

CADRef: Robust Out-of-Distribution Detection via Class-Aware Decoupled Relative Feature Leveraging

Supplementary Material

7. Experiment Details

7.1. Dataset Settings

In our experiments, we perform OOD detection across two dataset scenarios. In the small-scale dataset scenario, we employ CIFAR-10 and CIFAR-100 as ID datasets, with SVHN, LSUN-R, LSUN-C, iSUN, Texture, and Places serving as the corresponding OOD datasets. In the large-scale dataset scenario, we use ImageNet-1k as the ID dataset, with OOD detection conducted on six datasets: iNaturalist, SUN, Places, Textures, OpenImage-O and ImageNet-O. Additionally, we also conduct extended experiments on SSB-hard and Ninco to further discuss the impact of the logit-based component. Table 8 provides detailed information of the datasets utilized in both scenarios.

	Dataset	# of Classes	# of Samples	Size
ID	CIFAR-10	10	10000	32×32
	CIFAR-100	100	10000	32×32
	ImageNet-1k	1000	50000	224×224
OOD (small-scale)	SVHN	—	26032	32×32
	LSUN-R	—	10000	32×32
	LSUN-C	—	10000	32×32
	iSUN	—	8925	32×32
	Texture	—	5640	32×32
	Places	—	10000	32×32
OOD (large-scale)	iNaturalist	—	10000	224×224
	SUN	—	10000	224×224
	Places	—	10000	224×224
	Texture	—	5640	224×224
	OpenImage-O	—	17632	224×224
	ImageNet-O	—	2000	224×224
	SSB-hard	—	49000	224×224
	Ninco	—	5878	224×224

Table 8. Details of ID/OOD datasets. Note that the third column is the number of test samples. All OOD dataset samples are processed to match the size of the corresponding ID dataset samples.

7.2. Hyperparameter Settings

The detailed hyperparameters of all baseline methods are listed in Table 9. For most baseline methods, we keep the

same hyperparameters as the original paper. For methods that lack experiments on ImageNet-1k in the original work, we adopt the settings from the *ReAct* [34].

8. Detailed Experiment Results

For the ImageNet-1k benchmark, we conduct extensive experiments using ResNet-50, RegNet-8GF, DenseNet-201, ViT-B/16, Swin-B, ConvNeXt-B, and MaxViT-B models to complete Table 3. Besides the main text, we introduce three new baselines: *CIDER* [29], *PALM* [27], and *Mahalanobis* [21]. To address the performance drop observed with *PALM*'s original settings, we modify its learning rate from 0.001 to 0.01, maintaining all other configurations as in the original paper. All results are presented in Tables 10, 11, 12, 13, 14, 15, and 16, where *CARef* and *CADRef* consistently demonstrate superior performance across these models, highlighting the robustness of our methods.

9. Collapse of the feature norm

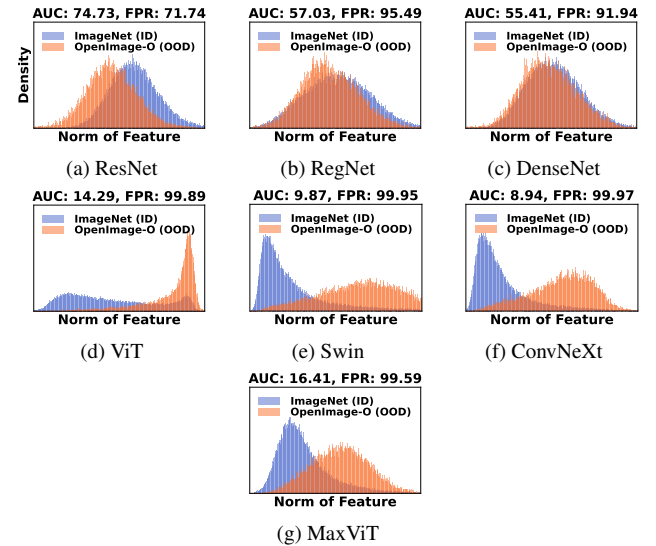


Figure 6. Comparison of feature ℓ_2 -norm across different models.

