

國立雲林科技大學
機器學習專案作業二

使用卷積神經網路進行影像分類預測

成員：

M11021009 吳宥霆

M11021028 劉軒瑋

M11021035 黃堉豪

M11021052 邱守燦

指導教授：

許中川 教授

日期：

2022 年 5 月 19 日

摘要

近年來，人工智慧的技術突飛猛進，深度學習的方法日新月異，影像辨識的應用十分廣泛，在實務上，卷積神經網路與其他延伸模型讓圖像與影像的分類與預測更加精確且快速。本研究應用卷積神經網路針對 CIFAR-10 資料集的十個類別進行分類，以及應用 VGG16 的卷積延伸模型對 New Rice 資料集中的台米 81 號至 83 號進行分類。經過模型訓練與優化，從分類報告的結果得知，CIFAR-10 資料集的預測準確率為 72%，New Rice 資料集的預測準確率為 97%。

關鍵字：稻米、影像辨識分類、卷積神經網路、VGG-16

一、緒論

1.1 動機

1.1.1 CIFAR-10 Dataset

CIFAR-10 數據集是通常用於訓練機器學習和計算機視覺算法的圖像的集合。也是機器學習研究中使用最廣泛的數據集之一。而且 CIFAR-10 含有許多真實世界中的物體，不僅噪點很多，而且物體的比例、特徵都不盡相同，這使得識別物體上帶來相當程度的困難。

1.1.2 New Rice Dataset

稻米為台灣最主要的農作物之一，也是世界上不可或缺的糧食種類之一，從糧食安全而言，稻米維持自給自足，能夠使人民安心且避免有糧食危機問題。而對於經濟而言，稻米的產量對國內經濟佔有一定程度的重要性，若能對稻米進行相關的預測分類，將糧食政策制定與經濟發展規劃有正面的影響。

1.2 研究目的

1.2.1 CIFAR-10 Dataset

選擇此資料集的目是，為練習與學習卷積神經網路相關技術的基礎運用，本組選用了 CIFAR-10 料集，並使用 CNN 卷積神經網路架構進行分類預測，投入的類別資料有飛機、交通工具、鳥、貓...等 10 種資料，並透過績效指標(Accuracy、Precision、Recall、f1-score)來衡量資料集辨識的準確性，從中了解卷積神經網路的運作。

1.2.2 New Rice Dataset

稻米為許多國家的經濟產物，而稻米不僅會影響一個國家的經濟規模，也會影響到全世界糧食上的供應問題，本組根據作業提供之 New Rice 資料集，裡面包含台農 81 號、台農 82 號、台農 83 號三種稻米影像資料，採用 Keras 的 VGG16 影像辨識技術進行分類，來分類台農稻米之影像。藉此希望可以提供地方政府或是農會等重要農產機構，一準確辨識與分類的技術應用，而且此一技術可適用的層面極廣，未來也可套用在其他產業或領域中。

二、資料集

2.1 資料集

2.1.1 CIFAR-10 Dataset 說明

CIFAR-10 資料集共有 10 個類別，每種類別分別包含 5000 張訓練資料與 1000 張測試資料，每張圖為 32*32 像素的彩色影像。

- 類別數量：10
- 類別標籤：Airplane、Automobile、Bird、Cat、Deer、Dog、Frog、Horse、Ship、Truck
- 訓練資料：50000 張
- 測試資料：10000 張

