

2024 하계 세미나

A Unified Model for Multi-class Anomaly Detection



Sogang University

Vision & Display Systems Lab, Dept. of Electronic Engineering



Presented By

김예슬

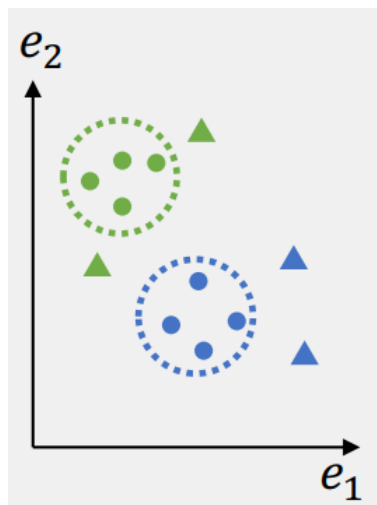
Outline

- Background
 - Anomaly detection
 - Multi-class anomaly detection
 - Reconstruction-based anomaly detection
- Zhiyuan You, et al. "A Unified Model for Multi-class Anomaly Detection" *Neural Information Processing Systems (NeurIPS 2022)*, 2022
 - Introduction
 - Proposed method
 - Experimental results
 - Conclusions

Background

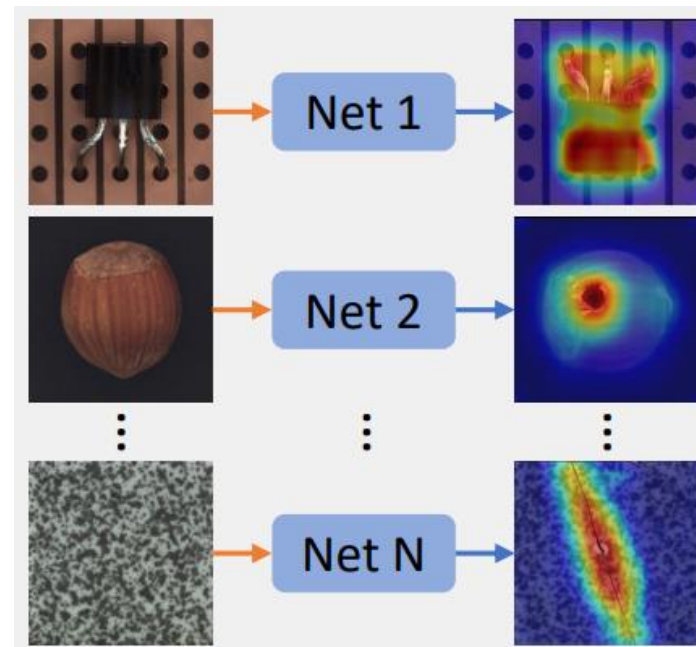
- Anomaly detection (AD)
 - 이상치 탐지 방법론들은 정상 데이터의 분포를 모델링하고 이를 기반으로 이상치를 탐지하기 때문에 정상 데이터의 compact boundary를 학습하는 것이 중요함
 - 기존의 이상치 탐지 방법론들은 여러 개의 object/class에 대하여 각각 개별적인 decision boundary를 학습하는 one-class-one-model 방식으로 작동함

