

스마트 팩토리를 위한 정상 객체 마스크를 사용하는 오토인코더 기반의 이상 동작 검출

홍상욱, 김형원*

충북대학교

swhong95@chungbuk.ac.kr, *hwkim@cbnu.ac.kr

Motion Anomaly Detection Based on Autoencoder With Normal Object Masking For Smart Factory

Sang Wook Hong, Kim Hyung Won*

Chungbuk National University

요약

오토인코더[1]는 입력 이미지를 인코딩한 후 디코딩하여 출력 이미지를 생성하며 이 출력 이미지가 입력 이미지와 같은 이미지가 되도록 오토인코더를 학습시킨다. 따라서 오토인코더는 비지도 학습이 가능한 장점을 가지며 많은 분야에 사용되고 있다. 이 오토인코더의 특성을 이용한 분야 중 Abnormal detection은 과적합이 일어난 오토인코더의 입출력의 차이를 통해 classification 또는 detection을 수행한다. 이 오토인코더가 새로운 객체를 인식하도록 훈련하기 위해서는 많은 데이터가 필요하며, 이는 일반적으로 데이터 증강[3] 기법을 이용해 이루어진다. 그러나 기존의 오토인코더 기반의 오류 검출 기술은 반복적인 패턴의 비디오 영상에 대해서만 학습 효과를 나타내며, 스마트팩토리의 작업자나 이동 로봇등의 불규칙적인 객체의 형상과 불규칙적인 동작에 대해서는 학습을 통해 정상으로 판정하기 매우 어려운 문제를 가지고 있다. 이를 개선하기 위해 본 논문에서는 정상 동작 마스크를 생성하는 보조 객체검출 CNN을 오토인코더 네트워크에 통합하여 정상 또는 이상 동작 판정의 정확도를 크게 향상하는 기술을 제안한다. 본 제안 기술은 객체검출 CNN 학습하여 판정하고자 하는 객체들을 검출하고 이들의 Bounding box를 마스크로서 오토인코더가 생성하는 Loss Map에 적용하여 최종 이상 동작 검출을 판정하는 기술이다. 제안 딥러닝 모델을 최적화 하였으며, 스마트팩토리에 적용하기 위한 초소형 임베디드 AI 가속기 모듈에 구현하였다. 또한 기존 오토인코더 기술과 오동작 검출 정확도 분석 결과를 제시한다.

I. 서론

오토인코더[1]는 입력을 인코딩하고 디코딩해 입력과 같은 출력을 만드는 DNN 네트워크이다. 이러한 오토인코더의 특성을 이용하여 과적합이 일어난 오토인코더의 입출력의 차이를 통해 Abnormal detection은 classification 또는 detection을 수행한다.

특정 물체를 이상 상태 감지 오토인코더에서 비정상이 아니라 정상 상황으로 판단하고 싶은 경우에 일반적으로 해당 데이터를 포함한 추가적인 학습 데이터를 사용하여 네트워크의 훈련을 다시 진행한다.[3] 하지만 이상 상태 감지 오토인코더는 다양한 상황들을 훈련시키는 경우 네트워크의 일반화 능력이 증가하여 정상 상황이 아닌 더 많은 입력 복구하게 되어 정확도가 감소한다. 이를 개선하기 위해 본 논문에서 보조 CNN을 이용한 실시간 Roi 변경 제안한다.

II. 본론

오토인코더[1]는 입력을 의미있는 특징들로 압축하는 인코더와, 이것을 다시 입력과 같게 복구하는 디코더로 구성되어있는 DNN 네트워크이다. 최초의 오토인코더는 fully-connected layer로 구성되어 많은 파라미터를 가져 훈련이 느리고 많은 메모리가 필요했지만 현재 많이 사용되는 Convolutional Autoencoder(CAE)(그림 1)[4]는 bottleneck 부분을 제외한 나머지 부분을 convolution layer로 구성하여 더 적은 메모리를 소모하고

더 빠른 훈련이 가능하다.

