

이산 웨이블릿 변환 기반의 대역 에너지를 활용한 음향 신호의 이상 탐지

최은혜*, 박형곤^o

Acoustic Anomaly Detection Based on Band Energy of Discrete Wavelet Transform

Eunhye Choi*, Hyunggon Park^o

요약

산업 현장에서 기계 설비의 이상 상태를 감지하고 모니터링하는 것은 중요하다. 음향 신호 기반의 설비 고장 감지 기술은 기존의 이미지/영상 기반이나 기타 별도 센서 기반의 이상 감지 기술에 비해 저비용으로 고장 감지를 수행할 수 있다는 장점이 있다. 하지만 센서로부터 수집된 대량의 음향 신호를 활용하기 위해서는 집약적인 신호 처리와 학습을 위해 많은 컴퓨팅 자원을 필요로 한다. 따라서 컴퓨팅 자원을 효율적으로 사용한 음향 신호 기반 이상 탐지 시스템에 대한 고려가 필요하다. 본 논문에서는 이산 웨이블릿 변환을 활용하여 음향 신호의 대역 에너지를 추출하는 방법과 이를 활용한 경량 이상 탐지 모델로 구성된 시스템을 제안한다. 이산 웨이블릿 변환 기반의 대역 에너지는 음향 신호의 특징을 효과적으로 압축하여 데이터 전처리 시간을 감소시키며 경량화된 이상 탐지 모델을 구축할 수 있음을 실험을 통해 검증하였다.

키워드 : 이산 웨이블릿 변환, 대역 에너지, 음향 신호, 이상 탐지

Key Words : discrete wavelet transform, band energy, acoustic signal, anomaly detection

ABSTRACT

It is important to detect and monitor abnormal conditions of machines in industrial environments. Acoustic signal-based anomaly detection technology has the advantage of performing fault detection at a low cost compared to conventional image/video-based or other sensor-based anomaly detection technologies. However, utilizing a large amount of acoustic signals collected from sensors requires a lot of computing resources for intensive signal processing and learning. Therefore, it is necessary to consider an acoustic signal-based anomaly detection system that utilizes computing resources efficiently. In this paper, we propose a method for extracting the band energy of acoustic signals using discrete wavelet transform and design a lightweight anomaly detection model. Experiments show that the band energy based on discrete wavelet transform can effectively compress the features of acoustic signals, reduce the data preprocessing time, and enable the construction of a lightweight anomaly detection model.

* 본 연구는 2022년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(NRF-2020R1A2B5B01002528)과 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(2021-0-00739), 그리고 2022학년도 이화여자대학교 대학원 장학금의 지원을 받아 수행되었습니다.

• First Author : Agency for Defense Development 연구원, Ewha Womans University, Department of Electronic and Electrical Engineering 학생(박사과정), eunhye.choi@ewha.ac.kr, 정회원

o Corresponding Author : Ewha Womans University, Department of Electronic and Electrical Engineering, Graduate Program in Smart Factory, hyunggon.park@ewha.ac.kr, 종신회원

논문번호 : 202309-080-C-RN, Received September 13, 2023; Revised November 25, 2023; Accepted December 20, 2023

1. 서론

기계 설비의 이상 상태 감지는 산업 현장에서 고장을 예측하여 사전에 대비함으로써 안전을 확보하고 생산 비용을 절감하는 데 중요한 역할을 한다. IoT와 네트워크 기술의 발전으로 방대한 양의 기계 설비의 데이터를 획득할 수 있게 되었고 이를 활용하여 설비의 이상 상태를 감지하는 연구가 활발히 진행되고 있다^[1]. 클라우드 컴퓨팅은 IoT 장비를 통해 획득한 데이터를 처리하여 이상 탐지 서비스를 제공할 수 있으나, 실시간 데이터 전송에 많은 대역폭이 필요하며 서비스 지연과 같은 문제가 발생할 수 있다. 엣지 컴퓨팅은 클라우드 컴퓨팅의 문제점을 완화할 수 있지만 엣지 노드의 컴퓨팅 및 메모리 자원이 제한적이기 때문에 효과적인 데이터 처리를 통한 경량 이상 탐지 시스템 구현이 중요하다.

