上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

SJTU

Kernel PCA for Out-of-Distribution Detection

Kun FANG, Qinghua TAO, Kexin LV,
Mingzhen HE, Xiaolin HUANG, Jie YANG

The Thirty-Eighth Annual Conference on
Neural Information Processing Systems

2024.11

饮 水 思 源 • 爱 国 荣 校

Kernel PCA for Out-of-Distribution Detection

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

- Out-of-distribution detection[1]
  - **In-Distribution (InD)**: data following the training distribution of neural networks
  - **Out-of-Distribution (OoD)**: data NOT from the training distribution
  - A bi-classification task: scoring function $S(\cdot)$, threshold s
  - Evaluation metrics: FPR with a 95% TPR, AUROC

$$D(\boldsymbol{x}) = \begin{cases} \text{InD}, & S(\boldsymbol{x}) > s, \\ \text{OoD}, & S(\boldsymbol{x}) < s. \end{cases}$$

- Related work
  - Base on logits [2], features [3], gradients [4]
  - Post hoc. v.s. training regularization [5]

[1] Yang, Jingkang, et al. "Generalized out-of-distribution detection: A survey." International Journal of Computer Vision (2024): 1-28.
[2] Weitang Liu, et al. Energy-based out-of-distribution detection. Advances in neural information processing systems, 33:21464–21475, 2020.
[3] Yiyou Sun, et al. Out-of-distribution detection with deep nearest neighbors. In International Conference on Machine Learning, pages 20827–20840. PMLR, 2022.
[4] Rui Huang, et al. On the importance of gradients for detecting distributional shifts in the wild. Advances in Neural Information Processing Systems, 34:677–689, 2021
[5] Hsu, et al. Generalized odin: Detecting out-of-distribution image without learning from out-of-distribution data. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 10951–10960, 2020.

Kernel PCA for Out-of-Distribution Detection

SHANGHAI JIAO TONG UNIVERSITY

Background

PCA for Out-of-Distribution Detection

- PCA learns **a subspace characterizing InD information** from the **training data (InD)**

  Projecting new data $\hat{\boldsymbol{x}}$ into the subspace and re-projecting back, we can obtain the **reconstruction error** $e(\hat{\boldsymbol{x}}) = ||\boldsymbol{U}_q \boldsymbol{U}_q^T(\hat{\mathbf{z}} - \boldsymbol{\mu}) - (\hat{\mathbf{z}} - \boldsymbol{\mu})||_2$

- An ideal case: InD data with a small $e(\hat{\boldsymbol{x}})$; OoD data with a large $e(\hat{\boldsymbol{x}})$.
- Existing works[6]:
  - Empirically verifying that PCA is insufficient in separating OoD and InD.
  - No further explorations on the reasons behind. A simple combination with other scores.

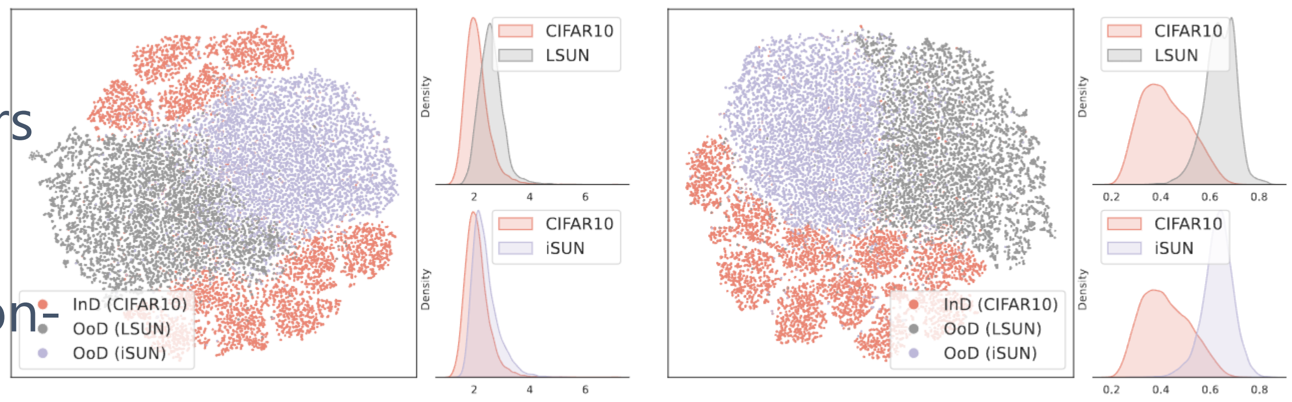
[6] Xiaoyuan Guan, et al. Revisit pca-based technique for out-of-distribution detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 19431–19439, 2023.

