

딥러닝 모델을 이용한 진동 데이터 기반 기계 이상 탐지 연구

박경욱, 최정훈, 서동준*

경북대학교

{pms8350, cjh34544, *dongjunsuh}@knu.ac.kr

A Study on the Detection of Mechanical Abnormalities Based on Vibration Data Using Deep Learning Model

Gyeonguk Park, Jeonghoon Choi, Dongjun Suh*

Kyungpook National University

요약

최근 원자재 가격 상승 및 노동비용 증가와 같은 경기 불황 속에서 다양한 제조 강국들이 기술적 경쟁력 강화를 위해 ICT 및 자동화 기술력 강화를 추구하고, 우리나라도 제조업 혁신과 스마트플랜트 정책 등 기술적 경쟁력을 강화하고 있다. 이에 따른 기술 발전으로 인한 산업 시스템 복잡성의 증가와 빅데이터, 사물인터넷, 인공지능 기술의 도입으로 사고 발생 가능성이 증가하고 있다. 이에 대응하기 위해 고장 예측과 시스템 이상 감지의 필요성이 대두되고 있다. 본 논문에서는 기계 진동 시스템에서 발생하는 이상을 탐지하는 딥러닝 기반 모델을 평가 방법중 하나인 ROC Curve에 따른 AUC 및 F1 Score, Accuracy를 통하여 비교 및 분석을 진행하였다. 실험결과 LSTM-Autoencoder 모델이 기계 이상탐지에 가장 적합함을 보였다.

I. 서론

최근에는 전 세계적으로 원자재 가격 상승 및 노동비용 증가와 같은 다양한 경기 불황 속에서도, 다양한 제조 강국들이 ICT 기술 융합 및 공장 자동화 등 기술적 경쟁력을 강화하고 있는 추세이다. 우리나라 또한 제조업 혁신과 스마트플랜트 관련 지원정책을 마련하여 이러한 흐름에 부합하고 있다 [1]. 특히, 빅데이터, 사물인터넷, 인공지능과 같은 첨단 기술들의 도입으로 데이터양과 다양성이 증가하면서 이를 신속하게 분석하고 처리하는 데에도 많은 도전이 따르고 있다 [2]. 이와 같은 기술 발전으로 인한 산업 시스템의 복잡성이 증가하면서 사고 발생의 가능성도 동시에 증가하고 있다.

산업 발전으로 인해 사고 확률 증가에 대한 이슈가 크게 부각 되면서, 고장 예측 및 시스템 이상 감지는 이러한 도전에 대응하기 위한 중요한 과제로 인식되고 있다 [3]. 특히, 기계 진동의 특성 패턴을 분석하여 고장 예측 시스템을 개발하는 것이 현재 제조업 분야에서 핵심적인 연구 주제로 부각 되고 있다. 이를 통해 설비기기의 운전 중 발생하는 진동에서 유용한 정보를 추출하고, 이상을 탐지하여 효율적인 유지보수 및 안전사고 예방에 기여할 수 있다.

따라서 본 논문에서는 기계 시스템의 이상을 탐지하기 위해 LSTM-Autoencoder 모델을 제안하고, Autoencoder, LSTM 모델과의 비교 분석을 진행하였다.

II. 본론

딥러닝 기반 기계 이상 탐지 분석 연구 수행을 하기 위해 사용된 데이터는 NSF I/UCR Center에서 제작한 NASA Bearing Dataset을 사용하였다 [4]. 데이터셋에는 엔진 내 여러 센서에서 4개의 베어링에서 측정하여 수집된 진동 및 소음 데이터를 포함하고 있고, 특정 간격에서 기록된 1초

진동 신호 스냅샷 데이터이다. 각 파일은 20,480개의 포인트로 구성되어 있으며 20kHz의 샘플링 속도로 설정되어 있다. 기계 시스템의 이상 탐지 모델의 성능향상을 위하여 데이터 정규화를 진행하였다. 또한 전체 데이터셋에 대하여 각각 70%, 30% 비율로 나누어 훈련 데이터셋과 테스트 데이터셋을 구성하였다.

