Video-Based Discrimination of Genuine and Counterfeit Facial Features Leveraging Cardiac Pulse Rhythm Signals in Access Control Systems

Murthad Al-Yoonus[1*] , Luqman Qader Abdulrahman[2] , Mudhafar Jalil Jassim Ghrabat[3]

[1] Department of Information Technology, Noble Technical Institute, Erbil 44001, Iraq
[2] College of Health Science, Hawler Medical University, Erbil 44001, Iraq
[3] Iraqi Commission for Computers and Informatics, the Informatics Institute for Postgraduate Studies, Baghdad 10081, Iraq

Corresponding Author Email: marthad.hussain@noble.edu.krd

ABSTRACT

Facial recognition technology, utilizing short-duration video, offers an accurate access control mechanism, capable of distinguishing genuine faces from fraudulent representations. However, the ability to detect concealed or artificially altered faces, as well as printed spoofs, remains a considerable challenge, even with sophisticated biometric recognition algorithms. This limitation can significantly undermine the performance of access control systems, rendering them susceptible to potential security breaches. In this study, we present an innovative technique that leverages the separation of Red, Green, and Blue (RGB) channels to discern between authentic faces and printed spoofs in color video recordings. This technique is deployed over a 20-second duration to effectively intercept and preclude security violations in access control systems. The proposed method was evaluated with a dataset from 20 participants, demonstrating commendable accuracy in detecting both real and spoof faces. Specifically, the technique exhibited remarkable precision in real face detection and counterfeit face detection. When applied to a 4-second video dataset from the same participants, comparable results were obtained. This method provides a significant advancement in the realm of face-controlled access systems by facilitating precise real-face detection via cost-effective video imaging, predicated on cardiac pulse rhythm signals. This methodology holds the potential to enhance the reliability of facial recognition systems, particularly in high-security environments.

1. INTRODUCTION

The accuracy of automated face recognition systems has seen substantial enhancements over recent decades. A myriad of advanced algorithms has been developed to overcome challenges associated with illumination, pose and expression variations, thereby enabling unconstrained face recognition across diverse applications [1]. However, despite these advancements, certain factors continue to impede the performance of access systems. These include issues pertaining to masked faces [2], printed faces, plastic surgery [3], make-up application [4], and spoofing faces [5]. For a face verification system to be effective in real-world scenarios, such as in securing access to restricted areas, a reliable system for distinguishing real and fake faces is essential.

Numerous approaches in the literature on fake face detection are based on the observation of texture properties in an image, comparing frames from a genuine person's face to spoofed face images.

In a recent paper, Kim et al. [2] proposed a method to distinguish a real face from a masked face using radiance measurements. This approach recovers the albedo from the reflectance of human facial skin and the reflectance of masked face materials (silicon, latex, or skinjell). Within a face verification system environment, Kim et al. simplified the computation of albedo from face reflectance, enabling the

measurement of radiance for the forehead region under 850 and 685 nm illumination.

In another paper, Chen et al. [4] studied the differences between a make-up face and a non-make up face, and designed a method to detect the presence of make up on the face by extracting the features from the color, shape and texture characteristics for the ROI of the right eye, left eye and mouth cropped from the input face. In his experiments he used SVM to classify the make-up face and nonmake up face. On the print photo attack database [6] in the IJCB 2011, Chakka et al. [7] found the performance of six spoofing printed face detection algorithms on that database in a competition for measuring 2D face spoofing attacks. Määttä et al. [8] detected the spoofing face from a single image-based microtexture analysis, and found that for spoofing face detection, local phase quantization with a Gabor wavelet-based descriptor was less efficient than (LBPs) local binary patterns. Furthermore, they demonstrated that using three LBP descriptors of different configurations was more efficient than using LBPs with a single configuration.

Määttä et al. [9] also introduced a fusion system leveraging two descriptors from Gabor wavelets obtained from the local blocks of a face image, LBPs, and a histogram of oriented gradients. The histogram calculated for all the blocks was concatenated for each descriptor, producing three feature vectors. The classification was carried out using a linear SVM, with the final result derived from the fusion of the match

scores of all three SVMs.

In another study aimed at detecting attacks from printed face photos, a fusion approach was employed which combined power spectrum and Local Binary Pattern (LBP) features, using a camera from an automated teller machine to compile the database [10]. These methods primarily relied on motion- and texture-based strategies, with facial movements leveraged to determine liveness.

Du et al. [11] undertook a project aimed at recognizing players' emotions during gameplay. This was achieved by measuring heart rate and detecting facial expressions in videos using continuous human emotion perception. In this work, a bidirectional long- and short-term memory (Bi-LSTM) network was deployed to learn the heart rate features, and a convolutional neural network (CNN) was trained to learn facial expression (FE) features. A self-organizing map back propagation (SOM-BP) was utilized to fuse these features and achieve an accurate match of the emotion.

Wang et al. [12] conducted a comprehensive review and comparative analysis of various methods for measuring heart rate from frontal face videos. The results of the survey indicate that using the heart rate signal extracted from facial skin as a source provides superior results, and that the independent component analysis (ICA) method outperforms others in signal extraction. Temko [13] introduced a novel algorithm, Wiener Filter and Phase Vocoder (WFPV), which employs a Wiener filter to reduce motion artifacts and a phase vocoder to refine heart rate estimates. Their heart rate estimation system was compared with existing algorithms on a database of 23 photoplethysmography (PPG) recordings.

