Classifying Texture Anomalies at First Sight

Andrei-Timotei Ardelean
Friedrich-Alexander-Universität Erlangen-Nürnberg
Erlangen, Germany
timotei.ardelean@fau.de

Tim Weyrich
Friedrich-Alexander-Universität Erlangen-Nürnberg
Erlangen, Germany
tim.weyrich@fau.de

ACM Reference Format:
Andrei-Timotei Ardelean and Tim Weyrich. 2024. Classifying Texture Anomalies at First Sight. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH Posters '24), July 27–August 01, 2024, Denver, CO, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3641234.3671071

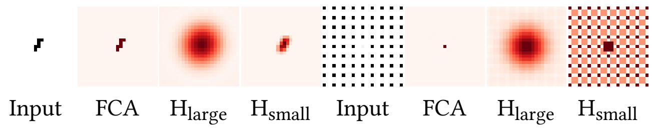

Figure 1: Anomaly localization: comparing FCA to histograms on two synthetic examples. H_{large} and H_{small} denote histograms with a large and small patch size, respectively.

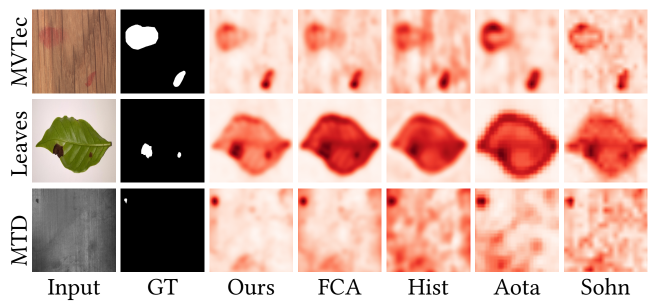

Figure 2: Qualitative comparison between our full method, just FCA, and prior work [Aota et al. 2023; Sohn et al. 2023].

1 INTRODUCTION

The problem of detecting and localizing defects in images has been tackled with various approaches, including traditional computer vision techniques, as well as machine learning. Notably, most of these efforts have been directed toward the normality-supervised setting of this problem. That is, these algorithms assume the availability of a curated set of *normal* images, known to not contain any anomalies. While this kind of data is easier to acquire than anomaly-annotated images, it is still costly or difficult to obtain in-domain data for certain applications.

We address the problem of anomaly detection and localization under a training-set-free paradigm and do not require any anomaly-free reference data. Concretely, we introduce a truly zero-shot method that can localize anomalies in a single image of a previously unobserved texture class. Then, we develop a mechanism to leverage additional test images, which may contain anomalies. Furthermore, we extend our analysis to also include a categorization of the anomalies in the given population through clustering. Importantly, we focus our attention on textures and texture-like images as we develop an anomaly detection method for structural defects, rather than logical anomalies. This poster summarizes our recent line of research on localization and classification of anomalies in real-world texture images [Ardelean and Weyrich 2024a,b].

2 APPROACH

Zero-shot. Our zero-shot method, named Feature Correspondence Analysis (FCA) [Ardelean and Weyrich 2024b], detects anomalies by finding the regions in an image that break the overall homogeneity of the texture. It is, at heart, a method for comparing the local pixel distribution in a patch with the global distribution of the image. Trivial distances between distributions, such as comparing moments or the EMD distance between histograms [Moritz et al. 2017], generally lead to anomaly maps with low fidelity due to the trade-off between context size and precision (Figure 1). Differently from other stationarity measures, we compute a mismatch score for **each** pixel in a patch instead of a holistic distance between the patch and the reference distribution (global statistic). This mismatch is computed by leveraging the fast algorithm for Wasserstein distance between empirical distributions [Elnekave and Weiss 2022]. That is, the elements of the distributions are sorted, creating a bijective mapping between values with the same rank. The absolute difference between the elements put in correspondence represents their contribution to the Wasserstein distance; we use this contribution as the mismatching score of a pixel in the context of a given patch. The final anomaly map is then obtained by averaging the scores from all patches that contain a certain pixel.

Blind anomaly localization. FCA localizes anomalies with high fidelity when applied to stationary textures. However, when the complexity of the global distribution increases, the method struggles to differentiate between genuine variance and true anomalies. When operating on a set of images, additional information can be extracted from this set, despite it being contaminated by anomalies. This unsupervised learning setting (blind anomaly detection) seeks to detect anomalies by using the common information in the given images to cooperatively uncover the outliers.

We adapt FCA to leverage the additional unlabeled data to distinguish between the normal variation in the images and the true anomalies. To this end, we employ a Variational Autoencoder (VAE), trained to reconstruct all features in the input images. Thanks to the training dynamics of the VAE, common features are reconstructed better than rare features, which are maily generated by

SIGGRAPH Posters '24, July 27–August 01, 2024, Denver, CO, USA

ACM ISBN 979-8-4007-0516-8/24/07
https://doi.org/10.1145/3641234.3671071

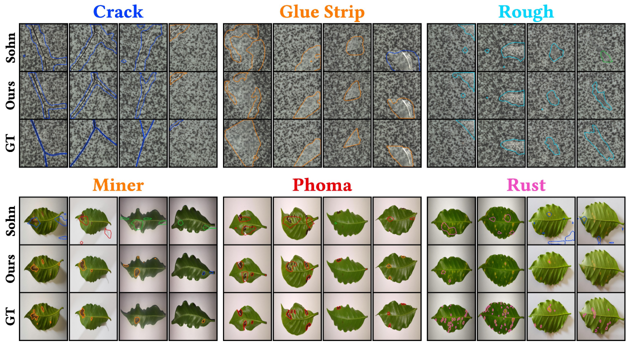

Figure 3: Qualitative evaluation of the blind localization (contours) and clustering (color). For more examples, see [Ardelean and Weyrich 2024a].

