

논문 2024-19-08

# 음향 데이터를 이용한 CNN 추론 윈도우 기반 산업용 직교 좌표 로봇의 고장 진단 기법

## (Failure Detection Method of Industrial Cartesian Coordinate Robots Based on a CNN Inference Window Using Ambient Sound)

조 현 태\*  
(Hyuntae Cho)

**Abstract** : In the industrial field, robots are used to increase productivity by replacing labors with dangerous, difficult, and hard tasks. However, failures of individual industrial robots in the entire production process may cause product defects or malfunctions, and may cause dangerous disasters in the case of manufacturing parts used in automobiles and aircrafts. Although requirements for early diagnosis of industrial robot failures are steadily increasing, there are many limitations in early detection. This paper introduces methods for diagnosing robot failures using sound-based data and deep learning. This paper also analyzes, compares, and evaluates the performance of failure diagnosis using various deep learning technologies. Furthermore, in order to improve the performance of the fault diagnosis system using deep learning technology, we propose a method to increase the accuracy of fault diagnosis based on an inference window. When adopting the inference window of deep learning, the accuracy of the failure diagnosis was increased up to 94%.

**Keywords** : Cartesian robot, Failure detection, Industrial robots, Deep learning, Coordinate, Anomaly, Abnormal, Anomalous situation

### 1. 서 론

로봇은 다수의 산업현장에서 위험하고 어렵고 힘든 작업을 노동자를 대신하여 생산성을 높이는데 활용되고 있다. 한국은, 중국, 일본, 미국, 독일과 함께 로봇 산업에서 빅 5 시장에 속하며, 2020년 국내 시장 규모는 42억 달러 (약 5조 5천억원)에 이르렀으며, 산업용 로봇의 시장은 전체 시장의 약 52%에 해당하였다 [1]. 산업 현장 또는 생산 라인에서의 로봇은 단독으로만 사용되는 것이 아니라 군집을 이루어 동작 및 작업을 하고 있다. 한 대의 로봇 고장은 제품의 불량 또는 전체 시스템의 고장으로 이루어질 수 있기 때문에 로봇의 고장을 조기에 진단 또는 예측하는 것은 전체 시스템 또는 제품에 대한 신뢰성을 향상시키고, 나아가 고장으로 인한 비용/시간에 따른 손실을 줄이기 위해서는 필수적인 작업이다. 최근 인공지능 및 ICT 기술 발전으로 사물 인터넷으로 로봇의 동작 데이터를 수집하여, 클라우드에서 데이터를 활용하여 인공 지능으로 고장을 진단하는 예측 시스템들에 대한 연구/개발/서비스 등이 이루어지고 있다 [2-4].

산업용 직교 좌표 로봇은 볼스크류, 클린룸, 벨트 드라이브, 리니어 모터, 에어 실린더 등의 다양한 타입이 있고, 직교 좌표 로봇의 특성상 대부분 슬라이딩 레일 (sliding rail)

의 형태로 구성되어 사용된다. 본 논문에서는 직교좌표 로봇의 슬라이딩 레일 소리 신호를 인식하여 고장을 진단하는 시스템을 제안한다. 본 논문에서는 정상적인 로봇의 소리 데이터만을 사용하고, 합성곱 신경망 (CNN: convolutional neural network) 기반의 비지도 오토인코더 (autoencoder)를 적용하여 고장을 진단한다. 시스템은 가장 먼저 전처리 알고리즘을 통해 노이즈를 제거하고 소리 신호를 가공한다. 이 과정에서 출력되는 로그 멜-에너지 스펙트로그램 (Log Mel-Energy Spectrogram) 신호는 설계된 합성곱 신경망 기반의 모델로 입력되어 이상 유무를 판단하게 된다. 이렇게 판단된 추론 결과는 슬라이딩 윈도우 큐 (sliding window queue)에 삽입되어 이상 여부에 대해 추가적으로 검증을 실시하도록 하여 고장 진단 성능을 높이도록 설계되었다. 실제 딥러닝 추론 슬라이딩 윈도우 기반의 재검증 방법을 추가하였을 때는 94% 이상의 정확도로 슬라이딩 레일 기반의 직교로봇의 고장을 진단한다.

본 논문의 구성은 다음과 같다. 2장에서는 고장 진단을 위한 머신러닝 기술들에 대한 연구를 소개하고, 3장에서는 산업용 로봇으로 활용되는 로봇 중 직교로봇의 고장을 진단하기 위한 방법을 제안한다. 4장에서 성능평가를 실시한 후, 5장에서는 향후 과제와 함께 결론을 맺는다.

\*Corresponding Author (marine@tu.ac.kr)

Received: Sep. 30, 2023, Revised: Oct. 30, 2023, Accepted: Dec. 15, 2023.

H. Cho: Tongmyong University (Assist. Prof.)

※ 이 논문은 2022학년도 동명대학교 교내학술연구비 지원에 의하여 연구되었음 (과제번호 2022B010).

## II. 관련연구

로봇의 고장을 진단하기 위해 보편적으로 이상 상태나 비정상적인 동작을 감지하는 기술 (anomaly detection)이 활용된다. 이러한 이상 상태나 비정상적인 동작을 감지하는 기술은 로봇의 고장 진단 이외에도 네트워크 침입 감지, 질병 진단, 사기 범죄 조기 감지, 기계류 고장 감지 등에 활용된다. 이상 상태를 감지하기 위한 기술은 인공지능 모델에서 훈련되어지는 데이터의 활용 방법에 따라 분류되어 질 수 있다 [5-8].

지도학습 (supervised learning)은 입력 데이터와 출력 데이터의 쌍으로 이루어진 데이터셋을 가지고 학습을 수행하며, 충분한 양의 정상데이터와 이상 상태의 데이터를 가지고 학습을 하여야 한다. 분류 (classification)의 기법으로 볼 수 있으나 이상 상태에 대한 종류가 많아지게 되면 다중 분류로 수행되지만 높은 정확도로 분류되기 어려운 단점이 있다. 또한 모든 종류의 이상 상태 데이터를 수집하는 것에는 많은 한계점이 있다 [9].

