Supplementary Material of “Anomaly Detection under Distribution Shift”

Tri Cao, Jiawen Zhu, and Guansong Pang

School of Computing and Information Systems, Singapore Management University

1. Full results on MNIST/MNIST-M, PACS and MVTec

The full experimental results on the MNIST/MNIST-M, PACS and MVTec datasets are presented in Tables 1, 2, and 3, respectively. These tables provide a detailed illustration of the performance of the considered methods and the proposed method GNL. It is observed that GNL consistently outperforms the compared methods on many subsets of these datasets, indicating the effectiveness and robustness of GNL.

2. Implementation details

2.1. Training

With MVTec, all images in MVTec are resized to 256x256. We take ResNet50 [1] as the backbone of the teacher encoder. The hyperparameters for PACS are the same as for MVTec. For MNIST/MNIST-M, all images are in their original scale, which are 28 × 28. We take ResNet18 [1] as the backbone of the teacher encoder.

With all datasets, the learning rate is set to 0.005 with a batch size of 16 and is optimized by Adam [2] optimizer with $\beta = (0.5, 0.999)$. The model is trained 20 epochs on MVTec, PACS, CIFAR-10 and 5 epochs on MNIST/MNIST-M dataset. The pseudo code of our training is shown in Algorithm 1.

Algorithm 1 The pseudo code of our training

for each batch $(ori, augs)$ in dataloader **do**
    $ens\_ori \leftarrow encoder(ori)$ ▷ Return a tuple with 3 embedded features from three residual encoder blocks, ordered from low-level features to abstract features
    $bn\_ori \leftarrow bn(ens\_ori)$ ▷ The feature at Bottleneck
    $des\_ori \leftarrow decoder(bn\_ori)$ ▷ Return a tuple with 3 reconstructed features from three residual decoder blocks, ordered from low-level features to abstract features
    $loss\_ori \leftarrow loss(ens\_ori, des\_ori)$
    $losses\_abs \leftarrow 0$
    $losses\_lowf \leftarrow 0$
    for each augmented image aug in $augs$ **do**
        $ens\_aug \leftarrow encoder(aug)$
        $bn\_aug \leftarrow bn(ens\_aug)$
        $des\_aug \leftarrow decoder(bn\_aug)$
        $loss\_abs \leftarrow loss(bn\_ori, bn\_aug)$
        $loss\_lowf \leftarrow loss(des\_ori[0], des\_aug[0])$
        $losses\_abs \leftarrow losses\_abs + loss\_abs$
        $losses\_lowf \leftarrow losses\_lowf + loss\_lowf$
    end for
    $losses\_abs \leftarrow losses\_abs/N$
    $losses\_lowf \leftarrow losses\_lowf/N$
    $sum\_loss \leftarrow alpha\_ori \times loss\_ori + alpha\_abs \times losses\_abs + alpha\_lowf \times loss\_lowf$
    Compute gradients of sum_loss with respect to the trainable parameters of the model
    Update the trainable parameters of the model using the optimizer
end for

2.2. Inference

For a given test sample x, our test time augmentation method performs the augmentation as follows:

$$\text{FDM}(\mathcal{C}, \mathcal{V}, \alpha) : \mathcal{C}_{\tau_i} = (1-\alpha)\mathcal{C}_{\tau_i} + \alpha\mathcal{V}_{\kappa_i}, \quad (1)$$

where $\{\mathcal{C}_{\tau_i}\}_{i=1}^{n}$ and $\{\mathcal{V}_{\kappa_i}\}_{i=1}^{n}$ are sorted values of embedded feature $\mathcal{C}$ and $\mathcal{V}$ in ascending order. Here, n represents the number of elements in vector $\mathcal{C}$ and $\mathcal{V}$. Note that $\mathcal{C}$ is the embedded feature of the test sample x, which plays the role of carrying the appearance information. $\mathcal{V}$ is the embedded feature of a normal sample randomly sampled from the training data, carrying the style information. In this way, the semantic information of the test sample is preserved, while its style information is pulled closer to the training data's style.

