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

[UCI]I804:24011-200000267159

2019년 2월

석사학위논문

딥 컨볼루션 뉴럴 네트워크 기반 심전도를 이용한 사용자 인식에 관한 연구

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 성 혁

조선대학교
CHOSUN UNIVERSITY

딥 컨볼루션 뉴럴 네트워크 기반 심전도를 이용한 사용자 인식에 관한 연구

A Study on User Identification using ECG based on Deep Convolutional Neural Network

2019년 2월 25일

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 성 혁

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

딥 컨볼루션 뉴럴 네트워크 기반 심전도를 이용한 사용자 인식에 관한 연구

지도교수 반 성 범

이 논문을 공학석사학위신청 논문으로 제출함.

2018년 10월

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 성 혁

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

김성혁의 공학석사학위논문을 인준함

위원장 조선대학교 교수 곽근창 (인)

위 원 조선대학교 교수 김판구 (인)

위 원 조선대학교 교수 반성범 (인)

2018년 12월

조선대학교 산업기술융합대학원

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

목 차

제1장 서론 ····· 1
제2장 딥 러닝 연구 ····· 4
제1절 딥 러닝 기술 및 구조 ····· 4
제2절 딥 러닝 기반 공개 네트워크 ····· 10
제3장 심전도를 이용한 사용자 인식 ····· 18
제1절 기준점 기반 사용자 인식 ····· 21
제2절 비 기준점 기반 사용자 인식 ····· 23
제4장 제안하는 DCNN 기반 심전도를 이용한 사용자 인식 시스템 ····· 25
제1절 심전도 신호 전처리 및 학습 데이터 변환 ····· 25
제2절 DCNN 네트워크 설계 ····· 28
제3절 실험결과 ····· 31
제5장 결론 ····· 35
참고문헌 ····· 36

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 목차

표 1-1. 개인 식별 요구사항 ····· 2
표 2-1. 공개 네트워크 기존 성능 ····· 16
표 3-1. SIEMENS Megacart에 의해 추출된 특징 ····· 21

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

그림 목차

그림 1.1 국내·외 바이오인식 시장 규모 전망 ····· 1
그림 1.2 생체신호 취득 방법 ····· 2
그림 2.1 심층 신경망 구조 ····· 4
그림 2.2 딥 러닝 및 머신러닝의 데이터 처리 방법 ····· 5
그림 2.3 신경세포 및 인공 신경망의 구조 ····· 6
그림 2.4 F. Rosenblatt의 퍼셉트론 구조 ····· 6
그림 2.5 AND, OR, XOR 연산 ····· 7
그림 2.6 MLP를 이용한 XOR 연산 ····· 7
그림 2.7 CNN 기본 구조 ····· 8
그림 2.8 컨볼루션 계층 & 맥스 풀링 과정 ····· 9
그림 2.9 ILSVRC 우승팀의 성능 ····· 10
그림 2.10 Inception-V1 구조 ····· 11
그림 2.11 Inception-V2, V3 layer 특성 ····· 12
그림 2.12 Inception-V2, V3 구조 ····· 12
그림 2.13 Inception-ResNet 메인구조 ····· 13
그림 2.14 ResNet구조의 Residual Learning ····· 14
그림 2.15 텐서플로우 그래프 구조 ····· 15
그림 2.16 기존 공개 네트워크 성능 및 실험성능 비교 ····· 17
그림 3.1 심전도 신호의 형태 ····· 18
그림 3.2 표준 사지유도 lead-I, II, III 구조 ····· 19
그림 3.3 심전도를 이용한 사용자 인식 시스템 ····· 19
그림 3.4 심전도 신호의 잡음 종류 ····· 20
그림 3.5 S. A. Israel 연구에 사용한 기준점 ····· 22
그림 3.6 G. Wubbeler 연구에 사용한 특징 추출 ····· 23
그림 3.7 S. Fatemian 연구에 사용한 특징 추출 ····· 24

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

그림 4.1 제안하는 DCNN 기반 심전도를 이용한 사용자 인식 시스템 ·····26
그림 4.2 심전도 신호 전처리 과정 ·····26
그림 4.3 2차원 영상 변환 과정 ·····27
그림 4.4 커플링 매트릭스 생성 과정 ·····27
그림 4.5 제안하는 DCNN 구조 ·····28
그림 4.6 경사 감소 학습 시 문제점 ·····29
그림 4.7 ReLU 활성화 함수 ·····29
그림 4.8 정규화의 필요성 ·····30
그림 4.9 MIT-BIH Normal Sinus Rhythm Databse의 lead-I 신호 ·····31
그림 4.10 커플링 매트릭스 네트워크 성능(18명) ·····32
그림 4.11 제안하는 구조 및 공개 네트워크 성능 비교(18명) ·····33
그림 4.12 커플링 매트릭스 네트워크 성능(47명) ·····34
그림 4.13 제안하는 구조 및 공개 네트워크 성능 비교(47명) ·····34

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

ABSTRACT

A Study on User Identification using ECG based on Deep Convolutional Neural Network

Kim, Sung Hyuck
Advisor : Prof. Pan, Sung Bum
Department of Software Convergence Engineering
Graduate School of Industry Technology Convergence, Chosun University

Biometrics technology is a technology that compares and judges biometrics information and signals presented using unique biometrics information and signals possessed by individuals. However, existing biometric technologies are exposed to various crimes. As a method to solve this problem, many researches are actively carried out by combining the electrocardiogram and the deep running technique resistant to forgery and modulation to the user recognition.

In this thesis, we propose DCNN-based ECG for user identification. The noise of the electrocardiogram signal is removed by using a bandpass filter and an moving average filter, and the signal is divided based on the R peak point. The input data of the DCNN is generated through the process of

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

preprocessing and converting the 2-D image using the divided 1-D electrocardiogram signal. The structure of DCNN is designed on the basis of VGGNet. For DCNN network design, weight initialization, activation function, normalization, cost and optimizer function are optimized for ECG data set.

In order to compare the performance of the proposed network structure, we compare the performance according to the input data, the performance comparison with the existing network, and the performance according to the number of data classes.

Experimental results showed that the performance of the proposed coupling matrix was 98.35% when using 18 ECG data and 97% when using 47 data. It is confirmed that this increase is 1.9% and 2.65%, respectively, when using 2D image.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제1장 서론

현대 사회에서 개인 신분을 식별하는 상황은 빈번히 발생한다. 기존 개인의 신분을 확인하는 방법은 신분증과 같은 소유기반 식별방법과 ID/PW 등의 지식기반 식별 방법이 있다. 그러나 기존 개인 신분 식별 방법은 도난 및 해킹, 개인정보 유출의 위험성이 있어 한계가 있다. 이에 따라 기존의 개인 신분 식별방법을 보완하기 위해 바이오인식 기술의 필요성이 증가하고 있다. 미국 시장조사업체에 따르면 바이오인식 시장의 규모가 2019년에는 150억 달러 규모에 이를 것이라는 전망을 내놨으며, 국내 바이오인식 규모 또한 증가하는 추세를 보이고 있다[1].

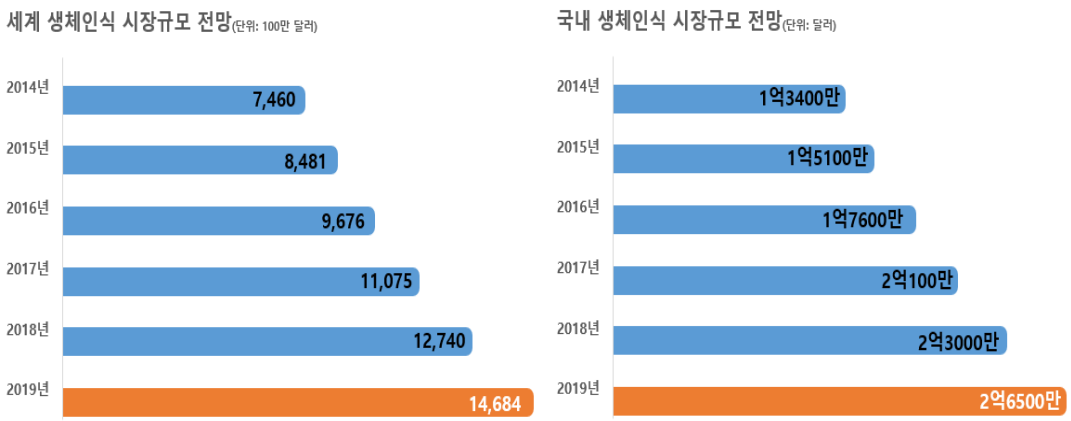


그림 1.1 국내 · 외 바이오인식 시장규모 전망

바이오인식 기술이란 개인의 고유한 생체정보 및 신호를 이용하여 등록, 저장한 후 제시한 생체정보 및 신호와 비교, 판단하는 기술이다[2-3]. 생체정보는 행동적 특징과 신체적 특징으로 나눠진다. 행동적 특징으로 얻을 수 있는 생체정보는 걸음걸이, 서명, 음성 등이 있으며, 신체적 특징으로 얻을 수 있는 생체정보는 지문, 홍채, 얼굴, 망막, 정맥 등이 있다. 현재 대중적으로 사용하고 있는 생체정보는 지문, 얼굴, 홍채로 표 1.1의 요구사항을 만족한다[4]. 그러나 이와 같은 생체정보를 이용한 신분 확인은 위조 지문을 이용한 금융사고, 얼굴 영상을 이용한 해킹 등과 같은 다양한 범죄의 위협에 노출되어 있다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 1-1 개인 식별 요구사항

특성	내용
보편성	누구나 갖고 있는 특성인가?
유일성	각 사람마다 구별할 수 있는 특성인가?
영구성	변하지 않고 변화시킬 수 없는 특성인가?
획득성	정량적으로 계측이 가능한 특성인가?
정확성	환경변화와 무관하게 높은 인식 성능을 지니고 있는 특성인가?
수용성	사용자의 거부감은 없는 특성인가?
기만성	우회적인 부정사용으로부터 안전한 특성인가?

이에 따라 주요 선진국은 위조 및 변조에 강한 생체신호를 활용한 연구에 대해 활발히 진행 중이다[5]. 생체신호의 종류로는 근전도, 뇌전도, 심전도 등이 있다. 근전도는 근육이 수축하거나 근육 활동에 의한 전기적인 변화를 기록한 신호로써 동작 인식, 질병 진단 등에 활용이 가능하다. 근전도를 취득하는 방법은 그림 1.1 (a)와 같으며, 각각의 근육에 신호 취득 장비를 부착하여 신호를 취득하기 때문에 불편한 점이 있다. 뇌전도는 신경계에서 뇌신경 사이에 신호가 전달 될 때 생기는 전기의 흐름으로 의도식별, 상황분석 등에 활용이 가능하다. 그러나 뇌전도는 그림 1.1 (b)와 같이 머리에 취득 장비를 착용하기 때문에 머리카락이 신호를 취득하는 데 방해가 된다. 심전도는 심장의 박동과 관련된 전위변화를 전압의 형태로 기록한 신호로 심장의 크기와 위치, 심장의 손상 여부 등 개인별로 고유한 특성이 있어 개인 식별, 질병 진단 등에 활용이 가능한 장점이 있다. 심전도는 그림 1.1 (c)와 같이 양쪽의 손목에서 취득이 가능하기 때문에 시계 및 팔찌의 소형 액세서리 형태로 대체가 가능하다[6-8].

(a) 근전도 측정 (b) 뇌전도 측정 (c) 심전도 측정

그림 1.2 생체신호 취득 방법

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

위조 및 변조에 강한 심전도를 이용한 사용자 인식에 관한 기존 연구는 신호의 잡음을 제거하는 전처리 과정과 특징 추출의 방법론에 대해 많은 연구가 진행되었다. 기존 심전도 신호를 이용한 사용자 인식 시스템은 사람이 특징 추출 알고리즘을 이용해 특징을 추출하는 수동적인 방법이었다. 그러나 심전도의 특징점을 잘못 정의하면 인식 성능을 떨어뜨리는 단점이 있다. 최근 딥 러닝 기반의 CNN(Convolutional Neural Network)은 데이터에서 특징을 자동으로 추출하며 기존 방법들에 비해 뛰어난 성능을 보인다.

본 논문에서는 DCNN(Deep Convolutional Neural Network)기반 심전도를 이용한 사용자 인식 시스템을 제안한다. 제안하는 DCNN 네트워크 설계 전, 기존 공개 네트워크인 Inception, ResNet 네트워크의 성능을 분석한다. 기존 공개 네트워크에 사용한 이미지넷 데이터셋을 사용하며, Inception 네트워크의 초기 버전인 Inception-V1의 성능은 86%의 성능을 나타냈으며, 버전이 올라갈수록 기존 성능과 같이 점차 증가함을 확인 할 수 있었다.

제안하는 DCNN 구조는 기존 공개 네트워크 보다 layer층이 깊지 않고 직관적으로 확인 할 수 있는 VGG 네트워크 기반으로 설계하였으며, 기존 공개 네트워크인 Inception 네트워크에서 적용한 작은 크기의 컨볼루션 필터를 사용한다. 전체적인 구조는 컨볼루션, 풀링의 형태로 되어있으며, 각각의 컨볼루션 연산후 정규화 과정을 진행하여 오버피팅을 방지한다. 실험에 사용한 데이터는 MIT-BIH Normal Sinus Rhythm Database, 18명, MIT-BIH Arrhythmia Database, 47명을 사용한다. 입력 데이터로 커플링 매트릭스를 사용하였을 때 각각 98.35%, 97%의 성능을 나타내며 2차원 영상을 사용하였을 때 보다 1.9%, 2.65% 증가함을 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서는 딥 러닝의 기술과 공개 네트워크 특징에 대해 설명하며, 기존 공개 네트워크의 성능 분석을 위해 직접 실험하여 비교한다. 3장에서는 심전도를 이용한 사용자 식별 시스템 관련 연구에 대해 설명하며, 4장에서는 제안하는 DCNN 기반 심전도를 이용한 사용자 인식 시스템에 대해 설명한다. 5장에서는 결론으로 논문을 마무리한다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제2장 딥 러닝 연구

제1절 딥 러닝 기술 및 구조

딥 러닝이란 컴퓨터가 스스로 사람처럼 학습할 수 있는 인공지능 기술로 이미지 분류, 음성 인식, 자연어 처리, 의학 분야 등에서 널리 사용되고 있다[9]. 딥 러닝은 계층 구조(layer structure)로 구성되면서 입력층과 출력층 사이에 여러 개의 은닉층(hidden layer)을 가지고 있는 심층 신경망(Deep Neural Network)이다.

