❒   1326

Implementation of 1D convolutional neural network for improvement remote photoplethysmography measurement

Riza Agung Firmansyah[1], Yuliyanto Agung Prabowo[1], Titiek Suheta[1], Syahri Muharom[1]
[1]Department of Electrical Engineering, Faculty of Electrical Engineering and Information Technology,
Institut Teknologi Adhi Tama Surabaya (ITATS), Surabaya, Indonesia

Article Info

ABSTRACT

Remote photoplethysmography (rPPG) for non-contact heart rate measurement has been widely developed and shows good development. However, motion artifact due to changes in illumination and subject movement is still the main problem. Especially when measurements are taken in real conditions. In these conditions, it will be vulnerable to rPPG signal readings with poor signal quality. So, in this paper, it is proposed to classify the signal quality using one dimensional convolutional neural network (1D CNN). The classification is carried out based on the extraction of the temporal features of the rPPG signal that has been obtained from the plane orthogonal to skin algorithm and the magnitude of the subject's movement when measured. The classification results are entered into a compensated network if the signal obtained shows moderate quality. The compensated network will provide a more accurate estimate of hr value. The test was carried out using a dataset of 10 subjects, each measured with 3 different types of illumination. In the experiments conducted, the system's performance showed an improvement compared to the POS algorithm alone. The experiment found that the mean absolute error measurement was 2.78, and the mean error was relative at 3.67%.

Corresponding Author:

Riza Agung Firmansyah
Department of Electrical Engineering, Faculty of Electrical Engineering and Information Technology
Institut Teknologi Adhi Tama Surabaya
Street of Arief Rachman Hakim no. 100, Sukolilo, Surabaya, Indonesia
Email: rizaagungf@itats.ac.id

1. INTRODUCTION

Measurement of vital signs has an important role in detecting people's health. These measurements are generally carried out using a measuring instrument that is in contact with the subject. In some conditions, such as the COVID-19 pandemic, the use of measuring instruments that are in direct contact with the subject is avoided as much as possible. So, we need vital signs measuring instrument that can be used without touching the subject. The existing non-contact measurements mostly use computer vision techniques. Measurement of body temperature and respiratory rate can be done based on thermal imaging [1]-[3]. Measurement of heart rate can use thermal imaging [4], [5], or red, green, and blue (RGB) color imaging [6], [7]. In this paper, focuses on heart rate measurement using RGB color imaging.

Remote photoplethysmography (rPPG) is a technique for measuring heart rate values based on changes in facial skin color [6], [7]. Changes in skin color occur due to changes in the volume of blood flowing due to the heart's pumping. This technique is quite widely developed with several variations. Remote photoplethysmograph performance is generally affected by the quality of the image captured by the camera. The image quality is strongly influenced by the lighting quality and the analyzed image's region of interest

(ROI). Lighting quality is affected by changes in illumination and ROI quality is affected by subject movement and motion artifacts. If these two problems arise, the heart rate signal obtained will be inaccurate. Based on these problems, the biggest challenge in rPPG research is the problem of resistance to changes in illumination and motion artifacts. So we need a way to improve measurement accuracy even though the image obtained still has interference.

Some rPPG technique uses light reflection techniques from the surface of the facial skin. The images used are RGB images with chrominance techniques [6], plane orthogonal to skin (POS) [7], parenchymal blood volume (PBV) [8], adaptive pulse projection (APP) [9] and several other techniques. To eliminate the dependence on light illumination, other image sources can be used other than RGB. The images used include near infra-red (NIR) images [10], [11], dan thermal imaging [5], [12]. However, NIR and thermal cameras are relatively expensive compared to RGB. Adaptive filters can cope well with changes in light illumination. The adaptive filters include adaptive pulse projection (APP) by [9], adaptive spatiotemporal homomorphic filtering (ASTHF) [13], ensemble empirical mode decomposition (EEMD) [14], and signal quality attenuation [15]. adaptive filter noise cancellation active noise control (ANC) on recursive least squares (RLS) can overcome motion artifacts of the subject [16]. The filter approach can produce good results but is only limited to one problem. The combination of the Spatial-spectral-temporal filter can produce good accuracy but must use two types of cameras, namely RGB and NIR [17]. The filter approach can produce good results but is limited to one problem.

Machine learning algorithm to improve the extracted signal using CNN can increase the reading accuracy [18]-[22]. The use of CNN sourced directly from RGB images can also be used to increase accuracy [23]-[25]. The use of machine learning is good enough to overcome all existing problems but is limited to datasets that have been trained. The use of machine learning can be used to determine the quality of the obtained rPPG signal [26]. With this technique, only signals with good quality can be output as measurement results. A compensated network can estimate signals with poor quality further [18]. Compensated networks only estimate the signal with poor quality. So that the resulting error value will be smaller and closer to the reference. Based on the description above, this study is proposed to combine the two types of machine learning, namely signal quality classification (SQC) and compensated networks.

