

공정 자동화를 위한 복원 모델 기반 산업 이상 탐지 연구

김은완, 황준화, 서동준*
경북대학교

{olligo20, jhhwang1907, *dongjunsuh}@knu.ac.kr

A Study on Industrial Anomaly Detection in Process Automation using Reconstruction-based Models

Eunwan Kim, Junhwa Hwang, Dongjun Suh *
Kyungpook National University

요 약

산업 자동화 공정은 인건비 감소, 생산성 안정화 등의 다양한 이점을 제공한다. 그러나 자동화 공정은 제조 불량, 기계 결함 등의 문제가 나타날 수 있어 현재까지는 인적 관리가 필요한 단계이다. 따라서 산업에서 이상을 탐지하는 문제는 공정이 자동화로 발전함에 따라 중요한 과제로 떠오르고 있다. 본 연구는 산업 공정에서 생산된 제품의 이상 탐지를 수행하기 위해 대표적인 산업 데이터로 알려진 Mvtec AD 데이터셋을 이용하여 이상 탐지 모델 개발 연구를 진행하였다. 우수한 이상 탐지를 위해 Cutpaste 모델에 Reconstruction 기반 모델을 결합하였으며, 제안된 모델에서 향상된 이상 탐지 성능을 확인하였다.

I. 서론

최근 산업 분야는 현대 기술의 발전으로 다양한 분야에서 공정 자동화가 이루어지고 있다. 공정 자동화는 생산성 안정화, 인건비 감소, 운영 효율성 등의 이점이 있다. 하지만 자동화된 공정에서 사람의 개입없이 공정의 안정성과 제품 품질을 유지하는 것은 여전히 해결되어야 할 과제로 남아있다. 또한 자동화에 사용되는 기계의 유지보수, 결함 관리 역시 사람에게 관리되고 있어, 완전한 공정 자동화를 위해서는 불량 검출 및 기계 결함 탐지 등 추가적인 연구가 필요하다.

이상 탐지는 정상 샘플에서 이상 샘플을 탐지해내는 기술로 정상 샘플의 기준에 따라 활용 분야가 다양하다. 산업 분야에서는 기계 소리 데이터나 불량 이미지 등을 이용하여 불량 검출, 기계 결함 탐지 등으로 활용이 가능하며 이에 따라 이상 탐지에 대한 많은 관심과 연구가 이루어지고 있다.

하지만 실제 데이터는 이상치의 비율이 현저히 낮아 데이터 불균형 현상이 발생하며, 실제 이상 샘플은 그 종류가 방대하여 이상을 탐지하기가 쉽지 않다. 따라서 데이터의 양이나 품질에 상관없이 좋은 성능의 모델을 구축하는 것이 중요하다. 해당 연구에서는 데이터 증강 기법을 활용하여 데이터 불균형 현상과 노이즈 견고성을 해결하였으며, 성능 향상을 위해 Reconstruction 기반 모델을 활용하였다.

본 연구는 대표적인 산업 데이터인 Mvtec AD 를 활용하여 산업 공정에서의 이상 탐지 연구를 진행하였다. 이상 탐지 모델은 Classification 모델 중 하나인 Cutpaste 를 베이스 모델로, 이상 탐지에서 좋은 성능을

보여주는 Reconstruction 기반 모델을 결합하여 사용하였다.

II. 본론

2.1 데이터셋

본 연구에는 산업 데이터의 특징을 고려하기 위해 대표적인 산업 데이터셋인 Mvtec AD 를 사용하였다. Mvtec AD 는 산업 이상 탐지 연구에 대한 데이터셋으로, 15 개의 클래스 및 5000 개의 고해상도 이미지로 구성되어 있다. 15 개 클래스는 알약, 케이블, 병 등 산업 공정에서 흔히 볼 수 있는 물체로 구성되어 있다. 또한, Mvtec AD 는 정상과 결함 데이터를 모두 포함하고 있으며, 모든 결함에 대해 픽셀 단위로 라벨링 되어 있어 성능 평가에 최적화되어 있는 데이터셋이다 [1].

제공되는 학습셋에는 결함 데이터가 포함되지 않아 정상 데이터로만 학습되기 때문에 입력의 정보가 누출되어 identical shortcut 문제가 발생하거나 노이즈에 취약할 수 있다는 한계가 있다. 따라서 데이터 증강 기법을 통해 정상 샘플에서 다양한 비정상 샘플을 생성하고, 이를 학습셋으로 사용하여 보완하였다. 실험에는 Mvtec AD 를 사용하여 모델 학습과 성능 평가를 진행하였다.

