An SoC-Based System for Real-time Contactless Measurement of Human Vital Signs and Soft Biometrics

Aminuddin Rizal*
Computer Engineering Department
Universitas Multimedia Nusantara
South Tangerang, Indonesia
*Email: Aminuddin.rizal@umn.ac.id

Kuan-Ting Chiang, Jia-Wei Lin, and Yuan-Hsiang Lin†
Electronic and Computer Engineering Department
National Taiwan University of Science and Technology
Taipei, Taiwan
†Email: linyh@mail.ntust.edu.tw

***Abstract*—Computer vision (CV) plays big role in our current society's life style. The advancement of CV technology brings the capability to sense human vital sign and soft biometric parameters in contactless way. In this work, we design and implement the contactless human vital sign parameters measurement including pulse rate (PR) and respiration rate (RR) and also for assessment of human soft biometric parameters i.e. age, gender, skin color type, and body height. Our designed system is based on system on chip (SoC) device which run both FPGA and hard processor while provides real-time operation and small form factor. Experimental results shows our device performance has mean absolute error (MAE) 2.85 and 1.46 bpm for PR and RR respectively compared to clinical apparatus. While, for soft biometric parameters measurement we got unsatisfied results on age and gender estimation with accuracy of 58% and 74% respectively. However, for skin color type and body height measurement we reach high accuracy with 98 % and 2.28 cm respectively on both parameters.**

Keywords*—*contactless measurement, remote imaging photopletysmography, soft biometric, system on chip, vital signs

I. INTRODUCTION

Nowadays, with the advancement of science and technology, a system are demanded to intelligently deliver solution for some problems. Especially for critical application such as biomedical and security, an accurate and credible device is a must. Over decades, the human vital sign and soft biometric parameters measurement are limited to contact based method which may not comfort and less flexible for long-term measurement [1].

As breakthrough, development of computer vision (CV) enables the human physical and physiological parameters measurement. This method offer non-contact measurement with high accuracy to compete the conventional method. Furthermore, in biomedical application, by using non-contact method offer high flexibility and continuous measurement assist the user decrease the risk of diseases [2].

On the other hand in processing stage, studies from [3] and [4] exposed System on Chip (SoC) is a platform which most suitable to handle the continuous of human vital signs and soft biometrics monitoring. SoC offers several benefits including stable operation performance, instant computation, and reconfigurable to adapt the application needs [5]. Low power and robust system are also becomes the advantages from using SoC for contactless vital signs and soft biometrics measurement.

In this paper, we propose a contactless system based on SoC platform and a consumer grade RGB camera as the sensor to measure human vital signs and soft biometric parameters. Our main goal is to make a comfortable and credible framework for long-term monitoring. Since our framework is a general purpose, however we have some target applications for our proposed framework in the future including medical instrumentation, robotic, smart assisting technology, exercise, security and automotive.