음향 신호는 진동이나 온도와 같은 신호 측정에 사용하는 접촉식 센서와 대조적으로, 마이크와 같은 비접촉식 센서를 활용하여 설비에 직접적인 영향을 미치지 않고도 측정할 수 있다. 이러한 방법은 비용이 저렴하며 시스템을 간편하게 구축할 수 있는 장점이 있어 고장 진단 데이터 수집에 유용하게 활용될 수 있다^[2]. 기존 음향 신호 기반의 고장 진단 연구들은 시간, 주파수 도메인에서 통계적인 특징을 활용하거나^[3,4] 시간-주파수 도메인의 정보를 사용하는 단시간 푸리에 변환(STFT), 웨이블릿 변환(Wavelet transform), 위그너 빌 분포(Wigner-Ville distribution) 등 그 변형을 특징 추출에 활용한다^[5,6]. 이 중에서 가장 대표적인 방법은 푸리에 변환(Fourier transform) 기반의 주파수 특성 추출 방법이다. 많은 연구에서는 원신호에 STFT(short time Fourier transform)를 적용하여 스펙트로그램(spectrogram)의 2차원 이미지로 변환하는 과정을 거친 후, 이미지 기반의 신경망 모델을 학습하여 이상 탐지에 활용하고 있다^[7]. 이는 기존의 방법들을 활용하여 손쉽게 학습 모델을 생성하여 접근할 수 있다는 장점이 있지만, 원신호를 이미지 데이터로 변환하는 전처리 과정에 상당한 시간이 소요되며, 모델이 변환된 2차원 데이터의 특성을 학습하는 데도 많은 시간이 필요하다는 단점이 있다. 따라서 효과적인 이상 탐지 모델을 구축하기 위해 데이터 처리 시간 및 학습 시간을 단축할 수 있는 특징 추출이 필요하다.

기계 설비에서 발생하는 환경 음향 신호(environmental audio)는 사람의 음성보다 더 넓은 범위에 가청 주파수가 걸쳐 있고, 고주파수 과도현상과 잡음을 동시에 포함하고 있는 특징을 갖는다^[8]. 하지만 STFT

기반의 신호 변환은 신호의 주파수 영역과 시간 영역 사이의 트레이드 오프가 존재하여 환경 음향 신호의 특성을 추출하는데 한계가 있다. 반면 웨이블릿 변환(Wavelet transform)은 웨이블릿 함수의 스케일링(scaling)과 이동(translation)을 통해 신호의 국지적인 특성을 파악할 수 있으며, 주파수 크기에 따라 유연한 윈도우 크기 설정이 가능하기 때문에 다중 해상도 분석이 가능하다. 따라서 고정된 해상도를 갖는 STFT보다 환경 음향 신호의 특성을 더 잘 파악할 수 있는 장점이 있다^[9]. 하지만 디지털 신호에 대한 이산 웨이블릿 변환(Discrete Wavelet Transform, DWT)을 통해 변환된 웨이블릿 계수는 음향 신호 길이만큼 계수가 추출되기 때문에 이를 이상 감지 모델 학습에 바로 적용하기 어렵다. 이에 따라 DWT로 추출된 계수에 대한 2차적인 특징 추출을 통해 입력 데이터의 압축이 필요하다.

원신호에서 추출된 특징은 SVM, KNN, MLP 등 지도학습 기반의 머신러닝 기술과 함께 사용되어 이상 탐지에 활용된다^[10,11]. 하지만 라벨링된 형태의 다양한 종류의 비정상 데이터를 획득하는 것이 실질적으로 어렵고 데이터양이 증가할수록 일반화 성능이 떨어질 수 있다^[12,13]. LOF나 GMM와 같은 비지도 학습을 사용할 수 있지만, 이는 계산 복잡도가 높아서 대규모 데이터에서의 활용에 제약이 존재하기^[14,15] 때문에 딥러닝을 활용하는 것이 더 적합하다.

본 논문에서는 웨이블릿의 필터뱅크 성질을 활용하여 각 주파수 대역에 포함된 음향 신호의 에너지를 음향 신호의 특징 벡터로 활용하는 오토인코더 기반의 이상 탐지 시스템을 제안한다. 제안한 시스템은 이산 웨이블릿 변환을 사용한 대역 에너지 특징 추출을 통해 데이터 전처리 시간을 단축하고 신호의 특성을 효과적으로 압축한다. 또한 압축된 음향 신호 특성을 기반으로 경량화된 오토인코더를 구성하여 정상 음향 신호의 복원 오차를 최소화하는 모델 학습을 통해 신뢰할 수 있는 경량의 이상 탐지 성능의 시스템을 구축한다. 음향 신호를 기반으로 한 이상 탐지 시스템 개발 목적으로 만들어진 ToyADMOS 데이터 세트^[16]를 사용하여 제안한 시스템이 데이터 전처리 시간과 학습 시간을 크게 단축하면서도 우수한 성능을 보이는 것을 검증하였다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 음향 신호 기반의 이상 탐지 시스템 구축을 위한 특징 추출 방법과 모델 구성에 관해서 기술한다. 3장에서는 기존 연구와 제안한 방식과의 비교실험을 통해 제안한 이상 탐지 시스템에 대한 검증 및 분석을 수행하고

4장에서 결론을 맺는다.

