Supplementary Materials of "Toward Generalist Anomaly Detection via In-context Residual Learning with Few-shot Sample Prompts"

Jiawen Zhu and Guansong Pang *

School of Computing and Information Systems, Singapore Management University

A. Dataset Details

A.1. Data Statistics of Training and Testing

We conduct extensive experiments on nine real-world Anomaly Detection (AD) datasets, including five industrial defect inspection dataset (MVTec AD [1], VisA [14], ELPV [4], SDD [12], AITEX [11]), two medical image datasets (BrainMRI [10], HeadCT [10]), and two semantic anomaly detection datasets: MNIST [8] and CIFAR-10 [7] under both one-vs-all and multi-class protocols [3].

To assess the Genealist Anomaly Dtection (GAD) performance, the full dataset of MVTec AD, including both training set and test set, is used as the auxiliary training data, on which AD models are trained, and they are subsequently evaluated on the test set of the other eight datasets without any further training. We train the model on the full dataset of VisA when evaluating the performance on MVTec AD. The few-shot normal prompts for the target data are randomly sampled from the training set of target datasets and remain the same for all models for fair comparison. Table 1 provides the data statistics of MVTec AD and VisA, while Table 2 shows the test set statistics of the rest datasets.

Dataset	Subset	Type	Original Training	Original Test	
			Normal	Normal	Anomaly
MVTec AD	Carpet	Texture	280	28	89
	Grid	Texture	264	21	57
	Leather	Texture	245	32	92
	Tile	Texture	230	33	83
	Wood	Texture	247	19	60
	Bottle	Object	209	20	63
	Capsule	Object	219	23	109
	Pill	Object	267	26	141
	Transistor	Object	213	60	40
	Zipper	Object	240	32	119
	Cable	Object	224	58	92
	Hazelnut	Object	391	40	70
	Metal_nut	Object	220	22	93
	Screw	Object	320	41	119
	Toothbrush	Object	60	12	30
VisA	candle	Object	900	100	100
	capsules	Object	542	60	100
	cashew	Object	450	50	100
	chewinggum	Object	453	50	100
	fryum	Object	450	50	100
	macaroni1	Object	900	100	100
	macaroni2	Object	900	100	100
	pcb1	Object	904	100	100
	pcb2	Object	901	100	100
	pcb3	Object	905	101	100
	pcb4	Object	904	101	100
	pipe_fryum	Object	450	50	100

Table 1. Data statistics of MVTec AD and VisA. When training GAD models with MVTec AD or VisA datasets, we utilize their complete datasets, including both training and test data. In contrast, for testing GAD models, only the test sets of MVTec AD or VisA are employed for inference.

A.2. Industrial Defect Inspection Datasets

MVTec AD [1] is a widely-used dataset that enables researchers to benchmark the performance of anomaly detection methods in the context of industrial inspection applications. The dataset includes over 5,000 images that are divided into 15 object and texture categories. Each category contains a training set of anomaly-free images, as well as a test set that includes images with both defects and defect-free images.

VisA [14] consists of 10,821 high-resolution color images (9,621 normal and 1,200 anomalous samples) covering 12 objects in 3 domains, making it the largest industrial anomaly detection dataset to date. Both image and pixel-level labels are provided. The anomalous images contain various flaws, including surface defects such as scratches, dents, color spots or crack, and structural defects like misplacement or missing parts.

ELPV [4] is a collection of 2,624 high-resolution grayscale images of solar cells extracted from photovoltaic modules. These images were extracted from 44 different solar modules, and include both intrinsic and extrinsic defects known to reduce the power efficiency of solar modules. In our study, we only use its test set for evaluation.

SDD [12] is a collection of images captured in a controlled industrial environment, using defective production items as the subject. The dataset includes 52 images with visible defects and 347 product images without any defects. In our

*Corresponding author: G. Pang (gspang@smu.edu.sg)

Dataset	Subset	Type	Test set	
			Normal	Anomaly
MNIST	0	Semantical	980	9020
	1	Semantical	1135	8865
	2	Semantical	1,032	8,968
	3	Semantical	1,010	8,990
	4	Semantical	982	9019
	5	Semantical	892	9108
	6	Semantical	958	9042
	7	Semantical	1028	8972
	8	Semantical	974	9026
	9	Semantical	1009	8991
	even_number	Semantical	4926	5074
CIFAR-10	airplane	Semantical	1000	9000
	automobile	Semantical	1000	9000
	bird	Semantical	1000	9000
	cat	Semantical	1000	9000
	deer	Semantical	1000	9000
	dog	Semantical	1000	9000
	frog	Semantical	1000	9000
	horse	Semantical	1000	9000
	ship	Semantical	1000	9000
	truck	Semantical	1000	9000
	animal	Semantical	6000	4000
ELPV	-	Texture	377	715
SDD	-	Texture	286	54
AITEX	-	Texture	564	183
BrainMRI	-	Medical	25	155
HeadCT	-	Medical	25	100

Table 2. Data statistics of seven AD datasets for inference. These datasets are exclusively used for inference purposes, hence only the details of the test sets are provided.

study, we only use its test set for evaluation.

AITEX [11] is a textile fabric database that comprises 245 images of 7 different fabrics, including 140 defect-free images (20 for each type of fabric) and 105 images with various types of defects. We only use its test set for evaluation.

A.3. Medical Anomaly Detection Datasets

BrainMRI [10] is a dataset for brain tumor detection obtained from magnetic resonance imaging (MRI) of the brain. In our study, we only use its test set for evaluation.

HeadCT [10] is a dataset consisting of 100 normal head CT slices and 100 slices with brain hemorrhage, without distinction between the types of hemorrhage. Each slice is from a different person, providing a diverse set of images for researchers to develop and test algorithms for hemorrhage detection and classification in medical imaging applications. In our study, we only use its test set for evaluation.

A.4. Semantic Anomaly Detection Datasets

MNIST [8] encompasses 70,000 grayscale images of handwritten digits. It serves as a semantic AD dataset in our work, where we utilize its original test set to construct test sets of one-vs-all and multi-class settings. Under the one-vs-all protocol, one of the ten classes is used as normal, with the other classes treated as abnormal; while under the multi-class protocol, images of even-number classes are treated as normal, with the images of the other classes are considered as anomalies. In this case, the category-level normal class label is set to 'even_number'.

CIFAR-10 [7] consists of 60,000 colour images in 10 classes, with 6,000 images per class. There are 50,000 training images and 10,000 test images. It serves as a semantic AD dataset in our work, where we utilize its original test set to construct test sets of one-vs-all and multi-class settings. Under the one-vs-all protocol, one of the 10 classes is used as normal, with the other classes treated as abnormal; while under the multi-class protocol, images of animal-related classes are treated as normal, with the images of the other classes are considered as anomalies. In this case, the category-level normal class label is set to 'animal'.

B. Implementation Details

B.1. Data Pre-processing

By default, for all CLIP-based models, including WinCLIP [6], CoOp [13], and `InCTRL`, we adopt the same CLIP implementation, OpenCLIP [5], and its public pretrained backbone ViT-B/16+ in our experiments. Our data preprocessing aligns with OpenCLIP across all datasets. Specifically, this involves channel-wise standardization using a predefined mean and standard deviation after scaling RGB images to the range of [0, 1], followed by bicubic resizing based on Pillow library. In addition, we resize the input resolution to 240×240 to match ViT-B/16+. This resizing is also applied to other baseline models for fair comparison, while retaining their original data preprocessing methods (if there are any).

