

**Anomaly detection -
one class novelty detection**

*Vision and Display System Lab.
Sogang University*

Outline

- Anomaly detection
 - One-class novelty detection
- OCGAN: one-class novelty detection using gans with constrained latent representations
- Experiments
- Conclusion

Anomaly detection

- Anomaly detection

- Normal sample과 abnormal sample을 구별해내는 문제

- 구분하고자 하는 대상에 따라 여러 분야로 나뉨

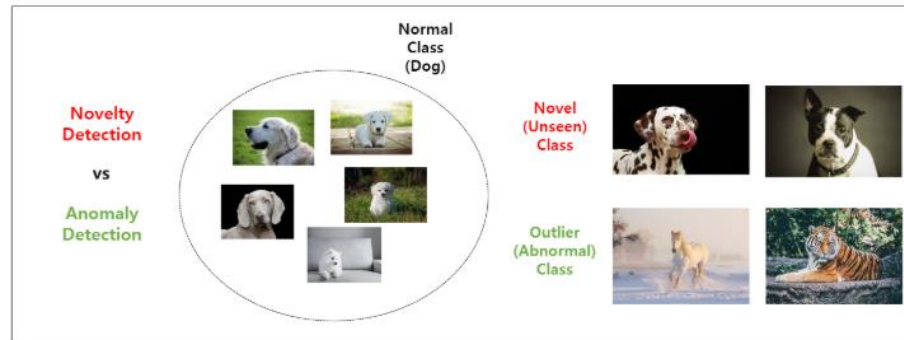
- 비정상 sample 정의에 따른 분류

- ⌘ Novelty Detection(Unseen)

- ✓ 지금까지 등장하지 않았지만 충분히 등장할 수 있는 sample을 찾아내는 분야

- ⌘ Outlier Detection(abnormal)

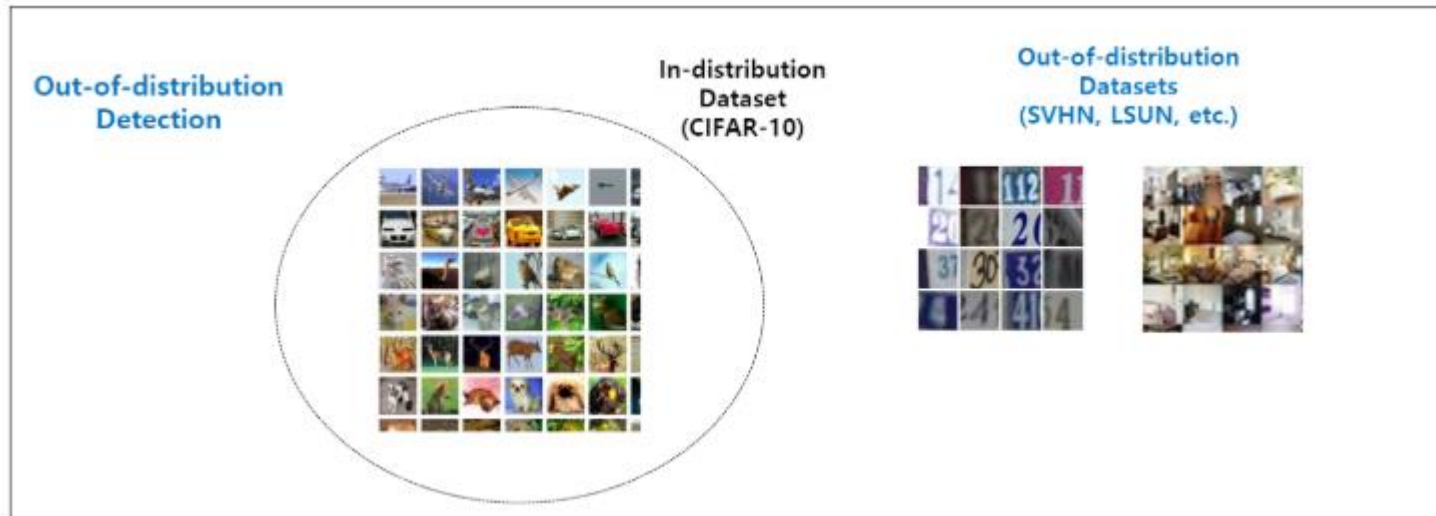
- ✓ 등장할 가능성이 거의 없는, 데이터에 오염이 발생했을 가능성이 있는 sample을 찾아 내는 분야



< Sample 예시 >

Anomaly detection

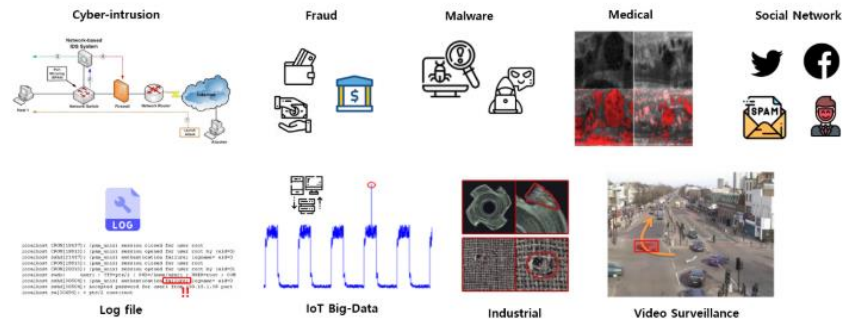
- Anomaly detection
 - Normal sample과 abnormal sample을 구별해내는 문제
 - 구분하고자 하는 대상에 따라 여러 분야로 나뉨
 - 정상 sample의 class 개수에 따른 분류
- ∴ Out-of-distribution Detection



