

로컬 패턴과 전역적 특성을 동시에 고려하는 sEMG 신호 분류 모델

박현우, 최수로, 이상철*

대구경북과학기술원

{hyeonu4995, choiyo21, sangchul.lee}@dgist.ac.kr

sEMG signal classification model with simultaneous consideration of local patterns and global characteristics

Hyeon-woo Park, Suro Choi, Sang-chul Lee*

Daegu Gyeongbuk Institute of Science and Technology

Abstract

본 논문은 EMG pattern recognition(PR)을 위해 gesture에 따른 표면 근전도 신호를 높은 정확도로 분류하는 CNN-Transformer 결합 모델을 제안한다. 제안하는 모델은 EMG PR에서 로컬 패턴뿐만 아니라 신호의 전역적인 특성을 고려하여 여러 gesture에 대해 높은 정확도를 유지 할 수 있다. 제안하는 모델의 성능을 평가하기 위해 ninapro의 DB2 dataset으로 사용한다. 기존의 대부분의 딥러닝 기반의 PR 모델은 ninapro의 DB2 dataset의 일부 gesture에 대해서만 실험을 진행하여 모델의 robustness를 가늠하기 어려웠다. 본 논문에서 모든 gesture 대해 일관적으로 높은 분류 정확도를 보인다. 제안하는 모델은 각 gesture에서 평균 86.2%의 분류 정확도를 달성하였으며, 기존 모델보다 4% 더 높은 성능을 보였다.

I. 서론

표면 근전도 신호(sEMG signal)는 골격근에서 발생하는 근육의 전기적 신호를 전극으로 측정하여 얻어진 신호이다. 이 신호는 근육 기능 평가나 재활 치료, 스포츠 과학 등의 다양한 분야에서 활용된다. 이러한 근전도 신호를 딥러닝 알고리즘을 바탕으로 다수의 gesture를 분류하는 방법이 다양하게 제안되었다. 대표적으로 로컬 패턴에 초점을 맞춘 dual view multi-CNN model [3], 전체적인 신호의 맥락까지 고려한 CNN-RNN 혼합 model [4] 등이 있다.

본 논문에서 제안하는 딥러닝 모델은 초기 layer에서 낮은 level의 특징을 추출하고 이후 layer를 거치면서 계층적으로 고차원적인 패턴을 인식한다. Max pooling 기법을 더해 연산 효율성을 높였다. 로컬 패턴에 대한 feature 학습뿐만 아니라 전역적 특성에 대한 고려가 필요하다고 판단되어 Transformer에 4개의 Multi-Attention head를 통해 sequence 간의 복잡한 맥락을 학습하도록 하여 전역적 특성에 대한 고려가 분류 정확도 향상에 유용함을 실험을 통하여 보인다.

성능 평가를 위하여 우리는 기존 방법들에서 주로 사용되는 데이터인 NinaPro DB2 dataset을 사용한다. 저자들이 알고 있는 한 기존 방법들은 일부 gesture를 이용하여 정확도를 평가하고 있으며, 모든 gesture에 대한 분류 성능을 보고하는 방법은 존재하지 않는다. 본 논문에서는 제안하는 모델이 모든 gesture에서 우수한 성능을 보임을 실험적으로 검증한다.

II. 본론

II-1 데이터 전처리

NinaPro DB2 dataset은 exercise B,C,D의 일상적인 동작과 손가락의 움직임 등을 포함한 총 49개의 gesture를 6회씩 반복하여 12채널의 표면 근전도 센서로 측정한 데이터이다. [5] 데이터는 정규화한 후, 20Hz-150Hz의 bandwidth 필터를 통과시켰다. 신호 데이터를 300ms 길이의 frame으로 나누면서 frame 간에 급작스러운 신호 변화를 방지하고자 10ms 길이의 stride로 인접한 frame 간에는 overlapping하도록 windowing을 하였다. 정답 label인 (,49)의 y data는 one-hot encoding하였다. 6회 행동 중 한번을 test, 나머지를 train dataset으로 나누어 학습과 검증을 진행했다.

II-1 모델 아키텍처

그림 1은 제안하는 모델 아키텍처를 나타낸 그림이다. 전처리한 sEMG signal에 대해 CNN과 Max Pooling을 합한 5개의 layer와 dropout, Transformer 인코더 단계에 4개의 multi-attention head를 배치하여 신호 sequence를 학습한다. Multie-head Transformer의 출력을 flatten단과 L2 정규화 및 dropout단을 통과시켜 분류를 진행한다.

