

딥러닝 기반의 모바일 얼굴 영상을 이용한 실시간 심박수 측정 시스템

지예림[†], 임서연[†], 박소연[†], 김상하[†], 동서연[†]

Deep Learning-based Real-time Heart Rate Measurement System Using Mobile Facial Videos

Yerim Ji[†], Seoyeon Lim[†], Soyeon Park[†], Sangha Kim[†], Suh-Yeon Dong[†]

ABSTRACT

Since most biosignals rely on contact-based measurement, there is still a problem in that it is hard to provide convenience to users by applying them to daily life. In this paper, we present a mobile application for estimating heart rate based on a deep learning model. The proposed application measures heart rate by capturing real-time face images in a non-contact manner. We trained a three-dimensional convolutional neural network to predict photoplethysmography (PPG) from face images. The face images used for training were taken in various movements and situations. To evaluate the performance of the proposed system, we used a pulse oximeter to measure a ground truth PPG. As a result, the deviation of the calculated root means square error between the heart rate from remote PPG measured by the proposed system and the heart rate from the ground truth was about 1.14, showing no significant difference. Our findings suggest that heart rate measurement by mobile applications is accurate enough to help manage health during daily life.

Key words: Remote Photoplethysmography, Heart Rate, Deep Learning, Mobile Application, Healthcare system

1. 서 론

심박은 사람의 생체, 정신적 상태를 반영하는 중요 생리적 신호로 물리적 상태, 정신 스트레스 등을 측정하는데 사용되는 중요한 지표이다. 일반적으로 심장 활동을 측정하기 위해 주로 사용되는 방법은 심장의 전기적 활동을 기록하는 심전도(electrocardiogram; ECG)나 혈류 변화를 광학센서를 이용하여

측정한 광전용적맥파(photoplethysmography; PPG)가 대표적이다[1,2]. 그러나 기존 방법들은 신체에 센서를 부착하여 신호를 획득하기 때문에 신호를 측정하는 동안에는 움직임을 최대한 제한하므로 일상생활에서 측정하는 데에는 어려움을 겪는다[3]. 이러한 문제를 해결하기 위해 최근 비접촉으로 심장 활동을 측정하는 원거리 광전용적맥파(remote PPG; rPPG) 방식이 시도되고 있다. rPPG는 심장 박동으로 인한

※ Corresponding Author : Suh-Yeon Dong, Address: (04310) Cheongpa-ro 47gil, Yongsan-gu, Seoul, Korea, TEL : +82-2-2077-7120, FAX : +82-2-710-9704, E-mail : sydong@sookmyung.ac.kr

Receipt date : Sep. 28, 2021, Revision date : Oct. 22, 2021
Approval date : Nov. 2, 2021

[†] Dept. of IT Engineering, Division of ICT Convergence Engineering, College of Engineering, Sookmyung Women's University

(E-mail : dpflawl@sookmyung.ac.kr)

(E-mail : imseoyeon0218@sookmyung.ac.kr)

(E-mail : skygreen97@sookmyung.ac.kr)

(E-mail : sanghakim36@gmail.com)

※ This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2021R1F1A1A1052389) and by the Center for Women In Science, Engineering and Technology (WISET) grant funded by the MSIT and the NRF.

얼굴 피부색의 변화를 감지해 심장 활동을 추정하는 일종의 컴퓨터 비전 기반 기술이다[4]. 해당 기술은 센서 부작이 어려운 신생아의 심장 모니터링, 손발 움직임에 제약이 있는 운전자의 상태 평가, 얼굴 스푸핑 탐지 등 다양한 분야에서의 활용 가능성을 시사해오며 발전해왔지만[5,6,7] 이러한 선행 연구들 또한 잘 통제된 실험실 환경에서 고해상도 카메라 혹은 웹캠을 사용해 수집된 영상으로 실험한 경우가 대부분으로, 일상생활에서의 적용 가능성은 여전히 어려운 상황이다[8].

얼굴 영상에서 심박 신호와 관련된 특징을 추출하는 방법은 크게 수작업 특징 방법과 딥러닝 방법으로 나뉜다. rPPG 초기 연구는 대부분 수작업 특징을 사용하는 영상 처리 기반의 방식으로, 얼굴 영상에서 피부색의 변화가 가장 잘 보이는 픽셀을 선택하고, 해당 영역에 포함된 노이즈를 제거하는 필터링 과정을 거치는 것이 일반적이었으며, 얼굴 영상에서 신호 대 잡음 비율(Signal-to-Noise Ratio; SNR)이 높은 PPG 신호를 얻기 위해 이미지 및 신호 처리에 관한 연구들이 주로 진행되었다[9,10,11]. 수작업으로 특징을 추출하는 방식은 피부 영역 검출, 색상 공간 변환, 특징 선택 및 노이즈 필터링 등의 여러 처리 단계가 요구되고[12], 일반화되기 어렵다는 단점이 존재하지만 딥러닝 방식은 얼굴 영상에서 모델이 자동으로 특징을 학습하고 추출한다는 점에서 장점이 있다. 또한 수작업 방식은 정해진 알고리즘에 따라 동작되므로 잡음에 예민하여 실제 환경에서의 적용이 어렵지만 딥러닝은 학습 단계에서 노이즈가 포함된 영상을 함께 학습하여 잡음에 견고하도록 설계될 수 있다[13]. 따라서 최근에는 딥러닝 기법을 이용해 기본적인 전처리만 거친 얼굴 영상에서 PPG 신호를 추출하는 방법으로 연구들이 진행되고 있다[14,15,16].