We also explore the challenges associated with using feature norm for OOD detection. According to the conclusion in [46], the feature norms of ID samples are generally larger than those of OOD samples. However, as shown in Figure 6, this conclusion holds only for ResNet. In RegNet and DenseNet, this method lacks discriminative power (with an

Method	Configurable Hyperparameters	CIFAR-10	CIFAR-100	ImageNet-1k
<i>MSP</i>	None	–	–	–
<i>MaxLogit</i>	None	–	–	–
<i>ODIN</i>	T : temperature scaling ϵ : perturbation magnitude	$T = 1000$ $\epsilon = 0.0014$	$T = 1000$ $\epsilon = 0.002$	$T = 1000$ $\epsilon = 0.005$
<i>Energy</i>	T : temperature scaling	$T = 1$	$T = 1$	$T = 1$
<i>GEN</i>	M : top classes used in truncated sum γ : exponential scale	$M = 10$ $\gamma = 0.1$	$M = 10$ $\gamma = 0.1$	$M = 100$ $\gamma = 0.1$
<i>ReAct</i>	p : percentile for rectification threshold	$p = 90$	$p = 90$	$p = 90$
<i>DICE</i>	p : sparsity parameter	$p = 0.9$	$p = 0.7$	$p = 0.7$
<i>ViM</i>	D : dimension of principal space	$D = 171$	$D = 171$	$D = \begin{cases} 256 & \text{for MaxViT} \\ 1000 & \text{for ResNet} \\ 512 & \text{for others} \end{cases}$
<i>ASH-S</i>	p : pruning percentage	$p = 95$	$p = 90$	$p = 90$
<i>OptFS</i>	K : number of intervals	$K = 100$	$K = 100$	$K = 100$
<i>Ours</i>	None	–	–	–

Table 9. Hyperparameter settings for different OOD detection methods.

AUROC of around 50%), and in other models, the conclusion is actually reversed.

9.1. Score and error distribution for other methods

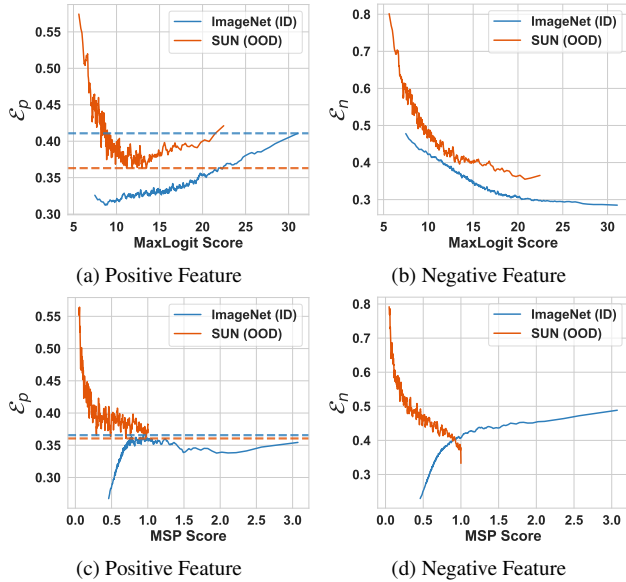


Figure 7. Score and error distribution of ID/OOD samples.

Figure 7 illustrates the score and error distributions for the *MSP* and *MaxLogit* methods. While the *MaxLogit* method demonstrates error trends consistent with the *Energy* and *GEN* methods, the *MSP* method exhibits a markedly different behavior. At higher S_{logit} values, *MSP* shows distinctly separated positive errors but significant overlap in negative errors, which fundamentally contradicts the *Error Scaling* component’s design principles. Consequently, integrating *MSP* as a logit-based component in *CADRef* leads to a detrimental impact, as substantiated by the results in Figure 5.