전술한 바와 같이 대부분의 경우에 고장 상태에 대한 데이터는 수집이 어려워, 고장 데이터의 수는 전체 데이터셋에 비해 매우 적다. 비지도학습 (unsupervised learning)은 입력과 출력의 쌍으로 이루어진 데이터를 활용하여 학습하는 것이 아니라 정상적인 특징들 (features) 내에서 어떠한 패턴 또는 구조를 찾아냄으로써 학습을 수행한다. 즉, 정상적인 신호 데이터에 대한 특징들을 파악하고 그 특징들과 다른 특징들이 발견되면 이상 상태로 분류하게 된다. 현실적으로 이상 상태 데이터에 라벨링 (labeling)을 수행하는 것이 어렵기 때문에 이러한 비지도 학습 기반의 이상 상태 감지가 더 용이하게 활용된다 [10, 11].

반-지도 학습 (semi-supervised learning)은 라벨링되지 않은 대용량을 데이터와 라벨링 된 소량의 데이터를 사용하는 하이브리드 방법이다. 즉, 정상적인 데이터를 획득하기는 쉽지만 이상 상태나 비정상 데이터를 획득하기 어려운 고장 진단과 같은 응용에 쉽게 활용될 수 있다. 반-지도 학습의 특징은 라벨링 되지 않은 데이터 세트 내의 대부분의 데이터가 정상 데이터라는 가정하고, 이 데이터 세트에서 머신러닝 모델을 구현한 후에, 소량의 라벨링 된 데이터를 사용하여 성능을 평가하고 모델 최적화를 수행한다 [12-15].

이상 상태 검출에 대한 대부분의 방식은 먼저 정상 동작을 모델링한 다음 이 모델을 활용하여 편차 (이상)를 식별하게 하는 것이다. 즉, 정상 상태의 데이터를 기반으로 특정 임계치를 결정하고, 새로운 입력 데이터에서 임계치를 초과한 값이 검출될 경우 이상 상태 (anomaly)로 결정짓는다. 이상 상태 감지로 많이 사용되고 있는 오토인코더 (autoencoder)는 그림 1 (a) 와 같이 입력 데이터의 내부 표현을 생성하는 방법을 학습하는 인코더 (encoder)와 이 내부 표현을 기반으로 원래 입력을 재구성 (복원)하려고 하는 디코더 (decoder)로 구성되어 있다 [16-21]. 그림 1 (b)에서는 입력과 출력된 복원의 차이를 기반으로 계산된 정상적인

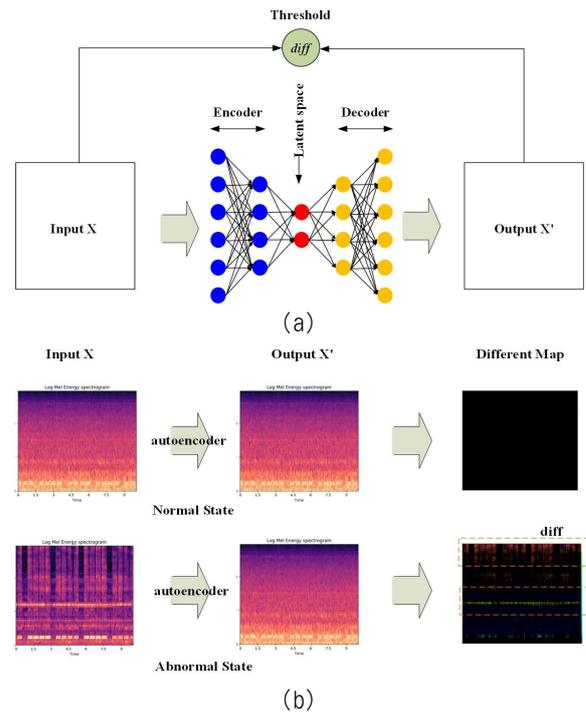


그림 1. 간단한 오토인코더 구조 및 예.  
(a) 오토인코더의 동작 개념, (b) 정상과 비정상 인식 예  
Fig. 1. Architecture and examples of a simple autoencoder.  
(a) conceptual overview of an autoencoder, (b) examples of normal and abnormal detection

데이터와 비정상적인 데이터의 재구성 에러의 맵을 보여준다. 이 맵을 기반으로 이상치 점수 (anomalous score)를 구하고, 임계치를 초과할 경우 이상 상태로 감지하는 것이다.

## III. CNN 추론 윈도우 기반의 고장 진단 시스템

본 절에서는 직교 좌표 로봇 시스템의 고장을 진단하기 위한 데이터 전처리, 딥러닝 모델, 및 딥러닝 추론 윈도우 적용 방법에 대해서 기술한다.

### 1. 데이터 전처리

이상 상태를 감지하기 위해 본 논문에서는 소리 데이터만을 사용하고, 전처리 과정을 통해 특성 추출 후에 딥러닝 모델의 입력으로 삽입한다. 그림 2는 소리 데이터를 Mel-Spectrogram으로 변경하는 과정을 보여준다. Mel-Spectrogram은 딥러닝을 이용하여 음성 인식, 음성 처리 등에서 많이 사용되는 소리 신호의 특징 추출 기법 중 하나이다. 고장을 감지하기 위해 본 논문에서 사용되는 소리 신호는 5개의 프레임으로 구성되고, 각각의 크기는 1024이며 각 프레임은 50% (hop length = 512)씩 중첩되게 구성된다. 즉, 하나의 프레임은 64ms 단위로 나뉘어 있으며 32ms단위로 중복되어 이어지게 된다. 이 단편화된 신호는 10초 길이로 연결되어 구성된다. 이 프레임들은 고속 푸리

