



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

2022년 8월

박사학위논문

보조 분류기를 결합한 멀티 스트림 삼
융합 네트워크 기반 근전도를 이용한
사용자 인식에 관한 연구

조선대학교 대학원

제어계측공학과

김진수

보조 분류기를 결합한 멀티 스트림 삼 융합 네트워크 기반 근전도를 이용한 사용자 인식에 관한 연구

A Study on Personal Recognition using Electromyogram based on
Multi-stream Siamese Fusion Network Combining
Auxiliary Classifier

2022년 8월 26일

조선대학교 대학원

제어계측공학과

김진수

보조 분류기를 결합한 멀티 스트림 삼 융합 네트워크 기반 근전도를 이용한 사용자 인식에 관한 연구

지도교수 반 성 범

이 논문을 공학박사학위신청 논문으로 제출함.

2022년 4월

조선대학교 대학원

제어계측공학과

김 진 수

김진수의 공학박사학위논문을 인준함

위원장 조선대학교 교수 김윤태 (인)

위원 조선대학교 교수 곽근창 (인)

위원 조선대학교 교수 최현식 (인)

위원 고려대학교 교수 정용화 (인)

위원 조선대학교 교수 반성범 (인)

2022년 6월

조선대학교 대학원

목 차

제1장 서론	1
제1절 연구 배경	1
제2절 연구 목적	4
제3절 연구 내용 및 방법	5
제2장 기존 근전도를 이용한 사용자 인식	8
제1절 도메인 정보 기반 핸드크래프트 특징 추출	8
제2절 신경망을 이용한 특징 추출	13
제3절 삼 구조를 이용한 데이터 쌍의 유사도 학습	17
제3장 제안하는 어텐션 메커니즘 기반 근전도를 이용한 사용자 인식	24
제1절 경험적 모드 분해를 이용한 특징 추출	27
제2절 보조 분류기를 결합한 멀티 스트림 삼 융합 네트워크	31
1. 합성곱 신경망 및 삼 네트워크 구조	31
2. 제안하는 보조 분류기를 결합한 삼 융합 네트워크	36
제4장 멀티 스트림 삼 융합 네트워크를 이용한 사용자 인식 실험 결과	47
제1절 근전도 공개 데이터베이스	47
제2절 사용자 인식 평가 방법	53
제3절 실험 결과 및 분석	54
제5장 결론	65
참고문헌	67

표 목차

표 1. 기존 사용자 인식 방법의 범죄 사례	1
표 2. 핸드크래프트 특징을 이용한 사용자 인식 연구	12
표 3. 신경망을 이용한 사용자 인식 연구	16
표 4. 삼 네트워크를 이용한 유사도 학습 연구	23
표 5. 하위 네트워크 각 스트림의 계층별 출력 결과	41
표 6. Decision 네트워크의 계층별 출력 결과	42
표 7. 근전도 공개 데이터베이스 개요	47
표 8. 사용자 인식 실험 단계별 구성	55
표 9. 학습된 모델 기반 새로운 사용자를 이용한 실험 결과	61

그림 목차

그림 1. 생체신호 측정 방법 및 파형 예시	2
그림 2. 수행한 동작에 따른 동일한 피험자의 근전도 파형	3
그림 3. 측정된 근전도의 불규칙성 예시	5
그림 4. 경험적 모드 분해 기반 특징 추출 예시	6
그림 5. 패턴과 근전도를 이용한 2-factor 사용자 인식 방법	9
그림 6. 고밀도 표면 근전도 측정 예시	10
그림 7. 문구 입력 과정에서 키스트로크와 근전도 측정 예시	11
그림 8. 얼굴 근육의 근전도 측정 예시	14
그림 9. 단일 동작 근전도 기반 사용자를 인식하기 위한 신경망 구조	14
그림 10. 2차원 합성곱 신경망 기반 삼 네트워크 예시	17
그림 11. 1차원 합성곱 신경망 기반 삼 네트워크 예시	18
그림 12. 장단기 메모리 기반 삼 네트워크 예시	20
그림 13. 합성곱 신경망-장단기 메모리 앙상블 기반 삼 네트워크 예시	21
그림 14. 신호의 비선형성 특성을 고려하지 않은 특징 추출 예시	23
그림 15. 제안하는 사용자 인식 방법 흐름도	24
그림 16. 동일한 동작을 수행한 피험자들의 근전도 파형	25
그림 17. 시간 및 주파수 도메인에서의 전처리에 따른 근전도 비교	26
그림 18. 체거름 과정을 통한 내재 모드 함수 분해 과정	29
그림 19. 근전도의 내재 모드 함수 및 잔차	30
그림 20. 1차원 데이터를 이용한 합성곱 연산	32
그림 21. 퓨샷 학습을 이용한 데이터 쌍 유사도 비교 결과 예시	34
그림 22. 삼 네트워크 구조 예시	35
그림 23. 사용자 인식을 위해 제안하는 멀티 스트림 삼 융합 네트워크 구조	37
그림 24. 설계한 하위 네트워크 구조	40
그림 25. 설계한 decision 네트워크 구조	42

그림 26. 보조 분류기를 이용한 데이터 융합	43
그림 27. GoogLeNet의 메인 분류기와 보조 분류기	44
그림 28. 제안한 보조 분류기를 이용한 어텐션 메커니즘 방법	45
그림 29. EMG(Dr. R. Khushaba) 데이터 세트 3의 손동작	48
그림 30. UCI의 sEMG for Basic Hand movements Data Set 손동작	49
그림 31. Ninapro DB2 손동작	49
그림 32. Ninapro DB5 손동작	50
그림 33. CU_sEMG DB의 측정 근육 위치	51
그림 34. CU_sEMG DB의 12개 손동작 및 근전도 파형	52
그림 35. 스트림 변화에 따른 사용자 인식 성능	56
그림 36. 멀티 스트림 삼 네트워크를 이용한 사용자 인식 성능	57
그림 37. 보조 분류기를 결합한 멀티 스트림 삼 융합 네트워크 기반 실험	58
그림 38. 보조 분류기를 이용한 어텐션 메커니즘 결과	59
그림 39. 제안한 사용자 인식 방법의 ROC 커브 결과	60
그림 40. 보조 분류기를 결합한 단일 스트림 삼 융합 네트워크 기반 실험	60
그림 41. 전이 학습 방법과 성능 비교 결과	62
그림 42. CU_sEMG DB를 이용한 기존 사용자 인식 방법의 실험 결과	63
그림 43. Ninapro DB2를 이용한 기존 사용자 인식 방법의 실험 결과	64

ABSTRACT

A Study on Personal Recognition using Electromyogram based on Multi-stream Siamese Fusion Network Combining Auxiliary Classifier

Kim, Jin Su

Advisor : Prof. Pan, Sung Bum, Ph. D.

Department of Control & Instrumentation

Graduate School of Chosun University

As information security becomes increasingly important in modern society, methods of user recognition use a fingerprint, face, and iris, which are accessible with less discomfort and have many other convenient advantages. However, they can be collected without consent and is prone to forgery due to lacking liveness. In addition, they are non-cancelable that cannot change the registered information. In order to overcome these problems, biosignals such as the electromyogram(EMG), electrocardiogram(ECG), and electroencephalogram(EEG) expressing an individual's unique physiological characteristics are introduced, which cannot be examined outside the body.

Among the biosignals, EMG is a signal that measures the microcurrent generated when a muscle moves and represents complex information reflecting not only physiological information of muscle tissue but also neuromuscular control information. Since the signal changes according to their respective gestures, the non-cancelable issue can be efficiently tackled for user recognition. Existing user recognition techniques using EMG use handcraft features or features generated from neural network-based deep learning structures. Each method has its drawbacks; while the handcraft features lack generality as it uses information extracted in a single domain, deep learning requires model retraining to register new users. In addition, the EMG has non-periodic and non-linear characteristics; thus, data irregularity issues arise, resulting in low performance.

This thesis proposes a user recognition method based on a multi-stream siamese fusion network combining an auxiliary classifier to resolve the abovementioned issues regarding user recognition using EMG. In the proposed method, features are extracted by an intrinsic mode function(IMF) using empirical mode decomposition(EMD) in consideration of the time-frequency domain characteristics of the signal. In order to solve the model retraining problem, the proposed method incorporates a model with a multi-stream siamese fusion network. The model consists of two sub networks and one decision network. The two sub networks are designed in a siamese structure to handle the model retraining problem and extract the compatibility function to calculate the similarity of the input data pairs. The decision network computes features from the compatibility function that comes from a convolutional neural network(CNN) organized by a stacking ensemble

structure. Finally, the auxiliary classifier utilizing euclidean distance(ED) is employed to solve the data irregularity problem. The structure incorporates a convolutional neural network and an auxiliary classifier to realize the attention mechanism.

The proposed multi-stream siamese fusion network combining the auxiliary classifier executed several experiments using CU_sEMG DB, which was built in multi-sessions. First, an experiment was conducted to identify the performance of the feature extraction method utilizing the time-frequency domain characteristics of EMG. As a result of the experiment, when the IMF 1-4 was applied for the multi-stream siamese fusion network without the auxiliary classifier, the average recognition accuracy was 92.01%, higher than the existing time-frequency domain characteristic. Second, in order to solve the data irregularity problem, a user recognition experiment was executed to verify the performance of the decision network combining the auxiliary classifier proposed in this thesis. User recognition using the proposed method classified 100 subjects with 94.35% accuracy, and the performance improved by 2.34% compared to the method without the auxiliary classifier. Next, to check the performance of the siamese network for alleviating the model retraining issue, the experiment proceeded using subjects that were not used for training. The result of the experiment using the EMG of new subjects illustrated that the model retraining problem was much relieved by showing an accuracy of 93.19%.

Finally, to verify the proposed method's superiority, using the single-session data of CU_sEMG DB and Ninapro DB2, the proposed method's performance was compared to the existing methods' performance. The existing

methods showed a significant decrease in accuracy when the number of subjects increased, indicating that the model was not adequately trained. On the contrary, the proposed method maintained similar performance as the number of subjects increased and outperformed the existing method, proving the superiority of the proposed method.

In short, a new user recognition method based on a multi-stream siamese network was proposed in the thesis. The proposed method incorporates an auxiliary classifier with a multi-stream siamese network to deal with the problems of user recognition methods using EMG. The next step will be the research for noise cancellation technologies of EMG and new loss functions to improve user recognition performance. Another research issue will be network optimization to improve computational speed.

제1장 서론

제1절 연구 배경

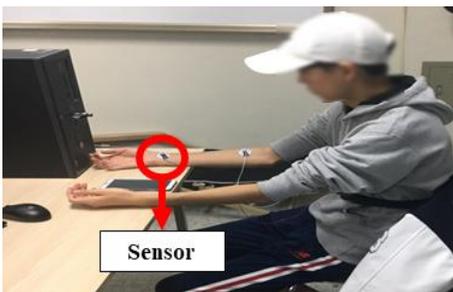
현대 사회에서 정보의 보안이 중요해짐에 따라 사용자를 인식하는 방법은 기존 아이디 및 비밀번호, 보안키를 이용한 방법에서 개인 신체에 있는 정보를 이용한 방법으로 변화되고 있다[1]. 개인 신체에 있는 지문, 얼굴, 홍채를 이용한 사용자 인식은 낮은 거부감과 사용하기 편리한 이점으로 인해 금융, 휴대폰 해제, 출입 통제 등과 같은 실생활과 밀접한 분야에서 많이 사용되고 있다. 그러나 지문, 얼굴, 홍채는 외형적으로 표출된 정보로 사용자 동의 없이 수집할 수 있는 위험성이 있으며, 자발적으로 등록된 정보를 변경할 수 없는 취소 불가능성이 있어 등록된 정보가 유출되면 보안에 취약한 단점이 있다[2][3]. 또한, 생동감이 없는 정보로 사진 및 렌즈 등을 이용해 위조가 가능하여 표 1과 같이 다양한 사고가 발생했다 [4].

표 1. 기존 사용자 인식 방법의 범죄 사례

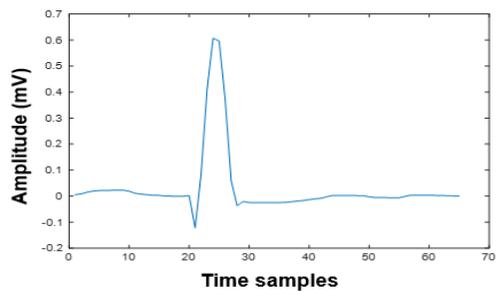
국가	세부 내용
미국	등록된 사용자의 얼굴 영상을 이용한 해킹
브라질	실리콘을 이용한 위조 지문으로 출입문 통과
한국	3D 프린터를 이용한 위조 지문으로 금융 해킹
일본	위조지문에 의한 전자여권 사고
러시아	사진에서 홍채를 복제한 해킹 사고

이러한 단점을 보완하기 위해 신체 외부에서 확인할 수 없으며 개인 고유한 생리적 특성을 표현하는 생체신호를 이용한 사용자 인식 연구가 진행되고 있다. 생

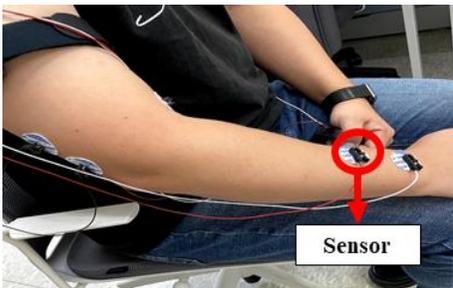
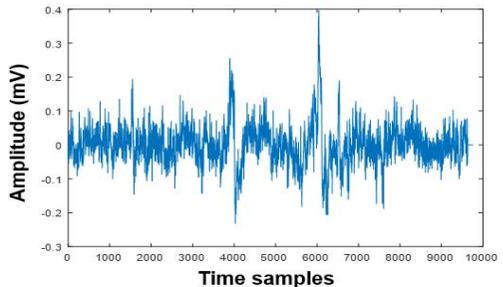
체신호는 사람의 신체 활동을 통해 발생하는 미세전류의 흐름을 측정할 정보를 의미한다. 이러한 생체신호는 사용자 인식으로 사용되기 위한 특성인 보편성(Universality), 특이성(Distinctiveness), 영구성(Permanence), 획득성(Collectability) 등을 만족한다[5]. 사람의 신체에는 다양한 생체신호가 있으며 근전도(EMG), 심전도(ECG), 뇌전도(EEG)를 이용한 사용자 인식 연구가 많이 진행되고 있다.



(a) 심전도 측정 방법 및 파형 예시



(b) 뇌전도 측정 방법 및 파형 예시



(c) 근전도 측정 방법 및 파형 예시

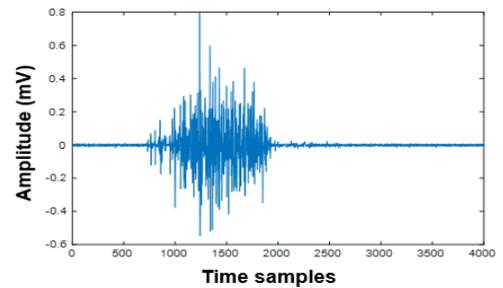


그림 1. 생체신호 측정 방법 및 파형 예시

심전도는 심장 박동에 의해 발생하는 생체신호로, 측정 방법 및 파형은 그림 1 (a)와 같다. 그러나 지문, 얼굴, 흉채와 동일하게 취소 불가능성이 있어 등록된 정보가 유출되면 사용자의 개인 정보 유출과 함께 금전적 손실로 이어질 수 있다. 뇌전도는 대뇌 활동 상태에 의해 발생하는 생체신호로, 측정 방법 및 파형은 그림 1 (b)와 같다. 뇌전도는 대뇌 활동에 의해 다른 형태의 신호 파형이 생성되어 취소 불가능성 문제를 해결할 수 있다. 그러나 신호를 증폭해 측정하여 머리카락 및 두개골, 두피에 의해 신호가 왜곡되고 잡음에 취약한 단점이 있다. 또한, 측정 장비를 사용함에 있어 다른 생체신호에 비해 거부감이 있다[6]. 근전도는 근육이 움직일 때 발생하는 신호를 측정하는 것으로, 측정 방법 및 파형은 그림 1 (c)와 같다. 인체의 다양한 근육에서 측정이 가능하며, 그림 2와 같이 수행한 동작에 따라 신호가 변화하는 특징이 있어 기존 사용자 인식의 문제점인 취소 불가능성 문제를 해결할 수 있다. 또한, 피부에 센서를 부착해 신호 측정이 가능하여 뇌전도에 비해 측정이 간편한 장점이 있다.

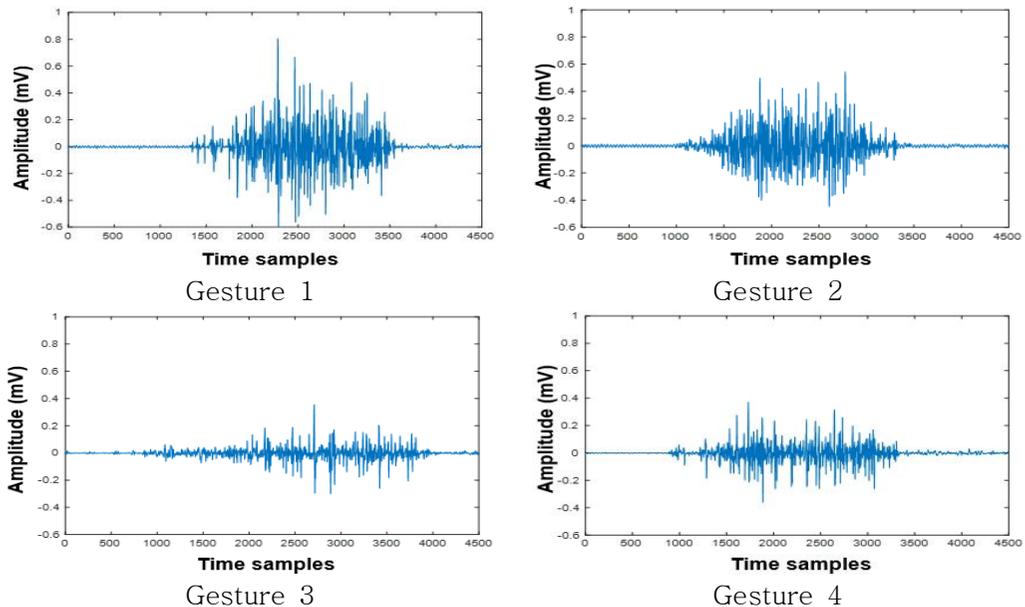


그림 2. 수행한 동작에 따른 동일한 피험자의 근전도 파형

제2절 연구 목적

근전도는 근육 조직의 생리학적 정보뿐만 아니라 신경근 제어 정보를 반영한 복합적인 정보를 나타낸다[7]. 근전도를 측정하는 방법은 침습적 방법과 비침습적 방법이 있다. 침습적 방법은 바늘 전극을 근육에 삽입하여 측정하기 때문에 통증이 적은 장점이 있다. 그러나 바늘 전극을 근육에 삽입할 때 발생하는 통증과 측정 장비의 거부감으로 인해 사용자 인식에 사용하기에는 어려움이 있다. 비침습적 방법은 피부에 전극을 부착해 측정하는 방법으로, 침습적 방법에 비해 비교적 편리하게 근전도를 측정할 수 있다. 비침습적 방법으로 측정한 근전도는 표면 근전도(sEMG, Surface EMG)로 불린다[8].

비침습적 방법으로 측정된 근전도의 특징 추출은 RMS(Root mean square), MAV(Mean absolute value), VAR(Variance), WL(Waveform length) 등과 같이 시간 및 주파수 도메인에서 핸드크래프트로 제작된 특징을 추출하는 방법[9]과 2차원 이미지로 변환하는 방법이 있다[10][11]. 그러나 핸드크래프트로 제작된 특징은 단일 도메인 정보를 이용함에 따라 범용성이 부족하며, 신호에 최적화된 특징을 추출하지 못할 수 있는 문제가 있다[6]. 2차원 이미지 변환은 단시간 푸리에 변환(STFT, Short time fourier transform), 연속 웨이블릿 변환(CWT, Continuous wavelet transform)과 같은 방법으로 진행되지만, 이미지 데이터로 인해 많은 연산량이 있는 단점이 있다[12].

근전도를 이용한 연구에서 딥러닝은 합성곱 신경망(CNN), 순환 신경망(RNN, Recurrent neural network), 장단기 메모리(LSTM, Long-short term memory) 등을 이용한다. 이러한 딥러닝 모델은 다수의 계층으로 구성되어 입력된 데이터에서 적응적으로 특징을 추출한다. 그러나 딥러닝은 모델 학습을 위해 대용량의 데이터가 필요한 단점이 있다. 또한, 입력된 데이터간 차이를 학습에 사용하는 지도 학습

방법으로, 딥러닝 기반 인식 시스템에 새로운 사용자가 추가되면 모델 재학습을 진행해야 함에 따라 프로세스가 낭비되는 문제가 있다[11].

또한, 모델을 학습하기 위해 사용하는 근전도는 시간의 흐름에 따라 주기가 없이 측정된 신호로, 그림 3과 같이 모든 피험자에게 동일한 근전도를 얻을 수 없는 비선형성(Nonlinear) 특성이 있다. 이러한 근전도를 이용할 경우 사용자 인식 성능이 낮아지는 데이터 불규칙성 문제가 발생한다[13].

