

1.

Number of defect classes: 8

Types of defect classes:

Bent wire

Cable swap

Combined

Cut inner insulation

Cut outer insulation

Missing cable

Missing wire

Poke insulation

圖片總數: 466

各資料夾圖片數量：
- train: 224 張圖片
- test: 150 張圖片
- ground_truth: 92 張圖片

總共讀入 466 張圖片，shape = (466, 1024, 1024, 3)

Distribution of training and test data:

各資料夾圖片數量：
- train: 224 張圖片
- test: 150 張圖片
- ground truth: 92 張圖片

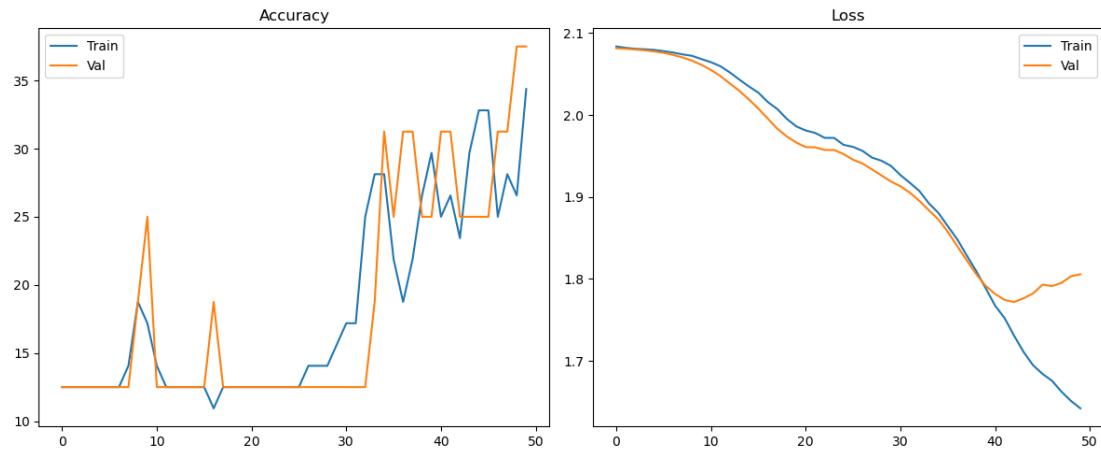
Image dimensions: shape = (466, 1024, 1024, 3)

2.

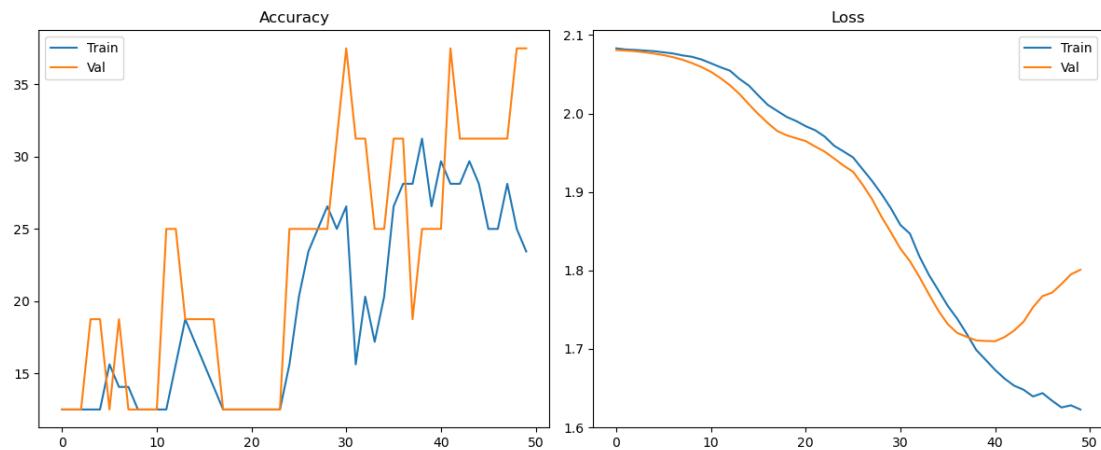
為提升模型表現，我嘗試了四種方法。第一是 Baseline CNN，使用三層卷積加上 ReLU 與 MaxPool，optimizer 為 Adam，learning rate 設為 $1e-3$ ，訓練 50 epochs，準確率表現約落在 30% 左右。第二是加入資料增強策略 (Data Augmentation)，包含隨機翻轉、旋轉與亮度抖動，有助於提升模型的泛化能力，但因資料不均仍有些波動，最終準確率略微提升。第三則採用 TorchVision 的 Pre-trained ResNet18，凍結卷積層僅微調分類層，搭配

AdamW optimizer 與較小 learning rate ($3e-4$)，性能穩定，效果比前兩者略好。第四種方法是在 ResNet18 上加入 CosineAnnealingLR scheduler，使學習率逐步下降以穩定收斂，並成功達到最佳驗證準確率。綜合比較下，ResNet18 + Scheduler 的表現最穩定，成功關鍵在於利用預訓練特徵與逐步調降學習率的訓練策略。

