Adaptation with pixel-level labels. Inspired by [22], we firstly calculate a pseudo-Huber loss, $\boldsymbol{\phi}(u)$, using the feature difference map, $\boldsymbol{d}(i,u)$.

$$\boldsymbol{\phi}(u) = ((\frac{1}{C}\sum_{i}^{C}|\boldsymbol{d}(i,u)|)^2 + 1)^{\frac{1}{2}} - 1. \tag{6}$$

The pseudo-Huber loss, $\boldsymbol{\phi}(u)$, is designed as a difference map, which is easy to train and extend. Then the reconstruction loss function with pixel-level labels is denoted as $\mathcal{L}_{px}$ and could be described as a "push-pull loss" as,

$$\mathcal{L}_{px} = \frac{1}{HW}\sum_{u}^{HW}(1-\boldsymbol{y}(u))\boldsymbol{\phi}(u) - \alpha\log(1-\exp(-\frac{1}{HW}\sum_{u}^{HW}\boldsymbol{y}(u)\boldsymbol{\phi}(u))), \tag{7}$$

where the first term pulls the reconstructed normal features to the extracted features, and the second term pushes the reconstructed anomalous features away from the original features, $\boldsymbol{y}(u)$ is the pixel-level label (0 for normal sample and 1 for anomaly) and α is a weight term.

Adaptation with image-level labels. Since anomaly samples could contain both anomalous and normal regions, simply treating all regions of anomaly samples as anomalous regions confuses the model. Considering that larger values in $\boldsymbol{\phi}(u)$ are more likely to be anomalous regions, we firstly collect k maximum values of $\boldsymbol{\phi}(u)$, then calculate their mean as the anomaly score of the image.

$$q = \frac{1}{k}\sum \mathtt{top\_k}(\boldsymbol{\phi}). \tag{8}$$

Then the image-level loss, $\mathcal{L}_{img}$, could be calculated as,

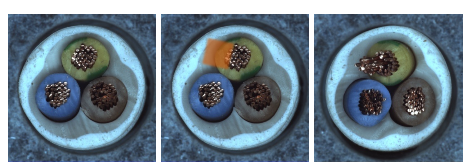
$$\mathcal{L}_{img} = (1-y)q - \alpha y\log(1-\exp(-q)), \tag{9}$$

where y is the image-level label (0 for normal sample and 1 for anomaly) and α is a weight term. In $\mathcal{L}_{img}$, the first term pulls the reconstructed features of normal samples towards the extracted features, while the second term pushes the reconstructed features of anomalies away from the extracted features.

4 Experiment

4.1 Dataset

MVTec-AD [4] is a multi-category, multi-defect, industrial anomaly detection dataset with 15 categories. The ground-truth includes both image labels and anomaly segmentation. In *normal-sample-only case*, we follow the original setting to use normal



(a) Normal sample (b) Confetti noise (c) Real anomaly

Fig. 3: **Synthetic anomalies** by adding confetti noise on normal samples.