To calculate the anomaly score, we use a similar method as in RD4AD. First, we calculate the anomaly maps of the test sample x at multi-level feature as follows:

$$\mathcal{M}^k = 1 - sim(\mathcal{P}^k, \mathcal{L}^k) \tag{2}$$

where $\mathcal{P}^k$ and $\mathcal{L}^k$ respectively are the embedded feature and the reconstructed feature of x at k^{th} encoding/decoding block in our method, and sim is a cosine similarity measure. Next, we increase the resolution of the feature maps $\mathcal{M}^k$ to match the input image size. To accomplish this, we use a bilinear up-sampling operation denoted as Ψ. We then accumulate these anomaly maps pixel-wise to generate a score map via:

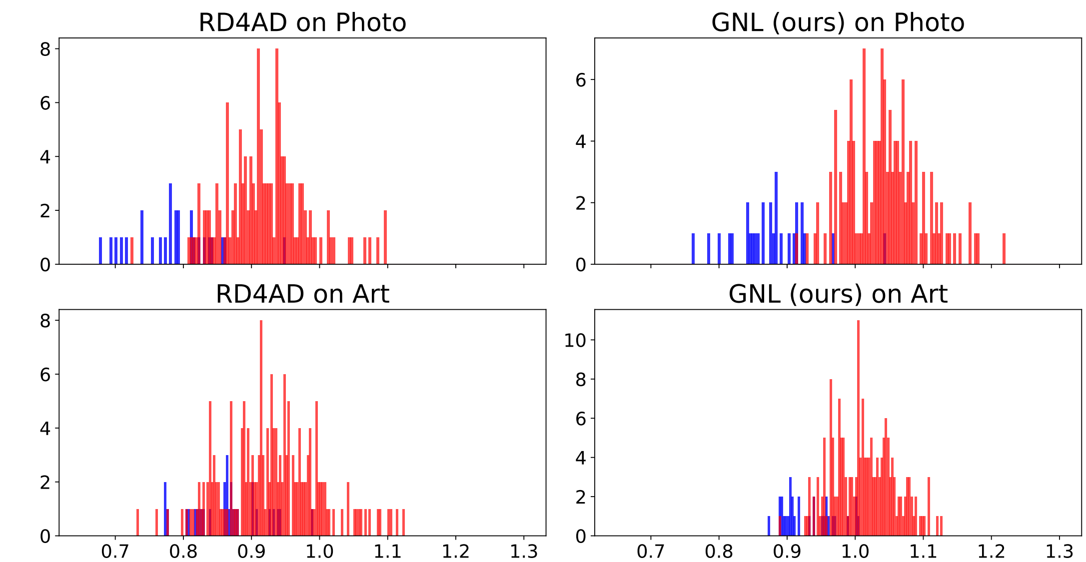
$$S_{AL} = \sum_{k=1}^{3} \Psi(\mathcal{M}^k) \tag{3}$$

Lastly, we choose the max value in S_{AL} as the anomaly score of x.

