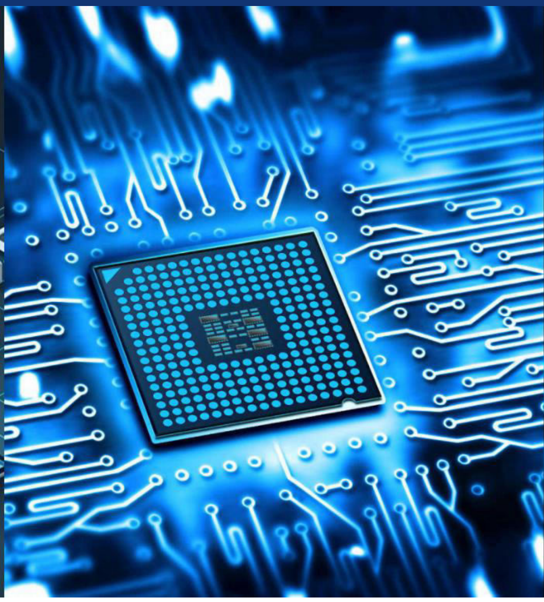
IJIRCCE

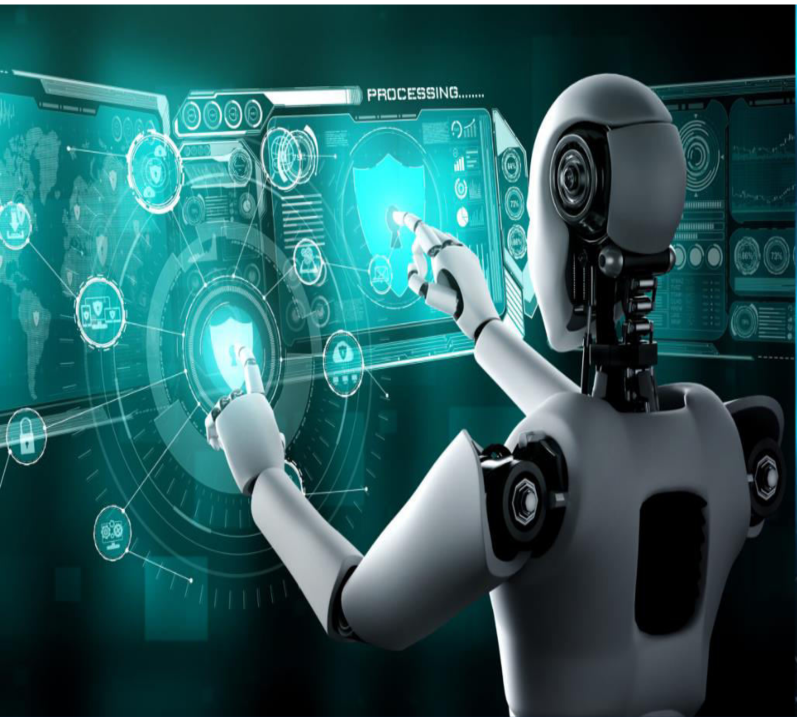
ISSN(O): 2320-9801

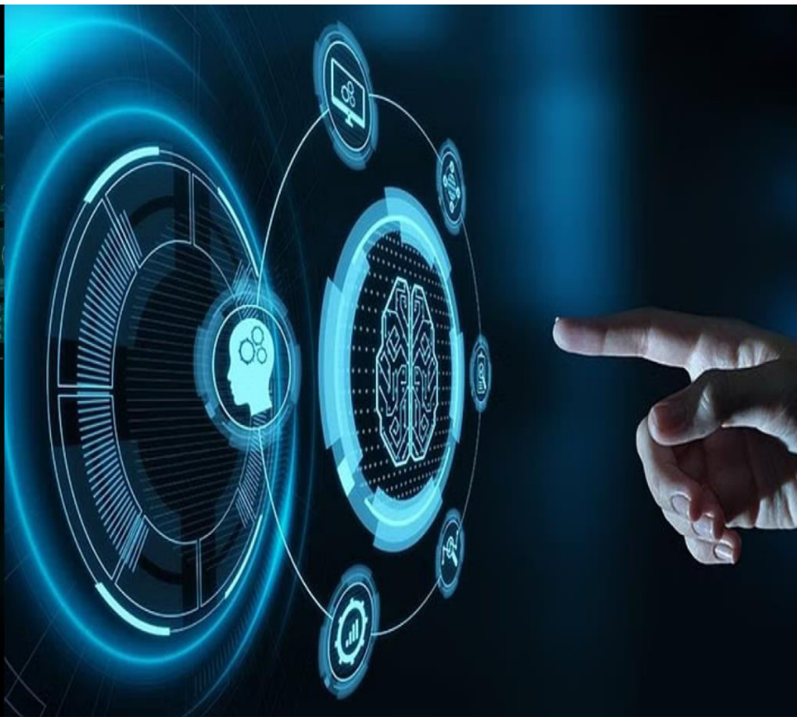
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.625

