

物件偵測系統

Object Detection System

陳宣蓉 周少宇 李宥樑 王姿婷 張勤振

國立聯合大學 資訊工程學系

苗栗市南勢里聯大二號

ccchang@nuu.edu.tw

摘要

要如何才能讓機器的判斷、邏輯趨近於人類呢？依照目前的各種研究與實作來看，擁有龐大的資料庫是不可或缺的。機器學習可以藉由這個龐大的資料庫，以人類觀點來做出預測，告訴我們這個事物為何。為了讓機器人能對於我們現實生活中的事物做出物件偵測並較準確的辨識出事物的名稱，我們必須先對於我們的機器做訓練。在給予龐大的資料的情況下，讓機器在訓練的過程中找出物件的特徵，當有新的資料時，機器透過已經完成的訓練模型，判斷該物件為何的機率最高，並給出預測結果。

關鍵詞：機器學習、物件偵測、資料庫、訓練模型、預測結果。

Abstract

How to make the machine approaching human-like capacity for judgment and logic? Currently, many research and implementation indicate that we should have huge database. Machine learning could use the database to execute detection which like the human's

view and tell us the thing. In order to make the machine identifying the things more accurately in real life. We need to training our machine. We give the huge data to machine and tell the object's feature to make the machine find itself during machine training. When we input the new data to the machine, the machine through the training model which has been completed to judgment the object and output the forecast result.

Keyword: machine learning, object detection, database, training model, forecast result

一 簡介

利用 Digits 進行數字的資料訓練，建立模型，並測試模型與預測資料，完成手寫數字的辨識。還有在 darknet 上執行 yolo 進行各種圖片的辨識，以及影像辨識。

(1) 開發環境：

Linux：

以桌面應用為主的 Linux Ubuntu 16.04LTS 64bit 發行版，是開源軟體。

CUDA 8.0 :

利用 GPU 作為 C 編輯器的開發環境。CUDA 可以相容 OpenCL 或者自家的 C 編譯器。無論是 CUDA C 語言或是 OpenCL，指令最終都會被驅動程式轉換成 PTX 代碼，交由顯示核心計算。

cuDNN 5.1 :

是一個為深層神經網路設計的 GPU 加速原式函式庫，提供經過調校的常式建置方法。(註明：參考文獻 [7])

DIGITS :

整合現有的 Deep Learning 開發工具，實現深度神經網路設計、訓練和可視化等任務變得簡單化。

Darknet :

開源的神經網路框架，由 C 和 CUDA 編寫，也可以幫助 CPU 和 GPU 進行運算。

(2) 系統架構：

使用者利用已經訓練好的模型將要預測的資料丟進去，產生結果。模型由大量已經經過訓練的資料產生。

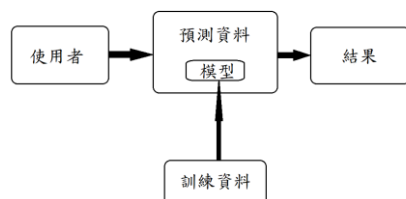


圖 2-1：系統架構圖。

二 系統功能

(1) 數字辨識：

對於手寫數字會自動辨識出來。

(2) 圖片辨識：

輸入圖片會使機器自動辨識出圖片上的物件，如果圖片上的物件的訓練模型被訓練過。

(3) 影像辨識：

將影片的影格做圖片辨識，快速輸出到畫面上。

三 專題理論

(1) 人工智慧：

人工智慧起初的目標是希望能讓電腦像人一般思考與學習，人工智慧的分類發誓可分為強人工智慧(Strong AI)和弱人工智慧(Weak AI)。(註明：參考文獻 [1])

(2) 機器學習：

機器學習是人工智慧的分支，機器學習是透過演算法，使用大量資料來進行訓練，訓練完成後會產生模型，未來當有新的資料，便可以使用訓練好的模型進行預測。(註明：參考文獻 [1])

(3) 類神經網路：

目的主要是去嘗試模仿人類的大腦的神經系統。人類神經元傳導運作很複雜，可以將其導成以下數學函式(1)。(註明：參考文獻 [1])

$y = \text{activation function}(x_1 * w_1 + x_2 * w_2 + x_3 * w_3 + b_1)$ 數學函式(1)

(4) 深度學習卷積神經網路：

卷積神經網路目前是深度神經網路領域發展的主力，在圖片的判斷上可以做到比人類更加精確。卷積運算的效果很類似濾鏡的效果，擷取了不同的特徵。經過卷積運算，產生多個影像，就好像將相片卷積起來。(註明：參考文獻 [1])

四 專題實作

(1) 數字辨識：

模型訓練完成後，開始做測試，測試完成後即可開始預測資料，我們使用圖 4-1 與圖 4-2 來做預測。預測完成後，圖 4-1 會如圖 4-3 預測此數字為 2 的機率 100%，圖 4-2 會如圖 4-4 預測此數字為 3 的機率 99.7%。



