圖3-2 ResNet50 架構**

* 將每一張照片做分類
* 辨識完的分類結果

**Accuracy : 99%**

**將靴子辨識成運動鞋**

**將運動鞋辨識成拖鞋**



**resnet50的結構**，可以分成幾個部分：

conv1：7x7卷積核(共1層)

conv2\_x：由三種卷積核(1x1,3x3,1x1)組成一個block，疊3次。(共9層)

conv3\_x：由三種卷積核(1x1,3x3,1x1)組成一個block，疊4次。(共12層)

conv4\_x：由三種卷積核(1x1,3x3,1x1)組成一個block，疊6次。(共18層)

conv5\_x：由三種卷積核(1x1,3x3,1x1)組成一個block，疊3次。(共9層)

output：由pooling+全連接層+softmax組成。共1層)

四、流程圖

辨識流程圖

進入畫面

拍照

選擇圖片

放入ImageView

上傳至伺服器的指定資料夾

使用ResNet架構進行辨識

辨識結果傳入資料庫

PHP抓取資料庫資料

手機端抓取PHP結果

結果放入TextView顯示

App使用流程圖

進入畫面

註冊

簡訊驗證

進行結帳

進行辨識

選擇商品

加入購物車

網頁流程圖

進入網站

選購商品

進入金流付款

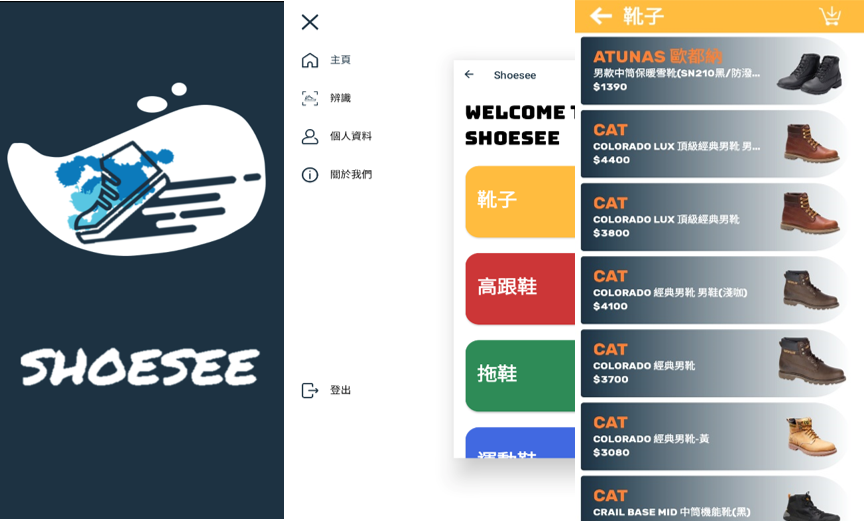
填寫購買資訊

加入購物車

查看訂單

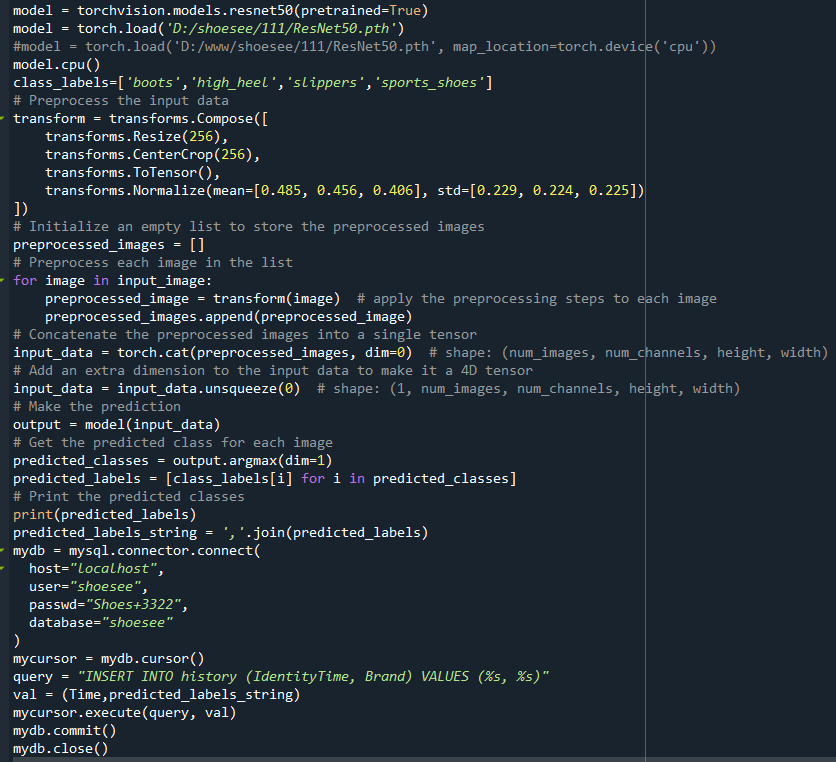
五、成果展示

app

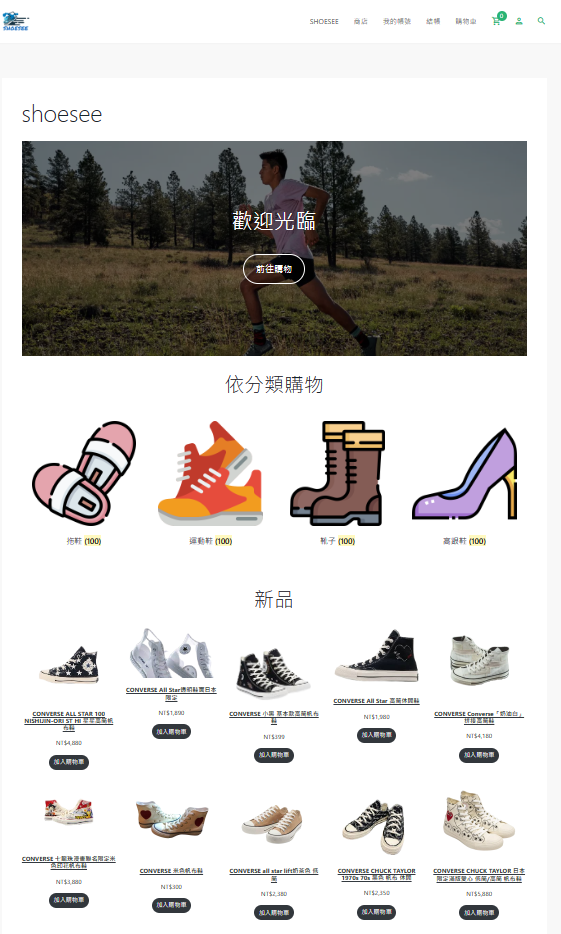




伺服器



網頁



# 五、遭遇問題與解決方案

Q1:在辨識裡常常有容易重複性錯誤的照片出現

A1:去觀察發現大多是因為切割而導致判斷錯誤，由於一開始並非等比例縮放而讓圖片常會有圖片切割過頭，改以比例縮放後情況有所感善了

Q2:在一開始篩寫ResNet辨識裡原本是使用自建模型，而後來改採用成的內建模型來去運用，但是想進行測試卻無法把訓練好的資料載入給他測試

A2: 後來是發現得使用torch.load函數加載它，不然是無法把訓練好的資料載入，否則就只會依照沒訓練得資料去跑。

六、工作分配與進度甘梯圖

* 陳振笙：資料蒐集、模型訓練及撰寫、報告撰寫、資料庫建置、APP撰寫、網頁撰寫
* 陳建安：資料蒐集、模型訓練及撰寫、報告撰寫、資料庫建置、伺服伺服器程式撰寫
* 葉士賢: 資料蒐集、模型訓練及撰寫、報告撰寫、資料庫建置、海報設計

**表1 進度甘梯圖**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 工作項目 | 6  月 | 7  月 | 8  月 | 9  月 | 10月 | 11月 | 12月 | 1  月 |
| 資料蒐集 |  |  |  |  |  |  |  |  |
| 模型訓練及撰寫 |  |  |  |  |  |  |  |  |
| 資料庫建置 |  |  |  |  |  |  |  |  |
| 伺服器程式撰寫 |  |  |  |  |  |  |  |  |
| APP撰寫 |  |  |  |  |  |  |  |  |
