



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](#)

공학석사 학위논문

딥러닝 기반의 멀티모달 감정 인식을 향
상을 위한 데이터 노이즈 및 신뢰도 연
구

**Data Noise and Reliability Study for
Improvement of Multi-Modal Emotion
Recognition Rate Based on Deep Learning**

2019년 2 월

서경대학교 대학원

전자컴퓨터공학과

윤 준 한

공학석사 학위논문

딥러닝 기반의 멀티모달 감정 인식을 향
상을 위한 데이터 노이즈 및 신뢰도 연
구

**Data Noise and Reliability Study for
Improvement of Multi-Modal Emotion
Recognition Rate Based on Deep Learning**

지도교수 김 진 현

2019 년 2 월

서경대학교 대학원

전자컴퓨터공학과

윤 준 한

딥러닝 기반의 멀티모달 감정 인식을 향
상을 위한 데이터 노이즈 및 신뢰도 연
구

**Data Noise and Reliability Study for
Improvement of Multi-Modal Emotion
Recognition Rate Based on Deep Learning**

이 논문을 석사학위 논문으로 제출함

2018 년 12 월

서경대학교 대학원

전자컴퓨터공학과

윤 준 한


윤 준 한의 석사학위 논문을 추천함


2019년 2월

지도교수 김 진 현

서경대학교 대학원

윤 준 한의 석사학위 논문을 인준함.

위 원 장 : 이 광 업 (인) 

위 원 : 박 태 립 (인) 

위 원 : 김 진 권 (인) 

2018 년 12 월 28 일

서 경 대 학 교 대 학 원

국 문 초 록

최근 딥러닝 모델을 이용한 감정 인식 연구 중 다양한 생체 신호를 분석하는 방법론이 조명되고 있다. 이때 감정 분류의 신뢰성에 가장 많은 영향을 끼치는 것은 학습 데이터의 종류에 따른 신뢰성 및 분류 방법이다. 생체 신호의 경우 데이터의 신뢰성이 노이즈 비율에 따라 결정되므로 노이즈 검출 방법이 우수할수록 신뢰도가 올라가며, 감정을 정의하는 방법론에 따라 그에 맞는 적절한 감정 평가 방법이 수반될 때 정확하게 감정을 분류할 수 있다. 본 논문에서는 Valence 및 Arousal 기준의 멀티모달 생체 신호에 대해, 생체 신호 데이터의 신뢰성을 검증하기 위한 고정 웨이블릿 기반의 노이즈 임계값 설정 알고리즘과 Valence-Arousal 값에 따른 신뢰도를 설정하는 법, 그리고 학습시 데이터 신뢰도와 Valence-Arousal 신뢰도를 반영하여 감정 인식률을 향상하는 방법을 제안한다. 웨이블릿 변환을 이용해 신호의 웨이블릿 성분을 추출 후, 해당 성분의 왜도와 첨도를 구하여 험펠 식별자를 통해 계산된 임계값으로 노이즈를 검출한 후, 원신호에 대한 노이즈 비율을 고려하여 데이터의 신뢰성을 평가하고 가중치로 환산한다. 더불어 감정 데이터 분류 시 Valence-Arousal 평면의 각 사분면 중앙값과의 유클리디언 거리를

가중치로 환산하여 학습데이터량 조절시 두 요소를 반영한다.
ASCERTAIN 데이터셋을 활용하여 나타난 감정 인식을 개선 정도
를 통해 제안된 알고리즘의 성능을 검증한다.

목차

I. 서론	1
II. 연구 배경	3
1. 기존의 감정 정의 및 분류에 관한 선행 연구	3
2. Valence-Arousal에 따른 감정 정의 및 분류	5
3. 감정 분류에 사용되는 주요 생체 신호	7
(1) 뇌전도(EEG)	7
(2) 심전도(ECG)	9
(3) 피부 반응(GSR)	11
4. 생체 신호별 노이즈 검출 관련 연구	12
5. 기계학습을 통한 감정 분류 선행 연구	14
III. 제안 방법	17
1. 데이터 신뢰성 평가	18
(1) 노이즈 검출	18
1) Stationary Wavelet Transform	18
2) 통계적 특성을 이용한 노이즈 검출	21
3) 노이즈 데이터 비율에 따른 데이터 신뢰성	23
2. Valence-Arousal 신뢰도 설정	24
3. 데이터 신뢰도와 Valence-Arousal 신뢰도 반영	25
IV. 실험 방법 및 결과 분석	27
1. 실험 준비 및 방법	27
(1) ASCERTAIN 데이터셋	27
(2) Stacked AutoEncoder	31
(3) 딥러닝 모델 설계	33
(4) 컴퓨팅 환경	35
2. 결과 및 분석	356
V. 결론 및 향후 연구 방향	38
참고문헌	40

그림 목차

그림 1 자기설문평가지	6
그림 2 기쁨(좌), 불쾌함(우) 감정 유발 콘텐츠.....	6
그림 3 Valence-Arousal 평면	7
그림 4 시계열에 따른 EEG 파형 예시, 빨간 동그라미는 노이즈를 의미 ...	8
그림 5 일반적인 EEG 측정 모습.....	9
그림 6 시계열에 따른 ECG 파형 예시, 빨간 동그라미는 노이즈를 의미..	10
그림 7 일반적인 ECG 측정 모습.....	10
그림 8 시계열에 따른 GSR 파형 예시, 빨간 동그라미는 노이즈를 의미..	11
그림 9 GSR 측정 모습.....	12
그림 10 기본 DBN 모델 구조 예시	15
그림 11 기본 CNN 모델 구조 예시	15
그림 12 기본 SAE 모델 구조 예시	16
그림 13 DWT 변환과정 예시(3레벨)	19
그림 14 SWT 변환과정 예시(3레벨).....	19
그림 15 모 웨이블릿 종류에 따른 파형 예시	21
그림 16 Valence-Arousal 평면에서의 각 사분면 중앙값(빨간점)과의 거리에 따른 비율 개념도.....	25
그림 17 ASCERTAIN 데이터 구성	28
그림 18 ASCERTAIN 무비클립 영역도	28
그림 19 K-means 클러스터링을 이용한 58명 데이터의 분류	29
그림 20 AutoEncoder 구조	31
그림 21 Stacked AutoEncoder 학습 과정	32
그림 22 퍼셉트론의 기본 모형.....	33
그림 23 딥러닝 모델 블록 다이어그램	35

표 목차

표 1 생체 신호별 특징 데이터	30
표 2 제안한 딥러닝 모델 구조	35
표 3 데이터 신뢰성 평가 방법에 따른 감정 인식률 비교.....	36
표 4 제안한 Valence-Arousal 가중치 적용여부에 따른 감정 인식률 비교...	37
표 5 제안한 방법을 모두 적용한 경우와 그렇지 않은 경우에 따른 감정 인식률 비교.....	37

I. 서론

컴퓨터 상호 작용(HCI)의 유연성 및 효율성을 개선하기 위한 목표는[1] 다양한 생활 환경 또는 작업 환경에서 필요로 하는 인간과 컴퓨터 간의 소통 관계를 향상하는 것과 일치한다. 이를 실현하기 위해 컴퓨터 즉, 지능형 HCI 시스템은 자기 적응 능력을 필요로 하며 [2], 인간의 의사소통에 대한 정확한 이해는 컴퓨터 에이전트가 올바른 피드백을 하기 위한 필수 요소이다[3]. 인간은 자신의 의견 및 생각을 구두 및 비언어적 형태로 이루어진 감정이란 것으로 전달하게 되며, 이러한 점으로 인해 컴퓨터는 인간의 감정을 인지하거나 혹은 예측도 할 수 있어야 한다.

인간의 감정을 정의하기 위한 다양한 방법 중 신경, 분노, 기쁨, 지루함과 같은 관점에서의 개별적인 척도를 포함시키거나, Valence-Arousal 평면이란 것을 이용할 수 있다[3-5]. Valence-Arousal 평면에서 Valence는 긍정과 부정의 정도를 나타내며, Arousal은 평온과 흥분 정도를 나타낸다[6-9]. 한가지 단점으로, 감정 분류를 할 때 각 사분면 내 Valence와 Arousal 값이 중앙에 가까워질수록 감정 정의에 대한 신뢰성이 모호할 수 있는 단점이 있다.

감정 인지, 분류에 대한 대표적인 접근법 중 하나로 EEG(Electroencephalogram), GSR(Galvanic skin response), ECG(Electrocardiogram), EMG(Electromyogram), PPG(Photoplethysmography) 등과 같은 전기 생리학적 정보를 측정하여 정서적 상태를 추론하는 방법이 있다. EMG를 제외한 다른 생체 신호들의 경우 자율 신경계가 담당하고 있으며, 감정 상태의 변화에 따라 해당 생체 신호들의 다양한 변화가 관찰

된다. 자율 신경계와 감정 상태의 관계에 대한 연구는 William James(1884)가 해당 관계의 패턴에 대한 연구를 내놓으며 시작되었고, 현재까지도 그 연구가 활발히 진행되고 있다.

측정된 생체 신호들은 전처리 과정을 통해 해당 정보를 나타낼 수 있는 특징 데이터로 가공되며, 이러한 특징들을 이용하여 기계학습 기반 모델을 통해 특정 감정 상태의 잠복 패턴을 밝힐 수 있다. 이때 데이터의 신뢰성에 따라 감정 분류의 정확도가 결정될 수 있고, 이는 곧 노이즈의 양과 직결되기에 노이즈 검출과정이 중요하다고 볼 수 있다. 웨이블릿 변환은 다양한 생체 신호들의 노이즈를 검출하는 대표적인 방법론이며 실제 그 성능을 입증하였으나, 변환 후 노이즈 성분을 정의하는 방법에 따라 검출 성능이 달라질 수 있다.

