

Revisit PCA-based technique for Out-of-distribution Detection

Xiaoyuan Guan^{1,3*} Zhouwu Liu^{1,3*} Wei-Shi Zheng^{1,3} Yuren Zhou^{1†}
Ruixuan Wang^{1,2,3†}

¹School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China

²Peng Cheng Laboratory, Shenzhen, China

³Key Laboratory of Machine Intelligence and Advanced Computing, MOE, Guangzhou, China

{guanxy36,liuzhw33}@mail2.sysu.edu.cn, wszheng@ieee.org, {zhouyuren,wangruix5}@mail.sysu.edu.cn

A. Dataset details

Dataset details in benchmark I and II

ImageNet-100. We use the following categories that are randomly sampled from ImageNet-1K to create ImageNet-100: n01986214, n04200800, n03680355, n03208938, n02963159, n03874293, n02058221, n04612504, n02841315, n02099712, n02093754, n03649909, n02114712, n03733281, n02319095, n01978455, n04127249, n07614500, n03595614, n04542943, n02391049, n04540053, n03483316, n03146219, n02091134, n02870880, n04479046, n03347037, n02090379, n10148035, n07717556, n04487081, n04192698, n02268853, n02883205, n02002556, n04273569, n02443114, n03544143, n03697007, n04557648, n02510455, n03633091, n02174001, n02077923, n03085013, n03888605, n02279972, n04311174, n01748264, n02837789, n07613480, n02113712, n02137549, n02111129, n01689811, n02099601, n02085620, n03786901, n04476259, n12998815, n04371774, n02814533, n02009229, n02500267, n04592741, n02119789, n02090622, n02132136, n02797295, n01740131, n02951358, n04141975, n02169497, n01774750, n02128757, n02097298, n02085782, n03476684, n03095699, n04326547, n02107142, n02641379, n04081281, n06596364, n03444034, n07745940, n03876231, n09421951, n02672831, n03467068, n01530575, n03388043, n03991062, n02777292, n03710193, n09256479, n02443484, n01728572, n03903868.

We use the following OOD test data in ImageNet-1K and ImageNet-100 benchmark. The selected categories are disjoint with ImageNet-1K.

iNaturalist: iNaturalist contains 675,170 training and validation images from 5089 natural fine-grained cate-

gories, including 13 major categories such as plants, insects, birds, and mammals. We randomly sample 10000 images that are disjoint from ImageNet-1k for evaluation.

SUN: The Scene UNDERstanding (SUN) contains 397 well-sampled categories to evaluate the performance of scene recognition algorithms. We randomly sample 10000 images that are disjoint from ImageNet-1k for evaluation.

Places: Places is a scene image dataset, which contains 10 million pictures and more than 400 different types of scene environments. We randomly sample 10000 images that are disjoint from ImageNet-1k for evaluation.

Textures: Describable Textures Dataset (DTD) is a texture dataset, including 5640 images, which can be divided into 47 categories according to human perception. We use the entire Textures dataset for evaluation.

Dataset details in CIFAR benchmark

For CIFAR benchmarks, we use the following OOD test dataset: **SVHN:** Street View House Number (SVHN) consists of the house numbers extracted from Google Street View images. We use the entire of its test set as OOD examples (26032 images).

Tiny ImageNet: Similar to ImageNet, Tiny ImageNet is an image classification dataset, which contains 200 categories, and each category contains 50 test images.

LSUN: LSUN is a scene understanding dataset, which mainly includes scene images of bedrooms, fixed houses, living rooms, classrooms, etc. We randomly sample 10000 images as out-of-distribution examples.

Textures: Describable Textures Dataset (DTD) is a texture dataset, including 5640 images, which can be divided into 47 categories according to human perception. We use the entire Textures dataset for evaluation.

iSUN: iSUN is a large-scale eye tracking dataset which contains 20608 natural scene images from the SUN database. We sample 8925 images with no conceptual overlap with CIFAR as out-of-distribution examples.