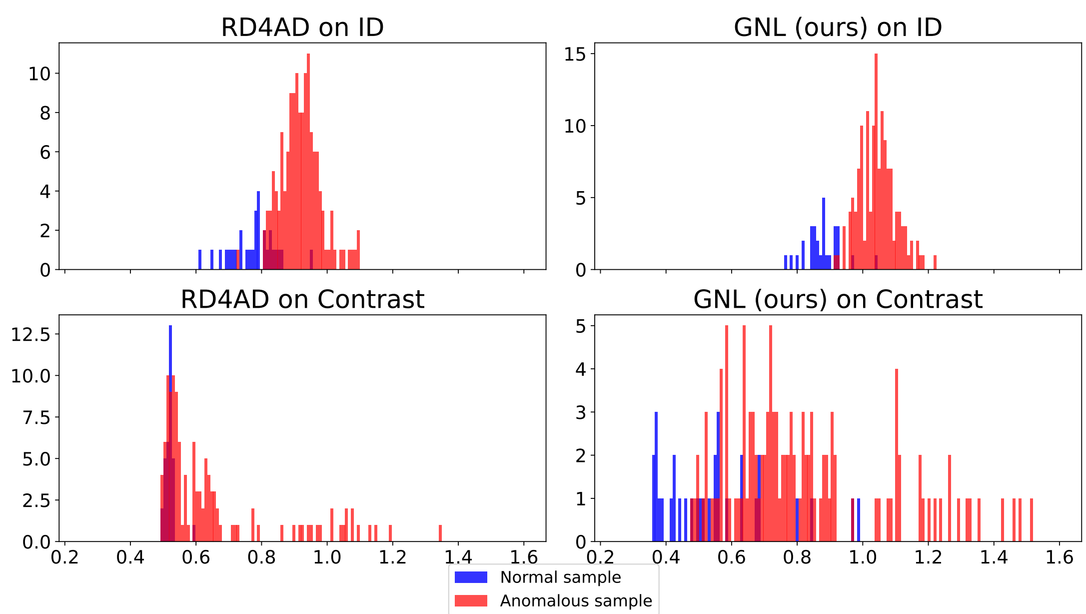
References

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 1

[2] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 1



(a) Anomaly scores on PACS ('elephant' as the normal class)



(b) Anomaly scores on MVTec (the 'Zipper' data)

Figure 1: Distribution of anomaly scores yielded by our method and RD4AD.

Class	0		1		2		3		4		5		6		7		8		9	
	ID	OOD	ID	OOD	ID	OOD	ID	OOD	ID	OOD	ID	OOD	ID	OOD	ID	OOD	ID	OOD	ID	OOD
Deep-SVDD	99.24	48.08	99.72	52.96	96.53	46.28	96.61	51.92	96.48	50.36	99.27	48.10	99.76	52.97	96.56	46.28	96.61	51.92	96.48	50.36
f-AnoGAN	99.41	54.04	99.85	56.24	96.52	49.60	95.08	52.97	96.81	50.71	**99.35**	54.04	**99.89**	56.31	96.36	49.60	95.08	52.97	96.81	50.71
KDAD	99.85	58.22	**99.88**	57.15	98.42	52.99	**99.06**	55.80	**98.38**	51.95	98.33	57.11	99.49	55.51	98.69	52.02	98.45	56.88	98.16	51.12
RD4AD	99.56	71.50	99.50	60.09	**99.11**	51.93	98.01	55.73	96.75	50.56	98.90	58.34	99.79	64.60	**99.21**	57.37	98.90	55.77	99.17	55.03
Augmix	99.78	70.83	98.64	62.84	97.66	53.28	98.41	58.29	97.51	54.25	97.12	62.15	98.83	62.33	98.32	61.44	97.38	55.91	98.98	54.74
Mixstyle	99.61	69.90	99.57	58.54	98.76	51.19	98.85	56.04	95.83	49.10	98.92	58.43	99.57	63.14	99.16	56.74	98.99	54.98	99.17	54.14
EFDM	98.09	69.60	99.36	59.33	98.39	50.74	98.86	55.53	95.79	49.96	98.81	58.30	99.59	62.92	99.09	56.66	**99.02**	55.09	**99.18**	54.19
Jigsaw	**99.90**	71.13	99.86	60.83	99.01	51.71	98.85	56.62	96.18	52.71	98.74	59.15	99.75	64.18	99.16	58.08	98.54	55.86	99.00	54.81
GNL (Ours)	99.40	**80.54**	99.55	**71.95**	96.52	**63.87**	95.92	**64.69**	95.05	**64.91**	94.48	**75.33**	97.93	**79.48**	97.46	**71.25**	94.69	**64.43**	98.13	**72.27**

Table 1: Full AUROC (%) results on MNIST/MNIST-M.

