

파킨슨병 진단 및 치료 보조를 위한 뇌전도 신호 분류 모델의 입력데이터 유형과 딥러닝 모델 아키텍처에 따른 분류 성능에 관한 연구

김시온, 김진모, 최지웅
대구경북과학기술원

slon11@dgist.ac.kr, jmkim@dgist.ac.kr, jwchoi@dgist.ac.kr

EEG classification performance on diagnosis and treatment of Parkinson's disease based on input types and deep learning model architectures

Sion Kim, Jinmo Kim, Ji-Woong Choi
Daegu Gyeongbuk Institute of Science and Technology (DGIST)

요약

파킨슨병의 EEG 신호를 분류하는 모델은 파킨슨병의 진단과 치료를 보조하는 데 사용될 수 있다. 이러한 분류 모델의 성능은 입력데이터와 모델 아키텍처의 유형에 영향을 받는다. 그러나 아직 파킨슨병 EEG 신호 분류 작업에 대해서 어떠한 입력과 모델의 조합이 가장 적합한지 비교한 연구는 부족하다. 따라서, 본 논문에서는 동일한 전처리 과정을 거치되, 서로 다른 시간 및 주파수 해상도를 갖는 입력데이터에 대해서 여러 모델 아키텍처 (CNN, RNN, ANN)를 적용하고, 모델 최적화 과정을 동일하게 통제함으로써 성능을 비교하였다. 결과적으로, 시간 및 주파수 해상도가 동시에 높은 입력데이터와 CNN의 조합이 가장 뛰어난 분류 정확도를 가짐을 확인하였다.

I. 서론

파킨슨병의 EEG (Electroencephalography) 신호 분류 모델은 파킨슨병의 진단과 치료를 보조하는데 활용될 수 있으며, 이를 위한 분류 모델은 다양한 입력데이터 유형과 인공지능 모델 아키텍처로 연구되고 있다 [1].

EEG 신호 분류 모델의 성능은 학습에 사용되는 데이터셋과 입력데이터의 유형 그리고 인공지능 모델 아키텍처에 크게 의존한다 [2]. 그러나 데이터셋마다 여러 입력데이터 유형과 모델 아키텍처의 조합을 실험하는 것은 많은 계산 비용과 시간이 드는 일이며, 모든 조합을 실험하는 것은 어려운 일이다. 이에 따라 입력데이터 유형과 모델 아키텍처의 조합 선정은 딥러닝 분야에서 통용되는 이미지와 CNN (Convolutional Neural Network)의 조합이나 시계열과 RNN (Recurrent Neural Network)의 조합과 같은 형태를 이용하거나, 기존 연구의 결과를 통해 선정하는 것이 일반적이다 [2].

하지만 이런 접근을 통한 입력데이터와 모델의 유형 선정에는 한계점이 있다. 딥러닝에서 통용되는 조합의 결과가 특정 도메인의 특성으로 인해 반드시 가장 좋은 성능으로는 연결되지 않을 수 있기 때문이다 [2]. 파킨슨병 환자의 EEG 데이터 분석에 대해 다양한 모델 아키텍처 간의 성능을 비교[3]하기도 하였으나, 이를 서로 다른 시간 및 주파수 해상도를 갖도록 하는 입력데이터 전처리 방식과 연결 지어서 분석한 사례가 부족하다.

따라서, 본 논문에서는 정상대조군을 비롯해 파킨슨병 환자의 약물 복용 유무에 따라 측정된 EEG 데이터에 대해서, 서로 다른 시간 및 주파수 해상도를 갖는 입력의 형태와 대표적인 딥러닝 모델 아키텍처의 유형에 따른 파킨슨병 환자의 질환과 약물 복용 유무에 대한 분류 성능을 비교하고자 한다.

