

National Tsing Hua University
11320IEEM 513600
Deep Learning and Industrial Applications
Homework 3

Name:

Student ID:

Due on 2025/04/10.

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.
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.)
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.

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.)
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?