본 논문에서 제안하는 딥러닝 기반 모델은 그림 1과 같다. 제안한 모델은 시계열 데이터 분석에 뛰어난 성능을 보이는 LSTM(Long Short-Term Memory) 신경망을 인코더와 디코더에 적용한 Autoencoder 모델이다. LSTM은 이전 시간 단계의 정보를 효과적으로 기억하고 활용하여 장기적인 의존성을 모델링 한다. 이를 가능하게 하는 주요 메커니즘 중 하나는 게이트 구조이며, 각 게이트는 입력, 삭제, 출력과 관련된 정보의 흐름을 조절하여 필요한 정보를 강조하고 불필요한 정보를 제어할 수 있다. 또한, LSTM은 그래디언트 소실 또는 폭발 문제에 대한 강건성을 가지고 있어, 긴 시퀀스 데이터를 처리할 때 안정적으로 작동하고 성능을 향상시킬 수 있다 [5]. Autoencoder는 입력 데이터에 적합한 내부 표현을 학습하고, 이를 활용하여 입력 데이터를 재구성한다. 이는 주로 데이터의 특징 추출과 차원 축소에 사용되며, 정상 데이터와 유사한 패턴을 가진 데이터는 작은 재구성 오차를 보이게 되지만, 이상 데이터와 같이 학습에 활용되지 않은 패턴이 포함된 데이터는 큰 복원 오차를 보이게 된다. 이러한 특성을 활용하여 시스템에서의 이상치 탐지에 적용하였다.

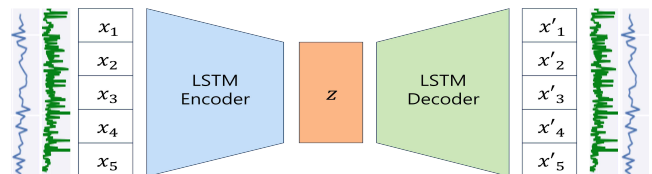


Figure 1 The Proposed Model

따라서, 본 논문에서는 LSTM과 Autoencoder를 효과적으로 결합한 LSTM-Autoencoder 모델을 제안하였다. LSTM의 장기 의존성 모델링 능력과 Autoencoder의 데이터 특징 추출 및 재구성 능력이 상호보완적으로 작용하여, 심층학습 기반의 이상 감지에 적합한 특징을 추출하여 학습 데이터에 없는 이상 패턴을 효과적으로 감지하고 탐지할 수 있다.

실험에서 사용한 손실 함수는 Autoencoder에서의 재구성 오차 (Reconstruction loss)에 대한 척도를 해석하기 용이하게 하는 MAE(Mean Absolute Error), 활성화 함수로 ReLU(Rectified Linear Unit)를 사용했다. 또한 그림 2에서는 딥러닝 모델의 성능 평가 지표인 ROC Curve와 AUC값을 나타낸다. ROC Curve 아래의 영역 넓이를 의미하는 AUC 값이 높을수록 이상 탐지 분류 성능이 좋은 모델을 나타낸다.

Table 1 F1 Score and Accuracy of Models

Model	Autoencoder	LSTM	LSTM Autoencoder
F1 score	0.8411	0.8724	0.9213
Accuracy	0.7237	0.8721	0.8996

실험 결과 그림 2와 표 1을 통해 제안된 LSTM-Autoencoder 모델이 기계 시스템 이상 탐지에서 다른 모델들보다 우수한 성능임을 확인하였다. AUC 값이 0.95, F1 Score 0.9213, Accuracy 0.8996 등의 평가 지표에서 높은 수치를 기록함으로써, 제안된 모델이 기계 이상 탐지 분류에 있어 가장 적합하다는 것을 확인할 수 있다.

