

表1: MVTec-ADにおけるピクセルレベルAUROCメトリクスに基づく異常局在化結果 [4]。

	Texture					Object										Mean
	Carp.	Grid	Leat.	Tile	Wood	Bot.	Cable	Caps.	Haze	Meta.	Pill	Screw	Toot.	Tran.	Zip.	
AnoGAN [31]	54	58	64	50	62	86	78	84	87	76	87	80	90	80	78	74
SCADN [36]	64.9	79.6	76.3	67.7	67.2	69.6	81.4	68.7	88.4	75.4	74.7	87.6	90.1	68.9	67.0	75.2
SSIM-AE [6]	87	94	78	59	73	93	82	94	97	89	91	96	92	90	88	86
VEVAE [21]	78	73	95	80	77	87	90	74	98	94	83	97	94	93	78	86
SMAI [20]	88	97	86	62	80	86	92	93	97	92	92	96	96	85	90	89
VAE-Grad [12]	74	96	93	65	84	92	91	92	98	91	93	95	99	92	87	89
P-Net [40]	57	98	89	97	98	99	70	84	97	79	91	100	99	82	90	89
KDAD [30]	95.6	91.8	98.1	82.8	84.8	96.3	82.4	95.9	94.6	86.4	89.6	96.0	96.1	76.5	93.9	90.7
Loc-Glo [34]	96	78	90	80	81	93	94	90	84	91	93	96	96	100	99	91
FCDD [22]	96	91	98	91	88	97	90	93	95	94	81	86	94	88	92	92
PSVDD [37]	92.6	96.2	97.4	91.4	90.8	98.1	96.8	95.8	97.5	98.0	95.1	95.7	98.1	97.0	95.1	95.7
SPADE [10]	97.5	93.7	97.6	87.4	88.5	98.4	97.2	99.0	99.1	98.1	96.5	98.9	97.9	94.1	96.5	96.0
ADTR(ours)	98.7	95.0	98.1	93.8	91.2	98.0	96.8	99.1	98.6	97.0	98.3	99.3	98.5	97.9	97.2	97.2
ADTR+(ours)	98.8	94.2	98.6	95.9	93.0	98.0	97.0	99.1	98.8	96.8	98.7	99.3	99.2	97.8	97.6	97.5

MVTec-ADにおける定性的結果は図4に示されています。当手法は、高い局所化精度で多様な異常を検出しています。特に、示された「Metal Nut」例では、異常が反転した正常サンプルであるにもかかわらず、テクスチャの乱れや色の変化のような明らかな視覚的異常がないにもかかわらず、当手法は「反転」異常を成功裡に検出しています。

異常局所化の定量的な結果は表1に示されています。当手法は、AnoGAN [31]、SCADN [36]、SSIM-AE [6]、VEVAE [21]、SMAI [20]、VAE-Grad [12]、P-Net [40]、KDAD [30]、Loc-Glo [34]、FCDD [22]、PSVDD [37]、SPADE [10] と比較されています。純粋な正常サンプルのみの場合、ADTR は最良のベースラインである SPADE [10] を 1.2% 上回っています。単純な合成異常のみの場合、ADTR+ の性能はさらに 0.3% 向上しています。

異常検出の定量的な結果は表2に示されています。当手法は、GANomaly [2]、SCADN [36]、ARNet [14]、SPADE [10]、KDAD [30]、PSVDD [37]、TS [5] と比較されています。ADTRは、正常サンプルのみを使用した場合、すべてのベースライン手法を大幅に上回っています（ $\geq 3.9\%$ ）。ADTR+の性能は、単純な合成異常データを使用することで0.5%向上します。

4.3 CIFAR-10における異常検出

異常検出能力をさらに検証するため、当モデルをCIFAR-10 [18]の無監督単一クラス分類タスクで評価しました。

設定。セクション4.2の設定と同じですが、以下の点のみ異なります。まず、画像と特徴マップのサイズはそれぞれ 32×32 と 8×8 です。次に、異常データが利用可能な場合、モデルは式(9)のイメージレベル損失 L_{img} で訓練され、 α と k はそれぞれ0.003と20に設定されます。

CIFAR-10における定量的な結果は表3に示されています。競合モデルには、VAE [3]、KDE [7]、AnoGAN [31]、LSA [1]、DSVDD [28]、OCGAN [26]、GradCon [19]、Loc-Glo [34]、TS [5]、GT [15]、KDAD [30]が含まれます。ADTRは、正常サンプルのみの場合のトレーニングにおいて、KDAD [30]を大幅な差（7.5%）で上回っています。