< Sample 예시 >

Anomaly detection

- Anomaly detection
 - Normal sample과 abnormal sample을 구별해내는 문제
 - 데이터 셋의 구성에 따라 anomaly detection의 분야가 달라짐
 - Supervised Anomaly Detection
 - ⊛ 이상 데이터와 정상 데이터의 label이 존재
 - Semi-supervised (One-Class) Anomaly Detection
 - ⊛ 정상 데이터에만 label이 존재
 - Unsupervised Anomaly Detection
 - ⊛ 모든 데이터가 label이 존재하지 않음



< Anomaly detection 적용 분야 >

Anomaly detection

- Supervised Anomaly Detection

- 정상 sample과 비정상 sample의 Data와 Label이 모두 존재하는 경우

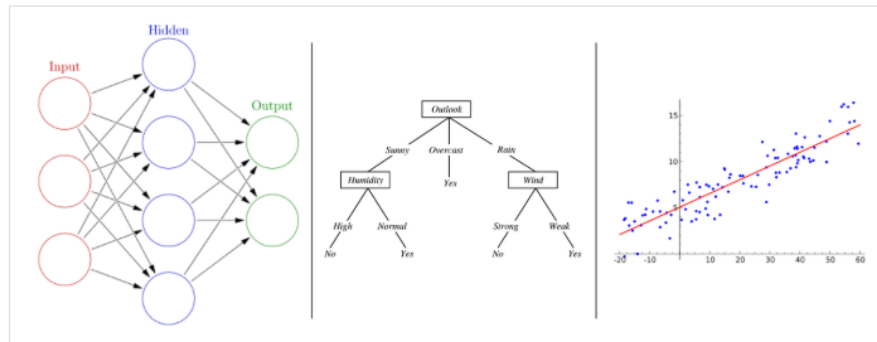
- 다른 방법 대비 정확도가 높음

- 그러나 현실적인 상황에서는 정상 sample보다 비정상 sample의 발생 빈도가 적어 class imbalance문제를 발생

- 또한 이상치의 기준이 명확하지 않아 이상치 label을 생성하기가 어렵고 시간과 비용이 많이 소모됨

- 훈련된 class나 데이터 유형이 아닌 이상치가 들어올 경우 모델 전체를 다시 훈련시켜야 함

※ 다중 회귀 분석, SVM, 의사결정나무, 인공신경망등의 기법이 존재



< Supervised 기법 예시 >

Anomaly detection

- Supervised Anomaly Detection

- Class imbalance problem

- Cifar10 데이터 셋의 0번 라벨을 정상으로 나머지를 비정상으로 구성한 anomaly detection 데이터셋을 생성

- ※ Train - Normal data: 5000, abnormal data: 45000

- ※ Test - Normal data: 1000, abnormal data: 9000

- VGG11 모델로 분류기 학습 후 성능 확인

- ※ Accuracy: 95.23%

- 전체 정확도로 보면 높은 성능을 보임

- Normal data에 대해서만 Accuracy를 확인해 보면 정상/비정상 데이터를 잘 구분하지 못하는 결과를 나타냄

- ※ Accuracy: 57.21%

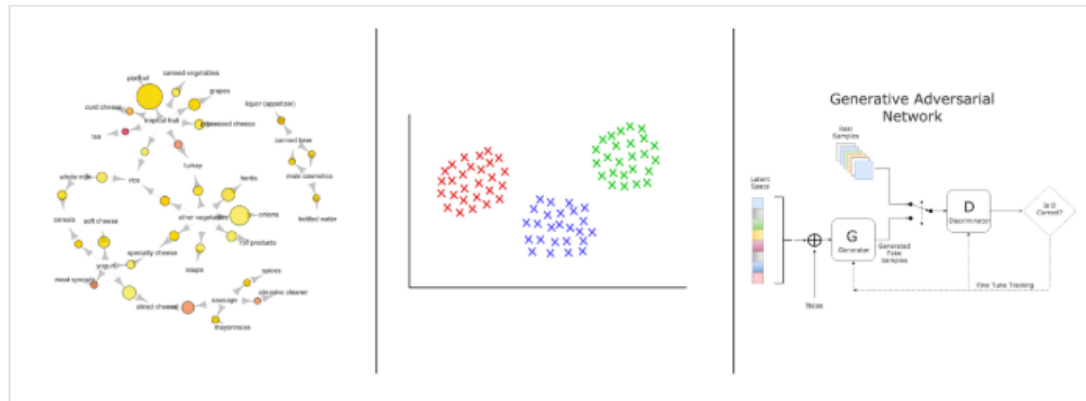
- 기존 현장에서의 Class imbalance 비율은 위의 실험에서의 비율보다 극단적이므로 좋은 성능을 기대하기 어려움

Anomaly detection

- Semi-supervised (One-Class), Unsupervised Anomaly Detection

- Data와 Label이 존재하지 않거나 정상 데이터의 Label만 존재하는 경우
- 정상 데이터를 이용하여 데이터가 본질적으로 가지고 있는 특징을 추출

