National Tsing Hua University

11220IEEM 513600

Deep Learning and Industrial Applications

Homework 3

Name: 黃家興 Student ID: 111003806

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:
   - I selected the "Transistor" from the dataset and the number of classes: 5
   - Types of classes: 'bent_lead', 'cut_lead', 'damaged_case', 'misplaced', 'good'
   - Number of images used in your dataset: 50 images
   - Distribution of training and test data: 40 for training and 10 for test
   - Image dimensions: 1024 * 1024* 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.)

The result showed as following, both adjusting the number of epochs from 50 to 200(Exp.1) and changing the batch size from 32 to 8 (Exp.2) individually can get significant improvement. Increasing the number of epochs allowed the model to undergo more training iterations, enabling it to better refine its representations. Initially, the model might not have converged completely within 50 epochs, especially considering the relatively small dataset size. By training for more epochs, the model has more opportunities to converge to a better solution. On the other hand, smaller batch sizes often lead to faster convergence and lower generalization error. With a batch size of 8, the model updates its parameters more frequently, allowing it to adjust its weights based on a more diverse set of samples at each iteration. This can help the model generalize better to unseen data, resulting in improved test accuracy.

	As is	To be	Training		Validation(Best)		Test
			Loss	Accuracy	Loss	Accuracy	Accuracy
	Original		1.4593	47.50%	1.6142	30.00%	30.00%
Exp. 1	epoch =50	epoch=200	1.0328	60.00%	1.1696	70.00%	70.00%
Exp. 2	batch size =32	batch size =8	1.2522	52.50%	1.2307	60.00%	60.00%
Exp. 3	ResNet18_IMAGENET1K_V1	ResNet50_IMAGENET1K_V2	1.4331	47.50%	1.4645	40.00%	40.00%
Exp. 4	learning rate =1e-3	learning rate =1e-4	1.6287	32.50%	1.6257	40.00%	40.00%

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.'

   The foundation of the long-tail concept plots different data values to their respective frequencies. The "head" of the distribution which includes a small number of high-frequency items and the long 'tail' of the distribution which includes huge numbers of less frequent items.

   (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.

   This paper[1] addresses the complexity of imbalanced data in industrial processes and proposes the DAC (Data Augmentation Classifier) method as a solution to improve classification performance for fault diagnosis. By utilizing data augmentation techniques and focusing on high-quality data generation, the DAC framework aims to mitigate the negative impact of data imbalance and enhance the accuracy of classification results in practical industrial applications. Case study of the imbalanced binary classification problem and the complex multiclassification problem more challenge then our case. They proposed methods still can get the better result. So, that should 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.)

   There are several strategies that can be employed to address this issue. At first, use data augmentation techniques to artificially create more examples of defects. Then use transfer learning to leverage pre trained models on similar datasets. GANs can also be used to generate synthetic defects to augment the training set. Additionally, one-class classification can be used to detect anomalies. These strategies can help improve the model's ability to detect anomalies in the absence of defect examples in the training set.

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.

   If we would like to apply the object detection technique for anomaly detection, then we need to annotate the images for training. Identify the coordinates of defect images. Dataset for segmentation models needs labeled with the class of object of each pixel in the image.

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

   YOLO-World and SAM have been pre-trained on many kind of datasets. That should be help to our datasets. We can add the MVTec AD dataset to fine turn the model then can get better result.

[1] X. Jiang and Z. Ge, "Data augmentation classifier for imbalanced fault classification," *IEEE Transactions on Automation Science and Engineering,* vol. 18, no. 3, pp. 1206-1217, 2020.