creative commons
COMMONS DEED

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

dCollection

조선대학교
CHOSUN UNIVERSITY

2022년 2월
석사학위논문

심전도를 이용한 LSTM과 2D-CNN의 앙상블기반 개인식별

조 선 대 학 교 대 학 원

전자공학과

이 진 아

조선대학교
CHOSUN UNIVERSITY

심전도를 이용한 LSTM과 2D-CNN의 앙상블기반 개인식별

Personal Identification Based on an Ensemble of LSTM and 2D-CNN from Electrocardiogram

2022년 2월 25일

조 선 대 학 교 대 학 원

전자공학과

이 진 아

조선대학교
CHOSUN UNIVERSITY

심전도를 이용한 LSTM과 2D-CNN의 앙상블기반 개인식별

지도교수 곽 근 창

이 논문을 석사학위신청 논문으로 제출함.

2021년 10월

조 선 대 학 교 대 학 원

전자공학과

이 진 아

조선대학교
CHOSUN UNIVERSITY

이진아의 공학석사학위논문을 인준함

위원장 조 선 대 학 교 교수 염 홍 기 인

위 원 조 선 대 학 교 교수 신 주 현 인

위 원 조 선 대 학 교 교수 곽 근 창 인

2021년 12월

조선대학교 대학원

조선대학교
CHOSUN UNIVERSITY

목 차

제1장 서론 ····· 1
제1절 연구 배경 ····· 1
제2절 연구 목적 ····· 3
제3절 연구 내용 ····· 5
제2장 관련 연구 ····· 6
제1절 심전도 신호의 전처리 방법 ····· 6
제2절 심전도기반 딥러닝을 이용한 개인식별 ····· 7
제3장 LSTM과 2D-CNN의 앙상블기반 개인식별 ····· 10
제1절 1D기반 LSTM ····· 10
제2절 시간-주파수 변환기반 2D-CNN ····· 15
제3절 제안된 앙상블기반 개인식별 ····· 31
제4장 실험 및 결과분석 ····· 38
제1절 심전도 데이터베이스 ····· 38
제2절 실험 및 결과분석 ····· 41
제5장 결론 ····· 51
참고문헌 ····· 53

조선대학교
CHOSUN UNIVERSITY

표 목차

표 4-1 심전도 데이터베이스

표 4.2 2D-CNN의 개인식별 정확도와 LSTM과 2D-CNN기반 앙상블한 개인식별 정확도 (CU-ECG DB)

표 4.3 2D-CNN의 개인식별 정확도와 LSTM과 2D-CNN기반 앙상블한 개인식별 정확도 (PTB DB)

표 4.4 기존 딥러닝기반 개인식별의 성능 비교 (CU-ECG DB)

표 4.5 기존 딥러닝기반 개인식별의 성능 비교 (PTB DB)

조선대학교
CHOSUN UNIVERSITY

도 목차

그림 2.1 평균 이동 필터 적용된 심전도 신호
그림 3.1 RNN의 기본 구조
그림 3.2 RNN의 그래디언트 소멸 문제점
그림 3.3 LSTM의 구조
그림 3.4 한 개의 LSTM 계층
그림 3.5 두 개의 LSTM 계층
그림 3.6 푸리에 변환과 STFT의 개념적인 차이
그림 3.7 STFT의 윈도우 길이(128,256,512)에 따른 분해능의 차이
그림 3.8 푸리에 변환 계산 과정 분석
그림 3.9 웨이블릿의 변환 영역
그림 3.10 T^2에 따른 모스 웨이블릿
그림 3.11 FSST에 의한 시간-주파수 표현
그림 3.12 WSST에 의한 시간-주파수 표현
그림 3.13 CNN의 기본 구조
그림 3.14 합성곱 연산 과정
그림 3.15 ReLU 함수
그림 3.16 최대 풀링 연산 과정
그림 3.17 GoogleNet의 구조
그림 3.18 인셉션 모듈
그림 3.19 보조 분류기
그림 3.20 VGG-19의 구조
그림 3.21 Residual Connection
그림 3.22 ResNet-101의 구조
그림 3.23 앙상블 모델 기반의 분류 과정
그림 3.24 제안된 앙상블기반 개인식별 흐름도
그림 3.25 전처리 과정을 통한 심전도 신호

조선대학교
CHOSUN UNIVERSITY

그림 3.26 59번 피험자의 심전도 신호
그림 3.27 검출된 R-피크 및 한 주기로 분할된 신호
그림 3.28 제안된 앙상블기반 개인식별 순서도
그림 4.1 CU-ECG 신호 측정을 위한 환경
그림 4.2 CU-ECG 데이터베이스의 심전도 신호의 예시
그림 4.3 CU-ECG의 LSTM기반 개인식별 정확도(잡음 제거 전)
그림 4.4 CU-ECG의 LSTM기반 개인식별 정확도(잡음 제거 후)
그림 4.5 PTB의 LSTM기반 개인식별 정확도
그림 4.6 CU-ECG 데이터베이스 신호의 피험자 시간-주파수 표현
그림 4.7 PTB 데이터베이스 신호의 피험자 시간-주파수 표현

조선대학교
CHOSUN UNIVERSITY

ABSTRACT

Personal Identification Based on an Ensemble of LSTM and from Electrocardiogram

Lee, Jin-A
Advisor : Prof. Kwak, Keun Chang, Ph. D.
Dept. of Electronic Engineering,
Graduate School of Chosun University

Recently, with the rapid development of artificial intelligence, technology has been applied to various fields to improve the quality of life and work, not only in human daily life, but also in medical care, security, and finance. However, there are cases where humans are threatened with security problems by exploiting this, and technologies that can authenticate only various individuals are expanding in order to defend against this. Previously, personal information such as passwords and OTP was protected, but there is a problem of loss or theft, so personal biometric information can be used to supplement important personal information. Biometrics is a physical and behavioral feature, and personal authentication technology using biometric information requires characteristics of universality, uniqueness, permanence, collectability, accuracy, acceptability, and circumvention.

조선대학교
CHOSUN UNIVERSITY

In this paper, an individual was identified by using an electrocardiogram(ECG) signal with the characteristics of excellent identification. A low-pass filter, a high-pass filter, and an average shift filter were applied to the signal as a preprocessing process, and the baseline was adjusted to zero. Also, the signal was divided at regular intervals based on the R-peak. The performance of the Long Short-Term Memory(LSTM) neural network consisting of one LSTM layer and two LSTM layers was compared and analyzed. In addition, the ECG signal was converted into the time-frequency domain for signal analysis, and expressed as two-dimensional images of Short-Time Fourier Transform(STFT), Scalogram, Fourier Synchrosqueezed Transform(FSST), and Wavelet Synchrosqueezed Transform(WSST). Deep models od CNN were used by GoogleNet, VGG-19, and ResNet-101. In this paper, the used database consists of the Chosun University(CU)-ECG database and the Physikalisch-Technische Bundesanstalt(PTB) database. Finally, the experimental results showed high performance from a minimum of 1.06% to a maximum of 3.75% and from a minimum of 0.8% to a maximum of 3.38% for CU-ECG databases and PTB database, respectively.

조선대학교
CHOSUN UNIVERSITY

제1장 서론

제1절 연구 배경

인공지능이 급속하게 발전되면서 좋은 성능으로 이루어 많은 주목을 받고 있다. 인공지능의 핵심적인 기술인 딥러닝이 등장하면서 인공지능 기술의 분야는 확대되고 다양한 분야에 적용함으로써 인간의 일상생활에 편리함을 가져다주고 있으며 의료, 농업, 보안, 금융, 자율 자동차 등에 적용되면서 삶과 일의 질을 향상시키고 있다. 기술이 계속해서 발전하고 확산됨에 따라 좋은 영향을 미치는 반면에 이를 악용하여 사이버보안으로 지능형 공격이 증가하면서 인간들을 위협하는 경우도 발생하고 있다[1]. 이러한 위협을 방어하기 위해서 다양한 개인인증 기술들이 연구되고 있다.

기존에는 패스워드, OTP 등을 사용하여 개인의 보안을 유지했지만, 분실 또는 도용의 위험이 있어서 최근에는 개인의 생체정보를 사용하여 개인의 중요한 정보들을 보안 하는 기술이 대두되고 있다. 이는 개인 정보를 편리하고 안전하게 유지할 수 있고 사용자의 신원을 확인하기 위해 사용되며 음성, 성별, 얼굴 또는 행동 등 생체정보를 활용한 개인인증 기술이 활발하게 발전되고 사용되고 있다[2-4]. 생체인식을 하기 위해서는 7가지의 특성이 필요하다. 기본적인 특성으로는 첫 번째 모든 사람이 갖고 있는지의 보편성, 두 번째는 동일한 특징이 아닌 고유의 특징을 가지고 있는지의 유일성, 세 번째는 시간에 따라 변하지도 않고 변화시킬 수 없는지의 영구성, 네 번째는 센서를 사용하여 생체정보를 쉽고 간단하게 획득할 수 있는 획득성이다. 생체정보를 사용하기 위해 신뢰하고 사용하기 위한 특성으로 다섯 번째는 정보를 받아들이고 처리하는 속도와 정확도 등이 포함된 정확성, 여섯 번째는 생체 정보의 측정 거부감이 들지 않을 정도의 접근성, 마지막 일곱 번째는 부정한 사용에도 안전한지의 기만성이 있다[4].

생체인식은 얼굴, 지문, 홍채 등과 같이 신체적인 특성과 음성, 걸음걸이, 심전도, 뇌파 등 행동적인 특징이 있다. 각 특성의 장단점으로 얼굴 인식은 얼굴의 특징들을 분석하여 인식하는 방법으로 쉽고 빠르게 인식이 가능한 장점이 있지만, 표정의 변화 및 조명에 의해 인식이 어려운 경우가 있다. 지문인식의 경우에는 적용 센서를 통해 지문 영상을 획득하여 인식하는 기술로, 보편적으로 사용되고 있

조선대학교
CHOSUN UNIVERSITY

다. 지문인식 기술을 간편하지만, 손의 상처와 이물질에 민감하고 위•변조가 가능하다는 단점이 있다. 하지만 얼굴인식과 지문인식은 현재 코로나 팬데믹으로 인해 마스크 착용이 필수가 되면서 공공장소와 같이 사람들이 많은 장소에서는 접촉식 보안 방법이 어려워지고 있다. 또한 행동적인 특성 중 음성인식은 음성으로부터 추출한 특성을 이용하는 기술로, 편리하고 안정성이 있지만 녹음된 파일과 소음에 영향을 받아 인식이 어려운 경우가 있다. 걸음걸이 인식은 사용자의 걸음걸이의 특성을 분석하여 인식하는 기술로, 다른 인식 기술보다 주변 환경의 영향을 더 많이 받는 단점이 있다. 이와 같이 신체 외부에서 표출되는 생체 정보들은 사용자 인식 오류, 위•변조 위험 등으로 개인적인 피해를 겪어 큰 문제가 되고 있다. 이러한 문제를 해결하기 위해서 신체 외부로 표출되는 정보가 아닌 신체 내부의 신호인 뇌전도(EEG, electroencephalogram), 근전도(EMG, electromyogram), 심전도(ECG, electrocardiogram) 등을 사용하여 개인을 식별하는 연구가 활발히 진행되고 있다.

조선대학교
CHOSUN UNIVERSITY

제2절 연구 목적

신체 내부의 신호 중 뇌전도는 두피에서 뇌의 전기적인 신호를 측정한 것으로 뇌-컴퓨터 인터페이스(BCI, brain-computer interface), 질병 진단 등에 활용되고 있다. 그러나 센서를 머리에 착용하는 형태로 신호를 측정하여 거부감이 있으며, 두개골을 지나면서 신호가 왜곡되는 문제가 발생한다. 근전도는 근육에서 발생하는 신호를 기록한 것으로 모션 인식, 질병 진단 및 재활 치료 등에서 활용되고 있다. 그러나 각 모션마다 신호를 설정해야 하고 정확한 근육 위치에 센서를 부착해야한다는 불편함이 있다.

심전도는 심장 박동의 전기적인 신호를 미세 전류에 의해 신호로 나타낸 것으로 P파(심방의 탈분극), QRS 복합파(심실의 탈분극), T파(심실의 재분극)로 구성되어 있다. 사람마다 심장의 위치, 크기, 성별 및 나이에 따라 다르게 나타나는 개인의 고유한 신호이다. 측정 방법은 표준 12 리드 유도법으로 양쪽 손목에서 신호 측정을 간편히 할 수 있으며, 질병을 진단하는 것뿐만 아니라 개인을 식별하는 것이 가능하다. 또한 심전도는 신체 내부에서 발생하는 신호이기 때문에 위•변조가 어려우며 환경의 변화에 영향을 받지 않는다는 장점이 있고, 살아 있는 모든 사람이 가지고 있는 신호이다[5]. 이처럼 뇌전도와 근전도, 심전도와 같이 생리학적 특징을 기반으로 하는 식별은 개인만의 신호로 보안 방법과 개인을 식별하기 위해 많은 장점들과 연구들이 활발히 진행되고 있지만 본 논문에서는 생체인식을 하기 위한 7가지의 특성을 모두 포함한 심전도를 사용하여 개인식별을 진행한다.

기존에는 심전도 신호의 특징을 추출하기 위해 Handcrafted 특징방법을 사용하여 심전도 신호의 진폭, 시간 간격(ST간격, PQ간격, QS간격), 피크(P, Q, R, S, T 최소 피크와 최대 피크), 각도 등에서 심전도의 특성점기반으로 특징을 추출하는 방법이다[6][7]. 이 방법은 머신러닝에 속하는 것으로 직접 특징을 추출하고 이를 기반으로 학습하는 방식으로 심전도 신호의 피크를 감지하는 것이 신호의 가변성이 높기 때문에 문제가 발생한다[8]. 또한 Handcrafted 특징방법은 잡음 제거 또는 특징 추출하는 과정에서 성능이 저하되는 문제가 있기 때문에 최근 딥러닝이 등장하면서 Handcrafted 특징방법의 문제점들을 해결하여 시계열의 성능을 향상시키는 LSTM과 이미지의 분류 성능이 뛰어난 2D-CNN을 사용한다. 딥러닝은 학습을 통해 특징을 추출하기 때문에 Handcrafted 특징방법과 같이 직접 추출할 필요 없

조선대학교
CHOSUN UNIVERSITY

다. Non-Handcrafted Fiducial 특징방법은 심전도 신호의 전반적인 형태를 기반으로 특성 점을 사용하지 않기 때문에 신호 분할을 위해 R-피크 검출이 필요하다.

본 논문에서는 심전도 신호를 Non-Handcrafted Fiducial 특징방법을 사용하기 위해 잡음 제거된 신호를 R-피크를 검출하여 일정한 간격으로 분할을 진행하였다. 분할된 신호를 이용하여 딥러닝 모델인 LSTM과 2D-CNN을 통해 시계열과 이미지의 성능을 향상시킬 수 있다. 하지만 단일 모델만을 사용할 경우보다 두 모델을 결합하는 제안된 앙상블 방법을 사용하여 더 높은 성능을 기대할 수 있다.

조선대학교
CHOSUN UNIVERSITY

제3절 연구 내용

본 논문에서는 심전도를 이용한 장단기 메모리와 합성곱 신경망의 앙상블기반으로 개인식별한다. 데이터베이스는 조선대학교에서 구축한 CU-ECG 데이터베이스는 100명의 심전도 신호를 포함하고 있으며, 데이터는 총 16,930개 중 학습 데이터로 13,546개, 검증 데이터로 3,384개를 사용하였다. 두 번째로 PTB 공개 데이터베이스를 사용한다. PTB 데이터베이스는 다양한 심장 질환을 가지고 있는 피험자와 정상 심전도 신호를 가진 피험자의 신호가 포함되어 있으며, 290명 중 211명의 데이터를 사용하였다. 총 25,320개 중 학습 데이터로 12,660개, 검증 데이터로 12,660개를 사용하였다. 원 신호는 신호를 측정할 때 다양한 잡음들을 포함되기 때문에 잡음 제거와 기저선 변동된 신호를 영점으로 조정 후 Non-Handcrafted Fiducial 특징방법으로 신호의 R-피크를 검출하여 일정한 간격으로 분할하였다. 1차원 시계열 신호를 분류하기 위해 한 개의 LSTM 계층과 두 개의 LSTM 계층을 비교 분석하여 사용한 성능을 확인하고, 더 높은 정확도를 나타내는 두 개의 LSTM 계층을 사용하였다. 계층을 쌓을수록 분류 성능 정확도를 향상시킬 수 있지만, 구조가 복잡해질 수 있는 문제점이 있다. 본 논문에서는 앙상블을 제안하기 때문에 복잡한 구조를 회피하기 위해 LSTM의 계층을 더 쌓지 않고 두 개의 계층을 사용하였다. 시간-주파수 표현으로 Short-Time Fourier Transform(STFT), Scalogram, Fourier Synchrosqueezed Transform(FSST), Wavelet Synchrosqueezed Transform(WSST)를 사용하여 이미지로 나타낸다. 또한 이미지 분류 성능이 우수한 합성곱 신경망의 전이학습 모델인 GoogleNet, VGG-19, ResNet-101을 사용하여 성능을 확인하여 LSTM과 2D-CNN의 각 모델들의 성능을 앙상블 방법을 통해 개인식별 성능이 향상된 것을 확인한다.

