International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82730

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Performance Comparison Between NSQI-Based Filtering and ML-Based Methods for Improving HR Prediction Using rPPG

Araadhye Bisht[1], Atharva Jakhetiya[2], Arpan Saraswat[3]
Dept. of Electronics and Communication Engineering, Delhi Technological University

Abstract: Remote photoplethysmography (rPPG) is a non- contact technique for estimating heart rate (HR) by analyzing subtle, blood-volume-induced color changes in skin captured by standard video cameras. While promising for applications in telemedicine and continuous health monitoring, practical rPPG implementation is severely hampered by susceptibility to motion artifacts, illumination variations, and skin-tone differences. In this study, we develop a comprehensive rPPG pipeline using the UBFC-rPPG dataset, employing standard face detection and sig- nal extraction algorithms (Green, CHROM, POS, LGI) to retrieve baseline pulse waveforms. To address the critical challenge of signal noise, we implement and rigorously compare two distinct performance optimization strategies: (1) a composite Normalized Signal Quality Index (NSQI)-based filtering framework designed to autonomously identify and discard corrupted signal segments, and (2) a supervised Machine Learning (ML) regression approach utilizing a 1D Convolutional Neural Network (1D-CNN) to map physiological signals directly to HR estimates. Our experimental results demonstrate that the NSQI filtering approach effectively reduces Mean Squared Error (MSE) and Mean Absolute Error (MAE) by selectively removing low-quality data. Conversely, the 1D-CNN approach exhibited significant generalization issues, failing to capture morphological signal features and instead acting as a mean regressor due to inherent dataset bias and normalization challenges.
Index Terms: Remote photoplethysmography, rPPG, Signal Quality Index, NSQI, Heart Rate Estimation, Deep Learning, 1D-CNN, Motion Artifacts.

I. INTRODUCTION

Heart rate (HR) Heart rate is a fundamental physiological parameter that has to be monitored during the assessment of the cardiovascular condition, stress, and general physical condition. Traditional HR monitoring methods, including electrocardiography and pulse oximetry, are performed using contact-based sensors. While they are highly accurate, both techniques require attachment to the body, which may provoke discomfort in case of long-term monitoring and could be impractical in sensitive situations such as neonatal intensive care or burn unit monitoring and large-scale public health screening. Emerging as a powerful, contactless alternative is remote photoplethysmography. The basic principle of rPPG depends on the fact that blood absorbs light differently compared to surrounding tissues. As the heart pumps blood, the volume of blood in the facial blood vessels changes periodically and creates small fluctuations of the intensity of the reflected ambient light. Thus, remote reconstruction of the cardiovascular pulse wave is made possible by capturing video of exposed skin, typically the face, and analyzing temporal variations of pixel intensity across color channels. Despite this high potential, the rPPG signal is very weak compared to contact PPG and contaminated by noise; hence, rigid and nonrigid head motion and fluctuations in ambient illumination, camera parameters, and facial expressions are the major sources of error. Such factors can easily mask the pulsatile signal, making accurate HR estimation problematic. Hence, developing robust methods that either clean noisy signals or identify and discard unreliable segments is vitally important in clinical translation.

Currently, improvements in rPPG robustness may be summarized under the use of either advanced signal processing techniques that employ color space transformations or purely data-driven Deep Learning methods. This paper aims to systematically compare the performance of these two paradigms in a unified pipeline.

Main contributions of this work are:

1) We implemented a modular rPPG pipeline integrating four various state-of-the-art extraction algorithms, namely Green, CHROM, POS, and LGI.
2) Development and evaluation of composite Normalised Signal Quality Index (NSQI) filtering system that integrates statistical, spectral and chaotic features defining signal usability.

3) We presented the design and analysis of a 1D-CNN regression model for direct HR estimation, pointing out some critical failure modes related to data bias in standard rPPG datasets.

