

시계열 특징을 갖는 선박용 공기 압축기 전류 데이터의 이상 탐지 알고리즘 적용 실험

이정형^{**}

* 목포해양대학교 기관시스템공학부 교수

Experimental Study on Application of an Anomaly Detection Algorithm in Electric Current Datasets Generated from Marine Air Compressor with Time-series Features

Jung-Hyung Lee^{**†}

* Professor, Division of Marine System Engineering, Mokpo Maritime University, Mokpo, Republic of Korea

요 약 : 본 연구는 선박용 공기압축기의 상태기반보전 시스템에 필요한 이상치 탐지 알고리즘 적용에 대한 실험적 연구로서 고장모사 실험을 통해 시계열 전류 센서 데이터를 이용한 이상탐지 적용 가능성을 확인하였다. 고장 유형 10개에 대해 실험실 규모의 고장 모사 실험을 수행하여 정상 운전데이터와 고장 데이터를 구축하였다. 실험 결과 구축된 이상탐지 모델은 시계열 데이터의 주기에 변화를 유발하는 이상은 잘 탐지하는 반면 미세한 부하 변동에 대한 탐지 성능은 떨어졌다. 또한 오토인코더를 이용한 시계열 이상탐지 모델은 입력 시퀀스의 길이와 초모수 조정에 따라 이상 탐지 성능이 상이한 것으로 나타났다.

핵심용어 : 공기압축기, 이상탐지, 상태기반보전, 오토인코더, 기계학습

Abstract : In this study, an anomaly detection (AD) algorithm was implemented to detect the failure of a marine air compressor. A lab-scale experiment was designed to produce fault datasets (time-series electric current measurements) for 10 failure modes of the air compressor. The results demonstrated that the temporal pattern of the datasets showed periodicity with a different period, depending on the failure mode. An AD model with a convolutional autoencoder was developed and trained based on a normal operation dataset. The reconstruction error was used as the threshold for AD. The reconstruction error was noted to be dependent on the AD model and hyperparameter tuning. The AD model was applied to the synthetic dataset, which comprised both normal and abnormal conditions of the air compressor for validation. The AD model exhibited good detection performance on anomalies showing periodicity but poor performance on anomalies resulting from subtle load changes in the motor.

Key Words : Air compressor, Anomaly detection, Condition-based maintenance (CBM), Autoencoder, Machine learning

1. 서 론

선박은 선체의 추진을 위한 주기관과 주기관의 원활한 운전을 위한 보조기계 그리고 운항 중 화물 관리를 위해 다양한 기계 설비가 탑재되어 있다. 이러한 장치와 설비는 내구연한이 있고 운전 중 기계구성품의 부식, 마모에 의해 노후화 과정을 거치게 되므로 성능 유지를 위한 보전을 필요로 한다. 선박 운항 관점에서 설비 보전(Maintenance)은 고장으로

로 인한 운항 지연이나 중단과 그로 인한 불필요한 자원 소모 방지에 일차적인 목적이 있겠으나, 궁극적으로는 고장 예방을 통한 선박의 감항능력(Seaworthiness) 유지에 보다 근본적인 목적이 있다. 2015년부터 2019년까지의 국내 사고유형별 해양사고 발생 현황에 따르면 취급이나 정비 부주의로 인한 기관 손상은 전체 사고의 약 32.1%로 가장 큰 비중을 차지하고 있다(Korean Maritime Safety Tribunal, 2019). 따라서 기관 고장 예방을 위해 수행되는 기존 정비전략의 재검토와 개선방안 도출이 필요한 시점이다.

고장예방과 성능유지를 위한 설비보전기법은 시간기반보

[†] jhlee@mmu.ac.kr, 061-240-7205

전(Time based maintenance, TBM)과 상태기반보전(Condition based maintenance, CBM)으로 구분할 수 있다. 시간기반보전은 설비를 구성하는 각 부품의 수명을 과거의 고장데이터의 분석을 통해 계산하여 정비시기를 설정하고 설정한 주기에 따라 정비를 수행하는 예방적 정비방식이다. 시간기반보전은 과잉정비로 인한 자원 낭비, 정상운전 중인 설비의 정비로 인한 고장, 돌발고장 발생 시 대응의 어려움 등의 단점이 있다. 이에 반해 상태기반보전은 상태 감시 기술에 근거하여 설비의 상태를 정확히 진단하므로 고장을 조기에 예측하고 이를 기반으로 정비시기를 결정하여 정비를 시행하는 방식이다.

전통적으로 선박은 기관 정비에 있어 계획정비제도(Planned maintenance system, PMS)라고 하는 시간기반보전 방식을 채택해 왔다. 최근 해운산업은 선원의 수급문제와 선박 운영 비용 문제 해결을 위해 자율운항선박 기술 도입을 고려하고 있다. 자율운항선박은 운항 자율화를 위해 기관실 각종 설비의 자동화와 유지보수의 지능화를 필수적으로 요구하고 있으며, 이에 따라 기관정비기법도 자연히 시간기반보전에서 상태기반보전방식으로의 이행이 필요하다.

