DSR – A dual subspace re-projection network for surface anomaly detection

Supplementary material

Vitjan Zavrtanik, Matej Kristan, and Danijel Skočaj

University of Ljubljana, Faculty of Computer and Information Science
{vitjan.zavrtanik, matej.kristan, danijel.skocaj}@fri.uni-lj.si

1 DSR Localization evaluation - MVTec

In Figure 1 results using both the commonly used pixel-wise AUROC and the pixel-wise AP metric are shown. The pixel-wise AP metric is more robust to class imbalance which is why it is more appropriate than the pixel-wise AUROC for localization evaluation in anomaly detection tasks.

Class	US[2]	RIAD[5]	PaDim[4]	DRÆM [6]	DSR
bottle	97.8 / 74.2	98.4 / 76.4	98.2 / 77.3	99.1 / 86.5	98.9 / 91.5
capsule	96.8 / 25.9	92.8 / 38.2	98.6 / 46.7	94.3 / 49.4	95.4 / 53.3
grid	89.9 / 10.1	98.8 / 36.4	97.1 / 35.7	99.7 / 65.7	99.6 / 68.0
leather	97.8 / 40.9	99.4 / 49.1	99.0 / 53.5	98.6 / 75.3	99.6 / 62.5
pill	96.5 / 62.0	95.7 / 51.6	95.7 / 61.2	97.6 / 48.5	93.4 / 65.7
tile	92.5 / 65.3	89.1 / 52.6	94.1 / 52.4	99.2 / 92.3	98.2 / 93.9
transistor	73.7 / 27.1	87.7 / 39.2	97.6 / 72.0	90.9 / 50.7	83.2 / 41.1
zipper	95.6 / 36.1	97.8 / 63.4	98.4 / 58.2	98.8 / 81.5	98.9 / 78.5
cable	91.9 / 48.2	84.2 / 24.4	96.7 / 45.4	94.7 / 52.4	96.7 / 70.4
carpet	93.5 / 52.2	96.3 / 61.4	99.0 / 60.7	95.5 / 53.5	95.5 / 78.2
hazelnut	98.2 / 57.8	96.1 / 33.8	98.1 / 61.1	99.7 / 92.9	99.2 / 87.3
metal nut	97.2 / 83.5	92.5 / 64.3	97.3 / 77.4	99.5 / 96.3	93.7 / 67.5
screw	97.4 / 7.8	98.8 / 43.9	98.4 / 21.7	97.6 / 58.2	98.5 / 52.5
toothbrush	97.9 / 37.7	98.9 / 50.6	98.8 / 54.7	98.1 / 44.7	99.5 / 74.2
wood	92.1 / 53.3	85.8 / 38.2	94.1 / 46.3	96.4 / 77.7	92.5 / 68.4
avg	93.9 / 45.5	94.2 / 48.2	97.4 / 55.0	97.3 / 68.4	96.2 / 70.2

Table 1: Results of (AUROC/AP) anomaly localization on MVTec dataset.

2 Qualitative examples - KSDD2

Qualitative examples of DSR on the KSDD2 dataset are shown in Figure 1 and Figure 2.

