深度學習課程報告 –

AOI image recognition

**學生：吳玫萱 RE6091054**

1. **最後成績(上傳截止)**

**Public set：**

****

1. **報告大綱**

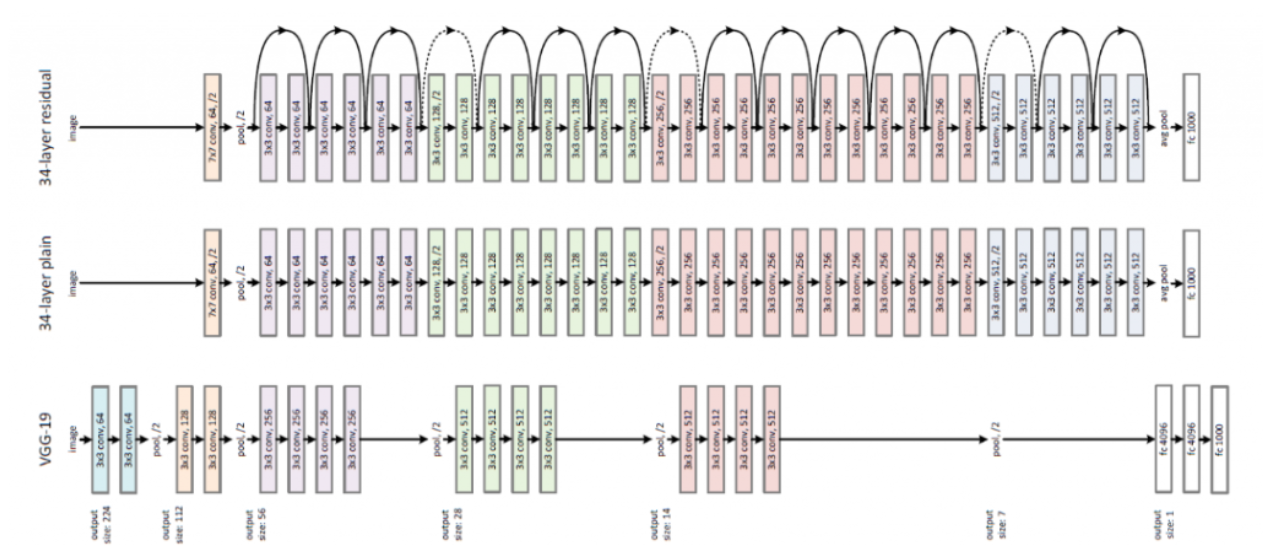
此次任務為針對AOI 影像資料的圖像分類，一共具有6類，分別是正常與其他五種瑕疵的圖像分類，其中AOI指的是Automated Optical Inspection ( 自動光學檢查，簡稱 AOI )，為高速高精度光學影像檢測系統，運用機器視覺做為檢測標準技術，可改良傳統上以人力使用光學儀器進行檢測的缺點，應用層面包括從高科技產業之研發、製造品管，以至國防、民生、醫療、環保、電力…等領域。該訓練賽是工研院發起的，工研院電光所投入軟性電子顯示器之研發多年，在試量產過程中，希望藉由 AOI 技術提升生產品質。目標是希望針對所提供的 AOI 影像資料，來判讀瑕疵的分類，藉以提升透過數據科學來加強 AOI 判讀之效能。

整個報告主要以ResNet進行嘗試，共嘗試ResNet18、34、50，三種ResNet架構，並嘗試不同參數，加入多種實驗設置，包含batch size調整、normalize與batch normalize 嘗試，更在ResNet50中使用變動的learning rate與Pre-Trained在ImageNet的參數作為初始化權重，以達到準確率的提升。

