

오토인코더를 활용한 MCT 이상탐지 알고리즘 개발

김민희 · 진교홍*

창원대학교

Autoencoder-based MCT Anomaly Detection Algorithm

Min-hee Kim · Kyo-hong Jin*

Changwon National University

E-mail : mini7961@changwon.ac.kr / khjin@changwon.ac.kr

요 약

제조 현장에서 설비의 이상이나 고장은 곧 제품의 결함을 발생시키는 요인이 된다. 최근 스마트팩토리 서비스의 확산에 따라 고장을 예측 및 방지를 위한 인공지능 기반의 애플리케이션에 대한 연구가 활발히 이어지고 있으나, 일반적으로 설비의 이상이나 고장 데이터는 정상에 비해 그 개수가 현저히 적어 분류 모델 개발에 큰 어려움이 있다. 본 논문에서는 오토인코더 모델의 입력과 출력 간 차이를 이용한 제조 현장의 MCT 이상탐지 알고리즘을 제안하고 성능을 분석하였다. 해당 알고리즘은 비정상 데이터가 존재하지 않는 MCT 제조 데이터에서 정상 데이터에 대한 특징만을 이용하여 비정상 탐지한다.

ABSTRACT

In a manufacturing fields, an abnormality or breakdown of equipment is a factor that causes product defects. Recently, with the spread of smart factory services, a lot of research to predict and prevent machine's failures is actively ongoing. However, there is a big difficulty in developing a classification model because the number of abnormal or failure data of the machine is severely smaller than normal data. In this paper, we present an algorithm for detecting abnormalities in an MCT at manufacturing work site depending on the differences between inputs and outputs of Autoencoder model and analyze its performance. The algorithm detects abnormalities using only features of normal data from manufacturing data of the MCT in which abnormal data does not exist.

키워드

Smart factory, Manufacturing data, Autoencoder, Anomaly detection, MCT

1. 서 론

제조업체의 가장 큰 자산이라 할 수 있는 생산 설비의 고장은 전체 생산라인의 가동 중단을 초래해 막대한 손실을 야기할 수 있다. 이에 생산설비의 주요 구성요소에 센서를 부착해 수집된 제조 데이터와 MES(Manufacturing Execution Systems)나 ERP(Enterprise Resource Planning) 시스템을 통해 얻은 공정 데이터를 활용하여 생산설비의 상태를 예측하는 인공지능 기반의 이상탐지 모델의 개발이 활발히 진행되고 있

다[1].

이러한 이상탐지 모델의 개발은 수집된 제조 데이터의 특성에 크게 좌우된다. 일반적으로 제조 데이터는 극심한 클래스 불균형(Class Imbalance)의 특징을 가진다[2,3]. 뿐만 아니라, 제조 데이터에는 생산설비의 정상 상태와 비정상 상태를 구분하는 정보가 포함되어 있지 않기 때문에, 지도학습(Supervised Learning)의 분류 알고리즘보다는 오토인코더(Auto-encoder)나 GAN(Generative Adversarial Networks)[4]과 같은 비지도학습(Unsupervised Learning)을 통해 이상치를 판별하는 방식이 더 적합하다[5].

MCT(Machining Center Tool)는 여러 공구

* corresponding author

를 자동으로 교체하면서 제품을 가공하는 공작기계이다. 본 논문에서는 MCT의 이상을 탐지하기 위한 오토인코더 기반의 이상탐지 알고리즘을 소개한다. 개발된 알고리즘은 CNN(Convolutional Neural Network)[6]을 기반으로 한 모델을 사용하여, 인코더와 디코더로 구성된 모델의 입출력 간차를 이용한다[7]. 수집된 특정 MCT 제조 데이터에 단순이동평균(SMA:Simple Moving Average)을 취해서 학습에 용이하도록 변환한 후 정상 데이터로만 모델을 학습시킨다. 그리고 새로운 데이터를 모델의 입력으로 사용했을 시 발생하는 재구성 손실값을 확인하여 설비의 이상을 탐지한다. 재구성 손실값에 대한 임계값을 설정하여 결과 정보를 부여하고, 이상탐지율을 확인한다.

본 논문의 구성은 다음과 같다. 2장에서는 MCT 이상탐지 알고리즘에 대해 설명하고 3장에서는 개발한 알고리즘의 성능을 분석하였다. 마지막으로 4장에서는 결론 및 향후 연구를 기술하였다.