II. RELATED WORK

The field of rPPG has evolved from simple single-channel analysis to complex mathematical modeling and deep neural networks:

A. Model-Based Signal Processing

Early rPPG research focused on the Green channel (G) method as described initially by Verkruysse et al. [8], based on the fact that hemoglobin has peak absorption in the green part of the spectrum. While simple, this kind of method is extremely sensitive to noise. Further methods used Blind Source Separation (BSS) techniques like Independent Component Analysis (ICA) to demix the pulse signal from noise sources; these methods, however, face challenging conditions in the case of non-stationary motion.

Notably, significant advances have been made possible by developing color-space transformations that are designed to suppress particular noise types. de Haan and Jeanne presented CHROM [3], which combines RGB channels linearly into orthogonal chrominance signals, wherein the effect of specular reflection is well dampened. Wang et al. suggested POS (Plane-Orthogonal-to-Skin) [4], which projects the RGB signal onto a plane orthogonal to a biologically defined skin-tone vector, resulting in an improved robustness compared to CHROM. In this respect, Pilz et al. recently proposed LGI (Local Group Invariance), where the modeling of motion artifacts as continuous transformation groups provides an invariant representation of the pulse.

B. Deep Learning Approaches

Motivated by the success of computer vision, deep learning methods have been actively explored in rPPG. Representative end-to-end architectures, such as DeepPhys [5], exploit convolutional attention networks with explicit motionand appearance-branch streams to extract physiological signals. PhysNet [6] applies three-dimensional (spatio-temporal) convolutions directly to raw volumes of video sequences to learn features. Although these solutions achieve state-of-the-art results on specific benchmarking datasets, they often consume high computational resources and rely on large-scale, diverse databases to generalize well to unseen light condition and skin tone variability.

III. METHODOLOGY

Our proposed system consists of a baseline signal extraction pipeline followed by two parallel optimization branches. The overall system flow is illustrated in Figure 1.

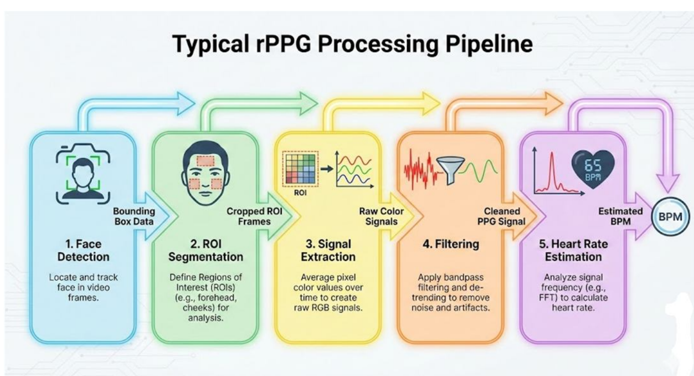


Fig. 1. Block diagram of the proposed rPPG system, showing face detection, ROI extraction, signal processing, and the two parallel optimization paths (NSQI Filtering and 1D-CNN).

A. Dataset and Preprocessing

We utilized the **UBFC-rPPG dataset**, comprising videos of subjects in a seated position facing a camera, synchronized with ground-truth PPG data from a finger pulse oximeter.

The preprocessing steps include:

- Face Detection ROI Selection: We used OpenCV's pre-trained Haar cascade classifier for face detection. To balance computational efficiency with signal strength, a heuristic Region of Interest (ROI) was defined as the central 60% of the detected face bounding box.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

- Spatial Averaging: The pixel values within the ROI were spatially averaged for each frame to yield raw RGB time-series signals.
- Temporal Filtering: A 4th-order Butterworth band- pass filter with cutoff frequencies of 0.7 Hz (42 BPM) and 3.5 Hz (210 BPM) was applied to remove low- frequency respiratory trends and high-frequency sensor noise, isolating the physiological HR range.