Volume 13, Issue 1, January 2025

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301117

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

HEART RATE ACCESSOR through LIGHT on FACE

Joyce A, Bavith Raj, Srivatsa H, Harish Kumar

UG Student, Department of Computer Science Engineering (Data Science), Presidency University, Bengaluru, India

UG Student, Department of Computer Science Engineering (Data Science), Presidency University, Bengaluru, India

UG Student, Department of Computer Science Engineering (Data Science), Presidency University, Bengaluru, India

Assistant Professor, Department of Computer Science Engineering (Data Science), Presidency University, Bengaluru, India

ABSTRACT: This research explores the application of Convolutional Neural Networks (CNNs) and Python libraries in determining heart rate (HR) by analyzing light variations on a subject's face captured via a webcam. By leveraging advancements in computer vision and signal processing, this work demonstrates an efficient, non-invasive method to estimate HR in real time. OpenCV, TensorFlow/Keras, NumPy, and SciPy are used for preprocessing, model development, and data analysis.

KEYWORDS: Heart rate monitoring, photoplethysmography (PPG), convolutional neural networks(CNN), non-contact HR measurement, facial analysis, machine learning, healthcare technology, fitness tracking, motion artifact mitigation, real- time Monitoring.

I. INTRODUCTION

Heart rate monitoring plays a pivotal role in various domains such as healthcare, fitness, and stress management. Traditional heart rate measurement techniques typically rely on contact-based sensors such as chest straps, smartwatches, or fingertip devices. While these methods are effective, they can be inconvenient or uncomfortable, particularly during prolonged use. Additionally, contact-based methods may not be suitable in situations where physical contact is impractical or poses a risk, such as during pandemic conditions or in neonatal care.

Recent advancements in remote photoplethysmography (rPPG) have introduced a paradigm shift in heart rate measurement. rPPG leverages the subtle color changes in the skin caused by blood flow variations to estimate heart rate. These variations are imperceptible to the naked eye but can be captured and analyzed using computer vision techniques. The face, being a highly vascularized and easily accessible region, is an ideal candidate for rPPG-based monitoring.

Heart rate variations, which can serve as indicators of physical fitness, emotional state, and cardiovascular health, are inherently tied to blood flow dynamics. These dynamics are reflected in the periodic color changes in the skin, particularly in the green channel of RGB images. Harnessing these changes using computational methods has gained traction in recent years, with researchers exploring a combination of signal processing and machine learning to refine accuracy.

In this research, we propose a CNN-based pipeline to estimate heart rate from videos of a subject's face. The approach utilizes state-of-the-art Python libraries to preprocess video data, extract signals, and train a neural network capable of identifying spatial and temporal patterns indicative of heart rate. This paper highlights the methodology, implementation, and challenges associated with this approach, as well as its potential applications in real-world scenarios.

II. METHODOLOGY

2.1 Data Collection: The foundation of any machine learning model is high-quality data. For this study, video data is collected using standard webcams, which are readily available and cost-effective. Participants are recorded under controlled conditions to minimize external factors such as variable lighting or excessive movement. Videos are

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

captured at a frame rate of 30 frames per second (FPS) with resolutions sufficient to preserve facial details. Diverse participants are included in the dataset to ensure that the model generalizes across different skin tones, facial structures, and age groups.

