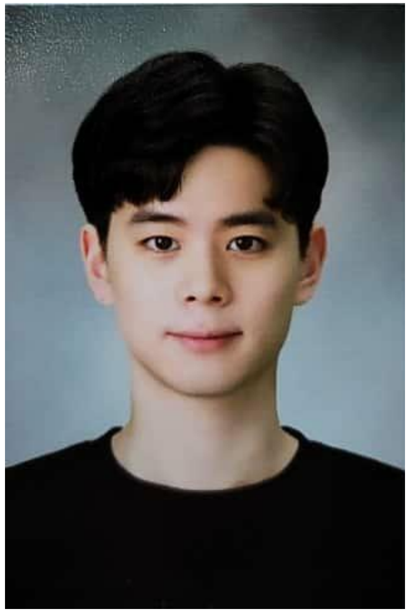
Contrastive Learning for Anomaly Detection

목충협

2023.05.26

발표자



- 목충협 (Chunghyup Mok)
  - ✓ 고려대학교 산업경영공학
  - ✓ Data Mining & Quality Analytics Lab.
  - ✓ 석박통합과정(2019.03 ~)
- 관심분야
  - ✓ Anomaly detection
  - ✓ Self-supervised learning
  - ✓ Contrastive learning
- E-mail : mokch@korea.ac.kr

Contents

Contrastive Learning

Definition

A contrast is a great difference between two or more things which is clear when you compare them.

Contrast : 여러 대상들을 비교할 때 나타나는 명확한 차이

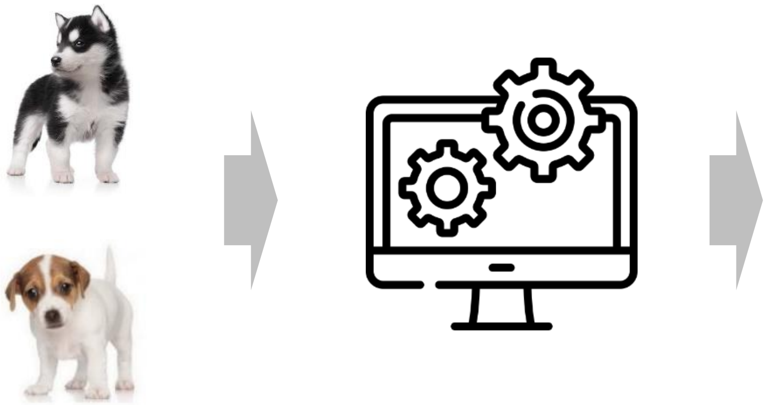
Contrastive Learning : 여러 대상들의 명확한 차이를 활용하는 학습

Contrastive Learning

Definition

데이터간 유사도를 정의하고 모델에 학습하는 방식

학습된 모델은 주어진 데이터들의 차이를 나타낼 수 있다



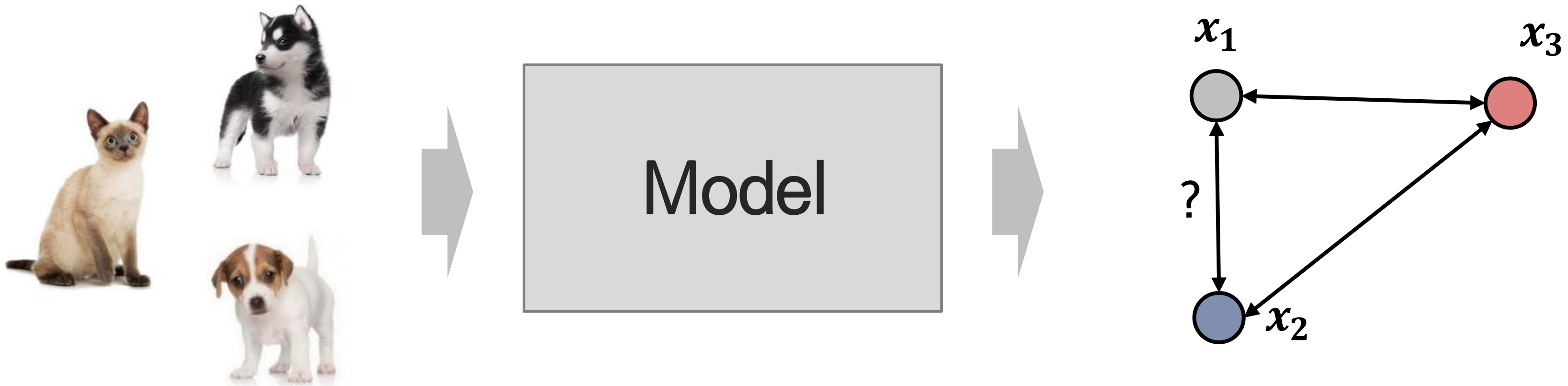
80% 비슷하다

Contrastive Learning

Definition

❖ Feature space에서의 거리 정의 (distance)

1. Non-negativity : $\mathbf{f}(\boldsymbol{x_1}, \boldsymbol{x_2}) \geq \mathbf{0}$
2. Identity of Discernible : $\mathbf{f}(\boldsymbol{x_1}, \boldsymbol{x_2}) = \mathbf{0} \iff \boldsymbol{x_1} = \boldsymbol{x_2}$
3. Symmetry : $\mathbf{f}(\boldsymbol{x_1}, \boldsymbol{x_2}) = \mathbf{f}(\boldsymbol{x_2}, \boldsymbol{x_1})$
4. Triangle inequality : $\mathbf{f}(\boldsymbol{x_1}, \boldsymbol{x_3}) \leq \mathbf{f}(\boldsymbol{x_1}, \boldsymbol{x_2}) + \mathbf{f}(\boldsymbol{x_2}, \boldsymbol{x_3})$



Data space
(Input space)