B. Baseline Signal Extraction

We implemented four established algorithms to extract the pulse waveform from the preprocessed RGB signals:

- Green (G): Utilizes the raw, filtered green channel trace.
- CHROM: Projects RGB signals onto a chrominance plane to remove specular components [3].
- POS: Projects signals onto a plane orthogonal to the skin vector to isolate the pulse [4].
- LGI: Uses geometric invariance principles to miti- gate motion artifacts.

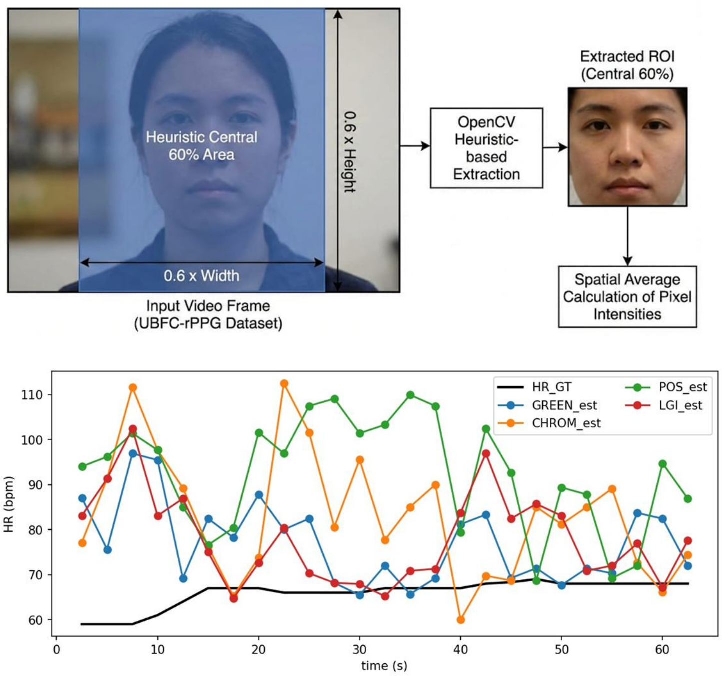


Fig. 2. (Top) Frame showing detected face and selected ROI. (Bottom) Comparison of extracted rPPG waveforms using different algorithms against ground truth.

C. Optimization Strategy 1: NSQI Filtering

The quality of extracted rPPG signals varies drastically over time due to subject movement. We developed a composite Normalized Signal Quality Index (NSQI) to assess the reliability of signal segments in 5-second sliding windows (2.5s overlap).

The composite index is derived from five diverse features:

- Signal-to-Noise Ratio (SNR): Calculated in the frequency domain as the ratio of power within a ±0.2 Hz band around the fundamental peak frequency versus the power in the rest of the physiological band.
- Skewness (SSQI) & Kurtosis (KSQI): Statistical moments used to detect signal asymmetry and impulsiveness typical of motion artifacts. Clean PPG signals generally exhibit positive skewness and moderate kurtosis.
- Entropy (ESQI): Shannon entropy is used to quantify signal complexity; noisy signals typically exhibit higher disorder.
- Zero-Crossing Rate (ZSQI): Measures the frequency of sign changes in the detrended signal, helping identify high-frequency noise.
- Attractor-Shape Metric: A novel metric analyzing the 3D phase-space reconstruction of the signal, measuring deviations from the expected quasi-periodic PPG attractor shape.

Each feature was min-max normalized. A composite score was calculated as a weighted average. Based on empirical analysis of the dataset, windows falling in the bottom 25th percentile of composite scores were categorized as ”Unfit” and excluded from final HR calculation.

D. Optimization Strategy 2: Deep Learning (1D-CNN)

As an alternative to traditional spectral peak detection, we designed a 1D Convolutional Neural Network (1D-CNN) to perform regression directly from the time-domain signal to a BPM value.

