

# 딥 러닝 기반의 이기종 무선 신호 구분을 위한 데이터 수집 효율화 기법

## An Efficient Data Collection Method for Deep Learning-based Wireless Signal Identification in Unlicensed Spectrum

최 재 혁<sup>\*\*</sup>

Jaehyuk Choi<sup>\*\*</sup>

### Abstract

Recently, there have been many research efforts based on data-based deep learning technologies to deal with the interference problem between heterogeneous wireless communication devices in unlicensed frequency bands. However, existing approaches are commonly based on the use of complex neural network models, which require high computational power, limiting their efficiency in resource-constrained network interfaces and Internet of Things (IoT) devices. In this study, we address the problem of classifying heterogeneous wireless technologies including Wi-Fi and ZigBee in unlicensed spectrum bands. We focus on a data-driven approach that employs a supervised-learning method that uses received signal strength indicator (RSSI) data to train Deep Convolutional Neural Networks (CNNs). We propose a simple measurement methodology for collecting RSSI training data which preserves temporal and spectral properties of the target signal. Real experimental results using an open-source 2.4 GHz wireless development platform Ubertooth show that the proposed sampling method maintains the same accuracy with only a 10% level of sampling data for the same neural network architecture.

### 요 약

최근 데이터 기반의 딥러닝 기술을 적용하여 비면허 대역의 다양한 통신 신호를 분류하는 연구가 활발히 수행되고 있다. 하지만, 복잡한 신경망 모델 사용을 기반으로 이뤄진 이러한 접근법은 높은 연산 능력을 필요로 하게 되어, 자원 제약적인 무선 인터페이스 및 사물인터넷(Internet of Things) 장비에서는 사용이 제약된다. 본 연구에서는 비면허 대역의 무선 이기종 기술을 인지하기 위한 데이터 기반의 접근 방법을 살펴보고, 신호의 특징 추출 및 데이터화의 효율화 문제를 다룬다. 구체적으로, 비면허 대역의 다른 종류의 무선 통신 기술을 구분하기 위해 수신 신호 강도 측정을 기반으로 한 시계열 데이터를 이용해 합성곱 신경망(Convolutional Neural Network, CNN) 모델을 학습시켜 신호를 분류하는 방법을 살펴본다. 이 과정에서 동일한 구조의 신경망 모델의 경량화를 위한 효율적 신호의 시계열 데이터 정보 수집시 주파수 대역의 특징을 함께 특징화하는 방법을 제안하고, 그 효과를 평가한다. Bluetooth 호환의 Ubertooth 장비를 이용한 실측 기반의 실험 결과는 제안된 샘플링 기법이 동일한 신경망에 대해서 10% 수준의 샘플링 데이터 이용만으로도 동일한 정확도를 유지함을 보여준다.

*Key words : Wireless identification, RSSI sampling, Unlicensed spectrum, CNN, Feature selection*

\* School of Computing, Gachon University

★ Corresponding author

E-mail : jchoi@gachon.ac.kr, Tel : +82-31-750-8657

※This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT) (No. NRF-2020R1A2C1013308).

Manuscript received Feb. 14, 2022; revised Mar. 20, 2022; accepted Mar. 22, 2022.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## I. 서론

Wi-Fi, Bluetooth, ZigBee 등의 통신 기술은 ISM (Industry-Science-Medical) 대역으로 대표되는 비면허 대역(licensed spectrum)을 사용하는 대중화된 무선 통신 기술이다. 높은 전송 속도, 낮은 전력 소모 등의 많은 장점으로 인해 일상생활, 산업 및 공공 분야에서의 해당 통신 기술에 대한 의존도와 활용도가 급격히 증가해 왔고, 이로 인해 ISM 비면허 대역은 이미 포화상태에 이르렀다. 환경적 변화와 주변 신호의 간섭에 쉽게 영향을 받는 특성으로 인해 비면허 대역에 공존하는 이기종 통신 기술 간의 상호 간섭 문제는 성능에 매우 큰 영향을 미치게 된다[1, 2].

