

基於深度學習技術提升 PCB 線路瑕疵檢測在 AOI 應用下的錯誤預警能力

陳旻奇、蔡沐霖、吳偉民、冷堂愷、劉晏誠、林鼎鈞

國立高雄大學電機工程學系

{a1105155|b1105129|b1115157|b1115112|a1095161|a1115162}@mail.nuk.edu.tw

摘要

本研究旨在基於深度學習技術，改善印刷電路板 (PCB) 中線路的瑕疵檢測問題，以及提升在自動光學檢測 (AOI) 應用中的錯誤預警能力。本次研究中使用黑白 PCB 資料集，包含六種常見瑕疵：copper、good、mousebite、open、pin-hole、short 及 spur。另外為了解決實際應用中可能發生的瑕疵資料集數量不足問題，我們團隊提出透過生成對抗網絡 (GAN) 擴充資料集，並設置多種評估指標以確保生成影像的可靠性與實用性。

模型訓練部分，本研究比較多個常見的神經網絡模型，包括 AlexNet、ConvNet、EfficientNet、GoogLeNet、MNASNet、MobileNet、ResNet、SqueezeNet、VGG、ViT 及 Xception，並調整學習率 (learning rate)、訓練輪數 (epoch)、批量大小 (batch size) 等參數，尋找最佳配置。對實驗中各模型的訓練結果進行觀察，結合過程中記錄的各種指標，進行綜觀的比較與分析，以進一步評估其在 AOI 應用中的實際效果。

關鍵字：深度學習、印刷電路板、自動光學檢測、瑕疵檢測、生成對抗網絡。

緒論

印刷電路板 (PCB) 是電子產業的關鍵元件，其品質直接影響產品性能。根據台灣電路板協會 (TPCA) 最新統計，2024 年上半年台商印刷電路板 (PCB) 全球產值達到新台幣 3722 億元，較 2023 年同期成長 6%。預計 2024 年全年產值將達 8337 億元，年增幅約 8.3%。

PCB 製造過程會遇到的問題十分繁雜，且製程中每個步驟都可能產生瑕疵，若未及時檢測，將導致產品效能下降甚至重大損失。因此，高效且精準的瑕疵檢測是重要的研究方向之一。隨著深度學習技術的快速發展，卷積神經網絡 (CNN) 已成為影像分類的主要技術。從 ResNet 到 MobileNet，再到新架構的 ViT (Vision Transformer) 提出，多種模型在結構設計與計算效率上都各具特色，且面對不同應用場景與數據特性時，表現可能存在不同的擅長區間，因此有必要對多種模型進行系統比較，以辨識其適用性與優勢。

本研究結合深度學習技術，針對 PCB 外觀瑕疵檢測的問題，訓練不同模型進行比較，再加上生成對抗網絡 (GAN) 技術的引入，可以通過數據庫增列解決資料分布不均的問題，以求進一步提高模型效能。

研究背景

卷積神經網絡 (Convolutional Neural Network) 在電腦視覺中佔有非常重要的地位，在影像辨識上的精準度可以達到超過人類的水準，是目前深度學習發展的一大主力。AlexNet [1] 為 2012 年被提出的卷積神經網絡，是同年的 ImageNet LSVRC 競賽的冠軍，造成很大的轟動，至此 CNN 廣泛地被研究。AlexNet 問世之前，深度學習冷門了

一段時間。因為 LeNet 在早期的小數據集可以取得好的表現，但在大的數據集上表現很差，因此許多人改鑽研其他的機器學習方法。而 AlexNet 的出現，可以說是時代的分水嶺，正式開啟了 CNN 的時代。