- Input: 5-second normalized windows of the signal extracted via the POS algorithm (chosen for its general robustness).
- Architecture: The network comprises three 1D convolutional blocks.
- Pooling Output: Each convolutional block is followed by Max Pooling to reduce dimensionality. The final feature map undergoes Adaptive Average Pooling before being fed into fully connected dense layers that output a single scalar HR value.
- Training: The model was trained using Mean Squared Error (MSE) loss with an Adam optimizer.

IV. RESULTS AND DISCUSSION

A. Baseline Performance

Initial evaluation of the four baseline algorithms confirmed that advanced color-space methods outperform raw channel analysis. POS and CHROM generally provided the most stable waveforms during stationary periods, closely matching the ground truth period (see Figure 2). However, all algorithms suffered significant degradation during periods of subject movement, as the fixed ROI included fluctuating background pixels, introducing substantial noise.

B. Evaluation of NSQI Filtering

Applying the NSQI-based filtering pipeline resulted in measurable improvements in overall system accuracy. By filtering out the lowest quality 25% of data windows, both MAE and MSE decreased compared to the unfiltered baseline.

Detailed correlation analysis between individual SQI metrics and the resulting HR estimation error revealed that SNR and the Attractor-Shape metric had the highest predictive power for signal quality on this dataset. Windows that scored high on these specific metrics consistently yielded accurate HR estimates.

However, it is crucial to note that NSQI is a selection mechanism, not a correction mechanism. While it successfully identifies and removes unreliable data, it does not improve the inherent quality of the signals extracted by the baseline algorithms. A persistent baseline error remained even in the ”Fit” data, primarily attributed to the limitations of the heuristic ROI tracking during significant motion.

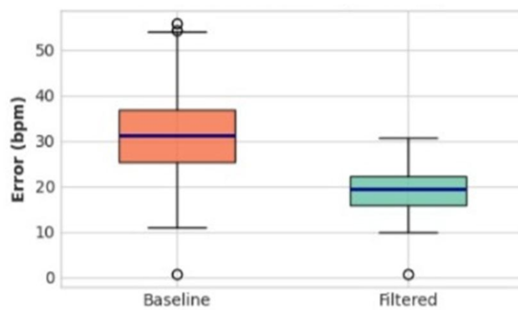


Fig. 3. Box plot comparing error in bpm before and after applying NSQI filtering, showing error reduction.

C. Evaluation of 1D-CNN Regression

The performance of the 1D-CNN regression model high- lighted critical challenges in applying supervised learning to rPPG without extensive data engineering.

While training loss converged, evaluation on the test set revealed a critical failure mode: the model predicted a nearly constant heart rate value (approximately 75 BPM) regardless of the input signal’s actual periodicity. This indicates that the network failed to learn the morphological features associated with the pulse wave and instead functioned as a mean regressor, minimizing global loss by predicting the average HR of the training dataset.We attribute this failure to two primary factors:

- Data Bias: The UBFC-rPPG dataset is heavily skewed toward resting heart rates (60-90 BPM), with very few examples of high or low rates. The model overfitted to this dominant distribution.
- Normalization Issues: The input POS signals exhibited significant amplitude variations based on subject skin tone and lighting. Standard min-max normalization over the entire dataset was insufficient; instance-based normalization (per window) is likely required for the network to learn scale-invariant features.