본 논문의 구성은 다음과 같다. 1장은 서론으로 시작하며, 2장에서는 관련 연구를 나타내고, 3장에서는 LSTM과 2D-CNN의 앙상블기반 개인식별, 4장에서는 실험 및 결과분석, 5장에서는 결론 순으로 마무리한다.

조선대학교
CHOSUN UNIVERSITY

제2장 관련 연구

제1절 심전도 신호의 전처리 방법

심전도 신호는 미세한 전류를 측정하는 생체 신호로 신호를 측정할 때 다양한 잡음들이 심전도 신호보다 크게 나타나는 경우가 있다. 심전도 신호는 0.05Hz부터 100Hz까지의 주파수 영역으로, 그로 인해 발생하는 잡음은 60Hz의 전력선 잡음, 근전도, 전극 접촉으로부터의 잡음 등에 의해 발생하는 고주파 잡음과 호흡, 움직임에 의한 기저선 변동 같은 저주파 잡음과 같이 다양한 유형의 잡음이 있다. 이로 인해 정확한 판단이나 진단이 어렵기 때문에 왜곡된 신호들을 제거하고 심전도 형태를 유지하기 위해 전처리 과정이 필요하다. 또한 심박수의 가변성의 영향을 줄이기 위해 진폭 정규화 과정이 필요하다. 심전도 신호의 잡음 제거는 차단 주파수 영역의 잡음들을 제거하기 위해 사용되고[9], 주파수 필터, 평균 이동 필터, 노치 필터, 칼만 필터, 베이지안 필터[10] 등이 있다. 주로 사용되는 저역 통과 필터와 고역 통과 필터는 차단 주파수 영역에서의 이하 또는 이상의 신호들을 통과시키는 필터로, 저주파 영역에서의 잡음 제거를 위해 높은 차단 주파수의 고역 통과 필터를 사용할 경우 심장 질환의 정보를 포함한 심전도 신호가 왜곡된 결과를 초래할 수 있다. 노치 필터는 특정 범위의 주파수 영역을 제거하는 필터로, 주로 신호 측정 중에 발생하는 60Hz의 전력선 잡음을 제거하기 위해 사용된다. 또한 미세한 잡음을 제거하기 위해 평균 이동 필터가 사용되며, 그림 2.1는 평균 이동 필터가 적용된 신호를 보여준다.

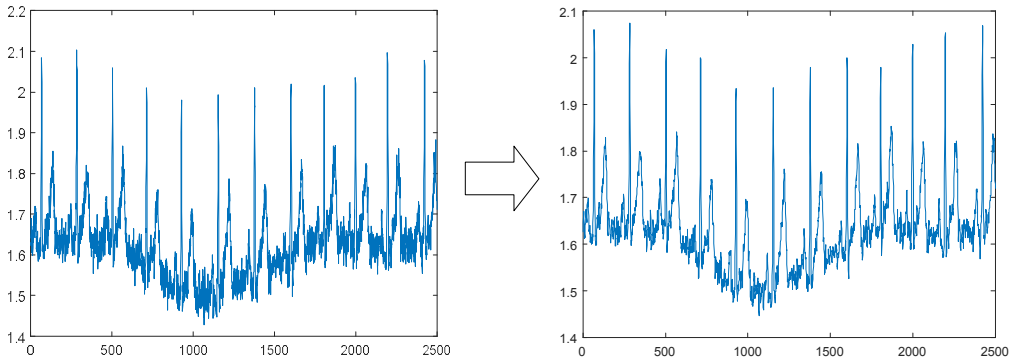


그림 2.1 평균 이동 필터 적용된 심전도 신호

조선대학교
CHOSUN UNIVERSITY

제2절 심전도기반 딥러닝을 이용한 개인식별

최근 다양한 분야에서 딥러닝을 사용하여 개인식별 연구가 활발하게 진행되고 있다. 심전도를 이용하여 개인 사용자 인식 관련 연구는 Biel[11]에 의해 최초로 연구가 진행되었다. 기존에는 머신러닝에 속하는 Handcrafted 특징방법을 사용하였다. Israel[12]은 심전도를 이용해 개인을 식별하기 위해 P, R, S 피크를 기반으로 최대 로컬 값을 찾고 경사를 추적하여 최소 곡률 반경의 위치를 찾아 시간 간격 특징으로 총 15개의 특징을 추출했다. Jahiruzzaman[13]는 웨이블릿 변환은 이미지 압축, 패턴인식 등과 같은 신호처리에 사용되는 시간 영역의 표현으로 특징 추출을 위해 Continuous Wavelet Transform(CWT)를 사용한다. MIT-BIH 부정맥 데이터베이스를 사용하여 ECG 신호에 CWT를 통해 암호화 기술을 적용하여 ID 매칭(Matching) 사용하여 개인 식별을 진행한다. Zhao[14]은 앙상블 경험적 모드 분해(EEMD, Ensemble Empirical Mode Decomposition)과 웰치 스펙트럼 분석을 사용하여 기반으로 IMFS(Intrinsic Mode Function Spectral) 특징을 추출하여 신호의 형태학적 정보, 스펙트럼 정보를 추출하기 위해 사용됐다.

하지만 Handcrafted 특징방법은 심전도 신호의 특징을 사람이 직접 추출하여 학습시키기 때문에 개인식별 성능이 저하되는 문제가 발생한다. 이와 반대로 딥러닝은 Handcrafted 특징방법의 문제점을 해결하고, 자동으로 특징을 추출하여 시계열과 영상, 텍스트 분야에서 분류, 회귀 널리 사용되고 있기 때문에 LSTM과 CNN과 같은 우수한 성능 결과를 보여주는 모델을 사용하여 개인식별 연구가 진행되고 있다. Labati[15]는 PTB 데이터베이스를 사용하여 CNN기반 Deep-ECG를 제안하여 전처리, CNN 특징추출, 식별 순으로 진행됐다. 전처리 과정에서 노치 필터, IIR(Infinite Impulse Response) 필터, 3차 고역 통과가 사용되고 CNN은 6개의 컨볼루션 계층, 드롭아웃 계층 1개, 완전 연결 계층과 소프트맥스 계층으로 구성되어 개인식별을 진행했다. Abdeldayem[16]은 심전도를 이용해 개인 식별하기 위해 다섯 가지의 접근 방식을 제안했다. 첫 번째는 심전도 신호의 주기적인 특징을 사용하여 신호를 구별하는 방법이다. 두 번째는 심전도 신호에 고정된 기간인 각 세그먼트의 지속 시간이 일정한 블라인드(blind) 분할하여 사용하는 방법으로 복잡한 계산을 낮추고 성능을 향상시킬 수 있다. 이 제안은 첫 번째의 심전도 주기적인 성질을 포함해야 한다. 세 번째는 잡음은 순환 정지하지 않기 때문에 잡음 제거하는 단계를 사용하지 않는 것이다. 네 번째는 신호의 주파수 영역인 파워 스펙

조선대학교
CHOSUN UNIVERSITY

트럼 밀도로 변환하여 2D-CNN을 사용하였다. 마지막 다섯 번째는 사용한 8개의 공개 데이터베이스를 하나의 데이터로 결합하여 사용하였다. Ciocoiu[17]는 심전도 신호를 전처리 과정으로 대역 통과 필터링을 통해 잡음을 제거하고 R피크 중심으로 일정한 시간으로 주기를 분할한다. 4가지 유형의 공간 표현인 CWT, Gramain Angular Field(GAF), 위상 공간 궤적(Phase-Space Trajectories), 반복 플롯(Recurrence Plots)을 통해 이미지로 변환한 후 3개의 합성곱 계층(Convolution layer), 활성화 함수 ReLU, 최대 풀링 계층과 완전 연결 계층과 소프트맥스 계층으로 구성된 CNN을 적용하여 심전도 기반 생체인식을 하여 정확도와 EER(Equal error rate)을 비교 분석을 하였다. Y. H. Byeon[18]은 다양한 시간-주파수 표현으로 전이학습의 CNN 모델을 사용하여 심전도 생체 인식 성능을 확인하였다. 시간-주파수 표현 방법으로 MFCC, spectrogram, log spectrogram, mel spectrogram, scalogram을 사용하여 4개의 전이학습 모델을 사용했다. G. H. Choi[19]는 ECG 신호를 스펙트로그램으로 변환하여 기존의 데이터 값을 유지하기 위해 bi-cubic기반의 2D 크기 조정을 통해 다차원 특징을 추출하는 개인식별 방법을 제안하였다. 개인식별을 위해 전처리 과정을 통해 잡음를 제거하고 P파, QRS 복합파, T파가 구성된 한 주기로 분할했다. 분할된 신호는 스펙트로그램으로 변환시켜 이미지의 크기를 1/2, 1/4 만큼 줄여 사용자를 인식했다. D. Jyotishi[20]는 ECG 신호를 이용하여 개인 식별하기 위해 LSTM의 셀의 출력을 합산하여 분류하는 방법를 제안했다. 제안한 모델은 비트 간의 변동을 고려하여 작게 분할할수록 비트의 변화를 잘 확인할 수 있고, 다양한 윈도우 길이에 따라 개인식별을 진행하였다. J. S. Kim[21]는 심전도 신호의 주기 정보를 이용하여 2D 커플링 이미지를 사용한 개인식별을 제안한다. 2D 커플링 이미지는 12개의 컨볼루션 계층, 6개의 최대 풀링 계층으로 구성된 CNN을 사용하여 심전도기반 사용자 인식 방법을 제안했다. M. Hammad[22]는 ECG기반 인증을 위한 종단 간 심층 신경망 모델을 제안했다. 첫 번째 모델은 4개의 합성곱 계층과 2개의 최대 풀링 계층, 2개의 완전 연결 계층, 1개의 최대 풀링 계층으로 구성된 1D-CNN을 설계하였다. CNN의 합성곱 연산은 시계열 또는 이미지 데이터에서 형태학적인 특징을 효율적으로 추출할 수 있다. 두 번째 모델은 2개의 합성곱 계층과 정규화 계층, Rectified Linear Unit(ReLU) 계층, 드롭아웃 계층인 첫 번째 계층과 2개의 정규화 계층, 2개의 ReLU 계층, 2개의 드롭아웃 계층, 2개의 합성곱 계층인 두 번째 계층의 출력을 병합하여 Attention의 입력으로 사용되는 ResNet-Attention을 설계했다. Attention은 2개의 Dense 계층, ReLU 계

조선대학교
CHOSUN UNIVERSITY

층, 소프트맥스 계층을 통해 사용자 인증 성능을 확인한다.

조선대학교
CHOSUN UNIVERSITY

제3장 LSTM과 2D-CNN의 앙상블기반 개인식별

제1절 1D기반 LSTM

1. 장단기 메모리(LSTM, Long Short-Term Memory)

LSTM은 Recurrent Neural Network(RNN)의 아키텍처이다. RNN은 출력 값이 입력으로 연결된 순환구조를 가진 신경망으로 그림 3.1과 그림 3.2는 각각 RNN의 기본 구조와 RNN의 그래디언트 소멸(gradient vanishing)을 보여준다. RNN은 기존 많은 타임 스텝을 가진 시퀀스를 사용할 경우 네트워크도 깊어질수록 학습시키기 위해 사용된 BPTT(Back Propagation Through Time)에서 활성화 함수 하이퍼볼릭 탄젠트(tanh)로 인해 -1부터 1사이의 값이 곱해지면서 chain rule에 의해 초기의 값들이 작아진다. 이로 인해 RNN은 그래디언트 소멸 문제로 데이터의 초기 입력 정보가 출력 결과에 영향을 못미쳐 정보가 소실되는 경우가 생기는 문제가 있다. 수식 3-1과 수식 3-2는 RNN의 계산 과정을 나타낸다. h_{t-1}는 이전 상태, x_t는 셀의 입력, h_t는 셀의 출력, w와 v는 가중치와 편향을 나타낸다.

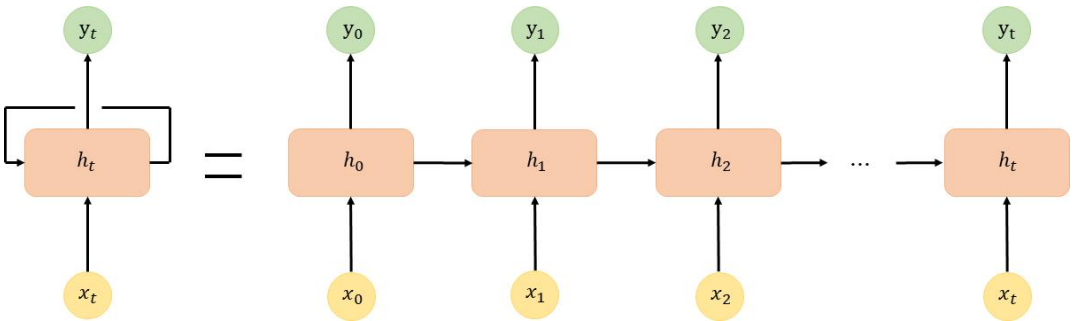


그림 3.1 RNN의 기본 구조

$$h_t = \tanh(w_x x_t + w_h h_{t-1} + v) \tag{3-1}$$

$$y_t = \tanh(W_y h_t + v) \tag{3-2}$$

조선대학교
CHOSUN UNIVERSITY

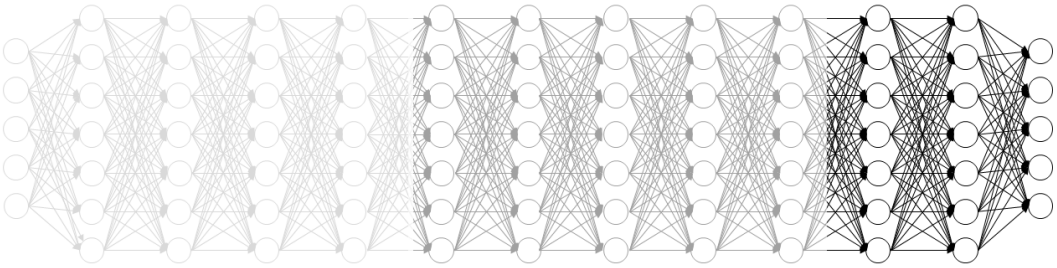


그림 3.2 RNN의 그래디언트 소멸 문제점

이러한 장기 의존성의 문제를 해결하기 위해 RNN보다 복잡한 구조를 가진 LSTM이 제안됐다. LSTM은 정보가 소실되지 않게 메모리 셀에 입력 게이트, 망각 게이트, 출력 게이트인 3개의 게이트로 구성되어 활성화 함수인 시그모이드는 0과 1사이의 값을 출력하므로 값에 따라 정보의 양을 결정하여 cell state에 정보를 추가 또는 제거할 수 있다. LSTM의 활성화 함수는 시그모이드(Sigmoid) 함수와 하이퍼볼릭 탄젠트 함수가 사용된다. Input gate는 새로운 정보를 cell state에 저장할지 결정하는 게이트이고, Forget gate는 과거 정보를 cell state에서 제거할지 결정하는 게이트이다. Output gate는 cell state에서 어떤 정보를 내보낼지 출력을 결정하는 게이트이다. 그림 3.3은 LSTM의 구조를 보여준다. 수식 3-3부터 수식 3-8까지는 LSTM 계산 수식으로 각 게이트의 출력 값을 나타낸다. h_{t-1}는 이전 상태, x_t는 셀의 입력, h_t는 셀의 출력, w와 v는 가중치와 편향을 나타낸다. 수식 3-9는 셀 상태의 업데이트 과정을 나타낸다.