- Normal ▲ Anomaly Boundary
- Class 1 ■ Class 2



< One-class data distribution >

→ Train on Normal Data → Infer to Detect Anomalies



< One-class-one-model scheme >

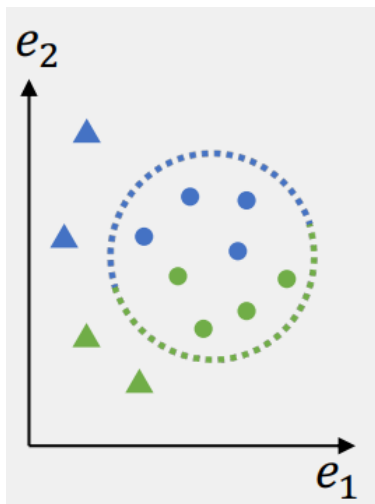
Background

- Multi-class anomaly detection

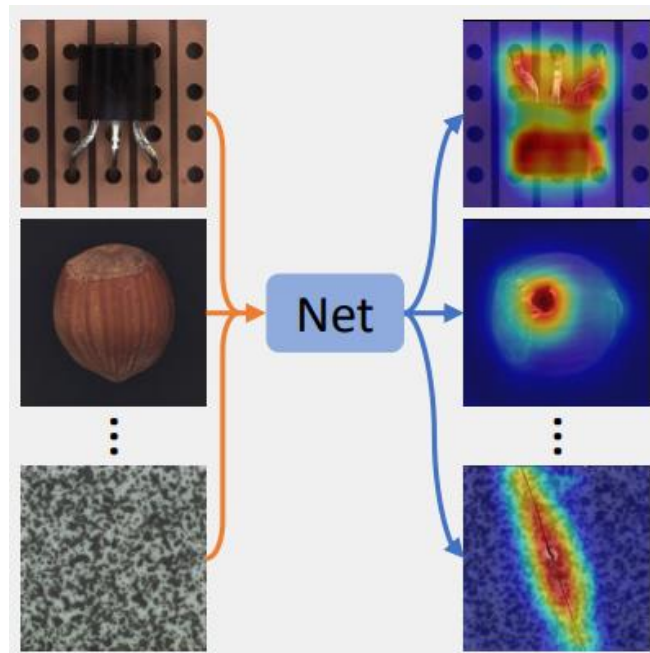
- 기존의 이상치 탐지 방법론들은 class의 개수가 증가하면 큰 memory가 소모되며, 다양한 특징을 가지는 정상 데이터의 분포 추정에 적합하지 않을 수 있음
- Multi-class 이상치 탐지는 여러 class의 정상 분포를 함께 추출하여 다양한 정상 데이터들이 하나의 boundary를 공유하도록 함으로써 위의 문제를 해결할 수 있음

→ Train on Normal Data → Infer to Detect Anomalies

● Normal ▲ Anomaly Boundary
 ■ Class 1 ■ Class 2



< Multi-class data distribution >



< Multi-class-one-model scheme >

Background

- Reconstruction-based anomaly detection
 - 이상치 탐지에서는 input을 복원하는 과정을 통해 정상 데이터의 분포를 학습하는 reconstruction 기반 방법론들이 널리 사용되어 왔음
 - 해당 방법론들은 정상 데이터의 분포를 학습한 모델은 정상 데이터로부터 구분되는 샘플을 잘 복원하지 못할 것이라는 가정 하에 복원 에러가 큰 샘플을 이상치로 탐지함
 - 하지만 reconstruction 기반 방법론에서는 identical shortcut 문제가 발생함
 - Identical shortcut: 모델이 input을 복제하여 output을 도출하는 것처럼 학습되어 normal/abnormal samples를 모두 잘 복원하고 이상치 탐지에 실패하는 문제
 - Multi-class anomaly detection에서는 정상 데이터의 분포가 더 복잡하기 때문에 identical shortcut 문제가 더 심하게 발생할 수 있음

A Unified Model for Multi-class Anomaly Detection

Introduction

- 기존 이상치 탐지 방법론들은 아래와 같은 한계점을 가짐
 - 각각 다른 object에 대한 이상치 탐지를 수행하기 위해 개별적인 모델을 구축해야 하므로 메모리 문제가 있으며, 다양한 특징을 가지는 정상 데이터에 적합하지 않을 수 있음
 - 널리 사용되는 reconstruction 기반 이상치 탐지 방법론들은 identical shortcut 문제로 인해 성능이 저하될 수 있음
- 본 논문에서는 이러한 한계점을 해결하기 위해 아래와 같은 기여점을 가지는 이상치 탐지를 위한 통합 구조인 UniAD를 제안함
 - UniAD는 multi-class 이상치 탐지 모델로 class 개수에 관계없이 이상치 탐지를 위한 하나의 통합 모델을 구축할 수 있으며, 그 성능이 separate case와 유사함
 - UniAD는 identical shortcut을 보완할 수 있는 Layer-wise Query Decoder, Neighbor Masked Attention, Feature Jittering을 제안 및 활용함

Proposed method

- 목표

- Identical shortcut 문제를 해결하고, reconstruction-based AD 방식으로 multi-class 이상치 탐지를 수행하는 것

- Overview

- 문제 탐색

- MLP, CNN, Transformer의 identical shortcut 발생 정도를 확인 및 비교함
- Transformer의 attention layer가 learnable query embedding으로 인해 identical shortcut 문제가 가장 적게 발생하는 것을 실험 및 이론적으로 확인함

- 문제 보완

- Attention layer의 identical shortcut을 개선하기 위해 query embedding의 사용을 극대화한 layer-wise query decoder을 제안함
- 비슷한 주변 정보의 복제로 인한 무분별한 reconstruction을 방지하기 위해 neighbor masked attention을 제안함
- Reconstruction의 성능 향상을 위해 denoising task를 위한 feature jittering을 활용함

Proposed method

- Comparison among MLP, CNN, and Transformer

- 본 논문에서는 AD의 backbone 모델인 MLP, CNN, Transformer를 활용한 reconstruction baselines을 구축하고 각 모델의 identical shortcut을 확인함

