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:
   - Number of defect classes.
   - Types of defect classes.
   - Number of images used in your dataset.
   - Distribution of training and test data.
   - Image dimensions.

1. 8 種
2. 'split_teeth', 'fabric_border', 'broken_teeth', 'fabric_interior', 'combined', 'rough', 'squeezed_teeth', 'good'
3. 80(train:64,test:16)
4. uniform distribution(每個類別的數量一樣)
5. Height:1024,Width:1024,Channel:3

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.)

原始超參數設定	
epoch	50
pre-trained model	resnet18
lr	0.001
optimizer	Adam

以下嘗試之超參數，都奠基在將 image resize 成 224X224(imagenet 輸入大小)

1. 更改 epoch 數(50->300) `Test accuracy is 68.75%`
2. 更改 epoch 數(50->300)+pretrained model 改為 VGG19 Test accuracy is 75.0%
3. 更改 epoch 數(50->300)+lr 0.001>0.005 `Test accuracy is 68.75%`
4. 更改 epoch 數(50->300)+Optimizer Adam->SGD `Test accuracy is 68.75%`

因為這個資料集的筆數較少，所以可能有點 train 不起來，但主要差異還是在 epoch 數量的提升會比較有助於更好的 performance。

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.'
   (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.

1. 長尾分布是一種頻率分布，前 20%的類別佔了大多數，但尾部比例較低的類別加總後仍不可忽視。
2. 本篇論文使用殘插融合機制，主分支用來識別所有圖片，使用另外兩個分支來識別中尾及尾部較少數類別，最後透過 additive shortcuts 相加。分為分為主支和分支，將"Good"視為 head，其餘視為 tail，分別計算 fusion 的 outputs 和各類別的 outputs 後，再加權兩者，透過特徵參數和殘差設計可以有效增加分類準確率。$\begin{cases} \mathcal{N}_{h+t}: head\ class \\ \mathcal{N}_t: tail\ class \end{cases}$

$$\mathcal{L}_{fusion} = loss(\mathcal{N}_{h+t}(x) + \mathcal{N}_t(x), y)$$

$$\mathcal{L}_{bracnch} = \sum_{i \in \{h+t,t\}} loss(\mathcal{N}_i(Sx_i), Sy_i)$$

$$\mathcal{L}_{all} = (1-\alpha)\mathcal{L}_{fusion} + \alpha\mathcal{L}_{branch}$$

https://arxiv.org/abs/2101.10633

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.)

Variational autoencoders

VAE 模型主要可以分成 encoder 和 decoder 兩個部分，encoder 將原始資料編碼成資料平均值與資料標準差進行參數化，並透過高斯分佈將平均值和變異數組成隱藏編碼，使其服從高斯分布，之後再將隱藏編碼生成回原始資料。在 test 時，當輸入值為其無法準確生成的資料時，即可以判斷是異常值。

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) (10 points) Why are these models suitable for fine-tuning for our custom dataset?

1. Object detection: train label 可能包含多個 target，需要標示這些 target 所屬的位置和相對應的類別。
   Segmentation: train label 需要包含圖片中的每個 pixel 屬於哪個類別

2. YOLO-World: pretrain 使用先透過文本的重要詞彙 embedding 後再針對輸入的圖像進行相對應匹配並輸出預測機率。在 fine-tune 時只要輸入該類別的瑕疵就能圈選出瑕疵所在位置。
   SAM:主要特色是在全景分割，也就是除了圖像中 target 的分割還能顯示其類別，因此適合對 fine-tune 到瑕疵檢測