Kernel PCA for Out-of-Distribution Detection

Motivation

Considering that PCA is linear, we propose

- The **non-linearity** in InD and OoD data hinders PCA from learning a suitable subspace.
- **Kernel PCA** [7] is introduced to leverage the non-linear kernel to learn *a subspace where the disparity between InD and OoD gets pronounced.*



(a) T-SNE of z and PCA reconstruction errors. (b) T-SNE of $\Phi(z)$ and KPCA reconstruction errors.

Challenges we face:

- How to find **appropriate kernels**?
- How to leverage KPCA in **large-scale data**?

(Storage and computation of the kernel matrix)

Solutions we propose:

- A kernel perspective on the KNN method[8]
- Explicit feature mappings to approximate kernels, avoiding computations on the kernel matrix

[7] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Kernel principal component analysis. In International conference on artificial neural networks, pages 583–588. Springer, 1997.
[8] Yiyou Sun, et al. Out-of-distribution detection with deep nearest neighbors. In International Conference on Machine Learning, pages 20827–20840. PMLR, 2022.

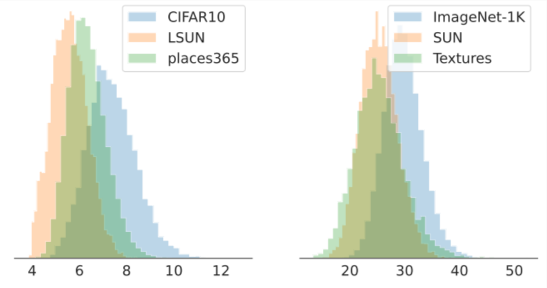
Kernel PCA for Out-of-Distribution Detection

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

Methodology

Non-linear kernel design

- **Cosine kernel**
  - Normalize the imbalanced feature norms
  - $k_{\cos}(\boldsymbol{z}_1, \boldsymbol{z}_2) = \frac{\boldsymbol{z}_1^T \boldsymbol{z}_2}{||\boldsymbol{z}_1||_2 \cdot ||\boldsymbol{z}_2||_2} = \phi_{\cos}^T(\boldsymbol{z}_1)\phi_{\cos}(\boldsymbol{z}_2)$



- **Cosine-Gaussian kernel**
  - l_2 distance on l_2-normalized features benefits OoD detection[8]
  - A Gaussian kernel preserves the l_2 distance

    $k_{\text{gau}}(\boldsymbol{z}_1, \boldsymbol{z}_2) = e^{-\gamma||\boldsymbol{z}_1 - \boldsymbol{z}_2||_2^2}$

Explicit feature mappings of kernels

- **Cosine kernel**

  l_2 normalization
  - $\Phi(\cdot) \triangleq \phi_{\cos}(\cdot)$
  - Computation complexity $\boldsymbol{\mathcal{O}(1)}$

- **Cosine-Gaussian kernel**

  l_2 normalization+ l_2 distance
  - Random Fourier Features[9] to approximate k_{gau}
  - $\Phi(\cdot) \triangleq \phi_{\text{RFF}}(\phi_{\cos}(\cdot))$
  - Computation complexity $\boldsymbol{\mathcal{O}(M), N_{tr} \gg M}$

$$\phi_{\text{RFF}}(\boldsymbol{z}) \triangleq \sqrt{\frac{2}{M}}[\phi_1(\boldsymbol{z}), \dots, \phi_M(\boldsymbol{z})]$$

$$\phi_j(\boldsymbol{z}) = \cos(\boldsymbol{z}^T \boldsymbol{\omega}_j + u_j), j = 1, \dots, M$$

[8] Yiyou Sun, et al. Out-of-distribution detection with deep nearest neighbors. In International Conference on Machine Learning, pages 20827–20840. PMLR, 2022
[9] li Rahimi and Benjamin Recht. Random features for large-scale kernel machines. Advances in neural information processing systems, 20, 2007.