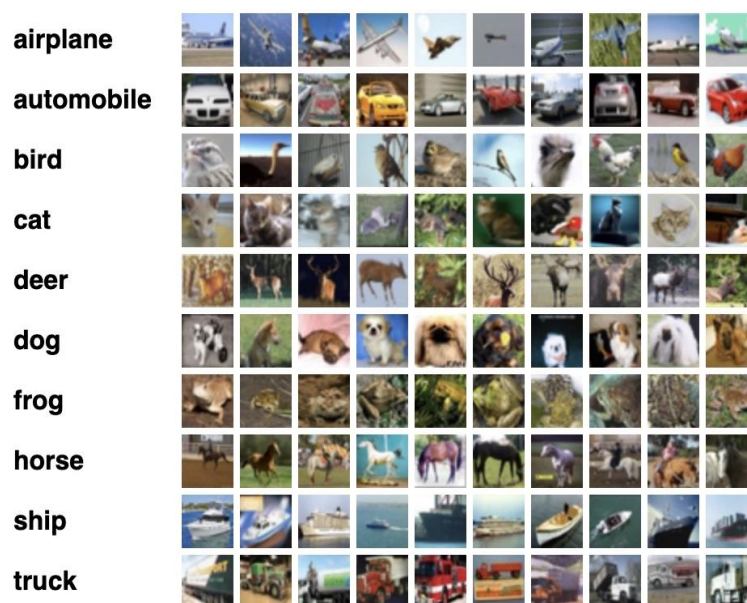





圖 1 CIFAR-10 資料集

2.1.2 New Rice Dataset 說明

New Rice 資料集共有 3 個類別，每種類別分別包含 450 張訓練資料與 50 張測試資料，每張圖為 4032*3024 像素的彩色影像。

- 類別數量：3
- 類別標籤：Tai-Nong No.81、Tai-Nong No.82、Tai-Nong No.83
- 訓練資料：1350 張
- 測試資料：500 張

