영상기반 비접촉식 PPG 신호 취득을 위한 3D-CNN 설계

김태완[1], 염찬욱[2], 곽근창[3]
[1]조선대학교 전자공학과 IT-Bio융합시스템전공 석사과정
[2]전남대학교 IT-Bio융합시스템농업교육연구단
[3]조선대학교 전자공학부 교수
gyd03002@naver.com, walt18@naver.com, kwak@chosun.ac.kr

Designing a 3D-CNN for Non-Contact PPG Signal Acquisition Based on Video Imaging

Tae-Wan Kim[1], Chan-Uk Yeom[2], Keun-Chang Kawk[3]
[1]Electronic Engineering, IT-Bio Convergence System, Chosun University
[2]Center for IT-Bio convergence system agriculture Chonnam University
[3]School of Electrical Engineering, Chosun University

요 약

생체 신호를 분석하여 사용자의 건강과 정신 상태를 예측하고, 관련 질병에 관해 예방하는 연구가 늘어나고 있다. 생체 신호 중 심박은 사람의 육체, 정신적인 상태를 반영하는 대표적인 신호이지만 기존의 접촉 패드를 통한 ECG나 광학 센서를 통한 PPG로 심박을 예측할 때는 구속적인 환경이 필요하여 일상적인 상황 속에 적용하기 어려웠다. 이러한 단점을 해결하고자 본 논문은 UBFC-RPPG 데이터셋의 동영상 프레임을 RGB 채널마다 다른 가중치를 적용하는 전처리를 하여 학습 데이터의 크기를 줄이면서 정확도를 높이고, 3D-CNN을 활용한 딥러닝으로 순간적인 영상에서도 PPG 신호를 예측할 수 있도록 1초 전처리 영상을 학습한 후, 신호를 예측하는 것을 목표로 한다. 이렇게 비접촉식으로 취득된 신호는 더 다양한 환경에서의 감정분류, 우울증 진단, 질병 감지 등 다양한 분야에 활용될 수 있다.

1. 서론

최근 사람의 건강과 정신 상태를 예측하여 대면할 수 없는 상황이나 고령의 환자의 경우 예방과 관리를 위해 자신의 건강 상태를 실시간으로 확인할 수 있는 헬스케어 시스템이 관심받고 있다. 이와 같은 분야의 연구 중 대표적인 것이 생체 신호를 분석하여 사람의 건강이나 정신적인 상태를 분석하는 일이다. 생체 신호로는 심장의 전기적인 활동을 기록하는 심전도(ECG, electrocardiogram), 두뇌의 신경세포들의 전기적 활동을 측정하는 뇌전도(EEG, electroencephalogram), 신경과 근육의 전기적 활동을 측정하는 근전도(EMG, electromyography) 등이 있다. 그 중이 심전도를 통해 분석되는 심박의 경우 사람의 상태를 반영하는 중요한 요소이지만 접촉 패드를 부착해야 하고, 그에 맞는 장치가 필요하다. 다른 심박 측정 방법으로 광학 센서를 통한 광용적맥파(PPG, photoplethysmography)의 경우 피부에 빛을 투과시켜 심장 박동으로 생겨나는 혈류의 변화를 광학 센서를 통해 측정하는 것으로 이는 접촉 패드 부착할 필요가 없는 점과 측정 장치 간소화의 이점이 있지만, 접촉 센서를 착용해야 해서 구속적인 단점을 해결하지 못했다[1]. 기존의 ECG와 PPG로 심박을 일상적으로 활용하기에는 어려움이 있어서, 최근 심박 신호를 비접촉으로 PPG 신호를 측정하는 방법들이 연구되고 있다[2]. 컴퓨터 비전 기술을 이용한 원격 PPG(rPPG, remote PPG)는 얼굴 피부색의 변화를 감지하여 심박과 관련된 신호로 변환하는 기술이며, 센서 부착이 어려운 환자의 심장 모니터링, 운전자의 상태 변화에도 사용 가능하고[3], 센서를 부착하는 것을 불편해하는 동물에게도 사용할 수 있다[4]. 초기에는 rPPG 신호를 직접 관심 영역(ROI)을 지정하고 필터링을 하여 신호를 얻었지만, 딥러닝 기술이 발전하면서 전처리 된 동영상을 모델에 입력하게 되면 학습하여 사용자의 생체 신호를 취득할 수 있게 되었다.

본 논문에서는 기존의 연구되고 있던 딥러닝을 사용한 PPG 취득을 순간적인 상황에서도 활용할 수 있도록, 짧은 시간의 영상에서 PPG 신호를 예측하는 1초 얼굴 영상을 3D-CNN으로 학습하여 생체 신호를 예측하는 것을 목표로 한다.

