

## 스펙트로그램 이미지를 이용한 CNN 기반 자동화 기계 고장 진단 기법

강경원<sup>1</sup>, 이경민<sup>2\*</sup>

<sup>1</sup>동명대학교 정보통신소프트웨어공학과

<sup>2</sup>동명대학교 학부교양대학

### CNN-based Automatic Machine Fault Diagnosis Method Using Spectrogram Images

Kyung-Won Kang<sup>1</sup>, Kyeong-Min Lee<sup>2\*</sup>

<sup>1</sup>Dept. of Information Communication & Software Engineering, Tongmyong University

<sup>2</sup>College of General Education, Tongmyong University

**요약** 소리 기반 기계 고장 진단은 기계의 음향 방출 신호에서 비정상적인 소리를 자동으로 감지하는 것이다. 수학적 모델을 사용하는 기존의 방법은 기계 시스템의 복잡성과 잡음과 같은 비선형 요인이 존재하기 때문에 기계 고장 진단이 어려웠다. 따라서 기계 고장 진단의 문제를 딥러닝 기반 이미지 분류 문제로 해결하고자 한다. 본 논문에서 스펙트로그램 이미지를 이용한 CNN 기반 자동화 기계 고장 진단 기법을 제안한다. 제안한 방법은 기계의 결함 시 발생하는 주파수상의 특징 벡터를 효과적으로 추출하기 위해 STFT를 사용하였으며, STFT에 의해 검출된 특징 벡터들은 스펙트로그램 이미지로 변환하여 CNN을 이용해 기계의 상태별로 분류한다. 그 결과는 제안한 방법은 효과적으로 결함을 탐지할 뿐만 아니라 소리 기반의 다양한 자동 진단 시스템에도 효과적으로 활용될 수 있다.

• 주제어 : 합성곱 신경망(CNN), 스펙트로그램, 단시간 푸리에 변환(STFT), 기계 고장진단, 심층 학습, 이미지 분류

**Abstract** Sound-based machine fault diagnosis is the automatic detection of abnormal sound in the acoustic emission signals of the machines. Conventional methods of using mathematical models were difficult to diagnose machine failure due to the complexity of the industry machinery system and the existence of nonlinear factors such as noises. Therefore, we want to solve the problem of machine fault diagnosis as a deep learning-based image classification problem. In the paper, we propose a CNN-based automatic machine fault diagnosis method using Spectrogram images. The proposed method uses STFT to effectively extract feature vectors from frequencies generated by machine defects, and the feature vectors detected by STFT were converted into spectrogram images and classified by CNN by machine status. The results show that the proposed method can be effectively used not only to detect defects but also to various automatic diagnosis system based on sound.

• Key Words : CNN(Convolution Neural Network), Spectrogram, STFT(Short Time Fourier Transform), Machine fault diagnosis, Deep learning, Image Classification

Received 05 June 2020, Revised 19 July 2020, Accepted 21 September 2020

\* Corresponding Author Kyeong-Min Lee, College of General Education, Tongmyong University, 428, Sinseon-ro, Nam-gu, Busan, Korea. E-mail: kmlee2k@tu.ac.kr

## I. 서론

산업 4.0의 통합 혁신에 힘입어 현대 산업은 전통적인 제조업에서 지능형 산업으로 새로운 혁명을 경험하고 있다[1, 2]. 기계장비는 이 혁명의 가장 중요한 역할 중 하나로써 생산을 지속해서 촉진하고 경제적 이익을 증가시키기 위해 진화하고 있다. 그러나 기계 시스템이 비정상적인 상태에서 지속해서 동작하면 예정에 없던 다운타임, 경제적 손실, 심지어는 치명적인 사고와 인명피해를 초래할 것이다[3, 4]. 전통적인 마모로 야기되는 고장을 위한 기계 고장 진단 시스템은 공구와 소재가 맞닿는 곳에서 발생하는 물리적인 현상을 센서를 활용해 데이터를 수집하고 수집된 정보를 통해 고장 여부를 분석하는 방법으로, 이 분야에 관한 연구들은 오래전부터 많이 진행되고 있다[5, 6]. 특히 밀링머신과 관련된 기계 고장 진단 시스템은 공구의 마모 상태를 실시간으로 감지하거나 선행 실험을 통해 학습된 데이터를 바탕으로 공구마모 시점을 예측하는 것에 초점을 맞추고 있다. 공구의 마모나 파손을 감지하거나 발생 시점을 예측하는 것은 공정의 문제 발생과 제품의 품질 저하를 방지할 수 있을 뿐만 아니라 효율적인 공정 운영을 가능하게 한다. 그러나 실제 시스템의 복잡성과 원하지 않는 잡음과 같은 비선형적인 요인에 의해 수학적 모델 기반 접근법은 정확성 측면에서 효과적이지 않다. 따라서 기계 고장 진단을 체계적으로 수행하기 위한 방법론이 요구되었다. 이 방법은 훈련 또는 학습이라는 과정을 통해 축적된 데이터를 해석하는 기법으로 인공신경망, 퍼지 논리 및 지원 벡터 시스템과 딥러닝 기반의 CNN 시스템 등이 있다[7-10]. 또한 입력 신호 측면에서는 시간 영역[11-13], 주파수 영역[14], 시간-주파수 영역[15, 16]에서의 연구가 되고 있다.