The proposed of this paper is combine signal quality classification and compensating networks. Signal quality classification (POS algorithm result) uses 1D convolutional neural network (CNN) to get 3 types of signals (good, poor, and bad). If the classification results show a good signal, the issued heart rate value is the same as the POS result. If the classification results show a bad signal, the system does not issue a heart rate value. If the classification results show a poor signal, the results from the POS are entered first into the compensating network. The compensation network will estimate the heart rate value generated by the POS.

So, it is expected that the measurement value shows accurate results. This paper comprises four section, where section 1 is the introduction part. Section 2 describes method about noise classification and filters design according to the proposed concept. The test results and analysis will be explained in section 3. Section 4 in this paper will describe the conclusions about this research.

2. METHOD

The proposed method starts with the face detection process to get a face image in RGB color. This step is carried out using a Haar cascade. After the face is detected, skin detection is carried out to separate the facial areas that are considered skin and non-skin. This step also eliminates the possibility of detecting other images other than faces. Estimating the heart rate value is carried out using the POS algorithm. The proposed method is shown in Figure 1.
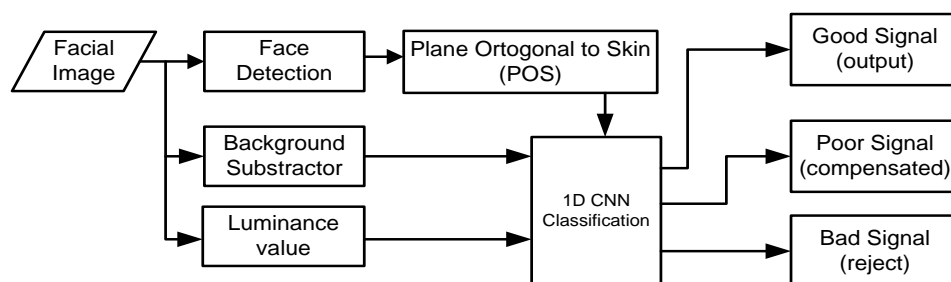

Figure 1. Proposed method

2.1. Face detection and ROI selection

Remote photoplethysmograph algorithm requires facial images to measure the subject's heart rate. In this paper, face detection is performed using the Haar cascade classifier. This Haar cascade classifier has a fairly fast detection speed than several other methods [27], [28]. Because of the speed of execution has a significant effect on the overall measurement results. Face detection produces landmark information that represents the position of the face in an image. The facial image parameter is shown in Figure 2.

Figure 2. Facial image parameters

The landmark information obtained is the anchor location of the landmark, the width of the face, and the height of the face. The face anchor locations in the image are represented in $face_x$ and $face_y$ notation, where x and y are the face anchor locations in the x and y axes of the image, respectively. Face width is denoted using $face_w$, and face height use $face_h$, which is presented in pixels.

$$roi_x = face_x + (0.2 face_w) \tag{1}$$

$$roi_y = face_y + (0.5\ face_h) \tag{2}$$

$$roi_w = roi_x + (0.6\ face_w) \tag{3}$$

$$roi_h = roi_y + (0.25\ face_h) \tag{4}$$

In measuring process, the rPPG performed in this paper does not use the entire face area. The area measured is focused on the area under the eyes and above the mouth. So, the face image obtained requires an ROI selection process. The desired ROI is obtained using (1)-(4), where ROI_{xy} is the anchor ROI, ROI_w is the width of the ROI, and ROI_h is the height of the ROI. To get the ROI value can be used (1)-(4). From the ROI determined, the ROI image extraction process is carried out based on the three-color channels R, G, and B, where the values of R(n), G(n), and B(n) are the average values of each intensity. RGB channel on the nth frame. This process is illustrated in Figure 3.
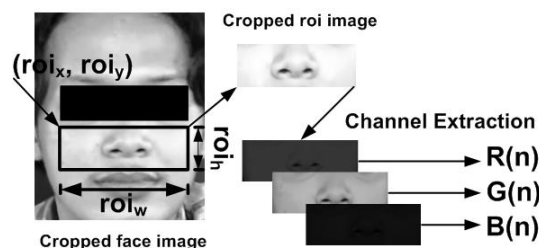

Figure 3. RGB extraction from ROI image

2.2. Plane orthogonal to skin

Plane orthogonal to skin (POS) is a method of extracting rPPG signals from changes in skin color for a certain duration. Changes in skin color occur due to changes in blood volume in facial skin tissues. POS is

carried out in several stages, including spatial averaging, temporal normalization, projection, tuning, and overlap-adding.