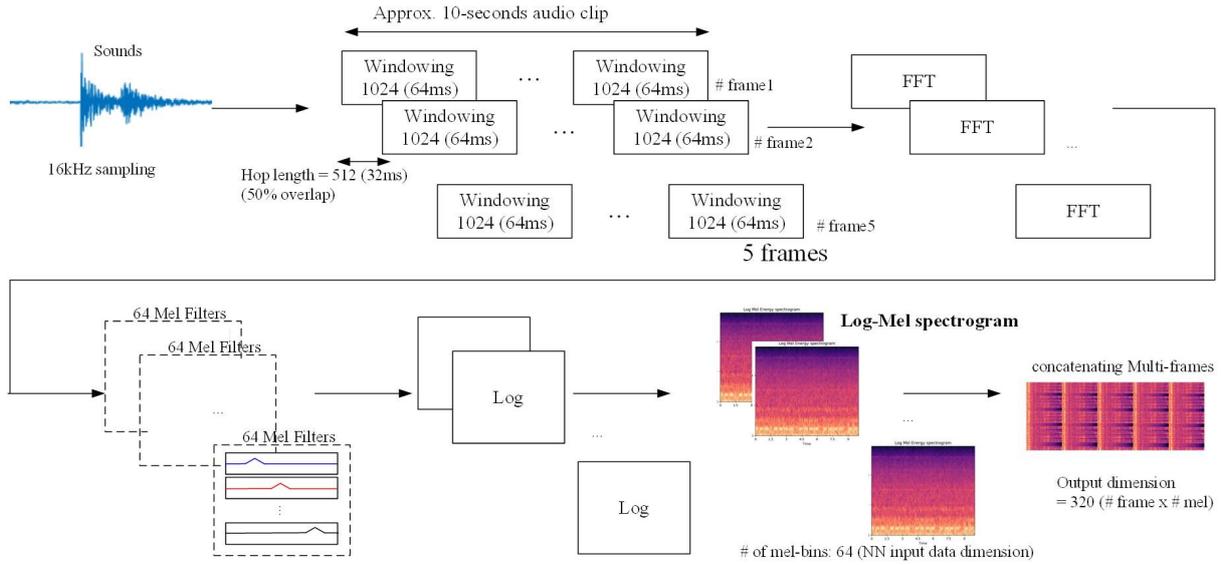


그림 2. 딥러닝 입력을 위한 로그 멜 스펙트로그램 추출  
Fig. 2. Extraction of Log Mel-Spectrogram for the input of deep learning

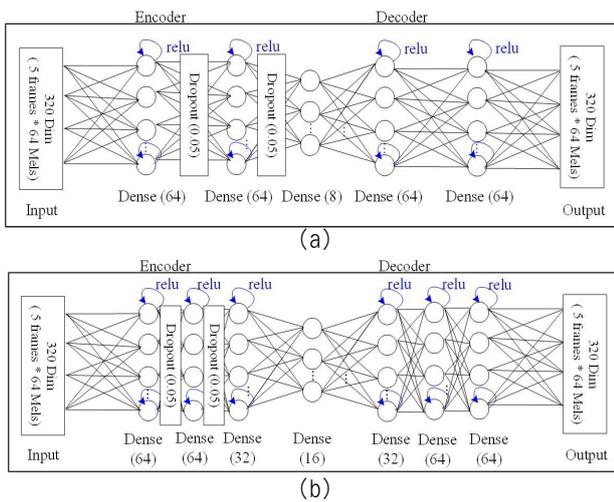


그림 3. 완전 연결층 기반의 오토인코더 모델:  
(a) MIMII 모델, (b) 확장된 MIMII 모델  
Fig. 3. Fully connected layer based autoencoder models:  
(a) MIMII model, (b) extended MIMII model

에 변환 (fast Fourier transform)과 64 Mel-필터 뱅크 과정을 거친 후, Y축을 로그 스케일로 변환 후 Log Mel-Spectrogram을 획득한다. 이렇게 획득된 5개의 Mel-Spectrogram은 단순히 직렬 연결되어 고장 진단을 위한 오토인코더의 입력으로 사용된다. 훈련과정과 추론과정에서 모두 이 소리 신호의 전처리 과정을 거치게 된다.

2. 딥러닝 모델 설계

본 절에서는 MIMII 데이터셋과 함께 제공되는 완전 연결층 (fully connected layers) 기반의 오토인코더 [22]와 함께 CNN 기반의 인코더 모델을 설계하고 이들에 대해 성능을 평가 및 분석하여 최적의 인코더를 도출한다. 오토인코더는

정상 데이터만을 가지고 학습을 진행하고, 테스트 단계에서는 정상 데이터와 비정상 데이터를 모두 활용하여 오토인코더 모델의 입력과 출력을 비교하여 재구성 에러 (reconstruction error)를 구한다. 설정된 임계치 이상일 경우 이상 상태 즉, 직교 좌표 로봇의 고장으로 판단하게 된다.

1.1 완전 연결층 기반 오토인코더

그림 3은 완전 연결층 기반의 오토인코더의 구조를 보여준다. 그림 3 (a) 모델은 히타치사에서 MIMII (malfunctioning industrial machine investigation and inspection) 데이터셋 [22-25] 과 함께 제공되는 기본 오토인코더의 구조이다. 입력과 출력의 차원은 320 차원이며, 인코더와 디코더는 각각 Dense 계층이 64-64 로 구성되어 있으며, 잠재 공간 (latent space)은 8개의 유닛으로 구성되어 있다. 각 레이어는 relu 활성화 함수를 사용한다.

그림 3 (b) 모델은 본 논문에서 기본 오토인코더 모델을 확장하여 인코더와 디코더를 3개의 층으로 구성하고 인코더 유닛의 개수를 64, 64, 32개로 구성하고 디코더는 32, 64, 64로 구성하였다. 그리고 잠재 공간에는 16개의 유닛을 가지도록 구성된다. (a) 모델과 마찬가지로 입력과 출력은 320개의 차원을 가지며 소리 신호의 데이터 전처리 한 결과가 딥러닝의 입력으로 삽입한다.

