National Tsing Hua University

Fall 2023 11210IPT 553000

Deep Learning in Biomedical Optical Imaging

Report

Author One1

謝瑋哲 清華大學動機系 碩士一年級 新竹 台灣

Student ID: 112033645

1. 介紹

本報告是基於清華大學光電所陳鴻文教授開設的深度學習於生醫光學影像之應用課程，這次的訓練資料是六種癌症組織細胞的分類，總資料近4000張的彩色醫學影像，而我多次尋ChatGPT，他都多次表明在較小的訓練資料集上，較簡單且淺層的訓練模型(尤其是resnet18)可能可以提升分類性能的有效性(圖1、2)‌，因此本報告我決定比較以淺層模型三個與深層的模型三個，來對比訓練效果以及最終時間與成果性價比。

1. 實驗
   1. ***資料***

資料主要有六類:腫瘤、基質、複合體、淋巴、碎片、黏膜，共3750張照片，分為:訓練2550張、驗證600張、測試600張。

* 1. ***模型***

模型皆來自pytorch模組的預設模型，初始權重也都是使用各模型的預設權重，總共訓練30個epoch，而比較模型淺層模型三個以及深層模型三個最為對比。

* + 1. *resnet18*

模型架構: 總共有18層，其中是17個卷積層(conv)與1層的全連接層(fc)，架構為兩個卷積層構成的殘差層基本塊(Residual Block)，兩個基本塊構成的殘差層Resnet塊。順序為卷積層、四個連續的Resnet塊，最後再接上一個全連接層。

* + 1. *AlexNet*

模型架構: 總共8層，5個卷積層和3個全連接層。前兩個卷積層後接最大池化層，第五個卷積層後接平均池化層，最後三個全連接層中的前兩層後接Dropout層。AlexNet使用ReLU作為激活函數，並在第一層和第二層卷積層後使用局部響應歸一化（Local Response Normalization）。

* + 1. *googlenet*

模型架構: GoogleNet 總共約有 22 層，每個 Inception 模組包括多個並行的捲積層和池化層。

* + 1. *ResNet50*

模型架構: 總共50層，由更多的殘差層基本塊構成。這些基本塊包含3個卷積層（不同於 ResNet18 的2個卷積層基本塊）。最後是全局平均池化層和全連接層。

* + 1. *EfficientNetB4*

模型架構:總共大概50層，架構使用了深度可分卷積和倒置殘差結構。包括多個倒置殘差塊，每個塊都具有一系列的卷積、激活和批量正規化操作。最後次全局平均池化層和全連接層。

* + 1. *DenseNet121*

模型架構:共有121層，由多個密集塊（Dense Block）和過渡層（Transition Layer）組成。每個密集塊內的層都與之前所有層直接相連。最後以全球平均池化層和一個全連接層結束。

1. 實驗結果與討論
   1. *訓練結果與分析*

在表1的訓練結果中，訓練效果最好的是 DenseNet121，最差的為AlexNet，雖然AlexNet毫不意外的是最差的，但他的模型是極其簡單的8層，也就是說在本組影線分類問題之上，隨意挑選模型即可達到一定程度上的分類，甚至我們如果考慮模型所花上的訓練成本，將準確度除上訓練的秒數作為簡單的效率評估，AlexNet反而是在最短時間得到最好的結果。

如果我們以30層為分隔點把模型分為兩組: 低層組與深層組，我們可以發現，在低層組中，層數與訓練時間是等比例的，層數越深，訓練得越久，而在高層組中，層數越深就不完全與時間成等比關係了，這可能是因為在某些深層模型中，透過參數共享和重複使用技術，即使層數增加，模型的參數總數和計算複雜度可能不會線性成長。這有助於在保持較高層數的同時，控制訓練時間和計算需求。，比較兩組結果，我們可以得出如果隨意選擇越深層的模型，更有可能訓練的比低層數模型訓練的好。

而如果加上所花上的時間而言，高層數的模型並未取得飛躍性的增長，甚至在選擇Resnet的模型架構之上，更深層的Resnet50效果竟不比Resnet18好，或許是因為更深的Resnet模型需要更多的資料來充分訓練。

最後，本次的實驗皆是使用IMAGENET1K\_V1作為預訓練權重，而這也導致了在圖3到圖8中，個模型的初始損失與準確度就有一定的水準，而大部分的模型都還是有在接下來的訓練之上多少提升一些準確度，但仔細觀察可以發現本次實驗超過90%的四個模型皆是在最一開始訓練時，準確度就已經來到了75%以上，所以初始權重也多少對最終訓練準確度有影響。

* 1. *結論*

總的來說，在本次實驗之中，隨意地選擇深層的模型確實更有可能比淺層模型較佳，但這是在隨意選擇模型架構的條件下，如果計算資源有限，或許選擇使用Resnet18就已經能取得不錯的結果，這也應證了ChatGPT在最一開始建議的模型，雖說生成式AI並不能全信，但在結果不確定的選擇上，也可以做為強而有力的參考。

1. 實驗圖表
   1. *結果表*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Name | test accuracy | test\_avg\_loss | Time cost | accuracy\_per\_second | loss\_per\_second | layer |
| ResNet18 | 93.0% | 0.307 | 148.5 秒 | 0.62626 % | 0.00207 | 18 |
| AlexNet | 85.0% | 0.543 | 95.5 秒 | 0.89005 % | 0.00568 | 8 |
| googlenet | 91.667% | 0.276 | 162.2 秒 | 0.56515 % | 0.0017 | 22 |
| DenseNet121 | 93.5% | 0.289 | 408.5 秒 | 0.22889 % | 0.00071 | 121 |
| ResNet50 | 89.833% | 0.276 | 434.3 秒 | 0.20685 % | 0.00063 | 50 |
| efficientnet\_b4 | 92.833% | 0.265 | 609.8 秒 | 0.15224 % | 0.00043 | 50 |

表1. 實驗結果

* 1. *結果圖*

以下為訓練後的accuracy and loss。

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

圖1、ChatGPT支持相對簡單，少層模型對話1

一張含有 文字, 螢幕擷取畫面, 字型, 文件 的圖片

自動產生的描述

圖2、ChatGPT支持相對簡單，少層模型對話2

一張含有 文字, 繪圖, 行, 圖表 的圖片

自動產生的描述

圖3、ResNet18 accuracy and loss

一張含有 文字, 繪圖, 行, 圖表 的圖片

自動產生的描述

圖4、AlexNet accuracy and loss

一張含有 文字, 繪圖, 行, 圖表 的圖片

自動產生的描述

圖5、googlenet accuracy and loss

一張含有 文字, 繪圖, 行, 圖表 的圖片

自動產生的描述

圖6、DenseNet121 accuracy and loss

一張含有 文字, 繪圖, 行, 圖表 的圖片

自動產生的描述

圖7、ResNet50 accuracy and loss

一張含有 文字, 繪圖, 行, 圖表 的圖片

自動產生的描述

圖8、EfficientNetB4 accuracy and loss