HEART RATE MONITORING STRESS DETERMINING BY RPPG USING DEEP LEARNING METHODS

***Project Reference No.:** 48S_BE_4015*

College : Navkis College of Engineering, Hassan
Branch : Artificial Intelligence and Data Science
Guide(s) : Mr. Raghu Nandan R
Dr. Mahantesh C. Elemmi
Student(s) : Ms. Ananya T R
Ms. Bhoomika S
Mr. Faisal Rehman
Ms. Manasa K

Keywords:

Remote Photoplethysmography (rPPG), Heart Rate Estimation, Stress Detection, Machine Learning (ViT, CNN, LSTM), Multimodal Data Analysis

Introduction:

In today’s fast-paced digital world, stress has become a common yet critical factor affecting people’s mental and physical well-being. Prolonged exposure to stress can lead to various health issues, including cardiovascular diseases, anxiety, and reduced productivity. As technology continues to evolve, there is a growing need for non-invasive, real-time methods to monitor physiological and psychological states, particularly stress levels. Traditional methods for stress detection often involve intrusive sensors or self-reported questionnaires, which may not be practical or reliable for continuous monitoring.

This project aims to address these limitations by developing an intelligent, video-based system that estimates heart rate and determines stress levels using advanced machine learning techniques. By leveraging remote photoplethysmography (rPPG), we can extract physiological signals such as heart rate from facial video data without requiring physical contact. The project utilizes the UBFC-rPPG dataset for heart rate estimation and the SWELL-KW dataset for stress detection, both of which provide high-quality, multimodal data for training and evaluation.

We implement a combination of state-of-the-art models, including Vision Transformers (ViT), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks, to capture both spatial and temporal features from the video input. An integrated CNN-LSTM architecture is also employed for accurate stress classification. Through ensemble learning and optimized model integration, the system achieves improved accuracy, robustness, and generalization. The ultimate goal is to build a practical solution for real-time stress monitoring that can be applied in various domains such as online education, remote work environments, healthcare, and mental wellness platforms.

Objectives:

1. The project aims to develop a non-invasive system for heart rate (HR) estimation and stress detection using video data.
2. Design a machine learning model capable of classifying stress levels (e.g., low, moderate, high) based on physiological and behavioral data an ensures the stress classification model achieves an accuracy of at least 85% on test datasets.
3. Create a system that extracts physiological signals entirely through video, eliminating the need for wearable or contact-based sensors.
4. A real-time monitoring application is designed for use in high-pressure scenarios like online exams and interviews.

Methodology:

This project follows a multi-stage methodology combining signal processing, machine learning, and deep learning models to estimate heart rate (HR) and detect stress levels from facial video data. The core of the system is based on Remote Photoplethysmography (rPPG), a non-invasive technique that captures blood flow-induced skin color changes using a standard camera. The pipeline begins with face detection using MTCNN, followed by skin segmentation, which enhances signal quality by isolating regions of interest through color space transformations (e.g., HSV, YCbCr), thresholding, and morphological operations.