表 1 New Rice 資料集

Tai-Nong No.81	Tai-Nong No.82	Tai-Nong No.83
		

三、方法

3.1 程式架構

3.1.1 CIFAR-10 Dataset

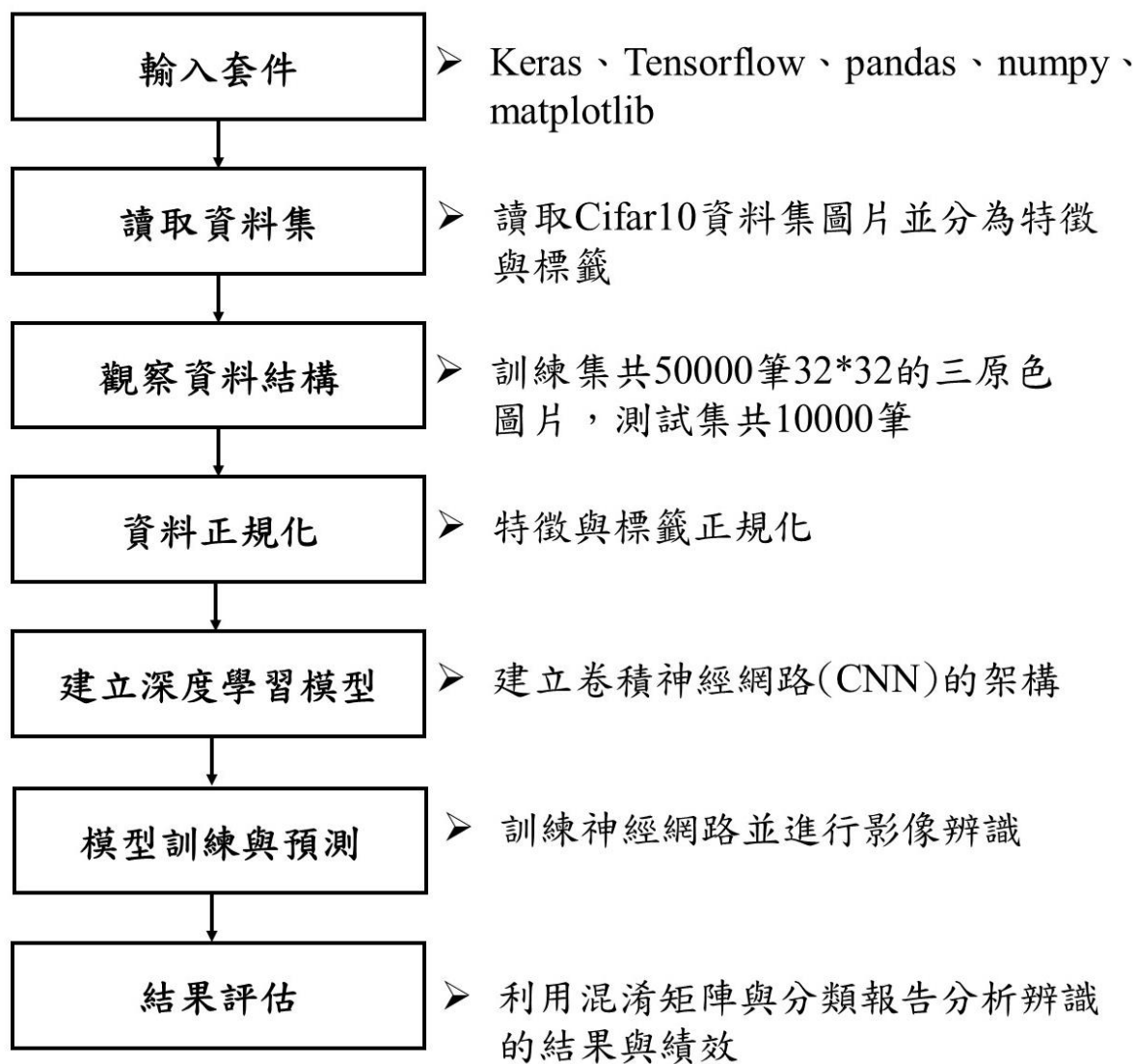


圖 2 CIFAR-10 Dataset 之程式架構流程圖及說明

3.1.2 New Rice Dataset

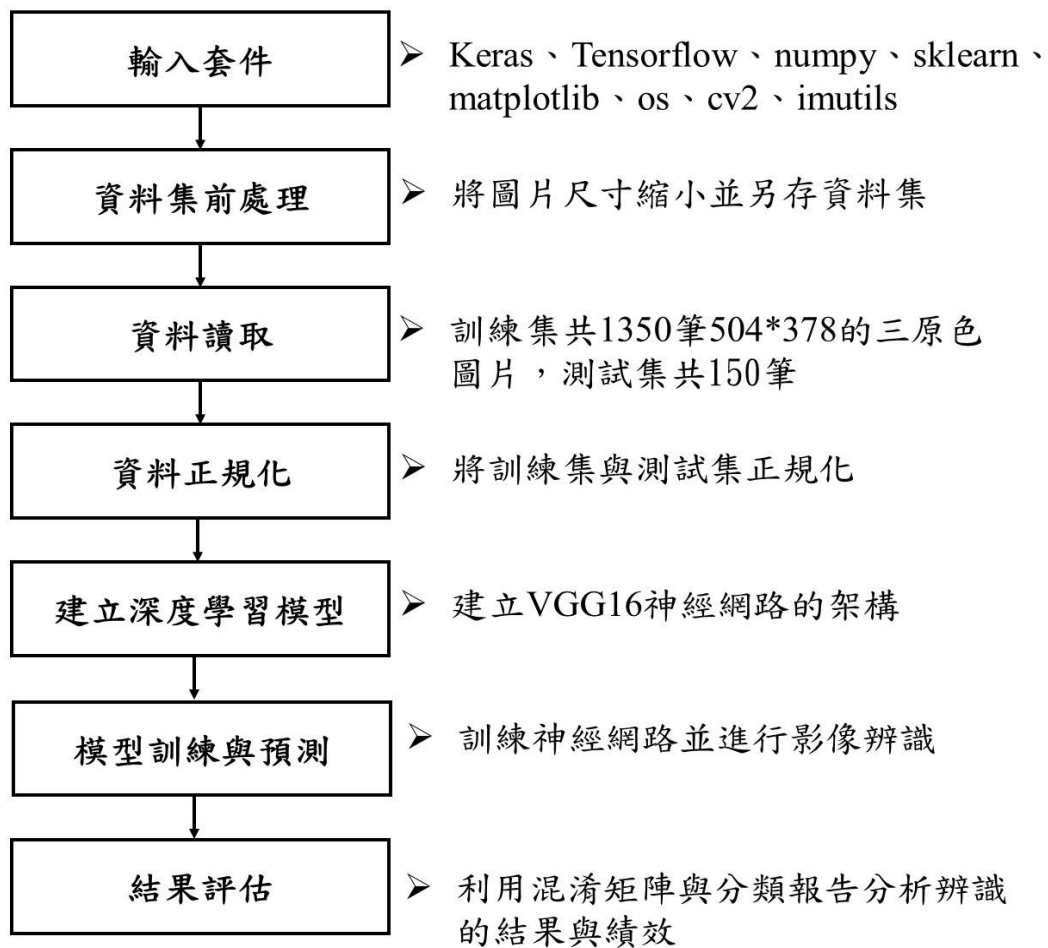


圖 3 New Rice Dataset 之程式架構流程圖及說明

3.2 程式執行方法

3.2.1 卷積神經網路（Convolutional Neural Network, CNN）

卷積神經網路（Convolutional Neural Network, CNN）是一種前饋式神經網路，它的人工神經元可以回應一部分覆蓋範圍內的周圍單元，對於大型圖像處理有出色表現。CNN 在影像識別方面的威力非常強大，許多影像辨識的模型也都是以 CNN 的架構為基礎去做延伸。其卷積神經網路由一個或多個卷積層和頂端的全連通層組成，同時也包括關聯權重和池化層。此一結構使其能夠利用輸入資料的二維結構，而且卷積神經網路需要考量的參數更少，使其成為深受眾人喜愛的深度學習架構之一。

3.2.2 VGG-16 卷積神經網路模型

VGG 是英國牛津大學 Visual Geometry Group 的縮寫，主要貢獻是使用更多的隱藏層，並且大量的圖片訓練，提高準確率至 90%。其架構為 16 層（13 個卷積層