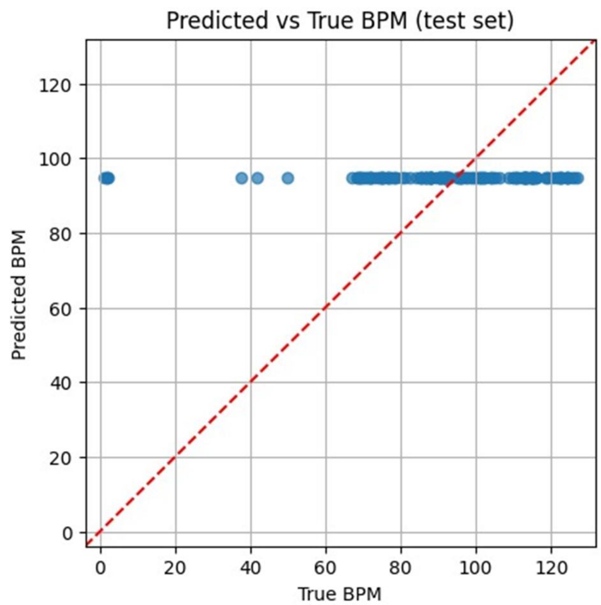


Fig. 4. Scatter plot of True BPM (x-axis) vs. Predicted BPM (y-axis) for the 1D-CNN model. The horizontal grouping of predictions illustrates the model's failure to generalize, predicting the dataset mean instead.

V. CONCLUSION AND FUTURE WORK

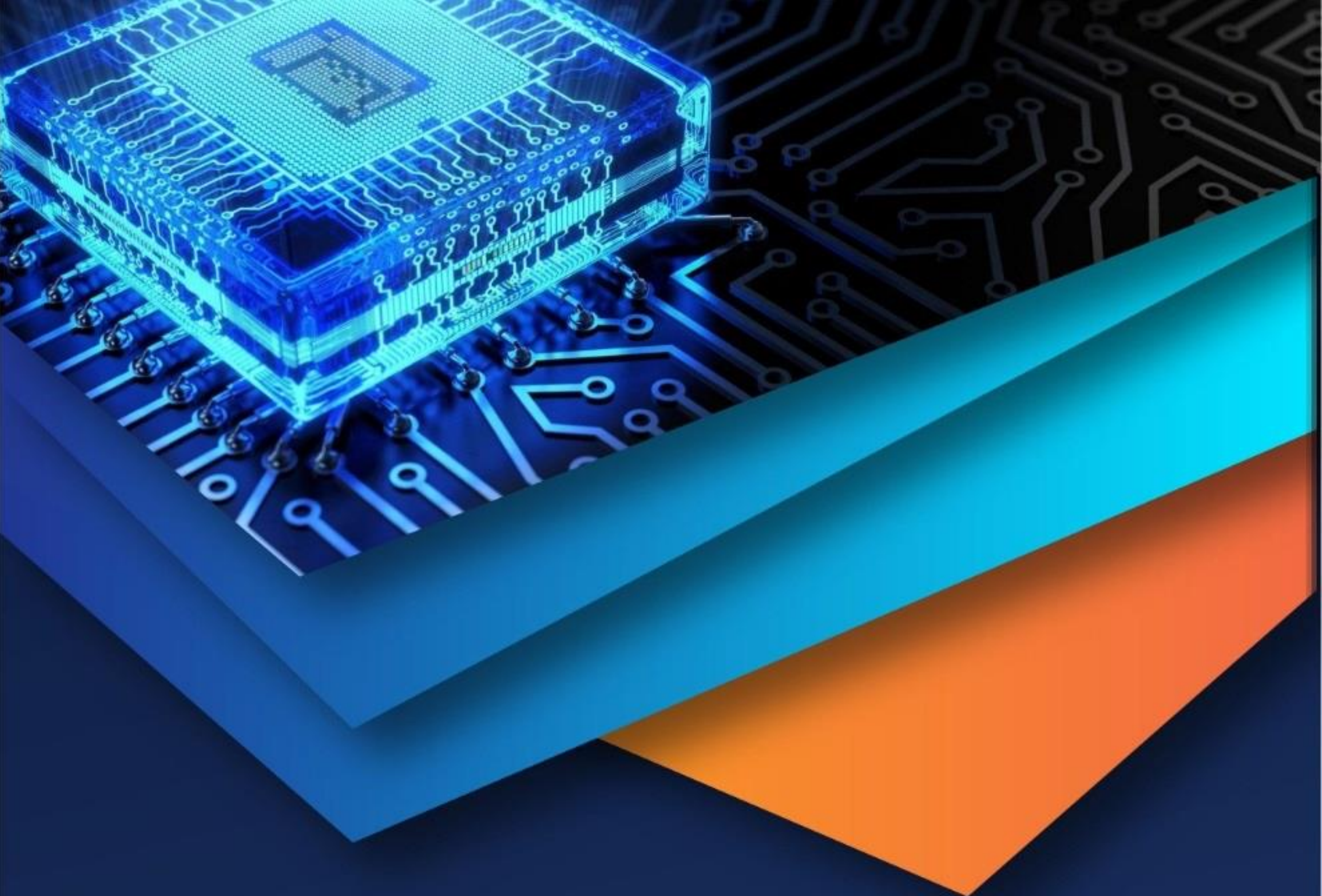
This paper presented a comparative study of signal quality filtering versus machine learning regression for improving rPPG accuracy. The findings demonstrate that NSQI-based filtering is a viable, interpretable strategy for enhancing reliability by discarding noisy data, though its ultimate performance is capped by the quality of the underlying signal extraction algorithms. Conversely, the 1D-CNN approach, while theoretically powerful, failed to generalize due to severe dataset bias and normalization challenges, highlighting the difficulty of training regression models on limited physiological data.