- 정상 sample들을 둘러싸는 discriminative boundary를 설정하고, 이 boundary를 최대한 좁혀 boundary 밖에 있는 sample들을 모두 비정상어로 간주
- 정상 데이터의 특징을 학습하므로 학습되지 않은 모든 이상치 데이터들을 구분할 수 있음
 - ※ 별도의 labeling작업이 필요하지 않음
- 모델 파라미터나 정상데이터의 구성에 따라 모델의 성능이 불안정함



< Unsupervised 기법 예시 >

OCGAN: One-class Novelty Detection Using GANs

- Background

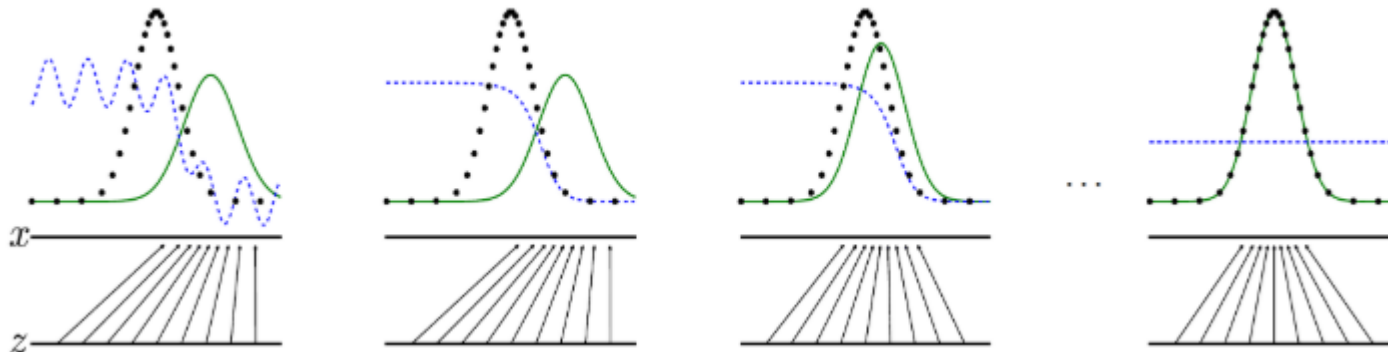
- GAN based anomaly detection

- 기존 딥러닝 모델은 입력에 대해 가장 높은 확률을 지닌 출력을 내도록 학습

- ※ 학습되지 않은 유사 데이터나 전체 데이터 셋의 분포를 고려하기 어려움

- GAN model은 학습 데이터셋의 distribution을 학습

- ※ Data distribution을 학습하기때문에 배우지 않은 데이터에 대해서도 In of distribution이라 판단되면 생성할 수 있음



< GAN이 분포를 학습하는 과정 >

OCGAN: One-class Novelty Detection Using GANs

- Background

- GAN based anomaly detection

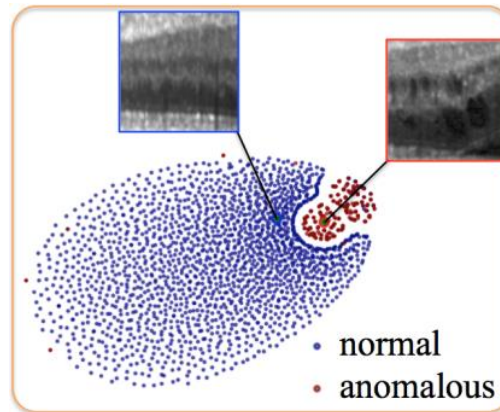
- 데이터의 분포를 학습하는 GAN의 성질을 이용하여 Unsupervised anomaly detection을 수행할 수 있음

- ※ 이미지는 $W \times H \times C$ 차원의 하나의 포인트라고 가정

- ※ 분포안에 data라고 판단하는 포인트에 대해서는 재구성이 가능

- ※ 분포 밖의 data에 대해서는 모델이 재구성하기 힘들기 때문에 그림이 뭉개지는 현상이 발생

- ※ 이때 입력 이미지와 생성된 이미지의 $l1$ loss를 통해 anomaly detection을 수행할 수 있음



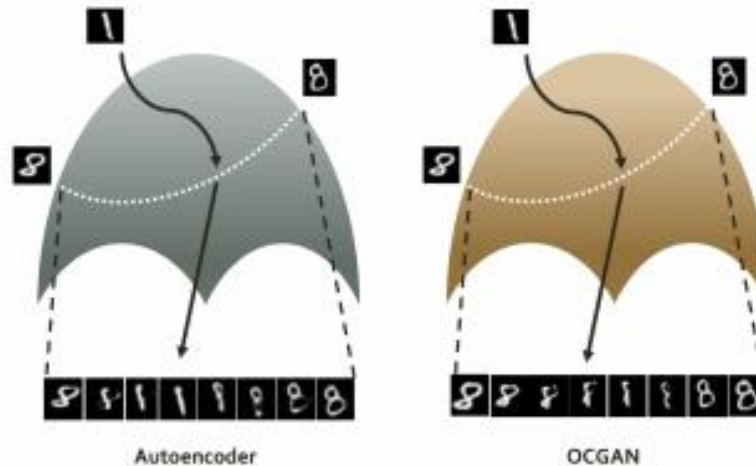
< Data의 분포 >

OCGAN: One-class Novelty Detection Using GANs

- Motivation

- Generative model problem

- VAE, GAN의 공통적인 부분은 입력 데이터셋을 대표 할 수 있는 하나의 vector로 mapping
 - ※ Latent vector의 사이즈에 따라 담을 수 있는 데이터셋의 분포가 달라짐
 - 만약 데이터셋의 분포보다 Latent vector의 차원이 크게 되면 Latent vector에 빈공간이 생길 수 밖에 없음



