

A. Appendix

A.1. Dataset

A.1.1. Defects and Variations

Various defects captured in RobustAD are captured in the table 1 We show the visualization of the defects in PCB subset in figure 2, Metal Parts subset in figure 1 and Pile of Packets in figure 3.

The detailed statistics on number of images per variation are captured in Table 6 , Table 7, Table 8 for Metal Parts, Pile of Packets and PCB datasets respectively. We refer to Source Domain (SD) and bg1 in metal parts refers to back ground variation1 and bg1 in metal parts refers to back ground variation2

A.1.2. Licenses of the Datasets and code

We have used the following datasets:

1. MVTec AD with CC BY-NC-SA 4.0 license

We have used the following code repositories

1. Anomalib <https://github.com/openvinotoolkit/anomalib> with Apache-2.0 license
2. <https://github.com/mala-lab/ADShift> license not provided
3. <https://github.com/zqhang/AnomalyCLIP> license not provided

A.2. Additional Results

The official GNL implemtation uses a image size of 256x256. We addtionally provide results here for GNL evaluated at 512x512 in the table 2.

GNL(512) has an improved performance over GNL for metal parts and PCB datasets where as it results in drop in accuracy for pile of packets this likely due to the very small defects metal parts and PCB are enhanced by increased image size. GNL(512) out performs PatchCore for PCB but does not improve over PatchCore for pile of Packets and PCB datasets.

We show the results in terms of AUPR in Table 3. We show the pixel level AUROC and and AUPR metrics in Table 4 and Table 5 respectively. We see similar trends in AUPR, pixel AUROC metrics as well. The pixel AUPR metrics is significantly better for GNL at both resolutions this likely due to the normalization done during Test Time Adaption results in enhancing the contrast required for better localization. Please note that MVTec AD-C is always evaluated at 256. Hence we do not include it for GNL(512) evaluation. Note that the Pile of Packets dataset is only defined for anomaly classification and not anomaly localization. Hence no pixel level metrics are provided for this dataset.

We also add the visualization of defect predictions by PatchCore model on Metal Parts sub-dataset in various tar-

get domains in figure 5. Note that we were not able to obtain the picture of same piece of metal plate in all domains. Hence the metal plates for bg1 and bg2 domains is different and have different ground truth masks. We can see scale changes result in most spurious detections in the background. A simple change in the background also breaks the model.

Table 1. Defect Types

Object	Defects
PCB	Missing components, scratches, soldering melts
Piled bags	Under load, overload, flipped
Metal Parts	Dents, chipping, porosity

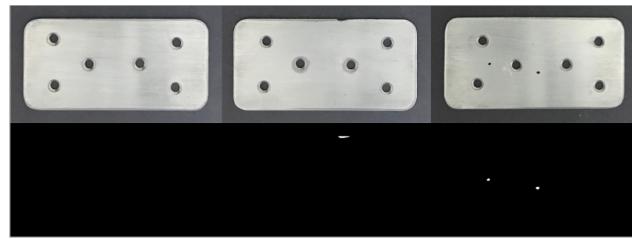


Figure 1. Defects in Metal Parts sub-dataset - top row(left to right) :Normal, Chipping, Porosity; bottom row(left to right) : Ground truth masks

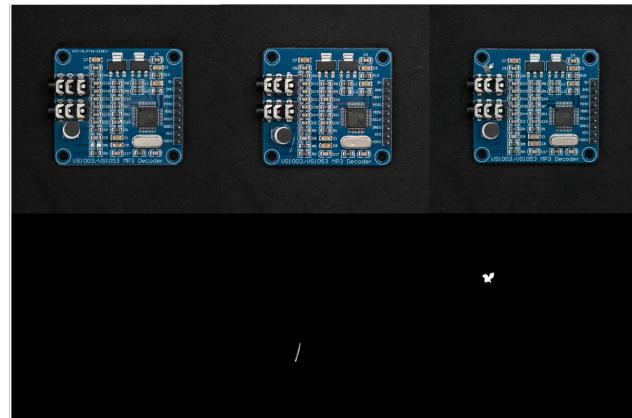


Figure 2. Defects in PCB sub-dataset - top row(left to right) : Normal, scratch, Soldering Melt; bottom row(left to right) : Ground truth masks



Figure 3. Defects in Pile of Packets sub-dataset - top row(left to right) :Normal, overload, under-load;



Figure 4. Left: Images of PCB captured at different times of the day. Right: Images of the pile of packets in various backgrounds with background artifacts

Table 2. Robust AD Benchmark on unsupervised models as measured in AUROC.