II. 음향신호 기반의 이상 탐지 시스템

2.1 시스템 모델

음향 신호 기반의 이상 탐지 시스템은 그림 1에서와 같이 마이크를 이용한 음향 신호의 획득, 음향 신호 데이터 전처리, 음향 신호 기반의 이상 탐지 모델, 이상 신호 판단 모듈로 구성된다.

마이크를 이용해 획득된 N 개의 데이터 포인트로 이루어진 정상 상태 음향 신호는 $s(t) \in \mathbb{R}^N$ 로 정의한다. 음향 신호 모델을 효과적으로 학습하기 위해 특징 추출 함수 $F: \mathbb{R}^N \rightarrow \mathbb{R}^D$ 를 이용하여 원신호 $s(t)$ 로부터 음향 신호 특징 벡터 $\mathbf{x} = \{x_1, x_2, \dots, x_D\}$ 를 추출한다. 여기서 $N \gg D$ 로 특징 추출 함수 F 를 통해 정상 상태 음향 신호를 저차원의 특징으로 압축하는 과정을 의미한다.

음향 신호의 이상 탐지 모델은 오토인코더를 사용한다^[19]. 오토인코더는 입력 데이터의 차원을 압축하는 인코더 부분과, 압축된 데이터를 복원하는 디코더 부분으로 구성되어, 정상 상태의 음향 신호의 특성을 학습하는 구조이다. 오토인코더의 입력을 $\mathbf{x} \in \mathbb{R}^D$, 오토인코더의 출력을 $\hat{\mathbf{x}} \in \mathbb{R}^D$ 라고 하면 오토인코더는 입력과 출력의 차이를 최소화하는 방향으로 가중치를 학습한다. 입력과 출력의 차이를 손실함수 $L(\cdot)$ 로 정의하면 변수 θ 로 표현되는 오토인코더의 목적함수는

$$\arg \min_{\theta} L(\mathbf{x}, \hat{\mathbf{x}}) = \sum_D \|x_d - \hat{x}_d\|^2 \quad (1)$$

로 표현된다. 본 논문에서는 L2-norm을 사용한다.

입력 신호의 정상 상태를 판별하기 위한 이상치 점

수 $A_{\theta}(s(t))$ 는 음향 신호 특징 벡터와 오토인코더 출력 벡터의 차이 평균으로 계산된다.

$$A_{\theta}(s(t)) = \frac{1}{D} \sum_D \|x_d - \hat{x}_d\|^2 \quad (2)$$

이때 계산된 $A_{\theta}(s(t))$ 값과 설정된 임계값 γ 와 비교를 통해 $A_{\theta}(s(t)) > \gamma$ 인 경우 이상 신호, 그 외의 경우 정상 신호로 판단되어 음향 신호의 이상 탐지를 수행한다.

2.2 이산 웨이블릿 변환 기반의 대역 에너지 특징 추출

본 연구에서는 음향 신호의 특징을 추출하기 위해 이산 웨이블릿 변환을 사용하여 음향 신호의 대역 에너지 특징을 추출한다. 이산 웨이블릿 변환은 신호의 길이가 $N=2^M$ 인 음향 신호 $s(t)$ 를 다양한 해상도를 가진 주파수 영역의 신호로 분해하며, 웨이블릿 함수 $\psi_{m,n}(t) = (1/\sqrt{2^m})\psi(2^{-m}t-n)$ 와 스케일링 함수 $\phi_{m,n}(t) = (1/\sqrt{2^m})\phi(2^{-m}t-n)$ 로 표현할 수 있다. 각 식에서 m 은 함수의 스케일을, n 은 위치를 의미한다.

$$s(t) = A_{M,n} \phi_{M,n}(t) + \sum_{m=1}^M \sum_{n=0}^{2^m-1} D_{m,n} \psi_{m,n}(t) \quad (3)$$

$A_{M,n} = \langle s(t), \phi_{M,n}(t) \rangle$ 은 스케일 M 일 때의 근사 계수(approximation coefficients)이고 $D_{m,n} = \langle s(t), \psi_{m,n}(t) \rangle$ 은 세부 성분 계수(detailed coefficients)이다. 이때 웨이블릿 함수와 스케일링 함수의 직교 성질은 이산 웨이블릿 변환 과정을 필터뱅크(filterbanks)로 구현 가능하게 한다^[20].