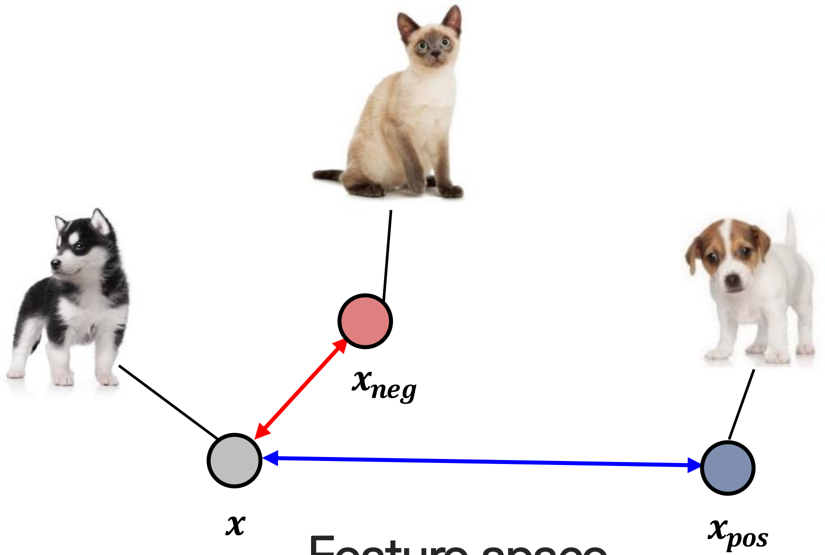
Feature space

Contrastive Learning

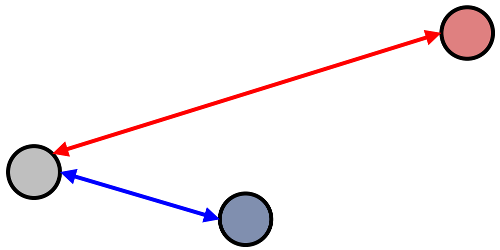
Definition

❖ 데이터간 유사도 정의

- Anchor (x) : 기준이 되는 sample
- Positive sample (x_{pos}) : anchor와 유사한 sample
- Negative sample (x_{neg}) : anchor와 유사하지 않은 sample



Feature space
(학습 전)



Feature space
(학습 후)

Contrastive Learning

Class Discrimination

❖ 데이터의 클래스를 활용하는 contrastive learning

1. Triplet loss
2. N-pair loss
3. Supervised contrastive loss

❖ 데이터간 유사도 정의

- Positive sample (x_{pos}) : anchor와 같은 class에 속하는 sample
- Negative sample (x_{neg}) : anchor와 다른 class에 속하는 sample

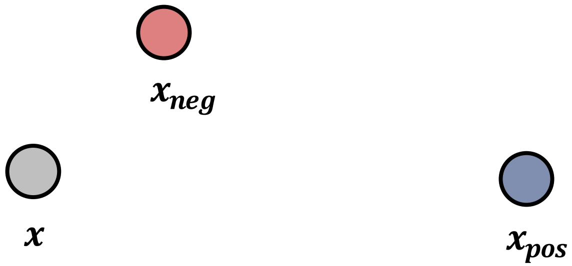
Contrastive Learning

Class Discrimination

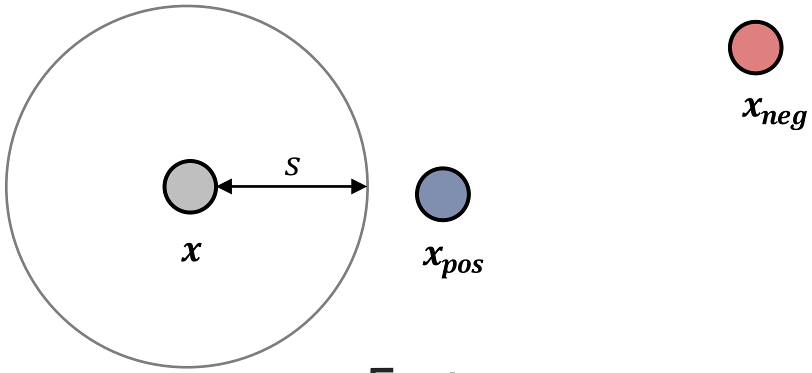
❖ Triplet loss (CVPR, 2015)

- Anchor와 하나의 positive sample과 가까워지도록
- Anchor와 하나의 negative sample과 멀어지도록
- 한 점으로 모이지 않기 위해 최소 거리 s (margin) 설정

$$\mathcal{L}_{triplet}(x,,x_{pos},x_{neg}) = \sum_{x \in X} \max(\|f(x) - f(x_{pos})\|^2 - \|f(x) - f(x_{neg})\|^2 - s, 0)$$



Feature space
(학습 전)



Feature space
(학습 후)

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823).

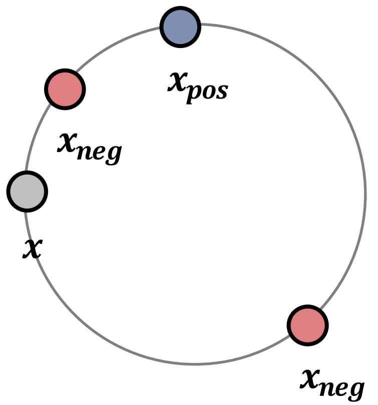
Contrastive Learning

Class Discrimination

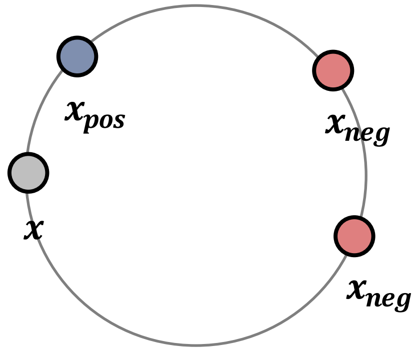
❖ N-pair loss (NeurIPS, 2016)

- 기존의 triplet loss를 확장하여 여러 개의 negative sample들을 고려
- Normalized feature space에서 cosine similarity 사용

$$\mathcal{L}_{N-pair}\left(x,,x_{pos},\{x_{neg}^{i}\}\right) = -\sum_{x \in X} log \frac{\exp\left(f(x)f(x_{pos})\right)}{\exp\left(f(x)f(x_{pos})\right) + \sum_i \exp\left(f(x)f(x_{neg}^{i})\right)}$$



Normalized feature space
(학습 전)



Normalized feature space
(학습 후)

Sohn, K. (2016). Improved deep metric learning with multi-class n-pair loss objective. Advances in neural information processing systems, 29.

