深度學習作業三

1. Download the MVTec Anomaly Detection Dataset from Kaggle. Select one type of product from the dataset. Document the following details about your dataset:
   - Number of defect classes: 3
   - Types of defect classes: broken_large、broken_small、contamination
   - Number of images used in your dataset: 292
   - Distribution of training and test data: Train (good=209)、
     Test (broken_large=20、broken_small=22、contamination=21、good=20)
   - Image dimensions: 900 x 900 (width x height)

2. 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.
   本實驗採用 Deep SVDD 進行 bottle 類別的異常檢測，總共實作四種方法。
   訓練皆採 batch size 32、epoch 20、learning rate 1e-4，優化器為 Adam。
   - 方法一使用簡單 CNN，AUC 0.5714。
   - 方法二改用 TorchVision 的預訓練 ResNet18 並修改分類層為 128 維特徵向量，AUC 顯著提升至 0.996。
   - 方法三在 ResNet18 架構下加入資料增強（隨機翻轉、旋轉、色偏與隨機裁切），進一步提升泛化能力，達到最高 Accuracy 95%、AUC 0.9889。
   - 方法四使用 Mahalanobis 距離取代歐式距離平方計算異常分數，雖 AUC 為 0.9881，但分類報告顯示誤判較多。

   因此方法三擁有最佳表現，主因為預訓練模型結合有效的資料擴增技術。

3. 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.
   i. Define what is 'long-tail distribution.'
      資料不平衡 long-tail distribution，指少部分類別(如:good)樣本極多，大部分類別（如:anomaly）樣本極少，導致模型過度偏頗學習，影響分類或偵測效能。

ii. 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.
NeurIPS 2021 論文《Generative Class-Agnostic Learning for Robust Anomaly Detection》提出的核心技術使用 GAN 模擬異常樣本，改善資料不平衡。
此法在 MVTec AD 可透過訓練生成器來製造多樣異常圖像，作為訓練時的輔助資料，減少模型對 'good' 類別的過度擬合，有效提升對少數異常類別的辨識能力與泛化能力。
具體做法，可在訓練初期用 good 資料訓練 GAN。接著，透過對 latent space 的擾動產生異常樣本（如瓶子變形、扭曲紋理）。最後，把這些合成樣本加入訓練集，與正常圖像共同訓練 anomaly detector（例如 Deep SVDD + ResNet）。

4. 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.
使用 Deep SVDD（Deep Support Vector Data Description）可有效處理 MVTec AD 中缺乏異常樣本的問題。由於訓練集中僅包含大量正常（good）樣本，透過訓練神經網路（Simple CNN 或 ResNet18）將正常樣本映射至特徵空間，並使其靠近一個學習到的中心向量 c。推論階段，若輸入圖像的特徵與中心距離過大，即視為異常。此方法無需異常樣本，能自動學習正常樣本的隱含分布，對於僅有正常訓練資料的無監督異常偵測任務非常合適。

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 and SAM.
   i. 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.
   YOLO-World: 需準備 bounding box 標註的資料集，每張圖像需標示異常區域的位置與異常類別。
   SAM: 需準備 pixel-level 標註的資料集，每張圖像對應到一張 segmentation mask，標記出異常區域的每個像素。
   ii. Why are these models suitable for fine-tuning for our custom dataset?
   YOLO-World 支援文字提示與多類別偵測，適合處理多樣化缺陷類型；SAM 則可進行高精度的即時分割，能精準框出瑕疵區域。在 MVTec AD 缺乏大量異常樣本的情況下，這些 foundation models 經過少量 fine-tuning 即可達到良好效果，特別適合應用於資料有限的視覺異常檢測任務。