Dautov et al. [14] implemented face detection using video plethysmography, with power spectral density (PSD) employed for heart rate monitoring. Nadrag et al. [15] utilized the Haar cascades method, initially proposed by Viola and Michel Jones, in combination with a fast Fourier transform (FFT) to identify the region of interest (ROI). Their paper presents a solution for measuring the heart rates of multiple individuals simultaneously, using object tracking to obtain a collection of face rectangles. The average color within the ROI is then used as the signal corresponding to the heart rate.

Zheng et al. [16] carried out heart rate predictions from facial videos with masks, using eye location as a reference point. Remote photoplethysmography (rPPG) was employed to extract signals from the face video, and a convolutional neural network (CNN) was utilized for heart rate detection.

Uppal et al. [17] conducted work on heart rate measurement using the brightness preserving bihistogram equalization (BBHE) technique. Their approach involves separating the captured image into red, blue, and green channels, with cheeks selected as the ROI. BBHE was applied to these regions, and heart rate measurement was performed using Principal Component Analysis (PCA).

Li et al. [18] proposed a framework that incorporates face tracking and the normalized least mean square adaptive filtering method. Discriminative response map fitting (DRMF) is employed to identify the ROI, and Welch's power spectral density estimation method is utilized for heart rate detection.

Gupta et al. [19] conducted heart rate monitoring using the Modeling and Bayesian Tracking (MOMBAT) method for detecting heart rate from face videos. Zheng et al. [20] detected heart rate using the symmetry substitution method, even in scenarios where facial details were lacking, such as during online classes. When the head is rotated approximately 30 to 40 degrees, causing partial loss of the ROI in the face,

information from the left and right cheeks is symmetrically copied.

A comprehensive review was conducted by Rouast et al. [21] in which remote heart rate measurements were taken using low-cost RGB face videos. In their analysis, they emphasize that future remote PPG (rPPG) algorithms should focus on balancing the trade-off between the amount of processed data and the complexity of the algorithm.

Although remote photoplethysmography (rPPG) is capable of detecting heart rates from facial videos, the accuracy can be compromised by head movement. To address this challenge, Wang et al. [22] proposed an anti-motion interference method termed T-distributed Stochastic Neighbor Embedding (T-SNE) Based Signal Separation (TSS). In this approach, TSS initially decomposes the observed color traces into pulse-related vectors and noise vectors using the T-SNE algorithm. The vector with the most significant spectral peak is then selected as the pulse signal for heart rate detection.

Ibrahim et al. [23] introduced a method to address the obstacles in heart rate estimation from cameras that capture videos from long distances. In their research, facial landmarks are computed using a cascaded regression mechanism. The region of interest (ROI) is selected based on these facial landmarks, specifically where minimal nonrigid motion is identified. Temporal photoplethysmography (PPG) signals are extracted from the ROI, and environmental illumination signals are removed using an independent component analysis (ICA) filter. This PPG signal is further processed by applying a series of temporal filters to exclude frequencies outside the range of interest before determining the heart rate.

Zhao et al. [24] proposed a novel method to extract pulse signals from ROIs across multiple scales. Their research constructs a facial ROI pyramid of multiple scale levels. Blood volume pulse (BVP) signals are then extracted, and the final pulse signal is computed via signal fusion, to which a convex combination is applied.

Špetlík et al. [25] proposed a two-stem convolutional neural network for heart rate estimation. Heart rate estimation is achieved remotely by tracking the peripheral circulation of the blood, that is, through a non-contact reflective photoplethysmographic (NrPPG) HR detection. Their system operates on two components, the extractor and the heart rate estimator. The extractor's role is to process an input image and output a single number. This extractor is applied to a sequence of images to obtain a sequence of scalar outputs and NrPPG signals. These signals are then input into the heart rate estimator to determine the heart rate.

Related work on spoof face detection relies on complex configurations and texture analysis to yield satisfactory results. In contrast, our proposed approach for detecting the liveness of real faces is based on the analysis of cardiac pulse rhythm signals, employing video imaging and RGB analysis. We conducted experiments with 20 subjects, each recording two videos—one of their real faces and one of a printed face. This research can be summarized as follows:

1)   A technique for a preprocessing method to detect the cardiac pulse signal from the green channel and red–green channel in real face videos based on color RGB separation and signal enhancements is proposed.

2)   Thresholds are set based on the difference between the max and min cardiac pulse signals.

3)   We propose a liveness detection technique for real faces and spoofing detection for fake printed faces based on the thresholds from the cardiac pulse signals yielded from the

green channel color of the ROI.

There are several proposed algorithms from existing works that measure the cardiac pulses from the face region only, with participants using video, that are non-contact [26, 27]. Some works use a thermal camera to monitor the change in the information in the thermal signal emitted from the vessels [28, 29]. In other works, a webcam camera was used to measure the cardiac pulses from the face using video imaging and separate the mixing channels; and FFT was used for the proper component after artifacts reduction from the separation channels [30, 31]. Another work used color magnification to amplify the color changes in facial skin [32].

There were artifact noise rejections, such as changes in the ambient lighting conditions and sudden motions of the participant [33]. We propose cardiac pulse signal detection using color RGB separation with artifact reduction, and we detect cardiac pulses from the green channel. FFT is used to obtain the frequency of the heart rate and then compares the results of cardiac pulses with a reference result from a commercial pulse oximeter, as shown in Figure 1.
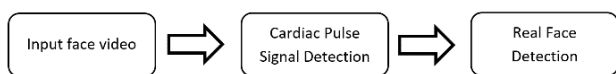

Figure 1. Real face detection based on cardiac pulse signal from input video of face

The rest of this paper is organized as follows. Section 2 describes the work, experimental setup and the database used for this study. Section 3 describes the method for cardiac pulse signal detection and artifact reduction for the green component and red-green component, and then describes the method for real face detection. Section 4 presents the experimental results and the effectiveness of the proposed work in real face detection and spoofing printed face detection, and then discusses the results. Finally, Section 5 concludes the paper and discusses future recommendations.