AlexNet 的特點

1. 模型架構：採用 Kernel Size 3x3、stride 為 2 的 Maxpooling。比 LeNet 的 meanpooling，更能保留重要的特徵。且 stride < size (2 < 3)，因此 pooling 可以重疊，能重複檢視特徵，避免重要的特徵被捨棄。
2. 輸入層尺寸較大，可以輸入 224x224 尺寸的彩色圖片。
3. kernel size：由於輸入層較大，第一層的 kernel Size 為 11x11、stride 為 4，可學習較大的局部特徵。
4. 激活函數：將 LeNet 使用的 Sigmoid 改為 ReLU，可以避免因為神經網路層數過深或是梯度過小，而導致梯度消失並且相較於早期使用的 Sigmoid/Tanh，ReLU 收斂速度較快。
5. 降低 overfitting 的方法：採用了 Dropout 以及資料擴增方法。

VGGNet

VGGNet 比 AlexNet 採用更深的網路 [2]，特點是重複使用一組 module，並改用小 kernel 替代 AlexNet 裡的大 kernel。

VGG 的特點

1. VGG Block: 由不同數量的 3x3 Convolutional kernel 以及 2x2 的 Maxpooling 組成。
2. 小 Convolutional kernel 代替大 Convolutional kernel: VGG 使用多個 3x3 kernel 代替 AlexNet 的 5*5 卷積核，可達到相同的 receptive field，同時減少參數量。

GoogLeNet 的特點 [3]

- 1.1x1 的 Convolution: 使用 1x1 卷積作為 dimension reduction 來減少計算量。藉由減少計算 bottleneck，可以增加網路深度和寬度。
2. Auxiliary network: 實驗發現在淺層的神經網路，分類更加準確，所以網路中間所產生的特徵圖有其分辨性，因此使用這些特徵圖做分類，增加額外的分支。好處為加強 gradient 的回傳以及額外的 regularization。

ResNet 特點 [4]

為了減輕越深的神經網路訓練難度，作者想到恆等映射 (Identity Mapping) 來解決。輸入等於輸出，所以不管網路再怎麼深，都不應該比什麼權重都沒乘來的差。然而實驗發現，嘗試用恆等映射的方式來訓練深層神經網路，幾乎無法在合理時間內找到足夠好的準確率。接著作者提出一種「殘差映射」(Residual Mapping) 訓練深層神經網路，作者優化「殘差」這件事情，會比直接學習恆等映射要來的容易。在大於 50 層的網路，作者提出 Bottleneck 的設計，來節省大量的參數。透過 1x1 的 conv 降維後再升維的技巧，架構可以兼顧更高維度與效能。SqueezeNet 架構設計策略 [5]

1. 將 3×3 filters 取代為 1×1 filters
2. 減少 3×3 filter 的 input channel 數
3. 在網絡的後期進行 Downsample, 使卷積層有較大的 activation map。減少 CNN 中的參數量, 同時保持 accuracy。與 AlexNet 相比, 使用 SqueezeNet 尺寸減少了 50 倍, 同時達到 AlexNet 的 top-1 和 top-5 accuracy。Xception [6]

Depthwise Separable Convolutions: 傳統的 CNN 由於 Filter 都是使用所有通道 (深度), 運算量會較多, 透過先進行 Depth-wise convolutions (通道分離) 再進行 Point-wise Convolutions (維度調整), 可以達成降低運算量的目的, 且與傳統的 CNN 相比, 運算量大幅減少, 並且 Performance 下降的幅度相當小, 使其在 Edge device 有相當大的優勢。作者基於通道及空間可分離的特性提出假設, 透過觀察後發現 Inception 介於兩種極端的中間 (CNN & Depthwise Separable), 從而將 Inception Module 改為 Depthwise Separable Convolution 並在 ImageNet 上有了小幅的提升, 也證明其效用較之 InceptionV3 在相同的參數量下, 能更有效的利用參數, 同時架構的架設也更為簡易。

MobileNet [7]

MobileNet v1 使用 Depthwise Separable Convolution 建構神經網路以及引入兩個超參數 Width Multiplier、Resolution Multiplier 讓開發人員可以依據自己的應用與資源選擇合適的模型。Depthwise Separable Convolution 包含了深度卷積 (depthwise convolution) 和逐點卷積 (pointwise convolution), 與一般的卷積操作不同, 核心理念就是把卷積過程的計算量減少。一般卷積的 filter 是用在所有的輸入通道上 (input channels), 而 depthwise convolution 則對每個輸入通道採用不同的 filter, 也就是說一個 filter 對應一個輸入通道。每個輸入通道做完 depthwise convolution 後, 再採用 pointwise convolution 將輸出融合, 其實就是使用 1×1 的卷積核做計算。如此的操作, 輸出和一般的卷積層差不多, 但是大大的降低了計算量及模型參數量。