OCGAN: One-class Novelty Detection Using GANs

- Motivation

- Generative model problem

- 이러한 빈공간에서 의도한 normal class가 아닌 의도하지 않은 class의 이미지가 학습되는 경우가 존재
 - 따라서 abnormal class에 대해서 재구성 loss가 낮아 normal class로 분류하는 문제가 발생
 - 이러한 문제를 해결하기 위해선 latent vector의 차원을 줄여 빈공간을 줄이면 됨
 - ※ 그러나 latent vector의 차원을 줄이면 재구성 성능이 떨어지게 됨

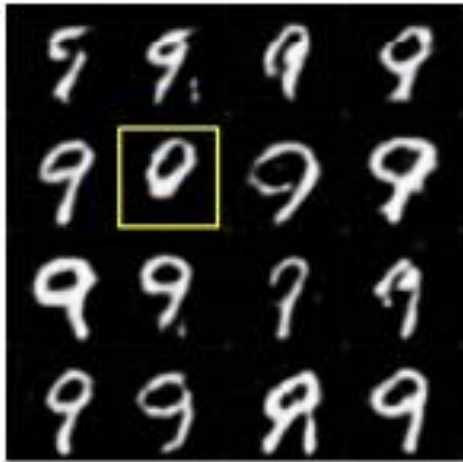


OCGAN: One-class Novelty Detection Using GANs

- Motivation

- Generative model problem

- 본 논문에서는 이러한 문제를 해결하고자 latent vector에서 재구성되는 이미지를 강제적으로 normal class로 구성되게 만드는 방법을 제안
 - 이를 통해 latent vector의 빈공간에 대해서도 최대한 normal class의 형태를 유지하면서 재구성되는 것이 가능



OCGAN: One-class Novelty Detection Using GANs

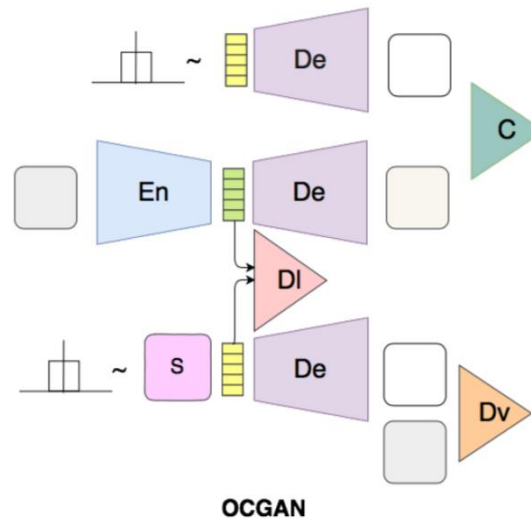
- Proposed Strategy

- Informative-negative Mining and Classifier

- Latent vector에서 재구성되는 이미지들을 normal class로 강제하기 위해서 2가지 과정이 필요

- ※ Latent vector에서 생성되는 이미지가 normal class가 아닌 이미지가 생성되는 빈공간을 찾아야함

- ※ 빈공간을 찾은 후 그 영역에서 생성되는 이미지를 강제하도록 만들어야 함



OCGAN: One-class Novelty Detection Using GANs

- Proposed Strategy

- Informative-negative Mining and Classifier

- 빈 공간을 찾기 위해 Informative-negative Mining 이라는 방법을 제안

- 먼저 normal/abnormal class를 구분하는 분류기를 훈련

- ※ Abnormal class를 구분하기 위해 처음 훈련을 normal 이미지에서 생성된 이미지(1)와 정규 분포에서 무작위로 생성된 이미지(0)를 통해 분류기 학습