Spatial averaging is a process to find the average pixel value in each red, green, and blue color channel in each n frame. In (5) is denoted by C(n). The number of frames needed in this study is 256, requiring 8 seconds of image capture duration. The camera used can take images with a frame rate of 30 FPS.

$$C(n) = [R(n), G(n), B(n)]^T \tag{5}$$

$$C_n = \frac{C_{(n)}}{u(C(n))} \tag{6}$$

Temporal normalization is shown by (6). This process is done by dividing the value of C(n) by the average value of C(n). After doing temporal normalization, the next step is projection. This process aims to obtain the Spn projection signal from the Cn signal multiplied by the POS projection matrix (7). Two projection signals are obtained, namely, S_1, which projects a combination of positive green and negative blue values, and S_2 signals, which project a combination of red, green, and blue, with positive double negative values.

$$S_p(n) = \begin{bmatrix} 0 & 1 & -1 \\ -2 & 1 & 1 \end{bmatrix} C_p(n) \tag{7}$$

$$h(n) = S_1(n) + \alpha . S_2(n) \; with \; \alpha = \frac{\sigma(S_1)}{\sigma(S_2)} \tag{8}$$

$$H_{n=0 \to l} = H_{n=0 \to l} + (h(n) - \mu(h_{n=0 \to l})) \tag{9}$$

The tuning process can be done using (8), where σ is the standard deviation. The value of hn results from tuning as many as n data. The value of n is the number of sliding window data set at 1.6 times the camera frame rate. After the h(n) value is obtained, overlap adding process is carried out based on (9). This process aims to obtain the result of the current hn with the previous frames along the sliding window used.

2.3. Frequency selection

Image quality greatly affects the quality of the rPPG signal reading results. The noise that appears can cause the quality of the rPPG signal to decrease. So the H signal obtained must be filtered using a band pass filter. The cut-off frequency used is adjusted to a normal human heart rate. Under normal conditions, the human heart rate ranges from 45 to 120 beats per second or the equivalent to 0.75 Hz to 2 Hz.

2.4. Feature extraction

The subject's movement during the measurement process is one of the most common disturbances. A large movement will impact the quality of the rPPG signal. One way to detect movement is to use a background subtractor. The background subtractor is used to count the number of pixels that change in value compared to the previous frame. So the number of pixels that change represents the amount of subject movement. The greater the subject's movement will be directly proportional to the number of pixels that change. The background subtractor is indicated by (10).

$$Bs(n) = \sum_{x=0, y=0}^{x=Iw, y=Ih} I_{gray}(x, y) \; (n) \tag{10}$$

Where Bs(n) is the number of pixels that change in the nth frame, I_w and I_h are the width and height of the image, respectively, I_{gray} is the grayscale image, and x y is the coordinates of the pixels in the image. Changes in light illumination during the measurement process are unavoidable when measuring in real conditions. These changes cause the input signal to have a change in intensity that is too large. So that the temporal normalization process cannot show optimal results, changes in light illumination that are read by the camera will directly impact changes in the RGB color components. According to Wu *et al.* [18], the illumination value will be directly proportional to the Y value in the YCrCb color space.

$$Y = 0.299R + 0.587G + 0.114B \tag{11}$$

2.5. Signal quality classification network

The appearance of noise can result in an error in reading the rPPG signal. This error can be minimized by applying a filter according to the type of noise. So that the noise that appears must be identified so that the appropriate filter can be used. In this research, the classification of noise types due to movement and changes

in light illumination will be carried out. The classification process is carried out using a 1D CNN. This classification process is carried out using the results of the extraction of the background subtractor feature and the light illumination change signal (the Y component in YCrCb). The 1D CNN network architecture is shown in Figure 4. This network has three output that identified quality of the signal. First output is good that presented good quality signal. This result directly used as measurement output. The second output is bad, that presented bad quality signal. This output directly rejected and measuremet output shown bad. The last output is poor, that represented poor quality signal. This signal has better quality than bad signal so can be fixed by compensated network.