EfficientNet [8]

複合縮放 (Compound Scaling): 網路深度、寬度、解析度是相互影響的, 因此作者認為應該要同時調整這三種縮放方法。在不同網路深度、解析度下寬度縮放的準確度比較, 作者得到了一個結論: 為了追求更高的準確率, 平衡網路深度、寬度、解析度是很重要的關鍵。作者使用一組固定的縮放係數 (就是複合係數 ϕ) 將網路模型均勻地縮放, 此超參數須自己定義, 是用於控制能使用多少資源在模型縮放上。該研究發現平衡這三種縮放能夠獲得更好的性能, 因為他們之間會相互影響: 對於解析度高的圖像, 使用更深的網路、獲取更大的感受野, 能夠有更好的效果, 同時增加網路寬度可以得到更細度的特徵。

作者受 MnasNet 的啟發, 利用多目標神經結構搜索產生高效能的模型 EfficientNet-B0, 使用的搜索空間與 MnasNet 相同, 但優化目標稍有不同: MnasNet 是優化延遲, 而 EfficientNet 則是優化 FLOPS。

Transformer [9]

Transformer 如今已成為熱門的神經網路架構, 並且大量的應用在自然語言 (NLP) 任務。它的成功追朔於 2017 年 Google 所提出的 Attention Is All You Need。這樣的重大突破使得 Google 團隊將這一套 Transformer 架構

中的 Encoder 抽離出來變成 Vision Transformer (ViT) 應用在影像分類技術上。此外它拋棄了 CNN 層並以 self-attention 計算做取代, 並在分類問題上取得不錯的成績。

GAN [10]

GAN 是由兩個神經網路所構成的, 一個叫做 Generator (生成器) 另一個叫做 Discriminator (判別器)。Generator 做的事情就是想辦法把一堆雜訊變成看起來很逼真的照片, 而 Discriminator 做的事情就是想辦法去判別說哪個是真的照片哪個是從 Generator 生成的假照片。所以在訓練的過程中, Generator 生成的照片會越來越逼真, 想辦法騙過 Discriminator, 而 Discriminator 也會想盡辦法去學習辨識真的和假的照片。而這樣兩個神經網路對抗的過程, 我們叫做 Adversarial Learning。所以我們的最終目標就是讓 Discriminator 完全無法區別哪個是真的照片哪個是假的照片。

GAN 的優點也是他的缺點, 它繞過了學 Representation 這個概念, GAN 的學習目標就是直接想辦法生出一堆看起來很像真實照片的假照片, 所以 GAN 的學習過程屬於隱式的, 也就是說他沒有學到真正的 Representation, 所以很難和文字、語音做整合, 但是也因為這樣 GAN 的訓練、推論速度比其他生成模型, 如 Diffusion model 來的快。另外很多學者並不喜歡 GAN 的架構, 作者用了很多技巧讓訓練變得更穩定, 但是沒有特別解釋為什麼。

資料準備

本研究使用 3000 張黑白一整塊的 PCB 影像作為資料集 [11], 每張影像均包含瑕疵 (defect) 與無瑕疵 (good) 兩種版本。根據資料集的特性, 對同一塊 PCB 影像中帶有瑕疵與無瑕疵的相同區域進行裁切, 提取出多個小區塊, 標註為瑕疵類別或無瑕疵類別。一張 PCB 影像可以裁切出多個小區塊, 經過資料處理後, 最終生成約 20000 張裁切的 PCB 區塊影像, 作為訓練分類模型的資料來源。