본 논문에서는 스마트폰에 내장된 카메라만을 이용해 비접촉으로 심박을 측정하는 시스템을 제안하고 실생활에 존재할 수 있는 다양한 조건(손 떨림, 머리 움직임, 빛 변동)에서 실험한 결과를 함께 제시하여 제안하는 시스템이 실제 환경에서도 잘 동작되는지 검증하였다. 이전 연구들에서도 스마트폰으로 촬영된 영상 데이터를 이용해 심박을 측정한 연구가 진행된 바 있지만[17,18], 여러 이미지 처리와 신호 처리를 거친다는 한계점이 존재한다. 본 연구에서는 입력 데이터의 전처리 단계를 최소화하기 위해 딥러

닝 모델 기반의 모바일 애플리케이션을 개발했다는 점에서 이전 연구들과 차별성을 가진다. 또한 기존에는 스마트폰은 데이터 수집용으로 사용되고, 서버에서 수집된 데이터를 받아 딥러닝 모델을 적용하는 것이 일반적이었지만[19,20,21] 본 논문에서는 수집된 데이터를 외부 클라우드나 서버로 전송하지 않고 스마트폰에서 데이터 수집과 컴퓨팅 작업을 동시에 수행하도록 하여 데이터를 보호하고, 네트워크로 인한 지연 시간을 없애 실시간 처리가 가능하도록 하였다. 본 연구에서 제안하는 시스템은 손가락 끝에 펄스 옥시미터로 동시에 측정된 값과 비교함으로써 성능을 평가하였다.

본 논문의 구성은 다음과 같다. 2장에서 관련 연구를 설명하고, 3장에서 제안하는 시스템의 전체 구성을 설명한 다음 학습모델을 구축하는 과정과 이를 기반으로 동작되는 모바일 애플리케이션의 세부 기능 구현을 서술하였다. 4장에서는 제안한 시스템의 성능 평가를 위해 다양한 측정 환경에 따른 실험 결과를 보여주고, 5장에서 결론을 맺는다.

2. 관련 연구

제안된 시스템의 주요 기능인 심박수 측정에 관한 연구에 초점을 두어 선행 연구를 살펴보고자 한다. ECG(electrocardiogram)는 가장 오랫동안 연구된 심박 측정 방식으로 심장의 전기적 활동을 기록하여 높은 정확도로 심박을 측정할 수 있어 주로 의료기관에서 사용된다. ECG 방식은 별도의 측정 장비를 구비해야 하며 몸에 부착된 센서로 인해 활동성 제약이 생겨 일상생활이나 외출상황에서는 활용되기 어렵다. 이러한 한계를 해결하고자 고감도 레이더 센서를 이용하여 심장 박동에 의해 변조되는 레이더 반사 신호로부터 심박 신호를 식별하는 연구들이 제안되었다[22, 23]. 접촉식 센서를 제거함으로써 일상생활에서의 활용가능성을 높였으나 레이더 센서가 설치되어 있는 지정된 공간에서만 측정이 가능하여 이동성 문제를 해결하지 못했다. 또 다른 측정 방식인 PPG(photoplethysmography)는 심장 박동으로 인해 변화하는 혈류량을 시계열 신호로 센싱함으로써 심박 신호를 추정한다. 이는 심장이 이완과 수축을 반복하면서 혈류량이 변화됨에 따라 변하는 빛의 반사량을 광학센서를 이용해 측정하는 방식이다. 상대적

으로 작은 크기의 광학 센서는 웨어러블 기기에 내장될 수 있어 이동성 문제를 해결하였지만[24], 피부가 센서에 접촉되어야 하므로 비대면 생체신호 측정 시스템으로는 적합하지 않다.

제안된 시스템에서 사용한 측정 방식인 rPPG(remote PPG)는 기존 PPG 방식의 원리를 기반으로 하되, RGB 카메라를 이용해 빛의 반사량의 변화로 인한 피부색의 변동을 포착하여 비접촉으로 심박을 측정하는 방식이다. 최근에는 스마트폰에 내장된 카메라 성능이 향상됨에 따라 스마트폰 전면카메라만을 이용해 미묘한 피부색의 변화를 포착할 수 있게 되면서 모바일 얼굴 영상으로부터 rPPG 신호를 측정하는 연구가 진행되었다. Kim의 연구[25]에서는 스마트폰 전면 카메라에서 추출한 얼굴 영상의 색상 성분을 분리하여 푸리에 변환을 적용해 심박을 포함한 생체 징후를 측정하였다. Qiao의 연구[26]에서는 피부색의 변화가 가장 잘 나타나는 녹색 채널의 평균 픽셀 값을 구하고 정제된 신호를 얻기 위한 노이즈 필터링, 독립 성분 분석 과정을 수행하여 rPPG 신호를 측정하는 시스템을 제안하였다. Lomaliza의 연구[17]에서는 스마트폰을 이용한 손떨림 문제를 해결하기 위해 배경 특징의 평균점을 사용해 얼굴 특징의 위치 조정을 통해 머리 움직임에 견고한 rPPG 신호 추출을 시도했다.

위에서 설명한 바와 같이 심박 측정 방식은 측정의 편리성을 증대시키고 측정 신호의 오차를 줄이는 연구들이 소개되며 발전되어 왔다. 가장 최근 제안된 rPPG 방식은 추가적인 센서나 웨어러블 기기 없이 스마트폰만을 이용해 비접촉으로 심박 신호 측정이 가능함을 보였다. 하지만 스마트폰을 이용한 기존

rPPG 연구들은 이미지와 신호 처리를 통해 얻어지는 수작업 특징을 기반으로 하는 연구들이 주를 이루고 있으며, 딥러닝 기법을 사용해 수작업 처리 과정을 최소화하고 잡음에 견고한 시스템을 개발한 연구는 거의 없었다. 본 연구에서는 실제 사용 환경에서 발생할 수 있는 다양한 시나리오를 포함한 데이터를 심층 신경망에 학습시켜 스마트폰에 탑재함으로써 실생활에 적용 가능한 실시간 심박수 측정 시스템을 제안한다.

3. 연구 방법

본 논문에서는 추가적인 하드웨어 장치 없이 오직 스마트폰 카메라에서 수집된 얼굴 영상을 사용해 실시간 심박 측정이 가능한 시스템을 제안한다. Fig. 1은 제안하는 시스템의 구성을 보여주며, 학습모델을 구축하는 부분(Fig. 1(a))과 학습된 모델을 모바일에 탑재해 심박을 추정하고 결과를 데이터베이스에 저장하는 부분(Fig. 1(b))으로 나뉜다.