圖 4-1：數字 2。



圖 4-2：數字 3。

圖 4-7，預測此數字為 2 的機率 100%。

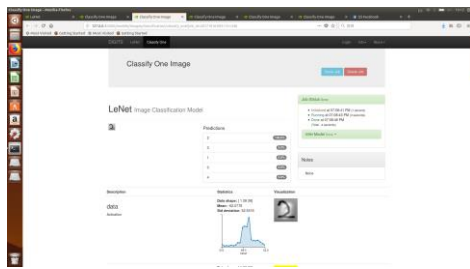


圖 4-3：數字 2 的機率。

圖 4-8，預測此數字為 3 的機率 99.7%。

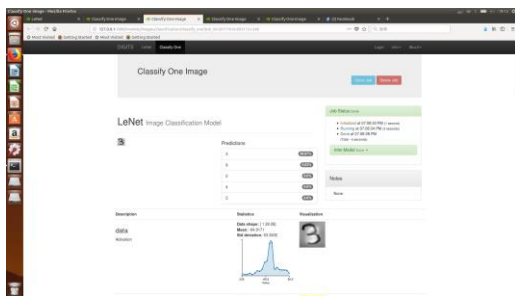


圖 4-4：數字 3 的機率。

(2) 圖片辨識：

圖片辨識會因為許多的原因而導致其結果有誤，如相近的物件，像圖 4-5。

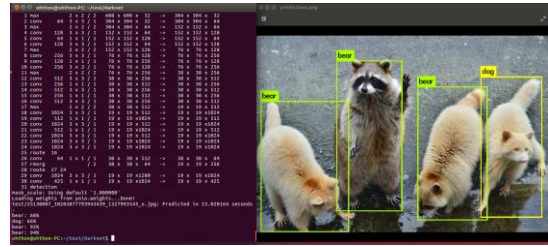


圖 4-5：浣熊圖片辨識。

大部分的浣熊被辨識成熊，這樣是不準確的，但對已經訓練好的神經網路來說，這種物種相似於熊，而被歸類成熊，他沒辦法區分熊跟浣熊，因為他沒有學習過浣熊，但是因為與熊相近所以，被辨識成熊。

另一個因素是訓練資料的數量，如果訓練的資料量太少，就會無法完整的讓神經網路訓練出針對該物件而產生的權重。

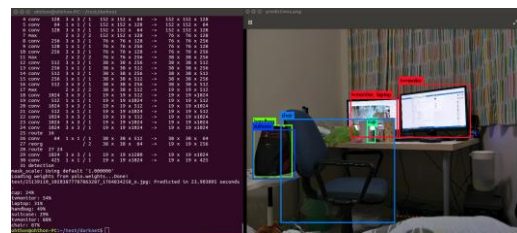


圖 4-6：家裡環境圖片辨識。

如圖 4-6 中有很多個杯子，但是只有一個杯子被辨識出來。

而標記的品質也是一個影響準確率的很大要素。如圖 4-7。

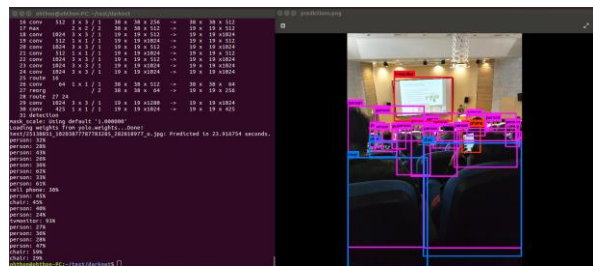


圖 4-7：演講廳後方拍攝並進行圖片辨識。

(3) 影像辨識：

因為硬體限制的關係，所以影像辨識的成果並沒有很好，影像的偵數只有每秒 2.5 偵，還有嚴重的延遲，將影片的影格做

圖片辨識，快速輸出到畫面上，一般人的肉眼大約在 30fps 以上的就分不太出來了，所以達到這個標準後，就比較不會有畫面延遲的問題。

這是經由 Webcam 截圖下來，如圖 4-8 與圖 4-9。

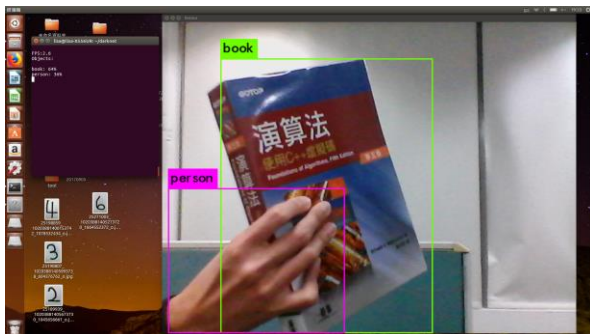


圖 4-8：攝像機畫面進行即時影像辨識。



圖 4-9：攝像機畫面進行即時影像辨識。

五 結論

在現今深度學習發展得越來越迅速，很多深度學習所辨識出的結果甚至可以超越人的肉眼辨識，在大環境的引導之下，尤其是硬體的進步讓深度學習成為最熱門的科技之一，機器學習的原理其實並不難，只是這世界上的事物有成千上萬種，必須不停的給機器學習吸收那些資料，產生龐大的資料庫，並讓機器有著可以處理這些龐大資料庫函式。才是我們一直要面對的課題。

參考文獻

[1] 林大貴，TensorFlow+Keras 深度學習人工智慧實務應用，初版，新北市，博碩

文化，民國一百零六年。

[2] GitHub, NVIDIA/DIGITS, <https://github.com/NVIDIA/DIGITS>。

[3] Hao, S., Deng, J., Fei-Fei, L.: Crowdsourcing annotations for visual object detection. In: AAAI(2012).

[4] Hung-yi Lee : Machine Learning, <http://speech.ee.ntu.edu.tw/~tlkagk/>, Fall, 2017.

[5] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi: You Only Look Once: Unified, Real-Time Object Detection. <http://pjreddie.com/yolo/>, 2013-2017

[6] J. Redmon. Darknet: Open source neural networks in c. <http://pjreddie.com/darknet/>, 2013-2017.

[7] NVIDIA Taiwan, cuDNN, <https://blogs.nvidia.com.tw/2014/09/accelerate-machine-learning-cudnn-deep-neural-network-library/>。

[8] Wikipedia, CUDA, <https://zh.wikipedia.org/wiki/CUDA>。

[9] Wikipedia, Ubuntu, <https://zh.wikipedia.org/wiki/Ubuntu>。