따라서 본 논문에서는 효과적인 결함 탐지를 위해 시간-주파수간의 특징을 추출할 수 있는 스펙트로그램 이미지를 이용한 CNN (Convolutional Neural Network) 기반 자동화 기계 고장 진단 기법을 제안한다.

## II. CNN 개요

CNN은 심층 신경망(DNN: Deep Neural Network)의 한 종류로, 하나 또는 여러 개의 컨볼루션 계층(convolution layer)과 풀링 계층(pooling layer), 플랫 계

층(fully connected layer)들로 구성된다. 이와 같은 구성은 영상처리를 수행하기에 적합하게 설계된 신경망이다. 그림 1은 CNN 구조의 예를 보여준다. 컨볼루션 계층은 입력 영상의 일부를 필터와 가중합하여 다음 계층에 전달하는 역할을 한다. 필터의 크기에 따라 출력 크기는 입력 이미지보다 작아지는데 이를 보완하기 위해 입력 이미지에 제로 패딩을 해 줄 수 있다. 풀링 계층은 일종의 서브 샘플링 과정이며, 대표적으로 맥스 풀링(max pooling)을 많이 사용한다. 맥스 풀링은 필터 내에서 가장 큰 값만 전달하고 나머지는 무시하여 샘플링을 수행한다. 마지막으로 플랫 계층은 일반적인 신경망의 은닉계층과 같은 구조로 되어 있으며, 컨볼루션 과정을 통해 추출한 특징을 식별하는 역할을 한다.

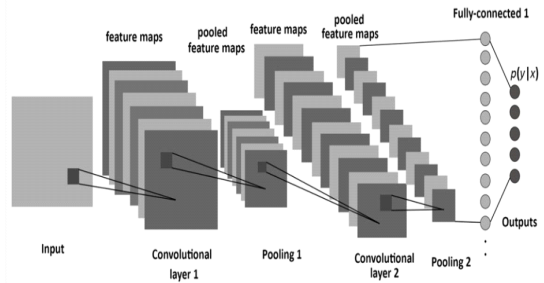


Fig. 1. Structure of CNN

## III. 제안한 고장진단 방법

본 논문에서는 밀링머신이 소재를 가공할 때 공구에서 발생하는 소리를 이용하여 공구에 결함이 있는지를 자동으로 탐지하는 방법을 제안한다. 일반적으로 소리 데이터는 주변 잡음에 영향을 많이 받기 때문에 신호의 잡음에 강인한 방법이 필요하다. 실험을 위한 데이터는 AE(Acoustic Emission) 센서를 이용하여 제조 공정에서 핸드 드릴을 이용해 공회전 상태의 공구 신호와 정상 상태의 공구와 마모된 비정상 상태의 공구를 이용해 강철판을 깎을 때 생성된 신호를 추출한 것이다. 신호의 샘플링 주파수는 44.1 kHz이며, 한 번의 처리를 위한 신호의 길이는 0.5s이다. 그림 2는 AE 센서로부터 얻은 신호의 기본적인 통계적 특성을 파악하기 위해 각각의 상태별로 1,000개의 샘플 신호들을 평균한 대표 신호의 시간영역 신호와 전력스펙트럼 밀도(Acoustic power spectral density, PSD) 신호를 나타낸

것이다. 그림 2(a)는 공회전상태의 도구 신호들을 평균한 것이며, 그림 2(b)와 그림 2(c)는 정상 상태의 공구와 마모된 비정상 상태의 공구에 의해 야기된 신호들을 나타낸 것이다. 대부분 에너지는 저주파에 집중되어 있으며, 정상 상태의 공구와 마모된 비정상 상태의 공구에 의해 야기된 전력스펙트럼 밀도의 시각적 형태는 유사함을 알 수 있다.