- 각 모델의 성능 결과 및 비교

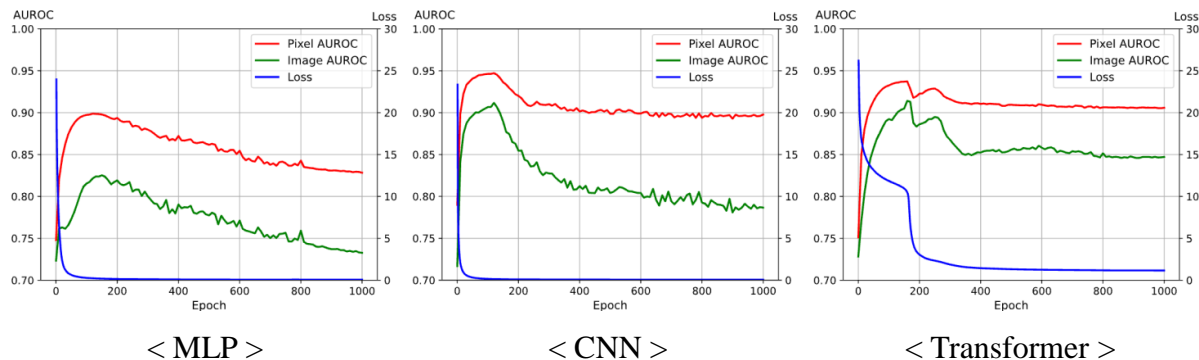
- MVtec-AD unified case에서 학습하면서 10 epoch마다 아래 3개의 지표를 측정함

- ※ **Loss**: 정상 데이터로 구성된 train 데이터셋에 대한 reconstruction loss

- ※ **Image AUROC**: 정상/이상치 데이터로 구성된 test 데이터셋에 대한 anomaly detection의 AUROC

- ※ **Pixel AUROC**: 정상/이상치 데이터로 구성된 test 데이터셋에 대한 anomaly localization의 AUROC

- 모든 모델이 학습이 진행됨에 따라 정상 데이터를 잘 복원하지만, 이상치 탐지 성능은 일정 수준 도달 이후 하락하는 것을 통해 identical shortcut을 확인할 수 있음



< MLP >

< CNN >

< Transformer >

Proposed method

- Comparison among MLP, CNN, and Transformer

- MLP, CNN, Transformer 기반 reconstruction baselines에 이상치 이미지를 input으로 넣은 결과, 각 모델이 이상치도 잘 복원하는 것을 통해 정성적으로도 identical shortcut이 발생함을 확인할 수 있음
- 그러나, 상대적으로 Transformer가 MLP와 CNN 보다 identical shortcut의 발생 정도가 적은 것으로 판단됨



< Anomaly >

< MLP >

< CNN >

< Transformer >

Proposed method

- Analysis on identical shortcut

- MLP의 fully-connected layer를 살펴보면, 아래와 같은 shortcut 학습이 가능함

- 1-layer: $y = x^+w + b$

- ※ y : output

- ※ x^+ : input feature

- ※ w : weight

- ※ b : bias

- Shortcut: $w \rightarrow$ identity matrix & $b \rightarrow 0$

- CNN의 경우 1×1 kernel로 구성된 convolutional layer가 fully-connected layer와 동일하며, $n \times n$ kernel은 1×1 kernel 보다 더 많은 파라미터와 더 큰 용량을 가지므로 shortcut을 학습할 가능성이 큼

- Query embedding을 사용한 Transformer의 attention layer을 살펴보면, query embedding q 가 정상 이미지 x^+ 와 유사도가 높은 shortcut 학습이 가능함

- 1-layer: $y = \text{softmax}(q(x^+)^T / \sqrt{C})x^+$

- ※ q : learnable한 query embedding

- ※ x^+ : normal sample의 feature

- ※ C : channel dimension

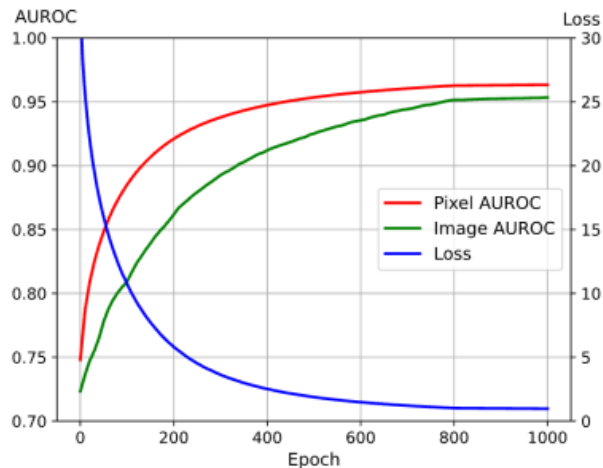
- Shortcut: $\text{softmax}(q(x^+)^T / \sqrt{C}) \rightarrow$ identity matrix