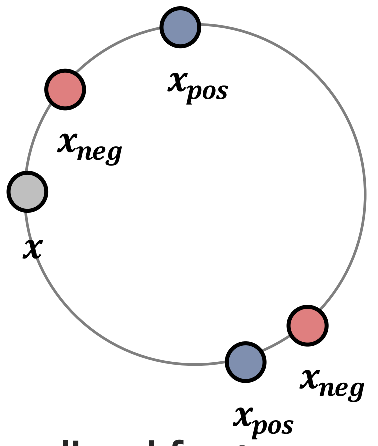
Contrastive Learning

Class Discrimination

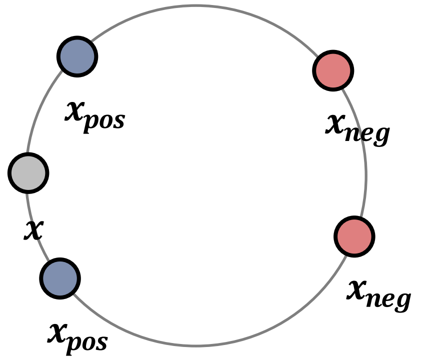
❖ Supervised contrastive learning (NeurIPS, 2020)

- 기존의 N-pair loss를 확장하여, 여러 개의 positive sample들에 대해 고려
- Positive / negative 기준은 class

$$\mathcal{L}_{supcon}(x,,\{x_{pos}^{i}\},\{x_{neg}^{j}\}) = -\sum_{x\in X}\frac{1}{N}\sum_{i=1}^{N} log\frac{\exp\left(f(x)f\left(x_{pos}^{i}\right)/\tau\right)}{\exp\left(f(x)f\left(x_{pos}^{i}\right)/\tau\right)+\sum_{j}\exp\left(f(x)f\left(x_{neg}^{j}\right)/\tau\right)}$$



Normalized feature space
(학습 전)



Normalized feature space
(학습 후)

Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., ... & Krishnan, D. (2020). Supervised contrastive learning. Advances in neural information processing systems, 33, 18661-18673.

Contrastive Learning

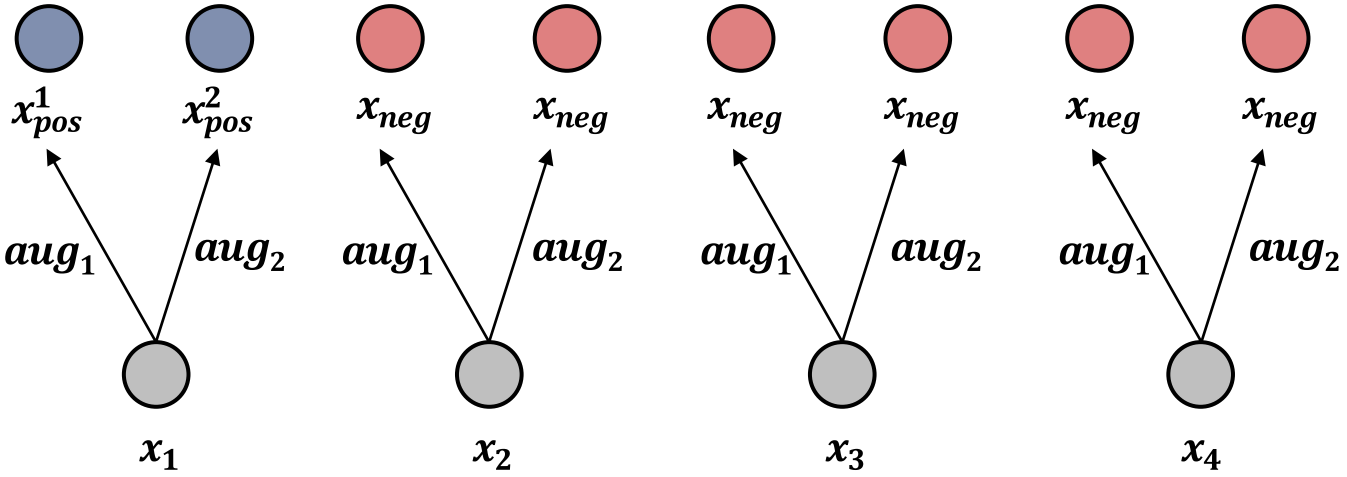
Instance Discrimination

❖ 데이터의 클래스를 활용하지 않는 contrastive learning

- SimCLR loss – InfoNCE (ICML, 2020)

❖ Augmentation 활용하여 유사도 정의

- Augmentation으로 positive/negative sample을 정의
- Positive sample : 같은 sample로 augmented 된 sample
- Negative sample : 다른 sample로 augmented 된 sample



Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020, November). A simple framework for contrastive learning of visual representations. In International conference on machine learning (pp. 1597–1607). PMLR.

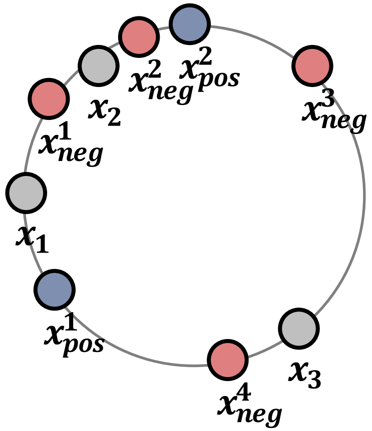
Contrastive Learning

Instance Discrimination

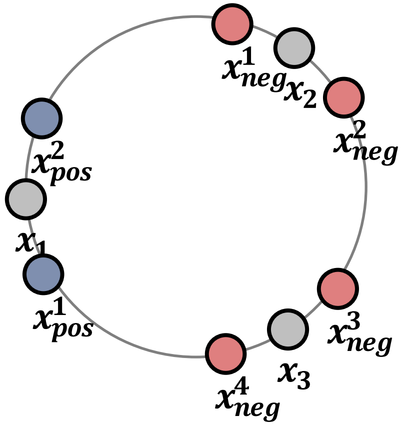
❖ SimCLR (ICML, 2020)

- Positive sample : 같은 sample로 augmented 된 sample
- Negative sample : 다른 sample로 augmented 된 sample

$$\mathcal{L}_{simCLR}(x_{pos}^1,, x_{pos}^2, \{x_{neg}^j\}) = -\sum_{x \in X} log \frac{\exp\left(f\left(x_{pos}^1\right) f\left(x_{pos}^2\right)/\tau\right)}{\sum_i \exp\left(f\left(x_{pos}^1\right) f\left(x_{neg}^i\right)/\tau\right)}$$



Normalized feature space
(학습 전)



Normalized feature space
(학습 후)

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020, November). A simple framework for contrastive learning of visual representations. In International conference on machine learning (pp. 1597-1607). PMLR.