Places365: It contains large-scale photographs of scenes

* Authors contributed equally

† Corresponding author

Table 1. Detailed results when using CIFAR-10 as ID data on seven common OOD benchmark datasets. \uparrow indicates larger values are better and \downarrow indicates smaller values are better

OOD datasets	Metrics	Out-of-distribution Detection Methods												
		MSP	Mahalanobis	ODIN	DICE	ViM	Energy	BATS	ReAct	DICE+ReAct	LogitNorm	Ours		
CIFAR-10 ResNet34	SVHN	FPR95	33.79	41.38	43.80	36.67	29.66	20.65	26.37	24.19	38.22	17.40	11.22	
		AUROC	94.18	94.10	86.52	90.64	95.06	95.25	94.75	94.40	90.64	96.96	98.02	
	Tiny-ImageNet	FPR95	32.13	86.48	14.21	14.79	54.27	12.85	21.25	14.72	15.09	3.52	3.03	
		AUROC	95.23	87.68	97.68	97.22	92.44	97.91	96.25	97.56	97.21	99.17	99.32	
	LSUN-R	FPR95	42.62	52.67	23.66	35.13	38.13	25.64	30.02	26.69	34.83	11.04	13.69	
		AUROC	92.31	92.54	93.70	92.79	94.00	94.13	93.72	94.16	92.79	97.92	97.59	
	LSUN-C	FPR95	20.47	92.99	7.03	6.67	49.00	5.32	10.67	6.72	6.67	0.48	0.34	
		AUROC	96.86	88.48	98.74	98.70	93.84	99.05	98.30	98.87	98.70	99.80	99.90	
	iSUN	FPR95	45.17	52.26	26.03	40.95	37.49	28.17	32.47	29.18	40.95	11.39	13.46	
		AUROC	91.71	92.41	93.18	90.92	93.84	93.46	92.90	93.46	89.92	97.86	97.56	
	Textures	FPR95	49.95	38.95	45.73	50.73	28.19	38.83	37.71	39.50	50.73	31.51	22.09	
		AUROC	89.66	94.24	83.19	86.60	94.94	88.91	91.71	89.33	87.17	94.71	96.18	
	Places365	FPR95	53.69	55.89	52.08	49.65	47.58	41.49	41.84	40.30	49.67	29.90	27.56	
		AUROC	87.84	90.19	82.55	84.70	90.78	88.26	90.27	89.55	84.72	94.31	94.54	
	Average	FPR95	39.69	60.09	30.37	33.51	40.62	24.71	28.62	25.90	33.74	15.03	13.06	
		AUROC	92.54	91.38	90.79	91.65	93.56	93.85	93.98	93.90	91.65	97.25	97.59	
	CIFAR10 WideResNet28-10	SVHN	FPR95	45.43	83.53	51.98	48.99	12.85	35.61	39.91	36.07	47.16	24.31	19.26
			AUROC	90.36	80.05	81.79	83.18	97.32	89.80	91.50	90.03	86.74	95.75	96.77
		Tiny-ImageNet	FPR95	24.61	99.38	8.77	0.84	27.71	8.70	17.47	10.28	0.62	1.58	1.39
			AUROC	96.31	43.41	98.13	99.74	94.61	98.45	96.90	98.18	99.80	99.47	99.56
LSUN-R		FPR95	40.76	95.76	20.99	31.51	20.21	25.95	29.04	27.25	31.31	9.51	10.64	
		AUROC	92.78	61.18	94.61	91.51	96.04	94.25	94.79	94.43	92.04	98.07	97.97	
LSUN-C		FPR95	16.10	98.95	4.57	0.12	12.68	3.53	9.57	4.53	0.15	0.27	0.20	
		AUROC	97.41	42.18	98.87	99.91	97.49	99.24	98.29	99.07	99.79	99.80	99.89	
iSUN		FPR95	44.69	94.42	25.30	35.15	20.29	30.27	33.48	31.32	35.65	10.69	12.98	
		AUROC	92.00	63.46	93.60	90.49	95.94	93.30	93.79	93.58	89.21	97.84	97.79	
Textures		FPR95	45.02	64.96	44.45	56.99	14.77	39.38	35.85	37.23	54.45	34.33	24.19	
		AUROC	40.73	80.71	82.64	78.33	96.46	88.30	92.48	90.35	84.33	91.98	94.19	
Places365		FPR95	45.54	81.56	46.67	54.24	42.07	36.81	35.20	37.12	56.25	26.86	25.60	
		AUROC	89.20	69.52	81.96	80.96	88.27	88.85	92.32	89.23	79.56	93.51	93.81	
Average		FPR95	37.45	88.37	28.98	32.56	21.51	25.75	28.66	26.26	32.82	15.36	13.32	
		AUROC	92.64	62.93	90.23	89.16	95.16	93.17	94.30	93.56	90.01	96.63	97.14	