Methods	iNaturalist		SUN		Places		Textures		OpenImage-O		ImageNet-O		Average	
	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓
<i>MSP</i>	88.42	52.73	81.75	68.58	80.63	71.59	80.46	66.15	84.98	63.60	28.61	100.00	74.14	70.44
<i>MaxLogit</i>	91.14	50.77	86.43	60.39	84.03	66.03	86.38	54.91	89.13	57.89	40.73	100.00	79.64	65.00
<i>ODIN</i>	87.00	52.33	86.57	53.49	85.30	58.64	86.51	46.08	86.65	52.76	43.63	98.65	79.28	60.32
<i>Energy</i>	90.59	53.96	86.73	58.28	84.12	65.43	86.73	52.30	89.12	57.23	41.79	100.00	79.85	64.53
<i>GEN</i>	92.44	45.76	85.52	65.54	83.46	69.24	85.41	59.24	89.31	60.44	43.59	100.00	79.95	66.70
<i>ReAct</i>	96.39	19.55	94.41	24.01	91.93	33.45	90.45	45.83	90.53	43.69	52.45	98.00	86.03	44.09
<i>DICE</i>	94.51	26.63	90.91	36.48	87.64	47.98	90.44	32.58	88.57	45.72	42.78	98.00	82.48	47.90
<i>ViM</i>	87.42	71.80	81.07	81.80	78.39	83.12	96.83	14.84	89.30	58.68	70.77	84.85	83.96	65.85
<i>ASH-S</i>	97.87	11.49	<u>94.02</u>	<u>27.96</u>	<u>90.98</u>	<u>39.83</u>	97.60	11.97	92.75	<u>32.77</u>	67.44	89.10	90.11	35.52
<i>OptFS</i>	96.88	16.79	93.13	35.31	<u>90.42</u>	<u>44.78</u>	95.74	23.08	<u>92.77</u>	37.68	59.94	97.20	88.15	42.47
<i>Mahalanobis</i>	63.97	93.84	52.19	97.43	51.85	97.30	90.10	43.49	69.56	87.64	80.90	66.55	68.10	81.04
<i>CIDER</i>	92.63	42.24	85.91	61.07	81.33	71.30	<u>97.71</u>	9.77	90.85	42.99	77.74	80.86	87.70	51.37
<i>PALM</i>	93.96	31.86	88.91	48.95	83.90	60.68	97.22	12.22	90.39	40.18	<u>79.66</u>	<u>75.85</u>	89.01	44.96
<i>CARef</i>	96.54	17.46	89.51	44.89	85.41	57.64	97.94	<u>10.15</u>	92.57	37.73	77.66	77.60	<u>89.94</u>	40.91
<i>CADRef</i>	<u>96.90</u>	<u>16.08</u>	91.26	39.23	87.80	51.12	97.14	12.60	93.93	32.69	68.38	92.35	89.24	<u>40.68</u>

Table 10. Results of OOD detection on ImageNet-1k benchmark with ResNet-50. ↑ indicates that higher values are better, while ↓ indicates that lower values are better. All values are percentages, with the best and second-best results being **highlighted** and underlined, respectively.

Methods	iNaturalist		SUN		Places		Textures		OpenImage-O		ImageNet-O		Average	
	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓
<i>MSP</i>	88.32	53.48	82.16	68.00	80.98	71.68	79.87	68.77	86.01	61.30	52.69	96.65	78.34	69.98
<i>MaxLogit</i>	90.01	51.89	86.32	57.56	84.02	64.45	82.71	61.24	89.59	51.53	57.64	95.80	81.71	63.75
<i>ODIN</i>	85.52	54.58	86.11	50.33	85.76	54.66	82.34	53.19	86.99	49.30	56.72	89.90	80.57	58.66
<i>Energy</i>	89.29	55.36	86.27	57.37	83.85	63.97	82.51	61.40	89.36	51.87	58.03	94.70	81.55	64.11
<i>GEN</i>	92.36	44.63	86.44	60.33	84.40	66.31	84.53	60.85	90.59	52.96	62.12	95.65	83.41	63.46
<i>ReAct</i>	96.04	21.72	94.85	24.55	91.71	36.38	87.26	56.80	87.96	47.03	61.91	91.35	86.62	46.30
<i>DICE</i>	88.83	56.48	84.19	62.09	80.05	74.33	80.13	64.18	81.65	62.26	52.65	94.00	77.92	68.89
<i>ViM</i>	90.88	58.04	85.31	70.72	82.12	75.84	<u>97.35</u>	12.23	91.95	48.58	80.85	71.40	88.08	56.13
<i>ASH-S</i>	96.49	18.52	91.00	<u>35.15</u>	86.84	<u>50.60</u>	97.12	13.40	91.07	<u>35.58</u>	73.71	74.80	<u>89.37</u>	38.01
<i>OptFS</i>	<u>96.06</u>	<u>20.96</u>	<u>92.03</u>	40.58	<u>88.30</u>	51.62	95.90	22.96	<u>92.33</u>	37.47	71.19	84.20	89.30	42.97
<i>Mahalanobis</i>	72.94	89.86	60.81	94.63	59.74	95.45	92.89	32.20	74.97	82.21	79.74	70.35	73.52	77.45
<i>CIDER</i>	90.62	51.76	85.59	61.69	80.00	73.79	97.63	9.75	89.10	50.64	<u>83.04</u>	<u>65.49</u>	87.66	52.19
<i>PALM</i>	85.12	65.34	80.80	71.64	73.89	80.24	97.23	<u>12.06</u>	83.88	59.89	83.33	62.80	84.04	58.66
<i>CARef</i>	93.07	41.97	86.21	57.28	81.48	72.04	97.26	15.00	90.21	47.77	81.38	70.00	88.27	50.68
<i>CADRef</i>	95.31	28.13	90.29	44.08	86.73	58.14	97.20	14.02	93.87	34.70	78.15	75.00	90.26	<u>42.34</u>