이상 상태 감지 오토인코더[2]는 외부의 입력을 통해 입력된 데이터를 판단하여 해당 입력이 정상상태인지 비정상인지 판단하는 DNN

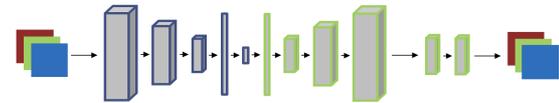


그림 1 CNN 오토인코더

네트워크이다. 기존의 classification을 이용한 이상 상태 감지는 정상 상황이 대부분을 차지하는 데이터 불균형 문제로 인하여 훈련에 제한사항이 다수 존재했다. 이를 해결하기 위해 정상 상황만을 사용한 비지도 학습이 가능한 오토인코더를 도입하게 되었다. 오토인코더를 정상 상황에 과적합 시킴으로서 생성 능력에 제한을 주어 네트워크가 정상 상황만을 복구 할 수 있게 훈련을 시키고 입력과 출력을 비교해 입력의 classification과 localization을 진행한다. 이상 상태 감지 오토인코더는 일반적인 classification과 같이 결과가 class로 나오는 것이 아닌 입력과 출력의 차이(loss)로 나온다. 이 loss를 계산하기 위한 Loss function(evaluation metrics)은 일반적으로 L1, L2, SSIM, PSNR 등을 사용한다. 이 차이(loss)를 thresholding하여 정상과 비정상을 구분한다.

이상 상태 감지 오토인코더가 학습하지 못한 특정 정상적인 객체를 비정상으로 판단하는 오류를 줄이기 위해서는 객체를 포함한 데이터에 대한

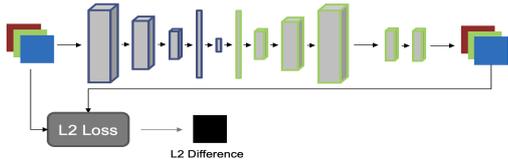


그림 2 이상 상태 감지 오토인코더

추가적인 학습이 필요하다. 그러나 이러한 데이터는 전체 훈련 데이터 중 극히 일부분을 차지하므로 데이터 불균형의 문제가 있다. 이를 해결하기 위해 데이터 증강을 사용하였을 시 오토인코더가 너무 많은 종류의 입력을 복구 하게 되거나 해당 물체가 훈련 데이터와 정확히 같은 위치에 있을 때만 정상 상황으로 판단하는 문제가 존재한다.

이 문제를 해결하기 위해 본 논문에서 비정상적으로 감지하는 것을 원하지 않는 물체로만 훈련된 보조 탐지 네트워크에서 생성된 Roi 마스크 사용을 제안한다. Roi 마스크를 사용하여 해당 물체를 loss map에서 마스킹해 가중치를 줌으로서 오토인코더의 과적합을 유지하면서 특정 종류의 물체는 복구하지 못하더라도 정상으로 판단하게 하여 정확도를 높인다.

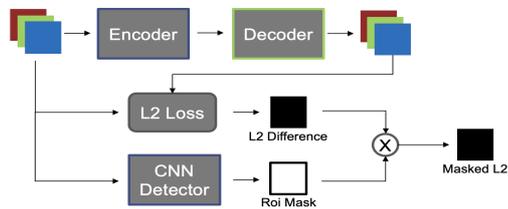


그림 3 보조 CNN 네트워크를 사용하는 이상 상태 감지 오토인코더

보조 CNN detector(그림 3)의 성능을 시험하기 위해 데이터 증강을 통해 정상으로 판단할 객체, 비정상적으로 판단할 객체를 이미지에 넣어 정상, 비정상, 비정상(정상객체 포함)(그림4) 3가지 테스트 데이터를 생성한다. 원본 데이터에도 소량의 정상객체(사람)가 포함되어 있다.