각 스케일 별로 이산 웨이블릿 변환을 적용할 때마

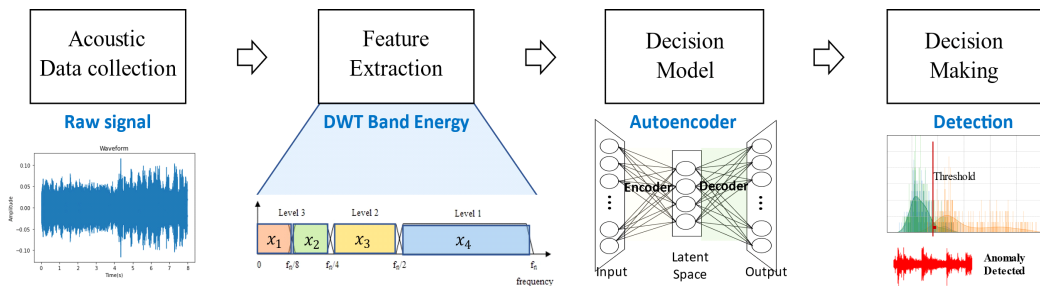


그림 1. 이산 웨이블릿 변환 기반의 음향 신호 이상탐지 시스템
Fig. 1. DWT-based acoustic anomaly detection system

다 스케일링 함수는 $(1/\sqrt{2})g(k)$ 로 표현되는 저주파 통과 필터(low-pass filter, LPF), 웨이블릿 함수는 $(1/\sqrt{2})h(k)$ 로 표현되는 고주파 통과 필터(high-pass filter, HPF)로 동작한다. m번째 분해 스케일에서의 근사 계수와 세부 성분 계수는

$$\begin{aligned} A_{m+1,n} &= \frac{1}{\sqrt{2}} \sum_k g(k) A_{m,2n+k} \\ &= \frac{1}{\sqrt{2}} \sum_k g(k-2n) A_{m,k} \end{aligned} \quad (4)$$

$$\begin{aligned} D_{m+1,n} &= \frac{1}{\sqrt{2}} \sum_k h(k) A_{m,2n+k} \\ &= \frac{1}{\sqrt{2}} \sum_k h(k-2n) A_{m,k} \end{aligned} \quad (5)$$

로 계산할 수 있다. 그림 2는 분해 수준 M=3인 경우 필터뱅크로 동작하는 이산 웨이블릿 변환의 과정을 보여준다.

원신호 $s(t)$ 의 대역 에너지 특징 신호는 분해 수준이 M인 경우 M 번째 스케일에서의 하나의 근사 계수와, 주파수 대역별로 존재하는 M개의 세부 성분 계수로부터 획득할 수 있다. 원신호의 에너지 $E(s(t))$ 는 근사 계수와 세부 성분 계수를 이용하여

$$E(s(t)) = (A_{M,n})^2 + \sum_{m=1}^M \sum_{n=0}^{2^{M-n}-1} (D_{m,n})^2 \quad (6)$$

로 계산할 수 있다. 따라서 원신호 대비 각 대역별로 포함된 대역 에너지 특징은 아래와 같이 계산된다.

$$x_0 = \frac{(A_{M,n})^2}{E(s(t))}, x_m = \frac{(D_{m,n})^2}{E(s(t))}, m = \{1, \dots, M\} \quad (7)$$

x_0 는 분해 수준이 M인 경우 근사 계수가 포함하는

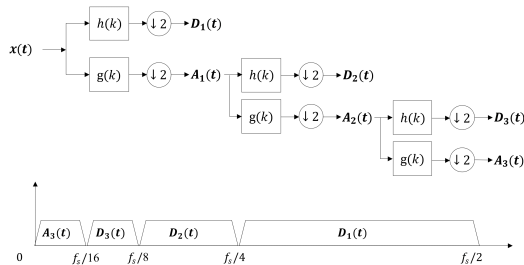


그림 2. 필터뱅크로 동작하는 이산 웨이블릿 변환
Fig. 2. DWT as filter bank (M=3)

대역의 전체 신호 대비 에너지를 의미하고, x_m 은 분해 수준이 m인 경우의 전체 신호 대비 해당 주파수 대역의 에너지를 의미한다. 따라서 음향 신호 특징 벡터는

$$\mathbf{x} = \{x_0, x_1, \dots, x_M\} \quad (8)$$

으로 표현할 수 있다.

III. 실험

3.1 실험 데이터

음향 신호에 기반한 이상 탐지 수행을 위해 ToyADMOS(anomaly detection in machine operating sounds) 데이터 세트를 사용한다^[16]. ToyADMOS는 음향 신호를 기반으로 기계의 이상 상태 탐지를 위해 만들어진 데이터 세트로, 소형 자동차(ToyCar), 소형 컨베이어 벨트(ToyConveyor), 소형 기차(ToyTrain)의 정상 상태에서의 음향 신호와 고의로 고장을 낸 후 획득한 비정상 음향 신호 데이터를 포함한다. 음향 신호는 마이크를 활용하여 48kHz로 샘플링 되었으며, 본 실험에서는 각 데이터 세트의 1번 ID 데이터를 16kHz로 리샘플링한 10초 길이의 신호를 사용하였다.

오토인코더는 정상 신호들을 이용하여 학습하였으며, 모델 평가 시 정상 데이터와 비정상 데이터의 비율을 표 1에서와 같이 1:1로 구성하여 평가를 수행했다.