II. 본론

파킨슨병 EEG 데이터 분석에 대한 입력데이터 유형으로는 각기 다른 시간 및 주파수 해상도를 갖는 시계열, PSD (Power Spectral Density), spectrogram, 그리고 scalogram 를 비교한다. 입력데이터의 PSD 를 얻기 위해 Welch's method 를 이용하며, 주기도 (Periodogram) 계산을 위한 윈도우 길이를 1 초 그리고 윈도우의 겹치는 구간을 0.5 초로 설정하였다. Spectrogram 을 위해서는 short-time Fourier transform 을 이용하고 푸리에 변환을 위한 각 윈도우 길이를 1 초 그리고 윈도우의 겹치는 구간을 0.75 초로 설정하여 계산하였다. Scalogram 은 Morlet wavelet transform 을 이용해 계산되었으며, 각 입력데이터 유형의 시간 및 주파수 해상도의 차이는 표 1 에 표기하였다. 딥러닝 모델 아키텍처에 따른 비교로는 파킨슨병 분류 모델에 주로 이용되는 CNN, RNN, 그리고 ANN (Artificial Neural Network) 아키텍처를 비교하였다 [1, 2].

모델의 분류 평가를 위한 파킨슨병 EEG 데이터로는 샘플링 주파수 512 Hz 로 측정된 9 명의 HC (Healthy Control)과 11 명의 파킨슨병 환자에 대해 3 분 동안 측정된 데이터를 이용하였으며, 11 명의 파킨슨병 환자에 대해서 약물 복용 시와 아닌 때를 각각 나누어 측정된 데이터를 이용하였다. 파킨슨병의 주요 증상인 rigidity 나 tremor 는 운동과 관련되었으므로, 총 31 개 전극 데이터 중 motor cortex 의 신호를 반영하는 C3 전극 신호를 구분하여 사용하였다. 각 3 분 길이의 데이터는 5 초 길이의 segment 로 나뉘어 분류 데이터셋 구성에 사용되었다. 데이터셋은 총 3 가지 클래스로 구성되었으며, 각각 HC (236 개), PD(Parkinson's Disease) - 약물 복용 (352 개), PD-약물 미복용 (313 개)으로 구성되었다.

표 1. 입력데이터 유형과 모델 아키텍처에 따른 분류 정확도

| Data type | Resolution | | Model type | Average accuracy (%) (5-fold) |
|-------------|----------------|----------------|------------|-------------------------------|
| | Time | Freq. | | |
| Time series | Highest | Lowest | CNN | 53.94 |
| | | | RNN | 41.40 |
| | | | ANN | 39.51 |
| PSD | Lowest | Highest | CNN | 74.25 |
| | | | RNN | 43.17 |
| | | | ANN | 81.91 |
| Spectrogram | Low | High | CNN | 70.59 |
| | | | RNN | 56.05 |
| | | | ANN | 59.26 |
| Scalogram | Highest | Middle | CNN | 83.35 |
| | | | RNN | 45.95 |
| | | | ANN | 55.94 |

표 2. Hyperparameters search space ranges

| Hyperparameter | Search space range |
|-------------------|--------------------|
| Number of layers | [2, 10] |
| Number of filters | [16, 64] |
| Number of nodes | [16, 64] |
| Learning rate | [0.0001, 0.1] |

각 입력데이터 유형과 모델 아키텍처로 구성된 모델은 동일한 조건에서의 비교를 위해 모델 최적화 과정을 통제하였다. 각 모델의 hyperparameter 는 동일한 search space 를 탐색하는 Bayesian optimization 을 통해 최적화되었고, 최적화된 hyperparameter 와 search space 는 표 2 에 표기하였다. 또한, 모델의 학습과 성능 평가를 위한 데이터셋이 특정 입력이나 모델 아키텍처 유형에 유리하지 않도록 k-fold cross-validation (k=5)을 이용하여 5-fold 평균 정확도를 비교하였다.

각 입력데이터 유형과 모델 아키텍처에 따른 5-fold 평균 분류 정확도는 표 1 과 같다. 분류 모델의 성능으로는 scalogram 과 CNN 이 가장 높은 정확도 (83.35%)를 보였고, 그 뒤를 따라 PSD 와 ANN (81.91%), PSD 와 CNN (74.25%), 그리고 spectrogram 과 CNN (70.59%)이 순차적으로 높은 분류 정확도를 보였다.