본 논문에서 제안하는 방법은 두 가지이다. 첫 번째는 생체 신호 데이터에 대해 웨이블릿 변환 후 왜도와 첨도를 구하여 햄펠 식별자(Hampel identifier)를 통해 계산된 임계값으로 노이즈를 검출하고, 데이터의 신뢰성을 평가하는 것이다. 두 번째는 평가된 데이터 신뢰성을 바탕으로 한 가중치와 Valence-Arousal 평면의 각 사분면 중앙값으로부터의 유클리디언 거리 가중치를 반영하여 감정 분류 정확성을 향상하는 것이다.

본 논문은 다음과 같이 구성된다. 제 2장에서는 제안한 두 가지 방법을 이해하기 위한 이론 및 배경들을 설명한다. 제 3장에서는 제안하는 방법들을 설명한다. 제 4장에서는 제안한 방법들에 대한 실험 및 평가를 통해 우수성을 검증한다. 제 5장에서는 결론을 내린다.

II. 연구 배경

감정을 정의하고 분류하는 방법은 오랜 시간 다양한 형태로 연구되어 왔다. 그러나 감정의 종류가 매우 다양하고 개인적인 편차가 커 현재까지도 활발한 연구가 진행 중에 있다. 이러한 감정의 정의 및 분류에 있어 어떠한 연구들이 있었는지 확인해 보고, 본 논문에서 사용할 Valence-Arousal 평면에 대해 소개한다.

인간의 모든 감정 상태는 생체 신호로 나타나게 되고, 이는 역으로 생체신호를 통해 감정 상태를 구분할 수 있음을 의미한다. 이 분야에서 전통적으로 쓰이는 주요 생체 신호들 중 본 논문에서 제안한 방법을 적용할 신호 및 노이즈 관련 연구를 간략히 소개한다.

추출된 생체 신호들을 바탕으로 감정 분류시 머신러닝 기법을 많이 사용한다. 지금까지 쓰인 다양한 머신러닝 모델에 대해 살펴보기로 한다.

1. 기존의 감정 정의 및 분류에 관한 선행 연구

Wilhelm wundt는 인간의 감정을 긴장과 이완, 기쁨과 불쾌, 흥분과 안정의 세 가지 영역으로 분류하였으며[10], Robert Plutchik는 슬픔과 기쁨, 공포 및 분노, 혐오와 수용, 놀람과 예상의 여덟 가지 항목으로 감정을 표현하고자 하였다. Wundt와 Plutchik가 제시한 감정 분류의 예에서 알 수 있듯이 큰 차이는 아니지만 아직도 인간

의 감정분류에는 서로 이견이 존재할 수 밖에 없다고 주장한다. 결론적으로 생체신호를 통해 인간의 감정을 분류하려는 시도는 방법론적인 문제보다는 무엇을 대상으로 할 것인가에 대한 감정의 근원적인 문제와 직면하게 되는 것이다. 즉, 어떤 감정을 어떤 기준으로 세분화하여 추출하는가에 대한 문제의 솔루션이 우선적으로 제시되어야 한다. 또 다른 문제점은 개인별로 경험과 지식이 다양할 뿐만 아니라 개인적인 감정의 종류와 그에 대한 반응 또한 매우 다양하여 일반화가 여의치 않다는 것이다. 즉, 특정 상황에서 어떤 사람은 해당 감정에 대해 매우 큰 반응을 보이지만 그러한 상황을 이미 반복적으로 경험한 사람 또는 관심조차 없는 사람의 경우 미미한 반응을 보이고, 생체신호의 반응 역시 그 변화가 크지 않을 수 있다. 이러한 여러 가지 어려움 때문에 감정에 대한 연구는 대부분 제한된 조건 아래 특정한 감정을 유발하는 자극을 주고 이에 대한 상반된 감정 유발 상태의 비교 및 분석이 많이 이용되고 있다. 대표적으로 Levenson은 여러 감정 상태에서의 자율신경계 반응에서 행복감 같은 긍정적 감정에서의 심박 수 및 피부 저항이 분노, 공포 등의 부정적 감성에서 보다 낮게 나옴을 주장하였다[11]. 그러나 해당 연구 결과들은 주로 특정 상태에서의 생체 신호 중에서도 극히 일부의 특성과 비교하여 통계적인 차이를 나타낸 사실만을 주장할 뿐, 슬픔, 분노, 기쁨 등의 감정 상태를 분류할 수 있는 생체 신호의 기준이나 획일화된 특성은 제시하지 못하고 있다.

2. Valence-Arousal에 따른 감정 정의 및 분류

일반적으로 감정은 물질적 요소와 인지적 요소의 비교적 복잡한 결합으로 정의한다. 물질적 요소는 생리적 변화에 따른 신체적 상태를 나타내고, 인지적 요소는 순수한 정신적 감정 상태를 의미한다 [12]. 물질적 요소는 육체적, 생리학적 시스템의 활동으로 정의되고 있으며, 교감 신경 및 부교감 신경의 활동에 의한 결과로써 정량화할 수 있다. 물질적 감정은 생리적으로 흥분 및 안정 상태로 나뉘며 감정 평가 시 감정 상태에 대한 분류를 단순한 두가지 관점으로 해석한다. 그러나 복잡한 감정 상태에 대해 물질적 해석만으로 평가하기는 어렵다. 따라서 인지적 요소를 추가하여 두 가지를 같이 고려한 감정의 해석은 다시 긍정적 감정과 부정적 감정으로 분류된다. 이 두 가지 항목으로 분류된 감정은 어떠한 절대적인 기준으로 분류하는 것이 아닌 상대적인 판단 기준을 사용한다. 하지만 이 방법의 경우 주관적인 평가를 통해 정성적 해석이 가능할 뿐 정량적인 해석은 불가능하다. 이에 따라 Lang은 감정의 속성을 다차원의 공간 상에서 정의하는 연구를 진행하였고, 이와 같이 구조적 측면을 통해 감정을 정의하는 방법으로 가장 널리 쓰이고 있는 것이 2차원 Valence-Arousal 평면이다. Valence는 긍정과 부정에 관한 감정의 정도를 나타내고, Arousal은 평온과 흥분의 정도를 의미한다. Valence와 Arousal은 Lang에 의해 고안된 SAM(Self-Assessment Manikin), 이른바 자기설문평가지를 이용해 측정하며, 그림 1과 같은 형태로 되어 있다.

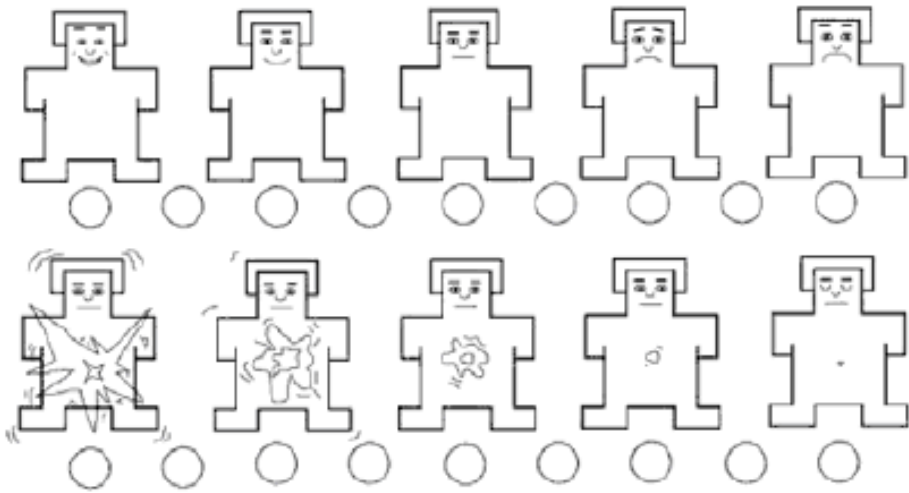


그림 1 자기설문평가지

상단 5개는 Valence의 정도를 뜻하며 오른쪽으로 갈수록 부정, 왼쪽으로 갈수록 긍정적임을 나타낸다. 하단 5개는 Arousal의 정도를 뜻하며 오른쪽으로 갈수록 평온 왼쪽으로 갈수록 흥분된 상태임을 뜻한다. 피실험자는 감정을 유발하는 그림 2와 같은 콘텐츠 시청 후, 해당 설문조사를 작성하여 자신의 감정 상태를 기록한다. 기록된 값은 K-means 클러스터링을 통해 4개 영역으로 나눔으로써 Valence-Arousal 평면을 생성할 수 있다.



그림 2 기쁨(좌), 불쾌함(우) 감정 유발 콘텐츠

그림 3은 Valence-Arousal 평면의 일반적인 형태이다. 이산적으로 분리된 보편적인 감정의 속성을 구분한 것으로, 크게는 네 가지 영역군 내에서 정의하고 있다. 그러나 이와 같은 형태의 분류는 사분면의 중심에서 멀어질수록 네 가지 영역군에 대한 경계가 모호해져 정확한 감정 분류가 힘들다는 단점이 있다.

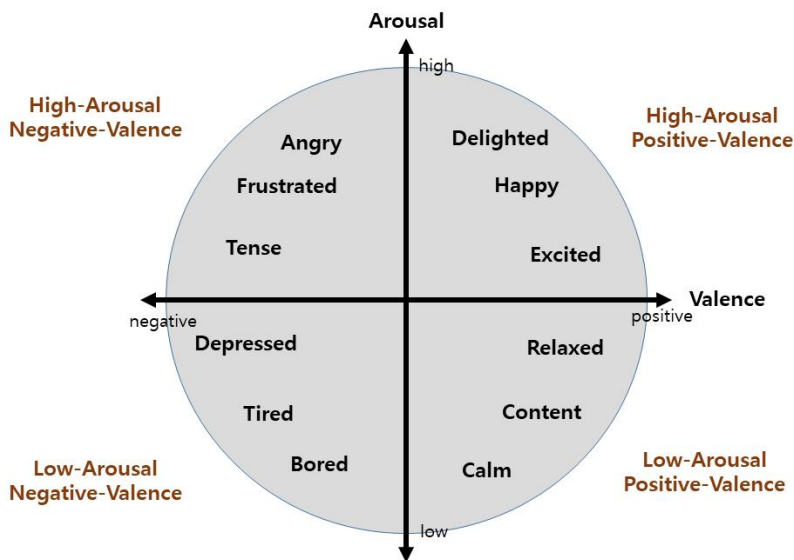


그림 3 Valence-Arousal 평면

3. 감정 분류에 사용되는 주요 생체 신호

(1) 뇌전도(EEG)

뇌의 전기적 활동을 나타내는 신호로, 다수의 전극을 이용하여 뉴

런에서의 이온 전류로 인한 전압 변동을 측정한다. 신경 과학, 인지 과학, 인지 심리학 신경 언어학 및 정신 생리학 연구에서 광범위하게 사용되며, 뇌의 변화를 추적하는 방법 중 가장 널리 쓰이고 있다. 일반적으로 측정된 신호의 스펙트럼 내용을 기반으로 감정을 예측한다.

