Supplementary Materials

A. More Results

Full Comparison on ImageNet Benchmark. Full comparisons with more competitors on the ImageNet benchmark are illustrated in Tab. 3, including the results of traditional methods [5, 8, 10, 13, 14, 21, 23] by fine-tuning CLIP visual encoder with labeled training samples following [11]. Compared with other traditional OOD detection methods that require fine-tuning, our ANTS is training-free and achieves better OOD detection performance on ImageNet-1k.

Results on OpenOOD Benchmark. The OpenOOD benchmark results are presented in Tab. 1, including the results of other methods requiring manually labeled training samples. Compared with other methods [9, 15, 19] that require manually labeled training samples, our ANTS surpasses other CLIP-based zero-shot OOD detection methods on both far-OOD and near-OOD tasks in the OpenOOD benchmark.

Table 1. Full results on OpenOOD Benchmark, where ImageNet is adopted as the ID dataset.

Methods	FPR95 ↓		AUROC ↑	
	Near-OOD	Far-OOD	Near-OOD	Far-OOD
Requires Manually Labeled Training Samples				
GEN [15]	–	–	78.97	90.98
AugMix [9] + ReAct [20]	–	–	79.94	93.70
RMDS [19]	–	–	80.09	92.60
AugMix [9] + ASH [4]	–	–	82.16	96.05
Zero Shot (No Training Required)				
MCM [16]	79.02	68.54	60.11	84.77
NegLabel [11]	68.18	27.34	76.92	93.30
EOE [2]	82.93	46.73	66.94	89.14
AdaNeg [27]	67.51	17.31	76.70	96.43
ANTS	60.98	15.38	82.15	96.50

Detailed OOD Detection Results of ANTS on OpenOOD Benchmark. The detailed OOD detection results of ANTS on the OpenOOD benchmark are shown in Tab. 2. We see that ANTS achieves good results in both Far-OOD and Near-OOD scenarios, demonstrating its strong adaptability to different OOD situations.

Near-OOD Detection. We alternate the use of ImageNet-10 and ImageNet-20 as ID and OOD datasets. The results are shown in Tab. 4. We see that ANTS significantly outperforms EOE and AdaNeg under the two experimental settings, demonstrating the advantage of our constructed adaptive negative text space.

Other Instantiations of Adaptive Weighted Score. As shown in Tab. 5, we consider various instantiations of the adaptive weighted score function. Specifically, we examine functions that employ only the ENS score, only the VSNL score, as well as logarithmic, exponential, and fractional functions. The results indicate that using the fractional function in conjunction with the ENS score and VSNL score yields the best results.

Table 2. Detailed OOD detection results of ANTS on the OpenOOD benchmark, where ImageNet is adopted as the ID dataset.

Near-/Far-OOD	Datasets	FPR95 ↓	AUROC ↑
Near-OOD	SSB-hard	62.73	81.97
	NINCO	59.23	82.32
	Mean	60.98	82.15
Far-OOD	iNaturalist	0.76	99.60
	Textures	16.50	96.78
	OpenImage-O	28.89	93.13
	Mean	15.38	96.50

Robustness to Domain Shift. To verify the robustness of ANTS to domain shift, we test several variants of ImageNet with different domain shifts. As shown in Tab. 6, ANTS outperforms MCM and NegLabels on the four different ID datasets, highlighting the remarkable robustness of our ANTS to domain shifts.

OOD Detection Results with Various ID Datasets. We compare the zero-shot OOD detection performance of ANTS and other methods on several ID datasets. As illustrated in Tab. 7, ANTS outperforms the other three methods across different ID datasets.

Robustness of ANTS to Noisy Negative Images. We investigate the impact of different rates of ID noise in negative images on the model OOD detection performance, as shown in Fig. 1. When the ID ratio of negative images is relatively low (less than 0.5), our ANTS achieves excellent OOD detection performance across different ID ratios. This demonstrates the robustness of ANTS against ID noise.

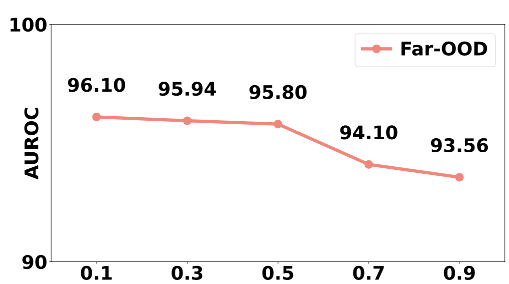


Figure 1. Robustness of ANTS to Noisy Negative Images

B. Case Study

We visualize the adaptive negative text space with some ID and OOD images for a case study, as shown in Fig. 2. We see that images from the ID dataset exhibit high similarity to the ID classes. For Far-OOD, as shown in Fig. 3, the images from the OOD dataset demonstrate high similarity to the ENS space. For Near-OOD, as shown in Fig. 4, the images from the OOD dataset show a high similarity to the VSNL space.