선박 기관에 대한 상태기반보전이 효과적으로 수행되기 위해서는 기관의 상태정보를 취득할 수 있는 다양한 센서와 이로 부터 얻어진 데이터의 처리, 알고리즘 적용 과정을 통한 고장 진단 및 예측 시스템의 개발이 필요하다. 해운산업은 조선 기자재 시장의 경직된 조달 구조와 선주의 보수성으로 위험부담을 안고 신기술 기자재의 도입을 꺼리는 경향이 있으며, 이로 인해 상태기반보전기술의 개발과 도입이 타 산업분야에 비해 더딘 편이다. 그럼에도 불구하고 최근 선박기관에 대한 상태기반보전정비의 실증 기술 개발이 시도되고 있어 고무적이라 할 수 있다.

이와 관련된 연구로서 박재철 등은 선박 엔진의 상태감시 기반 고장진단 기술에 관한 연구를 수행하였으며 선박 주기관에 대한 결합수 분석과 실운항 데이터 분석을 기반으로 한 유지보수 방법과 시점을 도출하였다(Park et al., 2019). 배영목 등은 선박 주기관의 상태 데이터에 대해 다변량 관리도라는 통계적 기법을 이용하여 관리 상한선을 도출하여 유지보수 관련 의사결정을 지원하는 연구를 수행하였다(Bac et al., 2018).

또한 선박의 기계설비의 대부분을 차지하는 회전 기계장치에 대한 연구는 진동 기반 센서 기반으로 연구가 진행되어 오고 있으며, 최근에는 고장진단과 예측에 있어 기계학습 알고리즘 기반 연구가 진행되었다. 이준혁 등은 연구용으로 공개된 IMS 베어링 결함 데이터를 기반으로 기계학습 알고리즘을 구현하고 회전기계장치의 베어링 고장 예측 기술 개발에 관한 연구를 수행하였다(Lee et al., 2019). 박지현

등은 회전기기의 진동신호를 이용하여 데이터 주도의 고장 진단 모형에 관한 연구를 진행하였으며 운전상태를 분류를 위한 기계학습 알고리즘과 잔존수명 예측 모델을 제시하였다(Park et al., 2020).

인공지능 기술을 활용한 데이터 기반의 상태기반보전 시스템에서 고장 진단 기능은 핵심 요소이며, 고장 진단 알고리즘의 성능은 수집 데이터에 포함된 이상치(Anomaly)를 얼마나 잘 탐지하느냐에 있다. 상기 선행 연구의 대부분은 현장 데이터를 기반으로 하였다고는 하나 수집 데이터의 대부분은 기계장치의 정상 운전 데이터이다. 따라서 정상과 고장(이상) 데이터 간 불균형으로 인해 고장 진단 알고리즘의 성능 저하의 문제가 있으며 현장 적용을 위한 진단 성능 검증이 어려운 문제점이 있다. 또한 기계의 상태 진단에 있어 고가의 진동 센서와 그로부터 수집되는 빅데이터의 처리, 선박 및 엔진의 고유 진동으로 인한 간섭 신호 제거 등을 위해 복잡한 알고리즘의 적용이 필요하다는 한계점을 지니고 있다.

본 연구의 목적은 선박의 대표적인 보조기계 중 공기압축기의 데이터 기반 고장 진단 시스템 구현에 필요한 기초적 연구를 수행하는 것이다. 본 연구에서는 상기한 기존 연구의 문제점을 해결하기 위해 실험실 규모로 구축된 공기압축기를 이용해 운전 데이터를 확보하였으며, 이를 기반으로 시계열 데이터의 이상치 탐지 알고리즘의 일종인 오토인코더(Autoencoder)를 적용하여 이상치 탐지 모델을 구축하였다. 또한 공기압축기의 다양한 고장 유형을 구현한 고장 모사 실험을 통해 고장 데이터를 수집하였으며 이를 이용하여 이상 탐지 성능을 검증하였다.

2. 실험방법

2.1 실험대상

실험대상은 선박용 공기압축기이다. 이는 다양한 기계요소가 구비되어 있고 잦은 운전 정지로 인해 고장 발생 가능성이 높을 뿐만 아니라 정비 항목이 많아 고장 진단 실험 및 검증에 적합하다고 판단하였다. 선박에서 공기 압축기는 주기관의 시동 및 제어에 필요한 압축공기 공급에 근본적인 설치 목적이 있으며 각종 보조기기의 제어뿐만 아니라 화물 유 탱크의 불활성기체 제조를 위한 질소 발생기에도 활용되는 등 중요한 보조기계 중 하나이다. 실험에 사용된 공기압축기는 왕복동식으로서 압축기와 압축기를 구동하는 원동기(Motor), 압축 공기가 저장되는 저장탱크, 자동 충전을 위한 제어기로 구성된다. 본 연구에서는 실험실 규모에서 데이터 취득이 용이한 이동식 공기압축기를 사용하였으며 제원은 아래와 같다.