1. **模型介紹**

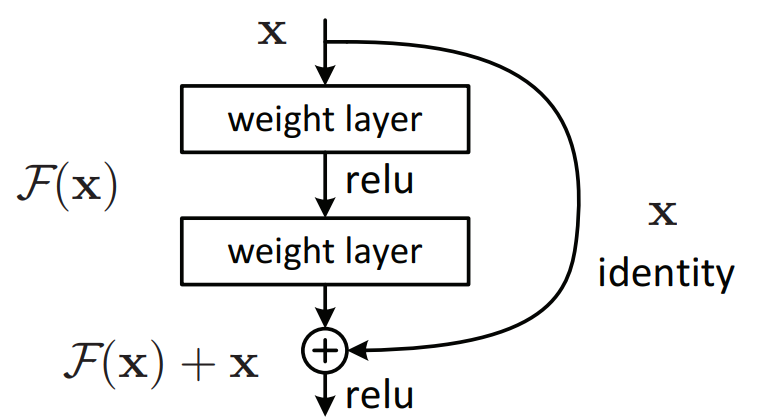
ResNet是一種以 residual block / residual learning為主架構的網路，其特色與主要貢獻為提出的residual learning使得深層網路更容易訓練，也開啟了各種超深網路的時代。

下圖一是ResNet和其他一般網路的比較圖。



圖一： ResNet和其他一般網路的比較圖

殘差學習解決了模型太深時產生的退化問題。對於一個堆積層結構（幾層堆積而成）當輸入為 x 時其學習到的特徵記為 H(x) ，現在我們希望其可以學習到殘差 F(x)=H(x)-x ，這樣其實原始的學習特徵是 F(x)+x 。之所以這樣是因為殘差學習相比原始特徵直接學習更容易。當殘差為0時，此時堆積層僅僅做了恆等映射，至少網絡性能不會下降，實際上殘差不會為0，這也會使得堆積層在輸入特徵基礎上學習到新的特徵，從而擁有更好的性能。殘差學習的結構如下圖二所示。這有點類似與電路中的「短路」，所以是一種短路連接（shortcut connection）。

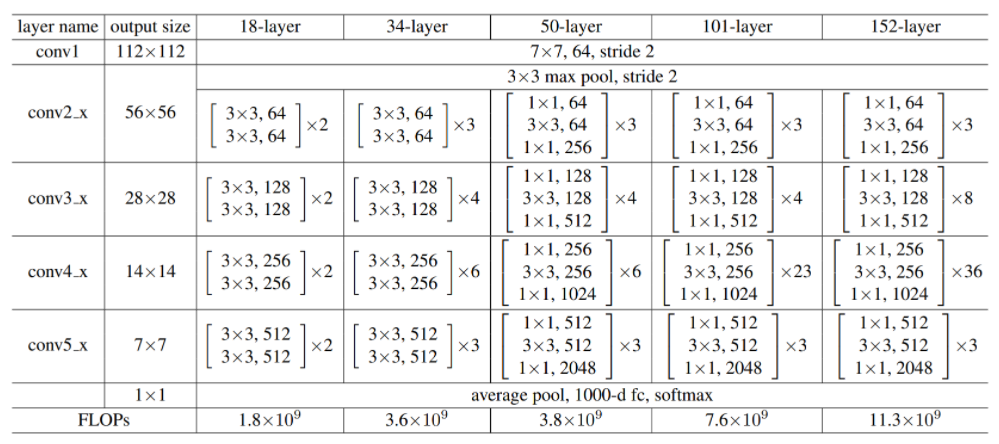


圖二：ResNet殘差學習結構圖

而不同深度的ResNet的差別有兩個：

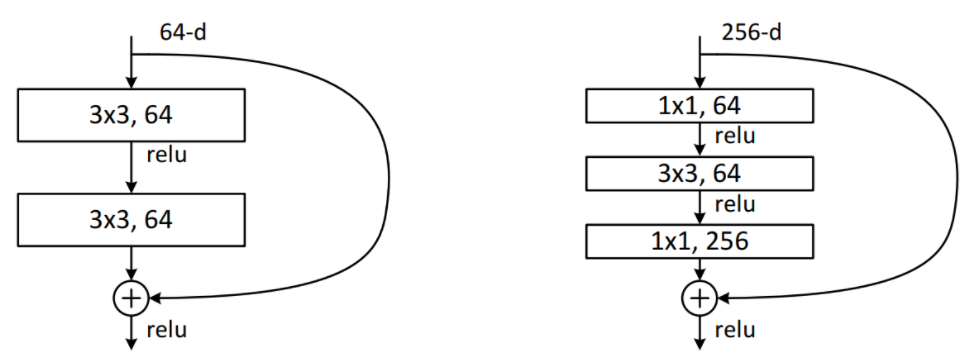
ResNet18、ResNet34使用一般的residual block，而ResNet50、ResNet101、ResNet152使用了expansion為4的bottleneck block。