조선대학교
CHOSUN UNIVERSITY

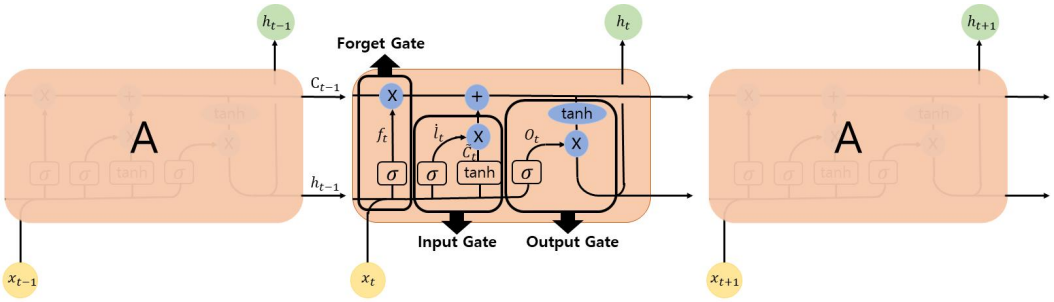


그림 3.3 LSTM의 구조

$$X = [h_{t-1,}\, x_t] \tag{3-3}$$

$$i_t = \sigma(w_i \bullet X + v_i) \tag{3-4}$$

$$\widetilde{C_t} = \tanh(w_C \bullet X + v_c) \tag{3-5}$$

$$f_t = \sigma(w_f \bullet X + v_f) \tag{3-6}$$

$$o_t = \sigma(w_o \bullet X + v_o) \tag{3-7}$$

$$h_t = o_t \times \tanh(C_t) \tag{3-8}$$

$$C_t = f_t \times C_{t-1} + i_t \times \widetilde{C_t} \tag{3-9}$$

조선대학교
CHOSUN UNIVERSITY

2. 심전도를 이용한 1D기반 LSTM

1차원 시계열 또는 시퀀스 신호를 분석하기 위해서는 순차적인 정보를 파악해야 하기 때문에 LSTM 신경망을 사용한다. LSTM은 시계열 또는 시퀀스 데이터를 입력하는 시퀀스 입력 계층, LSTM 계층은 시퀀스의 시간 스텝 사이의 장기 의존성을 학습 시키기 위한 계층, 클래스 레이블을 분류하기 위한 완전 연결 계층, 소프트맥스(softmax) 계층, 분류 계층으로 구성되어 있다. 은닉층의 수와 세 개의 셀로 구성된 유닛의 수가 증가할수록 분류 성능 정확도가 향상되기 때문에 LSTM 계층을 추가함으로서 LSTM 신경망의 심층구조를 한층 높일 수 있다. 그림 3.4와 그림 3.5는 각각 한 개의 LSTM 계층과 두 개의 LSTM 계층을 보여주고 있다.

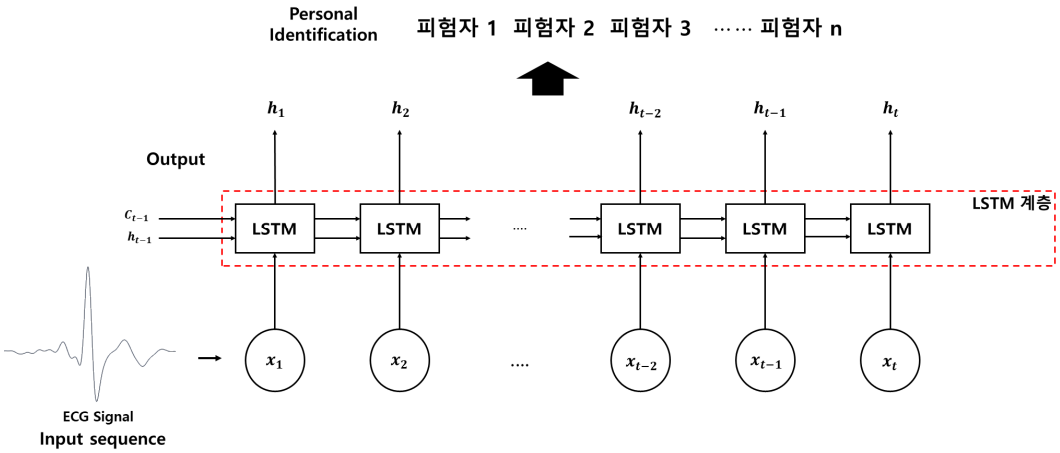


그림 3.4 한 개의 LSTM 계층

조선대학교
CHOSUN UNIVERSITY

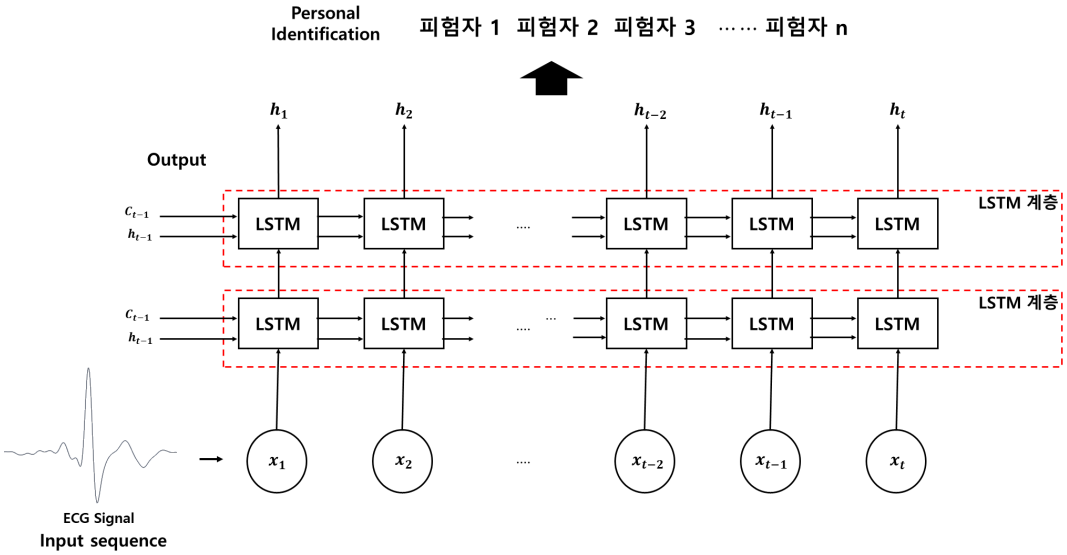


그림 3.5 두 개의 LSTM 계층

조선대학교
CHOSUN UNIVERSITY

제2절 시간-주파수 변환기반 2D-CNN

1. 시간-주파수 변환 방법

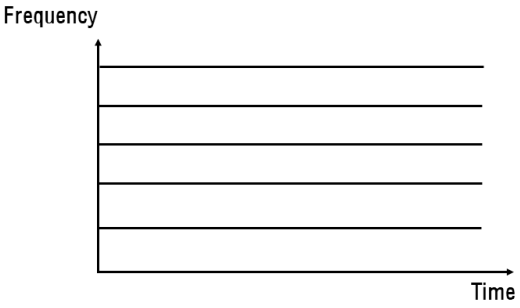
생리학적인 신호는 잡음의 영향을 크게 받기 때문에 신호 분석을 위해 시간-주파수 영역으로 변환하여 2차원 이미지 표현으로 나타낸다[23]. 시간-주파수 변환을 통해 심전도 신호는 STFT, Scalogram, CQT, FSST, WSST으로 변환된 2차원 이미지를 사용한다. 시간-주파수 변환으로 표현된 이미지는 이미지 분류에서 높은 성능을 나타내는 CNN을 통해 분류한다.

가. STFT

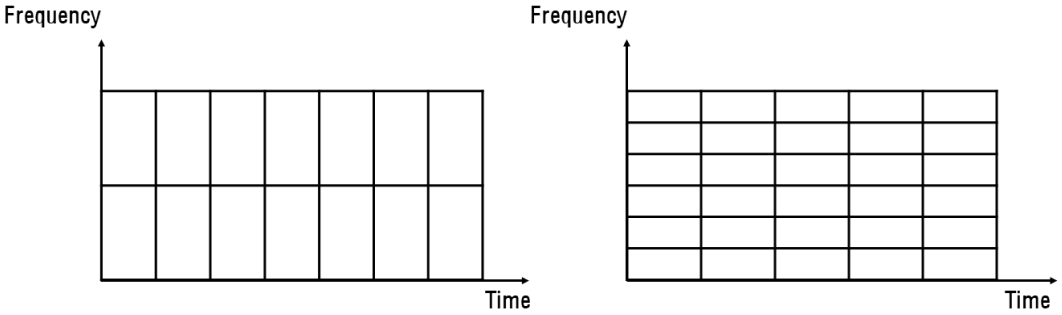
기존의 푸리에 변환(fourier transform)은 시계열 신호를 주파수로 분해하는 것으로 주파수 표현이다. 이는 신호에 존재하는 주파수를 분석할 수 있지만, 시간에 따른 변화가 고려되지 않는다. 이로 인해 시간에 대해 각 주파수가 존재하는 위치를 알 수 없다는 단점으로 기존의 푸리에 변환은 이미지를 분석하기 어려운 점이 있다[24]. 이를 보완하기 위해 STFT, DTFT 등이 연구되었다.

STFT는 시간에 따라 변하는 긴 신호를 동일한 길이로 짧게 분할하여 푸리에 변화에 적용시키는 것이다. 신호를 짧은 시간 단위로 분할을 할수록 특정 시간에 특정 주파수 존재를 파악할 수 있으며, 신호를 긴 시간 단위로 분할을 할수록 특정 주파수가 특정 시간 내에 존재하는지 파악할 수 있다. 분할하려는 윈도우의 너비가 작아질수록 시간 분해능이 좋고, 너비가 커질수록 주파수 분해능이 좋아진다. 그림 3.6(a)와 그림 2.7(b)는 푸리에 변환과 STFT의 개념적인 차이를 보여주고 그림 3.7(a)는 STFT의 윈도우 길이가 128일 때 분해능, 그림 3.7(b)는 STFT의 윈도우 길이가 256일 때 분해능, 그림 3.7(c)는 STFT의 윈도우 길이가 512일 때 분해능으로 윈도우 길이에 따른 분해능의 차이를 보여준다.

조선대학교
CHOSUN UNIVERSITY



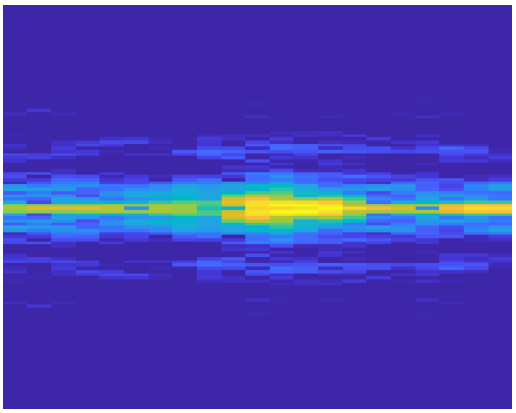
(a)



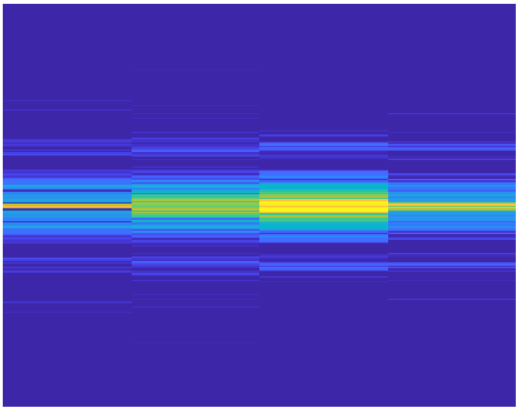
(b)

그림 3.6 푸리에 변환과 STFT의 개념적인 차이

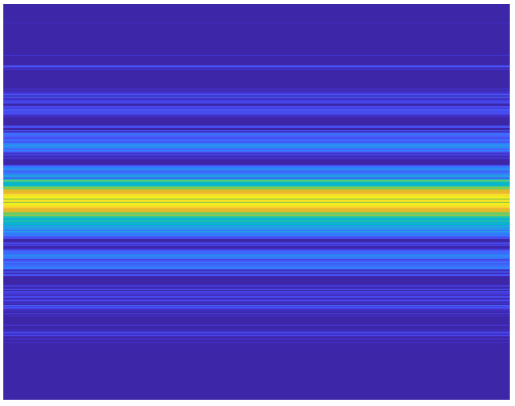
조선대학교
CHOSUN UNIVERSITY



(a) 윈도우 길이 128



(b) 윈도우 길이 256



(c) 윈도우 길이 512

그림 3.7 STFT의 윈도우 길이에 따른 분해능의 차이

조선대학교
CHOSUN UNIVERSITY

수식 3-10은 STFT의 신호를 분할하는 과정으로 신호와 움직이는 윈도우 함수를 사용하여 나타낸다. 수식 3-11는 STFT의 푸리에 변환 계산 과정을 나타낸다. 그림 3.8는 푸리에 변환 계산 과정 분석 보여주고, 시간 t에 대한 신호 $x(t)$, 윈도우 함수 $w(t)$, 윈도우 함수의 중점인 f를 보여준다[25].

$$x(f,t) = x(t)*w(t-f) \tag{3-10}$$

$$x(f,u) = \int_{-\infty}^{\infty} x(t)w(t-f)e^{-iut}dt \tag{3-11}$$

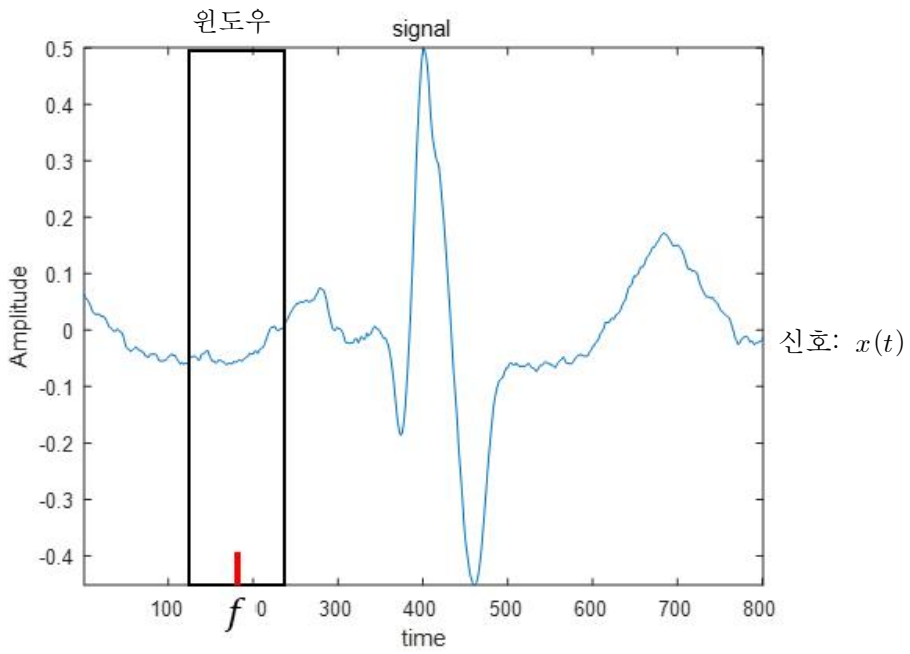


그림 3.8 푸리에 변환 계산 과정 분석

조선대학교
CHOSUN UNIVERSITY

나. Scalogram

Scalogram은 신호에 대해 연속 웨이블릿 변환의 절대값으로 scalogram을 표현하기 위해 연속 웨이블릿 변환을 계산하기 때문에 다음은 웨이블릿에 대해 설명한다. STFT는 기존 푸리에 변환의 한계를 보완한 것으로 윈도우의 길이에 따라 시간 또는 주파수에 대한 자세한 정보를 얻을 수 있지만, 고정된 윈도우의 길이로 동시에 시간과 주파수에 대한 정보를 얻기는 어렵다. 이러한 STFT 한계를 보완하기 위해 연구된 방법인 웨이블릿 변환(wavelet transform)은 그림 3.9는 연속 웨이블릿의 변환 영역으로 시간과 주파수 동시에 파악이 가능하다. 웨이블릿 변환은 고주파 영역의 신호에서 시간 분해능을 높이고 주파수 해상도을 낮추며 저주파 영역의 신호에서 주파수 분해능을 높이고 시간 해상도를 낮춤으로써 시간과 주파수에 대한 정보를 동시에 파악할 수 있기 때문에 불연속적인 신호를 분석하는 것이 효율적이다. 기존의 푸리에 변환은 시간 영역에서 무한한 사인 함수를 기본으로 사용한 반면 웨이블릿 변환은 시간 영역에서 한정되어 있는 웨이블릿 변환의 함수인 mother 웨이블릿을 사용하여 신호에 대한 스케일링과 시프팅을 통해 신호를 시간과 주파수 영역에서 분석할 수 있다[26]. 수식 3-12은 웨이블릿 변환(CWT, Continuous Wavelet Transform)의 공식을 나타낸다. a와 b는 mother 웨이블릿의 크기 조정, mother 웨이블릿을 시프팅, $h(t)$는 입력 신호, ψ는 mother 웨이블릿을 의미한다.