1.2 합성곱 신경망 (convolutional neural network) 기반 오토인코더

그림 4는 본 논문에서 설계 및 적용한 합성곱 기반의 신경망을 활용한 학습 모델을 보여준다. (a) 모델은 인코더에 합성곱 2개의 층과 디코더에 역합성곱 (de-convolution) 층이 2개 연결되어 구성된다. 인코더에서는 각 층에 각각 64개, 32개의 필터를 가진 합성곱 층이며, 디코더에는 각 층에 각각 32개, 64개의 필터를 가진 역합성곱 층이다. 그리고, 인코더에서는 3x2, 1x2 크기의 MaxPooling을 통해 합성곱층의 출력 크기를 줄이고, 디코더에서는 각층에 업샘플링 (upsampling)

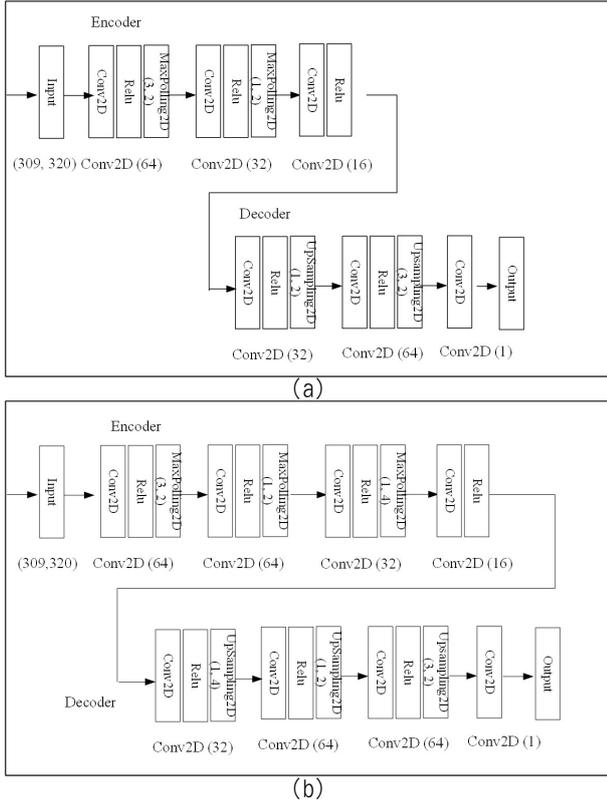


그림 4. CNN 기반의 오토인코더 모델:

(a) 기본 CNN 오토인코더, (b) 확장된 CNN 오토인코더

Fig. 4. CNN-based autoencoder models:

(a) simple CNN autoencoder, (b) extended CNN autoencoder

을 통해 역합성곱을 수행한다. 마지막 층에는 1개의 필터를 가지는 합성곱 층을 추가하여 출력 데이터를 생성한다 [26].

그림 4 (b) 모델은 그림 4 (a) 모델을 확장하여 인코더와 디코더가 3개의 층으로 구성되며, 각 층이 가지는 필터는 인코더는 64, 64, 32이고, 디코더는 32, 64, 64의 순으로 구성된다. 그림 4 (a) 모델과 마찬가지로 마지막 층에는 1개의 필터를 가지는 합성곱 층을 추가하여 출력 데이터를 생성한다.

이 오토인코더의 결과는 그림 1 (b)와 같이 입력 데이터와 출력데이터를 비교하여 재구성 에러를 수식 (1)을 이용하여 구한다.

$$ERR_{Reconst} = \frac{1}{M} \sum_{k=1}^M |x_k - \tilde{x}_k|^2. \quad (1)$$

이 수식에서 입력  $x_k$ 는 전처리후의 Mel-Spectrogram 입력 데이터를 의미하고  $\tilde{x}_k$ 는 딥러닝 결과로 나오는 출력 데이터를 의미한다.  $|\cdot|$ 는 벡터의 L1 놈 (norm)을 의미하고, M은 전처리후의 출력으로 나오는 개별 훈련데이터의 관측치 (observation) 개수를 나타낸다.

테스트 단계에서 재구성 에러가 이상 상태 결정 임계값을 초과하면 수식 (2)와 같이 이상 상태 또는 로봇의 고장으로 판단하게 된다.

$$\begin{cases} 1 & \text{if } ERR_{Reconst} \geq Decision\ Threshold \\ 0 & \text{otherwise} \end{cases}. \quad (2)$$

수식 (2)에서 이상 상태 결정을 위한 임계치는 수식 (3)과 같이 계산된다.

$$Decision\ Threshold = \frac{1}{N} \sum_{i=1}^N ERR_{NormalReconst}(i) + \sigma^2 + \alpha. \quad (3)$$

이 수식에서 입력  $ERR_{NormalReconst}(i)$ 는 정상 데이터와 수식 (1)을 통해 얻어진 값이며, N은 훈련데이터의 관측치 개수를 나타낸다.  $\sigma$ 는 표준편차를 의미하고,  $\alpha$ 는 실험적으로 획득된 최적의 상수값을 나타낸다.