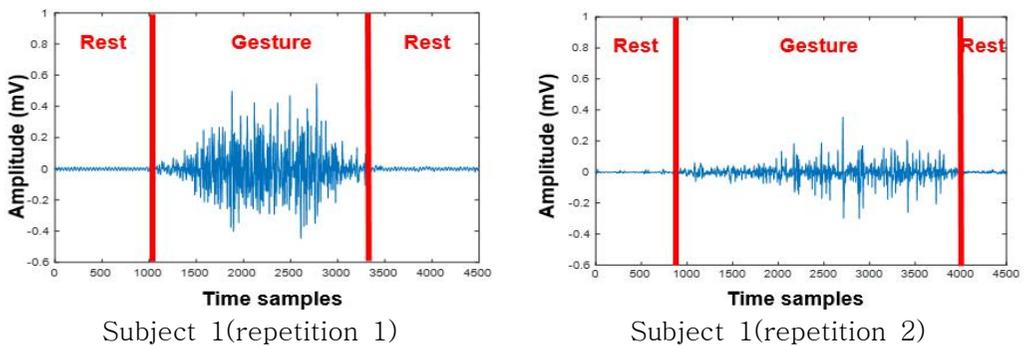


그림 3. 측정된 근전도의 불규칙성 예시

따라서 시간-주파수 도메인 특성을 고려한 특징 추출 방법을 사용하고, 새로운 사용자가 추가되어도 재학습이 필요 없는 딥러닝 모델 설계가 필요하다. 또한, 불규칙성 문제를 해결 가능한 알고리즘을 통해 높은 사용자 인식 성능을 나타내는 연구가 필요하다.

제3절 연구 내용 및 방법

근전도의 시간 도메인은 움직일 때 발생하는 신호의 직관적인 정보를 제공하며, 주파수 도메인은 잡음에 강인한 장점이 있다. 따라서 시간-주파수 도메인 분

석을 통한 특징 추출 방법을 사용해야 한다. 경험적 모드 분해는 데이터를 물리적으로 의미 있는 구성 요소로 분해하는 기술이며, 체거름 과정(Sifting process)을 반복 수행하여 데이터 중첩 없이 특징을 추출한다. 체거름 과정은 envelop 평균값 및 극값의 수 조건을 만족할 때까지 동작하며, 이를 통해 추출된 내재 모드 함수는 보유한 주파수 대역이 다르게 되어 데이터 일반화 오류가 발생하지 않는다. 본 논문에서는 경험적 모드 분해를 이용해 근전도를 처리하여 시간-주파수 도메인 특징을 추출한다.

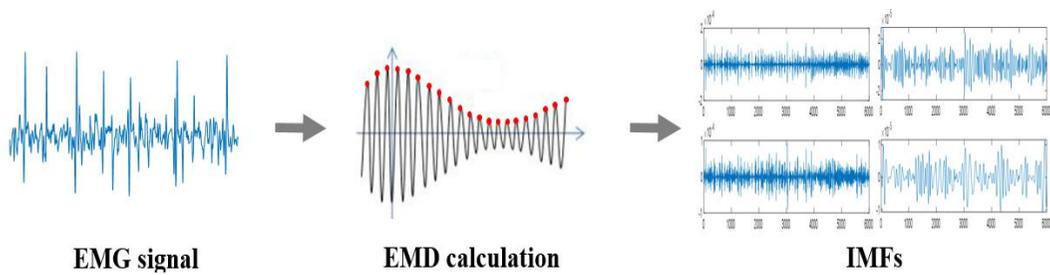


그림 4. 경험적 모드 분해 기반 특징 추출 예시

다음으로 기존 근전도를 이용한 연구는 합성곱 신경망, 장단기 메모리 등의 구조로 설계한 모델을 사용하였다. 입력 데이터에서 합성곱 신경망은 공간적 특징을, 장단기 메모리는 시간적 특징을 추출하며 이를 결합한 합성곱 신경망-장단기 메모리(CNN-LSTM) 및 다층으로 설계된 모델을 이용한 연구가 진행되었다. 그러나 이러한 구조의 모델은 시스템에 새로운 사람이 추가되면 모델을 재학습해야 하는 단점이 있다. 이를 해결하기 위해 동일한 계층 구조로 설계되고 가중치를 공유하는 삼 네트워크(Siamese network)를 기반으로 설계한 딥러닝 아키텍처가 연구되고 있다. 삼 네트워크는 입력된 데이터 쌍의 유사도를 학습 후 동일 여부를 반환하는 모델로, 기존 딥러닝 모델과 차별점이 있다. 따라서 본 논문에서는 딥러닝 모델 학습에 사용되지 않은 새로운 피험자가 추가되어도 딥러닝 모델을 재학습하지 않

고, 피형자의 데이터 일부를 이용해 인식이 가능한 삼 네트워크로 모델을 설계한다.

또한, 근전도의 비선형성 특성으로 인해 발생하는 근전도 불규칙성 문제를 해결하기 위해 어텐션 메커니즘을 사용한다. 어텐션 메커니즘은 입력된 데이터의 중요한 특징을 포착하여 가중치를 부여하는 방법으로, 근전도의 다양한 시간 간격으로 구성된 동작 상태 구간의 신호에 높은 가중치를 적용하여 데이터 불규칙성 문제를 해결할 수 있다. 그러나 기존 어텐션 메커니즘은 복잡한 신경망을 추가로 이용해야 하며 특징을 계산하기 위해 많은 파라미터가 필요한 단점이 있다[14]. 이에 본 논문에서는 보조 분류기(Auxiliary classifier)를 이용해 입력된 데이터에 어텐션 메커니즘을 적용하는 방법을 제안한다. 본 논문에서 제안하는 보조 분류기를 이용한 어텐션 메커니즘은 유클리디안 디스턴스(ED)를 이용한 방법으로, 벡터 연산을 진행함에 따라 파라미터가 적은 장점이 있다.

본 논문의 구성은 다음과 같다. 1장의 서론에 이어 2장에서는 기존 근전도를 이용한 사용자 인식 방법을 분석한다. 3장에서는 본 논문에서 제안하는 보조 분류기를 결합한 멀티 스트림 삼 용합 네트워크 기반 사용자 인식 방법에 대해 설명한다. 4장에서는 제안한 방법을 이용한 사용자 인식 실험 결과를 분석하고, 5장에서는 본 논문의 결론과 향후 연구 방향을 제시한다.

제2장 기존 근전도를 이용한 사용자 인식

제1절 도메인 정보 기반 핸드크래프트 특징 추출

근전도를 이용한 사용자 인식은 동일한 동작을 수행할 때 근육의 발달 정도와 활성도, 습관 등에 의해 발생하는 개인간 차이를 이용해 진행된다. 사용자를 인식하기 위한 근전도는 팔, 다리, 입술 등 인체의 다양한 근육에 발생하는 신호를 이용해 진행되고 있다. 근전도 데이터의 특징 추출은 시간 및 주파수 도메인에서 핸드크래프트로 생성한 특징을 이용한 방법과 신경망으로 생성한 특징을 이용한 방법이 있다.

핸드크래프트 특징은 계산을 위한 수식이 복잡하지 않아 낮은 연산량을 나타내며, 처리 속도가 빠른 장점이 있다. 이에 기존 근전도를 이용한 사용자 인식 연구에서 많이 사용되고 있다. 핸드크래프트 특징을 이용해 사용자 인식을 수행한 연구로 Yamaba[8]의 연구가 있다. Yamaba는 다수의 손동작으로 조합된 통과 제스처를 이용한 사용자 인식 방법을 제안했다. 11명의 피험자가 12개의 손동작을 수행할 때 DL-3100 및 DL-141 장비를 이용해 근전도를 1채널로 측정하고, 측정된 근전도 중 다른 동작보다 뚜렷하게 신호가 생성된 9개의 동작을 실험에 사용했다. 각 동작은 30회 반복 측정했으며, 측정된 근전도는 사용자 인식 성능을 향상하기 위해 상관 계수(Correlation coefficients) 및 교차 상관 함수(Cross-correlation functions)를 이용해 평균에서 벗어난 데이터를 제외하여 학습 데이터를 선별했다. 선별된 근전도는 10개의 세그먼트로 나누어 각 세그먼트에서 sum, mean, skewness, 표준편차(SD, Standard deviation) 등 11개의 핸드크래프트 특징을 추출하고, SVM(Support vector machine), DTW(Dynamic time warping)로 사용자를 분류했다.

Li[15]는 그림 5와 같이 스마트폰 화면에서 잠금 해제 패턴을 그릴 때 측정된 근전도로 2-factor 사용자 인식 연구를 진행했다. 근전도는 OpenBCI를 이용해 flexor digitorum superficialis에서 측정했으며, 10명의 피험자는 의자에 앉은 편안한 상태에서 2개의 잠금 해제 패턴을 수행했다. 측정된 근전도는 5Hz 고역 통과 필터(HPF, High pass filter) 및 60Hz 대역 저지 필터(NF, Notch filter)를 이용해 직류 오프셋, 기준선 변동, 전력선 잡음을 제거했다. 시간 도메인에서 MAV, VAR, WL, ZC(Zero crossing) 등 11개의 핸드크래프트 특징을 추출하고, OCSVM(One-class SVM) 및 LOF(Local outlier factor)를 이용해 98.2%의 정확도로 사용자를 인식했다.

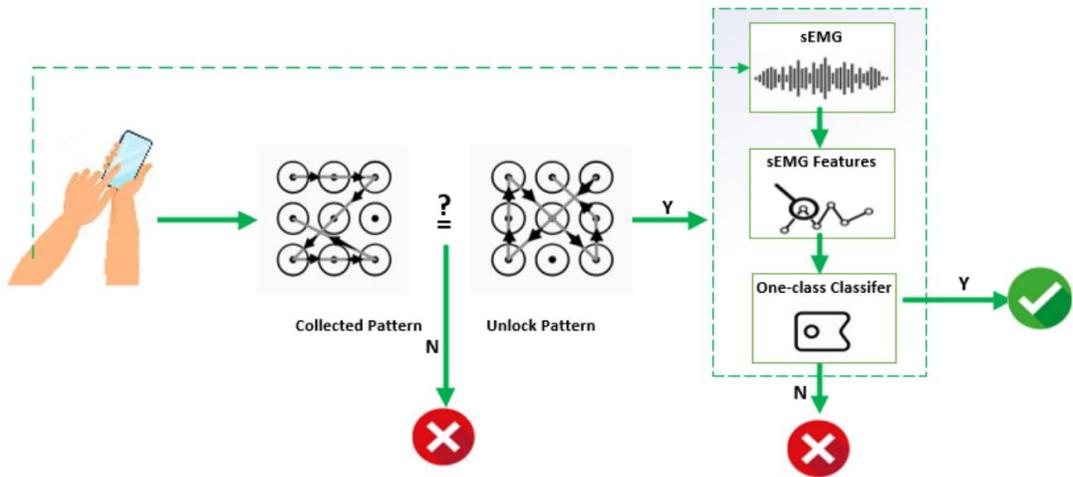


그림 5. 패턴과 근전도를 이용한 2-factor 사용자 인식 방법

Jiang[3]은 고밀도 표면 근전도(HD-sEMG, High density-sEMG)를 이용한 사용자 인식 연구를 진행했다. 고밀도 표면 근전도란 다수의 센서가 배치된 전극 어레이를 이용해 피험자의 근전도를 측정된 신호로, 기존 희소 표면 근전도(Sparse-sEMG)에 비해 많은 정보가 있다. 전극 어레이는 그림 6과 같이 피험자 22명의 오른쪽 손등에 배치했으며, TMSi SAGA64+ 장비를 이용해 64채널로 근전도를

측정했다. 피험자는 의자에 앉은 편안한 상태에서 8개의 손가락 운동을 수행했으며, 평균 9일 간격으로 2개의 세션(Session)에서 데이터를 구축했다. 측정된 근전도는 10-900Hz 대역의 대역 통과 필터(BPF, Band pass filter) 및 50Hz 대역 저지 필터를 이용해 잡음을 제거했다. 특징은 WL, FMD(Frequency median), SS(Spatial synchronization)로 추출했으며, 에너지 제약 조건(Energy constraint)을 이용해 개별 특징의 기여도를 동일하게 하였다. KNN(K-nearest neighbor)을 이용해 피험자를 분류했으며, 85.8%의 정확도를 보였다.

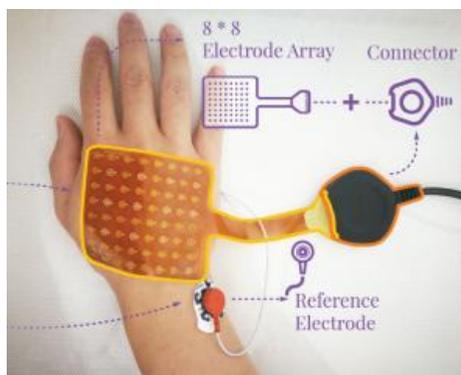


그림 6. 고밀도 표면 근전도 측정 예시

손 이외의 근육에서 측정된 근전도를 이용한 사용자 인식 연구는 Lee[16]의 연구가 있다. Lee는 보폭, 속도, 근육의 크기 등 개인의 보행 습관에 의해 나타나는 개인간 차이를 이용해 사용자 인식 연구를 진행했다. Biopac MP150을 이용해 sartorius, biceps femoris, tibialis anterior 등 왼쪽 다리 9개 근육 및 오른쪽 다리 2개 근육에서 근전도를 11채널로 측정했다. 피험자 12명은 복도를 걷는 동안 근전도를 측정했으며, 20-450Hz 대역 통과 필터 및 60Hz 대역 저지 필터를 이용해 근전도의 잡음을 제거했다. RMS, MAV, DF(Dominant frequency) 등 5개의 시간 및 주파수 도메인 특징을 추출하고, LDA(Linear discriminant analysis)를 이용해 93%의 정확도로 개인을 분류했다.

Khan[9]은 사람이 말을 할 때 측정된 스피치 근전도 데이터를 이용한 사용자 인식 연구를 진행했다. Biopac 측정 장비를 이용해 sternocleidomastoid에 센서를 부착하고, 10명의 피험자는 30분 동안 말을 하면서 신호를 측정했다. 측정된 근전도는 경험적 모드 분해로 신호를 분해하여 내재 모드 함수 1-2를 사용하는 방법으로 신호를 전처리했다. 특징은 MF(Mean frequency), PRMS(Peak to RMS), ZC, SSI(Simple square integral) 등 시간 및 주파수 도메인에서 15개의 핸드크래프트 특징을 추출했다. SVM, KNN, DT(Decision tree)를 이용해 피험자를 분류했으며, 95.3%의 정확도를 보였다.

Shreyas[17]는 14명의 피험자가 18자의 고정된 문구를 입력하는 환경에서 그림 7과 같이 근전도와 키스트로크를 측정하고, 이를 융합한 멀티모달 사용자 인식 연구를 진행했다. BioRadio 측정 장비를 이용해 8채널로 근전도와 키스트로크 정보를 측정했으며, 근전도는 왼손과 오른손의 extensor carpi ulnaris, anconeus, flexor carpi radialis, palmaris longus에 센서를 부착했다. 신호 내 잡음을 제거하기 위해 RMS를 이용했으며, PCA(Principal component analysis), UDP(Unsupervised discriminant projection), 및 CFA(Class-dependent feature analysis)로 신호에서 특징을 추출했다. 근전도와 키스트로크 정보는 유클리디안 디스턴스, KNN, SVM을 이용해 분류하고, 스코어 레벨 융합(Score level fusion)으로 두 결과를 융합해 사용자를 인식했다.



그림 7. 문구 입력 과정에서 키스트로크와 근전도 측정 예시

표 2. 핸드크래프트 특징을 이용한 사용자 인식 연구

저자	데이터	세션	채널	피험자 수	특징	분류기	성능
Yamaba et al.[8]	비공개	단일	1	11	SD, Skewness, Mean, etc. 11ea	SVM, DTW	FAR: 0%
Li et al. [15]	비공개	단일	1	10	MAV, VAR, WL, etc. 11ea	OCSVM, LOF	Acc: 98.2%
Jiang et al.[3]	비공개	멀티	64	22	WL, FMD, SS	KNN	Acc: 85.8%
Lee et al. [16]	비공개	단일	11	12	RMS, MAV, DF, IEMG, OS	LDA	Acc: 93%
Khan et al.[9]	비공개	단일	1	10	MF, PRMS, ZC, etc. 15ea	SVM, KNN, DT	Acc: 95.3%
Shreyas et al.[17]	비공개	단일	8	14	PCA, UDP, CFA	ED, KNN, SVM	FRR: 0.03%

기존 근전도의 핸드크래프트 특징을 이용한 사용자 인식 연구 방법은 표 2와 같으며, 시간 및 주파수 도메인으로 변환 후 특징 추출을 진행하고 있다. 시간 도메인은 근전도의 형태학적 특징을 사용하며, 주파수 도메인은 근전도의 도메인 변환 정보를 이용한다. 그러나 핸드크래프트로 추출된 특징은 단일 도메인 특성을 고려함에 따라 범용성이 부족하며, 최적화된 특징을 추출하지 못하는 문제가 있다 [6]. 이러한 문제를 해결하기 위해 딥러닝 모델을 이용해 입력된 데이터에서 특징을 추출하거나, 핸드크래프트 특징과 딥러닝 모델을 결합한 연구가 진행되고 있다.

제2절 신경망을 이용한 특징 추출

신경망은 기존 핸드크래프트 특징의 단점을 보완하기 위해 연구된 방법으로, 특징 엔지니어링을 수동으로 수행하는 기존 방법과 달리 데이터에서 최적의 특징을 적응적으로 추출한다[18]. 이러한 근전도 기반 신경망을 이용한 사용자 인식 연구로 Wong[19]의 연구가 있다. Wong은 리듬에 따라 동작하며 측정된 연속된 근전도를 이용한 사용자 인식 연구를 진행했다. 7명의 피험자는 MY0 암밴드를 이용해 8채널로 근전도를 측정했으며, 샘플링 레이트는 200Hz로 설정했다. 측정된 근전도는 3초 크기의 윈도우를 이용해 분할하고, 100Hz 저역 통과 필터(LPF, Low pass filter) 및 50~60Hz 대역 저지 필터를 이용해 신호의 잡음을 제거했다. 그 후 3개의 은닉 계층(30개, 20개, 10개)으로 구성된 인공 신경망(ANN, Artificial neural network)을 이용해 사용자를 인식했으며, 98.09%의 정확도를 나타냈다.

Shioji[20]는 Janken의 3가지 동작을 이용한 사용자 인식 연구를 수행했다. P-EMG plus 장비를 이용해 8채널의 근전도를 1,000Hz 샘플링 레이트로 손목에서 측정했다. 측정된 신호는 20Hz 고역 통과 필터, 60Hz 대역 저지 필터로 필터링하고, 수식을 이용해 이상치(Outliers)를 제거했다. 사용자를 인식하기 위한 합성곱 신경망은 3개의 1차원 합성곱 계층 및 완전 연결 계층으로 설계하고, 풀링 계층은 사용하지 않았다. 40%의 드롭아웃을 이용해 과적합을 방지했으며, 측정된 데이터는 128*8 크기로 모델의 입력 데이터로 사용하여 피험자 8명을 94.6%의 정확도로 인식했다.

Morikawa[21]는 얼굴의 근전도를 이용한 사용자 인식 연구를 진행했다. 그림 8과 같이 orbicularis, zygomatic, depressor anguli oris에 센서를 부착한 후 피험자가 모음 발음을 수행할 때 P-EMG plus 장비로 신호를 측정했다. 신호 내 잡음을 제거하기 위해 60Hz에 대역 저지 필터와 5~500Hz에 대역 통과 필터를 사용했고, 윈도우를 중첩해 신호를 분할하였다. 분할된 근전도로 4개의 1차원 합성곱 계

층 및 2개의 풀링 계층으로 설계된 모델을 학습했으며, 과적합을 피하기 위해 2개의 드롭아웃을 사용했다. 모델 학습의 에포크(Epoch)는 300, 학습률 0.001, 최적화 함수는 Adam을 사용했다.



그림 8. 얼굴 근육의 근전도 측정 예시

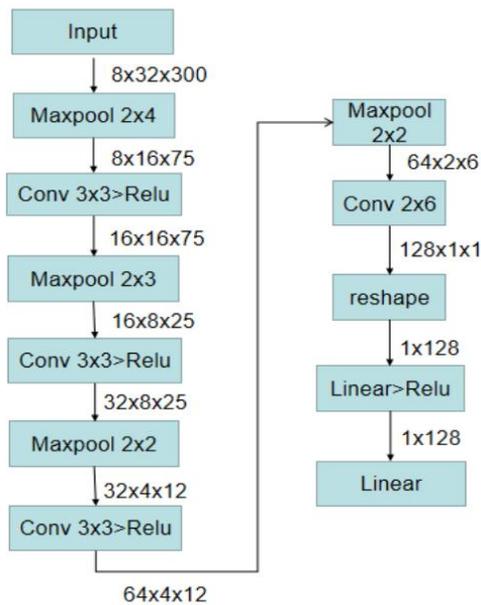


그림 9. 단일 동작 근전도 기반 사용자를 인식하기 위한 신경망 구조

Lu[22]는 단일 동작의 근전도를 이용해 사용자 인식 연구를 진행했다. MYO 암밴드를 이용해 21명의 피험자가 손 열기 동작을 수행할 때 근전도를 8채널로 측정했으며, 센서는 오른쪽 팔뚝의 동일한 위치에 배치했다. 샘플링 레이트는 200Hz 설정하고, 동작을 30회 반복했다. 측정된 근전도는 더 많은 특징을 사용하기 위해 32*300 크기의 연속 웨이블릿 변환 이미지로 변환했다. 그림 9와 같이 4개의 합성곱 계층과 풀링 계층으로 구성된 합성곱 신경망을 설계했으며, 처음 3개의 합성곱 계층은 3*3 크기의 필터, 4번째 합성곱 계층은 2*6 크기의 필터를 사용했다. 활성화 함수로 ReLU(Rectified linear unit)를 사용했으며, 21명을 99.2%의 정확도로 인식했다.

