

深度學習與產業應用 HW3

Q1.

- 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']
- Total number of images: 374
- Distribution: 224 training / 150 test
- Image dimensions: 1024 x 1024 pixels

Q2.

First try

- Data augmentation、batch size = 16、lr = 0.003

Second try

- Resize (128, 128)、Optimizer = SGD

Third try

- Resnet34、Optimizer = Adam

fourth try

- Regularization、loss function、Data augmentation、Optimizer = AdamW

第一次調整發現準確度雖然提升了一點點，但驗證的表現非常不穩；第二次可以看到訓練集的準確度大幅提升，loss 也下降得很好，但是在驗證集就有明顯過擬合現象；第三次讓訓練集的表現變得更平穩，稍微提升了模型準確度，不過仍有過擬合；第四次是表現得最好的一次，整體模型學習到後面 loss 可以穩定下降，減少一點 overfitting。我發現 batch size、resize、lr、model 的更換、正規化是讓模型越來越好的關鍵。

Q3

(1) Long-tail distribution 是指大多數樣本集中在某幾個常見類別，少數類別的樣本非常少。導致模型過度偏向樣本數多的類別，忽略罕見但重要的少數類別。

(2) 《DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data》

Dablain 等人提出一種用於解決資料不平衡的新方法。該方法結合了 SMOTE（可合成少數類樣本）與 encoder-decoder 架構（將影像壓縮後再還原），以產生高品質的合成影像。這些合成資料可用來擴增原本數量稀少的類別，幫助模型更充分學習少數類別的特徵，進而提升整體分類準確度。可以應用在 MVTec AD 資料集中，用來生成各種缺陷類別的合成影像，幫助模型學會辨識原本數量較少的缺陷類型。

Q4

當訓練資料僅包含”Good”影像時，可透過無監督方法進行異常偵測。例如使用 Autoencoder 或 VAE 模型，只學習正常樣本的特徵，當遇到異常影像時，模型無法準確重建，產生較高的重建誤差，即可視為異常。也可以使用特徵提取與距離分析，比較輸入影像與正常樣本在特徵空間中的差異。此類方法不需異常標註，適合應用這類缺乏異常樣本的情境中。

Q5

(1) 物件偵測-每張影像需要標示出缺陷的位置與類別，通常以邊界框來表示物件的位置，並搭配一個對應的類別標籤。

語意分割-每張影像中的每個像素都需標註其所屬類別，資料通常以遮罩圖的形式呈現，遮罩圖中的每個像素值代表其對應的類別編號。

(2) MVTec AD 缺陷多具空間特徵，適合用物件偵測或分割處理。YOLO 支援少樣本學習，SAM 可零樣本分割，都適合用於資料不平衡的瑕疵檢測任務。