

## 진동 데이터 이용 고장 진단 모델의 성능 향상을 위한 전력 섭동 기반 데이터 증강

곽민선\*, 박부견\*  
포항공과대학교\*

### Power Perturbation-Based Data Augmentation for Improvement of Fault Diagnosis using Vibration Data

Minseon Gwak\*, PooGyeon Park\*  
POSTECH\*

**Abstract** - 딥러닝을 이용한 고장 진단의 경우 학습에 사용된 데이터에 어떤 종류의 데이터가 포함되었는지에 따라 고장 진단 모델의 학습된 지식과 정확도가 결정된다. 센서의 위치 등 불확실한 요소가 존재하는 실제 공정에서 딥러닝 모델을 사용하기 위해서는 학습에 사용된 데이터셋과 다른 도메인의 데이터에 대한 처리가 요구되는 상황 또한 고려되어야 한다. 본 논문에서는 진동 신호를 사용하는 고장 진단 딥러닝 모델을 학습하는 과정에서 사용할 수 있는 데이터 증강 방법을 제시한다. 제안된 방법의 효과를 검증하기 위하여 오픈 데이터셋에서 학습 데이터의 측정 환경과 다른 환경에서 측정된 데이터를 테스트 데이터로 사용할 때 제안한 방법을 사용하여 학습한 모델과 사용하지 않고 학습한 모델의 정확도 성능을 비교한다.

클래스를 포함하고 있으며, 고장 클래스는 다시 해당 결함의 크기(Fault diameter)에 따라 나뉜다. 또한, 각 데이터셋은 네 가지 모터 부하(Motor load) 환경에서 측정되었다. 모터 부하가 바뀔 때 따라 일어나는 변화 중 가장 고장 진단에 영향을 주는 것은 모터의 속도이며, 각 부하 정도에 따라 작동하는 모터의 대략적 속도는 <표 2>와 같다.

<표 1> CWRU 데이터셋의 클래스 및 수집 환경 목록

Type	Fault diameter [mil]	Motor load [HP]
Normal	-	0, 1, 2, 3
Ball	7, 14, 21, 28	0, 1, 2, 3
Inner race	7, 14, 21, 28	0, 1, 2, 3
Outer race	7, 14, 21	0, 1, 2, 3

## 1. 서 론

고장 진단은 설비에 설치된 센서에서 측정되는 신호를 분석하여 설비의 작동 상태를 진단하는 과정이다. 고장 진단은 신호가 발생하는 과정을 직접 모델링하는 방식에서 보다 복잡한 비선형성을 모델링할 수 있고 개별 설비에 대한 경험적 지식 및 전문 지식이 덜 요구되는 딥러닝을 사용하는 방식으로 변화하고 있다 [1]. 딥러닝으로 얻어진 모델은 학습 데이터의 패턴을 학습하기 때문에, 학습 데이터의 질과 양이 충분할 때 좋은 성능의 모델을 얻을 수 있다.

고장 진단 모델 학습을 위한 충분한 양질의 데이터를 얻기에는 어려운 점이 있다. 먼저, 기계 설비의 경우 수명이 길고 설비에 치명적인 결함이 발생하기 전 대개 보수 조치가 취해지기 때문에 고장 데이터를 얻는 것이 어렵다 [2]. 또한, 측정하는 센서의 위치, 설비의 작동 조건 등 환경의 불확실성으로 인하여 학습 데이터를 수집할 때와 온전히 같은 상황에서 설비가 동작한다고 볼 수 없다. 학습 데이터의 특성에 과적합하게 학습된 모델의 경우 다른 환경에서의 데이터에 대한 정확한 진단이 불가능할 수 있다 [3].

고장 진단에서 가장 널리 사용되는 진동 데이터의 경우, 주파수 영역에서 정상과 고장의 특징이 뚜렷이 나타난다. 고장 진단 딥러닝 모델은 이러한 주파수 영역의 특징을 분류의 기준으로 학습한다 [4]. 본 논문에서는 주파수 영역의 특징의 중요성과 환경 변화에 따라 주파수 영역에 생기는 변화에 영감을 얻어, 주파수 영역에서 데이터에 변형을 가하여 학습 데이터를 생성하고 학습에 함께 사용하는 데이터 증강 방법을 제안한다.