B.2. Network Architectures

In our experiments, the parameters of visual encoder and text encoder of ViT-B/16+ are kept frozen. This model, while being similar in depth to ViT-B/16, increases the dimensions of image embeddings(768 $\rightarrow$ 896), text embedding(512 $\rightarrow$ 640) and input resolution (224×224 $\rightarrow$ 240×240). For the learnable components, to align with ViT-B/16+'s dimensions, the adapter ψ has input and output dimensions set to 896, including a 224-unit hidden layer with ReLU activation. The image-level anomaly classification learner η takes in-context image-level residual features $F_x$as input and yields a one-dimensional prediction, where η has two hidden layers with 128 and 64 units respectively. The holistic anomaly scoring model ϕ incorporates two hidden layers, projecting a 225-dimensional in-context residual map to generate a final single-dimensional anomaly score.

B.3. Details of Text prompts

The text prompts used in our work are based on the same main text and ensemble strategy to WinCLIP [6] except

Normal Examples	'*a photo of a flawless [c] for visual inspection.*' '*a cropped photo of a perfect [c].*' '*a blurry photo of the [c] without defect.*' '*a dark photo of the unblemished [c].*' '*a jpeg corrupted photo of a [c] without flaw.*'
Abnormal Examples	'*a photo of a [c] with flaw for visual inspection.*' '*a cropped photo of a [c] with damage.*' '*a blurry photo of the [c] with defect.*' '*a dark photo of the [c] with flaw.*' '*a jpeg corrupted photo of a [c] with defect.*'

Table 3. Examples of normal and abnormal text prompts used in InCTRL and WinCLIP. *[c]* represents a class label.

CIFAR-10 [14]. Table 3 provides several normal and abnormal examples of text prompts used in InCTRL (see [6] for the full list of the text prompts). The WinCLIP prompts fail to work for natural images like CIFAR-10, so the normal and abnormal text prompts of CIFAR-10 are designed as '*a photo of [c] for anomaly detection.*' and '*a photo without [c] for anomaly detection.*', respectively, which are used in both WinCLIP and InCTRL on CIFAR-10. Here, *[c]* represents a category-level label, *e.g.*, airplane.

B.4. Implementation of Comparison Methods

For the results of competing methods, we re-implemented SPADE, PaDiM, and WinCLIP, while using the official implementations of PatchCore, RegAD, and CoOp. Differing from SPADE's original $K = 50$ setup, we use $K = 2, 4, 8$ nearest neighbors to match the few-shot setting. For PaDiM, we select the wide_resnet50_2 model, pretrained on ImageNet, as the feature extractor. To ensure fair empirical comparison, we apply the same image prompts as used in InCTRL across all methods. All reported results are the average of three independent runs, each with a different random seed.

C. Detailed Empirical Results

C.1. Complexity of InCTRL vs. Other CLIP-based Methods

The number of parameters and per-image inference time for CLIP-based methods are shown in Table 4.

Our InCTRL and CoOp involve additional training on auxiliary data compared to the training-free method, *i.e.*, WinCLIP. This results in extra trainable parameters during the training phase. However, the extra time consumption leads to significant performance enhancements in InCTRL, and furthermore, the training can be taken offline, so its computation overhead is generally negligible.

Additionally, as Table 4 shows, InCTRL achieves faster inference compared to WinCLIP's multi-scale few-shot anomaly scoring approach. Although CoOp adopts a similar few-shot anomaly scoring method as WinCLIP, it gains better efficiency by avoiding the ensemble text prompt strategy in WinCLIP. Result indicates the effectiveness of the InCTRL framework in enhancing the base model's generalization ability.

Method	Number of Parameters	Inference Time (ms)
CoOp	6,400	197.3±5.6
WinCLIP	0	227.5±0.7
Ours (InCTRL)	334,916	81.7±1.4

Table 4. Number of Parameters and Per-image Inference time.

	Dataset	Brain Tumor MRI		LAG	
		AUR	PR	AUR	PR
Traditional Medical AD	**FPI**	0.831	0.789	0.543	0.556
	PII	0.843	0.805	0.610	0.607
	F-AnoGAN	0.825	0.743	0.842	0.775
	AEU	0.940	0.890	0.813	0.789
	AEU+DDAD	0.942	0.919	**0.860**	0.840
Generalist AD	**WinCLIP**	0.779	0.878	0.571	0.731
	Ours (InCTRL)	**0.951**	**0.968**	0.832	**0.880**

Table 5. Comparison with Traditional Medical AD Methods.

C.2. Comparison with Traditional Medical Anomaly Detection Methods

Even though our comparison is focused on detectors of similar generalist detection capabilities, we compare five recent AD methods specifically designed for medical images on two other medical datasets (Brain Tumor MRI[1] and LAG [9]) in Table 5. The results of traditional medical anomaly detection methods are from Cai *et al*. [2] based on **full-shot** one-class classification setting. It should be noted that better performance is gained if **extra anomaly image data** is used, but it would be very unfair comparison to our method that uses only **8-shot** normal images.

As shown in Table 5, our method outperforms all competing models on almost all cases, despite the fact that our method uses only 8-shot normal images for prompting and does not require any training on the medical data whereas the medical image AD methods require extensive training on a large set of normal medical images, indicating the superior generalized AD capability of our model.

C.3. Full Results on VisA and MVTec AD

Table 6 presents detailed comparison results of InCTRL against six SotA methods across each category of the VisA dataset. Overall, InCTRL markedly surpasses all competitors in every case within the three few-shot settings. We observe a general improvement in performance across all methods with an increase in the number of few-shot image prompts.

Similarly, Table 7 details the results of InCTRL and six SotA methods across each category of the MVTec AD dataset. InCTRL again consistently outperforms all baseline models in all few-shot settings.

[1]The dataset is available at https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset.