지난 십여 년간 통신/네트워크 분야에서 이러한 이기종 통신 기술간 상호 간섭 문제를 다루는 많은 연구가 수행되었고, 주요 연구 분야 중 하나로 실시간으로 주변의 이기종 통신 신호 탐지하는 문제가 주목받아 왔다. 이는 주변 상황에 맞추어 간섭 회피 등에 활용할 수 있기 때문이다. 하지만, 이기종 신호 탐지 문제는 무선 물리계층의 다른 신호 형상 사용, 상호 인지 및 협력적 제어 불가능 등의 이유로 매우 도전적인 요소를 내포하고 있다 [3].

최근, 이기종 무선 기술 공존 문제에 대해 데이터 기반의 딥러닝 기술을 적용한 연구가 활발해지고 있다[3]. 목표 주파수 대역에서 수집된 수많은 데이터를 기반으로 높은 정확도를 보이고 있으나, 복잡한 신경망 모델을 기반으로 이뤄진 이러한 접근법은 높은 연산 능력을 필요로 하게 되어, IoT 장비 등 마이크로프로세서 기반의 자원 제약적인 통신 장비에서는 사용이 제약된다.

본 연구에서는 합성곱 신경망(Convolutional Neural Network, CNN) 모델을 학습시켜 여러 이기종 신호를 분류하는 방법을 살펴보고, CNN구조의 경량화를 위한 효율적 신호의 샘플링(sampling) 기법 및 시계열 데이터 생성 방법을 제안한다. 실측 기반의 실험 결과는 제안된 샘플링 기법이 동일한 신경망에 대해서 10% 수준의 샘플링 데이터 이용만으로도 동일한 정확도를 유지함을 보여준다.

## II. 실험 설계

그림 1은 본 연구에서 수행한 RF 신호(radio frequency signal) 수신 세기 측정 기반의 스펙트럼 센싱 및 데이터 수집 과정의 전체 구성도를 나타낸다. (1) 첫 번째 단계로

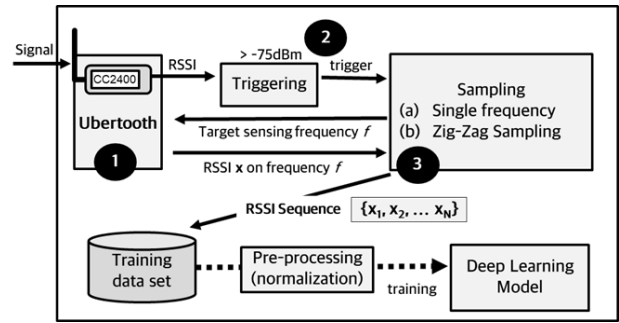


Fig. 1. Overview of our RF signal strength measurement and data collection process.

그림 1. RF 신호 세기 측정 및 데이터 수집, 신경망 훈련의 전체 과정

블루투스 기능의 Ubertooth [4]를 이용하여,  $\tau=128\mu s$  간 마다 대상 주파수/채널의 수신 신호 세기(RSSI)를 지속 측정, (2) 측정된 RSSI의 크기가 임계값 (실험에서는  $-75\text{dBm}$  이상)이 될 경우 (Triggering 컴포넌트) RSSI 시계열 데이터 생성 시작, (3) Sampling 컴포넌트에서  $N$ 개의 RSSI값  $x_i$ 를 수집하여 RSSI Sequence= $\{x_1, x_2, \dots, x_N\}$ 을 생성한다.

이러한 과정을 통해 Wi-Fi beacon, Wi-Fi video stream, 802.15.4 ZigBee 3가지 통신 기술에 대해 III장에서 설명하는 2개의 Sampling 방법을 통해 각각 43,000개 데이터를 수집하였고, CNN 신경망 모델의 훈련에 사용하였다. 이를 위해 그림 2의 2.4GHz 대역에서 Wi-Fi의 경우는 6번, ZigBee는 18번 채널인 2440MHz를 사용하여 전송을 수행하고, 모든 전송에 대해 Ubertooth는 2m거리에서 RSSI를 측정하였다.

