Calibrated Out-of-Distribution Detection with a Generic Representation Supplementary Material

Tomáš Vojíř*†, Jan Šochman†, Rahaf Aljundi‡, and Jiří Matas†

†CMP Visual Recognition Group, FEE, Czech Technical University in Prague
‡Toyota Motor Europe, Brussels, Belgium

Here we present the extended version of the Table 1 from the main paper for completeness.

			FPR at 95% TPR ↓ / AUROC ↑								Average
arch	pre-trained	classif	SVHN	MNIST	Textures	Places365	CIFAR-100	iNaturalist	TIN	LSUN	
Resnet50	ImageNet	LP	10.23 / 97.90	11.86 / 97.71	78.24 / 83.32	60.51 / 86.03	59.62 / 87.28	97.47 / 70.15	63.95 / 83.37	62.85 / 87.05	55.59 / 86.60
		NM	97.18 / 63.18	69.27 / 92.77	0.00 / 99.96	0.52 / 99.73	77.02 / 73.01	0.01 / 99.88	47.50 / 90.31	0.11 / 99.82	36.45 / 89.83
	CLIP	LP	10.00 / 98.21	13.55 / 97.35	75.98 / 76.83	76.92 / 76.24	62.70 / 85.53	98.80 / 56.03	72.95 / 78.31	64.58 / 86.76	59.44 / 81.91
		NM	46.26 / 92.75	4.77 / 97.77	0.00 / 99.95	0.00 / 99.99	81.30 / 68.87	0.01 / 99.99	8.24 / 98.29	0.00 / 100.00	17.57 / 94.70
VIT-B/16	ImageNet	LP	18.46 / 97.09	4.42 / 98.57	23.94 / 96.09	17.95 / 95.84	26.19 / 94.72	26.15 / 95.57	21.94 / 94.39	8.07 / 98.17	18.39 / 96.31
		NM	43.93 / 93.51	1.10 / 99.51	9.75 / 97.87	9.88 / 97.90	52.61 / 87.28	2.91 / 99.43	40.09 / 89.64	7.56 / 98.82	20.98 / 95.50
	CLIP	LP	1.57 / 99.08	0.02 / 99.85	18.55 / 96.37	57.12 / 82.90	25.70 / 94.77	30.45 / 94.97	38.19 / 87.69	33.09 / 93.86	25.59 / 93.69
		NM	0.07 / 99.74	3.09 / 97.69	0.00 / 100.00	0.00 / 100.00	63.59 / 84.56	0.00 / 100.00	0.92 / 99.63	0.00 / 100.00	8.46 / 97.70
VIT-L/16	ImageNet	LP	25.79 / 95.79	43.38 / 94.44	9.02 / 98.13	16.90 / 96.04	23.01 / 94.85	29.61 / 94.40	17.52 / 95.32	4.35 / 98.85	21.20 / 95.98
		NM	96.01 / 77.19	84.53 / 85.86	52.23 / 86.50	55.34 / 84.52	79.62 / 72.80	38.19 / 92.76	77.98 / 71.02	70.79 / 81.15	69.34 / 81.48
VIT-L/14	CLIP	LP	0.95 / 99.30	0.25 / 99.68	4.11 / 98.46	16.84 / 94.08	10.89 / 97.50	8.29 / 97.73	21.73 / 91.46	6.78 / 97.85	8.73 / 97.01
		NM	0.03 / 99.71	0.81 / 98.71	0.00 / 100.00	0.00 / 100.00	65.83 / 86.74	0.00 / 100.00	2.27 / 99.35	0.00 / 100.00	8.62 / 98.06

Table 1: Comparison of different architectures and pre-training data – OOD problems with mixed semantic and domain shifts. The measures are described in the main paper.