### 3. 딥러닝 추론 윈도우 기반의 고장 진단 기법

딥러닝 추론, 즉 고장 진단에 사용되는 데이터는 그림 2와 같이 5개의 10초 길이의 프레임들을 32ms 중첩되게 하여 5개의 프레임을 사용한다. 사용되는 10초 간격의 소리 신호 길이는 로봇의 고장을 진단하는데 사용되는 데이터의 길이가 짧게 때문에 잘못된 결정을 하는 문제들이 발생할 수 있다. 이를 보완하기 위해서 본 논문에서는 한번의 추론으로 고장을 진단하는 것이 아니라 여러 번의 추론을 수행하고 이 결과들을 이용하여 다시 한번 검증을 수행하여 고장을 진단한다. 각 딥러닝 추론 결과는 정상 또는 이상 상태를 나타내고 이 정보는 일정 길이의 슬라이딩 윈도우 큐에 저장된다. 저장된 추론 결과들을 기반으로 이상 상태가 임계치 이상인 경우 고장으로 판단한다. 슬라이딩 윈도우에서 사용되는 임계치는 수식 (4)와 같이 계산된다. Anomaly Event Count는 윈도우내에서 고장이라고 판단한 횟수를 나타내고 Total Event Count는 전체 윈도우의 크기 즉, 실시한 추론 횟수를 의미한다.

$$Threshold = \frac{Anomaly\ Event\ Count}{Total\ Event\ Count}. \quad (4)$$

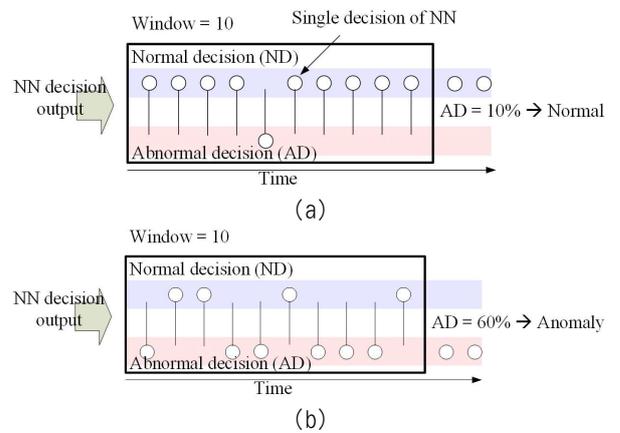


그림 5. 딥러닝 추론 윈도우 기반의 재검증 기법 예제:

(a) 정상 판단, (b) 이상 판단

Fig. 5. Example of a deep learning inference window-based re-verification method. (a) normal decision, (b) abnormal decision

그림 5는 딥러닝 추론 윈도우 큐를 이용하여 로봇의 상태가 정상인지 고장인지 판단하는 예제를 보여준다. 그림 5 (a)에서는 윈도우 크기가 10개이고, 고장 판별 임계치는 60%로 가정한다. 9개의 정상 추론과 1개의 이상 추론이 있기 때문에 10%에 해당하여 이 윈도우에서의 결과는 정상으로 판별된다. 그림 (b)에서는 6개의 이상 추론이 60%를 차지하고 있기 때문에 이 윈도우 큐 내에서는 고장 상태로 판별하게 된다.

#### IV. 성능평가

성능을 평가하기 위해 사용되는 데이터는 MIMII 데이터셋이다 [21-25]. 이 데이터셋에는 4개의 산업용 기계 데이터가 포함되어 있으나 본 논문에서는 직교좌표 로봇에 사용되는 슬라이딩 레일 (sliding rails)에 대한 소리 신호만을 추출하여 사용하였다. 슬라이딩 레일 사운드는 볼 스크류 타입과 벨트 타입으로 구성되어 있으며, 다양한 속도와 동작 모드, 벨트 재질, 주변환경, SNR 들에 대한 소리 데이터들을 포함하고 있다. 이 데이터들은 16kHz 샘플링 되었다. 데이터셋은 3개의 다른 소리 감도에서 녹음되고 4개의 다른 장치를 사용하였다. 소리에 대한 측정 환경 SNR은 0dB, -6dB, +6dB 이며, 장치의 ID는 00, 02, 04, 06으로 명명되어 사용된다. 훈련에는 69,420초의 정상 데이터가 사용되었으며, 성능을 평가하기 위해서 26,700초의 정상 데이터와 26,700초의 고장 데이터가 같은 비율로 사용되었다. 성능평가의 척도로 ROC (receiver-operating characteristic)의 아래 면적을 나타내는 수치로 분류 모델의 성능을 나타내는데 사용되는 AUC (area under the ROC curve)를 사용한다. AUC의 결과가 1에 가까울수록 완벽한 분류기로 판별하게 된다.

딥러닝 모델들의 성능 평가를 위해 사용되는 파라미터는 다음과 같다. 옵티마이저는 “adam”, batch size는 50, epoch는 50으로 설정되었다. 표 1은 환경 및 장치에 대해서 각 딥러닝 기법의 성능을 보여주며, 딥러닝 추론 윈도우는 적용되지 않았다. 표 1이 의미하는 바는 같은 딥러닝 기법을 사용하더라도 로봇의 소리가 녹음되는 환경과 녹음하는 장치에 따라서 성능이 달라질 수 있다는 것이다. 장치와 모델에 관계없이 SNR (signal to noise ratio)이 높은 수록 대체적으로 좋은 성능을 보여주고 있다. 하지만, 기본 CNN 인코더의 경우 장치 ID06은 SNR이 6dB인 상황보다 0dB에서 더 좋은 성능을 보여주었다. 녹음 장치에 따른 성능을 분석해보면 ID00의 경우 가장 훌륭한 성능을 보이고, 그 다음으로 ID04, ID02, ID06 순으로 높은 성능을 보여줄 수 있다. 즉, 산업용 로봇이 운용되는 환경에 의해서도 고장 진단의 성능이 영향을 받지만 로봇의 동작 소리를 녹음하는 장치에 의해서도 영향을 받음을 알 수 있다.

딥러닝 모델별 성능을 분석하기 위해 모델별로 평균을 하였을 경우 MIMII 데이터셋과 함께 제공되는 Baseline 오토 인코더의 경우 74%, 본 논문에서 확장한 Baseline 오토인코더의 경우 75%, 기본 합성곱 신경망의 경우 81.4%, 확장 합성곱 신경망의 경우 78.9%의 성능을 보여주었다.