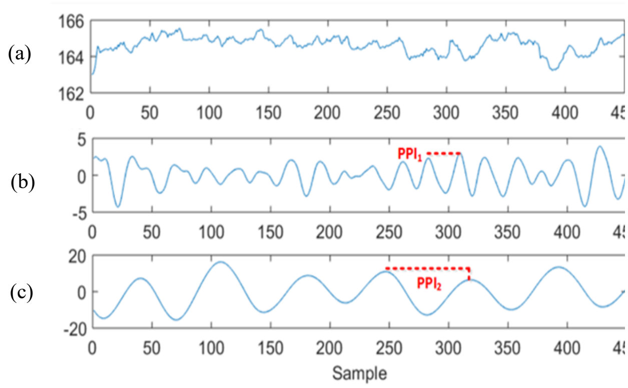


Figure 1 PR and RR signal

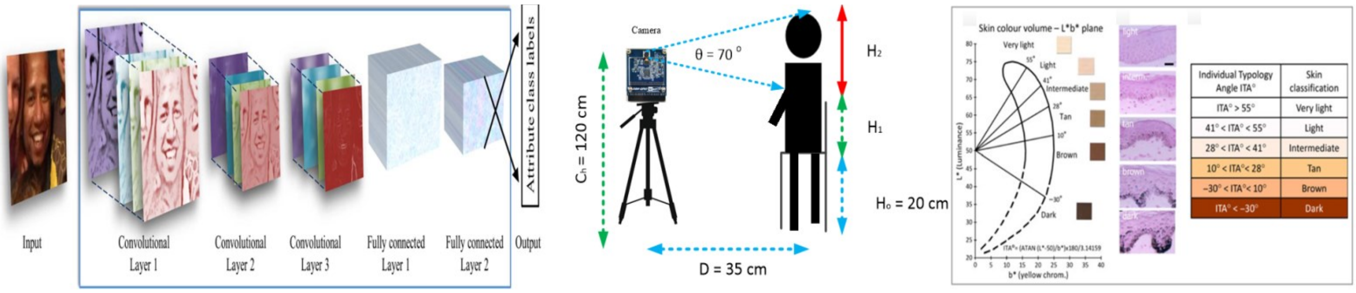
II. CONTACTLESS MEASUREMENT

A. Vital Sign Measurement

Vital sign measurement are usually by using photopletysmography (PPG) which refer to an electro-optical technique for invasively measuring the tissue blood volume pulses in the microvascular tissue bed underneath skin [6]. PPG by utilize camera as sensor known as remote imaging photopletysmography (rIPPG/rPPG). This method has more advantages as non-contact based method compared to radar such as low cost, small form factor, and less complexity. Recent years, different algorithm to rPPG has been proposed [7] and offers more vital sign parameters calculation like pulse rate (PR) , respiration rate (RR), blood pressure [8], heart rate variability (HRV) [9] and blood oxygen saturation (SpO_2) [10]. Basically there is two stages for rPPG to measure vital sign parameters, those stages are spatial and temporal processing. Spatial processing is a stage to convert pixel data into 1D signal, the example of raw signal of 450 frames (15 second) illustrated in Figure 1 (a). The raw signal already includes PR and RR signal. Therefore, two separate those two parameters we need temporal processing stage. In this system we only calculate pulse rate (PR) and respiration rate(RR).

B. Pulse Rate (PR)

Basic idea of rPPG is similar with conventional PPG device to detect the pulse, rPPG remotely sensing the subtle change in skin color to extract PPG signal. After temporal processing the raw signal that including filtering and windowing, clean PPG signal can be obtained and illustrated in Figure 1 (b). Pulse rate vale can be determined by calculating the time interval between two peaks (P). Equation (1) and (2) are used to calculate PR value:



(a) (b) (c)

Figure 2 (a) age and gender with CNN proposed by [14], (b) body height measurement, (c) skin color type based on ITA [16]

$$PPI_n = \frac{P_n - P_{n-1}}{fs} \tag{1}$$

$$PR_n = \frac{60}{PPI_{1_n}} \tag{2}$$

where PPI_n stands for a peak-to-peak interval in second, *fs* is the camera frame rate (fps), PR is the obtained pulse rate in beat per minute (BPM) unit.

C. *Respiration Rate (RR)*

Apart from conventional respiration rate measurement, the concept of measurement from rPPG for RR measurement is to detect small motion on face. In the analysis stage, an image amplification method like Eulerian Video Magnification (EVM) [11] with range amplification around the respiration rate frequency stand (0.3 – 1.5 Hz) can be used to determine whether camera can capture the respiration motion or not. The example of clean RR signal after temporal processing presented in Figure 1 (c). RR calculation is similar with PR parameter, we can refer equation (2) to determine signal peak-to-peak and equation (3) to calculate final RR.

$$RR_n = \frac{60}{PPI_{2_n}} \tag{3}$$

D. *Soft Biometric Measurement*

Soft biometric measurement intended to collect human physical information. This measurement is useful for some applications including security, person identification, and immigration [12]. The detail of soft biometric traits is well written in this paper [13]. Based on the paper soft biometric can be divided by two, facial and body soft biometric. In this paper we will cover three facial soft biometric parameters (age, gender and skin color type) and one body soft biometric parameter (body height).

E. *Age and Gender*

More research is conducted to improve accuracy of age and gender prediction. Key principle of age and gender prediction is to extract facial feature. One popular technique is convolutional neural networks (CNN). In this system we use network architecture proposed by this work [14] and illustrated in Figure 2 (a). This network constructed by three convolutional layers and two-fully connected layers with a small number of neurons. For age, we only use 4 classes "Teen", "Adult", "Mature" and "old".