			AUROC ↑ / OSCR ↑						Average
arch	pre-trained	classif	MNIST	SVHN	CIFAR10	CIFAR+10	CIFAR+50	TIN	
Resnet50	ImageNet	LP	91.82±1.45 / 90.84±1.51	68.37±2.24 / 51.22±2.95	89.19±2.82 / 84.21±2.24	94.30±2.10 / 90.92±2.78	92.53±1.13 / 89.19±1.69	88.52±0.99 / 81.27±2.57	87.46±1.79 / 81.28±2.29
		NM	76.84±2.92 / 62.59±1.78	50.51±3.02 / 20.93±2.84	81.57±0.63 / 71.14±2.01	89.57±0.44 / 82.22±1.40	88.82±2.17 / 81.50±0.66	80.30±0.80 / 70.20±2.84	77.94±1.66 / 64.76±1.92
	CLIP	LP	90.57±1.78 / 88.91±1.84	65.96±1.76 / 47.37±1.44	87.16±2.86 / 81.49±2.17	94.74±3.53 / 91.12±4.07	93.42±1.36 / 89.86±1.94	82.16±1.30 / 74.11±2.71	85.67±2.10 / 78.81±2.36
		NM	72.29±4.32 / 59.35±2.14	51.09±1.24 / 18.71±1.81	73.90±4.13 / 61.98±1.96	88.82±1.13 / 80.20±1.68	83.88±2.99 / 75.72±2.30	69.84±1.99 / 55.29±3.44	73.30±2.63 / 58.54±2.22
VIT-B/16	ImageNet	LP	93.10±2.06 / 91.82±2.14	65.59±1.32 / 46.49±3.12	96.07±0.94 / 93.25±1.45	98.09±1.95 / 96.27±2.46	97.19±1.55 / 95.35±2.03	97.07±0.14 / 93.05±1.31	91.19±1.33 / 86.04±2.09
		NM	84.78±2.01 / 74.61±0.64	50.19±1.56 / 18.93±1.09	93.66±0.39 / 88.89±1.04	97.24±0.90 / 94.32±1.76	97.11±0.39 / 94.18±1.10	87.69±1.65 / 82.68±2.82	85.11±1.15 / 75.60±1.41
	CLIP	LP	98.17±0.30 / 97.40±0.33	71.73±2.28 / 55.41±3.56	96.82±0.63 / 94.42±0.92	98.44±0.86 / 96.85±1.17	97.24±0.77 / 95.67±0.44	91.40±0.88 / 86.00±1.87	92.30±0.95 / 87.63±1.38
		NM	74.47±1.50 / 63.90±2.25	53.48±0.53 / 30.52±0.60	88.80±2.52 / 84.46±1.68	94.14±1.20 / 91.22±1.55	91.87±1.85 / 89.01±1.56	84.75±1.33 / 74.92±2.73	81.25±1.49 / 72.34±1.73
VIT-L/16	ImageNet	LP	92.48±2.07 / 91.16±2.30	69.20±3.08 / 50.64±5.40	95.97±1.07 / 93.17±1.32	98.45±1.15 / 96.60±1.64	96.97±0.87 / 95.01±1.41	97.87±0.14 / 94.93±1.18	91.82±1.40 / 86.92±2.21
		NM	82.82±2.25 / 72.94±0.62	50.62±1.43 / 22.02±2.44	81.15±1.24 / 76.73±2.32	85.88±1.38 / 83.09±1.94	86.99±0.72 / 84.16±1.16	78.54±2.92 / 75.35±3.77	77.67±1.66 / 69.05±2.04
VIT-L/14	CLIP	LP	97.40±0.57 / 96.76±0.61	78.32±0.65 / 65.61±1.10	98.28±0.60 / 97.08±1.00	99.32±0.38 / 98.50±0.50	98.15±0.15 / 97.36±0.22	94.64±0.62 / 91.27±1.66	94.35±0.50 / 91.10±0.85
		NM	84.44±1.28 / 76.86±1.36	57.73±1.69 / 41.94±2.03	90.93±1.31 / 88.61±1.05	93.79±1.40 / 92.12±1.90	93.46±0.55 / 91.78±0.62	89.92±1.53 / 84.23±2.15	85.05±1.29 / 79.26±1.52

Table 2: Comparison of different architectures and pre-training data – OOD problems with semantic shift only. The measures are described in the main paper.

*Corresponding author, vojirtom@fel.cvut.cz