針對數量較少的瑕疵類別, 透過資料增強技術進行平衡處理, 包括旋轉、翻轉、添加高斯雜訊以及模擬光源變化等操作, 以使各類別的樣本數量達到均衡。

最終, 模型的訓練任務為區塊分類, 分類模型需辨識出 6 種類別的瑕疵 (copper、mousebite、open、pin-hole、short、spur) 以及 1 類無瑕疵的區塊 (good), 以實現高準確率的瑕疵檢測目標。黑白 PCB 資料集。

研究方法

實驗結果

ViT

在 ViT 的模型訓練當中, 觀察 accuracy、loss 曲線:

1. 隨著 learning rate 增加, 模型收斂的速度也隨之增快, 若 learning rate 過低 (如 $lr = 1 \times 10^{-3}$), 會有可能遇到模型收斂困難的問題。
2. 模型訓練的準確度在 epoch10~20 的區間時就已經接近 100%, 訓練集準確度接近 100%, 但驗證集的準確率在收斂至約 95% 時就不再收斂。
3. 更改 batch size 參數, batch size 提升時整體 loss 皆有下降的趨勢, 準確率也有相對提升, 不過隨著 learning rate 提升, 改變 batch size 所造成的影響也隨之變小。

在 ViT 的模型訓練當中，觀察 confusion matrix 與其指標：

1. 對於同一個 class，有時候 Precision 會低於 Recall，有時 Precision 高於 Recall，表示模型漏檢或是誤報的現象並不固定。
2. 綜觀全部的 class，表現最差 class 大多都落在 open、short、spur，而 open 的表現最差，約落在 90%。

加入 GAN 之後的變化：

1. 對於整個訓練來說，隨著 epoch 增加的模型收斂速度有略為提升。
2. 在較低 learning rate 當中的表現有得到不錯的提升，到了 $lr=1 \times 10^{-5}$ 原本就很高的準確率也上升了 1~2%。

ResNet

在 ResNet 的模型訓練當中，觀察 accuracy、loss 曲線：

1. ResNet 在訓練中不論是有無使用 GAN 擴充的資料集，几乎都是呈現無法收斂的情況⁴，除了設定 batch size 為 32，learning rate 為 1×10^{-4} 或 2×10^{-5} 的情況有收斂完成⁵。
2. 在準確度上也對應了 loss 曲線的情況，因未成功收斂而導致準確度較差，在特定數據下才有準確率達到 0.97 的情況⁶。
3. 針對有使用 GAN 的資料集，在 loss 曲線有收斂完成的情況下，準確率較無使用 GAN 的資料集略低 0.01 左右。

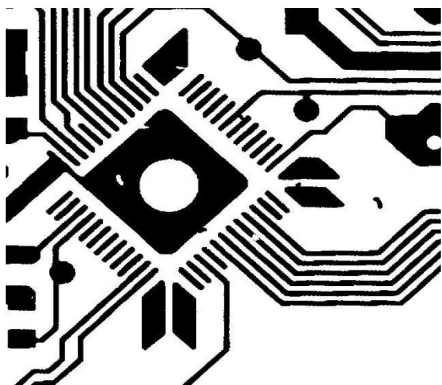


圖 1: 帶有瑕疵的黑白 PCB 影像

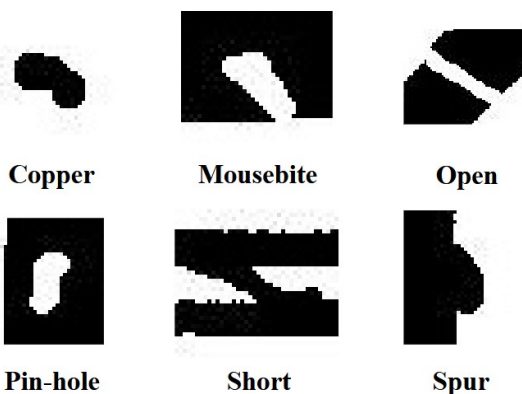


圖 2: 切割出來的 6 種 PCB 瑕疵區塊

在 ResNet 的模型訓練當中，觀察 confusion matrix 與其指標：

1. 對於同一個 class，在 loss 曲線有收斂完成的情況下，precision 與 recall 都達到 0.96 左右，而未完成收斂的情況則是 precision 明顯高於 recall。
2. 以 confusion matrix 綜觀全部的 classes，最差表現通常落在 open、good，而以 open 的情況又較差⁷。