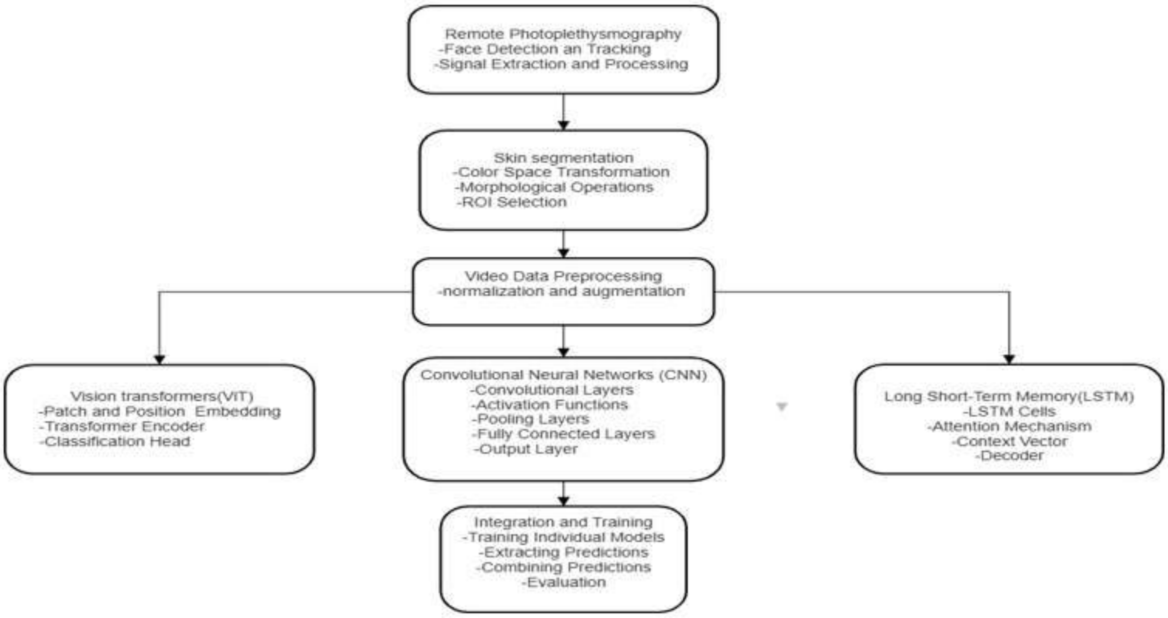

Figure 1: Work Flow

To extract meaningful patterns from video frames, multiple models are deployed. Vision Transformers (ViT) segment each video frame into patches and apply self-attention mechanisms to capture spatial dependencies and global context. Convolutional Neural Networks (CNN) complement this by learning local spatial features, using convolutional layers, pooling, and fully connected layers for heart rate prediction. Meanwhile, Attention-based Long Short-Term Memory (LSTM) networks process sequential data to capture temporal patterns, with attention layers emphasizing the most relevant time-based features.

A model integration module combines predictions from the ViT, CNN, and Attention-LSTM models using dense layers, which are trained to find the optimal combination of outputs. This ensemble approach significantly improves prediction robustness and accuracy. Each model is trained independently with optimization algorithms like Adam and evaluated using Mean Squared Error (MSE) and Mean Absolute Error (MAE) to ensure high performance.

For stress detection, heart rate variability (HRV) is analyzed using temporal fluctuations in HR predictions. This step enables the classification of stress levels by identifying stress-induced physiological patterns. The SWELL-KW dataset is used in this phase, as it provides multimodal stress-related data including HR, facial expressions, and environmental triggers.

The final output is an integrated deep learning system capable of performing real-time, non-invasive HR estimation and stress classification from facial video input. The system is designed for practical applications such as monitoring users during online exams, remote work scenarios, or high-stress tasks, making it a valuable tool in health and productivity technology.