Object Class	Domain	PADIM	PatchCore	RD4AD	GNL	GNL(512)
Metal Parts	SD	85.67 ± 2.03	87.92 ± 0.46	89.59 ± 1.35	89.54 ± 0.55	92.71± 0.88
	Lighting	96.31 ± 0.85	96.59 ± 0.09	96.77 ± 0.55	96.63 ± 0.15	97.61±0.30
	Position	63.28 ± 6.61	89.46 ± 0.32	87.43 ± 2.83	81.67 ± 3.07	87.77± 0.91
	Rotation	82.28 ± 6.41	88.27 ± 0.18	90.47 ± 0.84	86.95 ± 1.79	92.02 ± 0.61
	Scale	41.01 ± 3.89	80.88 ± 0.17	80.05 ± 2.95	61.64 ± 3.53	86.64± 1.27
	BG1	69.43 ± 9.62	74.87 ± 0.6	71.99 ± 5.98	90.31 ± 0.93	82.84± 1.31
	BG2	68.15 ± 11.97	83.5 ± 1.16	67.16 ± 2.8	77.41 ± 2.85	71.59± 1.91
(mean, std)		(72.30, 17.99)	(85.93, 6.94)	(83.35, 10.71)	(83.45, 11.41)	(87.31, 8.42)
(SD, ARD)		(85.67, -17.37)	(87.92, -4.09)	(89.59, -8.63)	(89.54, -8.41)	(92.71,-7.11)
Pile of Packets	SD	98.80 ± 0.70	99.98 ± 0.01	93.01 ± 0.93	99.91 ± 0.1	92.87± 1.45
	Lighting	65.70 ± 10.66	95.89 ± 0.15	85.41 ± 1.15	96.13 ± 1.82	89.01 ± 2.19
	Background	47.67 ± 1.78	51.23 ± 0.22	50.34 ± 0.50	51.63 ± 0.70	50.34 ± 1.43
	Rot + pos	53.41 ± 3.60	61.74 ± 0.75	74.85 ± 2.00	67.74 ± 2.75	49.68± 2.76
	Scale	51.34 ± 4.79	75.34 ± 0.24	55.13 ± 3.03	66.45 ± 4.12	59.05± 4.74
	Shadows	77.36 ± 12.68	84.31 ± 0.58	76.33 ± 1.26	95.26 ± 1.24	81.76±4.73
(mean, std)		(65.72, 19.56)	(78.08, 19.14)	(72.51, 16.73)	(79.52, 20.13)	(70.45, 19.70)
(SD, ARD)		(98.80, -39.7)	(99.98, -26.28)	(93.01, -24.60)	(99.91, -24.47)	(92.87, -26.90)
PCB	SD	99.99 ± 0.02	96.32 ± 0.14	99.74 ± 0.11	99.98 ± 0.01	99.92±0.09
	Lighting	10.41 ± 7.4	73.93 ± 0.65	57.36 ± 10.18	20.24 ± 6.62	71.50± 1.07
	AWB	6.96 ± 7.07	81.59 ± 0.49	66.58 ± 15.11	11.78 ± 3.93	83.11±4.00
	Rotation	31.79 ± 12.98	68.53 ± 0.87	54.65 ± 12.39	33.18 ± 2.80	53.47± 3.31
	Position	40.88 ± 9.32	67.18 ± 0.49	47.98 ± 9.16	42.49 ± 11.82	47.54± 3.45
	Shadows	60.68 ± 10.13	46.69 ± 0.50	70.56 ± 2.6	31.54 ± 7.35	62.62± 0.66
(mean, std)		(41.79, 34.76)	(72.38, 16.5)	(66.15, 18.37)	(39.87, 31.33)	(69.69., 19.50)
(SD, ARD)		(99.99, -69.85)	(96.32, -28.74)	(99.74, -40.32)	(99.98, -72.14)	(99.92, -36.27)