<표 2> CWRU 데이터셋의 모터 부하 환경 별 모터 속도

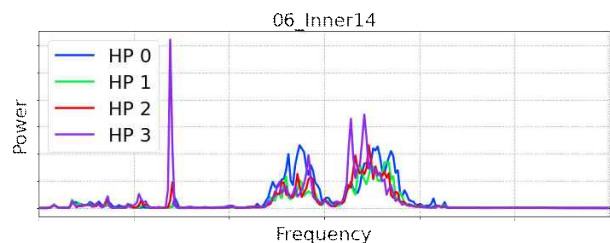
Motor load [HP]	Motor speed [rpm]
0	1797
1	1772
2	1750
3	1730

<그림 1>은 모터 부하 조건 별로 전체 시계열 데이터의 주파수 변환한 결과를 평균화하지 않고 그대로 한 그래프에 나타낸 것이다. 예시로, Inner race의 14 mil 고장 클래스의 데이터의 파워 스펙트럼은 <그림 1>과 같이 모터 부하 조건에 따라 달라진다. 같은 종류의 고장의 경우, 모터 부하 변화에 따라 모터의 속도가 달라지면서 주요 주파수의 위치는 거의 같으나 미세하게 조금씩 달라짐을 확인할 수 있다. 모터 부하가 달라짐에 따라 주요 주파수가 가지는 전력의 크기는 달라지나, 전력 변화에 규칙성은 확인되지 않는다.

## 2. 본 론

### 2.1 데이터 분석

본 장에서는 측정 환경에 따른 데이터의 주파수 영역에서의 변화를 살펴본다. 데이터셋으로는 Case Western Reserve University (CWRU) 베어링 결함 데이터셋 [5]을 활용하였다. 베어링은 많은 회전 운동을 포함한 설비의 고장 원인이 되는 부품으로, 해당 데이터셋은 고장 진단 연구의 대표 데이터셋으로 사용되고 있다. 해당 데이터셋은 <표 1>과 같이 정상 (Normal) 클래스를 비롯한 Ball, Inner race, Outer race의 세 가지 고장



<그림 1> 모터 부하 별 데이터 파워 스펙트럼 비교

#### 2.1.1 데이터 증강

본 논문에서는 모터 부하 변화에 따른 파워 스펙트럼 양상 변화에 규칙성이 관찰되지 않는 것에 근거를 두어, 개별 주파수 별로 전력에 확률적 변화를 가한 뒤 변형된 데이터를 시간 영역으로 변환하여 사용하는 데이터 증강 방법을 제안한다. 구체적으로, 클래스  $c$ 의 데이터  $x_c(t)$ 가 주어졌을 때,  $x_c(t)$ 의 주파수

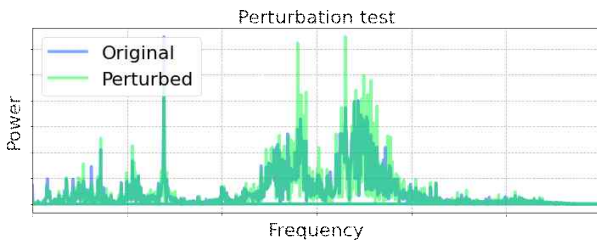
변환  $X_c(f)$ 에 대하여 주파수  $f$  별로 식 (1)과 같은 변형을 수행한다.

$$X_c(f) = \alpha X_c(f) \quad \text{for } f \in [0, \frac{f_s}{2}] \quad (1)$$

where  $\alpha \sim N(1, \sigma)$

식 (1)에서  $f_s$ 는 샘플링 주파수를 나타내며, 주파수 변환의 범위는 discrete Fourier transform을 수행할 때 Nyquist criterion에 의하여  $[0, \frac{f_s}{2}]$ 로 정의된다. 전력 변동을 수행하는 상수  $\alpha$ 는 정규분포  $N(1, \sigma)$ 를 따르는 값으로 정의되며, 표준 편차  $\sigma$ 는 hyperparameter이다.

<그림 2>는 데이터 변형을 수행한 결과 예시이다. 주파수 별로 개별적으로 랜덤 값을 생성하여 변형하여, 불규칙적으로 전력이 커지거나 작아짐을 확인할 수 있다. 변형한 데이터는 다시 시간 영역으로 변화하여 학습에 사용한다.



<그림 2> 주파수 영역에서의 데이터 변형 예시

### 3. 사례 연구

본 장에서는 모델 예측 정확도 변화를 통해 제안한 방법의 효과를 검증한다. 데이터셋은 앞서 분석을 진행한 CWRU 데이터셋을 사용하였으며 <표 1>의 정상 클래스 및 고장 클래스의 7, 14, 21 결합 크기 데이터를 사용하여 총 10 개 클래스에 대한 고장 진단 모델을 학습하였다. 각 클래스의 원본 시계열 데이터는 2024 샘플 단위로 분할하여 학습 및 테스트에 사용하였다. 학습 모델은 1D-CNN 레이어 6층으로 구성하였으며 각 레이어의 필터 크기는 7, 채널 수는 입력에 가까운 순으로 (8, 8, 16, 16, 32, 32)개로 설정하였다. 특성 추출 시 pooling layer는 사용하지 않았으며, 활성화 함수로 tanh 함수를 사용하였다. He initialization을 수행하였으며 최적화에 Adam optimizer를 사용하였다.