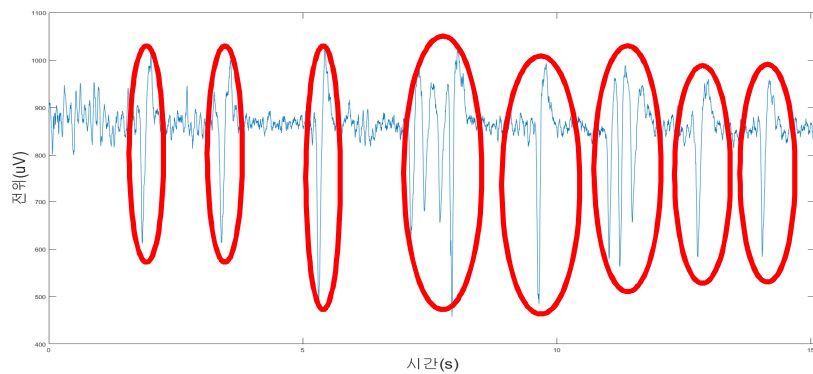


그림 4 시계열에 따른 EEG 파형 예시, 빨간 동그라미는 노이즈를 의미



그림 5 일반적인 EEG 측정 모습

(2) 심전도(ECG)

다수의 전극을 이용하여 각 심장 박동에 대해 탈분극 및 재분극으로 발생하는 피부의 미세한 전기적 변화를 측정하여 심장 근육의 전기 생리학적 패턴을 감지하는데 사용한다. 해당 패턴은 규칙적인 형태로 반복되며, 이를 통해 심장의 기능적 상태를 나타내는 심박수 및 산소포화도 등을 검출하는데 사용되고 있다. 감정 평가에는 주로 심박수를 사용한다.

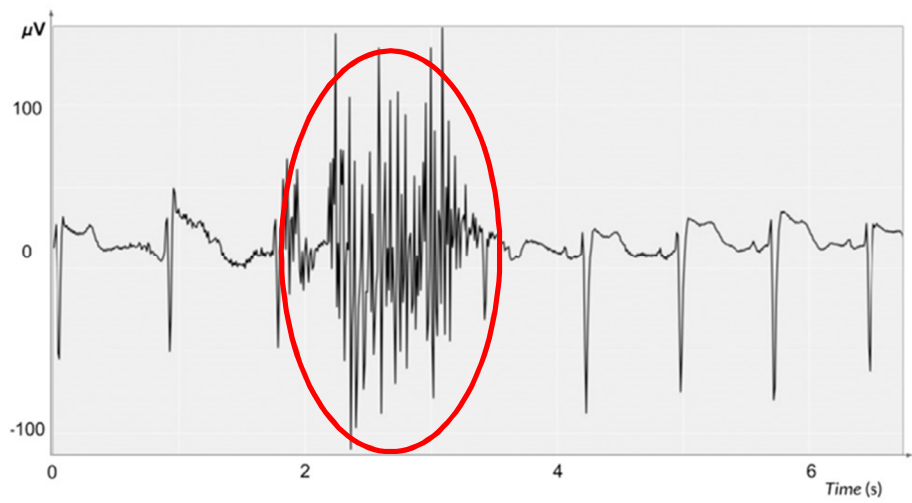


그림 6 시계열에 따른 ECG 파형 예시, 빨간 동그라미는 노이즈를 의미



그림 7 일반적인 ECG 측정 모습

(3) 피부 반응(GSR)

EDA(Electrodermal activity)라고도 불리우며 인체의 특정 상태에 따른 피부의 전기적 특성을 나타낸다. 피부 저항이 피부의 땀샘 상태에 따라 다를 것을 기반으로 하며, 자율 신경계와 교감 신경계가 크게 자극되면 땀샘 활동도 증가하여 피부 전도성을 증가시키고 두 전극을 통해 이를 측정한다. 신경계의 제어와 관련이 있기에 감정 반응에 대한 기준으로 사용할 수 있으며 ECG, EEG와 함께 감정 분류에 관한 연구에 많이 이용되고 있다.

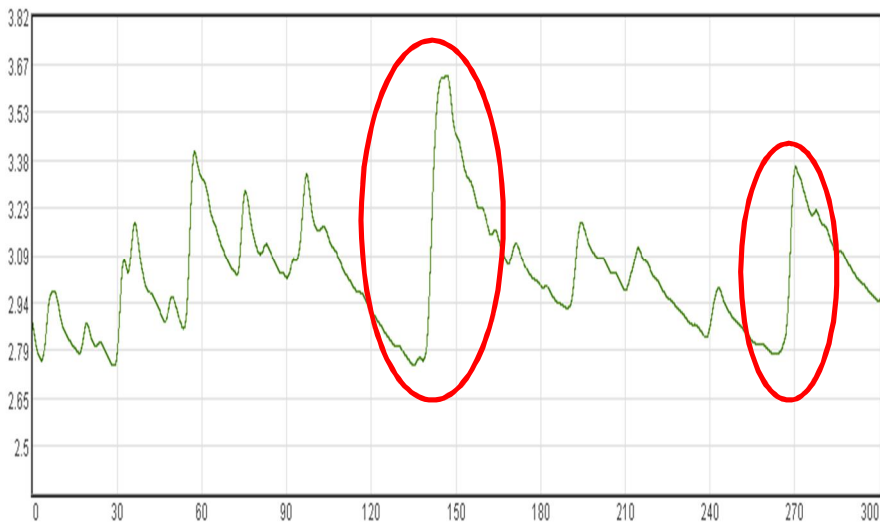


그림 8 시계열에 따른 GSR 파형 예시, 빨간 동그라미는 노이즈를 의미



그림 9 GSR 측정 모습

4. 생체 신호별 노이즈 검출 관련 연구

각 생체신호는 신호의 특징에 따라 여러 노이즈 제거 알고리즘을 가지고 있다. ICA(Independent Component Analysis)나 PCA(Principal Component Analysis)는 EEG의 노이즈 제거에 사용될 수 있는 대표 기술이다. 원시 데이터가 아닌 ICA를 통해 취득한 독립 성분에서의 부분 공간 필터링을 적용하는 방법이며[13], 신호 성분의 독립성에 대한 충분한 보장이 필요하다. PCA는 주성분 분석을 이용하여 적응 임계값을 설정함으로써 노이즈를 제거할 수 있지만[14], 실제 신호의 노이즈와의 관계가 정립되어 있지 않은 상태에서 제거한다는 단점

이 존재한다. ECG는 사비츠키-골레이 필터링 방법과 적응 필터 방법이 있다. 적응 필터는 노이즈가 존재하지 않는 참조 개념의 신호를 필요로 하며, 원시 신호와의 MSE(Mean Square Error)를 통해 노이즈를 제거할 수 있다. 사비츠키-골레이 필터링은 샘플 데이터 및 양사이드의 데이터들을 다항식으로 회귀시켜 만든 다항식으로 샘플 데이터의 값을 설정한다. 매번 다항식을 회귀시키며 계산할 필요가 없고, 특징적인 컨볼루션을 사용하여 찾아낼 수 있는 장점이 있다. 고차 다항식은 데이터가 급격히 변하는 구간을, 저차 다항식은 그렇지 않은 곳을 스무딩할 수 있다. GSR은 저역 통과 필터[15-17] 및 지수 평활화[15]가 있다. 그러나 이러한 비적응형 방법들은 돌연 나타나는 예상보다 큰 강도의 노이즈 대처에 미흡하며, 시계열의 전체가 모두 필터링 되어 원시 신호의 왜곡률을 높일 수 있다는 단점이 존재한다.

5. 기계학습을 통한 감정 분류 선행 연구

기계 학습 접근법은 정교한 생체 신호 획득 기기의 해상도 및 샘플링 주파수를 고려할 때, 신경 물리학 데이터의 방대한 양 분석을 쉽게 할 수도 있도록 한다[18-22]. 특히, 패턴 분류기는 상이한 형식의 생리학적 특징을 융합시킬 수 있다. 최근 Iacoviello et al.은 하이브리드 분류 프레임 워크를 만들기 위해 이산 웨이블릿 변환, PCA 및 SVM(Support Vector Machine)을 결합했다. Khezri et al.은 KNN 분류기를 통해 6개의 기본 감정을 인식하기 위해 GSR과 결합 된 3채널 EEG를 사용했다. Verma et al.은 EEG, EMG, ECG, GSR 및 EOG를 통합한 앙상블 분류 접근법을 개발했다. Mehmood와 Lee는 EEG, EMG, GSR, ECG, ERP로부터 정서적 지표를 추출하기 위해 독립적인 성분 분석을 사용했다. 이러한 고차원적인 신체적 특징을 추상화할 수 있다는 장점으로 인해 감정 분류를 위한 많은 학습 방법이 개발되었고 의미있는 결과들이 도출되었다. 이미 널리 사용되고 있는 딥러닝의 기본 요소에는 DBN(Deep Belief Network), SAE(Stacked AutoEncoder) 및 CNN(Convolutional Neural Network)가 있다.