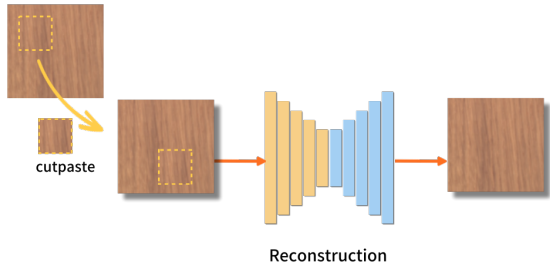
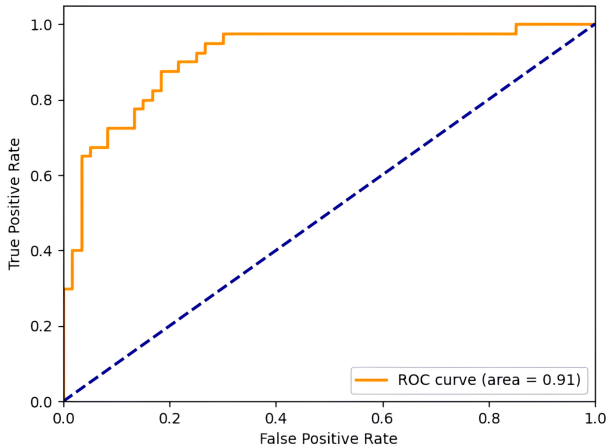
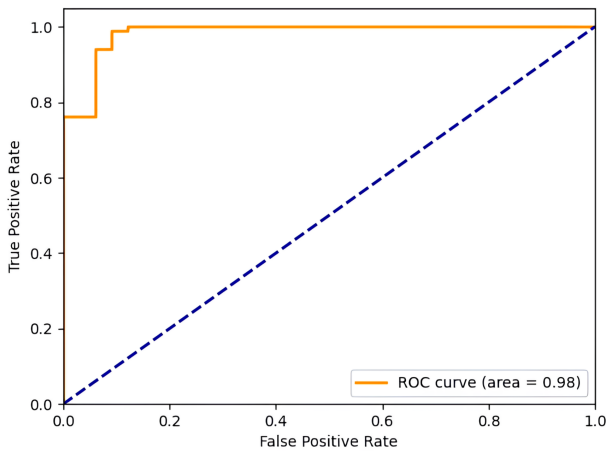


Fig 1. Framework of Reconstruction based model



(a) Cutpaste model



(b) Cutpaste + Reconstruction

Fig 2. Performance of Mvtec AD

2.2 이상 탐지 모델

본 논문에서는 기존 Classification 모델에 Reconstruction 기반 모델을 결합하여 Mvtec AD 산업 공정 이상 탐지 연구를 진행하였다. 비교 모델로 사용된 Cutpaste는 Classification 모델 중 하나로, cut & paste 데이터 증강 기법을 적용한 최신 모델이다. 이 모델은 정상 이미지를 자르고 붙여 비정상 이미지를 생성하는 데이터 증강을 통해 비정상이 섞인 학습셋을 구성하고 정상과 비정상을 분류하는 모델에 학습한다. 분류된 feature 분포를 통해 GDE(Gaussian Density Estimation)기법으로 이상 샘플을 탐지한다 [2]. 기존 Cutpaste 방법은 정상 이미지에서 비정상 이미지를 정교하게 생성할 수 있다는 점에서 우수하지만, 간단한

모델 구조로 구성되어 있다. 따라서 이상 탐지 성능을 보완을 위해 기존 Cutpaste 모델에 Reconstruction 기반 모델을 적용하였다.

Reconstruction 기반 이상 탐지 모델은 최근 이상 탐지 연구에서 좋은 성능으로 주목받고 있는 방법으로, 이미지를 Reconstruction 하는 과정을 통해 입력과 출력의 차이를 계산하고 이를 이용하여 이상 샘플을 탐지하는 기법이다. 별도의 라벨링을 필요로 하지 않으며, Reconstruction 하는 과정에서 모델이 데이터의 중요한 정보들만 압축적으로 학습할 수 있어 우수한 성능을 보여준다. 그림 1은 Reconstruction 기반 모델의 구조를 나타낸 것이다 [3].

본 연구에서는 Reconstruction 기반 이상 탐지 모델 중 Autoencoder 기반의 모델을 사용하였다. 정상과 비정상 이미지를 모두 학습하여 정상 이미지로만 학습 시 발생할 수 있는 identical shortcut 문제, 노이즈 견고성 등을 해결하여 모델 성능을 향상시켰다.

이상 탐지 성능 평가를 위해 AUC 평가 지표를 사용하였다. ROC curve의 아래 영역을 의미하는 AUC는 1.0에 가까울수록 높은 성능을 의미한다. 그림 2은 Mvtec AD 데이터셋에서 각 모델의 성능을 평가한 결과로 ROC Curve를 나타낸 것이다. 두 모델의 성능을 평가한 결과, 0.91에서 0.98로 제안된 모델에서 향상된 성능 결과를 확인하였다.

III. 결론

본 연구에서는 Reconstruction 기반 모델을 이용하여 공정 자동화를 위한 산업 이상 탐지 연구를 진행하였다. Mvtec AD 데이터셋을 사용하여 산업 데이터의 특징을 고려하였으며, 기존 모델을 보완하여 모델을 구현하였다. 그 결과 제안된 모델의 성능이 향상된 것을 확인할 수 있었다. 추후 연구에서는 다양한 산업 데이터들을 고려한 이상 탐지 모델의 성능 연구를 수행할 예정이다.

ACKNOWLEDGMENT

본 연구는 정부(과학기술정보통신부, 교육부)의 재원으로 한국연구재단의 지원을 받아 수행되었음 (과제번호: 2021R1A5A803165, 2021R1I1A3049503).

참고 문헌

- [1] Bergmann, Paul, et al. "A comprehensive real-world dataset for unsupervised anomaly detection." 2019, (<https://www.mvtec.com/company/research/datasets/mvtec-ad>)
- [2] Li, Chun-Liang, et al. "Cutpaste: Self-supervised learning for anomaly detection and localization." 2021.
- [3] Zavrtnik, Vitjan, et al. "Draem-a discriminatively trained reconstruction embedding for surface anomaly detection." 2021.