Design of Quaternion-based Neural Networks for Non-contact Heart Rate Monitoring

Yeong-Hun Yoon[1], Kwang-Kee Lee[2], Jung-Whan Lee[3], You-Sang Nam[3], Cheol-Soo Park[3*]
[1]*Department of Information Convergence, Kwangwoon University / Seoul, Korea*
[2]*TVSTORM / Seoul, Korea*
[3] *Department of Computer Engineering, Kwangwoon University / Seoul, Korea*
yyhosmium@gmail.com, kwangkeelee@gmail.com, hjn040281@gmail.com, unsc.odst59@gmail.com, parkcheolsoo@kw.ac.kr *Corresponding author

2022 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia) | 978-1-6654-6434-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCE-ASIA57006.2022.9954813

Abstract

Various researches have been conducted to estimate remote photoplethysmography from video streams, which are mostly based on a convolutional neural network model. Although these yielded meaningful performance, it takes long to train the model and produce the inference during the testing process since there are numerous weights and they are less scalable to the multiple tasks. Therefore, we propose a quaternion-based convolutional neural network model to design a lighter model with similar or better performance compared with the conventional model in the real domain. To achieve this goal, we extend the original model to the quaternion domain. The experimental results demonstrate that the new model is lighter keeping the similar performance as that of the conventional model.

Keywords: Quaternion Convolutional Neural Network, Photoplethysmography(PPG), remote sensing

1. Introduction

Photoplethysmography (PPG), which provides indirect measurement of heart activity through the changes in blood flow, has been measured in a contactless manner. The previous studies [1, 2] demonstrated significant performance of the non-contact sensing method compared with the contact one. However, the size of the previous models was too large and require lots of computational costs. In addition, scalability, conducting multiple tasks simultaneously, could be hardly expected.

In this paper, we propose a non-contact PPG prediction model based on quaternion CNN with solving the issues of the previous models. In the quaternion domain networks [3], correlation information is calculated among nodes, which could be the main difference from the real domain model. With this structure, the model could extract the features related to vital signs from video streams with high accuracy. We extend the existing real-domain model into a quaternion domain and demonstrate a similar performance with light-sized networks.

For the implementation of the proposed model, we applied quaternion-based neural networks to the rPPGNet [1] model. Additionally, a series of pipeline from the preprocessing process to the postprocessing process was constructed to compare the real domain model and the quaternion domain model.

2. Proposed Method

For an experiment to test the model`s performance, we implemented the pipeline shown below. The experiments made use of the PURE dataset.

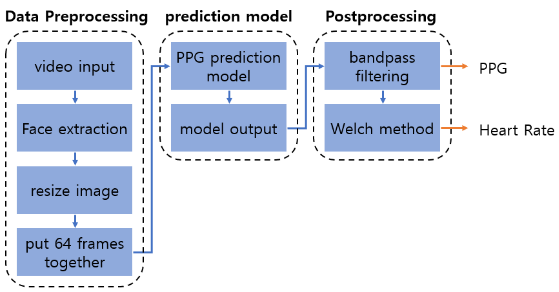


Figure 1: The flow chart of the proposed model

2.1. Data Preprocessing

In the data preprocessing process (Figure 1 Data Preprocessing), only the face part of the image is extracted and resized to 128x128 size. And 64 continuous face images processed in this manner are used as a single data (3x64x128x128, 3 is RGB three channels).

However, if a face detection method such as haar cascade is applied, the size of the detected face area(ROI) varies on a continual basis. When resizing, the shapes of the facial images become inconsistent as a result of this alteration. Therefore, the face was extracted from the image using MediaPipe (Figure 2) [4, 5] with a high detection rate and no changes to the ROI. Additionally, the performance was increased by removing the eyes and mouth from the face image [5].