Table 11. Results of OOD detection on ImageNet-1k benchmark with RegNet-8GF. ↑ indicates that higher values are better, while ↓ indicates that lower values are better. All values are percentages, with the best and second-best results being **highlighted** and underlined, respectively.

Methods	iNaturalist		SUN		Places		Textures		OpenImage-O		ImageNet-O		Average	
	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓
<i>MSP</i>	89.83	44.80	82.22	65.22	81.13	68.86	79.40	67.02	85.21	61.93	48.78	97.95	77.76	67.63
<i>MaxLogit</i>	92.12	40.68	85.91	55.85	83.83	62.12	83.41	58.31	88.31	55.05	53.35	97.35	81.16	61.56
<i>ODIN</i>	89.95	42.93	84.09	55.57	83.01	59.93	81.45	54.85	84.70	56.80	50.21	94.40	78.90	60.75
<i>Energy</i>	91.47	43.76	85.97	53.85	83.68	61.42	83.44	56.61	87.95	55.92	54.04	95.95	81.09	61.25
<i>GEN</i>	93.37	37.92	85.77	59.34	84.03	64.53	83.78	60.69	89.24	58.20	58.18	97.40	82.39	63.01
<i>ReAct</i>	88.65	51.99	89.55	51.38	85.92	61.57	81.49	60.83	74.19	68.20	50.97	91.10	78.46	64.18
<i>DICE</i>	91.73	39.18	86.76	48.13	82.29	60.72	83.85	53.28	81.62	59.35	48.94	93.60	79.20	59.04
<i>ViM</i>	82.69	85.18	73.99	91.50	72.87	91.43	94.66	26.13	87.21	66.06	76.58	75.64	81.33	72.66
<i>ASH-S</i>	96.18	20.51	90.58	39.42	87.54	50.40	93.81	26.13	92.28	37.42	64.97	88.95	87.56	43.80
<i>OptFS</i>	94.61	28.62	<u>90.08</u>	<u>46.12</u>	<u>86.18</u>	<u>56.28</u>	95.17	26.32	90.23	46.91	73.21	83.00	<u>88.25</u>	47.88
<i>Mahalanobis</i>	64.12	94.30	48.25	97.74	47.62	97.69	94.24	26.45	76.83	78.84	83.07	62.50	69.02	76.25
<i>CIDER</i>	89.17	61.90	81.61	74.77	77.25	81.31	96.95	12.91	90.51	47.23	<u>80.44</u>	<u>70.96</u>	85.99	58.18
<i>PALM</i>	89.44	60.74	81.08	82.11	75.91	86.09	94.08	26.72	85.99	63.67	76.92	82.10	83.90	66.90
<i>CARef</i>	92.94	36.18	84.16	60.12	78.68	72.13	96.43	18.99	87.01	54.50	80.08	73.60	86.55	52.59
<i>CADRef</i>	<u>95.41</u>	<u>24.59</u>	89.17	47.19	85.03	58.96	96.39	<u>16.68</u>	<u>91.57</u>	<u>42.85</u>	74.27	81.20	88.64	<u>45.25</u>

Table 12. Results of OOD detection on ImageNet-1k benchmark with DenseNet-201. ↑ indicates that higher values are better, while ↓ indicates that lower values are better. All values are percentages, with the best and second-best results being **highlighted** and underlined, respectively.