Contrastive Learning

❖ Contrastive learning의 효과

- 정확도와 강건함을 높일 수 있음

  → 기존의 cross entropy의 경우 label noise에 취약함

❖ 기존의 cross entropy loss에 추가하여 사용

- 더 좋은 representation learning을 위해 혼합하여 사용

  → Cross entropy loss + triplet loss / simCLR

❖ 사전학습 방식으로 사용

- 적절한 augmentation과 같이 활용하면 기존의 지도학습보다 더 좋은 성능을 보임

  → Supervised contrastive learning

Contrastive Learning for Anomaly Detection

Anomaly Detection

❖ 심각한 데이터 불균형

- 많은 양의 정상 데이터와 극소수의 이상 데이터로 구성됨

❖ 레이블 없음

- 레이블이 제한적이거나 없을 수 있음

❖ 이상 데이터의 다양성

- 하나의 클래스로 정의할 수 없음

❖ 사전 지식의 부족

- 학습 단계에 없었던 새로운 이상 데이터가 나타날 수 있음

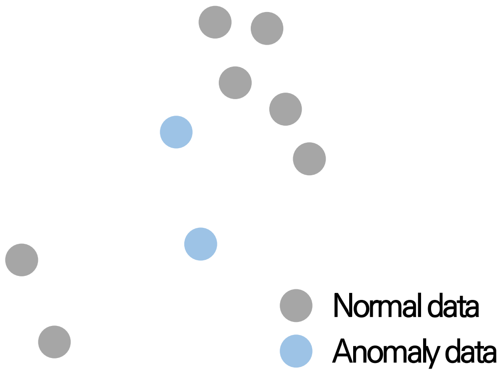
Contrastive Learning for Anomaly Detection

Deep one-class classification

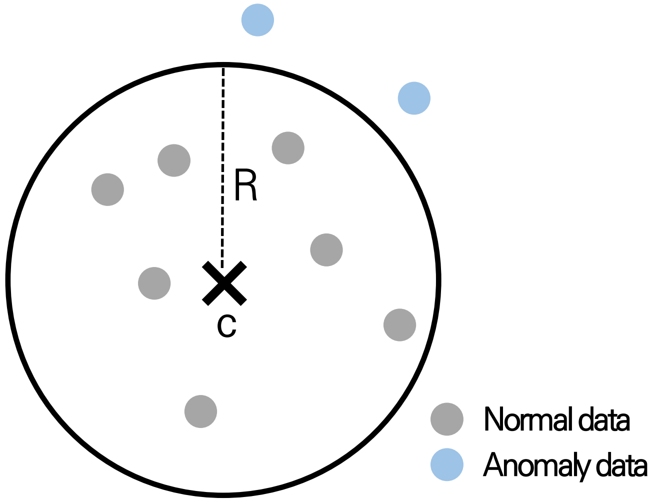
❖ Representation for anomaly detection

- 정상 데이터는 하나의 중심점(c)에 모이도록 학습
- 중심점으로부터 일정 거리(R) 이상 떨어져 있는 경우 혹은 밀도가 낮은 경우 이상으로 판단

  → SVDD / OCSVM 과 유사한 방식



Feature space
(학습 전)



Feature space
(학습 후)

Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., ... & Kloft, M. (2018, July). Deep one-class classification. In International conference on machine learning (pp. 4393-4402). PMLR.

Contrastive Learning for Anomaly Detection

Class Discrimination

❖ Classification-based anomaly detection for general data – GOAD (ICLR 2020)

- Augmentation을 이용하여 새로운 클래스(class 2, 3, 4) 생성

→ Class 정보를 활용하는 contrastive learning 적용

Class 1

Original data

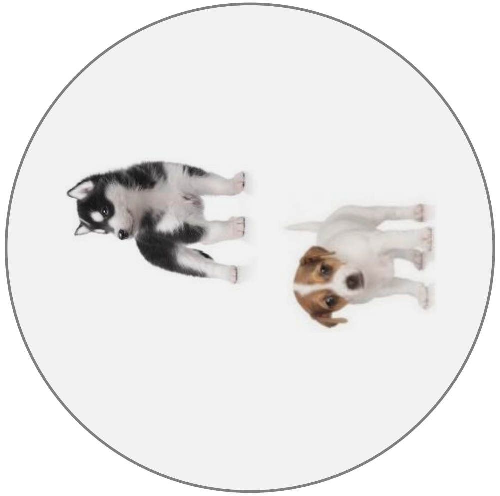
Class 2

Rotated data (90°)

Class 3

Rotated data (180°)

Class 4



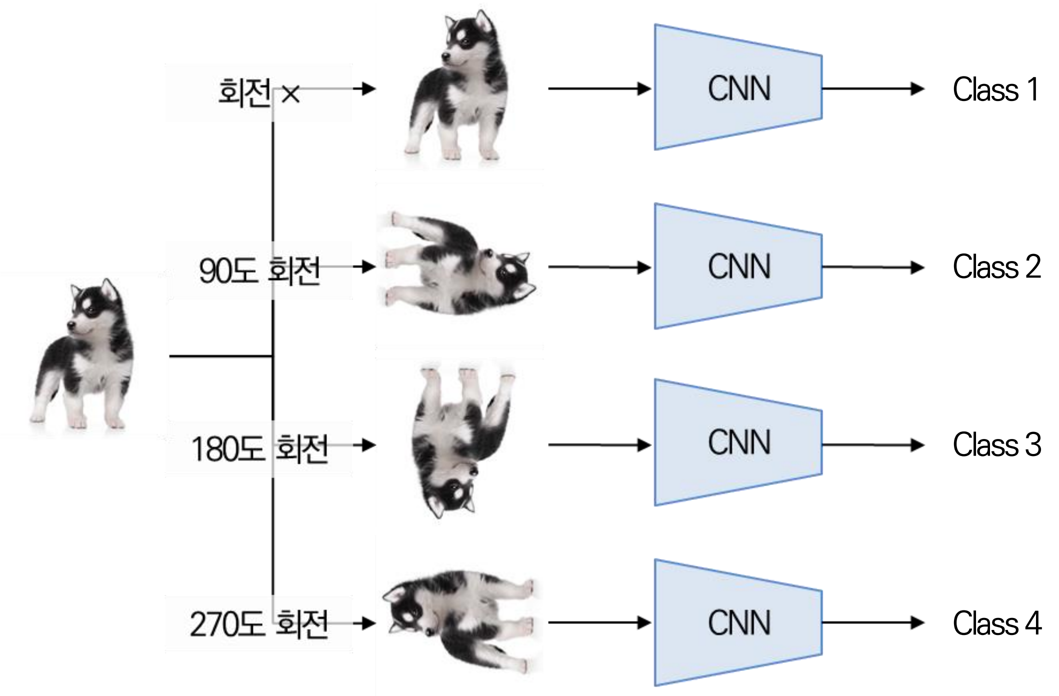
Rotated data (270°)