F. *Body Height*

Our body height estimation is based on upper body measurement, since we have only short distance. To have

$$\frac{H_{imsge}}{F} = \tan\theta = \frac{H_{Real}}{D} \tag{4}$$

$$F = \frac{H_{image}}{\tan\theta} = \frac{640}{\tan 70} = 238 \tag{5}$$

$$\frac{H_{image}}{F} = \frac{H_2}{D} \tag{6}$$

$$H_2 = \frac{(640 - \mathrm{ROI.y}) \times 35}{238} \tag{7}$$

$$\mathrm{BH} = H_0 + H_1 + H_2 \tag{8}$$

best result we use static measurement setup like shown in Figure 2 (b), we were setting camera height (C_h) is 120 cm and distance (D) to the subject is 30 cm. Illustrated in the same figure, body height (BH) are calculated from the sum of H_0, H_1, and H_2. Where H_0 is the height of chair, H_1 is linear regression constant to estimate middle body length with function of C_h and pixel-distance value, lastly H_3 is calculated by measurement using camera. So our main focus is to calculate H_2. We use simple optical physics principle to deal with this phenomenon and presented in Equation 4. Where F is the focus length of camera, and θ is the angle of view (AOV) of camera which is 70° known from camera datasheet [15]. Assume we have full image in 640 resolutions, so we can calculate the F value and results in Equation 5. Since we had obtained F value and set D in certain value, our next step to estimate H_{real}/H_2 by measuring captured H_{image} which represent y_0 position from face detection of region-of-interest (ROI). Equation 6 and 7 are used to measure H_2 with known variable. H_1 is known from a linear regression of camera height and pixel-distance value. Finally body height (BH) can be estimated using Equation 8.

G. *Skin Color Type*

$$\begin{bmatrix} X \\ Y \\ X \end{bmatrix} = \begin{bmatrix} 0.43 & 0.34 & 0.17 \\ 0.22 & 0.7 & 0.07 \\ 0.02 & 0.13 & 0.93 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{9}$$

$$L^* = \begin{cases} 116 \times \left(\frac{Y}{Y_n}\right)^{0.3}, & \frac{Y}{Y_n} > 0.008856 \\ 116 \times \left(\frac{Y}{Y_n}\right), & \frac{Y}{Y_n} \le 0.008856 \end{cases} \tag{10}$$

$$f(t) = \begin{cases} t^{0.3}, & t > 0.008856 \\ 7.78 \times t + \frac{16}{116}, & t \le 0.008856 \end{cases} \tag{11}$$

$$b^* = 200 \times \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right] \tag{12}$$

$$\mathrm{ITA} = \left[\arctan\left(\frac{(L^* - 50)}{b^*}\right) \right] \times \frac{180}{phi} - constant \tag{13}$$