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面的表格。



圖三：不同ResNet層數結構圖表

從上圖表中可以看到，對於ResNet18和ResNet34，第一個塊包含兩個具有64個filters的3x3 convolutional layers，並且ResNet18在第一層中包含兩個這樣的塊，而Resnet34具有三個。ResNet50，ResNet101和ResNet152塊的結構與ResNet18和ResNet34中的塊不同，這些塊稱為Bottleneck blocks。Bottleneck blocks可減少輸入中的通道數，然後再將其擴展回去。下圖四顯示了由ResNet18和ResNet34使用的標準Basic Block（左），以及由ResNet50，ResNet101和ResNet152使用的Bottleneck blocks。



圖四：Basic Block（左）與Bottleneck blocks（右）結構展示圖

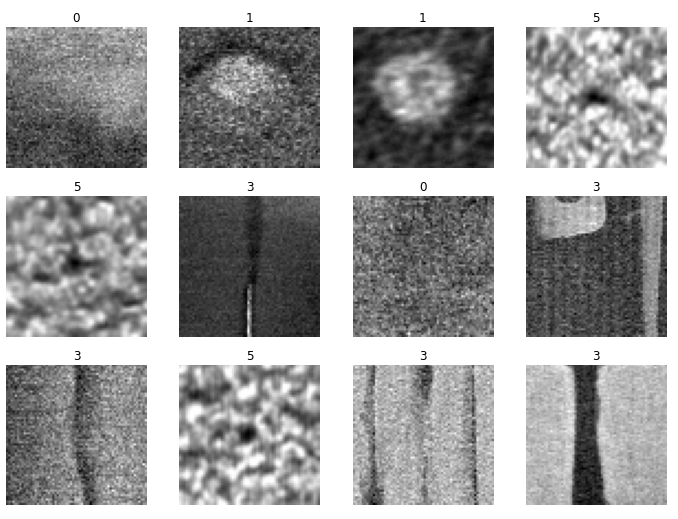
1. **實驗設置與實驗結果**

以ResNet進行嘗試，共嘗試ResNet18、34、50，三種ResNet架構，並嘗試不同參數，加入多種實驗設置，包含batch size調整、normalize與batch normalize 嘗試，更在ResNet50中使用變動的learning rate與Pre-Trained在ImageNet的參數作為初始化權重，以達到準確率的提升，此處僅展示最終效果模型結果。

最終結果模型為ResNet50，並使用變動的learning rate與Pre-Trained在ImageNet的參數作為初始化權重。

**資料集**

AOI影像資料，包含 6 個類別（正常類別 + 5 種瑕疵類別），以下為隨機抓出的12張範例圖。

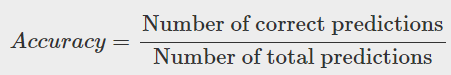
****

**模型相關參數**

* 損失：cross entropy
* 優化器：Adam
* 學習率：Resnet18、Resnet34為0.001，Resnet50的學習率會基於調度程序在模型訓練時動態更改學習率。
* 標準化方式：N，train、valid、test各自做一次標準化(減平均除標準差)；BN，針對每個Batch裡的照片做標準化
* Epoch：在Resnet18、Resnet34中皆跑10000個Epoch並存下所有Epoch中準確率最高的當下模型參數去做預測；在Resnet50中因為是取預訓練於ImageNet的參數作為初始權重Epoch僅需10個模型即收斂。
* 訓練使用機器：RTX 3090 24G x 4