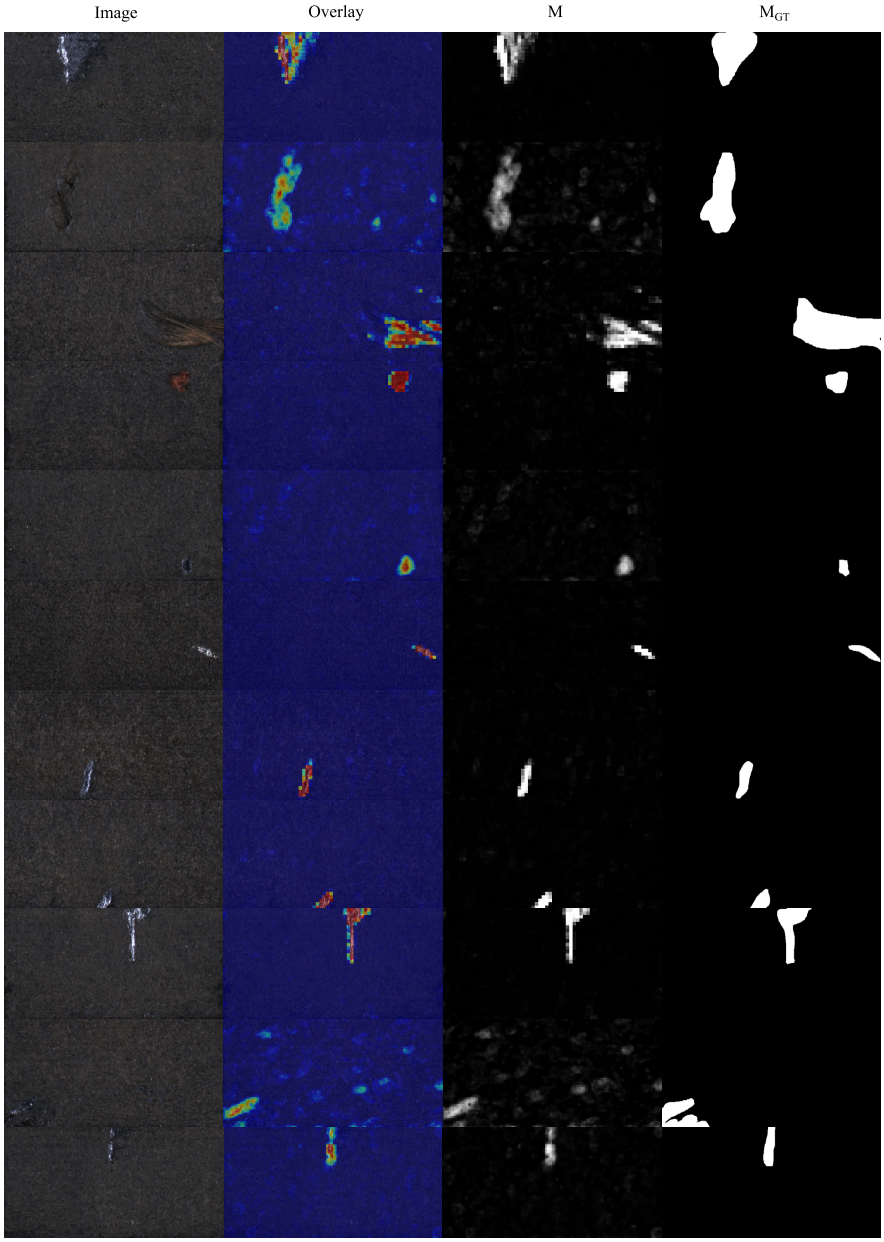


Fig. 1: Qualitative examples on the KSDD2 dataset [3]. The input image, overlaid DSR output map, DSR output map and the ground truth map are shown from left to right.