SqueezeNet

在 SqueezeNet 的模型訓練當中，觀察 accuracy、loss 曲線：

1. SqueezeNet 在不論是有無使用 GAN 進行擴充的資料集下，loss 曲線幾乎都有收斂⁸，除了在 batch size 為 8 與 learning rate 為 1×10^{-3} ，以本次實驗來說 batch size 極小 learning rate 極大的情況下，loss 曲線完全沒有收斂⁹，甚至是模型完全沒有學習到特徵。
2. 而在 learning rate 為 1×10^{-3} 的情況下，多數的 loss 曲線會在 epoch 為 10 左右出現較大的起伏¹⁰，進而影響到後面的模型收斂與最後的準確率。
3. 準確度上因為多數情況下 loss 曲線都有收斂，所以準確度普遍落在 0.96 左右¹¹，在 learning rate 為 1×10^{-3} ，並且最後有收斂的情況下，準確率在 0.69 左右¹²。
4. 針對有使用 GAN 的資料集，在 loss 曲線有收斂完成的情況下，準確率與無使用 GAN 的資料集無明顯差異。

在 SqueezeNet 的模型訓練當中，觀察 confusion matrix 與其指標：

1. 對於同一個 class，在 loss 曲線有收斂完成的情況下，precision 與 recall 都達到 0.96 左右，而未完成收斂的情況則是 precision 與 recall 數據持平，表示模型並沒有完整學習到特徵。
2. 以 confusion matrix 綜觀全部的 classes，最差表現通常落在 open、good，又以 open 的情況又較差¹³。

結論

透過實驗數據的綜合分析，得到以下結論：資料擴增的影響：引入 GAN 生成的瑕疵資料後，資料多樣性有明顯提升，尤其在初期訓練階段，有助於模型 Loss 的穩定收斂。我們也觀察到 GAN 擴增資料對於不同模型成效也不一，其中 ViT、MobileNetV1、GoogLeNet 等模型訓練成效都不錯，mAP 分別提升約 1% 至 2%。模型參數對表現的影響：

較小的 Batch size 有助於穩定訓練，但可能限制模型對細微瑕疵的 recall 表現。大批次大小則提升 recall 能力，但可能導致收斂速度變慢。Learning rate 對模型穩定性影響十分顯著，尤其是 VGG 模型，在使用極低 Learning rate（如 2×10^{-5} ）時，表現精度達到最佳，而過高的 Learning rate 則可能導致震盪以致無法收斂。

特殊情況的觀察：MNASNet 表現出較明顯的不穩定性，其 valid loss 波動顯著，且即使引入 GAN 資料，精度提升幅度有限。

整體模型表現：本實驗大中多數的模型表現較為穩定，且能適應 GAN 擴增資料的模型。而 VGG 模型則可能比較適用於極高精度的任務場景，MNASNet 則需進一步調整以提升穩定性。

未來，本研究可在以下方向展開更深入的探索：

1. 提升 GAN 資料的生成方法，提升其與實際資料特性的匹配程度，進一步增強模型對瑕疵的辨識能力。
2. 受限於訓練時長，本次模型訓練的參數選擇並不能做到全面涵蓋，可能需要探索更高效的訓練策略，以加速 Loss 收斂與提升模型精度的目的。
3. 再引入更多新的深度學習模型或技術，驗證這些模型在 PCD 瑕疵檢測領域的適用性。