Data subset		AUROC							AUPRC						
		SPADE	PaDiM	Patchcore	RegAD	CoOp	WinCLIP	Ours (InCTRL)	SPADE	PaDiM	Patchcore	RegAD	CoOp	WinCLIP	Ours (InCTRL)
2-shot	candle	0.840±0.057	0.876±0.017	0.935±0.013	0.455±0.042	0.943±0.019	0.974±0.007	0.916±0.006	0.889±0.033	0.839±0.015	0.940±0.021	0.461±0.038	0.950±0.015	0.974±0.006	0.920±0.008
	capsules	0.702±0.106	0.540±0.036	0.573±0.029	0.420±0.030	0.697±0.021	0.772±0.015	0.820±0.029	0.771±0.067	0.664±0.022	0.698±0.016	0.583±0.022	0.822±0.017	0.789±0.011	0.846±0.017
	cashew	0.952±0.033	0.674±0.024	0.915±0.010	0.811±0.021	0.933±0.008	0.943±0.008	0.978±0.007	0.954±0.021	0.814±0.018	0.959±0.005	0.879±0.019	0.971±0.004	0.974±0.006	0.990±0.007
	chewinggum	0.886±0.021	0.734±0.020	0.957±0.008	0.511±0.013	0.973±0.011	0.988±0.009	0.997±0.010	0.909±0.022	0.839±0.015	0.983±0.008	0.703±0.011	0.987±0.008	0.994±0.004	0.998±0.012
	fryum	0.846±0.025	0.678±0.039	0.811±0.024	0.592±0.035	0.899±0.008	0.791±0.037	0.953±0.011	0.912±0.019	0.801±0.021	0.853±0.017	0.743±0.026	0.957±0.005	0.894±0.019	0.976±0.009
	macaroni1	0.690±0.043	0.538±0.052	0.763±0.055	0.425±0.119	0.816±0.025	0.885±0.022	0.770±0.038	0.651±0.038	0.540±0.038	0.795±0.039	0.433±0.082	0.812±0.015	0.893±0.018	0.786±0.013
	macaroni2	0.586±0.024	0.614±0.022	0.560±0.054	0.476±0.023	0.774±0.047	0.663±0.018	0.726±0.023	0.548±0.020	0.606±0.039	0.597±0.031	0.502±0.021	0.774±0.034	0.678±0.020	0.734±0.015
	pcb1	0.883±0.049	0.667±0.028	0.769±0.074	0.653±0.025	0.521±0.032	0.831±0.046	0.937±0.039	0.863±0.043	0.647±0.018	0.684±0.059	0.656±0.016	0.582±0.024	0.831±0.032	0.945±0.024
	pcb2	0.783±0.065	0.600±0.031	0.822±0.012	0.453±0.082	0.673±0.067	0.655±0.025	0.660±0.020	0.789±0.031	0.606±0.021	0.848±0.009	0.476±0.065	0.683±0.033	0.664±0.037	0.669±0.018
	pcb3	0.761±0.034	0.709±0.029	0.803±0.016	0.463±0.027	0.709±0.022	0.759±0.016	0.734±0.018	0.776±0.023	0.656±0.018	0.821±0.012	0.523±0.023	0.728±0.020	0.771±0.021	0.716±0.030
	pcb4	0.885±0.079	0.671±0.087	0.952±0.021	0.763±0.038	0.778±0.043	0.895±0.063	0.966±0.027	0.893±0.057	0.672±0.031	0.955±0.016	0.756±0.025	0.773±0.034	0.891±0.057	0.965±0.021
	pipe_fryum	0.727±0.028	0.864±0.024	0.945±0.011	0.667±0.020	0.958±0.008	0.942±0.014	0.835±0.012	0.855±0.017	0.941±0.015	0.964±0.006	0.657±0.019	0.979±0.009	0.960±0.005	0.975±0.006
	MEAN	**0.795±0.045**	**0.680±0.042**	**0.817±0.028**	**0.557±0.053**	**0.806±0.023**	**0.842±0.024**	**0.858±0.022**	**0.818±0.031**	**0.719±0.027**	**0.841±0.023**	**0.614±0.037**	**0.835±0.019**	**0.859±0.021**	**0.877±0.016**
4-shot	candle	0.885±0.043	0.885±0.028	0.947±0.011	0.468±0.037	0.950±0.008	0.975±0.005	0.937±0.005	0.894±0.034	0.841±0.019	0.947±0.007	0.478±0.024	0.955±0.012	0.975±0.007	0.945±0.004
	capsules	0.711±0.091	0.549±0.044	0.716±0.023	0.446±0.022	0.766±0.018	0.813±0.022	0.859±0.005	0.772±0.077	0.666±0.030	0.805±0.014	0.582±0.027	0.863±0.010	0.828±0.013	0.864±0.010
	cashew	0.970±0.018	0.691±0.016	0.952±0.013	0.817±0.015	0.945±0.013	0.955±0.008	0.978±0.013	0.975±0.014	0.832±0.019	0.976±0.006	0.884±0.016	0.978±0.011	0.984±0.005	0.990±0.005
	chewinggum	0.909±0.024	0.838±0.027	0.973±0.006	0.535±0.011	0.975±0.005	0.991±0.007	0.998±0.006	0.911±0.018	0.898±0.029	0.989±0.003	0.716±0.013	0.988±0.006	0.996±0.003	0.998±0.007
	fryum	0.853±0.022	0.725±0.018	0.830±0.019	0.593±0.028	0.898±0.006	0.818±0.027	0.967±0.010	0.914±0.013	0.838±0.013	0.927±0.005	0.744±0.023	0.957±0.006	0.904±0.008	0.985±0.007
	macaroni1	0.714±0.037	0.631±0.043	0.767±0.037	0.456±0.096	0.811±0.019	0.897±0.022	0.776±0.031	0.672±0.025	0.602±0.024	0.789±0.023	0.459±0.048	0.809±0.012	0.901±0.007	0.812±0.012
	macaroni2	0.592±0.012	0.668±0.015	0.548±0.020	0.521±0.034	0.775±0.038	0.672±0.017	0.746±0.018	0.558±0.016	0.646±0.017	0.528±0.016	0.542±0.022	0.774±0.024	0.685±0.026	0.772±0.018
	pcb1	0.892±0.037	0.779±0.021	0.759±0.048	0.693±0.021	0.527±0.024	0.858±0.024	0.959±0.024	0.873±0.032	0.743±0.017	0.672±0.021	0.694±0.015	0.585±0.015	0.875±0.027	0.962±0.021
	pcb2	0.798±0.055	0.608±0.024	0.872±0.011	0.455±0.068	0.689±0.045	0.678±0.016	0.699±0.018	0.790±0.046	0.605±0.025	0.886±0.013	0.482±0.060	0.691±0.028	0.690±0.035	0.724±0.013
	pcb3	0.770±0.031	0.709±0.012	0.825±0.016	0.473±0.013	0.715±0.016	0.779±0.009	0.761±0.016	0.780±0.024	0.657±0.011	0.847±0.017	0.540±0.021	0.735±0.014	0.782±0.019	0.795±0.023
	pcb4	0.889±0.048	0.857±0.043	0.989±0.013	0.765±0.022	0.801±0.033	0.905±0.071	0.975±0.022	0.902±0.013	0.815±0.032	0.988±0.005	0.765±0.023	0.783±0.017	0.907±0.041	0.982±0.028
	pipe_fryum	0.745±0.025	0.885±0.019	0.937±0.006	0.668±0.012	0.962±0.010	0.958±0.011	0.869±0.011	0.866±0.017	0.947±0.015	0.971±0.007	0.653±0.018	0.982±0.032	0.968±0.006	0.989±0.005
	MEAN	**0.811±0.040**	**0.735±0.031**	**0.843±0.025**	**0.574±0.042**	**0.818±0.018**	**0.858±0.025**	**0.877±0.019**	**0.826±0.024**	**0.758±0.018**	**0.860±0.016**	**0.628±0.034**	**0.842±0.016**	**0.875±0.023**	**0.902±0.027**
8-shot	candle	0.899±0.036	0.903±0.022	0.966±0.009	0.781±0.031	0.960±0.003	0.980±0.007	0.955±0.006	0.925±0.012	0.858±0.016	0.968±0.004	0.744±0.026	0.964±0.013	0.980±0.008	0.956±0.005
	capsules	0.712±0.104	0.560±0.046	0.756±0.021	0.469±0.019	0.781±0.018	0.815±0.020	0.863±0.025	0.788±0.067	0.691±0.021	0.845±0.012	0.624±0.018	0.877±0.010	0.832±0.011	0.869±0.013
	cashew	0.977±0.026	0.787±0.023	0.958±0.017	0.519±0.013	0.934±0.015	0.961±0.004	0.979±0.015	0.983±0.010	0.882±0.017	0.978±0.008	0.681±0.015	0.972±0.007	0.987±0.006	0.990±0.005
	chewinggum	0.912±0.017	0.890±0.021	0.980±0.009	0.556±0.009	0.973±0.007	0.994±0.008	0.998±0.006	0.926±0.008	0.937±0.016	0.992±0.003	0.720±0.010	0.987±0.005	0.997±0.004	0.999±0.003
	fryum	0.857±0.023	0.767±0.015	0.858±0.017	0.743±0.022	0.901±0.006	0.826±0.023	0.970±0.010	0.919±0.016	0.848±0.015	0.938±0.005	0.846±0.020	0.958±0.007	0.916±0.018	0.990±0.006
	macaroni1	0.728±0.036	0.645±0.038	0.798±0.036	0.463±0.085	0.811±0.012	0.905±0.032	0.781±0.032	0.705±0.026	0.621±0.023	0.799±0.014	0.466±0.056	0.808±0.023	0.913±0.016	0.818±0.018
	macaroni2	0.598±0.018	0.683±0.013	0.547±0.024	0.543±0.039	0.775±0.035	0.677±0.015	0.751±0.017	0.578±0.023	0.651±0.013	0.549±0.012	0.527±0.018	0.775±0.029	0.689±0.026	0.773±0.008
	pcb1	0.897±0.032	0.812±0.016	0.753±0.045	0.531±0.023	0.540±0.025	0.904±0.024	0.964±0.022	0.886±0.026	0.767±0.019	0.659±0.018	0.548±0.022	0.610±0.028	0.877±0.013	0.967±0.004
	pcb2	0.804±0.047	0.638±0.019	0.907±0.011	0.509±0.079	0.695±0.042	0.690±0.013	0.696±0.019	0.815±0.032	0.613±0.026	0.912±0.013	0.563±0.065	0.696±0.034	0.696±0.026	0.727±0.017
	pcb3	0.777±0.028	0.749±0.021	0.839±0.019	0.466±0.019	0.716±0.017	0.789±0.006	0.798±0.029	0.795±0.021	0.718±0.012	0.855±0.009	0.478±0.015	0.733±0.012	0.791±0.013	0.796±0.028
	pcb4	0.894±0.045	0.878±0.039	0.993±0.011	0.703±0.023	0.813±0.029	0.918±0.160	0.977±0.028	0.918±0.012	0.838±0.033	0.993±0.004	0.634±0.021	0.808±0.033	0.918±0.009	0.975±0.014
	pipe_fryum	0.801±0.026	0.899±0.014	0.961±0.009	0.786±0.014	0.967±0.015	0.959±0.009	0.909±0.018	0.887±0.018	0.953±0.010	0.982±0.008	0.883±0.013	0.987±0.016	0.968±0.004	0.990±0.009
	MEAN	**0.821±0.042**	**0.768±0.032**	**0.860±0.026**	**0.589±0.040**	**0.822±0.021**	**0.868±0.020**	**0.887±0.021**	**0.844±0.031**	**0.781±0.024**	**0.873±0.022**	**0.643±0.032**	**0.848±0.020**	**0.880±0.021**	**0.904±0.025**