- ※ 완벽하지는 않지만 어느정도 구분을 해주는 weak classifier가 됨

- 그 후 Informative-negative Mining을 통해 빈공간을 검출

- ※ 훈련된 classifier의 파라미터를 고정

- ※ 입력하는 정규분포를 변수로 설정

- ※ 그 후 targets을 0으로 하고 분류기를 학습

- ✓ Backpropagation을 통해 기울기가 전달되지만 분류기의 변수들은 변하지 않음

- ✓ 맨 마지막 정규분포 입력만 기울기의 영향을 받아 변하게 됨

- ✓ 이 정규분포 입력은 분류기의 출력을 0으로 만드는 입력이 됨

- 즉 빈 공간 검출이 가능

OCGAN: One-class Novelty Detection Using GANs

- Proposed Strategy

- Informative-negative Mining and Classifier

- 찾아낸 빈 공간의 정규 분포 입력을 이용하여 GAN학습을 진행

- ※ GAN학습과정에 reconstruct loss가 포함되기 때문에 normal class의 형태를 띄게 강제로 훈련 됨

- ※ 만약 찾아낸 공간이 빈공간이 아니더라도 GAN학습에는 영향을 주지 못함



OCGAN: One-class Novelty Detection Using GANs

- Experimental Results

	0	1	2	3	4	5	6	7	8	9	MEAN
OCSVM [24]	0.988	0.999	0.902	0.950	0.955	0.968	0.978	0.965	0.853	0.955	0.9513
KDE [2]	0.885	0.996	0.710	0.693	0.844	0.776	0.861	0.884	0.669	0.825	0.8143
DAE [4]	0.894	0.999	0.792	0.851	0.888	0.819	0.944	0.922	0.740	0.917	0.8766
VAE [6]	0.997	0.999	0.936	0.959	0.973	0.964	0.993	0.976	0.923	0.976	0.9696
Pix CNN [26]	0.531	0.995	0.476	0.517	0.739	0.542	0.592	0.789	0.340	0.662	0.6183
GAN [23]	0.926	0.995	0.805	0.818	0.823	0.803	0.890	0.898	0.817	0.887	0.8662
AND [1]	0.984	0.995	0.947	0.952	0.960	0.971	0.991	0.970	0.922	0.979	0.9671
AnoGAN [23]	0.966	0.992	0.850	0.887	0.894	0.883	0.947	0.935	0.849	0.924	0.9127
DSVDD [19]	0.980	0.997	0.917	0.919	0.949	0.885	0.983	0.946	0.939	0.965	0.9480
OCGAN	0.998	0.999	0.942	0.963	0.975	0.980	0.991	0.981	0.939	0.981	0.9750

	PLANE	CAR	BIRD	CAT	DEER	DOG	FROG	HORSE	SHIP	TRUCK	MEAN
OCSVM [24]	0.630	0.440	0.649	0.487	0.735	0.500	0.725	0.533	0.649	0.508	0.5856
KDE [2]	0.658	0.520	0.657	0.497	0.727	0.496	0.758	0.564	0.680	0.540	0.6097
DAE [4]	0.411	0.478	0.616	0.562	0.728	0.513	0.688	0.497	0.487	0.378	0.5358
VAE [6]	0.700	0.386	0.679	0.535	0.748	0.523	0.687	0.493	0.696	0.386	0.5833
Pix CNN [26]	0.788	0.428	0.617	0.574	0.511	0.571	0.422	0.454	0.715	0.426	0.5506
GAN [23]	0.708	0.458	0.664	0.510	0.722	0.505	0.707	0.471	0.713	0.458	0.5916
AND [1]	0.717	0.494	0.662	0.527	0.736	0.504	0.726	0.560	0.680	0.566	0.6172
AnoGAN [23]	0.671	0.547	0.529	0.545	0.651	0.603	0.585	0.625	0.758	0.665	0.6179
DSVDD [19]	0.617	0.659	0.508	0.591	0.609	0.657	0.677	0.673	0.759	0.731	0.6481
OCGAN	0.757	0.531	0.640	0.620	0.723	0.620	0.723	0.575	0.820	0.554	0.6566

개인 연구 주제

- One class novelty detection

- Problem

- OCGAN에서는 latent space의 빈 공간에 대해 언급
 - 이러한 문제는 MNIST같은 간단한 데이터 셋에서는 의미 있는 접근일수 있음
 - ※ 결과에서도 MNIST 성능에서 유의미한 결과를 보임
 - 그러나 Cifar10 과 같은 분포가 큰 데이터에서는 약간의 성능 상승만 보임
 - ※ 분포가 큰 데이터 셋에 대해서는 빈 공간이 그만큼 적음
 - ※ 논문에서도 다른 데이터 셋에 비해 배경과 같은 노이즈가 많다고 언급



CIFAR10



COIL



FMNIST



MNIST

개인 연구 주제

- One class novelty detection

- Problem

- One class novelty detection task에서는 normal class의 feature를 잘 추출하는 것이 중요
 - 그러나 현실적인 데이터에서는 class 분류에 필요하지 않는 노이즈 성분들이 많이 들어가 있음
 - ※ 실제 잘 정렬된 데이터여도 모든 픽셀들이 class와 관련된 feature를 의미하지 않음
 - 실제 detection의 성능을 확인할 때 segmentation loss와 같이 훈련시키는 maskRCNN, RetinaMask등의 모델이 일반 객체 검출 모델보다 높은 성능을 보임
 - ※ 배경성분을 배제한 segmentation을 같이 학습 시킴으로써 class를 더 잘 구분할 수 있게 되고 그로인한 mAP성능이 증가하게 됨
 - 이를 해결하기 위해 Discriminator에 추가 정보를 입력영상과 함께 주어 성능을 높이는 연구가 진행되었지만 어느정도 한계가 있음
 - ※ Latent vector를 Discriminator에 같이 입력하여 Latent vector의 real/fake도 구분하도록 함

개인 연구 주제

- One class novelty detection

- 연구 목표

- 이미지 내에서 어떤 영역을 중요하게 봐야 하는지 자동으로 선정해주는 알고리즘 연구
 - 이를 통해 실제 class에 필요한 픽셀을 추출하고 그 분포만 학습시켜 성능향상 목표