參考文獻

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf
- [2] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Computational and Biological Learning Society*. 3rd International Conference on Learning Representations, 2015, pp. 1–14. [Online]. Available: <https://arxiv.org/pdf/1409.1556>
- [3] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [5] F. N. Iandola, M. W. Moskewicz, K. Ashraf, S. Han, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <1mb model size," *ArXiv*, vol. abs/1602.07360, 2016. [Online]. Available: <https://api.semanticscholar.org/CorpusID:14136028>
- [6] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1800–1807. [Online]. Available: <https://api.semanticscholar.org/CorpusID:2375110>
- [7] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *ArXiv*, vol. abs/1704.04861, 2017. [Online]. Available: <https://api.semanticscholar.org/CorpusID:12670695>
- [8] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *ArXiv*, vol. abs/1905.11946, 2019. [Online]. Available: <https://api.semanticscholar.org/CorpusID:167217261>
- [9] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," *ArXiv*, vol. abs/2010.11929, 2020. [Online]. Available: <https://api.semanticscholar.org/CorpusID:225039882>
- [10] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *arXiv preprint arXiv:1406.2661*, 2014. [Online]. Available: <https://doi.org/10.48550/arXiv.1406.2661>
- [11] Papers with Code, "Deeppcb dataset," 2024, accessed: 2025-01-24. [Online]. Available: <https://paperswithcode.com/dataset/deep-pcb>