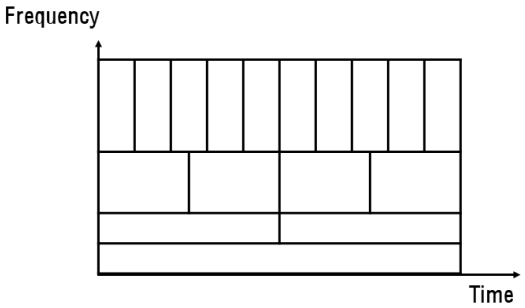


그림 3.9 웨이블릿의 변환 영역

$$C(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} h(t)\psi(\frac{t-b}{a})dt \qquad (3\text{-}12)$$

조선대학교
CHOSUN UNIVERSITY

mother 웨이블릿의 종류는 다양하고 종류에 따라 분석 결과도 다르게 표현되기 때문에 분석 유형마다 mother 웨이블릿의 종류가 다르게 사용된다. 연속 웨이블릿 변환을 위해 사용되는 mother 웨이블릿은 Morse, Morlet, Bump 등이 있다. Morse 웨이블릿은 시간과 주파수, 진폭을 가진 신호를 분석하기 유용하다[27]. 수식 3-13은 Morse 웨이블릿의 푸리에변환 수식으로 $P(\omega)$는 단위 단계, $a_{T,\gamma}$는 정규화 상수, γ과 T^2은 Morse 웨이블릿의 대칭을 나타내기 위한 매개변수, 시간과 대역폭의 곱을 나타낸다. β는 시간과 대역폭 곱보다는 감쇠 또는 압축 파라미터로 수식 3-14는 β와 γ가 매개 변수화로 표현된 Morse 웨이블릿 방정식이다. 두 개의 매개변수를 필요에 따라 변경하여 Morse 웨이블릿을 나타낼 수 있다[28]. T^2는 시간과 대역폭의 곱으로 시간에 따라 변하는 웨이블릿 지속 시간에 비례한다. 또한 지속 시간은 중앙 창에서 최대 피크 주파수가 윈도우 중앙에 위치하는 진동수로 수식 3-15는 최대 피크 주파수를 나타내고, 매개변수 γ은 시간에 따라 변하는 웨이블릿의 대칭을 제어한다.

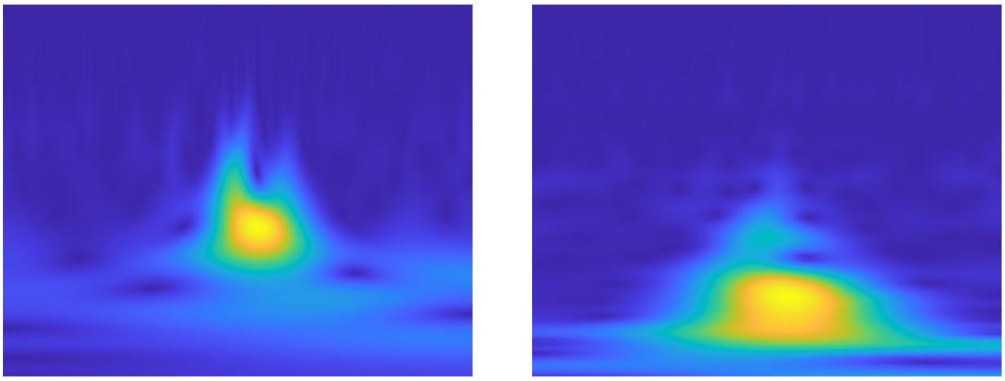
$$\Psi_{T^2,\gamma}(\omega) = P(\omega)a_{T,\gamma}\omega^{\beta}e^{-\omega^{\gamma}} \tag{3-13}$$

$$\Psi_{\beta,\gamma}(\omega) = P(x)k_{\beta,\gamma}\omega^{\beta}e^{-\omega^{\gamma}} \tag{3-14}$$

$$(\frac{T^2}{\gamma})^{\frac{1}{\gamma}} \tag{3-15}$$

그림 3.10는 γ=3일 때 T^2에 따른 Morse 웨이블릿을 보여주며, 그림 3.10(a)는 T^2=10일 때 Morse 웨이블릿 변환 결과를 나타내고, 그림 3.10(b)는 T^2=60일 때 Morse 웨이블릿 변환 결과를 나타낸다. Morse 웨이블릿 (3,10)과 Morse 웨이블릿 (3,60)을 보여줌으로써 Morse 웨이블릿(3,60)이 Morse 웨이블릿(3,10)보다 주파수 분해능이 높은 것을 파악할 수 있다.

조선대학교
CHOSUN UNIVERSITY



(a) T^2 = 10 (b) T^2 = 60

그림 3.10 T^2에 따른 Morse 웨이블릿

다. FSST

생리학적 신호, 음성 등과 같이 진동 모드인 신호들은 진폭 및 주파수 변조의 중첩으로 표현될 수 있다. 시간-주파수(TF, Time-Frequency) 분석은 수식 3-16과 같이 분석 신호의 합으로 표현한다. $f(x)$와 $\phi_n(t)$는 분석 신호의 $X_n(t)$의 시간-주파수 진폭과 위상, N은 분석 신호 수, j는 $\sqrt{-1}$이다. FSST는 스펙트럼 함수에서 사용된 STFT를 기반으로 선명한 시간-주파수 표현을 나타내며 원 신호와 동일한 시간 분해능을 유지하여 표현하는 변환 방법으로 사용된다[29]. 수식 3-17부터 수식 3-19은 FSST를 계산하는 과정이다. 그림 2.12는 FSST에 의한 시간-주파수 표현된 이미지를 보여준다.

$$f(x) = \sum_{n=1}^{N} f_n(t) = \sum_{n=1}^{N} X_n(t) e^{j2\pi\phi_n(t)} \tag{3-16}$$

$$H_g f(x,\eta) = \int_{-\infty}^{\infty} f(k) g(k-x) e^{-j2\pi\eta(k-x)} dk \tag{3-17}$$

조선대학교
CHOSUN UNIVERSITY

$$V_g f(x,\omega) = \int_{-\infty}^{\infty} H_g f(x,\eta)\delta(\omega - \Omega_g f(x,\eta))d\eta \tag{3-18}$$

$$\Omega_g f(x,\eta) = \frac{1}{j2\pi}\frac{\frac{\partial}{\partial x}H_g f(x,\eta)}{H_g f(x,\eta)} = \eta - \frac{1}{j2\pi}\frac{H_{\partial_g/\partial_x} f(x,\eta)}{H_g f(x,\eta)} \tag{3-19}$$

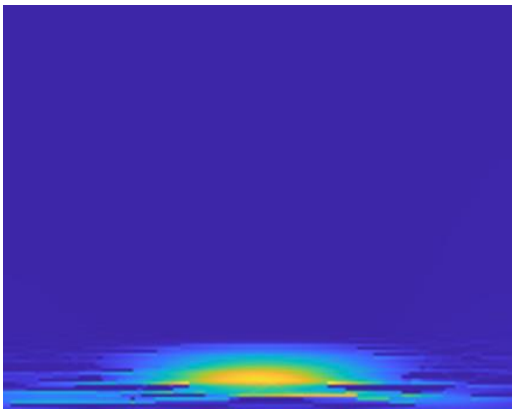


그림 3.11 FSST에 의한 시간-주파수 표현

라. WSST

주파수에서 신호 에너지를 재할당하는 시간-주파수 표현 방법으로 모 웨이블릿으로 인한 산란 효과를 보상한다. Synchrosqueezing은 시간-주파수 표현을 위해 재할당하는 다른 방법들과 다르게 시간 분해능을 유지하며 주파수 방향으로 에너지를 재할당한다[30]. 또한 synchrosqueezing은 연속 웨이블릿 변환과 관련하여 1차 도함수를 사용하며 연속 웨이블릿 변환은 가역성, Synchrosqueezing 변환은 연속 웨이블릿 변환의 속성에 상속하기 때문에 신호 재구성이 가능하다. WSST을 나타내는 알고리즘은 다음과 같다.

1. 수식 3-20과 같이 입력 신호의 연속 웨이블릿 변환을 구한다.

$$C(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} h(t)\psi(\frac{t-b}{a})dt \tag{3-20}$$

조선대학교
CHOSUN UNIVERSITY

2. Synchrosqueezing로 표현하기 위해서는 연속 웨이블릿 변환의 출력에서 순간 주파수 정보를 추출한다. 수식은 3-21을 나타내며, a와 b는 스케일링과 시프트 파라미터를 의미한다.

$$h_c = -i\frac{\frac{dC(a,b)}{db}}{C(a,b)} \tag{3-21}$$

3. 위상 변환은 일정한 영역에 연속 웨이블릿 변환을 압착시켜 순간 주파수 정보 값은 단일 값으로 재할당된다. 이로 인해 WSST는 선명한 해상도를 갖는 출력을 얻는다.

그림 3.12은 WSST에 의한 시간-주파수 표현된 이미지를 보여준다.

그림 3.12 WSST에 의한 시간-주파수 표현

조선대학교
CHOSUN UNIVERSITY

2. 합성곱 신경망(CNN, Convolutional Neural Network)

딥러닝은 기계학습 기술의 일종으로 인간의 뇌에서 뉴런의 구조와 유사하게 설계된 신경망이다. 이 신경망은 입력 계층, 은닉층, 출력 계층을 포함한 여러 층으로 구성된 심층 신경망을 의미한다. 심층 신경망은 2개 이상의 은닉층을 가진 신경망이다. 기존의 얕은 구조의 신경망은 복잡한 연산을 수행할 수 없고, 학습 과정에서 그래디언트 소멸 또는 오버피팅(overfitting)이 발생하지만 심층 신경망을 이용하여 이러한 문제들을 해결하여 학습이 가능하고 높은 성능을 기대할 수 있다.

CNN은 딥러닝의 아키텍처로 주로 이미지, 시계열 분야에서 가장 널리 사용되고 있는 신경망이다. CNN은 입력 데이터가 합성곱을 통해 특징이 추출됨으로써, 2차원 데이터를 분석 및 처리하기에 매우 적합한 아키텍처이다. CNN은 합성곱 계층, 활성화 함수인 ReLU 계층, 풀링 계층이 반복적으로 구성되며 그림 3.13은 CNN의 기본 구조를 보여준다.

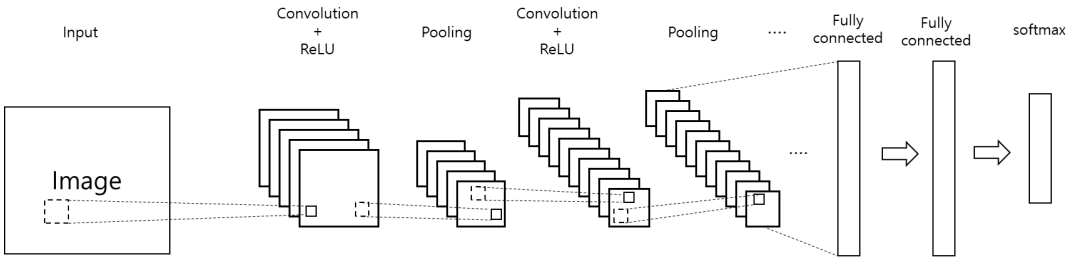


그림 3.13 CNN의 기본 구조

합성곱 계층은 합성곱 연산을 통해 입력 데이터로부터 특징을 추출하는 계층이다. 연산은 필터 크기의 이미지와 이동하는 필터의 각 원소들을 곱셈과 덧셈으로 수행하여 값을 출력한다. 패딩(padding)은 입력 데이터 주변에 특정한 값인 0을 채우는 방법으로 출력의 크기를 조정하는 과정으로 합성곱 연산을 통해 입력 데이터의 크기가 계속해서 작아지는 것을 방지한다. 스트라이드(stride)는 입력된 이미지에 필터를 적용하는 위치의 간격으로 필터가 스트라이드 값만큼 이동하면서 합성곱 연산이 수행된다. 그림 3.14는 합성곱의 연산 과정 보여준다.

조선대학교
CHOSUN UNIVERSITY

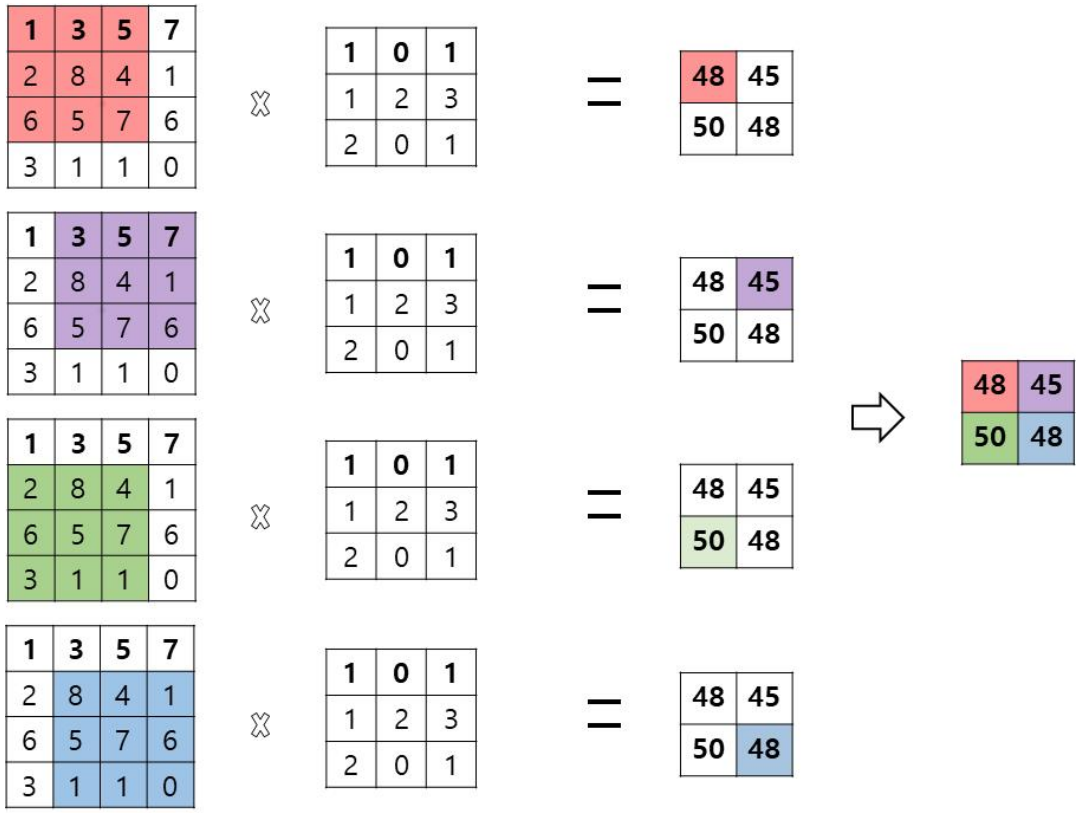


그림 3.14 합성곱 연산 과정

조선대학교
CHOSUN UNIVERSITY

활성화 함수는 비선형 함수로 시그모이드, ReLU, 계단 함수(step function), 하이퍼탄젠트 함수, 소프트맥스 등이 있다. 활성화 함수는 합성곱 계층과 풀링 계층 사이에 위치하고 주로 사용되는 활성화 함수는 ReLU 함수이다. Relu 함수는 0과 1으로 나타내며 음수 값은 0을 출력, 0보다 크면 입력 값 그대로를 출력하는 함수이다. 그림 3.15는 ReLU 함수를 보여주고 수식 3-22는 ReLU 함수의 수식을 나타낸다.

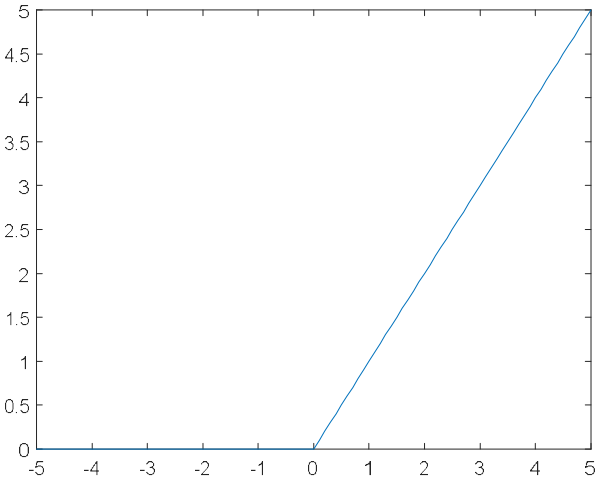


그림 3.15 ReLU 함수

$$f_x = \max(0, x), \quad f(x) = \begin{cases} 0 & (x < 0) \\ x & (x \geq 0) \end{cases} \tag{3-22}$$

조선대학교
CHOSUN UNIVERSITY

풀링 계층은 이미지 중요한 특징을 유지하면서 차원을 줄이는 계층으로 최대 풀링, 평균 풀링, L2-norm 풀링 등이 있다. 주로 최대 풀링이 사용되며 최대 풀링은 대상 영역에서 각 영역의 최댓값을 나타내는 방법이다. 그림 3.16은 최대 풀링의 결과를 보여준다.