Table 6. Fine-grained AUROC and AUPRC results(mean±std) on VisA datasets under various few-shot AD settings. Best results and the second-best results are respectively highlighted in **red** and **blue**.

C.4. Full Results on One-vs-all Setting of Sementic Anomaly Detection Datasets

Table 8 provides the detailed empirical comparison results of InCTRL against five SotA methods using the one-vs-all protocol on the MNIST dataset. Similarly, Table 9 details the performance of InCTRL relative to five SotA methods under the one-vs-all protocol on the CIFAR-10 dataset. InCTRL consistently surpasses all baseline methods in all few-shot settings on both datasets.

References

[1] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Mvtec ad–a comprehensive real-world dataset for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9592–9600, 2019. 1

[2] Yu Cai, Hao Chen, Xin Yang, Yu Zhou, and Kwang-Ting Cheng. Dual-distribution discrepancy with self-supervised refinement for anomaly detection in medical images. *Medical Image Analysis*, 86:102794, 2023. 3

[3] Tri Cao, Jiawen Zhu, and Guansong Pang. Anomaly detection under distribution shift. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 6511–6523, 2023. 1

[4] Sergiu Deitsch, Vincent Christlein, Stephan Berger, Claudia Buerhop-Lutz, Andreas Maier, Florian Gallwitz, and Christian Riess. Automatic classification of defective photovoltaic module cells in electroluminescence images. *Solar Energy*, 185:455–468, 2019. 1

[5] Gabriel Ilharco, Mitchell Wortsman, Ross Wightman, Cade Gordon, Nicholas Carlini, Rohan Taori, Achal Dave, Vaishaal Shankar, Hongseok Namkoong, John Miller, et al. Openclip. *Zenodo*, 4:5, 2021. 2

[6] Jongheon Jeong, Yang Zou, Taewan Kim, Dongqing Zhang, Avinash Ravichandran, and Onkar Dabeer. Winclip: Zero-/few-shot anomaly classification and segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19606–19616, 2023. 2, 3

[7] Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton. Cifar-10 (canadian institute for advanced research). 2009. *URL http://www. cs. toronto. edu/kriz/cifar. html*, 5, 2009. 1, 2

[8] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. 1, 2

[9] Liu Li, Mai Xu, Xiaofei Wang, Lai Jiang, and Hanruo Liu. Attention based glaucoma detection: A large-scale database and cnn model, 2019. 3

[10] Mohammadreza Salehi, Niousha Sadjadi, Soroosh Baselizadeh, Mohammad H Rohban, and Hamid R Rabiee. Multiresolution knowledge distillation for anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14902–14912, 2021. 1, 2

[11] Javier Silvestre-Blanes, Teresa Albero-Albero, Ignacio Miralles, Rubén Pérez-Llorens, and Jorge Moreno. A public