Table 1: Ablation and quantitative comparison to the recent baseline [Sohn et al. 2023]. PRO is a pixel-level anomaly localization metric, whereas NMI and F_1 evaluate clustering.

Method	MVTec textures			MTD			Leaves		
	PRO	NMI	F_1	PRO	NMI	F_1	PRO	NMI	F_1
Sohn *et* al.	90.84	0.67	0.71	66.51	0.18	0.35	75.25	0.46	0.63
Only FCA	96.92	0.76	0.78	72.15	0.18	0.51	48.10	0.33	0.44
FCA + CL	96.92	**0.81**	**0.81**	72.15	0.17	0.54	48.10	0.53	0.71
FCA + VAE	**97.50**	0.77	0.78	**75.53**	0.14	0.51	**77.20**	0.51	0.59
Ours full	**97.50**	0.79	**0.81**	**75.53**	**0.22**	**0.67**	**77.20**	**0.71**	**0.83**

anomalies. This concept has been widely used for stand-alone anomaly detection methods in a normality-supervised setting. Namely, by taking the L1 or L2 norm of the difference between original and reconstructed features, one directly gets an anomaly score for each pixel location. Our algorithm goes a step further and uses not only the magnitude of the residual features, but also their structure. Concretely, we apply our zero-shot anomaly localization on these residuals. The advantage of applying FCA on residuals rather than the original features comes from the homogenization of normal features through the VAE. This is supported by our experiments (see Table 1 and Figure 2), which show that the most significant improvements can be seen on more complex, heterogeneous textures.

Clustering. The localization of anomalies already makes the predictions both explainable and more actionable. A natural step forward in this direction is to further describe the discovered anomalies by identifying semantically distinct classes. In the case where no labels are available, this characterization can be formulated as a clustering problem. We are the first to explicitly make the connection between blind anomaly localization and anomaly clustering, splitting the problem in two complementary parts [Ardelean and Weyrich 2024a]. In essence, we leverage predicted anomaly maps to learn task-oriented features through contrastive learning.

Firstly, the predicted anomaly scores are used as weights for the initial feature maps (extracted using a pretrained WideResnet-50) to compute image-level descriptors, which attend to the anomalous regions in the image. Each image is matched with its k-nearest

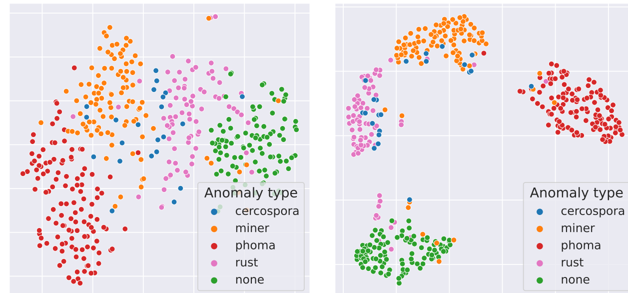

Figure 4: T-SNE projections of image-level descriptors; comparing the initial features (left) with the feature space obtained through contrastive learning (right).

neighbors for positive pairs selection and a random set from the bottom 0.5-quantile for negative pairs. Then, the mined positive and negative pairs are used to train a small convolutional network head on top of the initial feature extractor. The weights of this network are optimized using a contrastive loss function [Hadsell et al. 2006]. Finally, we recompute image-level descriptors using the learned features and employ Ward hierarchical clustering on these descriptors. As can bee seen in Table 1 and Figure 4, the contrastive learning is instrumental in obtaining a good clustering. The results of our full method are displayed in Figure 3.

3 CONCLUSION

Our work pushes the boundaries of texture analysis in the context of anomaly detection and clustering without using annotated data or any anomaly-free reference images. The results obtained in this challenging scenario sets a new baseline for data-scarce applications. Given the importance of textures to computer graphics applications, we foresee a great potential in the context of content creation from real-world photographs.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 956585.

REFERENCES

Toshimichi Aota, Lloyd Teh Tzer Tong, and Takayuki Okatani. 2023. Zero-Shot Versus Many-Shot: Unsupervised Texture Anomaly Detection. In *IEEE/CVF Winter Conference on Applications of Computer Vision.*

Andrei-Timotei Ardelean and Tim Weyrich. 2024a. Blind Localization and Clustering of Anomalies in Textures. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops.*

Andrei-Timotei Ardelean and Tim Weyrich. 2024b. High-Fidelity Zero-Shot Texture Anomaly Localization Using Feature Correspondence Analysis. In *IEEE/CVF Winter Conference on Applications of Computer Vision.*

Ariel Elnekave and Yair Weiss. 2022. Generating natural images with direct Patch Distributions Matching. In *European Conference on Computer Vision.*

R. Hadsell, S. Chopra, and Y. LeCun. 2006. Dimensionality Reduction by Learning an Invariant Mapping. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition.*

Joep Moritz, Stuart James, Tom S.F. Haines, Tobias Ritschel, and Tim Weyrich. 2017. Texture Stationarization: Turning Photos into Tileable Textures. *Computer Graphics Forum (Proc. Eurographics)* (2017).

Kihyuk Sohn, Jinsung Yoon, Chun-Liang Li, Chen-Yu Lee, and Tomas Pfister. 2023. Anomaly clustering: Grouping images into coherent clusters of anomaly types. In *IEEE/CVF Winter Conference on Applications of Computer Vision.*