3.1 3D 컨볼루션 신경망

얼굴 비디오 영상을 학습하기 위해 PhysNet[16] 모델의 구조를 차용하였다. 해당 모델은 3D 컨볼루션 신경망을 기반으로 한다. 일반적으로 단일 이미지에서 특징을 추출하는 목적으로 사용할 때에는 2D 컨볼루션 신경망이 사용되나, 여러 이미지 프레임으로 구성된 비디오의 경우에는 3D 볼륨 형태의 데이터가 입력되므로 연속되는 프레임 사이에 존재하는 시간적 특징을 효과적으로 추출하기 위해 3D 컨볼루션 신경망을 사용한다. 3D 컨볼루션 신경망과 유사

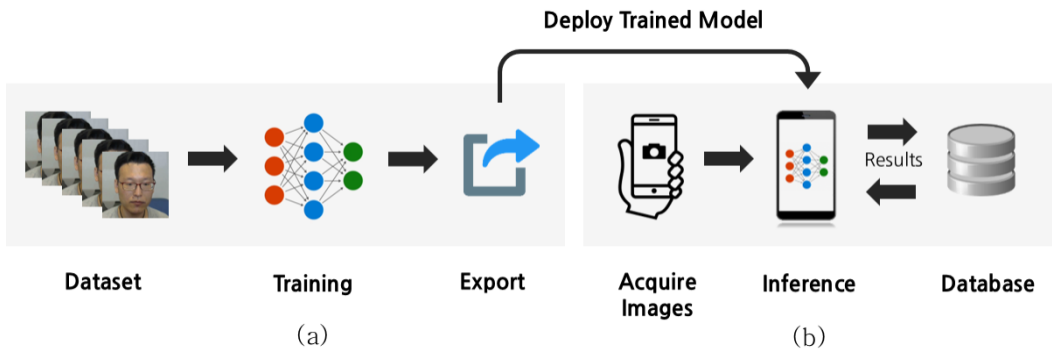


Fig. 1. Flowchart for deep learning-based mobile application system. (a) Model training and exportation for deployment to the smartphone and (b) Inference on the smartphone in real-time using the trained model.

한 기능을 하도록 2D 컨볼루션 신경망과 LSTM을 함께 사용하는 방법 또한 존재하지만, 시간적 정보를 장기 기억하는 것에 있어서는 3D 컨볼루션 신경망이 더 효과적이므로[16] 3D 컨볼루션 신경망 기반의 PhysNet을 기반으로 학습모델을 설계하였다.

얼굴 비디오에서 시계열 정보인 rPPG 신호를 추출하는 PhysNet 모델의 기본 구조는 Fig. 2와 같다. 모델 내부의 3D 컨볼루션 연산은 인접한 이미지 프레임을 연결하여 생성되는 입력에 3D 커널을 적용하므로 프레임 길이에 의해 연산량이 크게 좌우되기 때문에 학습모델은 GPU가 아닌 CPU 환경에서 구동되므로 연산량을 줄이기 위한 사전 설계가 필요하다. 본 연구에서는 얼굴 비디오에서 rPPG를 추정하는 데에 필요한 최소한의 프레임을 설정하고자 프레임 길이를 변경하며 성능을 확인하였고, 실험 결과 프레임 길이가 90 프레임보다 작아질 경우 성능이 크게 하락해 전체 프레임 길이(T)는 90으로 설정하였다. 각 프레임의 크기 또한 피부색 변화가 포착되는 수준까지의 해상도를 고려해 80×80로 조정하였다. 최종적으로 모델의 입력은 4×3×90×80(batch_size×channels×T×height×width)의 5차원 입력 텐서로 구성되어 제한된 모바일 환경에서도 연산이 가능하도록 모델 구축 단계에서 고려해 주었다.

3.2 3D 컨볼루션 신경망 학습

모델 학습을 위해 사용된 데이터셋은 VIPL-HR [27]이다. VIPL-HR은 107명의 피험자에 대해 다양한 실험 조건에서 촬영된 약 30초 길이의 얼굴 비디오와 생체 신호(SpO2, BVP, HR)가 함께 기록된 데이터셋이다. 실험 조건에는 안정적인 시나리오, 모션 시나리오, 말하는 시나리오, 어두운 시나리오, 밝은 시나리오, 장거리 시나리오, 운동 시나리오로 총 9가

지의 시나리오에서 데이터가 기록되어있다. 사용된 촬영 기기에는 웹캠 외에도 스마트폰이 포함되어 있어 본 연구에서 지향하는 실제 환경과 높은 유사성을 가지므로 학습 데이터셋으로 사용하였다. 이렇게 확보된 데이터셋은 총 2378개의 영상 데이터로, 1903개는 학습 데이터, 475개는 테스트 데이터로 사용되었다. 입력 영상의 크기 및 프레임 길이는 3.1 장에서 기술한 값으로 설정되었다.

학습 과정에서는 얼굴 영상에서 추정되는 rPPG 신호와 실제 심박 신호인 BVP(blood volume pressure) 신호의 차이를 최소화하는 데에 적합한 손실 함수가 설정되어야 하므로 손실 함수는 변수 간 선형 관계를 나타내는 대표적인 수치인 피어슨 상관계수를 활용하였다. 두 신호가 +1에 가까운 양의 선형 관계가 되는 방향으로 학습되어야 하므로 손실 함수는 1에서 피어슨 상관계수를 빼 음의 피어슨 상관계수(1)로 설정하여 학습과정 동안 손실이 0에 가까워지도록 하였다. 이외 학습 파라미터로 배치 크기는 4, 학습률은 0.0001, 최적화 함수는 Adam optimizer를 사용하였고, 25 에포크 이후엔 검증 손실값이 0.23에서 떨어지지 않아 학습을 조기종료하였다.

$$Loss = 1 - \frac{\sum_{i=1}^T x_i y_i - \sum_{i=1}^T x_i \sum_{i=1}^T y_i}{\sqrt{(T \sum_{i=1}^T x_i^2 - (\sum_{i=1}^T x_i)^2)(T \sum_{i=1}^T y_i^2 - (\sum_{i=1}^T y_i)^2)}} \quad (1)$$

식 (1)에서 T는 전체 신호의 길이인 프레임 길이를 나타내며, i는 각 비디오 프레임 순차적으로 가리킨다. x는 모델 예측 결과인 rPPG 신호이고, x_i는 각 프레임에 대한 rPPG 신호를 나타낸다. 마찬가지로 y는 ground truth 값인 BVP 신호이고, y_i는 각 프레임에 대한 BVP 신호를 의미한다.