Data subset		AUROC							AUPRC						
		SPADE	PaDiM	Patchcore	RegAD	CoOp	WinCLIP	Ours (InCTRL)	SPADE	PaDiM	Patchcore	RegAD	CoOp	WinCLIP	Ours (InCTRL)
2-shot	carpet	0.976±0.008	0.992±0.005	0.996±0.003	0.601±0.051	1.000±0.000	0.998±0.002	0.997±0.003	0.962±0.005	0.998±0.003	0.999±0.001	0.853±0.023	1.000±0.000	1.000±0.000	0.999±0.001
	grid	0.404±0.054	0.560±0.041	0.574±0.102	0.627±0.069	0.675±0.052	0.965±0.007	0.974±0.005	0.727±0.036	0.751±0.038	0.785±0.037	0.830±0.029	0.875±0.038	0.984±0.004	0.988±0.003
	leather	1.000±0.000	1.000±0.000	1.000±0.000	0.556±0.047	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000	0.835±0.016	1.000±0.000	1.000±0.000	1.000±0.000
	tile	0.940±0.011	0.947±0.013	0.987±0.012	0.807±0.027	0.999±0.002	0.997±0.003	1.000±0.000	0.978±0.007	0.980±0.005	0.996±0.004	0.925±0.013	0.998±0.002	0.999±0.002	1.000±0.000
	wood	0.932±0.008	0.982±0.005	0.985±0.009	0.921±0.011	0.966±0.008	0.993±0.005	0.996±0.003	0.979±0.004	0.995±0.002	0.995±0.005	0.976±0.014	0.989±0.006	0.998±0.003	0.998±0.002
	bottle	1.000±0.000	0.971±0.012	1.000±0.000	0.529±0.045	0.973±0.013	0.993±0.007	0.998±0.002	1.000±0.000	0.991±0.005	1.000±0.000	0.830±0.026	0.992±0.004	0.998±0.003	1.000±0.000
	capsule	0.736±0.027	0.741±0.026	0.666±0.068	0.593±0.073	0.823±0.024	0.806±0.068	0.860±0.081	0.935±0.015	0.945±0.010	0.911±0.021	0.861±0.017	0.959±0.016	0.945±0.042	0.960±0.063
	pill	0.759±0.023	0.601±0.059	0.829±0.022	0.693±0.036	0.895±0.022	0.888±0.014	0.893±0.024	0.944±0.018	0.889±0.031	0.965±0.009	0.923±0.015	0.980±0.009	0.978±0.006	0.973±0.011
	transistor	0.853±0.019	0.791±0.043	0.992±0.014	0.687±0.033	0.702±0.013	0.898±0.021	0.875±0.027	0.845±0.021	0.700±0.055	0.989±0.007	0.694±0.029	0.664±0.025	0.858±0.017	0.845±0.023
	zipper	0.905±0.033	0.767±0.024	0.978±0.016	0.602±0.049	0.877±0.018	0.955±0.013	0.954±0.010	0.974±0.010	0.875±0.021	0.994±0.005	0.869±0.018	0.968±0.011	0.979±0.009	0.988±0.009
	cable	0.767±0.046	0.525±0.057	0.805±0.025	0.513±0.057	0.724±0.067	0.862±0.011	0.884±0.013	0.968±0.015	0.693±0.063	0.893±0.018	0.674±0.023	0.777±0.052	0.904±0.009	0.920±0.009
	hazelnut	0.900±0.035	0.947±0.012	0.864±0.027	0.697±0.035	0.938±0.014	0.950±0.008	0.968±0.006	0.946±0.013	0.964±0.009	0.919±0.015	0.844±0.026	0.959±0.010	0.975±0.005	0.987±0.006
	metal_nut	0.760±0.062	0.613±0.035	0.851±0.053	0.632±0.038	0.942±0.009	0.930±0.011	0.942±0.028	0.876±0.042	0.876±0.016	0.965±0.008	0.878±0.017	0.987±0.006	0.984±0.006	0.985±0.010
	screw	0.531±0.024	0.501±0.046	0.531±0.027	0.654±0.041	0.659±0.027	0.780±0.019	0.803±0.043	0.766±0.011	0.751±0.017	0.747±0.033	0.843±0.022	0.867±0.018	0.903±0.013	0.917±0.022
	toothbrush	0.786±0.041	0.844±0.043	0.811±0.029	0.486±0.053	0.703±0.022	0.952±0.021	0.960±0.028	0.937±0.020	0.938±0.011	0.922±0.016	0.717±0.024	0.811±0.019	0.972±0.007	0.975±0.019
	MEAN	**0.817±0.054**	**0.785±0.025**	**0.858±0.034**	**0.640±0.047**	**0.858±0.016**	**0.931±0.019**	**0.940±0.015**	**0.922±0.023**	**0.890±0.015**	**0.939±0.012**	**0.837±0.034**	**0.922±0.007**	**0.965±0.007**	**0.969±0.004**
4-shot	carpet	0.981±0.006	0.994±0.004	0.994±0.005	0.604±0.046	0.998±0.002	0.997±0.003	1.000±0.000	0.996±0.005	0.998±0.002	0.998±0.002	0.855±0.025	0.999±0.001	0.999±0.001	1.000±0.000
	grid	0.476±0.046	0.556±0.049	0.565±0.077	0.660±0.043	0.769±0.035	0.978±0.004	0.998±0.003	0.714±0.038	0.752±0.042	0.781±0.032	0.859±0.025	0.894±0.021	0.985±0.004	0.999±0.001
	leather	1.000±0.000	1.000±0.000	1.000±0.000	0.593±0.037	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000	0.879±0.021	1.000±0.000	1.000±0.000	1.000±0.000
	tile	1.000±0.000	0.954±0.008	0.987±0.010	0.708±0.033	0.999±0.001	0.999±0.001	0.999±0.002	1.000±0.000	0.984±0.004	0.996±0.005	0.883±0.024	1.000±0.000	0.999±0.001	0.999±0.001
	wood	0.989±0.011	0.978±0.006	0.987±0.008	0.887±0.026	0.970±0.006	0.997±0.002	0.995±0.005	0.995±0.006	0.994±0.004	0.996±0.006	0.965±0.018	0.991±0.007	0.998±0.001	0.994±0.004
	bottle	1.000±0.000	0.976±0.008	1.000±0.000	0.656±0.035	0.976±0.011	0.992±0.006	1.000±0.000	1.000±0.000	0.993±0.005	1.000±0.000	0.865±0.031	0.993±0.005	0.997±0.004	1.000±0.000
	capsule	0.729±0.018	0.752±0.019	0.777±0.055	0.610±0.063	0.828±0.019	0.840±0.052	0.868±0.043	0.936±0.013	0.958±0.013	0.943±0.023	0.865±0.014	0.959±0.014	0.948±0.037	0.965±0.027
	pill	0.872±0.021	0.650±0.047	0.875±0.019	0.691±0.028	0.906±0.016	0.890±0.017	0.874±0.019	0.978±0.015	0.926±0.025	0.976±0.010	0.926±0.013	0.982±0.011	0.978±0.005	0.967±0.016