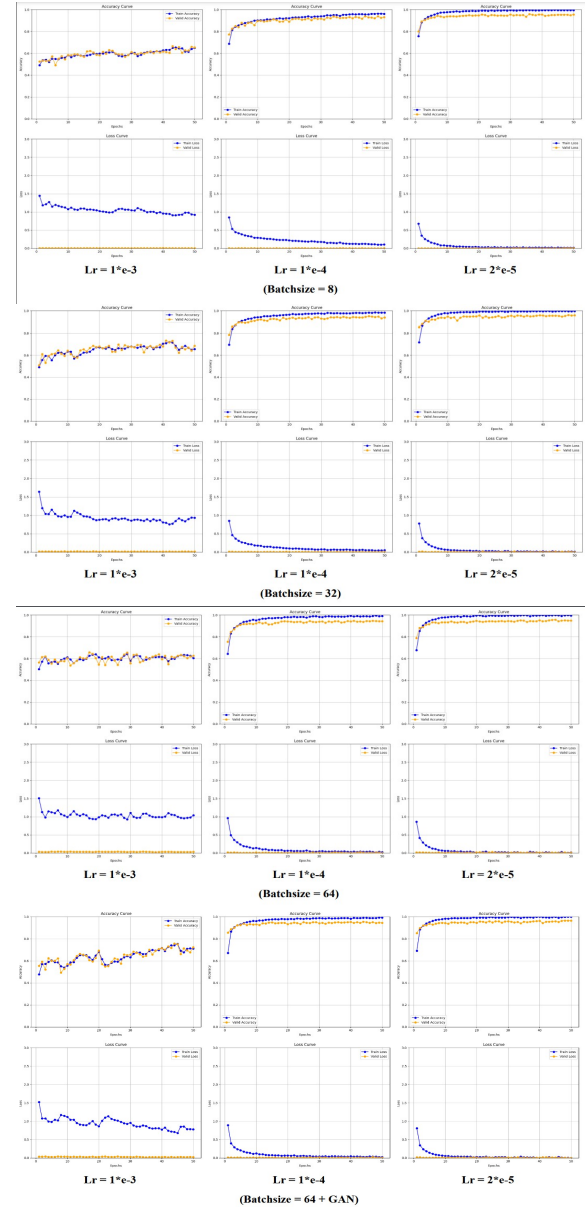


圖 3: ViT 訓練結果

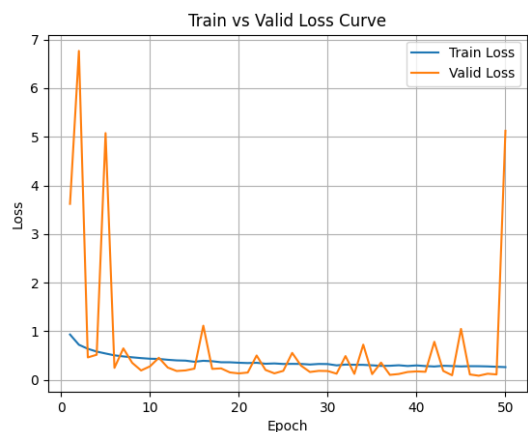


圖 4: ResNet 未收斂 loss 曲線

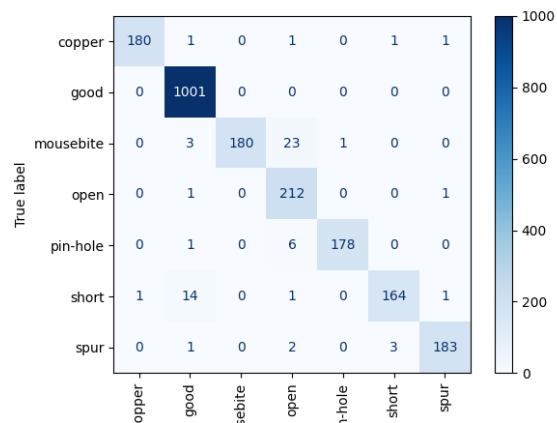


圖 7: ResNet Confusion Matrix

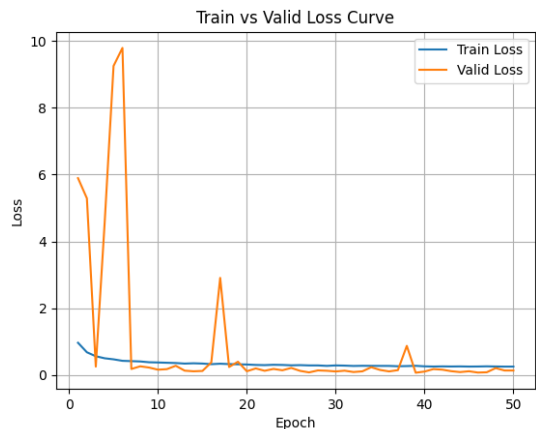


圖 5: ResNet 已收斂 loss 曲線

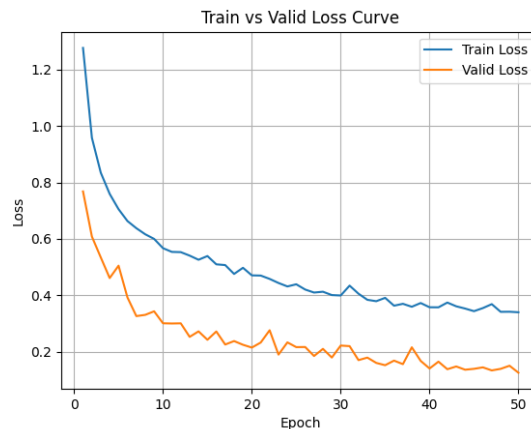


圖 8: SqueezeNet 已收斂 loss 曲線

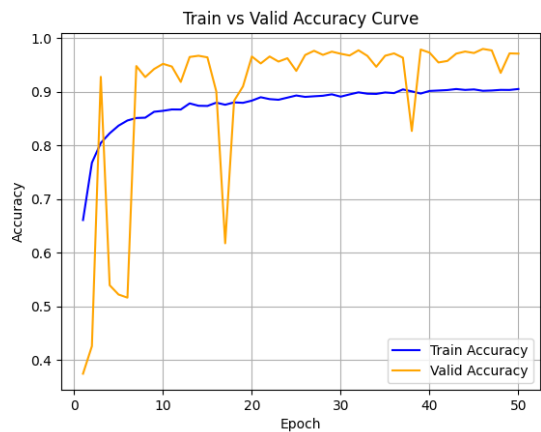