표 1. 환경 및 장치에 따른 고장 진단 성능평가

Table 1. Performance evaluation of failure detection by environment and devices

#device	SNR	Baseline (MIMII)	Extended Baseline	Basic CNN	Extended CNN
ID00	-6dB	0.858	0.803	0.956	0.923
	0dB	0.893	0.785	0.993	0.965
	6dB	0.991	0.946	1	0.991
ID02	-6dB	0.599	0.622	0.638	0.735
	0dB	0.741	0.776	0.897	0.851
	6dB	0.939	0.947	0.897	0.926
ID04	-6dB	0.618	0.752	0.607	0.657
	0dB	0.786	0.799	0.929	0.764
	6dB	0.851	0.868	0.88	0.81
ID06	-6dB	0.498	0.485	0.564	0.554
	0dB	0.469	0.525	0.717	0.57
	6dB	0.627	0.692	0.586	0.722
Avg.		0.739	0.75	0.814	0.789

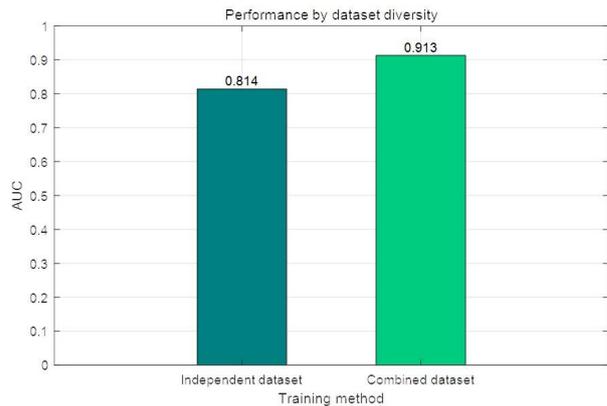


그림 6. CNN을 활용한 독립 데이터셋 훈련 및 통합 데이터셋 훈련의 성능 비교

Fig. 6. Performance comparison of independent dataset and combined dataset using CNN

CNN의 경우 좋은 성능을 내는 것으로 알려져 있지만 본 성능평가에서는 영상처리에서 사용될 때보다 낮은 성능을 나타냈다. 딥러닝이 좋은 성능을 내기 위해서는 훈련 데이터의 다양성이 중요하지만, 앞선 성능평가에선 SNR과 장치별로 구분하여 각각 훈련하고 성능을 평가하였다. 이를 보완하기 위해 2번째 성능평가에는 데이터셋을 모두 한번에 넣고 훈련을 시켜 성능을 평가하여 보았다. 평가에 사용된 딥러닝 모델은 기본 합성곱 신경망이며, 그림 6에서 보는 바와 같이 성능은 91.3%로 약 10%의 성능 향상이 됨을 알 수 있다.

그림 7은 CNN 모델에 따른 손실 그래프를 보여준다. 파란색은 훈련데이터에 대한 손실이며, 오렌지색은 테스트 데이터에 대한 손실을 의미한다. (a)는 기본 CNN의 예포크에 따른 손실 그래프이고 (b)는 확장 CNN 모델의 손실 그래프

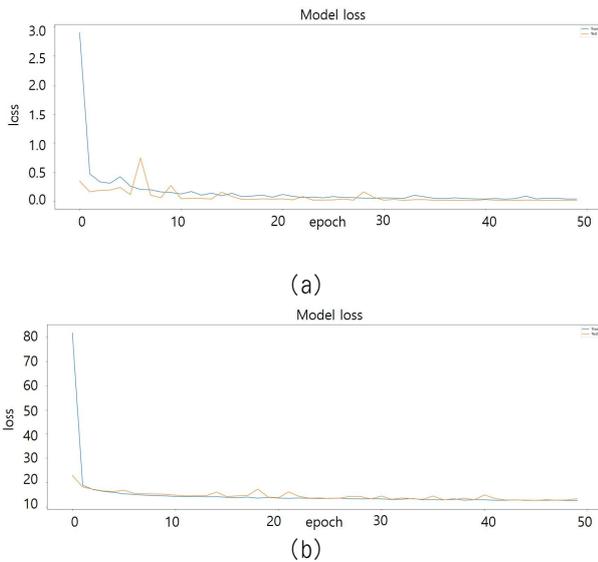


그림 7. CNN 모델에 따른 손실 그래프  
Fig. 7. Loss graphs by CNN models

표 2. 양자화 결과

Table 2. Quantization result

Quantization	AUC
TensorFlow Lite NNPACK	0.848624
Float32 model	0.848624
Float16 model	0.848520
Dynamic model	0.848687

이다. (a)의 경우 10번째 에포크에서 (b)의 경우는 3번째 에포크에서 손실이 더 이상 크게 감소하지 않음을 확인할 수 있다. 손실 감소가 지수적인 형태를 보이며, 구간에 따라 빠른 학습을 수행함을 의미한다.

본 논문에서는 임베디드 시스템 또는 모바일 폰에서 사용하기 위해 제안된 모델을 양자화를 수행하여 성능을 비교하였다. 표2는 양자화 결과를 보여준다. 양자화 수행 후, 성능은 약 6%가 하락하는 것을 확인할 수 있으나, 소요하는 메모리와 파일 사이즈는 2~4배까지 감소하여 경량 시스템에 사용할 경우 참고지표로 사용할 수 있다.

마지막으로 성능을 보다 높이기 위하여 슬라이딩 윈도우 기반의 고장 기법을 적용하였다. 그림 8은 딥러닝 추론 윈도우 기반의 고장 진단에 대한 성능 평가를 보여준다. 평가에 사용된 딥러닝 모델은 기본 CNN (모델3)이며, 30개의 윈도우 크기를 정하였고 고장 진단 기준을 위해 18개의 임계치를 사용하여 딥러닝 추정 결과의 60% 이상이 이상 상태라고 판단하면 직교로봇의 고장으로 최종 결정한다. 성능 평가 결과 고장 판단의 정확도는 약 94%로 독립적인 데이터셋만 사용했을 때보다 약 13%의 성능 향상, 통합 데이터셋을 사용했을 때보다는 약 3%의 성능 향상을 보여주었다. 이 성능의 결과는 윈도우의 크기와 이상상태를 판단하기 위한 이상 상태 추론 결과의 비율에 따라 달라질 수 있다.