	transistor	0.855±0.016	0.868±0.033	0.994±0.008	0.684±0.025	0.719±0.008	0.898±0.014	0.895±0.021	0.844±0.014	0.812±0.036	0.992±0.006	0.754±0.021	0.671±0.018	0.858±0.015	0.875±0.016
	zipper	0.913±0.027	0.759±0.017	0.979±0.013	0.551±0.027	0.881±0.021	0.965±0.010	0.964±0.008	0.976±0.008	0.872±0.015	0.994±0.005	0.854±0.016	0.961±0.016	0.980±0.007	0.990±0.006
	cable	0.818±0.043	0.550±0.042	0.822±0.032	0.510±0.044	0.751±0.046	0.865±0.008	0.888±0.021	0.904±0.026	0.732±0.048	0.908±0.022	0.663±0.018	0.781±0.034	0.912±0.006	0.920±0.008
	hazelnut	0.771±0.037	0.951±0.030	0.999±0.002	0.685±0.029	0.971±0.010	0.954±0.008	0.972±0.005	0.858±0.034	0.978±0.006	0.999±0.001	0.876±0.020	0.975±0.012	0.975±0.005	0.989±0.006
	metal_nut	0.734±0.048	0.741±0.021	0.944±0.036	0.701±0.032	0.943±0.006	0.948±0.009	0.946±0.024	0.971±0.033	0.923±0.013	0.987±0.004	0.880±0.015	0.987±0.005	0.986±0.005	0.989±0.008
	screw	0.531±0.019	0.504±0.038	0.537±0.026	0.599±0.034	0.720±0.018	0.800±0.013	0.815±0.031	0.760±0.008	0.760±0.014	0.751±0.028	0.806±0.026	0.864±0.014	0.915±0.009	0.920±0.016
	toothbrush	0.750±0.042	0.847±0.035	0.818±0.023	0.806±0.041	0.678±0.016	0.975±0.016	0.963±0.024	0.933±0.021	0.949±0.009	0.925±0.013	0.757±0.022	0.804±0.015	0.985±0.005	0.978±0.015
	MEAN	**0.828±0.044**	**0.805±0.018**	**0.885±0.026**	**0.663±0.032**	**0.874±0.017**	**0.940±0.021**	**0.945±0.018**	**0.924±0.015**	**0.909±0.013**	**0.950±0.013**	**0.846±0.026**	**0.924±0.008**	**0.968±0.008**	**0.972±0.006**
8-shot	carpet	0.987±0.004	0.994±0.004	0.993±0.006	0.601±0.045	1.000±0.000	0.998±0.002	1.000±0.000	0.991±0.006	0.998±0.002	0.998±0.002	0.858±0.032	1.000±0.000	0.999±0.001	1.000±0.000
	grid	0.477±0.044	0.612±0.042	0.652±0.053	0.679±0.036	0.759±0.033	0.986±0.006	0.999±0.001	0.718±0.035	0.808±0.033	0.828±0.027	0.862±0.024	0.891±0.017	0.987±0.005	0.999±0.001
	leather	1.000±0.000	1.000±0.000	1.000±0.000	0.646±0.028	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000	0.821±0.014	1.000±0.000	1.000±0.000	1.000±0.000
	tile	1.000±0.000	0.954±0.006	0.989±0.006	0.782±0.025	0.999±0.001	0.998±0.001	0.998±0.001	1.000±0.000	0.985±0.003	0.997±0.004	0.952±0.020	1.000±0.000	0.999±0.001	0.999±0.001
	wood	0.991±0.006	0.989±0.004	0.991±0.005	0.909±0.019	0.968±0.007	0.996±0.003	0.996±0.004	0.995±0.005	0.997±0.003	0.997±0.004	0.972±0.013	0.990±0.006	0.999±0.001	0.995±0.003
	bottle	1.000±0.000	0.987±0.007	1.000±0.000	0.659±0.027	0.976±0.010	0.994±0.003	1.000±0.000	1.000±0.000	0.996±0.004	1.000±0.000	0.868±0.025	0.993±0.006	0.998±0.002	1.000±0.000
	capsule	0.729±0.013	0.780±0.015	0.929±0.038	0.689±0.049	0.836±0.015	0.884±0.041	0.890±0.037	0.936±0.012	0.979±0.011	0.983±0.016	0.883±0.011	0.961±0.008	0.968±0.035	0.966±0.022
	pill	0.883±0.029	0.657±0.039	0.891±0.015	0.675±0.024	0.900±0.011	0.895±0.016	0.888±0.016	0.983±0.013	0.939±0.027	0.978±0.011	0.918±0.014	0.981±0.009	0.978±0.007	0.969±0.015
	transistor	0.889±0.026	0.887±0.041	0.997±0.005	0.689±0.018	0.742±0.019	0.898±0.013	0.903±0.019	0.877±0.019	0.864±0.047	0.995±0.005	0.784±0.016	0.712±0.021	0.861±0.013	0.884±0.013
	zipper	0.924±0.022	0.757±0.015	0.984±0.008	0.581±0.021	0.889±0.016	0.965±0.008	0.970±0.012	0.984±0.006	0.880±0.013	0.996±0.004	0.864±0.013	0.965±0.011	0.981±0.005	0.993±0.008
	cable	0.872±0.067	0.602±0.033	0.944±0.027	0.521±0.035	0.819±0.035	0.886±0.007	0.932±0.015	0.931±0.035	0.798±0.037	0.969±0.015	0.675±0.016	0.863±0.027	0.928±0.005	0.960±0.007
	hazelnut	0.784±0.033	0.953±0.026	1.000±0.000	0.744±0.024	0.968±0.012	0.958±0.007	0.975±0.007	0.869±0.030	0.985±0.007	1.000±0.000	0.889±0.021	0.971±0.010	0.980±0.009	0.987±0.008
	metal_nut	0.764±0.036	0.783±0.025	0.969±0.023	0.735±0.035	0.946±0.005	0.948±0.012	0.947±0.026	0.973±0.030	0.955±0.018	0.993±0.004	0.898±0.018	0.988±0.008	0.988±0.006	0.990±0.005
	screw	0.544±0.021	0.501±0.034	0.645±0.021	0.600±0.033	0.718±0.014	0.815±0.014	0.825±0.029	0.764±0.013	0.765±0.016	0.762±0.024	0.809±0.024	0.860±0.012	0.934±0.011	0.935±0.013
	toothbrush	0.756±0.036	0.849±0.027	0.850±0.016	0.603±0.075	0.681±0.018	0.978±0.012	0.975±0.022	0.934±0.019	0.956±0.008	0.941±0.011	0.766±0.041	0.818±0.016	0.990±0.003	0.982±0.013
	MEAN	**0.840±0.057**	**0.820±0.016**	**0.922±0.019**	**0.674±0.033**	**0.880±0.014**	**0.947±0.025**	**0.953±0.013**	**0.930±0.016**	**0.927±0.012**	**0.962±0.013**	**0.855±0.021**	**0.933±0.007**	**0.973±0.009**	**0.977±0.006**