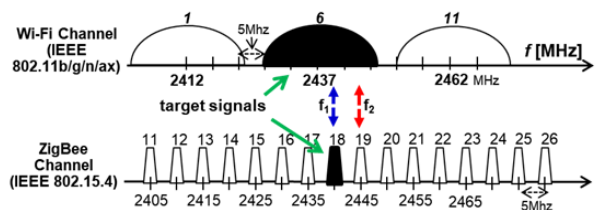


Fig. 2. IEEE 802.11 (Wi-Fi) and IEEE 802.15.4 WPAN (ZigBee) frequency in 2.4GHz unlicensed band.

그림 2. 2.4GHz 비면허 대역에서의 Wi-Fi, ZigBee 채널

기존의 일반적 접근법 [3]의 경우, 그림 1의 (3)의 시계열 데이터 생성을 위한 RSSI 측정은 고정된 단일 주파수를 이용하였다. 그림 3은 이러한 고정 주파수를 이용한 측정을 통해 Wi-Fi beacon, Wi-Fi video streaming, ZigBee에 대해 RSSI Sequence(시계열 데이터)를 도식화한 것이다.

이 결과에서 세 목표 신호별로 다른 형상을 확인할 수 있었고, 이는 블루투스 수신 모듈로 간단한 RF 신호 세기 (RSSI) 측정을 통해 데이터를 수집하고 이용하면 여러 이기종 신호 구분이 가능할 수 있다는 것을 보여준다. III장에서는 수집한 데이터를 이용하여 훈련시킬 딥러닝 모델과 효율적인 데이터 특성 추출 방법에 대해 논의한다.

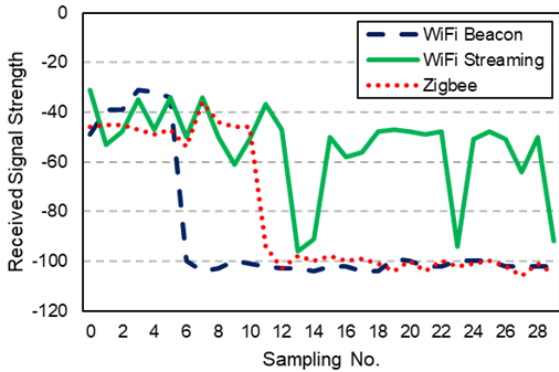


Fig. 3. RSSI patterns of three different cases: Wi-Fi beacon, Wi-Fi video streaming, and 802.15.4 ZigBee.

그림 3. Wi-Fi beacon, Wi-Fi video steaming, ZigBee의 3개 RF신호에 대한 RSSI 패턴

### III. 딥러닝 모델 및 Zig-Zag 샘플링 기반 데이터 특성 추출 알고리즘

본 장에서는 이기종 무선 신호 분류를 위한 신경망 모델 훈련을 위한 효율적 데이터 특성 선택(feature extraction/selection) 방법을 다룬다. 앞서 언급한 바와 같이 [3]을 비롯한 기존의 많은 연구는 대규모 시계열 데이터를 합성곱 신경망(CNN), 순환 신경망(RNN)등의 학습에 활용하여 분류를 수행하는 방식이다. 하지만, 이 방식의 경우, 고정된 주파수(예로, 그림 2에서  $f_1$ 로 고정)에서 RSSI 수집을 하였기 때문에, 사용된 특성(feature)에는 그림 3에서 보이듯이 해당 신호의 주파수 대역/스펙트럼 특성(spectral signature)는 제거되고 시간 축 정보(temporal signature)만 포함된다.