II. MCT 이상탐지 알고리즘

1. MCT 제조 데이터 및 공정 데이터

MCT 제조 데이터는 MCT 내부에 부착된 전력, 온도, 진동, 압력, 습도 센서를 통해 1초 간격으로 수집되는 센서 데이터이다. 그리고 MCT 공정 데이터는 설비명, 제품코드, 작업의 시작과 종료, 설비의 이상유무, 제품의 품질 등의 이벤트 정보를 기록한 데이터로 MES와 현장 작업자에 의해 기록된다. 표 1과 표 2는 MCT 제조 데이터와 MCT 공정 데이터의 예시이다.

MCT 머신에서 생산되는 한 제품의 공정은 3가지 단계로 이루어지고, 총 720초의 공정시간을 갖는다. MCT 제조 데이터에 포함된 다섯 종류의 센서 데이터를 상관분석하여 전력과 진동 데이터의 상관관계가 높다는 것을 확인한 다음, MCT 공정 데이터에 포함된 제품생산 구간 정보와 MCT 제조 데이터에서 반복적으로 나타나는 유사 패턴의 비교분석을 통해 전력과 진동 데이터가 이상탐지 모델의 학습에 적합하다고 판단하였다.

표 1. MCT 제조 데이터 예시

	일시	전력	온도	진동	유압	습도
1	2021-05-10 4:36:51 PM	2626.8	25.83	0.1	41.78	25.31
2	2021-05-10 4:36:52 PM	2414.7	25.82	0.1	41.02	25.33
3	2021-05-10 4:36:53 PM	2331.2	25.81	0.11	40.6	25.27
4	2021-05-10 4:36:54 PM	2440.2	25.83	0.12	40.2	25.28
5	2021-05-10 4:36:55 PM	2462	25.81	0.11	39.6	25.3

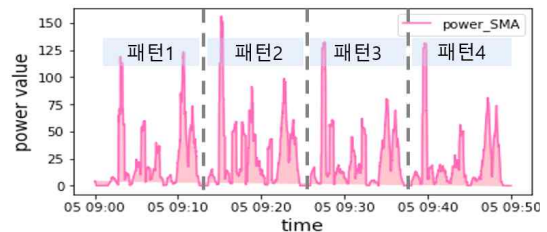
표 2. MCT 공정 데이터 예시

	일시	설비 상태	제품 코드	제품명	비고
1	2021-05-10 7:09 AM	end	360	A	자재 투입
2	2021-05-10 7:28 AM	start	360	A	-
3	2021-05-11 17:40 PM	end	360	A	-
4	2021-05-11 19:46 PM	start	365	B	-

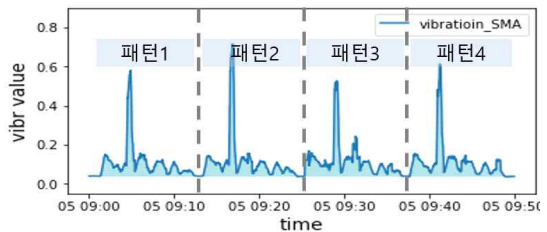
2. MCT 이상탐지 모델

MCT머신 이상탐지 알고리즘은 크게 전처리 단계, 학습 단계, 이상탐지 단계 순으로 진행된다.

전처리 단계는 단순이동평균(SMA: Simple Moving Average) 기법을 이용해서 데이터 값의 변동성을 줄이고, 선형 보간을 통해 시계열의 결측치를 제거한다. SMA의 사용 목적은 학습 모델이 공통된 속성을 찾거나 일반화에 용이하도록 하기 위함이며, 값의 변동성 및 무작위적인 변화량을 줄이고 시간 간격을 고정해주는 역할을 한다.



A. SMA 적용 후 전력데이터



B. SMA 적용 후 진동데이터

그림 1. SMA 처리한 전력 및 진동 데이터

그림 1은 실제 데이터에 단순이동평균(SMA)을 적용한 형태이다. 이 그래프에서 특정한 패턴이 반복되고 있음을 확인할 수 있다. 이러한 패턴을 추출하여 모델의 학습 데이터셋을 구성하기 위해, 유사 패턴들의 분포로부터 분포의 중심에 있는 대표 패턴을 생성하여 그 패턴과 실제 데이터를 대조해 보는 유사도 검사를 진행하였다.

