Workshop Challenge Report Template

Participant Information

• Name(s): 張景淵、楊峻銘、高煒軒

• Affiliation(s): National Tsing Hua University

• Contact Information: a0983102723@gmail.com

• Track: Category 2 — VLM Anomaly Challenge: Few-Shot Learning for Logical and Structural Detection

Abstract

The scarcity of training data in the field of anomaly detection poses a significant challenge, few-shot learning anomaly detection allows the model to maintain acceptable performance even in the absence of training data, which is addressed by our method that maintains acceptable performance even with limited data.

We employ the high-performance EfficientAD model, enhancing its performance through pretraining on the MVtec anomaly detection dataset. Modifications to EfficientAD have been made to improve its few-shot learning capabilities. The architecture has been adjusted to produce an image-level anomaly score, considering both the anomaly map of the testing data and the similarity between features of the reference and testing images.

Our training approach mirrors that of EfficientAD, with the added strategy of randomly re-selecting a category from the MVtec dataset every 20 epochs.

Introduction

• **Background**: Few-shot learning anomaly detection allows the model to maintain acceptable performance even in the absence of training data. This is especially important in the field of anomaly detection where training data is scarce

• **Challenge Description**: With the development of vision language models (VLMs), finding anomalies could reach an exciting new level, such as finding logical anomalies that need more than detecting structural defects.

Participants will create models using few-shot learning and VLMs to find and localize structural and logical anomalies. This shows that the models can handle structural defect detection and logical reasoning.

Methodology

Model Design

• **Approach:** We attempted to use the high-performance model, EfficientAD, and enhanced its performance by pretraining it on the MVtec anomaly detection dataset. Furthermore, we made some modifications to EfficientAD to improve its few-shot learning capabilities.

• **Architecture:** Based on EfficientAD, we modified the forward function to enable the model to produce an image-level anomaly score. This score takes into account the anomaly map of the testing data and similarity between features of reference image and testing image.

• **Training:** We essentially use the same training approach as we do for EfficientAD. The only difference is that we randomly re-select a category from the MVtec dataset every 20 epochs.

Dataset & Evalua2on

• **Dataset Utilization:** We use the Mvtec-ad as an additional dataset source for model training. Each image is formatted and pre-processed in accordance with the requirements of this competition.

• **Evaluation Criteria:** The overall performance did not achieve significant optimization, particularly with the image f1 scores showing almost no change, while the pixel f1 scores saw a slight improvement.

Results

• **Performance Metrics: F1-max**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | breakfast\_box | juice\_bottle | pushpins | screw\_bag | splicing\_connectors |
| Image-level | 0.7723 | 0.8339 | 0.7137 | 0.7821 | 0.7644 |
| Pixel-level | 0.1572 | 0.2582 | 0.0721 | 0.1758 | 0.1535 |

Discussion

• **Future Work:** Due to the insufficient time of this project, there are many tasks we haven't yet attempted, such as using different base pretrained models or incorporating additional datasets.

Conclusion

Overall, we experimented with the Efficient\_AD architecture and recorded the results for zero shot and few shot scenarios. Although there were no significant breakthroughs, we still manage to slightly improve the accuracy.

References

<https://github.com/openvinotoolkit/anomalib>

<https://openaccess.thecvf.com/content/CVPR2023/html/Jeong_WinCLIP_Zero-Few-Shot_Anomaly_Classification_and_Segmentation_CVPR_2023_paper.html>

<https://arxiv.org/abs/2303.14535>