Table 7. Fine-grained AUROC and AUPRC results(mean±std) on MVTec AD datasets under various few-shot AD settings. Best results and the second-best results are respectively highlighted in **red** and **blue**.

fabric database for defect detection methods and results. *Autex Research Journal*, 19(4):363–374, 2019. 1, 2

[12] Domen Tabernik, Samo Šela, Jure Skvarč, and Danijel Skočaj. Segmentation-based deep-learning approach for surface-defect detection. *Journal of Intelligent Manufacturing*, 31(3):759–776, 2020. 1

[13] Kaiyang Zhou, Jingkang Yang, Chen Change Loy, and Ziwei Liu. Conditional prompt learning for vision-language models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16816–16825, 2022. 2

[14] Yang Zou, Jongheon Jeong, Latha Pemula, Dongqing Zhang, and Onkar Dabeer. Spot-the-difference self-supervised pretraining for anomaly detection and segmentation. In *European Conference on Computer Vision*, pages 392–408. Springer, 2022. 1, 3

Data subset		AUROC						AUPRC					
		SPADE	Patchcore	RegAD	CoOp	WinCLIP	Ours (InCTRL)	SPADE	Patchcore	RegAD	CoOp	WinCLIP	Ours (InCTRL)
2-shot	0	0.772±0.174	0.803±0.028	0.615±0.037	0.508±0.031	0.801±0.026	0.970±0.004	0.959±0.036	0.950±0.003	0.923±0.013	0.927±0.004	0.954±0.004	0.995±0.000
	1	0.961±0.021	0.988±0.006	0.734±0.024	0.439±0.018	0.928±0.011	0.868±0.047	0.994±0.004	0.998±0.001	0.956±0.007	0.901±0.002	0.986±0.003	0.935±0.019
	2	0.670±0.055	0.643±0.035	0.658±0.062	0.570±0.053	0.723±0.016	0.860±0.035	0.946±0.010	0.933±0.008	0.935±0.016	0.923±0.005	0.943±0.004	0.969±0.005
	3	0.808±0.065	0.812±0.015	0.716±0.077	0.767±0.042	0.786±0.091	0.954±0.025	0.973±0.010	0.972±0.001	0.952±0.012	0.965±0.014	0.958±0.019	0.992±0.004
	4	0.747±0.036	0.774±0.024	0.319±0.023	0.599±0.026	0.756±0.086	0.965±0.026	0.959±0.007	0.962±0.004	0.870±0.020	0.939±0.016	0.952±0.017	0.995±0.003
	5	0.759±0.036	0.747±0.024	0.427±0.059	0.745±0.027	0.824±0.023	0.914±0.048	0.969±0.005	0.959±0.003	0.894±0.008	0.962±0.007	0.970±0.004	0.988±0.006
	6	0.725±0.045	0.611±0.047	0.341±0.036	0.428±0.042	0.878±0.018	0.779±0.027	0.959±0.005	0.934±0.009	0.873±0.008	0.902±0.005	0.982±0.002	0.952±0.004
	7	0.708±0.123	0.795±0.039	0.428±0.142	0.651±0.023	0.827±0.019	0.962±0.011	0.947±0.027	0.965±0.007	0.887±0.022	0.945±0.003	0.963±0.004	0.995±0.001
	8	0.800±0.102	0.629±0.045	0.588±0.078	0.344±0.025	0.712±0.031	0.702±0.136	0.970±0.016	0.928±0.011	0.936±0.015	0.873±0.005	0.948±0.005	0.943±0.043
	9	0.841±0.020	0.758±0.041	0.427±0.048	0.515±0.036	0.860±0.058	0.950±0.019	0.976±0.003	0.962±0.008	0.900±0.006	0.919±0.013	0.975±0.011	0.989±0.007
	MEAN	**0.779±0.024**	**0.756±0.004**	**0.525±0.030**	**0.557±0.006**	**0.810±0.008**	**0.892±0.009**	**0.965±0.004**	**0.956±0.001**	**0.913±0.006**	**0.926±0.003**	**0.963±0.001**	**0.975±0.004**
4-shot	0	0.854±0.049	0.874±0.003	0.624±0.045	0.509±0.018	0.813±0.014	0.974±0.003	0.978±0.007	0.982±0.001	0.926±0.009	0.928±0.003	0.956±0.003	0.995±0.001
	1	0.955±0.032	0.995±0.002	0.740±0.023	0.446±0.015	0.942±0.016	0.902±0.017	0.993±0.005	0.999±0.000	0.958±0.011	0.913±0.003	0.989±0.003	0.953±0.009
	2	0.658±0.044	0.712±0.044	0.690±0.058	0.581±0.035	0.752±0.029	0.869±0.071	0.944±0.008	0.946±0.011	0.941±0.007	0.927±0.005	0.949±0.005	0.962±0.015
	3	0.837±0.014	0.869±0.026	0.730±0.106	0.771±0.037	0.929±0.047	0.959±0.017	0.978±0.002	0.980±0.005	0.922±0.014	0.967±0.011	0.987±0.009	0.994±0.005
	4	0.790±0.033	0.856±0.051	0.335±0.034	0.608±0.022	0.841±0.017	0.972±0.022	0.919±0.081	0.976±0.011	0.875±0.023	0.944±0.009	0.968±0.003	0.995±0.002
	5	0.795±0.019	0.844±0.022	0.436±0.067	0.747±0.028	0.851±0.018	0.919±0.022	0.975±0.003	0.979±0.003	0.907±0.016	0.964±0.010	0.975±0.003	0.989±0.005
	6	0.798±0.041	0.750±0.125	0.348±0.039	0.435±0.035	0.914±0.023	0.790±0.018	0.969±0.008	0.960±0.023	0.879±0.003	0.918±0.006	0.987±0.003	0.967±0.011
	7	0.704±0.117	0.872±0.015	0.525±0.122	0.657±0.016	0.838±0.009	0.966±0.006	0.948±0.022	0.978±0.005	0.899±0.031	0.944±0.004	0.965±0.002	0.993±0.001
	8	0.856±0.018	0.728±0.021	0.602±0.087	0.356±0.027	0.734±0.001	0.711±0.038	0.979±0.003	0.953±0.004	0.939±0.017	0.879±0.009	0.952±0.000	0.963±0.006
	9	0.853±0.019	0.831±0.015	0.446±0.052	0.522±0.032	0.896±0.009	0.955±0.011	0.978±0.003	0.973±0.004	0.912±0.006	0.909±0.018	0.981±0.002	0.993±0.002
	MEAN	**0.810±0.009**	**0.833±0.009**	**0.548±0.053**	**0.563±0.004**	**0.851±0.010**	**0.902±0.016**	**0.966±0.008**	**0.972±0.011**	**0.916±0.013**	**0.929±0.002**	**0.971±0.002**	**0.980±0.007**
8-shot	0	0.855±0.038	0.916±0.013	0.631±0.039	0.521±0.014	0.833±0.015	0.978±0.011	0.979±0.006	0.988±0.001	0.932±0.005	0.931±0.002	0.960±0.003	0.995±0.001
	1	0.978±0.003	0.997±0.001	0.747±0.026	0.443±0.019	0.962±0.008	0.918±0.002	0.997±0.001	0.999±0.000	0.963±0.008	0.917±0.004	0.992±0.001	0.999±0.001
	2	0.661±0.010	0.803±0.055	0.651±0.063	0.584±0.032	0.790±0.030	0.870±0.014	0.944±0.002	0.963±0.013	0.947±0.006	0.929±0.002	0.956±0.006	0.965±0.002
	3	0.852±0.020	0.908±0.030	0.712±0.091	0.775±0.044	0.939±0.035	0.978±0.000	0.980±0.003	0.986±0.005	0.956±0.013	0.978±0.007	0.989±0.007	0.997±0.000
	4	0.808±0.012	0.885±0.039	0.353±0.038	0.608±0.025	0.856±0.017	0.978±0.003	0.971±0.002	0.982±0.006	0.930±0.019	0.953±0.005	0.971±0.003	0.997±0.000
	5	0.819±0.016	0.885±0.020	0.453±0.073	0.748±0.036	0.866±0.012	0.920±0.038	0.978±0.002	0.984±0.003	0.907±0.034	0.974±0.013	0.978±0.002	0.989±0.008
	6	0.834±0.005	0.850±0.019	0.352±0.025	0.469±0.039	0.925±0.008	0.855±0.036	0.975±0.001	0.977±0.003	0.881±0.002	0.925±0.005	0.988±0.001	0.990±0.006
	7	0.777±0.078	0.899±0.031	0.558±0.119	0.655±0.017	0.856±0.009	0.968±0.005	0.961±0.014	0.982±0.008	0.877±0.027	0.955±0.004	0.969±0.002	0.996±0.001
	8	0.870±0.020	0.767±0.023	0.555±0.085	0.354±0.023	0.744±0.005	0.777±0.011	0.981±0.003	0.958±0.006	0.880±0.016	0.885±0.008	0.953±0.001	0.968±0.002
	9	0.839±0.022	0.850±0.019	0.459±0.048	0.515±0.028	0.867±0.006	0.961±0.009	0.976±0.003	0.974±0.003	0.916±0.005	0.919±0.014	0.982±0.001	0.991±0.001
	MEAN	**0.829±0.009**	**0.876±0.004**	**0.547±0.063**	**0.567±0.007**	**0.867±0.007**	**0.920±0.003**	**0.974±0.002**	**0.979±0.001**	**0.919±0.018**	**0.937±0.004**	**0.974±0.001**	**0.989±0.001**

Table 8. Fine-grained AUROC and AUPRC results(mean±std) on One-vs-all protocol of MNIST datasets under various few-shot AD settings. Best results and the second-best results are respectively highlighted in **red** and **blue**.