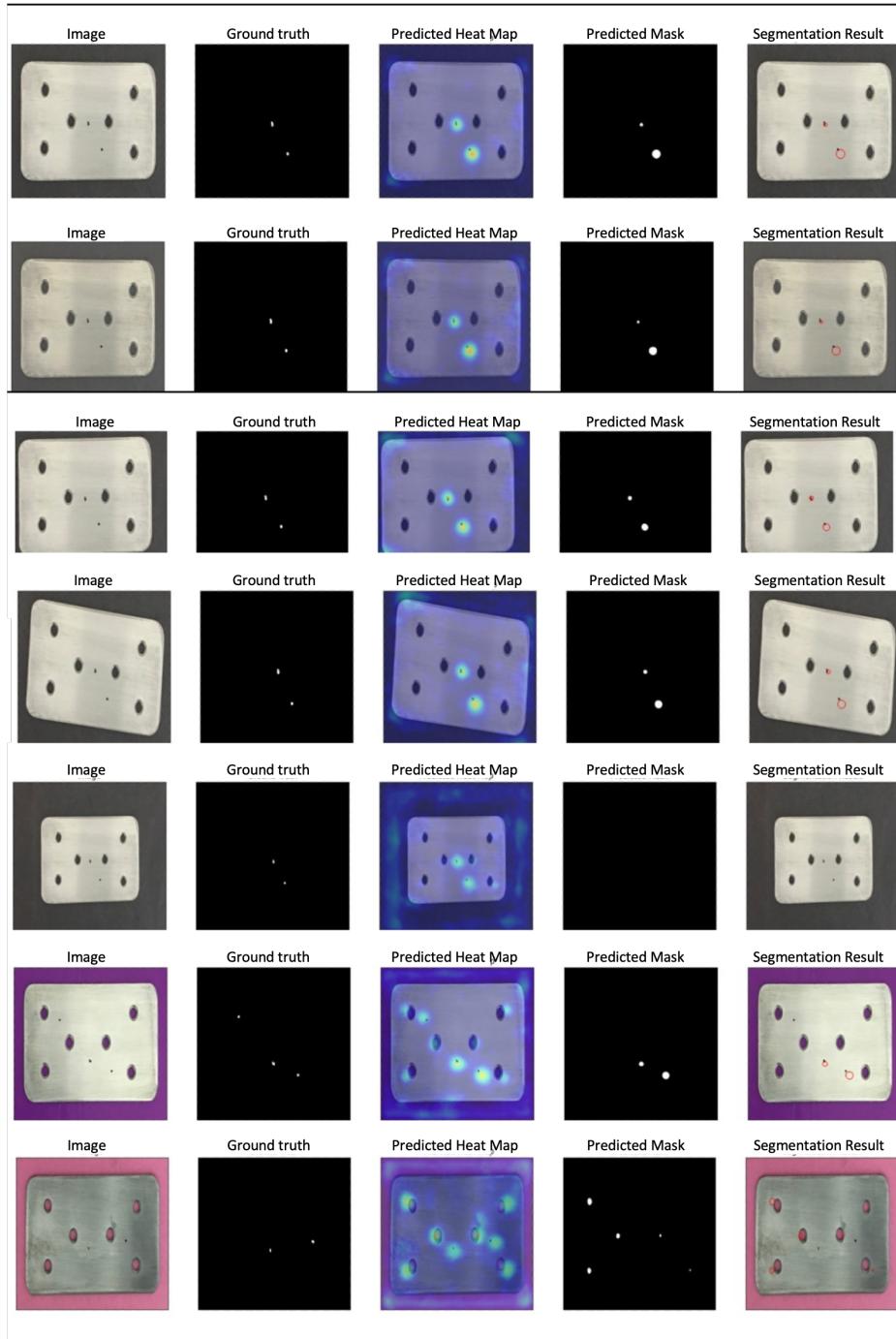


Figure 5. Visualization of defect predictions by PatchCore model on Metal Parts sub-dataset in various target domains - top to bottom :Source Domain, Lighting, Position, Rotation, Scale, background1, Background2; left to right : Image, Ground truth masks, Prediction overlaid on image, Predicted mask. Note that different instances of the metal plate are showed in background domains compared to other domains.

