

755

Supplementary Material

Table 4. For each configuration and metric: best-performing method and its margin over next-best in percentage points (Δ). Clear 1-Con wins (non-ties) highlighted in green.

Exp	Gallery Structure		Image Quality		ING Recall		OOG Recall		Accuracy		MDD, Acc	
	Images/ID	Dem	Probe	Gallery	Best	Δ	Best	Δ	Best	Δ	Best	Δ
1	Controlled	Single	M	M	1-Con	0.13	Score	0.03	Score	0.16	Score	0.86
2	Controlled	Single	Blur	M	1-Con	16.5	Score	0.3	1-Con	6.35	1-Con	4.5
			Res	M	1-Con	16.6	Score	0.3	1-Con	6.3	1-Con	2.35
			M	Blur	1-Con	6.4	Score	0.2	1-Con	4.85	1-Con	6.3
			M	Res	1-Con	14.4	Score	0.2	1-Con	5.25	1-Con	6.55
			Blur	Blur	Score	13.4	Gap	9.3	1-Con	1.65	1-Con	3.25
			Res	Res	Score	7.4	Gap	8.3	1-Con	2.35	1-Con	3.85
3	Semi- Controlled	Multi	M	M	1Con, Score	0.4	Gap	0.3	Score	0.05	Score	0.6
			Res	M	1-Con	15.4	Score	0.1	1-Con	5.9	1-Con	5.75
4	Uncontrolled	Single	M	M	1Con, Score	0.1	Gap	3.5	Gap	1.7	—	—
			Res	M	1-Con	16	Gap	0.2	1-Con	6.6	—	—
5	Uncontrolled	Multi	Res	M	1-Con	15.8	Score	0.1	1-Con	6.55	—	—
			S	M	—	—	All	0.0	—	—	—	—

Table 5. For each configuration and metric: worst-performing method and its margin over next-best in percentage points (Δ). Clear 1-Con losses highlighted in red.

Exp	Gallery Structure		Image Quality		ING Recall		OOG Recall		Accuracy		MDD, Acc	
	Images/ID	Dem	Probe	Gallery	Worst	Δ	Worst	Δ	Worst	Δ	Worst	Δ
1	Controlled	Single	M	M	Gap	0.29	1-Con	7.32	1-Con	3.45	1-Con	0.12
2	Controlled	Single	Blur4	M	Score	26.7	1-Con	3.8	Score	13.2	Gap	0.45
			Res18	M	Score	23.5	1-Con	4.0	Score	11.6	Gap	3.55
			M	Blur4	Score	32.9	1-Con	3.8	Score	16.35	Score	2.95
			M	Res18	Score	27.5	1-Con	3.9	Score	13.65	Score	0.4
			Blur4	Blur4	Gap	12.6	Score	59.0	Score	21.15	Score	8.15
			Res18	Res18	Gap	13.0	Score	27.5	Score	7.7	Score	14.2
3	Semi- Controlled	Multi	M	M	Gap	0.4	1-Con	6.7	1-Con	3.3	Gap	0.2
			Res18	M	Score	18.0	1-Con	3.6	Score	8.95	Score	0.55
4	Uncontrolled	Single	M	M	Gap	0.1	1-Con	1.9	1-Con	0.95	—	—
			Res18	M	Gap	4.7	1-Con	2.8	Gap	2.25	—	—
5	Uncontrolled	Multi	Res18	M	Score	13.2	1-Con	2.7	Score	6.55	—	—
			S	M	—	—	All	0.0	—	—	—	—

Table 6. Per-demographic and average performance metrics for the three classification methods.