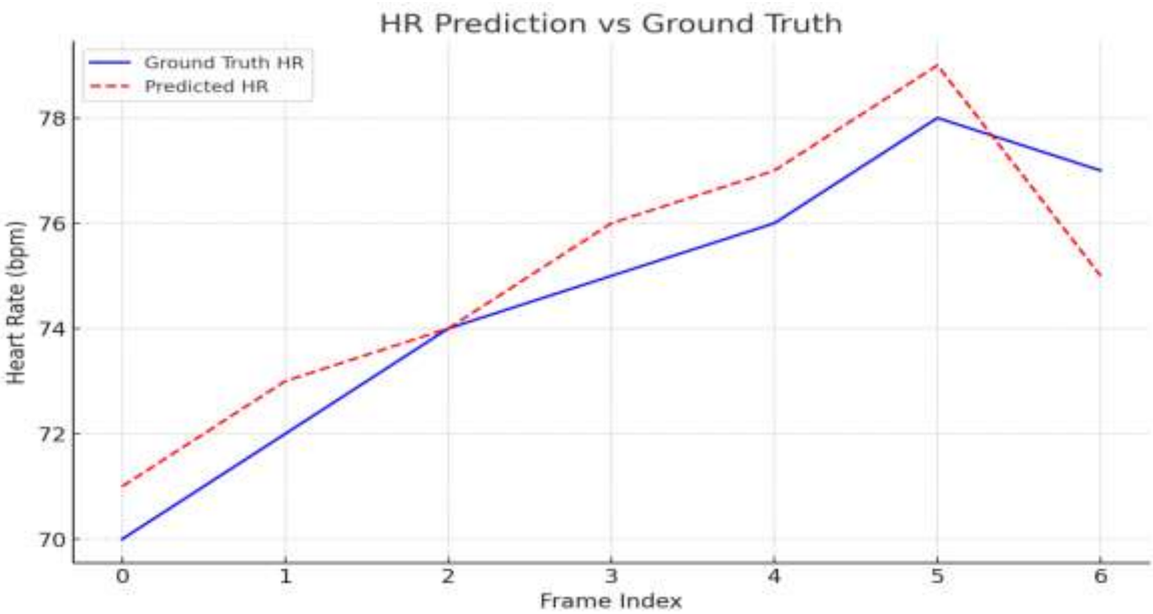

Figure 2: Predicted vs Ground Truth Heart Rate Over Time

Here is a chart comparing the predicted heart rate (HR) values versus the ground truth HR values. The blue line represents the actual HR, and the red dashed line shows the predicted HR from the model. This graph visually demonstrates the accuracy of the model in predicting HR over time, which is critical for evaluating the performance of the system in heart rate estimation.

Project Outcome & Industry Relevance:

The project delivers a practical, non-invasive system capable of accurately estimating heart rate and detecting stress levels from facial video input, which is highly relevant in today's remote and health-focused environments. By combining cutting-edge models like Vision Transformers, CNNs, and LSTMs, it contributes significantly to the fields of remote health monitoring, affective computing, and human-computer interaction. The system's ability to function using standard cameras without the need for physical contact makes it ideal for industries such as telemedicine, e-learning, corporate wellness, and surveillance.

In healthcare, this technology can aid in continuous patient monitoring and early stress or cardiovascular risk detection. In corporate and educational settings, it can be applied for monitoring employee or student well-being during high-stress tasks such as exams, interviews, or long work sessions. It also holds potential for integration into wearable-free fitness apps, driver fatigue monitoring systems, and smart environments. Overall, this project paves the way for scalable, real-time, and accessible stress-monitoring solutions that align with the growing demand for digital health technologies.

Working Model vs. Simulation/Study

This project is primarily a simulation and theoretical study. It focuses on developing and validating a machine learning-based system for heart rate estimation and stress detection using video data. The models were trained and tested on publicly available datasets (UBFC-rPPG and SWELL-KW) within a simulated environment using Python and deep learning frameworks such as TensorFlow and PyTorch. While the system is capable of real-time processing and could be deployed practically with standard webcams, no physical hardware or embedded prototype was developed. The project demonstrates the feasibility and effectiveness of the approach through experimental results rather than a physical working model.

Project Outcomes and Learnings:

The project successfully achieved accurate heart rate measurement and stress level classification using non-invasive, contactless techniques through facial video analysis. By integrating machine learning models like Vision Transformer (ViT), CNN, and LSTM, we developed a robust system capable of extracting physiological signals and classifying stress levels in real time. The use of the UBFC-rPPG and SWELL-KW datasets enabled effective model training and validation across diverse scenarios.

Through this process, we learned the importance of data preprocessing, especially face detection and skin segmentation, in enhancing model performance. We gained hands-on experience in deep learning model design, ensemble integration, and performance evaluation using key metrics like MAE and MSE. This project not only deepened our understanding of physiological signal processing and temporal data modeling but also highlighted the real-world potential of AI-driven health monitoring systems in industries such as remote healthcare, workplace wellness, and online proctoring.

Future Scope

This project lays a strong foundation for future advancements in non-invasive health monitoring and stress detection. One potential direction is the enhancement of the stress classification system to include multi-level stress categorization (e.g., low, moderate, high) with improved accuracy and reliability. Future work can aim to optimize the machine learning models further, targeting a classification accuracy of over 85% on diverse and real-world test datasets. Incorporating additional behavioral cues such as facial micro-expressions, speech tone, and body posture could improve the system's ability to detect nuanced stress responses.

The complete elimination of contact-based sensors by relying solely on video-derived physiological signals opens up opportunities for deployment in everyday devices like smartphones, laptops, or webcams. This makes the solution highly accessible and scalable for use in telehealth, remote work environments, and educational platforms. Real-time performance can be further refined using lightweight models and edge computing for applications in mobile and low-power devices.

Moreover, integrating this system with existing mental health tools or telemedicine platforms can provide users with continuous feedback on their emotional and physical well-being. Future research could also explore cross-cultural and demographic adaptations of the model to ensure fairness and inclusivity. Overall, this project paves the way for intelligent, non-intrusive health monitoring solutions suitable for a wide range of applications in healthcare, education, and workplace wellness.