<표 3>은 제안한 방법을 사용하지 않고 학습한 모델과 사용하여 학습한 모델의 성능을 비교한 표이다. 모든 결과는 각 test case에 대해 5번의 결과를 평균한 값이다. 제안 방법의 hyperparameter인  $\sigma$ 는 0.5로 설정하였다. 각 test case에서는 특정 motor load 데이터들만 학습 데이터에 사용하고 학습에 사용하지 않은 다른 특정 motor load 데이터들을 테스트 데이터로 사용하였다. 총 12 가지 test case 중 8 가지 test case에서 제안 방법을 사용하여 학습한 모델이 더 높은 정확도를 보였다.

제안 방법을 통해 정확도 증가를 보지 못한 test case 네 가지 중 세 가지는 3 HP의 motor load 데이터를 예측하는 경우였고 나머지 한 경우는 학습 데이터로 3 HP의 motor load 데이터를 사용한 경우였다. Motor load 3 HP 데이터의 파워 스펙트럼을 살펴보면, 주요 주파수의 전력이 10배 증가하는 등 극단적이며 뚜렷한 변화가 나타나는 경향이 있다. 난수를 이용한 제안한 데이터 증강 방법을 사용했을 때 motor load 3 HP에서 나타나는 경향과 반대 방향의 변화가 일어나는 등의 상황이 발생할 수 있으므로 제안 방법으로 예측 정확도가 증가하는 효과를 보는 것이 어려웠던 것으로 보인다. 이에 따라, 추후 연구로 모델의 주요 주파수를 특정하고 주요 주파수 별 전력 변동 방향 경우의 수에 대한 데이터 변형을 수행하여 데이터를 증강하는 방법이 연구될 필요가 있다. 더불어, 전력 크기 뿐 아니라 주요 주파수 위치를 변경하는 방법에 대한 추후 연구가 필요하다.

<표 3> 기본 모델과 제안 방법 사용 모델의 정확도 비교

Motor load (Train → Test)	Test accuracy [%]	
	Baseline	Augmentation
0 → 1	90.29	<b>100.00</b>
0 → 2	93.37	<b>99.03</b>
0 → 3	86.02	83.04
1 → 0	95.40	<b>98.15</b>
1 → 2	99.95	<b>100.00</b>
1 → 3	91.32	82.14
2 → 0	93.53	<b>95.98</b>
2 → 1	91.31	<b>97.93</b>
2 → 3	93.68	90.93
3 → 0	86.07	<b>87.17</b>
3 → 1	81.83	<b>84.66</b>
3 → 2	95.72	94.41

### 4. 결론

본 논문에서는 고장 진단의 대표 데이터셋 분석을 통해 데이터 수집 환경 변화에 따른 진동 데이터 특성 변화를 분석하고, 분석 결과를 바탕으로 설비에 대한 전문 지식 없이도 사용 가능한 데이터 증강 기법을 제안하였다. 사례 연구에서 학습 데이터와 테스트 데이터의 수집 환경이 다를 때 제안한 기법을 사용하여 학습할 경우, 환경 변화로 인해 데이터의 양상이 극단적으로 변하지 않는 경우에는 대부분 정확도가 증가하였다.

제안한 방법에서는 데이터의 전력을 변동시킴으로써 변형 데이터를 생성하였으나, 모터 속도 변화로 인한 주요 주파수 위치가 변화하는 것은 데이터 증강 과정에 고려되지 않았다. 따라서, 전력 변형과 더불어 주요 주파수 위치를 변화하는 증강 방법에 대한 추후 연구가 필요하다.

#### 감사의 글

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT, and Future Planning (2020R1A2C2005709)

#### [참고 문헌]

- [1] Zhang, L., Lin, J., Liu, B., Zhang, Z., Yan, X. and Wei, M., "A review on deep learning applications in prognostics and health management", IEEE Access, 7, pp.162415-162438, 2019
- [2] Lei, Y., Li, N., Guo, L., Li, N., Yan, T. and Lin, J., "Machinery health prognostics: A systematic review from data acquisition to RUL prediction", Mechanical systems and signal processing, 104, pp.799-834, 2018
- [3] Li, X., Zhang, W. and Ding, Q., "Cross-domain fault diagnosis of rolling element bearings using deep generative neural networks", IEEE Transactions on Industrial Electronics, 66(7), pp.5525-5534, 2018
- [4] Kim, Min Su, Jong Pil Yun, and PooGyeon Park., "An explainable convolutional neural network for fault diagnosis in linear motion guide", IEEE Transactions on Industrial Informatics, 17(6), pp.4036-4045, 2020
- [5] Smith, W.A. and Randall, R.B., "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study", Mechanical systems and signal processing, 64, pp.100-131, 2015