A Joint Representation Learning and Feature Modeling Approach for One-class Recognition

Pramuditha Perera and Vishal Patel

Introduction

One Class Classification(OCC) is an Extreme case in classification where knowledge of the classifier is limited to only a single class. Given training samples from a class, the one class classifier is expected to reject samples from any outside class. we learn generative features using the one-class data with a generative framework. We augment the learned features with the corresponding reconstruction errors to obtain augmented features. Then, we qualitatively identify a suitable feature distribution that reduces the redundancy in the chosen classifier space. Finally, we force the augmented features to take the form of this distribution using an adversarial framework.

Motivation

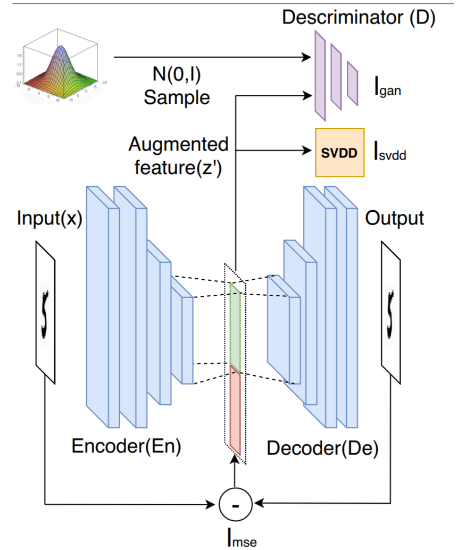
one-class modeling identifies the positive space in a given feature space. Objects appearing outside the positive space are identified as out-of-class samples. Redundant space could be identified as a part of the positive space in this approach. No guarantee that out-of-class samples will always get projected in the negative.

In **representation modeling**, a representation is learned from one class data. During inference, samples that are not represented well are identified as out of class samples. However, there is a possibility for out of-class samples to be well represented as well - Specially when the representation is generic.

Proposed Method

Modelling fails when out-of-class samples get projected inside the identified positive space. Provided that, latent space is smooth and each latent code inside the positive space corresponds to an in-class sample, failed cases can be identified considering the reconstruction error.

- Autoencoder network that is trained on reconstruction loss
  $l_{mse} = \|x - \hat{x}\|^2$ where, $\hat{x} = De(En(x))$
- Extend the latent space by appending MSE to the latent feature : Prevent out-of-class samples from entering positive space
  $l_{gan} = \mathbb{E}_{s\sim N(0,I)\in\mathbb{R}^{2k}}[logD(s)] + \mathbb{E}_{x\sim p_{z'}}[log(1 - D(z'))]$.
- Force extended latent features to follow a pre-determined distribution: Reduce redundant positive space and results in smooth latent codes
- Fit a one-class classifier on extended feature space



Results

We evaluated the effectiveness of the proposed method with recently published OCC methods by considering area under the ROC curve metric.

GTSRB STOP SIGN dataset

OCSVM [10]	67.5	*1.2*
KDE [7]	60.5	*1.7*
IF [7]	73.8	*0.9*
DCAE [19]	79.1	*3.0*
SDOCC [4]	77.8	*4.9*
DOCC [4]	80.3	*2.8*
Ours	**85.2**	*0.7*

MNIST

Class	AND*[6] (CVPR19)		OCGAN*[39] (CVPR19)		AE+SVDD		Ours	
0	99.3	*0.0*	**99.8**	*0.0*	96.8	*0.0*	99.6	*0.1*
1	**99.9**	*0.0*	**99.9**	*0.0*	99.3	*0.0*	98.8	*0.7*
2	95.9	*0.0*	94.2	*0.0*	83.4	*0.0*	**97.2**	*0.5*
3	**96.6**	*0.0*	96.3	*0.0*	86.8	*0.0*	95.5	*0.3*
4	95.6	*0.0*	**97.5**	*0.0*	92.4	*0.0*	95.7	*0.4*
5	96.4	*0.0*	**98.0**	*0.0*	75.8	*0.0*	96.3	*0.5*
6	**99.4**	*0.0*	99.1	*0.0*	93.1	*0.0*	98.8	*0.3*
7	98.0	*0.0*	**98.1**	*0.0*	92.6	*0.0*	95.7	*0.3*
8	95.3	*0.0*	93.9	*0.0*	88.9	*0.0*	**95.4**	*0.4*
9	**98.1**	*0.0*	**98.1**	*0.0*	93.7	*0.0*	97.7	*0.2*
Mean	**97.5**	*0.0*	**97.5**	*0.0*	90.2	*0.0*	97.1	*0.4*

CIFAR10

Class	AND*[6] (CVPR19)		OCGAN*[39] (CVPR19)		AE+SVDD		Ours	
Plane	73.5	*0.0*	**75.7**	*0.0*	55.2	*0.0*	66.4	*1.5*
Car	58.0	*0.0*	53.1	*0.0*	73.0	*0.0*	**78.5**	*0.6*
Bird	**69.0**	*0.0*	64.0	*0.0*	49.1	*0.0*	54.9	*0.6*
Cat	54.2	*0.0*	62.0	*0.0*	53.6	*0.0*	57.3	*0.6*
Deer	**76.1**	*0.0*	72.3	*0.0*	61.1	*0.0*	73.6	*0.1*
Dog	54.6	*0.0*	62.0	*0.0*	60.4	*0.0*	63.1	*0.4*
Frog	75.1	*0.0*	72.3	*0.0*	62.6	*0.0*	**80.8**	*0.1*
Horse	53.5	*0.0*	57.5	*0.0*	69.1	*0.0*	**72.0**	*1.1*
Ship	71.7	*0.0*	**82.0**	*0.0*	74.7	*0.0*	80.3	*0.6*
Truck	54.8	*0.0*	55.4	*0.0*	77.8	*0.0*	**79.9**	*1.0*
Mean	64.1	*0.0*	65.7	*0.0*	63.6	*0.0*	**70.7**	*0.7*