Table 3. OOD detection results by using ImageNet-1k as the ID dataset. ViTB/16 is used as the encoder.

Methods	INaturalist		SUN		Places		Textures		Average	
	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓
Training-required (or with Fine-tuning)										
MSP [8]	87.44	58.36	79.73	73.72	79.67	74.41	79.69	71.93	81.63	69.61
ODIN [13]	94.65	30.22	87.17	54.04	85.54	55.06	87.85	51.67	88.80	47.75
Energy [14]	95.33	26.12	92.66	35.97	91.41	39.87	86.76	57.61	91.54	39.89
GradNorm [10]	72.56	81.50	72.86	82.00	73.70	80.41	70.26	79.36	72.35	80.82
ViM [23]	93.16	32.19	87.19	54.01	83.75	60.67	87.18	53.94	87.82	50.20
KNN [21]	94.52	29.17	92.67	35.62	91.02	39.61	85.67	64.35	90.97	42.19
VOS [5]	94.62	28.99	92.57	36.88	91.23	38.39	86.33	61.02	91.19	41.32
NPOS [22]	96.19	16.58	90.44	43.77	89.44	45.27	88.80	46.12	91.22	37.93
ZOC [6]	86.09	87.30	81.20	81.51	83.39	73.06	76.46	98.90	81.79	85.19
CLIPN [24]	95.27	23.94	93.93	26.17	92.28	33.45	90.93	40.83	93.10	31.10
LSN [18]	95.83	21.56	94.35	26.32	91.25	34.48	90.42	38.54	92.26	30.22
LoCoOp [17]	93.93	29.45	90.32	41.13	90.54	44.15	93.24	33.06	92.01	36.95
ID-Like [1]	98.19	8.98	91.64	42.03	90.57	44.00	94.32	25.27	93.68	30.07
NegPrompt [12]	90.49	37.79	92.25	32.11	91.16	35.52	88.38	43.93	90.57	37.34
CoVer [26]	95.98	22.55	93.42	32.85	90.27	40.71	90.14	43.39	92.45	34.88
SCT [12]	95.86	13.94	95.33	20.55	92.24	29.86	89.06	41.51	93.27	26.47
LAPT [29]	99.63	1.16	96.01	19.12	92.01	33.01	91.06	40.32	94.68	23.40
Zero Shot (No Training Required)										
MCM [16]	94.59	32.20	92.25	38.80	90.31	46.20	86.12	58.50	90.82	43.93
EOE [2]	97.52	12.29	95.73	20.40	92.95	30.16	85.64	57.63	92.96	30.09
NegLabel [11]	99.49	1.91	95.49	20.53	91.64	35.59	90.22	43.56	94.21	25.40
OODD [25]	99.36	2.22	95.01	21.49	87.10	44.76	93.27	30.69	93.69	24.79
CLIPScope [7]	99.61	1.29	96.77	15.56	93.54	28.45	91.41	38.37	95.30	20.88
AdaNeg [28]	99.71	0.59	97.44	9.50	94.55	34.34	94.93	31.27	96.66	18.92
CSP [3]	99.60	1.54	96.66	13.66	92.90	29.32	93.86	25.52	95.76	17.51
ANTS	**99.75**	**0.54**	**98.77**	**5.43**	96.10	20.21	**96.38**	**18.52**	**97.75**	**11.20**

Table 4. Near OOD detection results.

Method	ImageNet-10 (ID) ImageNet-20 (OOD)		ImageNet-20 (ID) ImageNet-10 (OOD)	
	AUROC ↑	FPR95 ↓	AUROC ↑	FPR95 ↓
MCM [16]	98.71	5.00	97.87	17.40
NegLabel [11]	96.86	5.10	96.46	14.60
EOE [2]	99.09	4.20	98.10	13.93
AdaNeg [28]	99.17	3.00	97.25	14.40
ANTS (Ours)	**99.56**	**2.00**	**98.31**	**8.90**

Table 5. OOD detection results of other instantiations of Adaptive Weighted Score.