Proposed method

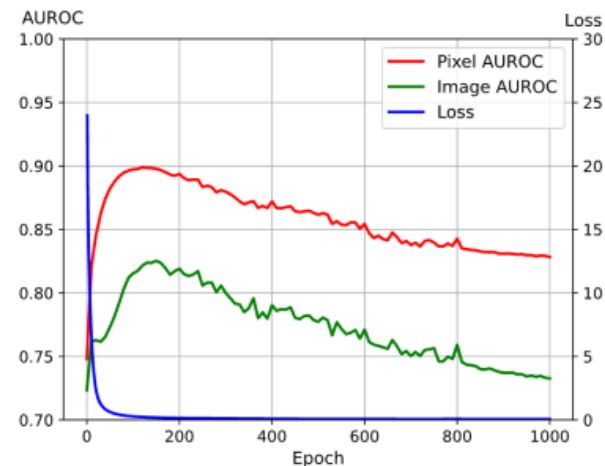
- Analysis on identical shortcut
 - MLP, CNN, Transformer를 활용한 reconstruction baselines 모두 이론적인 측면에서 모델이 identical shortcut을 학습할 가능성이 있음
 - 하지만, Transformer의 경우 attention layer의 shortcut이 learnable한 query embedding q 가 정상 이미지 x^+ 와 유사도가 매우 높은 것이라는 것을 고려하면 모델은 이상치 이미지를 잘 복원할 수 없음
 - 결론적으로 query embedding이 정상 데이터의 분포를 학습하는데 매우 중요한 역할을 한다는 것을 알 수 있음
 - 따라서, 본 논문에서는 이러한 사실에 입각하여 attention layer의 identical shortcut을 보완할 수 있는 multi-class 이상치 탐지를 제안함

Proposed method

- Relation between identical shortcut and unified case
 - MLP 기반 reconstruction baseline을 separate case에 대해 동일하게 학습한 후, unified case와 결과를 비교함
 - Separate case: 하나의 class에 대한 데이터셋만으로 독립적인 모델을 학습하는 것
 - Separate case에서는 지속적인 loss 감소와 AUROC 증가가 나타나는 것을 통해, unified case 난이도가 높고 이로 인해 identical shortcut이 발생함을 알 수 있음



< MLP - Separate case >



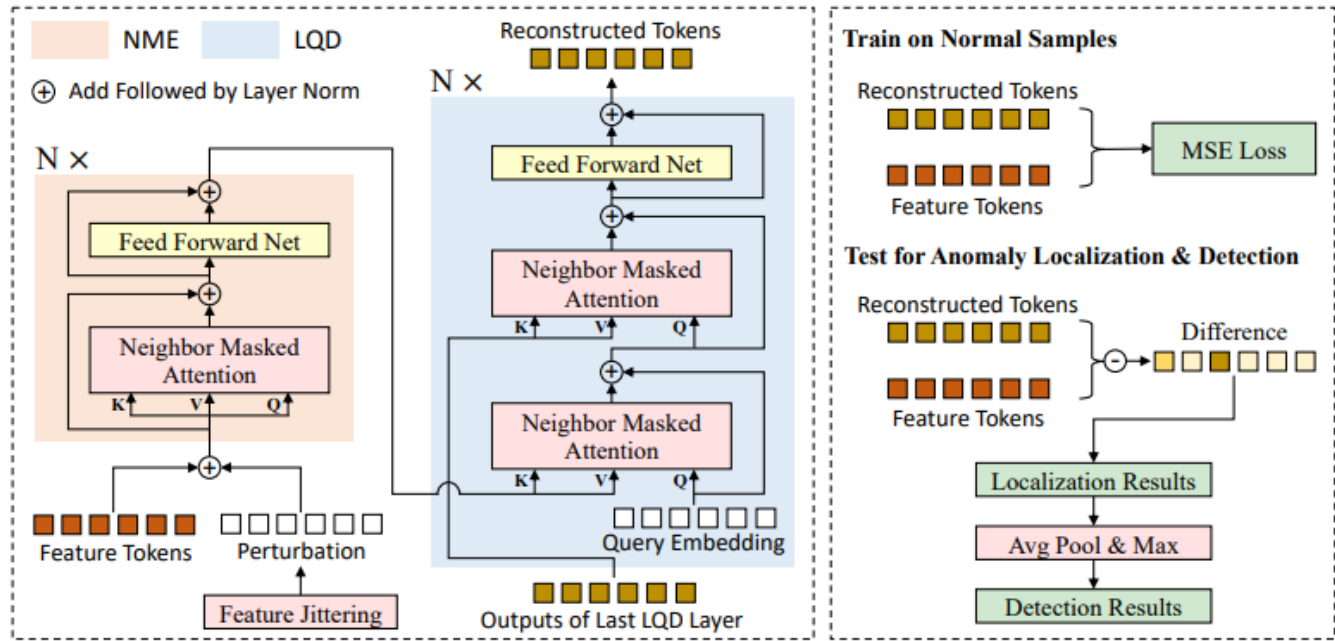
< MLP - Unified case >

Proposed method

- Improving feature reconstruction for unified anomaly detection

- Overview

- UniAD는 reconstruction 기반 multi-class 이상치 탐지 통합 구조로 Neighbor Masked Encoder (NME)와 Layer-wise Query Decoder (LQD)로 구성됨
- 기존의 Transformer 모델에 Feature Jittering, Neighbor Masked Attention, Query Embedding을 추가한 것



