HW3 113034512 黃冠毓

1. Select the type : wood
   - Number of defect classes : 5
   - Types of defect classes : color、combined、hole、liquid、scratch
   - Number of images used in your dataset : 68
   - Distribution of training and test data : 54 / 14
   - Image dimensions : Resized from (3, 1024, 1024) to (3, 32, 32)
2.

<table>
<tr><th>Parameters</th><th>Number of images</th><th>Pre-trained Model</th><th>Image dimensions</th><th>Epochs</th><th>Optimizer</th></tr>
<tr><td>Origin</td><td>68 (54 / 14)</td><td rowspan="2">Resnet18</td><td rowspan="3">(3, 32, 32)</td><td rowspan="4">100</td><td rowspan="5">Adam</td></tr>
<tr><td>1</td><td rowspan="4">98 (78 / 20)</td></tr>
<tr><td>2</td><td rowspan="3">Resnet34</td></tr>
<tr><td>3</td><td rowspan="2">(3, 128, 128)</td></tr>
<tr><td>4</td><td>250</td></tr>
</table>

Attempt	Train_Acc	Train_Loss	Val_Acc	Val_Loss
Origin	59.26%	1.3170	50.00%	1.9540
1	52.56%	1.2840	55.00%	1.7608
2	41.03%	1.5752	60.00%	1.5577
3	80.77%	0.7548	85.00%	0.6005
4	84.62%	0.5599	90.00%	0.4783

- Learning rate is 0.001，batch size is 32，Optimizer is Adam.
- The first attempt is to increase the size of the dataset while preventing data imbalance, without making excessive additions.
- The second attempt is to choose different pre-trained models, but it did not show significant improvement.
- The third attempt is to adjust the image size to change the image resolution.
- The fourth attempt is to increase the number of epochs for the model to learn more times.
- Among the methods mentioned above, the most significant effect is changing the image size, as this helps the model retain more details and information, allowing it to learn more complex features and patterns.

3.

I. Long-tail distribution usually refers to the phenomenon of class imbalance. Specifically, it means that certain classes or labels in the dataset occur with a very high frequency, while other classes or labels occur with a very low frequency. This leads to the model's inability to effectively recognize the minority classes during the learning process. It can affect the model's prediction accuracy, particularly when predicting the minority classes.

II. Paper: "MetaBalance: Improving Multi-Class Long-Tailed Learning through Meta-Sampled Decoupled Learning" (IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022)

This paper proposes a two-stage decoupled learning approach combined with meta-learning to address long-tail distribution issues:

i. **Representation Learning Stage**: Using a class-balanced sampling strategy to learn robust feature representations for all classes regardless of their sample sizes.

ii. **Classifier Learning Stage**: Applying a meta-learning framework that dynamically adjusts the sampling strategy based on model performance across all classes.

The MetaBalance method could be applied to our case: First train a balanced base model by oversampling rare defect classes; then fine-tune only the classifier using a meta-learning framework that dynamically shifts training focus to poorly performing defect types while applying intelligent weighting to prioritize rare defects.

4.

I. **Data Augmentation**: Enhance the few defect samples by applying transformations (e.g., rotation, scaling) to generate more varied defect examples.

II. **One-Class Classification**: Train the model using only "good" images to learn the normal pattern, and use this as a reference for detecting anomalies.

III. **Generative Models**: Utilize techniques like autoencoders or GANs to model normal data distribution, then detect anomalies as deviations from this learned distribution.

IV. **Transfer Learning**: Leverage pre-trained models from similar tasks to improve generalization on defect detection with limited defect data.

5.

I.

i. For Object Detection:

A. Labeled images with bounding boxes around defects/anomalies

B. Balanced dataset with sufficient examples of each defect type

C. Images representing various lighting conditions, angles, and environments

ii. For Segmentation:

A. Pixel-level masks that precisely outline the anomalous regions

B. Binary masks differentiating between normal and defective areas

C. Instance segmentation requires unique IDs for each separate anomaly instance

D. High-resolution images to capture fine details of defects

II.

i. They utilize transfer learning from large-scale pre-training, requiring less data for adaptation

ii. Their architectures are designed to identify visual patterns and features relevant to anomaly detection

iii. They offer flexibility in detecting objects of varying sizes and shapes, important for diverse defect types

iv. Models like YOLO-World and SAM have demonstrated strong zero-shot and few-shot learning capabilities