- 연구 내용

- Attention

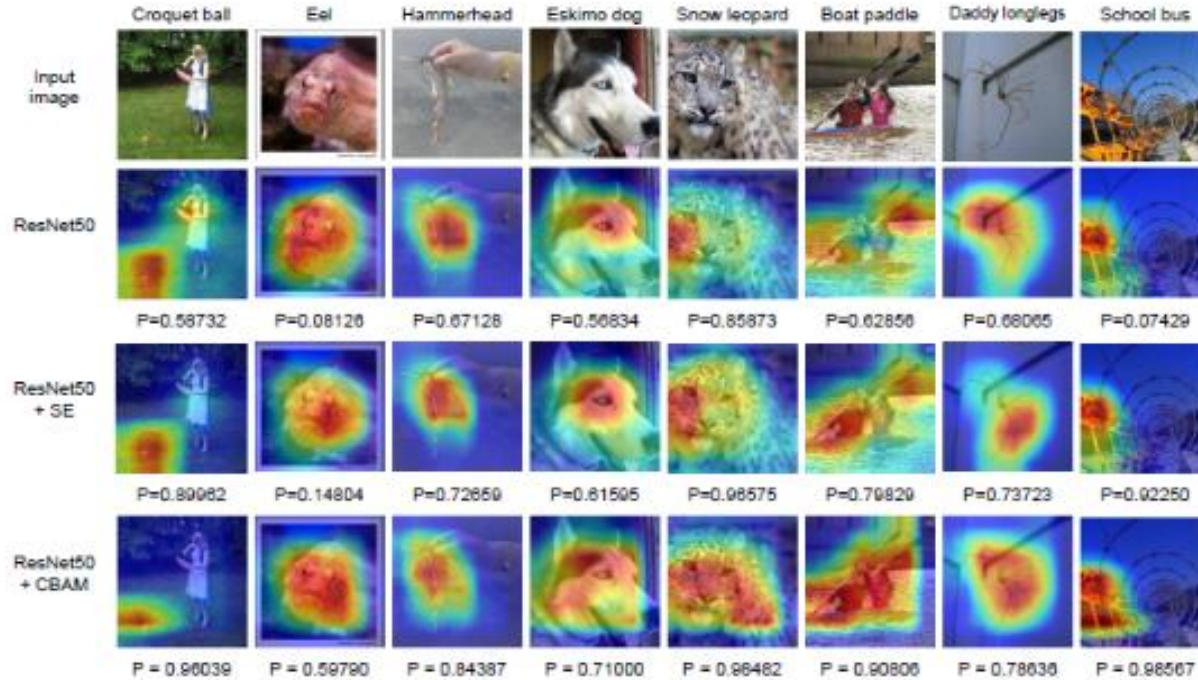
- ※ 자연어 처리에서 연구된 기법으로 어떤 단어에 더 집중해서 볼 것인지 자동으로 학습함
 - ※ 이를 이미지 영역에서 사용하려고 하는 시도들이 많이 연구되고 있음
 - ※ CBAM: Convolutional Block Attention Module
 - ✓ ECCV2018에서 발표된 논문
 - ✓ Conv연산에 Attention을 적용하여 어떤 채널을 더 중요하게 보고 어떤 픽셀을 더 중요하게 볼 것인지 자동으로 학습하여 성능을 높임

개인 연구 주제

- One class novelty detection

- 연구 내용

- Attention



개인 연구 주제

• 연구 내용

- 실제 background noise가 모델의 성능에 얼마나 영향을 끼치는지 확인하기 위해 실험 구성

- Segmentation data가 포함된 COCO dataset으로 실험 데이터셋 구성

※ 'airplane', 'car', 'bird', 'cat', 'sports ball', 'dog', 'bowl', 'horse', 'boat', 'truck' 10개의 class로 이루어진 데이터셋 구성

※ Detection용 bounding box로 object를 crop한 후 Cifar10 dataset과 같은 크기인 32x32로 resize

✓ Box 크기가 너무 작지 않은 50이상인 객체만 crop

✓ 종횡비의 크기가 너무 크지않은 0.5~2 사이에 객체만 crop

※ 같은 데이터의 segmentation mask로 이루어진 data 생성

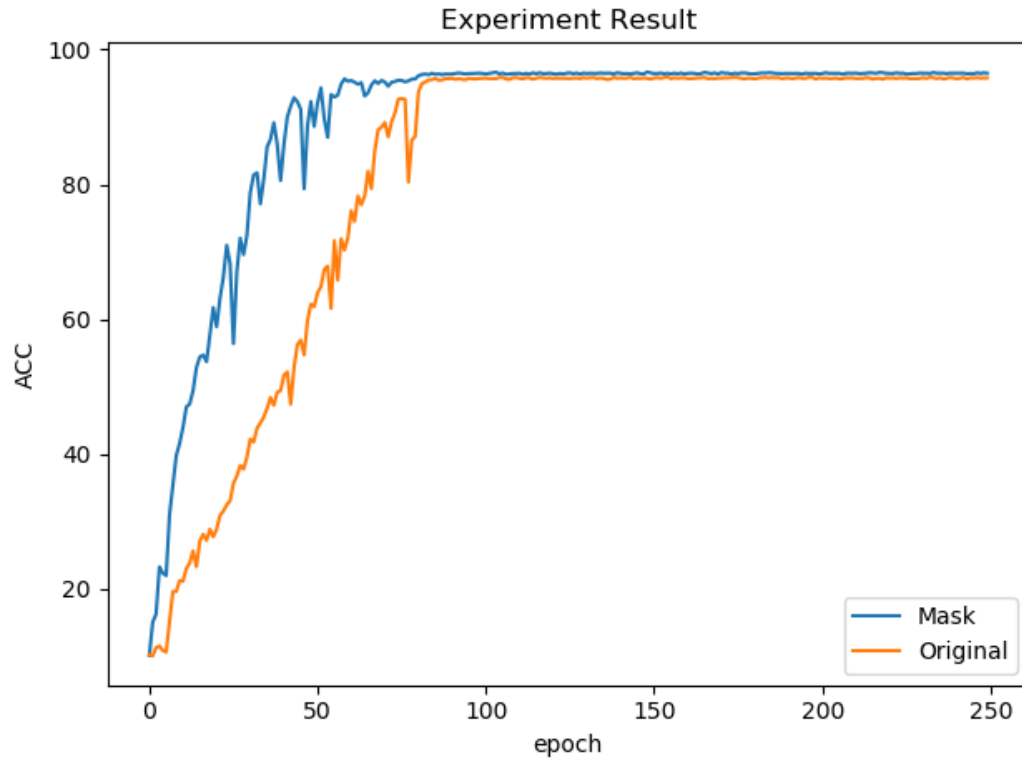
※ 각 550개씩 5500개의 데이터를 생성 후 5000/500 으로 train/test data를 split

