National Tsing Hua University

11320IEEM 513600

Deep Learning and Industrial Applications

Homework 3

Name: 黃皓澤 Student ID: 113034519

1.

Bottle dataset

Number of classes: 4

Types of classes: broken_large, broken_small, contamination, good

Number of images used in your dataset: 40 images (10 images * 4 categories)

Distribution of training and test data: train 80%, test 20%

Image dimensions: 900*900

2.

	Attempt 1	Attempt 2	Attempt 3	Attempt 4
Epochs	50	50	50	100
Batch Size	4	4	4	4
Learning rate	0.003	0.0001	0.0001	0.0001
Ptimizer	Adam	Adam	RMSprop	Adam
Accuracy	58.33%	70.83%	62.5%	75.32%

In Attempt 1, reducing the batch size to 4 improved the model's convergence. In Attempt 2, lowering the learning rate to 0.0001 further enhanced the results, suggesting that a slower learning process was beneficial. Attempt 3, which involved switching to a different optimizer, led to poorer outcomes, indicating that the original optimizer was better suited for this task. Finally, in Attempt 4, increasing the number of training epochs resulted in the best performance overall. These findings underscore the importance of careful hyperparameter tuning—including batch size, learning rate, optimizer choice, and epoch count—to achieve optimal model performance.

3.

(i) A long-tail distribution refers to a dataset where a small number of classes (head classes) contain the majority of samples, while most classes (tail classes) have very few samples. This creates a "long tail" in the frequency distribution, causing models to favor head classes and struggle to learn tail class features effectively.

(ii) Paper: CVPR 2020 "Balanced Group Softmax (BAGS)"

Method: Addressing the differences between classification and detection tasks, BAGS divides classes into groups (head/tail), computes softmax loss separately for each group to prevent head classes from dominating gradients, and dynamically adjusts grouping for balanced learning.

Application: For MVTec AD, BAGS groups "Good" samples and defect classes separately, optimizing their classifiers independently. This reduces bias toward the "Good" class and improves defect detection without requiring additional data augmentation, balancing efficiency and accuracy.

4.

When training anomaly detection models on MVTec AD (with only normal images), strategies focus on learning robust representations of normality. Pre-trained deep networks extract features from defect-free samples, with anomalies flagged as deviations from these patterns via clustering or density estimation. Reconstruction-based methods like autoencoders or memory-augmented networks enforce strict reproduction of normal data, where defects yield high reconstruction errors. Advanced frameworks like PatchCore utilize memory banks of normal feature patches for efficient anomaly scoring. Synthetic defects (e.g., artificial noise) or self-supervised tasks (e.g., rotation prediction) further enhance sensitivity to irregularities without labeled anomalies, prioritizing generalization to unseen defect types.

5.

(i) For object detection, datasets must include images with bounding boxes around defects, each annotated with class labels (e.g., "scratch," "crack"). For segmentation, pixel-level masks are required, precisely outlining defect regions. Both tasks demand diverse defect examples across variations in lighting, scale, and orientation to ensure robustness. Additional metadata (e.g., defect size, location) can enhance model adaptability.

(ii) Models like YOLO-World and SAM are ideal for fine-tuning because:

Pre-trained Backbones: They leverage features learned from large-scale datasets (e.g., COCO), enabling transfer learning even with limited defect data.

Architectural Flexibility: YOLO's efficient detection pipeline handles multi-scale defects, while SAM's prompt-driven segmentation excels in zero/few-shot adaptation to unseen defect shapes.

Modular Design: Output layers (e.g., detection heads, mask decoders) can be re-trained without overhauling the entire network, reducing computational costs.

These traits allow rapid adaptation to industrial defect patterns while maintaining high precision.