핸드크래프트 특징과 신경망을 같이 사용한 연구로 Kim[1]의 연구가 있다. Kim은 사람이 걸어 다닐 때 허벅지 근육에서 측정한 근전도를 이용해 사용자 인식 연구를 진행했다. 20명의 피험자는 rectus femoris, vastus medialis, vastus lateralis, semitendinosus에 센서를 부착한 후 복도를 걸어 다닐 때 Biopac Mp150 장비로 신호를 측정했다. 측정된 근전도에서 RMS, MAV, VAR, ZC, SSI 등 12개의 핸드크래프트 특징을 추출하고, 30개의 은닉층으로 구성된 인공 신경망을 이용해 사용자를 99.7% 정확도로 인식했다.

Shin[23]은 주먹 쥐기 동작을 수행할 때 측정한 근전도를 이용했다. 자체 제작한 장비를 이용해 10명의 피험자에게 2채널로 근전도를 측정했으며, 15Hz 고역 통과 필터, 450Hz 저역 통과 필터, 60Hz 대역 저지 필터를 이용해 잡음을 제거했다. Mean, VAR, ZC 등 5개의 시간 도메인 특징을 추출했으며, 설계한 인공 신경망으로 실험하여 95%의 정확도를 나타냈다.

Raurale[24]은 측정한 근전도 및 근전도 공개 데이터를 결합하여 피험자 65명의 근전도를 이용한 실험을 진행했다. 피험자는 의자에 앉아있는 편안한 상태로 MYO 암밴드 착용한 후 8채널로 근전도를 측정했다. 손 열기, 손 닫기 등 8개의 손 동작을 수행했으며, RSS(Root absolute sum square) 및 KFD(Kernel fisher

discrimination)를 이용해 특징을 추출했다. 추출된 특징은 LDA를 이용해 차원을 축소하고, MLP(Multi-layer perceptron) 및 RBF(Radial basis function)를 이용해 사용자를 분류했다. 다수결 의사 결정(Majority voting decision)을 위해 5번 반복 수행하며, 92.08%의 정확도로 사용자를 인식했다.

표 3. 신경망을 이용한 사용자 인식 연구

저자	데이터	세션	채널	피험자 수	특징	분류기	성능
Wong et al.[19]	비공개	단일	8	7	ANN	ANN	Acc: 98.09%
Shioji et al.[20]	비공개	멀티	8	8	CNN	CNN	Acc: 94.6%
Morikawa et al.[21]	비공개	멀티	3	5	CNN	CNN	Acc: 47.66%
Lu et al. [22]	비공개	단일	8	21	CWT, CNN	CNN	Acc: 99.2%
Kim et al. [1]	비공개	단일	4	20	RMS, MAV, VAR, etc. 12ea	ANN	Acc: 99.7%
Shin et al.[23]	비공개	단일	2	10	VAR, Mean, ZC, MF, length	ANN	Acc: 95%
Raurale et al.[24]	공개	멀티	8	65	RSS, KFD	MLP, RBF	Acc: 92.08%

신경망을 이용한 사용자 인식 연구 방법은 표 3과 같다. 근전도를 전처리 후 신경망을 이용한 방법 및 핸드크래프트 특징과 신경망을 동시에 이용한 방법으로 사용자 인식 연구가 진행되고 있다. 그러나 일반적인 신경망은 모델을 학습하기 위해 대용량의 데이터가 필요하며, 새로운 사용자가 추가되면 모델을 재학습해야 하는 단점이 있다.

제3절 삼 구조를 이용한 데이터 쌍의 유사도 학습

일반적인 신경망은 강력한 모델을 학습하기 위해 많은 수의 데이터가 필요하지만, 사용 가능한 데이터 수는 제한되어 있는 단점이 있다. 또한, 사용자 인식을 위해 등록된 사용자 그룹은 항상 새로운 사용자가 합류하고 일부 사용자가 떠나기 때문에 동적이므로, 사용자 그룹이 변경될 때 모델을 다시 학습시켜야 하는 재학습 문제가 발생한다. 이러한 단점을 해결하기 위한 삼 네트워크는 모델을 학습할 때 대용량 데이터 세트가 필요하지 않은 대표적인 퓨샷(Few-shot) 학습 방법이다. 삼 네트워크는 입력된 데이터 쌍의 유사도를 비교해 사용자를 인식하여 새로운 사용자가 추가되어도 모델을 재학습할 필요가 없는 장점이 있다[18].

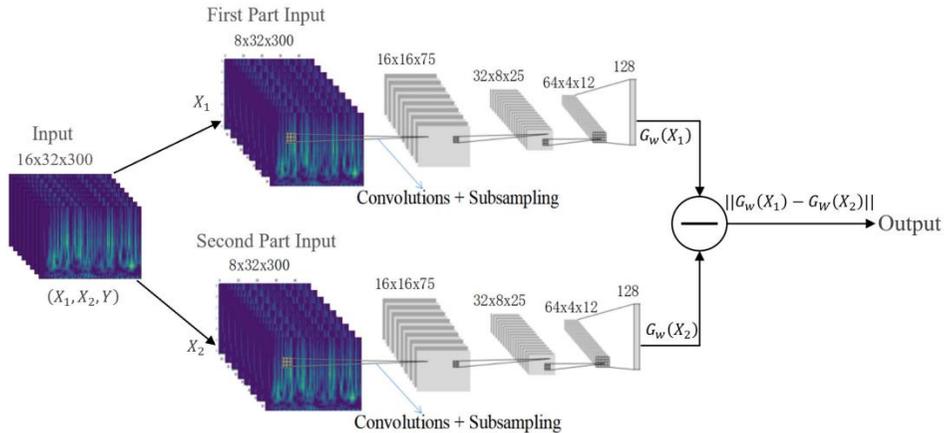


그림 10. 2차원 합성곱 신경망 기반 삼 네트워크 예시

Lu[11]는 모델을 학습할 때 사용하지 않는 피험자의 근전도를 이용한 삼 네트워크 기반 사용자 인식 실험을 진행했다. 21명의 피험자가 손 열기 동작을 수행할 때 MYO 암밴드로 근전도를 측정해 사용했다. 피험자는 8채널 센서를 오른팔의 radiohumeral joint 높이에 배치하고, 동작을 30회 반복했다. 사용자 인식을 위해

설계한 삼 네트워크는 그림 10과 같이 4개의 2차원 합성곱 계층 및 풀링 계층 등으로 구성했다. 활성화 함수로 ReLU를 사용했으며, 삼 네트워크 학습은 각 피험자의 데이터에서 무작위로 결합한 100개의 데이터 쌍을 사용했다. Positive와 negative 데이터 비율은 1:1로 구성했으며, 두 하위 네트워크(Sub network)의 출력은 유클리디안 디스턴스를 이용해 계산하였다. 삼 네트워크 모델을 학습한 후 새로운 4명의 피험자에게 근전도를 측정 한 후 사용자 인식 실험을 진행했으며, 97.76%의 정확도를 달성했다.

Fan[18]은 40명의 피험자가 스마트폰을 들고 사용할 때 측정 한 근전도로 삼 네트워크 기반 사용자 인식 연구를 진행했다. MY0 앰밴드를 이용해 8채널로 근전도를 측정했으며, 근전도 취득 센서의 회전에 대해 탄력적으로 작동하도록 데이터 증강 기법을 이용하였다. 측정 한 근전도 데이터는 정보를 손실하지 않기 위해 잡음 제거를 수행하지 않았다. 그림 11과 같이 필터 수가 다른 3개의 합성곱 계층으로 설계된 삼 네트워크를 이용했으며, 첫 번째 합성곱 계층은 채널간 특징을 학습하기 위해 8*1 크기의 필터를 사용했다. 활성화 함수로 ReLU를 사용했으며, 과적합을 방지하기 위해 드롭아웃을 사용했다. 두 개의 하위 네트워크의 출력 차이는 유클리디안 디스턴스를 이용해 계산했으며, 92.06%의 정확도를 달성했다.

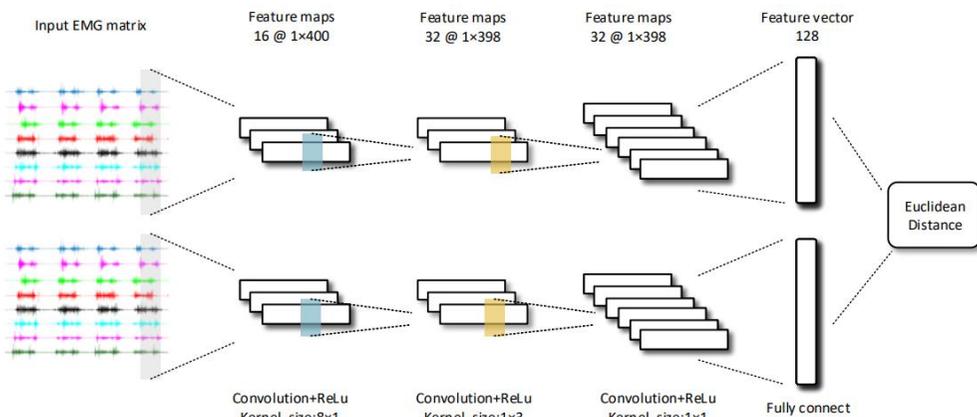


그림 11. 1차원 합성곱 신경망 기반 삼 네트워크 예시

Zhong[25]은 삼 네트워크를 이용해 장문 기반 사용자 인식 실험을 진행했다. 학습된 네트워크를 활용한 방법으로, 5개의 합성곱 계층과 5개의 풀링 계층으로 구성된 VGG-16을 삼 구조로 설계했다. 모델은 완전 연결 계층 2(fc2) 및 완전 연결 계층 3(fc3)을 전이 학습했다. 학습할 때 데이터 쌍은 모델에서 무작위로 선택하였으며, positive와 negative 비율은 1:2로 구성했다. 일상 환경에서 구축한 XJTU 장문 데이터를 이용해 실험했으며, 4.56%의 동일 오류율을 달성했다.

Maiorana[26]는 18개월의 시간 간격을 두고 멀티 세션(Multi-session)으로 측정한 뇌전도를 이용해 삼 네트워크 기반 사용자 인식 실험을 진행했다. 사용한 데이터는 256Hz 샘플링 레이트에서 작동하는 19채널 습식 전극이 있는 GALILEO BE Light 장비를 이용해 5개 세션으로 측정했다. 피험자 45명은 뇌전도를 측정할 때 눈을 감은 상태(EC, Eye close)를 유지했다. 데이터는 64Hz로 다운 샘플링하고, 8-30Hz 대역의 대역 통과 필터 및 CAR(Common average referencing)을 이용해 잡음을 제거했다. 전처리 후 데이터는 80%로 윈도우를 중첩해 분할했으며, 동일한 합성곱 신경망을 삼 네트워크 구조로 배치해 사용했다. 합성곱 신경망은 5개의 합성곱 계층과 4개의 최댓값 풀링 계층, 1개의 드롭아웃으로 설계했으며 두 개의 호환성 함수 사이의 거리는 유클리디안 디스턴스를 이용했다. 네트워크 학습은 교차 검증 접근 방식으로, 데이터에서 무작위로 선택한 30명 피험자의 데이터를 학습에 사용하고, 나머지 15명의 데이터를 테스트에 사용했다. 설계한 삼 모델로 30초 길이의 뇌전도를 이용한 결과, 7% 미만의 동일 오류율(Equal error rate)을 달성했다.

Abdulrazzaq[27]는 모델의 학습 속도를 줄이기 위해 가중치를 공유하는 삼 네트워크 기반 멀티모달 사용자 인식 실험을 진행했다. 얼굴 및 손바닥 정맥 이미지를 이용한 사용자 인식 연구로, 2개의 하위 네트워크의 입력으로 동일한 이미지를 사용하지 않고 얼굴-손바닥 정맥 이미지의 쌍을 이루어 사용하였다. 입력 데이터는 64*64 크기의 흑백 이미지를 사용했다. 하위 네트워크는 4개의 합성곱 계층과

4개의 풀링 계층으로 구성되고, 소프트맥스(Softmax)를 이용해 유사도를 계산하였다. 활성화 함수로 ReLU를 사용하여 50명의 피험자에 대해 99.33%의 정확도를 나타냈다.

Acien[28]은 모델을 학습하기 위한 사용자의 데이터 수가 적은 경우 적합한 방법인 삼 순환 신경망으로 사용자 인식 연구를 진행하였다. 키스트로크 정보를 이용한 사용자 인식 실험으로, 사용자가 키보드를 타이핑할 때 5개 특징(유지 시간, 입력간 대기 시간, 입력 대기 시간, 해제 대기 시간, 키 코드)을 신경망에 입력 데이터로 사용했다. 삼 순환 신경망은 128개의 히든 유닛으로 구성된 두 개의 장단기 메모리로 구성되며, 장단기 메모리 계층 사이에는 과적합을 피하기 위해 배치 정규화(Batch normalization) 및 드롭아웃을 사용했다. 모델의 유사도를 비교하기 위한 호환성 함수는 1×128 크기로, 유클리디안 디스턴스로 유사도를 비교했다. 모델을 학습할 때 사용한 데이터 쌍은 무작위로 선택하였으며, positive와 negative 비율은 1:1로 하였다. 실험은 피험자마다 데이터 쌍을 5개 이용해 실험했으며, 4.8%의 동일 오류율 성능을 확인했다.

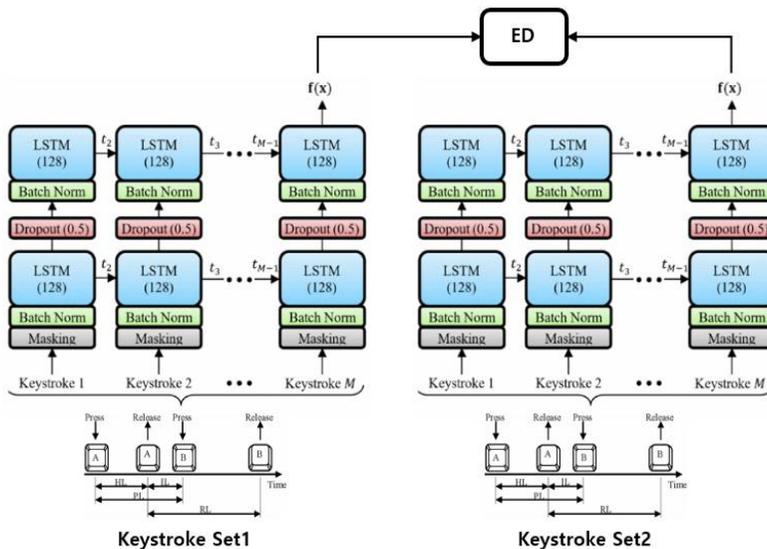


그림 12. 장단기 메모리 기반 삼 네트워크 예시

Chakladar [29]는 서명 및 뇌전도를 이용하는 멀티모달 삼 네트워크 기반 사용자 인식 연구를 진행했다. 서명 이미지는 155*200 크기로 리사이즈했으며, 뇌전도는 14채널로 측정해 대역 통과 필터(8-12Hz 대역)를 사용했다. 사용자 인식 데이터를 모델링하기 위한 positive 및 negative 데이터 쌍은 무작위로 생성했다. 설계한 삼 네트워크에서 합성곱 신경망은 서명 이미지에서 공간적 특징을 추출하고 장단기 메모리는 뇌전도의 시간적 관계 특징을 추출해 사용했다. 합성곱 신경망은 3개의 2차원 합성곱 계층으로 설계했으며, 장단기 메모리는 히든 유닛 75개 및 55개로 구성된 2개의 계층으로 설계했다. 두 인코더(합성곱 신경망 및 장단기 메모리)의 출력은 추가 특징 처리를 위해 공통 특징 공간에서 융합했다. 두 개의 하위 네트워크 출력은 유사도를 비교하기 위해 대조 손실 함수(Contrastive loss function)를 이용했다. 삼 모델의 다양한 파라미터는 학습을 통해 최적의 값을 선택했으며, 70명의 피험자를 98.57%의 정확도로 인식했다.

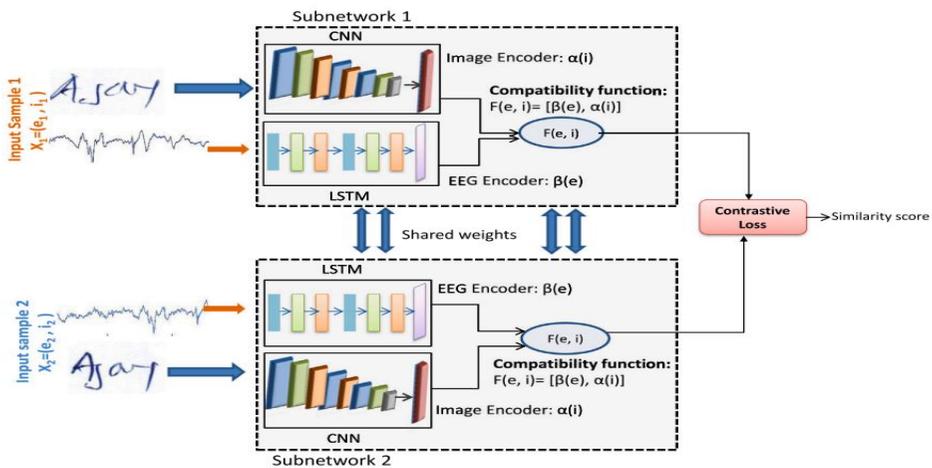


그림 13. 합성곱 신경망-장단기 메모리 앙상블 기반 삼 네트워크 예시

Zakeri [30]는 사용자가 말하는 동안의 입술 움직임 정보를 이용한 사용자 인식 연구를 진행했다. 삼 네트워크의 하위 네트워크 구조는 2개 스트림으로, 입술 이

미지와 이미지에서 추출한 24개의 랜드마크 특징점을 입력으로 사용하였다. 이미지는 3개의 합성곱 계층과 2개의 양방향 게이트 순환 유닛(GRU, Gated recurrent units) 계층으로 구성된 모델로 연산하였고, 랜드마크 특징점은 1개의 장단기 메모리 계층으로 연산하였다. 그 후 이를 결합하여 호환성 함수로 출력했다. 동일한 계층으로 구성된 삼 네트워크는 입력된 데이터 쌍의 유사도를 유클리디안 디스턴스로 계산하여, 11명의 피험자에 대해 95.41%의 정확도를 나타냈다.

Sriskandaraja[31]는 음성을 이용한 사용자 인식 시스템에서 위조된 음성 파일을 이용한 스푸핑을 방지하기 위해 위조된 음성을 감지하는 연구를 진행했다. 음성 데이터는 고속 푸리에 변환(FFT, Fast fourier transform)을 통해 864*400 크기의 log power sepectrum으로 생성하여 설계한 모델의 입력으로 사용하였다. 설계한 모델은 5개의 합성곱 레이어와 5개의 풀링 계층으로 구성되고, 과적합을 피하기 위해 드롭아웃 계층을 사용했다. 2개의 하위 네트워크 출력은 소프트맥스를 이용해 위조된 음성을 분류하였다. 학습에 사용한 positive 및 negative 데이터 쌍은 1:1 비율이며, 활성화 함수로 ReLU를 사용하였다.

표 4는 기존 삼 네트워크를 이용한 유사도 학습 연구를 정리한 내용으로, 기존 사용자 인식 연구는 모델 학습에 사용하지 않은 새로운 사용자를 적은 수의 데이터로 인식하기 위해 연구되어 왔다. 두 개의 하위 네트워크는 모델에 입력되는 데이터 쌍에서 동일한 크기의 호환성 함수를 생성하고, 유사도를 비교하였다. 유사도 비교는 유클리디안 디스턴스나 코사인 유사도 등 유사도 비교 방법 및 완전 연결 계층을 사용하였다.

기존 근전도를 이용한 사용자 인식 연구는 핸드크래프트 특징을 이용한 방법 및 신경망과 삼 네트워크를 이용한 방법 등 다양한 방법으로 진행되고 있다. 그러나 모델의 입력인 근전도는 비선형성 특성을 고려하지 않고 그림 14와 같이 모두 동일한 가중치로 특징을 추출하였다. 이러한 신호의 비선형성 특성을 고려하지 않은 방법은 사용자 인식 성능을 저하시키는 데이터 불규칙성 문제가 발생한다[13].

표 4. 삼 네트워크를 이용한 유사도 학습 연구

저자	데이터 종류	세션	피험자 수	특징 및 분류	유사도 비교	성능
Lu et al.[11]	근전도	단일	4	CWT, CNN	ED	Acc: 97.76%
Fan et al. [18]	근전도	단일	40	CNN	ED	Acc: 92.06%
Zhong et al. [25]	장문	단일	114	CNN	fc	EER: 4.56%
Maiorana et al.[26]	뇌전도	멀티	15	CNN	ED	EER: 7% under
Abdulrazzaq et al.[27]	얼굴, 정맥	단일	50	CNN	fc	Acc: 99.33%
Acien et al. [28]	키스트 로크	단일	1000	LSTM	ED	EER: 4.8%
Chakladar et al.[29]	서명, 뇌전도	단일	70	CNN, LSTM	ED	Acc: 98.57%
Zakeri et al. [30]	입술	단일	11	CNN, GRU	ED	Acc: 95.41%
Sriskandaraja et al.[31]	목소리	-	-	CNN	fc	위조 검출

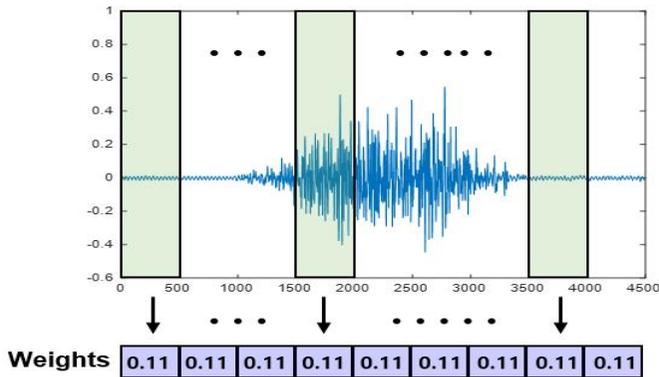


그림 14. 신호의 비선형성 특성을 고려하지 않은 특징 추출 예시

제3장 제안하는 어텐션 메커니즘 기반 근전도를 이용한 사용자 인식

본 장에서는 신호의 시간-주파수 도메인 특징을 추출하고, 학습된 모델 기반 새로운 사용자가 추가되면 재학습이 필요한 문제를 극복 가능한 방법을 설명한다. 또한, 근전도의 불규칙성 문제에 대한 해결 가능한 방법을 설명하고 제안한 방법이 적용된 사용자 인식 방법에 대해 설명한다. 그림 15는 제안하는 보조 분류기를 결합한 멀티 스트림 삼 용합 네트워크 기반 사용자 인식 방법의 흐름도를 나타낸다.

