

# Implicit Neural Representations를 이용한 이상 탐지

신은호, 전윤호

한밭대학교

20221055@edu.hanbat.ac.kr, yhjeon@hanbat.ac.kr

## Anomaly Detection using Implicit Neural Representations

Shin Eun Ho, Yunho Jeon

Hanbat Univ.

### 요약

제조 현장에서 불량 검출을 찾는 것이 중요하지만 인공지능을 이용한 탐지는 이상치 데이터의 부족으로 정답레이블이 없는 비지도 학습이 주로 사용된다. 본 논문에서는 재구성 방식인 Implicit Neural Representations(INR)이 이상 탐지에 사용될 가능성을 확인해 보고자 한다. 비교를 위해 기본적인 Convolutional Autoencoder(CAE)를 사용하여 비교해 보았고 CAE와 INR의 학습 과정에 차이로 인해 INR이 CAE보다 불량인 부분의 학습이 늦게 된다는 것을 확인하여 INR 이상 탐지에 활용될 가능성이 있음을 확인하였다.

## I. 서론

제조 현장에서는 제품의 불량을 탐지하고 찾아내는 것이 중요하다. 하지만 이상 데이터는 정상데이터보다 상대적으로 적기 때문에 비지도 학습방법 중 하나인 재구성 방식을 활용한 이상 탐지가 많이 활용되고 있다. 그중 CAE(Convolution Autoencoder)는 자주 쓰는 방법의 하나다. 하지만 CAE가 비정상인 데이터까지 잘 학습한다는 한계점이 존재한다[1]. 이러한 한계점을 해결하기 위해서 또 다른 이미지 재구성 방법 중 하나인 INR(Implicit Neural Representations)를 이용하는 방법을 생각했다.

본 논문에서는 INR이 이미지를 학습할 때 어떠한 과정으로 학습하는지 보고 CAE와 비교하여 INR이 이상 탐지를 할 수 있는지의 가능성을 확인하는 것이 목표이다.

## II. 재구성 방식을 활용한 이상 탐지 방식

### 2.1 CAE

CAE[2]는 Autoencoder를 변형한 모델로 합성곱 계층을 이용하여 이미지 처리가 잘 되는 특징이 있다. 입력 데이터의 패딩을 학습해서 최대한 유사한 데이터를 새롭게 생성하는 것이 목표이다.

CAE의 구조는 크게 보면 인코더, 디코더로 구성되어 있다. 인코더에서 입력 데이터가 합성곱과 pooling 연산을 통해 데이터의 공간적인 특징이 추출되면서 저차원의 벡터로 전환된다. 디코더에서는 인코더의 반대과정이 수행된다. 데이터가 다시 원본 데이터와 같은 고차원 벡터로 재구성된다.

CAE로 이상 탐지를 하는 아이디어는 CAE가 정상데이터의 잠재공간을 학습하고 이상 데이터가 들어오면 CAE가 학습된 정상데이터처럼 재구성하여서 생기는 원본 이미지 사이의 재구성 오류를 구해 임계값 이상이 나오는 데이터에 대해 이상치라고 판단하여 구한다.

### 2.2 INR

INR[3]은 신경망을 사용하여 데이터를 효과적으로 표현하는 개념이다. 기존 연구들은 이미지를 처리할 때 보통 RGB 값을 가지는 행렬로 표현하는데 컴퓨터가 이미지를 처리하는 방식을 함수로 생각하면 컴퓨터는 입력 신호인  $x, y$  좌표로 받아 R, G, B 각각의 값으로 출력하는 함수로 생각해 볼 수 있다. INR은 이러한 함수를 신경망으로 표현해 그 신경망을 학습하여 이미지를 효과적으로 재구성하는 것이 목표이다.

본 논문에서 사용한 INR 구조[4]는 SIREN (Sinusoidal Representation) 모델을 사용했다. INR도 크게 인코더와 디코더로 구성되어 있는데 인코딩은 이미지를  $I$ 라 하고  $I(x, y)$ 가 RGB 값을 반환하도록 한다. 파라미터를  $\theta$ 로 픽셀 위치를 이미지의 RGB 값으로 매핑하는 함수를  $f_{\theta}(x, y) = (R, G, B)$ 로 정의하고 이미지에  $f_{\theta}$ 를 일종의 왜곡 측정 하에 overfitting 하여 인코딩하는 방식이다. 디코딩은 인코딩에서 저장된 양자화된 가중치  $\theta$ 가 주어지면 단순히 함수  $f_{\theta}$ 를 모든 픽셀 위치에서 평가하여 이미지를 재구성한다.

INR 이상 탐지 아이디어는 INR이 이미지를 재구성할 때 이미지의 일부를 먼저 디코딩하거나 낮은 해상도 이미지를 디코딩함으로써 이미지를 점진적으로 디코딩하는 방식으로 학습 도중 적절하게 멈추면 불량이 있는 부분은 재구성이 제대로 되지 않아 그 재구성 오류를 확인하여 이상치 부분을 탐지하는 아이디어이다.

