National Tsing Hua University

11220IEEM 513600

Deep Learning and Industrial Applications

Homework 3

Name: 吳柏彥 Student ID: 112034525

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 ：toothbrush

Number of defect classes. ：2

Types of defect classes. ：defective、good

Number of images used in your dataset. ：2*10(num_image of each classes)

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.)

Epoch：50 lr：0.1 optimizer：adam model resnet：resnet18 weights：IMAGENET1K_V1 train loss：3.4341 train acc：50.0000% val loss：7.6183 val acc：50.0000% test acc：50%	Epoch：50 lr：0.1 optimizer：adam model resnet：resnet50 weights：IMAGENET1K_V1 train loss：4.3471 train acc：68.7500% val loss：78.4247 val acc：50.0000% test acc：50%	Epoch：50 lr：0.1 optimizer：adam model resnet：resnet50 weights：IMAGENET1K_V2 train loss：4.9882 train acc：56.2500% val loss：78.5655 val acc：50.0000% test acc：50%

Epoch：100 lr：0.1 optimizer：adam model resnet：resnet18 weights：IMAGENET1K_V1 train loss：5.0210 train acc：62.5000% val loss：0.5038 val acc：75.0000% test acc：100%	Epoch：100 lr：0.1 optimizer：adam model resnet：resnet50 weights：IMAGENET1K_V1 train loss：6.9809 train acc：43.7500% val loss：1.9522 val acc：50.0000% test acc：100%	Epoch：100 lr：0.1 optimizer：adam model resnet：resnet50 weights：IMAGENET1K_V2 train loss：5.0841 train acc：56.2500% val loss：5.2145 val acc：50.0000% test acc：100%
Epoch：150 lr：0.1 optimizer：adam model resnet：resnet18 weights：IMAGENET1K_V1 train loss：3.8201 train acc：75.0000% val loss：2.1714 val acc：75.0000% test acc：75%	Epoch：150 lr：0.1 optimizer：adam model resnet：resnet50 weights：IMAGENET1K_V1 train loss：6.1890 train acc：62.5000% val loss：3.8232 val acc：25.0000% test acc：100%	Epoch：150 lr：0.1 optimizer：adam model resnet：resnet50 weights：IMAGENET1K_V2 train loss：2.8111 train acc：81.2500% val loss：0.7001 val acc：75.0000% test acc：100%

發現調整參數 epoch 會明顯提高準確度，但可能是因為資料集太小的關係，所以在 epoch100 以上的時候，都出現 100%的準確度，這樣可能造成過度擬合的情況發生。在調整 model resnet 的時候發現 resnet50 必須要配權重為 IMAGENET1K_V2 才會有效提高訓練準確度，而利用權重為 IMAGENET1K_V1 時並不會提高訓練準確度。

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 refers to a statistical pattern where a small number of categories or values occur frequently, while the majority occur infrequently. It's important in fields like machine learning because it can affect model performance and decision-making.

   (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.

   Bag of Tricks for Image Classification with Convolutional Neural Networks

   該論文提出了多種方法來解決圖像分類中的數據不平衡問題。其中一個方法是使用類別加權（Class Weighting），即為不同類別的樣本分配不同的權重，使得模型更加關注罕見類別的樣本。另一個方法是使用過採樣（Oversampling），通過從少數類別中複製樣本或生成合成樣本，來平衡數據集中各類別的樣本數量。

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.)

    1. 使用生成對抗網絡（GANs）來合成缺陷圖像，以擴展訓練數據。
    2. 利用無監督學習方法，如自編碼器，來訓練模型對正常圖像的表示進行建模，從而檢測異常。
    3. 利用遷移學習，從其他數據集中的相似任務中學習特徵，然後微調到 MVTec AD 數據集。

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.

    物體檢測：需要準備帶有物體標註的圖像數據集，標註應該包括每個物體的類別標籤和邊界框（bounding box）信息。
    分割：需要準備帶有像素級別的標註的圖像數據集，標註應該將圖像中的每個像素分配給相應的類別，從而形成分割遮罩。

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

    模型適合進行微調，因為它們具有良好的泛化性能和在大型數據集上訓練的強大能力。通過在我們的自定義數據集上進行微調，可以使這些模型更好地適應我們的特定任務和數據分佈，從而提高異常檢測的準確性和效能。