- 압축기형식: 왕복동식, 3.5 HP
- 동력전달방식: V벨트
- 탱크용량: 40 L
- 공기 생산량: 245 L/min
- 작동전압: 220 V 단상

2.2 정상 및 고장 데이터 셋 구축

접이안이나 화물의 적양하 상황을 제외한 통상적인 선박 운항 상황에서 공기압축기는 공기의 소모와 충전이 주기적으로 이루어지는 특징이 있다. 만약 공기압축기 계통에 이상이 발생하면 작동 주기의 변화가 발생하고 이는 시계열 데이터에 반영된다. 본 연구에서 취득한 데이터는 단변량 시계열(Univariate time-series) 특징을 갖는 공기압축기의 전류 데이터이다.

본 실험에서는 공기압축기의 정상 운전과 10종의 고장모드에 대해 전류데이터를 취득하였다. 전류 측정은 원동기의 전원 1상에 대하여 관통형 변류기(Current transformer)를 설치하여 데이터수집장치(Data Acquisition Unit, DAQ)에서 20 Hz의 주기로 수집 및 처리 후 PC에 저장하였다. 측정된 전류값은 공기압축기의 부하 전류로서 RMS 형태로 변환된 값이다.

고장 모드는 사전에 진행된 결함수 분석(Fault tree analysis)을 통해 도출하였으며 주요 고장 유형을 동력전달계통의 고장, 압축기 및 원동기 고장, 공기 누설로 구분하였다. 정상 및 고장 시나리오 모두 압축공기가 일정하게 소모되는 상황을 구현하기 위해 2분을 주기로 탱크의 압축공기가 자동으로 배출, 차단되도록 하였다. 이에 따라 탱크 압력이 감소되고 공기압축기는 자동으로 운전, 정지가 반복되었다.

2.3 데이터 전처리

기계학습 기반 이상탐지 모델 구축을 위해 사용된 정상 운전 데이터는 단일 특징값(Single feature)을 가졌으나 그 수가 방대하므로(244,274개), 훈련데이터로 활용되기 위해서는 전처리가 필요하다. 또한 원동기의 전류 데이터는 기동-정지에 반복에 따른 값의 편차가 크며 고장과 무관한 이상치가 수집되기 때문에 이들 데이터는 Z-score법을 이용해 제거하였다. 이후 전처리 과정은 먼저 원시 데이터(Raw data)는 1초 단위로 재샘플링(Resampling)한 뒤 최대-최소 정규화를 통해 값의 범위를 0~1이 되도록 하였다. 전처리 후 데이터 개수는 12,220개였다.

2.4 학습 및 이상탐지 알고리즘 적용

이상탐지를 위해 적용한 알고리즘은 시계열 데이터의 이

상치 탐지에서 우수한 성능을 나타내는 오토인코더를 적용하였다. 오토인코더는 신경망을 이용해 임의의 입력값과 같은 출력값으로 재생성하도록 하는 비지도 학습(Unsupervised learning) 기반의 알고리즘이다. 오토인코더는 Fig. 1과 같이 인코더와 디코더로 구성되어 있으며 입력층으로 입력된 다차원 데이터가 은닉층으로 전달되면서 저차원으로 축소되며 이는 다시 출력층으로 전달되면서 입력데이터와 동일 차원의 데이터로 복원된다. 은닉계층을 중심으로 인코더와 디코더는 대칭구조를 이루고 있으며, 입력계층의 노드 수와 출력계층의 노드 수는 항상 동일한 특징을 가지고 있다.

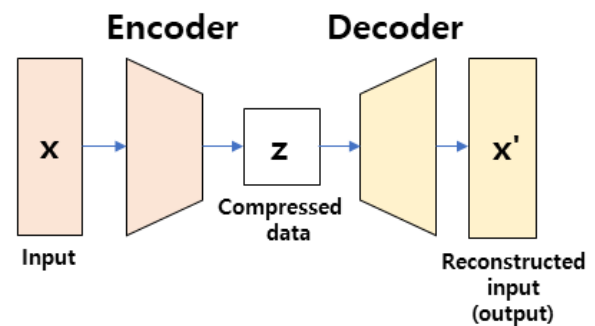


Fig. 1. Typical autoencoder model.

오토 인코더는 이를 구성하는 신경망 구조에 따라 적층형(Stacked), 변이형(Variational), 잡음제거형(Denoising) 등의 다양한 구조가 제안되고 있으나, 본 연구에서는 가장 단순한 형태의 합성곱 오토인코더(Convolutional autoencoder) 모델을 적용하였다. Fig. 2에 도시한 바와 같이 합성곱 기반 오토인코더는 인코더(Encoder) 부에 해당하는 2개의 1차원 합성곱층(Convolutional layer)과 디코더(Decoder) 부에 해당하는 2개의 합성곱 전치층(Convolution transpose)으로 구성되어 있으며 출력층을 제외한 모든 층에 적용된 활성화 함수는 Relu함수이다. 이와 같은 모델은 Keras에서 제공하는 Sequential 모델을 이용하여 구성하였다. 학습에 사용된 워크스테이션 사양과 소프트웨어 라이브러리 버전은 Table 1에 정리하였다.

