BAAD: Background-Aware Anomaly Detection and Localization Network

[1,3]Pichdara Po, [2,3]Hyungwon Kim
[1]*Department of Computer Science,* [2]*Department of Electronics Engineers,*
[3]*Chungbuk National University*
{ pichdara , hwkim } @chungbuk.ac.kr

Abstract

Anomaly detection has become one of the most crucial parts in the industrial production process of various products, especially to detect anomaly products in real-time. This has become one of the mainstream and important tasks in efficiency of the model in computer vision application. Reconstruction-based anomaly detection is one of the mainstream methods to detect anomaly, is prone to incorrectly predicting background noise as anomalous region. We propose a Background-Aware Anomaly Detection network which takes background information into account to eliminate falsely predicted regions which excel in real-life industrial image that have complex background. BAAD not only eliminates background noise, yet it is efficient and fast because of the Multi-task Reconstruction Network which is able to reconstruct both image and background mask within one network.

1. Introduction

Recent years have seen significant advancements in deep learning techniques, leading to improved performance across various computer vision tasks. Simultaneously, researchers have developed neural network architectures that are faster, more efficient, and require fewer resources. Anomaly detection seeks to identify data points that significantly deviate from normal patterns. It is crucial in industrial manufacturing for reducing labor costs and improving product quality. However, the wide range of anomalies makes it challenging to gather comprehensive data for supervised learning, shifting the focus towards unsupervised anomaly detection methods.

Existing unsupervised anomaly detection methods can be classified into embedding-based and reconstruction-based approaches. Embedding-based methods typically use teacher-student networks or pre-trained models with feature distance measurements to identify differences between normal and abnormal data distributions.

However, these methods overlook the complexity of real-world industrial data, often characterized by intricate background textures. Such textures, unlike ideal smooth surfaces, can significantly impact anomaly detection results. While embedding-based methods are more robust to background interference by extracting discriminative features, reconstruction-based methods struggle due to their reliance on comparing original and reconstructed images. Industrial image backgrounds are typically considered normal, and inaccurate reconstruction of these areas can lead to false anomaly detections.

To tackle these problems, a straightforward approach is to distinguish between foreground and background areas using foreground detection models, and then enhance the results with post-processing. However, it requires the utilization of two separate models to achieve this goal. Therefore, inspired by EfficientAD[1], we propose a Background-Aware Anomaly Detection and Localization Network (BAAD), which contains a simple yet effective and efficient multi-task reconstruction sub-network (MTRS) and a discriminative sub-network. The MTRS sub-network is employed to accomplish two tasks, including foreground detection and reconstruction. This MTRS is fast and efficient enough to reach a real time inference speed while maintaining a good accuracy. Finally, the discriminative sub-network is used to perform anomaly detection.

2. Related works
2.1 Anomaly Detection Tasks

Visual anomaly detection is a rapidly expanding field of research with a wide range of applications, including medical imaging, autonomous driving, and industrial inspection. These applications often have unique characteristics, such as the use of image sequences in surveillance datasets or the various modalities present in medical imaging datasets like MRI, CT, and X-ray.

The introduction of the MVTecAD[2] dataset has significantly advanced the development of methods for industrial applications. This dataset includes 15 distinct inspection scenarios, each with a separate training set and test set. The training sets consist exclusively of normal images, such as defect-free screws, while the test sets include both normal and anomalous images. This setup mirrors a common real-world challenge where the types and locations of potential defects are unknown during the development of an anomaly detection system. As a result, it is both challenging and essential for methods to perform effectively when trained solely on normal images.

2.2 Anomaly Detection Methods

Traditional computer vision algorithms have been effective in industrial anomaly detection for decades, often processing images within milliseconds. However, as Bergmann et al. [2] noted, these methods fail when conditions like well-aligned objects are not met. Deep learning-based methods, on the other hand, handle such challenges more robustly.

A successful approach has been to apply outlier detection and density estimation, these methods include Gaussian Mixture Models, Normalizing flow and KNN algorithm. Bergmann et al. [2] introduce a student-teacher (S-T) framework for anomaly detection, where a pretrained, frozen CNN (the teacher) guides student networks. The students, trained only on normal images, struggle to match the teacher's output on anomalous images, enabling anomaly detection. Generative models like autoencoders[3], [4] and GANs have been widely used for anomaly detection.

Recent autoencoder-based methods work by accurately reconstructing normal images while producing inaccurate reconstructions for anomalous ones, allowing anomalies to be detected by comparing the reconstruction to the input image.