III. 결론

본 논문에서는 기계 설비 중 베어링에서 측정된 NASA Bearing Dataset을 사용하여 기계 시스템에서의 이상 탐지를 위한 모델을 제안하고 기존 모델들과 비교 분석을 진행하였다. 분석 결과로 Autoencoder, LSTM 모델보다 제안한 LSTM-Autoencoder 모델이 우수한 성능을 보여주었다. 추후 연구에서는 산업 환경에서의 다양한 기계 진동 데이터셋에 제안된 모델을 적용할 계획이다. 더불어, 제안하는 모델을 실제 산업 환경에 적용할 수 있도록 성능 고도화 및 최적화 작업을 진행할 예정이다.

ACKNOWLEDGMENT

본 연구는 정부(과학기술정보통신부, 교육부)의 재원으로 한국연구재단의 지원을 받아 수행되었음 (과제번호: 2021R1A5A8033165, 2021R111A3049503)

참고 문헌

- [1] 고재진 외, “스마트공장을 위한 수직 통합패키지 개발”, KEIT PD 이슈리포트, 2021-5월호, 21(5), 2021.
- [2] 이은서, 배희철, 김현중, 한효녕, 이용귀. (2020). 미래 스마트 제조를 위한 인공지능 기술동향. 한국전자통신연구원.
- [3] 설문수. (2022). 중대산업사고 발생 원인과 통계 기반 중대산업사고 예방을 위한 시사점 연구. 한국위험물학회지, 10(2), 97-106, 10.31333/kihm.2022.10.2.97
- [4] NASA Bearing Dataset, <https://www.kaggle.com/datasets/vinayak123tyagi/bearing-dataset/data>
- [5] Sepp Hochreiter, Jürgen Schmidhuber; Long Short-Term Memory. Neural Comput 1997; 9 (8): 1735 - 1780.

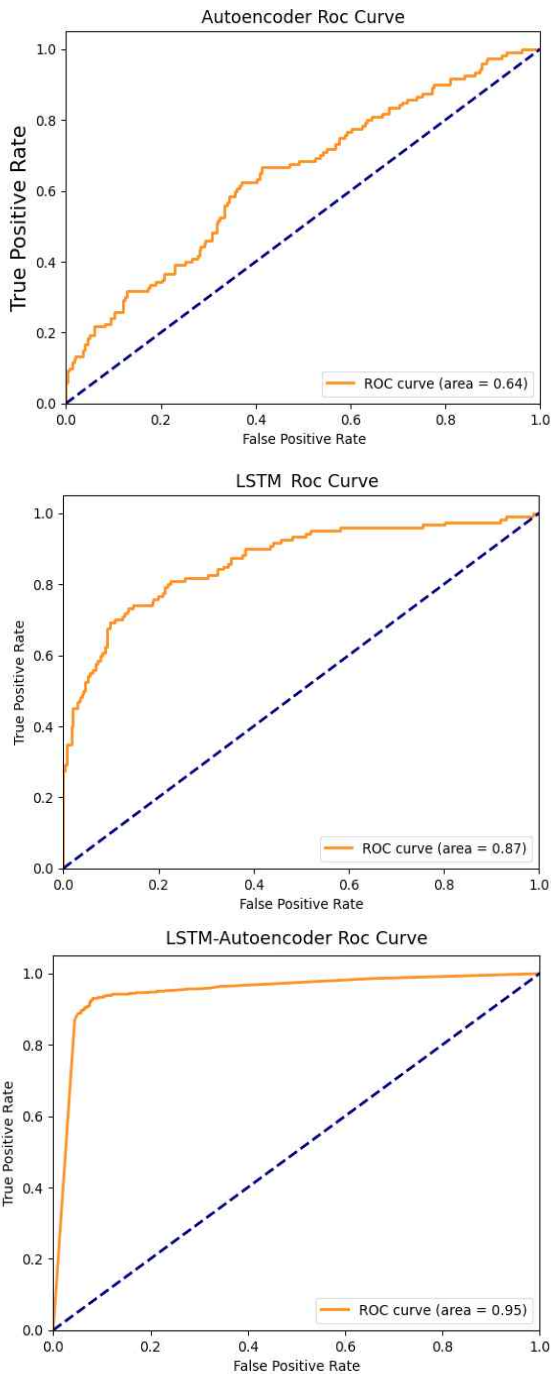


Figure 2 The Performance of Models