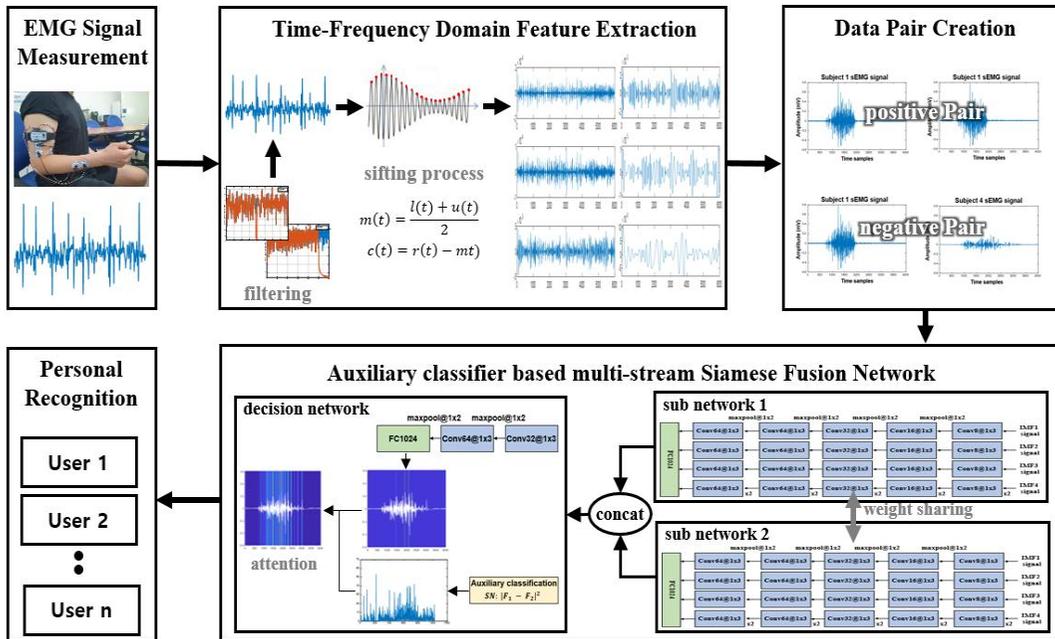


그림 15. 제안하는 사용자 인식 방법 흐름도

동작을 수행할 때 발생하는 근전도로 사용자를 인식하기 위해 신호에 포함되어 있는 대표적인 잡음들을 필터링 기술로 전처리한다. 그 후 데이터의 시간-주파수 도메인 특징을 추출하기 위해 경험적 모드 분해로 근전도를 처리하며, 내재 모드 함수 단위로 추출한다. 내재 모드 함수는 삼 모델을 학습하기 위해 positive와 negative로 데이터 쌍을 이루어 설계한 보조 분류기가 결합된 멀티 스트림 삼 용 합 네트워크의 입력 데이터로 사용한다. 설계한 모델은 재학습 문제를 해결하기 위해 삼 구조로 설계된 하위 네트워크와 신호의 불규칙성 문제를 해결하기 위해 보조 분류기가 결합된 decision 네트워크로 구성된다. 가중치를 공유하는 두 개의 하위 네트워크는 입력된 데이터 쌍에서 호환성 함수를 계산하고, 계산된 정보는 decision 네트워크로 전달한다. Decision 네트워크의 보조 분류기는 호환성 함수를 유클리디안 디스턴스로 연산하고, 합성곱 신경망으로 구성된 신경망과 결합하여 어텐션 메커니즘을 적용한다. 그 후 유사도를 계산하여 모델을 학습한다.

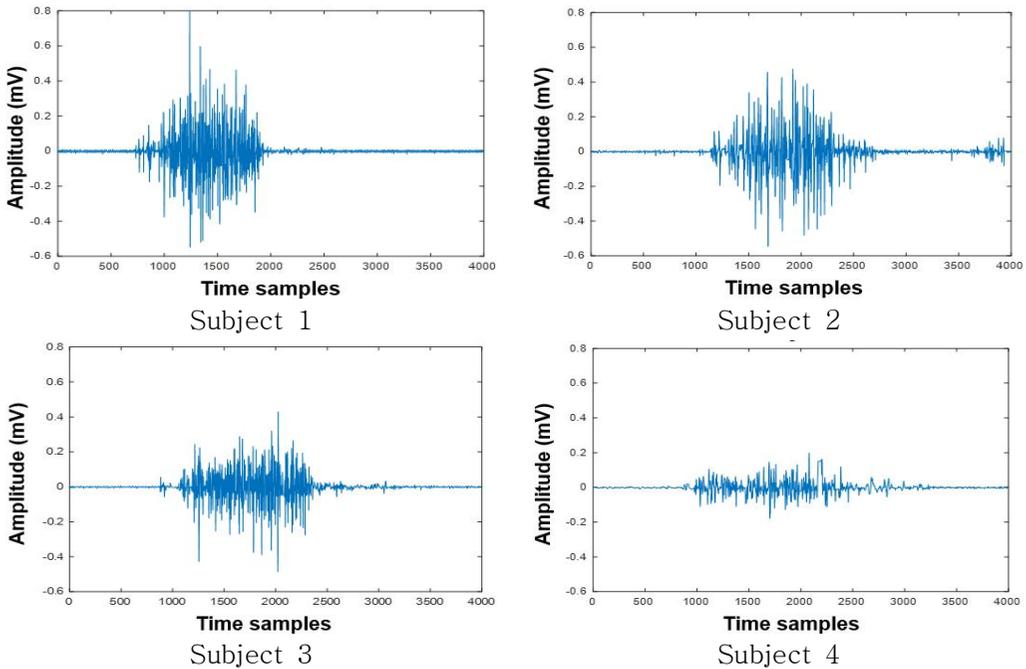
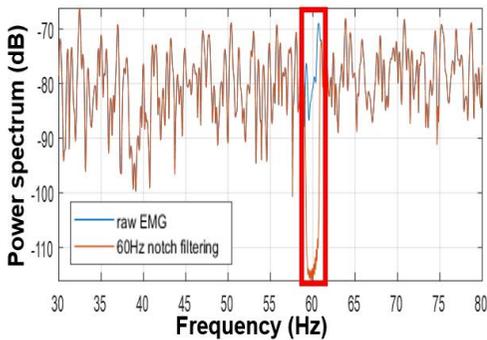


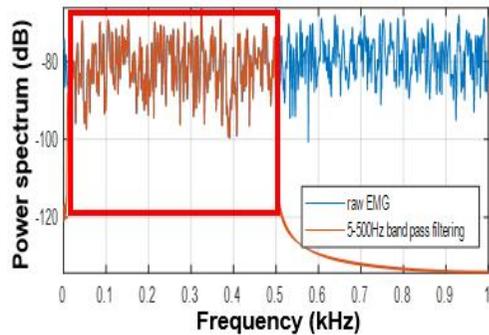
그림 16. 동일한 동작을 수행한 피험자들의 근전도 파형

사용자 인식을 위해 학습된 모델에 근전도가 입력되면 학습 과정과 동일하게 필터링을 이용해 신호를 전처리하고 내재 모드 함수로 특징을 추출한다. 그 후 데이터 쌍을 이루어 학습된 모델에 입력 데이터로 사용된다. 학습된 모델은 입력된 데이터 쌍의 유사도를 계산하여 최종 사용자 인식을 진행한다.

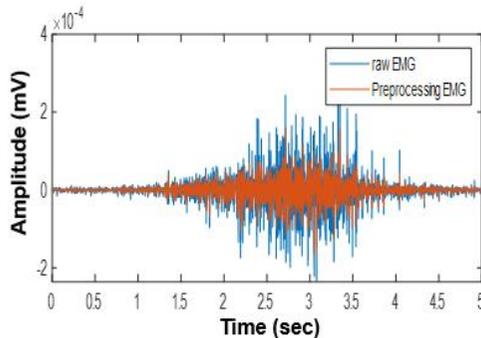
근전도를 이용한 연구는 동작 인식[12][19], 사용자 인식[11][18], 재활치료[32], 기기 제어[33] 등 다양한 분야에서 진행되고 있으며, 사용자 인식은 근육의 발달 정도와 활성도, 습관 등에 의해 발생하는 개인간 차이를 이용한다[34]. 그림 16은 동일한 동작을 수행할 때 서로 다른 피험자의 근전도를 나타내며, 피험자 간 서로 다른 신호 파형이 나타남을 확인할 수 있다.



(a) 대역 저지 필터가 적용된 근전도



(b) 대역 통과 필터가 적용된 근전도



(c) 전처리된 근전도

그림 17. 시간 및 주파수 도메인에서의 전처리에 따른 근전도 비교

근전도는 비침습적 방법으로 측정이 가능하여 통증이 없으며, 신호 측정이 편리한 장점이 있다. 그러나 특정 근육 위치에서 신호를 측정하기 위해 피부 표면에 전극을 부착하여 피부 임피던스가 발생한다. 또한, 측정된 근전도는 전력선 간섭, 백색 가우시안 잡음, 기준선 변동 등 다양한 요인에 의해 신호가 오염된다 [10][35]. 본 논문에서는 60Hz 대역에 대역 저지 필터를 사용해 전력선 간섭을 제거한다. 그리고 근전도에서 개인별로 중요한 정보가 많이 포함된 5-500Hz 대역에 대역 통과 필터를 이용해 다른 요인의 영향을 줄인다[36]. 그림 17은 근전도의 전처리 전·후에 따른 시간 및 주파수 도메인에서의 신호 변화를 나타낸다. 그림 17에서 파란색 선은 원시 근전도(Raw EMG), 빨간색 선은 잡음이 제거된 근전도(Preprocessed EMG)를 나타낸다. 근전도에 대역 저지 필터 및 대역 통과 필터를 적용한 후 원신호는 3.83% 손실되었으나, 잡음은 83.27% 제거되었다.

제1절 경험적 모드 분해를 이용한 특징 추출

경험적 모드 분해는 실제 응용 분야에서 발생하는 신호를 시간-주파수 도메인에서 분석 가능한 방법으로, 위상학적으로 정현파 신호의 진폭 및 주파수 특성을 변조하는 것과 동일하게 여러 진동 모드로 데이터를 적응적으로 분리한다 [37][38]. 수학적으로 식 1과 같이 내재 모드 함수 및 잔차(Residual)의 합은 원신호로 설명된다.

$$x(t) = \sum_{i=1}^n IMF_i + residual \quad (1)$$

Hilbert-Huang 변환을 통해 데이터의 국소 주파수와 진폭을 계산하며[38], 데이터에서 분해된 각 내재 모드 함수는 하나의 진동 모드를 가진다. 이러한 구조는 기존 필터뱅크 방법과 유사하지만, 내재 모드 함수는 미리 정해진 주파수 대역이 아닌 신호의 고유한 진동 구조를 기반으로 계산되어 대역이 제한되지 않는 차이가 있다[39]. 잔차는 단조 함수로 구성되며, 데이터의 경향성을 반영한다[40]. 낮은 차수의 내재 모드 함수는 입력된 신호의 고주파수 성분을 나타내고, 높은 차수의 내재 모드 함수는 저주파수 성분을 나타낸다[41]. 경험적 모드 분해는 체거름 과정을 통해 내재 모드 함수 및 잔차로 분해하며, 아래와 같이 5단계로 구성된다 [42].

- 1단계: 신호 $r_i(t)$ 에서 국소 최솟값 및 국소 최댓값 계산 (if $i=1, r_i(t) = x(t)$)
- 2단계: 국소 극값을 이용해 하부 envelop 및 상부 envelop 계산

$$\begin{aligned}
 l_i(t) &= [lmin_1, lmin_2, \dots, lmin_m] \\
 u_i(t) &= [lmax_1, lmax_2, \dots, lmax_m]
 \end{aligned}
 \tag{2}$$

- 3단계: 하부 envelop와 상부 envelop의 평균 $m_i(t)$ 계산

$$m_i(t) = (l_i(t) + u_i(t)) / 2
 \tag{3}$$

- 4단계: $r_i(t)$ 와 $m_i(t)$ 차이 $c_i(t)$ 를 계산

$$c_i(t) = r_i(t) - m_i(t)
 \tag{4}$$

- 5단계: 계산된 $c_i(t)$ 가 내재 모드 함수 조건을 만족하는지 확인

- 조건을 만족하지 않으면 $r_i(t) = c_i(t)$ 로 대입하여 1단계부터 재반복
- 조건을 만족하면 $c_i(t)$ 는 IMF_i 로 하며, $r_{i+1}(t)$ 를 계산하여 IMF_{i+1} 계산을 위해 1단계로 이동

$$r_{i+1}(t) = r_i(t) - c_i(t) \quad (5)$$

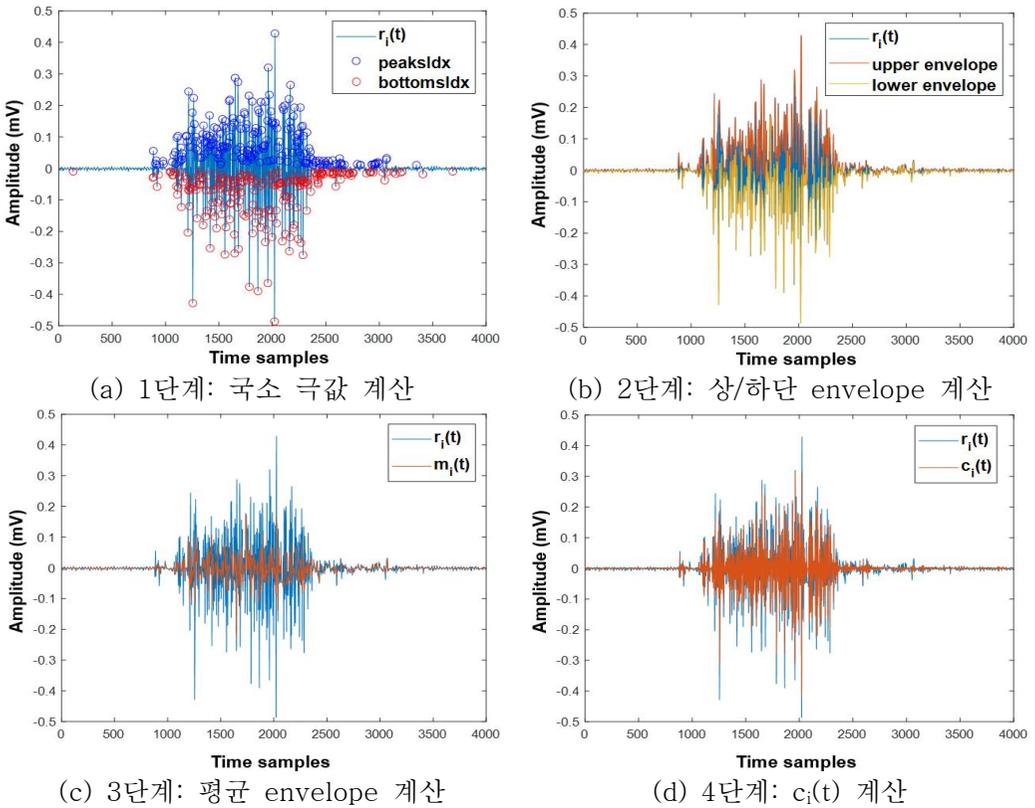
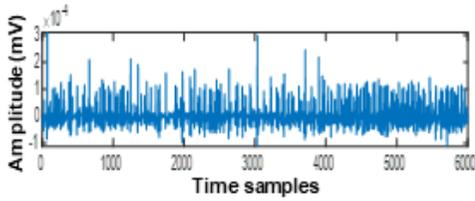


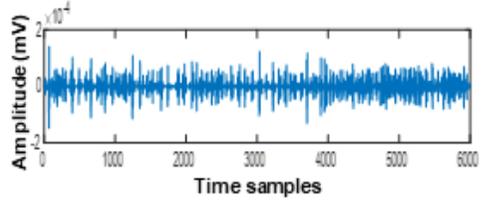
그림 18. 체거름 과정을 통한 내재 모드 함수 분해 과정

분해된 내재 모드 함수는 아래 두 가지 조건을 충족하는 함수로 정의된다. 그림 18은 체거름 과정을 진행할 때 근전도의 계산 과정을 보여주며, 그림 19는 근전도에 경험적 분해 모드를 적용하여 분해한 내재 모드 함수와 잔차를 나타낸다.

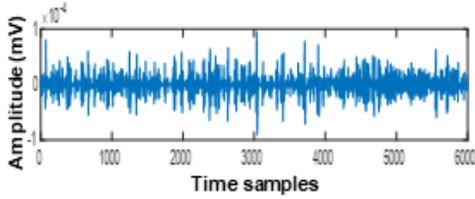
그림 19에서 내재 모드 함수 1-4에 근전도의 주요 정보가 많이 있음을 확인할 수 있다[16][43].



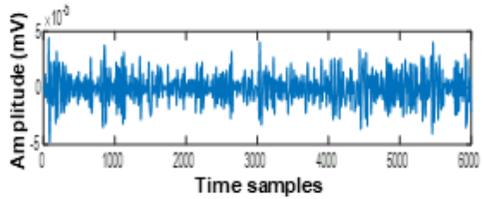
(a) 원시 근전도



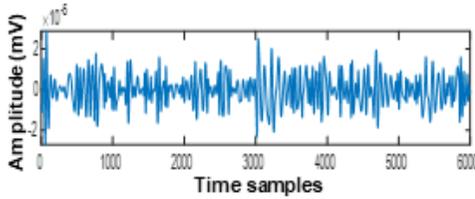
(b) 내재 모드 함수 1



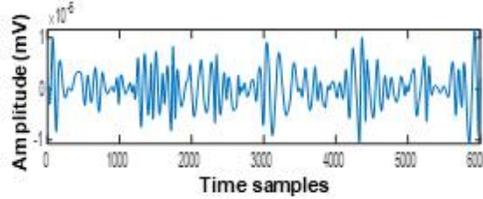
(c) 내재 모드 함수 2



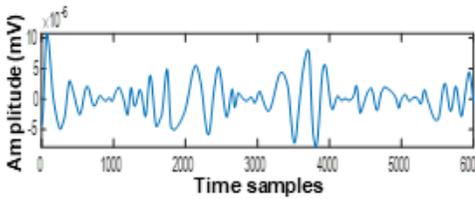
(d) 내재 모드 함수 3



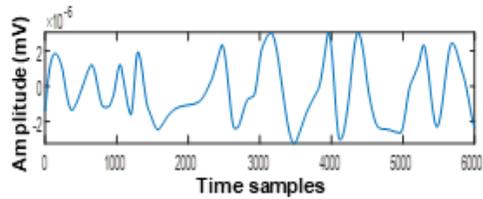
(e) 내재 모드 함수 4



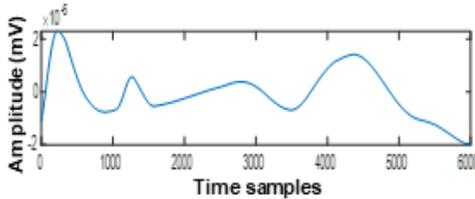
(f) 내재 모드 함수 5



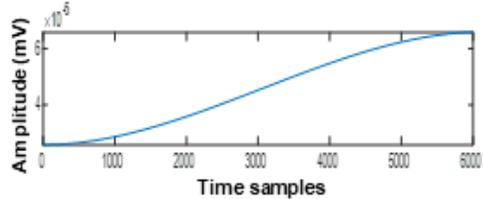
(g) 내재 모드 함수 6



(h) 내재 모드 함수 7



(i) 내재 모드 함수 8



(j) 잔차

그림 19. 근전도의 내재 모드 함수 및 잔차

- 조건 1: 극값의 수와 영점 교차의 수는 동일하거나 1개 이하의 차이가 발생함
- 조건 2: 하단 envelop와 상단 envelop의 평균값은 전체 시계열에서 0이어야 함

제2절 보조 분류기를 결합한 멀티 스트림 삼 용합 네트워크

근전도를 이용한 연구에서 신경망은 합성곱 신경망, 순환 신경망, 장단기 메모리 등 다양한 구조로 설계된 딥러닝 모델을 사용했다. 이러한 딥러닝 모델은 기존 핸드크래프트로 제작된 특징 추출 방법과 다르게 입력된 데이터에서 사전 파라미터 설정 없이 특징을 추출한다. 합성곱 신경망은 다수의 합성곱 계층 및 일반적인 신경망 계층으로 구성된다. 합성곱 계층은 입력 데이터에서 정해진 필터를 이용해 특징을 추출하며, 1차원 및 다차원 구조의 데이터 처리가 가능하다. 이에 영상 및 음성, 신호 처리 분야 등 다양한 분야에서 높은 성능을 나타낸다. 순환 신경망은 과거 출력 데이터를 재귀적으로 참조하여, 시간의 흐름에 따라 생성된 데이터를 처리하기 위한 모델이다. 구조는 one to many, many to one 등이 있으며, 음성 및 문자열 처리 분야 등에서 활용된다. 장단기 메모리는 순환 신경망의 단점을 보완하기 위해 연구된 모델로, 입력(Input) 및 망각(Forget) 게이트 등으로 구성된다. 이 중 망각 게이트를 이용해 이전 출력과 현재 입력된 데이터로 셀 상태를 계산하여 단점을 보완한다.