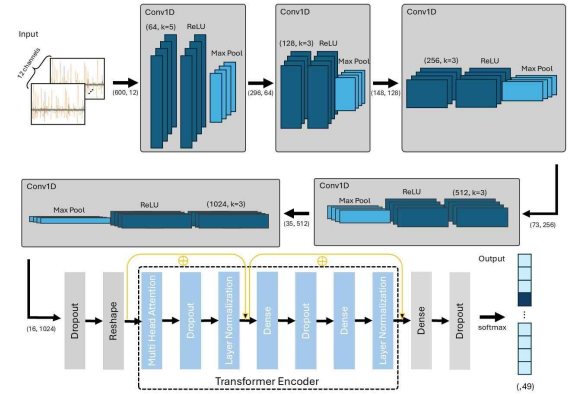


그림 1 모델 아키텍처

CNN과 Transformer의 결합을 통해 로컬 패턴과 전역적 관계를 동시에 고려하여 모델이 EMG 신호와 같이 복잡한 sequential data를 더 잘 이해할 수 있다. CNN을 통해 로컬 패턴을 추출하고, 이를 Transformer를 통해 임베딩하여 sequential data에서의 전역적 특성을 모델이 잘 학습할 수 있다. 다양한 normalization 기법을 통해 모델의 안정성과 일반화 성능을 높였다. Transformer encoder에서 4개의 attention head로 다양한 관점에서 sequence relation을 학습할 수 있고, feed-forward network로 비선형성을 추가하여서 더 높은 분류 정확도 향상을 달성하였다.

III. 실험

그림 2는 본 NinaPro DB2 dataset을 활용하여 학습 결과를 나타낸 그림이다. CNN과 Multi-head Transformer를 결합한 모델의 Loss와 Accuracy 모두 안정적으로 학습되었으며, 86.2%의 정확도를 나타낸다.

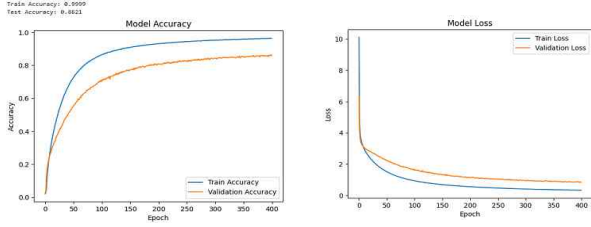


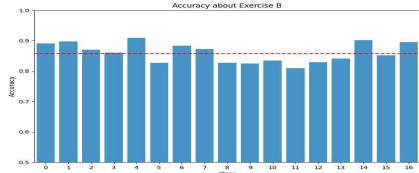
그림 2 모델 성능 및 정확도 평가

기존의 제안된 모델 중에서 NinaPro DB2 dataset의 모든 49개 gesture에 대해 분류한 모델만 비교하였다. DB2의 일부 데이터(exercise D를 제외한 40개의 gesture 등) [6]만 실험하여 분류 정확도를 실험한 연구들은 모델의 robustness가 부족하다고 판단하여 비교에서 제외하였다. 표 1에서 Input은 Ninapro DB2 dataset의 전처리 결과를 의미하며, Classifier는 각 모델의 메인 아키텍처를 의미한다. 본 연구에서 제안하는 모델은 기존 방법에 비해 4% 이상의 분류 정확도가 향상되었다..

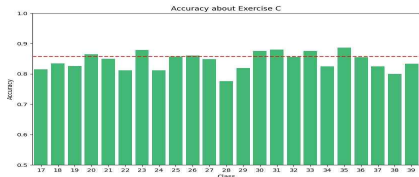
표 1 제안된 모델과 기존의 모델 분류 능력 비교

Work	Input	Classifier	Accuracy
Atzori et al. [3]	raw-sEMG Image	CNN	60.3%
Zhai et al. [4]	sEMG-spectrograms Image	CNN	78.7%
Wei et al. [1]	raw-sEMG Image	multi-view CNN	81.4%
Hu et al. [2]	raw-sEMG signal	CNN / RNN	82.2%
Our Work	raw-sEMG signal	CNN / Transformer	86.2%
Akram et al. [6]	sEMG-Map	CNN	87.9%

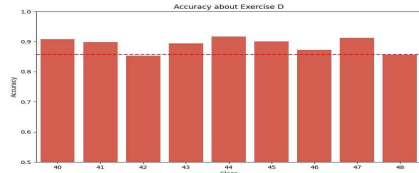
그림 3은 본 논문에서 제안하는 모델의 분류 정확도가 모든 gesture에 대해 균일하게 높은 것을 나타낸다. 이는 본 모델이 다수의 gesture에 대해서도 일관된 분류 성능을 지니며, 모델의 robustness가 큼을 의미한다.