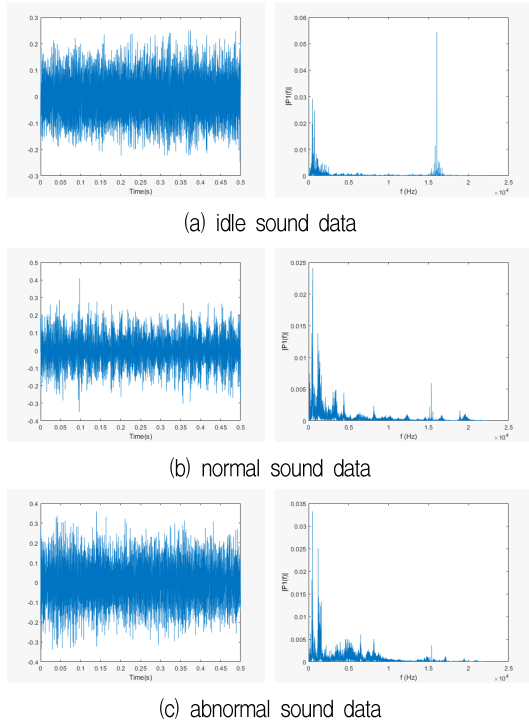


Fig. 2. Sound signals and their corresponding average power spectral density

딥러닝 기반 CNN를 이용해 특성 벡터를 추출하기 위해서는 1차 원신호에 대해 2차원 이미지화가 요구된다. 2차원 이미지를 구현하기 위해서 제안한 방법은 STFT(Short Time Fourier Transform)를 기반으로 하는 스펙트로그램(Spectrogram)을 사용한다.

AE 센서 데이터에 대한 스펙트로그램은 그림 3과 같다. 1차원 신호를 일정한 길이의 조각으로 나눈 다음, 이 조각에 대하여 푸리에 변환을 적용하여 수평축에는 조각의 시간 정보를 나타내고 수직축에는 주파수 성분의 크기를 데시벨 단위로 표현한 2차원 이미지들이다. 이 스펙트로그램은 시간-주파수 분석으로 비선형 신호의 특성을 추출하기 위해 적합하며, 노이즈가

있는 데이터를 분석에 유용한 도구이다.

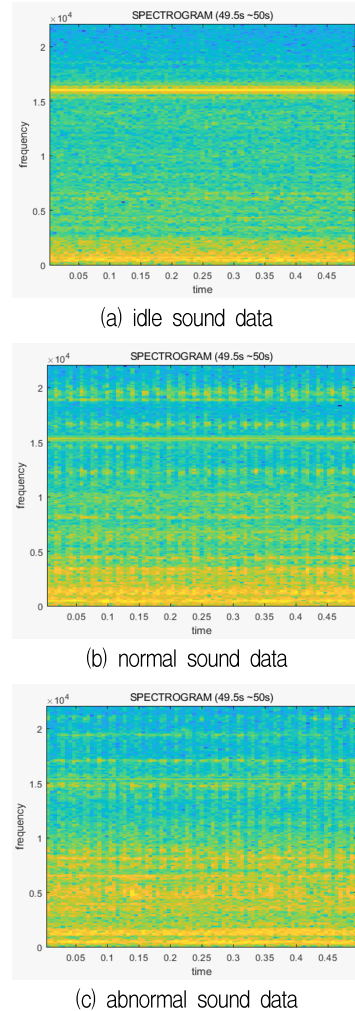


Fig. 3. Spectrogram of the Acoustic Emission

#### IV. 실험 결과 및 고찰

본 실험에서는 마이크를 사용하여 밀링머신의 공회전 상태에서 생성된 소리와 정상 상태의 공구와 마모된 비정상 상태의 공구를 사용하여 강철판을 깎을 때 생성된 소리를 각각 500s씩 수집하였다. 수집된 소리 데이터의 샘플링 주파수는 44.1 kHz이며, 길이는 0.5s이다. 데이터셋은 각각 상태별로 300개의 데이터로 구성된 총 900개의 데이터로 구성하였다.

