Supplementary Material: Towards Robot Anomaly Detection

Ken Miyamoto
miyamoto.ken@bc.mitsubishielectric.co.jp
Masato Tsuchiya
tsuchiya.masato@aj.mitsubishielectric.co.jp
Takashi Ota
ota.takashi@dx.mitsubishielectric.co.jp
Lee Teng-Yok
lee.teng-yok@ap.mitsubishielectric.co.jp

Information Technology R&D Center
Mitsubishi Electric Corporation
Kanagawa, Japan

This supplementary material describes evaluation results on another dataset. The dataset is synthesized by the same way as Section 4.1. Sequence5 in RISEdb [15] is used here as normal images and backgrounds of anomaly images. According to an anomaly image, an object in Pix3D [16] is randomly superposed on an image in sequence5 of RISEdb.: Section A gives performance comparison between PaDiM [6] and PatchCore [14]. Section B describes comparison between PatchCore-S, PathCore-L, and IFSC+PatchCore-S on the above dataset.

A Optimal Anomaly Detection

For clarifying an optimal algorithm under moving camera condition, we compare performance of PaDiM and PatchCore on another dataset. The result is shown in Table A. Image level metrics of PatchCore is higher than PaDiM, though pixel level metrics of PatchCore are inferior to PaDiM. We notice that an optimal model may be different by image feature variation in a dataset. We need further study to delve into the model selection of anomaly detection.

B Local Modeling by Image Feature Spatial Clustering

We show performance comparison between global and local modeling. A global model trains all normal images in sequence5 of RISEdb. PathCore-S and PatchCore-L are global models, as well as description in Section 4.4. Conversely, a local model trains a part of normal images in sequence5 of RISEdb, which are clustered by IFSC shown in Section 3.2. Only IFSC+PatchCore-S is local modeling.

The evaluation result is shown in Table B. Most metrics of IFSC+PatchCore-S are superior to PathCore-L, though IFSC+PatchCore-S consumes less memory than PatchCore-L. We consider that local modeling evolves better performance because of training image and pixel thresholds in addition to local memory banks.

Local Datasets	Training Frames	Test Frames	Models	AUROC (image)	F1 Score (image)	AUROC (pixel)	F1 Score (pixel)
(0)	445	490	PaDiM	0.807	0.724	**0.989**	**0.415**
			PatchCore	**0.828**	**0.759**	0.983	0.391
(1)	176	158	PaDiM	0.822	0.738	0.991	**0.453**
			PatchCore	**0.875**	**0.788**	**0.992**	0.421
(2)	225	208	PaDiM	0.699	**0.701**	0.966	**0.308**
			PatchCore	**0.733**	0.688	**0.969**	0.273
(3)	86	114	PaDiM	0.881	0.800	0.995	**0.460**
			PatchCore	**0.938**	**0.857**	**0.996**	0.436
(4)	102	98	PaDiM	**0.949**	0.870	**0.996**	**0.439**
			PatchCore	0.948	**0.872**	0.995	0.381
(5)	132	136	PaDiM	0.769	0.667	0.983	0.335
			PatchCore	**0.800**	0.667	**0.983**	**0.402**
(6)	45	52	PaDiM	**0.932**	0.857	**0.997**	**0.562**
			PatchCore	0.920	**0.875**	0.994	0.472
Total	1211	1256	PaDiM	0.810	0.739	**0.986**	**0.405**
			PatchCore	**0.839**	**0.763**	0.984	0.383

Table A: Comparison between PaDiM and PatchCore on dataset synthesized from sequence5 in RISEdb. "Total" describes weighted average of metrics. Test frames are used as the weight. Each local dataset is created by IFSC shown in Section 3.2

.

Datasets	Models	AUROC (image)	F1 Score (image)	AUROC (pixel)	F1 Score (pixel)
(0)	PatchCore-S	0.580	0.667	0.966	0.164
	PatchCoe-L	0.727	0.668	0.984	0.377
	IFSC+PatchCore-S	**0.828**	**0.759**	**0.983**	**0.391**
(1)	PatchCore-S	0.560	0.667	0.963	0.292
	PatchCoe-L	0.825	0.670	**0.992**	0.418
	IFSC+PatchCore-S	**0.870**	**0.779**	0.991	**0.421**
(2)	PatchCore-S	0.553	0.667	0.920	0.168
	PatchCoe-L	0.727	0.665	**0.972**	0.260
	IFSC+PatchCore-S	**0.747**	**0.710**	0.969	**0.318**
(3)	PatchCore-S	0.651	0.667	0.979	0.327
	PatchCoe-L	0.921	0.667	0.996	0.408
	IFSC+PatchCore-S	**0.959**	**0.867**	**0.996**	**0.435**
(4)	PatchCore-S	0.709	0.667	0.979	0.329
	PatchCoe-L	0.924	0.671	0.995	0.347
	IFSC+PatchCore-S	**0.932**	**0.884**	**0.995**	**0.375**
(5)	PatchCore-S	0.560	0.667	0.916	0.116
	PatchCoe-L	0.731	0.667	**0.982**	0.289
	IFSC+PatchCore-S	**0.800**	**0.673**	0.982	**0.405**
(6)	PatchCore-S	0.518	0.667	0.974	0.224
	PatchCoe-L	0.872	0.667	**0.995**	**0.520**
	IFSC+PatchCore-S	**0.925**	**0.880**	0.994	0.479
Total	PatchCore-S	0.582	0.667	0.955	0.206
	PatchCoe-L	0.779	0.668	**0.985**	0.357
	IFSC+PatchCore-S	**0.841**	**0.769**	0.984	**0.391**

Table B: Comparison between conventional anomaly detection by global modeling (*i.e*. PatchCore-S, PatchCore-L) and local modeling with DCS (*i.e*. IFSC+PatchCore-S). PatchCore-S limits embedding vectors in memory bank under 31360. PatchCore-L stores 125750 embedding vectors in memory bank. IFSC+PatchCore-L creates local models under DCS framework. A memory bank and thresholds are modeled locally. Embedding vectors in each memory bank are limited under 31360. These algorithms are evaluated on synthesized dataset, which combines sequence5 in RISEdb and Pix3D.