

웨이 체결 강도 음향 검사: 디노이저 및 분류 모델 학습 절차 설계 및 성능 평가

장인성*, 우민식**, 박호민***, 유대승***†

*부산대학교, **금오공과대학교, ***한국전자통신연구원

wkddlstjd2@gmail.com, alstlr0718@gmail.com, hominpark@etri.re.kr, ooseyds@etri.re.kr

Wedge Tightness Acoustic Test: Learning Procedure Design and Performance Evaluation of Denoiser and Classifier Model

Insung Jang*, Woo Min Sik**, Ho-Min Park***, Dae Seung Yoo***†

*Pusan National University, **Kumoh National Institute of Technology,

***Electronics and Telecommunications Research Institute

요약

발전기 내부의 고정자 코일을 고정하기 위해 사용하는 웨지의 조임 상태는 발전기의 고장과 직접 연관되어 있다. 전문 인력이 웨지를 타격하여 나온 음향을 듣고 판단하는 현재의 검사 방법을 대신하여 로봇이 웨지를 타격한 후 나온 음향을 이용한 딥러닝 기반 음향 검사를 하고자 한다. 특히, 본 논문은 잡음 및 잔향에 강건한 음향 검사를 위해 디노이저 및 분류 모델의 학습 절차를 설계하고 학습한 후 성능을 평가하였다. 클린 데이터에 잡음 및 잔향을 합성하여 데이터셋을 구축하였으며 이를 로그 멜 스펙트로그램으로 변환하여 이용하였다. 평가한 디노이저 모델은 DVUNET, DUNET, DVAE, DAE의 총 4 가지이며, 분류 모델로는 ResNet-50 을 이용하였다. 합성한 잔향의 강도에 따라 4 가지 경우에 대해 실험을 진행하였으며, 성능 평가 지표로 디노이저 모델은 MSE, 분류 모델은 F1 Score 를 이용하였다. 디노이저 모델로 U-Net 기반을 사용한 경우가 AutoEncoder 기반을 사용한 경우보다 좋은 결과를 보였다.

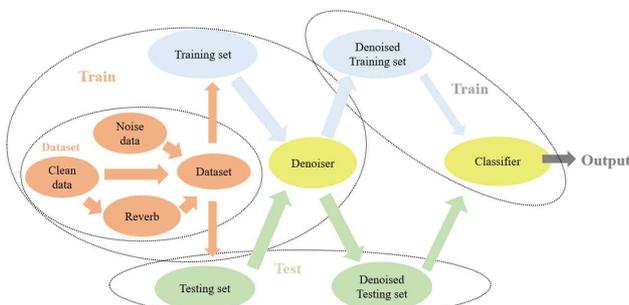
I. 서론

발전기 내부의 고정자 코일을 고정하기 위해 웨지를 이용한다. 웨지 조임 상태를 나타내는 웨지 체결 강도는 발전기의 고장과 직접 연관되어 있어 매우 중요하다. 현재 사용되는 웨지 체결 강도를 평가하는 방법은 전문 검사 인력을 통해 웨지를 타격하여 나온 소리를 듣고 판단하는 타음 기반 청음 방식[1]이다. 이 과정에서 발전기를 정지하고 회전자 및 고정자를 탈거해야 하므로 비용 및 시간이 많이 소요된다. 이러한 문제를 해결하기 위해 본 논문은 로봇을 이용하여 내부 부품을 탈거하지 않고 웨지를 타격하여 녹음한 음향을 바탕으로 불량 여부를 판단하는 딥러닝 기반 디노이저 및 분류 모델들의 학습 절차를 설계하고 성능을 평가한다.

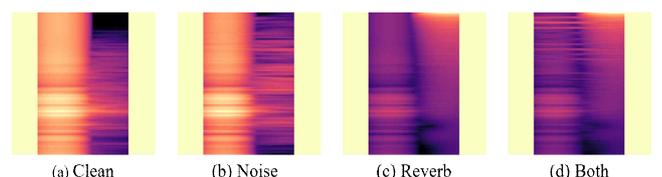
<그림 1>은 모델 학습 및 평가를 위해 설계한 평가 방법의 구조 및 진행순서를 나타낸다. 현장 상황에 강건한 모델을 얻기 위해 클린 데이터에 잡음 및 잔향을 합성하여 데이터셋을 구축하였다. 디노이저 모델을 통해 잡음 및 잔향을 제거한 후 이를 분류 모델을 이용하여 판단한다.

II. 본론

본 논문에서 사용한 클린 데이터는 실제 웨지 타음 검사에서 획득한 잡음 및 잔향이 없는 음향으로, Caution 320 개, Fault 440 개, Normal 320 개의 총 1080 개이다. 잡음 데이터는 구글 오디오셋[2]의 진동, 메커니즘, 기계식 팬, 도구 카테고리에서 50 개를 선정하였다. 클린 데이터에 직접 잡음과 잔향을 추가하여 데이터셋을 제작하였다. 잡음은 각 클린 데이터마다 잡음 데이터에서 SNR(Signal to Noise Ratio)을 0dB~40dB 사이에서 무작위로 결정한 10 개를 선정하여 합성하였다. 잔향은 클린 데이터에서 음량의 비율을 0%, 25%, 50%, 75%로 추출하여 최초 음향 발생 50ms 뒤에 합성하였다. 제작한 데이터셋은 Train:Valid:Test = 7:2:1 비율로 사용하였다.

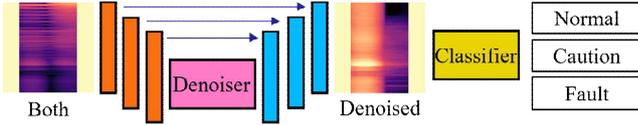


<그림 1> 전체 구성도



<그림 2> 오디오 별 로그 멜 스펙트로그램

<그림 2>는 (a), (b), (c), (d) 순서대로 클린, 잡음 합성, 잔향 합성, 잡음과 잔향 합성 음향을 로그 멜 스펙트로그램으로 변환한 결과이다.