To simulate real-world applications, additional datasets are collected under varied lighting conditions and with participants performing mild activities such as speaking or nodding. These scenarios introduce challenges such as motion artifacts and uneven illumination, which are addressed during preprocessing and model training. Each participant's heart rate is simultaneously measured using a reference device such as a pulse oximeter, ensuring that the ground truth is available for model validation.

2.2 Preprocessing: Preprocessing is a critical step in extracting meaningful signals from video data. The primary stages of preprocessing include:

1. **Face Detection:** The first step involves detecting the subject's face in each video frame. OpenCV's Haar cascades or Dlib's facial landmark detection algorithms are employed for this purpose. These algorithms provide accurate and efficient face detection, even under suboptimal conditions. Detection is refined by focusing on stable landmarks such as the forehead, reducing noise from regions like the mouth or jawline that are prone to motion artifacts.
2. **Region of Interest (ROI):** Once the face is detected, a specific region of interest (ROI) is selected for further analysis. The forehead is typically chosen as it is less affected by facial movements and provides a consistent area for signal extraction. Advanced algorithms ensure dynamic adjustment of the ROI to account for slight positional shifts in the face.
3. **Signal Extraction:** For each frame, the average RGB values within the ROI are computed. These values form a temporal signal that reflects subtle changes in skin color due to blood flow. The green channel is often emphasized due to its higher sensitivity to hemoglobin variations.
4. **Signal Denoising:** The raw signal often contains noise from environmental factors, camera artifacts, and motion. A band-pass filter is applied to isolate frequencies corresponding to typical heart rate ranges (0.7 to 4 Hz, equivalent to 42 to 240 beats per minute). This filtering removes irrelevant components while preserving the core signal.
5. **Normalization:** The filtered signal is normalized to remove baseline drift and enhance the relative amplitude of the heart rate component. Techniques such as z-score normalization and detrending ensure that the signal is centered and scaled appropriately for analysis.

2.3 CNN Architecture: Convolutional Neural Networks (CNNs) are employed to model spatial-temporal patterns within the preprocessed data. The CNN architecture is designed with the following layers:

- **Input Layer:** Accepts sequences of ROI images as input. Each sequence represents a short video segment, allowing the model to learn from both spatial patterns in individual frames and temporal dynamics across frames.
- **Convolutional Layers:** Extract spatial features from the images. These layers use small convolutional kernels to detect patterns indicative of blood flow. Multiple layers are stacked to capture increasingly abstract features.
- **Pooling Layers:** Reduce the spatial dimensions of feature maps, retaining essential information while reducing computational complexity. Both max pooling and average pooling are tested to optimize feature retention.
- **Recurrent Layers (Optional):** Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) layers capture temporal dependencies within the signal, improving the model's ability to identify periodic patterns corresponding to heart rate. These layers enable the integration of information over time, enhancing robustness against noise.
- **Fully Connected Layers:** Combine spatial and temporal features to predict heart rate or reconstruct a clean rPPG signal for post-processing. Dropout layers are included to mitigate overfitting.

2.4 Implementation: The pipeline is implemented using Python, leveraging the following libraries:

- **OpenCV:** Used for face detection, ROI extraction, and video frame manipulation.
- **TensorFlow/Keras:** Provides a framework for building, training, and deploying the CNN model. Custom layers are defined to accommodate the unique temporal nature of rPPG signals.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- **NumPy and Pandas:** Facilitate data manipulation, storage, and analysis. These libraries are critical for handling large datasets and performing numerical computations efficiently.
- **SciPy:** Enables signal processing tasks such as filtering and Fourier Transform analysis. Functions such as butter and filtfilt are employed for precise band-pass filtering.

Additional tools, such as Matplotlib and Seaborn, are used for data visualization and result interpretation. Libraries like Scikit-learn assist in splitting datasets and evaluating performance metrics.

2.5 Training and Validation: The dataset is divided into training, validation, and test sets. To enhance model robustness, data augmentation techniques are applied, such as varying brightness levels, introducing random noise, and simulating motion blur. Synthetic data is also generated using GANs to further expand the dataset.
The model is trained using the Adam optimizer, with mean squared error (MSE) serving as the primary loss metric. Hyperparameter tuning is performed to identify the optimal learning rate, batch size, and number of epochs. Early stopping and model checkpointing are employed to prevent overfitting and ensure convergence.

III. RESULTS

The proposed pipeline demonstrates promising results across multiple performance metrics. The Mean Absolute Error (MAE) between the predicted heart rate and ground truth values, obtained using a reference pulse oximeter, is consistently below 5 beats per minute in controlled conditions. This level of accuracy is comparable to that of contact-based devices.

Visualizations of the reconstructed rPPG signals reveal a strong correlation with ground truth measurements. Frequency spectra derived from the signals exhibit distinct peaks corresponding to the subject's heart rate, validating the effectiveness of the preprocessing and CNN-based approach.

In real-time applications, the pipeline achieves a processing speed of 30 FPS, ensuring minimal latency. However, performance varies under challenging conditions such as poor lighting or significant motion artifacts. These limitations underscore the need for further refinements to enhance robustness.

IV. DISCUSSION

The results highlight the potential of CNN-based rPPG for real-time, non-contact heart rate monitoring. However, several challenges remain. Lighting variations, skin tone differences, and motion artifacts significantly affect signal quality. Advanced techniques such as domain adaptation, transfer learning, and attention mechanisms could address these issues.

Furthermore, the current implementation relies on controlled environments. Expanding the dataset to include diverse real-world scenarios, such as outdoor settings and varying activity levels, would improve model generalizability. Integration with other physiological monitoring techniques, such as respiratory rate estimation, could further enhance the utility of the system.

V. CONCLUSION

This research presents a novel approach to heart rate monitoring using CNNs and Python-based tools. The non-invasive nature of the system, combined with its real-time capabilities, makes it a promising alternative to traditional methods. The ability to remotely and reliably estimate heart rate introduces a range of potential applications, from fitness tracking to telemedicine, without the need for expensive equipment.
One of the major strengths of this approach is its adaptability. With continued advancements in camera technology and computational power, this method can be seamlessly integrated into existing consumer devices such as smartphones and laptops. Such integration would democratize access to heart rate monitoring, empowering individuals to monitor their cardiovascular health independently and frequently.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Future research should aim to address current limitations by expanding the system's robustness against real-world challenges. This includes improving performance in diverse lighting conditions, accommodating a broader range of skin tones, and mitigating the effects of motion artifacts. The incorporation of multi-modal data, such as audio or thermal imaging, could further enhance the system's accuracy and reliability.

In conclusion, the combination of CNNs, advanced preprocessing techniques, and Python-based.

REFERENCES

[1] Allen, J. (2007). Photoplethysmography and its application in clinical Physiological measurement. Physiological Measurement, 28(3), R1–R39. [1] Kumar, A., Kaur, H., & Singh, S. (2021). Heart rate monitoring using Facial video: A systematic review. Sensors, 21(8), 2766.
[2] Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering, 32(3), 230–236.
[3] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for Image recognition. Proceedings of the IEEE Conference on ComputerVision and Pattern Recognition (CVPR), 770–778.
[4] Kwon, S., Kim, H., & Park, K. S. (2012). Validation of heart rateExtraction using video imaging on a built-in camera system of a smartphone. Engineering in Medicine and Biology Society (EMBC), 2174–2177.
[5] 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.
[6] Liu, J., Cao, T., & Qian, W. (2021). Robust heart rate estimation from Facial videos using deep learning. Biomedical Signal Processing and Control, 66, 102430.
[7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
[8] Li, X., Chen, J., Zhao, G., & Pietikäinen, M. (2014). Remote heart rate Measurement from face videos under realistic situations. Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 4264–4271.
[9] Chen, W., & McDuff, D. J. (2018). DeepPhys: Video-based Physiological measurement using convolutional attention networks. Proceedings of the European Conference on Computer Vision (ECCV), 349–365.
[10] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic Optimization. arXiv preprint arXiv:1412.6980.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details