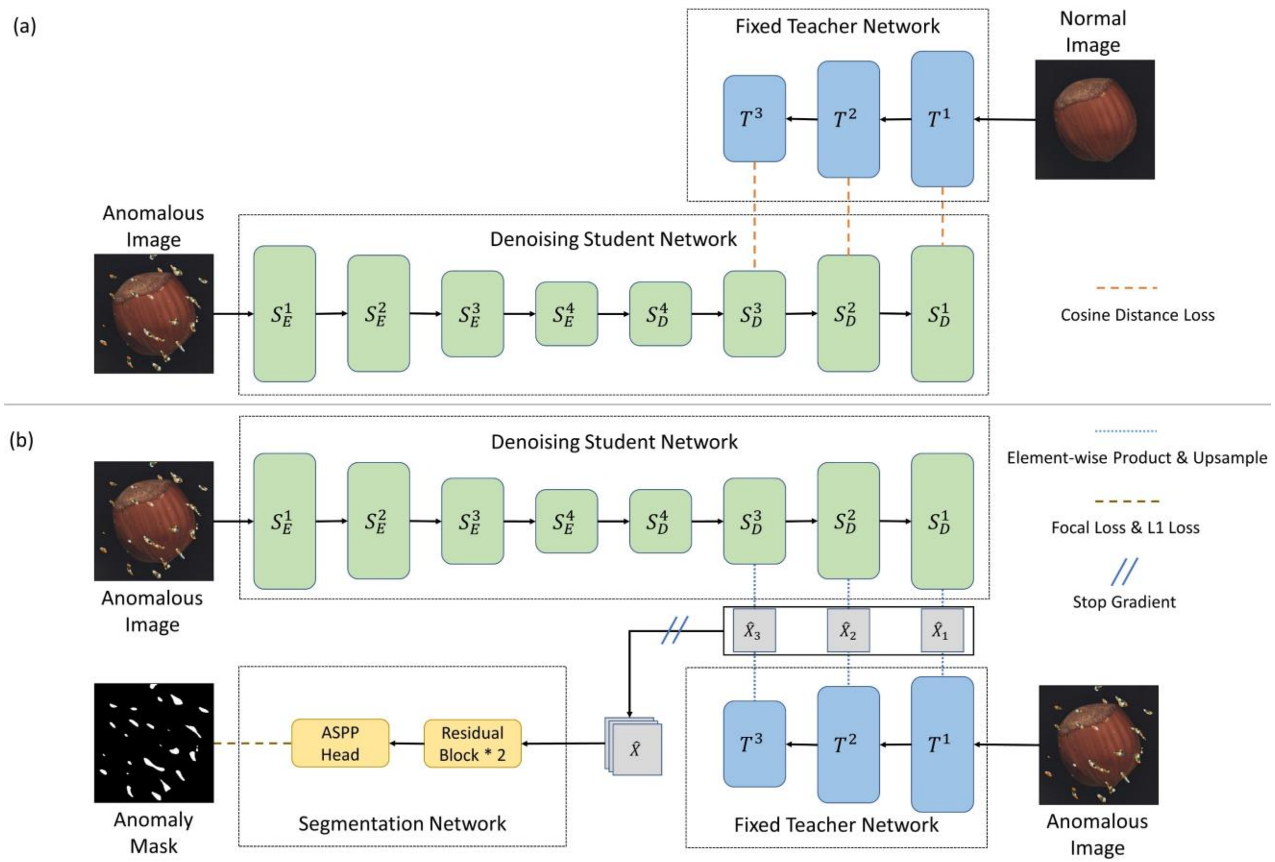
TUE-AM-373

DeSTSeg: Segmentation Guided Denoising Student-Teacher for Anomaly Detection

Xuan Zhang, Shiyu Li, Xi Li, Ping Huang, Jiulong Shan, Ting Chen

Presenter: **Xuan Zhang**
Department of Computer Science and Technology,
Tsinghua University, Beijing, China
x-zhang18@mails.tsinghua.edu.cn

- We propose a model called **DeSTSeg** for visual anomaly detection.



Highlights

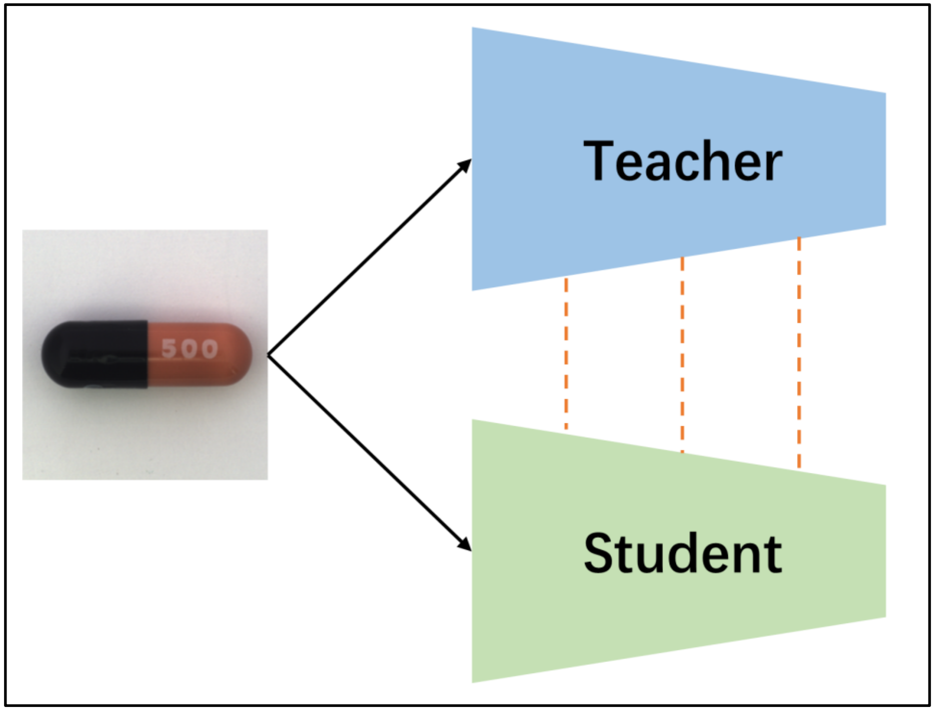
- Denoising student encoder-decoder
  - Explicitly generate different feature representations from the teacher with anomalous inputs
- Segmentation network
  - Adaptively fuse the multi-level feature similarities
- High performances on benchmark dataset

Problem Formulation

- Characteristic of anomalies
  - Limited anomaly samples
  - Long-tailed distribution

- Formulation of vision anomaly detection
  - Training: use normal data only
  - Inference: localize anomaly pixels

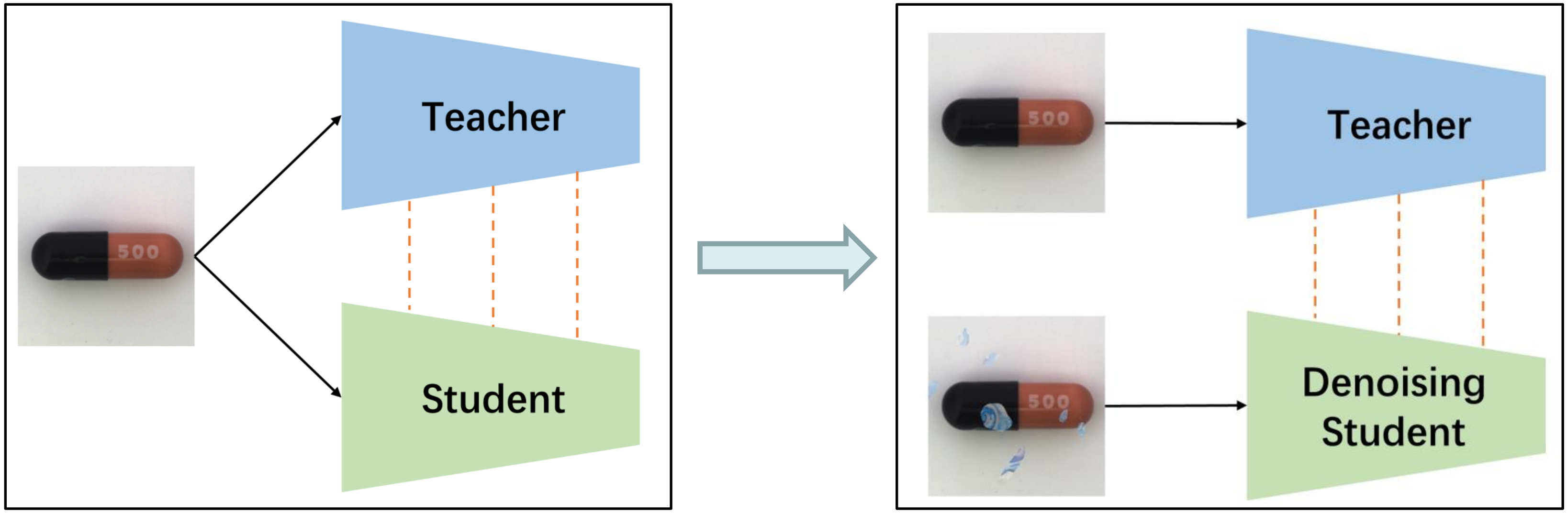
Baseline: STPM[1]