전체 데이터 세트는 학습용 90%, 검증용 10%의 비율로 구분하였으며 학습을 통해 구축된 모델의 성능은 검증용 데이터 세트를 통해 검증되도록 하였다. 학습은 학습계수(Learning rate) 0.001, 배치크기(Batch size) 128로 설정하여 총 50번의 epoch 동안 이루어졌으며, 5회 연속 오차 값의 개선이 없을 경우 학습이 조기 종료되도록 하였다. 또한 Adam optimizer를 통해 최적화되도록 하였으며, 손실함수는 입력과 복원의 오차의 차에 대한 평균제곱오차(Mean Square Error, MSE)를 사용하였다.

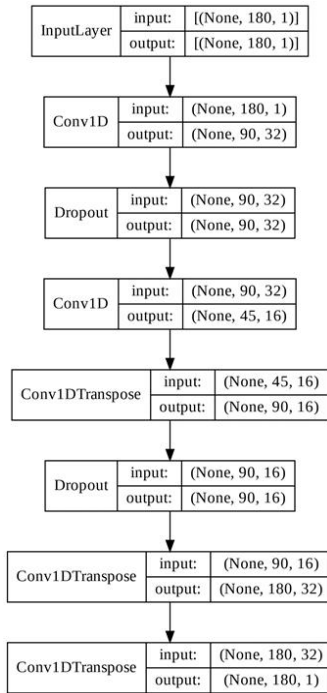


Fig. 2. Structure of a 1-D convolutional autoencoder.

Table 1. H/W and S/W specifications of workstation

Device/Library	Model/Version
CPU	Intel(R) Xeon W-2225 @4.6GHz
RAM	32G
GPU	NVIDIA GeForce RTX 2080 Ti 11GB (Cuda version: 10.1)
OS	Ubuntu 18.04
Python	3.7
Keras	2.1.5
Tensorflow	2.1.0

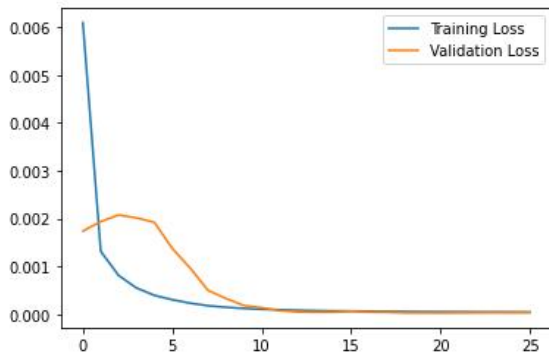


Fig. 3. Training and validation loss during training.

Fig. 3은 학습 과정 중 손실함수 값의 곡선을 나타낸 것이다. 훈련 손실(Training loss)은 훈련 초기에 최저치에 도달하고 있고, 검증 손실(Validation loss)은 8 epoch까지는 훈련 손실함수 값을 상회하여 차이를 보이고 있다. 그러나 학습이 진행됨에 따라 10 epoch 이후에는 두 손실함수 값은 0에 수렴하였다.

3. 실험 결과 및 고찰

3.1 고장모드 별 데이터 관찰

Fig. 4는 실험 대상 공기압축기의 정상 운전 중 1 사이클에 대한 시간에 따른 전류값의 변화를 나타낸 시계열 그래프이다. 여기서 1사이클은 압축공기가 사용처에서 소비되어 공기압축기의 기동-운전-자동정지에 이르는 기간으로 정의하였다. ①은 공기 압축기 기동 전 대기 상태이며 0에 가까운 값을 나타낸다. 압축공기가 소모됨에 따라 ② 시점에서 압축기의 압력스위치의 작동에 의해 자동 기동된다. 이때 전류값은 기동부하의 영향으로 피크치가 관찰된다. 기동 직후 전류치는 순간적으로 감소하여 정상 운전 시(③)에는 약 0.65~7 A를 나타내고 있으며 탱크 압력 상승에 따라 미세하게 우상향하는 모습을 보여주고 있다. 이후 공기압축기는 계속 운전되어 정해진 탱크 용량까지 보충된다. 탱크의 압력이 압력스위치의 설정치 상한값에 이르면 압축기는 정지하며 다음 기동까지 대기하게 된다(④).

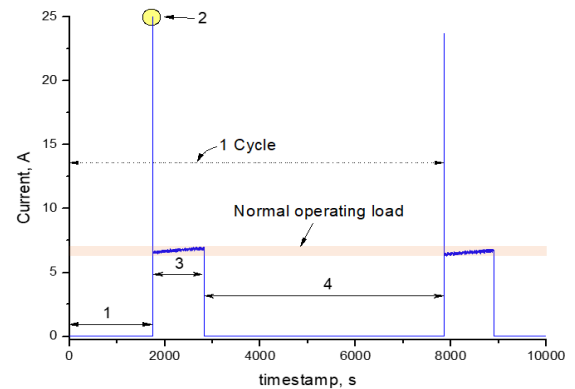


Fig. 4. Typical time-current curve of air compressor during normal operation.