딥러닝 모델 아키텍처 유형으로 CNN 이 입력 유형과 관계없이 높은 정확도 (평균 70.53%)를 보였다. PSD 를 제외한 모든 입력데이터 유형에서 RNN, ANN 대비 CNN 의 성능이 가장 높은 정확도를 보였으며, RNN 보다는 항상 좋은 분류 정확도를 기록하였다. 반면에 RNN 은 모든 입력 유형에 대해 전반적으로 낮은 정확도 (평균 46.64%)를 보였다. ANN 은 PSD 를 제외한 입력 유형에 대해서 CNN 과 비교하여 낮은 성능(평균 52.68%, PSD 제외)을 보였지만, PSD 에 대해서 큰 성능 개선 (다른 입력 유형 대비 최대 42.40% 개선)을 보였다. 이를 통해 CNN 과 RNN 의 분류 성능은 입력 유형의 시간 및 주파수 해상도에 따른 영향을 적게 받으며, CNN 이 전반적으로 높은 분류 정확도를 보임을 확인하였다. 반면에 ANN 은 입력 유형에 따라 모델 성능의 차이가 큼을 확인하였다.

입력데이터 유형으로는 시간 해상도가 원래의 시계열 데이터와 동일하지만 주파수 해상도 또한 갖도록 변형된 scalogram 이 가장 높은 정확도를 기록하였고 PSD 와 spectrogram 이 그 뒤를 따른다. 반면에 시계열의

경우, 모든 모델 아키텍처에 대해서 전반적으로 낮은 정확도를 보이고 최대 정확도 53.94%를 기록하며 다른 입력데이터 유형과는 큰 차이를 보였다. 이를 통해 주파수 특징을 모델에 직접적으로 제공하는 것이 모델의 분류 성능을 개선시킬 수 있고, 특히 높은 시간과 주파수 해상도를 모두 제공하였을 때 개선효과가 커짐을 확인하였다.

III. 결론

본 논문에서는 파킨슨병 분류 모델에 대해, 서로 다른 시간 및 주파수 해상도를 갖는 입력데이터 유형과 딥러닝 모델 아키텍처에 따른 분류 정확도를 동일한 조건에서 비교하였다. 그 결과로 주파수 영역의 정보를 입력으로 직접적으로 제공하는 것과 CNN 아키텍처를 이용하는 것이 전반적으로 높은 분류 성능을 가짐을 보였다. 특히 시간 및 주파수 해상도가 모두 높은 입력 유형과 CNN 아키텍처를 이용하는 것이 가장 우수한 성능을 가짐을 확인하였다. 따라서, 이후 파킨슨병 분류 모델에 대한 입력데이터 유형과 모델 아키텍처 선정 시, 높은 시간 및 주파수 해상도를 갖는 입력데이터 유형과 CNN 모델 아키텍처를 고려하는 것이 유리할 수 있다.

ACKNOWLEDGMENT

이 논문은 과학기술정보통신부의 재원으로 한국연구재단의 지원 (No. RS-2023-00266171)과 과학기술정보통신부가 지원하는 한국뇌연구원(KBRI)을 통한 KBRI 기초연구 프로그램 (23-BR-04-04)의 지원을 받아 수행된 연구임.

참고 문헌

- [1] Maitin, A. M., Romero Muñoz, J. P., & García-Tejedor, A. J. (2022). Survey of machine learning techniques in the analysis of EEG signals for Parkinson's disease: A systematic review. *Applied Sciences*, 12(14), 6967.
- [2] Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of neural engineering*, 16(3), 031001.
- [3] Shi, X., Wang, T., Wang, L., Liu, H., & Yan, N. (2019, November). Hybrid Convolutional Recurrent Neural Networks Outperform CNN and RNN in Task-state EEG Detection for Parkinson's Disease. In 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (pp. 939-944). IEEE.