2. 관련 연구

Unakafov의 경우 rPPG를 통한 심박수 예측을 진행하였는데[2], 사람의 얼굴 중 관심 영역(ROI)을 지정하고 해당 영역의 RGB 신호를 적절한 전처리 과정과 iPPG(Imaging Photoplethysmography) 추출 기술로 RGB 신호를 PPG신호로 변환하였다. 추출 기술은 초록색 신호의 파장이 헤모글로빈이 가장 잘 흡수하여 심박과 밀접한 관계가 있음을 적용한 기술로, 취득된 신호는 필터링 과정과 주파수 변환(DFT, CWT)을 적용하여 분당 심박수(BPM)를 예측하였다. 또한 rPPG 기술을 영장류에도 적용하였다[4]. 사람의 경우는 접촉 센서를 통해서 생체 신호를 취득할 수 있지만, 동물의 경우에는 고도의 훈련을 받지 않은 이상 접촉 패드를 장착하는 과정을 지시하기 어렵다. 그래서 이전까지 동물에게는 센서를 몸속에 침습시키거나 마취를 진행하고 취득하였는데, rPPG 기술을 적용함으로 패드나 센서 장착의 어려움을 해결할 수 있었다.

Jaromir의 경우, CNN을 활용하여 얼굴 ROI 영역을 검출하도록 하고, 해당 영역의 RGB 원시 신호를 추출하였다. 추출된 신호에서 PPG 추출 기법(G, ICA, POS 등)으로 필터링 후 예측된 심박과 LSTM을 통해 원시 신호를 학습 시킨 후, 예측된 신호를 바탕으로 예측된 심박을 비교하였다[5].

3. 환경 및 방법

3.1 3D-CNN 네트워크

기존의 연구되고 있던 rPPG를 실시간으로 사용할 수 있도록 1초 단위의 얼굴 영상을 3D CNN을 사용해 시계열적인 특징을 학습하였다.

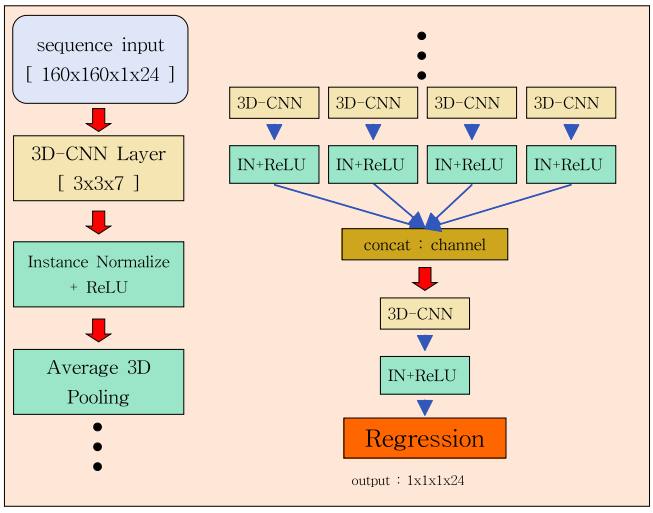

그림 1. 영상 데이터를 학습하기 위한 3D-CNN 네트워크

3D 컨볼루션과 채널별로 인스턴스 정규화(Instance Normalize)를 하고, 마지막에 채널별로 독립적으로 컨볼루션과 정규화 후 1차원으로 합쳐서, 채널별로 얻어진 값을 한꺼번에 컨볼루션을 진행하였다. 위 방법으로 채널끼리 가지고 있는 정보를 공유할 수 있도록 하여 모델의 채널 수를 낮추었고, 전체적인 모델의 학습 파라미터 수를 낮추었다.

3.2 이미지 전처리

효과적인 학습을 위해서 동영상을 프레임으로 전환하고 전처리를 진행하였다. 배경을 제외하고 얼굴 이미지에 집중하기 위해서 이미지를 얼굴 중심으로

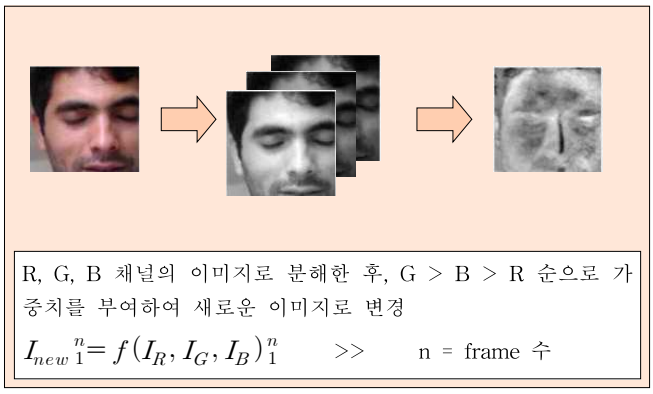

그림 2. 이미지 전처리 과정

편집하고, PPG 신호를 추출할 때, RGB의 경우 초록색 파장이 심박에 대한 정보를 많이 포함하고 있어서, 각 차원의 비율을 조정하여 이미지를 전처리하고, 과적합을 막기 위해 데이터 증감을 사용하였다. 이미지는 그림 2와 같이 R, G, B 채널로 나누고, 가중치를 부여하여 하나의 채널을 가지는 이미지로 변환하였다. 이는 학습의 시간을 단축하고, 정확도를 높여주었다.