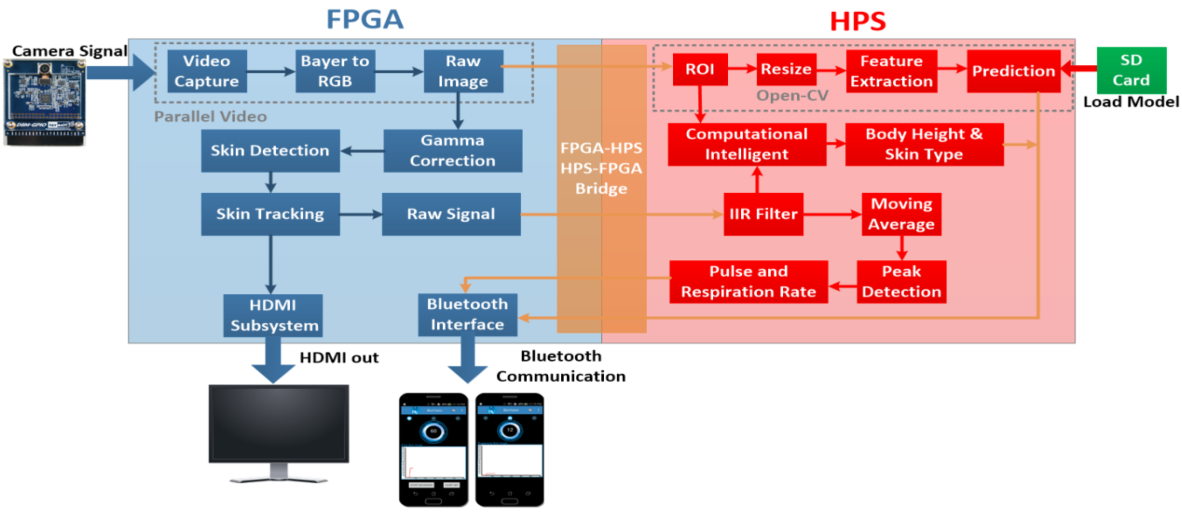


Figure 3 Main system block diagram

Basically, we can use several color space to obtain skin color type classification. In this work, we use individual typology angle (ITA) algorithm to deal with this problem. ITA proved more accurate compared to others algorithm like YCbCr and RGB [16]. Since ITA works in CIE Lab color space, a color conversion is needed. However ITA only need L and b component of CIE Lab color channel. We use Equation (9) – (13) to calculate ITA from an RGB camera. Constant in Equation (13) is illumination condition, since original CIE Lab equation is based on D65 illumination.

Illustrated in Figure 2 (c), the correlation between ITA value with skin type classification. The ITA based method allowed to distinguish skin type into 6 classes: very light > 55° > light > 41° > intermediate > 28° > tan > 100° > brown > -30° > dark.

III. Methods

In this work, we use a system on chip (SoC) processor as single and main processor to work. We use Cyclone V SoC which armed with FPGA and Hard Processor System (HPS) / ARM architecture. HPS consisting of processor, peripherals and memory interfaces tied seamlessly with the FPGA fabric using a high-bandwidth interconnect backbone Figure 3 presents the block diagram of our proposed framework, which can be divided into four major stages: (1) image processing on FPGA, (2) signal processing on HPS, and lastly (3) parameters calculation for both vital sign and soft biometric. In our system, for input device we use a consumer grade RGB CMOS camera to capture image continuously at 30 fps. While for output devices consist of HDMI display and Bluetooth module for smartphone communication. We synthesized FPGA part of SoC to deal with input/output (I/O) interfaces and pre-image processing from raw image data. And also, we utilize the HPS side to handle signal processing, parameters computation and prediction.

A. Image Processing on FPGA

Usually image acquisition performance in general CPU achieved unstable fps, therefore inaccurate measurement especially in vital signs parameters occur. The acceleration performance brought by FPGA is ideal for our proposed framework to improve accuracy and enable long-term measurement.

At the beginning, FPGA makes communication with CMOS camera to request a frame. And then we convert the data from camera which is a Bayer pattern to RGB data and formed raw image. Since the rPPG signal extracted from the skin, we do skin detection and tracking. We select region of interest (ROI) were cheeks area. And lastly, we use green channel data of the ROI to generate rPPG signal with averaging pixels inside ROI area by using Equation 14. Where ROI_{area} is the pixel area of ROI, $Intensity_n(x_i,y_j)$ is the intensity of green channel located in (x_i,y_j), and n is the current frame number. Green channel has better blood absorption compared red and blue channel [3].

$$Signal_{raw}[n] = \frac{1}{ROI_{area}} \sum_{i=1}^{20} \sum_{j=1}^{20} Intensity_n(x_i, y_j) \tag{14}$$

Simultaneously, we send an RGB frame from captured image to HPS side by using FPGA-to-HPS (F2H) bridge at speed 25 MHz.

B. Signal Processing on HPS

The raw signal itself was already contain information related with PR and RR. However, we need separate the two parameters from the raw signal. We built two 8 order digital bandpass filter with cutoff frequency is 0.8 to 3.4 Hz for PR (in the scope of normal pulse rate 48 to 204 bpm) and 0.05 to 0.5 Hz for RR (according to normal respiration rate 6 to 30). Equation 15 is mathematic operation for filtering stage, where P and Q refer to the feedforward and feedback filter order respectively which are set as 8 orders in this work. The b_i and a_j are the feedforward and feedback filter coefficients. x[n] is the input signal and y[n] is corresponding filter output across each frame. Afterwards, we conduct 10 samples moving average to avoid false positive peak detection by using Equation 16 with SMA is the output signal, y_n is the input signal from IIR filter, n presents as an average order. Once after the clean signal of PR and RR obtained, peak detection algorithm can be applied.

$$y[n] = \frac{1}{a_0}\left(\sum_{i=0}^{P} b_i \cdot x[n-i] - \sum_{j=1}^{Q} a_j \cdot y[n-j]\right) \tag{15}$$

$$\mathrm{SMA} = \frac{y_1 + y_2 + \cdots + y_n}{n} \tag{16}$$

C. Parameter Calculation on HPS

Once after clean signal were obtained for both PR and RR parameter, we can calculate the parameters by applying equation as explained in section before. Where we firstly define the peak-to-peak interval for each signal, and then calculate the PR and RR. In vital signs parameter measurement, every parameter was calculated for every second (30 frames).