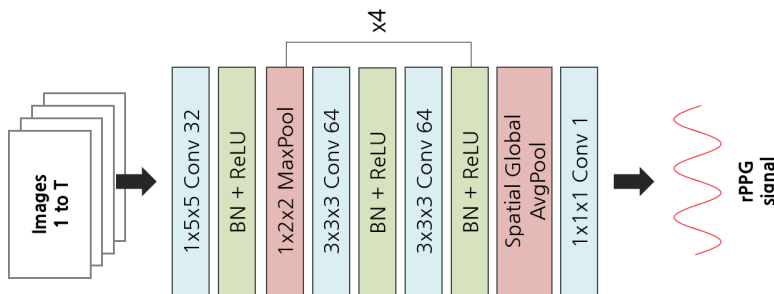


Fig. 2. Architecture of the PhysNet-3DCNN.

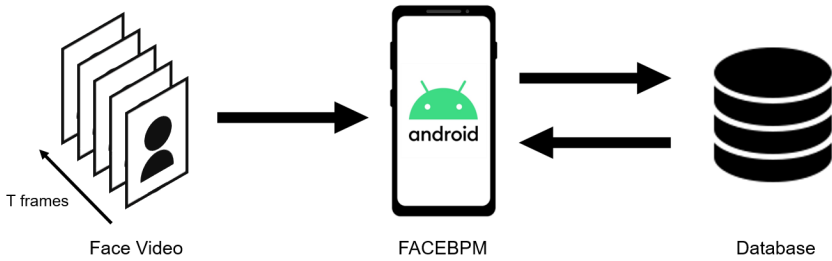


Fig. 3. Mobile application process.

최종적으로 학습된 모델의 매개변수는 모바일 내에서의 연산을 위해 저장되어 추출되고, 추출된 모델은 학습 모드가 아닌 추론 모드로 설정해주었다. 추론 모드에서는 드롭아웃을 비활성화하고, 배치 정규화가 실행되지 않도록 하여 학습 시 저장된 분산과 표준편차를 사용하도록 설정된다. 추론 모드로 변환된 모델은 모바일 프로젝트에 탑재되어 심박수 측정을 위한 모듈로써 동작하게 된다.

3.3 딥러닝 기반 심박수 측정 모바일 애플리케이션 구현

학습된 모델이 스마트폰에서 실시간으로 데이터를 받아 추론할 수 있도록 안드로이드를 이용해 모바일 애플리케이션을 구현했다. 애플리케이션의 동작 흐름은 Fig. 3과 같다. 스마트폰의 전면 카메라를 통해 촬영한 얼굴 영상이 모바일 애플리케이션에 탑재한 모델의 입력으로 들어가고, 계산된 심박수는 SQLite를 사용해 디바이스 내부 데이터베이스에 저장된다.

모바일 애플리케이션을 실행하면 Fig. 4(a)에서 보이는 화면에서 이름, 나이, 성별, 현재 상태를 입력하고, 심박수 측정이 시작된다. 입력받은 사용자 정보는 개인 맞춤 심박 구간을 계산하고 제공하기 위해 데이터베이스에 저장된다. 심박수 측정 시 전면 카메라가 구동되어 사용자의 얼굴이 포함된 90개의 이미지 프레임을 입력으로 받아 심박수를 추정한다(Fig. 4(b)). 이때 실시간 영상에서 얼굴 인식을 위해 Google ML Toolkit을 사용하였고, 카메라 화면에 가이드라인을 제시해 사용자의 편의성과 측정 결과의 정확성을 높였다.

내부적으로 입력 영상은 pytorch_android_torchvision 라이브러리[28]를 이용해 텐서로 생성된다. 안드로이드 내에서 생성된 텐서는 1×3×80×80 4차원

텐서 형식인데, 모바일에 탑재된 PhysNet 모델의 입력은 1×3×90×80×80 5차원 텐서이므로 4차원 텐서 리스트를 5차원 텐서로 변환하는 것이 필요하다. 그러나 파이토치 모바일은 5차원 텐서 생성을 지원하지 않으므로 5차원 텐서를 생성하는 파이썬 스크립트를 별도로 만들어 안드로이드 내에서 모듈로 구현해 변환을 진행했다. 변환된 5차원 텐서는 모델의 입력으로 들어가 연산이 수행된다.

측정된 rPPG 신호는 피크를 검출하여 PPI(Peak to Peak Interval)를 계산한다. PPI는 이전 peak에서 현재 peak까지의 시간 간격(peak(i+1)-peak(i), rPPG 신호 상의 심장 박동 간격을 의미한다. 심박수(HR)는 앞서 구한 PPI를 이용한 수식(2)을 사용해 bpm(beats per minute) 단위로 계산된다. 심박수는 정확성을 높이기 위해 3번을 연이어 측정하며, 최종 심박수는 측정된 3개 심박수 평균으로 기록된다. 최종 심박수는 Fig. 4(a)에서 입력받은 사용자 정보와 함께 디바이



Fig. 4. Screen shots of the mobile application: (a) requiring user information, (b) capturing face images, and (c) displaying heart rate prediction result.

