

Out-of-Distribution Detection Using Union of 1-Dimensional Subspaces

Alireza Zaeemzadeh, Niccolò Bisagno, Zeno Sambugaro, Nicola Conci,
Nazanin Rahnavard, Mubarak Shah

Joint work between University of Central Florida and University of Trento

<https://github.com/zaeemzadeh/OOD>
https://github.com/mmlab-cv/OOD_video

Out-of-distribution Detection

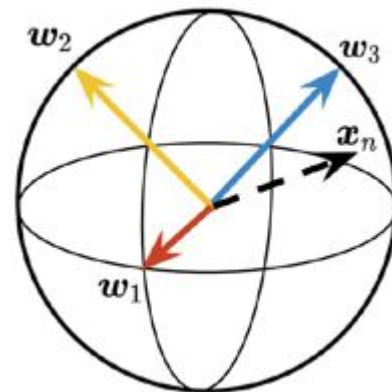
The goal of **out-of-distribution (OOD)** detection is to handle novel situations.

How can we make the in-distribution samples more distinguishable from OOD samples with unknown distribution?



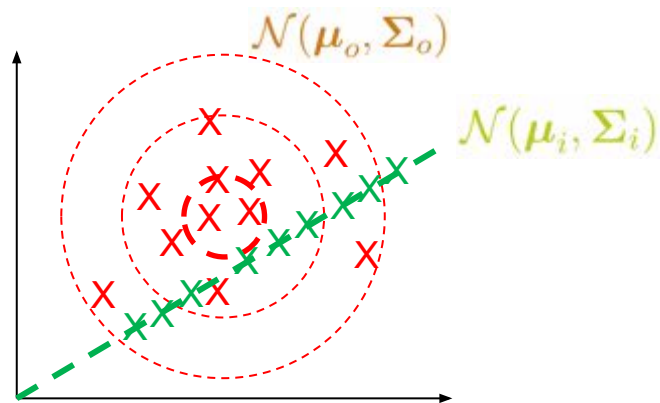
OOD Detection Using Union of 1D Subspaces

- Distribution-agnostic minimization of error
- No need for OOD samples to tune hyperparameters
- Does not rely on sample generation or reconstruction
- Easy to implement
- Modality-independent



Distribution-Agnostic Minimization of Error Probability

- Consider a binary classification problem:
- If one of the classes is degenerate, the detection can be performed with zero error probability.



Distribution-Agnostic Minimization of Error Probability

p_i : The probability that the sample belongs to $\mathcal{N}(\mu_i, \Sigma_i)$

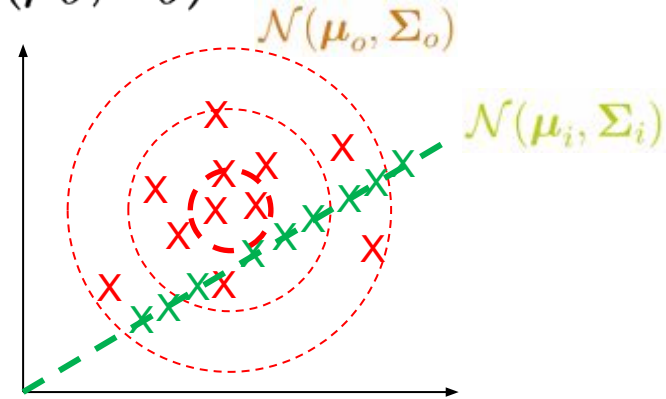
p_o : The probability that the sample belongs to $\mathcal{N}(\mu_o, \Sigma_o)$

$$\Delta = \mu_i - \mu_o$$

$$B = \frac{1}{8} \Delta^T \left(\frac{\Sigma_i + \Sigma_o}{2} \right)^{-1} \Delta + \frac{1}{2} \ln \left(\frac{\det\left(\frac{\Sigma_i + \Sigma_o}{2}\right)}{\sqrt{\det(\Sigma_i) \det(\Sigma_o)}} \right)$$