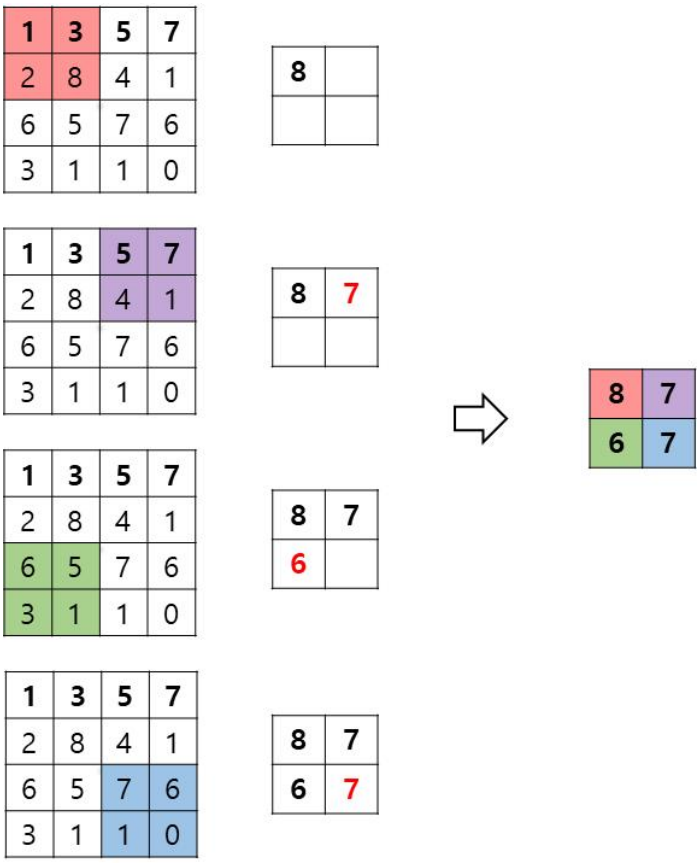


그림 3.16 최대 풀링 연산 과정

완전 연결 계층은 뉴런이 이전 계층의 뉴런과 연결된 계층으로 1차원 형태로 이미지를 분류하기 위해 사용된다. 소프트맥스 계층은 최종 분류 결과를 확률로 나타낸 것으로 출력의 값들의 합이 항상 1인 되는 특성을 갖는다. 이를 통해 CNN은 기존 신경망에 합성곱 계층과 풀링 계층을 추가하여 이미지를 분류하는데 우수한 성능을 보인다.

조선대학교
CHOSUN UNIVERSITY

3. 2D 변환기반 CNN

STFT, Scalogram, FSST, WSST와 같이 시간-주파수 변환으로 표현된 이미지는 이미지 분류에 높은 성능을 보여주는 CNN을 사용한다. CNN은 직접 설계하여 성능을 확인하는 방법과 CNN 모델을 포함한 전이학습(transfer learning)을 사용하여 성능을 확인하는 방법이 있다. CNN을 기반으로 한 딥러닝 모델을 훈련 시키기 위해서는 많은 데이터의 수가 필요하며 데이터에 맞게 레이어의 개수, 하이퍼파라미터 등을 수정하고 복잡한 모델의 경우에는 학습 시간이 오래 걸린다는 단점이 있다. 따라서 CNN의 모델인 전이학습을 사용한다. 전이학습은 기존에 만들어진 모델을 이용해 새로운 데이터를 학습시키는 방법으로, 데이터 수가 작을 경우에도 사용이 가능하고 AlexNet, GoogleNet, VGG, ResNet, SqueezeNet 등이 있다. 2D 변환기반 CNN으로 GoogleNet, VGG-19, ResNet-101을 사용한다. GoogleNet은 학습이 가능한 계층의 수가 22개인 심층 네트워크로 9개의 인셉션 모듈을 포함한 심층 신경망이며, 그림 3.17과 같이 구성되어 있다. 또한 인셉션 모듈을 포함하여 출력이 연결된 병렬 컨볼루션 필터로 구성되어 있다[31]. 그림 3.18은 인셉션 모듈로 1*1 컨볼루션, 1*1 컨볼루션 + 3*3 컨볼루션, 1*1 컨볼루션 + 5*5 컨볼루션, 3*3 max pooling + 1*1 컨볼루션으로 4가지 연산이 수행된다. 이 연산들은 채널의 수 조절, 파라미터 수를 감소시켜줌으로서 필요한 연산량을 줄여주기 위해 사용된다. GoogleNet과 같이 심층 신경망의 경우 그래디언트 소멸 문제가 생기기 때문에 방지하기 위해 훈련 과정 중에 계산된 손실 값을 추가하는 과정으로 그림 3.19와 같이 보조 분류기가 포함되어 있다. 보조 분류기는 훈련 과정 중에만 사용되며 테스트 과정에서는 사용되지 않는다.

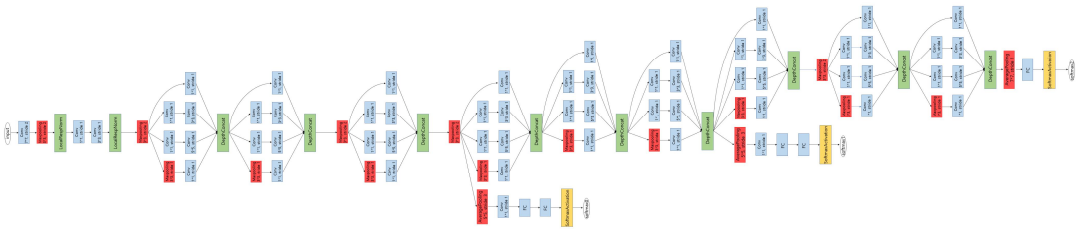


그림 3.17 GoogleNet의 구조

조선대학교
CHOSUN UNIVERSITY

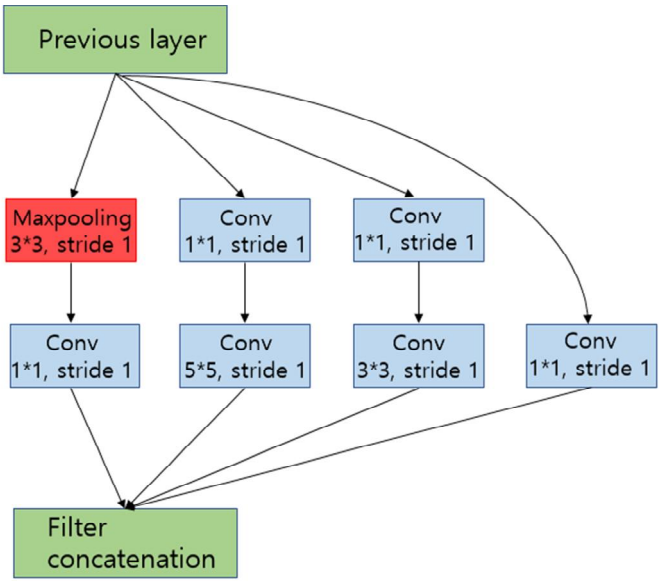


그림 3.18 인셉션 모듈

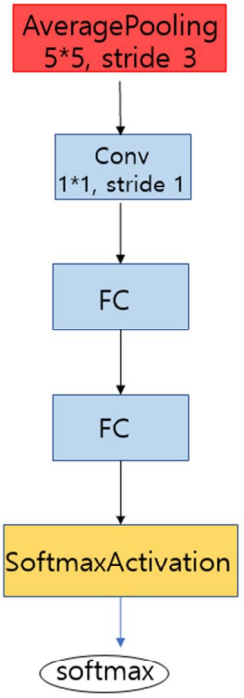


그림 3.19 보조 분류기

조선대학교
CHOSUN UNIVERSITY

VGG-19는 합성곱 계층과 최대 풀링 계층으로 구성된 5개의 블록으로 19개의 계층으로 구성되어 있으며 컨볼루션 연산을 위해 가장 작은 3*3 필터의 크기를 사용한다[32]. 그림 3.20은 VGG-19의 구조를 보여준다.

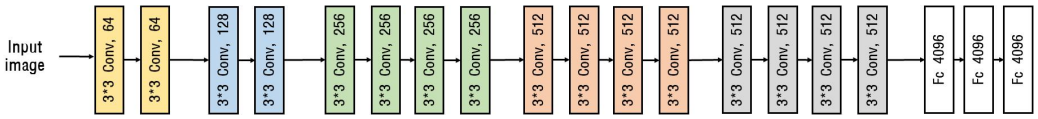


그림 3.20 VGG-19의 구조

ResNet은 104개의 컨볼루션 계층을 포함하며 33개의 계층 블록으로 구성되어 있다. 블록은 1*1 컨볼루션, 3*3 컨볼루션, 1*1 컨볼루션으로 병목 계층을 사용하여 연산을 줄일 수 있다. 또한 그래디언트 소멸을 해결하기 위해 Residual Connection을 추가한다. Residual Connection을 통해 출력 값에 x를 더하므로 $P(x) = F(x) + x$로 나타낸다. 이로 인해 그래디언트 소멸의 문제가 해결되고 신경망의 계층을 깊게 쌓아 더 나은 성능을 얻을 수 있다[33]. 그림 3.21은 Residual Connection을 보여주고, 그림 3.22는 ResNet-101의 구조를 보여준다.

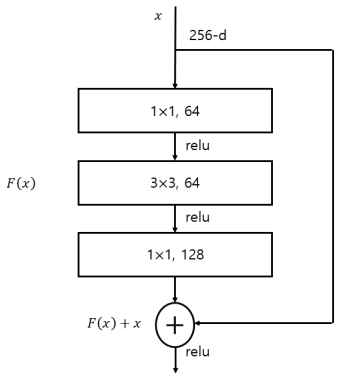


그림 3.21 Residual Connection

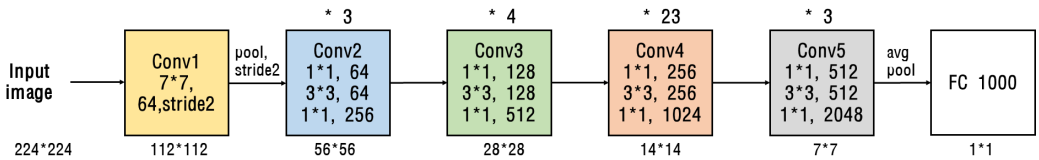


그림 3.22 ResNet-101의 구조

조선대학교
CHOSUN UNIVERSITY

제3절 제안된 앙상블기반 개인식별

다양한 모델을 통해 학습시키고 그 중 높은 성능을 가진 모델을 사용하여 각 모델의 성능 비교를 할 수 있다. 하지만 높은 성능을 가진 단일 모델만 사용할 경우 다른 모델들의 성능이 향상될 수 있는 경우를 놓치는 한계점이 있다. 이러한 한계점을 위해 다양한 모델의 결합으로 성능 향상 시킬 수 있다. 앙상블은 여러 모델의 결합을 통해 성능을 향상 시키기 위한 기법으로 단일 모델의 성능보다 더 높은 성능을 얻을 수 있다. 딥러닝 모델과 앙상블의 장점을 이용하여 계산, 표현, 통계 등 측면에서 향상시킬 수 있다[34]. 앙상블 방법은 배깅(Bagging), 스태킹(Stacking), 부스팅(Boosting)이 있고, 그린 3.23과 같이 동일한 입력에 대해 각 모델의 출력 값들을 투표 방식, 평균, 최댓값, 곱셈 등 사용하여 앙상블을 통해 최종 예측 결과를 확인할 수 있다. 각 모델의 출력값의 중앙값, 최솟값, 최댓값, 평균값으로 도출된 모델들의 값을 평균, 덧셈 또는 곱셈을 통해 앙상블의 결과값을 나타낸다.

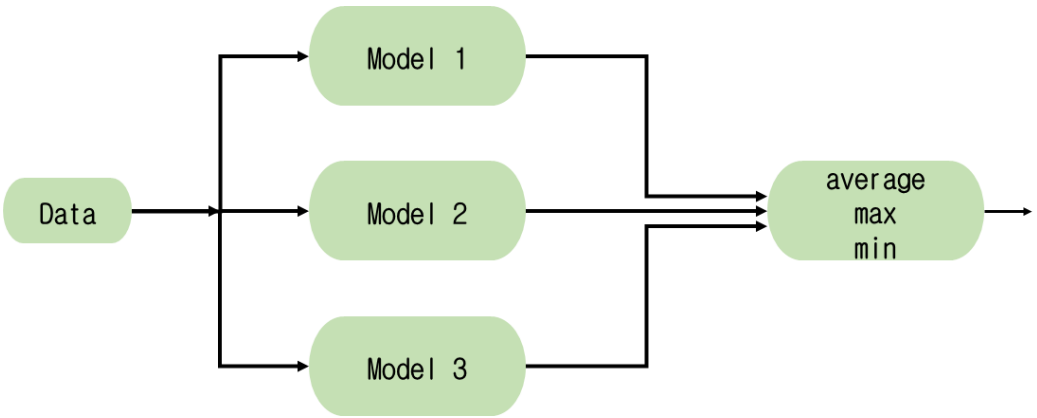


그림 3.23 앙상블 모델 기반의 분류 과정

조선대학교
CHOSUN UNIVERSITY

본 논문에서는 딥러닝 모델 LSTM과 2D-CNN을 통해 개인을 식별한다. 분류 성능 정확도를 높이기 위해 은닉층 수와 유닛의 수를 증가시켜 LSTM 계층을 추가하여 LSTM의 심층을 높인다. 또한 1차원 심전도 신호를 2차원으로 변환시키기 위해 시간-주파수 표현 방법을 사용하여 사전에 학습된 모델 GoogleNet, VGG-19, ResNet-101 세 개의 CNN 모델을 사용한다. 단일 모델의 성능을 높이기 위해 두 개의 모델의 출력 값을 결합하는 앙상블 방법을 사용한다. 앙상블은 동일한 입력에 대해 각 모델의 출력 값을 이용하여 최종 예측 결과 값을 도출하여 개인식별한다.

심전도를 이용하여 개인식별하기 위해서 그림 3.24와 같이 신호 전처리 과정 단계, 특징 추출 및 학습 단계, 개인식별 단계인 세 단계로 나눌 수 있다.