실험은 Matlab을 이용하였으며, 입력 스펙트로그램의 크기는 420x560이며, 구현한 CNN은 구조는 그림 4와 같이 총 3개의 컨볼루션 계층과 2개의 풀랫 계층으

로 구성하였다. 각각의 컨볼루션 계층은 컨볼루션, 배치 정규화(batch normalization), ReLU, 맥스 풀링으로 구성되어 있다. 그리고 마지막 풀넷 계층인 fully connected layer는 softmax를 이용해 현재 밀링머신이 공회전 상태인지 정상 상태인지 비정상 상태인지를 확률로 나타낸다. CNN 네트워크 학습을 위한 주요 옵션은 표 1과 같다. CNN 네트워크를 위한 주요 파라미터는 표 2와 같으며, 학습은 Adam 옵티마이저를 사용해서 학습시켰으며, 배치 사이즈는 10으로 설정하였으며, 총 10 epoch까지 학습을 진행하였다.

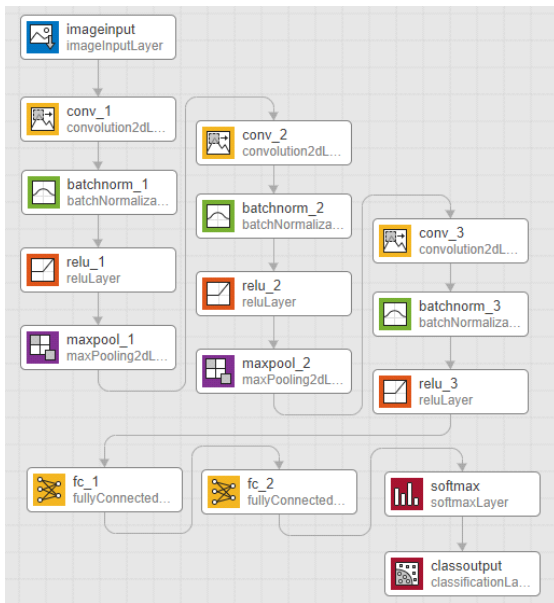


Fig. 4. CNN Architecture

Table 1. Learning Options for CNN

Field	Value
Optimizer	adam optimizer
Mini-batch size	10
Shuffling	every-epoch
Epoch	10
Learning rate	0.001

그림 5는 본 논문에서 적용된 CNN의 정확도와 손실률을 나타낸 것으로, 입력 데이터의 70%는 훈련 셋(training set)으로 사용하였으며, 30%는 검증 validation set)으로 사용하였다. 제한된 입력 데이터에 대해서는

정확도가 100% 수렴하였으며, 반복 테스트에서도 거의 100% 수렴함을 알 수 있었다.

Table 2. Key Parameters for CNN Network

Network	Parameters
convolution layer 1	8개 3x3x3 컨볼루션
batch normalization layer 1	채널 8개
ReLU layer 1	-
maxpooling layer 1	2x2 맥스 풀링
convolution layer 2	16개 3x3x8 컨볼루션
batch normalization layer 2	채널 16개
ReLU layer 2	-
maxpooling layer 2	2x2 최댓값 풀링
convolution layer 3	32개 3x3x16 컨볼루션
batch normalization layer 3	채널 32개
ReLU layer 3	-
fully-connected layer 1	128
fully-connected layer 2	3
softmax	-
classification layer	3 classes

