National Tsing Hua University

11220IEEM 513600

Deep Learning and Industrial Applications

Homework 3

Name: 單允辰　　　　　　　Student ID: 112034518

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

   Dataset name: zipper
   Number of defect classes: 8
   Types of defect classes: broken_teeth, combined, fabric_border, fabric_interior, good, rough, split_teeth, squeezed_teeth
   Number of images used in your dataset: 8*10(num_image of each class)
   Distribution of training and test data: (0.8, 0.2)
   Image dimensions: (3, 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.)

Epochs: 50 lr: 0.001 batch size: 32 optimizer: adam model resnet: resnet18 weights: IMAGENET_1K_V1 train loss: 1.9481 train acc: 20.3125% val loss: 2.1785 val acc: 6.2500% test acc: 25.0%	Epochs: 200 lr: 0.001 batch size: 32 optimizer: adam model resnet: resnet18 weights: IMAGENET_1K_V1 train loss: 1.7273 train acc: 37.5000% val loss: 2.1013 val acc: 12.5000% test acc: 31.25%

Epochs: 50 lr: 0.001 batch size: 64 optimizer: adam model resnet: resnet18 weights: IMAGENET_1K_V1 train loss: 2.1401 train acc: 20.3125% val loss: 2.6799 val acc: 12.5000% test acc: 12.5%	Epochs: 200 lr: 0.001 batch size: 64 optimizer: adam model resnet: resnet18 weights: IMAGENET_1K_V1 train loss: 1.7031 train acc: 37.5000% val loss: 2.0052 val acc: 18.7500% test acc: 43.75%
Epochs: 50 lr: 0.001 batch size: 32 optimizer: adam model resnet: resnet50 weights: IMAGENET_1K_V2 train loss: 1.9245 train acc: 23.4375% val loss: 2.1543 val acc: 18.7500% test acc: 25%	Epochs: 200 lr: 0.001 batch size: 32 optimizer: adam model resnet: resnet50 weights: IMAGENET_1K_V2 train loss: 1.6850 train acc: 37.5000% val loss: 1.8627 val acc: 25.0000% test acc: 50.0%
Epochs: 50 lr: 0.001 batch size: 64 optimizer: adam model resnet: resnet50 weights: IMAGENET_1K_V2 train loss: 1.7851 train acc: 37.5000% val loss: 1.9817 val acc: 25.0000% test acc: 50.0%	Epochs: 200 lr: 0.001 batch size: 64 optimizer: adam model resnet: resnet50 weights: IMAGENET_1K_V2 train loss: 1.7904 train acc: 37.5000% val loss: 1.9347 val acc: 18.7500% test acc: 43.75%

從結果可以發現，其餘參數不變下，調高epochs通常都可以訓練出表現較好的模型。而Batch size的大小對於模型表現的影響不大。另外，在model resnet與weights的部分，選擇resnet50與IMAGENET_1K_V2權重的模型表現較resnet18與IMAGENET_1K_V2權重更好。不過，無論參數如何調整，模型的test acc都很低，推測原因是資料集大小較少或資料不平衡。

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.'
   長尾分布是指 data set 中，少數類別擁有大量樣本(頭部)，其餘多數類別的樣本較少(尾部)，導致類別間的樣本不平衡，呈現長尾的分布。
   (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.

   《Mixup: Beyond Empirical Risk Minimization》，提出了一種名為Mixup的新方法，用於解決資料不平衡的問題。其核心思想是在訓練過程中，將兩個不同類別的樣本混合起來，生成一個新的樣本，其標籤是這兩個樣本的加權平均。通過這種方式，Mixup不僅增加了訓練數據的多樣性，還有助於平衡不同類別的樣本分佈，從而減輕了資料不平衡帶來的問題。

   參考資料: Mixup: Beyond Empirical Risk Minimization

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.)
   - Data augmentation: 人工生成「好」圖像的變化，對圖像進行旋轉、翻轉、縮放、添加噪聲或應用其他變換，以模擬可能的缺陷場景。
   - 合成缺陷圖像：利用合成技術生成缺陷圖像，以擴充訓練集。
   - 遷移學習：利用從其他相關任務或領域的先前訓練模型進行遷移學習。使用來自相似任務的先前知識，加速模型的訓練過程並提高性能。

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.
   在object detection需要準備包含圖像和對應標籤（bounding box）的資料集，其中標籤應該指示每個圖像中缺陷的位置和類別。對於分割，需要準備圖像和像素級別的標籤，其中標籤指示每個像素屬於哪個類別（正常或缺陷）。
   (ii) (10 points) Why are these models suitable for fine-tuning for our custom dataset?
   這些模型由於其預訓練知識、遷移學習能力、豐富的特徵表示以及靈活性，因此適合在我們的定製資料集上進行微調。利用這些優勢進行微調，可提高異常偵測的準確性和性能。