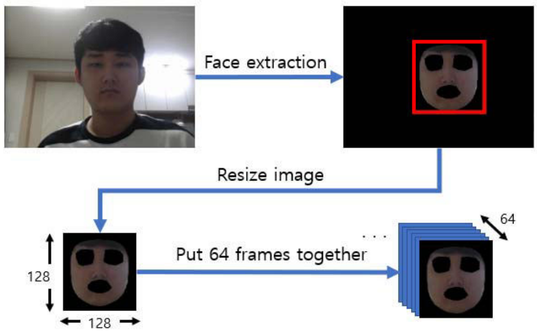
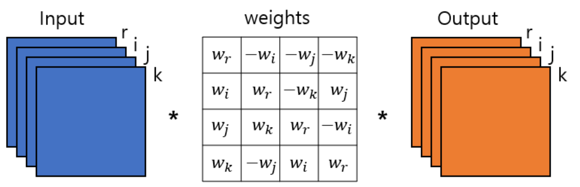


Figure 2: Data Preprocessing of input images using mediapipe

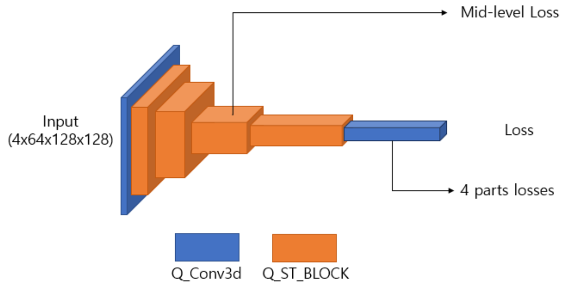
2.2. Model Configuration

For comparison, the original rPPGNet [1] was built initially. According to the paper [1], the Adam optimizer was used, with the learning rate set to 1e-4 and the epoch set to 15. During the experiment, it was confirmed that these hyperparameter settings were correct.

In addition, the quaternion domain modified rPPGNet was built. rPPGNet`s layers are all altered to quarternion-based convolution layers. The quaternion convolution procedure is carried out using the Hamilton product, and the weights are presented in Figure 3 (a).Unlike the original model, the skin segmentation and attention processes were eliminated, and all layers were converted to quaternion domain layers (Figure 3 (b)). The Adam optimizer was applied, and after multiple learning cycles, the optimal hyperparameter values were discovered (learning rate=5e-4, epoch=12). And since it is a quaternion domain, the input requires four values (0, R, G, B) rather than three values (R, G, B). We tried using alternative values instead of (0, R, G, B), but the results were poor.



(a)



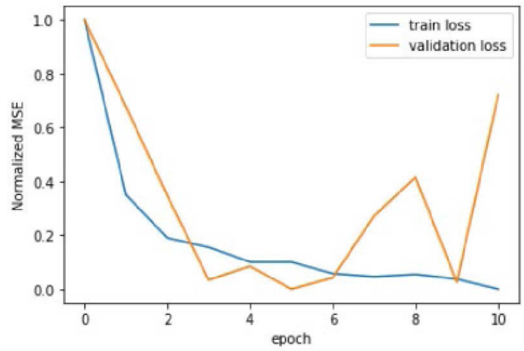
(b)

Figure 3: (a)Quaternion convolution calculation, (b)Quaternion-based rPPGNet

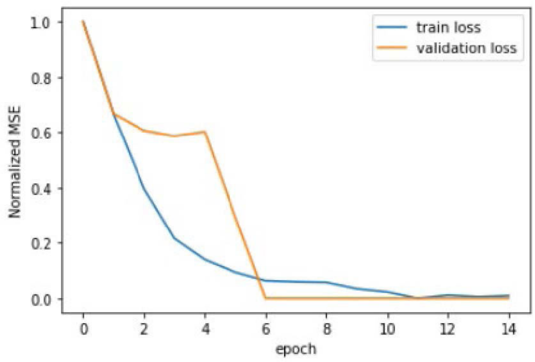
2.3. Model Training

MSE loss was used, and model learning was conducted with an early stopping option. Because the quaternion model output four signals, only the PPG signal of the first index was used for evaluation.

When loss is the lowest but epoch is low, as seen in figure 1 (a), learning is frequently incomplete. Even if the loss increases slightly in this situation, it is preferable to use higher epoch`s weights.



(a)



(b)

Figure 4: (a)Learning curve of quaternion model, (b)Learning curve of the original model

2.4. Postprocessing

Postprocessing was carried out in accordance with the procedure outlined in the postprocessing section of Figure 1.