| 網頁撰寫 |  |  |  |  |  |  |  |  |
| 參考資料整理 |  |  |  |  |  |  |  |  |
| 專題LOGO |  |  |  |  |  |  |  |  |
| 版面配置 |  |  |  |  |  |  |  |  |
| 介面美化 |  |  |  |  |  |  |  |  |
| 測試與回報 |  |  |  |  |  |  |  |  |
| 海報製作 |  |  |  |  |  |  |  |  |

七、結論與未來展望

在這次的專題裡面我們嘗試了各種模型，也調適了不一樣的數據來去運行，從簡單的模型到複雜的模型都有所接觸，了解到簡單的模型所跑的結果不一定會比較差，複雜的模型也不一定會比較好，模型的選擇會依照不同的數據跟不同的數量都會有所改變。

而我們的數據量不多，再嘗試ResNet151後，其結果與ResNet50其實並不會差距過大，而後確定以ResNet50為最終模型。

最後是希望我們在之後可以增加鞋子辨識種類，也一併把資料庫圖片數量擴增，然後幫App增加金流系統，也為網頁新增辨識系統，使整個專題更加完善

而目前在伺服器端是可以上傳圖片，但是由於安全性等問題並未成功執行python辨識，未來會更加學習伺服器的運作來解決此問題

八、參考文獻

訓練模型與伺服端

1. [Residual Leaning: 認識ResNet與他的冠名後繼者ResNeXt](https://medium.com/%E8%BB%9F%E9%AB%94%E4%B9%8B%E5%BF%83/deep-learning-residual-leaning-%E8%AA%8D%E8%AD%98resnet%E8%88%87%E4%BB%96%E7%9A%84%E5%86%A0%E5%90%8D%E5%BE%8C%E7%B9%BC%E8%80%85resnext-resnest-6bedf9389ce)

2. [ResNet及其变种的结构梳理、有效性分析与代码解读 - 知乎专栏](https://zhuanlan.zhihu.com/p/54289848)

3. [[Day 10] 從tensorflow.keras 開始的ResNet 生活 - iT 邦幫忙](https://ithelp.ithome.com.tw/m/articles/10264843)

4. [在Android Studio 3.x版開發Android系統的開發記事-如何在Android上傳檔案(中) - iT 邦幫忙::一起幫忙解決難題，拯救 IT 人的一天 (ithome.com.tw)](https://ithelp.ithome.com.tw/articles/10203278)

App

1. https://www.youtube.com/@CodingwithT
2. https://www.youtube.com/@eddydn71
3. https://www.youtube.com/@CalvinHuo1004

網頁

1. https://www.youtube.com/@wpdog
2. https://www.youtube.com/@ProgressBarTW
3. https://frankknow.com/what-is-wordpress/