이에, 그림 4와 같은 동일한 무선 신호 분류를 위한 이기종 신호들의 주파수 점유 특성을 고려한 RSSI 수집 방법(Sampling method)과 이를 통한 데이터 feature를 이용한 데이터 생성 방안을 제안한다. 기존 고정된 주파수를 이용하는 대신, 두 개의 주파수  $f_1, f_2$ (그림 2 참조)를 Zig-Zag 방식으로 번갈아가면서 RSSI 값을  $N$ 개 측정하여 데이터를 형성한다. 즉, 크기  $N$ 의 한 데이터 중

홀수 번호 feature는  $N/2$ 개는  $f_1$ 에서 측정된 RSSI값이고, 짝수 번호 feature는  $N/2$ 개는  $f_2$ 에서 측정되게 된다. 두 주파수간 차이로  $|f_1 - f_2| = 5\text{MHz}$ 를 이용하는 경우, 20~40MHz의 대역폭을 사용하는 Wi-Fi는 그림4에서 보듯  $f_1, f_2$  두 주파수 모두에서 유사한 RSSI값이 추출된다. 반면, 그림 5에서 볼 수 있듯이, ZigBee는 대역폭이 2MHz로 좁기 때문에  $f_1, f_2$  중 하나의 주파수에서만 높은 RSSI값이 추출된다. Wi-Fi Beacon과 Wi-Fi Streaming은 전송 시간 및 전송 빈도의 차이가 존재하고 이는 수집된 RSSI sequence에 반영된다. 그림 5는( $f_1, f_2$ ) = {2440, 2445}MHz를 이용한 측정(그외는 그림 3 측정과정과 동일 조건)한 결과이다.

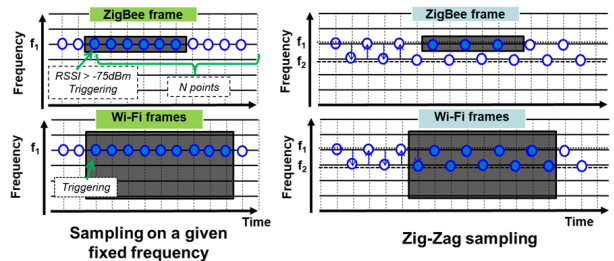


Fig. 4. Zig-Zag sampling method alternating between two frequencies  $f_1$  and  $f_2$ .

그림 4. 지그재그 RSSI 수집 방식: 두 개의 주파수를 차례로 번갈아 이용하며 N개의 RSSI를 수집

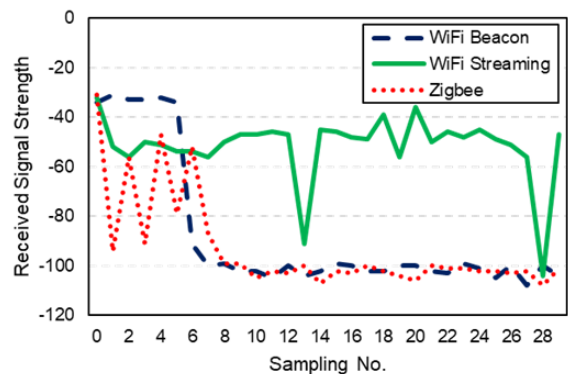


Fig. 5. RSSI patterns with Zig-Zag sampling.

그림 5. Zig-Zag 샘플링 사용 시의 변화된 RSSI 패턴

앞서 설명한 방식으로  $N$ 개의 RSSI를 수집하여 데이터를 생성하고 수집된 데이터를 그림 6에 도식한 두 가지 다른 CNN 구조의 모델의 학습에 사용하였다. 입력층(input layer)의 크기를  $N=100, 50, 20, 10$  형태로 100에서 10까지 감소시키며 4가지 네트워크를 구성하였고, 다음 장에서 각 모델별로 Zig-Zag 샘플링으로 수집하여 주파수 특성을 반영한 효과를 검증한다.