스 내부 데이터베이스에 저장되고 Fig. 4(c)에서 출력된다.

$$HR[bpm] = \frac{60}{(peak(i+1) - peak(i))} \quad (2)$$

심박수 측정이 끝나면 자동으로 심박수 측정 기록 조회 화면으로 전환된다(Fig. 4(c)). 심박수 측정 기록 조회를 통해 디바이스 내부 데이터베이스에 접근해 사용자의 심박수 측정 기록을 불러온다. 사용자는 심박수, 상태, 측정 날짜가 포함된 자신의 심박수 측정 기록을 확인할 수 있고, 최근 5일간의 일일 평균 심박수 기록은 기록 상단에 시각화되어 제공된다.

심박수뿐만 아니라 개인마다 특화된 심박수 구간을 제공하여 계산된 심박수가 정상 심박수 범위에 있는지, 건강 상태를 판단할 수 있는 구간을 함께 제공한다. 심박수 구간은 사전에 입력받은 사용자의 나이, 성별, 상태를 기반으로 개개인마다 다르게 계산되어 제공된다. 안정 시 심박수 구간은 미국 질병 통제 예방센터의 성별, 연령에 따른 안정 시 심박수 구간[29]을, 활동(운동) 시 심박수 구간은 미국 스포츠 의학회(ACSM)에서 권장한 최대 심박수의 백분을 범위[30]를 활용하였다. 최대 심박수는 공식 (3)을 사용해 계산했다[31]. 계산된 심박수는 제공된 맞춤 심박수 구간에 해당되는 상태에 따라 아이콘 색상이 결정된다. 안정되거나 정상 상태라면 초록색, [정상 범위, 정상 범위+20] 구간은 노랑색, 그 외의 범위는 위험 상태라고 판단하여 빨간색으로 상태를 보여준다.

$$HR_{MAX} = 220 - age \quad (3)$$

age는 사용자의 나이를, HR_{MAX} 는 최대 심박수를 의미한다.

4. 실험

제안하는 시스템이 일상에 존재할 수 있는 4가지 실험 조건(움직임과 조명 변화가 없음, 손 떨림, 머리 움직임, 조명 변화가 있음)에서도 잘 동작하는지 검증하기 위해 실험을 진행하였다. 얼굴 영상은 Galaxy A51 5G(SM-A516N) 스마트폰을 사용했고, 카메라 해상도는 2944×2208 픽셀, 데이터 샘플링 속도는 30 fps로 설정했다. 기준값을 측정하기 위해서는 맥박 산소포화도 측정기기 CMS50D-BT(SHENZ, China)를 사용하였다. 실험 환경은 Fig. 5와 같이 구성되어 본 연구에서 제안하는 모바일 애플리케이션과 맥박

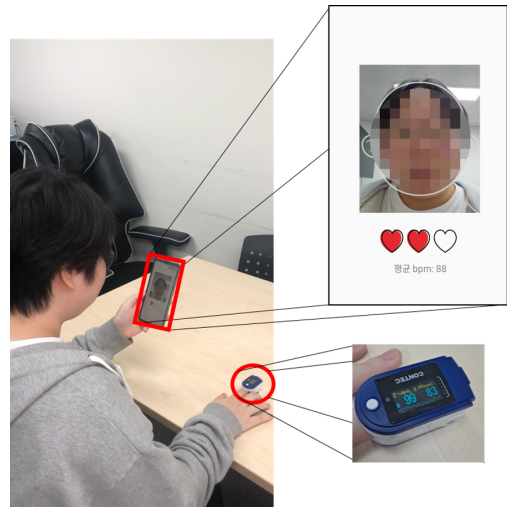


Fig. 5. System evaluation environment; heart-rate measurement of the proposed system and pulse oximeter.

산소포화도 측정기기를 동시에 구동하여 심박수 측정값을 비교해보았다.

실험에는 20~25세의 9명의 건강한 성인이 참여하였다. 피험자들은 의자에 앉은 상태에서 한쪽 손으로는 직접 스마트폰을 얼굴을 향해 들고 다른 손에는 손가락 산소포화도 측정기를 착용해 rPPG와 PPG 값을 동시에 측정하였다. 피험자는 4개의 서로 다른 조건에서 각각 3분 동안 실험을 진행했다. 4개의 조건은 정적인 자세에서 조명 변화가 없는 상황(Baseline), 손을 움직이는 상황(M1), 머리를 움직이는 상황(M2), 주변 조명 빛이 변하는 상황(L1)으로 구성하여 진행했다.

Table 1에서는 전체 피험자의 차례로 rPPG와 PPG로부터 계산된 심박수의 평균, 최소, 최대값을 나타내며 얼굴로부터 측정되는 rPPG 값이 손가락으로부터 측정된 PPG 값보다 약간 높게 측정되는 것을 알 수 있다. 머리를 움직이는 상황(M2)에서는 오히려 정적인 상황(Baseline) 보다 두 값의 차이가 근소하였다. Fig. 6는 Baseline 환경에서 안면 PPG와 손가락 끝 PPG를 사용해 얻은 심박수에 대한 산점도표를 통해 측정값의 분포 및 선형관계를 확인하였다. 대부분의 데이터는 양의 선형관계에 가까운 것으로 보이나, 선형관계에서 벗어난 이상치도 확인할 수 있다. 이상치를 보인 피험자를 확인해 본 결과, 메이크업을 한 상태의 피험자들의 측정값들이 이상치를 보

Table 1. Comparison of HR measurement between rPPG (proposed) and PPG(groundtruth).