공기압축기의 동력계통 이상과 공기 누설 등의 고장은 원동기의 부하에 반영되며, 이때 부하는 전류의 크기와 주기로 표현된다. Fig. 5는 10종의 고장 모드에 대한 고장 모사 실험으로부터 취득한 전류의 시계열 그래프를 나타낸 것이다. x축은 시간, y축은 전류(부하)의 크기를 나타내며 화살표

시계열 특징을 갖는 선박용 공기 압축기 전류 데이터의 이상 탐지 알고리즘 적용 실험

는 고장 모사 실험이 중단된 시점을 나타낸다. 고장모드 1, 2, 3, 6, 9는 고장에 의해 압축기(모터)가 계속 운전되는 조건으로 운전 중 전류값이 0을 나타낸 것은 고장 부여가 원인이 아니며 안전 상 실험을 강제로 중단한 것이다.

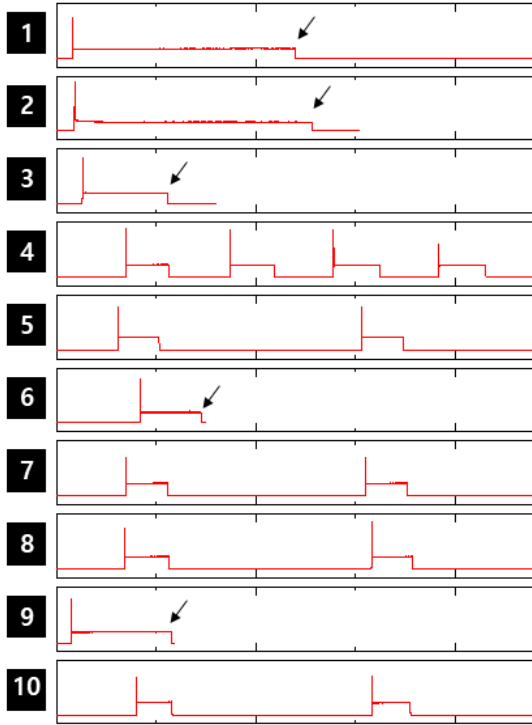


Fig. 5. Time-current curves showing temporal patterns with different fault modes of air compressor.

고장 종류 별 그래프에 대한 관찰 결과를 정리하면 아래와 같다.

1) V-belt 이탈: 기동과 동시에 운전 부하는 정상 상태에 비해 낮은 상태로 유지된다. 구동체(모터)와 피동체(압축기)를 연결하는 벨트가 이탈된 상태로 운전되어 압축기가 구동되지 않으므로 압축공기가 충전되지 않는다.

2) 압축공기의 대량 누설: 누설량이 충전량보다 많은 상태로서 압축기가 기동되어 운전되나 누설량이 과다하여 충전이 되지 않으며 압축기는 계속 운전된다(Long-run).

3) 압축공기의 미량 누설: 정상운전 상태와 운전 패턴은 유사하나 기동과 정지가 단 시간 내에 반복되는 이른바 쇼트 사이클링(short cycling)이 관찰된다.

4) 압축공기의 소량 누설: 고장모드 2에 비해 누설량이 많지는 않으나 충전량이 누설량에 미치지 못하여 운전이 계속되는 고장 형태이다. 고장모드 2에 비해 운전부하는 약간 높은 상태가 관찰된다.

5) 모터 측 베어링의 손상: 정상 운전상태와 유사한 운전 패턴을 보이고 있으나 운전 중 부하가 정상 운전보다 높게 나타난다.

6) V벨트의 장력 부족: 원동기의 동력이 피구동체인 압축기로 제대로 전달되지 못하므로 압축기의 운전이 계속되며, 운전 중 전류값에서 소폭의 요동이 관찰되는 것이 특징이다.

7) V벨트의 장력 과다: 운전 패턴은 정상운전과 차이가 없으나 운전 중 부하가 정상 운전에 비해 높게 나타나고 있다.

8) 압축기 측 베어링의 손상: 모터 측 베어링 손상 모드와 유사한 패턴을 보이고 있다.

9) 압력스위치의 고장: 공기압축기의 제어가 불가능한 상태로 탱크가 과충전되는 상황이다. 일정 압력 이상 충전 시 안전밸브가 작동하여 탱크의 공기를 분출하게 된다. 안전밸브의 개방에 의해 탱크의 압력이 일정값 이하로 떨어지면 안전밸브는 다시 닫히게 되고 충전이 진행된다. 압력이 안전밸브 작동 조건에 이르게 되면 다시 안전밸브는 작동되어 탱크의 압축공기는 대기 중으로 방출된다. 결과적으로 이러한 과정이 지속되므로 공기압축기는 계속 운전된다.

10) 공기압축기의 윤활유 부족: 윤활유를 제거한 상태에서 운전한 것으로 운전 패턴은 정상 운전상태와 유사하나 운전 중 부하는 정상에 비해 약간 큰 값을 나타내고 있다.

