

III 期末報告-骰子圖像辨識

108034538 杜彥陵

摘要

本研究針對網路上所提供之骰子圖像數據集進行數值的識別，運用卷積神經網路的方式來建立模型，透過此方法來對圖像進行辨識。

一、研究動機與目的

在現在科技如此進步的社會，對於數值、圖像及聲音等等的識別，已經不需要透過人類來執行，皆以人工智慧的方式來做處理，以降低判讀錯誤的可能性，以此提高辨識的準確率。因此本研究以透過建立卷積神經網路的模型對骰子圖像進行辨識，以作為卷積神經網路的應用例子。

二、文獻回顧

1. 圖型識別(Pattern recognition)

圖型識別(Pattern recognition)，就是通過電腦用數學技術方法來研究圖型的自動處理和判讀。我們把環境與客體統稱為「圖型」。隨著電腦技術的發展，人類有可能研究複雜的資訊處理過程。資訊處理過程的一個重要形式是生命體對環境及客體的辨識。對人類來說，特別重要的是對光學資訊（通過視覺器官來獲得）和聲學資訊（通過聽覺器官來獲得）的辨識。這是圖型識別的兩個重要方面。市場上可見到的代表性產品有光學字元辨識、語音辨識系統。

電腦辨識的顯著特點是速度快、準確性高、效率高，在將來完全可以取代人工錄入。

2. 卷積神經網路(CNN)

卷積神經網路(Convolutional Neural Network, CNN)是一種前饋神經網路，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。

卷積神經網路由一個或多個卷積層和頂端的全連通層（對應經典的神經網路）組成，同時也包括關聯權重和池化層(pooling layer)。這一結構使得卷積神經網路能夠利用輸入資料的二維結構。與其他深度學習結構相比，卷積神經網路在圖像和語音辨識方面能夠給出更好的結果。這一模型也可以使用反向傳播演算法進行訓練。相比較其他深度、前饋神經網路，卷積神經網路需要考量的參數更少，使之成為一種頗具吸引力的深度學習結構。

3. 遷移學習(Transfer Learning)

遷移學習是屬於機器學習的一種研究領域。它專注於存儲已有問題的解決模型，並將其利用在其他不同但相關問題上。比如說，用來辨識汽車的知識（或者是模型）也可以被用來提升識別卡車的能力。計算機領域的遷移學習和心理學常常提到的學習遷移在概念上有一定關係，但是兩個領域在學術上的關係非常有限。

三、研究方法

本研究針對骰子的圖像進行圖像的辨識，透過建立卷積神經網路的模型，利用卷積神經網路對圖像識別的技術來判定圖像的骰子數值，並計算模型學習的準確率。

四、案例研究

本研究針對網路上所提供之骰子圖像數據集運用卷積神經網路的方式對圖像進行辨識。

首先，先將骰子數據集依造各自的點數分類在不同的資料夾中，並依造數值為資料夾命名，以此作為數據的標籤，以利後續的資料處理。第二，將圖像重新以 128 乘以 128 的維度做整理，並將各個數值的圖像以 75% 及 25%，各為 300 張及 100 張分成訓練數據集及測試數據集。第三，透過標準化將所有的圖像值縮小到介於 0 與 1 之間。第四，將數據集打亂以避免卷積神經網路偷吃步。第五，建立卷積神經網路模型，模型結構如圖一所示。第六，在多次的參數及

層數的修改調整中，皆對該模型進行訓練(訓練次數為十五次)及測試，以挑選出準確率較高的模型最為最終模型。

CNN parameter	Value
Number of convolution layers	3
Filter size	3X3
Number of pooling layers	2 max-pooling
Pooling filter size	2X2
Number of feature maps	3/6/9
Number of fully connected layers	1
Number of nodes in the fully connected layer	6
Activation function	ReLU
Regularization method	Dropout
Classification function of the output layer	Softmax
Loss function	sparse_categorical_crossentropy

(圖一) 卷積神經網路模型

五、結論

1. 研究結果

透過調整過濾器(Filter)的數量、激活函數(activation function)的種類、模型的卷積層(convolution layer)的層數等，來比較模型在該參數組合下的準確率。最後，在多次的調整後(調整過程如圖二)，發現模型沒有過度擬合，且準確率相較於其他參數組合之模型高，所以以圖一之模型結構作為最終模型，並求得最終的準確率為 93.33%。

改善步驟	Accuracy
原始Model	0.61
原始Model+activation(sigmoid->relu)	0.16->0.61
原始Model+filter(3->6)	0.61->0.89
原始Model+一層Conv2D	0.61->0.91

(圖二) 調整參數之過程

本研究也另外找尋了其他十二張新的骰子圖像，並將其放入上方之模型進行辨識，發現其準確率只有 33%，因此本研究認為雖然可以透過上述調整參數的方式來提高模型的準確率，但由於本數據集之圖像較為簡單與單一，所以讓模型在學習時不具有多樣性，無法舉一反三，以至於造成新圖像的測試準確率不高。

2. 研究時遇到的困難及如何克服

在本次的期末報告中，因為本身沒有學過程式也沒有程式相關的基礎，所以沒有辦法理解程式碼的涵義，使得需要花費較多的時間在了解程式用語，才能開始撰寫程式碼。另外，因為一開始在建立模型時，建立了比較簡單且相對不嚴謹的模型，所以準確率並不高，而在調整模型層數及參數時，因為很難去界定設定多少數值會使得準確率提高，所以需要多次的嘗試及修改模型。後來我透過上網搜尋相關資訊及書籍資料來做參考，利用網路上提供的類似案例及助教教學資料作為輔助完成這次的期末報告。

六、未來發展

因為本次的期末報告所使用的資料集較為簡單，皆為黑白的圖像且背景及骰子相似度很高，所以在訓練模型及測試模型時，能達到較高的準確率，但若是替換其他較為複雜的數據集後，可能會使得準確率下降。為了避免這類問題發生可以提高訓練及測試模型的複雜度，讓卷積神經網路能學習到更多樣的數據。另外也可以透過更多的前處理，例如旋轉圖像增加數據數量，或是將圖像灰階化，讓圖像以黑與白兩色呈現，提高圖像中的顏色落差，以避免模型學錯重點，並提高模型準確率。

七、參考資料

1. 圖型識別. (2019, Dec 6).In Wikipedia, the free encyclopedia. Retrieved Dec. 31, 2019, from <https://zh.wikipedia.org/wiki/模式识别>
2. 卷積神經網路. (2019, Aug 13).In Wikipedia, the free encyclopedia. Retrieved Dec. 31, 2019, from <https://zh.wikipedia.org/wiki/卷积神经网络>
3. 遷移學習. (2019, May 18).In Wikipedia, the free encyclopedia. Retrieved Dec. 31, 2019, from <https://zh.wikipedia.org/wiki/迁移学习>
4. Dataset
<https://github.com/tomitomi3/DiceRecognitionDatasetForML>

5. Model

助教上課內容(W6-CNN)

Keras Documentation <https://keras.io/layers/convolutional/>

Tensorflow <https://www.tensorflow.org/>

6. New image

https://cdn.clipart.email/a4671e739e665dc3dbe557549d36b5d8_dice-clipart-1-clipartfest-3-cliparting-com-picturesque-learnfreeme_2453-1588.jpeg