Condition	Method	Mean HR (std)	Min HR	Max HR
Baseline	rPPG	89.97 (7.87)	73	108
	PPG	82.77 (8.83)	67	100
M1	rPPG	90.16 (9.14)	71	108
	PPG	81.94 (9.26)	64	98
M2	rPPG	87.06 (10.43)	67	111
	PPG	83.16 (9.02)	68	99
L1	rPPG	89.61 (9.82)	74	104
	PPG	81.13 (8.00)	68	98

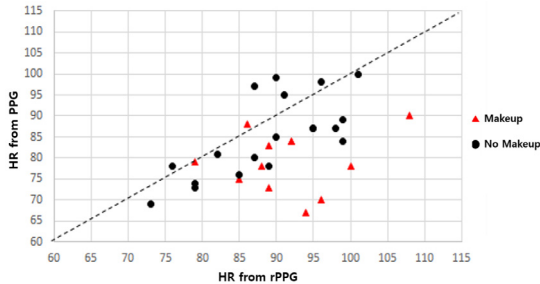


Fig. 6. Distribution of HR estimates with and without makeup.

이는 경우가 대부분이었다. 메이크업은 얼굴 영상에서 혈액 산소 포화도 추출을 어렵게 하기 때문에 rPPG 측정에 부정적인 영향을 미치는 것으로 알려져 있어[32] 메이크업 유무에 따라 rPPG 성능이 크게 좌우된다는 것을 본 연구를 통해서도 확인할 수 있었기 때문에 추가적으로 메이크업 유무에 따라 Makeup 그룹과 No Makeup 그룹으로 나누고, 전체 그룹에 대하여 상관관계와 RMSE 분석을 진행해 PPG 측정에 부정적인 영향을 끼치는 요인이 있는 환경과 없는 환경에서의 시스템 성능을 비교해 보았다.

Table 2는 4개의 실험 조건에서 rPPG와 PPG로부터 계산된 심박수 사이의 상관계수와 RMSE를 보여준다. Makeup 그룹과 No Makeup 그룹 간의 비교에

Table 2. Evaluation under different noisy conditions: correlation coefficient and Root Mean Square Error (RMSE) between HR from rPPG and PPG.

		Makeup	No Makeup	ALL
Correlation Coefficient	Baseline	0.16	0.71	0.46
	M1	0.12	0.53	0.23
	M2	-0.38	0.53	0.22
	L1	0.05	0.33	0.30
RMSE	Baseline	15.38	7.73	11.31
	M1	19.53	9.08	14.08
	M2	16.36	9.99	12.84
	L1	14.12	12.36	13.07

서는, Makeup 그룹에서 모든 조건에서 상관관계가 거의 보이지 않지만 No Makeup 그룹은 정적인 상황(Baseline)에서 0.71로 높은 상관관계를 보였다. 또한 손을 움직이는 상황(M1)과 머리를 움직이는 상황(M2)에서도 각각 0.53으로 양의 상관관계가 있는 것으로 나타났다. 전체 그룹으로 보았을 때는 No Makeup 그룹보다는 전반적으로 낮아진 결과가 나타났다. Baseline 조건에서 0.46으로 가장 높은 양의 상관관계를 보였고, 다른 조건에서는 0.2~0.3 사이의 상관관계를 나타냈다. rPPG와 PPG로부터 계산되는 HR의 차이를 확인할 수 있는 지표인 RMSE를 살펴보면, 전체 그룹(ALL)을 기준으로 Baseline에서는 11.31, M1에서는 14.08, M2에서는 12.84, L1에서는 13.07로, 다양한 환경 조건들 간의 RMSE의 편차는 약 1.14로 큰 차이를 보이지 않는다. 반면, Makeup 그룹의 평균 RMSE는 16.35, No Makeup 그룹의 평균 RMSE는 9.79로 Makeup 그룹에 의해 전체 피험자의 평균 RMSE가 증가되었음을 확인하였다. 또한 PhysNet과 같이 입력 영상에서 스스로 특징을 학습하는 end-to-end 모델을 활용한 연구에서 제시된 RMSE와 실험 결과(RMSE)를 비교하였다(Fig. 7). DeepPhys[14]는 13.8, I3D[15]는 15.9로 M1을 제외한 모든 실험 조건에서 기존 연구보다 실제 심박 신호와의 오차가 더 낮은 것으로 나타났으며, 가장 오차가 컸던 M1에서도 I3D보다 더 낮은 결과를 보여 제안된 시스템에서 측정된 심박수는 신뢰할 수 있는 결과임을 알 수 있다. DeePhys, I3D 모델은 학습 데이터와 동일한 데이터셋에서 일부를 테스트 데이터로 사용해 측정된 결과이며, 제안한 시스템은 스마트

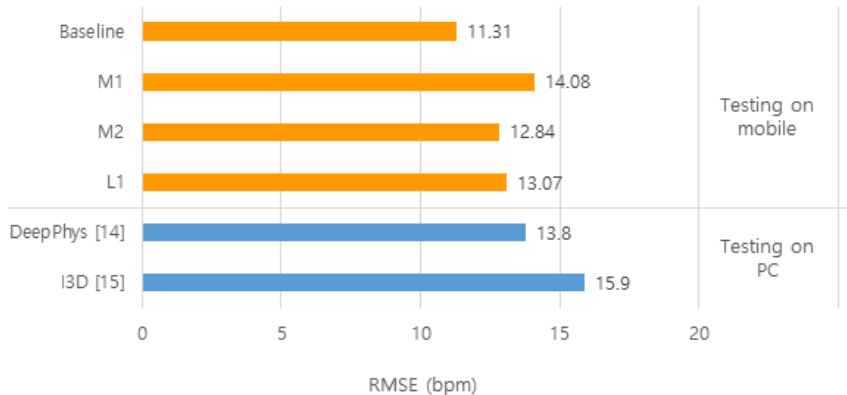


Fig. 7. Comparison of RMSE results of end-to-end models trained with VIPL-HR dataset.

폰에서 수집한 데이터로 테스트했다는 점에서 차이가 있다. 따라서 정확한 비교 분석은 아니지만 이를 통해 탑재된 모델이 기존 학습 데이터와 양상이 다른 실제 데이터에도 안정적으로 구동될 수 있음을 확인할 수 있다.