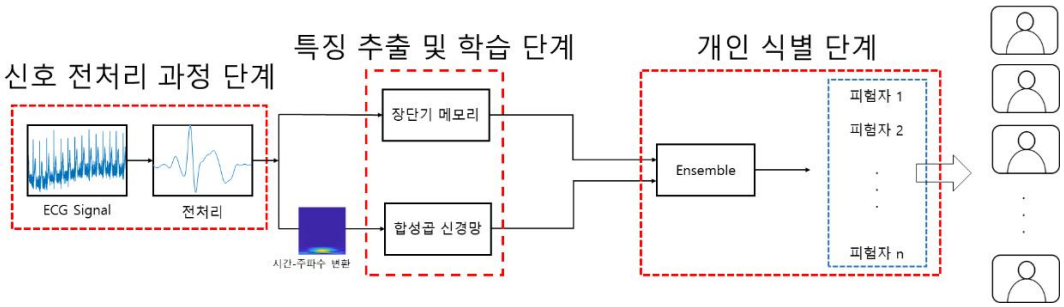
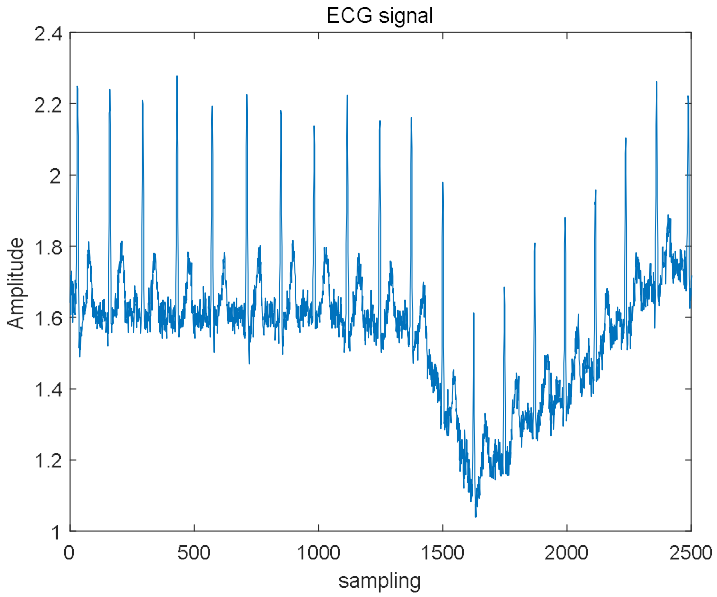


그림 3.24 제안된 앙상블기반 개인식별 흐름도

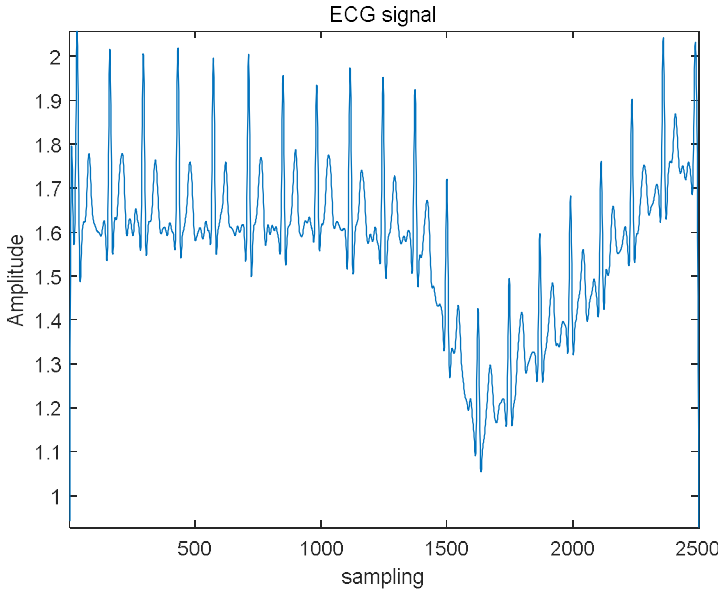
첫 번째 신호 전처리 과정 단계는 심전도 신호를 측정 시 다양한 잡음으로 인해 신호가 왜곡되기 때문에 심전도 형태 유지를 위해 잡음들을 제거하는 과정이 필요하다. 원 신호에 저역 통과 필터, 고역 통과 필터, 평균 이동 필터를 적용하여 잡음을 제거한 후 기저선 변동된 신호를 영점으로 조정한다. 심전도는 측정 시 주로 발생하는 호흡, 전극과 피부와의 마찰, 근육에서 발생한 잡음, 전극과 전력선 접촉의 잡음 등이 포함되어 측정된다. 이러한 잡음들은 저역 통과 필터에 의해 그림 3.25(a)에서 그림 3.25(b)와 같이 제거된다. 잡음이 제거된 그림 3.25(b)처럼 신호의 기저선이 일정하지 않고 기저선이 신호의 x축이 위 또는 아래로 이동하는 것을 확인할 수 있다. 기저선 변동은 피험자의 호흡, 땀 또는 움직임으로 인해 발생하는 저주파 진동으로 이러한 변동은 전극과 피부 사이의 측정자의 임피던스의 변화로 인해 발생한다. 기저선은 심전도 신호의 특징점을 검출하는 기준으로 기저선 변동을 제거하지 않으면 심전도 신호의 형태학적 특성을 파악하기 어렵기 때문에 기저선 잡음 제거 과정이 필요하다. 그림 3.25(c)는 기저선이 영점 기준으로 조정

조선대학교
CHOSUN UNIVERSITY

된 신호를 보여주고, 그림 3.25(d)는 표준화한 신호를 보여준다.

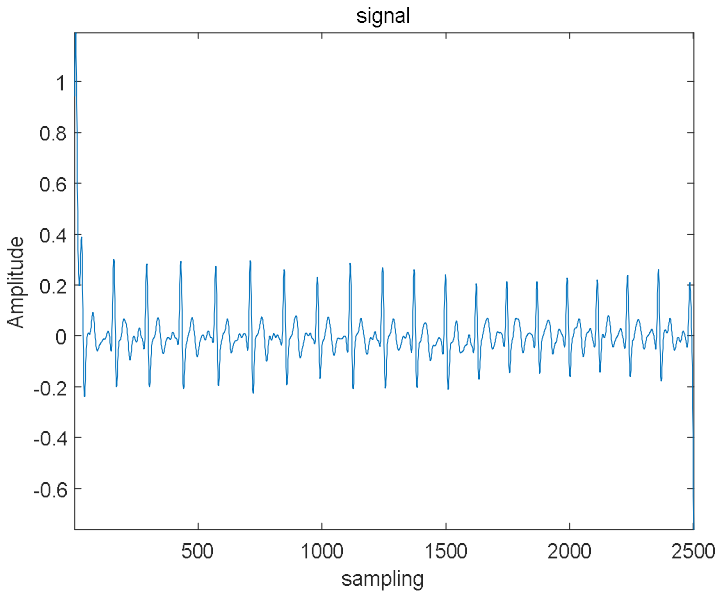


(a) 원신호

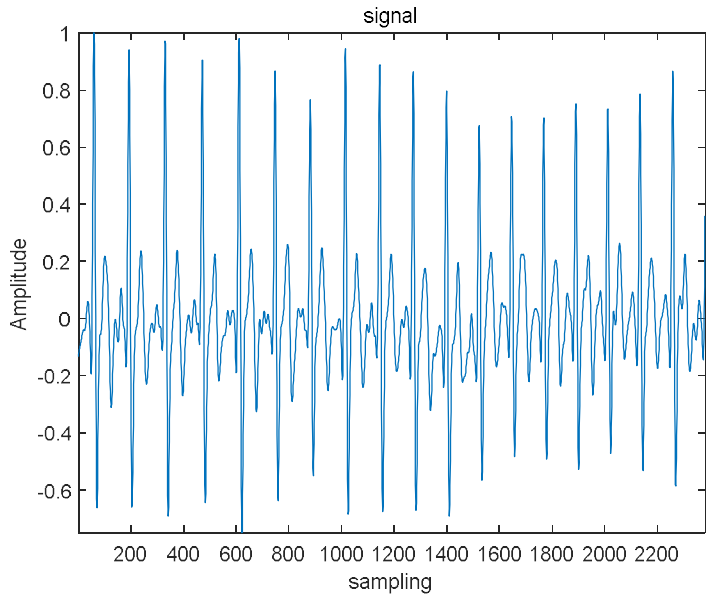


(b) 필터링

조선대학교
CHOSUN UNIVERSITY



(c) 기저선 조정



(d) 표준화

그림 3.25 전처리 과정을 통한 심전도 신호

조선대학교
CHOSUN UNIVERSITY

그림 3.26은 한 사람의 4개의 신호로 동일한 피험자의 경우에도 기저선이 변동되어 기저선 조정 없이 동일한 피험자라고 보기 어렵다는 것을 보여준다. 심전도 신호는 그림 3.27(a)과 같이 P파, QRS 복합파, T파로 구성되며 다양한 주기를 포함한다. 심전도 신호의 특징들을 추출하기 위해 분할과정을 진행한다. 그림 3.27(b)와 같이 R-피크를 검출하여 R-피크 중심으로 일정하게 한 주기로 분할된다.

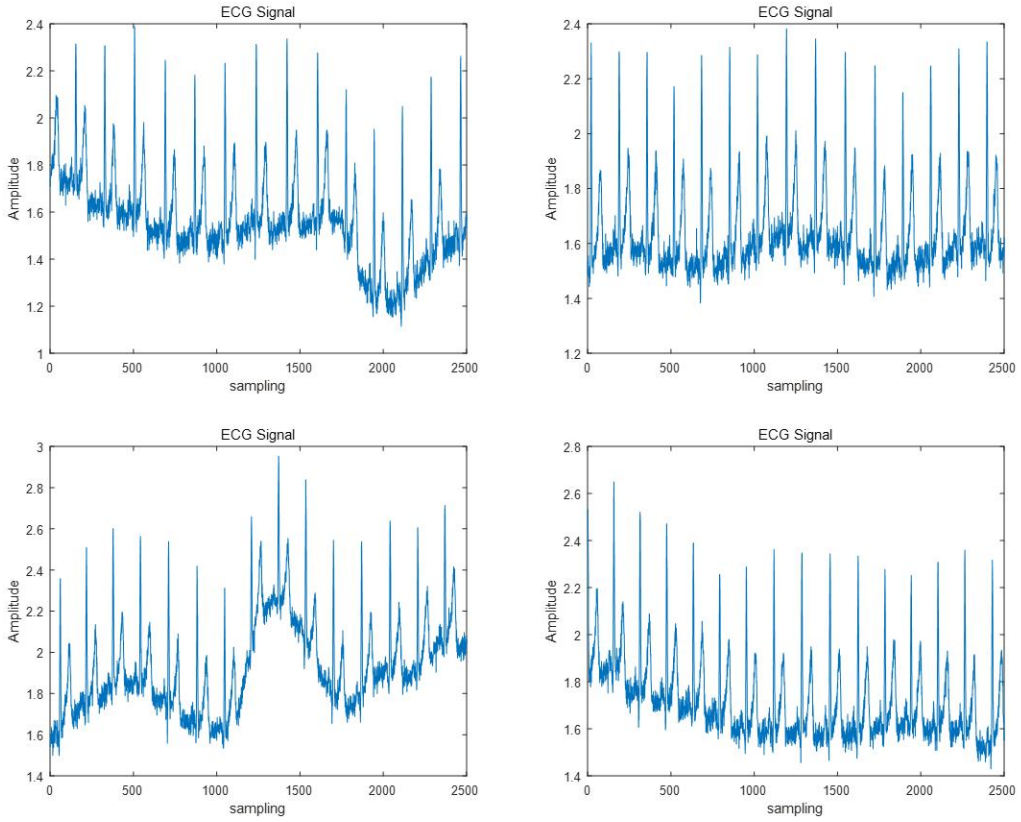
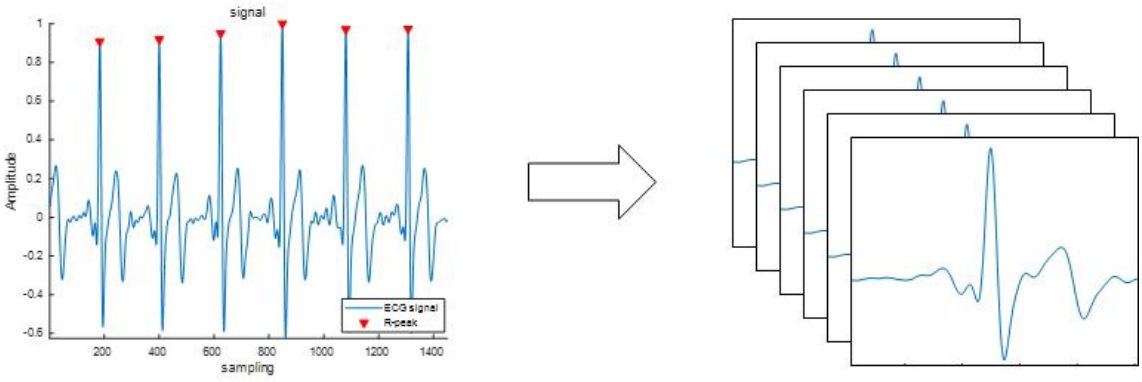


그림 3.26 59번 피험자의 심전도 신호

조선대학교
CHOSUN UNIVERSITY



(a) (b)

그림 3.27 검출된 R-피크 및 한 주기로 분할된 신호

두 번째 특징 추출 및 학습 단계에서는 분류 성능 결과에 영향을 줌으로써 LSTM 신경망과 1차원 심전도 신호를 2차원 이미지로 변환시키기 위해 시간-주파수 표현 방법을 통해 사전에 학습된 모델 GoogleNet, VGG-19, ResNet-101을 사용한다. 각 이미지의 크기는 224*224로 축소되며 학습 방법으로 Adam을 사용했다. Adam은 이동 평균과 모멘텀을 모두 유지함으로써 손실 함수의 최소값을 구할 수 있다. 초기 학습률과 에포크, 미니배치사이즈는 모델에 맞게 설정하여 개인식별 성능을 확인한다.

세 번째 개인식별 단계에는 특징추출 및 학습 단계에서 학습시킨 개인식별 성능 결과를 이용하여 단일 모델의 성능보다 여러 모델의 성능을 합쳐 더 높은 성능을 기대할 수 있는 앙상블을 사용한다. 앙상블은 모델들의 출력 값을 곱셈을 통해 앙상블의 최종 결과 값을 나타낸다. 그림 3.28은 사용한 4개의 변환 방법과 3개의 모델 중 대표적으로 Scalogram과 GoogleNet을 이용한 제안된 앙상블기반 개인식별 순서도를 보여준다. 두 개의 데이터베이스는 전처리 과정을 통해 LSTM와 CNN의 출력 값을 결합하여 피험자를 분류한다.

조선대학교
CHOSUN UNIVERSITY

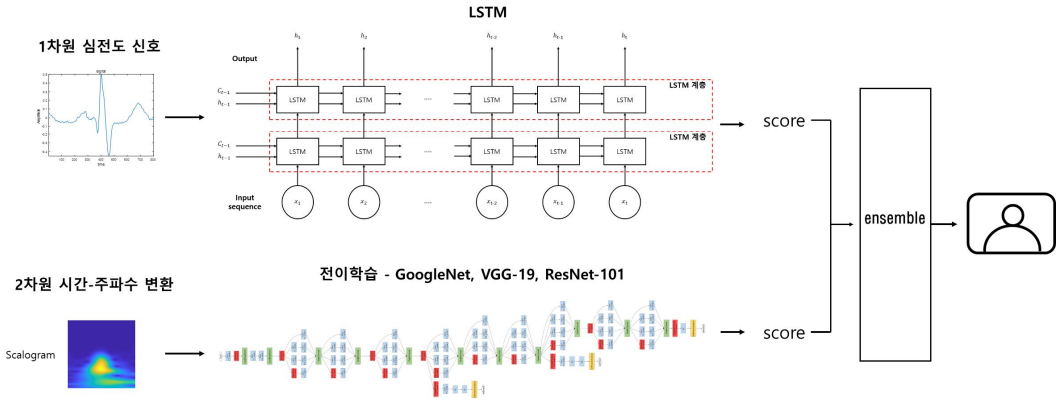


그림 3.28 제안된 앙상블기반 개인식별 순서도

조선대학교
CHOSUN UNIVERSITY

제4장 실험 및 결과분석

제1절 심전도 데이터베이스

본 논문에서는 심전도를 사용하여 개인 식별하기 위해 표 4-1과 같이 조선대학교에서 구축한 CU-ECG 데이터베이스와 PTB 공개 데이터베이스를 사용하였다.

표 4-1 심전도 데이터베이스

데이터베이스	피험자 수	데이터 설명
CU-ECG[35]	100	• 조선대학교에서 구축한 데이터 • 23세 ~ 34세의 남성 89명과 여성 11명 포함 • 심전도 lead-I 신호에서 의자에 편안하게 앉아있는 자세로 측정 • 1회에 10초로 기록되며 총 60회 취득 • 샘플링 속도 500kHz • 측정 장치 - 신호 오실로스코프 Keysight - MS09104와 마이크로컨트롤러 Atmega8 - 습식 전극 • 각 피험자의 데이터 60개씩 구성 • 여러 주기의 심전도 신호 → R-피크 포인트를 중심으로 일정하게 한 주기로 분할 → 분할된 신호는 약 1~2초로 구성 • 학습 데이터 80%, 검증 데이터 20% • 총 16,930개의 데이터 사용 - 학습 데이터 13,546개 - 테스트 데이터 3,384개

조선대학교
CHOSUN UNIVERSITY

PTB[36]	290	• 독일의 국립 계측 연구소 PTB에서 수집한 데이터 • 17세 ~ 87세의 남성 209명과 여성 81명 포함 • 다양한 심장 질환을 가지고 있는 피험자, 건강한 피험자의 신호 포함 • 12개의 표준 리드와 3개의 Frank 리드로 측정 • 샘플링 속도 1000 samples/s • 전체 290명 신호 중 R-피크 감지가 어려운 79명 제외하여 211명의 데이터 사용 • 211명의 클래스로 각 120개의 데이터로 구성 • 학습 데이터 60개, 검증 데이터 60개 사용

조선대학교
CHOSUN UNIVERSITY

그림 4.1은 CU-ECG 데이터베이스 측정을 위한 환경을 나타내고 그림 4.2는 CU-ECG 데이터베이스의 심전도 신호의 일부를 나타낸다.