1. 합성곱 신경망 및 삼 네트워크 구조

합성곱 신경망은 다수의 계층으로 구성된 인공 신경망의 한 종류로, 3차원 텐

서(높이, 너비, 채널)를 이용해 일반 신경망에 비해 높은 성능을 나타낸다. 신경망은 풀링 계층과 합성곱 계층으로 구성되며, 풀링 계층은 채널 수 변경 없이 데이터 크기를 축소하는 역할을 수행한다. 합성곱 계층은 특징을 추출하는 기능인 필터(또는 커널)라고 불리는 행렬과 필터값을 비선형 값으로 변경하는 활성화 함수로 이루어진다. 필터는 입력된 데이터의 처음부터 끝까지 이동하며, 필터와 겹치는 부분의 데이터와 필터값을 합성곱 연산한다. 2차원 데이터에서는 왼쪽 위부터 오른쪽 아래까지 순차적으로 이동한다. 1차원 데이터에 대한 합성곱 계층의 연산은 그림 20과 같으며, 2차원 데이터의 필터 연산과 유사하게 데이터의 처음부터 끝까지 순차적으로 이동하며 데이터와 필터의 연산을 수행한다. 필터는 일반적으로 3*3 크기 또는 5*5 크기를 사용한다.

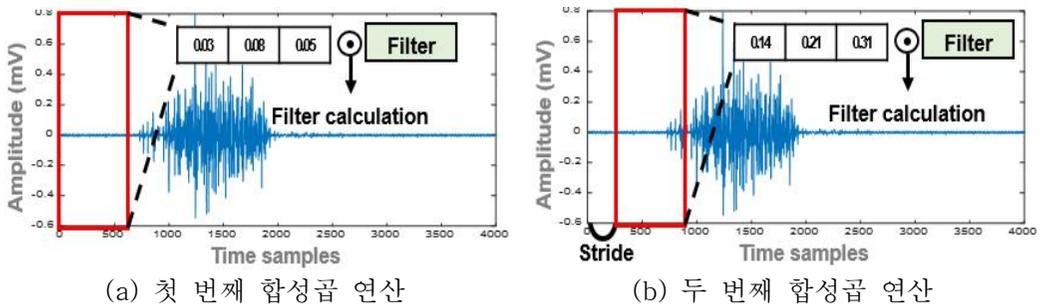


그림 20. 1차원 데이터를 이용한 합성곱 연산

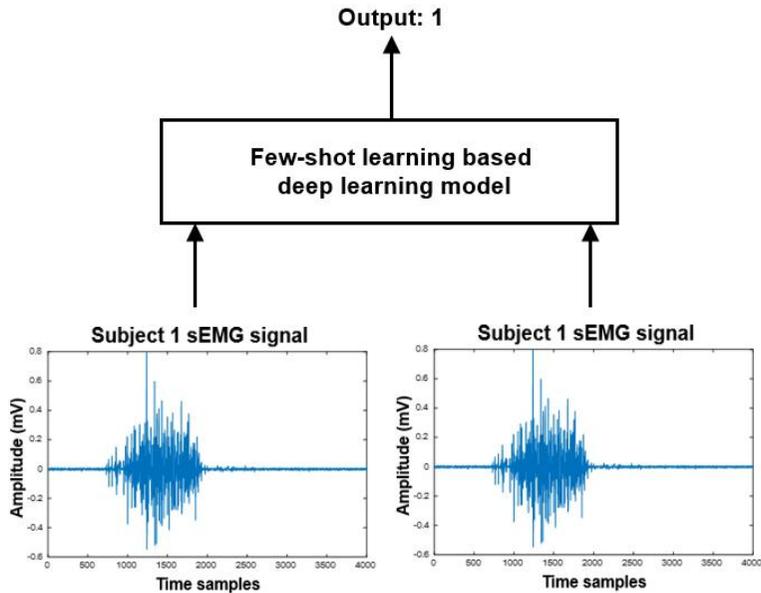
기본적인 필터의 이동 범위는 1차원 데이터에서 1개의 샘플링 데이터로 설정되며, 스트라이드(Stride)를 이용해 사용자가 변경할 수 있다. 입력된 데이터에서 합성곱 연산을 수행하면 입력보다 데이터 크기가 작아지는 특징이 있다. 다수의 합성곱 계층을 이용해 연산을 수행하면 출력되는 데이터 크기는 너무 작아지는 문제가 발생한다. 합성곱 연산 후에도 데이터 크기가 유지되도록 하는 파라미터는 패딩(Padding)으로, 패딩은 합성곱 연산을 수행하기 전 데이터 가장자리에 지정된 개수의 폭만큼 데이터를 추가한다. 입력된 데이터에서 합성곱 연산 후 출력되는

데이터 크기는 식 6으로 계산할 수 있다. 식에서 OH 는 합성곱 연산 후 출력되는 데이터 크기, H 는 입력 데이터 크기, FH 는 필터 크기를 의미한다.

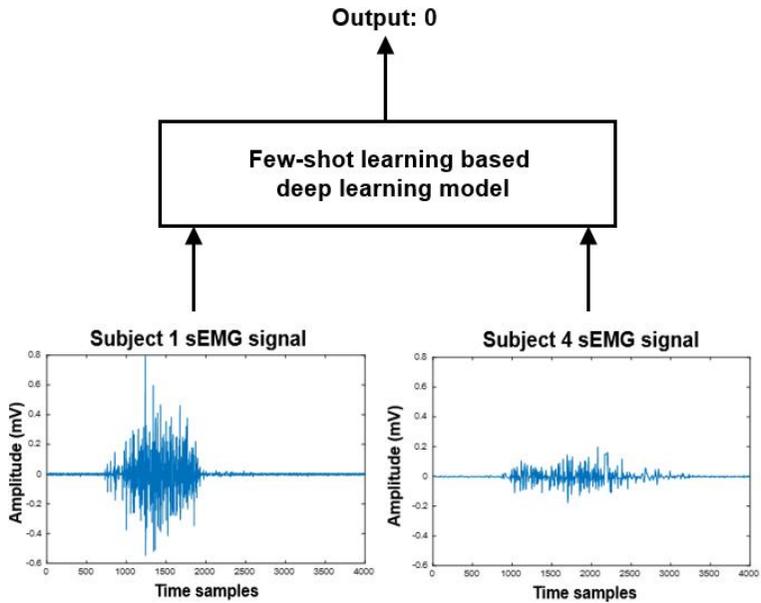
$$OH = \frac{H + 2 \times padding \times FH}{stride} + 1 \quad (6)$$

합성곱 신경망 및 장단기 메모리와 같은 딥러닝 모델은 사용자 인식을 위해 데이터의 특징 추출과 분류 작업에 많이 사용되고 있다. 그러나 이러한 딥러닝은 모델을 학습할 때 대용량의 라벨이 지정된 데이터가 필요하며, 짧은 세션에서 측정된 근전도로 구성된 데이터를 이용하면 모델을 학습하지 못할 수 있다. 거대한 데이터 세트로 모델을 학습할 때 데이터에 올바르게 라벨을 지정하면 많은 시간 및 비용이 들 수 있으며, 여러 샘플을 사용할 수 있는 경우에도 모든 데이터에 적응적인 특징을 계산하는 것은 계산 비용이 많아지는 단점이 있다. 또한, 새로운 사용자가 사용자 인식 응용 프로그램에 추가될 때마다 사용자 인식을 위해 모델을 재학습하는 것은 항상 가능하거나 실용적인 것은 아니다[29]. 결과적으로 더 적은 수의 데이터 샘플로 모델을 학습할 수 있고, 모델을 재학습시키지 않고도 새로운 사용자에게 확장할 수 있는 다른 접근 방식이 필요하다. 이러한 단점들을 해결하기 위해 동일한 가중치를 공유하는 삼 네트워크 모델 기반 퓨샷 학습 방법을 이용한 연구가 진행되고 있다.

퓨샷 학습은 적은 데이터만을 가지고 좋은 성능을 얻기 위한 방법들을 다루는 연구 분야로, 처음 마주하는 데이터 및 적은 데이터에 대해 일반화된 성능 확보가 가능한 방법이다. 입력 데이터의 유사도를 학습하는 방법으로, 두 개의 데이터로 구성된 데이터 쌍으로 딥러닝 연산을 통해 그림 21과 같이 유사도를 비교하여 동일하면 1, 동일하지 않으면 0으로 계산한다.



(a) 데이터 쌍이 유사한 경우



(b) 데이터 쌍이 유사하지 않은 경우

그림 21. 퓨샷 학습을 이용한 데이터 쌍 유사도 비교 결과 예시

상 네트워크는 유사도 메트릭(Metric)을 학습하기 위해 그림 22와 같이 동일한 구조를 가진 두 개의 하위 네트워크와 유사도를 계산하는 비용 모듈(Cost module)

로 구성된다. 두 개의 하위 네트워크들은 가중치를 서로 공유하며, 비용 모듈로 연결된다. 비용 모듈은 특징 표현을 기반으로 입력된 데이터 쌍의 유사도를 계산한다[44].

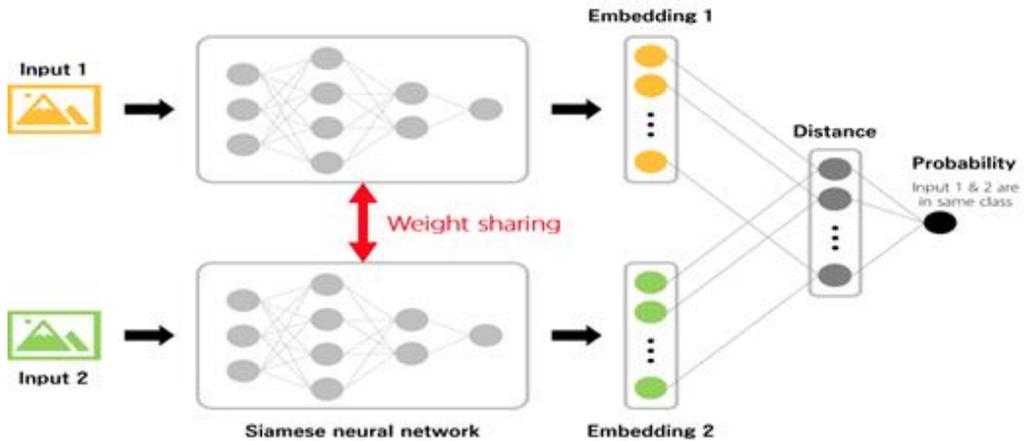


그림 22. 삼 네트워크 구조 예시

삼 네트워크는 데이터 쌍과 라벨 정보를 입력으로 사용한다. 데이터 쌍은 두 개의 동일한 하위 네트워크를 각각 통과하여 호환성 함수를 계산한다. 비용 모듈은 두 개의 하위 네트워크에서 출력된 함수를 이용해 식 7과 같이 유클리디안 디스턴스로 데이터의 유사도를 계산한다. 식에서 x_1 , x_2 는 입력되는 데이터 쌍, $F(x_1)$ 는 1번째 하위 네트워크에서 출력된 호환성 함수, $F(x_2)$ 는 2번째 하위 네트워크에서 출력된 호환성 함수를 의미한다.

$$D = |F(x_1) - F(x_2)|^2 \quad (7)$$

두 개의 하위 네트워크를 제어하는 파라미터 벡터에 대한 손실 함수(Loss function)는 식 8과 같이 라벨과 비용 모듈의 유사도 정보를 이용해 계산한다. 두

개의 하위 네트워크 파라미터 벡터는 계산된 손실 함수를 이용해 확률적 기울기 방법으로 업데이트한다[45]. 식에서 I 는 데이터 쌍이 동일한 클래스에 속하는지 라벨을 나타내는 함수로 동일하면 1, 동일하지 않으면 0으로 설정된다.

$$L(x_i, x_j, I) = I \times D^2 + (1 - I) \times \max(0, 1 - D)^2 \quad (8)$$

삼 네트워크는 입력된 데이터 쌍의 유사도를 반환하는 모델로, 기존 분류를 위한 신경망과 달리 소프트맥스 계층을 사용하지 않아도 된다. 이와 같은 방법으로 학습에 사용되지 않은 새로운 피험자가 모델에 추가되어도, 네트워크를 재학습하지 않고 사용자 인식이 가능하다. 이에 클래스가 매우 많은 경우거나 모델을 학습할 때 모든 클래스의 데이터를 사용할 수 없는 경우 적합한 방법이다[11].

2. 제안하는 보조 분류기를 결합한 삼 융합 네트워크

본 논문에서 제안하는 멀티 스트림 삼 융합 네트워크는 그림 23과 같으며, 2개의 하위 네트워크와 1개의 decision 네트워크로 구성된다. 제안한 모델에서 보조 분류기를 결합한 decision 네트워크는 하위 네트워크에 스택킹 양상블 구조로 설계한다. 하위 네트워크는 적은 수의 데이터를 이용해 모델 학습이 가능하며, 새로운 사용자가 시스템에 추가되어도 네트워크 재학습이 필요하지 않도록 삼 구조로 설계한다. Decision 네트워크는 근전도의 불규칙성 문제를 해결하기 위한 보조 분류기와 신경망으로 구성된다. 설계한 모델의 입력 데이터는 경험적 모드 분해를 이용해 추출한 내재 모드 함수 특징을 이용한다. 데이터는 positive 및 negative 데이터 구조로 데이터 쌍을 이루어 사용한다. 설계한 모델에서 2개의 하위 네트워크는 동일한 계층 구조를 이루며, 호환성 함수를 추출하기 위해 동일한 가중치로

모델을 학습한다. 하위 네트워크의 출력은 decision 네트워크의 입력 데이터로 사용되며, decision 네트워크는 보조 분류기와 신경망을 이용해 호환성 함수를 연산한다. 그 후 이를 결합하여 어텐션 메커니즘을 기반으로 입력된 데이터 쌍의 유사도를 비교한다.

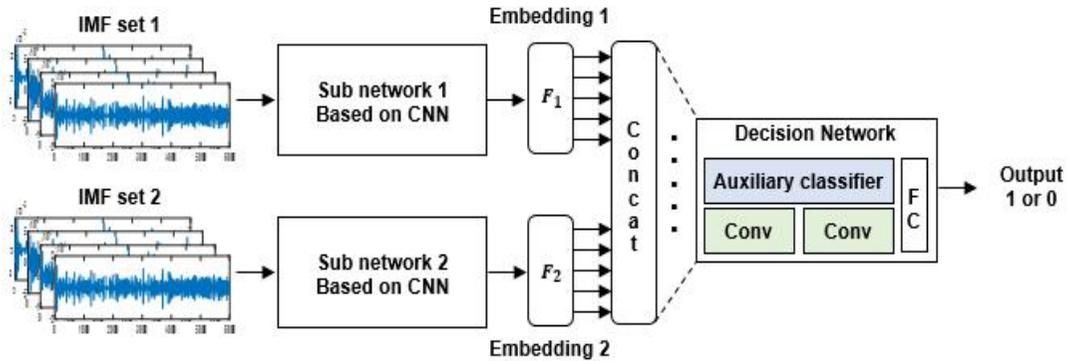


그림 23. 사용자 인식을 위해 제안하는 멀티 스트림 삼 용합 네트워크 구조

설계한 삼 용합 네트워크는 딥 인코더를 사용해 내재 모드 함수들의 특징을 추출하고, 데이터 쌍의 유사도를 이용해 모델을 학습하도록 설계하였다. 설계한 삼 용합 네트워크는 아래와 같이 동작한다.

- 1단계: 내재 모드 함수 세트로 구성된 데이터 쌍($IMFs_1, IMFs_2$)은 설계한 멀티 스트림 삼 용합 네트워크의 각 스트림에 입력 데이터로 사용

$$\begin{aligned}
 IMFs_1 &= [IMF1_1, IMF2_1, IMF3_1, IMF4_1] \\
 IMFs_2 &= [IMF1_2, IMF2_2, IMF3_2, IMF4_2]
 \end{aligned}
 \tag{9}$$

- 2단계: 딥 인코더(α)는 컨볼루션 계층을 이용해 입력된 데이터를 임베딩 공간 (η)으로 변환

$$\begin{aligned}
 \alpha_1 : IMF1 &\rightarrow \eta_1 \\
 \alpha_2 : IMF2 &\rightarrow \eta_2 \\
 \alpha_3 : IMF3 &\rightarrow \eta_3 \\
 \alpha_4 : IMF4 &\rightarrow \eta_4
 \end{aligned}
 \tag{10}$$

- 3단계: 임베딩 공간 정보를 1차원 벡터에 투영하기 위해 선형 계층(L)을 적용함. 이는 호환성 함수(F)로 출력

$$L : [\eta_1, \eta_2, \eta_3, \eta_4] \rightarrow F
 \tag{11}$$

- 4단계: 두 개의 하위 네트워크에서 생성된 호환성 함수(F_1, F_2)는 설계한 decision 네트워크의 신경망(DN)으로 특징(dF)을 추출

$$DN : [F_1, F_2] \rightarrow dF
 \tag{12}$$

- 4-1단계: 또한, 두 개의 호환성 함수(F_1, F_2)는 어텐션 메커니즘을 위해 보조 분류기로 연산

$$SN = |F_1 - F_2|^2
 \tag{13}$$

- 5단계: decision 네트워크의 보조 분류기로 계산된 SN 과 신경망으로 계산된 dF 를 결합하여 유사도($Matching\ score$)를 계산

$$L: [SN, dF] \rightarrow \text{Matching score} \quad (14)$$

- 6단계: positive와 negative 데이터 쌍으로 알려진 학습 데이터의 라벨(*Label*)과 유사도로 손실 함수(*Loss*)를 계산

$$Loss = Label - \text{Matching score} \quad (15)$$

- 7단계: 유사한 데이터 쌍의 경우 손실 함수가 작은 값을 표현하고, 유사하지 않은 데이터 쌍의 경우 큰 손실 함수를 표현하도록 모델 학습 진행

모델 재학습 문제를 해결하기 위해 설계한 2개의 하위 네트워크 구조는 그림 24와 같다. 설계한 하위 네트워크는 4개의 스트림으로 구성된 모델과 1차원 벡터에 투영하기 위한 완전 연결 계층으로 구성되어 있다. 변동성 있는 근전도에서 사용자를 구분할 수 있는 특징을 추출하기 위해 각 스트림은 10개의 합성곱 계층 및 4개의 풀링 계층으로 구성되며, 내재 모드 함수 세트에서 1개의 내재 모드 함수를 입력 데이터로 사용해 임베딩 공간으로 변환한다. 그 후 선형 계층을 통해 결합되며, 두 개의 하위 네트워크는 가중치를 공유하며 모델을 학습한다.



그림 24. 설계한 하위 네트워크 구조

표 5. 하위 네트워크 각 스트림의 계층별 출력 결과

Layer(type)	Output shape
Input(input layer)	(1, 8000, 1)
conv_1(convolutional layer)	(1, 8000, 8)
conv_2(convolutional layer)	(1, 8000, 8)
pool_1(maxpooling layer)	(1, 4000, 8)
dropout_1(dropout layer)	(1, 4000, 8)
conv_3(convolutional layer)	(1, 4000, 16)
conv_4(convolutional layer)	(1, 4000, 16)
pool_2(maxpooling layer)	(1, 2000, 16)
dropout_2(dropout layer)	(1, 2000, 16)
conv_5(convolutional layer)	(1, 2000, 32)
conv_6(convolutional layer)	(1, 2000, 32)
pool_3(maxpooling layer)	(1, 1000, 32)
dropout_3(dropout layer)	(1, 1000, 32)
conv_7(convolutional layer)	(1, 1000, 64)
conv_8(convolutional layer)	(1, 1000, 64)
pool_4(maxpooling layer)	(1, 500, 64)
dropout_4(dropout layer)	(1, 500, 64)
conv_9(convolutional layer)	(1, 500, 64)
conv_10(convolutional layer)	(1, 500, 64)
FC_1(fully connected layer)	(1, 1024)

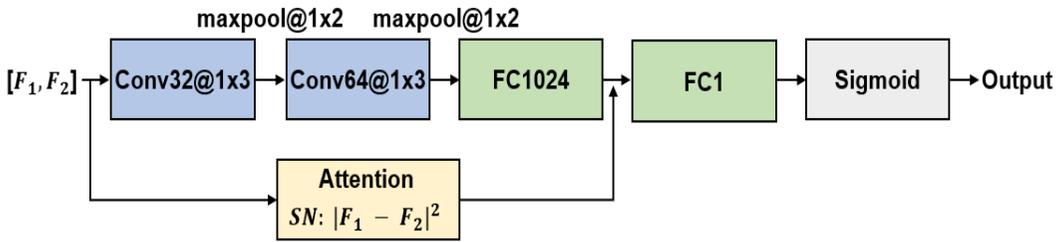


그림 25. 설계한 decision 네트워크 구조

표 6. Decision 네트워크의 계층별 출력 결과

Layer(type)	Output shape
Input(input layer)	(1, 2048, 1)
conv_1(convolutional layer)	(1, 2048, 32)
pool_1(maxpooling layer)	(1, 1024, 32)
dropout_1(dropout layer)	(1, 1024, 32)
conv_2(convolutional layer)	(1, 1024, 32)
pool_2(maxpooling layer)	(1, 512, 32)
FC_1(fully connected layer)	(1, 1024)
FC_2(fully connected layer)	(1, 1)

근전도의 불규칙성 문제를 해결하기 위해 설계한 decision 네트워크 구조는 그림 25와 같다. 설계한 네트워크는 2개의 합성곱 계층 및 2개의 풀링 계층 등으로 구성된 신경망과 데이터 불규칙성 문제를 해결하기 위한 보조 분류기로 구성되어 있다. 신경망은 스택킹 앙상블 구조로 하위 네트워크와 결합되며, 입력된 호환성 함수에서 합성곱 계층을 이용해 특징을 추출한다. 보조 분류기는 어텐션 메커니즘을 위해 두 개의 호환성 함수를 유클리디안 디스턴스로 연산한다. 유클리디안 디스턴스를 이용한 어텐션 메커니즘은 1차원 벡터를 연산함에 따라 기존 방법에 비해 파라미터가 적은 장점이 있다. 그 후 신경망의 출력과 보조 분류기의 출력을

결합한 후 시그모이드(Sigmoid)로 데이터 쌍의 유사도를 계산한다.