Dataset	Dog				Elephant				Giraffe			
Domain	Photo	Art	Cartoon	Sketch	Photo	Art	Cartoon	Sketch	Photo	Art	Cartoon	Sketch
Deep-SVDD	43.25	55.60	42.99	38.00	47.47	53.65	40.86	37.09	36.39	53.59	38.44	37.26
f-AnoGAN	46.30	54.06	42.90	42.90	67.36	44.09	**79.99**	34.11	64.96	51.13	47.30	51.79
KDAD	**76.15**	62.50	40.38	41.53	91.78	50.52	76.75	15.53	87.91	**55.67**	**65.53**	63.79
RD4AD	70.39	67.16	47.77	53.57	**92.07**	58.89	65.81	61.20	76.82	46.20	53.57	46.73
Augmix	70.36	64.33	47.08	53.80	83.36	58.83	66.31	67.16	63.75	48.38	51.44	45.70
Mixstyle	72.63	65.61	48.46	52.99	86.97	60.72	65.93	63.69	74.72	48.42	55.64	43.79
EFDM	71.81	67.06	46.95	57.05	85.46	60.80	67.12	64.61	77.64	47.21	58.27	41.96
Jigsaw	47.37	44.29	40.43	38.38	62.27	60.40	56.61	47.80	68.59	51.27	45.32	46.81
GNL (Ours)	76.13	**70.04**	**57.75**	**59.35**	90.86	**66.20**	74.83	**67.80**	**88.27**	53.80	54.21	**64.11**

Table 2: Full AUROC (%) results on PACS (Part I).

Dataset	Guitar				Horse				House				Person			
Domain	Photo	Art	Cartoon	Sketch	Photo	Art	Cartoon	Sketch	Photo	Art	Cartoon	Sketch	Photo	Art	Cartoon	Sketch
Deep-SVDD	41.79	55.20	44.47	39.51	43.07	53.39	39.24	38.19	38.89	52.69	39.92	44.92	35.25	49.83	42.70	41.40
f-AnoGAN	42.82	34.65	56.25	**96.94**	51.72	50.40	39.76	33.28	58.76	53.17	49.47	**94.21**	97.45	63.57	51.25	93.19
KDAD	77.19	53.79	82.26	67.13	**85.41**	51.99	51.69	43.31	**98.76**	91.12	65.53	64.54	**100.00**	**74.39**	**56.31**	**64.00**
RD4AD	76.62	59.35	76.71	49.48	64.15	59.18	46.93	53.24	93.52	76.29	79.92	61.65	96.85	60.39	51.66	59.53
Augmix	63.68	**60.15**	67.86	55.23	64.22	56.07	48.66	46.52	95.62	74.49	80.38	79.37	93.46	61.26	51.03	57.22
Mixstyle	68.56	59.98	77.22	47.19	55.57	54.88	50.70	51.91	94.97	74.85	76.81	64.21	94.19	62.03	51.72	60.43
EFDM	65.35	56.99	77.79	47.78	61.05	56.75	52.82	53.51	92.47	73.34	80.64	63.45	95.49	61.70	51.45	61.04
Jigsaw	55.63	51.59	61.96	95.72	55.45	50.44	42.57	40.13	70.28	52.29	81.26	88.34	75.71	57.54	48.69	77.90
GNL (Ours)	**85.37**	57.68	**82.53**	45.33	77.59	**60.20**	**63.64**	**69.71**	97.55	**91.14**	**88.79**	75.42	97.95	60.31	53.97	55.01

Table 3: Full AUROC (%) results on PACS (Part II).