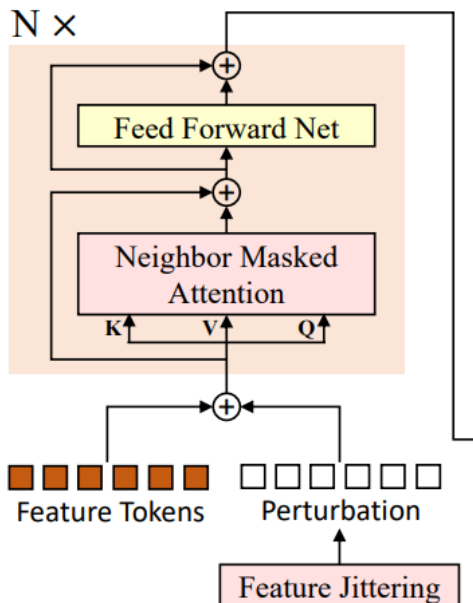
< UniAD overview >

Proposed method

- Improving feature reconstruction for unified anomaly detection

- Neighbor Masked Encoder (NME)

- 이미지를 입력으로 받으면, CNN으로 feature tokens를 추출하고 해당 tokens에 feature jittering을 통해 도출된 perturbation을 추가하여 input을 구축함
 - NME의 각 layer는 neighbor masked attention module과 feed-forward network로 구성되어 있으며, NME는 위에서 구축된 input으로부터 encoder embedding을 도출함



< Neighbor Masked Encoder (NME) >

Proposed method

- Improving feature reconstruction for unified anomaly detection

- Feature Jittering

- Fixed pre-trained backbone을 통해 이미지로부터 feature tokens를 추출함

- ※ Pre-trained backbone: ImageNet에서 사전 학습된 EfficientNet-b4

- ※ Feature tokens: backbone의 1~4번째 block에서 도출된 feature maps를 동일한 크기로 resize한 후, channel dimension을 기준으로 concatenate하여 도출함

- 본 연구에서는 모델이 denoising task로부터 정상 데이터의 분포를 학습하도록 하기 위해 도출된 feature tokens에 perturbation을 추가하여 input으로 사용함

- Neighbor Masked Attention (NMA)

- 본 연구에서는 input을 복제하는 identical shortcut을 예방하기 위해 neighbor masked attention (NMA)를 제안함

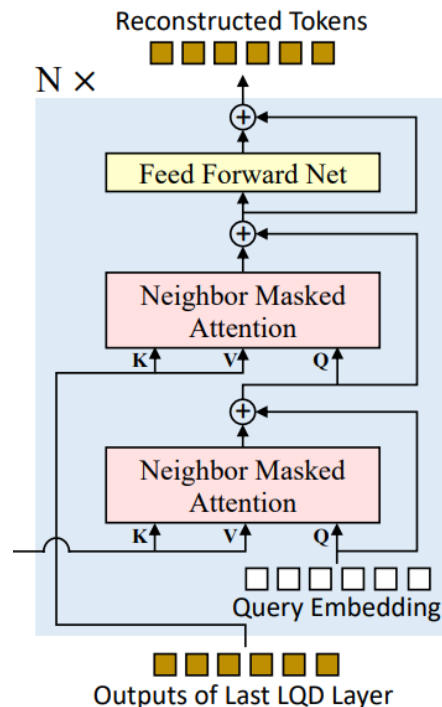
- NMA는 attention map을 계산할 때 특정 영역의 neighbor tokens를 masking하여 유사한 정보를 가진 주변 tokens의 정보를 reconstruction에 사용하지 않음으로써 단순히 input을 거의 복제하여 output을 도출하는 문제를 방지할 수 있음

Proposed method

- Improving feature reconstruction for unified anomaly detection

- Layer-wise Query Decoder (LQD)

- 본 논문에서는 query embedding이 identical shortcut을 예방할 수 있다는 점을 기반으로 모든 layer에서 query embedding을 수행하는 LQD를 제안함
- LQD의 각 layer는 NMA와 feed-forward network로 구성되어 있음