4. 실험

4.1 UBFC-RPPG 데이터셋

모델 학습 데이터셋은 UBFC-RPPG[6]을 활용하였다. 이는 8명의 참가자가 카메라 앞에서 움직이지 않는 상태에서 촬영된 1~2분 영상이다. 위 데이터 중 학습 시간과 데이터의 품질(움직임 등)을 고려하여 총 5명의 데이터를 사용했다. 학습, 검증, 테스트 데이터는 각각의 실험자들의 동영상 프레임을 75%, 10%, 15%로 나누어서 학습하였다.

4.2 실험 결과

그림 3와 같이 센서를 통해 얻은 PPG신호와 모델

을 통해 학습하여 예측한 PPG신호를 겹쳐서 비교하기 위해 신호를 1초마다 잘라서 파형의 모양을 분석하였다.

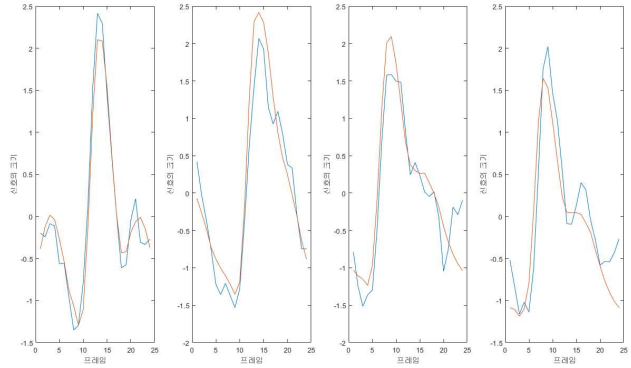

그림 4 학습된 네트워크로 예측된 PPG(파란색), 센서 PPG(주황색)

결과적으로 파형이 어긋나는 부분이 존재하지만, 파형의 주기성과 피크값이 크게 차이가 나지 않게 출력되었다. 하지만 신호 예측을 진행하면 결과별로 정확성이 차이가 있다는 것을 그림 4에서 확인할 수 있었다.

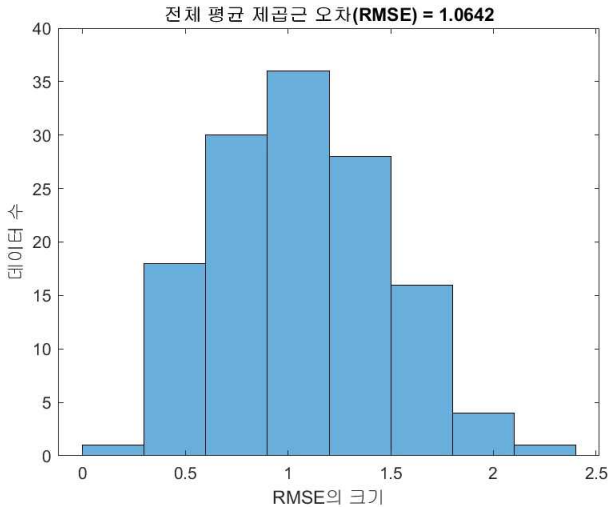

그림 4 예측과 실제 값의 평균제곱오차(RMSE)의 히스토그램

5. 결론

본 논문에서는 도보 중, 운전 중과 같은 구속적인 환경에서 생체 신호를 취득할 수 없는 경우에 활용할 수 있는 영상 기반 비접촉 실시간 PPG 신호 취득을 목표로 연구가 진행되었다. 컴퓨터 비전 기반의 생체 신호 취득과 같이 복잡한 과정을 거치지 않고도, 딥러닝 기술을 사용하여 간단한 이미지 전처리로 PPG 신호 예측을 할 수 있었다.

해당 기술을 응용하여 비대면 상황이나 스마트카 등에서 감정분류나 우울증 검출, 산소 포화도 예측과 같은 분야에 응용 가능하여, 다양한 헬스케어 시스템에 응용할 수 있을 것으로 기대된다.

감사의글

이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2017R1A6A1A03015496). 또한 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신산업진흥원의 지원을 받아 수행된 헬스케어 AI 융합 연구개발 사업임(No. S1601-20-1041)

참고문헌

[1] 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.

[2] Unakafov AM. “Pulse rate estimation using imaging photoplethysmography”, Biomedical Physics & Engineering Express, 2018, 4(4),045001

[3] 지예림, 임서연, 박소연, 김상하, 동서연 “딥러닝 기반의 모바일 얼굴 영상을 이용한 실시간 심박수 측정 시스템”, 한국멀티미디어학회논문지, 제 24권 제 11호, pp. 1481-1491, 2021

[4] Unakafov AM, Moeller S, Kagan I, Gail A, Treue S, Wolf F, “Using imaging photoplethysmography for heart rate estimation in non-human primates”, PLoS ONE, 2019, 14(1) e0211518

[5] Jaromir Przybyło, “A deep learning approach for remote heart rate estimation”, Biomedical Signal Processing and Control, April 2022, 74, 103457

[6] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, J. Dubois, "Unsupervised skin tissue segmentation for remote photoplethysmography", Pattern Recognition Letters, 2019, colume 124, 89-90