Kernel PCA for Out-of-Distribution Detection

Experiments – OoD detection

- Comparisons with the KNN method[8]
  - Better performance, cheaper complexity
- Comparisons with regularized PCA[6]
  - Better performance, indicating superior non-linearity

method	OoD data sets								AVERAGE	
	iNaturalist		SUN		Places		Textures			
	FPR↓	AUROC↑	FPR↓	AUROC↑	FPR↓	AUROC↑	FPR↓	AUROC↑	FPR↓	AUROC↑
MSP	54.99	87.74	70.83	80.86	73.99	79.76	68.00	79.61	66.95	81.99
+ PCA [8]	51.47	88.95	67.64	82.71	71.20	80.87	60.53	85.86	62.71	84.60
+ CoP	**50.84**	**89.21**	**67.35**	**82.81**	**70.96**	**81.08**	**59.96**	**86.21**	**62.28**	**84.83**
+ CoRP	**43.70**	**91.70**	**61.79**	**85.43**	**66.67**	**83.07**	**45.67**	**91.86**	**54.46**	**88.02**
Energy	55.72	89.95	59.26	85.89	64.92	82.86	53.72	85.99	58.41	86.17
+ PCA [8]	50.36	91.09	54.19	87.55	64.13	84.00	29.33	92.59	49.50	88.81
+ CoP	**45.13**	**92.15**	**52.33**	**88.01**	**61.49**	**84.96**	**29.13**	92.57	**47.02**	**89.42**
+ CoRP	**26.85**	**95.15**	**40.38**	**90.76**	**51.26**	**87.35**	**12.11**	**97.17**	**32.65**	**92.61**
ReAct	20.38	96.22	24.20	94.20	33.85	91.58	47.30	89.80	31.43	92.95
+ PCA [8]	10.17	97.97	18.50	95.80	27.31	93.39	18.67	95.95	18.66	95.76
+ CoP	13.30	97.44	19.80	95.37	29.92	92.64	**15.90**	**96.51**	19.73	95.49
+ CoRP	10.77	97.85	18.70	95.75	28.69	93.13	**12.57**	**97.21**	**17.68**	**95.98**
BATS	42.26	92.75	44.70	90.22	55.85	86.48	33.24	93.33	44.01	90.69
+ PCA [8]	29.66	94.49	38.11	90.03	51.70	87.25	13.46	97.09	33.23	92.56
+ CoP	**27.14**	**94.87**	**34.36**	**91.96**	**47.68**	**87.87**	**11.97**	**97.33**	**30.29**	**93.01**
+ CoRP	**18.74**	**96.31**	**28.02**	**93.49**	**41.41**	**89.78**	**9.45**	**97.79**	**24.41**	**94.34**
ODIN	47.66	89.66	60.15	84.59	67.89	81.78	50.23	85.62	56.48	85.41
Mahalanobis	97.00	52.65	98.50	42.41	98.40	41.79	55.80	85.01	87.43	55.47
ViM	68.86	87.13	79.62	81.67	83.81	77.80	14.95	96.74	61.81	85.83
DICE	26.66	94.49	36.08	90.98	47.63	87.73	32.46	90.46	35.71	90.92
DICE+ReAct	20.08	96.11	26.50	93.83	38.34	90.61	29.36	92.65	28.57	93.30
NNGuide	25.73	95.12	37.18	91.21	46.97	88.67	27.70	92.30	34.39	91.82

