h_da
HOCHSCHULE DARMSTADT
UNIVERSITY OF APPLIED SCIENCES

Few/Zero-Shot Anomaly Detection on Textile Surfaces

Patrick Wessely

Referentin: Prof. Dr. Elke Hergenröther | Korreferent: Prof. Dr. Andreas Weinmann

Motivation

In textile manufacturing, efficiently detecting fabric defects is essential for minimizing production costs and maintaining high-quality standards. Traditional anomaly detection (AD) methods based on handcrafted features or basic edge detection techniques struggle to handle the diverse range of fabric patterns, textures, and machinery in this industry. As a result, these methods often fail to reliably identify defects in complex and variable manufacturing environments, which can lead to increased waste and lowered product quality. To address these limitations, researchers have increasingly turned to deep learning (DL)-based methods that use neural networks to automatically learn intricate patterns and detect defects with high accuracy. However, the widespread application of DL-based models is hindered by their reliance on extensive labeled datasets, which are not only costly to produce but are also impractical given the infrequent occurrence of defects.

In response to these data limitations, recent research has explored unsupervised and, more recently, zero-shot and few-shot learning approaches for anomaly detection. These methods aim to reduce the dependency on labeled data by either training on unlabeled "normal" data or using only minimal labeled samples. Specifically, zero-shot and few-shot techniques have shown potential for adapting pre-trained neural networks to detect anomalies in previously unseen or new types of fabrics with little to no additional data. This study explores these emerging zero-shot and few-shot techniques, particularly focusing on models like CLIP [1], which integrates both image and text data for flexible and rapid classification. By testing CLIP-based algorithms for textile defect detection and localization, this research seeks to offer a scalable, adaptable solution that is both reliable and efficient across different fabric types and manufacturing setups.

Research Questions

Challenges in applying zero-shot and few-shot methods for AD include ensuring robustness and maintaining fast inference speeds, which are crucial for real-time defect detection on textile surfaces. This study aims to address the following research questions in the context of vision-language models, such as CLIP, repurposed for AD:

- How robust is a pre-trained vision-language model like CLIP when adapted to a zero-shot AD setting?
- Can fine-tuning few-shot models yield better results when compared to the zero-shot AD approach?
- Could a few-shot model fine-tuned on domain specific data be used in a cross-dataset AD setting for defect detection on unseen data?
- Does increasing the size of the image given to the zero-shot and fewshot models increase the model performance?

Methodology

In this study, five anomaly detection models, namely WinCLIP, WinCLIP+ [2], AnomalyCLIP [3], SDP, and SDP+ [4], were trained and evaluated using both zero-shot and few-shot techniques across three image sizes: 240x240, 480x480, and 600x600 pixels. The evaluation centered on testing how effectively these models could detect defects on two datasets: the simpler charmeuse fabric dataset, characterized by low defect variety and higher visibility of anomalies, and the more complex C05 fabric dataset, which presents a more challenging range of defect types and greater variation in texture. Zero-shot models were evaluated directly on test sets without training, while few-shot models were trained on limited data to improve adaptability to new patterns. This testing allowed for insights into how different model architectures and image sizes impact defect detection performance under each setting, especially in handling high-variability data.

The study's key methodological contribution was its cross-dataset testing approach, designed to assess the models' ability to generalize knowledge across different datasets, thus creating a new form of zero-shot testing from domain-adapted few-shot models. By training the few-shot models on one dataset and testing on the other, the study aimed to determine if a model fine-tuned on specific domain data could detect defects on entirely new, unseen data without additional training. To further enhance image classification performance, the SDP+ and AnomalyCLIP models were modified by leveraging their pixel-level anomaly maps. Specifically, rather than relying on direct image-level predictions, an anomaly score was computed from the pixel-level outputs using the maximum score of the anomaly map to classify images as defective or clean.