Data subset		AUROC						AUPRC					
		SPADE	Patchcore	RegAD	CoOp	WinCLIP	Ours (InCTRL)	SPADE	Patchcore	RegAD	CoOp	WinCLIP	Ours (InCTRL)
2-shot	airplane	0.822±0.053	0.566±0.065	0.439±0.043	0.437±0.014	0.917±0.005	0.938±0.006	0.974±0.009	0.915±0.014	0.865±0.011	0.886±0.005	0.988±0.002	0.993±0.001
	automobile	0.929±0.013	0.499±0.041	0.681±0.026	0.800±0.007	0.932±0.002	0.945±0.005	0.991±0.002	0.898±0.012	0.946±0.004	0.969±0.003	0.992±0.000	0.995±0.001
	bird	0.721±0.093	0.583±0.040	0.413±0.019	0.424±0.003	0.925±0.002	0.931±0.011	0.952±0.019	0.923±0.010	0.880±0.016	0.886±0.003	0.990±0.000	0.992±0.001
	cat	0.748±0.087	0.567±0.010	0.516±0.027	0.359±0.003	0.942±0.000	0.940±0.002	0.956±0.019	0.918±0.002	0.914±0.012	0.875±0.002	0.992±0.000	0.992±0.001
	deer	0.869±0.037	0.738±0.034	0.657±0.021	0.436±0.006	0.928±0.005	0.943±0.002	0.978±0.008	0.956±0.006	0.933±0.012	0.884±0.006	0.990±0.001	0.993±0.000
	dog	0.716±0.071	0.529±0.034	0.588±0.024	0.533±0.028	0.842±0.003	0.856±0.009	0.947±0.014	0.911±0.017	0.933±0.003	0.914±0.002	0.978±0.000	0.981±0.001
	frog	0.832±0.026	0.791±0.018	0.650±0.013	0.288±0.045	0.920±0.004	0.931±0.002	0.974±0.004	0.967±0.004	0.946±0.005	0.867±0.008	0.990±0.000	0.992±0.000
	horse	0.787±0.102	0.621±0.027	0.483±0.015	0.762±0.026	0.916±0.001	0.936±0.007	0.962±0.021	0.930±0.007	0.896±0.015	0.957±0.006	0.987±0.000	0.990±0.001
	ship	0.894±0.020	0.601±0.028	0.519±0.027	0.549±0.005	0.951±0.001	0.956±0.006	0.986±0.002	0.929±0.008	0.898±0.004	0.926±0.002	0.992±0.000	0.996±0.001
	truck	0.906±0.058	0.529±0.071	0.391±0.054	0.679±0.013	0.978±0.002	0.976±0.004	0.988±0.008	0.912±0.008	0.878±0.009	0.947±0.005	0.997±0.001	0.998±0.001
	MEAN	**0.823±0.014**	**0.602±0.009**	**0.534±0.005**	**0.527±0.011**	**0.925±0.001**	**0.935±0.002**	**0.971±0.003**	**0.926±0.002**	**0.909±0.003**	**0.911±0.002**	**0.990±0.001**	**0.992±0.000**
4-shot	airplane	0.831±0.038	0.578±0.025	0.474±0.036	0.471±0.008	0.919±0.002	0.942±0.011	0.974±0.007	0.918±0.007	0.886±0.007	0.905±0.004	0.989±0.001	0.995±0.004
	automobile	0.929±0.018	0.538±0.025	0.728±0.018	0.808±0.005	0.933±0.000	0.948±0.022	0.991±0.002	0.909±0.006	0.951±0.002	0.971±0.003	0.992±0.000	0.995±0.007
	bird	0.754±0.048	0.621±0.023	0.424±0.020	0.437±0.005	0.927±0.001	0.933±0.005	0.959±0.009	0.933±0.005	0.874±0.012	0.892±0.001	0.991±0.000	0.992±0.001
	cat	0.760±0.075	0.584±0.021	0.529±0.022	0.356±0.002	0.942±0.001	0.945±0.012	0.959±0.015	0.923±0.005	0.912±0.009	0.872±0.002	0.992±0.000	0.992±0.003
	deer	0.905±0.020	0.789±0.013	0.597±0.013	0.455±0.005	0.932±0.003	0.956±0.003	0.984±0.004	0.965±0.002	0.920±0.006	0.891±0.002	0.990±0.001	0.994±0.000
	dog	0.665±0.066	0.526±0.050	0.554±0.017	0.527±0.024	0.841±0.002	0.866±0.019	0.938±0.015	0.911±0.017	0.914±0.005	0.911±0.003	0.978±0.000	0.984±0.007
	frog	0.862±0.013	0.828±0.011	0.513±0.004	0.296±0.039	0.923±0.001	0.935±0.009	0.979±0.003	0.973±0.002	0.920±0.003	0.871±0.005	0.991±0.000	0.993±0.005
	horse	0.846±0.040	0.694±0.013	0.524±0.012	0.763±0.018	0.916±0.001	0.939±0.012	0.974±0.009	0.946±0.003	0.912±0.010	0.957±0.005	0.987±0.000	0.987±0.002
	ship	0.903±0.027	0.663±0.034	0.496±0.006	0.557±0.007	0.953±0.001	0.958±0.016	0.988±0.003	0.942±0.007	0.904±0.003	0.930±0.001	0.993±0.000	0.991±0.006
	truck	0.901±0.049	0.567±0.053	0.498±0.015	0.699±0.012	0.981±0.002	0.980±0.009	0.987±0.007	0.920±0.013	0.889±0.002	0.951±0.003	0.997±0.001	0.994±0.003
	MEAN	**0.836±0.006**	**0.639±0.010**	**0.534±0.002**	**0.537±0.005**	**0.927±0.001**	**0.940±0.010**	**0.973±0.002**	**0.934±0.003**	**0.908±0.001**	**0.915±0.003**	**0.990±0.000**	**0.992±0.004**
8-shot	airplane	0.847±0.023	0.590±0.015	0.528±0.034	0.481±0.005	0.922±0.001	0.945±0.002	0.978±0.004	0.921±0.005	0.902±0.005	0.909±0.005	0.990±0.000	0.993±0.000
	automobile	0.935±0.012	0.594±0.007	0.770±0.012	0.813±0.007	0.934±0.000	0.950±0.004	0.992±0.002	0.924±0.001	0.945±0.002	0.972±0.001	0.992±0.000	0.996±0.000
	bird	0.711±0.025	0.632±0.008	0.456±0.016	0.440±0.004	0.928±0.001	0.941±0.003	0.951±0.005	0.935±0.001	0.866±0.009	0.892±0.001	0.991±0.000	0.993±0.000
	cat	0.768±0.063	0.606±0.022	0.533±0.023	0.361±0.002	0.943±0.001	0.951±0.003	0.961±0.012	0.928±0.005	0.887±0.007	0.874±0.003	0.992±0.000	0.993±0.001
	deer	0.920±0.009	0.821±0.002	0.628±0.018	0.452±0.003	0.932±0.001	0.963±0.003	0.987±0.001	0.971±0.000	0.923±0.003	0.895±0.002	0.991±0.000	0.993±0.000
	dog	0.733±0.020	0.576±0.006	0.565±0.023	0.525±0.005	0.842±0.001	0.874±0.029	0.954±0.005	0.926±0.002	0.897±0.014	0.911±0.002	0.978±0.000	0.984±0.006
	frog	0.877±0.008	0.850±0.004	0.525±0.007	0.311±0.026	0.923±0.001	0.937±0.002	0.982±0.002	0.977±0.001	0.949±0.002	0.901±0.007	0.991±0.000	0.995±0.000
	horse	0.843±0.009	0.722±0.015	0.529±0.014	0.763±0.024	0.916±0.000	0.941±0.001	0.973±0.002	0.952±0.003	0.891±0.011	0.956±0.003	0.987±0.000	0.995±0.000
	ship	0.929±0.017	0.711±0.030	0.501±0.008	0.556±0.008	0.954±0.001	0.968±0.004	0.991±0.003	0.951±0.005	0.950±0.005	0.929±0.002	0.993±0.000	0.996±0.001
	truck	0.930±0.019	0.618±0.029	0.512±0.009	0.720±0.009	0.982±0.001	0.984±0.001	0.991±0.003	0.932±0.006	0.904±0.001	0.956±0.002	0.998±0.000	0.998±0.000
	MEAN	**0.849±0.006**	**0.672±0.006**	**0.555±0.008**	**0.542±0.005**	**0.928±0.001**	**0.945±0.002**	**0.976±0.001**	**0.942±0.002**	**0.911±0.001**	**0.920±0.003**	**0.990±0.000**	**0.994±0.001**

Table 9. Fine-grained AUROC and AUPRC results(mean±std) on One-vs-all protocol of CIFAR-10 datasets under various few-shot AD settings. Best results and the second-best results are respectively highlighted in **red** and **blue**.