For age and gender prediction system, we just simply using inference of CNN by employing pre trained CNN model. Training stage was done offline in a PC, a 1000 random images for each age class were feed to the age prediction model. While 2000 images collected for each gender class used as input for gender prediction model. The inference can be used after the ROI image (face region) was resized to 227 x 227 pixels.

In the other hand, body height estimation skin color type classification done by using mathematic operation explained in section *II F* and *G* respectively We use same ROI area used in age and gender classification.

All parameter calculated in HPS were passed back to FPGA. Concurrently, FPGA will display captured image from camera and send the measurement results to smartphone through Bluetooth communication. Android application was developed to take over the task for displaying the results from measurement. Android application will display the results after measurement was taken, the display separated into 3 fragment, like seen in Figure 4, where PR value, RR value, and soft biometric parameters displayed on different fragment

IV. Experiment and Results

To validate our system performance, we conduct separate experiment for both vital sign measurement and soft biometric prediction. In vital sign measurement case, 5 subjects recruited to participate in this study. The overall experiment setup is seen in Figure 4. In the experiment, the subject was required to face the camera at the distance around 35 cm with the only light source from the indoor parallel incandescence lamps. For verification, the subject wearing ECG based heart rate and respiration rate from FDA approved tool (we use GE-Dash 3000 in this work). During measurement, the subject needs to keep static for 60 second. While for soft-biometric part we use 100 collected images as testing input from various source of database including self-collection database.

TABLE I. Results

Parameter	Value
Speed	Represented by fps = 33ms
FPGA Utilization	11699 of 41910 (28%) LE
Power	6 Watt
PR MAE	2.85 BPM
RR MAE	1.46 BPM
Age Accuracy	58 %
Gender Accuracy	74 %
Skin Color Accuracy	98 %
Body Height Accuracy	± 2.28 cm
HPS Utilization	200 % of dual core CPU

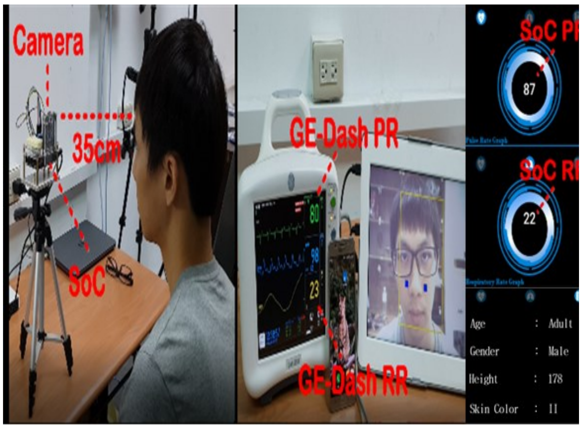


Figure 4 Experiment and Results

Table 1 provides the results of the experiment and SoC utilization. We use mean absolute error (MAE) to see the framework performance for PR and RR measurement. As well as for soft biometric parameter measurement in Table 1, we use accuracy in percentage to gives concrete measurement results. Hardware utilization gives information about how many resources were used from the SoC specification in our designed framework.

V. Discussion and Future Work

Our designed system has lower power consumption with maximum power is 6 Watt, the maximum load happens on CNN inference stage with 200% of dual core CPU usage. On the FPGA side, we utilized only 28% of total logic element (LE). However, the FPGA achieve standard for rPPG measurement with offer 30 fps image processing in 640 x 480 pixels resolution.

Compared with clinical apparatus, our vital sign parameters measurement results were close. Variables like light condition and skin color are affecting the measurement. In addition, a high signal-to-noise ratio (SNR) of respiration rate signal can be obtained if the subject has strong respiration movement.

We obtained unsatisfying results from age and gender prediction. From our experiment we got 58 % accuracy prediction for 4 age class, while 74 % accuracy for gender class. The error may be came from the overfitting of small dataset in training stage. And also the dataset used in training stage were not the same with image captured from our camera used in the system. In contract, our skin color type classification and body height measurement achieve high accuracy with 98 % and ± 2.28 cm respectively.

Proposed system has many prospect futures work including medical instrumentation, robotic, smart assisting technology, exercise, security, and automotive.

VI. Conclusion

In this paper, we proposed a contactless real-time human vital sign (pulse and respiration rate) and soft biometric (age, gender, skin color type and body height estimation) parameters measurement based on SoC platform. The results show the satisfactory performance by adopting the proposed framework. Moreover, proposed system provides for long-term and convenient measurement environment.

ACKNOWLEDGMENT

This work was supported in part by a research grant from the Ministry of Science and Technology under Grant MOST 106-3011-E-011-001 -. And further was funded by UMN for publication.

REFERENCES

[1] A. Rizal, Y. Lin and Y. Lin, "Contactless vital signs measurement for self-service healthcare kiosk in intelligent building," 2018 3rd International Conference on Intelligent Green Building and Smart Grid (IGBSG), Yi-Lan, 2018, pp. 1-4.

[2] H. Liu, Y. Wang, and L. Wang, "A review of non-contact, low cost physiological information measurement based on photoplethysmographic imaging," 34th Annual International Conference of IEEE EMBS, Sept.

[3] Y. C. Lin, G. Y. Lin and Y. H. Lin, "Development of a FPGA-based contactless pulse rate detection system," 2016 2nd International Conference on Intelligent Green Building and Smart Grid (IGBSG), Prague, 2016.

[4] Y. C. Lin, G. Y. Lin, J. W. Lin, Y. L. Tsai and Y. H. Lin, "An image-based pulse rate detection system using an FPGA platform," 2017 IEEE/SICE International Symposium on System Integration (SII), Taipei, 2017.

[5] Y. C. Lin, N. K. Chou, G. Y. Lin, M. H. Li, Y. H. Lin, "A Real-Time Contactless Pulse Rate and Motion Status Monitoring System Based on Complexion Tracking," Viik J, ed. Sensors, Basel, Switzerland. 2017.

[6] A. B. Hertzman and C. R. Spealman, "Observations on the finger volume pulse recorded photoelectrically," Amer. J. Physiol. Meas., vol. 119, pp. 334-335, 1937.

[7] W. Wang, A. C. den Brinker, S. Stuijk and G. de Haan, "Algorithmic Principles of Remote PPG," in *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1479-1491, July 2017.

[8] A. Secerbegovic, J. Bergsland, P. S. Halvorsen, N. Suljanovic, A. Mujcic and I. Balasingham, "Blood pressure estimation using video plethysmography," *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, Prague, 2016, pp. 461-464.

[9] K. Alghoul, S. Alharthi, H. Al Osman and A. El Saddik, "Heart Rate Variability Extraction From Videos Signals: ICA vs. EVM Comparison," in *IEEE Access*, vol. 5, pp. 4711-4719, 2017.

[10] D. Shao *et al*., "Noncontact Monitoring of Blood Oxygen Saturation Using Camera and Dual-Wavelength Imaging System," in *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 6, June 2016.

[11] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William Freeman. 2012. Eulerian video magnification for revealing subtle changes in the world. ACM Trans. Graph. 31, 4, Article 65 (July 2012), 8 pages.

[12] Dantcheva, A., Velardo, C., D'Angelo, A., and Dugelay, J.-L. (2011). Bag of soft biometrics for person identification. Multimedia Tools and Applications, 51(2), 739–777.

[13] S. Tiwari, A. Singh and S. K. Singh, "Can face and soft-biometric traits assist in recognition of newborn?," 2012 1st International Conference on Recent Advances in Information Technology (RAIT), Dhanbad, 2012, pp. 74-79.

[14] G. Levi and T. Hassncer, "Age and gender classification using convolutional neural networks," 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Boston, MA, 2015, pp. 34-42.

[15] M, W., CY, W., duPlessis JL, and A, R. (2015). Fitzpatrick skin type, individual typology angle, and melanin index in an african population: Steps toward universally applicable skin photosensitivity assessments. JAMA Dermatology, 151(8), 902–903.

[16] Terasic Inc. (2018). D8M-GPIO User Manual.