위의 결과로부터 본 연구에서 제안하는 시스템의 성능은 움직임보다는 조명 환경에 영향을 많이 받고, 환경 조건보다는 메이크업 여부에 더 큰 영향을 받는다는 것을 알 수 있다. 결과적으로, 얼굴 영상을 측정하는 동안에는 가능한 한 움직임과 조명 변화는 최소화하고 메이크업을 하지 않은 상태여야 최적의 심박 측정이 가능함을 알 수 있다.

5. 결 론

본 논문에서는 PC 환경에서 학습한 3D 컨볼루션 신경망 모델을 모바일에 탑재함으로써 별도의 장치나 센서 없이 실시간으로 심박 측정이 가능한 모바일 애플리케이션을 개발하여 실생활에서 편리하게 사용 가능한 심박 측정 시스템을 제안하였다. 스마트폰 전면 카메라를 통해 얻은 사용자의 얼굴 영상을 이용해 모바일에 탑재된 딥러닝 모델을 통해 실시간으로 심박을 계산한다. 제안한 시스템의 성능을 평가하기 위해 손가락 산소포화도 측정기를 사용해 동시에 측정해본 결과, 다양한 환경 조건들 간의 RMSE의 편차가 약 1.14로 큰 차이를 보이지 않아 일상생활 속에서 심박수 측정이 가능함을 입증하였다. 추가적으로 사용자의 화장 여부에 따른 성능 변화 또한 관찰, 분석함으로써 적정 사용 환경에 대한 가이드를 제시하

여 사용자가 보다 정확하게 해당 시스템을 사용할 수 있도록 안내할 수 있도록 하였다. 본 연구에서 제안하는 시스템은 비대면 헬스케어 시스템의 필요성이 높아진 언택트 환경에서 비접촉식으로 시간과 장소에 관계없이 심박수를 측정할 수 있으므로 심혈관 질환 분야를 포함한 전반적인 헬스케어에 유용하게 활용될 것으로 기대된다.

REFERENCE

- [1] J.A. Dawson, C.O. Kamlin, and C. Wong, "Changes in Heart Rate in the First Minutes after Birth," *Archives of Disease in Childhood. Fetal and Neonatal Edition*, Vol. 95, No. 3, pp. 177-181, 2010.
- [2] N.V. Hoan, J. Park, S. Lee, and K. Kwon, "Real-time Heart Rate Measurement based on Photoplethysmography using Android Smartphone Camera," *Journal of Korea Multimedia Society*, Vol. 20, No. 2, pp. 234-243, 2017.
- [3] X. Chen, J. Cheng, R. Song, Y. Liu, R. Ward, and Z. J. Wang, "Video-Based Heart Rate Measurement: Recent Advances and Future Prospects," *IEEE Transactions on Instrumentation and Measurement*, Vol. 68, No. 10, pp. 3600-3615, 2019.
- [4] W. Verkruysse, L.O. Svaasand, and J.S. Nelson, "Remote Plethysmographic Imaging using Ambient Light," *Optics Express*, Vol.

- 16, No. 26, pp. 21434–21445, 2008.
- [5] M. Van Gastel, B. Balmaekers, S.B. Oetomo, and W. Verkruyse, “Near-continuous Non-contact Cardiac Pulse Monitoring in a Neonatal Intensive Care Unit in Near Darkness,” *Proceedings of Optical Diagnostics and Sensing XVIII*, Vol. 10501, 2018.
- [6] P. Huang, B. Wu, and B. Wu, “A Heart Rate Monitoring Framework for Real-World Drivers Using Remote Photoplethysmography,” *IEEE Journal of Biomedical and Health Informatics*, Vol. 25, No. 5, pp. 1397–1408, 2021.
- [7] C. Yao, S. Wang, J. Zhang, W. He, H. Du, and J. Ren, et al., “rPPG-Based Spoofing Detection for Face Mask Attack using Efficientnet on Weighted Spatial-Temporal Representation,” *Proceeding of IEEE International Conference on Image Processing (ICIP)*, pp. 3872–3876, 2021.
- [8] D. Qiao, F. Zulkernine, R. Masroor, R. Rasool, and N. Jaffar, “Measuring Heart Rate and Heart Rate Variability with Smartphone Camera,” *Proceeding of 22nd IEEE International Conference on Mobile Data Management (MDM)*, pp. 248–249, 2021.
- [9] G. de Haan and V. Jeanne, “Robust Pulse Rate from Chrominance-based rPPG,” *IEEE Transactions Biomedical Engineering*, Vol. 60, No. 10, pp. 2878–2886, 2013.
- [10] W. Wang, S. Stuijk, and G. de Haan, “A Novel Algorithm for Remote Photoplethysmography: Spatial Subspace Rotation,” *IEEE Transactions Biomedical Engineering*, Vol. 63, No. 9, pp. 1974–1984, 2016.
- [11] W. Wang, A.C. den Brinker, S. Stuijk, and G. de Haan, “Algorithmic Principles of Remote PPG,” *EEE Transactions on Biomedical Engineering*, Vol. 64, No. 7, pp. 1479–1491, 2017.
- [12] A. Dasari, S. Ananyananda, L.A. Jeni, and C.S. Tucker, “Evaluation of Biases in Remote Photoplethysmography Methods,” *NPJ Digital Medicine*, Vol. 4, No. 1, p. 91, 2021.
- [13] M. Sabokrou, M. Pourreza, X. Li, M. Fathy, and G. Zhao, “Deep-HR: Fast Heart Rate Estimation from Face Video Under Realistic Conditions,” *Expert Systems with Applications*, Vol. 186, Article 115596, 2021.
- [14] W. Chen and D. McDuff, “DeepPhys: Video-based Physiological Measurement using Convolutional Attention Networks,” *Proceeding of the European Conference on Computer Vision (ECCV)*, pp. 349–365, 2018.
- [15] J. Carreira and A. Zisserman, “Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6299–6308, 2017.
- [16] Z. Yu, X. Li, and G. Zhao, “Remote Photoplethysmograph Signal Measurement from Facial Videos using Spatio-temporal Networks,” *Proceedings of the British Machine Vision Conference*, pp. 1–12, 2019.
- [17] B.P. Yan, C.K. Chan, and C.K. Li, “Resting and Postexercise Heart Rate Detection From Fingertip and Facial Photoplethysmography Using a Smartphone Camera: A Validation Study,” *JMIR mHealth and uHealth*, Vol. 5, No. 3, e33, 2017.
- [18] J.-P. Lomaliza and H. Park, “Improved Heart-Rate Measurement from Mobile Face Videos,” *Electronics*, Vol. 8, No. 6, pp. 663, 2019.
- [19] V. Jindal, “Integrating Mobile and Cloud for PPG Signal Selection to Monitor Heart Rate during Intensive Physical Exercise,” *Proceeding of International Conference on Mobile Software Engineering and Systems (MOBILE Soft)*, pp. 36–37, 2016.
- [20] G. Zhang, Z. Mei, Y. Zhang, X. Ma, B. Lo, and D. Chen, “A Non-invasive Blood Glucose Monitoring System Based on Smartphone PPG Signal Processing and Machine Learning,” *IEEE Transactions on Industrial Informatics*, Vol. 16, No. 11, pp. 7209–7218, 2020.

