

National Tsing Hua University
11320IEEM 513600
Deep Learning and Industrial Applications
Homework 3

113034506 李家欣

1.

- (1) Number of defect classes: 4
- (2) Types of defect classes: ['broken_large', 'broken_small', 'contamination', 'good']
- (3) Number of images used in your dataset: 50
- (4) Distribution of training and test data: Skewed Distribution(good 類別佔 40%)
- (5) Image dimensions: (50, 900, 900, 3)



2.

● **調整方向**

- (1) 不同 backbone: Resnet18(預設)、Resnet50、Resnet101
- (2) 資料增強: 預設, 新增
- (3) 解凍策略: 解凍 fc(預設)、解凍 fc+layer3+layer4
- (4) Batch size: 32(預設)、8

● **最佳模型**

Model: Resnet18 + 資料增強 + 解凍多層 + batch size 8

optimizer: AdamW

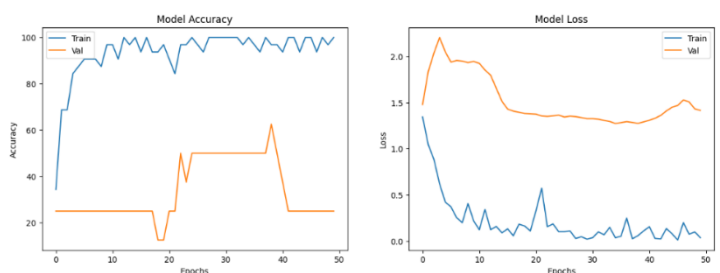
Learning Rate: 0.0001

Weight Decay: 0.01

Test accuracy: 62.5%

● **成功關鍵**

- (1) 相較 ResNet18，ResNet50 有更深的結構、更多參數，能夠學到更細緻的瑕疵特徵；另外 ResNet101 雖然深，但針對小資料集容易 overfitting，故最終測試下來 ResNet50 是一個表現最好的折衷方法
- (2) 資料增強(採用不同角度、光照、大小的瑕疵)可以改善資料集數量少的問題，以此降低 overfitting
- (3) 解凍三層使得模型更能學習到細節的瑕疵特徵，增加泛化能力
- (4) Batch: 初始設定為 32，但考慮到資料集數量少(總共 50 張)，可能導致每



個 epoch 更新次數少或沒有，反而造成學習不佳，故選擇較小的 batch size

● 其他模型效果(編號是對應.ipynb 的檔名)

編號	backbone	資料增強	解凍層數	batch size	準確度
1	resnet18	x	fc	32	25.00%
2		V	fc	32	37.50%
3		x	fc + layer3 + layer4	32	25.00%
4		x	fc	8	50.00%
5		V	fc + layer3 + layer4	32	37.50%
6		x	fc + layer3 + layer4	8	37.50%
7		V	fc	8	25.00%
8		V	fc + layer3 + layer4	8	25.00%
9	resnet50	x	fc	32	25.00%
10		V	fc	32	25.00%
11		x	fc + layer3 + layer4	32	25.00%
12		x	fc	8	50.00%
13		V	fc + layer3 + layer4	32	25.00%
14		x	fc + layer3 + layer4	8	50.00%
15		V	fc	8	50.00%
16		V	fc + layer3 + layer4	8	62.50%
17	resnet101	x	fc	32	25.00%
18		V	fc	32	25.00%
19		x	fc + layer3 + layer4	32	25.00%
20		x	fc	8	50.00%
21		V	fc + layer3 + layer4	32	25.00%
22		x	fc + layer3 + layer4	8	50.00%
23		V	fc	8	50.00%
24		V	fc + layer3 + layer4	8	25.00%

3.1 Long-Tail Distribution

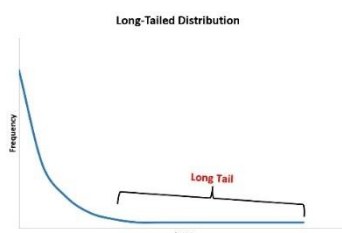
資料集中少數類別樣本數量多，而多數類別數量少的情況

3.2 論文方法

- 使用 VQGAN 來生成影響，改善分布
- 將原始和生成影像做比較，獲得潛在的異常區域(圖像、區塊、像素等向量指標來量化異常)
- 以多層次異常指標訓練分類器，做為評估指標

這篇論文實驗使用變化很細微的零件瑕疵資料集，其中 MVTec AD 資料集成功提高一般準確率與 ZFN 約束下到 95.69% 和 87.93%。

Source: ArnaudBougaham1 · MohammedElAdoui1 · IsabelleLinden2 · BenoîtFréna(2022) Composite



score for anomaly detection in imbalanced real-world industrial dataset

4.

- (1) **Feature Embedding + Nearest Neighbor:** 利用預訓練模型抽取正常特徵，異常樣本因特徵不同無法匹配，e.g. PatchCore、PaDiM
- (2) **Autoencoder 重建誤差檢測:** 模型學習還原正常圖像，異常圖像重建失敗產生高誤差，e.g. Variational AE、Denoising AE
- (3) **Contrastive Learning:** 透過比對相似與不相似樣本特徵，學習具區分性的空間，與正常樣本距離較遠者視為異常。
- (4) **One-Class Learning:** 僅學習正常類型特徵邊界，落在邊界外視為異常，e.g. One-Class SVM
- (5) **GAN:** 生成器學習產生正常樣本，異常樣本因無法準確重建，進而找出異常區域的差異
- (6) **資料增強:** 人工加入雜訊、模糊、mask 等模擬異常以強化模型辨識能力

5.1 dataset

(1)物件檢測:

- 圖片檔
- 標註檔: 包含邊界框座標、標籤

(2)圖像分割

- 圖像
- 遮罩、像素標記檔: 每個像素都有標籤，或是用顏色區分

5.2

(1) **移學習能力強:** YOLO world 及 SAM 模型已在大規模資料集 (e.g. COCO、ImageNet) 上進行過預訓練，擁有豐富的特徵表達學習能力，故僅需少量樣本就能進行微調

支援自定義類別與標註格式，使用者可以透過 labelimg 等軟體進行人工標註，以應對不同應用做針對性訓練。

(2) **支援自定義類別與標註格式:** 使用者可以透過 labelimg 等軟體進行人工標註，以適應更特殊的應用場域

(3) **減少標註成本:** YOLO world 及 SAM 具備“用點選或框選提示就能分割”半自動標註流程，可加速自定義資料集