3. Proposed Method

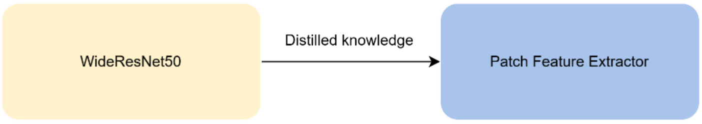
3.1 Lightweight Student-Teacher

The student-teacher approach is a machine learning technique where a "teacher" model, typically a pre-trained and more complex model, is used to guide the training of a "student" model. The student model is usually simpler and faster than the teacher. EfficientAD is one of the state-of-the-art approaches in the field of anomaly detection and localization. It is designed for rapid and accurate visual anomaly detection, and it excels in real-time applications where speed is crucial. it features a lightweight feature extractor called Patch Descriptor, which could extract features from input images under a millisecond on moder GPUs. The Patch Descriptor follows the student-Teacher approach which uses WideResNet-50 pre-trained on ImageNet dataset as the teacher and train to minimize the output feature similarity. It also features Autoencoder model integration for capturing a global map for Logical anomaly task.

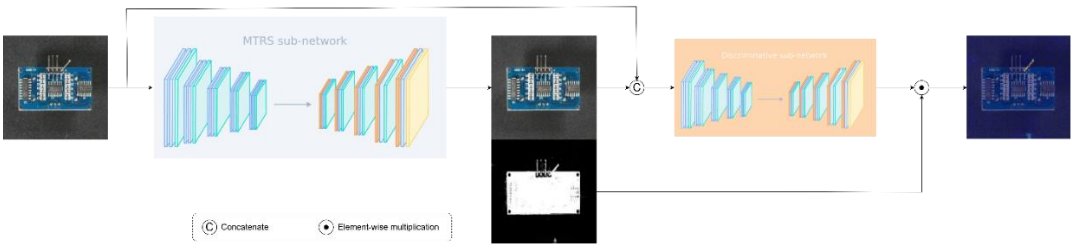
Inspired by EfficientAD's Student-Teacher architecture, we implemented the same approach for our MTRS sub-network. Pre-trained WideResNet-50 will be used to distill knowledge to a Patch Descriptor, which then will be used as our feature extractor for the MTRS to reconstruct the image and background mask.

3.2 Multi-task Reconstruction Network

A Multi-task Reconstruction Network is designed to simultaneously perform multiple reconstruction tasks. MTRS sub-network is designed to reconstruct both input image and background mask in a single network. MTRS share early layers to capture common underlying patterns among different tasks, then later layers are tailored to specific reconstruction. MTRS is proved to have improved performance while also being efficiency due to the shared parameters.



(a)



(b)

Figure 1: (a) Knowledge distillation from a pre-trained WideResNet-50 (b) Overview of the proposed BAAD network. BBAD comprises two sub-networks. The MTRS sub-network is responsible for simultaneous foreground detection and image reconstruction, while the discriminative sub-network serves for final anomaly detection.

4. Conclusion and Future works

In this study, Background-Aware Anomaly Detection and Localization network solves the problem of falsely predicted background region as anomalous region due to the complexity of the background. BAAD incorporates MTRS network which is a multitask reconstruction network capable of reconstructing both image and background mask to eliminate background noise at fast speed.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant for RLRC funded by the Korea government (MSIT) (No. 2022R1A5A8026986, RLRC) and was also supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2020-0-01304, Development of Self- learnable Mobile Recursive Neural Network Processor Technology). It was supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2023-2020-0-01462, Grand-ICT) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation). This research was supported by National R&D Program through the National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT (No. 2020M3H2A1076786, System Semiconductor specialist nurturing). **The corresponding Author is HyungWon Kim.**

References

[1] K. Batzner, L. Heckler, and R. König, "EfficientAD: Accurate Visual Anomaly Detection at Millisecond-Level Latencies," 2023, doi: 10.48550/ARXIV.2303.14535.

[2] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, "MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA: IEEE, Jun. 2019, pp. 9584–9592. doi: 10.1109/CVPR.2019.00982.

[3] P. Bergmann, K. Batzner, M. Fauser, D. Sattlegger,

and C. Steger, “Beyond Dents and Scratches: Logical Constraints in Unsupervised Anomaly Detection and Localization,” *Int. J. Comput. Vis.*, vol. 130, no. 4, pp. 947–969, Apr. 2022, doi: 10.1007/s11263-022-01578-9.

[4] C. Baur, B. Wiestler, S. Albarqouni, and N. Navab, “Deep Autoencoding Models for Unsupervised Anomaly Segmentation in Brain MR Images,” in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11383, A. Crimi, S. Bakas, H. Kuijf, F. Keyvan, M. Reyes, and T. Van Walsum, Eds., in Lecture Notes in Computer Science, vol. 11383. , Cham: Springer International Publishing, 2019, pp. 161–169. doi: 10.1007/978-3-030-11723-8_16.