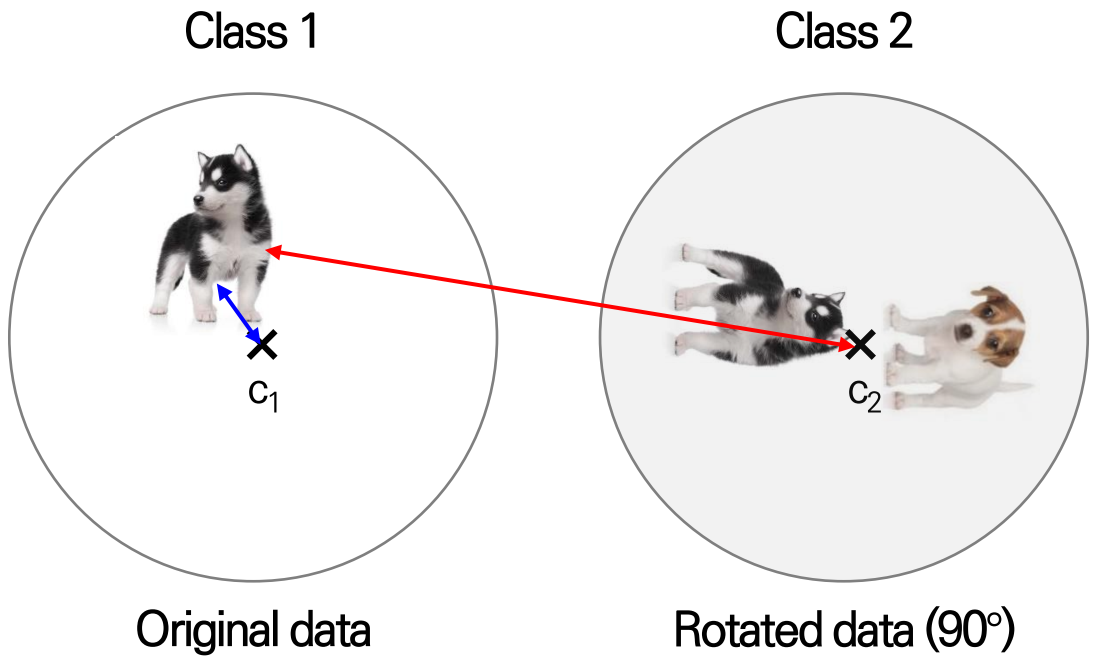
Contrastive Learning for Anomaly Detection

Class Discrimination

❖ Classification-based anomaly detection for general data – GOAD (ICLR 2020)

- Cross entropy loss + triplet loss
- Positive sample : 해당 클래스의 중심점
- Negative sample : 다른 클래스 중 가장 가까운 클래스의 중심점



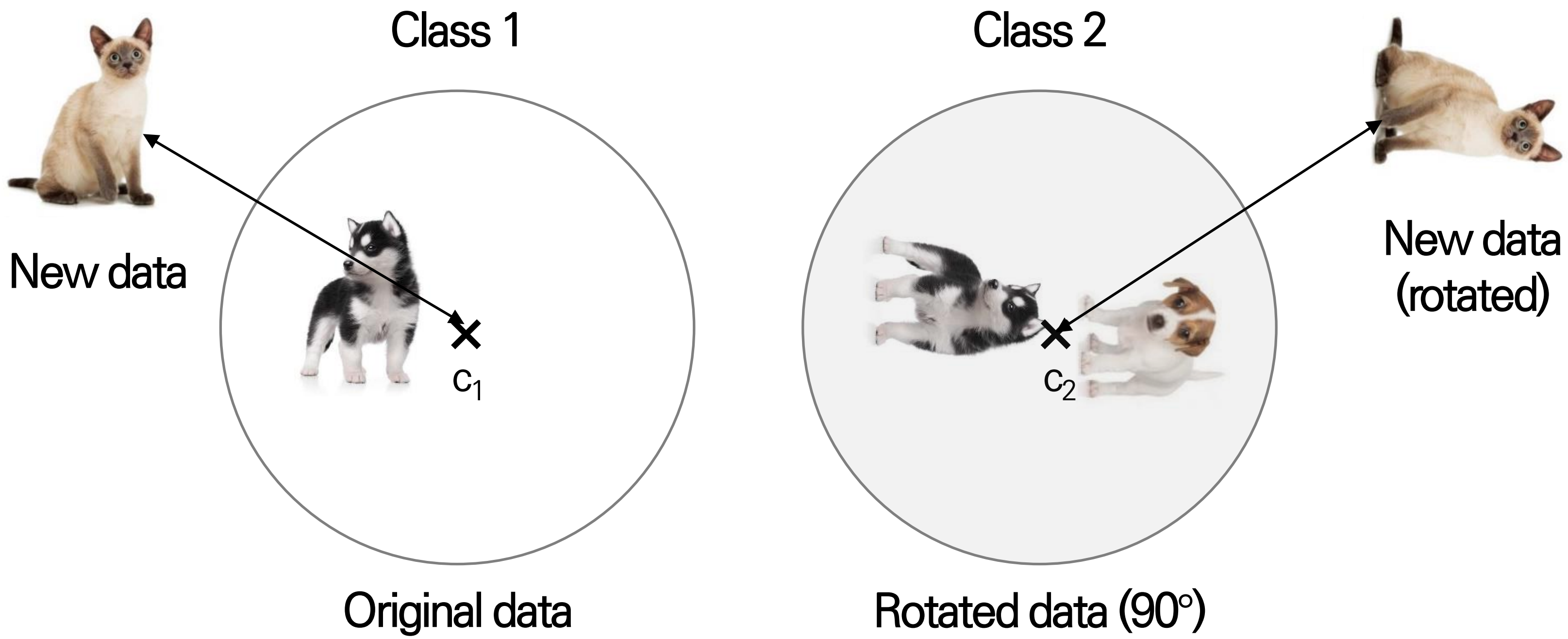


Contrastive Learning for Anomaly Detection

Class Discrimination

❖ Classification-based anomaly detection for general data – GOAD (ICLR 2020)