- 생성된 데이터셋을 이용 일반 데이터와 배경이 제거한 마스킹 된 데이터를 이용

※ classification task 수행 후 성능평가

개인 연구 주제

- 연구 내용
 - 실험 결과



개인 연구 주제

- 연구 내용

- 실험 결과

- VGG11 모델에서 실험

- Seed를 고정하여 초기 파라미터 값 고정

- 모델의 수렴 속도 및 정확도가 향상

- ※ Acc 80% 달성 epoch - mask: 25, original: 64

- ※ 모델 정확도 - mask: 96.672%, original: 95.247%

- Masking 된 이미지를 사용하는 것이 모델이 class의 특징을 더 빠르게 배울 수 있음을 확인

- ※ Back ground 부분이 noise로 간주된다는 것을 검증

- 이를 이용하여 class에 관련된 픽셀을 추출할 수 있으면 One class novelty detection task의 성능이 높아질 것이라 가정

개인 연구 주제

- 연구 내용

- Proposed model

- 기존 GAN기반 One class novelty detection model에서 네트워크를 추가함

- ※ Attention map generator

- ✓ Generator의 feature layer들을 받아 픽셀의 중요도를 나타내는 attention map을 형성

- ※ Classifier

- ✓ Mask generator가 의미 있는 픽셀을 잡을 수 있도록 guide하는 역할

- ※ GAN

- ✓ 이미지와 attention map을 곱한 새로운 이미지의 분포를 배우는 GAN

개인 연구 주제

- 연구 내용

- Proposed model

- 데이터의 분포를 배우는 GAN을 학습시키면서 attention map 형성
 - 형성된 attention map과 normal 이미지를 element-wise 곱을 한 데이터와 abnormal 이미지에
도 같은 과정을 거친 이미지를 구분하는 classifier를 훈련
 - ※ Attention map을 곱하지 않은 normal/abnormal 이미지도 같이 훈련
 - ※ 이를 통해 attention map이 class를 구분하는 데에 의미 있는 픽셀을 고르게 됨
 - attention map과 normal 이미지의 분포를 학습하는 새로운 GAN을 만들어 의미 있는 픽셀의
분포만 학습
 - ※ 기존 학습된 GAN과 새로운 GAN의 reconstruct loss를 이용하여 One class novelty
detection task 수행

개인 연구 주제

- 연구 결과

- Cifar10

- OCGAN보다 높은 결과를 확인

	PLANE	CAR	BIRD	CAT	DEER	DOG	FROG	HORSE	SHIP	TRUCK	MEAN
OCSVM [24]	0.630	0.440	0.649	0.487	0.735	0.500	0.725	0.533	0.649	0.508	0.5856
KDE [2]	0.658	0.520	0.657	0.497	0.727	0.496	0.758	0.564	0.680	0.540	0.6097
DAE [4]	0.411	0.478	0.616	0.562	0.728	0.513	0.688	0.497	0.487	0.378	0.5358
VAE [6]	0.700	0.386	0.679	0.535	0.748	0.523	0.687	0.493	0.696	0.386	0.5833
Pix CNN [26]	0.788	0.428	0.617	0.574	0.511	0.571	0.422	0.454	0.715	0.426	0.5506
GAN [23]	0.708	0.458	0.664	0.510	0.722	0.505	0.707	0.471	0.713	0.458	0.5916
AND [1]	0.717	0.494	0.662	0.527	0.736	0.504	0.726	0.560	0.680	0.566	0.6172
AnoGAN [23]	0.671	0.547	0.529	0.545	0.651	0.603	0.585	0.625	0.758	0.665	0.6179
DSVDD [19]	0.617	0.659	0.508	0.591	0.609	0.657	0.677	0.673	0.759	0.731	0.6481
OCGAN	0.757	0.531	0.640	0.620	0.723	0.620	0.723	0.575	0.820	0.554	0.6566

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	mean.
0.791.	0.533.	0.632.	0.608.	0.723.	0.610.	0.742.	0.580.	0.801.	0.743.	0.673.