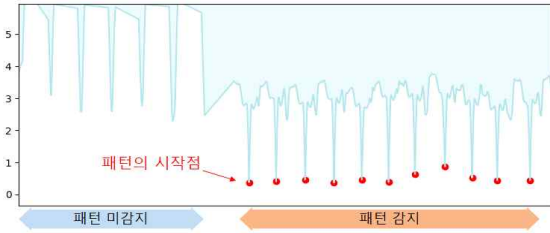


그림 2. 유사도 비교

그림 2와 같이, 대표 패턴과 실제 패턴과의 차이를 계산하여 낮은 값을 도출하면 유사하다고 판단한다. 일정한 시간 구간동안 가장 차이가 적은 부분이 패턴의 시작점이며, 그 시작점들로부터 720초의 공정시간을 추출하였다. 이 과정을 통해 정상만이 담긴 유사한 특정 패턴 데이터 집단을 획득하였다.

학습 단계에서는 CNN을 적용한 오토인코더에 전처리 과정에서 획득한 MCT 전력, 진동 데이터의 정상 패턴 집단을 사용하였다. 정상 패턴을 반복적으로 학습시키고, 출력이 입력과 비슷해질 수 있도록 손실값을 최소화하면서 복원을보다 이상 탐지율에 초점을 두었다. 학습에 사용한 오토인코더의 구조는 그림 3과 같다.

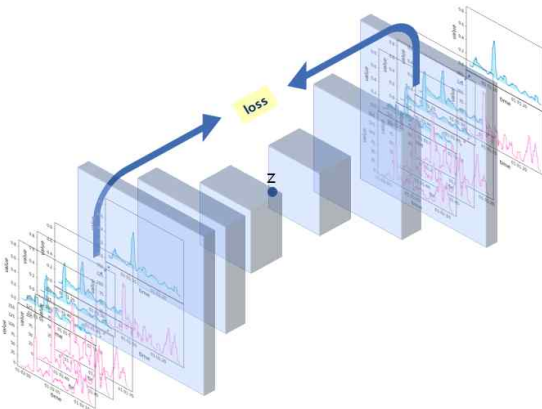


그림 3. 오토인코더의 구조

마지막으로 이상탐지 단계에서는 재구성 손실값을 확인하여 특정 임계값보다 크면 이상, 작으면 정상으로 판단한다. 이상탐지 단계에서 사용한 테스트 데이터는 월별 제조 데이터이다. 테스트 데이터는 720초 크기의 윈도우를 1초 단위로 슬라이딩하면서 학습된 모델의 입력으로 주어진다.

그림 4에서 차이가 비교적 낮은 구간이 정상 구간이고, 높이 치솟은 구간이 임계값보다 큰 이상 구간이다.



그림 4. 재구성 손실값

다음 그림은 제안된 알고리즘의 순서도이다.

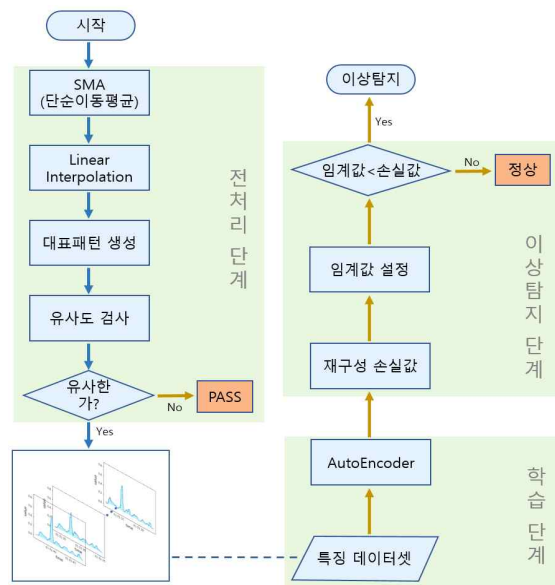


그림 5. 알고리즘 순서도

III. 성능 평가

성능 평가 지표로는 accuracy(정확도), recall(재현율), precision(정밀도), f1-score(조화평균)를 사용하였으며, 최종 성능의 판단은 f1-score로 확인했다.

단순이동평균(SMA)을 적용할 때, 설정하는 윈도우의 크기에 따라 성능의 차이가 표 3과 같이 나타났다. 윈도우의 크기를 60초로 설정하여 진행했을 때 91.4%의 가장 높은 성능을 보였고, 잠재 벡터에 따른 성능 차이를 표 4에 나타냈다. 표 4에서 잠재 벡터 차원은 30으로 설정했을 때 91.6%의 성능을 보였다.