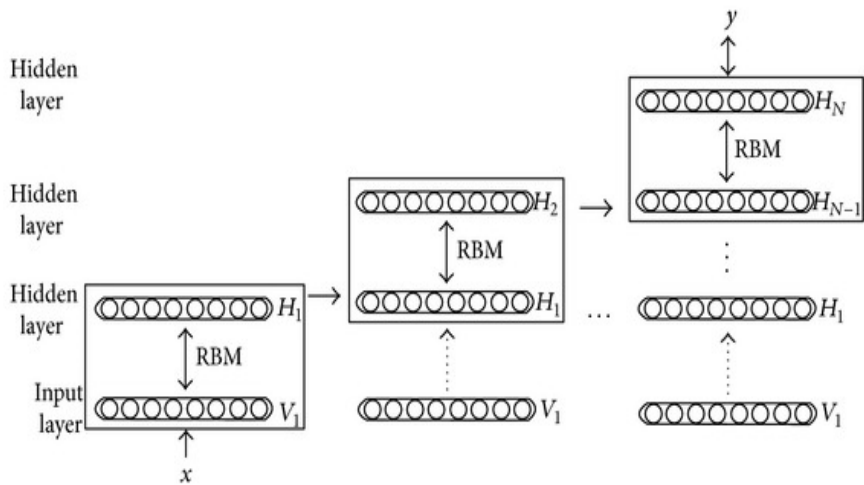


그림 10 기본 DBN 모델 구조 예시

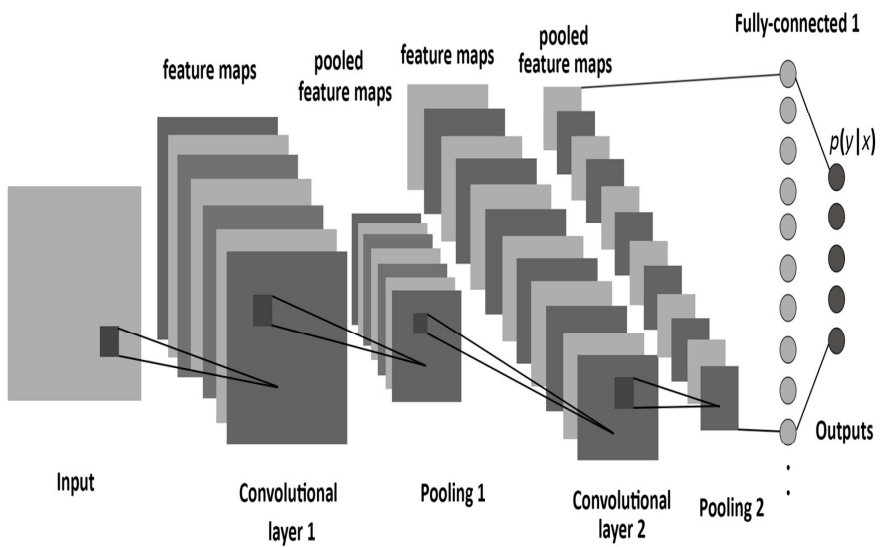


그림 11 기본 CNN 모델 구조 예시

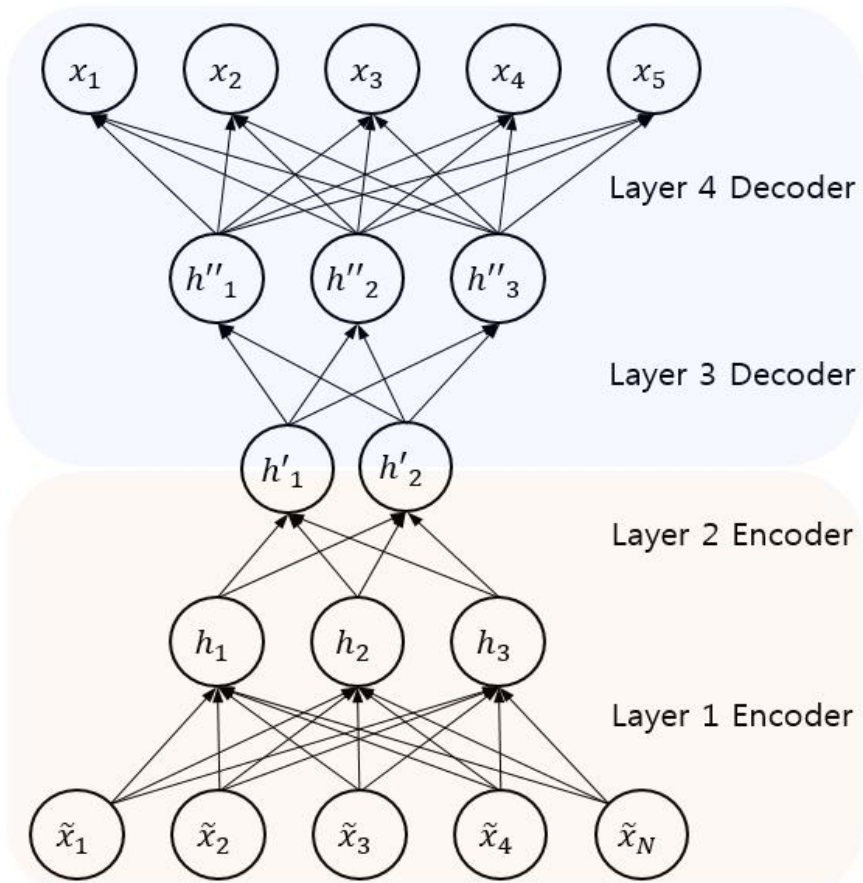


그림 12 기본 SAE 모델 구조 예시

특히 Wand와 Shang은 표준 DBN을 채택하여 비 숙련 사전 훈련을 기반으로 생리적 생체 데이터로부터 피처를 추출하고 Arousal, Valence 및 선호도의 수준을 평가하기 위해 3개의 심층 분류기를 만들었다[23].

III. 제안 방법

데이터 신뢰성은 일반적으로 검출된 노이즈의 양에 영향을 받는다. 따라서 노이즈 검출의 정확성이 데이터 신뢰성과 직결되는 것으로 해석할 수 있다. 웨이블릿 변환은 다양한 생체 신호들의 노이즈를 검출하는 대표적인 방법론이며 실제 그 성능이 입증되었으나, 변환 후 노이즈 성분을 정의하는 임계값 설정 방법에 따라 검출 성능이 달라질 수 있다. 생체 신호의 경우 노이즈가 없을 때 각 주기의 신호 패턴이 유사하므로 이를 이용하여 통계적 특성을 이용한 왜도, 첨도, 햄펠 식별자(Hampel identifier)를 통해 노이즈를 검출할 수 있다. 또한, 기존의 Valence-Arousal 평면을 이용한 네 개 카테고리, 즉 Low Valence-Low Arousal, High Valence-Low Arousal, Low Valence-High Arousal, High Valence-High Arousal의 단순한 감정 분류법은 평면의 중심에 가까워질수록 실질적 감정 속성의 경계가 모호해지는 단점으로 인해 정확성이 떨어진다. 따라서 앞서 언급한 노이즈 검출 방법을 이용해 추출한 데이터 신뢰성 가중치와, Valence-Arousal 평면에서 중심으로부터의 유클리디언 거리를 통해 산출한 가중치를 반영하여 감정을 평가하는 방법을 제안한다.

1. 데이터 신뢰성 평가

(1) 노이즈 검출

제안한 노이즈 검출 기법은 다음의 두 과정으로 설명할 수 있다.

- 입력 신호에 대해 SWT(Stationary Wavelet Transform) 수행.
- 웨이블릿 계수에 대해 왜도와 첨도 신호를 구하고 Hampel 식별자를 이용해 노이즈 검출.

1) Stationary Wavelet Transform

DWT(Discrete Wavelet Transform)는 신호에서 시계열 영역의 영향을 직접적으로 받으므로 링잉 효과(Ringing effect)와 같은 예상 외의 여러 노이즈를 유발할 수 있다[24]. 이를 해결하고자 시계열과 상관없이 DWT의 다운샘플링 과정을 생략한 SWT가 등장하였다. 그림 13과 그림 14는 각각 원신호 $x[n]$ 에 대한 DWT와 SWT의 변환과정을 보여준 것이다

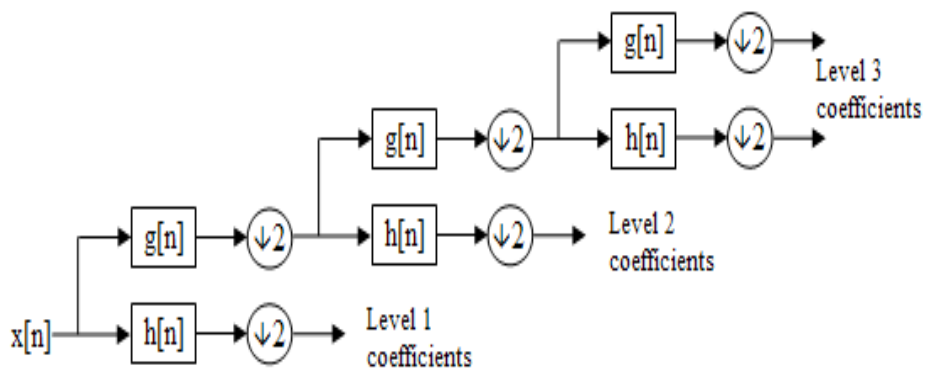


그림 13 DWT 변환과정 예시(3레벨)

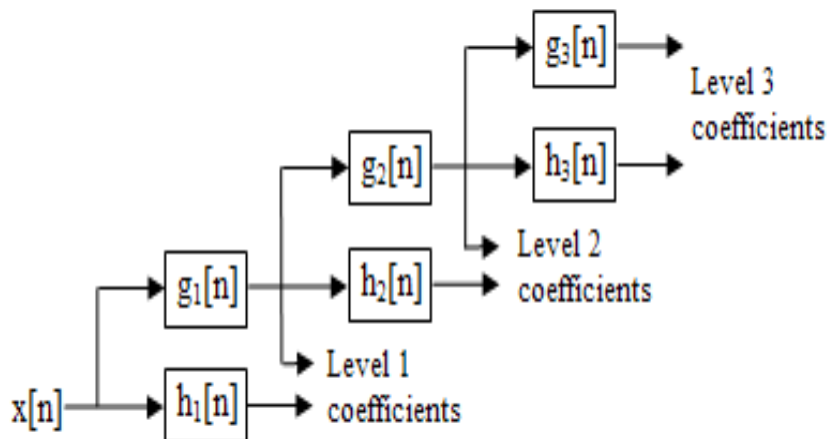


그림 14 SWT 변환과정 예시(3레벨)

SWT는 다음 식을 통해 원본 신호의 Approximation Coefficients와 Detail Coefficients를 추출할 수 있다.