Table 3. Robust AD Benchmark on unsupervised models as measured in AUPR(%).

Object Class	Domain	PADIM	PatchCore	RD4AD	GNL	GNL(512)
Metal Parts	SD	78.52 ± 5.32	82.24 ± 0.34	81.98 ± 2.75	84.54 ± 0.46	88.88 ± 1.01
	Lighting	94.49 ± 1.17	94.10 ± 0.19	94.72 ± 0.42	94.92 ± 0.23	96.22 ± 0.63
	Position	61.75 ± 6.47	87.68 ± 1.22	83.26 ± 4.05	78.39 ± 4.86	87.86 ± 1.05
	Rotation	76.22 ± 10.30	87.20 ± 0.28	88.85 ± 1.64	83.48 ± 3.75	91.07 ± 0.86
	Scale	50.91 ± 1.94	74.89 ± 2.10	68.04 ± 5.20	52.19 ± 5.30	84.51 ± 1.42
	BG1	66.70 ± 15.38	76.13 ± 6.51	59.53 ± 6.61	88.73 ± 0.76	78.57 ± 2.35
	BG2	64.04 ± 15.98	75.51 ± 0.64	62.34 ± 3.70	75.20 ± 4.12	65.92 ± 4.59
(mean, std)		(70.38, 14.08)	(82.54, 7.43)	(76.96, 13.66)	(79.64, 13.72)	(84.72, 9.92)
(SD, ARD)		(78.52, -12.16)	(82.54, -3.37)	(81.98, -9.34)	(84.54, -8.15)	(88.88, -6.44)
Pile of Packets	SD	97.85 ± 1.21	99.57 ± 0.01	91.60 ± 0.88	99.86 ± 0.18	91.35 ± 1.36
	Lighting	65.04 ± 14.14	94.36 ± 0.18	87.19 ± 1.02	95.95 ± 1.68	88.28 ± 2.89
	Background	51.62 ± 3.77	47.61 ± 1.20	40.98 ± 0.74	42.64 ± 0.43	41.40 ± 1.14
	Rot + pos	53.62 ± 10.14	52.78 ± 1.34	74.39 ± 3.84	61.62 ± 3.69	44.85 ± 2.84
	Scale	51.54 ± 7.17	57.23 ± 0.50	55.10 ± 3.54	59.04 ± 9.10	52.54 ± 5.59
	Shadows	76.89 ± 10.23	71.91 ± 0.85	74.94 ± 1.44	95.02 ± 1.46	82.00 ± 6.16
(mean, std)		(66.09, 18.46)	(70.58, 22.05)	(70.70, 19.32)	(75.69, 24.23)	(66.74, 22.91)
(SD, ARD)		(97.85, -38.10)	(99.57, -34.80)	(91.60, -25.07)	(99.86, -29.01)	(91.35, -29.54)
PCB	SD	99.58 ± 0.03	92.55 ± 0.30	99.24 ± 0.18	99.97 ± 0.03	99.87 ± 0.19
	Lighting	28.53 ± 2.00	71.54 ± 0.60	58.31 ± 12.75	30.98 ± 4.49	66.99 ± 1.05
	AWB	20.42 ± 0.86	73.73 ± 0.73	57.03 ± 17.97	20.78 ± 1.29	77.54 ± 3.01
	Rotation	50.57 ± 14.73	69.51 ± 0.80	48.70 ± 10.47	34.56 ± 2.24	47.39 ± 2.87
	Position	52.88 ± 15.64	66.69 ± 0.65	47.92 ± 9.55	45.52 ± 10.36	41.29 ± 3.31
	Shadows	56.37 ± 10.28	46.94 ± 1.12	65.85 ± 2.69	35.87 ± 2.84	53.92 ± 0.20
(mean, std)		(51.39, 27.68)	(70.16, 14.61)	(62.84, 19.03)	(44.61, 28.27)	(64.50, 21.78)
(SD, ARD)		(99.58, -57.82)	(92.55, -26.87)	(99.24, -43.68)	(99.97, -66.43)	(99.87, -42.44)
MVTec AD-C	SD	94.18 ± 1.15	99.51 ± 0.05	99.13 ± 0.11	99.06 ± 1.12	-
	Brightness	93.02 ± 1.50	98.93 ± 0.09	98.25 ± 0.27	98.77 ± 1.03	-
	Contrast	89.08 ± 3.22	98.92 ± 0.09	97.79 ± 0.27	98.52 ± 1.92	-
	Defocus Blur	93.26 ± 1.63	98.85 ± 0.05	99.04 ± 0.13	98.86 ± 0.13	-
	Gaussian Noise	90.08 ± 2.00	97.67 ± 0.13	95.85 ± 0.36	97.73 ± 2.2	-
(mean, std)		(91.92, 2.21)	(98.76, 0.78)	(98.01, 1.33)	(98.59, 0.52)	-
(SD, ARD)		(94.18, -2.82)	(99.51, -0.92)	(99.13, -1.40)	(99.06, -0.59)	-

