

**National Tsing Hua University**  
**11220IEEM 513600**  
**Deep Learning and Industrial Applications**  
**Homework 3**

Name: 曾聖閔

Student ID: 112034564

**Due on 2024/04/11.**

**Note: DO NOT exceed 3 pages.**

1. (10 points) Download the MVTec Anomaly Detection Dataset from Kaggle ([here](#)). Select one type of product from the dataset. Document the following details about your dataset:

Wood

- Number of defect classes. 6
- Types of defect classes. Color, combined, good, hole, liquid, scratch
- Number of images used in your dataset. 247+78
- Distribution of training and test data. Long tailed
- Image dimensions. 1024\*1024

2. (30 points) Implement 4 different attempts to improve the model's performance trained on the dataset you choose in previous question. Ensure that at least one approach involves modifying the pre-trained model from TorchVision. Summarize the outcomes of each attempt, highlighting the best performing model and the key factors contributing to its success. You may also need to describe other hyperparameters you use in your experiment, like epochs, learning rate, and optimizer. (Approximately 150 words.)

我調整了優化器為SGD，learn rate=0.01，batch size=16，epochs=10，得到的ACC僅只有16%；調整為ADAM，learn rate=1E-3，batch=64, epochs=100，得到ACC為25%

；調整為ADAMW，learn rate=1e-3，batch size =32, epochs=50，得到acc=41%。

以最後一個的結果改用resnet50，ACC最多只到25%。由於訓練的資料不夠多，導致訓練模型的效果不是很好，或許可以微調參數或所使用的調整器，但礙於colab的流量不夠，無法重複微調。

3. (20 points) In real-world datasets, we often encounter long-tail distribution (or data imbalance). In MVTec AD dataset, you may observe that there are more

images categorized under the 'Good' class compared to images for each defect class. (Approximately 150 words.)

(i) (5 points) Define what is 'long-tail distribution.'

長尾分佈指的是一種統計現象，即隨著事件或觀察的罕見程度增加，其發生頻率迅速減少。在數據集的語境中，這意味著某些類別的實例數量很多（分佈的「頭部」），但其他類別的實例數量很少（分佈的「尾部」）。

(ii) (15 points) Identify and summarize a paper published after 2020 that proposes a solution to data imbalance. Explain how their method could be applied to our case.

AdaBelief Optimizer: Adapting Stepsizes by the Belief in Observed Gradients。

該論文提出了一種自適應優化器，根據梯度觀察的統計信息動態調整學習率，旨在應對不平衡的數據集挑戰。

AdaBelief優化器可以應用於我們的情況，通過在訓練期間動態調整學習率。通過根據觀察到的梯度信念調整學習率，該優化器可以有效地在訓練過程中更多地關注罕見的類別（在這種情況下是缺陷類別），從而減輕數據不平衡的影響。這種自適應方法有助於模型更好地學習少數類別的表示，提高其在真實世界圖像中檢測缺陷的能力。

4. (20 points) The MVTec AD dataset's training set primarily consists of 'good' images, lacking examples of defects. Discuss strategies for developing an anomaly detection model under these conditions. (Approximately 100 words.)

- 無監督學習：使用無監督學習技術，如自編碼器或基於密度的方法，來學習數據的分佈，然後將異常視為與訓練集分佈不同的樣本。
- 弱監督學習：利用少量已標記的缺陷樣本，結合無監督或半監督技術，以便模型可以更好地識別異常。
- 遷移學習：利用在其他數據集上訓練過的模型的知識，通過微調或特徵提取來適應 MVTec AD數據集的特性，從而改善異常檢測性能。

5. For the task of anomaly detection, it may be advantageous to employ more sophisticated computer vision techniques such as object detection or segmentation. This approach will aid in identifying defects within the images more accurately. Furthermore, there are numerous open-source models designed for general applications that can be utilized for this purpose, including YOLO-World ([website](#)) and SAM ([website](#)). (Approximately 150 words.)

(i) (10 points) To leverage these powerful models and fine-tune them using our dataset, it is necessary to prepare specific types of datasets. What kind of data should be prepared for object detection and for segmentation.

(ii) 要利用這些強大的模型並使用我們的數據集進行微調，需要準備特定類型的數據集。

對於物體檢測，應準備帶有物體標註的圖像數據集；對於分割，應準備帶有像素級

別標註的圖像數據集。(10 points) Why are these models suitable for fine-tuning for our custom dataset?

這些模型適合於我們的自定義數據集進行微調，因為它們在一般圖像識別任務上已經表現出色，具有良好的泛化性能。通過微調，可以將這些模型進一步調整為我們的特定任務，從而提高異常檢測的準確性和效能。