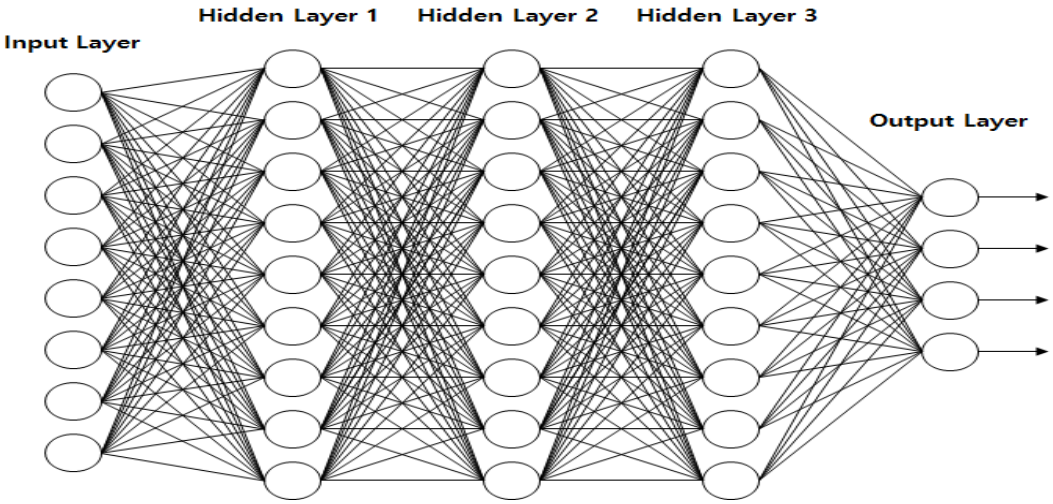


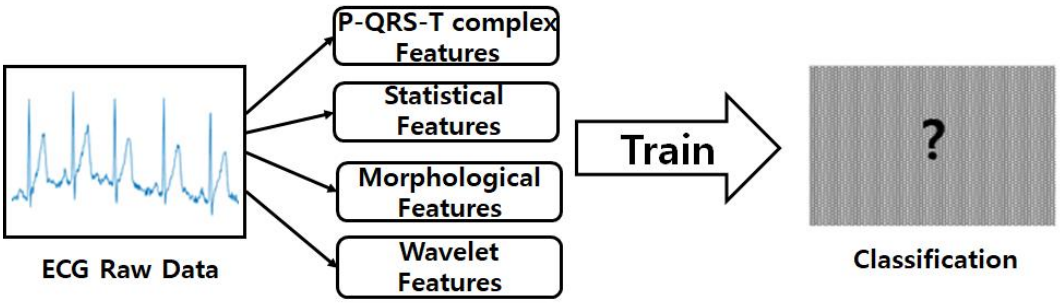
그림 2.1 심층 신경망 구조

딥 러닝은 머신러닝의 일종으로 머신러닝의 용어는 1959년 A. Samuel의 연구에서 처음 등장하였다. A. Samuel은 머신러닝의 개념을 명시적으로 프로그램을 작성하지 않고 컴퓨터에 학습할 수 있는 능력을 부여하기 위한 연구 분야라고 정의하였다[10]. 1997년 T. Mitchell은 머신러닝을 컴퓨터 프로그램이 특정한 작업 T를 수행할 때 성능 P만큼 개선되는 경험 E를 보이면, 그 컴퓨터 프로그램은 작업 T와 성능 P에 대해 경험 E를 학습했다로 머신러닝 개념을 재정의하였다[11]. 즉 기계가 특정 작업을 좀 더 정확하게 수행하도록 학습을 시키는 방법을 머신러닝이라고 정의한다.

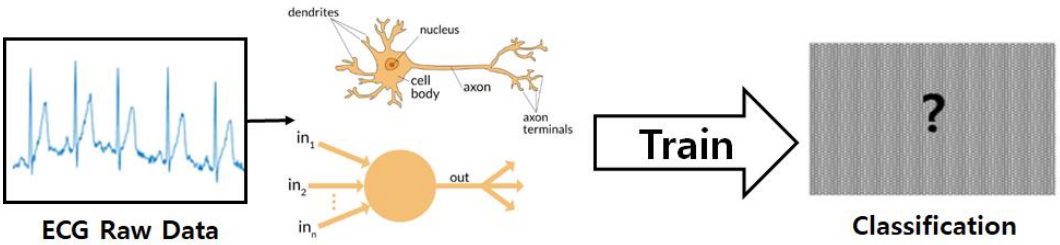
Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

머신러닝과 딥 러닝은 컴퓨터에서 데이터를 처리하려는 목적은 같지만 접근하는 방식에서 차이가 있다. 그림 2.2와 같이 머신러닝은 학습시키고자 하는 데이터의 특징을 사람이 직접 추출하여 학습하지만, 딥 러닝은 인간의 뇌에서 일어나는 의사결정 과정을 모방한 인공 신경망 구조를 통해 데이터 자체를 컴퓨터가 학습한다.



(a) 머신러닝을 통한 심전도 분류



(b) 딥 러닝을 통한 심전도 분류

그림 2.2 딥 러닝 및 머신러닝의 데이터 처리 방법

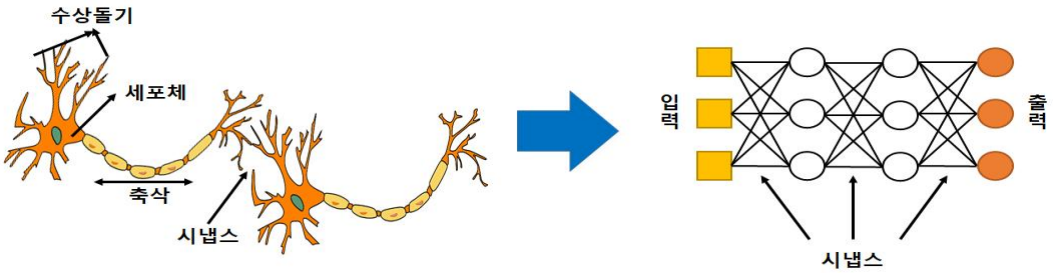
딥 러닝에서 사용되는 인공 신경망은 1943년 W. S. McCulloch에 의해 처음 등장하였다[12]. 생물학적 신경망을 단순화하여 논리, 산술, 기호 연산을 구현할 수 있게 하였으며, TLU(Threshold Logic Unit)로 정의하였다. 즉 인공 신경망은 인간의 뇌 구조를 모방하여 수학적 모델로 나타낸 방식이라고 할 수 있다.

생물학적 신경망과 인공 신경망의 구조는 그림 2.3과 같다. 생물학적 신경망에서 수상돌기는 인공 신경망에서 입력이며, 축삭은 출력, 세포체는 노드의 역할을 하게 된다. 신경세포의 입력 값은 다수이고 다수의 신경세포가 연결되어있는 부분

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

을 인공 신경망에서 시냅스라고 부른다. 여러 신경세포로부터 전달되어 온 신호들은 합산되어 하나의 출력 값을 만들며, 합산된 값이 임계 값 이상이면 출력 신호가 생기고 이하이면 출력 신호가 생기지 않는다.



(a) 신경세포 (b) 인공 신경망

그림 2.3 신경세포 및 인공 신경망의 구조

인공 신경망의 등장 이후 1958년 F. Rosenblatt의 연구에 의해 퍼셉트론(perceptron)알고리즘이 등장하였다. 퍼셉트론은 인공 신경망의 한 종류로서 다수의 신호를 입력으로 받아 하나의 신호를 출력한다. 그림 2.4와 같이 입력(input) 신호가 노드에 보내질 때 각각 고유한 가중치(weight)를 곱한 후, 모두 합한 뒤 활성화 함수를 적용해서 그 값이 0보다 크면 1(뉴런 활성화), 0보다 작으면 -1(뉴런 비활성화)을 출력하는 선형 분류기의 구조 형태이다[13].

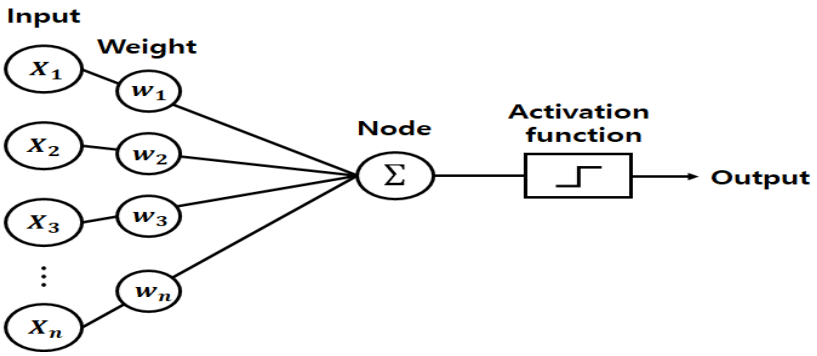
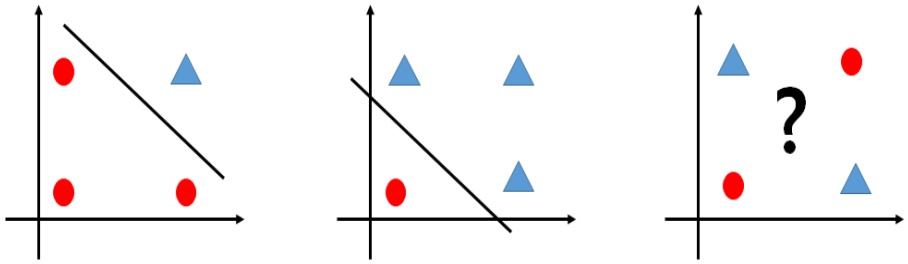


그림 2.4 F. Rosenblatt의 퍼셉트론 구조

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

퍼셉트론의 등장으로 인해 딥 러닝의 연구가 더욱더 가속화될 것이라고 예상하였지만, 이런 기대와 달리 1969년 M. Minsky, A. Seymour가 발표한 책에서 퍼셉트론의 한계를 수학적으로 증명함으로써 문제에 직면하게 된다[14]. 당시 퍼셉트론을 이용해 AND, OR 연산과 같은 선형적인 문제해결이 가능하였지만, XOR 연산과 같은 비선형적인 문제 해결이 불가능한 단점이 한계로 지적되었다.



(a) AND (b) OR (c) XOR

그림 2.5 AND, OR, XOR 연산

기존 퍼셉트론이 선형 분류기라는 한계에 의해 XOR 문제를 해결할 수 없었지만 이후 1986년 MLP(Multi-Layer Perceptron)와 back-propagation의 등장으로 인해 해결되었다. 그림 2.6과 같이 MLP에 은닉층이라는 중간 layer를 추가하여 선형 분류 판별 선을 여러 개 그리는 효과를 얻음으로써 기존의 XOR 문제를 해결하였다.

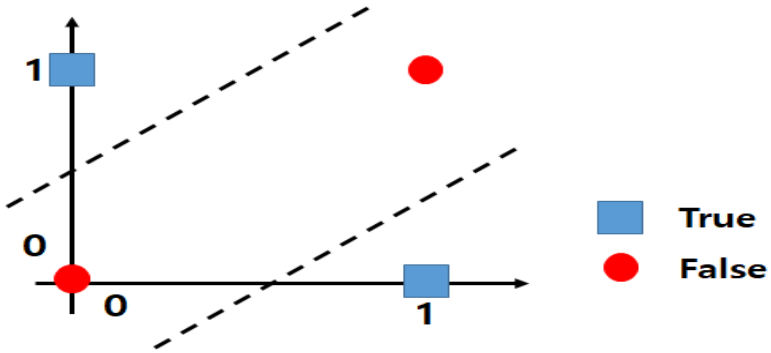


그림 2.6 MLP를 이용한 XOR 연산

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

그러나 MLP는 파라미터의 개수가 많아지면 학습을 통해서 구하고자 하는 최적의 가중치와 바이어스의 값을 찾아내기 어려운 문제점이 있었다. 이러한 문제를 해결하기 위해 back-propagation을 사용하며, 학습은 feed-forward와 back-propagation의 단계로 구성된다. Feed-forward 단계에서 학습 데이터는 신경망을 통해 학습한 후 출력 단에서 error와 cost function 값을 구한다. 신경망이 충분히 학습되지 못한 경우 오차가 크게 발생한다. Back-propagation 단계에서는 출력부터 반대 방향으로 순차적으로 편미분을 수행해가면서 가중치와 바이어스 값을 갱신한다. 학습 데이터에 대해서 반복적으로 이 과정을 거치며, 가중치와 바이어스는 학습 데이터에 최적화된 값으로 바뀌게 된다[15].

1990년 Y. Lecun은 MLP와 back-propagation을 컨볼루션과 결합하여 필기숫자(MNIST)를 인식하는 논문을 발표하였으며[16], 1998년 Y. Lecun은 LeNet-5라는 CNN의 시초가 되는 신경망 구조를 제안하며 인공 신경망의 기대를 높였다[17].

CNN은 딥 러닝의 한 종류로 신경망으로 구성되며, 이미지 인식 분야에서 많이 사용되고 있다. Y. Lecun이 1989년 발표한 논문에서 처음 소개되었으며, 사람의 필기체를 인식하기 위한 방법으로 CNN을 적용시켰다[18]. 이후 컴퓨터 성능 및 빅데이터 처리의 중요성으로 인해 CNN을 활용한 연구 분야가 증가하였다. CNN의 기본적인 구조는 크게 컨볼루션 계층, 풀링 계층, 완전 연결 계층으로 나뉘며 그림 2.7과 같다.