Table 4. Robust AD Benchmark on unsupervised models as measured in pixel AUROC(%).

Object Class	Domain	PADIM	PatchCore	RD4AD	GNL	GNL(512)
Metal Parts	SD	91.95 ± 9.49	96.70 ± 0.20	98.30 ± 0.21	97.27± 0.23	97.83± 0.31
	Lighting	92.94 ± 8.69	97.95± 0.15	98.95± 0.09	97.24± 0.11	98.33± 0.17
	Position	92.30 ± 3.18	98.03 ± 0.08	98.59 ± 0.32	93.17±0.45	98.40 ± 0.17
	Rotation	91.37 ± 7.99	97.94 ± 0.05	98.99 ± 0.10	95.52 ± 0.39	97.88± 0.35
	Scale	86.80 ± 2.94	97.60 ± 0.06	98.95± 0.33	92.28 ±0.63	98.50± 0.27
	BG1	90.89 ± 8.07	97.98 ± 0.08	96.44 ± 1.19	99.64 ±0.03	99.31 ± 0.05
	BG2	90.71 ± 6.88	97.80 ± 0.12	97.81 ± 0.23	99.24 ± 0.11	98.78± 0.11
	(mean, std)	(91.00,2.01)	(97.71, 0.47)	(98.29, 0.92)	(96.34, 2.83)	(98.43, 0.51)
(SD, ARD)		(91.95, -1.34)	(96.70, 0)	(98.30, -0.39)	(97.27, -1.81)	(97.83, 0)
PCB	SD	94.40 ± 0.45	97.42 ± 0.06	98.39 ± 0.14	88.61 ± 0.46	96.69±0.45
	Lighting	83.83 ± 2.18	93.82 ± 0.14	96.43 ± 0.35	89.55 ±0.35	96.55± 0.39
	AWB	82.10 ± 2.98	95.07 ± 0.08	97.37 ± 0.22	90.04 ± 0.67	97.34± 0.26
	Rotation	81.58± 1.98	96.74 ± 0.03	96.23 ± 0.30	85.27±0.42	96.03 ± 0.45
	Position	78.74 ± 2.23	89.94 ± 0.26	95.05 ± 0.40	81.21± 0.62	96.75 ± 0.33
	Shadows	86.67 ± 1.97	93.34 ± 0.06	96.97 ± 0.24	84.53± 0.67	95.54± 0.39
(mean, std)		(84.55, 5.49)	(94.39, 2.70)	(96.74, 1.13)	(86.53, 3.45)	(96.48, 0.62)
(SD, ARD)		(94.40, -11.82)	(97.42, -3.64)	(98.39, -1.98)	(88.61, -2.97)	(96.69, -0.39)
MVTec AD-C	SD	96.32 ± 0.36	97.20 ± 0.04	97.47± 0.10	97.95 ± 1.92	-
	Brightness	90.99 ± 1.44	94.56±0.04	92.80±0.53	97.21± 2.29	-
	Contrast	81.24 ± 3.31	96.24± 0.05	95.79± 0.13	97.30± 2.24	-
	Defocus Blur	94.89 ± 0.64	96.78± 0.04	97.01± 0.08	97.60 ± 1.94	-
	Gaussian Noise	78.95 ±	94.30± 0.07	92.76 ± 0.24	92.51± 8.52	-
	(mean, std)	(88.48, 7.94)	(95.82, 1.31)	(95.17, 2.26)	(96.51, 2.26)	-
(SD, ARD)		(96.32, -9.80)	(97.20, -1.73)	(97.47, -2.88)	(97.95, -1.80)	-

Table 5. Robust AD Benchmark on unsupervised models as measured in pixel AUPR(%).