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n - k] \quad (1)$$

$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n - k] \quad (2)$$

$$C_{2^j}^{2^j k+p} = \left\langle y(t), 2^{-j/2} \Phi \left(\frac{t-p}{2^j} - k \right) \right\rangle \quad (3)$$

$$d_{2^j}^{2^j k+p} = \left\langle y(t), 2^{-j/2} \Psi \left(\frac{t-p}{2^j} - k \right) \right\rangle \quad (4)$$

식 (1)과 (2)는 각각 Approximation Coefficients와 Detail Coefficients로 이루어진 신호를 나타내고 있다. 식(3)과 (4)는 각 계수들을 구하는 식을 의미하며, Φ 와 Ψ 는 각각 Approximation, Detail 함수이고, j 와 k 는 모 웨이블릿의 스케일링 정도와 이동변환 정도를 나타내고. $p \in 0, \dots, 2^j - 1$ 은 이산적인 상황의 모든 쉬프팅 범위를 의미한다. 제안한 알고리즘의 경우 $j \in 1, \dots, 9$ 로 설정하였고, 모 웨이블릿은 ECG와 GSR은 db4[26-27], EEG는 sym9[25]를 이용함으로써 각 신호의 특성에 맞게 선택하였다. 그림 15는 모 웨이블릿을 종류를 보여주고 있다.

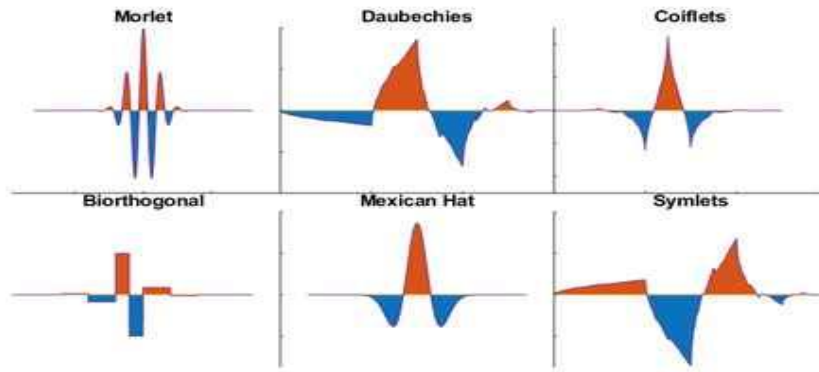


그림 15 모 웨이블릿 종류에 따른 파형 예시

2) 통계적 특성을 이용한 노이즈 검출

확률분포에서의 왜도는 대칭에서 어느 한쪽으로 치우쳐진 정도를 의미하고, 첨도는 뾰족한 정도를 나타낸다. 이는 신호에서도 그대로 활용할 수 있으며, 데이터의 표준 편차를 기반으로 다음과 같이 구할 수 있다.

$$STD = \sqrt{1/n \sum_{i=1}^n (x_i - E[x])} \quad (5)$$

$$SKEW = \frac{1/n \sum_{i=1}^n (x_i - E[x])^3}{STD^3} \quad (6)$$

$$KURTO = \frac{1/n \sum_{i=1}^n (x_i - E[x])^4}{STD^4} \quad (7)$$

위 식에서 $E[x]$ 는 각 신호값 x_i 들의 평균을 의미하고, STD는 표준

편차이다. 노이즈가 존재하지 않을 때 주기별로 신호 파형이 유사한 특징을 가지므로 앞서 제시한 통계적 특성들 역시 주기마다 대략적인 일치성을 보인다. 노이즈가 존재할 경우 각 주기별 파형이 변경되어 결과적으로 통계적 특성 역시 변화하므로, 침도와 왜도 같은 통계적 특성을 통해 노이즈를 충분히 검출할 수 있다. 이후 해당 두가지 통계 데이터에서 햄펠 식별자를 이용하여 노이즈 임계값을 다음과 같이 구할 수 있다.

$$\text{Med} = \text{Median}(\text{signal}) \quad (8)$$

먼저 위의 식과 같이 왜도와 침도 신호에 대한 윈도우별 국소중앙값들을 구한다.

$$\text{Mad} = \text{Median}(\text{abs}(\text{signal} - \text{Med})) \quad (9)$$

식 (8)에서 나온 중앙값을 통해 식 (9)처럼 왜도와 침도 신호에 대해 중앙값 편차를 구한 후 절대값을 취한다. 이후 식 (10)과 식 (11)를 통해 표준편차 추정값 SD를 계산한다.

$$\text{Param} = \frac{1}{\sqrt{2} \text{erf}^{-1}(\frac{1}{2})} \quad (10)$$

$$\text{SD} = \text{Param} * \text{Mad} \quad (11)$$

식 (12)와 같이 추정 표준편차 SD보다 3 이상 차이를 보이는 중앙

값 절대 편차에 해당하는 왜도 또는 첨도 데이터를 노이즈로 간주한다.

$$\text{abs}(\text{signal} - \text{Med}) > 2 * \text{SD} \quad (12)$$

3) 노이즈 데이터 비율에 따른 데이터 신뢰성

실험을 통해 임의로 정한 수치인 노이즈 데이터 비율 5%부터 40% 까지를 데이터 신뢰성 0~1로 표준화한다. 이때 노이즈 비율이 5% 이하일 때 신뢰성 1에 해당하고, 40% 이상일 때 신뢰성 0으로 간주한다.

2. Valence-Arousal 신뢰도 설정

Valence-Arousal 데이터 분류시 일반적으로 Valence-Arousal 평면의 사분면 영역을 활용한다. 그러나 서로 다른 사분면의 데이터에서 각 사분면 중심과의 거리가 멀 경우 기계학습 모델 관점에서 추상화된 데이터 특성의 분류 기준이 모호해질 수 있으며, 감정 인식률이 떨어질 위험성을 내포하게 될 수 있다. 이에 대한 해결책으로 Valence-Arousal 값이 각 사분면 중앙값에 가까울수록 반영 가중치를 적용할 필요가 있다. 가중치는 식 (13)의 중앙값으로부터의 유클리디언 거리에 따라 정하며, 중앙값에 가까울수록 0, 멀어질수록 1로 표준화 한다. 그림 16은 제안한 개념을 그래프로 표현한 것이다.

$$\|VA_{sample} - VA_{center}\| = \sqrt{\|VA_{sample}\|^2 + \|VA_{center}\|^2 - 2VA_{sample}VA_{center}} \quad (13)$$

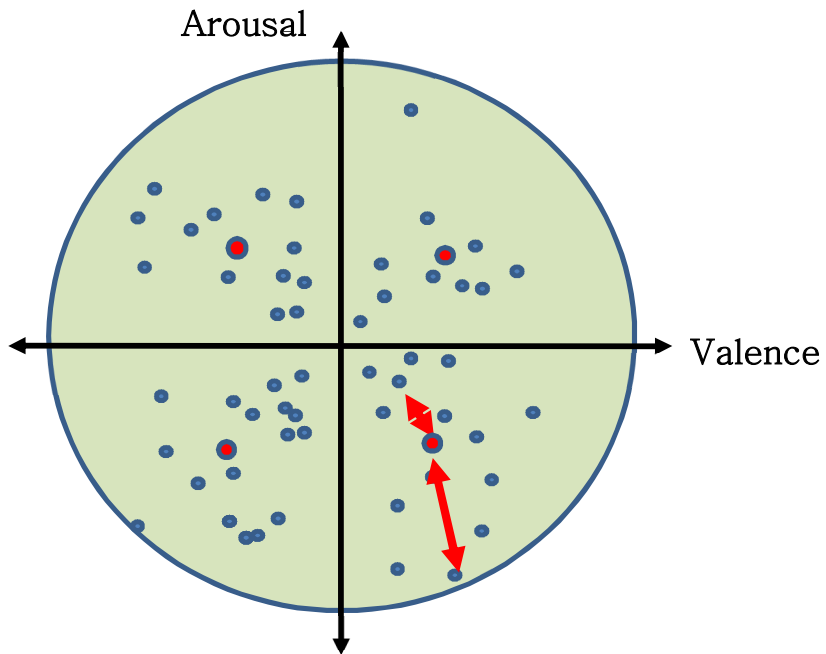


그림 16 Valence-Arousal 평면에서의 각 사분면 중앙값(빨간점)과의 거리에 따른 비율 개념도

3. 데이터 신뢰도와 Valence-Arousal 신뢰도 반영

딥러닝 기반의 솔루션은 일반적으로 학습데이터의 패턴에 기반한다. 따라서 목표한 방향으로 데이터를 학습하고자 할 경우 방향에 맞는 학습데이터량이 그렇지 않은 경우보다 상대적으로 많을수록 해당 방향의 패턴에 맞게 학습될 확률이 크다. 따라서 데이터 신뢰도와 Valence-Arousal 신뢰도를 통해 신뢰도가 높은 데이터를 판별하여 학습데이터량을 조절한다면 개선된 감정인식률을 얻을 수 있다. 먼저 데이터 신뢰도와 Valence-Arousal 신뢰도는 하나의 최종 신뢰도

로 나타내기 위해서는 다음 식 (14)를 이용할 수 있다.

$$\text{FinalQuality} = \frac{\text{DataQuality} + \text{VAQuality}}{2} \quad (14)$$

위 식과 같이 두 신뢰도의 평균을 최종 신뢰도로써 활용하며 그 값의 범위 역시 0~1이다. 각 학습데이터셋에 대한 최종 신뢰도 비율이 정해졌으면 학습데이터셋의 양을 조절한다. 예를 들어 최종 신뢰도 비율이 0.7일 경우 해당 데이터셋 수를 7배 복제 증가시키고, 1.0일 경우 10배 복제 증가시킨다. 이를 통해 최종 신뢰도가 높은 데이터를 더 많이 학습시켜 감정인식률 향상을 추구한다.