Table 1: **Anomaly localization results under pixel-level AUROC metric on MVTec-AD [4].**

	Texture					Object										Mean
	Carp.	Grid	Leat.	Tile	Wood	Bot.	Cable	Caps.	Haze.	Meta.	Pill	Screw	Toot.	Tran.	Zip.	
AnoGAN [31]	54	58	64	50	62	86	78	84	87	76	87	80	90	80	78	74
SCADN [36]	64.9	79.6	76.3	67.7	67.2	69.6	81.4	68.7	88.4	75.4	74.7	87.6	90.1	68.9	67.0	75.2
SSIM-AE [6]	87	94	78	59	73	93	82	94	97	89	91	96	92	90	88	86
VEVAE [21]	78	73	95	80	77	87	90	74	98	94	83	97	94	93	78	86
SMAI [20]	88	97	86	62	80	86	92	93	97	92	92	96	96	85	90	89
VAE-Grad [12]	74	96	93	65	84	92	91	92	98	91	93	95	99	92	87	89
P-Net [40]	57	98	89	97	98	99	70	84	97	79	91	100	99	82	90	89
KDAD [30]	95.6	91.8	98.1	82.8	84.8	96.3	82.4	95.9	94.6	86.4	89.6	96.0	96.1	76.5	93.9	90.7
Loc-Glo [34]	96	78	90	80	81	93	94	90	84	91	93	96	96	100	99	91
FCDD [22]	96	91	98	91	88	97	90	93	95	94	81	86	94	88	92	92
PSVDD [37]	92.6	96.2	97.4	91.4	90.8	98.1	96.8	95.8	97.5	98.0	95.1	95.7	98.1	97.0	95.1	95.7
SPADE [10]	97.5	93.7	97.6	87.4	88.5	98.4	97.2	99.0	99.1	98.1	96.5	98.9	97.9	94.1	96.5	96.0
ADTR(ours)	98.7	95.0	98.1	93.8	91.2	98.0	96.8	99.1	98.6	97.0	98.3	99.3	98.5	97.9	97.2	97.2
ADTR+(ours)	98.8	94.2	98.6	95.9	93.0	98.0	97.0	99.1	98.8	96.8	98.7	99.3	99.2	97.8	97.6	97.5

Qualitative results on MVTec-AD are shown in Fig. 4. Our approach successfully detects different kinds of anomalies with high localization accuracy. Especially, for the shown “Metal Nut” example, where the anomaly is a flipped normal sample, our approach detects the “flip” anomaly successfully though there are no obvious vision anomalies like texture disorder nor color change.

Quantitative results of anomaly localization are given in Tab. 1. Our approach is compared with AnoGAN [31], SCADN [36], SSIM-AE [6], VEVAE [21], SMAI [20], VAE-Grad [12], P-Net [40], KDAD [30], Loc-Glo [34], FCDD [22], PSVDD [37], SPADE [10]. With pure normal samples, ADTR stably outperforms the best baseline, SPADE [10], by 1.2%. With merely simple synthetic anomalies, the performance of ADTR+ is further improved by 0.3%.

Quantitative results of anomaly detection are shown in Tab. 2. Our approach is compared with GANomaly [2], SCADN [36], ARNet [14], SPADE [10], KDAD [30], PSVDD [37], TS [5]. ADTR considerably exceeds all baseline methods ($\geq 3.9\%$) with only normal samples. The performance of ADTR+ is improved by 0.5% with simple synthetic anomalies.

4.3 Anomaly Detection on CIFAR-10

To further validate the anomaly detection ability, we evaluate our model in the unsupervised one-class classification task of CIFAR-10 [18].

Setup. The setup is the same as that in Sec. 4.2 except the followings. First, the sizes of the image and feature map are 32×32 and 8×8 , respectively. Second, in anomaly-available case, the model is trained with the image-level loss, \mathcal{L}_{img} , in Eq. (9), where α and k are selected as 0.003 and 20, respectively.

Quantitative results on CIFAR-10 are shown in Tab. 3. The competitors include: VAE [3], KDE [7], AnoGAN [31], LSA [1], DSVDD [28], OCGAN [26], GradCon [19], Loc-Glo [34], TS [5], GT [15], KDAD [30]. ADTR surpasses KDAD [30] by a great margin (7.5%) when training in normal-sample-only case. In