圖 6: ResNet Accuracy 曲線

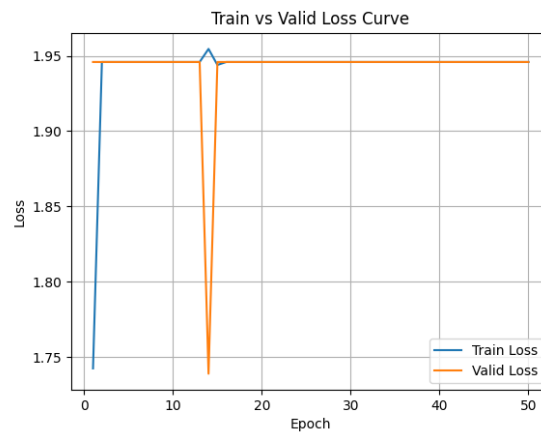


圖 9: SqueezeNet 之 loss 曲線表示模型無學習到特徵

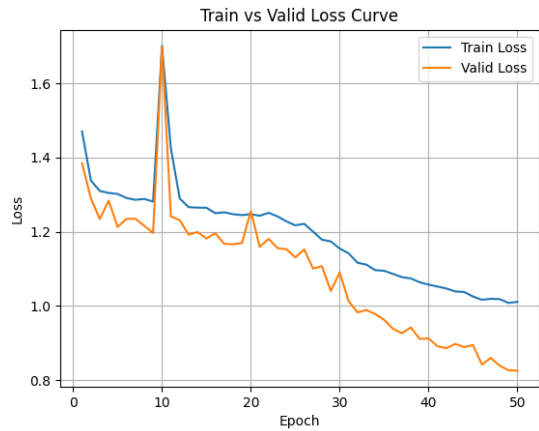


圖 10: SqueezeNet 之 loss 曲線收斂但較不穩

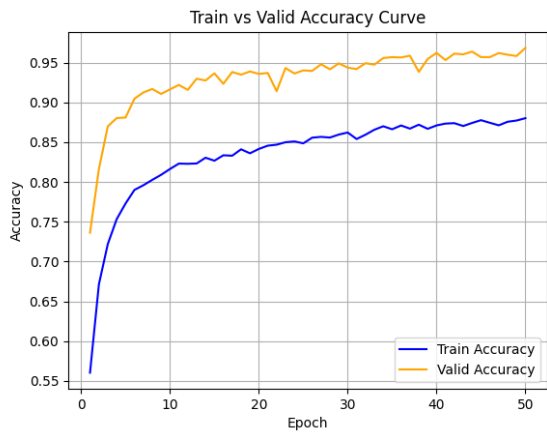


圖 11: ResNet Accuracy 曲線

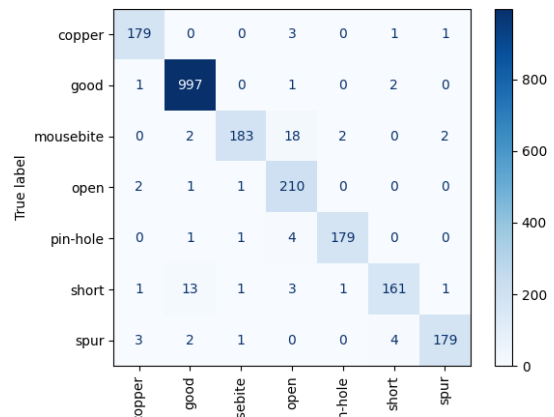


圖 13: ResNet Confusion Matrix

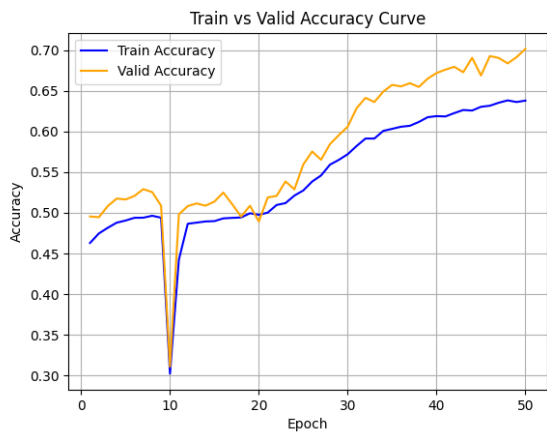


圖 12: ResNet Accuracy 曲線