

딥러닝 모델 경량화 기술을 활용한 실시간 심박수 측정

- 원격 광혈류측정 방식 기반 -

김종민¹, 표정식², 김세진², 박우영³, 김형신²

서울대학교 수학교육과¹, 서울대학교 데이터사이언스대학원², 서울대학교 건설환경공학부³

Real-time Heart Rate Monitoring using Deep Learning Model Lightening Techniques

- Based on Remote Photoplethysmography -

Jongmin Kim¹, Jeongsik Pyo², Sejin Kim², Woo Young Park³, and Hyung-Sin Kim²

Department of Mathematics Education, Seoul National University, Korea¹

Graduate School of Data Science, Seoul National University, Korea²

Department of Civil and Environmental Engineering, Seoul National University, Korea³

hyungkim@snu.ac.kr

Abstract

The importance of remote health monitoring increased during the Covid-19 pandemic. In this context, remote photoplethysmography (rPPG), a video-based heart rate monitoring, is one of the rising technologies, since heart rate is a significant and important vital sign for healthcare. However, various studies on rPPG using deep learning algorithms have recently been published without considering application on edge devices. Adopting temporal shift-convolutional attention network (TS-CAN) algorithm, this paper presents the possibility of a lightweight rPPG model which is optimized for edge devices. Thus, achieving less computational cost without sacrificing accuracy, lightweighting made it possible to run on an edge device environment.

Keywords: telehealth, remote photoplethysmography, heart rate, temporal shift module, attention, edge device

1. 연구 배경

심박수는 의료전문가들이 의학적 상태 전반을 확인하는 중요한 신체 신호이다. 심박수를 측정하는 기존 방법으로는, 빛으로 심박수의 변화를 측정하는 광혈류 (photoplethysmography, PPG) 측정 방법이 있다. 우리의 귀나 손가락 끝 같이 신체에서 얇은 부분은, 빛의 일정량이 통과하면, 해당 부분이 빨갛게 변하는데, 광혈류측정은 이러한 원리를 토대로, 피부에서 반사된 빛의 양을 분석해서 혈액의 흐름 및 심박수를 측정한다.

한편, 코로나19의 등장으로 원격 의료에 대한 관심이 의료진 및 환자들뿐만 아니라 대중에게도 높아졌다. 기존의 PPG 방법은 심박수를 측정해주는 특정 접촉식 기기가 있는 현장에서만 측정이 가능하기 때문에, 이에 대한 원격 광혈류측정 (remote photoplethysmography, rPPG) 방법이 등장하였다.

rPPG는 영상의 얼굴 프레임 영역에서 혈류량에 의한 피부의 미세한 색상 변화를 RGB 신호로 추출하여 심박수를 예측한다. 초기의 rPPG 방법은 머리의 움직임, 조명 변화에 따른 잡음 (noise)에 민감하게 반응하여 정확한 측정이 어려웠다 [1]. 이러한 문제를 해결하기 위해서 최근에는 딥러닝 기술이 적용되고 있다. 하지만 현재의 rPPG 연구는 경량화까지는 많은 고려가 이루어지지 않고 있다.

본 연구에서는 딥러닝을 활용한 rPPG 방식의 모델 경량화 과정을 다루었다. 모델 경량화를 통해, rPPG 측정이 edge device에서도 가능해지므로 개인 의료정보 보안성이 높아질 것으로 기대된다.

2. 연구 방법

본 연구의 모델은 크게 두 개의 모듈로 이루어진다. 첫 번째는 데이터 영상에서 얼굴에 해당하는 영역만 탐지하여 추출하는 face

detection 모듈이고, 두 번째는 추출된 영상으로부터 혈류량의 변화를 RGB 신호로 입력 받아, 심박수를 예측하는 TS-CAN 모듈이다. 이는 그림 1과 같이 진행된다.

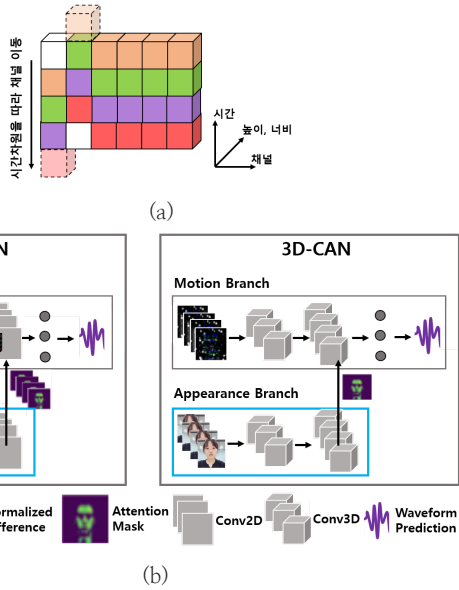


[그림 1. rPPG를 통한 심박수 측정 과정]

Face detection 모듈은 residual neural network (ResNet) 으로 구현하였다. ResNet은 신경망의 레이어가 깊어질수록 데이터 학습이 어려워지는 기존의 convolutional neural network (CNN) 의 문제를 해결한 알고리즘이다. 원본 영상에서 얼굴 프레임만 추출하기 위하여 사용하였다.