with 365 scene categories. There are 100 images per category in the test set, we use the entire test set for evaluation.

B. Proof for $\mathbf{U}_k(h(\mathbf{x}) - \boldsymbol{\mu})$ coincide with $\mathbf{M}(h(\mathbf{x}) - \boldsymbol{\mu})$

Denote by $\mathbf{z} = h(\mathbf{x})$ the feature representation for input \mathbf{x} , and $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_K\}$ the eigenvectors for principal directions, which is also can be viewed as orthogonal bases of principal coordinate system. View \mathbf{U}_k as a transform operator in the original feature space (*fill the dimensions that were not involved in the transformation*), we have

$$\mathbf{U}_k(h(\mathbf{x}) - \boldsymbol{\mu}) = \sum_{i=1}^k \langle \mathbf{z} - \boldsymbol{\mu}, \mathbf{u}_i \rangle \cdot \mathbf{u}_i + \sum_{j=k+1}^K 0 \cdot \mathbf{u}_j, \quad (1)$$

and after back projection, we have

$$\mathbf{M}(h(\mathbf{x}) - \boldsymbol{\mu}) = \mathbf{U}_k^T \mathbf{U}_k (h(\mathbf{x}) - \boldsymbol{\mu}) \quad (2)$$

$$= \sum_{i=1}^k \langle \mathbf{z} - \boldsymbol{\mu}, \mathbf{u}_i \rangle \cdot \langle \mathbf{u}_i, \mathbf{u}_i^T \rangle \cdot \mathbf{u}_i^T + \sum_{j=k+1}^K 0 \cdot \langle \mathbf{u}_j, \mathbf{u}_j^T \rangle \cdot \mathbf{u}_j^T \quad (3)$$

$$= \sum_{i=1}^k \langle \mathbf{z} - \boldsymbol{\mu}, \mathbf{u}_i \rangle \cdot \mathbf{u}_i^T + \sum_{j=k+1}^K 0 \cdot \mathbf{u}_j^T \quad (4)$$

$$= \mathbf{U}_k^T (h(\mathbf{x}) - \boldsymbol{\mu}) \quad (5)$$

C. Detailed experimental results

We report the detailed performance for all seven test OOD datasets for models trained on CIFAR-10 and CIFAR-100 respectively in Table 1 and Table 2.

Table 2. Detailed results when using CIFAR-100 as ID data on seven common OOD benchmark datasets. \uparrow indicates larger values are better and \downarrow indicates smaller values are better