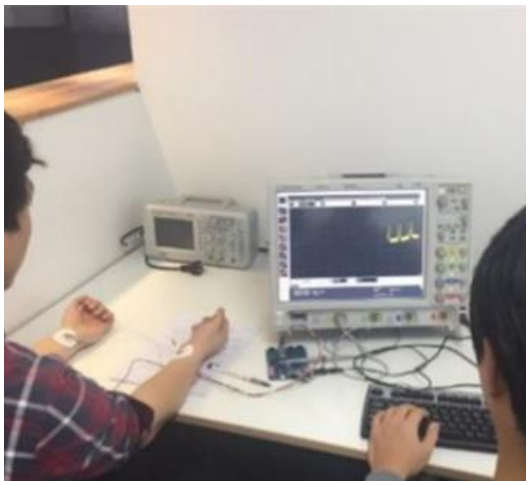


그림 4.1 CU-ECG 신호 측정을 위한 환경

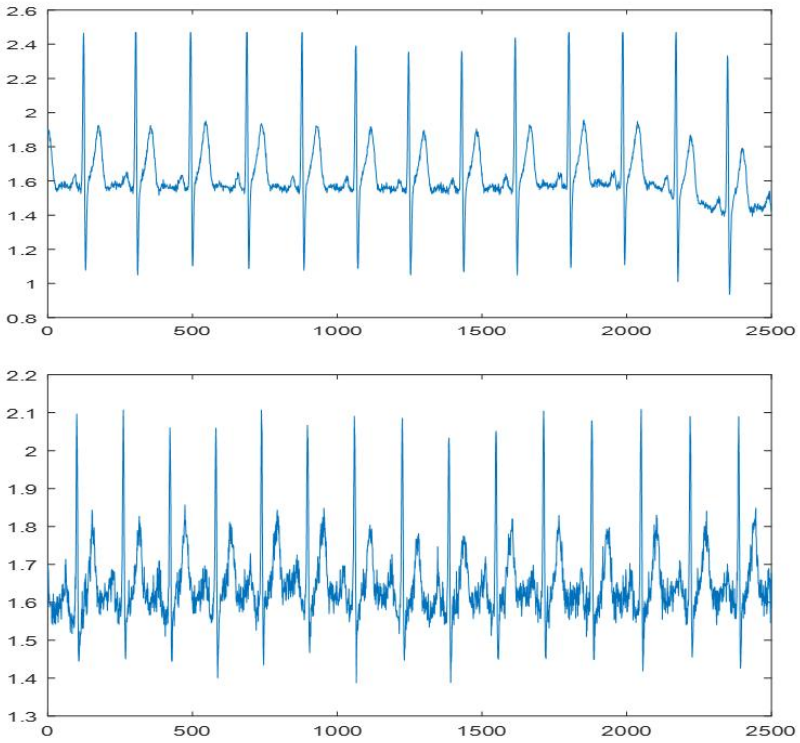


그림 4.2 CU-ECG 데이터베이스의 심전도 신호의 예시

조선대학교
CHOSUN UNIVERSITY

제2절 실험 및 결과분석

본 절에서는 심전도 신호로부터 LSTM과 2D-CNN을 이용하여 개인식별의 성능 결과를 분석한다. 1차원 심전도 신호를 분석하기 위해서 LSTM을 사용하고, 한 개의 LSTM 계층과 두 개의 LSTM 계층의 성능을 비교 분석한다. 또한 LSTM의 성능 향상을 위해 중요한 과정으로 그림 4.3은 심전도 신호의 잡음 제거 전의 성능을 보여주고, 그림 4.4는 잡음 제거 후의 성능으로 잡음 제거 전과 후의 성능을 비교 분석한다. 실험은 초기 학습률은 0.01, 예측 값과 실제 값의 오차 값을 최소화하기 위해 최적화 함수는 Adam을 사용하고, 에포크는 30, 50, 60, 100, 미니배치사이즈는 128로 반복하여 성능을 확인하였다. 또한 LSTM 계층에는 100개의 은닉 유닛을 사용한다. 잡음을 제거한 성능이 제거하지 않은 성능보다 높기 때문에 잡음을 제거한 심전도 신호를 사용하였다. CU-ECG 데이터베이스로부터 잡음 제거한 후의 LSTM기반 개인식별 정확도는 에포크 100, 미니배치사이즈를 128로 설정하였을 때 두 개의 LSTM 계층이 95.12%로 가장 높은 성능을 확인하였다.

조선대학교
CHOSUN UNIVERSITY

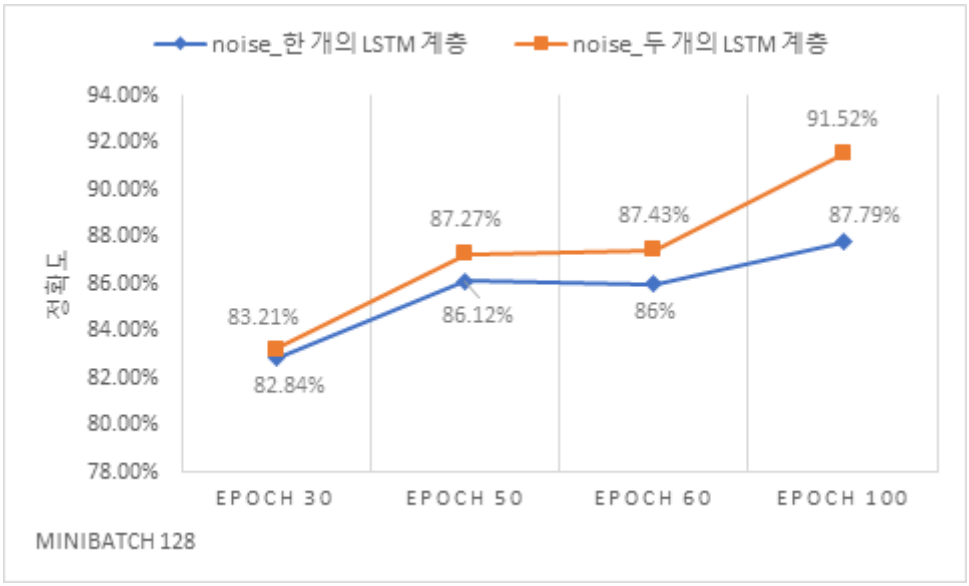


그림 4.3 CU-ECG의 LSTM기반 개인식별 정확도(잡음 제거 전)

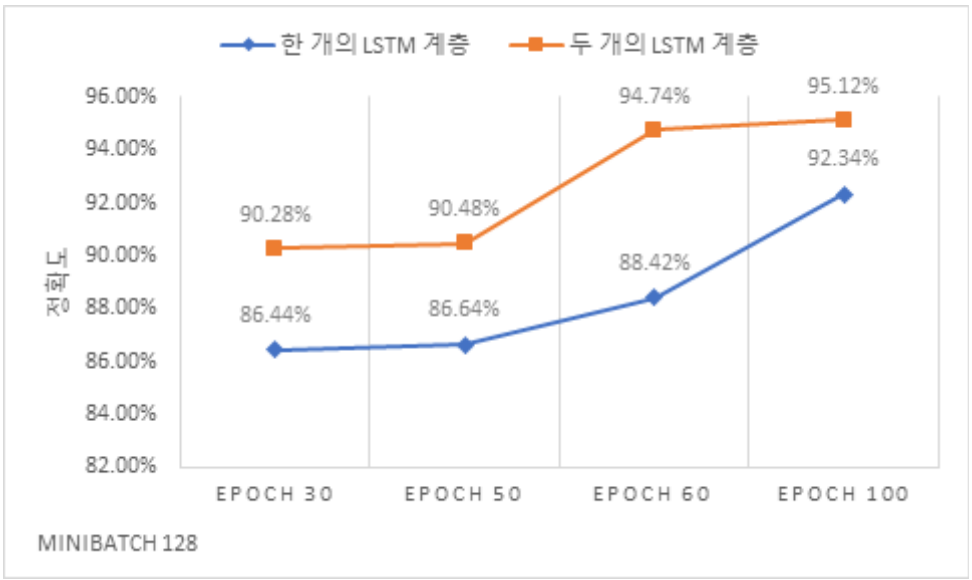


그림 4.4 CU-ECG의 LSTM기반 개인식별 정확도(잡음 제거 후)

조선대학교
CHOSUN UNIVERSITY

또한 PTB 데이터베이스도 CU-ECG 데이터베이스 실험과 동일하게 잡음 제거 후의 데이터를 사용하여 실험은 초기 학습률은 0.01, 최적화 함수는 Adam, 에포크는 30, 50, 60, 100, 미니배치사이즈는 128로 성능을 확인하였다. CU-ECG 데이터베이스, PTB 데이터베이스로부터 LSTM의 에포크와 반복 횟수의 값을 증가함에 따라 성능의 향상을 확인할 수 있다. 그림 4.5는 PTB 데이터베이스의 LSTM기반 개인식별 정확도를 보여주고, 정확도는 에포크 100, 미니배치사이즈 128로 설정하였을 때 92.34%로 가장 높은 성능을 확인하였다.

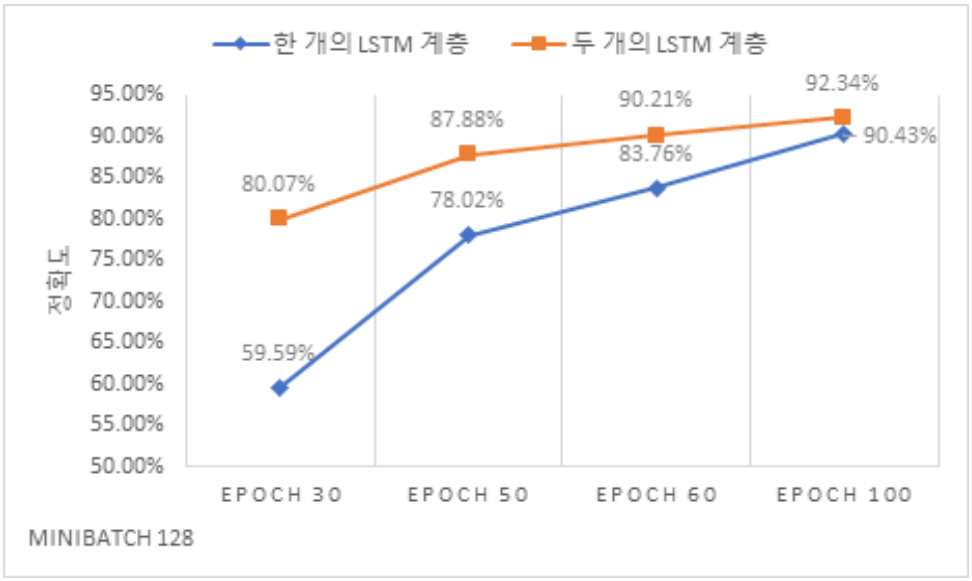


그림 4.5 PTB의 LSTM기반 개인식별 정확도

조선대학교
CHOSUN UNIVERSITY

심전도 신호와 같은 생리학적인 신호의 경우 다양한 잡음에 영향을 받기 때문에 이를 2차원 이미지로 변환시켜 CNN을 사용하여 개인 식별한다. 그림 4.6과 그림 4.7은 각각 CU-ECG와 PTB 데이터베이스에 포함된 임의의 피험자에 대해 STFT, Scalogram, FSST, WSST로 변환된 이미지를 보여주고 있다.

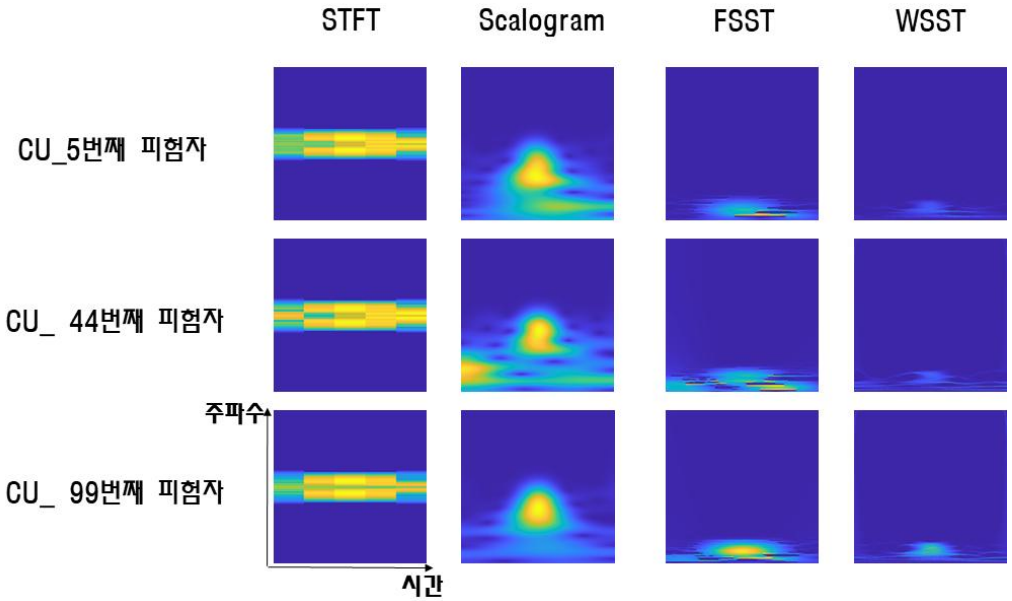


그림 4.6 CU-ECG 데이터베이스 신호의 피험자 시간-주파수 표현

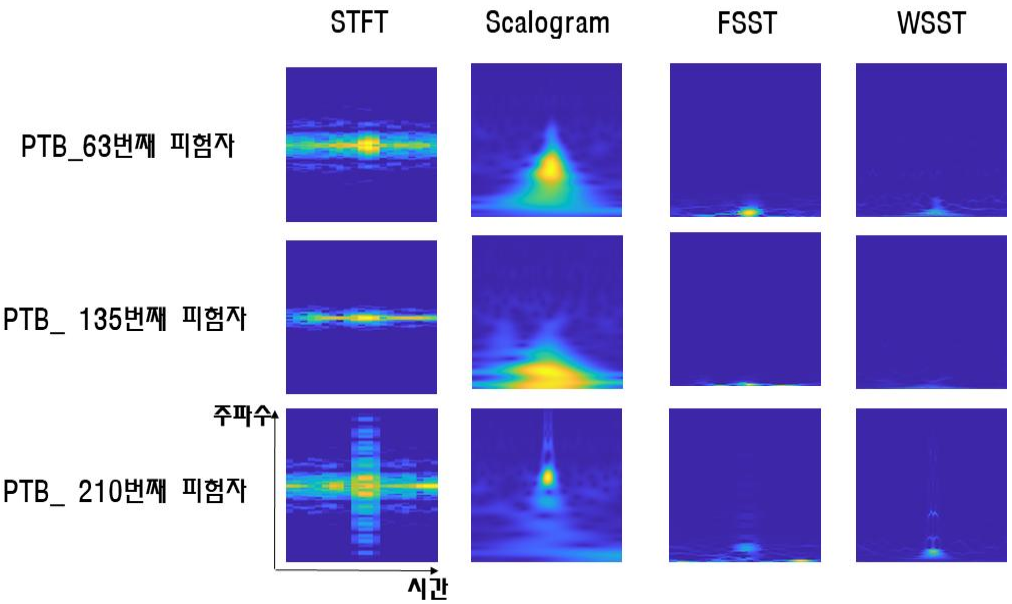


그림 4.7 PTB 데이터베이스 신호의 피험자 시간-주파수 표현

조선대학교
CHOSUN UNIVERSITY

변환된 이미지는 2D-CNN을 통해 개인식별 정확도를 확인한다. 2D-CNN은 GoogleNet, VGG-19, ResNet-101인 전이학습모델을 사용하였다. 데이터는 224*224으로 변경하여 입력으로 사용된다. CU-ECG 데이터베이스 실험을 위해 GoogleNet은 초기 학습률은 1e-4, 최적화 함수는 Adam, 에포크는 30, 미니배치사이즈는 64로 설정하였고, VGG-19는 초기 학습률은 1e-4, 최적화 함수는 Adam, 에포크는 20, 미니배치사이즈는 32로 설정하였다. ResNet-101은 VGG-19와 동일하게 진행하였다. 또한 단일 성능으로 LSTM의 결과와 CNN의 성능 결과를 확인한 것보다 제안된 앙상블을 통해 심전도를 이용한 개인식별의 정확도를 확인한다. 표 4.2는 CU-ECG 데이터베이스의 2D-CNN기반 개인식별 정확도와 LSTM과 2D-CNN기반 앙상블의 개인식별 정확도를 나타낸다. GoogleNet은 FSST에서 96.25%의 가장 높은 성능을 나타내며, VGG-19는 WSST에서 95.12%로 가장 높은 성능을 나타낸다. ResNet-101는 STFT에서 97.67%로 가장 높은 성능을 나타낸다. 시간-주파수 표현으로 나타낸 STFT, Scalogram, FSST, WSST을 전이학습 모델인 GoogleNet, VGG-19, ResNet-101을 통해 LSTM과 2D-CNN기반으로 앙상블한 개인식별 정확도를 확인한다. GoogleNet은 FSST에서 가장 높은 성능을 나타내며, 앙상블한 결과는 2.33%가 향상되었다. 또한 VGG-19는 WSST에서 가장 높은 성능으로, 단일 모델의 경우보다 3.4%가 향상되었고, ResNet-101는 STFT에서 가장 높은 성능으로, 단일 모델의 경우보다 1.06%가 향상되었다.