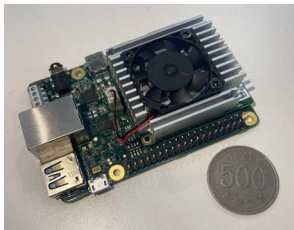
실시간 혈류량 변화에 따른 심박수 예측모델은 TS-CAN이다. 이미지 특징 추출에 2D 컨볼루션 (convolution) 연산을 이용하는데, 영상에서 특징 추출 시, 시간상의 정보 (temporal information) 를 고려하기 위해서는 3D 컨볼루션 연산을 해야한다. 하지만, 경량화가 필요한 환경에서 3D 컨볼루션 연산은 전력제한 및 계산 복잡도가 높아 적합하지 않다[2]. 이러한 문제를 해결하기 위해, 그림 2(a)와 같이 시간차원을 따라 채널을 이동시키는 temporal shift (TF) 모듈을 사용하였다. 그림 2(b)에서 알 수 있듯이, 2D 컨볼루션 연산으로 3D 컨볼루션 연산

효과를 얻을 수 있으며, attention mask를 통해 이상치를 제거하여 얼굴 프레임에 더욱 초점을 맞춘다.



[그림 2. TS-CAN 개념도]
(a) TF모듈 개념도, (b) TS-CAN과 3D-CAN 비교

Pytorch나 TensorFlow의 라이브러리를 사용한 기존 연구들과 달리, 본 연구에서는 TensorFlowLite의 라이브러리를 활용하여 추가적 경량화를 했으며, 이를 통해 edge device인 Google Coral Dev Board 환경에서의 구현도 가능하게 하였다.



[그림 3. Coral Dev Board]

3. 연구 결과

(1) 모델 경량화

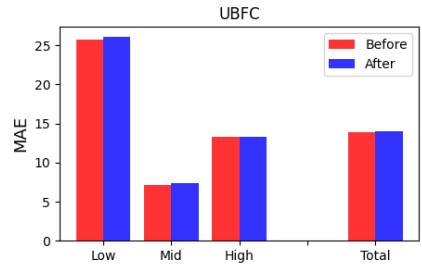
face detection 모델의 경량화에 대해서는 기존 연구된 바가 많으므로 정확한 변인 통제를 위해 blood volume pulse를 예측하는 TS-CAN 모듈에 대해 경량화를 수행하였다. 경량화를 하지 않은 TS-CAN 모듈의 용량은 2.34MB 였으나 경량화된 모델의 용량은 0.75MB로서 기존 대비 32.07% 수준이다.

(2) 경량화 전후 성능 비교

본 연구에서는 AFRL dataset으로 pre-train된 모델[2]을 경량화하여 Coral Dev Board에 이식한 뒤, UBFC dataset과 VIPL-HR-V2 dataset에 대하여 성능을 측정하였다. dataset의 sample들을 ground truth bpm에 따라 low(bpm 77 미만), middle(77이상 90미만), high(90이상) group으로 나누어 각 그룹별 MAE 및 dataset 전체에 대한 MAE를 분석하였다.

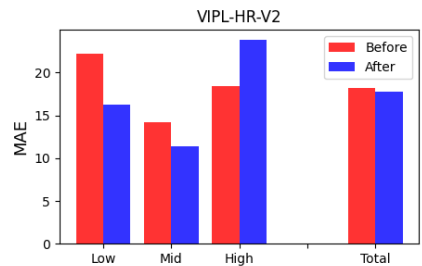
그림 4에 나타난 바와 같이 UBFC dataset에 대하여 테스트한 결과, MAE는 경량화 전 13.94, 경량화 후 14.06으로 경량화하였음에도 비슷한 성능을 유지하였다. 다만 BPM 77 미만

low group에 대해서는 BPM을 과대평가하는 경향을 보여 MAE가 변환 전후 모두 약 26으로 높았다. Low group을 제외하면 MAE는 변환 전 11.73, 변환 후 11.81로 나타났다.



[그림 4. 경량화에 따른 모델의 성능 변화(MAE) : UBFC]
(좌) 경량화되지 않은 모델의 성능, (우) 경량화된 모델의 성능

그림 5에 나타난 바와 같이 VIPL-HR-V2 dataset에 대하여 테스트한 결과, MAE는 경량화 전 18.24, 경량화 후 17.81로서 경량화에 따른 성능 하락은 크지 않았으나 정확도가 낮다. 구체적으로 확인해본 결과, 조명이 어두우면서 움직임이 큰 경우 error가 상당히 크게 나타났다. 이는 본 모델이 심박에 따라 얼굴의 색이 미세하게 변화하는 것으로부터 inference 하는데, 어두운 조명아래에서 크게 움직이는 경우 얼굴에 그늘지는 영역이 달라지면서 정확도에 나쁜 영향을 미치는 것으로 보인다.



[그림 5. 경량화에 따른 모델의 성능 변화(MAE) : VIPL-HR-V2]
(좌) 경량화되지 않은 모델의 성능, (우) 경량화된 모델의 성능

(3) Inference delay 분석

평균적으로 영상 길이 대비 inference time은 40.39%였다. 따라서 실시간 스트리밍 inference도 가능하다. 한편, 전체 inference time 중 ① face detection + pre-process, ② pulse prediction, ③ post-process의 각 단계가 차지하는 비중은 평균적으로 각각 47.59%, 48.81%, 3.60%였다.

4. Acknowledgements

This work was supported by Create-Pioneering Researchers Program through Seoul National University.

5.참고 문헌

[1] Wang, Wenjin, et al. "Algorithmic principles of remote PPG." *IEEE Transactions on Biomedical Engineering* 64.7 (2016): 1479–1491.

[2] Liu, Xin, et al. "Multi-task temporal shift attention networks for on-device contactless vitals measurement." *Advances in Neural Information Processing Systems* 33 (2020): 19400–19411.