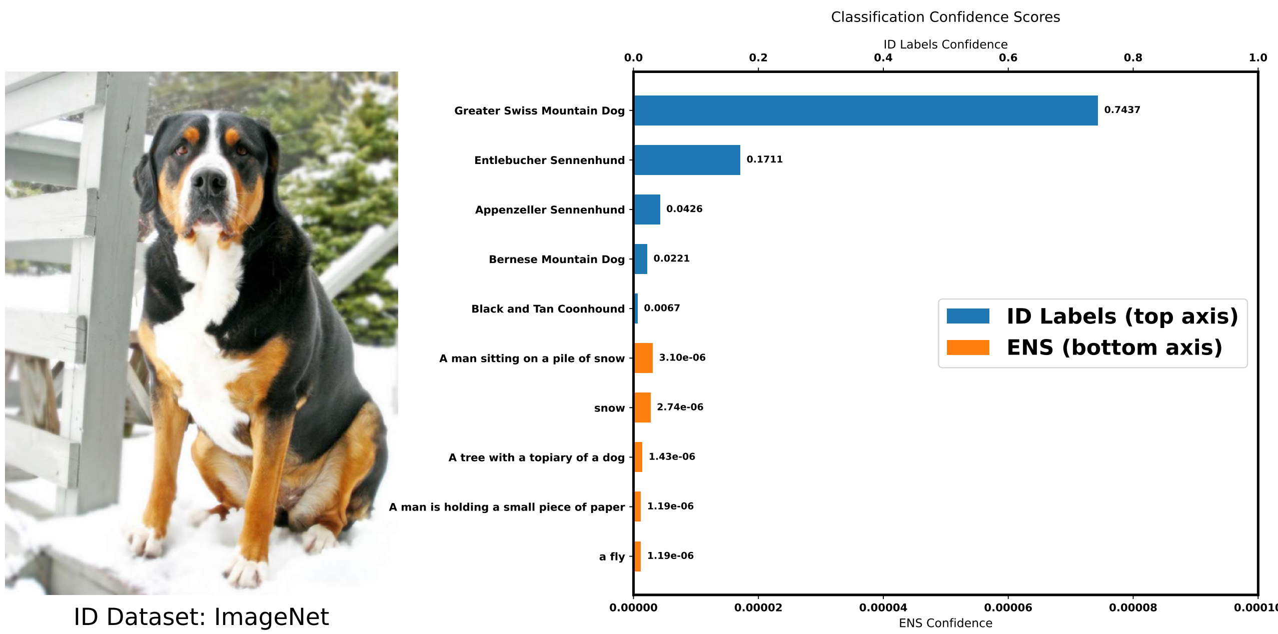
Methods	FPR95 ↓		AUROC ↑	
	Near-OOD	Far-OOD	Near-OOD	Far-OOD
Only ENS	66.23	21.70	77.58	95.81
Only VSNL	62.36	31.15	79.10	92.46
Logarithmic	61.58	21.04	79.60	95.83
Exponential	62.08	22.80	78.98	95.68
Fractional	**59.23**	**19.96**	**82.32**	**96.40**

Table 6. The robustness of zero-shot OOD detection to domain shift. The VITB/16 CLIP encoder is used.