2.6. Compensated network

After the signal quality classification results are obtained, 3 results will be obtained. If the signal quality is good, then the POS algorithm results are used as measurement results. If the signal quality is poor, then the POS algorithm results are not issued. However, if the classification results indicate a moderate signal, the results of the POS algorithm are entered into the compensated network to estimate the correct value. The results of this network will change the value of the POS algorithm to a larger or smaller value according to the results of the compensated network. The architecture of the compensated network is shown in Figure 5.
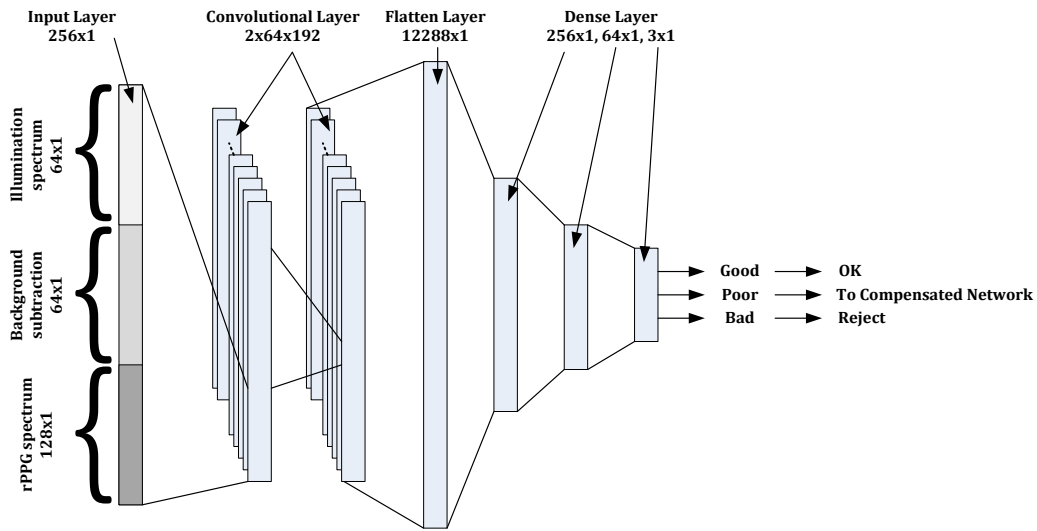

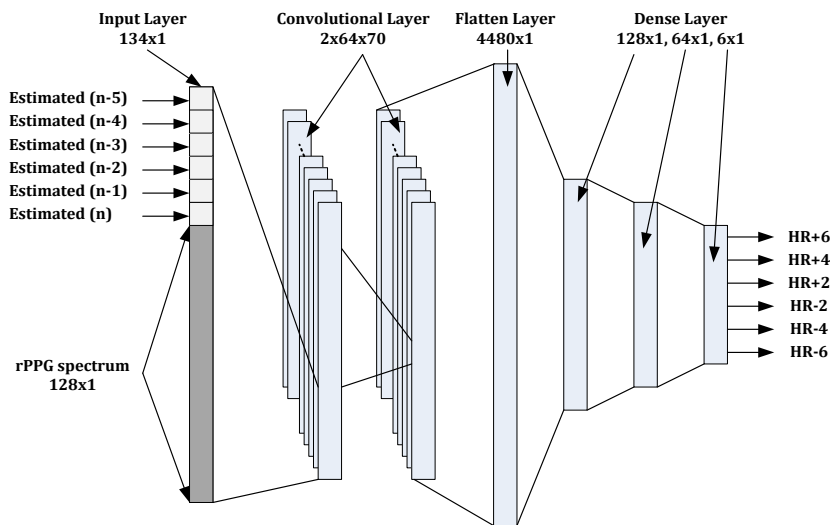Figure 4. Proposed signal quality classification network

Figure 5. Proposed compensated network

2.7. Training and validation dataset

Datasets for training and validation are created using the Logitech C920 camera with a resolution of 640×480. The frame rate obtained ranges from 15 to 18 frame per-second (FPS). Pulse oximeter CMS50E is used as comparison data. The number of videos captured is 80 with 4 different illumination variations. Each video is 512 frames long with the duration adjusting its frame rate. The results of the reading of the fingertip pulse oximeter are used as label data in the training and validation process. The dataset was taken from 20 subjects aged 19-21 years. There are 18 male subjects and 2 female subjects with brown skin color.

2.8. Testing dataset

The dataset for testing is made with the same specifications as the dataset for training and validation. However, the number of subjects in the training dataset is less, namely 30 videos from 10 different subjects. So that each subject is taken a video with 3 variations of light illumination.