2. WORK DESCRIPTION

In this study, we enrolled 20 participants with varying demographic characteristics, including 16 males and 4 females, aged between 18-35 years, and from different ethnic backgrounds, including Malay, Arabic, Pakistani, and Chinese participants. To ensure diversity, participants were selected based on their varying skin colors. The video recording sessions were conducted at different times of the day and were recorded using a basic mobile camera. The videos were captured in color (24-bit RGB with 3 channels×8 bits/channel) at 30 frames per second (fps) with a pixel resolution of 720×404. The videos were saved in AVI format on a laptop for analysis in MATLAB. The lighting conditions during video capture were maintained as natural as possible, and the participants were asked to maintain a neutral facial expression while sitting in a comfortable posture. The video content focused on the participants' faces and upper bodies. All participants were seated on a chair in front of the mobile camera at a distance of approximately 0.7 m from the camera.

Two 20-second videos were recorded for each participant. The first video was captured while the participants were instructed to maintain a neutral facial expression and a steady gaze at a mobile camera. In the second 20-second video, participants printed faces on an A4 size paper. In addition, we recorded 4-second videos for all participants to demonstrate the capability of our method with reduced processing time. The region of interest (ROI) for all participants was selected as the small rectangular-shaped area of the nose and cheeks between the eyes and the mouth, excluding the regions covered by facial hair for some participants. The ROI selection was based on the Viola and Jones algorithm for nose detection, and 3/4th of the detected nose was added to both sides of the nose coordinates to yield our ROI, as shown in Figure 2 [34]. The RGB signals were calculated from the 24-bit color images, and the thresholds for signal identification were set based on the average pixel intensities in the ROI.
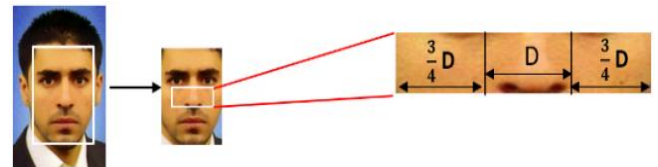

Figure 2. ROI selection for RGB extraction

2.1 RGB channels decomposition

In this study, the source signal of interest is the cardiovascular pulse wave that propagates throughout the body. For the real face from the first video, volumetric changes in the ROI of the facial blood vessels during the cardiac cycle modified the path length of the incident ambient light such that the subsequent changes in the amount of reflected light indicated the timing of the cardiac pulse events, while there were no changes in the amount of reflected light for the printed face on the A4 size paper. Therefore, from this feature of changes, according to the cardiovascular events from real human facial skin, we set some thresholds to recognize the real face from the printed face. By recording a video of the facial region with a mobile camera, the RGB color sensors pick up a mixture of the reflected signal. Each color sensor records a mixture of the original source signals with slightly different weights. To obtain the red, blue, and green weights for every frame, we summed all the pixel values separately in the ROI of each channel, resulting in 600 frame yields with r(t), g(t), and b(t), respectively. Next, three component signals were recovered by applying an enhancement technique to the RGB traces. This technique allowed for distinguishing between the real face and the printed face based on the resulting signals. Specific thresholds were set based on the cardiovascular events from real human facial skin, and these were compared to the RGB traces obtained from the printed face. Figure 3(a) shows face detection and the cropped ROI.
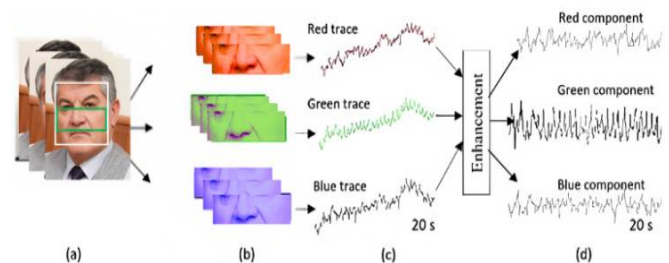

Figure 3. (a) The region of interest (ROI) is automatically cropped after being detected; (b) The ROI decomposed into the RGB channels; (c) The raw RGB traces. An enhancement is applied on the RGB traces to recover; (d) Three component signals

There are three RGB channels from the decomposed ROI (Figure 3(b)), and (Figure 3(c)) shows a three component signals obtained from the RGB channels. Then, three component signals were recovered by an applied enhancement on the RGB traces (Figure 3(d)).

3. DETECTION METHODOLOGY

Postprocessing of both the videos for real faces and printed faces was performed using demo software written in MATLAB version 2017b (The MathWorks, Inc.). Specifically, we used MATLAB's Signal Processing Toolbox to process and analyze the video data. This toolbox provides a comprehensive set of algorithms and tools for signal processing and data analysis. An overview of the main steps in this method to obtain the blood pulse from the real face video is outlined in Figure 3, and both signals from both videos are analyzed to set some thresholds to distinguish between the real face and fake face.