이상의 관찰을 통해 고장 모드 구현에 의한 운전 중 전류 신호의 패턴을 파악할 수 있었다. 고장 모드 별 전류 데이터는 전반적으로 정상 운전과 유사한 경향을 보이고 있지만 운전 주기와 운전 중 부하 크기에 차이를 보이고 있다. 따라서 이 같은 특징을 응용한 이상 탐지 모델의 구축이 가능함을 알 수 있다. 원동기 기동 시 발생하는 피크 값은 고장 모드 별로 패턴을 나타내지 않아 이상 탐지에 유용하지 않은 값으로 판단되었으며 전처리 과정에서 제거하였다.

본 연구에서 취득한 전류 센서값은 단일 특징값을 갖는 시계열 데이터로서 이상 탐지에 유효한 정보이지만, 고장 모드의 식별(분류)에는 유용하지 않은 정보임에 주목할 필요가 있다. 즉, 원동기에서 얻을 수 있는 다양한 센서 데이터 중 전류의 RMS값은 단독적으로 고장 패턴 분류에 적합하지 않은 특징이라 할 수 있다. 따라서 고장 패턴 분류를 위해서는 추가적인 특징값 선정이 필요할 것으로 판단된다.

3.2 이상 기준치 설정

상태기반보전시스템에서는 정비 시기 결정을 위해 이상값 인지를 위한 기준치 선정이 중요하다. 현재 선박의 AMS(Alarm monitoring system)에 적용된 공기압축기의 정보는 압축공기의 상한값, 하한값, 공기압축기 이상으로 단순하게 구성되어 있다. 따라서 이러한 정보를 직접적으로 상태기반보전에 적용하기에는 유용성이 떨어진다고 볼 수 있다. 또

한 정보 발생을 위한 상한, 하한의 임계치는 기관사의 경험이나 제조사에 의해 선정되고 있어 객관성이 결여되어 있고, 운전 당시의 상태는 전혀 고려되지 않는 실정이다. 이러한 임계치 설정에 있어 기존에는 정적(Static) 기준치를 적용하였으나 상태기반보전시스템에서는 상태에 기반의 동적(Dynamic) 기준치 설정이 필요하다. 본 연구에서는 공기압축기의 상태기반보전에 필요한 이상 판단의 기준치 설정을 위해 비지도학습 기반의 오토인코더를 적용하였다.

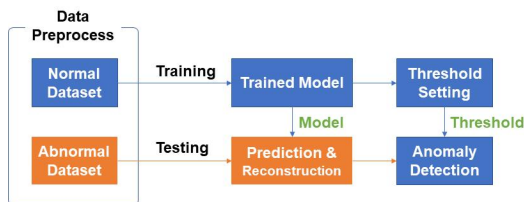


Fig. 6. Anomaly detection process using autoencoder.

오토인코더 기반의 이상탐지는 재구성 오차(Reconstruction error)의 계산과 임계치(Threshold) 설정에 의해 이상 여부를 판정하게 된다(Cook et al., 2019). 재구성 오차는 복원 오차라고도 하며 입력과 출력 사이의 오차이며 오토인코더의 학습은 이 재구성오차가 최소화되도록 진행된다. 재구성 오차는 이상치 점수(Anomaly score)로 작동하게 되며 임계치와의 비교를 통해 이상여부를 결정할 수 있다. 즉, 정상 데이터로 학습시킨 모델은 정상 데이터에 대해서는 낮은 재구성 오차를 나타내지만, 이상 데이터에 대해서는 높은 재구성 오차를 나타낸다. 따라서 정상 데이터를 대상으로 모델을 학습시키고 재구성 오차에 대한 임계치를 결정하게 되며 이를 기준으로 입력 데이터의 이상/정상 여부를 판정할 수 있는 것이다. Fig. 6은 상기한 오토인코더를 이용한 이상 탐지 과정을 정리하여 나타냈다.

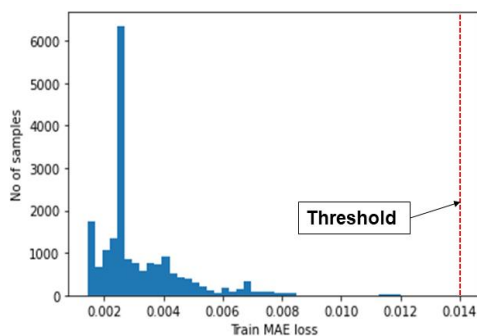


Fig. 7. Histogram of reconstruction error on normal training dataset.

Fig. 7은 합성곱 기반 오토인코더 모델의 학습 후 훈련 데이터에 대한 재구성 오차의 분포를 나타낸 것이다. 재구성

오차는 평균절대오차(Mean absolute error)를 이용하여 계산한 값으로 입출력 간 오차 분포는 그래프의 왼쪽에 치우치고 있음을 알 수 있다. 본 실험에서 이상 기준치 설정을 위한 임계치는 재구성 오차의 최대값(0.014)으로 설정하였다. 이렇게 얻어진 임계치는 이상 데이터에 학습 모델에 주입하여 얻은 재구성 오차에서 이상 여부를 판단하는 기준치로 작동하게 된다. Fig. 8은 180초 단위의 첫 번째 입력시퀀스에 대해 복원된 출력을 나타낸 그래프이다. 이는 공기 압축기의 정상 운전 중의 데이터를 포함한 것으로 입력(청색)에 대해 복원된 출력(오렌지색)이 형상적으로는 근접하게 복원되었으나 약간의 차이를 나타냄을 알 수 있다. 오토인코더의 출력이 완벽하게 입력을 복원하였다고 해서 반드시 이상탐지 성능이 좋은 것은 아니다. 이는 반대로 말하면 주어진 데이터의 유용한 특징을 학습한 것이 아니기 때문이다.