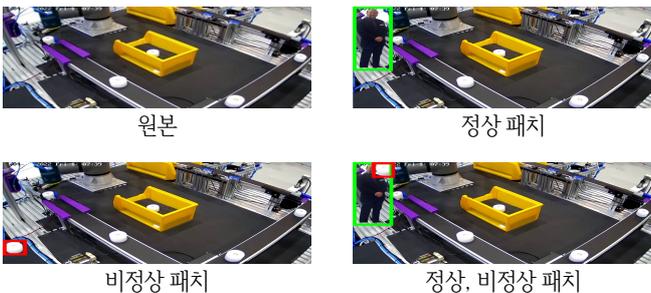


그림 4 데이터셋

그림4의 구조를 가진 네트워크를 이용해 L2 loss에 마스킹을 한 결과이다. Autoencoder는 VAE[6] 기반의 오토인코더, 보조 네트워크는 YOLOv5n(object detection)[7], Mask R-CNN(Instance segmentation)[5][7], Deeplab v3(semantic segmentation)[5][8]을 사용했다.

III. 결론

표1의 결과를 확인할 시 보조 네트워크가 있는 경우들이 정확도가 증가한다는 것을 확인할 수 있다. 하지만 YOLOv5n같이 너무 작은 모델은 detection 정확도 문제로 마스킹이 제대로 동작하지 않는다는 문제가 있었고, Deeplab v3는 정상 객체에 대한 마스크가 객체의 크기보다 작게 나온다는 문제점이 발견되었다. 이 둘과 다르게 Instance segmentation은 정상 객체를 비교적 정확히 분할하여 가장 정확도가 크게 증가하였다.

	원본(정상)	정상 패치	비정상 패치	정상비정상패치	평균
Vanilla	0.87	0.02	0.97	1.0	0.72
YOLOv5n	0.89	0.00	0.97	1.0	0.72
MASKR-CNN	0.95	0.77	0.96	0.96	0.91
DeepLab v3	0.94	0.43	0.97	0.99	0.83

표 1 보조 네트워크 모델별 정확도

ACKNOWLEDGMENT

This work was supported by Regional Leading Research Center (RLRC) of the National Research Foundation of Korea (NRF) grant funded by the Korean government(MSIT)(No. 2022R1A5A8026986), and supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2020-0-01304, Development of Self-learnable Mobile Recursive Neural Network Processor Technology). It was also supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Communication Technology Research Center support program (IITP-2022-2020-0-01462) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

The correspondence author is Pf.Kim Hyungwon.

참고 문헌

- [1] Dor Bank, Noam Koenigstein, Raja Giryes, "Autoencoders", 2020
- [2] Raghavendra Chalapathy, Sanjay Chawla, "DEEP LEARNING FOR ANOMALY DETECTION: A SURVEY", 2019
- [3] Connor Shorten & Taghi M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning:", 2019
- [4] Jonathan Masci, Ueli Meier, Dan Cireşan, and Jürgen Schmidhuber, "Stacked Convolutional Auto-Encoders for Hierarchical Feature Extraction", 2011
- [5] Abdul Mueed Hafiz, Ghulam Mohiuddin Bhat, "A Survey on Instance Segmentation: State of the art", 2020
- [6] Jinwon An, Sungzoon Cho, Variational Autoencoder based Anomaly Detection using Reconstruction Probability
- [7] Glenn Jocher, Alex Stoken, Jirka Borovec, NanoCode012, Ayush Chaurasia, TaoXie, Liu Changyu, Abhiram V. Laughing, tkianai, yxNONG, Adam Hogan, lorenzomamma, AlexWang1900, Jan Hajek, Laurentiu Diaconu, Marc, Yonghye Kwon, oleg, wanghaoyang0106, Yann Defretin, Aditya Lohia, ml5ah, Ben Milanko, Benjamin Fineran, Daniel Khromov, Ding Yiwei, Doug, Durgesh, and Francisco Ingham. "ultralytics/yolov5", AWS, Supervise.ly and YouTube integrations, Apr. 2021.
- [8] Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick, "Mask R-CNN", 2017
- [9] Liang-Chieh Chen, George Papandreou, Florian Schroff, Hartwig Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation", 2017