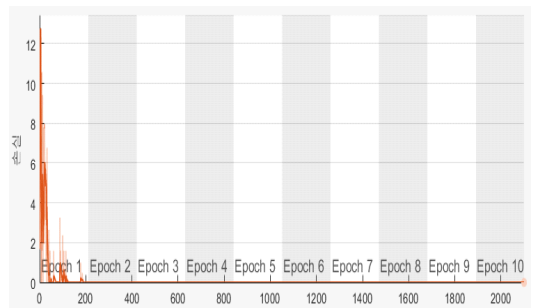
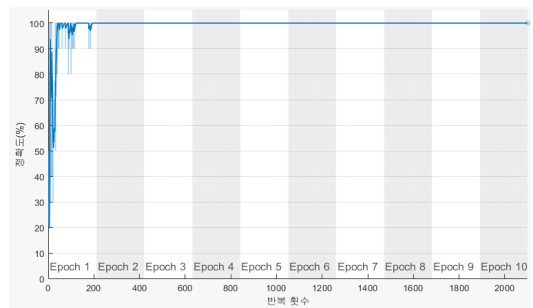


Fig. 5. Learning Result

제안한 밀링머신의 공회전상태 및 정상 상태, 비정상 상태의 데이터 900개의 대해 10번의 반복 수행하였을 경우 혼돈 행렬은 그림 6과 같다. 수행 결과 9번은 그림 6의 (a)와 같이 검출률이 100%이며, 한 번의 경우는 그림 6 (b)와 같이 검출률이 99.9%로 평균 99.99%를 나타냈었다. 정상 상태의 공구와 마모된 비정상 상태의 공구에 의해 야기된 신호의 경우 상호 유사한 점이 존재하지만, CNN의 특성상 거의 정확하게 분류함을 알 수 있었다.

Output Class \ Target Class	abnormal	idle	normal	
abnormal	300 33.3%	0 0.0%	0 0.0%	100% 0.0%
idle	0 0.0%	300 33.3%	0 0.0%	100% 0.0%
normal	0 0.0%	0 0.0%	300 33.3%	100% 0.0%
	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%

(a)

Output Class \ Target Class	abnormal	idle	normal	
abnormal	300 33.3%	0 0.0%	0 0.0%	100% 0.0%
idle	0 0.0%	300 33.3%	1 0.1%	99.7% 0.3%
normal	0 0.0%	0 0.0%	299 33.2%	100% 0.0%
	100% 0.0%	100% 0.0%	99.7% 0.3%	99.9% 0.1%

(b)

Fig. 6. Confusion matrix by ten iterations

표 3은 기존의 방법과 제안한 방법의 분류 정확도를 비교한 것으로 제안한 방법의 우수성을 확인할 수 있었다.

Table 3. Summary of the result of the conventional and proposed method

	P.K. Kankar et al.[13]	K.Y. Lee et al.[16]	Proposed Method
feature extraction method	time serial + statistical	DWT + PCA	spectrogram
Classifier	ANN	ANN	CNN
Accuracy	64.20%	94.52%	99.99%

## V. 결론

본 논문에서 제안한 스펙트로그램 이미지를 이용한 CNN 기반 자동화 기계 고장 진단 기법은 소리를 이용한 특징 벡터 인식 기반 공구의 결합 진단에서 우수한 성능을 제시하였다. CNN의 우수한 장점과 소리의 특성을 효과적으로 추출하기 위해 STFT를 이용한 스펙트로그램 이미지를 활용함으로써 분류기의 높은 성능을 나타낼 수 있음을 확인하였다. 특히, 정상 상태의 스펙트럼과 비정상 상태의 스펙트럼의 시각적 유사성이 존재하지만 CNN의 우수한 성능 때문에 정확히 분류할 수 있었다.

제시한 방법은 다른 고장 진단 시스템에서 사용될 수 있으며, 공구의 정상 및 비정상적인 마모된 소리를 분리할 수 있으면 진동 및 기타 음향 방출 데이터에도 활용할 수 있을 것이다.

## REFERENCES

- [1] J. Wan, S. Tang, D. Li, S. Wang, C. Liu, H. Abbas, A. V. Vasilakos, "A manufacturing big data solution for active preventive maintenance," *IEEE Transactions on Industrial Electronics*, vol. 13, no. 4, pp. 2039-2047, 2017.
- [2] Y. Lei, F. Jia, J. Lin, S. Xing, S. X. Ding, "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3137-3147, 2016.
- [3] J. Jiao, M. Zhao, J. Lin, K. Liang, "Hierarchical discriminating sparse coding for weak fault feature extraction of rolling bearings," *Reliability Engineering*

& System Safety, vol. 184, pp. 41-54, 2019.