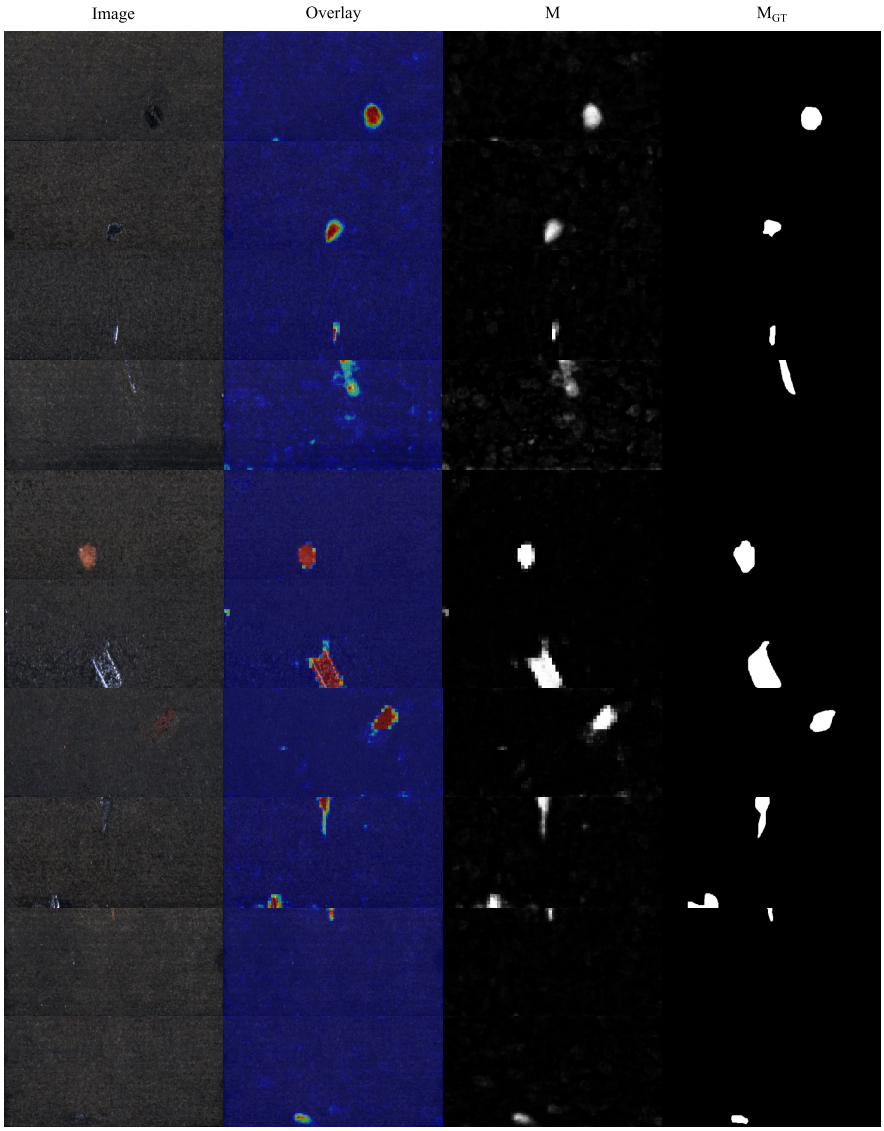


Fig. 2: Qualitative examples on the KSDD2 dataset [3]. The input image, overlaid DSR output map, DSR output map and the ground truth map are shown from left to right.

3 Qualitative examples - MVTec

Qualitative examples on the MVTec dataset [1] are shown in Figures 3, 4, 5 and 6. In Figure 8 examples are shown, where the output of DSR fails to exactly match the ground truth, sometimes due to the ambiguity of the anomalies and the ground truth labels.

4 Generated training anomaly examples

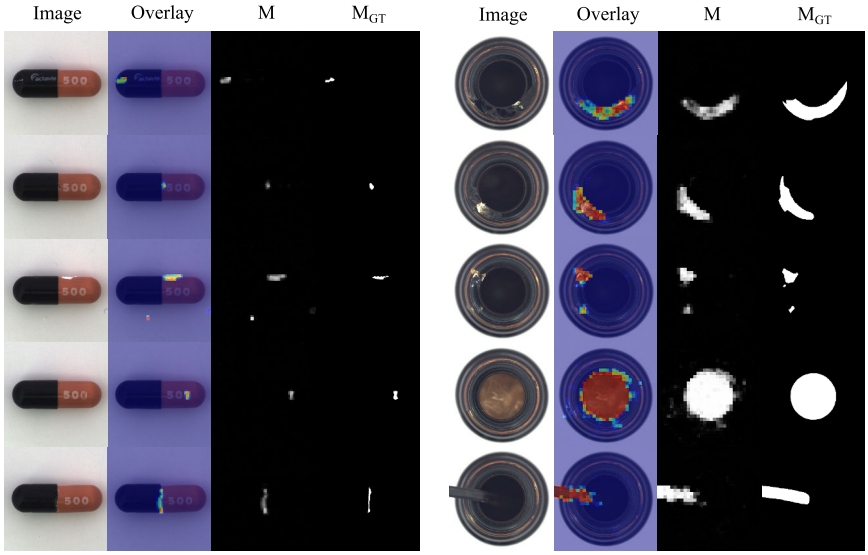
5 DSR - Reference implementation

A reference implementation of DSR is also accompanying the paper:

https://github.com/VitjanZ/DSR_anomaly_detection

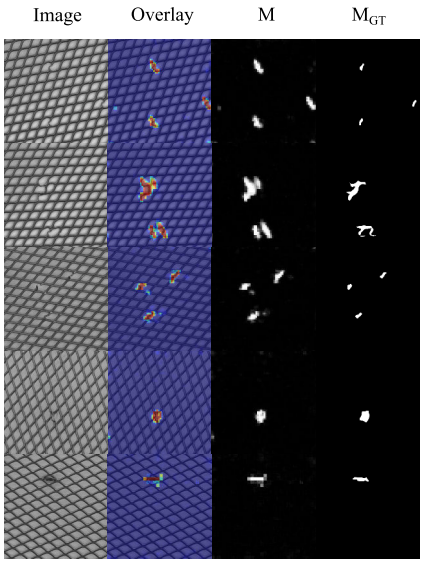
References

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 Conference on Computer Vision and Pattern Recognition*, pages 9592–9600, 2019. 4, 5, 6, 7, 8
2. Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Uninformed students: Student-teacher anomaly detection with discriminative latent embeddings. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4183–4192, 2020. 1
3. Jakob Božič, Domen Tabernik, and Danijel Skočaj. Mixed supervision for surface-defect detection: from weakly to fully supervised learning. *Computers in Industry*, 129:103459, 2021. 2, 3
4. Thomas Defard, Aleksandr Setkov, Angelique Loesch, and Romaric Audigier. Padim: A patch distribution modeling framework for anomaly detection and localization. In *International Conference on Pattern Recognition*, pages 475–489. Springer, 2021. 1
5. Vitjan Zavrtanik, Matej Kristan, and Skočaj Danijel. Reconstruction by inpainting for visual anomaly detection. *Pattern Recognition*, 2020. 1
6. Vitjan Zavrtanik, Matej Kristan, and Danijel Skočaj. Draem - a discriminatively trained reconstruction embedding for surface anomaly detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 8330–8339, October 2021. 1