Exp	Gallery Structure		Image Quality		Group	ING Recall			OOG Recall			Accuracy			
	Images/ID	Dem	Probe	Gallery		1-Con	Score	Gap	1-Con	Score	Gap	1-Con	Score	Gap	
1	Controlled	Single	M	M	BLF	99.79	99.44	99.32	93.59	99.18	99.68	96.69	99.31	99.50	
					BLM	100	99.94	99.97	92.06	99.74	99.65	96.03	99.84	99.81	
					WHF	99.47	99.82	98.35	89.76	97.68	97.38	94.62	98.75	97.87	
					WHM	99.97	99.50	99.91	91.68	99.88	99.65	95.83	99.69	99.78	
					Avg	99.81	99.68	99.39	91.77	99.12	99.09	95.79	99.40	99.24	
2	Controlled	Single	Blur	M	BLF	77.20	45.40	54.30	96.80	99.90	100	87.00	72.65	77.15	
					BLM	93.10	59.40	80.20	96.80	100	99.90	94.95	79.70	90.05	
					WHF	90.10	34.80	74.50	94.00	99.70	98.80	92.05	67.25	86.65	
					WHM	91.10	39.10	76.60	95.60	100	99.90	93.35	69.55	88.25	
					Avg	87.90	44.70	71.40	95.80	99.90	99.60	91.85	72.30	85.50	
					Res	BLF	75.80	47.00	52.40	96.60	99.90	99.90	86.20	73.45	76.15
						BLM	91.60	58.60	76.90	95.90	100	99.90	93.75	79.30	88.40
						WHF	89.80	39.20	73.70	94.50	99.60	98.70	92.15	69.40	86.20
						WHM	91.80	43.40	79.30	95.50	100	99.90	93.65	71.70	89.60
						Avg	87.20	47.10	70.60	95.60	99.90	99.60	91.40	73.50	85.10
						Blur	BLF	84.80	44.70	64.10	96.70	100	100	90.75	72.35
					BLM		97.00	65.00	88.40	96.20	100	100	96.60	82.50	94.20
			WHF	92.10	35.10		78.70	94.40	99.70	98.80	93.25	67.40	88.75		
			WHM	94.60	37.80		83.10	96.30	100	99.90	95.45	68.90	91.50		
			Avg	92.10	45.70		78.60	95.90	99.90	99.70	94.00	72.80	89.15		
			Res	BLF	83.40		48.70	61.40	97.10	100	99.90	90.25	74.35	80.65	
				BLM	96.00		65.60	85.50	95.50	100	99.90	95.75	82.80	92.70	
				WHF	92.20		41.00	78.60	94.50	99.70	99.00	93.35	70.35	88.80	
				WHM	94.80		43.50	83.20	95.90	100	99.90	95.35	71.75	91.55	
				Avg	91.60		49.70	77.20	95.80	99.90	99.70	93.70	74.80	88.45	
				Blur	BLF		73.30	97.40	56.90	91.40	11.80	99.60	82.35	54.60	78.25
			BLM		89.40		98.70	78.30	90.10	18.00	99.50	89.75	58.35	88.90	
			WHF		86.30	96.30	74.90	88.90	42.70	98.50	87.60	69.50	86.70		
			WHM		85.50	95.50	73.90	89.80	51.30	99.50	87.65	73.40	86.70		
			Avg		83.60	97.00	71.00	90.00	31.00	99.30	86.80	64.00	85.15		
			Res		BLF	73.70	93.10	56.40	92.20	34.10	99.80	82.95	63.60	78.10	
					BLM	87.10	94.90	75.50	91.10	53.90	99.70	89.10	74.40	87.60	
					WHF	86.80	90.10	74.70	90.40	77.40	98.50	88.60	83.75	86.60	
					WHM	87.50	86.50	76.60	90.70	89.10	99.60	89.10	87.80	88.10	
					Avg	83.80	91.20	70.80	91.10	63.60	99.40	87.45	77.40	85.10	
Blur	BLF	99.90			99.80	99.60	93.20	98.20	99.60	96.55	99.00	99.60			
	BLM	100	100		100	92.80	99.30	99.50	96.40	99.65	99.75				
	WHF	99.60	100	98.60	90.20	97.60	97.40	94.90	98.80	98.00					
	WHM	100	99.80	99.90	92.30	99.90	99.80	96.15	99.85	99.85					
	Avg	99.90	99.90	99.50	92.10	98.80	99.10	96.00	99.35	99.30					
	Res	BLF	78.10	55.50	55.30	96.70	99.90	99.90	87.40	77.70	77.60				
		BLM	94.10	71.70	82.10	96.60	99.90	100	95.35	85.80	91.05				
		WHF	91.30	43.60	75.50	94.80	99.50	98.80	93.05	71.55	87.15				
		WHM	93.60	52.90	82.60	96.20	100	100	94.90	76.45	91.30				
		Avg	89.30	55.90	73.90	96.10	99.80	99.70	92.70	77.85	86.80				
		4	Uncontrolled	Single	M	M	BLM	100	100	99.90	94.20	96.10	99.60	97.10	98.05
	Res				BLM		87.90	71.90	67.20	96.90	99.70	99.90	92.40	85.80	83.55
S	All				91.30		62.30	75.50	97.20	100	99.90	94.25	81.15	87.70	
5	Uncontrolled	Multi	Res	M	All	—	—	—	100	100	100	100	100	100	
			S		All	—	—	—	100	100	100	100	100	100	

Table 7. Distributional separation (EMD) between ING and OOG distributions for each classification method. Higher values indicate better separation. Δ shows percentage change from baseline (Exp. 1).

Exp	Gal Comp		Degradation		1-Consistency		Score-Threshold		Gap-Threshold	
	N Imgs	N Dems	Prb	Gal	EMD	Δ	EMD	Δ	EMD	Δ
1	2	1	—	—	2.982	—	0.399	—	0.376	—
2	2	1	BL	—	3.452	+15.8%	0.145	-63.6%	0.126	-66.3%
			RE	—	3.359	+12.6%	0.146	-63.3%	0.127	-66.1%
			—	BL	3.555	+19.2%	0.159	-60.1%	0.141	-62.4%
			—	RE	3.494	+17.2%	0.159	-60.3%	0.140	-62.7%
			BL	BL	2.306	-22.7%	0.171	-57.2%	0.148	-60.5%
			RE	RE	2.405	-19.4%	0.170	-57.3%	0.148	-60.6%
3	2-17	4	—	—	3.057	+2.5%	0.409	+2.3%	0.385	+2.6%
			RE	—	3.499	+17.3%	0.150	-62.3%	0.132	-64.9%
4	1-108	1	—	—	3.370	+13.0%	0.383	-4.0%	0.364	-3.2%
			RE	—	3.844	+28.9%	0.129	-67.7%	0.112	-70.1%
5	1-108	6	RE	—	3.761	+26.1%	0.149	-62.7%	0.132	-64.9%