Supplementary Material for Anomaly Detection and Localization Using Attention-Guided Synthetic Anomaly and Test-Time Adaptation

Behzad Bozorgtabar[1,2]
behzad.bozorgtabar@epfl.ch
Dwarikanath Mahapatra[3]
dwarikanath.mahapatra@inceptioniai.org
Jean-Philippe Thiran[1,2]
jean-philippe.thiran@epfl.ch

[1] École Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland
[2] Lausanne University Hospital (CHUV)
Lausanne, Switzerland
[3] Inception Institute of AI (IIAI)
Abu Dhabi, UAE

In this supplemental draft, we provide further details on the datasets (Section A) and additional qualitative results (Section B).

A Further Details on the Datasets

MVTec AD dataset [1] is composed of 15 categories (5 textures and 10 object categories) representing 88 different anomalies. The examples of anomaly-free images, and various types of anomalous images per category are shown in Figure 1.

NIH dataset [3] is composed of 112,120 frontal-view X-ray images labeled either as normal or with one or more of the 14 classes of thoracic diseases from 30,805 patients. We split patient-level data without patient ID overlap among the subsets (training and test set). However, we filter out abnormal X-rays from the training set, leaving us with 50,500 normal training X-ray images (28,212 male and 22,288 female patients). The test set contains a subset with anomaly localization annotations, which we use for evaluating the anomaly localization methods trained on anomaly-free training images. In particular, the test subset contains bounding box annotations (weak labels) for 880 X-ray images (503 for male and 377 for female patients) with 984 labeled bounding boxes of 8 abnormalities.

B Additional Qualitative Results

We use t-SNE [2] feature visualization to visualize the distributions of learned features from our model trained on the anomaly-free samples of the MVTec AD dataset. Here, in Figure 2, we illustrate the learned features of the backbone before the projection head on four

© 2022. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

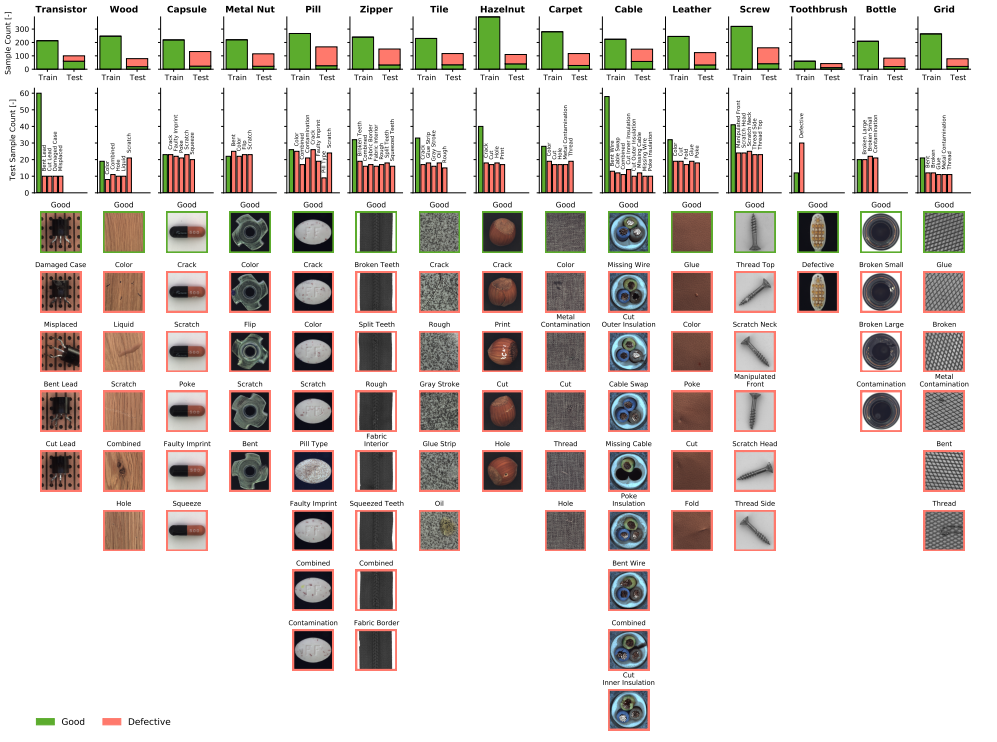

Figure 1: **Visual insight into the MVTec AD dataset**. The MVTec AD dataset is composed of 15 categories (5 textures and 10 object categories) representing 88 different anomalies. The examples of anomaly-free images and various types of anomalous images per category are shown.

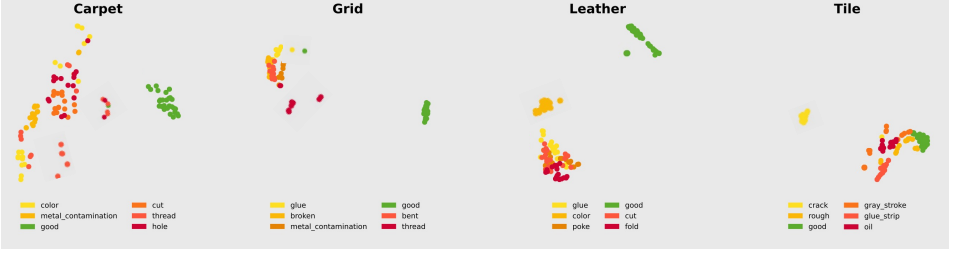

Figure 2: **The t-SNE visualization on the MVTec AD dataset**. The t-SNE visualization of the learned features (before the projection head). The green dots represent nominal features for four categories. The visualization for each category is also presented by anomaly type, each with a different color. **Zoom** in for the **best view**.

categories from the MVTec AD dataset. The t-SNE projection results demonstrate well-separated feature distribution (normal vs. anomaly) in these four categories.

References

[1] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Mvtec ad–a comprehensive real-world dataset for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9592–9600, 2019.

[2] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[3] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2097–2106, 2017.