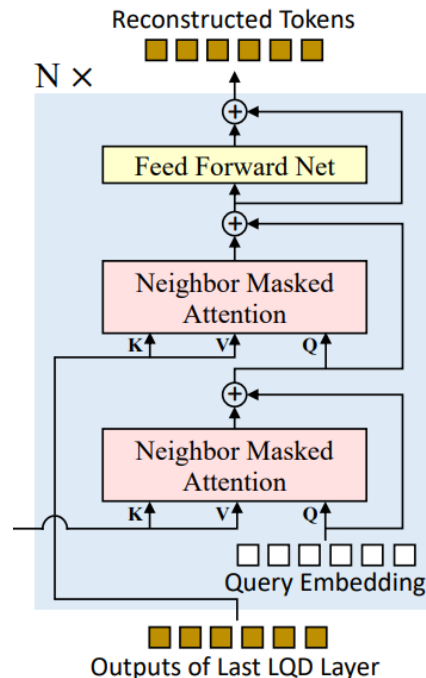
< Layer-wise Query Decoder (LQD) >

Proposed method

- Improving feature reconstruction for unified anomaly detection

- Layer-wise Query Decoder (LQD)

- LQD의 각 layer은 NMA를 통해 learnable query embedding이 encoder embedding과 결합된 후, 이전 layer의 output과 다시 결합됨
- LQD의 마지막 layer에서 도출된 output은 linear projection을 통해 feature tokens의 reconstruction 결과인 reconstructed tokens로 최종 변환됨



< Layer-wise Query Decoder (LQD) >

Experimental results

- Datasets and metrics

- Datasets

- 본 논문에서는 아래 두 데이터셋을 기반으로 여러 개의 class에 대한 하나의 이상치 탐지 모델을 구축하는 unified case에서 모델을 평가함

- ※ MVTEC-AD: 총 15개의 class 각각에 대한 이미지 이상치 탐지 데이터셋으로 구성되어 있음

- ✓ 각 데이터셋은 이미지와 픽셀에 대한 anomaly label을 포함함

- ※ CIFAR-10: 총 10개의 class에 대한 이미지로 구성되어 있음

- Metrics

- AUROC(Area Under the Receiver Operating Characteristic curve)

- ※ Anomaly detection 및 localization의 성능을 평가하는 지표로 사용됨

Experimental results

- Anomaly detection on MVTec-AD

- Unified case에서 UniAD와 6개의 baselines의 성능을 비교함

- Quantitative results

- MVTec-AD의 15개 class에 대해 separate 및 unified case로 각각 평가함

- Baselines은 separate case에 대해 성능이 높은 반면, unified case에 대해서는 성능이 하락함

- 반면, UniAD는 separate case와 unified case 모두에서 높은 성능을 보임

☼ 평균 AUROC는 96.6% (separate case)와 96.5% (unified case)로 나타남

Category	US [6]	PSVDD [49]	PaDiM [11]	CutPaste [26]	MKD [38]	DRAEM [53]	Ours	
Object	Bottle	84.0 / 99.0	85.5 / 98.6	97.9 / 99.9	67.9 / 98.2	98.7 / 99.4	97.5 / 99.2	99.7 ± 0.04 / 100
	Cable	60.0 / 86.2	64.4 / 90.3	70.9 / 92.7	69.2 / 81.2	78.2 / 89.2	57.8 / 91.8	95.2 ± 0.84 / 97.6
	Capsule	57.6 / 86.1	61.3 / 76.7	73.4 / 91.3	63.0 / 98.2	68.3 / 80.5	65.3 / 98.5	86.9 ± 0.73 / 85.3
	Hazelnut	95.8 / 93.1	83.9 / 92.0	85.5 / 92.0	80.9 / 98.3	97.1 / 98.4	93.7 / 100	99.8 ± 0.10 / 99.9
	Metal Nut	62.7 / 82.0	80.9 / 94.0	88.0 / 98.7	60.0 / 99.9	64.9 / 73.6	72.8 / 98.7	99.2 ± 0.09 / 99.0
	Pill	56.1 / 87.9	89.4 / 86.1	68.8 / 93.3	71.4 / 94.9	79.7 / 82.7	82.2 / 98.9	93.7 ± 0.65 / 88.3
	Screw	66.9 / 54.9	80.9 / 81.3	56.9 / 85.8	85.2 / 88.7	75.6 / 83.3	92.0 / 93.9	87.5 ± 0.57 / 91.9
	Toothbrush	57.8 / 95.3	99.4 / 100	95.3 / 96.1	63.9 / 99.4	75.3 / 92.2	90.6 / 100	94.2 ± 0.20 / 95.0
	Transistor	61.0 / 81.8	77.5 / 91.5	86.6 / 97.4	57.9 / 96.1	73.4 / 85.6	74.8 / 93.1	99.8 ± 0.09 / 100
	Zipper	78.6 / 91.9	77.8 / 97.9	79.7 / 90.3	93.5 / 99.9	87.4 / 93.2	98.8 / 100	95.8 ± 0.51 / 96.7
Texture	Carpet	86.6 / 91.6	63.3 / 92.9	93.8 / 99.8	93.6 / 93.9	69.8 / 79.3	98.0 / 97.0	99.8 ± 0.02 / 99.9
	Grid	69.2 / 81.0	66.0 / 94.6	73.9 / 96.7	93.2 / 100	83.8 / 78.0	99.3 / 99.9	98.2 ± 0.26 / 98.5
	Leather	97.2 / 88.2	60.8 / 90.9	99.9 / 100	93.4 / 100	93.6 / 95.1	98.7 / 100	100 ± 0.00 / 100
	Tile	93.7 / 99.1	88.3 / 97.8	93.3 / 98.1	88.6 / 94.6	89.5 / 91.6	99.8 / 99.6	99.3 ± 0.14 / 99.0
	Wood	90.6 / 97.7	72.1 / 96.5	98.4 / 99.2	80.4 / 99.1	93.4 / 94.3	99.8 / 99.1	98.6 ± 0.08 / 97.9
Mean	74.5 / 87.7	76.8 / 92.1	84.2 / 95.5	77.5 / 96.1	81.9 / 87.8	88.1 / 98.0	96.5 ± 0.08 / 96.6	