First, an automated face detector was used to detect the face within the video frames and localize the measurement region of interest (ROI) for every video frame. In this work, we used the face detector and nose detector based on the Viola and Jones algorithm for object detection [19], the algorithm returns the x- and y-coordinates along with the height and width that yields a box around the face for each face detected in the first fame for the input videos. From this box, we crop the center of the box by nose detection as an ROI for the subsequent calculations stated in Figure 2. The same coordinates of the selected region (ROI) are used for the entire sequence of frames in the input video.

To avoid face detection errors, if the face was not detected in the first frame, the face box coordinates from the next frame were used. If more than one face was detected, then the method takes the face box coordinates that were the closest to the box from the next frame. Some thresholds are set from the calculations for the ROI of all the frames, and our demonstration software shows a green box around the face with the label "Real Face" if the face is real in the input video, and shows a red box around the face with the label "Fake face" if the face is a spoof printed face on paper in the second input video.

3.1 Cardiac pulse signal detection and enhancement

A sum over all the pixel weights for every RGB channel in the ROI for each frame yields three components, r(t), g(t) and b(t) for red, green and blue, respectively, from all the video frames. The observed changes in the component plotting are based on the cardiovascular change events, as stated in Figure 4(a) for the g(t) component; always, the g(t) component is the best component for a cardiac pulse rate measurement [15].

The RGB sensors pick up other sources of fluctuation noise in the light with the cardiac pulse signals during video recordings of the participants due to artifacts such as changes in ambient lighting conditions and sudden motions of the participant, as plotted by the high amplitude changes of the red line in Figure 5(a). In this work, we used RGB sensors to detect cardiovascular changes in video recordings of participants. To process the RGB signals and reduce artifacts, we developed the following algorithm:

(1) Partition the green component of the acquired signal (20 sec video) into 40 partitions, as in Figure 5(a).

(2) Calculate the mean value of each partition, using the formula below:

$$\bar{X} = \frac{\sum_{i=1}^{n} pg(t)_i}{n} \tag{1}$$

where, g(t)=Green Component, pg(t)=partition.

(3) Shift the samples of every partition to the mean level, using the equation:

$$sg(t) = pg(t) - \bar{X} \tag{2}$$

where, sg(t)=Shifted partition.

(4) Repeat steps 2-3 above for all the partitions and then recombine the partitions to yield a cardiac pulse signal with a reduction in artifacts using the Eq. (3) and shown in Figure 5(b).

$$G(t)=[sg(t)1\ sg(t)2\ sg(t)3\ldots\ldots\ldots sg(t)60] \tag{3}$$

where, G(t)= recombined partitions.

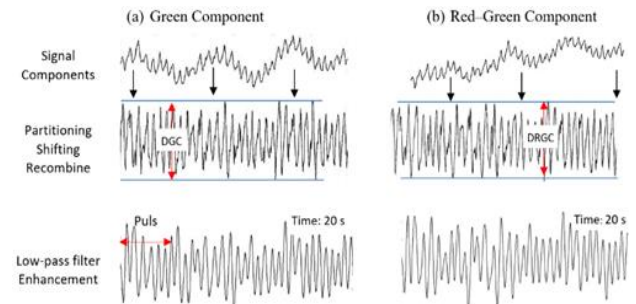(5) Use low-pass filter of (5 Hz) for the final enhancement as shown in Figure 5(c).

Figure 4. Real face video: (a) Green component and (b) Red–green component with enhancements
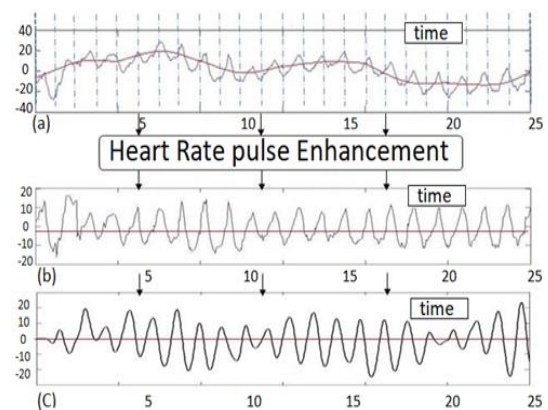

Figure 5. (a) Partitioning the signal from the green component and then (b) Shifting all partitions to the mean level; (c) Final Enhancement with low-pass filter

Therefore, we have provided a detailed explanation of the steps involved in our algorithm. By partitioning the green component into 40 partitions and shifting the samples to the mean level, we were able to reduce the effects of artifacts and obtain a cardiac pulse signal with enhanced accuracy. We also applied a low-pass filter of 5 Hz for the final enhancement.

We believe that our algorithm is effective in reducing artifacts and enhancing the accuracy of the cardiac pulse signal.

We have carefully selected the parameter values based on our analysis of the data.

In this work, there are four extraction components for each video, the red component, green component, blue components and (red + green) component. We use all the steps in Figure 5 of the enhancement of all the components (real face video and printed face video) for all the subjects, and then from this enhancement and artifact reduction, we can set accurate thresholds to distinguish between the real face and printed face.

3.2 Thresholds

The first threshold (T1) was set at a difference of 12 for the green components, while the second threshold (T2) was set at a difference of 18 for the red + green components. If the difference between the max and min weights for a vector exceeded these thresholds, we considered it to be a real face; otherwise, we classified it as a printed face. We also set an additional threshold of 100, which triggers a retest if the differences exceed this value, as shown in Table 1.

In addition, we also set a third threshold (T3) for the green components based on the number of pulses of any 4 sec with amplitudes greater than 12, with a minimum of 3 pulses required to classify a signal as a real face.