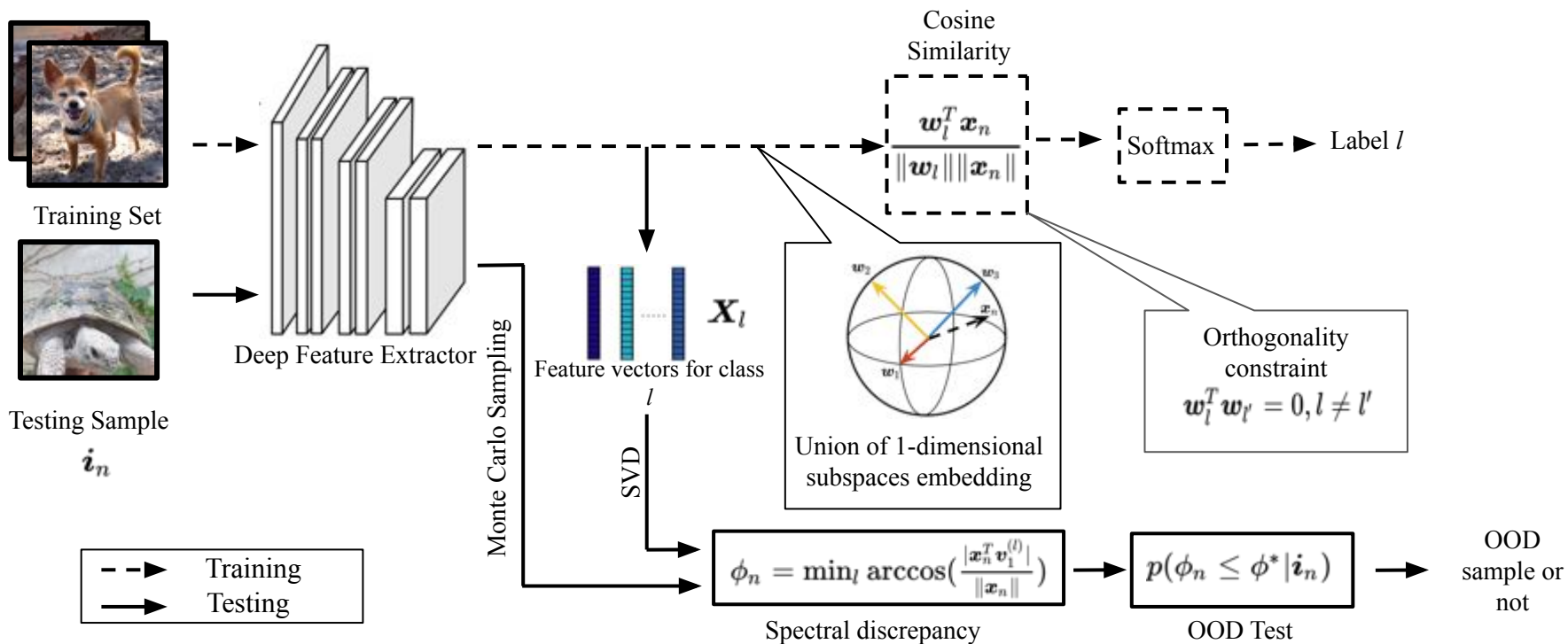
The classification error probability can be upper bounded by

$$p_e \leq \sqrt{p_i p_o} e^{-B}$$



The probability of error can be decreased by making the distribution of the known classes as compact as possible (e.g. *1D subspace*).

OOD Detection Using Union of 1-Dimensional Subspaces

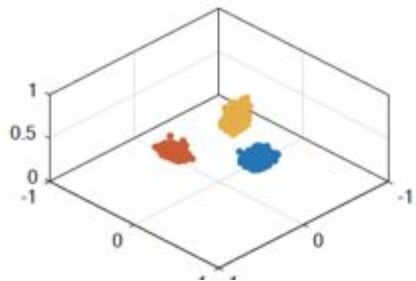


Experiments

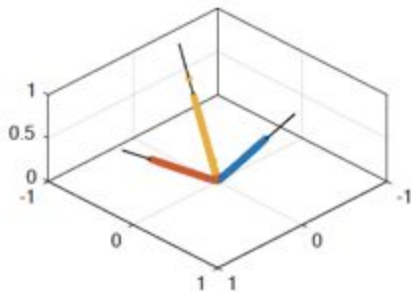
3-dimensional visualization of the features extracted from CIFAR10