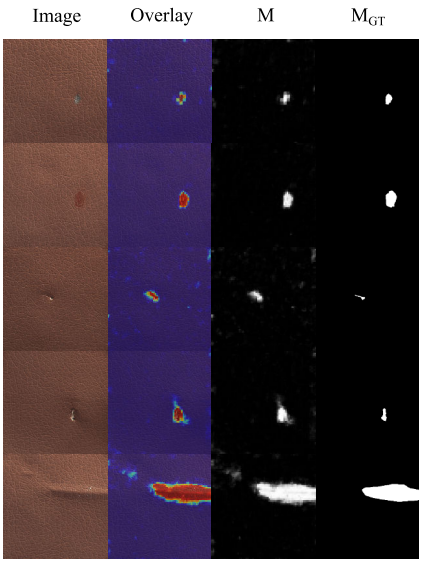


(a) Capsule

(b) Bottle

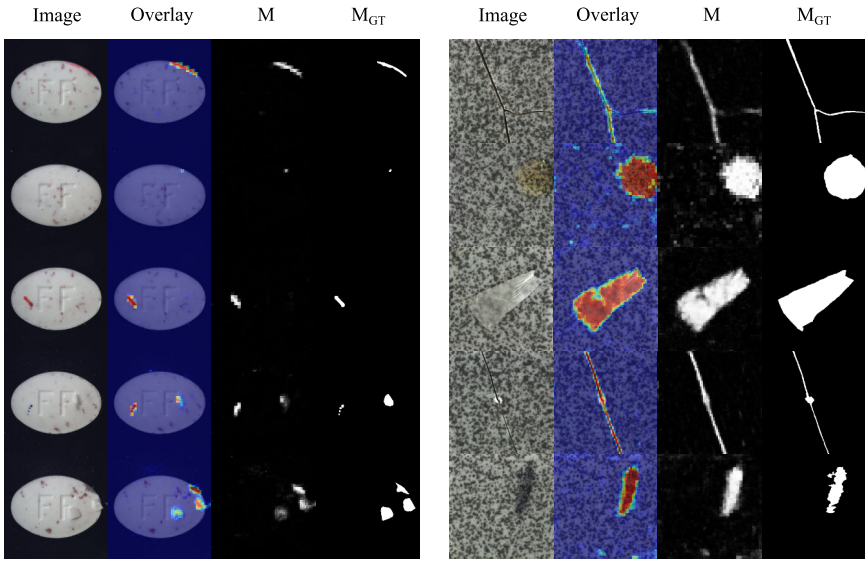


(c) Grid

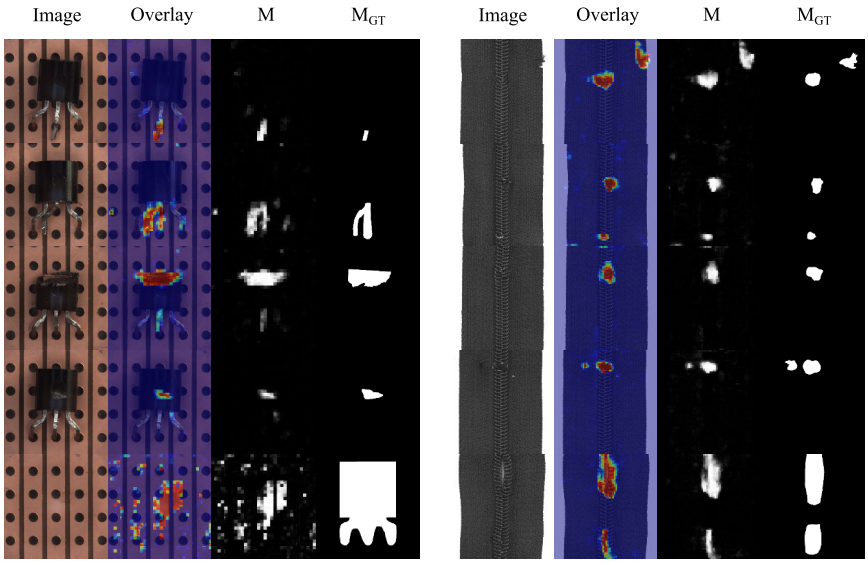


(d) Leather

Fig. 3: Qualitative examples on the MVTec dataset [1]. The input image, overlaid DSR output map, DSR output map and the ground truth map are shown from left to right.

