

Table 1: Performance on image datasets. The best score is marked in bold.

Methods	MNIST-USPS (K=1200)				MNIST-Invert (K=500)			
	Recall@K	ROCAUC	Rec Diff	Time(s)	Recall@K	ROCAUC	Rec Diff	Time(s)
FairOD	12.35±1.13	50.00±0.28	11.56±0.64	29.57	7.52±0.74	50.40±0.20	8.26±1.27	20.25
DCFOD	12.63±0.33	50.09±0.27	8.99±0.83	710.33	6.95±0.91	50.54±0.54	7.23±2.02	1277.31
FairSVDD	15.62±1.52	58.33±1.18	13.75±2.56	768.79	12.41±0.76	49.67±3.98	12.46±2.12	843.12
VFAE	17.43±0.25	50.49±0.38	23.80±0.34	50.07	10.99±0.67	56.15±1.61	10.54±1.74	63.46
MCM	39.75±0.23	78.80±1.02	55.81±0.80	417.09	25.35±0.56	80.96±0.49	80.13±1.41	752.36
NSNMF	39.16±0.84	65.38±0.58	62.90±3.84	28.53	51.79±0.61	74.21±0.34	51.07±1.79	18.97
Recontrast	64.29±3.18	83.46±3.77	41.16±5.63	116.75	64.22±1.60	85.13±5.19	56.50±11.23	117.15
FADIG	67.19±0.33	91.28±0.46	3.77±2.18	121.97	71.82±0.63	97.99±0.07	9.78±3.10	60.42

Table 2: Performance on tabular datasets. The best score is marked in bold.

Methods	COMPAS (K=350)				CelebA (K=5000)			
	Recall@K	ROCAUC	Rec Diff	Time(s)	Recall@K	ROCAUC	Rec Diff	Time(s)
FairOD	16.56±2.12	50.09±1.28	7.97±1.23	4.18	8.93±0.14	49.94±0.12	0.68±0.56	78.92
DCFOD	16.08±1.94	49.55±1.21	9.81±1.76	115.86	9.66±0.69	49.92±0.14	7.83±1.26	2517.68
FairSVDD	15.33±2.10	52.68±5.29	11.57±4.06	6.81	10.19±0.50	58.40±1.02	10.95±1.93	243.17
VFAE	16.67±0.16	52.23±0.06	18.67±0.18	9.60	8.62±0.07	48.11±0.49	10.00±0.09	45.09
MCM	21.10±0.54	50.97±0.43	6.29±2.66	38.12	11.03±0.38	46.23±3.46	26.15±9.31	640.12
NSNMF	22.92±0.32	57.97±0.66	36.78±1.71	7.69	10.91±0.54	50.45±0.30	8.04±1.33	1927.55
FADIG	34.38±0.36	61.45±0.47	5.97±4.34	19.88	11.96±0.49	59.43±0.42	4.72±1.26	48.93

Table 3: Performance on tabular datasets. The best score is marked in bold.

Methods	ACSIIncome (K=2000)				CelebA (K=5000)			
	Recall@K	ROCAUC	Rec Diff	Time(s)	Recall@K	ROCAUC	Rec Diff	Time(s)
FairOD	16.56±2.12	50.09±1.28	7.97±1.23	4.18	8.93±0.14	49.94±0.12	0.68±0.56	78.92
DCFOD	16.08±1.94	49.55±1.21	9.81±1.76	115.86	9.66±0.69	49.92±0.14	7.83±1.26	2517.68
FairSVDD	15.33±2.10	52.68±5.29	11.57±4.06	6.81	10.19±0.50	58.40±1.02	10.95±1.93	243.17
VFAE	16.67±0.16	52.23±0.06	18.67±0.18	9.60	8.62±0.07	48.11±0.49	10.00±0.09	45.09
MCM	21.10±0.54	50.97±0.43	6.29±2.66	38.12	11.03±0.38	46.23±3.46	26.15±9.31	640.12
NSNMF	22.92±0.32	57.97±0.66	36.78±1.71	7.69	10.91±0.54	50.45±0.30	8.04±1.33	1927.55
FADIG	34.38±0.36	61.45±0.47	5.97±4.34	19.88	11.96±0.49	59.43±0.42	4.72±1.26	48.93

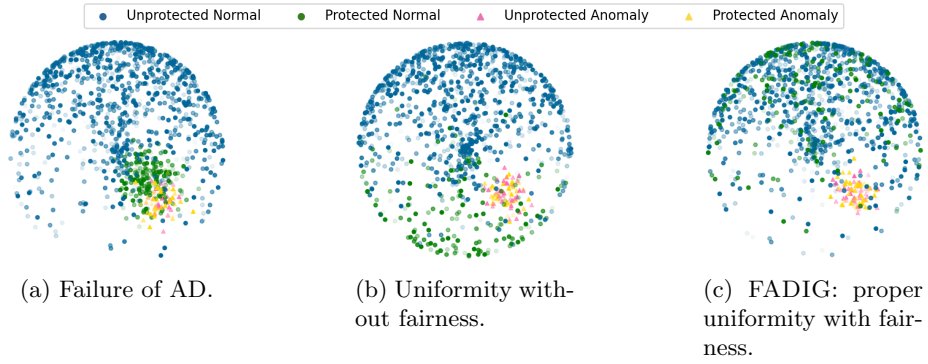


Figure 1: Illustrations of uniformity. The blue and green circles denote normal examples from the unprotected group and protected group respectively. The pink and yellow triangles denote anomalies from the unprotected group and protected group respectively. The three subfigures illustrate three different projections from the same data set. With projection (a), many existing AD methods overly flag the examples from the protected groups (green circles) as anomalies (triangles). In projection (b), traditional contrastive regularization methods encourage uniformity but do not consider group fairness. In (c), our FADIG ensures group fairness while maintaining proper uniformity.