기존 보조 분류기는 출력단 결과의 데이터를 융합하기 위해 사용되어 왔다. 다수의 분류기를 이용해 추출된 특징을 분류하고, 분류된 결과를 스코어 레벨 융합, 등급 레벨 융합(Rank level fusion), 결정 레벨 융합(Decision level fusion)하였다. 스코어 레벨 융합은 그림 26과 같이 분류기를 통해 출력된 결과의 점수를 이용해 계산하는 방법이며, 등급 레벨 융합은 순위, 결정 레벨 융합은 빈도를 이용해 계산한다. 데이터 융합은 단일 단계의 분류기를 이용한 방법에 비해 성능 향상이 가능한 방법으로, 기존 보조 분류기를 이용한 연구는 성능 향상을 위해 사용하였다[46].

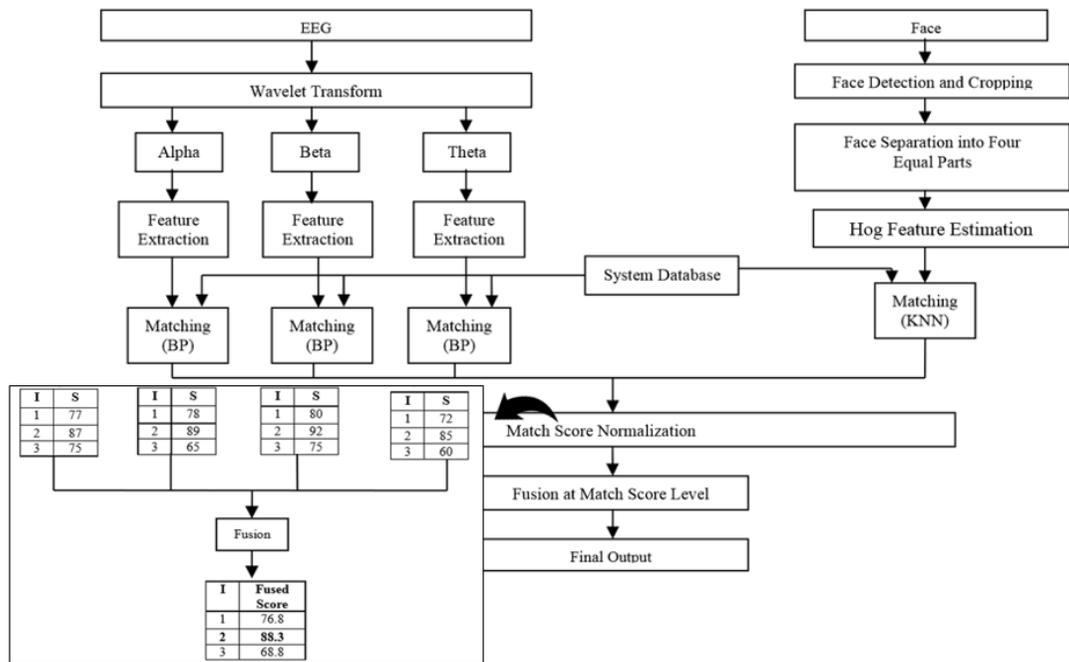


그림 26. 보조 분류기를 이용한 데이터 융합

신경망의 학습은 출력단의 손실률을 계산하여 파라미터를 업데이트하는 역전파 과정을 통해 진행된다. 기존 학습 과정에서 데이터를 비선형으로 변경하기 위한

활성화 함수로 시그모이드와 ReLU를 사용하였으나, 파라미터의 변화가 없어지는 기울기 소실 문제 및 죽어가는 ReLU(Dying ReLU) 문제가 발생한다. 이러한 문제를 보완하고자 GoogLeNet에서는 기존 활성화 함수와 보조 분류기를 결합한 기울기 소실 문제 해결 방법을 처음으로 제안하였다. 이는 네트워크의 손실률을 계산할 때 출력단의 분류기를 이용함과 동시에 중간 레이어에 결합한 보조 분류기의 출력을 함께 사용하는 방법이다. GoogLeNet은 그림 27과 같이 1개의 메인 분류기와 2개의 보조 분류기로 구성된다. 각 보조 분류기의 손실률은 0.3배의 가중치와 메인 분류기는 0.4배의 가중치를 이용해 최종 손실률을 계산하여 기울기 소실 문제를 해결하기 위해 보조 분류기를 사용하였다[47][48].

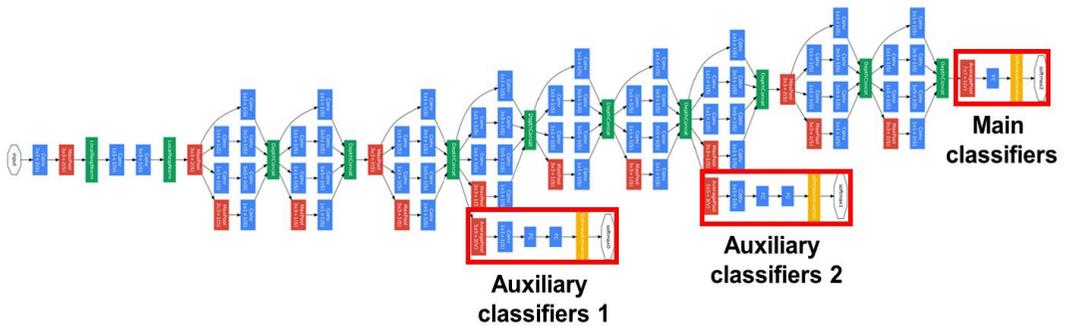


그림 27. GoogLeNet의 메인 분류기와 보조 분류기

어텐션 메커니즘은 가중치를 통해 학습 데이터 내부의 중요성을 모델링하는 방법으로, 출력 결과를 예측할 때 히든 계층을 참고하여 데이터의 중요한 구간에 가중치를 부여하는 방법이다[49]. 이에 다양한 시간 간격으로 구성되는 근전도의 비선형성 특성으로 인해 성능이 낮아지는 데이터 불규칙성 문제를 동작 상태 구간의 신호에 높은 가중치를 적용하는 방법으로 해결할 수 있다. 그러나 기존 어텐션 메커니즘은 신경망을 추가로 설계해야 하며 모든 인덱스에 대한 관계를 정의해야 함에 따라 많은 파라미터 연산이 필요한 단점이 있다[14].

이러한 단점을 보완하기 위해 본 논문에서는 보조 분류기를 이용해 입력된 데이터에 어텐션 메커니즘을 적용하는 방법을 제안한다. 제안한 보조 분류기를 이용한 어텐션 메커니즘은 유클리디안 디스턴스로 구성된 보조 분류기를 이용해 1차원 벡터 연산을 수행하는 방법으로 기존 방법에 비해 연산량이 적은 장점이 있다. 제안한 방법은 두 개의 하위 네트워크에서 연산된 호환성 함수를 이용해 decision 네트워크의 보조 분류기로 입력된 데이터를 비교하며, 이를 통해 데이터 쌍에서 사용자를 인식하기 위해 중요한 구간을 탐색한다. 그 후 식 16과 같이 decision 네트워크의 신경망으로 계산된 특징과 결합하여 입력된 데이터에 어텐션 메커니즘을 적용한다. 식에서 dF 는 decision 네트워크의 신경망으로 계산된 특징이며, SN 은 어텐션 메커니즘을 위해 보조 분류기로 연산된 점수를 의미한다.

$$Attention = dF * SN \quad (16)$$

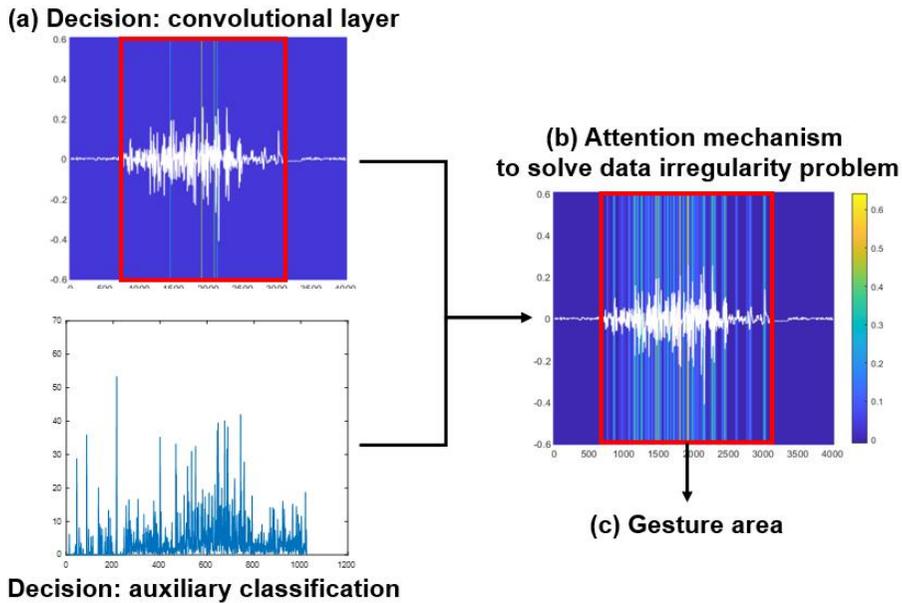


그림 28. 제안한 보조 분류기를 이용한 어텐션 메커니즘 방법

그림 28은 데이터 불규칙성 문제를 해결하기 위해 본 논문에서 제안한 보조 분류기를 이용한 어텐션 메커니즘의 Grad-CAM 결과를 표현한다. Grad-CAM은 파란색부터 노란색으로 구성된 컬러바를 이용해 입력된 데이터의 가중치를 표현하며, 상단 수치인 노란색에 가까울수록 높은 가중치를 나타낸다. Decision 네트워크의 신경망에서 출력된 데이터(그림 28에서 (a))는 동작 구간의 신호에 가중치가 집중되지 않음을 확인할 수 있다. 그러나 어텐션 메커니즘을 위해 보조 분류기를 이용해 연산된 점수를 결합한 데이터(그림 28에서 (b))는 동작 구간의 신호(그림 28에서 (c))에 가중치가 높게 부여됨을 확인할 수 있다. 이에 상대적으로 중요도가 높은 근전도의 동작 구간의 신호에 적응적으로 높은 가중치를 부여하여 데이터 불규칙성 문제를 해결할 수 있다.

제안한 모델은 과적합을 피하기 위해 하위 네트워크의 각 스트림은 4개의 드롭아웃 계층을 사용하며, decision 네트워크는 1개의 드롭아웃 계층을 사용한다. 네트워크의 활성화 함수로 ReLU를 사용한다.

제4장 멀티 스트림 삼 용합 네트워크를 이용한 사용자 인식 실험 결과

제1절 근전도 공개 데이터베이스

근전도 공개 데이터베이스는 표 7과 같이 손가락, 손, 손목 및 발 동작의 신호를 측정하였다. CU_sEMG DB를 제외한 공개 데이터베이스는 40명 이하의 피험자가 단일 세션에서 동작을 수행할 때 구축되었다.

표 7. 근전도 공개 데이터베이스 개요

구분	데이터명	채널(개)	인원(명)	동작	세션
1	EMG(Dr. R. Khushaba): Set 3	2	8	손가락 동작 14개 (각 동작 6회 반복)	단일
2	sEMG for basic hand movement data set	2	5	손 동작 6개 (각 동작 30회 반복)	단일
3	Ninapro DB2	12	40	손과 손목 동작 49개 (각 동작 6회 반복)	단일
4	Ninapro DB5	16	10	손과 손목 동작 52개 (각 동작 6회 반복)	단일
5	EMG dataset in lower limb data set	4	22	걷기 동작 3개	단일
6	CU_sEMG DB	2	200	손과 손목 동작 12개 (각 동작 30회 반복)	멀티

EMG(Dr. R. Khushaba) 데이터[50]는 6개 세트로 나뉜다. 세트 1[51]은 8명의 피험자가 10개 손동작을 수행할 때 2채널로 측정된 데이터이다. 세트 2[52]는 8명

의 피험자가 15개 손동작을 수행할 때 8채널로 측정된 데이터, 세트 3[53]은 그림 29와 같이 8명의 피험자가 14개 손동작을 수행할 때 8채널로 측정된 데이터이다. 세트 4[54]는 손이 절단된 사람의 근전도를 측정된 데이터, 세트 5[55]는 팔 방향에 대한 영향을 기록한 데이터, 세트 6[56]은 사지 위치에 따른 영향을 기록한 데이터이다. 세트 1-3 및 6은 신호 내 잡음을 없애기 위한 전처리 과정으로 대역 통과 필터와 대역 저지 필터를 적용했으며, extensor carpi ulnaris, extensor digitorum, extensor carpi radialis longus 등에서 신호를 취득하였다.

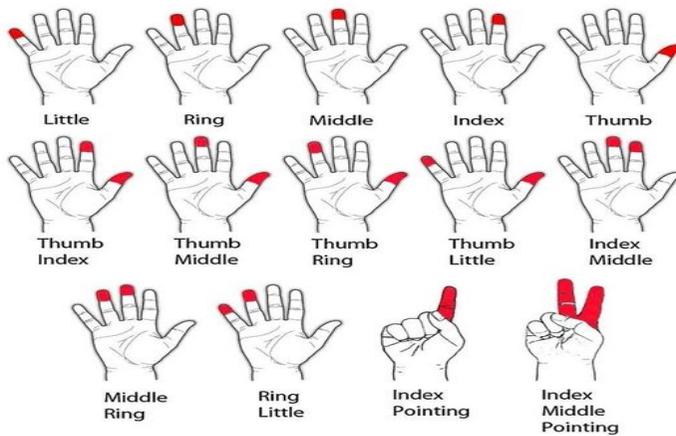


그림 29. EMG(Dr. R. Khushaba) 데이터 세트 3의 손동작

UCI의 sEMG for Basic Hand movements Data Set[43][57]은 그림 30과 같이 6가지 손동작을 수행할 때 20~22세 사이의 피험자 5명에게 취득한 데이터이다. Extensor carpi radialis 및 extensor carpi ulnaris에서 2채널로 근전도를 측정했으며, 대역 통과 필터 및 대역 저지 필터를 이용해 잡음을 제거했다. 데이터는 피험자가 단일 세션에서 동작을 반복 수행하면서 측정된 신호로 구성된다.

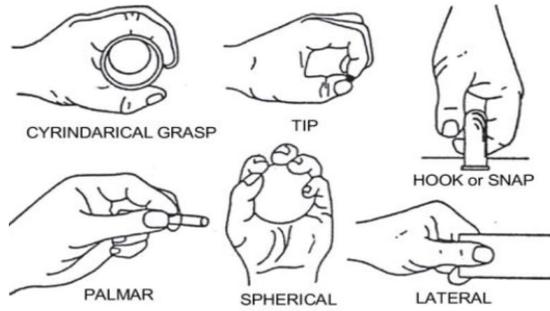


그림 30. UCI의 sEMG for Basic Hand movements Data Set 손동작

Ninapro DB2[58][59]는 그림 31과 같이 40명의 피험자가 수행한 49개의 손동작에 대한 근전도로 구성된다. 데이터는 3개 세트로 구성되며, 세트 1은 기본적인 손가락과 손목을 펼치거나 회전하면서 측정된 근전도로 구성된다. 세트 2는 물체를 집거나 잡는 동작을 수행하면서 취득한 신호이다. 세트 3은 단일 또는 복수 개의 손가락을 구부리면서 취득한 데이터이다. 모든 세트에서 손동작은 6회 반복 동작했으며, 각 동작은 5초 동안 지속했다. 동작과 동작 사이에는 근피로를 피하기 위해 3초의 휴식을 진행했다.



그림 31. Ninapro DB2 손동작

Ninapro DB5[60][61]는 그림 32와 같이 10명의 피험자가 수행한 52개의 손 동작에 대한 근전도로 구성된다. 데이터는 3개 세트로 구성되며, 세트 1은 손가락의 기본적인 동작을 수행하며 근전도를 측정했다. 세트 2 및 세트 3은 Ninapro DB2의 세트 1 및 세트 2와 동일한 동작으로 구성된다. 데이터베이스의 모든 손 동작은 5초 동안 지속했으며, 동작 사이에 3초의 휴식을 진행하여 근피로를 피하였다.



그림 32. Ninapro DB5 손동작

손 이외의 근육 활동을 측정한 데이터로 UCI의 EMG dataset in Lower Limb Data Set[62]가 있다. 무릎 근방 근육의 행동을 관찰하기 위한 데이터로, 22명의 피험자(정상인 11명, 환자 11명)가 3가지 동작(앉기, 서기, 걷기)을 수행하며 근전도를 측정했다. 센서는 rectus femoris, vastus medialis, semitendinous, biceps femoris에 부착했다.

기존 대부분의 근전도 공개 데이터베이스는 피험자의 손 동작을 인식하기 위해 구축되었으며, 데이터베이스를 구축할 때 수행한 동작은 손가락과 손목을 이용해 일상생활에서 많이 사용되는 동작을 선택하였다. 그러나 40명 이하의 피험자가 단

일 세션에서 신호를 측정하여 신호의 변동성을 고려할 수 없는 문제가 있다. 이에 본 논문에서 실험은 실생활 환경에서 발생 가능한 신호의 변동성을 고려하기 위해 멀티 세션으로 피험자들에게 근전도를 측정할 데이터베이스를 이용하였다. 사용한 데이터베이스는 CU_sEMG DB[63][64]로, 조선대학교 IT연구소에서 구축한 근전도 공개 데이터베이스이다. 데이터는 취득 전 각 피험자에게 실험 목적과 절차, 수행 동작에 대해 설명하고, 실험 전·후 카페인 및 음주 여부, 건강 상태, 수면 상태, 데이터 취득 시 불편 사항 등을 조사하였다. 근전도는 Biopac MP160 장비를 이용해 평균 24.69세(19-70세) 피험자 200명(남성 98명, 여성 102명)의 오른팔에서 취득하였다. 데이터는 2채널로, 그림 33과 같이 피험자의 palmaris longus 및 extensor digitorum에 Ag/AgCl 센서를 부착하였다. 센서 위치는 손동작 시 근육 변화량이 많은 지점을 선택하였다.

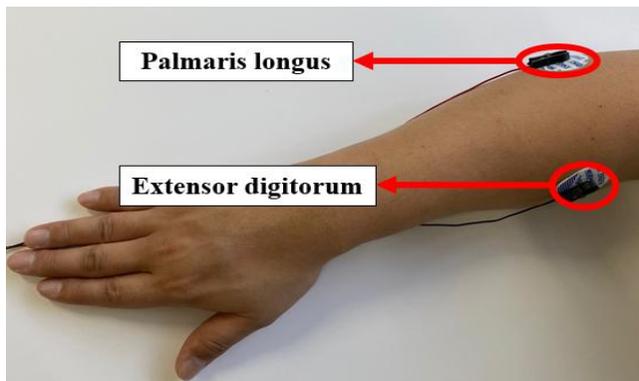


그림 33. CU_sEMG DB의 측정 근육 위치

피험자는 의자에 앉은 편안한 상태에서 각 동작을 수행했으며, 각 세션에서 10회 반복하며 연속적으로 취득하였다. 동작 반복 시 최소 1초 이상 동작을 유지하도록 하여, ‘시작(휴식 상태)-손동작 수행-종료(휴식 상태)’ 순서로 진행하였다. 신호는 2,000Hz 샘플링 레이트, 16bit ADC 해상도로 취득했으며, 피험자는 조선대학교 학생 및 연구원, 일반인을 대상으로 선정하였다.

근전도 측정 시 피험자는 그림 34와 같이 12개 손동작을 수행하였다. 동작은 정적 손동작(Static gesture)과 동적 손동작(Dynamic gesture)으로 구분되며, 일상생활에서 많이 사용하는 동작으로 선정하였다. 정적 손동작은 단일 움직임을 수행하는 동작으로 구성되며, 동적 손동작은 연속된 움직임을 수행하는 동작으로 구성된다. CS_sEMG DB는 근전도의 변동성을 고려하기 위해 총 3개 세션으로 취득했으며, 세션 사이에는 하루 이상의 시간 간격을 두었다. 각 피험자는 1개 동작에 대해 3개 세션으로 취득하여, 총 30개(10회 반복*3개 세션)의 근전도를 취득하였다. 따라서 각 피험자는 12개 동작을 수행하여 총 360개(12개 동작*10회 반복*3개 세션)의 근전도 데이터를 취득하였다.