及 3 個全連接層），而 VGG 的結構非常簡潔，整個網絡都使用了同樣大小的卷積核尺寸（3*3）和最大池化尺寸（2*2）。到目前為止，VGG 依然經常被用來提取圖像特徵。因為提供了非常好的初始化權重，所以被應用在許多地方。

四、實驗

4.1 前置處理

4.1.1 CIFAR-10 Dataset

將訓練集的圖片一致規格化為浮點數，並除以 255 將圖片正規化；測試資料集使用 One-Hot-Encoding 將圖片類別的屬性轉換為數值。

4.1.2 New Rice Dataset

使用 OpenCV 與 imutils 套件讀取原始資料夾中圖片檔，並使用 resize 函數將原始圖片從 4032x3024 轉換成 504x378，最後匯出成 Train 與 Test 的資料集。

4.2 實驗設計

4.2.1 CIFAR-10 Dataset

- （1）套件載入：載入協助資料分析的 Pandas、Numpy 套件，並載入 Keras、Tensorflow 套件以利建構卷積神經網路、最後載入繪圖套件 matplotlib。
- （2）資料讀取：讀取訓練與測試資料集。
- （3）建立卷積神經網路：利用 keras 的 Sequential、Dense、Flatten、Conv2D、MaxPooling2D 套件建立卷積神經網路，卷積層與全連接層的激活函數以'Relu'、輸出層（output layer）的激活函數以'Softmax'為主，損失函數（loss function）利用 categorical_crossentropy 進行運算，優化器設定 Adam，最後使用分類問題的評估績效（Accuracy）來檢測模型的成效。