Object Class	Domain	PADIM	PatchCore	RD4AD	GNL	GNL(512)
Metal Parts	SD	9.80 ± 1.88	12.72 ± 0.40	10.62 ± 2.64	43.92 ± 1.20	42.03 ± 1.63
	Lighting	11.17 ± 1.22	14.83 ± 0.62	21.40 ± 1.69	41.10 ± 0.77	45.17 ± 2.04
	Position	1.63 ± 2.51	16.41 ± 0.85	15.79 ± 4.84	31.67 ± 2.84	39.34 ± 2.14
	Rotation	4.25 ± 3.48	14.85 ± 0.69	16.29 ± 2.81	28.97 ± 3.01	43.98 ± 1.98
	Scale	0.23 ± 0.20	9.40 ± 0.76	8.77 ± 3.89	24.41 ± 2.95	32.15 ± 2.07
	BG1	7.38 ± 4.37	12.71 ± 3.53	7.29 ± 2.02	51.74 ± 2.59	48.05 ± 2.11
	BG2	4.28 ± 4.46	8.33 ± 0.55	4.51 ± 0.90	42.46 ± 1.97	28.83 ± 1.40
	(mean, std)	(5.53, 4.08)	(12.75, 2.97)	(12.10, 5.94)	(37.75, 9.66)	(39.93, 7.05)
PCB	(SD, ARD)	(9.80, -5.20)	(12.72, -1.29)	(10.62, -1.88)	(43.92, -8.5)	(42.03, -4.3)
	SD	0.27 ± 0.06	5.82 ± 0.10	5.15 ± 1.32	3.88 ± 1.16	11.83 ± 0.71
	Lighting	0.17 ± 0.02	33.10 ± 0.27	12.56 ± 5.45	6.31 ± 1.11	20.45 ± 0.91
	AWB	0.07 ± 0.01	10.34 ± 0.23	5.76 ± 2.90	5.79 ± 1.36	17.30 ± 2.33
	Rotation	0.15 ± 0.02	36.27 ± 0.06	3.71 ± 1.11	1.95 ± 0.27	9.57 ± 1.33
	Position	0.11 ± 0.02	28.32 ± 0.13	3.76 ± 2.48	0.84 ± 0.06	16.22 ± 1.97
	Shadows	0.26 ± 0.05	0.86 ± 0.01	5.38 ± 2.08	4.46 ± 0.60	12.93 ± 1.37
	(mean, std)	(0.17, 0.08)	(19.12, 15.24)	(6.05, 3.30)	(3.87, 2.14)	(14.72, 4.00)
MVTec AD-C	(SD, ARD)	(0.27, -0.12)	(5.82, -0.99)	(5.15, -0.57)	(3.88, -0.99)	(11.83, -0.45)
	SD	33.87 ± 2.53	51.55 ± 0.13	50.54 ± 0.53	71.24 ± 8.24	-
	Brightness	33.62 ± 2.62	53.49 ± 0.13	49.63 ± 0.54	68.95 ± 10.74	-
	Contrast	21.68 ± 3.63	54.32 ± 0.17	46.59 ± 0.66	67.80 ± 10.00	-
	Defocus Blur	38.17 ± 2.58	54.74 ± 0.12	52.76 ± 0.61	69.26 ± 9.00	-
	Gaussian Noise	21.89 ± 2.55	48.09 ± 0.20	40.25 ± 0.62	60.78 ± 16.7	-
	(mean, std)	(29.85, 7.58)	(52.44, 2.72)	(47.96, 4.84)	(67.61, 4.01)	-
	(SD, ARD)	(33.87, -6.10)	(51.55, -0.86)	(50.54, -3.79)	(71.24, -4.54)	-