To reduce noise and generate a proper PPG signal, a 0.5~4.0 Hz 2nd order bandpass filter is applied [6] to the model`s output. The heart rate can also be estimated by applying the Welch method or Fast Fourier Transform to the PPG signal and selecting the Hz with the highest power [5].

3. Results

Table 1 shows the results obtained. Considering the error of the HR measurement method, the performance is comparable.

MAE and RMSE HR assessment measures, while Pearson Correlation Coefficient is a PPG assessment measure.

Table 1: Test results of the proposed model

model \ data	movement X	movement O
	MAE RMSE PCC	MAE RMSE PCC
Original	2.909 9.045 0.812	5.958 18.489 0.744
Quaternion	0.969 5.222 0.791	7.142 15.095 0.696

We classified the data used for learning and testing into two categories, one without movement and one including movement. Through this, we tried to see whether the quaternion model is affected by the movement in two cases.

As a result, when there was no movement, it outperformed the conventional model. And in the case of including movement, almost comparable performance was demonstrated considering the error in the method of obtaining HR in PPG.

4. Conclusion

This research conducted an experiment in which the real domain model was converted to a quaternion-based model and its performance was compared to that of the conventional model. As a result, it was discovered that the quaternion model could yield comparable or better performance while requiring less computational cost than the original model.

In this paper, there is a limitation in that the test was conducted with only one model. Following that, the test will be conducted with more models. We have found that the quaternion-based CNN model works better in the previous article [7], and we will focus on a test of whether this is also in the PPG and HR prediction models. Additionally, there is a limitation that the characteristic of the quaternion`s four inputs could not be used appropriately by using them as (0, R, G, B) because it could not discover four suitable values for them.

In future studies, we will see if the quaternion model can outperform the conventional models when the size of the quaternion model is identical. Furthermore, we will investigate whether the correlation among network nodes and the quaternion domain model has a positive effect on multi-task learning. Also, we will discover four suitable values for quaternion input.

5. Acknowledgement

This research was supported by the MSIT (Ministry of Science and ICT), under the National Program for Excellence in SW (2017-0-00096), supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation), and the Institute of Information & Communications Technology Planning & Planning & Evaluation (IITP) grant funded by the Korean government (MSIT). (No. 2021-0-00900, Adaptive Federated Learning in Dynamic Heterogeneous Environment).

References

[1] Z. Yu, W. Peng, X. Li, X. Hong, and G. Zhao, "Remote Heart Rate Measurement from Highly Compressed Facial Videos: An End-to-End Deep Learning Solution With Video Enhancement", *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, pp. 151-160, July 2019.

[2] W. Chen, M. Daniel, "DeepPhys: Video-Based Physiological Measurement Using Convolutional Attention Networks", *Computer Vision – ECCV 2018*, Springer International Publishing, pp. 356-373, October 2018.

[3] T. Parcollet, M. Ravanelli, M. Morchid, G. Linarès, C. Trabelsi, R. De Mori and Y. Bengio, "Quaternion Recurrent Neural Networks", *International Conference on Learning Representations (ICLR) 2019*, ICLR, September 2018.

[4] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. Chang, M. Yong, J. Lee, W. Chang, W. Hua, M. Georg and M. Grundmann, "MediaPipe: A Framework for Perceiving and Processing Reality", *Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR) 2019*, CVPR, 2019.

[5] G. Boccignone, D. Conte, V. Cuculo, A. D'Amelio, G. Grossi, R. Lanzarotti and E. Mortara, "pyVHR: a Python framework for remote photoplethysmography", *PeerJ Computer Science vol.8*, PeerJ, p.e929, April 2022.

[6] X. Liu, J. Fromm, S. Patel, D. McDuff, "Multi-Task Temporal Shift Attention Networks for On-Device Contactless Vitals Measurement", *Proceedings of the 34th International Conference on Neural Information Processing Systems*, Curran Associates Inc., p.12, 2020.

[7] X. Zhu, Y. Xu, H. Xu, and C. Chen, “Quaternion Convolutional Neural Networks”, *Computer Vision – ECCV 2018*, Springer International Publishing, pp. 631-647, 2018.