- 새로운 데이터가 들어오면, 동일한 augmentation을 적용함
- 각 class의 중심점과의 거리를 이용하여 이상 탐지 진행



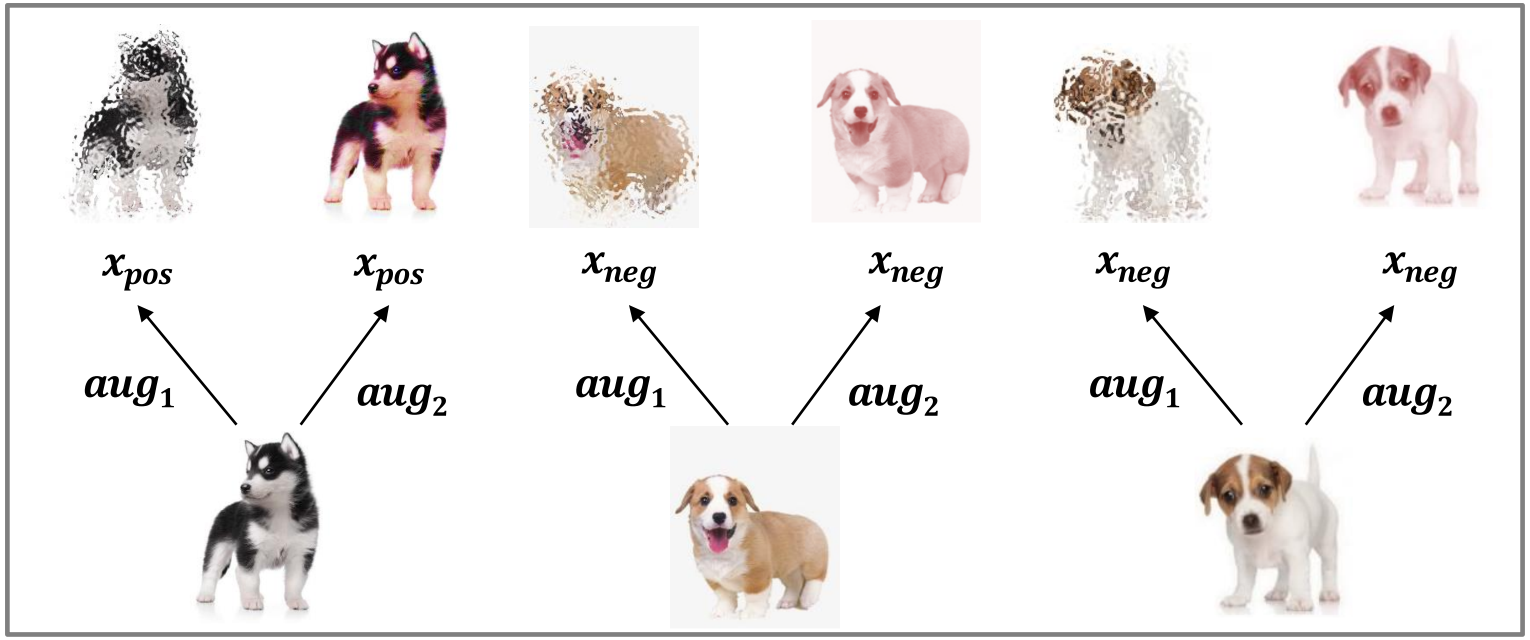
Contrastive Learning for Anomaly Detection

Instance Discrimination

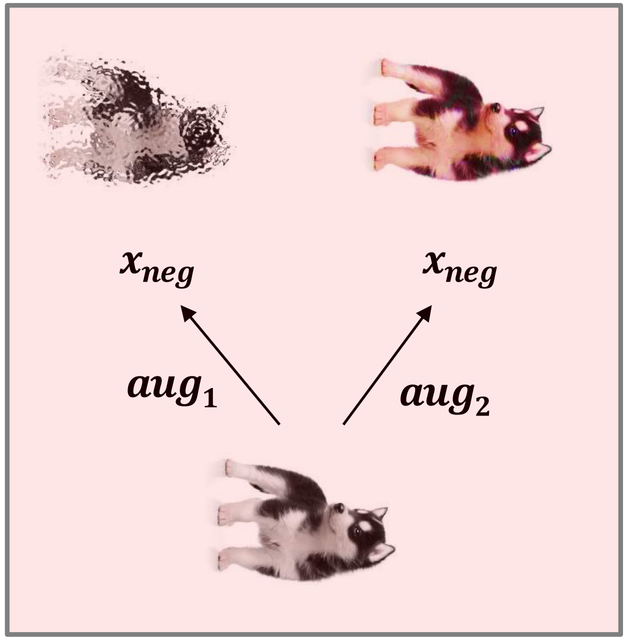
❖ Learning and evaluating representations for deep one-class classification (ICLR, 2021)

- 2-stage framework 제안
  1. Anomaly detection에 적합한 representation을 학습

     → SimCLR에 negative sample을 추가하여 학습에 사용



SimCLR



추가된 negative sample

Contrastive Learning for Anomaly Detection

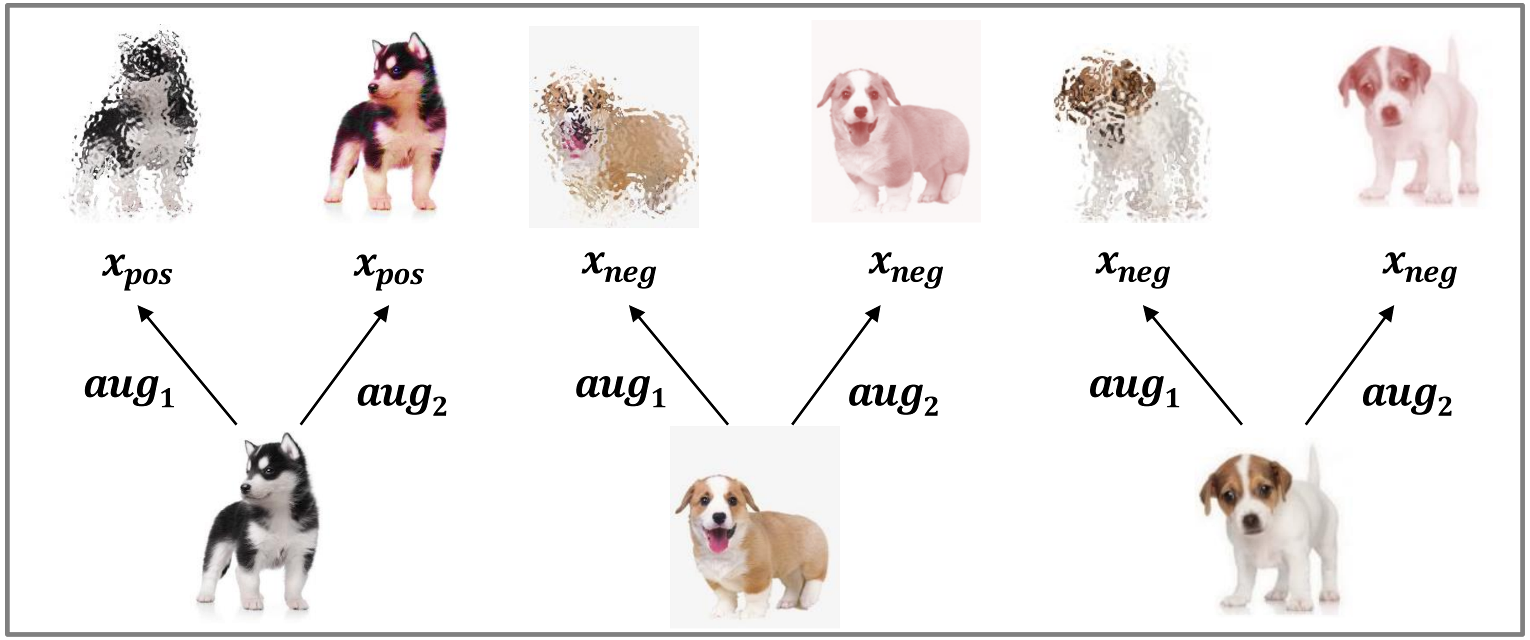
Instance Discrimination

❖ Learning and evaluating representations for deep one-class classification (ICLR, 2021)