개인 연구 주제

- 연구 결과

- Cifar10

- OCGAN보다 높은 결과를 확인

	PLANE	CAR	BIRD	CAT	DEER	DOG	FROG	HORSE	SHIP	TRUCK	MEAN
OCSVM [24]	0.630	0.440	0.649	0.487	0.735	0.500	0.725	0.533	0.649	0.508	0.5856
KDE [2]	0.658	0.520	0.657	0.497	0.727	0.496	0.758	0.564	0.680	0.540	0.6097
DAE [4]	0.411	0.478	0.616	0.562	0.728	0.513	0.688	0.497	0.487	0.378	0.5358
VAE [6]	0.700	0.386	0.679	0.535	0.748	0.523	0.687	0.493	0.696	0.386	0.5833
Pix CNN [26]	0.788	0.428	0.617	0.574	0.511	0.571	0.422	0.454	0.715	0.426	0.5506
GAN [23]	0.708	0.458	0.664	0.510	0.722	0.505	0.707	0.471	0.713	0.458	0.5916
AND [1]	0.717	0.494	0.662	0.527	0.736	0.504	0.726	0.560	0.680	0.566	0.6172
AnoGAN [23]	0.671	0.547	0.529	0.545	0.651	0.603	0.585	0.625	0.758	0.665	0.6179
DSVDD [19]	0.617	0.659	0.508	0.591	0.609	0.657	0.677	0.673	0.759	0.731	0.6481
OCGAN	0.757	0.531	0.640	0.620	0.723	0.620	0.723	0.575	0.820	0.554	0.6566

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	mean.
0.791.	0.533.	0.632.	0.608.	0.723.	0.610.	0.742.	0.580.	0.801.	0.743.	0.673.

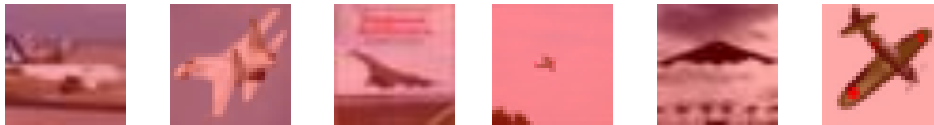
개인 연구 주제

- 연구 결과

- Attention을 적용하기 전 기본 모델과 적용한 후 모델의 출력을 Grad-cam 통해 확인

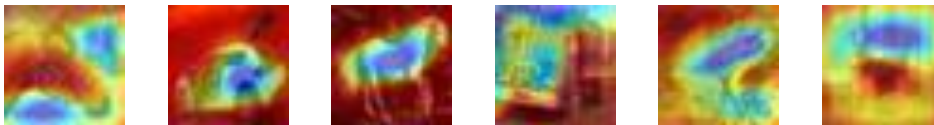
- 기본 모델 grad cam결과

※ Normal data



✓어떤 부분이 중요한 부분인지 잡지 못하고 있음

※ Abnormal data



✓OOD의 데이터도 중요하다 보는 경우가 많음

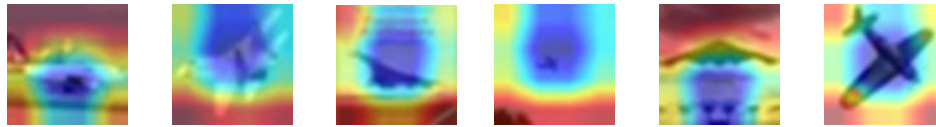
개인 연구 주제

- 연구 결과

- Attention을 적용하기 전 기본 모델과 적용한 후 모델의 출력을 Grad-cam 통해 확인

- Attention 적용 모델 grad cam결과

※ Normal data



✓기본 모델 출력과 다르게 중요한 부분을 잡고 있음

※ Abnormal data



✓OOD데이터를 중요하다 잡는 경우가 적음

개인 연구 주제

- 연구 결과

- Attention mask 분석

- MNIST의 데이터는 간단하기 때문에 의미 있는 픽셀을 직관적으로 확인하기 쉬움
 - 의미 있는 픽셀들을 잘 보는지 확인하기 위해 Attention mask 확인



- ※ 의미 있는 픽셀들을 잘 포착하는 모습을 보임
 - ※ Attention 네트워크가 원하는 역할을 했다고 볼 수 있음

개인 연구 주제

- 추후 계획

- Classifier를 추가한 모델에 대해서는 성능을 많이 측정하지 않음

- 0, 1 label에 대해서는 2~10%내 성능향상을 보임

- ※ 전체 cifar10 데이터셋에 대해 성능 측정을 할 예정

- Sigmoid 출력으로 나오는 Attention map에 temperature scale을 적용하여 더 극명한 mask를 만든 후 모델에 적용해 볼 예정

- ※ 중요한 픽셀을 더 선명히 보고 중요하지 않은 픽셀을 보지 않음

- 처음 학습을 시작할 때의 Attention 네트워크는 masking을 잘 생성하지 못함

- ※ 따라서 마스킹 된 이미지의 학습을 담당하는 generator와 discriminator가 초기 학습되지 않은 데이터의 분포까지 학습

- ※ 이런 문제를 해결하고자 어느정도 학습시킨 후 마스킹 된 이미지의 학습을 담당하는 generator와 discriminator를 초기화 하여 다시 학습시켜 안정화된 마스킹 이미지의 분포를 학습시켜 볼 예정

- ※ 또는 mask generator를 미리 학습하고 진행할 예정