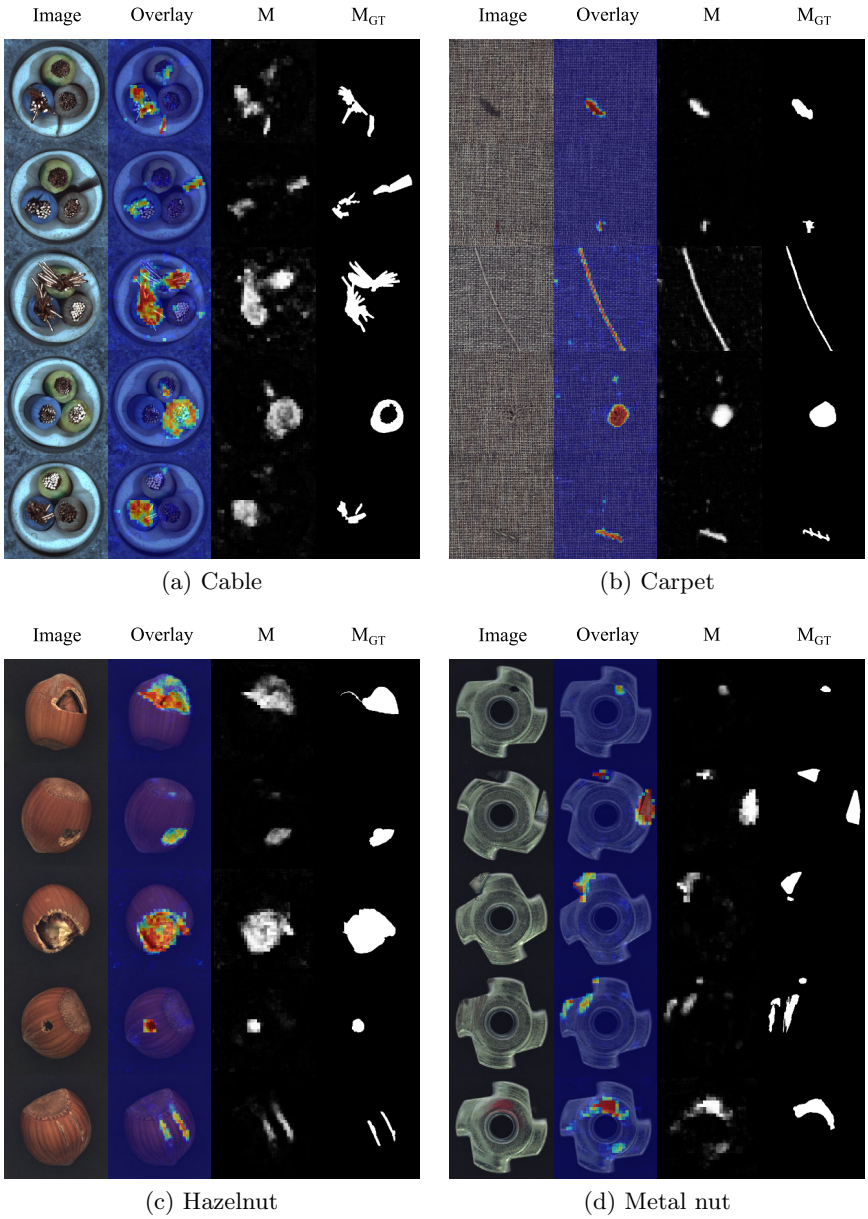


(a) Pill (b) Tile



(c) Transistor (d) Zipper

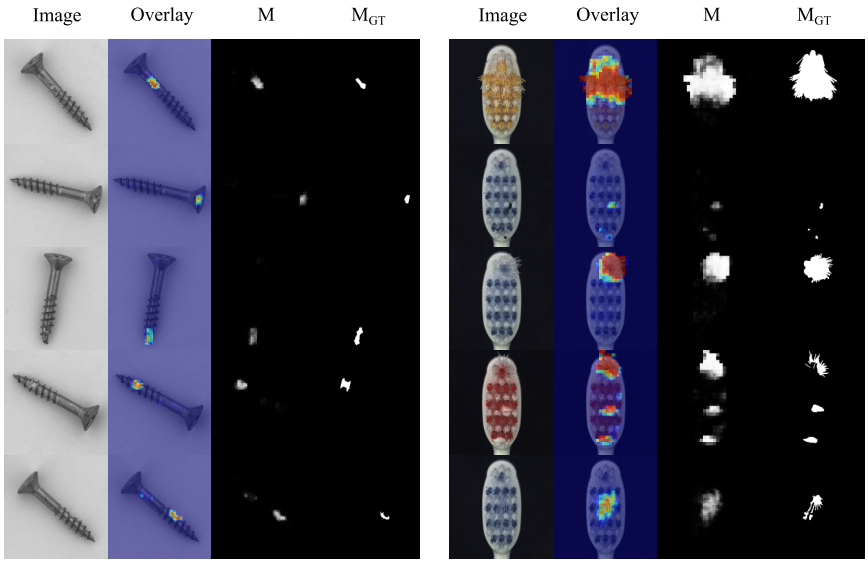
Fig. 4: Qualitative examples on the MVTec dataset [1]. The input image, overlaid DSR output map, DSR output map and the ground truth map are shown from left to right.