Future work will focus on addressing the identified limitations:

1) Dynamic ROI Tracking: Replacing the fixed heuristic ROI with facial landmark tracking to maintain skin coverage during motion.
2) ML Data Augmentation: Implementing temporal re- sampling techniques to artificially balance the dataset across a wider range of heart rates (e.g., 40-140 BPM) to prevent mean-regression overfitting.
3) Instance Normalization: Applying Z-score normalization on a per-window basis to help the CNN learn features independent of signal amplitude.

REFERENCES

[1] S. Nakamura et al., "A Review of Photoplethysmography for Remote Physiological Monitoring," IEEE Trans. Biomed. Eng., vol. 68, no. 9, pp. 2896-2908, 2020.
[2] C.-H. Cheng, K.-L. Wong, et al., "Deep Learning Methods for Remote Heart Rate Measurement: A Review," Sensors, vol. 21, no. 18, 6296, 2021.
[3] G. de Haan and V. Jeanne, "Robust Pulse Rate from Chrominance-based rPPG," IEEE Trans. Biomed. Eng., vol. 60, no. 10, pp. 2878-2886, 2013.
[4] J. Wang et al., "Exploiting Spatial Redundancy of Facial Skin for Remote Photoplethysmography," IEEE Trans. Biomed. Eng., vol. 67, no. 5, pp. 1735-1744, 2020.
[5] W. Chen and D. McDuff, "DeepPhys: Video-Based Physiological Mea- surement Using Convolutional Attention Networks," in ECCV, 2018.
[6] Z. Yu, X. Li, and G. Zhao, "Remote Photoplethysmograph Signal Measurement from Facial Videos Using Spatio-Temporal Networks (PhysNet)," in BMVC, 2019.
[7] M.Z. Poh, N.C. Swenson, and R.W. Picard, "A Wearable Sensor for Unobtrusive, Long-term Assessment of Electrodermal Activity," IEEE Trans. Biomed. Eng., vol. 57, no. 5, pp. 1249-1258, 2010.
[8] W. Verkruysse, L.O. Svaasand, and J.S. Nelson, "Remote plethysmo- graphic imaging using ambient light," Opt. Express, vol. 16, no. 26, pp. 21434-21445, 2008.
[9] S. Bobbia et al., "Unsupervised Skin Tissue Segmentation for Remote Photoplethysmography," Pattern Recognit. Lett., vol. 95, pp. 71-81, 2017.
[10] M. Elgendi, I. Martinelli, and C. Menon, "Optimal signal quality index for remote photoplethysmogram sensing," npj Biosensing, 2024.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY