iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52812

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Heart Rate Estimation using Hermite Transform Video Magnification and Deep Learning

M. Jaya Joshna[4], Y. Shashank Varma[2], M. Ikram Mirza[3], P. Sai Kumar[4], Dr. Maznu Shaik[5]
Department of Electronic and Communication Engineering, Vidya Jyothi Institute of Technology, CB post, Aziz Nagar Gate, Hyderabad, Telangana, India

Abstract: ***Monitoring of heart rate can be used in many medical and sports applications. Lack of portability and connection problems make traditional monitoring methods difficult to use outside of clinical environments. The computer vision techniques have shown that some physiological variables as heart rate can be measured without contact. Video magnification is one of these approaches used for the detection of the pulse signal. In this paper we propose a new strategy to magnify motion in a video sequence using the Hermite transform. In addition, a deep learning technique is implemented to estimate the beat-by-beat pulse signal. We trained the system and validated our results using an electronic pulse monitoring device. Our approach is compared with the classical video magnification using a Gaussian pyramid. The results show a better enhancement of spectral information from the color changes allowing an accurate estimation of the instantaneous beat by beat pulse than the Gaussian approach.***
Keywords: heart rate, video magnification, Hermite transform, deep learning.

I. INTRODUCTION

Monitoring of heart and respiratory rate can be used in many areas including medical, sports and daily activities applications. Traditional methods using contact sensors have proven their effectiveness but have important drawbacks related to the complexity of their use and the lack of portability. For example, the electrocardiogram (EEG) requires adhesivegel patches that can cause skin irritation and discomfort and the oximetry sensors can cause pain over a long period of time. In addition, outside of clinical environments, these measures restrict the movement in many activities, such as sports training and prevent daily activivties in monitored people as the elderly. The monitoring of vital signals as using non-contact techniques such as computer vision has been developed in the last ten years. For a review of the non-contact techniques using computer vision see [1]. Video motion magnification is one of these methods that allows to enlarge the motion in a video sequence. Wu et al. [8] and Wadhwaet al. [7] used a Eulerian video magnification technique to amplify the imperceptible movements in a sequence, their approach tracks motion using a Laplacian decomposition and apply a band pass filter to extract the chest movements produced by the respiration. Recently, Breva et al. [1], use a decomposition based on the Hermite transform (HT) that allows a better reconstruction in the final magnified video decreasing the reconstruction error. In addition, Wu et al [8] use a Gaussian pyramid to track changes in chrominance for each pixel over time and magnify the pulse signal. Recently, authors in [5] validated video magnification approaches for heart rate estimation by comparing results with standard medical devices. This paper proposes a new method to estimate the instantaneous beat by beat pulse signal using the Hermite transform and deep learning. First, a magnification technique amplifies the changes in luminance using the Hermite transform and then, a convolutional neural network (CNN) is trained to estimate the beat-by-beat pulse signal. For the experiments we collected data video using a standard camera. Results are validated with respect to a personal pulse monitoring device. The rest of the sections are arranged as follows: Section II describes the proposed approach, Section III explains the experiments and the protocol, Section IV discusses results and finaly, conclusion is presented.

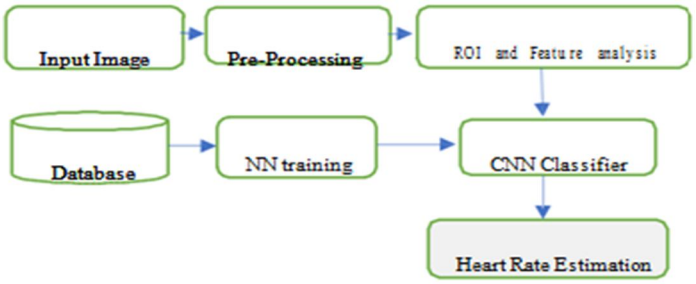


Fig. 1: Block diagram

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

II. VIDEO PROCESSING AND ANALYSIS

This section describes an overview of the proposed approach. First, we implement a video magnification method that amplify changes in chrominance using the Hermite trans- form that performs a spatial decomposition of the images allowing to eliminate most of the spatial noise. To detect the motion components to amplify, we use a band-pass filter with fixed cutting frequencies covering a sufficiently wide range for a set of healthy subjects at rest. Then, we choose a suitable region including the arm wrist. The amplified signal from each 10-by-10 patches is extracted and the maximum energy patch signal is chosen. Then a window profile is built using all the temporal signals to train the CNN. The overall system is shown in Fig. 2.

A. Video Magnification

The HT is an image multi-resolution decomposition model, used on the well-known analysis and synthesis processes, inspired in the human vision system (HVS) [3]. It consists on the decomposition of the image (analysis) usinga shifting Gaussian window and then expanded into a family of polynomials that are orthogonal to the window function.

Then, a reconstruction step (synthesis) using the same HT family basis is performed. The proposed magnification method uses the Eulerian approach [8] that consists of amplifying chromatic components in an image sequence. The procedure considers the following steps:

1) A multi-resolution spatial decomposition of low frequency components using the HT is performed.
2) A broad-band temporal filter is applied to the sequence.The cutting frequencies are chosen to cover a sufficiently wide range including the spectral information
3) A magnification of retained components of the chromatic temporal changes in the sequence is applied.
4) A reconstruction of the image sequence by the HT synthesis procedure is carried out.
5) The magnified sequence is added to the original one. *B. ROI detection and profile extraction*

After video magnification, a raw temporal signal is extracted from the average chrominance for each of the 10- by-10 pixels patches in the sequence. Hence, we divide the image in a grid obtaining 3600 patches. One final patchis selected from the maximal energy signal of the suitable region. The region of interest (ROI) is chosen taken a neighborhood of the selected patch. For each pixel on the ROI a temporal variation signal is extracted. Then, a gray level image is formed by stacking the signals. Hence, we associate for each window a heart rate reference value to use for training the CNN.

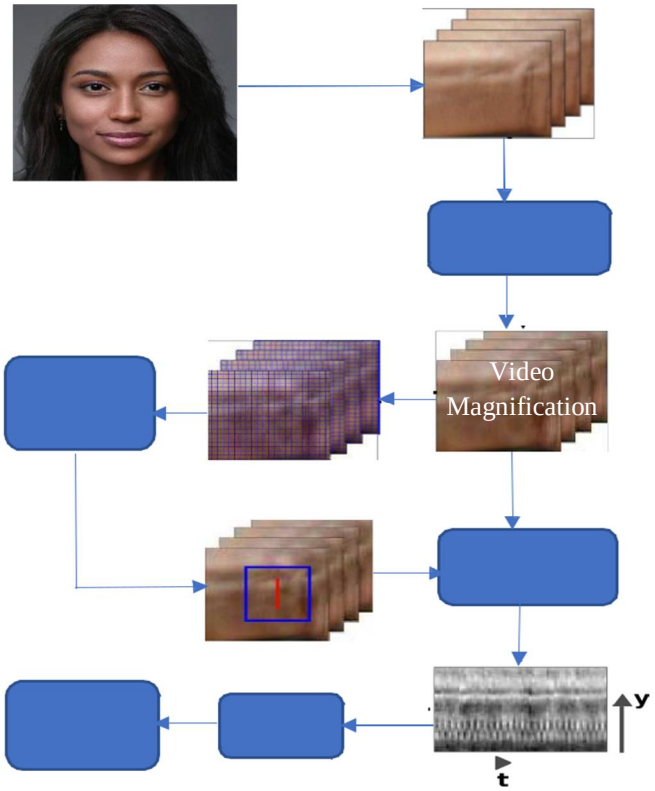


Fig.2:Block diagram of the heart rate estimation system.

©IJRASET: All Rights are Reserved

B. Deep Learning-Based Contactless Heart Rate Measurement Methods:

Contactless methods involve the utilization of a video camera and image processing algorithms. Recently, deep learning methods have been used to improve the performance of conventional contactless methods for heart rate measurement. Photoplethysmography(PPG) is a physiological measurement method that is used to detect volumetric changes in blood in vessels beneath the skin (measure heart rate (HR), respiratory rate, and blood pressure). The physics of rPPG is similar to contact-based PPG.

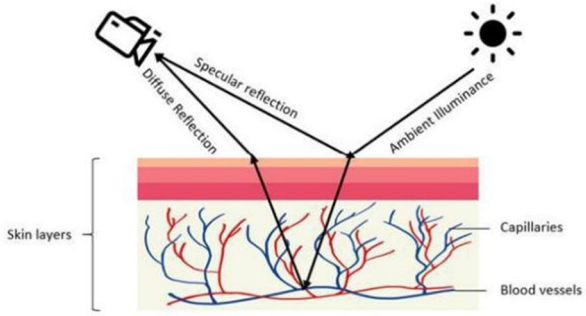


Fig 3: Illustration of rPPG generation

In rPPG(Remote Photoplethysmography) methods, the light-emitting diode in contact-based PPG methods is replaced with ambient illuminance, and the photodetector is replaced with a video camera. The light reaching the camera sensor can be separated into static (DC) and dynamic (AC) components. DC Components – corresponds to static elements including tissue, bone, and static blood.AC Components - corresponds to the variations in light absorption due to arterial blood volume changes.

C. Estimation of the heart rate

Convolutional neural network (CNN) is a deep learning method based on the nature of visual perception, mostly applied in image processing [2], [4]. Typically, CNN is employed for regression and classification problems

[4]. It is mainly constituted by three types of layers: (i) convolutional, (ii) pooling, and (iii) fully connected. Convolutional layers aim to compute feature representations of the input, pooling layers aim to reduce the resolution of feature maps, and fully-connected layers aim to perform high-level reasoning [2]. At the end of the CNN, it requires an output layer to perform regression or classification tasks.

In this regard, we propose the use of a simple CNN for heart pulse estimation enhanced with magnification information, as described above. For implementation purposes, the profile image is resized (20×30) and introduced as input to the CNN. Then, a convolutional layer with 25 filters of size (12×2), followed by a rectified linear unit layer, is placed to compute feature representations of the input image. Finally,a fully connected layer of size 1 and a regression layer are considered with the intention to perform high-level reasoningfor heart pulse estimation.

III. RESULTS

We have used a personal contact pulse monitoring device to measure the reference pulse signal in beat per minute(bpm) while recording video with a standard RGB cameraat a rate of 30 frames per second (see Fig. 3). We use a sampling frequency widely large to cover all the spectral frequency of the signal.

We have recorded three healthy subjects and 3 sessions per participant. Subject stayed intrest in normal conditions(they held normal conversation with the team acquisition members) for 2 minutes. In the magnification process, the applied broad- band temporal filter have cutting frequencies from 0.833-10583 *Hz*(50-95 bpm) which allow us to cover a sufficiently wide range including the spectral information for the conditions of acquisition. In all experiments we used a magnification factor of 50.

We trained one CNN for each subject. The data collection for each subject contains 348 instances representing the 5 s windows with their respective references. Particularly, we divided the data collection in 70% (244 instances) training set and 30% (104 instances) testing set, with samples randomly chosen. For each subject, we employed data from three trials.Notice that temporal information was considered at each sample in the dataset, since it considers a profile image that synthesizes 5s of the signal. For training, a stochastic gradient descent algorithm with initial learningrate of 0.01, momentum 0.90, mini-batch size of 64 and *L*2regularization was performed with $\lambda = 0.0001$. The same training parameters of the CNN were used for each subject.

Fig 4(a): Displaying the heart rate through the input face image as 91.7 bpm.

Fig 4(b): Displaying the heart rate through the input face image as 79.5 bpm.

Fig 4(c): Displaying the heart rate through the input face image as 100.1 bpm.

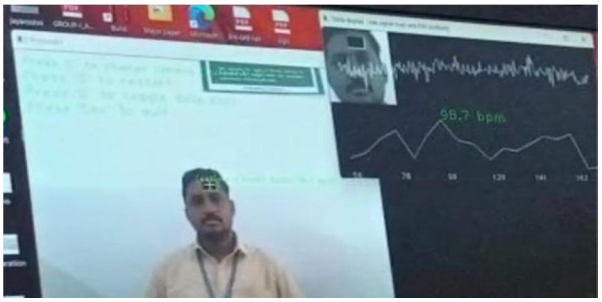


Fig 4(d): Displaying the heart rate through the input face image as 98.7 bpm.

Comparing the heart rate values with contact(sensors) and without contact.

without contact	0	91.7	85.3	78.6	74.9
with contact 1	0	89.6	83.7	75.9	71.2
with contact 2	0	88.1	83.9	77.4	72.3
	0	10	20	30	40

Table 1: Compairing values of fig 4(a) with and without contact.

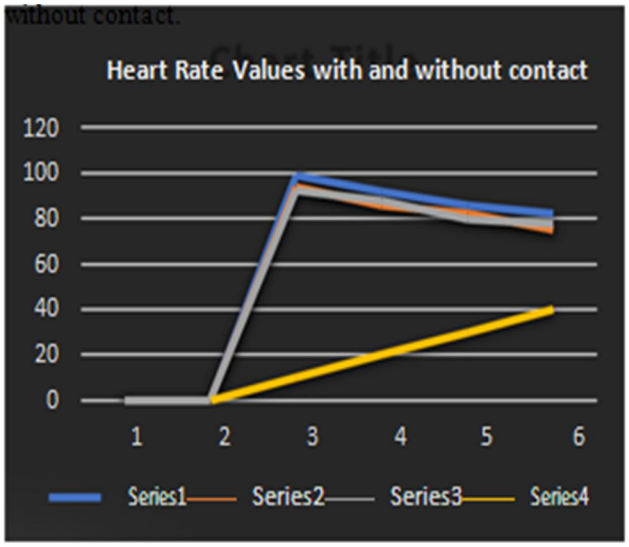


Fig 5: Graph of fig 4(a) with an without contact.

without contact	0	79.5	70.8	88.6	81.9
with contact 1	0	74.9	66.4	80.9	79.2
with contact 2	0	72.6	61.8	84.8	75.8
	0	10	20	30	40

Table 2 : Here , we are going to compare both the outputs of heartrate using sensor with contact and without contact.

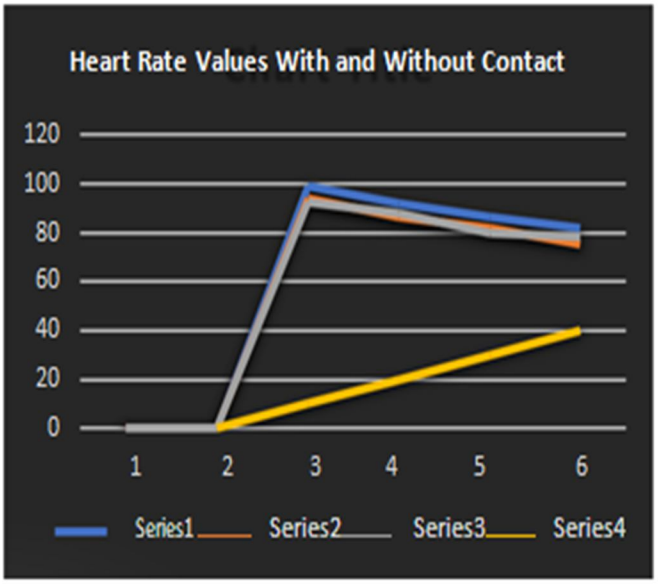


Fig 5: Graph of fig 4(b) with and without contact.

without contact	0	100.1	91.7	85.9	81.9
with contact 1	0	93.6	85.9	81.4	74.9
with contact 2	0	92.5	87.8	79.8	77.9
	0	10	20	30	40

Table 3: Compairing values of fig 4(c) with and without contact.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

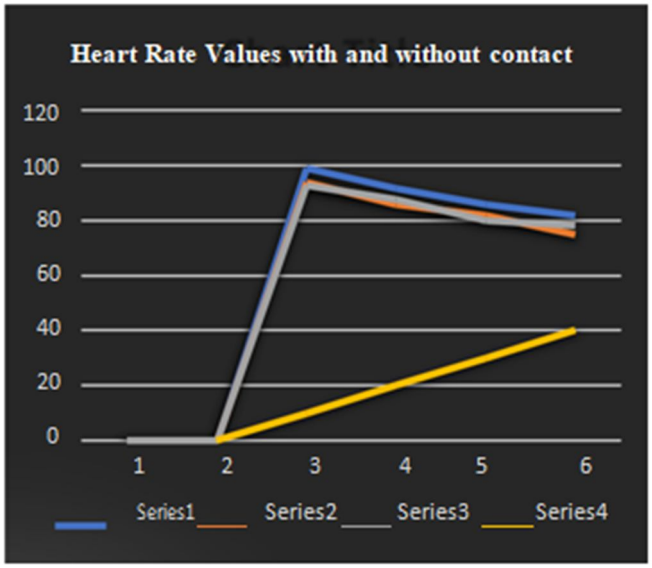


Fig 6: Graph of fig 4(c) with and without contact.

The most critical parameters in the magnification process are the cutting frequencies of the temporal filtering. Normally, these parameters must be tuned for each patient. In our approach we fixed a wide range of frequencies for all the subjects without additional refinements.

This approach is limited to controlled conditions of acquisitionsuch as illumination, motion constraints of the subjects and manual selection of the ROI. In this sense, we avoid registration and detection processes.

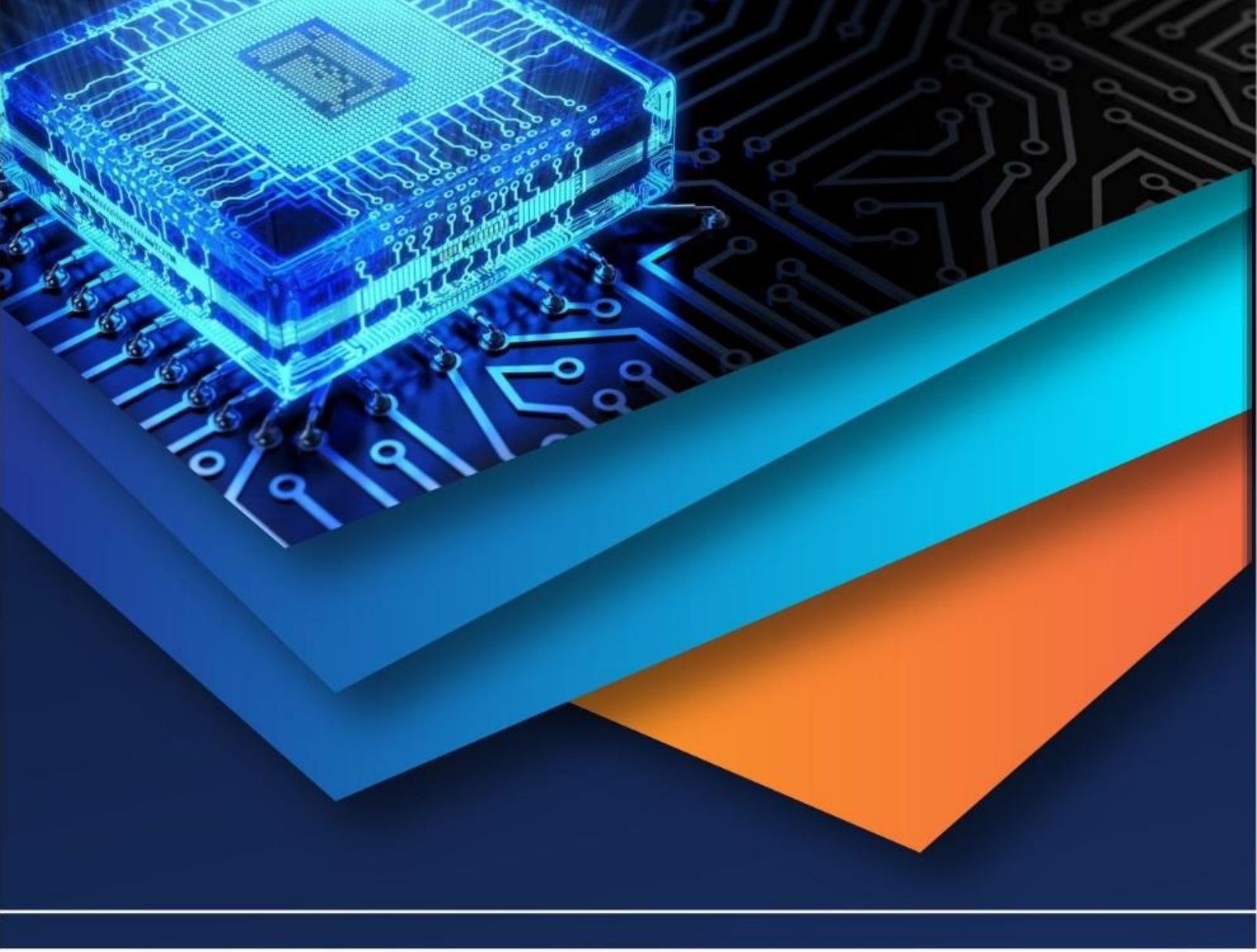
IV. CONCLUSION

In this work we have proposed a new magnification method using the Hermite transform and a CNN to estimate the instantaneous beat by beat pulse signal. The magnification technique amplifies the changes in chrominance using the Hermite transform to extract a temporal signal. Then, a CNN is trained to estimate the heart rate. We tested our approach using image sequences and beatby beat signal pulse acquired from a standard RGB cameraand a personal contact pulse monitoring device as reference.

REFERENCES

[1] Strong robustness heart rate estimation using discrete fourier transform and personality heart rate characteristic by Xiangze Li and Baoming Pu - 2019
[2] A Metabolic Rate Estimation Model Based on Heart Rate and Respiratory Rate byHexiang Zhang, Tanqiu Li, Tao Wang, Kun Shang - 2021
[3] Heart Rate Estimation using PPG signal during Treadmill Exercise by Youngsun Kong and Ki Chon - 2021
[4] A Multi-Spectral Database for NIR Heart Rate Estimation by Michal Rapczynski, Chen Zhang, Ayoub Al-Hamadi and Gunther Notni – 2018.
[5] Noncontact Heart Rate Detection Method Based on Zekun Chen, Yunzxue Liu and ZhuroanCai - 2022
[6] Jean-Bernard Martens. The Hermite Transform-Theory. IEEE Transactions on Acoustics, Speech and Signal Processing, 38(9):1595–1606, 1990.
[7] K. Nogueira, O.A. Penatti, and J.A. dos Santos. Towards better exploiting convolutional neural networks for remote sensing scene classification. Pattern Recognition, 61:539 – 556, 2017.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)