Methods	iNaturalist		SUN		Places		Textures		OpenImage-O		ImageNet-O		Average	
	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓
<i>MSP</i>	88.18	51.42	80.88	66.65	80.37	68.49	82.96	60.47	84.81	59.84	58.79	90.85	79.33	66.29
<i>MaxLogit</i>	85.27	52.14	76.28	66.86	75.03	69.19	81.64	56.57	81.53	58.72	54.31	88.94	75.68	65.40
<i>ODIN</i>	71.50	94.86	56.56	95.31	56.12	95.49	66.32	92.18	62.79	94.79	55.43	93.35	61.45	94.33
<i>Energy</i>	79.27	63.93	70.16	72.81	68.39	74.35	79.24	58.45	76.44	64.83	52.71	87.00	71.04	70.23
<i>GEN</i>	92.91	40.05	85.03	61.06	<u>83.49</u>	64.22	87.97	50.69	89.36	52.53	67.44	88.90	84.37	59.58
<i>ReAct</i>	85.99	65.07	78.93	72.38	77.48	73.84	84.62	57.09	84.21	64.58	66.69	87.00	79.65	69.99
<i>DICE</i>	74.54	90.41	65.23	94.51	64.79	93.05	77.34	84.04	77.28	83.01	70.75	86.05	71.65	88.51
<i>ViM</i>	97.18	12.55	83.99	56.77	81.48	58.98	88.92	<u>46.48</u>	92.44	38.11	76.29	83.25	86.72	49.36
<i>ASH-S</i>	6.68	99.99	16.70	99.73	18.32	99.65	24.13	99.13	14.04	99.84	28.46	99.40	18.06	99.62
<i>OptFS</i>	89.94	55.59	84.19	66.36	82.70	68.39	86.46	56.60	88.20	59.86	71.85	89.15	83.89	65.99
<i>Mahalanobis</i>	<u>95.97</u>	<u>18.69</u>	85.18	<u>60.07</u>	84.03	<u>60.71</u>	87.66	52.93	<u>92.26</u>	<u>40.44</u>	77.03	86.55	87.02	<u>53.23</u>
<i>CIDER</i>	91.74	47.68	83.69	71.43	82.41	72.99	88.76	51.14	90.45	52.16	<u>78.89</u>	<u>81.77</u>	85.99	62.86
<i>PALM</i>	86.97	59.15	74.47	76.94	72.04	78.38	88.03	43.85	86.46	61.12	81.24	77.85	81.53	66.21
<i>CARef</i>	94.07	36.59	84.83	66.92	83.03	68.36	<u>89.28</u>	52.57	91.54	50.32	78.31	88.14	86.84	60.48
<i>CADRef</i>	93.81	38.45	<u>85.12</u>	65.85	83.28	67.91	89.59	49.78	91.52	50.44	78.12	87.85	<u>86.91</u>	60.05

Table 13. Results of OOD detection on ImageNet-1k benchmark with ViT-B/16. ↑ indicates that higher values are better, while ↓ indicates that lower values are better. All values are percentages, with the best and second-best results being **highlighted** and underlined, respectively.