表 2 CIFAR-10 CNN 模型架構表

Layer (type)	Output Shape	Param#
Conv2D (Conv2D)	(None,32,32,32)	896
Max_pooling2D (Max_Pooling2D)	(None,16,16,32)	0
Conv2D_1 (Conv2D)	(None,16,16,64)	18496
Max_pooling2D_1 (Max_Pooling2D)	(None,8,8,64)	0
Conv2D_2 (Conv2D)	(None,8,8,128)	73856
Max_pooling2D_2 (Max_Pooling2D)	(None,4,4,128)	0
flatten (Flatten)	(None, 2048)	0
dense (Dense)	(None, 256)	1049088
dense (Dense)	(None, 10)	5130
Total params : 1,147,466		
Trainable params : 1,147,466		
Non-trainable : 0		

4.2.2 New Rice Dataset

- (1) 套件載入：載入讀取資料的 OS 套件，並載入 Keras、Tensorflow 套件以利建構卷積網路 VGG-16。
- (2) 正規化：使用 ImageDataGenerator 進行正規化
- (3) 建構卷積 VGG-16 網路：使用 Keras 載入預訓練之 VGG-16 的模型，Input Shape 設定為 504*378*3，Weights 為 imagenet，激活函數在網路層為 Relu、輸出層為 Softmax，最後輸出分類類別為 3。
- (4) 模型訓練：使用 RMSprop 作為優化器，並使用 categorical_crossentropy 評估 loss，訓練次數 (epochs) 設定 20 次，批次更新的數量 (batch-size) 設定為訓練 50 次、測試 20 次。
- (5) 績效評估：使用 confusion_matrix 與 classification_report 函數，以混要矩陣與評估指標 (精確率 Precision、召回率 Recall、F1 分數 F1-score) 評估模型分類的績效。

表 3 New Rice Dataset VGG-16 模型架構表

Layer (type)	Output Shape	Param#
vgg16 (Functional)	(None, 15, 11, 512)	14714688
flatten (Flatten)	(None, 84480)	0
dense (Dense)	(None, 256)	21627136
Dense_1 (Dense)	(None, 3)	771
Total params : 36,342,595		
Trainable params : 36,342,595		
Non-trainable : 0		

4.3 實驗結果

4.3.1 CIFAR-10 Dataset 結果預測績效評估

利用精確率 (Precision)、召回率 (Recall)、F1 分數 (F1-score) 進行分類績效評估，評估 CIFAR-10 模型的準確度。

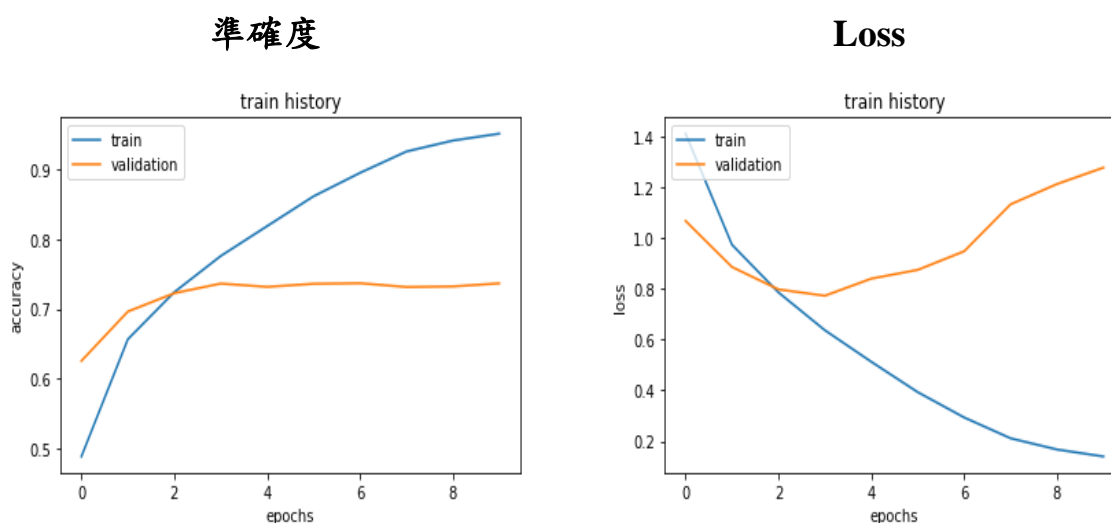


圖 4 CIFAR-10 準確度與損失函數

表 4 CIFAR-10 分類績效評估總表

類別	Precision	Recall	F1-score
0	0.76	0.77	0.77
1	0.84	0.83	0.83
2	0.61	0.64	0.62
3	0.50	0.56	0.53
4	0.69	0.67	0.68
5	0.67	0.53	0.59
6	0.72	0.86	0.78
7	0.83	0.75	0.79
8	0.86	0.82	0.84
9	0.80	0.81	0.80