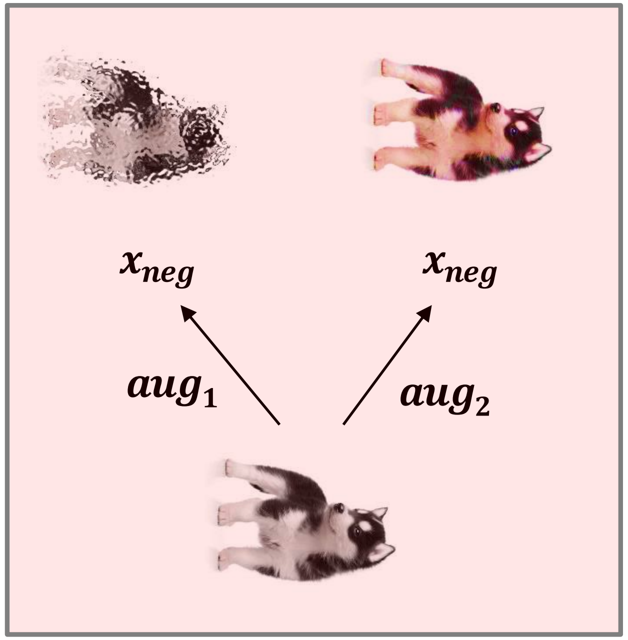
- 2-stage framework 제안

    1. Anomaly detection에 적합한 representation을 학습

        → SimCLR에 negative sample을 추가하여 학습에 사용



SimCLR



추가된 negative sample

Contrastive Learning for Anomaly Detection

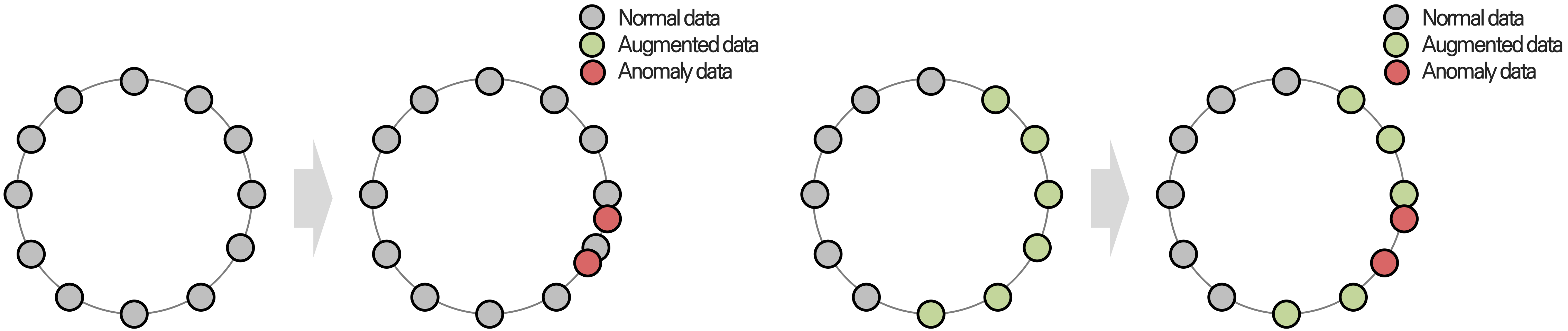
Instance Discrimination

❖ Learning and evaluating representations for deep one-class classification (ICLR, 2021)

- 2-stage framework 제안

  1. Anomaly detection에 적합한 representation을 학습

     → SimCLR에 negative sample을 추가하여 학습에 사용



SimCLR 로 학습된 feature space

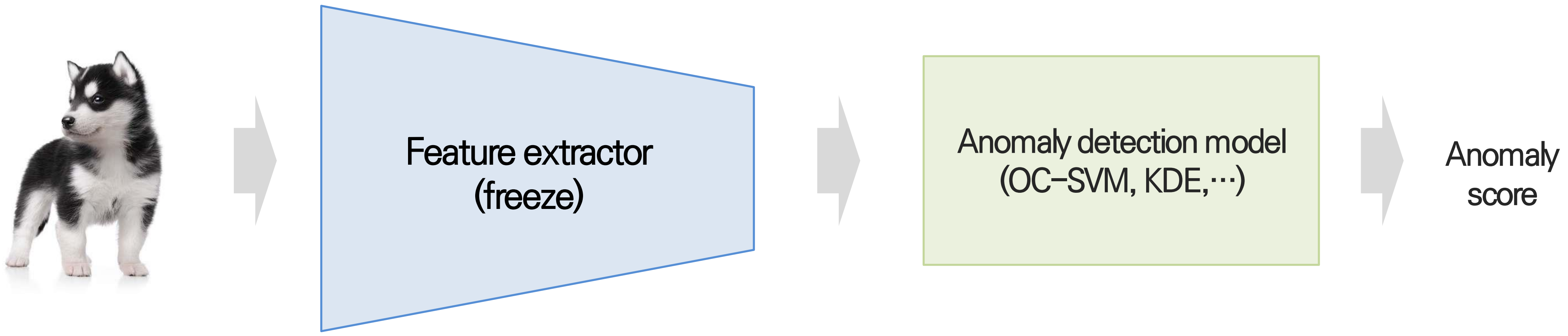
Negative sample을 추가하여 학습된 feature space

Contrastive Learning for Anomaly Detection

Instance Discrimination

❖ Learning and evaluating representations for deep one-class classification (ICLR, 2021)

- 2-stage framework 제안
  1. Anomaly detection에 적합한 representation을 학습
     → SimCLR에 negative sample을 추가하여 학습에 사용
  2. 간단한 anomaly detection 모델을 사용하여 이상 탐지