OOD datasets	Metrics	Out-of-distribution Detection Methods											
		MSP	ODIN	Mahalanobis	Energy	ViM	ReAct	BATS	DICE	DICE+ReAct	LogitNorm	Ours	
CIFAR100 ResNet34	SVHN	FPR95	81.31	85.55	98.81	76.73	75.80	43.10	60.48	53.65	48.18	64.65	24.80
		AUROC	77.61	76.83	54.62	81.52	82.31	92.22	89.50	89.97	91.19	88.69	95.08
	Tiny-ImageNet	FPR95	71.07	63.32	97.83	60.44	72.18	54.40	57.77	62.40	59.99	18.99	60.34
		AUROC	84.04	85.57	49.77	86.76	85.86	87.19	87.51	86.11	87.01	96.68	86.49
	LSUN-R	FPR95	74.39	38.18	97.38	58.02	37.70	41.60	38.64	85.83	84.17	93.39	26.96
		AUROC	82.08	92.96	53.35	88.40	93.40	91.58	92.96	76.80	78.80	69.28	95.08
	LSUN-C	FPR95	76.52	68.48	99.48	61.12	89.49	52.61	65.69	32.29	32.01	10.57	49.64
		AUROC	80.65	84.16	36.04	85.88	73.78	87.32	84.84	93.49	93.71	98.22	89.65
	iSUN	FPR95	75.99	41.92	94.64	61.82	38.51	41.95	40.80	86.01	82.23	94.39	28.13
		AUROC	80.96	91.84	58.99	86.95	93.01	91.29	92.29	77.83	79.65	68.36	94.75
	Textures	FPR95	81.72	71.19	75.67	80.73	49.47	53.31	59.65	68.17	66.74	81.56	35.80
		AUROC	76.90	79.28	76.95	77.17	89.33	87.65	86.59	82.49	83.96	74.45	92.53
	Places365	FPR95	79.81	78.74	97.17	78.06	78.09	70.80	75.13	82.90	80.26	80.30	79.83
		AUROC	77.30	75.54	51.33	76.72	78.17	79.73	78.96	76.31	78.04	77.59	76.73
Average	FPR95	77.26	63.91	94.43	68.13	63.03	51.11	56.88	67.32	64.81	63.41	43.64	
	AUROC	79.93	83.74	54.43	83.99	85.12	88.14	87.52	83.99	84.62	81.90	90.04	
CIFAR100 WideResNet28-10	SVHN	FPR95	81.83	88.35	54.47	81.09	49.55	76.69	79.61	72.91	84.17	27.40	59.00
		AUROC	77.54	74.20	86.96	78.36	90.69	85.43	84.21	80.51	79.67	95.76	90.16
	Tiny-ImageNet	FPR95	53.78	39.11	99.55	35.53	81.43	45.81	47.30	9.96	22.51	14.66	46.75
		AUROC	89.34	92.87	32.75	93.62	83.80	90.60	90.57	97.69	95.60	97.43	89.49
	LSUN-R	FPR95	80.09	57.99	45.73	78.67	49.04	67.95	67.03	85.69	97.85	82.51	50.05
		AUROC	77.38	86.38	90.28	80.34	90.31	84.54	85.15	79.72	53.43	79.00	89.40
	LSUN-C	FPR95	64.57	54.16	99.59	46.02	68.64	54.87	60.83	6.21	6.58	4.80	44.67
		AUROC	86.67	90.64	26.74	92.25	86.55	89.83	87.03	98.61	98.53	99.07	91.87
	iSUN	FPR95	81.62	58.69	45.46	80.96	44.45	67.93	68.89	88.65	98.51	84.67	46.16
		AUROC	76.98	86.06	90.26	79.49	90.82	84.75	84.86	78.08	54.78	77.78	90.22
	Textures	FPR95	80.66	75.09	33.58	78.09	33.00	66.26	70.01	77.29	72.80	69.43	37.73
		AUROC	77.19	78.09	91.82	78.16	93.14	85.85	84.58	76.52	74.21	80.46	91.82
	Places365	FPR95	78.41	76.63	93.51	79.50	82.91	75.64	77.29	82.64	92.81	75.12	78.19
		AUROC	78.40	78.34	59.38	77.97	76.32	81.36	80.47	76.15	62.69	81.06	79.60
Average	FPR95	74.42	64.29	66.99	68.55	58.43	65.02	67.28	60.48	67.89	51.23	51.79	
	AUROC	80.50	83.80	68.31	82.88	87.37	86.05	85.27	83.90	74.13	87.22	88.94	