IV. 실험 방법 및 결과 분석

멀티모달 생체 신호 데이터셋 ASCERTAIN[28]과 앙상블 형태의 Stacked AutoEncoder[29] 딥러닝 모델을 이용하여 제안한 방법에 따른 감정인식률 성능을 검증한다. 먼저 ASCERTAIN에서 제공하는 데이터 신뢰성 지표를 이용한 경우와 제안한 데이터 신뢰성 지표, 그리고 데이터 신뢰를 이용하지 않은 경우의 감정 인식률 성능을 비교한다. 그 다음 데이터 신뢰성 지표를 고정하고 Valence-Arousal 가중치의 적용 여부에 따른 감정 인식률 성능을 비교한다. 마지막으로 제안한 방법 모두를 적용한 경우와 그렇지 않은 경우를 비교하여 감정 인식률 성능을 검증한다.

1. 실험 준비 및 방법

(1) ASCERTAIN 데이터셋

21명의 여성과 37명의 남성, 평균 연령 서른 살의 총 58명이 36편의 감정 유발용 무비 클립(Movie Clip)을 시청하면서 측정한 생리적 신호를 기록한 데이터셋이다. 각 무비 클립은 약 51~127초 사이의 분량으로 이루어져 있으며, Valence-Arousal 평면의 사분면에 무비 클립 18와 같이 배포되어 있다.

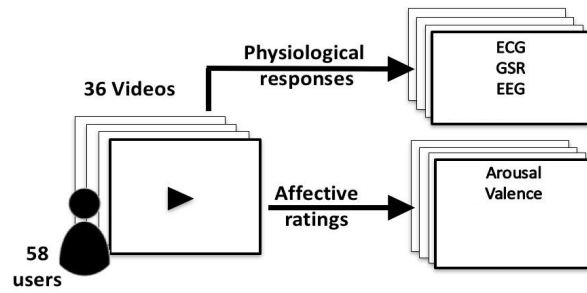


그림 17 ASCERTAIN 데이터 구성

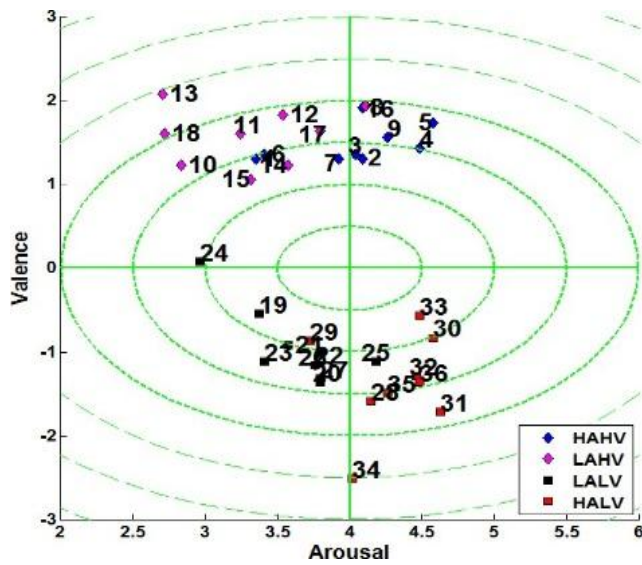


그림 18 ASCERTAIN 무비클립 영역도

각 클립을 본 후, 참여자들은 Valence-Arousal 등급에 레이블을 붙였다. 본 실험에서는 EEG, ECG, GSR을 사용하며, 표 1과 같이

ASCERTAIN에서 제공하는 생체신호별 특징 데이터들을 기계학습의 입력값으로 활용한다. 본 실험에 사용할 생체 신호인 EEG는 92차원, ECG는 32차원, GSR은 31차원으로 이루어져 있으며, EEG의 각 주파수별 PSD(Power Spectral Density)는 따로 추가하였다. 또한, 무작위로 선별한 40명은 학습데이터로, 18명은 테스트 데이터로 활용하였고, ASCERTAIN 데이터셋에서 제공하는 데이터 신뢰성 지표의 경우 방법과 마찬가지로 0~1로 맵핑 후 활용하였다. ASCERTAIN 데이터셋의 각 58명에 대한 Valence-Arousal 평면은 그림 19와 같으며, 해당 수치들은 데이터셋의 성격에 따라 변할 수 있다. K-means 클러스터링을 통해 구성되었다. 이후 학습데이터 라벨링시 샘플들이 속해있는 해당 사분면이 라벨링 기준이 된다.

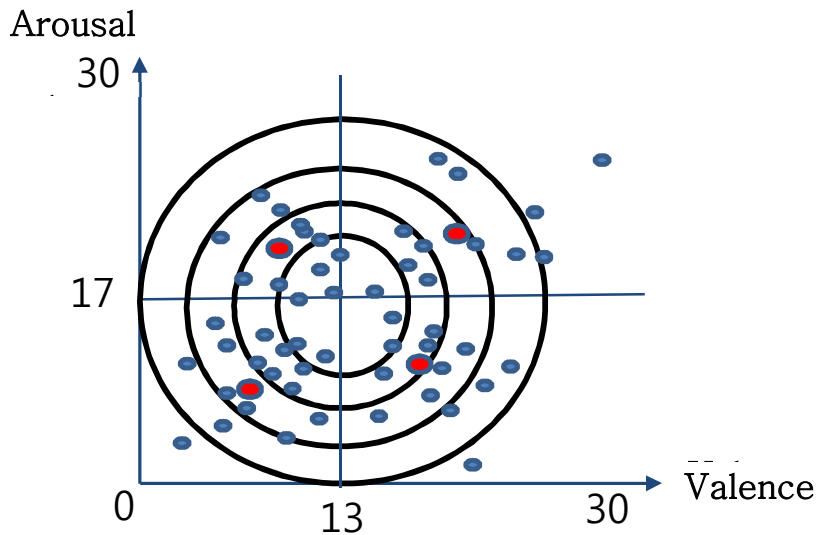


그림 19 K-means 클러스터링을 이용한 58명 데이터의 분류

표 1 생체 신호별 특징 데이터

Modality	Extracted Features
ECG(32)	Ten low frequency ([0-2.4]Hz) power spectral densities(PSDs), Four very slow response ([0-0.04]Hz) PSDs, IBI, HR and HRV statistics.
GSR(31)	Mean skin resistance and mean of derivative, mean differential for negative values only (mean decrease rate during decay time), Proportion of negative derivative samples, number of local minima in the GSR signal, average rising time of the GSR signal, spectral power in the [0-2.4] Hz band, zero crossing rate of skin conductance slow response ([0-0.2] Hz), zero crossing rate of skin conductance very slow response ([0-0.08] Hz), mean SCSR and SCVSR peak magnitude.
EEG(92)	Average of first derivative, proportion of negative differential samples, mean number of peaks, mean derivative of the inverse channel signal, average number of peaks in the inverse signal, delta frequency PSDs([0~4]Hz), theta frequency PSDs([4~7]Hz), alpha frequency PSDs([8~15]Hz), beta frequency PSDs([16~31]Hz), gamma frequency PSDs([31~]Hz).

(2) Stacked AutoEncoder

기계학습을 이용한 감정 분류기의 경우 감정을 대변하는 근원적인 특성을 추출하기 위해 생리학적 특징으로부터 더욱 추상화된 특징이 필요하다. 실제 상용화된 프로그램에서 AutoEncoder는 사전학습에서 이미 우수한 성능을 입증하였다[29]. AutoEncoder는 입력과 출력 레이어가 같은 피드 포워드 아키텍처로 이루어져 있으며, 학습 종료 후 가운데 히든 레이어의 출력값을 피쳐로써 활용할 수 있다 [29]. 학습은 다음의 그림 21과 같은 과정을 거치게 된다.

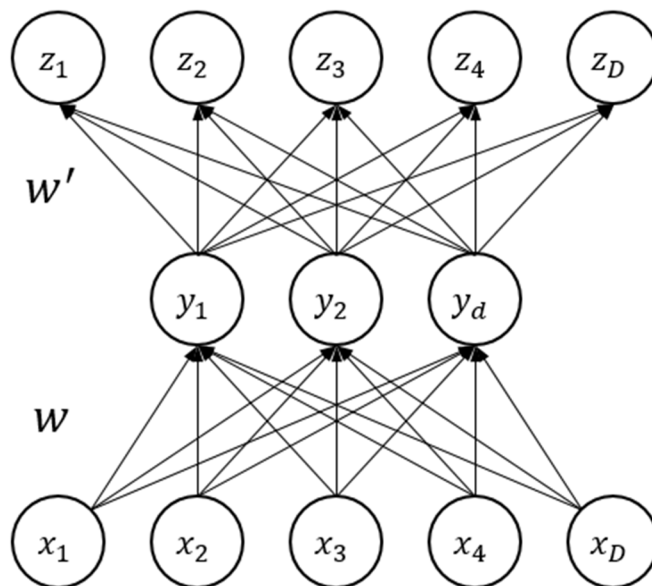


그림 20 AutoEncoder 구조

기본적으로 AutoEncoder는 히든 레이어의 추상화 특징이 잘 학습되는 것이 목표이므로, 입력값과 출력값과의 에러율이 최소가 될 때까지 학습 시킨다. 에러율 계산은 MSE(Mean Square Error) 또는 Cross-entropy 방식을 선택적으로 사용 가능하며 본 논문에서는 MSE를 사용하였다. Stacked AutoEncoder는 AutoEncoder를 계층적으로 쌓아 좀 더 깊은 추상화가 일어나도록 하는 모델이다[29]. 그림 22와 같이 첫 번째 계층의 AutoEncoder가 학습이 종료되면, Encoder 부분의 히든레이어가 입력이 되도록 하고 더 적은 노드 수를 가진 다음 히든레이어에 대한 추상화 학습을 진행한다. 이를 계속 진행하여 깊은 추상화 작업이 완료되면 Fine Tuning을 거쳐 학습을 완료한다.