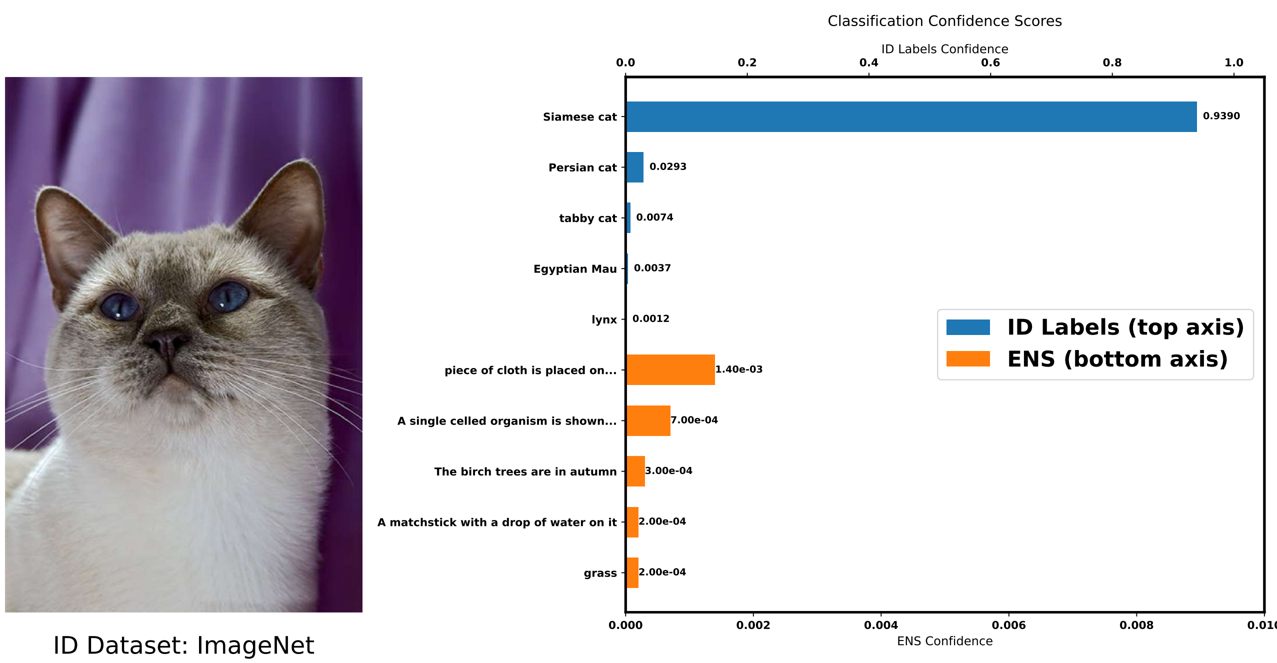
ID Dataset	Methods	OOD datasets INaturalist		SUN		Places		Textures		Average	
		AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓
ImageNet-S	MCM	87.74	63.06	85.35	67.24	81.19	70.64	74.77	79.59	82.26	70.13
	NegLabel [11]	99.34	2.24	94.93	22.73	90.78	38.62	89.29	46.10	93.59	27.42
	ANTS (Ours)	99.75	0.80	98.62	6.24	94.17	20.95	93.39	26.08	96.48	13.52
ImageNet-A	MCM [16]	79.50	76.85	76.19	79.78	70.95	80.51	61.98	86.37	72.16	80.88
	NegLabel [11]	98.80	4.09	89.83	44.38	82.88	60.10	80.25	64.34	87.94	43.23
	ANTS (Ours)	98.70	3.55	96.06	21.36	89.38	32.61	93.12	25.55	94.31	20.77
ImageNet-R	MCM [16]	83.22	71.51	80.31	74.98	75.53	76.67	67.66	83.72	76.68	76.72
	NegLabel [11]	99.58	1.60	96.03	15.77	91.97	29.48	90.60	35.67	94.54	20.63
	ANTS (Ours)	99.32	1.31	98.17	10.80	93.80	20.54	94.40	22.84	96.42	13.88
ImageNet-V2	MCM [16]	91.79	45.90	89.88	50.73	86.52	56.25	81.51	69.57	87.43	55.61
	NegLabel [11]	99.40	2.47	94.46	25.69	90.00	42.03	88.46	48.90	93.08	29.77
	ANTS (Ours)	99.50	1.32	98.45	7.98	94.28	22.01	95.93	19.75	97.04	12.77

Table 7. Zero-shot OOD detection performance with different ID datasets. The VITB/16 CLIP encoder is used.

ID Dataset	Methods	OOD datasets INaturalist		SUN		Places		Textures		Average	
		AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓
CUB-200-2011	MCM [16]	98.24	9.83	99.10	4.93	98.57	6.65	98.75	6.97	98.66	7.09
	EOE [2]	99.98	0.07	100.00	0.01	99.92	0.28	100.00	0.00	99.98	0.09
	NegLabel [11]	99.96	0.18	99.99	0.02	99.90	0.33	99.99	0.01	99.96	0.13
	ANTS (Ours)	99.96	0.03	99.99	0.00	99.83	0.02	100.00	0.00	99.95	0.01
STANFORD-CARS	MCM [16]	99.77	0.05	99.95	0.02	99.89	0.24	99.96	0.02	99.89	0.08
	EOE [2]	99.99	0.00	99.99	0.01	99.97	0.11	100.00	0.00	99.99	0.03
	NegLabel [11]	99.99	0.01	99.99	0.01	99.99	0.03	99.99	0.01	99.99	0.01
	ANTS (Ours)	100.00	0.00	100.00	0.00	99.97	0.00	100.00	0.00	99.99	0.00
Food-101	MCM [16]	99.78	0.64	99.75	0.90	99.58	1.86	98.62	4.04	99.43	1.86
	EOE [2]	99.98	0.06	100.00	0.00	99.97	0.14	99.01	2.61	99.74	0.70
	NegLabel [11]	99.99	0.01	99.99	0.01	99.99	0.01	99.60	1.61	99.90	0.40
	ANTS (Ours)	99.99	0.01	100.00	0.00	99.98	0.05	99.72	0.13	99.92	0.05
Oxford-IIIT Pet	MCM [16]	99.38	2.85	99.73	1.06	99.56	2.11	99.81	0.80	99.62	1.70
	EOE [2]	100.00	0.00	99.99	0.01	99.97	0.14	99.97	0.11	99.98	0.07
	NegLabel [11]	99.38	2.85	99.73	1.06	99.56	2.11	99.81	0.80	99.62	1.70
	ANTS (Ours)	100.00	0.00	100.00	0.03	99.97	0.03	99.97	0.00	99.99	0.02