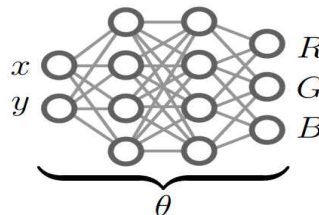


그림 1. INR 구조

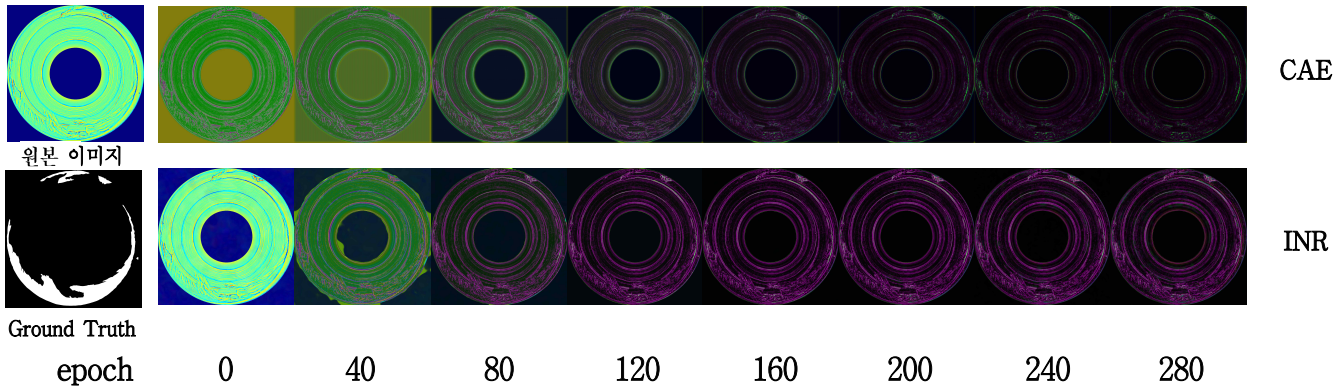


그림 2 (좌) 상단 이미지는 원본, 하단 이미지는 Ground Truth (우) 원본과 재구성 이미지 차이를 표시

### III. 실험 결과 및 분석

#### 3.1 데이터셋

사용한 데이터는 BTAD(BeanTech Anomaly Detection)[5]을 사용하였다. BTAD는 세 가지 산업용 제품의 RGB 이미지이다. 정상인 학습 이미지와 정상과 불량 두 가지가 있는 테스트용 이미지, 테스트 데이터 중 불량인 이미지의 Ground Truth가 있다. Product 1은 이미지 크기가 1600x1600인 이미지 400장, Product 2는 이미지 크기가 600x600인 이미지 1000장, Product 3은 이미지 크기가 800x600인 이미지 399장이 포함되어 있다. 제품 중 Product 1을 사용 하였고 test 데이터 중 이상치가 있는 0번째 data를 사용했다.

#### 3.2 실험 방법

실험 방법으로는 CAE와 INR 각각 같은 이미지를 재구성하도록 한가지의 이미지로만 학습하였다. 그림 2의 좌상단 이미지는 학습에 쓴 원본 이미지이고 좌하단 이미지는 불량인 부분을 표시한 Ground Truth이다. CAE는 입력 값으로 원본 이미지를 넣었고 INR은 입력값으로 원본 이미지의 2D 픽셀 위치를 좌표로 표현한 x, y 값을 사용했다. CAE와 INR의 학습이 어떻게 되는지 비교를 위해 학습 epoch 따라 원본과의 차이를 비교하였다. 원본 이미지와 재구성 이미지 차이는 절댓값을 사용하여 이미지 차이를 구할 때 음수 값이나 큰 값이 결과의 변화를 주지 않게 했다.

#### 3.3 실험 결과

실험 결과로 그림 2의 오른쪽 이미지들은 학습에 따른 원본과 재구성의 이미지 차이를 시각화한 이미지이고 위쪽이 CAE이고 아래쪽이 INR이다. 이미지의 유의미한 변화를 보기 위해 학습 수는 40단위로 비교하여 결과를 확인하였다.

학습 과정에 대해서 CAE는 불량인 부분도 초기부터 학습이 되는 반면 INR은 불량인 부분이 CAE보다 늦게 학습이 되는 것을 그림 2에서 볼 수 있다.

### IV. 결론

본 논문에서는 비지도 학습 중 하나인 CAE와 INR의 학습 과정을 비교하여 INR을 활용한 이상 탐지 가능성을 확인해 봤다. 학습 과정에서 불량인 부분도 초기부터 학습되는 CAE와 달리 INR은 이미지에서 불량인 부분이 학습이 늦게 되는 과정이기 때문에 학습 수에 따른 이상 탐지를 기대할 수 있을 것 같다.

향후 연구에서 INR을 사용하여 효과적으로 이상 탐지를 할 수 있는 모델을 만들 예정이다.

### 참 고 문 헌

- [1] Gong, D., Liu, L., Le, V., Saha, B., Mansour, M. R., Venkatesh, S., and Hengel, A. V. D. "Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection." Proceedings of the IEEE/CVF International Conference on Computer Vision pp. 1705-1714, 2019.
- [2] Guo, Xiaojie, Liu, Xinwang, Zhu, En, Yin, Jianping, Yu, Yizhou, and Huang, Junzhou. "Stacked Convolutional Auto-Encoders for Hierarchical Feature Extraction." International Journal of Computer Vision, vol. 121, no. 2, pp. 336-356, 2017.
- [3] Sitzmann, Vincent, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and GordonWetzstein . "Implicit neural representations with periodic activation functions." Advances in Neural Information Processing Systems, vol. 33, pp. 7462-7473, 2020.
- [4] Dupont, E., Goliński, A., Alizadeh, M., Teh, Y. W., & Doucet, A.. "Coin: Compression with implicit neural representations." ICLR Neural Compression Workshop, 2021.
- [5] P. Mishra, R. Verk, D. Fornasier, C. Piciarelli, G.L. Foresti "VT-ADL: A Vision Transformer Network for Image Anomaly Detection and Localization" 30th IEEE/IES International Symposium on Industrial Electronics (ISIE) Kyoto, Japan, June 20-23, 2021.