Methods	iNaturalist		SUN		Places		Textures		OpenImage-O		ImageNet-O		Average	
	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓
<i>MSP</i>	87.00	51.68	80.53	66.52	80.69	67.79	78.46	66.20	81.78	62.13	58.38	88.40	77.81	67.12
<i>MaxLogit</i>	79.86	55.48	72.23	67.86	71.97	69.54	73.98	61.93	71.08	65.52	56.00	85.75	70.85	67.68
<i>ODIN</i>	58.88	90.31	51.06	92.18	48.57	93.24	64.36	86.61	52.42	92.10	56.00	91.70	55.22	91.02
<i>Energy</i>	68.57	72.55	63.08	78.54	62.48	78.92	69.58	65.95	60.23	76.13	55.13	83.65	63.18	75.96
<i>GEN</i>	92.69	32.94	85.02	56.61	84.06	59.95	85.61	49.93	87.18	48.66	67.77	83.90	83.72	55.33
<i>ReAct</i>	88.72	59.29	81.46	69.10	80.87	70.79	84.07	57.83	85.90	61.58	70.45	80.85	81.91	66.57
<i>DICE</i>	23.13	97.70	42.48	91.57	33.91	94.74	73.15	63.22	45.55	86.66	56.50	86.65	45.79	86.76
<i>ViM</i>	93.60	44.56	80.12	71.10	77.96	71.82	83.96	64.64	92.10	46.36	76.03	83.80	83.96	63.63
<i>ASH-S</i>	10.69	99.81	20.18	99.36	21.37	99.59	18.41	98.65	11.94	99.84	32.52	98.85	19.18	99.35
<i>OptFS</i>	90.71	54.98	84.86	67.93	<u>83.94</u>	68.63	85.10	61.68	90.34	51.19	73.28	85.75	84.71	65.03
<i>Mahalanobis</i>	93.99	<u>36.30</u>	84.19	71.72	83.02	71.88	86.21	64.31	<u>92.41</u>	44.39	76.74	85.35	86.09	62.32
<i>CIDER</i>	93.35	38.72	85.45	66.91	83.69	68.95	88.67	52.15	92.25	43.46	<u>78.83</u>	84.24	<u>87.04</u>	59.07
<i>PALM</i>	92.79	43.35	77.62	69.17	73.68	71.91	88.17	50.39	91.54	46.40	80.95	82.95	84.12	60.70
<i>CARef</i>	93.57	39.78	84.83	66.72	83.25	68.68	88.74	<u>49.86</u>	92.55	<u>43.52</u>	78.57	83.35	86.92	58.65
<i>CADRef</i>	<u>93.77</u>	37.87	<u>85.12</u>	<u>64.23</u>	83.57	<u>66.71</u>	89.08	47.38	92.32	46.63	78.76	<u>81.45</u>	87.10	<u>57.38</u>

Table 14. Results of OOD detection on ImageNet-1k benchmark with Swin-B. ↑ indicates that higher values are better, while ↓ indicates that lower values are better. All values are percentages, with the best and second-best results being **highlighted** and underlined, respectively.

Methods	iNaturalist		SUN		Places		Textures		OpenImage-O		ImageNet-O		Average	
	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓
<i>MSP</i>	89.69	44.35	79.28	63.37	78.41	67.02	73.77	65.76	83.40	58.48	52.85	92.45	76.23	65.24
<i>MaxLogit</i>	83.32	57.91	69.83	72.02	68.44	74.92	65.72	70.88	74.08	68.10	51.18	92.10	68.76	72.66
<i>ODIN</i>	63.62	81.28	47.54	91.28	47.73	90.28	43.43	93.49	52.75	88.20	51.48	91.45	51.09	89.33
<i>Energy</i>	57.94	90.10	51.17	94.53	50.85	92.84	52.87	87.26	54.62	88.78	50.00	93.00	52.91	91.09
<i>GEN</i>	<u>94.86</u>	<u>25.69</u>	84.78	52.70	83.04	59.18	80.10	55.21	89.37	46.11	64.02	90.50	82.69	54.90
<i>ReAct</i>	87.94	68.80	78.51	78.02	77.07	79.42	76.45	71.57	83.30	73.30	65.61	90.10	78.15	76.87
<i>DICE</i>	20.82	98.55	41.27	93.95	34.27	96.24	71.72	69.92	47.62	88.50	54.99	91.45	45.12	89.77
<i>ViM</i>	93.63	40.84	85.00	59.57	82.17	61.76	86.79	52.41	91.36	45.33	66.46	91.50	84.24	58.57
<i>ASH-S</i>	3.79	99.91	10.77	99.56	14.87	99.02	33.22	96.20	17.33	98.67	41.80	96.00	20.30	98.23
<i>OptFS</i>	92.20	46.02	85.96	61.69	85.10	64.02	86.11	54.32	91.02	49.68	69.97	91.30	85.06	61.17
<i>Mahalanobis</i>	94.20	34.23	87.62	58.74	86.25	<u>60.61</u>	89.58	47.84	93.00	40.45	74.35	91.05	87.50	55.49
<i>CIDER</i>	93.78	37.16	86.78	60.33	84.55	65.51	<u>92.12</u>	<u>33.97</u>	<u>93.52</u>	<u>35.82</u>	<u>79.50</u>	<u>84.51</u>	<u>88.37</u>	<u>52.88</u>
<i>PALM</i>	95.65	24.74	84.64	<u>56.83</u>	79.85	63.13	93.90	25.09	93.95	31.52	83.19	77.75	88.53	46.51
<i>CARef</i>	94.81	29.79	<u>87.37</u>	58.04	85.40	63.09	90.04	45.09	93.39	38.82	76.70	89.70	87.95	54.09
<i>CADRef</i>	94.50	33.08	86.84	60.16	84.92	65.09	89.27	49.17	92.85	44.85	76.49	89.15	87.48	56.92