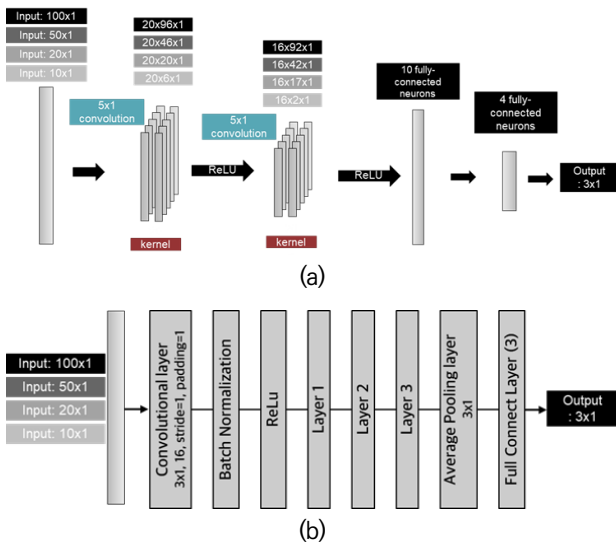


Fig. 6. CNN architecture: (a) simple CNN, (b) ResNet [5].  
그림 6. 합성곱 신경망 구조: (a) 기본 CNN, (b) ResNet [5]

#### IV. 성능 평가

본 장에서는 제안한 Zig-Zag 샘플링 기법을 통한 데이터 수집 방식으로 그림 6의 두 CNN구조를 이용하였을 경우에 RSSI Sequence 크기  $N$ , 즉 입력층 크기에 따른 이기종 신호 구분의 정확도를 평가했다. 정확도는 3개의 다른 전송을 정확히 분류한 비율을 나타낸다. 총 43,000개의 데이터를  $k$ -fold 교차검증( $k$ -fold cross validation)으로  $k=5$ 에 대해 평균을 구하였다.

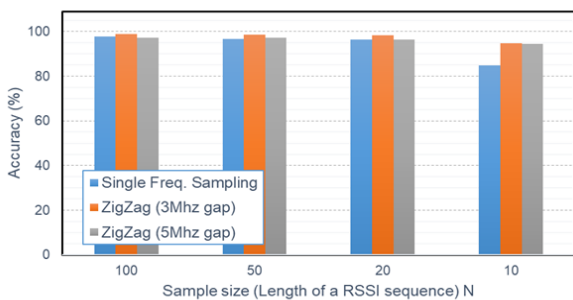


Fig. 7. Classification accuracy using ResNet [5] with Zig-Zag sampling for data sizes  $N=100, 50, 20, 10$ .  
그림 7. Zig-Zag 샘플링의 데이터 길이  $N=100, 50, 20, 10$ 에 대한 ResNet [5]의 분류 정확도

그림 7, 8은 각각 두 CNN 구조에 대해 단일 주파수 샘플링(2440MHz), Zig-Zag 샘플링을 이용할 경우의 정확도를 나타낸다. Zig-Zag 샘플링은 두 종류의 주파수 차이값 3MHz (2440-2443MHz), 5MHz (2440-2445 MHz)을 이용해 비교하였다.

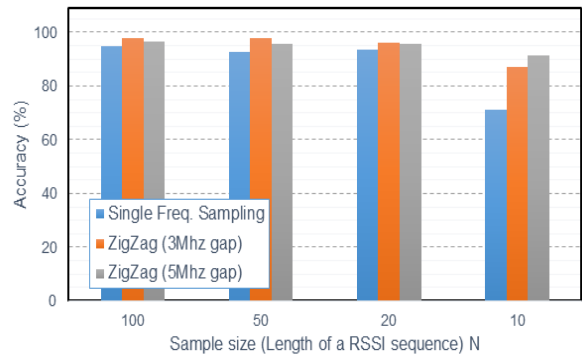


Fig. 8. Classification accuracy using simple CNN with Zig-Zag sampling for data sizes  $N=100, 50, 20, 10$ .  
그림 8. Zig-Zag 샘플링의 데이터 길이  $N=100, 50, 20, 10$ 에 대한 기본 CNN의 분류 정확도

그림 7 ResNet[5]의 경우  $N=100$ 을 이용할 경우 세 방법이 각각 97.7%, 98.9%, 97.2%의 높은 정확도로 신호를 구분했다. 그림 8의 simple CNN의 경우도 93%, 98.3%, 96.6%의 정확도를 나타냈다. 샘플 크기를  $N=20$ 으로 줄여도 두 CNN 모델의 정확도에 큰 변화가 없이 높은 정확도로 신호를 분류하는 결과도 보여주고 있다. 이는 그림 3, 5를 볼 수 있는 ZigBee, Wi-Fi Beacon 신호별 전송 시간의 차이에 따라 시간적 특성이 20개의 RSSI값에 내제(embedding)됨으로, 정확한 구분이 가능한 것으로 파악된다. 기존 연구인 [3]의 경우는 입력계층 크기를 210개로 고정해서 사용하는데, 210개는 매우 보수적인 값으로 판단된다.