[4] M. Zhao, X. Jia, "A novel strategy for signal denoising using reweighted SVD and its applications to weak fault feature enhancement of rotating machinery," Mechanical Systems and Signal Processing, vol. 94, pp. 129-147, 2017.

[5] B. Robert, J. Antoni, "Rolling element bearing diagnostics," Mechanical System and Signal Processing, vol. 25, no. 2, pp. 485-520, 2011.

[6] T. H. Ioutas, G. Sotiriades, I. kalaitzoglou, V. Kostopouls, "Condition monitoring of a single stage gearbox with artificially induced gear cracks utilizing on-line vibration and acoustic emission measurements," Applied Acoustics, vol. 70, issue 9, pp. 1148-1159, 2009.

[7] C. Chen, B. Zhang, G. Vachtsevanos, M. Orchard, "Machine condition prediction based on adaptive neuro-fuzzy and high-order particle filtering," IEEE Trans. Industrial Electronics, vol. 58, issue 9, pp. 4353-4364, 2011.

[8] Y. S. Wang, Q. H. Ma, Q. Zhu, L. Zhao, "An intelligent approach for engine fault diagnosis based on hibert-huang transform and support vector machine," Applied Acoustics, vol. 75, pp. 1-9, 2014.

[9] M. Gan, C. Wang, C. A. Zhu, "Construction of hierarchical diagnosis network based on deep learning and its application in te fault pattern recognition of rolling element bearings," Mechanical Systems and Signal Processing, vol. 72-73, no. 11, pp. 7067-7075.

[10] W. Sun, R. Zhao, R. Yan, S. Shao, X. Chen, "Convolutional discriminative feature learning for induction motor fault diagnosis," IEEE Transactions on Industrial Informatics, vol. 13, no. 3, pp. 1350-1359, 2017.

[11] B. Sreejith, A. K. Verma, A. Srividya, "Fault diagnosis of rolling element bearing using time-domain features and neural networks," IEEE Region 10 and the Third International Conference on Industrial and Information Systems, pp. 1-6, 2009.

[12] T. Ince, S. Kiranyaz, L. Eren, M. Askar, M. Gabbouj, "Real-time motor fault detection by 1D convolutional neural networks," IEEE Transactions on Industrial Electronics, vol. 63, no. 11, pp. 7067-7075, 2016.

[13] P. K. Kankar, S. C. Sharma, S. P. harsha, "Fault

diagnosis of ball bearings using machine learning methods," Expert Systems with Applications, vol. 38, issue 3, pp. 1876-1886, 2011.

[14] D. Z. Li, W. Wang, F. Ismail, "An enhanced bispectrum technique with auxiliary frequency injection for induction motor health condition monitoring," IEEE Transactions on Instrumentation & Measurement, vol. 64, no. 10, pp. 2679-2687, 2015.

[15] J. D. Zheng, H. Y. Pan, X. L. Qi, X. Q. Zhang, Q. Y. Liu, "Enhanced empirical wavelet transform based time-frequency analysis and its application to rolling bearing fault diagnosis," Acta Electronica Sinica, vol. 46, no. 2, pp. 358-364, 2018.

[16] K. M. Lee, C. Vununu, K. S. Moon, S. H. Lee, K. R. Kwon, "Automatic machine fault diagnosis system using discrete wavelet transform and machine learning," Journal of Korea Multimedia Society, vol. 20, no. 8, pp. 1299-1311, 2017.

---

## 저자 소개

---

### 강 경 원 (Kyung-Won Kang)



1996년 2월 : 부경대학교  
전자공학과(공학사)  
1998년 2월 : 부경대학교  
전자공학과(공학석사)  
2002년 2월 : 부경대학교  
전자공학과(공학박사)  
2006년~2014년 : ㈜휴케스트

2018년~현재 : 동명대학교 정보통신학과 조교수  
관심 분야 : 영상 신호 처리, 머신러닝

### 이 경 민 (Kyeong-Min Lee)



2000년 : 동의대학교  
컴퓨터공학과(공학사)  
2002년 : 동아대학교  
전자공학과(공학석사)  
2019년 : 부경대학교  
IT융합응용공학과(공학박사)  
2010년~현재 : 동명대학교

학부교양대학 교수  
관심 분야 : 멀티미디어 신호처리, 머신러닝