Training Set

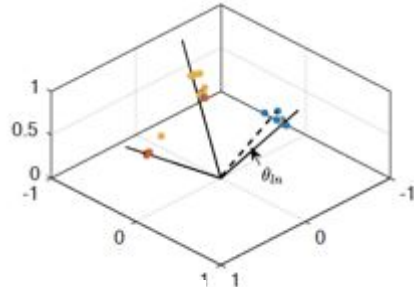
w/o Cosine similarity



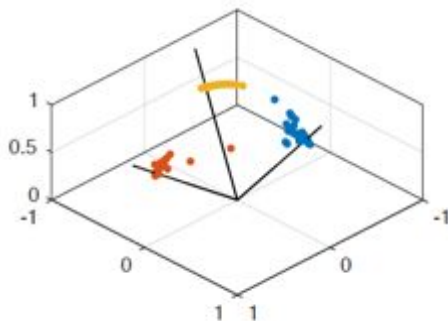
w/ Cosine similarity



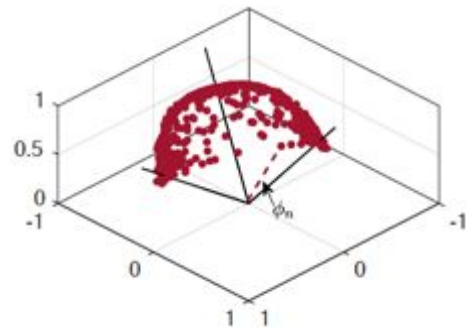
w/ Cosine similarity (normalized)



Test Set



In-distribution



Out-of-distribution

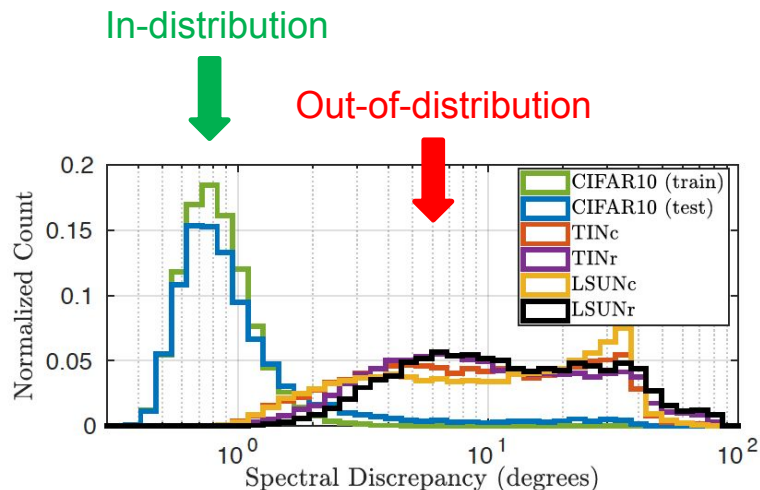
Experiments: Image Classification

The proposed spectral discrepancy measure can effectively distinguish between the ID (CIFAR10) and OOD samples.

AUROC: Area under receiver operating characteristic curve

AUPR: Area under precision recall curve

FPR/TPR: False/True Positive Rate



Training dataset	OOD dataset	FPR at 95% TPR	Detection Error	AUROC	AUPR In	AUPR Out
		↓	↓	↑	↑	↑
Softmax. Pred. [12]/OLTR [24]/ Ours						
CIFAR10	TINc	38.9/25.6/9.0	21.9/14.8/6.8	92.9/91.3/98.1	92.5/93.2/98.2	91.9/88.3/98.1
	TINr	45.6/28.8/7.6	25.3/15.8/6.2	91.0/90.3/98.5	89.7/92.3/98.6	89.9/87.1/98.4
	LSUNc	35.0/21.3/2.8	20.0/13.0/3.7	94.5/92.9/99.4	95.1/94.4/99.4	93.1/90.8/99.4
	LSUNr	35.0/21.7/3.4	20.0/13.2/3.8	93.9/92.6/99.3	93.8/94.4/99.4	92.8/90.0/99.3

OOD detection can be improved by making the known classes as compact as possible.

Experiments: Video Classification

Our method can generalise for multiple modalities (images, videos)

New baseline for OOD detection on action recognition in videos on UCF 50 and UCF 101 datasets

Training dataset	OOD dataset	FPR at 95% TPR	Detection Error	AUROC	AUPR In	AUPR Out
		↓	↓	↑	↑	↑
		SoftMax. Pred. (Baseline) [12]/		SoftMax. Pred. (Orthogonal Subs.) [12]/ Ours		
UCF50	UCF51	86.3/82.44/ 71.6	36.8/36.1/ 30.0	66.0/68.3/ 75.7	89.8/ 90.1 /74.3	25.6/27.8/ 72.5
HMDB25	HMDB26	82.0 /85.2/84.5	41.8/44.5/ 40.8	59.7/56.4/ 61.9	88.9 /87.6/65.4	20.4/19.7/ 56.6

Conclusions

- We can improve the OOD detection performance by engineering the distribution of in-distribution samples
- Our approach does not have hyperparameters, does not need extra information, and can be easily applied to existing methods with minimal change.
- Our approach is modality-independent and can be applied to different input types

Thank You!

<https://github.com/zaemzadeh/OOD>

https://github.com/mmlab-cv/OOD_video