조선대학교
CHOSUN UNIVERSITY

표 4.2 2D-CNN의 개인식별 정확도와 LSTM과 2D-CNN기반 앙상블한 개인식별 정확도(CU-ECG DB)

CU-ECG	시간-주파수 표현	2D-CNN 검증 정확도	LSTM * 2D-CNN 앙상블 검증 정확도
GoogleNet	STFT	95.09%	98.37%
	Scalogram	95.69%	97.87%
	FSST	96.25%	98.58%
	WSST	95.48%	98.73%
VGG-19	STFT	94.86%	97.90%
	Scalogram	94.18%	97.93%
	FSST	94.52%	97.93%
	WSST	95.12%	98.52%
ResNet-101	STFT	97.67%	98.73%
	Scalogram	97.25%	98.73%
	FSST	96.04%	98.43%
	WSST	95.24%	98.49%

조선대학교
CHOSUN UNIVERSITY

PTB 데이터베이스 실험은 이전과 동일하게 GoogleNet, VGG-19, ResNet-101인 전이학습 모델을 사용하였다. 또한 초기 학습률, 최적화 함수, 에포크, 미니배치 사이즈도 동일하게 설정하여 진행하였다. 표 4.3은 PTB-ECG 데이터베이스의 2D-CNN기반 개인식별 정확도와 LSTM과 2D-CNN기반 앙상블의 개인식별 정확도를 나타낸다. GoogleNet은 STFT에서 98.04%의 가장 높은 성능을 나타내며, VGG-19는 Scalogram에서 96.99%로 가장 높은 성능을 나타낸다. ResNet-101는 FSST에서 98.47%로 가장 높은 성능을 나타낸다. GoogleNet은 STFT에서 가장 높은 성능을 나타내며, 앙상블한 결과는 1.56%가 향상되었다. 또한 VGG-19는 Scalogram에서 가장 높은 성능으로, 단일 모델의 경우보다 2.31%가 향상되었고, ResNet-101는 FSST에서 가장 높은 성능으로, 단일 모델의 경우보다 1.32%가 향상되었다.

조선대학교
CHOSUN UNIVERSITY

표 4.3 2D-CNN의 개인식별 정확도와 LSTM과 2D-CNN기반 앙상블한 개인식별 정확도 (PTB DB)

PTB-ECG	시간-주파수 표현	2D-CNN 검증 정확도	LSTM * 2D-CNN 앙상블 검증 정확도
GoogleNet	STFT	98.04%	99.60%
	Scalogram	97.13%	99.09%
	FSST	97.32%	99.65%
	WSST	97.50%	99.53%
VGG-19	STFT	96.48%	99.34%
	Scalogram	96.99%	99.30%
	FSST	95.89%	99.27%
	WSST	96.20%	99.55%
ResNet-101	STFT	98.44%	99.76%
	Scalogram	96.00%	99.23%
	FSST	98.47%	99.27%
	WSST	98.29%	99.77%

조선대학교
CHOSUN UNIVERSITY

표 4.4는 CU-ECG 데이터베이스의 기존 딥러닝기반 개인식별과 제안한 앙상블을 통해 개인식별을 비교한 결과를 보여준다. Y. H. Byeon[18]은 5개의 시간-주파수 표현 변환 방법을 사용하여 4개의 전이학습을 사용한 결과 스펙트로그램을 사용했을 때 94.03%로 가장 높은 성능을 확인했다. G. H. Choi[19]는 심전도 신호를 스펙트로그램으로 변환하고 이미지의 크기를 1/4만큼 축소하였을 때 96.40%로 가장 높은 개인식별 성능을 확인했다. 기존 논문과 비교 결과 제안된 LSTM과 2D-CNN의 앙상블 성능이 기존 딥러닝기반 개인식별 성능보다 향상된 것을 확인할 수 있다.

표 4.4 기존 딥러닝기반 개인식별의 성능 비교(CU-ECG DB)

저자	방법	피험자 수	정확도
Y. H. Byeon[18]	2D-CNN	100	94.03%
G. H. Choi[19]	2D-CNN	100	96.40%
제안된 LSTM과 2D-CNN의 앙상블		100	98.73%

조선대학교
CHOSUN UNIVERSITY

표 4.5는 PTB 데이터베이스의 기존 딥러닝기반 개인식별과 제안한 앙상블을 통해 개인식별을 비교한 결과를 보여준다. D. Jyotishi[20]는 제안한 LSTM을 사용하여 개인식별한 결과는 97.29%의 성능을 확인했고, J. S. Kim[21]는 2D 커플링 이미지를 2D-CNN을 통해 98.45%의 성능을 확인했다. M. Hammad[22]는 ECG기반 인증을 위해 1D-CNN을 제안한 결과 98.85%의 성능으로 기존 논문과 비교 결과 제안된 LSTM과 2D-CNN의 앙상블 성능이 기존 딥러닝기반 개인식별 성능보다 향상된 것을 확인할 수 있다.

표 4.5 기존 딥러닝기반 개인식별의 성능 비교(PTB DB)

저자	방법	피험자 수	정확도
D. Jyotishi[20]	LSTM	290	97.29%
J. S. Kim[21]	2D-CNN	52	98.45%
M. Hammad[22]	1D-CNN	290	98.85%
제안된 LSTM과 2D-CNN의 앙상블		211	99.77%

조선대학교
CHOSUN UNIVERSITY

5장 결론

본 논문은 심전도를 이용한 LSTM과 2D-CNN의 앙상블기반으로 개인식별을 진행했다. 심전도기반 개인식별은 사용자의 심전도 신호를 저장된 등록 사용자의 심전도와 비교하여 사용자를 인식, 인증하는 기술이다. 심전도는 사람마다 심장의 위치, 크기, 성별 및 나이에 따라 다르게 나타나는 고유한 신호를 나타내기 때문에 개인의 심전도 신호의 특성을 사용하여 개개인을 90% 이상 구별할 수 있다. 심전도는 개인식별뿐만 아니라 심장 관련 질병을 예측하고 진단하는 의료 분야에 가장 널리 사용되고 있다. 심전도 신호를 이용하여 개인을 식별하는 것 동시에 심정지, 부정맥 등과 같은 심장질환을 분석하여 원격으로 확인이 가능한 심전도기반으로 헬스 모니터링이 가능한 기술을 기대할 수 있다.

심전도 신호는 미세한 전류를 측정하는 생체 신호로 다양한 잡음들과 함께 나타난다. 이로 인해 정확한 판단이나 진단이 어렵기 때문에 왜곡된 신호들을 저역통과 필터, 고역 통과 필터, 평균 이동 필터를 통해 제거한다. 또한 변동된 기저선은 심전도 신호의 특징점을 검출하는 기준으로 기저선을 영점으로 맞추지 않으면 심전도 신호의 형태학적 특성을 파악하기 어렵다. 이를 위해 전처리 과정으로 잡음 제거와 기저선 변동 잡음을 진행했다. 잡음이 제거된 1D 심전도 신호를 LSTM 신경망에 적용하고, 1차원 심전도 신호를 시간-주파수 변환에 의해 사전에 학습된 모델을 사용했다. 각 모델의 출력 값들을 결합하여 앙상블기반으로 더 높은 성능을 기대할 수 있다. 실험을 진행하기 위해 사용된 데이터베이스는 두 개로 CU-ECG 데이터베이스와 PTB 데이터베이스를 사용했다. CU-ECG 데이터베이스는 조선대학교에서 구축한 데이터베이스로 남성 89명, 여성 11명으로 총 100의 피험자가 포함되어 있고, 공개 데이터베이스인 PTB 데이터베이스는 남성 209명, 여성 81명으로 총 290명의 피험자 중 R-피크 검출이 어려운 79명을 제외한 211명을 사용했다.

CU-ECG 데이터베이스와 PTB 데이터베이스는 LSTM 실험 결과로 epoch가 100, minibatchsize가 128로 설정하였을 때 95.12%, 92.34%로 가장 높은 성능을 확인하였다. 시간-주파수로 변환된 Scalogram, STFT, FSST, WSST는 2차원 이미지로 사전학습된 모델 GoogLeNet, VGG-19, ResNet-101을 사용하여 진행했다. 실험결과로 CU-ECG 데이터베이스는 GoogLeNet에서 FSST가 97.25%, VGG-19에서 WSST가 95.12%, ResNet-101에서 STFT가 97.67%로 높은 성능을 확인했다. LSTM 성능과 CNN의 성능

조선대학교
CHOSUN UNIVERSITY

을 앙상블 모델을 통해 최소 1.06%에서 최대 3.75%로 성능이 향상된 것을 확인했다. PTB 데이터베이스는 GoogLeNet에서 STFT가 98.04%, VGG-19에서 Scalogram이 96.99%, ResNet-101에서 FSST가 98.47%로 높은 성능을 확인했다. 또한 각 모델의 출력인 score 값들을 곱셈을 통해 최대 값을 사용하는 앙상블 방법으로 장단기 메모리 성능과 합성곱 신경망의 성능을 앙상블 모델을 통해 최소 0.8%에서 최대 3.38%로 성능이 개선된 것을 확인했다. 향후 심전도를 이용한 개인식별뿐만 아니라 감정에 따라 개개인의 심전도 신호가 달라지기 때문에 심전도기반 감정분류 방법을 연구할 계획이다.

조선대학교
CHOSUN UNIVERSITY

참고문헌

[1] D. Jeong, “Artificial intelligence security threat, crime, and forensics: Taxonomy and open issues”, IEEE Access, Vol. 8, pp. 184560-184574, 2020.

[2] 김진우, 이필규, “적응적 딥러닝 학습 기반 영상 인식”. 한국인터넷방송통신학회 논문지, Vol. 18, No. 1, pp. 113-117, 2018.

[3] 윤일영, “바이오와 보안의 융합 생체인식 기술”, 융합연구정책센터, 2018.

[4] 김도현, “[이슈분석]최근 생체인식 산업 동향과 시사”, 정보통신기획평가원, 2021.

[5] A. Barros, D. Rosário, P. Resque, E. Cerqueira, “Heart of IoT: ECG as biometric sign for authentication and identification”, 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), pp. 307-312, 2019.

[6] M. Ingale, R. Cordeiro, S. Thentu, Y. Park, N. Karimian, “ECG biometric authentication: A comparative analysis”, IEEE Access, Vol. 8, pp. 117853-117866, 2020.

[7] Q. Zhang, D. Zhou, X. Zeng, “HeartID: A multiresolution convolutional neural network for ECG-Based biometric human identification in smart health applications”, IEEE Access, Vol. 5, pp. 11805-11816, 2017.

[8] L. Wieclaw, Y. Khoma, P. Fałat, D. Sabodashko, V. Herasymenko, “Biometrie identification from raw ECG signal using deep learning techniques”, 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), pp. 129-133, 2017.

[9] J. Y. Lee, “The motion artifact reduction using periodic moving average Filter”, Journal of the Korea Society of Computer and Information, Vol. 17, No. 4, pp. 75-82, 2012.

조선대학교
CHOSUN UNIVERSITY

[10] R. Vullings, B. D. Vries, J. W. M. Bergmans, “An adaptive kalman filter for ECG signal enhancement”, IEEE Transactions on Biomedical Engineering, Vol. 58, No. 4, pp. 1094-1103, 2011.

[11] L. Biel, O. Pettersson, L. Philipson, P. Wide, “ECG analysis: a new approach in human identification”, IEEE Transactions on Instrumentation and Measurement, Vol. 50, No. 3, pp. 808-812, 2001.

[12] S. A. Israel, J. M.Irvine, A. Cheng, M. D. Wiederhold, B. K. Wiederhold, “ECG to identify individuals”, Pattern Recognition, Vol. 38, No. 1, pp. 113-142, 2005.

[13] M. Jahiruzzaman, A. B. M. A. Hossain, “ECG based biometric human identification using chaotic encryption”, 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), pp. 1-5, 2015.

[14] Z. Zhao, L. Yang, D. Chen, Y. Luo, “A human ECG identification system based on ensemble empirical mode decomposition”, Sensors, Vol. 13, No. 5, pp. 6832-6864, 2013.

[15] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, F. Scotti, “Deep-ECG: Convolutional neural networks for ECG biometric recognition”, Pattern Recognition Letters, Vol. 126, pp 78-85, 2019.

[16] S. S. Abdeldayem, T. Bourlai, “A novel approach for ECG-Based human identification using spectral correlation and deep learning”, IEEE Transactions on Biometrics, Behavior, and Identity Science, Vol. 2, No. 1, pp. 1-14, 2020.

[17] I. B. Ciocoiu, N. Cleju, “Off-Person ECG biometrics using spatial representations and convolutional neural networks”, IEEE Access, Vol. 8, pp. 218966-218981, 2020.

[18] Y. H. Byeon, K. C. Kwak, “Pre-Configured Deep Convolutional Neural Networks with Various Time-Frequency Representations for Biometrics from ECG Signals”, Applied Sciences, Vol. 9, No. 22, 2076-3417, 2019

[19] G. H. Choi, E. S. Bak, S. B. Pan, "User Identification System Using

조선대학교
CHOSUN UNIVERSITY

2D Resized Spectrogram Features of ECG," IEEE Access, Vol. 7, 34862-34873, 2019.

[20] D. Jyotishi, S. Dandapat, "An LSTM-Based Model for Person Identification Using ECG Signal", IEEE Sensors Letters, Vol. 4, No. 8, 1-4, 2020.

[21] J. S. Kim, S. G. Kim, S. B. Pan, "Personal recognition using convolutional nearal network with ECG coupling image", Journal of Ambient Intelligence and Humanized Computing, Vol. 11, 1923-1932, 2020.

[22] M. Hammad, P. Pławiak, K. Wang, U. R. Acharya, "ResNet-Attention model for human authen ticationusing ECG signals", Expert Systems, Vol. 38, No. 6, 2021.

[23] Y. H. Byeon, S. B. Pan, K. C. Kwak, "Intelligent deep models based on scalograms of electrocardiogram signals for biometrics", Sensors, Vol. 19, No. 4, 2019.

[24] S. Chikkerur, A. N. Cartwright, V. Govindaraju, "Fingerprint enhancement using STFT analysis", Pattern Recognition, Vol. 40, No. 1, pp. 198-211, 2007.

[25] K. El-Shennawy, "Communication theory and signal processing for transform coding", Bentham Science Publishers, 2014.

[26] J. W. Lee, H. W. Lee, C. S. Yoo, "Selection of mother wavelet for bivariate wavelet analysis", Journal of Korea Water Resources Association, Vol. 52, No. 11, pp. 905-916, 2019.

[27] S. C. Olhede, A. T. Walden, "Generalized morse wavelets", IEEE Transactions on Signal Processing, Vol. 50, No. 11, pp. 2661-2670, 2002.

[28] J. M. Lilly, S. C. Olhede, "Higher-order properties of analytic wavelets", IEEE Transactions on Signal Processing, Vol. 57, No. 1, pp. 146-160, 2009.

[29] T. Oberlin, S. Meignen, V. Perrier, "The fourier-based synchrosqueezing transform", 2014 IEEE International Conference on

조선대학교
CHOSUN UNIVERSITY

Acoustics, Speech and Signal Processing (ICASSP), pp. 315-319, 2014.

[30] A. Kumar, C. P. Gandhi, Y. Zhou, G. Vashishtha, R. Kumar, J. Xiang, "Improved CNN for the diagnosis of engine defects of 2-wheeler vehicle using wavelet synchro-squeezed transform (WSST)", Knowledge-Based Systems, Vol. 208, 2020.

[31] A. Lumini, L. Nanni, "Deep learning and transfer learning features for plankton classification", Ecological Informatics, Vol. 51, pp. 33-43, 2019.

[32] N. Habib, M. M. Hasan, M. M. Reza, M. M. Rahman, "Ensemble of cheXNet and VGG-19 feature extractor with random forest classifier for pediatric pneumonia detection", SN Computer Science, Vol. 1, No. 6, 2020.

[33] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition", 2016 IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.

[34] G. Haralabopoulos, I. Anagnostopoulos, D. McAuley, "Ensemble deep learning for multilabel binary classification of user-generated content", Algorithms, Vol. 13, No. 4, 2020.

[35] Y. H. Byeon, J. N. Lee, S. B. Pan, K. C. Kwak, "Multilinear eigenECGs and fisherECGs for individual identification from information obtained by ad electrocardiogram sensor", Symmetry, Vol. 10, No. 10, 2018.

[36] R. Bousseljot, D. Kreiseler, A. Schnabel, "Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet", Biomedizinische Technik, pp. 317, 1995.