

Homework 3 郭盈均 312704004

1. The details about my dataset: Screw

- Number of classes: 6
- Types of classes:
Good, manipulated front, scratch head, scratch neck, thread side, thread top
- Number of images used in your dataset: 120 (20 pics per class)
- Distribution of training and test data (train data: 96, valid data: 24)
- Image dimensions: (channels, height, width) = (3, 1024, 1024)

2. 下表為原始的部分參數設計，下面會一次調整一個參數，共調整四項參數，觀察準確率的變化：

	before
Learning rate	0.001
n_epochs	50
Batch size	32
resize	32*32
Pre-trained model	Resnet18
weight	'IMAGENET1K_V1'
Test Accuracy	29.17%

(1)	before	after
Pre-trained model	Resnet18	regnet_y_16gf
Test Accuracy	29.17%	41.67%

(2)	before	after
Resize	32*32	512*512
Test Accuracy	29.17%	45.83%

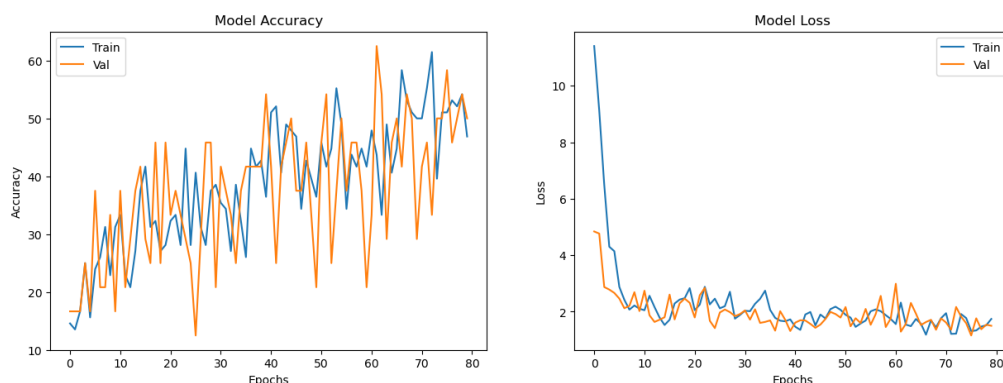
(3)	before	after
Batch size	32	16
Test Accuracy	29.17%	33.3%

(4)	before	after
Learning rate	0.001	0.01
Test Accuracy	29.17%	37.5%

四項參數在調整後都有比原先設計之參數表現較佳，其中有較顯著改善的是

Resize 的大小與 Pretrained model 的選擇。在 Resize 的部分，因為原本大小是 1024*1024，因此 32*32 的設定太小，使得模型學不起來重要特徵，此外 Regnet model 可以使模型性能升高，然而單一參數的調整無法讓學習收斂，接著下方將四個參數同時調整，並且加上 Data Augmentation，觀察性能表現變化。

	After
Learning rate	0.01
n_epochs	80
Batch size	16
resize	512*512
Pre-trained model	regnet_y_16gf
weight	'IMAGENET1K_V1'
Test Accuracy	62.5%



3.

(1) 長尾分佈是指在統計學和機器學習中，某一標籤數量太多，其餘標籤數量太少，造成資料不均衡的問題，在機器學習中被稱為「長尾問題」。

(2) Deep Imbalanced Data Learning Approach for Video Anomaly Detection(2022) 該論文提出了一種深度不平衡資料學習方法，用於視頻異常檢測。該方法融合了兩個互補的深度學習模型，分別是區分性深度殘差網絡 (DRN) 通道和生成性深度回歸長短期記憶 (LSTM) 通道。通過加權回歸模型進行不平衡資料學習，該方法旨在提高對於不平衡資料的學習效果。若應用於本次題目上，對於影像異常偵測應該也有不錯的表現。

4.利用 Auto Encoder 應用於異常檢測：

當使用 Auto Encoder 進行異常檢測時，通常的做法是將正常的訓練資料（即沒有異常的資料）用於訓練 Auto Encoder 模型。Auto Encoder 被訓練來學習將輸入資料壓縮成一個潛在的表示，然後再將這個表示解碼回原始的輸入資料。在訓練完成後，如果 Auto Encoder 成功地學習了正常資料的特徵，則它將能夠很好地

重建這些正常資料。

在進行異常檢測時，我們將測試資料通過 Auto Encoder，並計算重建資料與原始資料之間的差異。如果測試資料與訓練資料類似，則重建的資料將與原始資料相似，差異較小；但如果測試資料包含異常，則由於自動編碼器未見過這種異常，其重建資料與原始資料之間的差異會增大。因此，通過設定一個閾值，我們可以根據差異的大小來檢測是否存在異常。

5.

(1) 對於目標檢測和分割任務，我們需要準備特定類型的資料集：

目標檢測：對於目標檢測，需要標記每個圖像中存在的目標物體的位置和類別信息。因此，我們需要準備帶有 bounding box 標籤的訓練圖像，每個 bounding box 應標示出圖像中的目標物體以及相應的類別。

分割：對於分割任務，我們需要為每個像素標記其所屬的對象或區域。因此，我們需要準備帶有像素級標籤的訓練圖像，其中每個像素被分配到相應的類別（對象或背景）。

(2) 這些模型通常在大型資料集上進行了預訓練，學習到了豐富的特徵表示，可以有效地捕捉圖像中的信息。此外，YOLO-World 和 SAM 等開源模型經過廣泛的驗證和改進，在許多不同的應用中都表現良好，因此適合用於我們的自定義資料集進行微調。因此，透過微調這些模型，我們可以充分利用它們在目標檢測和分割任務中的強大功能，以提高我們的異常檢測效果。

參考資料/

chatGPT

Medium/用 AI 做一些故障預測的小經驗(AutoEncoder 介紹)

Medium/ Anomaly Detection Techniques

IEEE/ Deep Imbalanced Data Learning Approach for Video Anomaly Detection(2022)