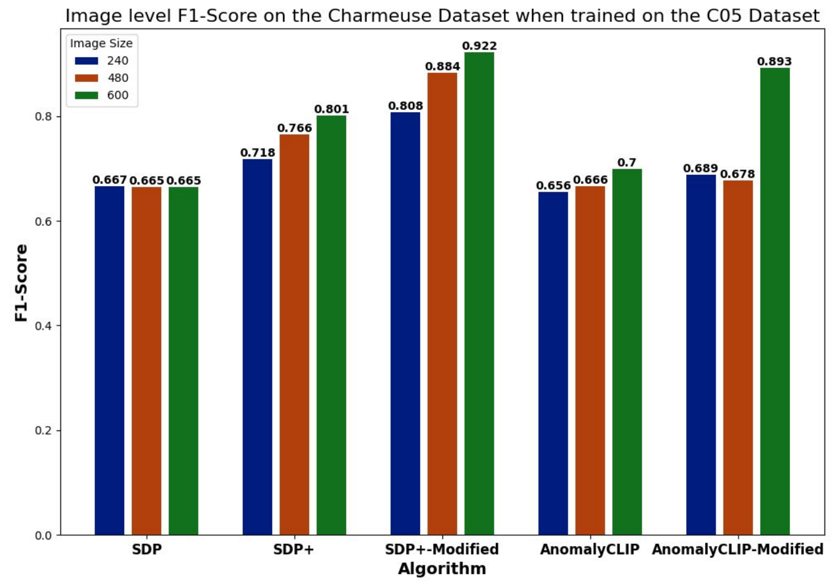

Figure 1: Results from the Cross-Dataset evaluation

Results

The WinCLIP approach and its extension, WinCLIP+, were excluded from this study due to poor runtime performance, allowing the focus to shift to three key models: SDP, SDP+, and AnomalyCLIP. Results showed that while the zero-shot model, SDP, adapted with a CLIP backbone, was ineffective for anomaly detection, the few-shot models achieved robust performance. CLIP model's general pre-training on publicly available, non-domain-specific images did not translate well to this industrial context. In contrast, the few-shot models, SDP+ and AnomalyCLIP, were fine-tuned with nominal and defective examples and displayed strong classification and localization capabilities, achieving high accuracy even on the more challenging C05 dataset. The SDP+ model excelled in pixel-level defect localization, achieving high dice scores across datasets. This reliable segmentation suggested that using anomaly maps for image-level classification could enhance performance, leading to the successful modification strategy that further improved defect detection accuracy.

Additionally, a cross-dataset evaluation demonstrated that fine-tuned few-shot models could function as zero-shot models with embedded domain knowledge when adapted to novel datasets. When trained on Charmeuse and evaluated on C05, the modified SDP+ model, as shown in Figure 1, displayed impressive image-level classification, achieving a F1 of 0.922 and significantly outperforming the original zero-shot SDP model. This cross-dataset approach effectively transformed a few-shot into a domain-informed zero-shot model.

Conclusion

This study demonstrates that zero-shot and few-shot approaches can address the limitations of data dependency in anomaly detection. Among the models tested, SDP+ excelled in both few-shot learning and cross-dataset transfer, especially when modified. The ability to transfer fine-tuned models to unseen datasets highlights the potential of few-shot training to generate zero-shot models with embedded domain knowledge. Additionally, larger image sizes generally improved performance though returns diminished beyond 480x480 pixels. These findings suggest that fine-tuning on specific data and selecting optimal image sizes are key to enhancing zero- and few-shot models' adaptability and effectiveness in detecting anomalies in novel industrial settings.

Literature

[1] Alec Radford et al. Learning Transferable Visual Models From Natural Language Supervision. 2021.

[2] Jongheon Jeong, Yang Zou, Taewan Kim, Dongqing Zhang, Avinash Ravichandran, and Onkar Dabeer. WinCLIP: Zero-/Few-Shot Anomaly Classification and Segmentation. 2023.

[3] Qihang Zhou, Guansong Pang, Yu Tian, Shibo He, and Jiming Chen. AnomalyCLIP: Object-agnostic Prompt Learning for Zero-shot Anomaly Detection. 2024.

[4] Xuhai Chen, Jiangning Zhang, Guanzhong Tian, Haoyang He, Wuhao Zhang, Yabiao Wang, Chengjie Wang, and Yong Liu. CLIP-AD: A Language-Guided Staged Dual-Path Model for Zeroshot Anomaly Detection. 2024.