Figure 3: Generated images (inlier: CIFAR-10 Cat)

Appendices

A   More Experimental Results or Details

Figure 3 shows the qualitative examples of the synthesized images produced by our model trained on treating CIFAR-10 Cat as inlier class while all the others are outliers.

Table 2 presents the detailed AUROC (%) value for our experiments on CIFAR-10 dataset. Each row indicates the results of taking a specific class as inlier and all the others as outliers.

Table 3 presents the detailed AUROC (%) value for our experiments on CIFAR-100 dataset. Noting that for each task we take one superclass as inlier and all the others as outliers.

Table 4 presents the detailed AUROC (%) value for our experiments on ImageNet-30 dataset. Noting that for each task we take one class as inlier and all the others as outliers.

Table 5 is the detailed AUROC (%) value of our ablation study on model designs. $+\mathcal{G}$ means that model integrates CSI [32] method with our generator $\mathcal{G}$. And $+\mathcal{E}$ means that model integrates CSI [32] method, our generator $\mathcal{G}$, and our encoder $\mathcal{E}$ (thus reaching our full model).

B   Hyperparameters to Balance Objectives

We set $\mathcal{L}_{\text{CSI}}, \mathcal{L}_{\text{D}}, \mathcal{L}_{\text{CSI-S}}, \mathcal{L}_{\text{LD}}, \mathcal{L}_{\text{REC}}, \mathcal{L}_{\text{LREC}}, \mathcal{L}_{\text{JOINT}}$ to have the same weight during training. Please note that, we have attempted to conduct experiments for adjusting those weights and find that having all objectives equally weighted already provides favorable results.

Table 2: AUROC (%) of semi-supervised anomaly detection on CIFAR-10 dataset

Class	DeepSVDD [29]	OCGAN [26]	Geom [11]	Rot+Trans [16]	GOAD [3]	CSI [32]	Ours
0	61.7	75.7	74.7	77.5	77.2	89.9	**91.6**
1	65.9	53.1	95.7	96.9	96.7	**99.1**	**99.1**
2	50.8	64	78.1	87.3	83.3	93.1	**94.4**
3	59.1	62	72.4	80.9	77.7	86.4	**88.4**
4	60.9	72.3	87.8	92.7	87.8	93.9	**95.3**
5	65.7	62	87.8	90.2	87.8	93.2	**97.7**
6	67.7	72.3	83.4	90.9	90.0	95.1	**95.9**
7	67.3	57.5	95.5	96.5	96.1	98.7	**98.8**
8	75.9	82	93.3	95.2	93.8	**97.9**	**97.9**
9	73.1	55.4	91.3	93.3	92.0	95.5	**95.9**
Mean	64.8	65.7	86	90.1	88.2	94.3	**95.1**

Table 3: AUROC (%) of semi-supervised anomaly detection on CIFAR-100 dataset

	0	1	2	3	4	5	6	7	8	9	
CSI [32]	86.3	84.8	88.9	85.7	93.7	81.9	91.8	83.9	91.6	**95.0**	
Ours	**86.5**	**87.1**	**91.2**	**87.0**	**95.3**	**86.3**	**92.9**	**86.7**	**93.0**	**95.0**	
	10	11	12	13	14	15	16	17	18	19	Mean
CSI [32]	**94.0**	90.1	**90.3**	81.5	94.4	85.6	83.0	**97.5**	95.9	95.2	89.6
Ours	93.0	**91.4**	90.0	**84.1**	**95.3**	**87.8**	**84.8**	97.0	**97.0**	**96.2**	**90.9**

Table 4: AUROC (%) of semi-supervised anomaly detection on ImageNet-30 dataset

	0	1	2	3	4	5	6	7	8	9	
CSI [32]	**85.9**	99.0	**99.8**	90.5	95.8	**99.2**	96.6	83.5	92.2	84.3	
Ours	85.3	**99.6**	99.5	**95.8**	**96.4**	98.5	**96.8**	**91.7**	**96.9**	**86.9**	
	10	11	12	13	14	15	16	17	18	19	
CSI [32]	**99.0**	94.5	97.1	87.7	96.4	84.7	**99.7**	75.6	95.2	**73.8**	
Ours	98.8	**95.5**	**98.4**	**92.1**	**97.3**	**88.7**	**99.7**	**80.9**	**96.1**	69.5	
	20	21	22	23	24	25	26	27	28	29	Mean
CSI [32]	94.7	95.2	**99.2**	98.5	**82.5**	**89.7**	82.1	97.2	82.1	97.6	91.6
Ours	**95.4**	**97.2**	98.6	97.1	78.9	88.3	**86.1**	**98.1**	**86.3**	**98.8**	**93.0**

Table 5: Ablation study on our model designs, based on the AUROC (%) of semi-supervised anomaly detection on CIFAR-10 dataset

Class	CSI [32]	$+\mathcal{G}$	$+\mathcal{E}$ (Proposed model)
0	89.9	90.9	91.6
1	99.1	99.3	99.1
2	93.1	94.1	94.4
3	86.4	87.6	88.4
4	93.9	95.3	95.3
5	93.2	93.2	93.7
6	95.1	95.3	95.9
7	98.7	99.0	98.8
8	97.9	97.6	97.9
9	95.5	95.8	95.9
Mean	94.3	94.8	95.1