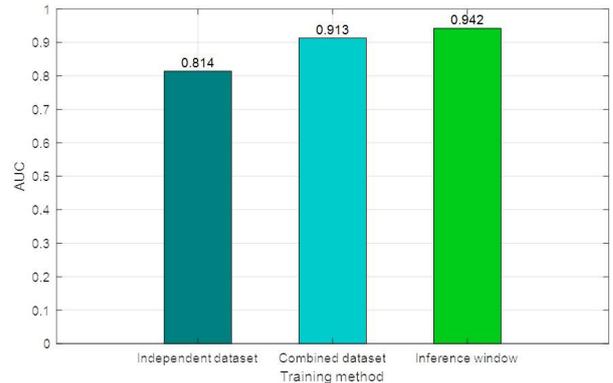


그림 8. 딥러닝 추론 윈도우 기반의 고장 진단 성능 평가  
Fig. 8. Performance evaluation of the deep learning inference window-based failure detection method

표 3. 기존 기술 [27] 과의 성능평가

Table 3. Performance comparison with previous works

Method	Accuracy (%)
Baseline (AE)	67.09
Baseline (MobileNetV2)	65.43
AE (sequence-level autoencoder)	86.01
BC (binary classification)	80.64
AE ensemble	86.12
BC ensemble	89.28
AE+BC ensemble (mix)	89.06
CNN AE	81.4
CNN AE (combined data)	91.3
CNN AE + Inference Window	94.2

또한, 본 논문에서는 기존의 연구에서 제시한 성능평가 지표와의 비교 분석을 수행하였다. [26]에서는 이상 상태 감지 성능을 분석하기 위해 다양한 음향 데이터와 알고리즘을 적용하고 결과를 분석하였다. 본 논문에서는 그 중 직교로봇과 관련된 슬라이딩 레일과 관련된 음향 데이터에 대한 결과만을 비교 분석한다. 표 3은 [27]에서 제공하는 성능 지표와 본 논문에서 제안한 기법의 성능을 비교한 표이다. [26]에서 분석한 결과는 이진 분류 앙상블 (BC ensemble) 시스템이 89.28%로 가장 우수한 성능을 보여주었고, 그 다음은 시퀀스 레벨 오토인코더와 이진 분류 앙상블을 함께 사용하였을 경우가 89.06%로 비슷한 결과를 보여주었다.

본 논문에서 제안한 기법의 경우 94.2%의 정확도로 시퀀스 레벨 오토인코더와 이진 분류 기법의 성능보다 더 우수함을 알 수 있다. 다만, 두 성능 평가의 차이점은 본 논문에서는 모든 MIMII 데이터를 한번에 학습하여 모델을 훈련시킨 것이고, [26]에서는 장치별, SNR을 달리하여 학습한 후 평균한 결과를 보여주는 것이다. 향후, [26]에서 제안하는 시퀀스 레벨 오토인코더와 이진 분류 앙상블 기법에 본 논문에서 제시한 추론 윈도우 기법을 사용하면 더욱 훌륭한

성능이 나타날 것으로 예상된다.

## V. 결론

본 논문에서는 산업용 직교로봇의 고장을 진단하기 위해 정상 데이터만을 사용하여 학습시킨 후, 고장을 진단하는 비지도학습인 오토인코더 기술에 대하여 소개하고, 상대적으로 낮은 성능을 보완하기 위한 접근법들을 소개하였다. 먼저, 이상 상태를 감지하기 위해 사용되는 정상 데이터의 다양성이 고장 진단의 성능에 밀접하게 관계되어 있다는 것을 보였으며, 신뢰성 있는 성능을 보장하기 위해 딥러닝 추론 윈도우 기반의 고장 진단 기법을 제안하였다. 제안된 기법을 산업용 직교 로봇의 고장 진단에 사용하기 위해 볼 스크류 타입과 벨트 타입으로 구성된 다양한 데이터를 사용하여 성능 평가를 실시하였다. 제안된 기법은 산업용 직교로봇의 고장 진단이외에도 다양한 소리 기반의 이상 상태를 감지하는데 활용될 수 있을 것으로 사료된다.

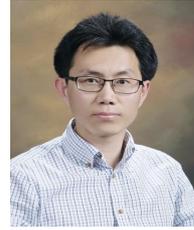
향후 과제로는 실제 산업용 직교 로봇이 운영되는 환경에서 훈련을 시키고, 고장을 진단하는 실증에 대한 연구와 개발된 딥러닝 모델을 컴퓨팅 파워가 상대적으로 약한 직교로봇에 직접 이식하는 연구를 포함하고 있다.