Table 1. Detection thresholds

Real Face	Printed Face	Re-Test
DGC>12	DGC<12	DGC>100
DRGC>18	DRGC<18	DRGC>100

*DGC: Difference between max and min in the green component.
*DRGC: Difference between max and min in the red + green component.

The specific threshold values used in our study were informed by previous research that investigated the cardiovascular events that occur in real human facial skin during the cardiac cycle [12]. We also considered factors such as the video capture equipment and lighting conditions. To validate our threshold setting method, we tested it on a sample of 20 participants with varying skin colors and ages, and compared the results to the ground truth obtained from manual inspection of the video data. The three final results were fused to the decision-maker for real face detection. For real face detection, the values of the results must be more than our thresholds. The three final results were fused to the decision-maker for real face detection.

4. EXPERIMENTAL RESULTS

4.1 Cardiac pulse signals

The green component extraction from the 600 frames for 20 s shows the heart pulses but is not accurate enough with the different peaks, and enhancement is still needed to yield clear pulse signals. The first stage of the results is based on our five-step enhancement method for the cardiac pulse signal by partitioning the green component and red-green component into 40 partitions and using steps 2, 3 and 4 and then recombining again, as shown in Figure 4.

The results of the obtained signals after step 4 show the accurate difference measurements between the max and min in the green and red-green components. We can clearly count the pulses in any 4 sec from the obtained signal of step 5 for

the green component. Real faces were detected based on the three results, DGC, DRGC and Pulse.

To compare the results of the cardiac pulse rate obtained from our method, we used a commercial pulse oximeter as a reference to measure the pulses in 30 sec, during the same time we recorded video for the participant's face. Figure 6 shows the use of FFT for the signal obtained from the enhanced green component. The frequency with a high peak obtained from FFT was clearly close to the reference frequency.

Table 2 shows the measurement of the cardiac pulse rate of five selected subjects. The results obtained from video imaging closely matched when compared to the reference pulse oximeter measurements.
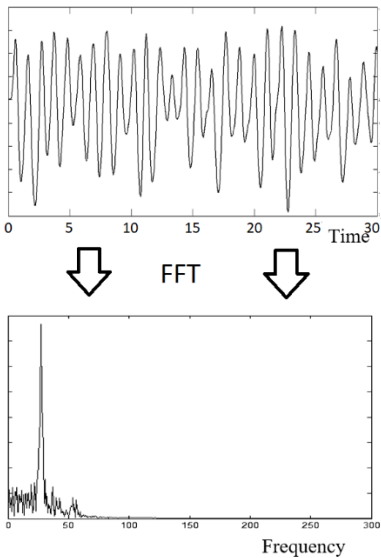
Figure 6. Use the FFT to the green component and the high peak in frequency domain for the heart rate

Table 2. Cardiac pulses rate measurements

From Video (FFT)	Pulse Oximeter
55	58
72	76
63	63
88	86
83	81

4.2 Real face detection

Figure 7 shows the real face detection and fake printed face detection results for one subject with both videos (20 s) for the real face video and the printed face video.

Figure 7. Real face and printed face detection for one participant

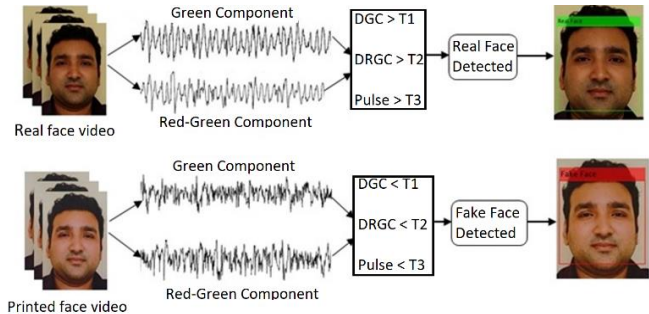Table 3 shows the results of the difference between the max and min in the green component (DGC), the difference between the max and min in the red–green component (DRGC) and the cardiac pulses in the first 4 sec. Table 3 shows the results of real face videos, and Table 4 shows the DGC and DRGC for the videos of printed faces on A4 size paper. The results in both tables are for 20 different subjects. The three final results for every subject are fused to the decision-maker based on the thresholds, and then used to decide the final decision for the detection of liveness in the real face or the spoof in the printed face.

Table 3. Results of real face videos (600f-20 Sec)

Thresholds (T1=12, T2=18, T3=3)				
Subjects	DGC	DRGC	Pulses	Fusion	Real Face
1	20	35	5		Detected
2	21	34	5		Detected
3	19	41	5		Detected
4	29	42	6		Detected
5	20	37	5		Detected
6	19	39	6		Detected
7	23	40	4		Detected
8	22	37	6	DGC>T1	Detected
9	18	39	5	&	Detected
10	18	27	6	DRGC>T2	Detected
11	26	32	4	&	Detected
12	17	31	4	Pulse>T3	Detected
13	20	39	5		Detected
14	22	41	4		Detected
15	23	34	6		Detected
16	20	38	5		Detected
17	25	47	4		Detected
18	20	36	5		Detected
19	21	29	6		Detected
20	29	60	4		Detected

Table 4. Results of printed face videos (600f-20 sec)