Cosine Distance Loss

Improved Student

- Apply constraints on anomalous data



Synthetic Anomaly Generation[2]


Origin Image

+


Anomaly Mask

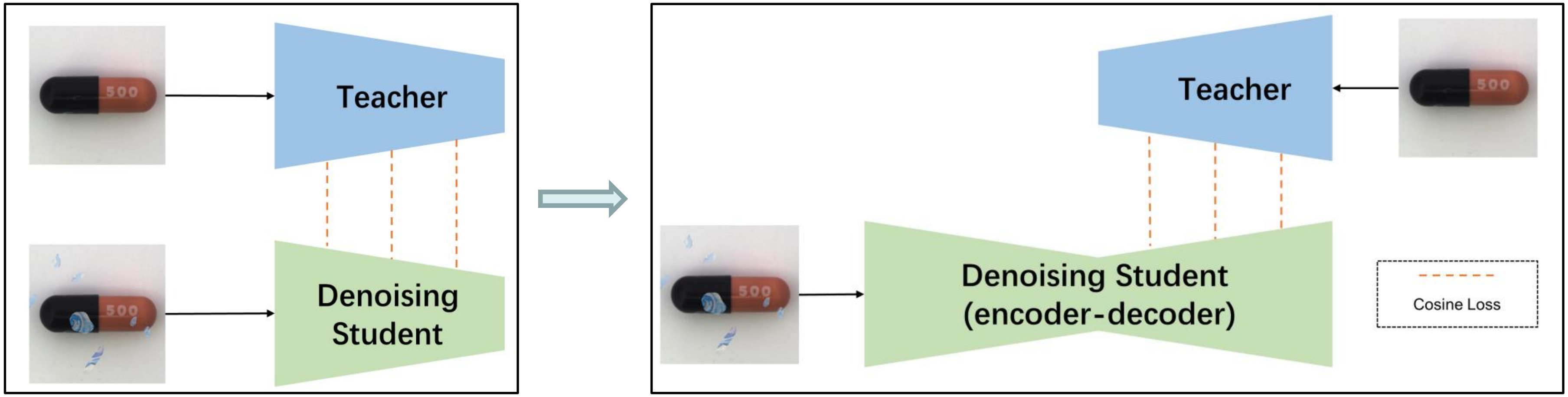

Anomaly Source Image

=


Generated Anomaly Image

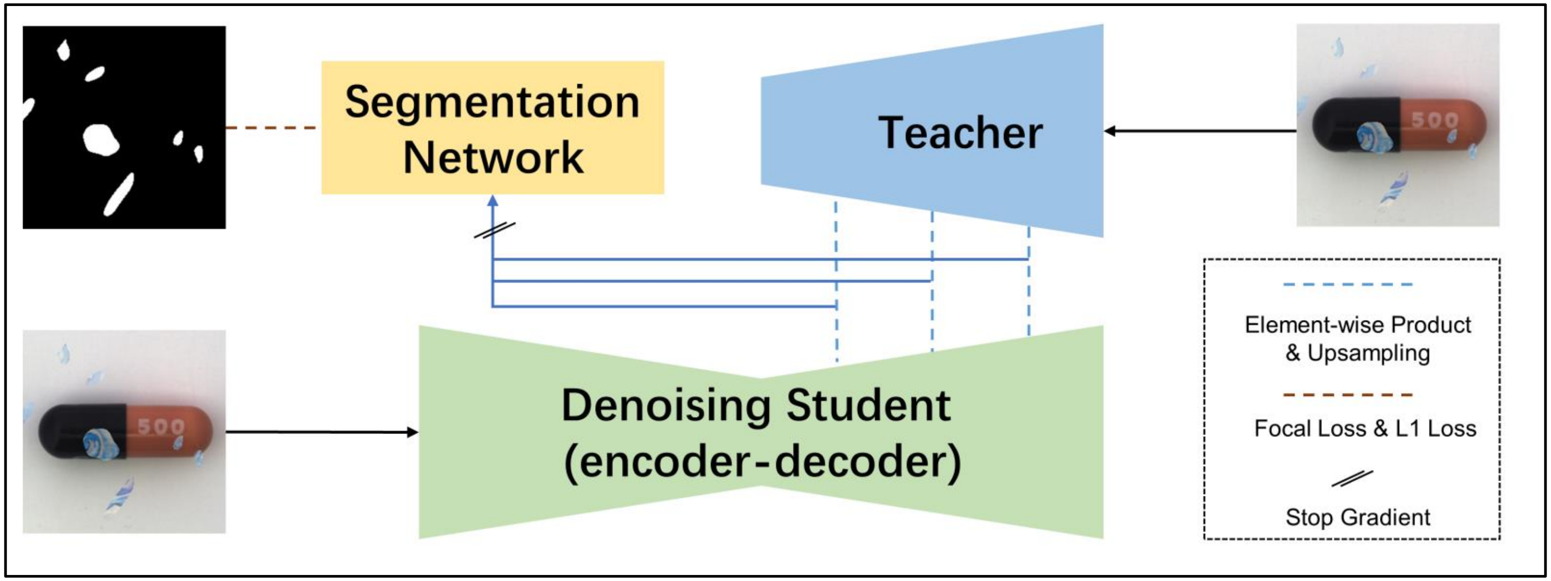
Improved Student

- Use encoder-decoder architecture to better support the feature denoising task