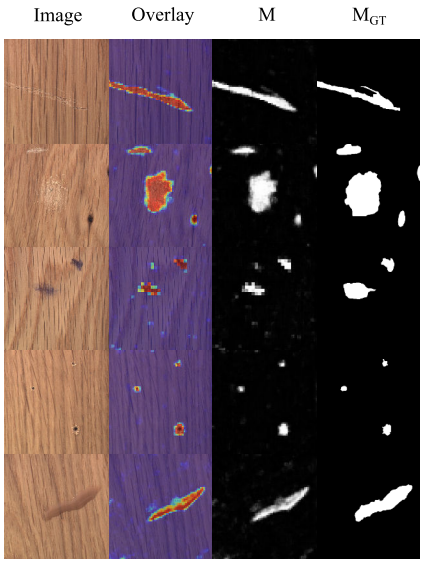
(a) Cable (b) Carpet

(c) Hazelnut (d) Metal nut

Fig. 5: Qualitative examples on the MVTec dataset [1]. The input image, overlaid DSR output map, DSR output map and the ground truth map are shown from left to right.



(a) Screw (b) Toothbrush



(c) Wood

Fig. 6: Qualitative examples on the MVTec dataset [1]. The input image, overlaid DSR output map, DSR output map and the ground truth map are shown from left to right.

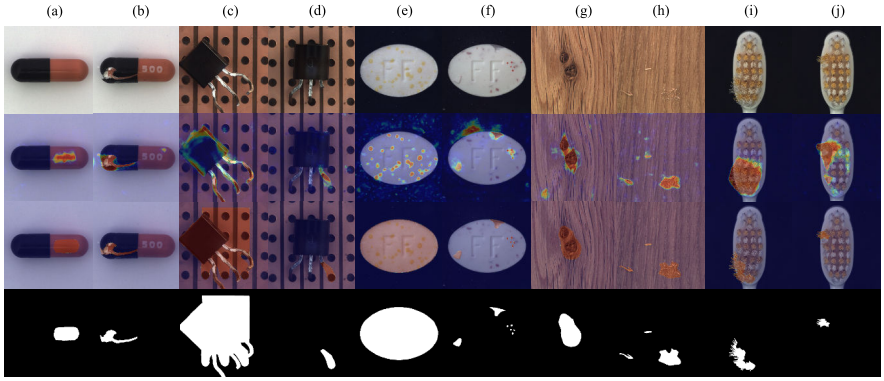


Fig. 7: DSR localization examples on the MVTec dataset. From top to bottom, the input image, the overlay of the DSR segmentation output, the overlay of the ground truth mask and the ground truth mask are shown.

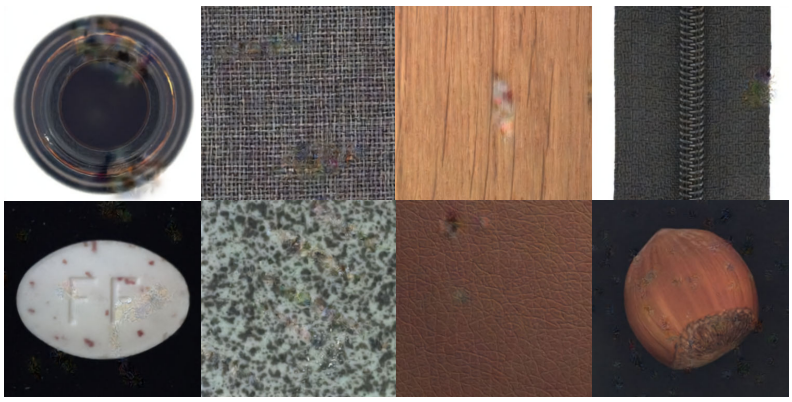


Fig. 8: Reconstructed anomalous training examples generated by sampling the quantized feature space.