

방향 제어 BCI 를 위한 EEG 오류 관련 전위 기반 강화학습 모델

김도연, 이재호*
덕성여자대학교

ehdus2022@duksung.ac.kr, *izeho@duksung.ac.kr

EEG-Based Reinforcement Learning Model with Error-Related Potentials for Intentional Direction Control in Brain-Computer Interfaces

Kim Do Yeon, Jaeho Lee
Duksung Woman's Univ.

요 약

본 연구에서는 Self-Attention 을 적용한 GRU 네트워크와 오류 관련 전위(ErrP) 보상 체계를 결합한 강화학습 BCI(뇌-컴퓨터 인터페이스) 방향 제어 시스템을 제안한다. EEG 신호는 32 채널 캡으로 512 Hz 샘플링한 뒤 0.5-40 Hz 밴드패스 필터링, ICA 아티팩트 제거, Z-score 정규화를 거쳐 전처리한다. Attention-GRU 는 시계열 데이터 중 핵심 시점을 강조하여 상태(State) 벡터를 생성하고, Double DQN 에이전트는 이 벡터를 입력받아 좌·우 방향(Action)을 결정한다. 행동 후 300-500ms 구간에서 ErrP 검출 시 -1, 미검출 시 +1 의 보상을 부여하며 연속 오류에는 지수 감소 페널티를 적용한다. ErrP 기반 보상은 별도 보상 함수 설계 없이 즉각적·암묵적 피드백을 제공하여 학습 목표 설정 부담을 제거하고 학습 수렴 속도와 제어 성능을 동시에 개선한다. 또한 명시적 입력 없이 동작하여 사용자 피로도를 낮추고 인터페이스 자연성을 확보함으로써 실용적 BCI 시스템으로의 확장 가능성을 확인하였다.

I. 서 론

BCI 는 뇌와 컴퓨터가 직접 통신하여 기기를 조작하는 기술로, 재활, VR 인터페이스, 커뮤니케이션 보조 시스템에서 다양한 응용 가능성을 가진다.[1] 특히, 저비용과 높은 시간분해능으로 뇌의 전기적 활동을 기록하는 뇌파(EEG)가 BCI에서 많이 사용된다. 그러나 EEG 신호는 낮은 신호 대 잡음비(SNR)과 개개인 및 세션 간 비정상성(non-stationarity) 문제로, 안정적인 문제로 특징 추출과 분류 성능이 저해된다. 또한, 기존 BCI 시스템은 오류 발생 시 버튼 클릭이나 음성 명령 등 명시적 입력을 요구하여 실시간 제어 흐름을 방해하는 한계가 있다. 본 연구는 Self-Attention 기반 GRU 구조로 EEG 시계열 중 핵심 시점만을 강조하여 비정상성 문제를 완화하고, ErrP 를 자동 보상으로 활용하는 강화학습 체계를 결합함으로써 명시적 입력 없이도 좌·우 방향 제어를 수행하는 자동화된 BCI 시스템을 제안한다.[2]

II. 본론

EEG 신호로부터 의미 있는 특징을 추출하기 위해 전통적으로는 밴드파워, 주파수 스펙트럼, CSP(Common Spatial Pattern) 같은 통계 및 스펙트럼 기반 기법이 사용되어 왔다. 이러한 방법은 계산이 비교적 간단하지만, 시공간적 상호작용을 충분히 반영하지 못하고 잡음에 취약하다는 한계가 있다. 이후 CNN 과 LSTM 같은 딥러닝 모델이 도입되며 비선형 패턴 학습 능력이

향상되었으나, CNN 은 시간 축상의 장기 의존성 처리에, 순환 신경망(RNN 계열)은 시퀀스 내 잡음 구간에 대한 민감도 제어에 어려움을 겪었다.

Self-Attention 메커니즘은 시계열 전체를 대상으로 쿼리, 키, 가치 연산을 수행한 뒤, 각 시점의 상대적 중요도에 따라 가중치를 부여한다. 이로써 잡음이 많은 구간을 자동으로 배제하고 핵심 시점의 특징을 부각시킬 수 있어, 비정상성이 심한 EEG 데이터 처리에 효과적이다.[3]

오류 관련 전위(ErrP)는 사용자가 시스템 오류를 지각할 때 나타나는 300~500ms 구간의 전위 패턴이다. 선행 연구들은 ErrP 를 외부 입력 없이도 피드백 신호로 활용하여 분류 정확도를 높이거나, 주관적 평가를 자동 보상으로 전환하는 방식을 제안해 왔다.[4,5] 이러한 ErrP 기반 보상 설계는 사용자의 인지를 실시간 강화학습에 통합할 수 있는 가능성을 보여준다.