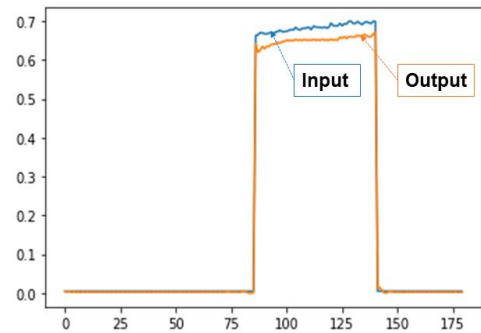


Fig. 8. Input sequential data vs reconstructed input (output) of first sample data.

3.3 이상탐지 가능성 검토

본 실험에서는 공기압축기의 정상 운전 데이터를 기반으로 오토인코더를 이용하여 입력 시퀀스의 길이(180초, 240초 - 각각 모델 A, B로 명명)를 변수로 하여 2개의 이상 탐지 모델을 구축하였다. 구축된 이상탐지 모델은 실제 운전 및 고장 데이터를 이용해 제작한 인공 데이터 셋(Synthetic dataset)을 통해 검증하였다. 검증 데이터의 시나리오에서는 정상 운전 중이던 공기압축기가 2020년 12월 11일 11시 05분경 고장모드 3(압축공기의 미량누설)에 의한 고장이 발생하여 쇼트사이클 운전을 한 후 11시 25분경 정상으로 복구되었다. 이후 공기압축기는 다시 정상 운전을 시작하였으며, 12시 25분 경 기동과 함께 V-벨트가 이탈된 상황을 가정하였다. 이와 같이 정상운전과 2종의 고장 신호가 포함된 데이터를 구축된 모델에 주입하여 이상탐지 가능성을 검토하였다.

Fig. 9는 검증시나리오에 대한 공기압축기 운전 전류의 시계열 데이터이며 a와 b는 각각 이상탐지모델 A와 B를 적용한 이상탐지 결과를 보여주고 있다. 음영으로 표시된 부분은

시계열 특징을 갖는 선박용 공기 압축기 전류 데이터의 이상 탐지 알고리즘 적용 실험

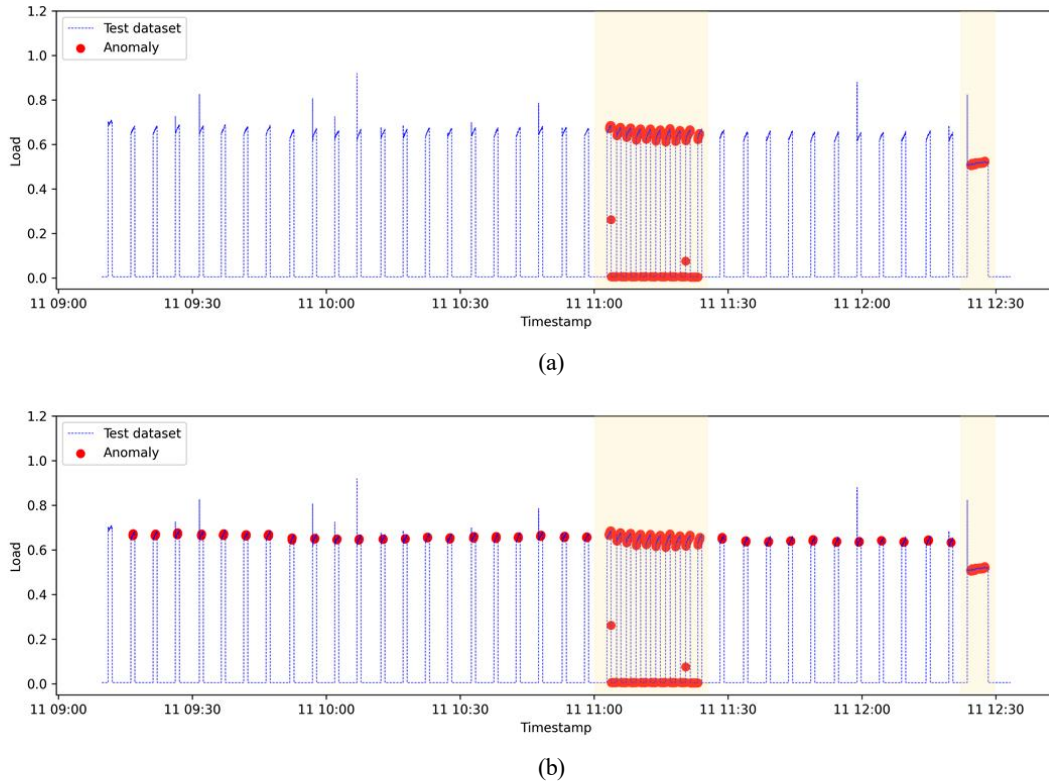


Fig. 9. Result of anomaly detection for synthetic dataset (normal + abnormal operation) of the air compressor: (a) anomaly detection model A; (b) anomaly detection model B.