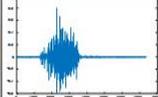
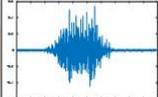
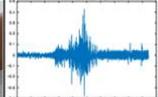
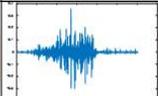
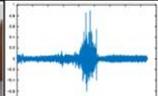
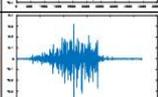
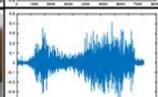
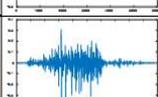
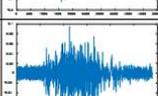
Exercise A (Static gestures)				Exercise B (Dynamic gestures)			
1	All fingers extended (paper)			1	Fist→paper → scissors		
2	All fingers folded (fist)			2	Wrist left/right rotation (palm horizontal)		
3	Extend index and middle fingers (scissors)			3	Wrist up/down rotation (palm horizontal)		
4	Extend all fingers except thumb			4	Wrist left/right rotation (palm vertical)		
5	Intersection of index and middle fingers			5	Wrist rotation (clockwise)		
6	Finger of middle and thumb folds						
7	Finger of little and thumb folds						

그림 34. CU_sEMG DB의 12개 손동작 및 근전도 파형

취득자 및 피험자의 잘못된 근전도 데이터베이스 구축 방식으로 인해 발생하는 영향을 피하기 위해 CU_sEMG DB는 한 개 이상의 근전도가 잘못된 경우 해당 피험자를 제외하였다. 피험자 200명 중 58%인 116명의 근전도가 올바르게 취득되었으며, 84명은 1) 블루투스 연결 끊김(Device loses Bluetooth connection), 2) 근육 수축 및 이완 불완전(Incompletion of muscle contraction and relaxation), 3) 반복 횟수 부족(Insufficient number of repetitions), 4) 근육이 활성화되지 않음(Muscle not activated)의 이유로 제외하였다. 제외된 피험자는 실험에 사용하지 않는다.

제2절 사용자 인식 평가 방법

실험 성능은 정확도(Accuracy), 재현율(Recall), 정밀도(Precision)를 이용해 평가하며, 위와 같은 평가 지표는 TP(True positive), TN(True negative), FP(False positive), FN(False negative)을 이용해 계산할 수 있다. 정확도는 true를 학습한 모델이 true로 예측한 경우와 false를 학습한 모델이 false를 예측한 경우에 대한 성능으로, 식 17과 같이 계산할 수 있다. 재현율은 true를 학습한 모델이 true로 분류한 비율을 뜻하며, 식 18과 같이 TP와 FN으로 계산할 수 있다. 정밀도는 학습된 모델이 true로 분류한 것 중에서 실제 true의 비율을 의미하며, 식 19와 같이 TP와 FP로 계산할 수 있다.

- TP: 실제 true인 데이터를 학습한 모델이 true로 올바르게 식별한 경우
- FP: 실제 false인 데이터를 학습한 모델이 true로 잘못 식별한 경우
- TN: 실제 false인 데이터를 학습한 모델이 false로 올바르게 식별한 경우
- FN: 실제 true인 데이터를 학습한 모델이 false로 잘못 식별한 경우

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

$$Recall = \frac{TP}{TP + FN} \quad (18)$$

$$Precision = \frac{TP}{TP + FP} \quad (19)$$

제3절 실험 결과 및 분석

본 절에서는 표 8과 같이 네트워크의 스트림 개수에 따른 성능 분석, 시간-주파수 도메인 특징 추출에 따른 인식 성능 분석, 근전도의 불규칙성 문제를 해결하기 위한 보조 분류기의 인식 성능 분석, 모델 재학습 문제를 해결하는 삼 모델의 인식 성능을 분석 실험을 진행한다. 또한, 기존 사용자 인식 방법과 인식 성능을 비교 분석하여 제안하는 사용자 인식 방법의 우수성을 설명한다. 실험은 올바르게 측정된 116명의 근전도 데이터가 있는 CU_sEMG DB에서 100명의 피험자에게 측정된 근전도 데이터를 이용해 신경망을 학습하였다. 남은 피험자 16명의 근전도 데이터는 신경망 학습에 사용하지 않았다. 이는 학습된 모델에 새로운 사용자가 추가될 때 모델을 재학습하지 않고 인식이 가능한지 확인하기 위한 모델 재학습 문제를 해결하는 실험(단계 4) 데이터로 사용했다. 설계한 모델은 RMSProp 및 Momentum 기법을 합친 Adam 최적화 방법을 이용해 학습률 0.001, 에포크 100, 배치 크기 128로 설정해 학습을 진행하였다.

표 8. 사용자 인식 실험 단계별 구성

단계	실험 종류	사용한 데이터	실험 결과	
			전(%)	후(%)
1	스트림 개수에 따른 멀티 스트림 모델의 성능 확인	CU_sEMG DB 100명(멀티 세션)	78.71 (단일 스트림)	87.62 (멀티 스트림)
2	내재 모드 함수 특징을 이용한 사용자 인식	CU_sEMG DB 100명(멀티 세션)	86.85 (원시 근전도)	92.01 (내재모드함수 1-4)
3	데이터 불규칙성을 해결하기 위한 보조 분류기를 이용한 실험	CU_sEMG DB 100명(멀티 세션)	92.01 (미사용)	94.35 (사용)
4	모델 재학습 문제를 해결하기 위한 새로운 피험자 데이터를 이용한 실험	새로운 CU_sEMG DB 16명(멀티 세션)	-	93.19 (새로운 피험자)
5	기존 연구와 성능 비교	CU_sEMG DB 100명(단일 세션) Ninapro DB2 40명	핸드크래프트 특징, 합성곱 신경망, 삼 네트워크를 이용한 방법과 비교	

첫 번째로 멀티 스트림 네트워크를 설계하기 위해 스트림 개수에 따른 성능을 분석하였다. 실험은 스트림 개수를 1개부터 6개로 변경하며 진행했다. 모델의 입력 데이터는 원시 근전도 데이터를 중복해 사용하였다. 근전도는 모델 학습에 세션 1에 측정된 데이터를 사용하고 테스트에 세션 2와 세션 3에서 측정된 데이터를 사용했다. 스트림 변화에 따른 사용자 인식 실험 결과는 그림 35와 같다. 실험 결과, 스트림이 증가할 때 사용자 인식 성능이 향상되었으며, 스트림이 4개일 때 87.62%로 가장 높은 정확도를 나타냈다. 이는 단일 스트림에 비해 8.91% 성능이 향상되었으며, 스트림 개수가 5개 이상일 때 학습 데이터에 과적합으로 인해 인식 성능이 감소됨을 확인하였다. 따라서 향후 멀티 스트림을 이용한 실험은 4개의 스트림을 이용해 진행하였다.

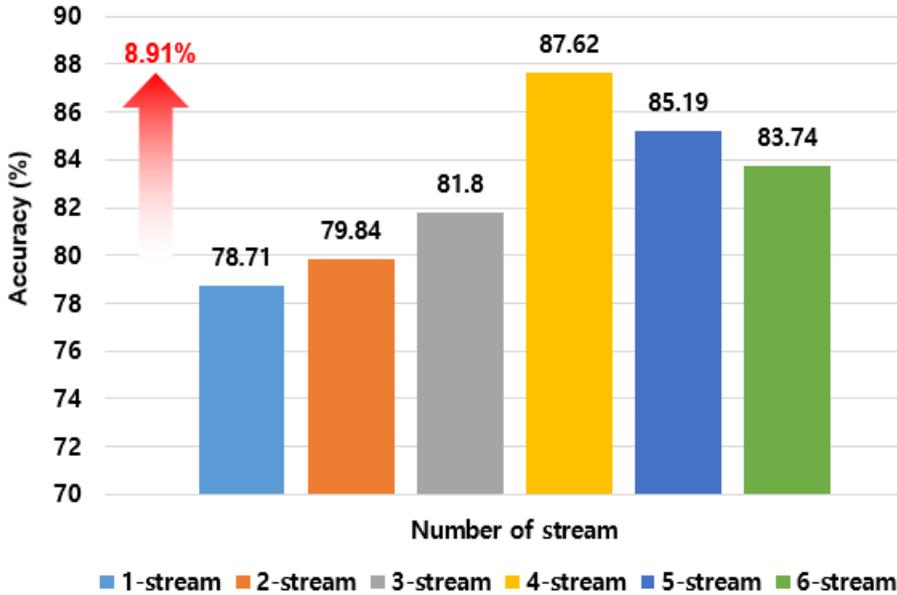


그림 35. 스트림 변화에 따른 사용자 인식 성능

신호의 시간-주파수 도메인 특성을 사용하기 위해 본 논문에서 사용한 내재 모드 함수 특징을 이용한 사용자 인식 성능을 분석했다. 내재 모드 함수 특징을 이용한 사용자 인식 실험은 그림 24와 같은 하위 네트워크의 입력 데이터로 내재 모드 함수의 데이터 쌍을 사용해 특징을 추출하였다. 두 개의 하위 네트워크의 유사도 비교는 코사인(Cosine), 맨하탄(Manhattan) 유사도 방법에 비해 우수한 성능을 나타낸 유클리디안 디스턴스를 이용해 진행했다[18]. 실험 결과는 그림 36과 같다. 삼 네트워크를 멀티 스트림 구조로 설계하고, 내재 모드 함수 1-4를 이용하면 100명의 피험자에게 평균 92.01%의 인식 정확도를 나타냈으며, 내재 모드 함수 1-3과 잔차를 이용한 방법에 비해 1.83%, 내재 모드 함수 1-3과 핸드크래프트 특징[1]을 이용한 방법에 비해 2.21% 성능이 향상되었다. 또한, 스펙트로그램을 이용한 방법에 비해 3.21%, 원시 근전도 데이터를 중복해 사용한 방법에 비해 5.16% 정확도가 높음을 확인하였다. 이는 경험적 모드 분해를 이용해 근전도를 처리하

고, 근전도에서 중요한 성분이 포함된 내재 모드 함수 1-4를 이용하여 높은 사용자 인식 성능을 나타냈다[16][43].

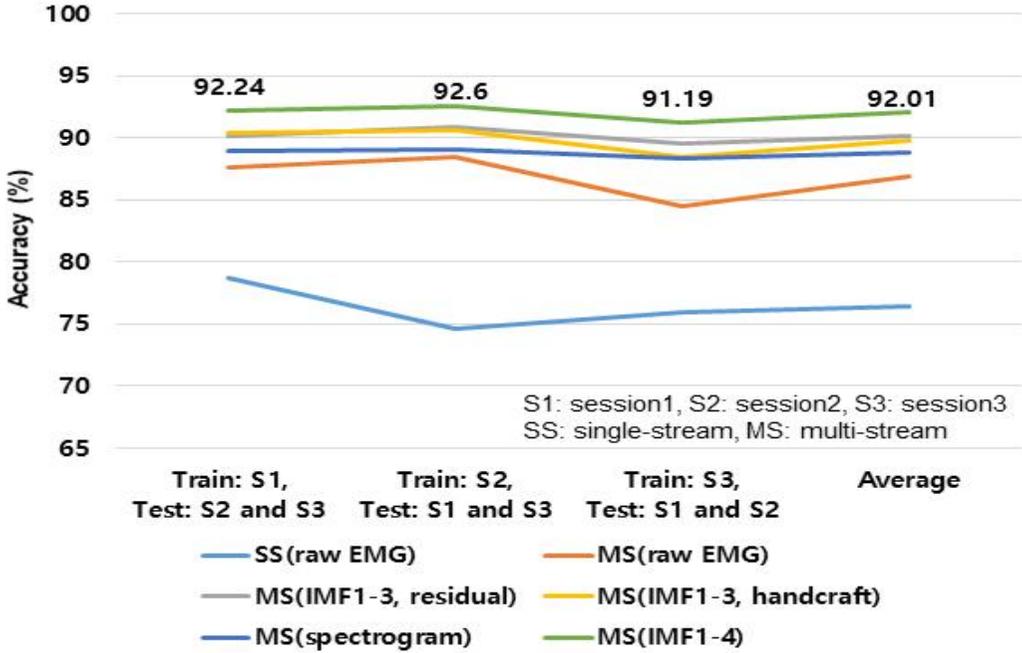


그림 36. 멀티 스트림 삼 네트워크를 이용한 사용자 인식 성능

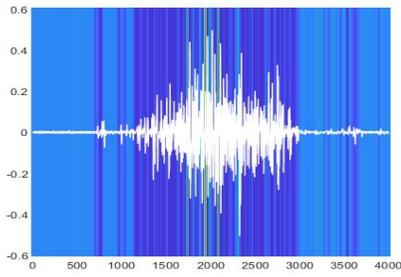
데이터 불규칙성 문제를 해결하기 위해 본 논문에서 제안한 보조 분류기가 결합된 decision 네트워크를 이용한 모델의 성능을 확인하기 위해 사용자 인식 실험을 진행하였다. 실험은 그림 23과 같은 네트워크와 내재 모드 함수 1-4를 이용하였다. Decision 네트워크가 없는 모델은 유클리디안 디스턴스를 이용해 유사도를 비교했으며, 실험 결과는 그림 37과 같다. 제안한 방법을 이용한 사용자 인식은 100명의 피험자를 94.35%의 정확도로 분류하였으며, 보조 분류기를 사용하지 않은 멀티 스트림 삼 네트워크 방법에 비해 2.34% 성능이 향상되었다. 이는 decision 네트워크를 스택킹 앙상블 구조로 결합하고, 보조 분류기를 이용해 근전도의 동작 상태 구간 신호에 높은 가중치를 적용하여 성능이 향상되었다.



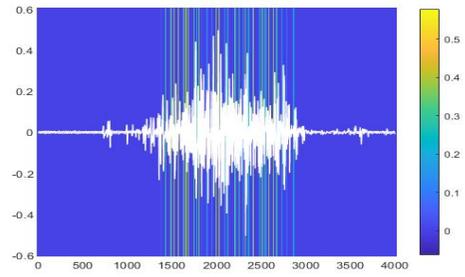
그림 37. 보조 분류기를 결합한 멀티 스트림 삼 용합 네트워크 기반 실험

그림 38은 보조 분류기를 이용한 어텐션 메커니즘의 Grad-CAM 결과를 나타낸다. 그림에서 가중치를 표현하는 컬러바(color bar)는 파란색부터 노란색으로 구성되며, 상단 수치인 노란색에 가까울수록 사용자 인식을 위한 모델에서 높은 가중치로 사용된다. 그림 38 (a)와 같은 신호는 동작을 수행하지 않은 구간의 가중치가 높게 계산되어 낮은 사용자 인식 성능을 나타냈다. 이는 보조 분류기를 이용한 어텐션 메커니즘으로 그림 38 (b)와 같이 동작을 수행한 구간의 근전도에 높은 가중치를 부여하였다. 다른 피험자들 또한 그림 38 (c)-(h)와 같이 제안한 보조 분류기를 이용한 방법으로 동작 구간의 근전도에 높은 가중치를 부여하여 인식 성능이 향상되었다.

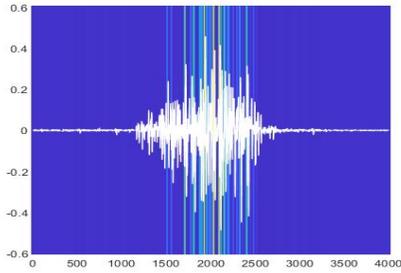
그림 39는 CU_sEMG DB의 멀티 세션 데이터로 본 논문에서 제안한 사용자 인식 방법을 실험한 ROC(Receiver operating characteristic) 커브 결과를 나타낸다. 실험 결과, AUC(Area under curve)는 98.08%의 정확도를 나타냈으며, 보조 분류기를 사용하지 않은 방법에 비해 1.44% 높은 성능이 나타남을 확인하였다.



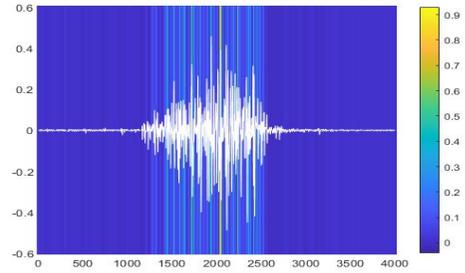
(a) Subject 1(before attention)



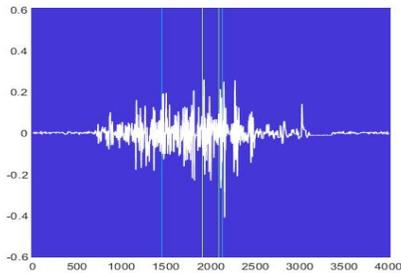
(b) Subject 1(after attention)



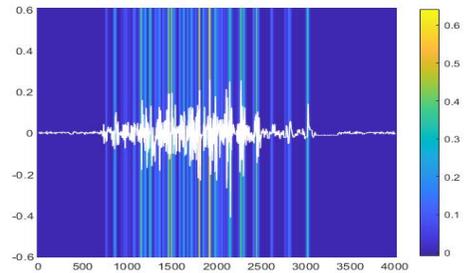
(c) Subject 2(before attention)



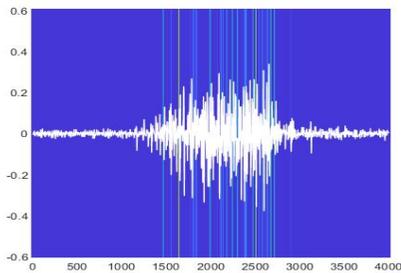
(d) Subject 2(after attention)



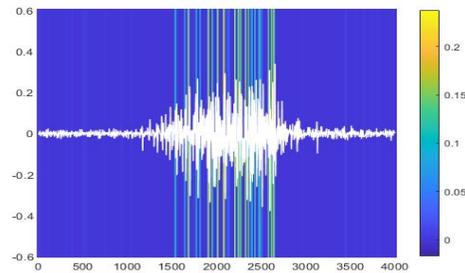
(e) Subject 3(before attention)



(f) Subject 3(after attention)



(g) Subject 4(before attention)



(h) Subject 4(after attention)

그림 38. 보조 분류기를 이용한 어텐션 메커니즘 결과

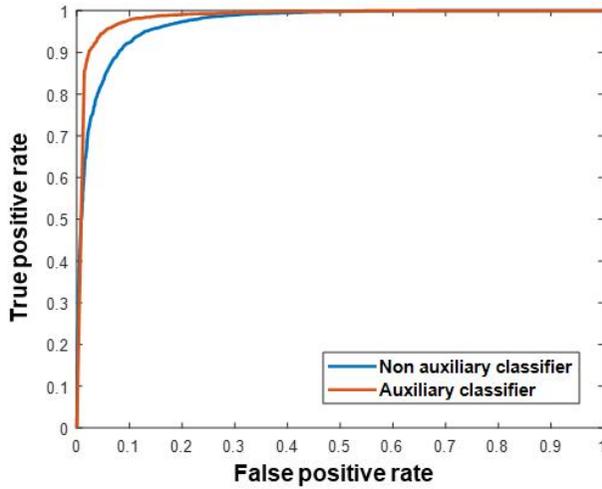


그림 39. 제안한 사용자 인식 방법의 ROC 커브 결과



그림 40. 보조 분류기를 결합한 단일 스트림 삼 용합 네트워크 기반 실험

또한, 단일 스트림에서 보조 분류기가 결합된 decision 네트워크의 성능을 확인하였다. 실험은 단일 스트림을 기반으로 보조 분류기 유·무에 따른 성능을 비

교 실험했으며, 실험 결과는 그림 40과 같다. 논문에서 제안한 보조 분류기를 이용한 어텐션 메커니즘을 이용한 경우 92.64%의 정확도로, 보조 분류기가 없는 삼 네트워크에 비해 16.2%의 정확도가 향상됨을 확인하였다.

표 9. 학습된 모델 기반 새로운 사용자를 이용한 실험 결과

Method	Accuracy(%)	Recall(%)	Precision(%)
Experiment 1 (registration: session 1) (test data: session 2 & 3)	93.33	88.75	97.71
Experiment 2 (registration: session 2) (test data: session 1 & 3)	91.60	91.53	91.66
Experiment 3 (registration: session 3) (test data: session 1 & 2)	94.65	95.56	93.86
Average	93.19	91.94	94.41

기존 딥러닝의 모델 재학습 문제를 해결하기 위해 본 논문의 삼 네트워크 성능을 확인하기 위한 실험으로, 모델을 학습할 때 사용하지 않은 피험자의 근전도 데이터로 모델을 재학습하지 않고 사용자 인식 실험을 진행하였다. 보조 분류기가 결합된 멀티 스트림 삼 융합 네트워크를 이용한 실험은 3개 세션에서 측정된 피험자 16명의 근전도를 모두 사용하였다. 데이터는 다른 세션의 데이터와 쌍을 이루어 실험을 진행했으며, 실험 결과는 표 9와 같다. 실험 결과, 모델을 학습할 때 사용하지 않은 16명의 피험자를 93.19%의 정확도로 인식하였으며, 91.94%의 재현율과 94.41%의 정밀도 성능을 나타냈다. 전이 학습 방법과 성능을 비교한 결과는 그림 41과 같다. 실험 결과, 전이 학습을 진행한 경우와 1.32% 성능이 차이남에

따라, 기존 학습된 모델에 새로운 사용자가 추가되면 재학습이 필요한 문제가 해결됨을 확인하였다.