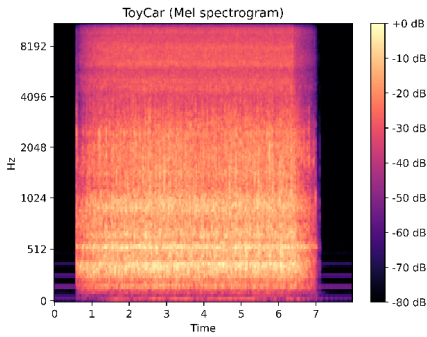
표 1. 실험용 음향 신호의 구성
Table 1. Audio signals for the dataset

Machine Type	Training data		
	Normal	Normal	Abnormal
ToyCar	1085	265	264
ToyConveyor	1445	355	400
ToyTrain	1080	270	270

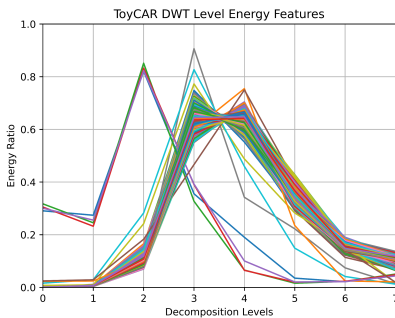
3.2 실험 변수 설정

3.2.1 비교 시스템

비교 대상으로 고려한 시스템은 IEEE ASSP에서 주최한 DCASE2020 챌린지의 “기계 상태 모니터링을 위한 비지도 학습의 음향 이상 신호 탐지”에서 제시된 시스템^[21]으로 관측된 음향 신호를 STFT 기반의 로그 멜-스펙트로그램(mel-spectrogram)의 2차원 이



(a) 로그 멜-스펙트로그램 (Log mel-spectrogram)



(b) DWT 대역 에너지 특징 (DWT band energy features)

그림 3. ToyCar 샘플에 대한 음향 신호 입력 비교
Fig. 3. Comparison of acoustic features for ToyCar sample

미지 형태로 변환하여 오토인코더를 학습하는 구조를 갖는다.

실험에서 로그 멜-스펙트로그램을 획득하기 위해 FFT 프레임은 64ms, 프레임 간격은 50%의 크기로 설정되었으며, 128개의 멜 필터를 사용했다. 로그 멜-스펙트로그램으로 변환된 입력은 309×640의 행렬로 표현된다. 그림 3의 (a)는 ToyCar 정상 데이터 샘플의 로그-멜 스펙트로그램을 보여준다.

오토인코더의 입력과 출력의 차원은 D=640이며, 인코더로 4개의 은닉층, 하나의 잠재공간, 디코더로 4개의 은닉층을 갖는다. 각 은닉층은 128개의 유닛을 보유하며 잠재공간은 8개의 유닛으로 구성된다. 활성화 함수로는 렐루(ReLU) 함수가, 최적화를 위해 이담(Adam)이 사용되었다. 학습률은 $\alpha = 0.001$, 배치 크기는 512, 100 에포크(epoch) 이후 학습이 종료된다.

3.2.2 제안한 시스템

제안한 이산 웨이블릿 변환 기반의 음향 탐지 시스템은 분해 수준 M=7을 사용했다. 따라서 음향 신호 특징 벡터 x 는 1×8인 벡터로 표현된다. 그림 3의 (b)

는 ToyCar 전체 정상 데이터 샘플의 음향신호 특징 벡터를 보여준다. x축은 각 분해 수준을 의미하고, y축은 분해 수준 별 포함된 에너지 비율을 보여준다.

비교 모델인 DCASE 2020과 동일한 성능 평가를 위해 오토인코더는 동일한 구조를 사용했다. 이산 웨이블릿 변환 기반의 대역 에너지 특징 추출을 통해 학습 데이터의 차원이 압축되었기 때문에 오토인코더의 입력과 출력의 차원은 D=8로 설정되었으며, 이에 따라 각 은닉층은 8개의 유닛을 보유한다. 잠재공간은 4개의 유닛을 갖는다. 활성화 함수와 최적화도 동일한 형태를 사용하고, 데이터가 압축됨에 따라 학습률은 $\alpha = 0.0001$, 배치 크기는 20으로 변경하여 실험을 수행했다.

3.2.3 성능 평가

(1) 성능 평가 지표

음향 신호 기반의 이상 탐지 시스템은 일반적으로 분류에서 많이 사용되는 AUC(area under ROC curve)와 pAUC(partial AUC)를 사용했다.