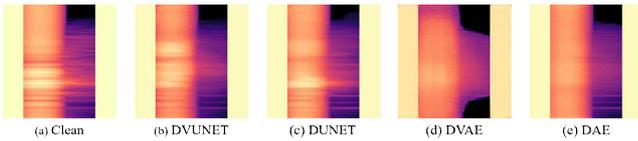


<그림 3> 모델 구조

전체 모델 구조는 <그림 3>과 같다. 디노이저 모델은 잡음 및 잔향을 제거한 로그 멜 스펙트로그램을 재구축한다. 분류 모델은 이를 이용하여 클래스를 예측한다.

사용한 디노이저 모델은 DVUNET[3], DUNET[4], DVAE[5], DAE[6]로 총 4 가지이며, 이 모델들은 기존의 Convolution 커널에 패딩을 추가함으로써 연산량을 줄여 특징을 효율적으로 추출하는 Dilated Convolution 을 적용한 것이다. DVUNET 과 DUNET 은 U-Net 구조를, DVAE 와 DAE 는 AutoEncoder 구조를 이용하였다. DVUNET 은 U-Net 의 Linear bottleneck 을 VAE 의 Variational bottleneck 으로 변경한 모델이다. 분류 모델은 ResNet-50[7]을 사용하였다.

모델 학습 시 배치 사이즈는 8, 옵티마이저는 AdamW, 학습률은 0.0003 으로 설정하여 200 epoch 간 학습하던 중 10 epoch 동안 모델의 성능 개선이 없는 경우 조기 종료하여 모델을 획득하였다. 각 모델들의 성능은 MSE, F1 score 평가 지표를 사용해 평가하였다.



<그림 4> 디노이저 모델 별 잡음 및 잔향 제거 결과

<그림 4>는 잔향의 비율을 75%로 설정하였을 때 디노이저 모델을 이용하여 잡음 및 잔향을 제거한 결과이다. 디노이저 모델로 U-Net 구조를 사용하였을 때는 클린 음향에 가깝게 재구축되었으나, AutoEncoder 구조를 이용한 경우 클린 음향의 특징이 많이 소실되었다.

Ratio of Reverb		DVUNET	DUNET	DVAE	DAE
0%	MSE	0.0022	0.0010	0.0168	0.0037
	F1	0.9991	1	0.2413	0.3007
25%	MSE	0.0022	0.0014	0.0163	0.0034
	F1	0.9936	0.9936	0.2413	0.4616
50%	MSE	0.0021	0.0006	0.0162	0.0033
	F1	0.9706	0.9881	0.2413	0.3992
75%	MSE	0.0023	0.0011	0.0166	0.0034
	F1	0.9642	0.9716	0.2434	0.4252

<표 1> 디노이저 및 분류 모델의 성능

<표 1>의 모든 결과에서 디노이저 모델로 DUNET 을 사용한 경우가 가장 좋은 성능을 보였다.

DVUNET 의 경우 Variational bottleneck 에서 잡음 및 잔향에 대한 표현력이 감소하였다. DVAE 및 DAE 의 경우 skip connection 의 부재로 인코더의 저수준 특징이

디코더로 전달되지 않아 클린 음향으로의 재구축에 실패하였고, 분류를 위한 정보가 소실되어 낮은 F1 score 를 보였다.

잔향의 비율을 달리하여 성능을 평가한 결과, 잔향이 커질수록, 분류 모델의 F1 score 가 감소하였다. 이를 통해 잔향이 청음 기반 검사에 부정적인 영향을 주며, 본문에서 사용한 잔향 합성 방법이 유효함을 확인하였다.

III. 결론

본 논문에서는 로봇 기반 웨지 체결 강도 타음 검사를 위한 딥러닝 모델의 학습 절차를 설계하였다. 클린 데이터에 잔향과 잡음을 추가하여 생성한 데이터셋을 이용하여 디노이저 및 분류 모델을 학습한 후 성능을 평가하였다. 디노이저 모델로 U-Net 기반을 사용한 경우 저수준의 정보를 전달할 수 있어 AutoEncoder 기반보다 우수한 성능을 보였다.

본 연구가 웨지 체결 강도 검사뿐만 아니라 다른 기계 고장 청음 검사에서도 활용될 수 있을 것으로 예상된다.

ACKNOWLEDGMENT

이 연구는 울산시-ETRI 2 차 공동협력사업(24AB1610, 제조 혁신을 위한 주력산업 지능화 기술 개발 및 산업현장에서의 사람-이동체-공간 자율협업지능 기술 개발) 및 2024 년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음(2023-0-00041).

참 고 문 헌

- [1] 오준석 외 4 명, “발전기 고장자 웨지 건전성평가를 위한 신호분석 기법”, 한국소음진동공학회논문집 제 30 권 제 5 호, pp. 441~447, 2020
- [2] <https://research.google.com/audioset/>
- [3] E. J. Nustede and J. Anemüller, “Towards speech enhancement using a variational U-Net architecture”, EUSIPCO, 2021
- [4] Olaf R., Philipp F., Thomas B., “U-Net: Convolutional Networks for Biomedical Image Segmentation”, MICCAI, 2015
- [5] Diederik P Kingma, Max Welling, “Auto-Encoding Variational Bayes”, ICLR, 2014
- [6] Rumelhart, D.E., Hinton, G.E., Williams, R.J., “Learning Internal Representations by Error Propagation” Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1, 1986
- [7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, “Deep Residual Learning for Image Recognition”, CVPR, 2016