Table 6. Statistics of the Metal Parts dataset

Class	Perturbation	Domain	Total Images
Normal	None	source	300
Anomaly	None	source	100
Normal	Lighting	Target1	120
Anomaly	Lighting	Target1	80
Normal	Position	Target2	120
Anomaly	Position	Target2	80
Normal	Rotation	Target3	120
Anomaly	Rotation	Target3	80
Normal	Scale	Target4	120
Anomaly	Scale	Target4	80
Normal	background1	Target5	60
Anomaly	background1	Target5	40
Normal	background2	Target6	60
Anomaly	background2	Target6	40
Total			1400

Table 7. Statistics of the Pile of packets dataset

Class	Perturbation	Domain	Total Images
Normal	None	source	300
Anomaly	None	source	150
Normal	Lighting	Target1	120
Anomaly	Lighting	Target1	90
Normal	Background, Box Color	Target2	240
Anomaly	Background, Box Color	Target2	180
Normal	Position, Rotation	Target3	120
Anomaly	Position, Rotation	Target3	90
Normal	Scale	Target4	120
Anomaly	Scale	Target4	90
Normal	Shadow	Target4	120
Anomaly	Shadow	Target4	90
Total			1710

Table 8. Statistics of the PCB dataset

Image Class	Perturbation Type	Domain	Total Images
Normal	None	source	300
Anomaly	None	source	150
Normal	Lighting	Target1	120
Anomaly	Lighting	Target1	90
Normal	AWB	Target2	120
Anomaly	AWB	Target2	60
Normal	Rotation	Target3	80
Anomaly	Rotation	Target3	60
Normal	Position	Target4	80
Anomaly	Position	Target4	60
Normal	Shadows	Target5	80
Anomaly	Shadows	Target5	60
Total			1260

	Train(Normal/Anomaly)	Val	Test	Target1	Target2	Target3	Target4	Target5	Target6
PCB	144/60	36/15	120/75	120/90	120/60	80/60	80/60	80/60	NA
Metal Parts	144/41	36/10	120/50	120/80	120/80	120/80	120/80	60/40	60/40
Piled Bags	148/56	32/19	120/75	120/90	240/180	120/90	120/90	120/90	NA

Table 9. Dataset distribution across all sub-domains

Set	Perturbation Type	Perturbation Value	Class	Count
Train	None	None	Normal	180
Train	None	None	Missing Component	25
Train	None	None	Scratch	25
Train	None	None	Soldering Melt	25
Test	None	None	Normal	120
Test	None	None	Missing Component	25
Test	None	None	Scratch	25
Test	None	None	Soldering Melt	25
			Subtotal	450
Target 1	Lighting	Morning/Afternoon/Evening	Normal	120
Target 1	Lighting	Morning/Afternoon/Evening	Missing Component	30
Target 1	Lighting	Morning/Afternoon/Evening	Scratch	30
Target 1	Lighting	Morning/Afternoon/Evening	Soldering Melt	30
			Subtotal	210
Target 2	Position	(-10%, 10%)	Normal	80
Target 2	Position	(-10%, 10%)	Missing Component	20
Target 2	Position	(-10%, 10%)	Scratch	20
Target 2	Position	(-10%, 10%)	Soldering Melt	20
			Subtotal	140
Target 3	Rotation	(-5%,5%)	Normal	80
Target 3	Rotation	-(5%,5%)	Missing Component	20
Target 3	Rotation	(-5%,5%)	Scratch	20
Target 3	Rotation	(-5%,5%)	Soldering Melt	20
			Subtotal	140
Target 4	Shadow	Light Alternate Position	Normal	80
Target 4	Shadow	Light Alternate Position	Missing Component	20
Target 4	Shadow	Light Alternate Position	Scratch	20
Target 4	Shadow	Light Alternate Position	Soldering Melt	20
			Subtotal	140
Target 5	White Balancing	Morning/Afternoon/Evening + AWB	Normal	120
Target 5	White Balancing	Morning/Afternoon/Evening + AWB	Missing Component	20
Target 5	White Balancing	Morning/Afternoon/Evening + AWB	Scratch	20
Target 5	White Balancing	Morning/Afternoon/Evening + AWB	Soldering Melt	20
			Subtotal	180
			Total	1260

Table 10. PCB dataset details