Dataset	Carpet					Leather					Grid					Tile					Wood				
Corruption	ID	Brighness	Contrast	Blur	Noise	ID	Brighness	Contrast	Blur	Noise	ID	Brighness	Contrast	Blur	Noise	ID	Brighness	Contrast	Blur	Noise	ID	Brighness	Contrast	Blur	Noise
Deep-SVDD	54.74	33.03	53.77	44.86	49.96	64.12	35.16	38.07	63.73	39.08	84.13	91.23	74.02	85.55	76.19	74.86	41.88	45.67	70.53	68.25	89.21	75.96	54.21	89.65	76.75
f-AnoGAN	65.46	42.00	39.26	17.29	19.17	81.15	38.90	75.24	60.60	57.19	86.06	24.37	30.43	23.11	23.36	58.88	57.06	59.59	47.53	46.94	86.12	50.35	41.02	32.92	32.39
KDAD	76.59	74.70	52.68	77.44	79.44	94.29	95.41	83.22	98.99	95.55	53.38	51.54	30.74	47.65	52.77	91.70	93.81	64.86	90.45	91.87	89.24	82.49	28.51	85.23	81.17
RD4AD	98.75	98.30	96.35	98.60	98.18	**100.00**	**100.00**	96.31	**100.00**	**100.00**	**100.00**	**100.00**	97.63	99.92	98.19	99.09	99.46	99.22	99.37	99.76	**99.39**	**98.65**	98.66	**99.15**	**98.95**
Augmix	98.53	98.39	96.80	98.14	97.93	99.14	99.85	95.38	96.59	99.26	95.27	96.52	98.19	99.28	95.63	97.39	97.59	96.01	95.85	97.36	94.68	90.96	95.94	91.02	95.00
Mixstyle	98.68	98.30	97.00	98.58	98.45	**100.00**	**100.00**	96.55	**100.00**	**100.00**	99.58	99.28	98.36	**99.97**	**98.25**	99.25	**99.67**	99.43	99.52	99.67	99.27	98.36	97.78	98.95	98.92
EFDM	98.70	98.46	96.71	98.52	98.19	**100.00**	**100.00**	96.28	**100.00**	**100.00**	99.53	99.30	**98.55**	99.44	97.44	99.40	99.63	99.43	**99.65**	**99.77**	99.15	98.04	97.78	98.77	98.65
Jigsaw	96.91	95.02	95.21	97.33	94.60	93.77	95.70	82.86	98.30	92.32	80.20	81.70	74.94	79.50	78.89	80.52	71.76	73.51	72.01	72.13	90.38	87.31	78.33	83.95	86.81
GNL (Ours)	**99.48**	**99.23**	**99.20**	**99.52**	**99.11**	**100.00**	**100.00**	**99.93**	**100.00**	99.90	98.86	98.16	97.66	98.00	95.80	**99.59**	**99.67**	**99.76**	98.85	99.59	98.68	97.69	**98.77**	98.36	98.33
Dataset	Bottle					Hazelnut					Cable					Capsule					Pill				
Deep-SVDD	78.25	60.40	59.84	89.13	76.90	84.79	31.93	16.11	85.04	51.36	73.41	50.45	47.96	73.84	70.41	58.06	62.48	68.83	64.15	47.90	64.59	39.90	36.05	60.48	49.87
f-AnoGAN	93.25	49.37	45.08	27.70	27.78	64.96	50.91	58.76	35.87	38.59	55.36	77.98	54.80	54.46	55.98	59.59	43.12	68.65	48.22	62.70	80.22	52.87	52.31	67.69	63.27
KDAD	99.71	93.76	87.43	99.55	**99.02**	98.38	94.32	88.03	97.97	95.36	90.55	81.28	53.23	91.28	91.78	80.67	81.18	51.66	80.65	74.58	78.99	73.31	49.17	79.69	75.23
RD4AD	**100.00**	**100.00**	98.47	**100.00**	83.33	99.96	**100.00**	99.21	99.95	99.58	96.24	96.63	90.48	95.00	96.23	97.45	88.63	81.44	95.34	76.51	**96.90**	84.08	86.05	96.12	76.44