AUC는 FPR(false positive rate)과 TPR(true positive rate)로 그려지는 ROC 곡선 하단의 면적을 의미하며, AUC 값이 클수록 분류 성능이 뛰어난 것을 의미한다. pAUC는 부분 AUC로 특정 FPR 구간에서의 ROC 곡선의 면적을 의미한다. 특히 이상 탐지에서는 신뢰할 수 있는 시스템을 평가하기 위해 낮은 FPR에서 높은 TPR을 달성하는 것이 좋다. 동일한 척도에서 비교를 위해 FPR이 [0, 0.1] 구간에서 표준화된 pAUC를 계산하고 비교를 수행했다.

(2) 성능 비교 결과

실험은 총 10번 수행했으며, 각 독립적인 실험 수행의 평균값을 이용해 비교했다.

표 2에서 오토인코더 학습을 위해 사용하는 모델의 입력 크기를 비교해 보면, 로그 멜-스펙트로그램을 사용하는 DCASE 2020 모델의 입력 데이터 차원은 640인 반면에 DWT 기반의 대역 에너지를 사용하는 제안한 모델의 입력 데이터는 8차원이다. 따라서 제안한 모델은 DCASE 2020 모델 대비 80배로 압축된 크기의 입력 데이터를 사용하기 때문에 경량화된 오토인코더를 구성할 수 있다. 또한 압축된 저차원의 데이터를 학습하기 때문에 은닉층 유닛수가 감소하여 오토인코더의 학습 가능 변수(trainable parameters)도 DCASE 2020 모델 대비 99.71% 감소한다. 압축된 입력 데이터 및 경량화된 모델을 사용했을 때 DCASE

표 2. 비교 시스템과 제안한 시스템의 변수 설정값
Table 2. Parameters settings for the baseline and proposed system

		DCASE 2020 ^[10]	Proposed
Acoustic feature		Log mel-spectrogram	DWT based band energy
Input size per file		313x640	1x8
AE architecture	Input layer	1 input layer (640 units)	1 input layer (8 units)
	Encoder	4 hidden layers (128 units per layer)	4 hidden layers (8 units per layer)
	Latent space	1 hidden layer (8 units)	1 hidden layer (4 units)
	Decoder	4 hidden layers (128 units per layer)	4 hidden layers (8 units per layer)
	Output layer	1 output layer (640 units)	Output layer (8 units)
Trainable parameters		267,928	788

2020 시스템과 AUC, pAUC 측면에서 분류 성능 평가 비교는 각각 그림 4, 그림 5와 같다.

제안한 시스템은 로그 멜-스펙트로그램을 사용한 DCASE 2020 시스템 대비 99% 수준의 압축된 입력 데이터를 사용했음에도 불구하고 AUC 측면에서 ToyCar의 경우 3%의 성능 증가, ToyConveyor는 동일한 성능, ToyTrain의 경우 1%의 성능 감소만을 보였다. pAUC 측면에서는 DCASE 2020 시스템 대비

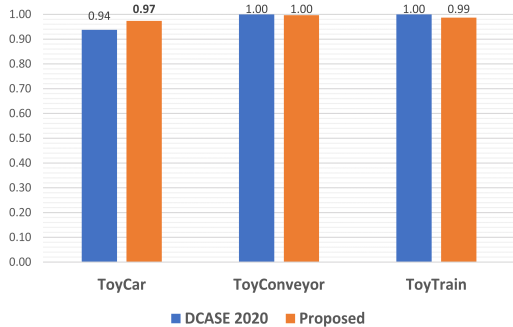


그림 4. 평균 AUC 성능 비교
Fig. 4. Comparison of average AUC

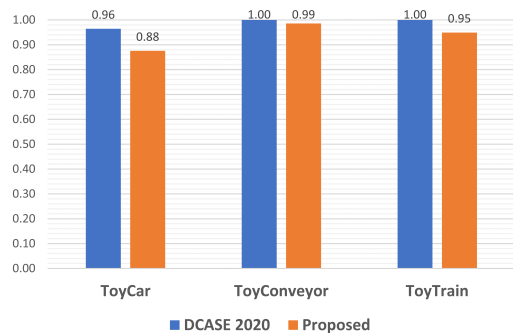


그림 5. 평균 pAUC 성능 비교
Fig. 5. Comparison of average pAUC

하여 ToyCar는 8%, ToyConveyor는 1%, ToyTrain의 경우 5%의 성능 감소세를 보였지만, 이는 압축된 입력 데이터 및 모델 측면에서 비교해 보았을 때 미미한 수준으로 판단할 수 있다. 이는 DWT 기반의 대역 에너지 특징이 시간-주파수 해상도가 고정된 로그 멜-스펙트로그램과 다르게 다중 해상도 분석을 통해 주파수 대역별 신호 특징을 효과적으로 추출하기 때문이다. 실험을 통해 압축된 데이터 수준 대비 음향 신호의 이상 탐지 성능을 유지함을 확인함으로써 제안한 방법이 효과적으로 음향 신호의 이상 특징을 추출함을 검증했다.