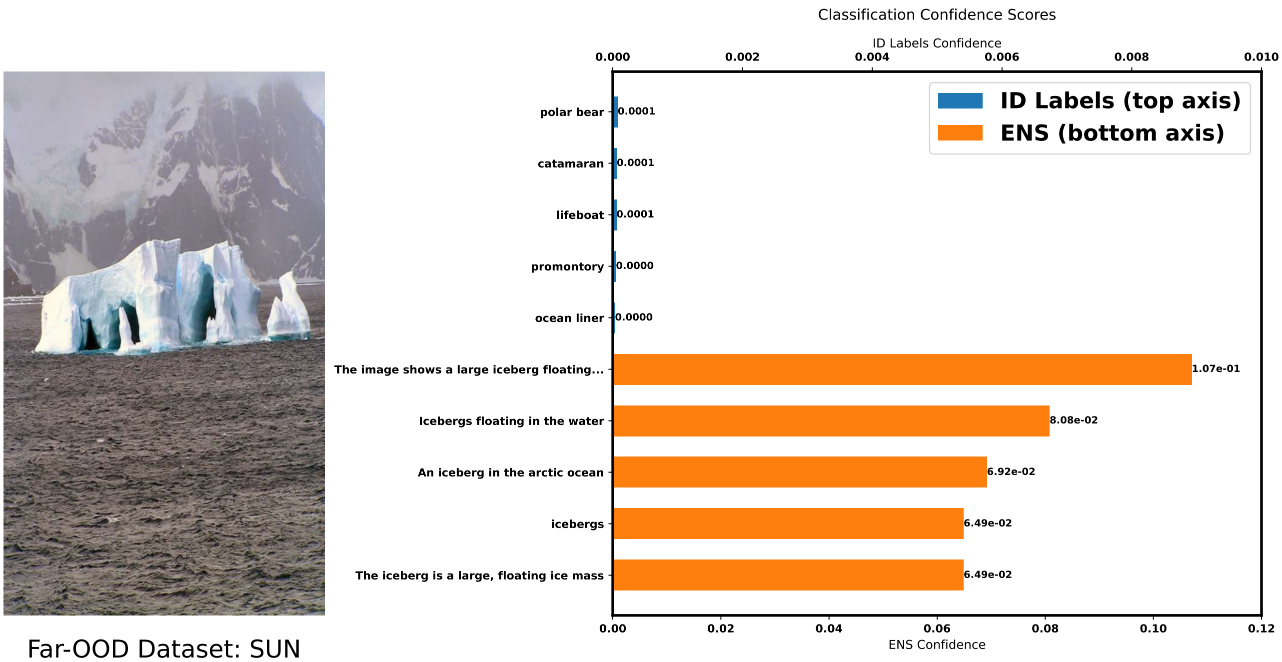


(a)

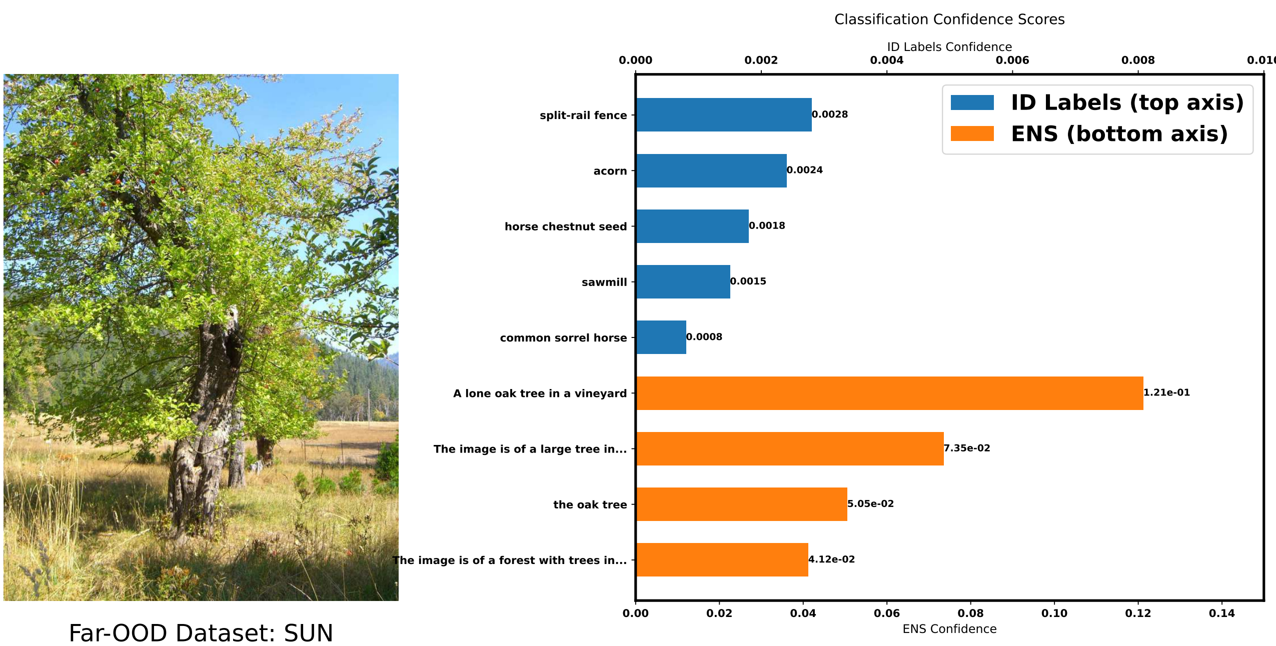


(b)

Figure 2. Example cases of classification confidence scores on ImageNet ID samples.

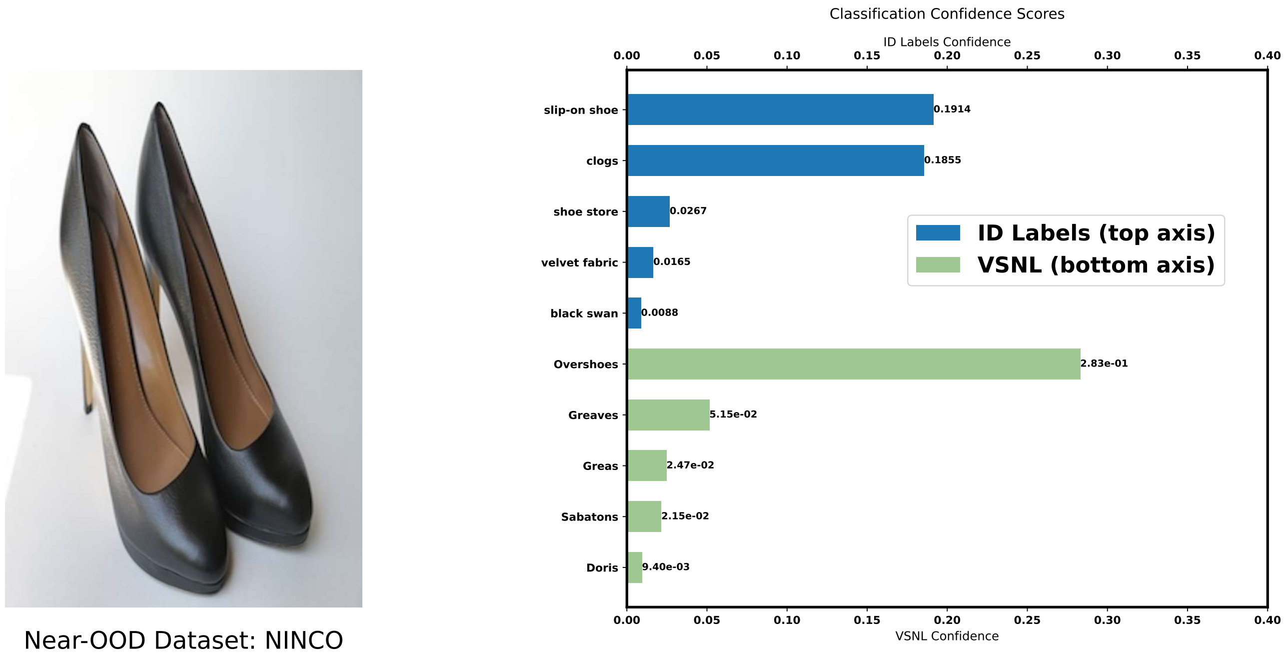


(a)

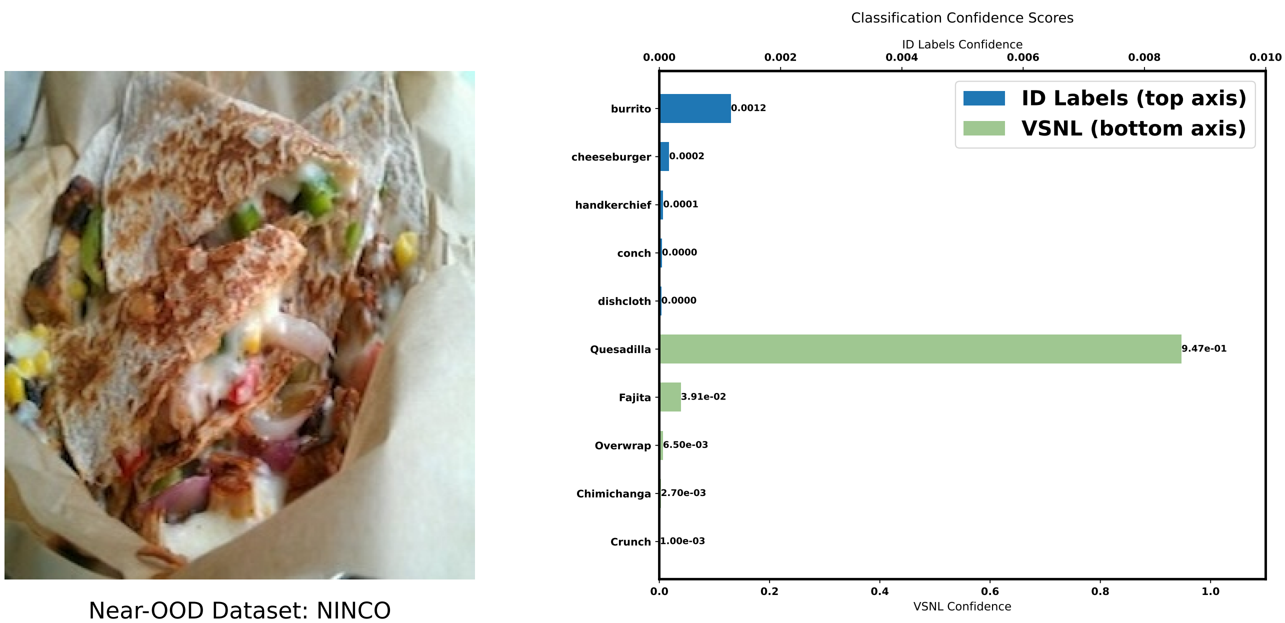


(b)

Figure 3. Example cases of classification confidence scores on Far-OOD samples of the SUN dataset.



(a)



(b)

Figure 4. Example cases of classification confidence scores on Near-OOD samples of the NINCO dataset.

References

[1] Yichen Bai, Zongbo Han, Bing Cao, Xiaoheng Jiang, Qinghua Hu, and Changqing Zhang. Id-like prompt learning for few-shot out-of-distribution detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17480–17489, 2024.

[2] Chentao Cao, Zhun Zhong, Zhanke Zhou, Yang Liu, Tongliang Liu, and Bo Han. Envisioning outlier exposure by large language models for out-of-distribution detection. *arXiv preprint arXiv:2406.00806*, 2024.

[3] Mengyuan Chen, Junyu Gao, and Changsheng Xu. Conjugated semantic pool improves ood detection with pre-trained vision-language models. *arXiv preprint arXiv:2410.08611*, 2024.

[4] Andrija Djurisic, Nebojsa Bozanic, Arjun Ashok, and Rosanne Liu. Extremely simple activation shaping for out-of-distribution detection. *arXiv preprint arXiv:2209.09858*, 2022.

[5] Xuefeng Du, Xin Wang, Gabriel Gozum, and Yixuan Li. Unknown-aware object detection: Learning what you don't know from videos in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13678–13688, 2022.

[6] Sepideh Esmaeilpour, Bing Liu, Eric Robertson, and Lei Shu. Zero-shot out-of-distribution detection based on the pre-trained model clip. In *Proceedings of the AAAI conference on artificial intelligence*, pages 6568–6576, 2022.

[7] Hao Fu, Naman Patel, Prashanth Krishnamurthy, and Farshad Khorrami. Clipscope: Enhancing zero-shot ood detection with bayesian scoring. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 5346–5355. IEEE, 2025.

[8] Dan Hendrycks and Kevin Gimpel. A baseline for detecting misclassified and out-of-distribution examples in neural networks. *arXiv preprint arXiv:1610.02136*, 2016.

[9] Dan Hendrycks, Norman Mu, Ekin D Cubuk, Barret Zoph, Justin Gilmer, and Balaji Lakshminarayanan. Augmix: A simple data processing method to improve robustness and uncertainty. *arXiv preprint arXiv:1912.02781*, 2019.

[10] Rui Huang, Andrew Geng, and Yixuan Li. On the importance of gradients for detecting distributional shifts in the wild. *Advances in Neural Information Processing Systems*, 34:677–689, 2021.

[11] Xue Jiang, Feng Liu, Zhen Fang, Hong Chen, Tongliang Liu, Feng Zheng, and Bo Han. Negative label guided ood detection with pretrained vision-language models. *arXiv preprint arXiv:2403.20078*, 2024.

[12] Tianqi Li, Guansong Pang, Xiao Bai, Wenjun Miao, and Jin Zheng. Learning transferable negative prompts for out-of-distribution detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17584–17594, 2024.

[13] Shiyu Liang, Yixuan Li, and Rayadurgam Srikant. Enhancing the reliability of out-of-distribution image detection in neural networks. *arXiv preprint arXiv:1706.02690*, 2017.

[14] Weitang Liu, Xiaoyun Wang, John Owens, and Yixuan Li. Energy-based out-of-distribution detection. *Advances in neural information processing systems*, 33:21464–21475, 2020.

[15] Xixi Liu, Yaroslava Lochman, and Christopher Zach. Gen: Pushing the limits of softmax-based out-of-distribution detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 23946–23955, 2023.

[16] Yifei Ming, Ziyang Cai, Jiuxiang Gu, Yiyou Sun, Wei Li, and Yixuan Li. Delving into out-of-distribution detection with vision-language representations. *Advances in neural information processing systems*, 35:35087–35102, 2022.

[17] Atsuyuki Miyai, Qing Yu, Go Irie, and Kiyoharu Aizawa. Locoop: Few-shot out-of-distribution detection via prompt learning. *Advances in Neural Information Processing Systems*, 36, 2024.

[18] Jun Nie, Yonggang Zhang, Zhen Fang, Tongliang Liu, Bo Han, and Xinmei Tian. Out-of-distribution detection with negative prompts. In *The Twelfth International Conference on Learning Representations*, 2024.

[19] Jie Ren, Stanislav Fort, Jeremiah Liu, Abhijit Guha Roy, Shreyas Padhy, and Balaji Lakshminarayanan. A simple fix to mahalanobis distance for improving near-ood detection. *arXiv preprint arXiv:2106.09022*, 2021.

[20] Yiyou Sun, Chuan Guo, and Yixuan Li. React: Out-of-distribution detection with rectified activations. *Advances in Neural Information Processing Systems*, 34:144–157, 2021.

[21] Yiyou Sun, Yifei Ming, Xiaojin Zhu, and Yixuan Li. Out-of-distribution detection with deep nearest neighbors. In *International Conference on Machine Learning*, pages 20827–20840. PMLR, 2022.

[22] Leitian Tao, Xuefeng Du, Xiaojin Zhu, and Yixuan Li. Non-parametric outlier synthesis. *arXiv preprint arXiv:2303.02966*, 2023.

[23] Haoqi Wang, Zhizhong Li, Litong Feng, and Wayne Zhang. Vim: Out-of-distribution with virtual-logit matching. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4921–4930, 2022.

[24] Hualiang Wang, Yi Li, Huifeng Yao, and Xiaomeng Li. Clipn for zero-shot ood detection: Teaching clip to say no. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1802–1812, 2023.

[25] Yifeng Yang, Lin Zhu, Zewen Sun, Hengyu Liu, Qinying Gu, and Nanyang Ye. Oodd: Test-time out-of-distribution detection with dynamic dictionary. *arXiv preprint arXiv:2503.10468*, 2025.

[26] Boxuan Zhang, Jianing Zhu, Zengmao Wang, Tongliang Liu, Bo Du, and Bo Han. What if the input is expanded in ood detection? *Advances in Neural Information Processing Systems*, 37:21289–21329, 2024.

[27] Jinsong Zhang, Qiang Fu, Xu Chen, Lun Du, Zelin Li, Gang Wang, Shi Han, Dongmei Zhang, et al. Out-of-distribution detection based on in-distribution data patterns memorization with modern hopfield energy. In *The Eleventh International Conference on Learning Representations*, 2022.

[28] Yabin Zhang and Lei Zhang. Adaneg: Adaptive negative proxy guided ood detection with vision-language models. *arXiv preprint arXiv:2410.20149*, 2024.

[29] Yabin Zhang, Wenjie Zhu, Chenhang He, and Lei Zhang. Lapt: Label-driven automated prompt tuning for ood detection with vision-language models. *arXiv preprint arXiv:2407.08966*, 2024.