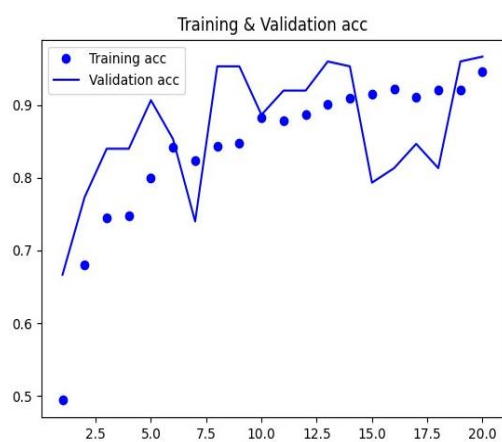
Prediction	0	1	2	3	4	5	6	7	8	9
label										
0	816	18	37	18	11	4	3	12	35	46
1	29	826	9	5	3	3	2	8	23	92
2	77	7	620	82	84	34	29	46	4	17
3	35	19	84	565	57	114	38	41	16	31
4	23	5	82	78	641	18	31	105	8	9
5	22	8	70	222	49	534	14	58	11	12
6	18	17	76	90	43	28	695	11	11	11
7	21	5	30	27	55	39	3	802	3	15
8	90	53	17	16	7	3	4	2	782	26
9	41	91	10	16	3	1	6	19	20	793

圖 5 CIFAR-10 混淆矩陣圖

4.3.2 New Rice Dataset 分類績效評估

利用精確率（Precision）、召回率（Recall）、F1 分數（F1-score）進行分類績效評估，評估 New Rice 模型的準確度。

準確度



Loss

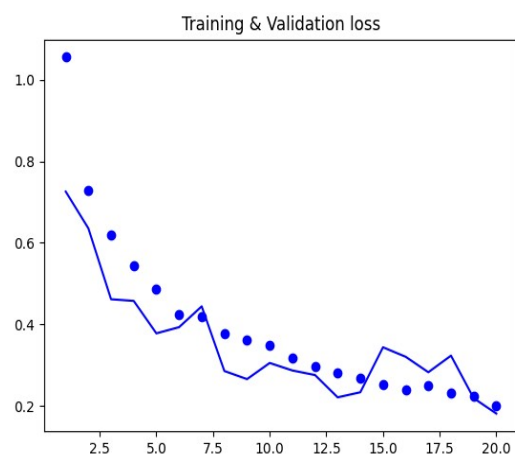


圖 6 New Rice 準確度與損失函數

表 5 New Rice 分類績效評估總表

類別	Precision	Recall	F1-score
0	0.96	1.00	0.98
1	0.96	0.94	0.95
2	0.98	0.96	0.97

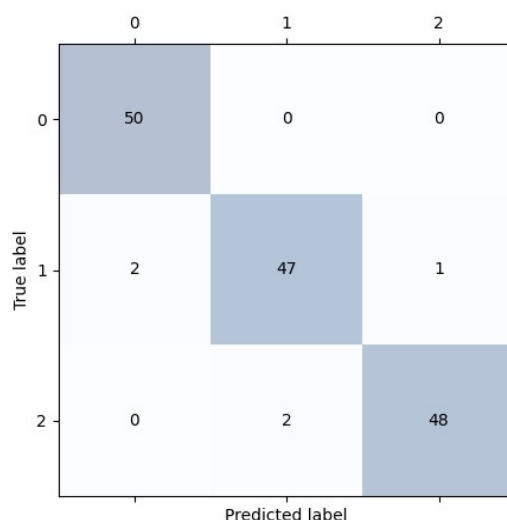


圖 7 New Rice 混淆矩陣圖

五、結論

在 CIFAR-10 資料集中，包含了 10 種不同類別的物件，在該物件識別任務中預測正確數量為 7074 張，錯誤張數為 2926 張，準確率達 72%，透過表 4 和圖 5 綜合評估，該分類器可預測出大部分圖像之物件。