시스템의 계산 복잡도를 평가하기 위해 그림 6에서는 DCASE 2020 모델과 제안한 모델의 데이터별 전처리 시간과 모델의 학습 시간을 비교하였다(로그 스케일). DCASE 2020 데이터 전처리 시간(Baseline Gen), 제안한 시스템의 전처리 시간(Proposed Gen)을 비교하면 제안한 시스템이 평균적으로 6.4배 이상 빠른 것을 확인할 수 있다. 또한 비교 모델의 학습 시간(DCASE 2020 Train)과 제안한 모델의 학습 시간(Proposed Train)을 비교하면 제안한 시스템이 평균적

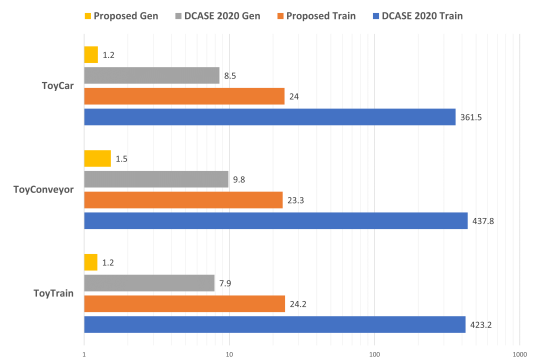


그림 6. 시스템 계산 복잡도 비교
Fig. 6. Comparison of computational complexity

으로 17배 빠른 것을 확인할 수 있다.

따라서 제안한 이산 웨이블릿 변환 기반의 대역 에너지 특징 추출을 활용하면 효과적으로 음향 신호의 특징을 추출하여 우수한 성능의 경량 이상 탐지 시스템을 구축할 수 있음을 확인했다.

IV. 결 론

본 논문에서는 컴퓨팅 자원을 효율적으로 사용하는 이산 웨이블릿 변환 기반 음향 신호 이상 탐지 시스템을 제안하였다. 이산 웨이블릿 변환을 활용하여 음향 신호의 대역 에너지를 추출하여 특징 벡터로 구성하고, 이를 기반으로 이상 탐지 모델을 구축했다. 추출된 특징 벡터는 원 음향 신호를 효과적으로 압축함으로써 데이터 전처리 시간을 단축시킨다. 또한 추출된 특징 벡터 기반으로 경량화된 오토인코더 모델을 구성하여 단시간 학습을 통해 이상 탐지 모델 구축이 가능함을 실험을 통해 검증했다. 따라서 제안한 방법은 제한된 컴퓨팅 자원을 갖는 IoT 엣지 노드에 효과적으로 활용될 수 있다. 본 논문은 실험실 환경에서 획득한 음향 신호에 대해 수행되었기 때문에 실제 환경 노이즈가 포함된 음향 신호에 대한 이상 탐지에 효과적으로 적용할 수 있는 방법에 관한 추가적인 연구가 필요하다.

References

- [1] G. Pang, C. Shen, L. Cao, and AVD. Hengel, "Deep learning for anomaly detection: A review," *ACM Computing Surv. (CSUR)*, vol. 54, no. 2, pp. 1-38, 2021. (<https://doi.org/10.1145/3439950>)
- [2] J. Grebenik, C. Bingham, and S. Srivastava, "Acoustic diagnostics of electrical origin fault modes with readily available consumer-grade sensors," *IET Electric. Power Appl.*, vol. 13, no. 12, pp. 1946-1953, 2019. (<https://doi.org/10.1049/iet-epa.2019.0232>)
- [3] A. Qurthobi, R. Maskeliūnas, and R. Damaševičius, "Detection of mechanical failures in industrial machines using overlapping acoustic anomalies: A systematic literature review," *Sensors*, vol. 22, no. 10, pp. 3888, 2022. (<https://doi.org/10.3390/s22103888>)
- [4] Q. Pan, R. Zhou, J. Su, T. He, and Z. Zhang, "Automatic localization of the rotor-stator rubbing fault based on acoustic emission method and higher-order statistics," *J. Mechanical Sci. and Technol.*, vol. 33, pp. 513-524, 2019. (<https://doi.org/10.1007/s12206-019-0104-9>)
- [5] G. Zheyu, L. Jing, W. Xiufeng, and L. Yuhe, "Grinding burn detection based on cross wavelet and wavelet coherence analysis by acoustic emission signal," *Chin. J. Mechanical Eng.*, vol. 32, no. 1, pp. 1-10, 2019. (<https://doi.org/10.1186/s10033-019-0384-0>)
- [6] W. FQ and G. Meng, "Feature extraction based on the 3D spectrum analysis of acoustic signals to identify rotor malfunction," *Int. J. Advanced Manufacturing Technol.*, vol. 28, pp. 1146-1151, 2006. (<https://doi.org/10.1007/s00170-004-2470-3>)
- [7] S. C. Lee and S. Y. Lee, "딥러닝, 소음진동분야에도 유용한가?," *J. KSNVE*, vol. 30, no. 1, pp. 7-13, Jan. 2020. (<https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE09292347>)
- [8] S. Waldekar and G. Saha, "Wavelet transform based mel-scaled features for acoustic scene classification," in *INTERSPEECH*, vol. 2018, pp. 3323-3327, 2018. (<https://doi.org/10.21437/Interspeech.2018-2083>)
- [9] J. Lin, "Feature extraction of machine sound using wavelet and its application in fault diagnosis," *NDT & e Int.*, vol. 34, no. 1, pp. 25-30, 2001. ([https://doi.org/10.1016/S0963-8695\(00\)00025-6](https://doi.org/10.1016/S0963-8695(00)00025-6))
- [10] M. Altaf, M. Uzair, M. Naeem, A. Ahmad, S. Badshah, J. A. Shah, and A. Anjum, "Automatic and efficient fault detection in rotating machinery using sound signals," *Acoustics Australia*, vol. 47, pp. 125-139, 2019. (<https://doi.org/10.1007/s40857-019-00153-6>)
- [11] R. P. da Cruz, F. V. da Silva, and A. M. F. Fileti, "Machine learning and acoustic method