(a) Exercise B(gesture 수: 17개)



(b) Exercise C(gesture 수: 23개)



(a) Exercise D(gesture 수: 9개)

그림 3 각 Exercise에 대한 gesture 분류 정확도

그림 4를 통해 모델에서 5개의 Conv1D layer 중 1개를 제거하였을 때, 원 모델보다 정확도나 loss가 불안정해짐을 확인할 수 있다. Multi-head Transformer 인코더를 제거하면 정확도가 약 15% 정도 감소하였다. 이는 sequential data의 전역적 특성을 모델이 제대로 학습하지 못했기 때문인

것으로 추정된다.

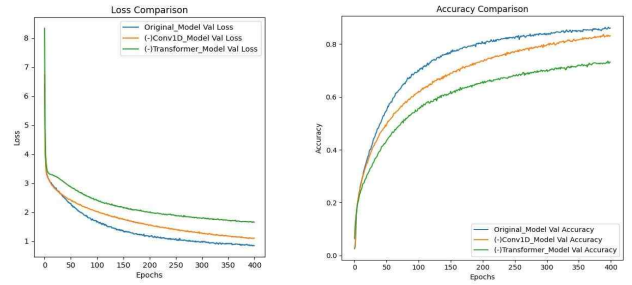


그림 4 Ablation study

IV. 결론

표면 근전도 신호(sEMG)를 통한 pattern recognition을 위한 효과적으로 수행하기 위하여 CNN과 Transformer encoder를 결합하여 로컬 패턴과 전역적 관계를 동시에 고려하는 모델을 제안하였다. 제안하는 방법은 NinaPro DB2 dataset의 모든 49개의 gesture 분류에서 평균 86.2%의 정확도를 달성하여 기존 모델 대비 약 4% 향상된 성능을 보였다. 또한, Ablation study를 통해 sequential data에서 Multi-head Transformer 인코더를 통한 전역적 특성 학습의 중요도를 확인하였고, 이것이 분류 정확도 향상에 기여함을 확인하였다. 향후, 우리는 EEG와 ECG 등의 다양한 생체신호에서도 전역적 특성을 고려한 모델 설계가 분류 정확도 향상에 기여할 수 있는지 추가적인 연구를 진행할 계획이다.

ACKNOWLEDGMENT

본 논문은 2024년도 과학기술정보통신부에서 지원하는 DGIST UGRP(2024020023), 대구경북과학기술원(No. 24-ET-02, 2024020023), 현대차 정몽구 재단 장학생(최수로)으로서 지원을 받아 수행된 연구임.

참고 문헌

- [1] W. Wei, Q. Dai, Y. Wong, Y. Hu, M. Kankanhalli, and W. Geng, "Surface-electromyography-based gesture recognition by multi-view deep learning," IEEE Transactions on Biomedical Engineering, vol. 66, pp. 2964 - 2973, 2019.
- [2] Y. Hu, Y. Wong, W. Weti, Y. Du, M. Kankanhalli, and W. Geng, "A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition," PLoS One, vol. 13, Oct. 2018.
- [3] M. Atzori, M. Cognolato, and H. Müller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," Frontiers in Neurobotics, vol. 10, 2016.
- [4] X. Zhai, B. Jelfs, R. H. M. Chan, and C. Tin, "Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network," Frontiers in Neuroscience, vol. 11, 2017.
- [5] M. Atzori et al., "Electromyography data for non-invasive naturally controlled robotic hand prostheses," Scientific Data, vol. 1, pp. 1 - 13, 2014.
- [6] Akram Fatayer, Wenpeng Gao, and Yili Fu, "sEMG-Based Gesture Recognition Using Deep Learning From Noisy Labels," IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 9, 2022.