Segmentation Network

- Fuse the multi-level features adaptively



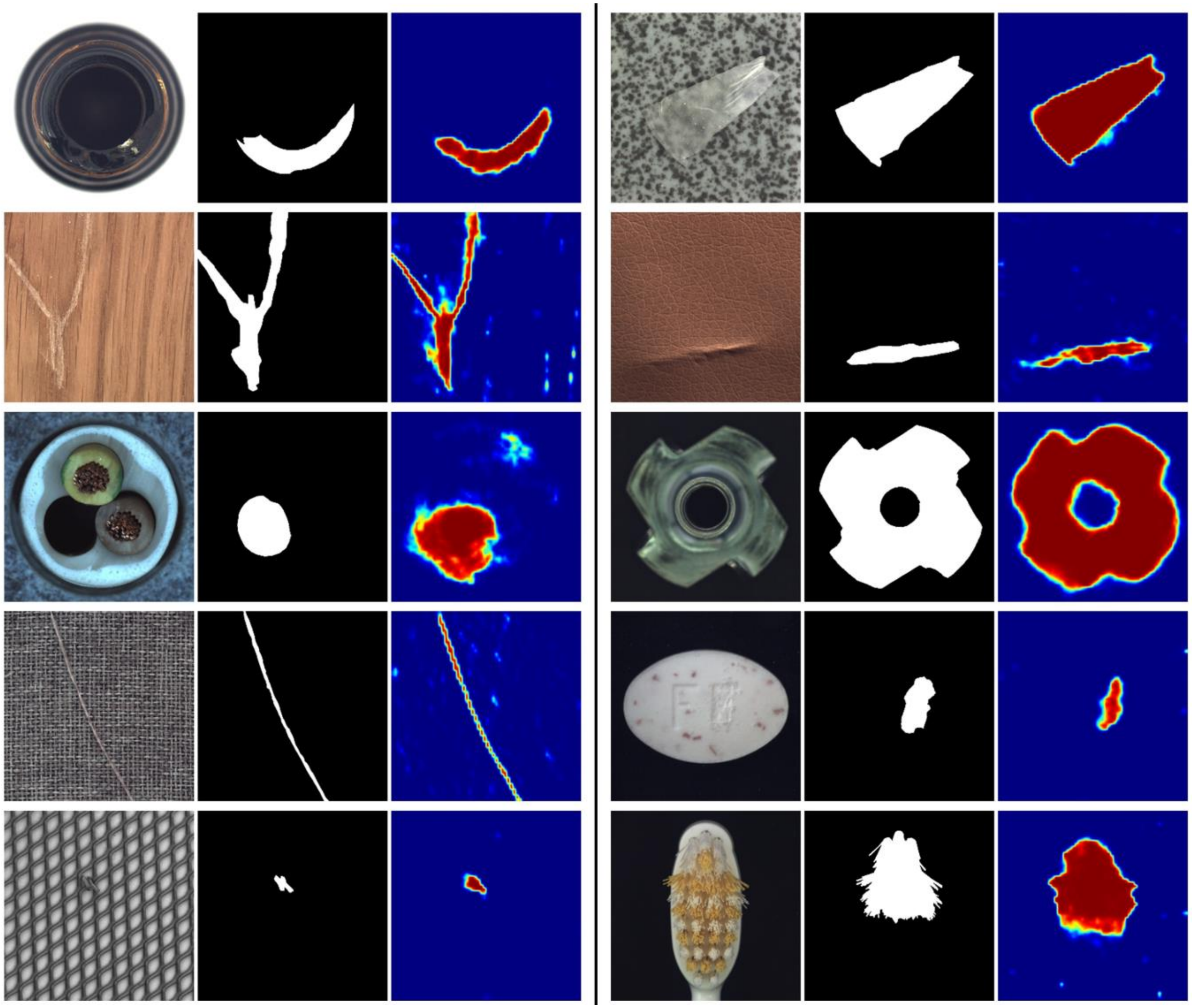
Experimental Setup

- Dataset
  - MVTec AD Dataset[3]
  - 10 object + 5 texture
- Metrics
  - Image-level AD: AUC
  - Pixel-level AD: AUC, **AP**
  - Instance-level AD: we propose a metric, **IAP** (instance average precision)