Table 15. Results of OOD detection on ImageNet-1k benchmark with ConvNeXt-B. ↑ indicates that higher values are better, while ↓ indicates that lower values are better. All values are percentages, with the best and second-best results being **highlighted** and underlined, respectively.

Methods	iNaturalist		SUN		Places		Textures		OpenImage-O		ImageNet-O		Average	
	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓	AU↑	FP↓
<i>MSP</i>	89.04	49.30	82.99	61.46	81.95	64.91	84.33	57.32	86.25	55.89	57.62	92.10	80.36	63.50
<i>MaxLogit</i>	90.54	37.76	80.88	55.30	78.52	61.36	84.81	45.18	82.97	51.04	45.37	93.05	77.18	57.28
<i>ODIN</i>	75.26	75.28	60.07	83.23	55.45	86.82	75.86	68.35	64.37	82.93	56.56	92.15	64.60	81.46
<i>Energy</i>	88.79	43.36	76.62	62.14	73.14	68.26	82.78	46.37	76.69	59.99	40.03	94.65	73.01	62.46
<i>GEN</i>	94.41	28.58	87.36	50.14	85.44	56.12	89.07	42.70	91.18	40.67	65.53	89.10	85.50	51.22
<i>ReAct</i>	80.15	68.51	62.81	81.64	58.01	86.14	74.11	63.78	66.68	76.46	40.92	96.90	63.78	78.91
<i>DICE</i>	80.47	66.61	67.78	76.60	63.51	83.42	75.38	60.53	63.90	75.45	28.18	98.10	63.20	76.78
<i>ViM</i>	<u>95.26</u>	27.62	80.81	63.22	75.83	67.38	<u>91.30</u>	41.21	93.06	37.38	75.79	84.55	85.34	53.56
<i>ASH-S</i>	47.01	90.66	65.40	80.89	65.26	84.48	65.48	73.95	49.06	89.75	39.32	98.05	55.26	86.30
<i>OptFS</i>	81.08	63.02	79.33	68.12	78.40	71.79	77.54	66.22	79.82	65.97	54.35	89.95	75.09	70.84
<i>Mahalanobis</i>	93.42	37.82	85.24	67.94	83.65	69.50	88.05	59.79	92.25	43.33	<u>74.92</u>	89.00	86.25	61.23
<i>CIDER</i>	92.20	42.15	85.38	64.35	84.20	67.99	86.39	60.76	90.58	48.55	72.00	87.30	85.13	61.85
<i>PALM</i>	76.82	61.79	64.24	74.81	60.14	78.74	76.63	61.32	72.94	68.00	64.48	<u>86.85</u>	69.21	71.92
<i>CARef</i>	95.11	<u>26.30</u>	<u>87.61</u>	54.21	<u>85.49</u>	59.71	91.18	<u>38.62</u>	93.30	<u>35.97</u>	73.52	89.70	<u>87.70</u>	<u>50.75</u>
<i>CADRef</i>	95.35	24.23	87.82	<u>52.19</u>	85.68	<u>58.26</u>	91.38	37.06	<u>93.29</u>	35.75	72.87	89.45	87.73	49.49

Table 16. Results of OOD detection on ImageNet-1k benchmark with MaxViT-T. ↑ indicates that higher values are better, while ↓ indicates that lower values are better. All values are percentages, with the best and second-best results being **highlighted** and underlined, respectively.