Thresholds (T1=12, T2=18, T3=3)				
Subjects	DGC	DRGC	Pulses	Fusion	Printed Face
1	6	8	2		Detected
2	4	7	0		Detected
3	7	6	1		Detected
4	5	5	2		Detected
5	5	9	0		Detected
6	5	8	1		Detected
7	5	8	1		Detected
8	5	5	0	DGC<T1	Detected
9	5	4	0	&	Detected
10	5	8	0	DRGC<T2	Detected
11	6	11	1	&	Detected
12	6	11	2	Pulse<T3	Detected
13	5	7	1		Detected
14	6	9	2		Detected
15	6	10	2		Detected
16	4	11	0		Detected
17	6	12	2		Detected
18	4	8	1		Detected
19	6	6	2		Detected
20	7	13	1		Detected

All the results of DGC and DRGC obtained from the real face videos are clearly more than T1=12 and T2=18, and the results of DGC and DRGC obtained from the printed face videos are sharply less than T1=12 and T2=18, as shown in Figure 8 and Figure 9, respectively.
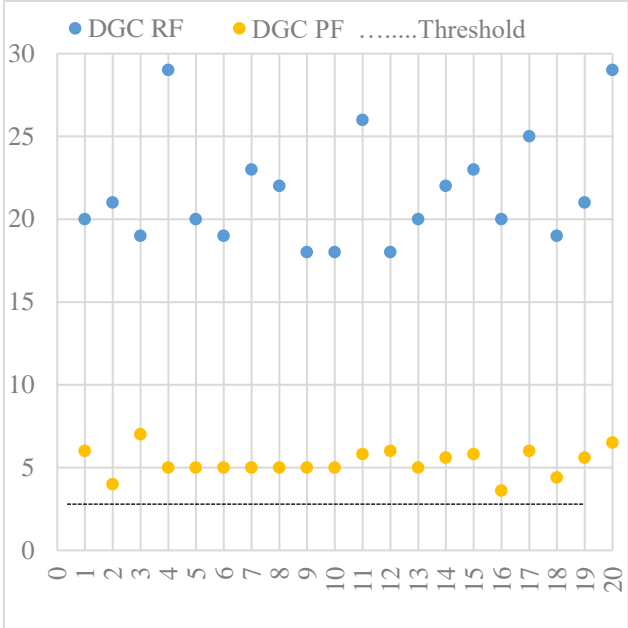

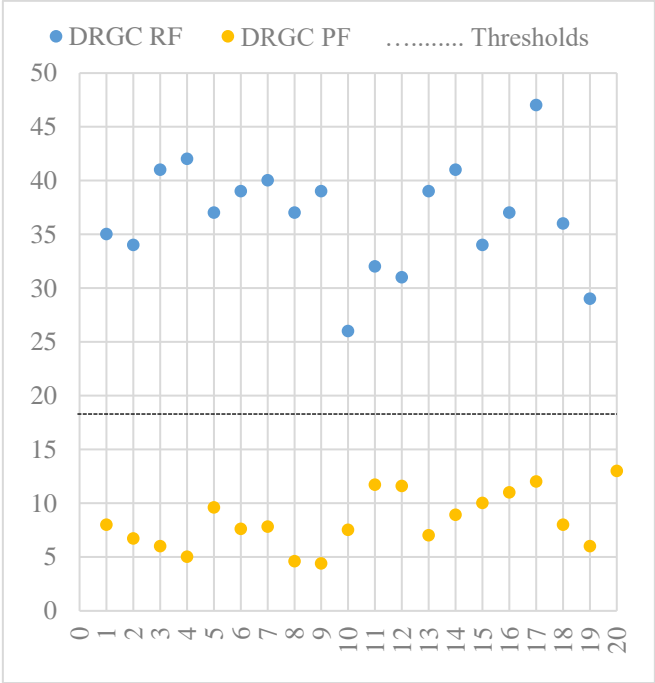**Figure 8.** Results of DGC for the real face (RF) videos and printed face (PF) videos

Figure 9. Results of DRGC for the real face (RF) videos and printed face (PF) videos

4.3 The results of the proposed method within the (4s) video

The show the capability of this proposed work for real face detection within short-time video recording (4 sec) only. We recorded a 4 s video for the same participants within the same field of view. The final results were accurate enough to detect the liveness in the real face and to detect the spoof in the printed face based on the three obtained results (DGC, DRGC and pulse) from the short cardiac pulse signals within 120 frames. Table 5 shows the results of real face detection and spoof face detection within the 4 sec video recording.

Table 5. Results of real face detection within short time video recording (4 sec)/Thresholds (T1=12, T2=18, T3=3)

Subjects	Real Face Videos			Printed Face Videos		
	DGC	DRGC	Pulses	DGC	DRGC	Pulses
1	20	35	5	6	7	0
2	20	36	6	4	5	0
3	22	40	5	6	7	0
4	19	31	6	5	5	0
5	20	36	5	4	5	0
6	19	31	7	4	7	0
7	22	40	4	5	9	0
8	17	27	6	4	5	0
9	16	24	5	5	6	0
10	18	26	5	5	6	0
11	26	31	4	5	9	0
12	18	26	6	5	11	0
13	17	22	5	5	3	0
14	19	38	4	5	6	0
15	21	31	6	7	6	1
16	20	37	5	4	7	0
17	18	29	4	5	4	0
18	19	33	5	5	9	0
19	17	27	5	6	6	0
20	17	46	4	6	11	1
	DGC>T1	DRGC>T2	Pulses>T3	DGC<T1	DRGC<T2	Pulses<T3
	Real Face Detected			Printed Face Detected		

5. DISCUSSION

The main use of this proposed work for liveness detection in face recognition systems is to access secure areas. Fake printed face detection must be rapidly improved to enhance facial biometric systems. We detected liveness in the real faces based on cardiac pulse rhythms. We extracted the cardiac pulse signal using the method detailed in Sections 2 and 3.