2.9. Evaluation metrics

The performance of the proposed method is compared with other methods used as a reference. In this paper, we will use four metrics. To represent the divergence between the prediction and the reference result, the mean absolute error (MAE) (12) metric will be used, and the degree of deviation is represented by the mean error rate (MER) (13).

$$MAE = \frac{1}{K}\sum_{i=1}^{K}|y_i - \dot{y}_i| \tag{12}$$

$$MER = \frac{1}{K}\sum_{i=1}^{K}\frac{|y_i - \dot{y}_i|}{y_i} \; x \; 100\% \tag{13}$$

$$tpf = t_f(n) - t_f(n-1) \tag{14}$$

The value of y is the predicted heart rate value, and the value of y is the reference value/ground truth of the pulse oximeter. K is the amount of data analyzed.

3. RESULTS AND DISCUSSION

In this section, the results of the experiments that have been carried out are explained. The experiment started with a training dataset for noise classification, and a compensated network. After obtaining the appropriate model, it is continued with experiments using data testing.

3.1. Signal quality classification network training result

The dataset obtained is trained using a 1D CNN network with the architecture shown in Figure 4. The training process was carried out 5 times with various filter values. This test aims to find the model with the highest validation accuracy. Where the model will be carried out for the validation and testing process. Training result of each filter shown in Table 1.

Table 1. SQC network training result with filter variance

No	filter	epoch	loss	val loss	accuracy (%)	val accuracy (%)
1	8	14648	0.0912	0.1171	76.01	67.44
2	16	16971	0.0335	0.0685	94.70	84.65
3	32	18034	0.0128	0.0613	99.37	88.83
4	64	15049	0.0025	0.0543	99.98	89.31
5	128	7538	0.1689	0.1729	43.30	45.11

Data was trained using 1D CNN architecture in Figure 4. A model indicates the greatest validation accuracy value with a filter of 64. The accuracy obtained was 89.31%. The validation loss obtained is also the lowest, which is 0.0543. So that the model will be used in heart rate measurement testing.

3.2. HR measurement result

The compensated network training process is carried out with the same procedure as the SQC Network. This network using 1D CNN Architecture in Figure 5. Models with a filter size of 64 provide the best validation accuracy. In the HR measurement test used, this model. Compensated network training performance is shown in Table 2.

3.3. HR measurement result

The test results of heart rate measurements are shown in Table 3. Tests were carried out on 10 subjects, each taken 4 times. Each measurement produces 20 measurement data. The measurement results of each subject were calculated as the mean absolute error (MAE), relative mean error (MER), and average time per frame (tpf). The POS algorithm is compared with the POS algorithm with a signal quality classification network added and a compensated network using 1D CNN (POS+1DCNN).

Based on these data, the application of the signal quality classification network and compensated network reduced the mean absolute error from 3.89 to 2.78. The relative mean error also decreased from 5.14% to 3.67%, so the deviation was smaller. However, the time per frame required for the signal quality classification network and the compensated network is about 10 ms longer. This causes the frame rate to drop to 16-17 FPS compared to the POS algorithm, which reaches 19-20 FPS.

Table 2. Compensated network training result with filter variance

No	filter	epoch	loss	val loss	accuracy (%)	val accuracy (%)
1	8	21921	0.0806	0.1151	80.68	69.30
2	16	19457	0.0323	0.0865	96.88	77.67
3	32	17848	0.0427	0.1007	94.39	73.48
4	64	22503	0.0060	0.0676	99.98	84.18
5	128	30364	0.0078	0.0758	96.99	83.58

Table 3. rPPG POS and POS+1DCNN (64 filter) result

No Subject	POS			POS + 1DCNN		
	MAE	MER (%)	tpf (ms)	MAE	MER (%)	tpf (ms)
Subject 1	4.94	6.00	51.00	2.78	3.67	63.00
Subject 2	4.05	6.00	51.00	2.97	4.00	63.00
Subject 3	4.35	5.33	50.00	1.19	4.67	59.00
Subject 4	3.85	5.33	49.00	2.86	4.00	65.00
Subject 5	4.32	6.00	51.00	2.39	3.67	65.00
Subject 6	2.34	5.67	52.00	2.00	1.67	61.00
Subject 7	1.96	5.00	53.00	1.19	1.33	58.00
Subject 8	4.97	4.67	50.00	1.29	2.33	63.00
Subject 9	3.82	3.67	52.00	0.23	1.67	58.00
Subject 10	4.74	3.67	52.00	2.90	2.67	58.00
Mean	3.89	5.14	**51.10**	2.78	3.67	61.30