method	OoD data sets								AVERAGE	
	iNaturalist		SUN		Places		Textures			
	FPR↓	AUROC↑	FPR↓	AUROC↑	FPR↓	AUROC↑	FPR↓	AUROC↑	FPR↓	AUROC↑
Standard Training										
MSP[54]	54.99	87.74	70.83	80.86	73.99	79.76	68.00	79.61	66.95	81.99
ODIN[55]	47.66	89.66	60.15	84.59	67.89	81.78	50.23	85.62	56.48	85.41
Energy[57]	55.72	89.95	59.26	85.89	64.92	82.86	53.72	85.99	58.41	86.17
GODIN[82]	61.91	85.40	60.83	85.60	63.70	83.81	77.85	73.27	66.07	82.02
Mahalanobis[45]	97.00	52.65	98.50	42.41	98.40	41.79	55.80	85.01	87.43	55.47
KNN[77]	59.00	86.47	68.82	80.72	76.28	75.76	11.77	97.07	53.97	85.01
CoP (ours)	67.25	83.41	75.53	79.93	82.48	73.83	8.33	98.29	58.40	83.86
CoRP (ours)	50.07	89.32	62.56	83.74	72.76	78.91	9.02	98.14	**48.60**	**87.53**
Supervised Contrastive Learning										
MSP[54]	64.96	86.23	53.55	87.20	57.80	85.54	73.99	74.14	62.57	83.28
ODIN[55]	65.08	86.28	53.79	87.21	58.04	85.56	74.22	74.15	62.78	83.30
Energy[57]	48.13	91.28	49.57	88.54	54.40	86.90	70.66	75.83	55.69	85.64
SSD[73]	57.16	87.77	78.23	73.10	81.19	70.97	36.37	88.52	63.24	80.09
KNN[77]	30.18	94.89	48.99	88.63	59.15	84.71	15.55	95.40	38.47	90.91
CoP (ours)	29.85	94.79	44.99	90.62	56.77	86.19	10.28	97.35	**35.47**	**92.24**
CoRP (ours)	23.61	95.86	41.07	91.25	53.52	87.27	10.23	97.04	**32.11**	**92.86**

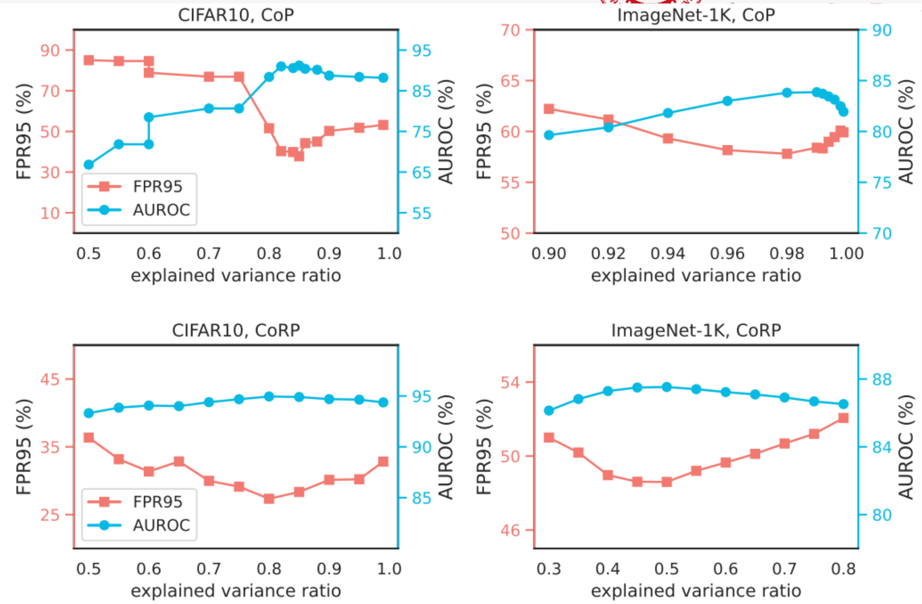
method	time and memory complexity	time consuming (ms, per sample)	storage
KNN	$O(N_{tr})$	≈ 15.59	≈ 20 GiB
CoP	$O(1)$	≈ 0.035	≈ 22 MiB
CoRP	$O(M)$	≈ 0.086	≈ 29 MiB

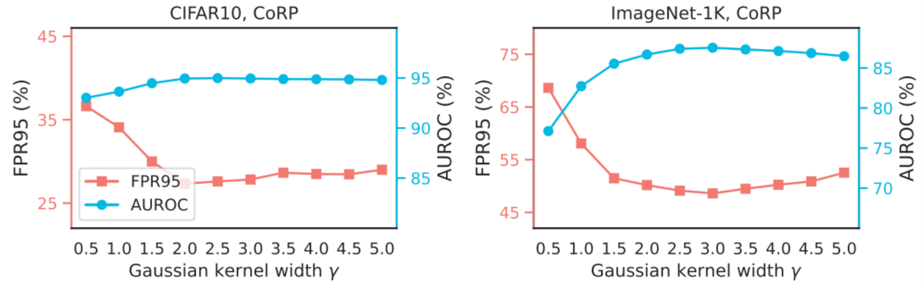
[6] Xiaoyuan Guan, et al. Revisit pca-based technique for out-of-distribution detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 19431–19439, 2023.
[8] Yiyou Sun, et al. Out-of-distribution detection with deep nearest neighbors. In International Conference on Machine Learning, pages 20827–20840. PMLR, 2022.

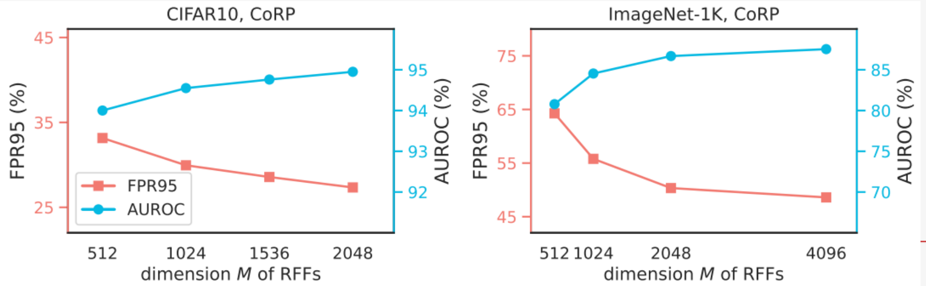
Kernel PCA for Out-of-Distribution Detection

Experiments

- Ablation studies: **effect of more kernels**
  - Cosine、Gaussian、Laplacian、Polynomial
  - Cosine-Laplacian、Cosine-Polynomial
- Sensitivity analysis: **effect of involved hyper-parameters**
  - Explained variance ratio
  - Gaussian kernel width
  - Number of RFFs







kernel	OoD data sets								AVERAGE	
	iNaturalist		SUN		Places		Textures			
	FPR↓	AUROC↑	FPR↓	AUROC↑	FPR↓	AUROC↑	FPR↓	AUROC↑	FPR↓	AUROC↑
PCA (no kernels)	95.46	52.01	97.98	44.86	97.99	45.19	46.22	87.77	84.41	57.46
Polynomial	96.03	53.07	98.26	42.84	97.85	45.02	95.50	47.96	96.91	47.22
Laplacian	94.65	50.25	94.68	50.29	95.28	49.80	94.66	50.34	94.82	50.17
Gaussian	94.46	50.83	95.17	50.33	94.80	50.46	95.09	50.80	94.88	50.60
Cosine (CoP)	67.25	83.41	75.53	79.93	82.48	73.83	8.33	98.29	58.40	83.86
Cosine-Polynomial	54.10	84.48	75.97	75.04	82.82	69.01	59.15	83.27	68.01	77.95
Cosine-Laplacian	76.18	77.95	77.54	76.70	84.47	70.16	11.97	97.57	62.54	80.60
Cosine-Gaussian (CoRP)	50.07	89.32	62.56	83.74	72.76	78.91	9.02	98.14	48.60	87.53

Kernel PCA for Out-of-Distribution Detection

Conclusion

- KPCA learns a subspace where the disparity between InD and OoD is pronounced.

- **Effective kernels** under the OoD detection task
  - Cosine kernel
  - Cosine-Gaussian kernel

- Resolving the challenge of KPCA in large-scale data
  - **Explicit feature mappings**
  - Significantly reduced time and memory complexity. $(\mathcal{O}(N_{tr}) \to \mathcal{O}(M),\ N_{tr} \gg M)$

上海交通大學

SHANGHAI JIAO TONG UNIVERSITY

Thanks!

饮 水 思 源　爱 国 荣 校