- [21] M.J. Lee, Y.M. Kim, and Y.M. Lim, "Masked Face Temperature Measurement System Using Deep Learning," *Journal of Korea Multimedia Society*, Vol. 24, No. 2, pp. 208-214, 2021.
- [22] S. Wang, A. Pohl, T. Jaeschke, M. Czaplík, M. Kony, and S. Leonhardt, et al., "A Novel Ultra-Wideband 80 GHz FMCW Radar System for Contactless Monitoring of Vital Signs," *Proceeding of Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4978-4981, 2015.
- [23] N. Malešević, V. Petrovic, M. Belic, C. Antfolk, V. Mihajlovic, and M. Jankovic, "Contactless Real-time Heartbeat Detection via 24 GHz Continuous-Wave Doppler Radar Using Artificial Neural Networks," *Sensors*, Vol. 20, No. 8, pp. 2351, 2020.
- [24] C. Yang, C. Veiga, J.J. Rodríguez-Andina, J. Fariña, A. Iñiguez, and S. Yin, "Using PPG Signals and Wearable Devices for Atrial Fibrillation Screening," *IEEE Transactions on Industrial Electronics*, Vol. 66, No. 11, pp. 8832-8842, 2019.
- [25] D. Kim, J. Kim, and K. Lee, "Real-time Vital Signs Measurement System using Facial Image Data," *Journal of Broadcast Engineering*, Vol. 26, No. 2, pp. 132-142, 2021.
- [26] D. Qiao, F. Zulkernine, R. Masroor, R. Rasool, and N. Jaffar, "Measuring Heart Rate and Heart Rate Variability with Smartphone Camera," *Proceeding of IEEE International Conference on Mobile Data Management (MDM)*, pp. 248-249, 2021.
- [27] X. Niu, H. Han, S. Shan, and X. Chen, "VIPL-HR: A Multi-Modal Database for Pulse Estimation from Less-Constrained Face Video," *Proceedings of the Asian Conference on Computer Vision*, pp. 562-576, 2018.
- [28] PyTorch(2019), <https://pytorch.org/mobile/home> (accessed September 17, 2021).
- [29] Y. Ostchega, K.S. Porter, J. Hughes, C.F. Dillon, and T. Nwankwo, "Resting Pulse Rate Reference Data for Children, Adolescents, and Adults: United States, 1999-2008," *National Health Statistics Reports*, Vol. 41, pp. 1-16, 2011.
- [30] K. Norton, L. Norton, and D. Sadgrove, "Position Statement on Physical Activity and Exercise Intensity Terminology," *Journal of Science and Medicine in Sport*, Vol. 13, No. 5, pp. 496-502, 2010.
- [31] N. Zhu, J.R. Suarez-Lopez, S. Sidney, B. Sternfeld, P.J. Schreiner, M.R. Carnethon, et al., "Longitudinal Examination of Age-Predicted Symptom-Limited Exercise Maximum HR," *Medicine and Science in Sports and Exercise*, Vol. 42, No. 8, pp. 1519-1527, 2010.
- [32] W. Wang and C. Shan, "Impact of Makeup on Remote-PPG Monitoring," *Biomedical Physics and Engineering Express*, Vol. 6, No. 3, 035004, 2020.



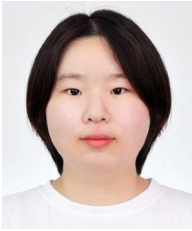
지 예 립

2021년 3월~현재 숙명여자대학교 IT공학과 학석사 연계 과정 재학
 관심분야: 머신러닝/딥러닝, 모바일 헬스케어, 생체 신호처리, 감성공학



김 상 하

2021년 3월~현재 숙명여자대학교 IT공학과 학석사 연계 과정 재학
 관심분야: 머신러닝/딥러닝, 헬스케어, 생체 신호처리



임 서 연

2021년 9월~현재 숙명여자대학교 IT공학과 학석사 연계 과정 재학
 관심분야: 머신러닝/딥러닝, 헬스케어, 생체 신호처리



동 서 연

2010년 2월 한국과학기술원 전기 및 전자공학과(공학사)
 2011년 8월 한국과학기술원 전기 및 전자공학과(공학석사)
 2016년 2월 한국과학기술원 전기 및 전자공학과(공학박사)



박 소 연

2020년 9월~현재 숙명여자대학교 IT공학과 학석사 연계 과정 재학
 관심분야: 머신러닝/딥러닝, 헬스케어, 생체 신호처리

2016년 3월~2016년 7월 한국과학기술원 정보전자연구소 박사후연구원
 2016년 8월~2018년 8월 한국과학기술연구원 의공학연구소 바이오닉스연구단 박사후연구원
 2018년 9월 숙명여자대학교 IT공학과 조교수
 관심분야: 뇌-컴퓨터 인터페이스, 모바일 헬스케어, 머신러닝/딥러닝, 생체 신호처리 등