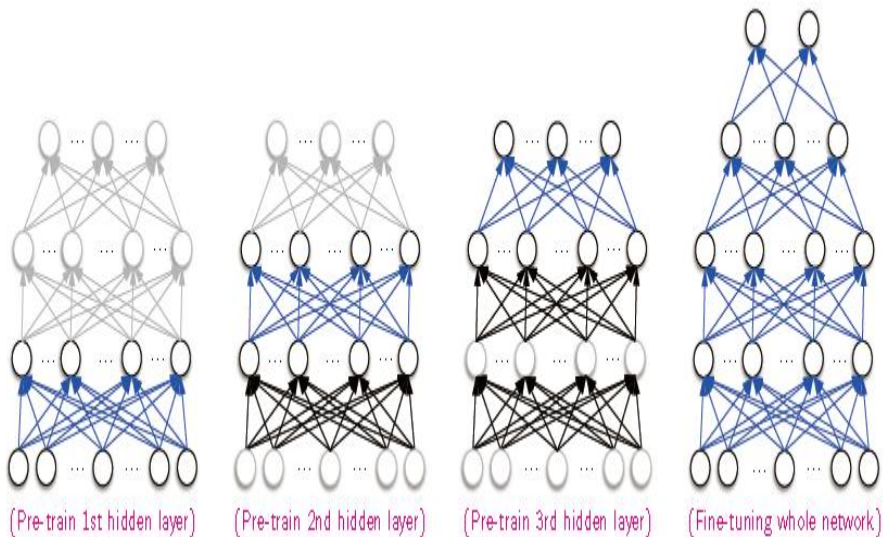


그림 21 Stacked AutoEncoder 학습 과정

(3) 딥러닝 모델 설계

Stacked AutoEncoder를 통해 추상화된 각 생체신호의 특징들은 차원이 낮은 다음 레이어에서 MLP(Multi Layer Perceptron) 형태로 결합하며, 이는 그림 22와 같다.

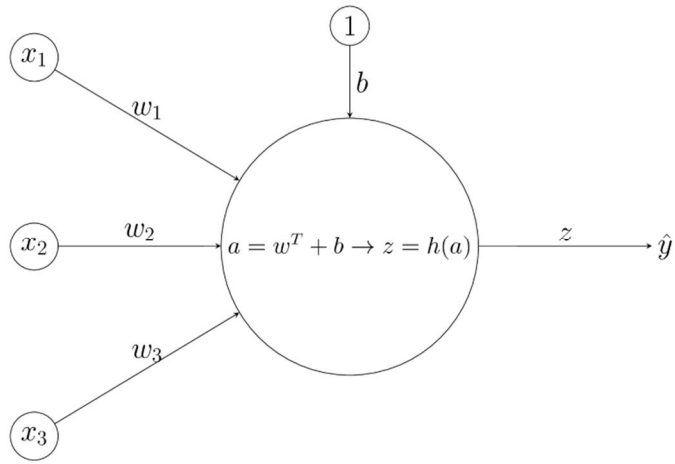


그림 22 퍼셉트론의 기본 모형

입력이 x 벡터이고 가중치를 w 벡터, 바이어스를 b 라고 하면, 활성화 함수 입력값 a 는 다음과 같다.

$$a = \sum_{i=1}^k w_i x_i + b = w^T x + b \quad (15)$$

이에 식 (15)를 이용해 활성화 함수 h 를 이용한 최종 출력값 z 은 식 (16)과 같다.

$$z = h(a) = h(w^T x + b) \quad (16)$$

이후 최종적으로 Softmax 분류기[30]를 통해 Valence-Arousal 면의 사분면으로 분류한다. Softmax 분류기는 SVM(Support Vector Machine)과 더불어 가장 많이 쓰이는 분류기 중 하나로, Logistic regression 분류기의 다중 클래스에 대한 일반화 개념이라고 볼 수 있다. 각 클래스의 점수를 입력받아 0~1 범위 정규화된 확률값들을 리턴한다. 즉, 분류하고 싶은 클래스의 수 만큼 출력 노드의 크기가 결정되며, 가장 큰 출력 값을 부여받은 클래스가 확률이 가장 높은 것으로 이용된다. Softmax 함수는 다음의 식과 같이 표현될 수 있다.

$$f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}} \quad (17)$$

식 (17)의 j 는 클래스를 의미하며, f_j 는 j 번째 클래스의 점수를 나타낸다. k 는 클래스의 전체 갯수를 의미하며, 전체 클래스의 합 $\sum_k e^{z_k}$ 은 항상 1로 고정이다. 그림 23은 제안하는 딥러닝 모델에 대한 블록 다이어그램을 보여주고, 표 2는 모델의 각 파트에 대한 상세 구조 정보를 나타낸다.

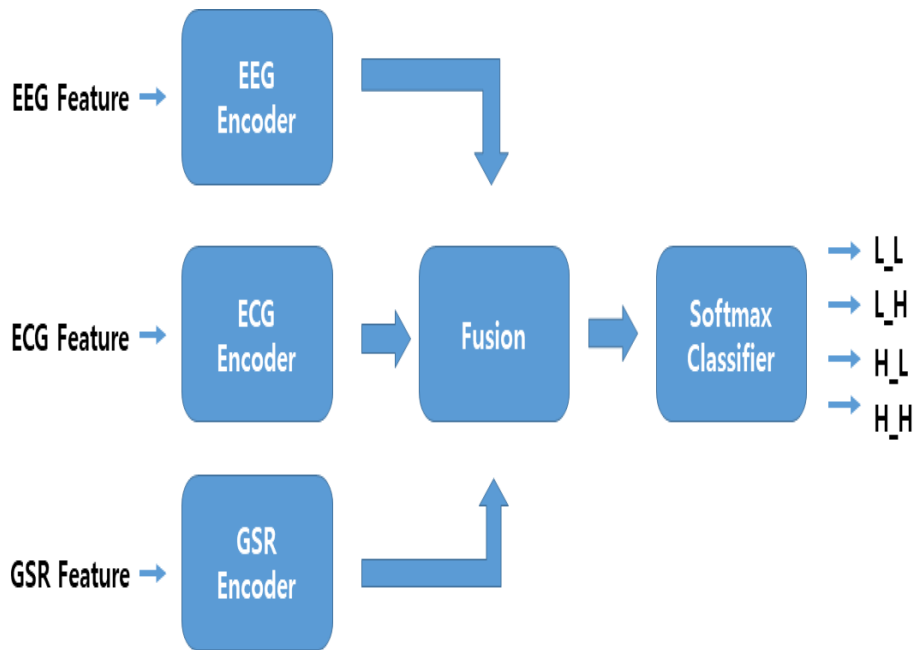


그림 23 딥러닝 모델 블록 다이어그램

표 2 제안한 딥러닝 모델 구조

Type	Input size	Output size	Hidden Layers
EEG Encoder	92	20	5
ECG Encoder	32	10	4
GSR Encoder	31	10	4
Fusion MLP	20	10	3

(4) 컴퓨팅 환경

본 논문에서 제안한 실험은 CPU가 Intel Core i7-7700 3.60GHz, 메모리 16GB, GPU의 경우 Nvidia GeForce GTX 1080Ti를 사용하였다.

2. 결과 및 분석

표 2는 ASCERTAIN에서 제공하는 데이터 신뢰성 지표와 제안한 데이터 신뢰성 지표를 적용했을 때의 감정 인식률을 비교한 것이다. L_L, L_H, H_L, H_H는 각각 Valence와 Arousal의 Low와 High를 나타낸 것이다. 이때 Valence-Arousal에 따른 가중치는 둘 다 미적용한 상태이다. 제안한 데이터 신뢰성 지표가 ASCERTAIN에서 제공된 지표를 사용했을 때보다 평균 약 10% 개선되었음을 알 수 있다.

표 3 데이터 신뢰성 평가 방법에 따른 감정 인식률 비교

	Proposed method	ASCERTAIN method
L_L	73%	68%
L_H	75%	65%
H_L	72%	61%
H_H	70%	64%

표 3은 데이터 신뢰성 평가 지표는 제안한 방법으로 고정하고 Valence-Arousal에 따른 가중치의 적용 여부에 따른 감정 인식률 성능을 비교한 것이다. Valence-Arousal에 따른 가중치를 적용한 경우가 그렇지 않은 경우보다 평균 약 3.5% 가량 개선되었음을 알 수 있다.

표 4 제안한 Valence-Arousal 가중치 적용여부에 따른 감정 인식을 비교

	Applied method	Unapplied method
L_L	76%	73%
L_H	80%	75%
H_L	75%	72%
H_H	73%	70%

표 4는 제안한 데이터 신뢰성에 따른 가중치 및 Valence-Arousal 가중치 모두 적용하였을 경우와 그렇지 않은 경우를 비교한 것이다. 평균 약 11.5%로 감정인식률이 개선되었음을 볼 수 있다.