그림 41. 전이 학습 방법과 성능 비교 결과

제안한 방법의 우수성을 확인하기 위해 기존 근전도를 이용한 사용자 인식 방법과 비교 실험을 진행했다. 비교 실험은 핸드크래프트 특징을 이용한 방법[1]과 합성곱 신경망을 이용한 방법[22], 삼 네트워크를 이용한 방법[11][18]으로 구성했으며, CU_sEMG DB의 단일 세션 데이터로 진행하였다. 실험 결과는 그림 42와 같다. 기존 사용자 인식을 위해 핸드크래프트 특징 및 합성곱 신경망을 이용한 방법은 피험자 수가 증가할 때 정확도가 큰 폭으로 감소했으며, 핸드크래프트 특징을 이용한 실험은 피험자 100명에 대해 73.4%의 정확도로 가장 낮은 성능을 나타냈다. 합성곱 신경망을 이용한 기존 방법은 피험자 100명에 대해 85.55%의 정확도로 낮은 사용자 인식 성능을 나타냈다. 삼 네트워크를 이용한 기존 방법은 100명의 피험자에 대해 90.36% 및 91.94%의 정확도를 나타내어, 근전도와 같이 모델을 학

습하기 위한 데이터가 부족할 때 삼 구조가 적합함을 다시 한번 확인하였다. 본 논문에서 제안한 보조 분류기가 결합된 멀티 스트림 삼 용합 네트워크를 이용한 사용자 인식 방법은 100명의 피험자를 멀티 세션에서 94.35%, 단일 세션에서 95.33%의 정확도를 나타냈다. 이는 보조 분류기를 이용해 어텐션 메커니즘을 적용하여 성능이 향상되었으며, 비교 실험을 통해 제안한 방법의 우수성을 확인하였다.

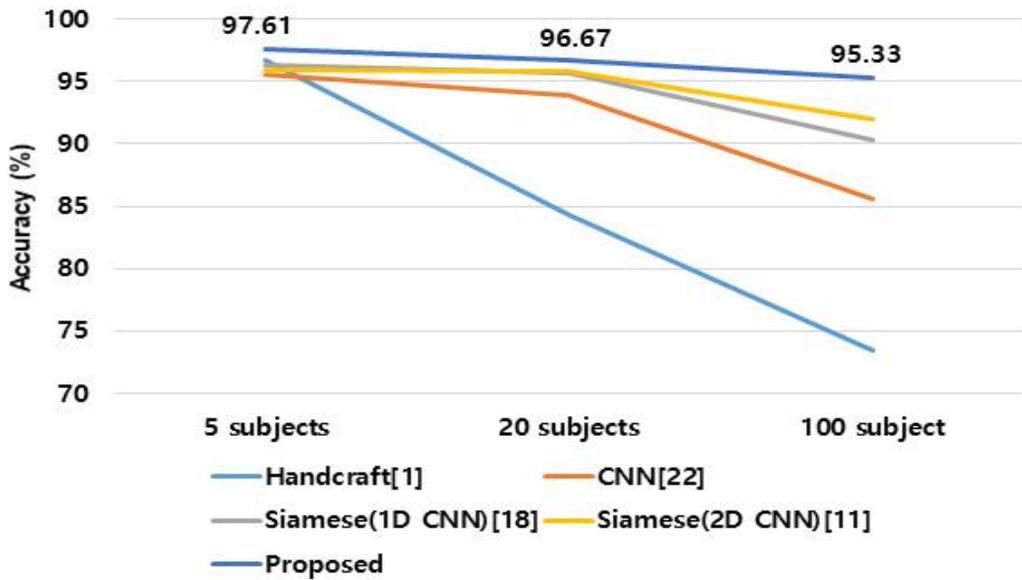


그림 42. CU_sEMG DB를 이용한 기존 사용자 인식 방법의 실험 결과

또한, 제안한 방법의 범용성을 확인하기 위해 Ninapro DB2로 비교 실험을 진행하였다. Ninapro DB2는 모델 학습에 각 동작의 4개 데이터를 이용하고, 테스트에 2개 데이터를 이용했다. 실험 결과, 본 논문에서 제안한 보조 분류기가 결합된 멀티 스트림 삼 용합 네트워크를 이용한 사용자 인식 방법은 40명의 피험자에 대해 98.41%의 정확도를 나타냈다. 이는 CU_sEMG DB를 이용한 경우와 동일하게 제안한 방법의 성능이 가장 높게 나타남을 확인하였다.

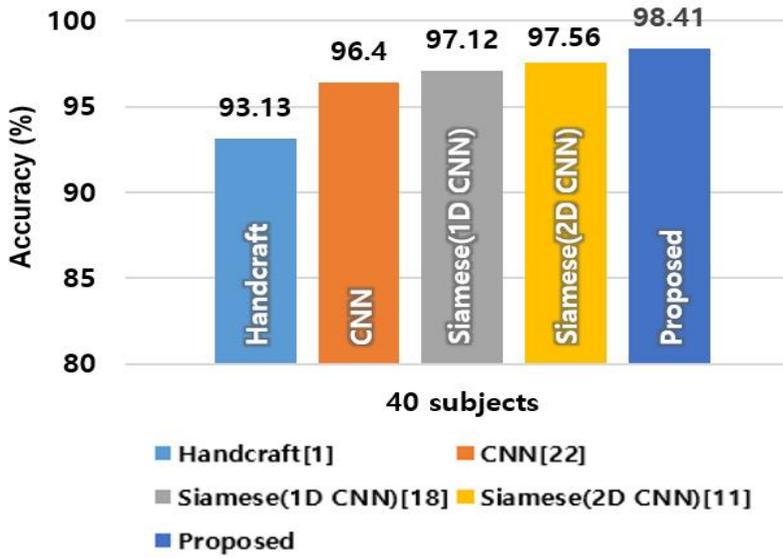


그림 43. Ninapro DB2를 이용한 기존 사용자 인식 방법의 실험 결과

제5장 결론

기존 근전도를 이용한 사용자 인식 연구는 단일 도메인에서 분석된 특징을 이 용함에 따라 범용성이 부족하며, 새로운 사용자가 추가되면 모델 재학습이 필요한 문제가 발생한다. 또한, 근전도의 비선형성 특성으로 인해 데이터 불규칙성 문제 가 발생한다. 본 논문에서는 이러한 문제를 해결하기 위해 보조 분류기를 결합한 멀티 스트림 삼 용합 네트워크 기반 사용자 인식 방법을 제안하였다. 제안된 방법 에서 특징 추출은 근전도의 시간-주파수 도메인 특성을 고려하여 내재 모드 함수 특징을 생성하는 경험적 모드 분해를 이용하였다. 사용자 인식을 위해 제안한 모 델은 재학습 문제를 해결하기 위해 삼 구조로 설계된 2개의 하위 네트워크와 보조 분류기가 결합된 1개의 decision 네트워크로 구성했다. Decision 네트워크의 보조 분류기는 데이터 불규칙성 문제를 해결하기 위해 유클리디안 디스턴스로 어텐션 메커니즘을 적용했다.

실생활 환경에서 발생 가능한 신호의 변동성을 고려하기 위해 멀티 세션에서 측정된 CU_sEMG DB로 실험을 진행했다. 본 논문에서 제안한 보조 분류기를 결합한 멀티 스트림 삼 용합 네트워크는 피험자 100명을 94.35%의 정확도로 인식하여, 보 조 분류기를 사용하지 않은 방법에 비해 2.34% 높은 성능을 보였다. 이는 어텐션 메커니즘을 이용해 근전도 동작 구간의 신호에 높은 가중치를 적용하여 성능이 향 상되었다. 또한, 모델 학습에 사용하지 않은 새로운 피험자를 93.19% 정확도로 인 식함에 따라 모델 재학습 문제가 해결됨을 확인하였다. 마지막으로, 제안한 방법 의 우수성을 확인하기 위해 기존 연구 방법과 성능을 비교했다. 실험 결과, 본 논 문에서 제안한 방법은 가장 높은 정확도를 나타냈으며, 피험자 수가 증가할 때에 도 유사한 성능을 나타내어 제안한 방법의 우수성을 증명했다.

본 논문에서는 보조 분류기를 이용한 어텐션 메커니즘으로 데이터 불규칙성 문 제를 해결하여 사용자 인식 성능을 향상시켰다. 그러나 실생활에 적용되기 위해서

는 보다 높은 인식 성능이 필요하다. 이에 향후 과제로 근전도의 잡음 제거 기술과 새로운 손실함수를 연구할 계획이다. 또한, 연산 속도 향상을 위해 네트워크 최적화에 대해 연구를 진행할 계획이다.

참고문헌

- [1] S. H. Kim, J. H. Ryu, B. H. Lee, and D. H. Kim, "Human identification using EMG signal based artificial neural network," *Journal of the Institute of Electronics and Information Engineers*, vol. 53, no. 4, pp. 142-148, 2016.
- [2] M. G. Kim, "A study on user recognition system based on ensemble convolutional neural networks using synthetic electrocardiogram generation," Ph.D. dissertation, Chosun University, Republic of Korea, 2019.
- [3] X. Jiang, K. Xu, X. Liu, C. Dai, D. A. Clifton, E. A. Clancy, M. Akay, and W. Chen, "Cancelable HD-sEMG-based biometrics for cross-application discrepant personal identification," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 4, pp. 1070-1079, 2020.
- [4] J. S. Kim and S. B. Pan, "A study of EMG-based biometrics," *J. Internet Serv. Inf. Secur.*, vol. 7, no. 2, pp. 19-31, 2017.
- [5] C. Ashby, A. Bhatia, F. Tenore, and J. Vogelstein, "Low-cost electroencephalogram based authentication," In Proc. Int. IEEE/EMBS Conf. Neural Eng., pp. 442-445, 2011.
- [6] J. S. Kim, M. G. Kim, and S. B. Pan, "Two-step biometrics using electromyogram signal based on CNN-LSTM networks," *Appl. Sci.*, vol. 11, no. 15, p. 6824, 2021.
- [7] P. Shenoy, K. J. Miller, B. Crawford, and R. P. N. Rao, "Online electromyographic control of a robotic prosthesis," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 3, pp. 1128-1135, 2008.
- [8] H. Yamaba, S. Usuzaki, K. Takatsuka, K. Aburada, and T. Katayama, "On a user authentication method to realise an authentication system using s-EMG," *Int. J. Grid and Util. Comput.*, vol. 11, no. 5, pp. 725-734, 2020.
- [9] M. U. Khan, Z. A. Choudry, S. Aziz, S. Z. H. Naqvi, A. Aymin, and M.

- A. Imtiaz, “Biometric authentication based on EMG signals of speech,” In Proc. Int. Conf. Electr. Commun. Comput. Eng., 2020.
- [10] J. M. Kim, G. H. Choi, J. S. Kim, and S. B. Pan, “User recognition using electromyogram 2D spectrogram images based on CNN,” *Journal of Korean Institute of Information Technology*, vol. 19, no. 1, pp. 107–117, 2021.
- [11] L. Lu, J. Mao, W. Wang, G. Ding, and Z. Zhnag, “A study of personal recognition method based on EMG signal,” *IEEE Trans. on Biomed. Circuits Syst.*, vol. 14, no. 4, pp. 681–691, 2020.
- [12] D. H. Kisa, M. A. Ozdemir, O. Guren, and A. Akan, “EMG based hand gesture classification using empirical mode decomposition time-series and deep learning,” In Proc. Med. Technol. Cong., 2020.
- [13] C. Savur and F. Sahin, “Real-time american sign language recognition system using surface EMG signal,” In Proc. IEEE Int. Conf. Mach. Learn. Appl., pp. 497–502, 2015.
- [14] J. Ku, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” In Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 7132–7141, 2018.
- [15] Q. Li, P. Dong, and J. Zheng, “Enhancing the security of pattern unlock with surface EMG-based biometrics,” *Appl. Sci.*, vol. 10, no. 2, p. 541, 2020.
- [16] M. R. Lee, J. H. Ryu, and I. C. Youn, “Biometric personal identification based on gait analysis using surface EMG signals,” In Proc. IEEE Int. Conf. Comput. Intell. App., pp. 318–321, 2017.
- [17] V. Shreyas, J. X. Felix, C. Benjamin, and S. Marios, “Electromyograph and keystroke dynamics for spoof-resistant biometric authentication,” In Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 109–118, 2015.
- [18] B. Fan, X. Liu, X. Su, P. Hui, and J. Niu, “EmgAuth: an EMG-based smartphone unlocking system using siamese network,” In Proc. IEEE Int. Conf. Pervasive Comput. Commun., 2020.

- [19] A. M. H. Wong, M. Furukawa, H. Ando, and T. Maeda, “Dynamic hand gesture authentication using electromyography (EMG),” In Proc. Int. Symp. Syst. Integr., pp. 300–304, 2020.
- [20] R. Shioji, S. I. Ito, M. Ito, and M. Fukumi, “Personal authentication and hand motion recognition based on wrist EMG analysis by a convolutional neural network,” In Proc. Int. Conf. Soft Comput. Intell. Syst. Int. Symp. Adv. Intell. Syst., pp. 1172–1176, 2018.
- [21] S. Morikawa, S. I. Ito, M. Ito, and M. Fukumi, “Personal authentication by lips EMG using dry electrode and CNN,” In Proc. Int. Conf. Internet Things Intell. Syst., pp. 180–183, 2018.
- [22] L. Lu, J. Mao, W. Wang, G. Ding, and Z. Zhang, “An EMG-based personal identification method using continuous wavelet transform and convolutional neural networks,” In Proc. Biomed. Circuits Syst. Conf., 2019.
- [23] S. H. Shin, J. H. Jung, M. G. Kang, and Y. T. Kim, “A study on EMG signal acquisition modules and artificial neural networks for personal authentication,” In Proc. Int. Conf. Comput. Sci. Comput. Intell., pp. 1458–1459, 2018.
- [24] S. A. Raurale, J. Mcallister, and J. M. D. Rincon, “EMG biometric systems based on different wrist–hand movements,” *IEEE Access*, vol. 9, pp. 12256–12266, 2021.
- [25] D. Zhong, Y. Yang, and X. Du, “Palmpoint recognition using siamese network,” *Biometric Recognition*, vol. 10996, pp. 48–55, 2018.
- [26] E. Maiorana, “EEG-based biometric verification using siamese CNNs,” *New Trends in Image Analysis and Processing*, vol. 11808, pp. 3–11, 2019.
- [27] H. I. Abdulrazzaq and N. F. Hassan, “Modified siamese convolutional neural network for fusion multimodal biometrics at feature level,” In Proc. Scient. Conf. Comput. Sci., pp. 12–17, 2019.
- [28] A. Acien, A. Morales, R. V. Rodriguez, and J. Rierrez, “TypeNet:

- scaling up keystroke biometrics,” In Proc. IEEE Int. Joint Conf. Biom., 2020.
- [29] D. D. Chakladar, P. Kumar, P. P. Roy, D. P. Dogra, E. Scheme, and V. Chang, “A multimodal-siamese neural network(mSNN) for person verification using signatures and EEG,” *Inf. Fusion*, vol. 71, pp. 17–27, 2021.
- [30] A. Zakeri and H. Hassanpour, “WhisperNet: deep siamese network for emotion and speech tempo invariant visual-only lip-based biometric,” In Proc. Int. Conf. Signal Process. Intell. Syst., 2021.
- [31] K. Sriskandaraja, V. Sethu, and E. Ambikairajah, “Deep siamese architecture based replay detection for secure voice biometric,” In Proc. Interspeech, pp. 671–675, 2018.
- [32] E. J. A. Mireles, F. J. R. Sanchez, and O. G. Salazar, “EMG patterns induced in upper limb by haptic guidance for diagnosis and treatment evaluation,” In Proc. Int. Conf. Electr. Eng. Comput. Sci. Automatic Control, 2015.
- [33] A. A. Adewuyi, L. J. Hargrove, and T. A. Kuiken, “An analysis of intrinsic and extrinsic hand muscle EMG for improved pattern recognition control,” *IEEE Trans. Neural Syst. Rehabilitation Eng.*, vol. 24, no. 4, pp. 485–494, 2015.
- [34] J. S. Kim, J. M. Kim, M. G. Kim, and S. B. Pan, “Electromyogram signal based personal recognition using multi-stream CNN and empirical mode decomposition,” *Journal of Korean Institute of Information Technology*, vol. 20, no. 2, pp. 149–156, 2022.
- [35] X. Song, S. Guo, B. Gao, and W. Wang, “Motion recognition of the bilateral upper-limb rehabilitation using sEMG based on ensemble EMD,” In Proc. IEEE Int. Conf. Mechatron. Autom., pp. 1637–1642, 2014.
- [36] Y. Zhang, P. Xu, P. Li, K. Duan, Y. Wen, Q. Yang, T. Zhang, and D. Yao, “Noise-assisted multivariate empirical mode decomposition for

- multichannel EMG signals,” *Biomed. Eng. Online*, vol. 16, p. 107, 2017.
- [37] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis,” In *Proc. R. Soc. Lond.*, pp. 903–995, 1998.
- [38] X. Mengying, Y. Xiaoli, X. Chenli, and Y. Bin, “EMG signal processing and application based on empirical mode decomposition,” *Math. Comput. Sci.*, vol. 4, no. 6, pp. 99–103, 2019.
- [39] X. Zhang and P. Zhou, “Filtering of surface EMG using ensemble empirical mode decomposition,” *Med. Eng. Phys.*, vol. 35, no. 4, pp. 537–542, 2013.
- [40] B. Lee and S. K. Yoo, “Amplitude and phase analysis of the brain evoked potential about performing a task related to visual stimulus using empirical mode decomposition,” *Korean Journal of the Science of Emotion & Sensibility*, vol. 18, no. 1, pp. 15–26, 2015.
- [41] Z. Sun, X. Xi, C. Yuan, Y. Yang, and X. Hua, “Surface electromyography signal denoising via EEMD and improved wavelet thresholds,” *Math. Biosci. Eng.*, vol. 17, no. 6, pp. 6945–6962, 2020.
- [42] N. E. Huang, X. Chen, M. T. Lo, and Z. Wu, “On Hilbert spectral representation: a true time–frequency representation for non-linear and nonstationary data,” *Advances in Adaptive Data Analysis*, vol. 3, pp. 63–93, 2011.
- [43] C. Sapsanis, G. Georgoulas, A. Tzes, and D. Lymberopoulos, “Improving EMG based classification of basic hand movements using EMD,” In *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 5754–5757, 2013.
- [44] G. Koch, R. Zemel, and R. Salakhutdinov, “Siamese neural networks for one-shot image recognition,” *ICML Deep Learning Workshop*, vol. 2, 2015.

- [45] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," In Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 539–546, 2005.
- [46] M. F. Zanuy, "Data fusion in biometrics," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 20, no. 1, pp. 34–38, 2005.
- [47] C. Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, "Deeply-supervised nets," In Proc. Int. Conf. Artif. Intell. Stat., pp. 562–570, 2015.
- [48] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," In Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 1–9, 2015.
- [49] Y. Hu, Y. Wong, W. Wei, Y. Du, M. Kankanhalli, and W. Geng, "A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition," *Plos One*, vol. 13, no. 10, p. 0206049, 2018.
- [50] EMG datasets repository, Available online: <http://www.rami-khushaba.com/electromyogram-emg-repository.html> (accessed on 21 February 2022).
- [51] R. N. Khushaba, S. Kodagoda, M. Takruri, and G. Dissanayake, "Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 10731–10738, 2012.
- [52] R. N. Khushaba and S. Kodagoda, "Electromyogram (EMG) feature reduction using mutual components analysis for multifunction prosthetic fingers control," In Proc. Int. Conf. Control Autom. Robot. Vis., pp. 1534–1539, 2012.
- [53] R. N. Khushaba, S. Kodagoda, D. Liu, and G. Dissanayake, "Muscle computer interfaces for driver distraction reduction," *Comput. Methods Programs Biomed.*, vol. 110, no. 2, pp. 137–149, 2013.
- [54] A. H. A. Timemy, R. N. Khushaba, G. Bugmann, and J. Escudero,

- “Improving the performance against force variation of EMG controlled multifunctional upper-limb prostheses for transradial amputees,” *IEEE Trans. Neural Syst. Rehabilitation Eng.*, vol. 24, no. 6, pp. 650–661, 2016.
- [55] R. N. Khushaba, A. H. A. Timemy, S. Kodagoda, and K. Nazarpour, “Combined influence of forearm orientation and muscular contraction of EMG pattern recognition,” *Expert Syst. Appl.*, vol. 61, no. 1, pp. 154–161, 2016.
- [56] R. N. Khushaba, M. Takruri, J. V. Miro, and S. Kodagoda, “Towards limb position invariant myoelectric pattern recognition using time-dependent spectral features,” *Neural Netw.*, vol. 55, pp. 42–58, 2014.
- [57] UCI machine learning repository, Available online: <http://archive.ics.uci.edu/ml/datasets/sEMG+for+Basic+Hand+movements>(accessed on 21 February 2022).
- [58] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, A. G. M. Hager, S. Elsig, G. Giatsidis, F. Hassetto, and H. Muller, “Electromyography data for non-invasive naturally-controlled robotic hand prostheses,” *Sci. Data*, vol. 1, p. 140053, 2014.
- [59] Ninaweb, Available online: <http://ninapro.hevs.ch/node/17>(accessed on 21 February 2022).
- [60] S. Pizzolato, L. Tagliapietra, M. Cognolato, M. Reggiani, H. Muller, and M. Atzori, “Comparison of six electromyography acquisition setups on hand movement classification tasks,” *Plos One*, vol. 12, no. 10, p. 0186132, 2017.
- [61] Ninaweb, Available online: http://ninapro.hevs.ch/DB5_DoubleMyo (accessed on 21 February 2022).
- [62] UCI machine learning repository, Available online: <http://archive.ics.uci.edu/ml/datasets/emg+dataset+in+lower+limb>(accessed on 21 February 2022).
- [63] J. S. Kim, C. H. Song, E. Bak, and S. B. Pan, “Multi-session

surface electromyogram signal database for personal identification,” *Sustainability*, vol. 14, no. 9, p. 5739, 2022.

- [64] IT Research Institute of Chosun University, Available online: <http://www.chosun.ac.kr/user/indexSub.do?codyMenuSeq=842922712&siteId=riit>(accessed on 10 January 2022).