< Anomaly detection results with AUROC metric on MVTec-AD >

Experimental results

- Anomaly localization on MVTec-AD

- Quantitative results

- Anomaly detection의 정량적 결과와 유사하게, baselines은 separate case에 대해 성능이 높은 반면, unified case에 대해서는 성능이 하락함

- 반면, UniAD는 separate case와 unified case 모두에서 높은 성능을 보임

※ 평균 AUROC는 96.6% (separate case)와 96.8% (unified case)로 나타남

Category		US [6]	PSVDD [49]	PaDiM [11]	FCDD [28]	MKD [38]	DRAEM [53]	Ours
Object	Bottle	67.9 / 97.8	86.7 / 98.1	96.1 / 98.2	56.0 / 97	91.8 / 96.3	87.6 / 99.1	98.1 ± 0.04 / 98.1
	Cable	78.3 / 91.9	62.2 / 96.8	81.0 / 96.7	64.1 / 90	89.3 / 82.4	71.3 / 94.7	97.3 ± 0.10 / 96.8
	Capsule	85.5 / 96.8	83.1 / 95.8	96.9 / 98.6	67.6 / 93	88.3 / 95.9	50.5 / 94.3	98.5 ± 0.01 / 97.9
	Hazelnut	93.7 / 98.2	97.4 / 97.5	96.3 / 98.1	79.3 / 95	91.2 / 94.6	96.9 / 99.7	98.1 ± 0.10 / 98.8
	Metal Nut	76.6 / 97.2	96.0 / 98.0	84.8 / 97.3	57.5 / 94	64.2 / 86.4	62.2 / 99.5	94.8 ± 0.09 / 95.7
	Pill	80.3 / 96.5	96.5 / 95.1	87.7 / 95.7	65.9 / 81	69.7 / 89.6	94.4 / 97.6	95.0 ± 0.16 / 95.1
	Screw	90.8 / 97.4	74.3 / 95.7	94.1 / 98.4	67.2 / 86	92.1 / 96.0	95.5 / 97.6	98.3 ± 0.08 / 97.4
	Toothbrush	86.9 / 97.9	98.0 / 98.1	95.6 / 98.8	60.8 / 94	88.9 / 96.1	97.7 / 98.1	98.4 ± 0.03 / 97.8
	Transistor	68.3 / 73.7	78.5 / 97.0	92.3 / 97.6	54.2 / 88	71.7 / 76.5	64.5 / 90.9	97.9 ± 0.19 / 98.7
	Zipper	84.2 / 95.6	95.1 / 95.1	94.8 / 98.4	63.0 / 92	86.1 / 93.9	98.3 / 98.8	96.8 ± 0.24 / 96.0
Texture	Carpet	88.7 / 93.5	78.6 / 92.6	97.6 / 99.0	68.6 / 96	95.5 / 95.6	98.6 / 95.5	98.5 ± 0.01 / 98.0
	Grid	64.5 / 89.9	70.8 / 96.2	71.0 / 97.1	65.8 / 91	82.3 / 91.8	98.7 / 99.7	96.5 ± 0.04 / 94.6
	Leather	95.4 / 97.8	93.5 / 97.4	84.8 / 99.0	66.3 / 98	96.7 / 98.1	97.3 / 98.6	98.8 ± 0.03 / 98.3
	Tile	82.7 / 92.5	92.1 / 91.4	80.5 / 94.1	59.3 / 91	85.3 / 82.8	98.0 / 99.2	91.8 ± 0.10 / 91.8
	Wood	83.3 / 92.1	80.7 / 90.8	89.1 / 94.1	53.3 / 88	80.5 / 84.8	96.0 / 96.4	93.2 ± 0.08 / 93.4
Mean	81.8 / 93.9	85.6 / 95.7	89.5 / 97.4	63.3 / 92	84.9 / 90.7	87.2 / 97.3	96.8 ± 0.02 / 96.6	