표 5 제안한 방법을 모두 적용한 경우와 그렇지 않은 경우에 따른 감정 인식을 비교

	Applied method	Unapplied method
L_L	76%	68%
L_H	80%	65%
H_L	75%	61%
H_H	73%	64%

V. 결론 및 향후 연구 방향

본 논문에서는 Valence와 Arousal로 라벨링 된 멀티모달 생체 신호 데이터에 대해 데이터의 신뢰성을 검증하기 위한 웨이블릿 기반의 노이즈 임계값 설정 알고리즘 및 감정 평가 시 데이터 신뢰도와 Valence-Arousal 값에 따른 가중치를 부여하여 감정 인식률을 향상하는 방법을 제안하였다.

신호의 성분 분해시 이미 우수한 성능을 검증한 웨이블릿 변환을 이용해 신호의 웨이블릿 성분을 추출 후, 해당 성분의 통계적 데이터인 왜도와 첨도를 구하여 헵펠 식별자를 통해 계산된 임계값으로 노이즈를 판별하였다. 이후, 원신호에 대한 노이즈 비율을 고려하여 데이터의 신뢰성을 평가하고 이를 가중치로 환산하였다.

더불어 감정 데이터 분류 시 Valence-Arousal 평면을 이용한 이산적인 감정분류법의 단점을 개선하고자 평면 중앙값과의 유클리디언 거리를 가중치로 환산하였다.

감정 인식률에 대한 종합 평가 시 각 요소에 대한 개별적인 성능 및 모두 적용시 성능을 검증하고자 ASCERTAIN 데이터셋을 활용하였고, 감정 인식률 개선 정도를 통해 제안된 알고리즘의 성능이 우수한 것을 확인하였다.

추후 웨이블릿 변환을 이용한 노이즈 검출 방법에서 생체 신호별 적정 모 웨이블릿 선택에 따른 노이즈 검출 성능 연구가 필요할 것으로 생각되며 최소 노이즈 비율과 최대 노이즈 비율을 반복적 실

험을 통해서 찾는 것이 아닌 자동화 방법에 대한 연구 역시 병행되어야 한다고 본다. 또한 딥러닝 모델에서 입력 데이터의 특징 추상화 시 적절한 추상화 정도에 따른 모델링 구조에 대한 고찰도 필요할 것으로 보인다.

참고문헌

- [1] M. Soleymani, J. Lichtenauer, T. Pun, M. Pantic, “A multi-modal database for affect recognition and implicit tagging”, *IEEE Trans. Affect. Comput.* Vol. 3 pp. 42–55, 2012.
- [2] G. Fanelli, J. Gall, H. Romsdorfer, T. Weise, L. Van Gool, “A 3-D audio-visual corpus of affective communication”, *IEEE Trans. on Multimedia* Vol. 12 pp. 591–598. 2010.
- [3] Z. Yin, J. Zhang, “Operator functional state classification using least-square support vector machine based recursive feature elimination technique”, *Comput. Methods Prog. Biomed.* Vol. 113, pp. 101–115. 2014.
- [4] J. Kim, E. Andre, “Emotion-specific dichotomous classification and feature-level fusion of multichannel biosignals for automatic emotion recognition”, in: *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems* 2008.
- [5] H. Lee, A. Shackman, D. Jackson, R. Davidson, “Test-retest reliability of voluntary emotion regulation”, *Psychophysiology* Vol. 46 pp. 874–879. 2009.

- [6] O. AlZoubi, S.K. D'Mello, R.A. Calvo, "Detecting naturalistic expressions of non- basic affect using physiological signals", *IEEE Trans. Affect. Comput.* Vol. 3 pp. 298-310. 2012.
- [7] L. Brown, B. Grundlehner, J. Penders, "Towards wireless emotional valence de- tection from EEG", *in: Proceedings of Engineering in Medicine and Biology Society, EMBC*, 2011.
- [8] R.J. Davidson, "Affective neuroscience and psychophysiology: toward a synthesis", *Psychophysiology* Vol. 40, pp. 655-665, 2003.
- [9] Z. Yin, J. Zhang, "Identification of temporal variations in mental workload using locally-linear-embedding-based EEG feature reduction and support-vec tor-machine-based clustering and classification technique", *Comput. Methods Prog. Biomed.* Vol. 115, pp. 119-134. 2014.
- [10] 정기삼, 이병채, 하은호, 김동윤, 김동선 "감성 완구 제작을 위한 감 성분석 시스템 개발에 관한 연구", *한국감성과학회:학술대회지*, 한국감성과학회 01 추계 학술 대회 논문집, pp.239-244, 2001.
- [11] R.W. Levenson, P. Ekma, and W.V. Friesen, "Voluntary facial action generates emotion-specific autonomic nervous systemactivity", *Psychophysiology*, Vol.27, pp.363-384, 1990.

- [12] Lang PJ, Bradley, M.M&Cuthbert,B, NNIMH, “Center for the Study of Emotionand Attention”, *international Affective Picture System(IAPS) Technical Manual and Affective Ratings*, 1995.
- [13] Senthil, R. Arumuganathan, K. Sivakumar, and C. Vimal, “Removal of ocular artifacts in the EEG through wavelet transform without using an EOG reference channel”, *Int.J. Open Problems Compt. Math.*, Vol.1, No.3, 2008.
- [14] Mrs V Baby Deepa ,Dr P Thangaraj, Dr S Chitra, “Investigating the performance improvement by sampling techniques in EEG data”, *International Journal on Computer Science and Engineering*, Vol. 2, pp. 2025–2028, 2010.
- [15] J. Hernandez et al. “Call center stress recognition with person-specific models.” *Proceedings of the Affective Computing and Intelligent Interaction, Memphis*, 2011.
- [16] M. Poh et al. “A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity”. *IEEE Trans. Biomed. Eng.*, 2010.
- [17] M. Poh et al. “Continuous monitoring of electrodermal activity during epileptic seizures using a wearable sensor”. *Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2010.

- [18] J. Christensen , J. Estepp , G. Wilson , C. Russell , “The effects of day-to-day variability of physiological data on operator functional state classification”, *NeuroImage*, Vol. 59 pp. 57–63, 2012.
- [19] F. Laurent , M. Valderrama , M. Besserve , M. Guillard , J.P. Lachaux , J. Martinerie , G. Florence, “Multimodal information improves the rapid detection of mental fatigue”, *Biomed. Signal. Process. Control*. Vol. 8, pp. 400–408, 2013.
- [20] Y. Shin , S. Lee , M. Ahn , H. Cho , S.C. Jun , H.-N. Lee, “Simple adaptive sparse representation based classification schemes for EEG based brain-computer interface applications”, *Comput. Biol. Med.* Vol. 66, pp. 29–38, 2015.
- [21] D. Iacoviello , A. Petracca , M. Spezialetti , G. Placidi , “A real-time classification algorithm for EEG-based BCI driven by self-induced emotions”, *Comput. Methods Prog. Biomed.* Vol. 122, pp. 293–303, 2015.
- [22] R.M. Mehmood , H.J. Lee , “A novel feature extraction method based on late positive potential for emotion recognition in human brain signal patterns”, *Comput. Electr. Eng.* Vol. 53, pp. 444–457, 2015.
- [23] S. Jirayucharoensak , S. Pan-Ngum , P. Israsena , “EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation”, *Scientific World J.* Vol. 1, pp. 1–10, 2014.

- [24] R. R. Coifman and D. L. “Donoho. in Wavelets and statistics”. Vol. 103, pp. 281–299, 1995.
- [25] D. P. Subha, P. K. Joseph, R. Acharya, C. M. Lim, “EEG Signal Analysis: A Survey”, *Journal of Medical Systems*, Vol. 34, No. 2, pp 195–212, 2010.
- [26] Chmelka and J. Kozumplik, “Wavelet-based Wiener Filter for electrocardiogram signal denoising,” *Comput. Cardiol*, vol. 32, pp. 771–774, 2005.
- [27] Maarten Jansen,”wavelet Thresholding and Noise Reduction”, *Katholieke Universiteit Leuven Faculteit Toegepaste Wetenschappen Arenbergkasteel*, 2000.
- [28] R. Subramanian et al., “ASCERTAIN: Emotion and Personality Recognition using Commercial Sensors,” *IEEE Transactions on Affective Computing*, pp. 1–14, 2016.
- [29] P. Baldi, “Boolean AutoEncoder,” pp. 37–50, 2012.
- [30] R. Memisevic and C. Zach, “Gated softmax classification,” *Advances in Neural*, pp. 1–9, 2010.

Abstract

Data Noise and Reliability Study for Improvement of Multi-Modal Emotion Recognition Rate Based on Deep Learning

Junhan Yoon

Dept. of Electronics and Computer Engineering

The Graduate School of Seokyeong University

Recently, a methodology for analyzing a variety of bio-signals in an emotion recognition study using a deep learning model is being studied. At this time, the reliability and classification method according to the type of learning data are the most influential to the reliability of the emotion classification. In the case of bio-signals, reliability of data is determined according to the noise ratio. Therefore, reliability of the noise detection method increases as the method of detecting the noise increases, and the emotion can be correctly classified according to the method of defining the emotion. In this paper, we propose a fixed wavelet-based noise threshold setting algorithm for verifying the reliability of bio-signal data for multi-modal bio-signals based on Valence and Arousal standards, a method for setting reliability according to Valence- We propose a method to improve emotion recognition rate by reflecting reliability and Valence-Arousal

reliability. The wavelet component of the signal is extracted using the wavelet transform, the degree of kurtosis of the component is obtained, the noise is detected at the threshold value calculated through the hamburg identifier, the reliability of the data is evaluated in consideration of the noise ratio of the original signal Converted to a weight. In addition, when classifying the emotional data, the Euclidean distance from the median of each quadrant of the Valence-Arousal plane is converted to a weight, and the two factors are reflected when adjusting the amount of learning data. The performance of the proposed algorithm is verified through the degree of improvement of emotion recognition rate using ASCERTAIN dataset.