실제 고장이 발생한 부분(True anomaly)이며, 이상탐지모델에 의해 이상이 탐지된 부분은 빨간 색 점으로 표시하였다.

모델 A는 2개의 실제 고장에 대해 이상탐지가 잘 되었고 판단되나, 모델 B는 모델 A와 같이 실제 고장에 대한 이상 탐지만만 아니라 정상 운전 데이터에 대해서도 이상이라고 탐지한 부정오류(False negative) 문제를 보여주고 있다. 다른 8개의 고장 모드에 대해서도 인공 데이터셋을 검증한 결과, 공기누설과 같은 주기에 변화가 유발되는 고장 모드는 이상 탐지가 잘 이루어지는 반면 부하의 미세한 변화가 관찰되는 베어링 고장 등의 유형은 탐지하지 못하였다.

본 연구에서는 초모수(Hyperparameter)의 최적화를 통한 모델의 성능 개선은 고려하지 않았다. 초모수는 오토인코더로 학습시킨 모델의 이상 탐지 성능에 영향을 미치는 요소이므로 향후 추가적인 연구가 필요할 것으로 판단된다(Nicolau and McDermott, 2018). 또한 비지도학습 기반의 이상탐지 모델 특성 상 정답 레이블이 없으므로 Accuracy, Recall, Precision, F1-score 등을 통한 성능 평가를 적용하기 어렵다는 한계가 있다. 따라서 이상탐지 모델의 성능 평가를 위해 데이터의 이상 정도를 측정할 수 있는 성능지표 개발이 필요할 것이다.

4. 결 론

본 연구는 선박용 공기압축기의 상태기반보전 시스템에 필요한 이상치 탐지 알고리즘을 적용한 실험적 연구로서 고장모사 실험을 통해 얻은 시계열 전류 센서 데이터에서 오토인코더 알고리즘을 활용한 이상탐지의 가능성을 확인하였다. 고장 유형 10종에 대해 실험실 규모의 고장 모사 실험을 수행하였으며 정상 운전데이터와 고장 데이터를 구축하였다. 정상 운전데이터를 이용하여 비지도기반의 오토인코더를 이용해 이상탐지모델을 구축하였으며 실제 고장 데이터를 이용하여 이상 탐지 가능성을 검증하였다. 실험 결과, 구축된 이상탐지모델은 시계열 데이터의 주기에 변화를 유발하는 고장은 잘 탐지하는 반면 미세한 부하 변동에 대한 탐지 성능은 떨어졌다. 오토인코더를 이용한 시계열 데이터의 이상탐지 모델은 입력 시퀀스의 길이와 초모수 조정에 따라 이상 탐지 성능이 상이하므로 향후 이의 최적화를 통한 성능 향상이 중요한 과제로 남는다.

감사의 글

본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 사회맞춤형 산학협력 선도대학(LINC+) 육성사업의 연구결과입니다.

References

- [1] Bae, Y. M., M. J. Kim, K. J. Kim, C. H. Jun, S. S. Byeon, and K. M. Park(2018), A Case Study on the Establishment of Upper Control Limit to Detect Vessel's Main Engine Failures using Multivariate Control Chart, Journal of the Society of Naval Architects of Korea, Vol. 55, No. 6, pp. 505-513.
- [2] Cook, A. A., G. Misirlı, and Z. Fan(2019), Anomaly detection for IoT time-series data: A survey, IEEE Internet of Things Journal, Vol. 7, No. 7, 6481-6494.
- [3] Korean Maritime Safety Tribunal(2019), Investigation report of Maritime Safety Judgement.
- [4] Lee, J. H., S. Y. Yoo, S. Shin, D. H. Kang, S. S. Lee, and J. C. Lee(2019), Fault diagnosis of bearings using machine learning algorithm, Journal of Advanced Marine Engineering and Technology, Vol. 43, No. 6, pp. 455-462.
- [5] Nicolau, M. and J. McDermott(2018), Learning neural representations for network anomaly detection, IEEE transactions on cybernetics, Vol. 49, No. 8, pp. 3074-3087.
- [6] Park, J. C., J. S. Jang, and Y. H. Jo(2019), A Study for the Development of Fault Diagnosis Technology Based on Condition Monitoring of Marine Engine, Proceedings of the Korean Institute of Navigation and Port Research Conference, pp. 230-231.
- [7] Park, J., M. Song, and S. Ahn(2020), Developing the Fault Diagnostics and Prognostics Model of a Rotating Machinery, Journal of the Korean Operations Research and Management Science Society, Vol. 45, No. 1, pp. 25-38.

Received : 2021. 02. 08.

Revised : 2021. 02. 23.

Accepted : 2021. 02. 25.