applied to leak detection and location in low-pressure gas pipelines,” *Clean Technol. and Environ. Policy*, vol. 22, pp. 627-638, 2020.
(<https://doi.org/10.1007/s10098-019-01805-x>)

[12] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, “Applications of machine learning to machine fault diagnosis: A review and roadmap,” *Mechanical Syst. and Signal Process.*, vol. 138, p. 106587, 2020.
(<https://doi.org/10.1016/j.ymssp.2019.106587>)

[13] J. Jiao, M. Zhao, J. Lin, and K. Liang, “A comprehensive review on convolutional neural network in machine fault diagnosis,” *Neurocomputing*, vol. 417, pp. 36-63, 2020.
(<https://doi.org/10.1016/j.neucom.2020.07.088>)

[14] X. Gu, L. Akoglu, and A. Rinaldo, “Statistical analysis of nearest neighbor methods for anomaly detection,” *Advances in NIPS*, vol. 32, 2019.
(<https://doi.org/10.48550/arXiv.1907.03813>)

[15] J. Wu, F. Yang, and W. Hu, “Unsupervised anomalous sound detection for industrial monitoring based on ArcFace classifier and gaussian mixture model,” *Applied Acoustics*, vol. 203, p. 109188, 2023.

[16] Y. Koizumi, S. Saito, H. Uematsu, N. Harada, and K. Imoto, “ToyADMOS: A dataset of miniature-machine operating sounds for anomalous sound detection,” *IEEE WASPAA*, pp. 313-317, Oct. 2019.
(<https://doi.org/10.1109/WASPAA.2019.8937164>)

[17] P. Baldi, “Autoencoders, unsupervised learning, and deep architectures,” in *Proc. ICML Wkshp. Unsupervised and Transfer Learning. JMLR Wkshp. and Conf. Proc.*, pp. 37-49, 2012.
(<https://dl.acm.org/doi/abs/10.5555/3045796.3045801>)

[18] S. Mallat, “A theory for multiresolution signal decomposition: The wavelet representation,” *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 11, no. 7, pp. 674-693, 1989.
(<https://doi.org/10.1109/34.192463>)

[19] Y. Koizumi, Y. Kawaguchi, K. Imoto, T. Nakamura, Y. Nikaido, R. Tanabe, H. Purohit, K. Suefusa, T. Endo, M. Yasuda, and N. Harada, “Description and discussion on DCASE2020 challenge task2: Unsupervised anomalous sound detection for machine condition monitoring,” *arXiv preprint arXiv:2006.05822*, 2020.
(<https://doi.org/10.48550/arXiv.2006.05822>)

최 은 혜 (Eunhye Choi)



2012년 2월 : 이화여자대학교 전
자공학과 졸업
2014년 2월 : 이화여자대학교 전
자공학과 석사
2014년 2월~현재 : 국방과학연
구소 선임연구원
2022년 3월~현재 : 이화여자대
학교 전자공학과 박사과정

<관심분야> 이상 탐지, 인공지능, 머신러닝, 멀티에이
전트 시스템

[ORCID:0000-0002-6208-4030]

박 형 곤 (Hyunggon Park)



2004년 2월 : 포항공과대학교 전
자전기공학과 졸업
2006년 3월 : University of
California, Los Angeles
(UCLA) M.S.
2008년 12월 : University of
California, Los Angeles
(UCLA) Ph.D.

2010년~현재 : 이화여자대학교 전자전기공학과 교수
<관심분야> 멀티에이전트 시스템 최적화, 머신러닝,
인공지능, 게임이론

[ORCID:0000-0002-5079-1504]