**評估標準**

官方評估方式採用計算與實際值的相符正確率（Accuracy）。公式如下：



**實驗結果**

下表為所有實驗結果表，準確率部分展示的是valid下所得準確率結果，且是記錄所有Epoch中最高的數值。

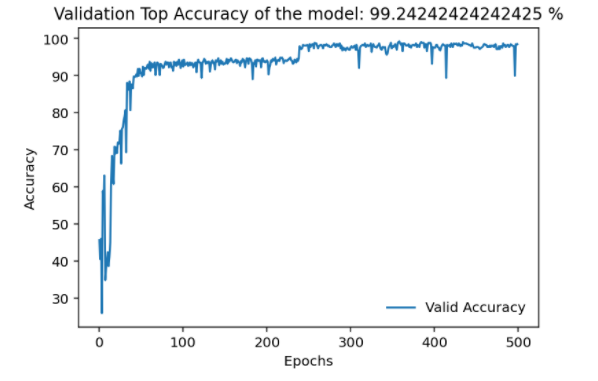
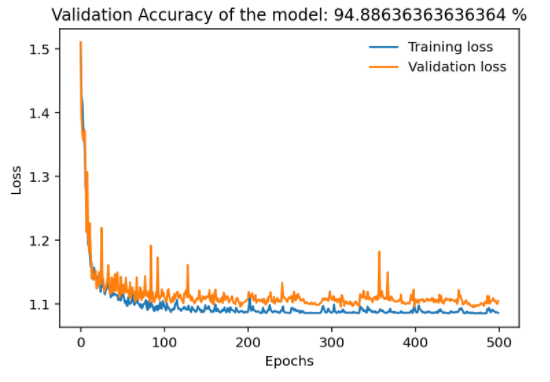
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型 | 批次大小 | 標準化 | 學習率 | 準確率(%) |
| Resnet18 | 12 | N | 0.001 | 94.8863 |
| Resnet34 | 12 | N | 0.001 | 99.2424 |
| Resnet18 | 12 | BN | 0.001 | 98.8636 |
| Resnet34 | 12 | BN | 0.001 | 98.6742 |
| Resnet34 | 256 | BN | 0.001 | 94.8863 |
| Resnet34 | 512 | N | 0.001 | 98.6742 |
| Resnet50 | 64 | N | 變動 | 99.31 |

#準確率為valid下所得結果，記錄所有Epoch中最高的做紀錄。

下列為部分模型訓練過程視覺化展示圖，左圖為損失的改變情形，右圖則準確率的改變情形。

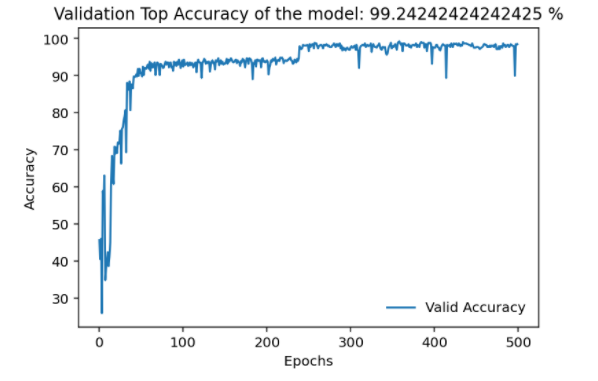
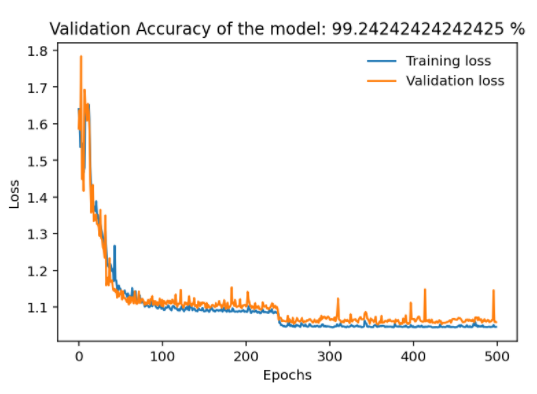
Resnet18

