National Tsing Hua University

11220IEEM 513600

Deep Learning and Industrial Applications

Homework 3

Name: 陳奕均 Student ID: 112034604

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:

   Zipper

   - Number of defect classes: 7
   - Types of defect classes: broken teeth, combined, fabric border, fabric interior, rough, split teeth, squeezed teeth
   - Number of images used in your dataset: 391
   - Distribution of training and test data: long-tail distribution
   - 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.)

   在沒更改任何參數的情況下，能夠得到的最高 test accuracy 是44.44%，然而在試著更動batch size之後卻發現，無論設為多少都會讓 performance 下降，同樣地更改 training 和validation的split point也是一樣的結果，也嘗試了將optimizer設為SGD(lr=0.01, momentum=0.9)，卻也無法得到超越原參數的 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.'

   long-tail distribution 指的是只有少數項目發生頻率高，大部分項目則都是低頻率的現象，是在現實生活中常見的分配。

   (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.

Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning (2021) 中提出一個以Python寫成的工具包，專門用於處理機器學習中的不平衡數據集。該工具包提供了一系列基於採樣、敏感學習和集成方法的算法，能夠幫助用戶有效地處理不平衡數據集，並改善模型的性能。

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.)

在training data中缺乏defect sample的情況下，發展異常檢測模型的策略之一是採用無監督學習技術。這包括使用自編碼器或生成對抗網絡（GANs）等方法，這些模型能夠去學習怎麼對正常數據進行重構，並透過偵測與正常不同的地方來達到檢測異常的目的。

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.

對於object detection，我們需要準備包含圖像及其對應的ground truth的數據集，ground truth應包含給出空間答案的bounding box和類別答案的label。對於segmentation，我們需要準備包含圖像及其對應的像素級別的標籤（pixel mask）的數據集。

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

- 這些模型已經在大量數據集上進行了訓練和測試，因此它們具有較強的通用性，可以適應各種不同類型的數據和任務，並且也在廣泛的應用中得到驗證和優化。
- 我們可以利用這些模型提供的預訓練權重作為初始權重，從而加速訓練過程並提高自己模型的性能。