Figure 6 shows a change in the HR value of subject 1. It can be seen that the HR value of the measurement results tends to be lagging against the reference HR. As seen at 10.95 seconds, the reference HR rose rapidly from 77 to 79, but the measurement HR was still downtrend and responded at 13.87 seconds, or 2.92 seconds too late. This could be due to the POS algorithm taking the average value of the last few frames (according to the sliding window). So this weakness is difficult to overcome when using the POS algorithm. However, the error that occurs is quite small and within tolerance limits.
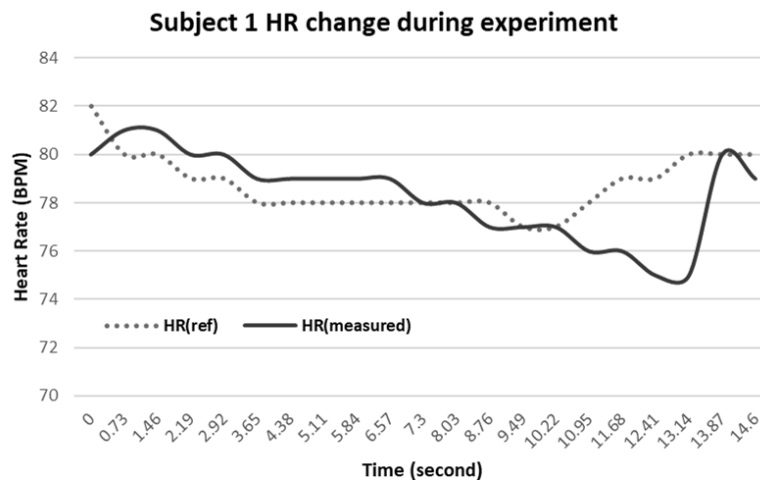

Figure 6. Subject 1 HR change during experiment

The last data presented in this paper is the distribution of measurement error values using the bland-Altman plot. A total of 450 measurement results from all subjects showed the error value had met the 95% limits of agreement. The measurement results show that there are only 22 data that are outside the 95% area. From these data in Figure 7, the system can be quite accurate.

Figure 7. Bland-Altman plot HR measurement result
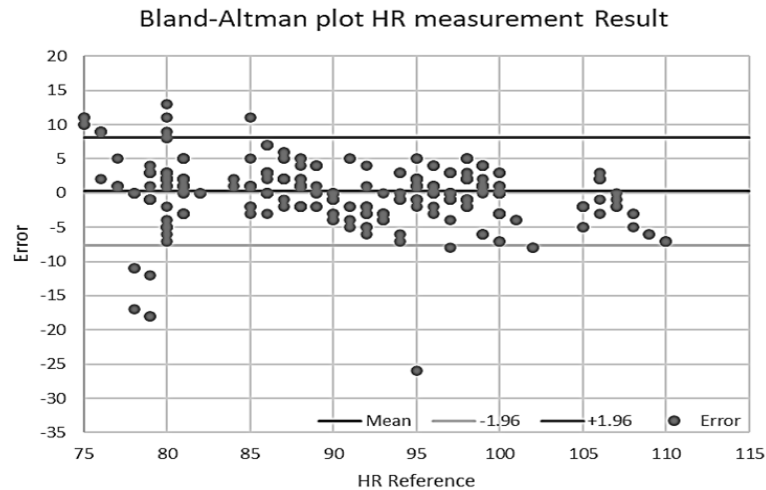
4.   CONCLUSION

Based on testing on 10 subjects, it shows that the 1D CNN network that was created succeeded in classifying signal quality well. The estimation process with the compensated network also reduced the error value. The mean absolute error decreased from 3.89 to 2.78, and the mean relative error value also decreased from 5.14% to 3.67%. The system has also met the 95% limits of agreement so that the measurement results are quite accurate.

ACKNOWLEDGEMENTS

REFERENCES

[1]   P. Jakkaew and T. Onoye, "Non-contact respiration monitoring and body movements detection for sleep using thermal imaging," *Sensors*, vol. 20, no. 21, pp. 1–14, Nov. 2020, doi: 10.3390/s20216307.
[2]   R. A. Firmansyah, Y. A. Prabowo, and T. Suheta, "Thermal imaging-based body temperature and respiratory frequency measurement system for security robot," *Przeglad Elektrotechniczny*, vol. 98, no. 6, pp. 126–130, 2022, doi: 10.15199/48.2022.06.23.
[3]   H. Elphick, A. Alkali, R. Kingshott, D. Burke, and R. Saatchi, "Thermal imaging method for measurement of respiratory rate," in *7.1 Paediatric Respiratory Physiology and Sleep*, Sep. 2015, p. PA1260, doi: 10.1183/13993003.congress-2015.pa1260.
[4]   R. A. Firmansyah, Y. A. Prabowo, and T. Suheta, "Design of a non-contact heart rate meter using thermal imaging for a security robot (in Indonesian)," *Jurnal Media Informatika Budidarma*, vol. 6, no. 1, p. 320, Jan. 2022, doi: 10.30865/mib.v6i1.3416.
[5]   C. B. Pereira, M. Czaplik, V. Blazek, S. Leonhardt, and D. Teichmann, "Monitoring of cardiorespiratory signals using thermal imaging: A pilot study on healthy human subjects," *Sensors (Switzerland)*, vol. 18, no. 5, p. 1541, May 2018, doi: 10.3390/s18051541.
[6]   G. De Haan and V. Jeanne, "Robust pulse rate from chrominance-based rPPG," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2878–2886, Oct. 2013, doi: 10.1109/TBME.2013.2266196.
[7]   W. Wang, A. C. Den Brinker, S. Stuijk, and G. De Haan, "Algorithmic principles of remote PPG," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1479–1491, Jul. 2017, doi: 10.1109/TBME.2016.2609282.
[8]   G. De Haan and A. Van Leest, "Improved motion robustness of remote-PPG by using the blood volume pulse signature," *Physiological Measurement*, vol. 35, no. 9, pp. 1913–1926, Sep. 2014, doi: 10.1088/0967-3334/35/9/1913.
[9]   Q. V. Tran, S. F. Su, W. Sun, and M. Q. Tran, "Adaptive pulsatile plane for robust noncontact heart rate monitoring," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 9, pp. 5587–5599, Sep. 2021, doi: 10.1109/TSMC.2019.2957159.

[10] P.-W. Huang, C.-H. Lin, M.-L. Chung, T.-M. Lin, and B.-F. Wu, "Image based contactless blood pressure assessment using Pulse Transit Time," in *2017 International Automatic Control Conference (CACS)*, Nov. 2017, vol. 2017-Novem, pp. 1–6, doi: 10.1109/CACS.2017.8284275.

[11] Y. Nakayama, G. Sun, S. Abe, and T. Matsui, "Non-contact measurement of respiratory and heart rates using a CMOS camera-equipped infrared camera for prompt infection screening at airport quarantine stations," in *2015 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, CIVEMSA 2015*, Jun. 2015, pp. 1–4, doi: 10.1109/CIVEMSA.2015.7158595.

[12] T. Negishi *et al.*, "Contactless vital signs measurement system using RGB-thermal image sensors and its clinical screening test on patients with seasonal influenza," *Sensors (Switzerland)*, vol. 20, no. 8, p. 2171, Apr. 2020, doi: 10.3390/s20082171.

[13] X. Liu, X. Yang, D. Wang, and A. Wong, "Detecting pulse rates from facial videos recorded in unstable lighting conditions: an adaptive spatiotemporal homomorphic filtering algorithm," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–15, 2021, doi: 10.1109/TIM.2020.3021222.

[14] R. Song, J. Li, M. Wang, J. Cheng, C. Li, and X. Chen, "Remote photoplethysmography with an EEMD-MCCA method robust against spatially uneven illuminations," *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13484–13494, 2021, doi: 10.1109/JSEN.2021.3067770.

[15] H. Gao, X. Wu, J. Geng, and Y. Lv, "Remote heart rate estimation by signal quality attention network," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 2022, vol. 2022-June, pp. 2121–2128, doi: 10.1109/CVPRW56347.2022.00230.

[16] D. Cho, J. Kim, K. J. Lee, and S. Kim, "Reduction of motion artifacts from remote photoplethysmography using adaptive noise cancellation and modified HSI model," *IEEE Access*, vol. 9, pp. 122655–122667, 2021, doi: 10.1109/ACCESS.2021.3106046.

[17] S. Kado, Y. Monno, K. Yoshizaki, M. Tanaka, and M. Okutomi, "Spatial-spectraloral fusion for remote heart rate estimation," *IEEE Sensors Journal*, vol. 20, no. 19, pp. 11688–11697, Oct. 2020, doi: 10.1109/JSEN.2020.2997785.

[18] B.-F. Wu, Y.-C. Wu, and Y.-W. Chou, "A compensation network with error mapping for robust remote photoplethysmography in noise-heavy conditions," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–11, 2022, doi: 10.1109/TIM.2022.3141149.

[19] R. Song, S. Zhang, C. Li, Y. Zhang, J. Cheng, and X. Chen, "Heart rate estimation from facial videos using a spatiotemporal representation with convolutional neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7411–7421, Oct. 2020, doi: 10.1109/TIM.2020.2984168.

[20] H. Hwang and E. C. Lee, "Non-contact respiration measurement method based on rgb camera using 1d convolutional neural networks," *Sensors,* vol. 21, no. 10, p. 3456, May 2021, doi: 10.3390/s21103456.