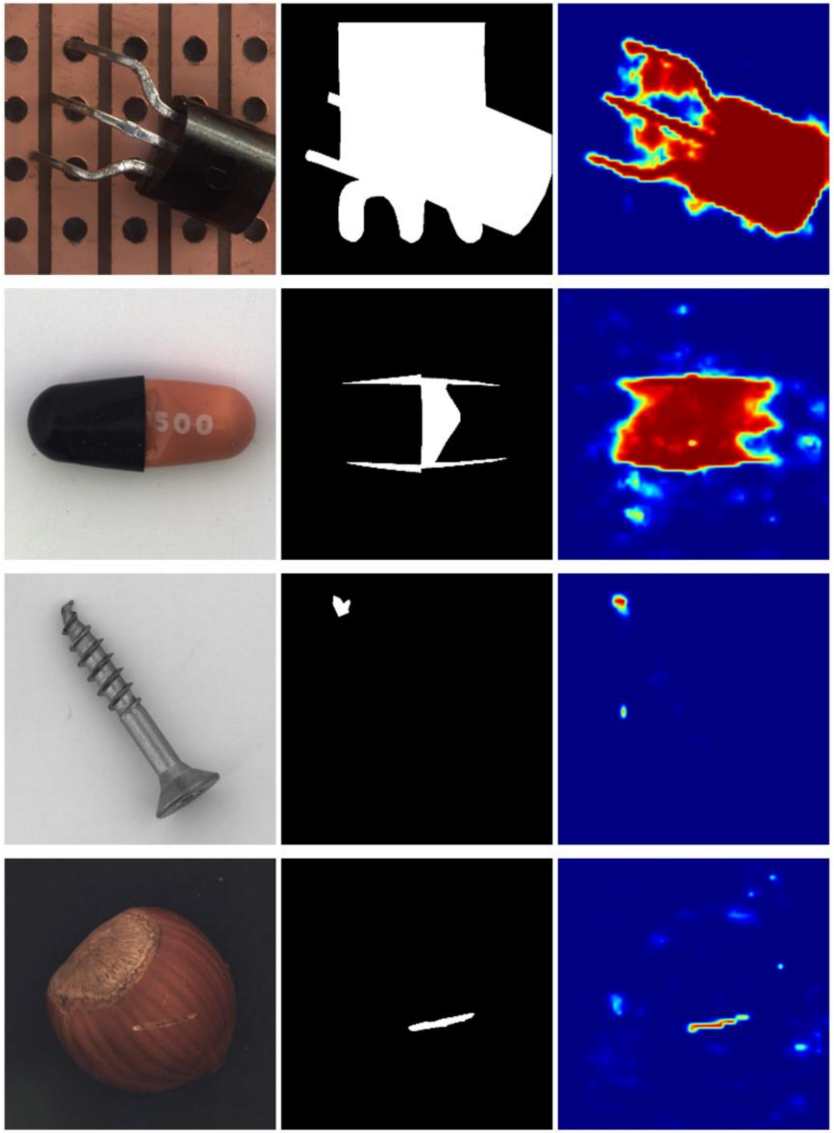
Quantitative Results

Task and metric	Ours	SOTA
Image-level AUC	**98.6**±0.4	98.5 (PatchCore[4])
Pixel-level AUC	97.9±0.3	**98.4** (PatchCore[4])
Pixel-level AP	**75.8**±0.8	70.2 (DSR[5])
Instance-level IAP	**76.4**±1.0	71.5 (DRAEM[2])
Instance-level IAP@90	**57.8**±1.8	51.7 (DRAEM[2])
Instance-level PRO	94.4±0.4	**94.5** (PatchCore[4])

Visualization Examples



Failure Cases



Ablation Studies

- **den**: use denoising student network (checked) or use origin student network (unchecked)
- **ed**: use encoder-decoder student (checked) or use decoder student (unchecked)
- **seg**: use segmentation network (checked) or use the product of cosine distances (unchecked) for anomaly localization

Exp.	den	ed	seg	img (AUC)	pix (AP)	ins (IAP)
1				94.8	52.9	55.8
2	✓			93.4	49.6	53.9
3		✓		95.4	53.3	57.7
4			✓	97.3	70.1	71.8
5	✓	✓		94.5	54.0	58.5
6	✓		✓	97.3	70.9	72.3
7		✓	✓	97.7	69.7	71.2
8	✓	✓	✓	**98.6**	**75.8**	**76.4**

Summary

- We design DeSTSeg for anomaly detection and localization.
  - Denoising student-teacher network
  - Encoder-decoder student network architecture
  - Segmentation network

- We prove the effectiveness of our method.
  - DeSTSeg achieves high performance on MVTec AD dataset
  - Detailed ablations are discussed

References

1. Wang G, Han S, Ding E, et al. Student-teacher feature pyramid matching for anomaly detection [C]//Proceedings of the British Machine Vision Conference (BMVC). 2021: 1-14.
2. Zavrtanik V, Kristan M, Skočaj D. Draem-a discriminatively trained reconstruction embedding for surface anomaly detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2021: 8330-8339.
3. Bergmann P, Fauser M, Sattlegger D, et al. Mvtec ad–a comprehensive real-world dataset for unsupervised anomaly detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019: 9592-9600.
4. Roth K, Pemula L, Zepeda J, et al. Towards total recall in industrial anomaly detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022: 14318-14328.
5. Zavrtanik V, Kristan M, Skočaj D. Dsr–a dual subspace re-projection network for surface anomaly detection[C]//Proceedings of the European Conference on Computer Vision (ECCV). Springer, 2022: 539-554.

Thanks

Paper

Code