Augmix	99.26	99.29	98.52	98.78	91.15	93.10	97.15	99.56	94.17	99.31	87.21	84.70	87.35	87.72	88.12	97.50	**94.86**	86.97	93.62	**87.64**	95.67	90.15	91.79	94.45	**86.58**
Mixstyle	**100.00**	**100.00**	**99.53**	**100.00**	81.90	99.96	99.99	99.38	99.82	99.77	96.73	96.43	90.92	95.84	96.42	97.26	89.27	83.37	95.37	76.13	96.72	86.29	86.44	96.04	78.40
EFDM	99.95	**100.00**	99.26	99.97	77.80	**100.00**	99.96	99.37	99.93	99.64	96.42	96.16	92.35	95.71	96.95	**97.88**	91.04	83.73	**95.84**	76.48	96.85	88.22	88.31	**96.26**	78.37
Jigsaw	76.54	74.47	70.82	79.36	76.15	82.29	81.08	45.57	83.67	86.65	64.59	57.96	57.01	62.53	57.96	62.60	50.99	53.05	61.34	61.84	55.21	57.95	56.15	53.79	57.42
GNL (Ours)	99.76	99.71	99.36	99.87	97.83	**100.00**	**100.00**	**100.00**	**100.00**	**99.87**	**96.82**	**97.29**	**97.43**	**97.23**	**97.69**	93.30	91.53	**89.65**	89.79	79.46	96.63	**90.94**	**94.41**	95.40	84.22
Dataset	Transistor					MetalNut					Screw					Toothbrush					Zipper				
Deep-SVDD	70.79	65.42	61.58	70.96	68.29	56.06	52.10	73.12	63.83	45.45	35.23	95.08	3.18	24.37	35.66	96.94	46.94	67.22	96.11	79.44	64.47	45.80	51.42	50.11	51.16
f-AnoGAN	78.04	25.33	52.67	28.11	28.00	58.21	42.04	74.62	66.62	62.76	80.13	92.83	0.00	0.00	0.00	95.56	56.11	33.61	2.50	9.17	78.22	48.21	53.31	57.02	59.25
KDAD	88.64	86.61	64.89	90.94	87.14	81.35	86.87	79.36	80.26	84.12	73.10	89.42	77.43	56.80	35.36	92.59	85.56	55.00	93.06	91.76	93.34	86.91	94.18	92.64	**95.50**
RD4AD	96.34	96.17	90.58	94.36	93.97	**100.00**	**100.00**	99.43	**99.87**	92.86	98.43	97.73	95.65	97.82	84.08	99.26	89.17	97.69	99.44	98.58	97.76	98.70	84.61	97.97	55.39
Augmix	91.94	91.26	84.26	89.97	88.36	98.61	98.74	95.68	96.06	92.95	97.98	91.74	95.37	97.25	55.18	99.81	**96.67**	**100.00**	**100.00**	**100.00**	**98.22**	98.55	**95.79**	97.92	90.35
Mixstyle	95.53	95.74	88.82	93.75	93.85	**100.00**	**100.00**	99.41	**99.87**	92.83	**98.63**	**97.99**	95.44	**98.00**	85.09	98.98	88.80	**100.00**	99.91	80.56	98.18	**98.96**	84.25	**98.47**	53.59
EFDM	96.29	95.69	89.78	93.93	94.32	**100.00**	**100.00**	99.43	99.85	**93.53**	98.39	97.85	95.33	97.70	84.51	98.98	88.61	**100.00**	**100.00**	87.69	98.11	98.73	85.29	98.24	56.02
Jigsaw	69.50	68.50	64.47	66.26	69.60	60.58	65.46	65.45	65.33	51.91	53.83	68.66	64.15	60.24	61.15	81.76	85.56	68.70	80.74	73.80	60.96	58.22	67.95	63.93	67.73
GNL (Ours)	**97.47**	**97.00**	**95.06**	**97.58**	**96.69**	99.98	99.75	**99.90**	99.75	89.46	93.53	97.90	**96.47**	94.97	**90.11**	**100.00**	95.95	99.63	99.26	97.22	95.80	96.69	94.72	97.96	86.25

Table 4: Full AUROC (%) results on MVTec.