Unsupervised Surface Anomaly Detection with Diffusion Probabilistic Model

Xinyi Zhang[1,*] Naiqi Li[1,*,†] Jiawei Li[2] Tao Dai[3,†] Yong Jiang[1] Shu-Tao Xia[1,4]

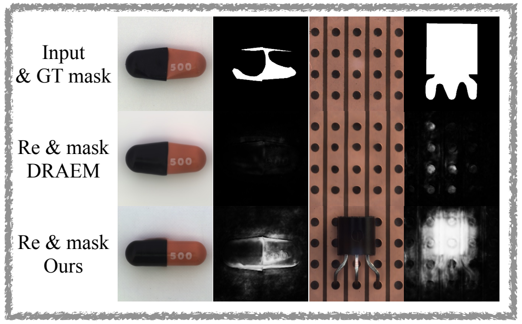
[1] Tsinghua Shenzhen International Graduate School, Tsinghua University
[2] Huawei Manufacturing [3] Shenzhen University
[4] Research Center of Artificial Intelligence, Peng Cheng Laboratory

ICCV23 PARIS

Motivation

- Reconstruction-based models rely on the assumption that anomaly regions are more difficult to reconstruct. However, many of them can also well reconstruct the anomalies, falling into **"direct copy"**.
- An anomalous instance may correspond to multiple normal patterns, but most current methods ignore this fact.

- We propose ***DiffAD***, a method based on the latent diffusion model, inspired by its ability to generate high-quality and diverse images.
- We further propose the noisy condition embedding and interpolated channels.



Method

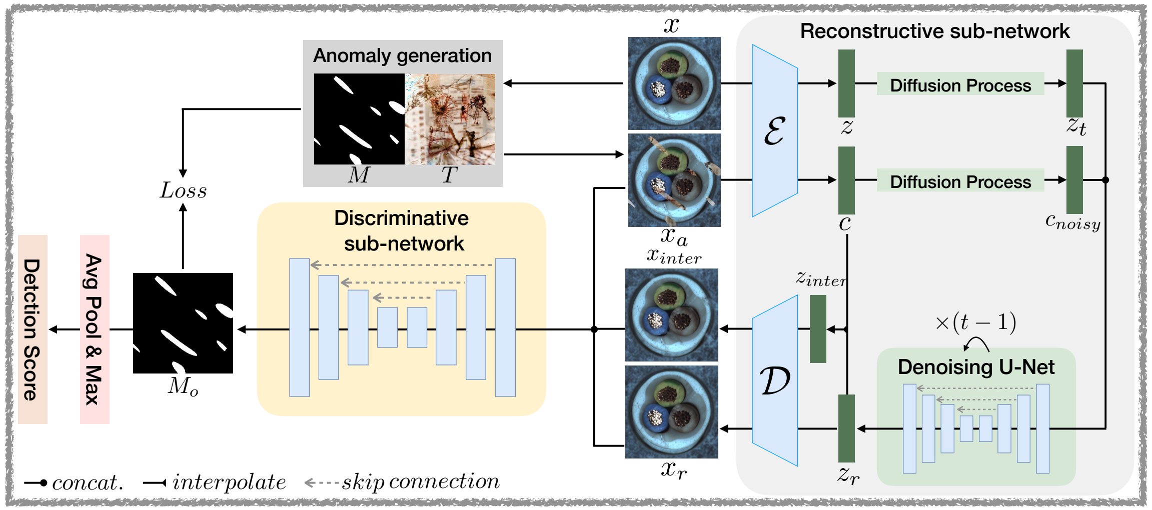
- Our method is composed of a reconstructive sub-network and a discriminative sub-network.
- **The reconstructive sub-network**: a latent diffusion model accompanied by the noisy condition embedding, learning the distribution of normal samples. The noisy condition embedding helps to instruct sample generation while avoiding the diffusion model excessively relying on the condition.

$$\mathbf{c}_{noisy} = \sqrt{\bar{\alpha}_t}\mathbf{c} + \sqrt{1-\bar{\alpha}_t}\epsilon$$

$$L_{LDM} = \mathbb{E}_{z,\epsilon\sim\mathcal{N}(0,1),t}\left[\|\epsilon - \epsilon_\theta(z_t, t, \mathbf{c}_{noisy})\|_2^2\right].$$

- **The discriminative sub-network**: a U-net-based segmentation network, taking the concatenation of the reconstruction, anomalous input and the interpolated intermediate states as input. Since some subtle differences in the normal pixels between the reconstructed and original images are inevitable, the interpolated channels help model recognize diversity during reconstruction and distinguish real anomalies.

$$\mathbf{z}_{inter} = \lambda\cdot\mathbf{c} + (1-\lambda)\cdot\mathbf{z}_r$$



Main Results

- Reducing **"direct copy"**: we measure the degree of direct copy with PSNR on the GT anomaly-masked regions (A higher PSNR indicates a severer direct copy), and use FID score to evaluate the reconstruction quality.

Method	PSNR(dB)	FID
DiffAD	36.73	69.2
DRAEM	38.49	121.7

- Results for anomaly detection and localization

Method	Det.(AUROC)	Loc.(AUROC / AP)
DiffAD	98.7	**98.3 / 74.6**
DRAEM	98.0	97.3 / 68.4
RIAD	91.7	94.2 / 48.2
PaDim	95.3	97.4 / 55.0
PatchCore	**99.0**	98.1 / 63.1