Contrastive Learning for Anomaly Detection

Instance Discrimination

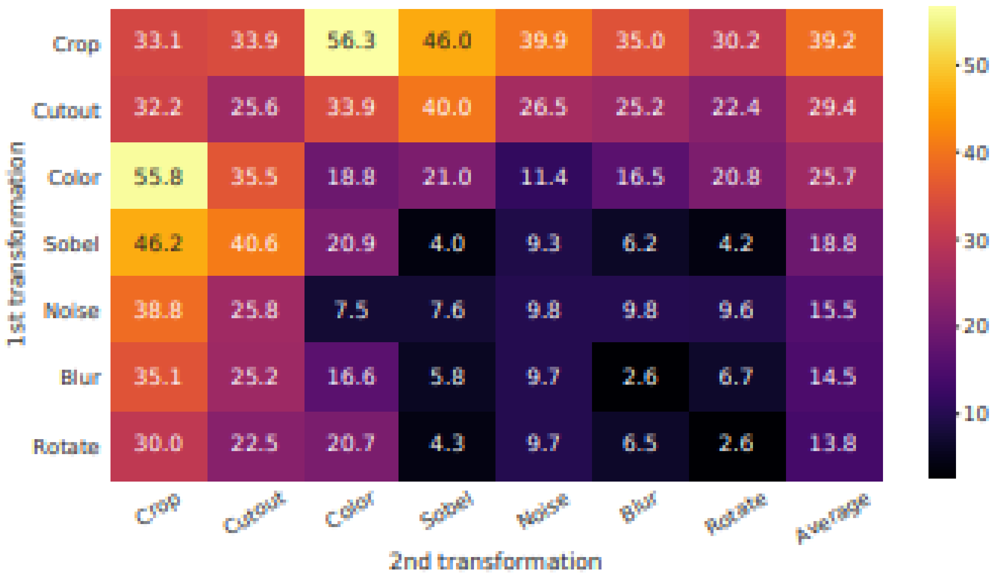
❖ CSI: Novelty Detection via Contrastive Learning on Distributionally Shifted Instances (NeurIPS 2020)

- One-class OOD detection

1. Augmentation의 종류에 따라 positive / negative sample 정의

→ SimCLR에서 augmentation을 적용했을 때, 성능이 향상하는 경우 positive sample로 정의

→ SimCLR에서 augmentation을 적용했을 때, 성능이 감소하는 경우 negative sample로 정의



Contrastive Learning for Anomaly Detection

Instance Discrimination

❖ CSI: Novelty Detection via Contrastive Learning on Distributionally Shifted Instances (NeurIPS 2020)

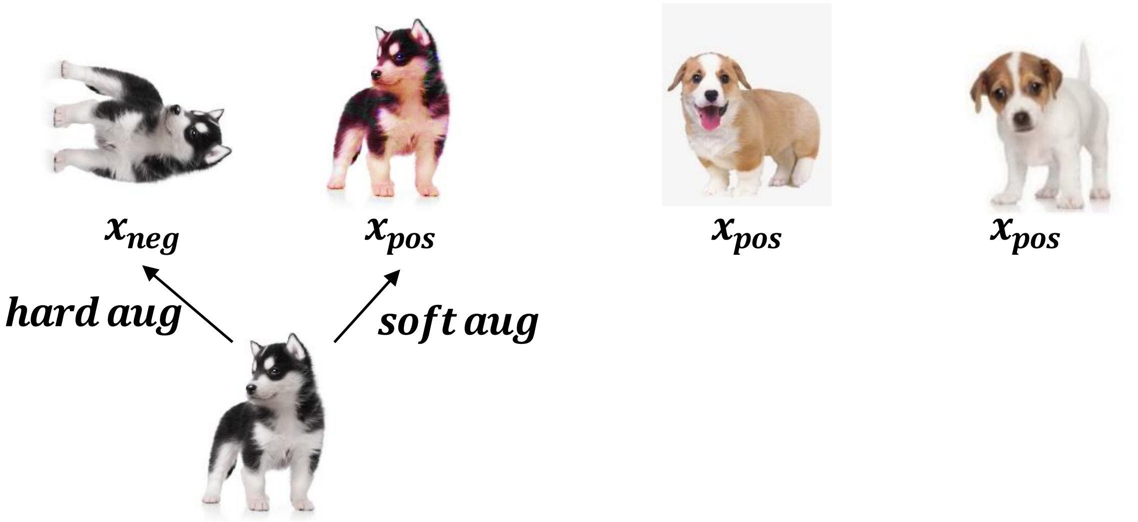
- One-class OOD detection

1. Augmentation의 종류에 따라 positive / negative sample 정의

   → SimCLR에서 augmentation을 적용했을 때, 성능이 향상하는 경우 positive sample로 정의

   → SimCLR에서 augmentation을 적용했을 때, 성능이 감소하는 경우 negative sample로 정의

2. 나머지 정상 샘플들을 positive sample로 정의

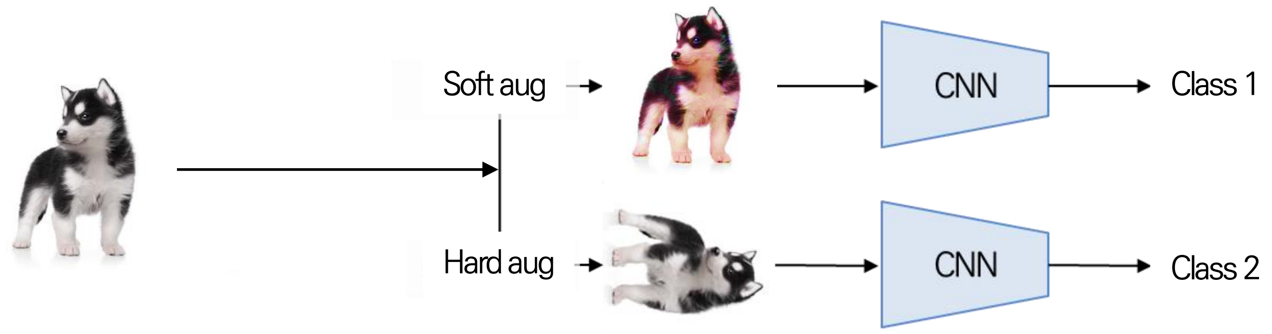


Contrastive Learning for Anomaly Detection

Instance Discrimination

❖ CSI: Novelty Detection via Contrastive Learning on Distributionally Shifted Instances (NeurIPS 2020)

- One−class OOD detection
  - SimCLR loss + Cross entropy loss



Contrastive Learning for Anomaly Detection

❖ Contrastive learning은 강건하고 좋은 representation을 학습할 수 있다.

❖ Positive / negative sample을 잘 정의하면 anomaly detection 에도 적용할 수 있다.

- Hard augmentation → negative sample
- Normal sample / soft augmentation → positive sample

❖ Augmentation에 의존적이기 때문에, augmentation을 적용하기 힘든 데이터에는 적용하기 어렵다.

감사합니다.