한편,  $\Lambda$ 의 크기를 10으로 줄이게 되면 동일한 CNN구조에 대해 단일 주파수 샘플링 기법의 정확도는 84.9% (ResNet), 69.2%(simple CNN)로 급격히 저하되는 것을 볼 수 있다. 이와는 다르게, Zig-Zag Sampling을 사용한 데이터의 경우는 평균 94.6%(ResNet), 90.1% (simple CNN)로 높은 정확도를 유지했다. Shortcut (skip) connection을 사용하지 않은 플레인(plain) 구조의 simple CNN에 비해 ResNet의 정확도가 높게 나타났다.

이는 간단한 Zig-Zag 샘플링으로 데이터 특징 추출시 신호의 주파수의 특성까지 고려되었음을 나타내고, 이러한 데이터의 처리 효율화는 정확도를 유지하며 신경망의 경량화를 도모함을 함축하고 있다. 실제, 위의 접근으로  $N=100$  사용할 경우 대비,  $N=50, 20, 10$ 의 사용시, 모델 사용 연산 시간이 각각 80.7%, 57.8%, 50.6% 정도 필요한 것으로 모델 경량화에 큰 효과가 있음을 나타낸다.

## V. 결론

본 연구를 통해 시간 및 주파수 대역의 무선 신호를 특성을 데이터에 반영될 수 있게 하는 간단한 데이터 수집 및 추출 방법을 제안했다. 성능 평가 결과는 데이터의 선택의 효율화를 통해 정확도는 유지하며 CNN 구조의 경량화가 가능함을 보여 주었다. 이는 딥러닝 분류 기능을 연산 기능이 제약적인 사물인터넷 장비 등에 on-device 형태로 적용하는데 중요한 의미를 갖는다. 향후 연구를 통해 시계열 데이터에 특화되고 복잡도가 높지 않은 RNN(Recurrent Neural Network), LSTM(Long Short-Term Memory) 모델을 이용한 이기종 신호 분류 문제를 다룰 계획이다. 또한, 제안하는 Zig-Zag 샘플링 기법의 주파수 선택 알고리즘을 개선하여 성능향상도도모할 계획이다.

## References

- [1] S. Siddhartha, S. Tschimben, and K. Gifford. "Towards Enhancing Spectrum Sensing: Signal Classification Using Autoencoders," *IEEE Access*, no.9, pp. 82288-82299, 2021. DOI: 10.1109/ACCESS.2021.3087113
- [2] G. Simone, A. Mahmood, and M. Gidlund. "Real-time interference identification via supervised learning: Embedding coexistence awareness in IoT devices," *IEEE Access*, no.7 pp.835-850, 2018. DOI: 10.1109/ACCESS.2018.2885893
- [3] Y. Su, et al. "Interference source identification for IEEE 802.15. 4 wireless sensor networks using deep learning," in *Proc. IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp.1-7, 2018. DOI: 10.1109/PIMRC.2018.8580857
- [4] M. Ossmann, D. Spill. "Project Ubetooth," <https://greatscottgadgets.com/ubetoothone/>
- [5] H. Kaiming, et al. "Deep residual learning for image recognition," in *Proc. IEEE conference on computer vision and pattern recognition*, pp. 770-778, 2016. DOI: 10.1109/CVPR.2016.90

## BIOGRAPHY

### Jaehyuk Choi (Member)



2003 : BS degree in Material Science and Engineering, Seoul National University.  
2008 : PhD degree in Computer Science and Engineering, Seoul National University.

2008~2011 : Postdoctoral Researcher, University of Michigan

2017~2018 : Visiting Scholar, UC San Diego

2011~2022 : Associate Professor, Gachon University