N， batch\_size = 12



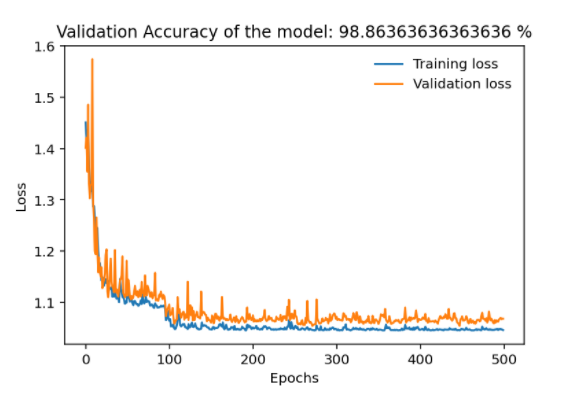
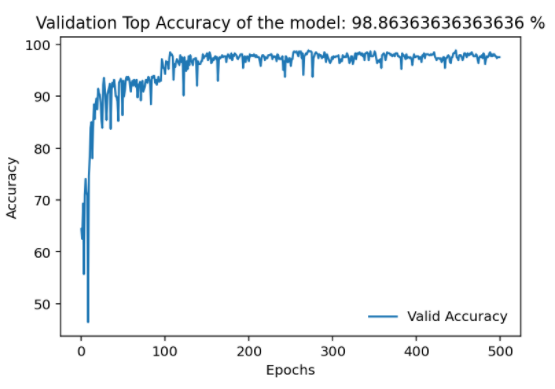
Resnet34

N， batch\_size = 12



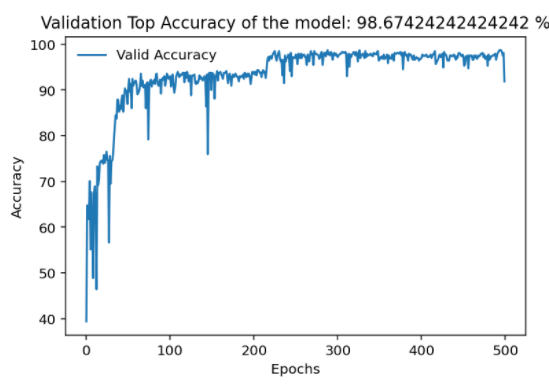
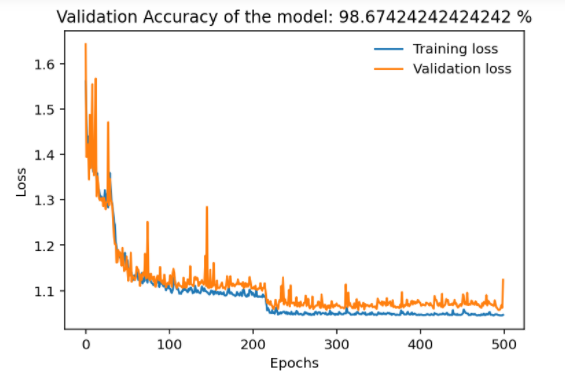
Resnet18

BN ，batch\_size = 12

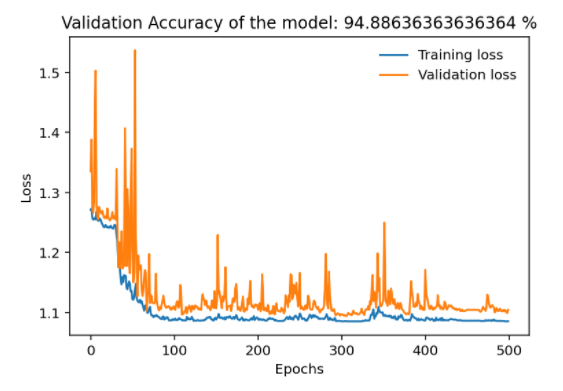
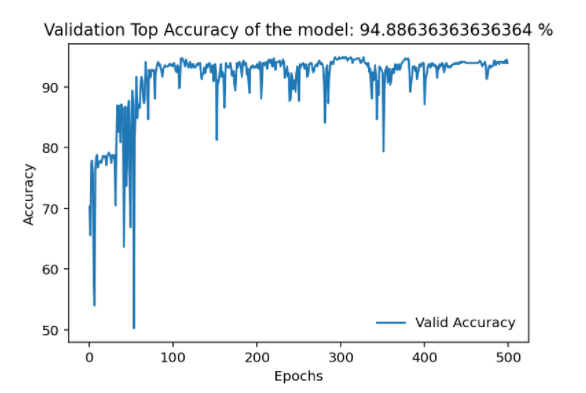
Resnet34

BN ，batch\_size = 12



Resnet34

BN ，batch\_size = 256

ResNet34

N，batch\_size = 512