## References

- [1] Korea Institute for Robot Industry Advancement, "Survey on Robot Industry 2021", 2021. (in Korean)
- [2] KITECH, "AI-based Failure Diagnostic Device and Method," Korea patent No. 10-2289212-0000, 2021. 08. 06. (in Korean)
- [3] D. T. Hoang, H. J. Kang, "A Survey on Deep Learning Based Bearing Fault Diagnosis," *Neurocomputing*, Vol. 335, No. 28, pp. 327-335, 2019. (in Korean)
- [4] H. Cho, H. Choi, H. Kim, K. Nam, J. Ryu, K. Ha, "Feature Selection for Unmanned Surface Vehicle Fault Diagnosis Research and Experimental Verification," *Journal of Institute of Control, Robotics and Systems*, Vol. 28, No. 6, 542-550.
- [5] R. Chalapathy, S. Chawla, "Deep Learning for Anomaly Detection: A Survey," arXiv preprint arXiv:1901.03407, 2019.
- [6] V. Chandola, A. Banerjee, V. Kumar, "Anomaly Detection, A Survey," *ACM Computing Surveys*, Vol. 41, Iss. 3, No. 15, pp. 1-58, 2009.
- [7] F. Mattia, P. Galeone, M. Simoni, E. Ghelfi, "A Survey on Gans for Anomaly Detection," arXiv preprint arXiv:1906.11632.
- [8] V. Wilmet, S. Verma, T. Redl, H. Sandaker, Z. Li, "A Comparison of Supervised and Unsupervised Deep Learning Methods for Anomaly Detection in Images," arXiv preprint arXiv:2107.09204.
- [9] N. Goernitz, M. M. Kloft, K. Rieck, U. Brefeld, "Toward Supervised Anomaly Detection," *Journal Of Artificial Intelligence Research*, Vol. 46, pp. 235-262, 2013.
- [10] J. Yoon, K. Sohn, C. L. Li, S. Arik, C. Y. Lee, T. Pfister, "Self-supervise, Refine, Repeat: Improving Unsupervised Anomaly Detection," *Transactions on Machine Learning Research*, arXiv:2106.06115, 2022.
- [11] B. Zong, Q. Song, M. R. Min, W. Cheng, C. Lumezanu, D. Cho, H. Chen, "Deep Autoencoding Gaussian Mixture Model for Unsupervised Anomaly Detection," *International Conference on Learning Representations*, 2018.
- [12] L. Ruff, R. A. Vandermeulen, N. Görnitz, A. Binder, E. Müller, K.-R. Müller, M. Kloft, "Deep Semi-Supervised Anomaly Detection," arXiv:1906.02694, 2019.
- [13] S. Kcay, A. Atapour-Abarghouei, T. P. Breckon, "GANomaly: Semi-supervised Anomaly Detection via Adversarial Training," *Computer Vision - ACCV 2018, Lecture Notes in Computer Science*, Vol 11363. Springer, Cham., 2018.
- [14] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, X. Chen, "Improved Techniques for Training GANs," *International Conference on Neural Information Processing Systems*, (NeurIPS2016), <https://arxiv.org/abs/1606.03498>, 2016.
- [15] D. Li, D. Chen, L. Shi, B. Jin, J. Goh, S. K. Ng, "MAD-GAN: Multivariate Anomaly Detection for Time Series Data with Generative Adversarial Networks," arXiv:1901.04997, 2019.
- [16] C. Zhou, R. C. Paffenroth, "Anomaly Detection with Robust Deep Autoencoders," In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 665-674, 2017.
- [17] M. Sakurada, T. Yairi, "Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction," *International Workshop on Machine Learning for Sensory Data Analysis (MLSDA)*, pp. 4-11, 2014.
- [18] Z. Chen, C. K. Yeo, B. S. Lee, C. T. Lau, "Autoencoder-based Network Anomaly Detection," In *2018 Wireless Telecommunications Symposium (WTS)*, pp. 1-5, 2018.
- [19] A. A. Pol, V. Berger, C. Germain, G. Cerminara, M. Pierini, "Anomaly Detection with Conditional Variational Autoencoders," *18th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 1651-1657, 2019.
- [20] O. I. Provotar, Y. M. Linder, M. M. Veres, "Unsupervised Anomaly Detection in Time Series Using LSTM-Based Autoencoders," In *2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT)*, pp. 513-517, 2019.
- [21] S. Yan, H. Shao, Y. Xiao, B. Liu, J. Wan, "Hybrid Robust Convolutional Autoencoder for Unsupervised Anomaly Detection of Machine Tools Under Noises," *Robotics and Computer-Integrated Manufacturing*, Vol. 79, pp. 102441, 2013.
- [22] H. Purohit, R. Tanabe, K. Ichige, T. Endo, Y. Nikaido, K. Suefusa, Y. Kawaguchi, "MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection," arXiv preprint arXiv:1909.09347, 2019.
- [23] H. Purohit, R. Tanabe, K. Ichige, T. Endo, Y. Nikaido, K. Suefusa, Y. Kawaguchi, "MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection," in *Proc. 4th Workshop on Detection and Classification of Acoustic Scenes and Events (DCASE)*, 2019.

- [24] Y. Kawaguchi, K. Imoto, Y. Koizumi, N. Harada, D. Niizumi, K. Dohi, R. Tanabe, H. Purohit, T. Endo, "Description and Discussion on DCASE 2021 Challenge Task 2: Unsupervised Anomalous Sound Detection for Machine Condition Monitoring under Domain Shifted Conditions," in arXiv e-prints: 2106.04492, 2021.
- [25] R. Tanabe, H. Purohit, K. Dohi, T. Endo, Y. Nikaido, T. Nakamura, Y. Kawaguchi, "MIMII DUE: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection with Domain Shifts due to Changes in Operational and Environmental Conditions," 2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA). IEEE, pp. 21-25, 2021.
- [26] H. Cho, "CNN-based Autoencoder and Post-Training Quantization for On-Device Anomaly Detection of Cartesian Coordinate Robots," IEEE Computing and Communication Workshop and Conference (CCWC), 2024.
- [27] I. Kuroyanagi, T. Hayashi, Y. Adachi, T. Yoshimura, K. Takeda, T. Toda, "Anomalous Sound Detection With Ensemble of Autoencoder and Binary Classification Approaches," Technical Report, Detection and Classification of Acoustic Scenes and Events, 2021.

### Hyuntae Cho (조현태)



2003 Computer and Information Engineering from Korea Maritime and Ocean University (B.S.)

2005 Computer Engineering from Pusan National University (M.S.)

2011 Computer Engineering from Pusan National University (Ph.D.)

2021~2023 School of Digital Media Engineering, Tongmyong University (Assist. Prof.)

2023~Dept. of Digital Content, Tongmyong University (Assist. Prof.)

Career:

2018~2021 CEO & Founder of Soricare Inc.

2012~2021 Research Professor at Center for Integrated Smart Sensors, KAIST

2016~2021 CTO & Co-founder of Luftsens Inc.

2011~2012 Research Professor at Pusan National University

Field of Interests: Embedded system, A.I. for embedded systems, In vitro diagnostic systems, IoT, Smart Sensors

Email: marine@tu.ac.kr