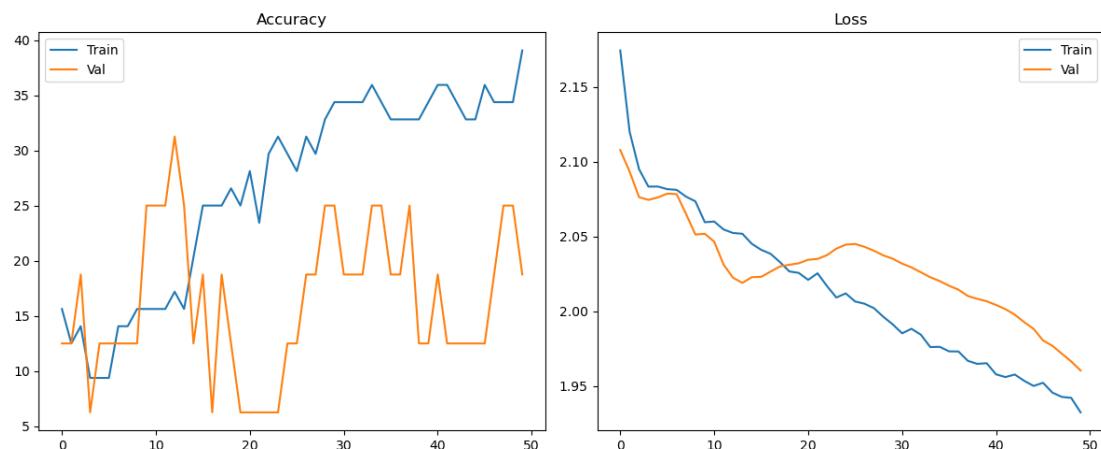
Baseline CNN:



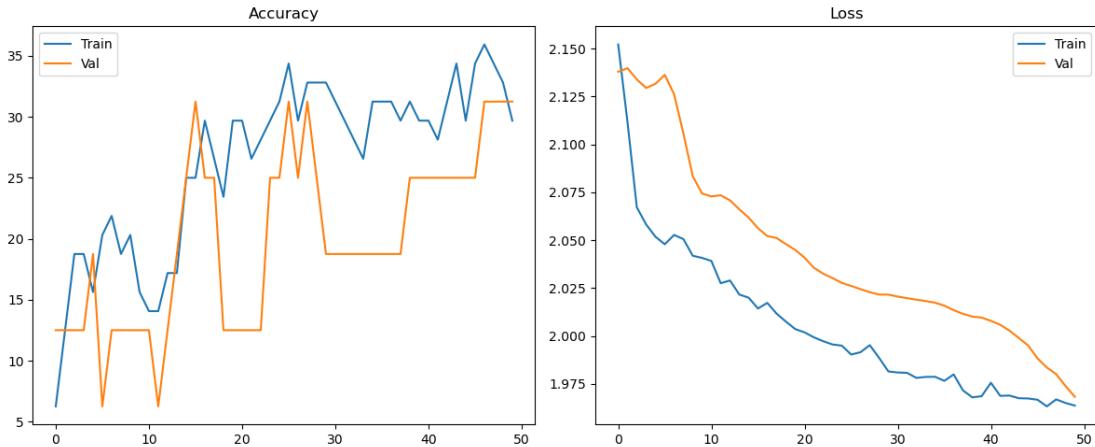
Data Augmentation:



Pretrained ResNet18:



ResNet18 + Learning Rate Scheduler:



3.

3-1

所謂的長尾分布，就是在資料集中，只有少數幾個類別（通常是「正常」或「背景」）有大量的圖片，而其他類別（像是各種缺陷）數量都很少。這會導致模型在訓練時偏向那些資料比較多的類別，進而忽略少數類別，造成辨識效果不佳。

3-2

我參考了 2021 年 CVPR 上的一篇論文《Distribution-Balanced Loss for Long-Tailed Visual Recognition》。這篇研究主要是針對分類時資料不平衡的問題，提出了一種新的 loss function，讓模型在訓練時能夠更平均地學到各個類別的特徵。它透過加權的方式，把出現次數少的類別強化，避免模型只專注在主流類別上。

如果應用在 MVTec AD 的情境，比如「good」類別資料很多，缺陷類別資料很少，我們就可以採用這種 loss function，讓模型在學習時對缺陷類別給予更多注意力，提升整體的判別效果。

4.

當訓練資料幾乎只有正常 (good) 樣本而缺乏缺陷影像時，可採用「異常檢測 (anomaly detection)」策略來建立模型。一種常見做法是使用自編碼器 (Autoencoder)，讓模型學習正常樣本的重建特徵，若未來輸入為缺陷樣本，重建誤差通常會較高，可作為判別依據。另一種方法是使用類似 One-Class SVM 或 Deep SVDD 等僅使用正常樣本訓練的模型。此外，也可利用預訓練模型

提取特徵後，以距離或密度為基礎進行異常判斷。這些方法不需要缺陷標籤，適用於資料不平衡的情況。

5.

5-1

若要微調 YOLO-World 進行物件偵測，需準備包含「邊界框（bounding boxes）」與「類別標籤」的資料，格式如 COCO 或 YOLO 格式皆可。對於 SAM 這類分割模型，則需提供每個物件的「像素遮罩（segmentation mask）」以及對應標籤，常見格式為 COCO segmentation 或 PNG 掩膜圖。這些標註可清楚表示缺陷位置與範圍。

5-2

這些模型已在大型通用資料集上預訓練，具備強大的特徵提取能力，因此即使在本地小型資料上也能有效遷移學習。YOLO-World 擅長偵測多類型目標，適合找出局部缺陷；SAM 可提供精細的區域分割，有助於精準定位異常區域。這讓我們在標註有限的情況下，仍能獲得高準確度的異常檢測結果。