본 시스템은 EEG 획득·전처리, Attention-GRU 특징 추출, 강화학습 에이전트, ErrP 기반 보상 모듈로 구성된다. 먼저 32 채널 EEG 캡을 통해 뇌파를 실시간으로 수집한다. 수집된 신호는 0.5-40Hz 밴드패스 필터링을 적용해 주요 뇌파 대역을 추출하고, ICA(Independent Component Analysis)를 통해 눈 깜빡임이나 근전도 아티팩트를 제거한다. 마지막으로 각 채널별로 Z-score 정규화를 적용하여 세션 간, 피험자 간 변동성을 최소화한다.

전처리된 EEG 전처리된 시퀀스는 GRU 레이어에 입력되며, 내부의 Multi-Head Self-Attention 블록이 상대적 중요도를 계산해 핵심 시점을 강조함으로써 비정상성이 심한 EEG 에서도 안정적 상태 표현을

생성한다. 각 타임스텝별로 계산된 attention 가중치는 소프트맥스 연산을 통해 정규화되며, 이를 바탕으로 최종 상태 벡터가 생성된다. 이 상태 벡터는 이후 강화학습 에이전트의 관측값으로 사용된다.

Attention-GRU 가 산출한 상태 벡터를 기반으로 에이전트는 좌/우 행동을 선택한다. 강화학습 에이전트는 Double DQN 구조를 사용하며, ϵ -greedy 정책으로 행동을 선택한다. 타겟 네트워크를 주기적으로 업데이트하여 학습의 수렴 속도를 높이고 발산을 방지한다. 행동 직후 300-500 ms 구간에서 ErrP 가 검출되면 -1, 미검출되면 +1 을 보상으로 부여하며, 연속 오류 발생 시 지수 감쇠 페널티를 적용해 과도한 부정 보상 누적을 방지한다. ErrP 는 사용자가 오류를 인지할 때 자연 발생하는 전위 패턴이므로 별도 보상 함수 설계 없이 인간 두뇌 신호 그대로를 보상으로 활용할 수 있다. 이는 강화학습 에이전트의 추측 정확도를 높이며, 학습 수렴 속도를 높인다.[5]

III. 결론

본 논문에서는 뇌파 기반 방향 제어를 실현하기 위해, Self-Attention 메커니즘을 통합한 GRU 네트워크와 오류 관련 전위(ErrP)를 보상 신호로 활용하는 강화학습 기반 BCI 시스템을 제안하였다. 제안한 시스템은 EEG 의 비정상성 문제를 완화하고, 사용자의 명시적 입력 없이도 실시간으로 방향 명령을 생성할 수 있다는 점에서 기존 접근 방식과 차별성을 가진다. 또한, 시계열 데이터의 핵심 정보를 효과적으로 추출하고, 자동화된 피드백 구조를 갖추어 학습 효율성과 제어 안정성을 동시에 확보하였다. 향후에는 다중 행동에 대한 확장 적용과 온라인 학습 기법을 적용함으로써, 실제 환경에서도 적용 가능한 실용적인 BCI 제어 시스템으로 발전시킬 수 있을 것으로 기대된다.

참 고 문 헌

- [1] Dongseon Kim, & DooHan Yoo (2024). Study on the Effectiveness of Person-Centered Communication Intervention Using EEG. Journal of the Korean Gerontological Society, 44(3), 381-400.
- [2] B. Mu, K. Yin, C. Niu, and C. Yu, "EEG-based study of error-related potentials in cerebellar regions," IEEE Joint International Information Technology and Artificial Intelligence Conference, vol. 10, pp. 806-810, Jun. 2022.
- [3] P. Núñez et al., "Characterizing Non-stationarity in Alzheimer's Disease and Mild Cognitive Impairment by Means of Kullback-Leibler Divergence," Springer, Cham, 2018, pp. 574-578.
- [4] T. Luo, Y. Fan, J. Lv, and C. Zhou, "Deep reinforcement learning from error-related potentials via an EEG-based brain-computer interface,"

Bioinformatics and Biomedicine, pp. 697-701, Dec. 2018.

- [5] A. X. Fidêncio, C. Klaes, and I. Iossifidis, "Error-Related Potentials in Reinforcement Learning-Based Brain-Machine Interfaces," Frontiers in Human Neuroscience, vol. 16, Jun. 2022,