

**Table S1.** ImageNet-based FDs. Columns 1 and 2 split the models by dataset and augmentation technique (Aug): no augmentation (None), ADA, APA, and DiffAugment (DiffAug). Columns 3-9 display the FDs computed using seven ImageNet-trained feature extractors: InceptionV3 (Incept), ResNet50 (Res), InceptionResNetV2 (IRV2), DenseNet121 (Dense), SwAV, DINO, and Swin Transformer (Swin).  $\downarrow$  indicates that a lower value is preferable. The underlined boldface type represents the best performance per dataset.  $\dagger$  denotes decreased performance compared to no augmentation.

Dataset	Aug	Fréchet Distances (ImageNet) $\downarrow$						
		Incept	Res	IRV2	Dense	SwAV	DINO	Swin
ChestXray-14	None	5.01	4.16	2.79	14.02	1.07	299.13	4.76
	ADA	3.56	3.11	2.37	11.52	0.66	187.16	3.69
	APA	7.03 $\dagger$	7.97 $\dagger$	3.34 $\dagger$	20.09 $\dagger$	1.32 $\dagger$	407.05 $\dagger$	7.49 $\dagger$
	DiffAug	<b>3.07</b>	<b>2.65</b>	<b>1.46</b>	<b>8.82</b>	<b>0.50</b>	<b>170.84</b>	<b>3.19</b>
SLIVER07	None	8.72	9.04	4.74	14.02	2.40	640.32	30.37
	ADA	7.34	6.79	4.41	12.65	1.99	478.61	31.09 $\dagger$
	APA	8.07	8.22	4.40	12.92	2.26	585.60	27.17
	DiffAug	<b>4.62</b>	<b>4.33</b>	<b>1.95</b>	<b>6.47</b>	<b>1.53</b>	<b>321.32</b>	<b>23.83</b>
MSD	None	7.09	5.05	10.40	<b>9.43</b>	0.57	<b>422.74</b>	85.76
	ADA	<b>7.00</b>	<b>5.00</b>	<b>10.01</b>	11.33 $\dagger$	1.22 $\dagger$	475.41 $\dagger$	52.46
	APA	8.29 $\dagger$	5.60 $\dagger$	13.90 $\dagger$	11.61 $\dagger$	<b>0.49</b>	493.23 $\dagger$	33.43
	DiffAug	8.80 $\dagger$	10.04 $\dagger$	13.58 $\dagger$	13.63 $\dagger$	3.30 $\dagger$	538.61 $\dagger$	<b>29.77</b>
ACDC	None	67.05	51.60	73.51	87.22	5.90	2888.53	127.74
	ADA	28.34	21.21	26.91	35.96	3.82	1350.34	70.71
	APA	42.05	33.44	46.20	55.06	4.53	1807.38	82.91
	DiffAug	<b>21.42</b>	<b>16.05</b>	<b>20.04</b>	<b>29.23</b>	<b>3.55</b>	<b>1040.24</b>	<b>54.31</b>

**Table S2.** RadImageNet-based FDs. Columns 1 and 2 split the models by dataset and augmentation technique (Aug): no augmentation (None), ADA, APA, and DiffAugment (DiffAug). Columns 3-6 display the FDs computed using four RadImageNet-trained feature extractors: InceptionV3 (Incept), ResNet50 (Res), InceptionResNetV2 (IRV2), and DenseNet121 (Dense).  $\downarrow$  indicates that a lower value is preferable. The underlined boldface type represents the best performance per dataset.  $\dagger$  denotes decreased performance compared to no augmentation.

Dataset	Aug	Fréchet Distances (RadImageNet) $\downarrow$			
		Incept	Res	IRV2	Dense
ChestXray-14	None	<b><u>0.03</u></b>	0.15	<b><u>0.08</u></b>	0.04
	ADA	0.13 $\dagger$	0.27 $\dagger$	0.19 $\dagger$	0.08 $\dagger$
	APA	0.06 $\dagger$	0.13	<b><u>0.08</u></b>	0.08 $\dagger$
	DiffAug	0.06 $\dagger$	<b><u>0.10</u></b>	0.09 $\dagger$	<b><u>0.03</u></b>
SLIVER07	None	0.07	0.22	0.24	0.13
	ADA	<b><u>0.03</u></b>	<b><u>0.13</u></b>	0.15	<b><u>0.07</u></b>
	APA	0.04	<b><u>0.13</u></b>	<b><u>0.12</u></b>	0.08
	DiffAug	0.08 $\dagger$	0.23 $\dagger$	0.22	0.14 $\dagger$
MSD	None	0.05	0.13	<b><u>0.13</u></b>	<b><u>0.08</u></b>
	ADA	<b><u>0.04</u></b>	<b><u>0.11</u></b>	0.15 $\dagger$	0.13 $\dagger$
	APA	0.05	0.13	0.16 $\dagger$	<b><u>0.08</u></b>
	DiffAug	1.55 $\dagger$	4.42 $\dagger$	1.40 $\dagger$	1.23 $\dagger$
ACDC	None	0.29	0.76	0.61	0.65
	ADA	<b><u>0.11</u></b>	0.37	<b><u>0.29</u></b>	0.26
	APA	0.15	<b><u>0.35</u></b>	<b><u>0.35</u></b>	0.35
	DiffAug	0.13	0.61	<b><u>0.29</u></b>	<b><u>0.21</u></b>