< Anomaly localization results with AUROC metric on MVTec-AD >

Experimental results

- Anomaly localization on MVTec-AD

- Qualitative results

- 아래 case에 해당하는 이미지에 대하여 UniAD의 정성적 결과를 도출함

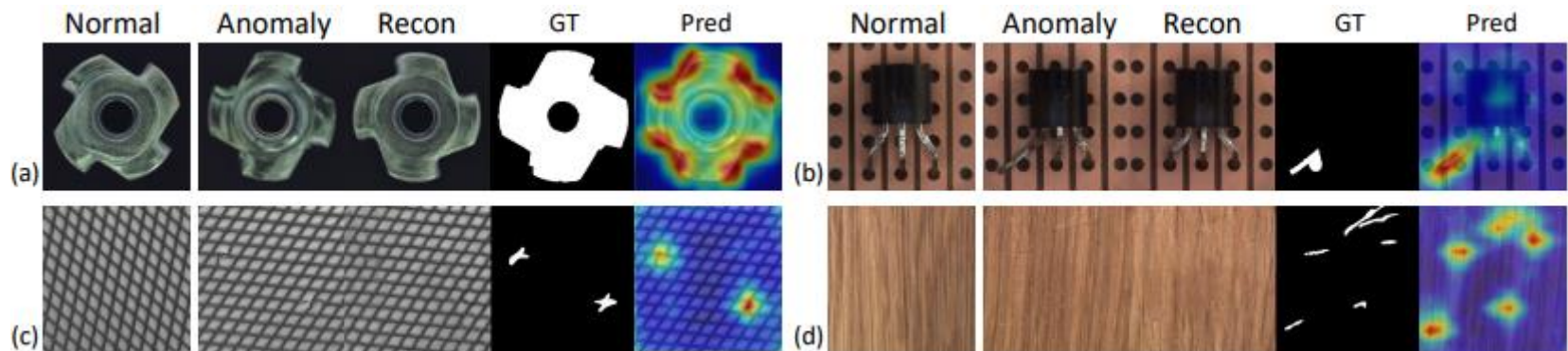
- ※ (a) global structural anomalies

- ※ (b) local structural anomalies

- ※ (c) scattered texture perturbation

- ※ (d) multiple texture scratches

- UniAD가 이상치 이미지를 정상 이미지처럼 복원하는 것을 통해 UniAD에서 identical shortcut 문제가 발생하지 않은 것을 확인할 수 있음



< Qualitative results for anomaly localization on MVTec-AD >

Experimental results

- Anomaly detection on CIFAR-10

- Quantitative results

- Normal Indices: 10개의 class 중 어떤 class를 정상으로 볼 것인지를 설정하는 컬럼
 - 총 10개의 class에 대하여 5개의 class를 정상으로, 5개의 class를 이상치로 구성한 총 4개의 조합에 대하여 multi-class 이상치 탐지 성능을 비교함
 - 본 논문에서 제안한 UniAD가 모든 조합에서 월등하게 높은 성능을 도출함
 - Baselines 중 US와 FCDD는 unified case에서 50%대의 성능을 보임
 - ※ 즉, 이상치 탐지 능력이 상실된 것과 같음
 - Out-of-distribution 모델 Outlier Exposure (OE)를 FCDD에 적용했을 때, multi-class 이상치 탐지 성능이 크게 향상되었지만, UniAD보다는 낮은 성능을 보임

Normal Indices	US [6]	FCDD [28]	FCDD+OE [28]	PANDA [34]	MKD [38]	Ours
{01234}	51.3	55.0	71.8	66.6	64.2	84.4 ± 0.02
{56789}	51.3	50.3	73.7	73.2	69.3	80.9 ± 0.02
{02468}	63.9	59.2	85.3	77.1	76.4	93.0 ± 0.03
{13579}	56.8	58.5	85.0	72.9	78.7	90.6 ± 0.09
Mean	55.9	55.8	78.9	72.4	72.1	87.2 ± 0.03

< Anomaly detection results with AUROC metric on CIFAR-10 >

Conclusions

- 본 논문은 multi-class anomaly detection을 위한 통합 모델인 UniAD를 제안함
- Unified case에서는 identical shortcut 문제가 발생함을 확인함
- Identical shortcut 문제의 발생 정도가 가장 적은 Transformer 모델을 backbone 모델로 사용함
- Transformer 모델에 layer-wise query decoder, neighbor masked attention module 및 feature jittering을 추가하여 identical shortcut 문제를 해결함
- MVTec-AD와 CIFAR-10 데이터셋에서 기존 방법들을 능가하는 성능을 보여주며, 다양한 class에서의 이상 탐지를 효과적으로 수행할 수 있음을 입증함

감사합니다!