The strongest color change information in the ROI corresponding to the cardiac events was in the green component and red-green component. For some participants, the forehead is covered by a scarf or hair, and the chin region is covered by a beard. A small rectangular-shaped part of the nose and cheeks was selected as an ROI to reduce the computation time and complexity, and this ROI was used to avoid noise from eye blinking and from mouth motions. The experiments were performed at different times of day with different illumination ranges. We did not address how the proposed algorithm will work in a low lighting environment. Due to the small range of artifacts in our experiments from the motion and illumination changes, our proposed methodology can easily overcome these artifacts.

For further artifacts from rapid head movement, alternative artifact rejection could be used for improvements in future works [35, 36]. In addition, for this work, the video recording time was not short enough, future work still needs further reduce the video recording time to enable detection of real face liveness with less time.

6. CONCLUSIONS

We have demonstrated a simple low-cost new methodology for detecting liveness in real faces based on cardiac pulse rhythms obtained from the video imaging of human faces using a simple mobile camera for data collection in normal daylight illumination. The technique showed the ability to detect spoofing faces, including printed face photos and masked faces. Furthermore, we have shown that this method can easily work within a short recording video (4s). This technique is under development and extension to use in robust, accurate face recognition access systems. Creating a real-time detection system based on multiple vital sign parameters, such as cardiac pulse rhythms and respiratory movements, to detect liveness in real faces, and the use of better classifications will be the next topics for future work.

REFERENCES

[1] Weigelt, S., Koldewyn, K., Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: A review of behavioral studies. Neuroscience & Biobehavioral Review, 36(1): 1060-1084. https://doi.org/10.1016/j.neubiorev.2011.12.008

[2] Kim, Y., Na, J., Yoon, S., Yi, J. (2009). Masked fake face detection using radiance measurements. Journal of the Optical Society of America A, 26(4): 760-766. https://doi.org/10.1364/JOSAA.26.000760

[3] Bhatt, H.S., Bharadwaj, S., Singh, R., Vatsa, M. (2013). Recognizing surgically altered face images using multi-objective evolutionary algorithm. IEEE Transactions on Information Forensics and Security, 8(1): 89-100. https://doi.org/10.1109/TIFS.2012.2223684

[4] Chen, C., Dantcheva, A., Ross, A. (2013). Automatic facial makeup detection with application in face recognition. In 2013 International Conference on Biometrics (ICB), Madrid, Spain, pp. 1-8. https://doi.org/10.1109/ICB.2013.6612994

[5] Maatta, J., Hadid, A., Pietikäinen, M. (2012). Face spoofing detection from single images using texture and local shape analysis. IET Biometrics, 1(1): 3-10. https://doi.org/10.1049/iet-bmt.2011.0009

[6] Anjos, A., Marcel, S. (2011). Counter-measures to photo attacks in face recognition: A public database and a baseline. In 2011 International Joint Conference on Biometrics (IJCB), Washington DC, USA, pp. 1-7. https://doi.org/10.1109/IJCB.2011.6117503

[7] Chakka, M.M., Anjos, A., Marcel, S., Tronci, R. (2011). Competition on counter measures to 2-D facial spoofing attacks. In 2011 International Joint Conference on Biometrics (IJCB), Washington DC, USA, pp. 1-6. https://doi.org/10.1109/IJCB.2011.6117509

[8] Määttä, J., Hadid, A., Pietikainen, M. (2011). Face spoofing detection from single images using micro-texture analysis. In 2011 International Joint Conference on Biometrics (IJCB), Washington DC, USA, pp. 1-7. https://doi.org/10.1109/IJCB.2011.6117510

[9] Määttä, J., Hadid, A., Pietikainen, M. (2012). Face spoofing detection from single images using texture and local shape analysis. IET Biometrics, 1(1): 3-10. https://doi.org/10.1049/iet-bmt.2011.0009

[10] Kim, G., Eum, S., Suhr, J.K., Kim, D.I. (2012). Face liveness detection based on texture and frequency analyses. In 2012 5th IAPR International Conference on Biometrics (ICB), New Delhi, India, pp. 67-72. https://doi.org/10.1109/ICB.2012.6199760

[11] Du, G.L., Long, S.Y., Yuan, H. (2020). Non-contact emotion recognition combining heart rate and facial expression for interactive gaming environments. IEEE Access, 8: 11896-11906. https://doi.org/10.1109/ACCESS.2020.2964794

[12] Wang, C., Pun, T., Chanel, G. (2018). A comparative survey of methods for remote heart rate detection from frontal face videos. Frontiers in Bioengineering and Biotechnology, 6: 33. https://doi.org/10.3389/fbioe.2018.00033

[13] Temko, A. (2017). Accurate heart rate monitoring during physical exercises using PPG. IEEE Transactions on Biomedical Engineering, 64(9): 2016-2024. https://doi.org/10.1109/TBME.2017.2676243

[14] Dautov, R., Savur, C., Tsouri, G. (2018). On the effect of face detection on heart rate estimation in video plethysmography. 2018 IEEE Western New York Image and Signal Processing Workshop (WNYISPW), Rochester, NY, USA, pp. 1-5. https://doi.org/10.1109/WNYIPW.2018.8576439

[15] Nadrag, C., Poenaru, V., Suciu, G. (2018). Heart rate measurement using face detection in video. In IEEE International Conference on Communications COMM, Bucharest, Romania, Bucharest, Romania, pp. 131-134. https://doi.org/10.1109/ICComm.2018.8484779

[16] Zheng, K., Ci, K.Y., Li, H., Shao, L., Sun, G.M., Liu, J.H., Cui, J.L. (2022). Heart rate prediction from facial video with masks using eye location and corrected by convolutional neural networks. Biomedical Signal Processing and Control, 75: 103609. https://doi.org/10.1016/j.bspc.2022.103609

[17] Uppal, G., Prakash, N.R., Kalra, P. (2017). Heart rate measurement using facial videos. Advances in Computational Sciences and Technology, 10(8): 2343-2357.

