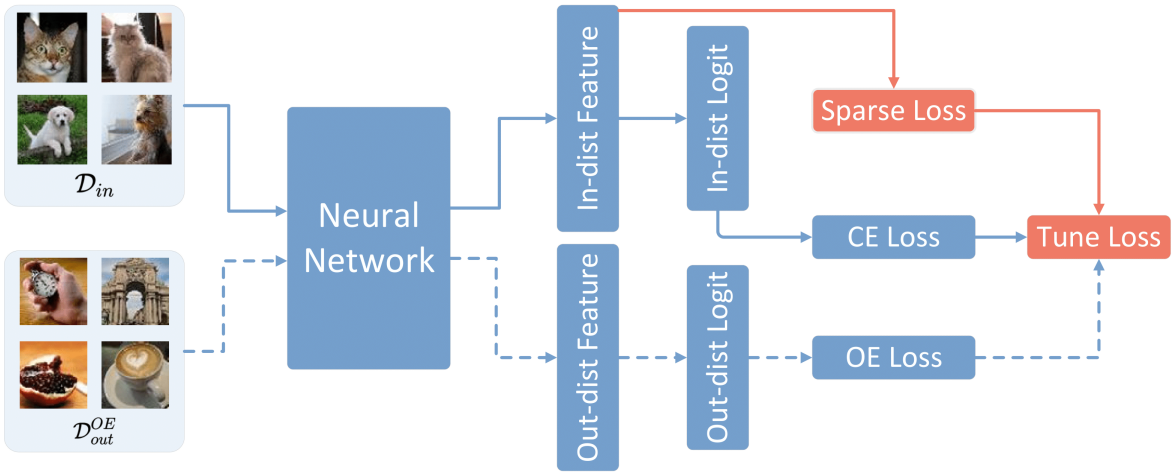
Sparsity-Regularized Out-of-distribution Detection

This repository is the implementation of Improving Energy-based OOD Detection by Sparsity Regularization by Qichao Chen, Wenjie Jiang, Kuan Li and Yi Wang. This method is a simple yet effective for improve Energy-based OOD Detection. Our code is modified from energy_ood.



Requirements

It is tested under Ubuntu Linux 18.04 and Python 3.7 environment, and requries some packages to be installed:

- PyTorch 1.4.0
- torchvision 0.5.0
- numpy 1.17.2

Training Pretrained Models

Please download the datasets in folder

```
./data/
```

Training pretrained classifier

```
python baseline.py cifar10
python baseline.py cifar100
```

Pretrained models are provided in folder

```
./CIFAR/snapshots/
```

Testing and Fine-tuning

Evaluate the pretrained model using energy-based detector

```
python test.py --model cifar10_wrn_pretrained --score energy
python test.py --model cifar100_wrn_pretrained --score energy
```

Fine-tune the pretrained model

```
python tune.py cifar10 --save ./snapshots/tune_sr
python tune.py cifar100 --save ./snapshots/tune_sr
```

Testing the detection performance of fine-tuned model

```
python test.py --model cifar10_wrn_s1_tune --score energy
python test.py --model cifar100_wrn_s1_tune --score energy
```

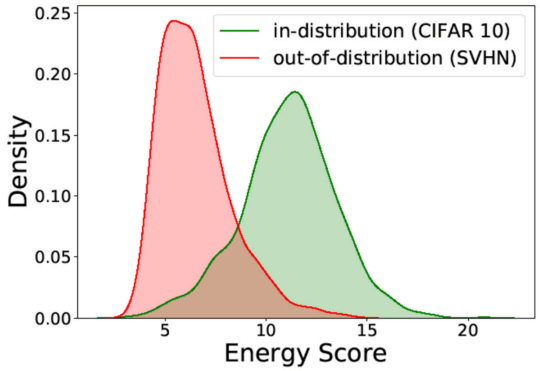
Results

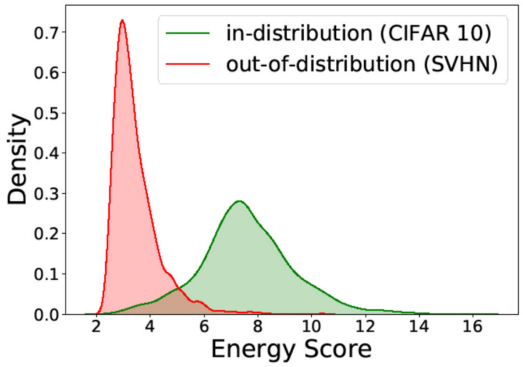
Our model achieves the following average performance on 6 OOD datasets:

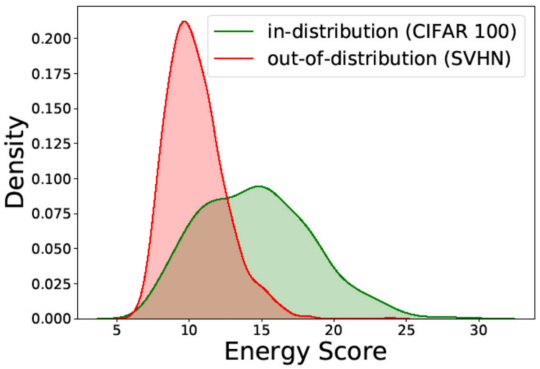
1. MSP vs energy score with and without fine-tuned on CIFAR-10

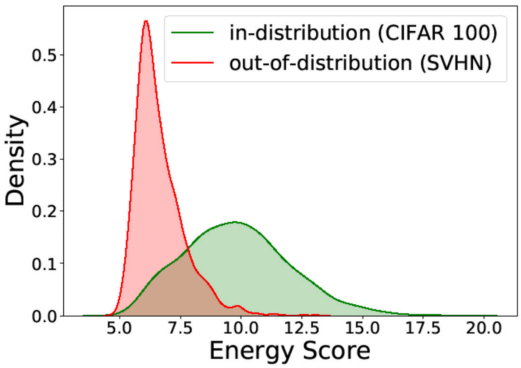
Model name	FPR95
MSP	51.35%
ODIN	35.59%
Mahalanobis	37.08%
EBD	33.01%
SR (Ours)	19.19%

2. CIFAR-10 (in-distribution) vs SVHN (out-of-distribution) Score Distributions









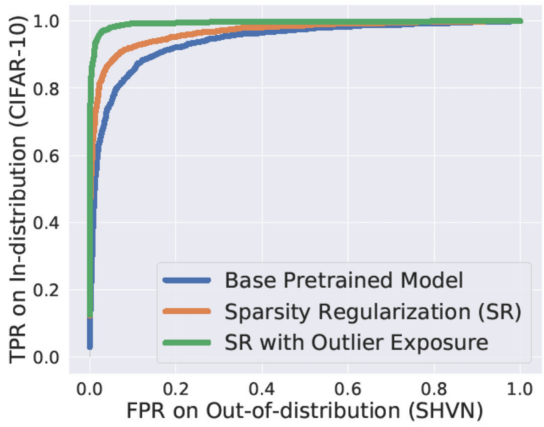
3. Performance among different baselines for WideResNet

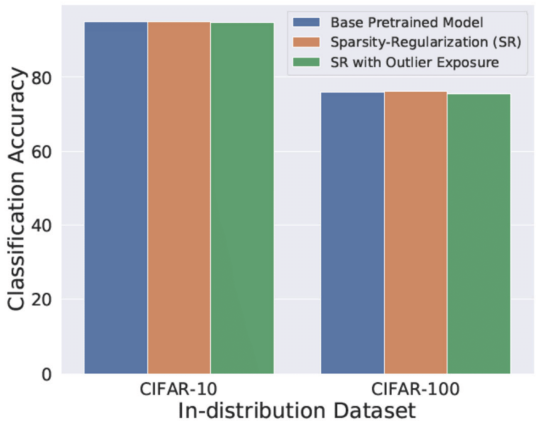
CIFAR-10:

Method	FPR95
Baseline	34.92%
Outlier Exposure	8.53%
Energy	3.32%
SROE (Ours)	4.15%

CIFAR-100:

Method	FPR95
Baseline	71.86%
Outlier Exposure	56.57%
Energy	49.28%
SROE (Ours)	23.84%





Outlier Datasets

These experiments make use of numerous outlier datasets. Links for less common datasets are as follows, 80 Million Tiny Images Textures, Places365, LSUN-C, LSUN-R, iSUN and SVHN.

Our **tiny** dataset available at here

in in out out out out

Citation

```
 @article{chen2022sparsity,
      title={Improving Energy-based Out-of-distribution Detection by Sparsity
Regularization},
      author={Chen, Qichao and Jiang, Wenjie and Li, Kuan and Wang, Yi},
      journal={Pacific-Asia Conference on Knowledge Discovery and Data Mining},
      year={2022}
 }
```