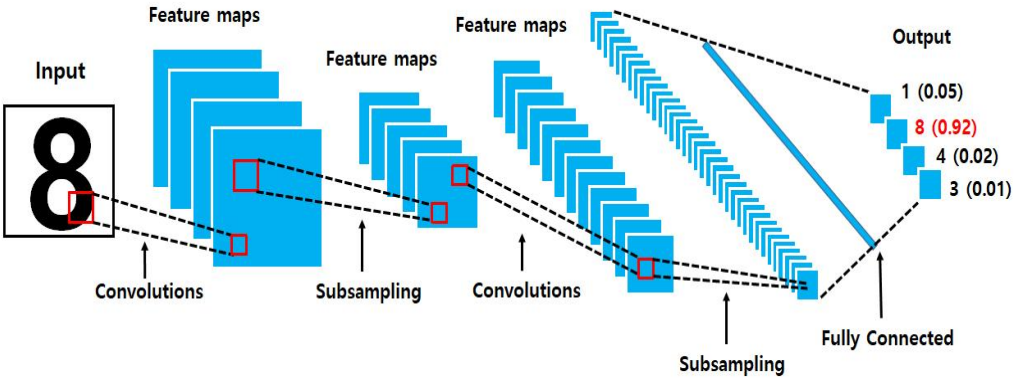
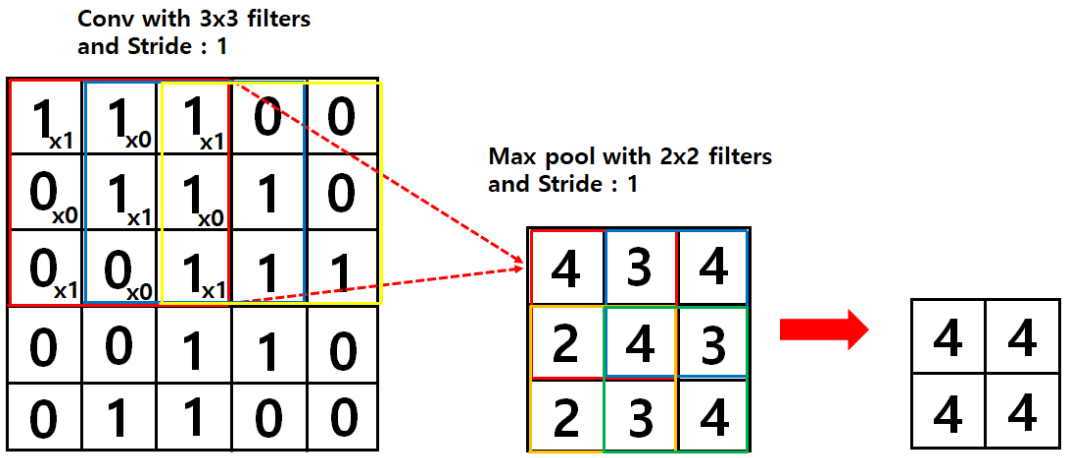


그림 2.7 CNN 기본 구조

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

컨볼루션 계층에서는 입력 데이터로부터 특징을 추출하는 역할을 하며, 특징을 추출하는 기능을 하는 필터와 필터의 값을 비선형 값으로 바꾸어주는 활성화 함수로 이루어진다. 필터를 입력 데이터에 적용하는 방법은 그림 2.8 (a)와 같이 5x5의 입력 데이터가 있다고 가정할 때 필터의 크기는 3x3인 경우, 필터를 좌측 상단에서부터 옮겨가며 특징을 추출한다. 필터를 적용하는 간격 값을 스트라이드(stride)라고 하며 값을 1로 설정하였을 경우 오른쪽으로 한 칸씩 움직여 가며 특징을 추출한다. 필터를 적용해서 얻은 결과를 특징 맵(feature map)이라고 한다. 추출된 특징 맵의 모든 특징을 사용하지 않고 인위적으로 크기를 줄이는 작업을 서브 샘플링 또는 풀링이라고 한다. 이 방법에는 맥스 풀링, 평균 풀링 등 다양한 방법이 존재하며 맥스 풀링이 가장 많이 사용된다. 맥스 풀링은 특징 맵을 MxN의 크기로 잘라낸 후, 그 안에서 가장 큰 값을 뽑아내는 방법이다. 그림 2.8 (b)와 같이 3x3의 특징 맵에서 2x2 맥스 풀링 필터를 스트라이드 값을 1로 하여 그림 2.8 (c)와 같은 결과 값이 나오는 것을 확인할 수 있다. 샘플링 과정을 통해 전체 데이터 사이즈가 줄어들기 때문에 연산에 들어가는 연산량이 적어지며, 이로 인해 오버피팅을 방지할 수 있는 장점이 있다.



(a) Input data (b) Feature map (c) Max pooing

그림 2.8 컨볼루션 계층 & 맥스 풀링 과정

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

마지막으로 완전 연결 계층은 이전의 모든 layer가 연결되어있는 뜻이며, 최종적으로 이미지를 분류하는 역할을 한다. 완전 연결 계층은 그림 2.7과 같이 1차원 데이터만 입력받을 수 있기 때문에 3차원 데이터를 평탄화해서 입력해야 한다.

제2절 딥 러닝 기반 공개 네트워크

CNN 기반의 공개 네트워크의 성능을 분석하기 위해 ILSVRC(Imagenet Large Scale Visual Recognition Competition)에서 성능이 입증된 공개 네트워크를 사용한다. ILSVRC는 이미지 인식 연구를 위한 이미지넷(ImageNet)측에서 개최하는 대회이며, 이미지 분류를 위한 데이터는 이미지넷 데이터로 학습 데이터 120만장, 테스트 데이터 5만장, 총 1000개의 클래스로 이루어져있다. 성능 평가는 Top-5 Error Rate를 이용하여 성능을 확인하며, 그림 2.9는 매년 ILSVRC 대회의 우승팀과 Top-5 Error Rate, layer의 개수를 보여준다[19].

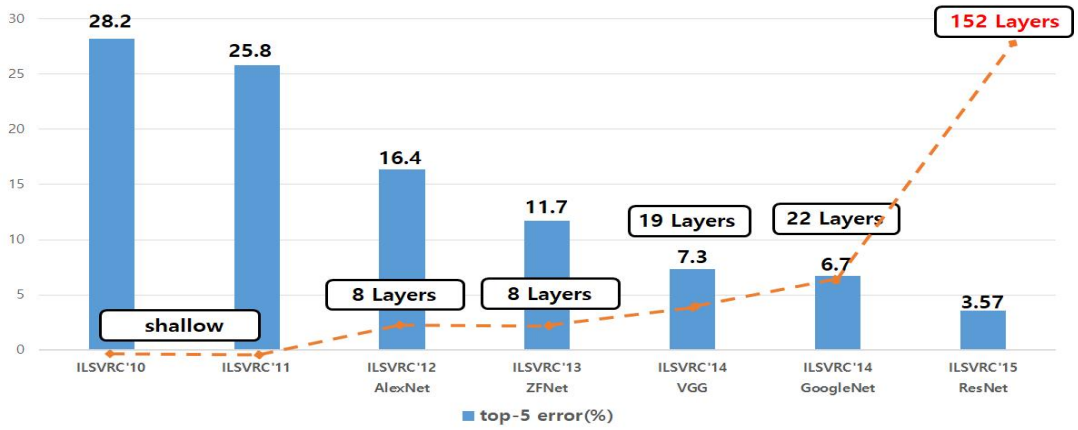


그림 2.9 ILSVRC 우승팀의 성능

공개 네트워크의 기존 성능 분석을 위해 사용한 네트워크는 크게 Inception과 ResNet 네트워크로 나눠진다. Inception 네트워크는 Inception-V1, Inception-V2, Inception-V3, Inception-V4, Inception-ResNet-V2를 사용하였으며, ResNet 네트워크는 ResNet-V1-50, ResNet-V1-101, ResNet-V1-152, ResNet-V2-50, ResNet-V2-101, ResNet-V2-152를 사용한다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

Inception 네트워크는 그림 2.10과 같이 인셉션 모듈(inception module)에서 다른 네트워크들과 차이가 있다. 대부분 큰 컨볼루션 필터를 이용하여 처리하는 방식과 달리 작은 컨볼루션 필터 여러 개를 한 층에서 구성하는 방식을 취하며 연산속도를 향상 시킨다.

Inception-V1은 인셉션 모듈을 이용한 방식을 처음으로 도입하였다. 그림 2.10과 같이 layer 초반에는 인셉션 모듈의 효과가 없기 때문에 적용하지 않으며, 이 영역을 Stem영역이라 부른다. 일반적인 네트워크는 출력 layer 부분이 소프트맥스 계층 한 층으로 구성되어있지만, Inception-V1은 그림 2.10과 같이 앞 단 layer에서 2개의 소프트맥스 계층이 추가된, 총 3개의 소프트맥스 계층의 형태를 취한다. 이는 중간 layer에서도 back-propagation을 수행하여 가중치를 갱신하는 역할을 하며, 망이 깊어질 때 발생하는 vanishing 문제를 해결한다[20].

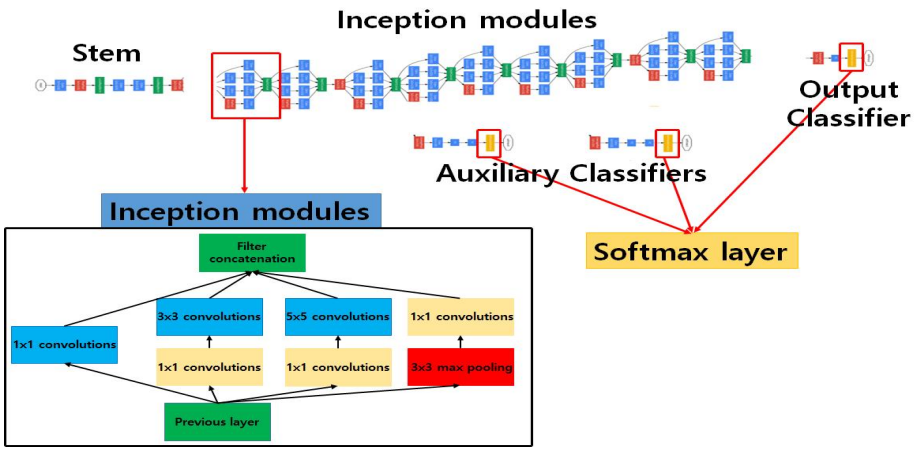


그림 2.10 Inception-V1 구조

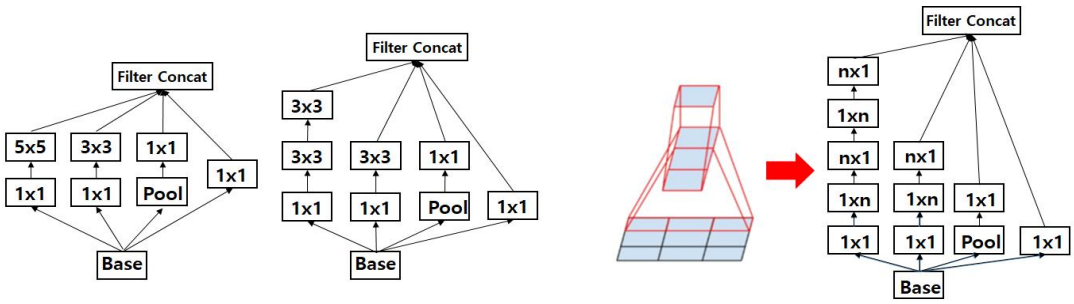
Inception-V2는 Inception-V1에 batch normalization을 적용한 네트워크이며, Inception-V2의 구조에 RMSProp, Label Smoothing, Factorized 7x7, BN-Auxiliary 등을 결합시켜가며 성능을 발전시켜나간다. 이와 같은 방법들이 모두 적용된 네트워크가 Inception-V3이며, 네트워크의 구조가 Inception-V2와 동일한 형태이다[21].

Inception-V2는 Inception-V1에서 사용했던 인셉션 모듈의 구조를 변경한다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

그림 2.11과 같이 5x5 필터를 3x3 필터 2개로 변경하는 구조를 선택하며, 이와 같은 경우 5x5 컨볼루션 연산 한번은 3x3 컨볼루션 연산보다 약 2.78배의 비용이 더 들어간다. 또한 크기가 같은 2개의 layer를 하나의 5x5로 변환하는 것과, 3x3 필터 2개로 변환하는 것의 비용을 계산하면 약 28%의 비용을 감소할 수 있다.



(a) 기존 인셉션 모듈 (b) 변형된 인셉션 모듈 (c) 비대칭 구조

그림 2.11 Inception-V2, V3 layer 특성

다음은 비대칭 컨볼루션을 사용한 방법으로, NxN 형태의 컨볼루션 연산을 1xN, Nx1의 형태로 변형한다. 이와 같은 방식으로 연산속도 및 연산량을 33% 줄인다. Inception-V2 네트워크부터는 Inception-V1에서 중간 layer에 사용했던 소프트맥스 계층이 효과가 없다고 판단하여 맨 앞 layer의 소프트맥스 계층은 제거하게 된다[22].

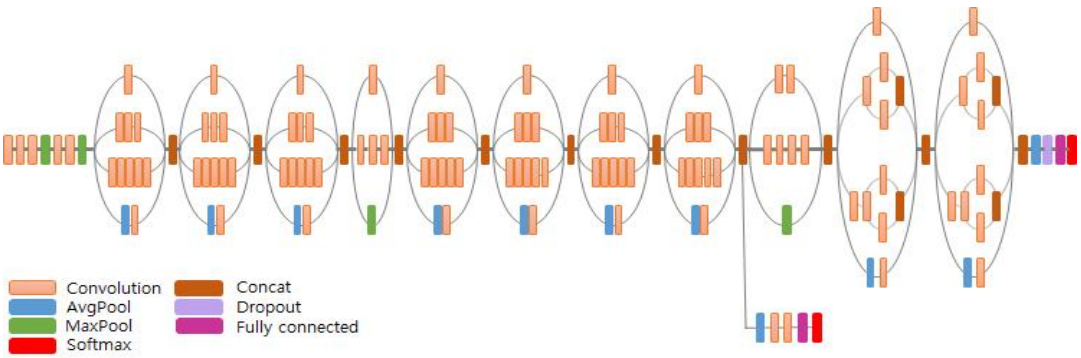


그림 2.12 Inception-V2, V3 구조

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

Inception-V4는 stem 영역의 9개 layer를 이용해 299x299x3의 입력 데이터를 35x35x384 크기의 특징 맵으로 만든다. 이후 여러 개의 Inception-A, Inception-B, Inception-C를 수행하며, 인셉션 모듈 중간에는 특징 맵의 크기를 줄이기 위한 Reduction-A와 Reduction-B의 구조를 취한다. 기본구조는 이전 Inception구조와 크게 다르지 않지만, stem영역은 이전 Inception구조가 단순한 컨볼루션과 풀링의 조합이었던 반면에 Inception-V4는 복잡한 구조가 되었다. 특히 특징 맵의 크기를 줄이는 부분이 이전 Inception구조와 달리 다양한 형태를 취하고 있다[23].