[18] Li, X.B., Chen, J., Zhao, G.Y., Pietikainen, M. (2014). Remote heart rate measurement from face videos under realistic situations. In 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, pp. 4264-4271. https://doi.org/10.1109/CVPR.2014.543

[19] Gupta, P., Bhowmick, B., Pal, A. (2020). MOMBAT: Heart rate monitoring from face video using pulse modeling and Bayesian tracking. Computers in Biology and Medicine, 121: 103813.

https://doi.org/10.1016/j.compbiomed.2020.103813

[20] Zheng, K., Ci, K.Y., Cui, J.L., Kong, J.P., Zhou, J. (2020). Non-contact heart rate detection when face information is missing during online learning. Sensors, 20(24): 7021. https://doi.org/10.3390/s20247021

[21] Rouast, P.V., Adam, M.T., Chiong, R., Cornforth, D., Lux, E. (2018). Remote heart rate measurement using low-cost RGB face video: A technical literature review. Frontiers of Computer Science, 12(5): 858-872. https://doi.org/10.1007/s11704-016-6243-6

[22] Wang, H.Q., Yang, X.Z., Liu, X.N., Wang, D.L. (2022). Heart rate estimation from facial videos with motion interference using T-SNE-based signal separation. Biomedical Optics Express, 13(9): 4494-4509. https://doi.org/10.1364/BOE.457774

[23] Ibrahim, N., Tomari, R., Zakaria, W.N.W., Othman, N. (2018). Analysis of non-invasive video based heart rate monitoring system obtained from various distances and different facial spot. Journal of Physics: IOP Publishing, 1049(1): 012003. https://doi.org/10.1088/1742-6596/1049/1/012003

[24] Zhao, C.C., Han, W.R., Chen, Z., Li, Y.Q., Feng, Y.Q. (2020). Remote estimation of heart rate based on multi-scale facial ROIs. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, pp. 278-279. https://doi.org/10.1109/CVPRW50498.2020.00147

[25] Špetlík, R., Franc, V., Matas, J. (2018). Visual heart rate estimation with convolutional neural network. In British Machine Vision Conference, Newcastle, UK, pp. 3-6.

[26] Kranjec, J., Beguš, S., Geršak, G. (2014). Non-contact heart rate and heart rate variability measurements: A review. Biomedical Signal Processing and Control, 13: 102-112. https://doi.org/10.1016/j.bspc.2014.03.004

[27] Aarts, L.A.M., Jeanne, V., Cleary, J.P., Lieber, C., Stuart Nelson, J., Oetomo, S.B., Verkruysse, W. (2013). Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit-A pilot study. Early Human Development, 89(12): 943-948.

https://doi.org/10.1016/j.earlhumdev.2013.09.016

[28] Garbey, M., Sun, N., Merla, A., Pavlidis, I. (2007). Contact-free measurement of cardiac pulse based on the analysis of thermal imager. Biomedical Engineering, IEEE Transactions, 54(8): 1418-1426. https://doi.org/10.1109/TBME.2007.891930

[29] Hamedani, K., Bahmani, Z., Mohammadian, A. (2016). Spatio-temporal filtering of thermal video sequences for heart rate estimation. Expert Systems with Applications, 54: 88-94. https://doi.org/10.1016/j.eswa.2016.01.022

[30] Poh, M.Z., McDuff, D.J., Picard, R.W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Optics Express, 18(10): 10762-10774. https://doi.org/10.1364/OE.18.010762

[31] Al-Yoonus, M., Alhabib, M.H., Al-Dabagh, M.Z.N., Abdullah, M.F.L. (2020). Motion artifacts reduction in cardiac pulse signal acquired from video imaging. International Journal of Electrical and Computer Engineering, 10(6): 5687-5693. https://doi.org/10.11591/ijece.v10i6

[32] Liu, C., Torralba, A., Freeman, W.T., Durand, F., Adelson, E.H. (2005). Motion magnification. ACM Transactions on Graphics (TOG), 24(3): 519-526.

https://doi.org/10.1145/1073204.1073223

[33] Krishnan, R., Natarajan, B., Warren, S. (2010). Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data. IEEE Transactions on Biomedical Engineering, 57(8): 1867-1876. https://doi.org/10.1109/TBME.2009.2039568

[34] Viola, P., Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, HI, USA, pp. 511-518. https://doi.org/10.1109/CVPR.2001.990517

[35] Wang, Z., Wong, C.M., Cruz, J.N., Wan, F., Mak, P.I., Mak, P.U., Vai, M.I. (2014). Muscle and electrode motion artifacts reduction in ECG using adaptive Fourier decomposition. In 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), CA, USA, pp. 1456-1461. https://doi.org/10.1109/SMC.2014.6974120

[36] Ansari, S., Belle, A., Hobson, R., Najarian, K. (2011). Reduction of periodic motion artifacts from impedance plethysmography. In IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW), Atlanta, GA, pp. 540-547. https://doi.org/10.1109/BIBMW.2011.6112427