在本研究中使用臺農稻米資料集與 CIFAR-10 資料集，當作分類預測實驗資料集，並分別使用 CNN 與 VGG16 預訓練模型，再分別建構分類器模型，以完成分類預測之任務。在臺農稻米資料集中，進行分類的預測任務，期望能藉由 VGG16 模型來判別出臺農稻米的種類，該集料集包含了臺農 81 號、臺農 82 號以及臺農 83 號三種不同種類的稻米，在該實驗中預測正確之數量為 145 張，錯誤張數為 5 張，準確率高達 97%，透過表 5 和圖 7 綜合評估，該分類器可用於臺農稻米分類任務中，判別出大部分的稻米種類。

參考文獻

[1] Convolutional Neural Network (CNN) | TensorFlow Core

<https://www.tensorflow.org/tutorials/images/cnn>

[2] Day 08：CNN 模型設計 - iT 邦幫忙

<https://ithelp.ithome.com.tw/articles/10192028>

[3] [資料分析&機器學習] 第 5.1 講: 卷積神經網絡介紹(Convolutional Neural Network)

<https://medium.com/jameslearningnote/%E8%B3%87%E6%96%99%E5%88%86%E6%9E%90-%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-%E7%AC%AC5-1%E8%AC%9B-%E5%8D%B7%E7%A9%8D%E7%A5%9E%E7%B6%93%E7%B6%B2%E7%B5%A1%E4%BB%8B%E7%B4%B9-convolutional-neural-network-4f8249d65d4f>

[4] 實作 Tensorflow (3)：Build First Convolutional Neural Network (CNN)

https://www.ycc.idv.tw/tensorflow-tutorial_3.html

[5] tensorflow 教學(4) -- 建置一個 CNN 網路分辨手寫辨識字

<https://darren1231.pixnet.net/blog/post/332753859-tensorflow%E6%95%99%E5%AD%B8----%E5%BB%BA%E7%BD%AE%E4%B8%80%E5%80%8Bcnn%E7%B6%B2%E8%B7%AF%E5%88%86%E8%BE%A8%E6%89%8B%E5%AF%AB%E8%BE%A8%E8%AD%98>

[6] VGG16 and VGG19 - Keras

<https://keras.io/api/applications/vgg/>

[7] [Day 09] 從 tensorflow.keras 開始的 VGG Net 生活 (第二季)

<https://ithelp.ithome.com.tw/articles/10265331>

[8] Keras 手動搭建 VGG 卷積神經網絡識別 ImageNet1000 種常見分類

<https://codertw.com/%E7%A8%8B%E5%BC%8F%E8%AA%9E%E8%A8%80/656351/>

[9] 【遷移學習】案例：Keras 基於 VGG 對五種圖片類別識別

<https://chowdera.com/2021/08/20210816030426672w.html>

[10] 【Day 29】 Google ML - Lesson 13 – 以混淆矩陣(confusion matrix)分析 ML 模型好壞,可評估的三種指標 Accuracy, Precision, Recall

<https://ithelp.ithome.com.tw/articles/10222697>

[11] sklearn.metrics.classification_report

https://scikit-learn.org/stable/modules/generated/sklearn.metrics.classification_report.html

[12] 淺談機器學習的效能衡量指標 (1) -- 準確率(Accuracy)、精確率(Precision)、召回率(Recall)

<https://ithelp.ithome.com.tw/articles/10228941>