Inception-ResNet은 V1과 V2가 있으며 Inception구조에 ResNet을 도입하여 Inception-ResNet이라 불린다. Inception-ResNet-V1의 stem영역은 비교적 단순한 구조를 취하지만, V2의 경우 Inception-V4와 동일한 구조의 형태이다. 필터의 개수는 Inception-ResNet-V1에 비해 Inception-ResNet-V2가 더 많다.

기존의 ResNet과 달리 GoogLeNet 팀은 Inception구조의 특성을 고려하여, ResNet구조를 변형한다. 기존의 ResNet구조는 컨볼루션-컨볼루션 형태이지만, 1x1 컨볼루션-컨볼루션으로 변형하여 ResNet구조에 맞게 특징 맵의 차원을 맞춰주면서 연산량을 적정선으로 유지한다. Inception-ResNet의 메인구조는 그림 2.13과 같으며, 전체적으로 Inception-V4와 Inception-ResNet-V2는 75개의 layer로 동일하다.

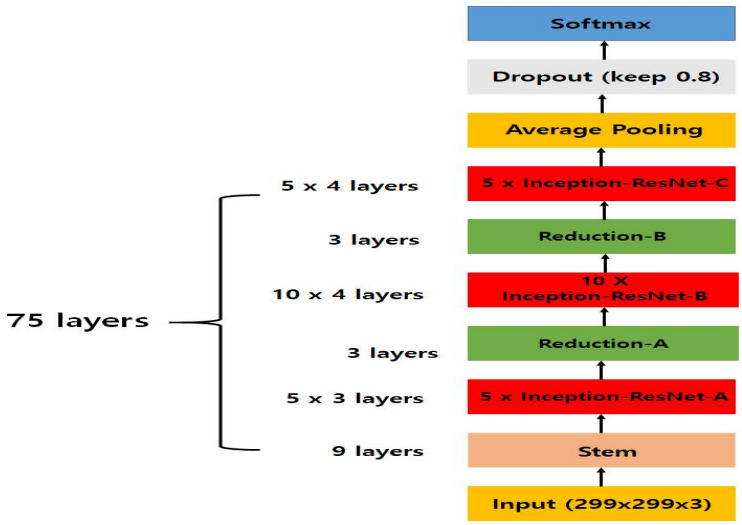


그림 2.13 Inception-ResNet 메인구조

Collection @ chosun

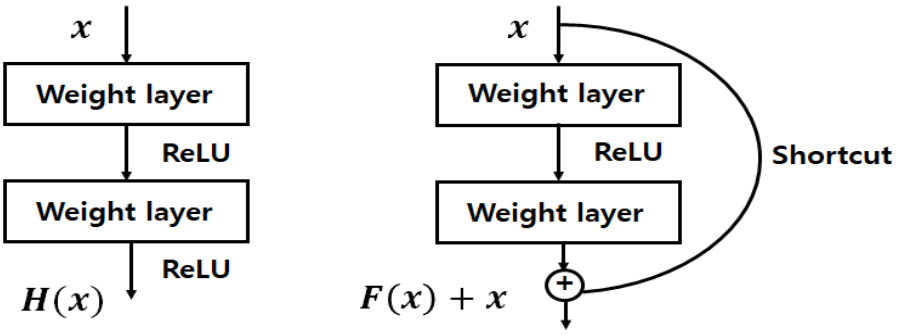
조선대학교
CHOSUN UNIVERSITY

ResNet 네트워크인 ResNet-50, ResNet-101, ResNet-152의 각 숫자는 layer의 개수를 뜻한다. ResNet 팀은 네트워크의 layer를 100개 이상으로 깊게 하면서, 깊이에 따른 학습 효과를 얻을 수 있는 방법으로 residual learning을 발표하였다. 기본적인 CNN 구조는 그림 2.14 (a)와 같이 입력 x를 받아 2개의 가중치 layer를 거쳐 최적의 H(x)를 얻는 것을 목표로 한다. 그러나 H(x)를 얻는 것이 목표가 아니라 출력과 입력의 차를 얻을 수 있도록 학습한다면, 2개의 가중치 layer는 H(x)-x를 얻도록 학습되어야 한다. 학습되는 과정을 식(2.1)로 표시하며, 최종 출력 값은 식(2.2)이 된다[24].

$$F(x) = H(x) - x \tag{2.1}$$

$$H(x) = F(x) + x \tag{2.2}$$

위와 같은 방법으로 입력에서 바로 출력으로 연결되는 shortcut 연결이 생기게 되며, 파라미터가 없이 바로 연결이 되는 구조이기 때문에 연산량 관점에서는 덧셈이 추가되는 것 외에는 차이가 없다. 또한 몇 개의 layer를 건너뛰면서 입력과 출력이 연결되기 때문에 깊은 신경망도 쉽게 최적화가 가능하고 학습할 수 있기 때문에 정확도를 개선할 수 있다.



(a) 기본적인 CNN 구조 (b) Residual Learning

그림 2.14 ResNet구조의 Residual Learning

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

공개 네트워크의 성능 분석 실험을 위해 텐서플로우 및 텐서플로우 슬림을 사용한다. 텐서플로우는 구글에서 만든 오픈소스 라이브러리이며, 그림 2.16과 같이 데이터 플로우 그래프 방식을 사용한다. 데이터 플로우 그래프는 수학 계산과 데이터의 흐름을 노드와 엣지를 사용한 방향 그래프로 표현한다. Python과 C++을 통해 구현 가능하며, 윈도우와 리눅스 운영체제 모두 사용 가능하다. 또한 CPU, GPU 모드로 동작이 가능하여 GPU의 성능에 따라 연산 속도를 빠르게 처리할 수 있다 [25].

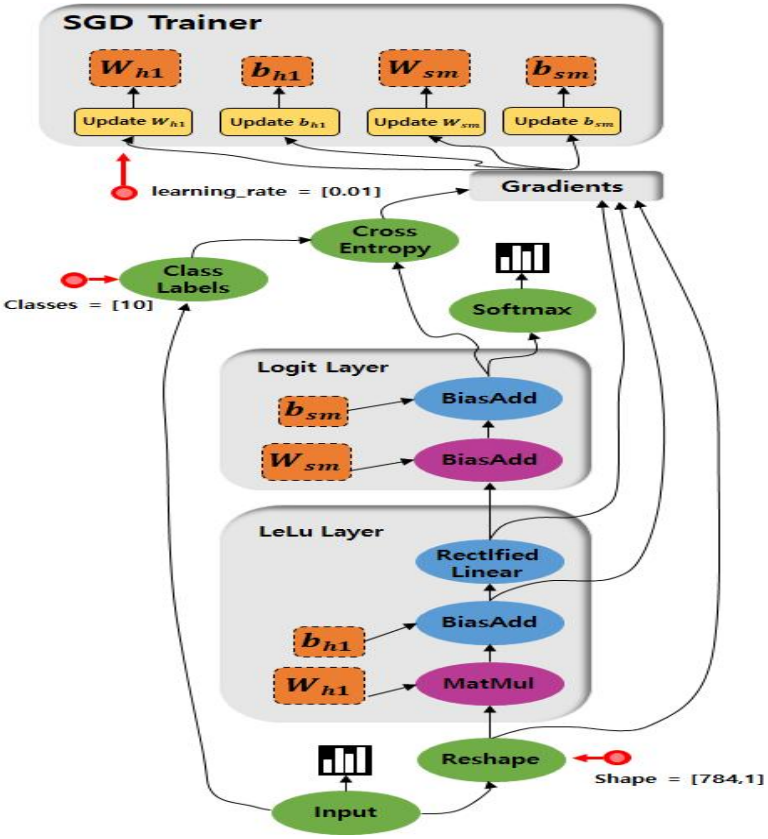


그림 2.15 텐서플로우 그래프 구조

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

텐서플로우 슬림은 텐서플로우의 복잡한 모델을 정의, 학습 및 평가할 수 있는 경량 라이브러리로 학습 데이터를 TFRecord format 형식으로 변환하여 학습시킨다. 학습 데이터를 한 장씩 읽어서 학습에 사용될 경우 파일 입, 출력에 많은 시간이 소요되는데 TFRecord format 형식으로 변환 시 학습 데이터를 바이너리 형태로 만들기 때문에 입, 출력에 걸리는 시간을 단축시킬 수 있는 장점이 있다. 학습 데이터 또한 공개 데이터 셋인 MNIST, Cifar10, ImageNet 등 4가지 종류의 데이터 셋을 제공하여 쉽게 접근할 수 있으며, 이미지넷 데이터 셋으로 학습되어진 공개 네트워크를 제공하기 때문에 전이학습 방법을 사용하여 자신의 데이터 셋으로도 학습시킬 수 있는 장점이 있다[26].

공개 네트워크의 기존 성능은 표 2.1과 같으며 Inception과 ResNet 네트워크 각각 버전이 올라갈수록 성능이 올라가는 것을 확인할 수 있다. Inception 네트워크에선 Inception-ResNet-V2의 성능이 95.3%로 가장 높으며, ResNet 네트워크에선 ResNet-V2-152의 성능이 94.1%로 가장 높은 것을 확인할 수 있다. 성능은 Top-5 Accuracy를 사용하여 평가하며, Top-5 Accuracy란 가장 높은 확률로 추정된 5개의 클래스에 테스트 데이터가 있을 확률을 나타낸다.

표 2-1 공개 네트워크 기존 성능

Model	Top-5 Accuracy
Inception-V1	89.6%
Inception-V2	91.8%
Inception-V3	93.9%
Inception-V4	95.2%
Inception-ResNet-V2	95.3%
ResNet-V1-50	92.2%
ResNet-V1-101	92.9%
ResNet-V1-152	93.2%
ResNet-V2-50	92.8%
ResNet-V2-101	93.7%
ResNet-V2-152	94.1%

심전도 데이터 셋을 사용하여 공개 네트워크의 전이학습 방법을 사용하기 전, 공개 네트워크의 기존 성능을 본 실험의 컴퓨터 환경과 비교하기 위해 직접 이미지넷 데이터 셋을 학습시켜 성능을 비교하였다. 성능비교는 기존 공개 네트워크와 같이 Top-5 Accuracy를 사용하여 비교하였으며, 기존 성능에 비해 최대 5%의 오차

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

성능을 보임을 확인하였다. Inception 네트워크의 초기 버전인 Inception-V1의 성능은 86%의 성능을 나타냈으며, 버전이 올라갈수록 기존 성능과 같이 점차 증가함을 확인할 수 있었다. Inception 네트워크에서 Inception-ResNet-V2의 성능이 91.4%로 가장 높았으며, 기존 성능에 비해 3.9% 감소하였다. ResNet 네트워크 또한 layer의 수를 늘려감에 따라 성능이 증가함을 확인할 수 있었고, ResNet-V2-152의 성능이 91.8%로 가장 높은 성능을 나타냈다. 그러나 ResNet 네트워크 또한 기존 성능에 비해 감소함을 확인할 수 있었다. 본 실험의 컴퓨터 환경에 맞추어 진행하였기 때문에 성능이 다소 떨어짐을 알 수 있었다. 딥 러닝 기반의 텐서플로우는 cpu와 gpu, 둘 다 동작이 가능하지만 gpu는 많은 양의 데이터 처리가 가능하고 병렬처리에 특화되어있기 때문에 gpu 환경이 매우 중요함을 알 수 있었다. 또한 Inception 네트워크와 ResNet 네트워크 모두 버전이 올라감에 따라 layer, 파라미터 수, 처리해야할 연산 등이 증가하기 때문에 기존 성능에 비해 감소함을 확인할 수 있었다.

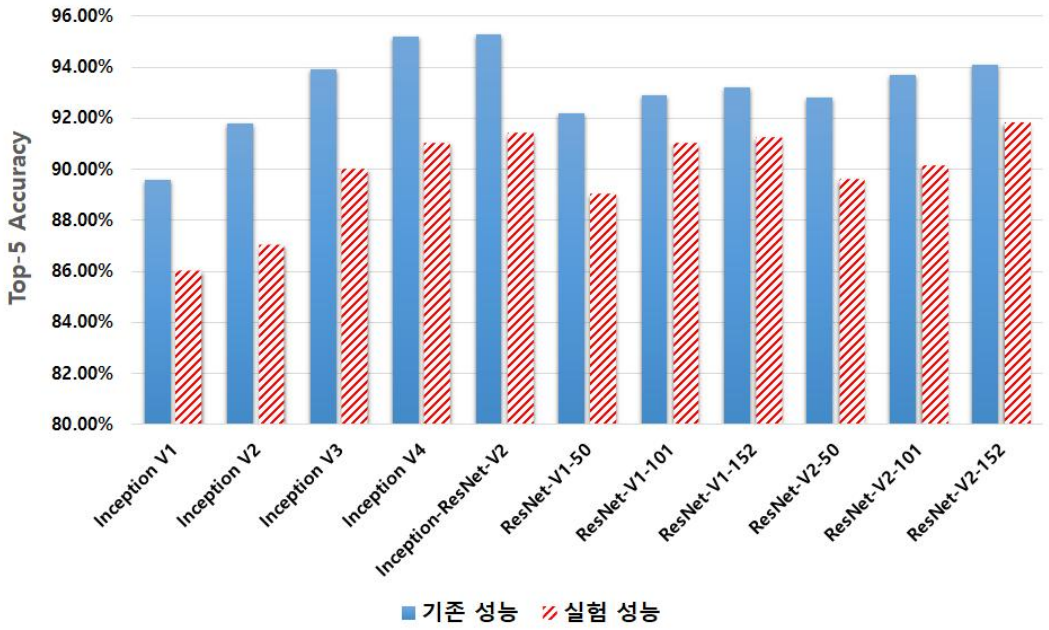


그림 2.16 기존 공개 네트워크 성능 및 실험성능 비교

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제3장 심전도를 이용한 사용자 인식

심전도는 심장의 수축과 이완 시 발생하는 전기적 활동을 해석한 신호로서 질병 진단, 생체 인식 등에 사용된다. 그림 3.1은 심전도 신호의 한 주기 파형을 나타내며, 심방의 탈분극을 나타내는 P파, 심실의 탈분극을 나타내는 QRS 복합체, 심실의 재 분극을 나타내는 T파로 구성되어있다[27]. 심전도는 심장의 크기, 위치, 신체적 조건에 의해 개인별로 파형의 형태가 고유성을 가지고 있어 차세대 바이오인식 기술로 급부상하고 있다[28].

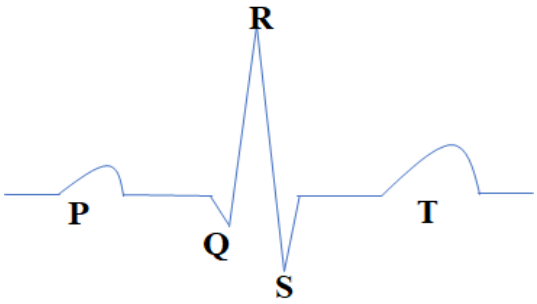


그림 3.1 심전도 신호의 형태

심전도 신호를 취득하는 방법은 그림 3.2와 같다. lead-I은 오른팔과 왼팔, lead-II는 오른팔과 왼쪽 다리, lead-III는 왼팔과 왼쪽 다리에 측정 장비를 부착하여 취득한다. 심전도 신호는 국제 표준 유도법에 의해 총 12-lead 측정법이 있으며, 12-lead는 3개의 표준 사지유도인 lead-I, II, III, 3개의 증폭유도인 aVR, aVL, aVF 그리고 6개의 흉부유도인 V1, V2, V3, V4, V5, V6로 구성되어 있다. 증폭유도는 표준 사지유도에 의해 취득된 신호를 벡터 합성을 시켜 구한다. 즉 aVR 신호는 lead-I과 II의 벡터 합, aVL은 lead-I과 III의 벡터 합, aVF는 lead-II와 III의 벡터 합에 의해 구해진다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

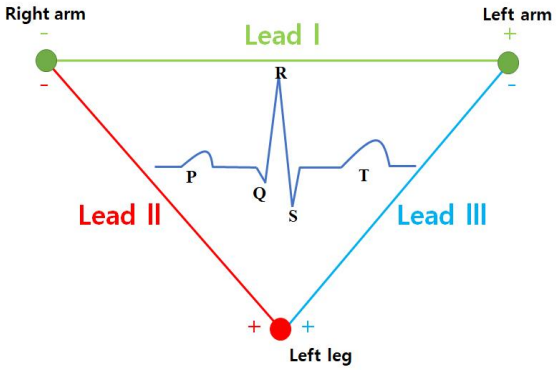


그림 3.2 표준 사지유도 lead-I, II, III 구조

심전도를 이용한 사용자 인식 과정은 그림 3.3과 같이 전처리 과정, 특징 추출, 분류 세 가지 단계로 나뉜다. 전처리 과정은 심전도 신호의 다양한 잡음을 제거하는 과정이다. 잡음의 종류에 그림 3.4와 같이 기저선 변동 잡음, 전력선 잡음, 근 잡음 등이 있다[29-31]. 기저선 변동 잡음은 호흡에 의해 발생하며, 특징점 검출의 기본 조건이 되므로 기저선 변동 잡음 제거는 필수적이다. 전력선 잡음은 측정 장비에 의해 발생하며, QRS 복합체와 QT 간격을 측정할 수 없는 경우가 발생할 수 있기 때문에 신호를 해석하는데 있어 문제점이 발생할 수 있다. 근 잡음은 신체의 움직임으로 인해 발생하며, 대부분 QRS 복합체에 발생하기 때문에 잡음 제거 과정에서 QRS 복합체의 신호 왜곡에 주의해야 한다.

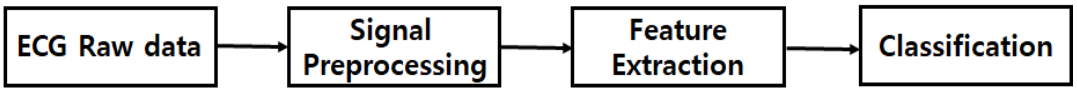
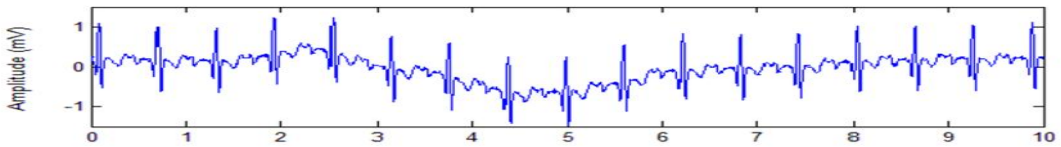


그림 3.3 심전도를 이용한 사용자 인식 시스템

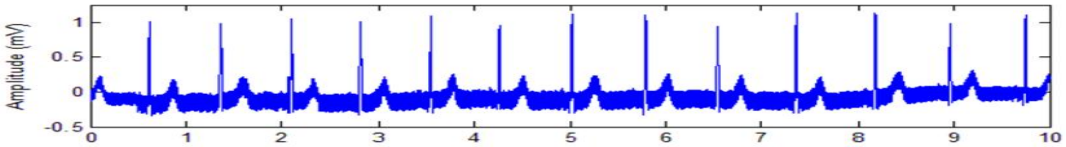
특징 추출은 전처리 과정이 완료된 심전도 신호에서 사용자를 분류하기 위한 특징을 추출하는 과정으로서 기존의 특징 추출 방법으로는 시간 영역, 주파수 영역, 위상 영역의 특징 추출 방법을 사용하였다[32-34]. 최근에는 딥 러닝을 사용하여 특징 추출과 학습을 동시에 진행하는 방법을 사용한다[35].

Collection @ chosun

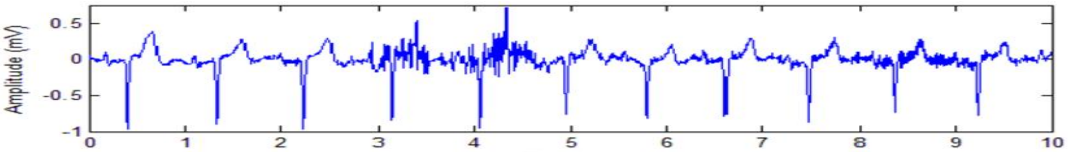
조선대학교
CHOSUN UNIVERSITY



(a) 기저선 변동 잡음



(b) 전력선 잡음



(c) 근 잡음

그림 3.4 심전도 신호의 잡음 종류

심전도 신호를 이용한 사용자 인식에 관한 기존 연구들은 심전도 파형의 P, QRS 복합체, T파와 파형간의 여러 가지 조합으로 이루어진 특징 값을 이용하는 기준점 방식과 파형 자체의 형태학적 특성을 이용한 위상 공간 분석을 통해 기준점을 검출하지 않는 비 기준점 방식으로 구분된다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제1절 기준점 기반 사용자 인식

심전도를 이용한 사용자 인식 시스템에서 L. Biel의 연구는 각 개인별로 심전도 신호가 고유성을 지니고 있음을 보였다. 20~ 55세 사이의 여성과 남성, 20명으로부터 12-lead를 사용하여 신호를 취득하였으며, 취득 장비는 SIEMENS Megacart를 사용하였다. 취득 장비에 의해 총 360(30x12-lead)개의 특징을 얻을 수 있으며, 30개의 특징은 표 3.1과 같다. 30개의 특징을 줄이기 위해 전극 부착이 간단한 lead-I을 선택하며, 상관 행렬 등의 실험을 통해 QRS 복합체와 T파의 진폭, P파의 기간 등과 같은 파형의 특성을 포함하는 10개의 특징 추출 값으로 100%의 인식률을 나타냈다[36].

표 3-1 SIEMENS Megacart에 의해 추출된 특징

Features	
P wave onset	R wave amplitude
P wave duration	S wave amplitude
QRS wave onset	R' wave amplitude
QRS wave duration	S' wave amplitude
Q wave duration	ST segment amplitude
R wave duration	2/8 ST segment amplitude
S wave duration	3/8 ST segment amplitude
R' wave duration	T+ wave amplitude
S' wave duration	T-wave amplitude
P+ wave duration	QRS wave area
QRS wave deflection	T wave morphology
P+ wave amplitude	R wave notch existence
P-wave amplitude	Delta wave confidence
QRS wave peak to peak amplitude	ST segment slope
Q wave amplitude	T wave onset

T. W. Shen은 lead-I을 사용하여 개인 인식에 관한 연구를 진행하였다. 기존 2차원 바이오인식과 달리 lead-I인 심전도 신호의 1차원 신호를 이용하여 사용자 인식의 가능성을 증명하였다. 심전도 데이터는 공용 데이터베이스인 MIT-BIH 데이터베이스를 사용하며 총인원은 20명으로 구성되어있다. 심전도 신호의 기저선 변동 잡음, 전력선 잡음 등을 제거하기 위해 전처리를 수행한 후, 7개의 특징을 추출한다. 파형의 탐지가 쉬운 QRS 복합체에서 대부분의 특징을 추출하였으며 분류

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

단계에서는 템플릿 매칭과 DBNN의 방법을 적용한다. 템플릿 매칭의 성능은 95%, DBNN의 성능은 80%의 성능을 각각 달성하였으며, 템플릿 매칭의 결과 값을 DBNN에 적용하였을 경우 100%의 인식 성능을 보였다[37].

정확한 사용자 인식을 위해 S. A. Israel은 전처리, 특징 추출, 분류 3단계로 나누어 실험을 진행하였다. 심전도 데이터는 22~48세 사이의 남성과 여성의 심전도를 직접 취득하였다. 전처리 과정에서 1.1-40Hz에서의 신호의 정보를 얻기 위해 0.06 및 60Hz의 잡음을 제거하였고, 특징 추출을 위해 기준점을 P, R, T파에 기반을 둔다. 여기에 4개의 기준점을 추가하며 그림 3.5와 같다. 기준점을 기반으로 총 15개의 특징을 추출한다. 분류 성능은 선형 판별 분석을 사용하여 29명에 대해 100%의 인식을 보였다[38].

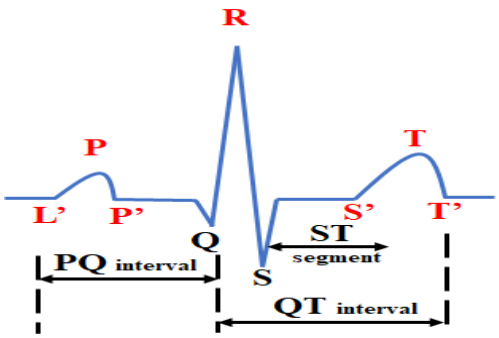


그림 3.5 S. A. Israel 연구에 사용한 기준점

O. Boumbarov은 직접 취득한 9명의 심전도 데이터를 사용한다. 기저선 변동 잡음을 제거하기 위해 하이패스 필터를 사용하며 잡음 제거 후 PCA와 LDA를 사용하여 특징을 추출한다. 분류 과정에서 신경망 기법인 HMM-GMM(Hidden Markov Model with Gaussian Mixture Model), HMM-SGM(Hidden Markov Model with Single Gaussian Model), CRF(Conditional Random Field)를 사용하였다. 실험 결과 9명에 대해 86.1%의 인식 성능을 보였다[39].

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제2절 비 기준점 기반 사용자 인식

파형의 진폭, 기울기와 같은 특징점을 찾는 기준점 방법과 달리 비 기준점 방법은 파형 형태에 대한 신호를 변환하는 방법과 차원감소, 영역변환 등을 통해서 특징 추출을 하는 방식이다.

G. Wubbeler은 lead-I, lead-II, lead-III의 조합으로부터 특징을 추출하여 2차원의 벡터를 구성하였다. 사용자 인식을 위한 심전도 데이터는 PTB 공개 데이터 74명의 심전도 신호를 사용하며, 전처리 과정으로 75Hz의 차단 주파수를 갖는 로우 패스 필터를 적용하였다. 파형의 위치와 추출을 위해 임계값을 적용하고, 두 개의 벡터 사이를 계산하여 분류한다. 분류 결과 74명에 대해 99%의 인식 성능을 보였다[40].

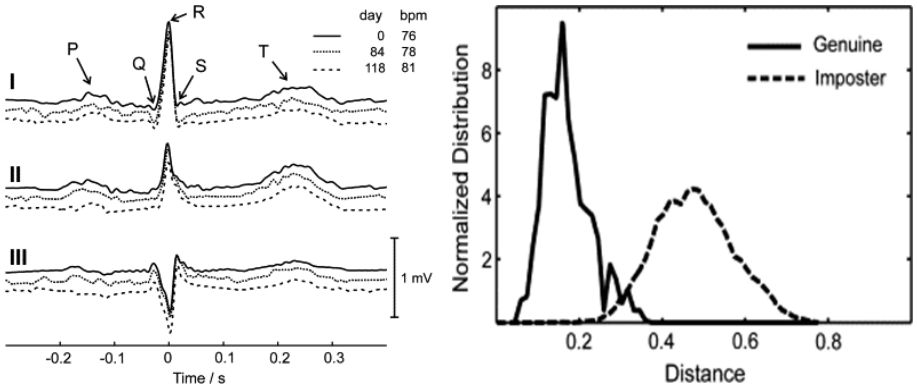


그림 3.6 G. Wubbeler 연구에 사용한 특징 추출

I. Odinak은 심전도 신호를 시간과 주파수 영역으로 변환하여 사용자를 인식하는 방법을 제시하였다. 신호 취득 장비는 Biopac TEL-100이며, 하이패스 필터 0.5Hz, 로우패스 필터 500Hz를 사용하여 신호를 취득하였다. 총 269명의 신호를 취득하였으며 1kHz로 다운 샘플링 되었다. 전력선 잡음 제거를 위해 노치 필터 60Hz를 사용하였다. 시간, 주파수 영역으로 변환하기 위해 스펙트로그램을 적용하였으며, 특징값은 원 신호와 후보 신호사이의 엔트로피가 특정 임계값보다 큰 경우를 선택한다. log-likehood를 유사도 측정을 위해 사용하며 269명에 대한 인식 성능은 76.9%를 보였다[41].

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

S. Fatemian은 단일 lead 심전도 분류를 위해 새로운 웨이블릿 변환 방법을 제시한다. 이는 잡음 제거와 특징 추출에 적용되며, 템플릿 생성을 위해 심전도 신호의 표본화, 정규화 등을 수행한다. 생성된 템플릿의 크기를 줄이기 위해 상관관계 분석 방법을 적용하며 총 13명에 대한 인식 성능은 99.6%를 달성하였다[42].

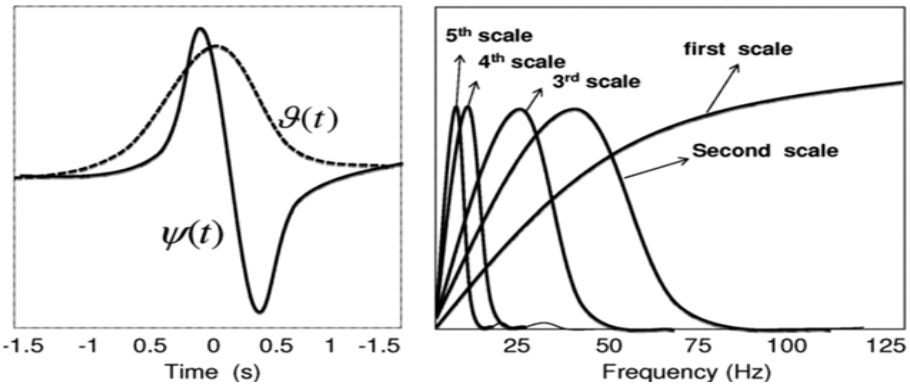


그림 3.7 S. Fatemian 연구에 사용한 특징 추출

M. M. Tawfik은 직접 취득한 22명의 lead-I 심전도 데이터를 사용하여 대표적인 형태 검출 방법인 이산 웨이블릿 변환을 이용하여 특징을 추출한다. 밴드패스 필터 1-40Hz의 주파수 범위로 신호의 잡음을 제거하며, 인공 신경망에 back-propagation을 적용한 머신러닝 기법을 사용하여 분류하였다. 22명에 대한 인식 성능은 99.09의 성능을 보였다[43].

심전도의 기준점과 비 기준점을 이용한 사용자 인식은 현재도 많은 연구가 진행되고 있다. 본 논문에서는 잡음 제거와 한 주기의 신호로 분할하기 위해 기준점 기반의 방법을 적용하고, 특징 추출 및 분류 단계에서 비 기준점 방식인 딥 러닝 기반의 cnn을 이용하여 사용자 인식에 관한 연구를 진행한다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제4장 제안하는 DCNN 기반 심전도를 이용한 사용자 인식 시스템

제안하는 DCNN 구조는 기존 공개 네트워크 분석을 통해 설계한다. 기존 공개 네트워크는 이미지넷 데이터에 맞게 설계된 구조이므로 상대적으로 layer층이 깊다. 적은 수의 클래스를 분류하는 심전도 데이터셋에 맞게 네트워크를 설계하기위해 기존 공개 네트워크에 비해 상대적으로 layer층이 깊지 않은 VGG 네트워크 기반으로 설계한다. 따라서 기존 공개 네트워크에 비해 직관적인 구조를 가지고 연산 속도를 감소시키는 장점이 있다. 또한 기존 공개 네트워크 분석을 통해 Inception 네트워크에서 적용한 작은 크기의 컨볼루션 필터를 사용하여 연산량을 줄인다. 제안하는 DCNN의 전체적인 구조는 컨볼루션, 맥스 풀링의 형태로 되어있으며, 각각의 컨볼루션 연산후 정규화 과정 진행하여 오버피팅을 방지하며, 가중치 초기화, 활성화 함수, cost, optimizer 함수 등을 사용하여 심전도 데이터 셋에 맞게 구조를 설계한다.

제1절 심전도 신호 전처리 및 학습 데이터 변환

제안하는 DCNN 기반 심전도를 이용한 사용자 인식 시스템의 흐름도는 그림 4.1과 같다. 심전도 원 신호는 전력선 잡음, 근 잡음, 기저선 변동 잡음 등 다양한 잡음이 섞여있어 사용자 인식 성능에 영향을 미친다. 전처리 과정에서는 밴드패스 필터, 평균 이동 필터를 사용하여 잡음을 제거한다. 잡음 제거 후 심전도 한 주기의 입력 데이터를 만들기 위해 R 피크점을 찾아 분할한다.

분할된 신호는 입력 데이터에 따른 학습 성능을 확인하기 위해 2차원 영상과 커플링 매트릭스로 각각 변환하며, 영상 사이즈는 각각 100x100을 사용한다. 제안하는 DCNN 구조는 컨볼루션-컨볼루션-풀링 구조인 VGG 네트워크 기반으로 설계하였고, 심전도 데이터 MIT-BIH Normal Sinus Rhythm Database, 18명과 MIT-BIH Arrhythmia Database, 47명의 데이터를 학습시켜 분류 성능을 확인한다[44]

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

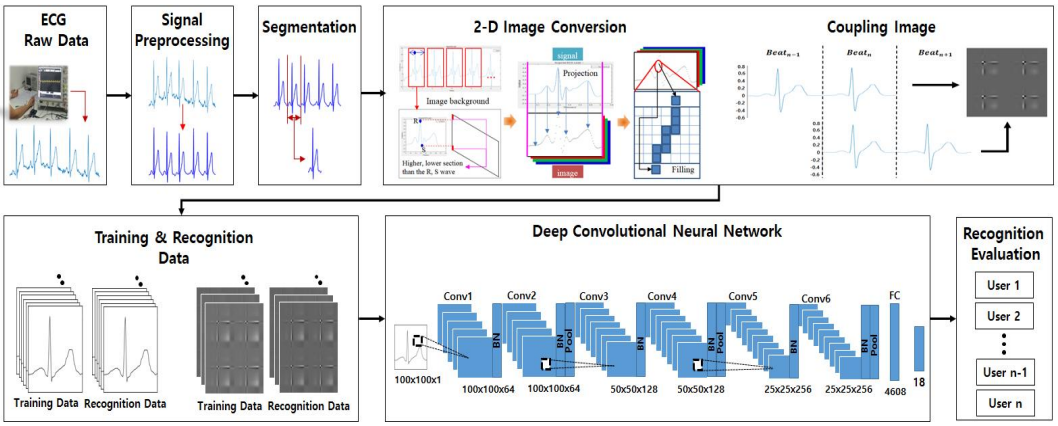
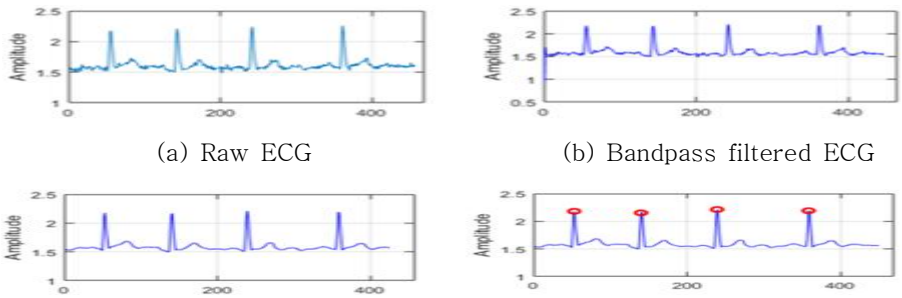


그림 4.1 제안하는 DCNN 기반 심전도를 이용한 사용자 인식 시스템

그림 4.2는 본 논문에서 적용한 심전도 신호의 전처리 과정이다. 심전도 신호의 취득 장비에 의해 발생되는 60Hz 전력선 잡음, 측정자의 움직임에 의해 발생하는 근 잡음, 호흡에 의한 기저선 변동 잡음을 밴드패스 필터를 이용하여 차단 주파수를 설정하고 제거하였으며 그림 4.2 (b)와 같다. 밴드패스 필터링 후 신호에 남아있는 미세한 잡음을 제거하기 위해 그림 4.2 (c)와 같이 평균이동 필터를 적용한다. 학습 데이터로 심전도의 P, QRS 복합체, T파로 구성된 한 주기를 사용하기 때문에 팬 톰킨스 알고리즘을 이용하여 R 피크점을 검출한다. 다른 구간에 비해 가파른 경사를 가지는 R파의 특징과 일정한 주기로 반복되는 특징을 이용하여 R 피크점을 검출하며 그림 4.2 (d)와 같다.



(a) Raw ECG (b) Bandpass filtered ECG

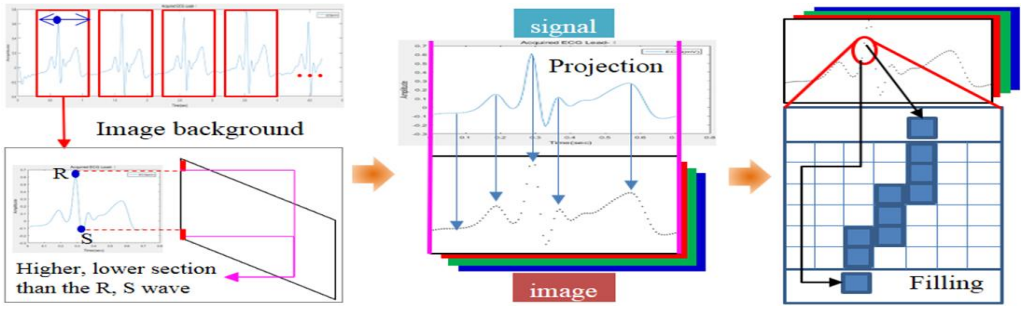
(c) Moving average filtered ECG (d) R wave detected ECG

그림 4.2 심전도 신호 전처리 과정

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

본 논문에서 학습 데이터는 한 주기의 2차원 영상과 coupling matrix를 사용한다. 2차원 영상을 변환하는 과정은 그림 4.3과 같다. 영상 배경을 생성하기 위해 그림 4.3 (a)와 같이 영상의 높이와 폭을 일정하게 정규화 하여 생성하며, 생성된 영상 배경에 한 주기의 심전도 픽셀 값을 투영하는 과정은 그림 4.3 (b)와 같다. 마지막으로 그림 4.3 (c)와 같이 픽셀 사이에 데이터를 채워 2차원 영상을 만든다[45].



(a) 영상 배경 생성 (b) 심전도 픽셀 투영 (c) 데이터 채우기

그림 4.3 2차원 영상 변환 과정

심전도를 이용한 coupling matrix 생성 과정은 그림 4.4과 같다. 본 논문에서는 coupling matrix를 심전도 신호 주기 3개($Beat_{x-1}$, $Beat_x$, $Beat_{x+1}$)를 사용해 표현한다. $Beat_x$를 중심으로 $Beat_{x-1}$과 $Beat_x$, $Beat_x$와 $Beat_{x+1}$, 2개 주기가 포함된 $DualBeat_a$와 $DualBeat_b$를 식(4.1)과 같이 계산한다. $Beat_x$를 중첩해서 사용하는 이유는 같은 주기의 QRS 복합체가 만나는 구간으로 각 개인의 고유의 값을 표현하기 위해 사용한다. 그 후 식(4.2)와 같이 $DualBeat_a$, $DualBeat_b$를 계산하여 2차원 coupling matrix를 생성한다[46].

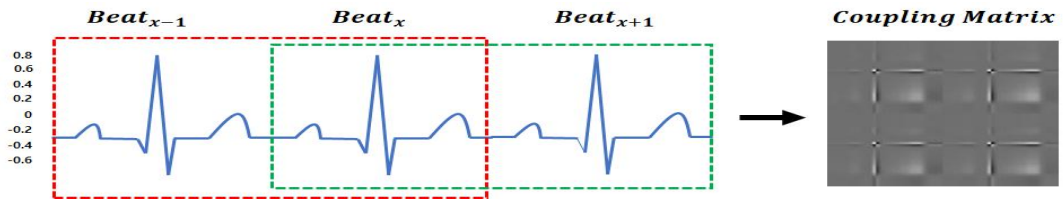


그림 4.4 커플링 매트릭스 생성 과정

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

$$DualBeat_a = [Beat_{x-1}(1), Beat_{x-1}(2), \cdot\cdot\cdot, Beat_x(M-1), Beat_x(M)]$$
$$DualBeat_b = [Beat_x(1), Beat_x(2), \cdot\cdot\cdot, Beat_{x+1}(M-1), Beat_{x+1}(M)] \quad (4.1)$$

$$CM = DualBeat_a \times DualBeat_b^T \quad (4.2)$$

제2절 DCNN 네트워크 설계

본 절에서는 심전도 데이터에 최적화된 DCNN 네트워크를 설계한다. VGG 네트워크를 기반으로 설계하며, VGG 네트워크는 다른 공개 네트워크에 비해 직관적이고 간단한 구조이며 파라미터 값들을 변경해가면서 실험을 진행할 수 있는 장점이 있다. 또한 심전도 이미지는 상대적으로 단순한 이미지이기 때문에 layer를 무작정 깊게 할 필요가 없다. 제안하는 DCNN 구조는 그림 4.5와 같다.

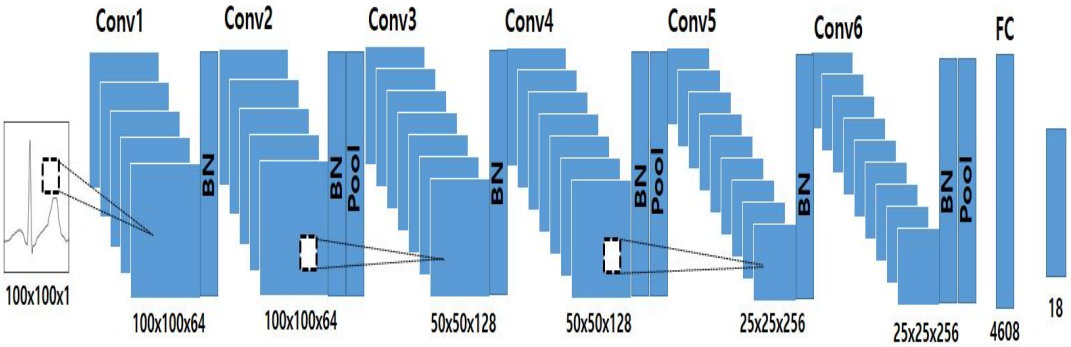


그림 4.5 제안하는 DCNN 구조

1. 가중치 초기화

CNN에서 가중치는 커널 또는 필터로 표현하며, 커널 그룹은 단일 컨볼루션 계층을 형성한다. 경사 감소(Gradient descent)기반의 학습 시 back-propagation이 진행되는 과정에서 경사가 0에 근접하거나 발산할 경우 학습이 진행되지 않는 문제점이 생기게 된다. 가중치가 너무 크면 값이 분산하고, 작은 경우 0에 지나치게 수렴되는 문제가 발생하며 그림 4.6과 같다. 이러한 문제를 해결하기 위해 제안하

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

는 DCNN 구조에서는 Xavier initialization을 사용하여 가중치를 초기화한다. 또한 3x3의 작은 커널 사이즈를 사용하여 경사 감소의 문제점을 해결한다.

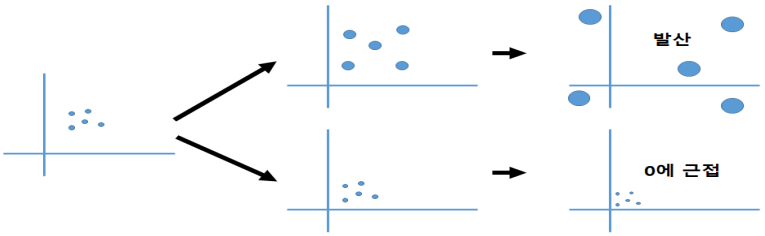


그림 4.6 경사 감소 학습 시 문제점

2. 활성화 함수

활성화 함수는 네트워크에서 커널 가중치의 출력 값을 정의한다. CNN 구조에서는 비선형 활성화 함수를 사용하며, ReLU, LReLU, ELU등이 많이 사용되고 있다. 본 구조에서는 비선형 함수이며, 미분을 해도 경사 감소 문제가 발생하지 않는 ReLU 함수를 사용한다. ReLU 활성화 함수는 그림 4.7과 같으며 식(4.3)과 같다.

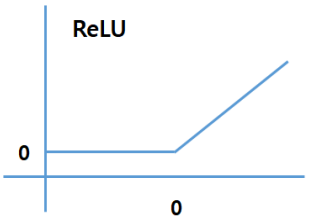


그림 4.7 ReLU 활성화 함수

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \tag{4.3}$$

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

3. 정규화

Regularization은 normalization이라고도 부르며 학습 단계에서 오버피팅을 방지하기 위해 사용한다. 그림 4.8과 같이 데이터 셋이 정규화가 잘 되어있어도 layer를 지나면서 처음에는 유지가 될 수 있지만, 흩어짐이 누적되면서 전체 결과에 영향을 줄 수 있다. 이러한 상황을 방지하기 위해 제안하는 DCNN 네트워크에서는 그림 4.5와 같이 입력 데이터뿐만 아니라 배치 단위로 매 layer에서 batch normalization을 수행한다. 추가적으로 오버피팅을 해결하기 위해 dropout 방법을 사용한다. dropout이란 임의의 확률로 일부 뉴런을 학습에서 제외하는 방법으로 테스트 시에는 학습 시 제외하였던 뉴런을 다시 사용할 수 있으며 본 실험에서는 0.7의 값을 설정하였다.

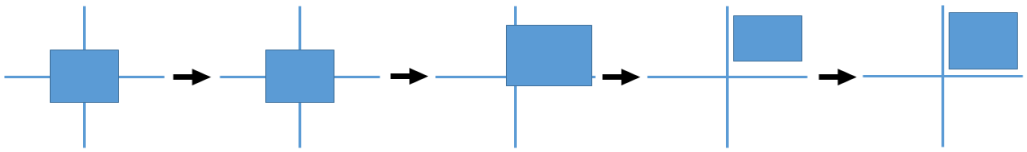


그림 4.8 정규화의 필요성

4. Cost, Optimizer 함수

Cost 함수는 학습이 잘 되는지, 주어진 학습 데이터 샘플과 예상 출력 간의 차이를 나타내는 척도이다. 딥 러닝에서는 기본적으로 cross-entropy 함수를 사용하며 식(4.4)와 같다. n은 학습 데이터의 수, y는 예측 값, a는 출력 layer로부터의 실제 값을 타나낸다. cost 값을 최소화하기 위해 optimizer 함수를 사용하며 본 실험에서는 Adam optimizer 함수를 사용한다.

$$C=-\frac{1}{n}\sum[y\ln a+(1-y)\ln(1-a)] \tag{4.4}$$

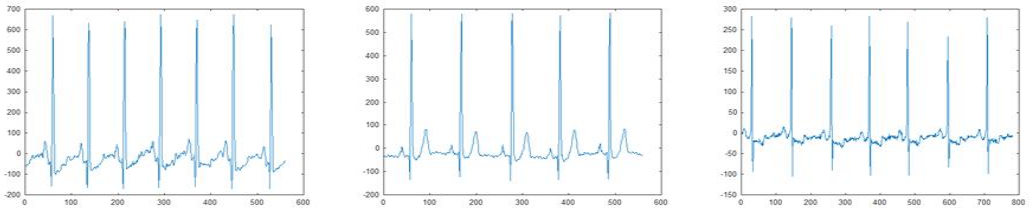
Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제3절 실험결과

본 절에서는 DCNN 기반 심전도를 이용한 사용자 인식 시스템의 성능을 확인하기 위해 공개 네트워크의 전이학습방법을 사용하였을 때의 성능과 직접 구현한 네트워크의 성능을 비교 분석한다. 네트워크의 입력 데이터는 MIT-BIH Normal Sinus Rhythm Database와 MIT-BIH Arrhythmia Database을 사용하여 각 성능을 비교하며, 두 데이터 셋 모두 학습 데이터는 개인당 450장, 인식 데이터는 50장을 사용한다.

MIT-BIH Normal Sinus Rhythm Database는 남성 5명, 여성 13명, 총 18명의 인원으로 구성되어 있으며 평균 나이는 20~50세까지 다양하게 분포되어있다. 취득된 인원 모두 심장 질병이 없는 정상적인 사람이며 편안한 자세를 유지한 상태에서 샘플링 주파수는 128Hz로 취득되었다. MIT-BIH Arrhythmia Database는 초당 360 샘플로 취득되었으며, 부정맥을 가지고 있는 남성 25명, 여성 22명, 총 47명의 인원으로 구성되어있다. 평균 나이는 남성의 경우 32~89세, 여성의 경우 23~89세로 분포되어있다. 그림 4.9는 MIT-BIH Normal Sinus Rhythm Database의 취득 인원 18명중 일부의 lead-I 신호를 나타내며, 개인마다 심전도 신호의 모양이 다름을 확인할 수 있으며, 고유성을 가짐을 알 수 있다.



(a) Subject 1 (b) Subject 2 (c) Subject 3

그림 4.9 MIT-BIH Normal Sinus Rhythm Database의 lead-I 신호

그림 4.10과 그림 4.11은 MIT-BIH Normal Sinus Rhythm Database를 사용하였을 때 기존 공개 네트워크의 2차원 영상과 커플링 매트릭스의 성능 및 제안하는 구조의 성능을 나타내며, 그림 4.12와 그림 4.13은 MIT-BIH Arrhythmia Database를 학습 데이터로 사용하였을 때의 성능을 나타낸다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

1. MIT-BIH Normal Sinus Rhythm Database 실험결과

그림 4.10은 공개 네트워크의 전이학습 방법을 사용하였을 때 각 성능을 나타낸다. 전이학습이란 이미지넷과 같은 대형 데이터 셋을 기반으로 학습되어진 네트워크를 사용하여 자신이 가지고 있는 데이터 셋에 맞게 재학습 시키는 방법을 말한다. 전이학습은 크게 두 가지 방법으로 구분된다. 이미지넷 데이터로 학습되어진 네트워크에서 분류 계층만 제거하여 나머지 부분을 새로운 데이터 셋의 특징을 추출하는데 사용하는 방법과, 마지막 분류 계층만 재학습하는 것이 아니라 전체 네트워크의 가중치를 재조정하여 학습하는 방법이 있다. 본 논문에서는 tf-slim에서 제공하는 네트워크의 체크포인트를 사용하여 실험을 진행하였다.

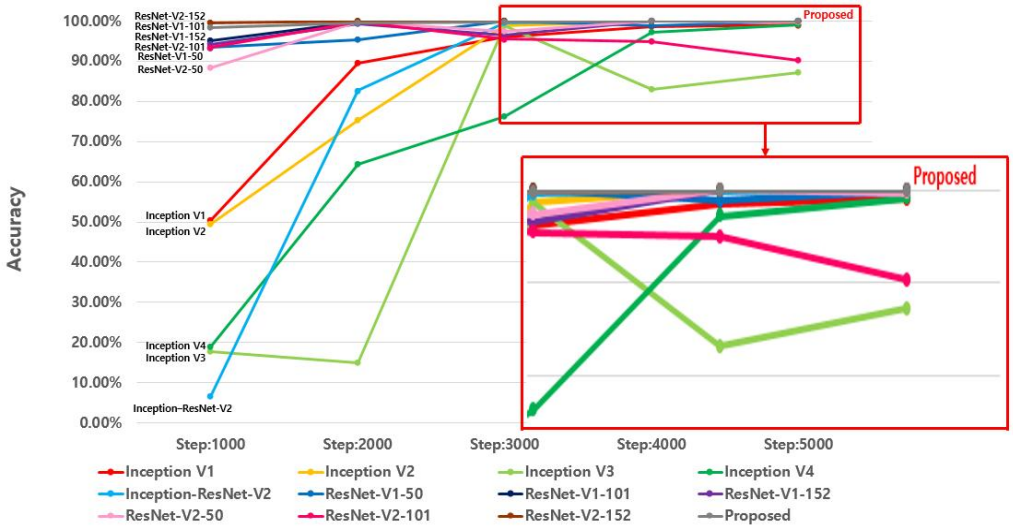


그림 4.10 커플링 매트릭스 공개 네트워크 성능(18명)

그림 4.10은 공개 네트워크와 제안하는 DCNN 구조에 커플링 매트릭스를 학습 데이터로 사용하였을 때 학습 반복 횟수 별 각 성능을 나타낸다. 학습 반복 횟수를 늘려감에 따라 성능이 향상되는 것을 확인 할 수 있으며, Inception 네트워크와 달리 ResNet 네트워크는 대부분 비슷한 성능을 보임을 확인할 수 있었다. 또한 제안하는 DCNN구조는 기존 공개 네트워크에 비해 layer층이 깊지 않음에도 불구하고 우수한 성능을 보임을 확인할 수 있었다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

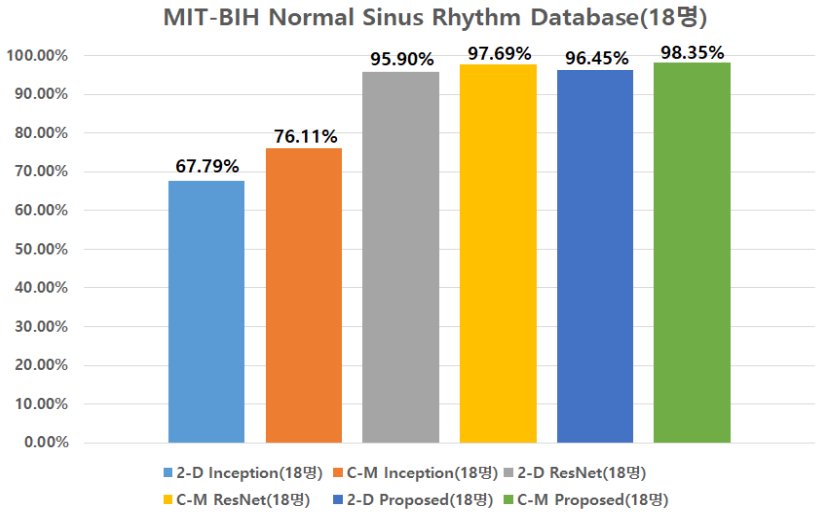


그림 4.11 제안하는 구조 및 공개 네트워크 성능 비교(18명)

그림 4.11은 제안하는 구조 및 공개 네트워크의 성능을 나타낸다. 비교를 위해 Inception 네트워크와 ResNet 네트워크의 평균 성능과 비교한다. 또한 2차원 영상과 커플링 매트릭스를 입력 데이터로 사용하여 입력 데이터에 따른 성능을 비교한다. 제안하는 DCNN 구조에 2차원 영상과 커플링 매트릭스를 입력 데이터로 사용하였을 때 성능은 96.45%, 98.35%를 나타냈다. 공개 네트워크와 제안하는 구조의 성능에서 확인할 수 있듯이 2차원 영상보다 커플링 매트릭스를 사용하였을 때 Inception 네트워크는 8.32%, ResNet 네트워크는 1.79%, 제안하는 구조는 1.9% 증가함을 확인하였다.

2. MIT-BIH Arrhythmia Database 실험결과

그림 4.12와 그림 4.13은 클래스 수를 47명으로 증가하여 실험을 진행한 결과이며, 기존 실험과 마찬가지로 2차원 영상과 커플링 매트릭스를 학습 데이터로 사용하였을 때 공개 네트워크의 성능과 제안하는 DCNN 구조의 성능을 나타낸다.

커플링 매트릭스를 공개 네트워크와 제안하는 DCNN 구조에 학습 데이터로 사용하였을 때 기존 실험과 마찬가지로 2차원 영상을 사용하였을 때 보다 성능이 증가함을 확인할 수 있었다. 그러나 기존 실험과 달리 클래스 수를 증가하여 실험을 진행하였기 때문에 성능이 감소함을 확인하였다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

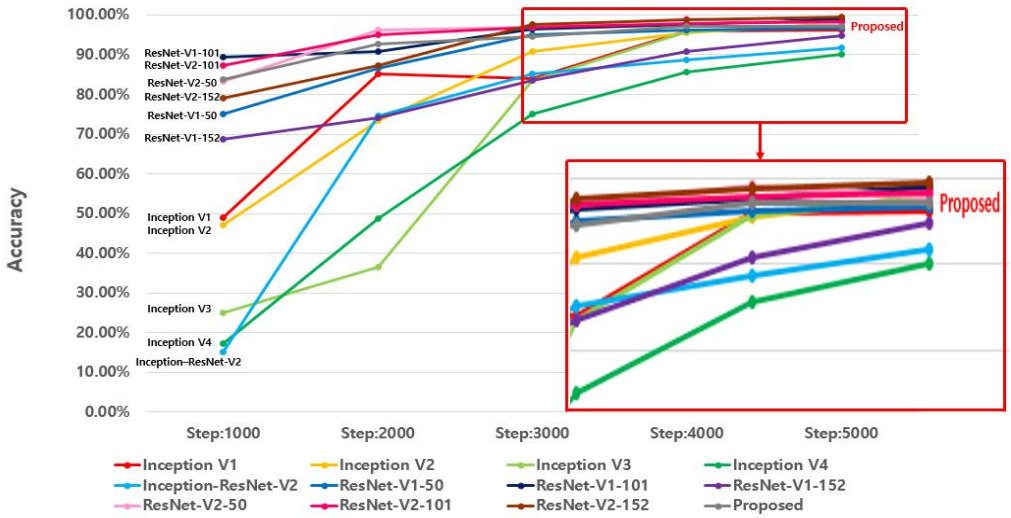


그림 4.12 커플링 매트릭스 공개 네트워크 성능(47명)

그림 4.13은 제안하는 DCNN 구조의 성능과 공개 네트워크의 성능을 나타낸다. 제안하는 DCNN 구조에 2차원 영상과 커플링 매트릭스를 입력 데이터로 사용하였을 때 성능은 각각 94.35%, 97%를 나타냈다. 기존 실험과 마찬가지로 커플링 매트릭스를 학습 데이터로 사용하였을 때 성능이 2.65% 증가함을 확인하였다.

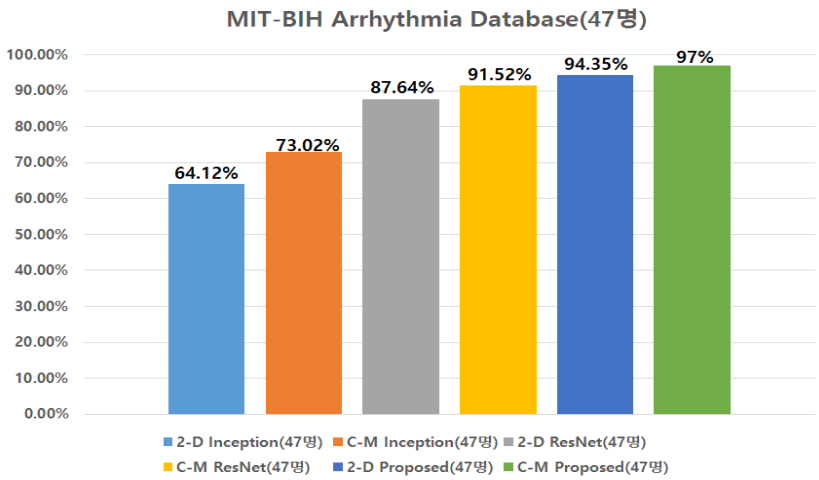


그림 4.13 제안하는 구조 및 공개 네트워크 성능 비교(47명)

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제5장 결론

기존 바이오인식인 얼굴, 지문, 홍채 등을 이용하는 방식은 현대 사회에서 다양한 범죄에 악용되고 있다. 이를 보완하기 위해 주요 선진국은 위조 및 변조에 강한 생체신호를 활용한 연구에 대해 활발히 진행 중이다. 그 중 심전도는 개인마다 고유한 특성을 가지고, 취득하기 쉬운 장점이 있어 차세대 바이오인식으로 주목받고 있다. 또한 최근 많은 분야에서 사용 중인 딥 러닝을 심전도와 접목하여 사용자 인식에 관한 연구가 진행되고 있다.

본 논문에서는 공개 네트워크인 Inception, ResNet 네트워크의 실험 및 분석을 통해 제안하는 DCNN 구조를 설계하였다. 제안하는 DCNN 구조는 적은 수의 클래스를 분류하기 때문에 layer층을 기존 공개 네트워크처럼 깊게 설계하지 않으며, 상대적으로 layer층이 깊지 않은 VGG 네트워크 기반으로 설계하였다.

제안하는 DCNN의 전체적인 구조는 컨볼루션-풀링의 단순한 구조로 이루어져있어 직관적인 구조이고 기존 공개 네트워크인 Inception 네트워크에서 적용한 작은 크기의 컨볼루션 필터를 제안하는 DCNN 구조에 적용하여 연산속도를 감소하였다. 또한 오버피팅을 방지하기 위해 각각의 컨볼루션 연산 후 정규화 과정을 진행하며, 가중치 초기화, 활성화 함수, cost, optimizer 함수 등을 사용하여 심전도 데이터 셋에 맞게 구조를 설계하였다.

실험결과 제안하는 구조의 성능은 18명의 심전도 데이터를 사용하였을 때 커플링 매트릭스의 성능이 98.35%로, 2차원 영상보다 1.9% 증가하였으며 마찬가지로 클래스 수를 증가하여 47명의 심전도 데이터를 사용하였을 때 커플링 매트릭스의 성능이 2.65% 더 우수함을 확인하였다.

향후에는 심전도 데이터의 수를 늘려 보다 많은 사용자를 분류하는 실험을 진행할 예정이다. 또한 성능 향상을 위해 네트워크의 구조를 변경하여 실험을 진행하며, 심전도 데이터를 이용해 새로운 입력 데이터를 만드는 방법에 관해 진행할 것이다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

참고문헌

[1] Frost&Sullivan, https://ww2.frost.com/

[2] J. L. Wayman, "Technical Testing and Evaluation of Biometric Identification Devices," In Biometrics: Personal identification in networked society, pp. 345-368, 1999.

[3] R. De luis-garcia, et al., "Biometric Identification Systems," Signal Processing, Vol. 83, Issue 12, pp. 2539-2557, 2003.

[4] A. K. Jain, et al., "An Introduction to Biometric Recognition," IEEE Trans. on Circuits and Systems for Video Technology, Vol. 14, No. 1, pp. 4-19, 2004.

[5] S. Gutta, and Q. Cheng, "Joint Feature Extraction and Classifier Design for ECG-based Biometric Recognition," Journal of Biomedical and Health Informatics, Vol.20, Issue 2, pp. 460-468, 2016.

[6] N. Kunju, et al., "EMG Signal Analysis for Identifying Walking Patterns of Normal Healthy Individuals," Indian Journal of Biomechanics, Vol. Special Issue, pp. 118-122, 2009.

[7] M. W. Rahman and L. Gavrilova, "Emerging EEG and Kinect Face Fusion for Biometric Identification," Computational Intelligence, pp. 1-8, 2017.

[8] Jin Su Kim and Sung Bum Pan, "A Study on EMG-based Biometrics," Journal of Internet Services and Information Security, Vol. 7, No. 2, pp. 19-31, 2017.

[9] Y. LeCun, et al., "Deep Learning," Nature, Vol. 521, pp. 436-444, 2015.

[10] A. Samuel, "Some Studies in Machine Learning using the Game of Checkers," IBM Journal of Research and Development, Vol. 3, pp. 211-219, 1959.

[11] T. Mitchell, "Machine Learning," McGraw Hill, Vol. 45, pp. 870-877, 1997.

[12] W. S. McCulloch and W. Pitts, "A Logical Calculus of the Ideas

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

Immanent in Nervous Activity," Bulletin of Mathematical Biophysics, Vol. 5, pp. 115-133, 1943.

[13] F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," Psychological Review, Vol. 65, pp. 386-408, 1958.

[14] https://en.wikipedia.org/wiki/Perceptrons_(book)

[15] D. E. Rumelhart, et al., "Learning Representations by Back-Propagating Errors," Nature, Vol. 323, pp. 533-536, 1986.

[16] Y. LeCun, et al., "Handwritten Digit Recognition with a Back-propagation network," Advances in Neural Information Processing Systems, Vol. 2, pp. 396-404, 1990.

[17] Y. LeCun, et al., "Gradient-based Learning Applied to Document Recognition," Proceedings of the IEEE, Vol. 86, No. 11, pp. 2278-2324, 1998.

[18] Y. LeCun, et al., "Backpropagation Applied to Handwritten Zip Code Recognition," Neural Computation, Vol. 1, Issue. 4, pp, 541-551, 1989.

[19] O. Russakovsky, et al., "Imagenet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, Vol.3, No. 115, pp. 211-252, 2015.

[20] C. Szegedy, et al., "Going Deeper with Convolutions," Computer Vision and Pattern Recognition, pp. 1-9, 2015.

[21] C. Szegedy, et al., "Rethinking the Inception Architecture for Computer Vision," Computer Vision and Pattern Recognition, pp. 2818-2826, 2016.

[22] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," Proc. of Machine Learning Research, Vol. 37, pp. 448-456, 2015.

[23] C. Szegedy, et al., "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," Proc. AAAI, pp. 1-3, 2017.

[24] K. He, et al., "Deep Residual Learning for Image Recognition,"

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

Coumputer Vision and Pattern Recognition, pp. 770-778, 2016.

[25] Tensorflow, https://www.tensorflow.org/

[26] TF-Slim, https://github.com/tensorflow/models/tree/master/research/slim

[27] E. K. Zaghouani, et al., "ECG based Authentication for E-healthcare Systems: Towards a Secured ECG Features Transmission," Wireless Communications and Mobile Computing Conference, pp. 1777-1783, 2017.

[28] 김정준, 이승민, 류강수, 이종학, 박길흠, "곡률기반 기준점 검출을 이용한 계층적 심전도 신호 개인인증 알고리즘," 멀티미디어학회 논문지, Vol. 20-3, pp. 465-473, 2017.

[29] M. Merone, et al., "ECG Databases for Biometric Systems: A Systematic Review," Expert System with Application, Vol. 67, pp. 189-202, 2017.

[30] F. A. Elhaj, et al., "Arrhythmia Recognition and Classification using Combined Linear and Nonlinear Features of ECG Signals," Computer Methods and Programs in Biomedicine, Vol. 1, No. 2, pp. 52-63, 2016.

[31] G. M. Friesen, et al., "A Comparison of the Noise Sensitivity of Nine QRS Dection Algorithms," IEEE Trans. on Biomedical Engineering, Vol. 37, pp. 85-98, 1990.

[32] N. Venkatesh and S. Jayaraman, "Human Electrocardiogram for Biometrics using DTW and FLDA," International Conference on Pattern Recognition, pp. 3838-3841, 2010.

[33] Z. Zhao, et al., "A Human ECG Identification System based on Ensemble Empirical Mode Decomposition," Sensors, Vol. 13, No. 5, pp. 6832-6864, 2013.

[34] S. L. Lin, et al., "Individual Identification based on Chaotic Electrocardiogram Signals during Muscular Exercise," IET Biometrics, Vol. 3, No. 4, pp. 257-266, 2014.

[35] R. D. Labati, et al., "Deep-ECG: Convolutional Neural Networks for ECG Biometric Recognition," Pattern Recognition Letters, 2018.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

[36] L. Biel, et al., "ECG Analysis: A New Approach in Human Identification," IEEE Transactions on Instrumentation and Measurement, Vol. 50, No. 3, pp. 808-812, 2001.

[37] T. W. Shen, et al., "One-lead ECG for Identity Verification," IEEE EMBS/BMES, Vol. 1, pp. 62-63, 2002.

[38] S. A. Israel, et al., "ECG to Identify Individuals," Pattern Recognition, Vol. 38, pp. 133-142, 2005.

[39] O. Boumbarov, et al., "ECG Personal Identification in Subspaces using Radial Basis Neural Networks," Workshop on Intelligent Data Acquisition and Advanced Computing Systems, pp. 446-451, 2009.

[40] G. Wubbeler, et al., "Verification of Humans using the Electrocardiogram," Pattern Recognition Letters, Vol. 28, pp. 1172-1175, 2007.

[41] I. Odinak et al., "ECG Biometrics: A Robust Short-time Frequency Analysis," Proc. IEEE Int. Workshop Inf. Forensics Security, pp. 1-6, 2010.

[42] S. Fatemian, D. Hatzinakos, "A New ECG Feature Extractor for Biometric Recognition," Proc. 16th Int, Conf. Digital Signal Processing , pp. 1-6, 2009.

[43] M. M. Tawfik, et al., "Human Identification using QT Signal and QRS complex of the ECG," Online J. Electron. Elect, Vol. 3, No. 1, pp. 383-387.

[44] A. L. Goldberger, et al., "PhysioBank, PhysioToolkit and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," Circulation, Vol. 101, No. 23, pp. e215-e220, 2000.

[45] 김성혁, 최규호, 반성범, "InceptionV3 기반 심전도를 이용한 사용자 인식," 한국전자거래학회 & 한국스마트미디어학회 2018 춘계학술대회 논문집, Vol. 7, No. 1, pp. 162-163, 2018.

[46] X. Zhai, C. Tin, "Automated ECG Classification using Dual Heartbeat Coupling based on Convolutional Neural Network," IEEE Access, Vol. 6, pp. 27465-27472, 2018.

Collection @ chosun