표 3. SMA 처리의 윈도우 크기에 따른 성능 비교(잠재 벡터: 15)

SMA	accuracy	recall	precision	f1-score
15s	0.9895	0.6796	0.8799	0.7669
30s	0.9913	0.7702	0.8723	0.8180
60s	0.9955	0.9388	0.8912	0.9143
90s	0.9932	0.9642	0.8067	0.8784
120s	0.9185	0.9674	0.2332	0.3758

표 4. 잠재 벡터 차원에 따른 성능 비교(윈도우 크기: 60초)

z dim	accuracy	recall	precision	f1-score
5	0.9954	0.9335	0.8905	0.9115
10	0.9952	0.9369	0.8822	0.9087
15	0.9955	0.9388	0.8912	0.9143
30	0.9956	0.9417	0.8919	0.9161
60	0.9948	0.9441	0.8627	0.9015

표 5는 윈도우 크기 60초, 잠재 벡터 차원 30을 적용한 이상탐지 모델에 전체 기간에 대한 성능을 확인한 것이다. 기간A, B, C, D는 월별 데이터이며, 그 결과 약 93.53% 수준의 성능을 보였다.

표 5. 윈도우 크기 60초, 잠재 벡터 차원 30을 적용한 모델의 기간별 성능

Test	accuracy	recall	precision	f1-score
기간A	0.9868	0.9834	0.9068	0.9436
기간B	0.9956	0.9417	0.8919	0.9161
기간C	0.9852	0.9874	0.9188	0.9519
기간D	0.9382	0.9891	0.8455	0.9116
전체 기간	0.9840	0.9838	0.8913	0.9353

IV. 결론 및 향후 연구

본 논문에서는 MCT에서 수집한 제조 데이터와 대표적인 비지도 학습 방법 중의 하나인 오토인코더를 활용하여 MCT의 이상 상태를 감지하는 알고리즘을 개발하였다.

전처리 단계에서는 데이터의 변동성을 줄이는 단순이동평균(SMA) 기법을 사용해서 값을 변화시켰으며, 대표 패턴을 생성하여 실제 데이터와 유사도를 비교했다. 이를 통해 대표 패턴과 유사한 형태를 가진 정상 패턴 집단을 획득하고, 학습 단계에서 정상 패턴 집단을 사용해 CNN 기반의 오토인코더 딥러닝 모델을 학습시켰다. 최종적으로 MCT에 대해 f1-score 93.53%의 이상탐지 성능을 확인했다.

개발한 이상탐지 알고리즘은 실제 MCT에 적용해볼 예정이며, 계층의 수, 필터의 크기 등의 하이퍼파라미터 최적화를 통해 성능을 향상시킬 계획

이다. 개발한 모델은 실시간 이진분류만 가능하지만, 더 나아가 여러 제품 구분 및 예지 보전을 위한 딥러닝 모델로 발전시키고자 한다.

Acknowledgement

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 지역지능화혁신인재양성(Grand ICT연구센터) 사업의 연구결과로 수행되었음. (IITP-2021-2016-0-00318)

References

- [1] J. H. Lee, J. H. Kim, J. B. Hwang, and S. S. Kim, "A Study on Fault Detection of Cycle-based Signals using Wavelet Transform," in *Journal of the Korea Society for Simulation*, Vol. 16, No. 4, pp.13-22, 2007.
- [2] T. Lee, L. K. Lee, and C. Kim, "Performance of Machine Learning Algorithms for Class-Imbalanced Process Fault Detection Problems," in *IEEE Transactions on Semiconductor Manufacturing*, Vol. 29, No. 4, pp. 436-445, 2016.
- [3] S. Barua, M. Islam, X. Yao, and K. Murase, "MWMOTE-majority weighted minority oversampling technique for imbalanced data setlearning," in *IEEE Transaction on Knowledge and Data Engineering*, Vol. 26, No. 2, pp. 405-425, 2014.
- [4] I. Goodfellow, J. P. Abadoe, M. Mirza, B. Xu, D. W. Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," in *Communications of the ACM*, Vol. 63, No. 11, pp. 139-144, 2020.
- [5] S. Akcay, A. A. Abarghouei, T. Breckon, "GANomaly: Semi-supervised Anomaly Detection via Adversarial Training," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 11363 LNCS, pp. 622-637, 2019.
- [6] R. Chauhan, K. Ghanshala, and R. Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition," *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, pp 278-282, 2018.
- [7] K. Qian, Y. Zhang, S. Chang, X. Yang, and M. H. Johnson, "AUTOVC: Zero-Shot Voice Style Transfer with Only Autoencoder Loss," 2019.