[21] B. F. Wu, P. W. Huang, D. H. He, C. H. Lin, and K. H. Chen, "Remote photoplethysmography enhancement with machine leaning methods," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, Oct. 2019, vol. 2019-October, pp. 2466–2471, doi: 10.1109/SMC.2019.8914554.

[22] N. A. Akbar, A. Muneer, S. M. Taib, and S. Mohamed Fati, "Measuring accuracy towards facial video heart-rate estimation using haar-cascade and CNN method," in *2022 International Conference on Decision Aid Sciences and Applications, DASA 2022*, Mar. 2022, pp. 1607–1611, doi: 10.1109/DASA54658.2022.9764992.

[23] M. Artemyev, M. Churikova, M. Grinenko, and O. Perepelkina, "Robust algorithm for remote photoplethysmography in realistic conditions," *Digital Signal Processing: A Review Journal*, vol. 104, p. 102737, Sep. 2020, doi: 10.1016/j.dsp.2020.102737.

[24] O. Perepelkina, M. Artemyev, M. Churikova, and M. Grinenko, "HeartTrack: Convolutional neural network for remote video-based heart rate monitoring," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 2020, vol. 2020-June, pp. 1163–1171, doi: 10.1109/CVPRW50498.2020.00152.

[25] M. Kopeliovich, Y. Mironenko, and M. Petrushan, "Architectural tricks for deep learning in remote photoplethysmography," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, Oct. 2019, pp. 1688–1696, doi: 10.1109/ICCVW.2019.00209.

[26] K. Kido *et al.*, "A novel CNN-based framework for classification of signal quality and sleep position from a capacitive ECG measurement," *Sensors*, vol. 19, no. 7, p. 1731, Apr. 2019, doi: 10.3390/s19071731.

[27] N. Singh and R. M. Brisilla, "Comparison analysis of different face detecting techniques," in 2021 *Innovations in Power and Advanced Computing Technologies (i-PACT)*, Nov. 2021, pp. 1–6, doi: 10.1109/i-PACT52855.2021.9696583.

[28] K. Dang and S. Sharma, "Review and comparison of face detection algorithms," in *2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence*, Jan. 2017, pp. 629–633, doi: 10.1109/CONFLUENCE.2017.7943228.

BIOGRAPHIES OF AUTHORS

Riza Agung Firmansyah ⓘ 🅶 🆂🅲 ◗ is Lecturer at Department of Electrical Engineering, Institut Teknologi Adhi Tama Surabaya, Indonesia. He holds a magister's degree in Electronics Engineering with a specialization in robotics and computer vision. His research areas are robotics, artificial intelligence, computer vision, and robot vision. He is the head of the robotics Lab at Institut Teknologi Adhi Tama Surabaya. He has filed a number of patents and intellectual property rights on his innovative ideas and has been awarded with one national patent and five intellectual property rights on computer programs. He can be contacted at email: rizaagungf@itats.ac.id.

Yuliyanto Agung Prabowo ⒾⒼ ㏒ ↻ he is Lecturer at Department of Electrical Engineering, Institut Teknologi Adhi Tama Surabaya, Indonesia. He received a Bachelor Engineering and Master Engineering degree in Electrical Engineering with a specialization in Control System Engineering from the Sepuluh Nopember Institut of Technology, Indonesia. His research areas are modeling, nonlinear systems, intelligent control, automation, and robotics. He is head of Control System Laboratory at Institut Teknologi Adhi Tama Surabaya. He can be contacted at email: agungp@itats.ac.id.

Titiek Suheta Ⓘ Ⓖ ㏒ ↻ she is Lecturer at Department of Electrical Engineering, Institut Teknologi Adhi Tama Surabaya, Indonesia. He received his Bachelor's Engineering and Master's Engineering degrees at Institut Teknologi Adhi Tama Surabaya, Indonesia. Her research areas are electrical distribution, energy management, and electrical optimation. She has several international and national journals about her research. She can be contacted at email: hita@itats.ac.id.

Syahri Muharom Ⓘ Ⓖ ㏒ ↻ he is Lecturer at Department of Electrical Engineering, Institut Teknologi Adhi Tama Surabaya, Indonesia. He received the Bachelor Engineering and Master Engineering degree in Electrical Engineering at Institut Teknologi Sepuluh Nopember Surabaya with a specialization in Robot Vision, Image Processing, and Microelectronics Systems. He is the head of Microcontroller and Computer Laboratory at Institut Teknologi Adhi Tama Surabaya. He can be contacted at email: syahrimuharom@itats.ac.id.