Personalized Worker Physiological Load Assessment Using Multimodal Wearable PPG Analysis and Activity Recognition

Olena Litovska , Myroslav Mishchuk * , Olena Pavliuk

Posted Date: 10 March 2026

doi: 10.20944/preprints202603.0657.v1

Keywords: photoplethysmography; human activity recognition; heart rate estimation; wearable sensors; machine learning; accelerometer signals; motion artifact removal; KID-PPG; physiological monitoring

 doi:10.20944/preprints202603.0657.v1

Article

Personalized Worker Physiological Load Assessment Using Multimodal Wearable PPG Analysis and Activity Recognition

Olena Litovska [1], Myroslav Mishchuk [2,3,*] and Olena Pavliuk [2,3]

[1] Department of Artificial Intelligence Systems, Lviv Polytechnic National University, 79000 Lviv, Ukraine
[2] Department of Automated Control Systems, Lviv Polytechnic National University, 79000 Lviv, Ukraine
[3] Department of Distributed Systems and Informatic Devices, Silesian University of Technology, 44-100 Gliwice, Poland
* Correspondence: Myroslav.Mishchuk@polsl.pl; Tel.: +48-511-646-535

Abstract

Wearable sensors enable continuous monitoring of human activity and physiological state, with applications in workplace health monitoring, occupational safety, sports performance analysis, and rehabilitation. However, effective use of these devices requires specialized data-processing algorithms and machine-learning (ML) methods. This work proposes a methodology for assessing workers' physiological load using ML, combining accelerometer and photoplethysmography (PPG) signal processing. An ensemble Random Forest algorithm is used for activity recognition, while the KID-PPG deep learning model is applied for heart rate (HR) estimation. A personalized physiological load assessment framework normalizes effort indices against demographic-group-specific distributions defined by sex, age, and activity intensity. The methodology was validated on the PPG-DaLiA dataset comprising 15 participants with diverse demographic profiles across eight daily activities. Experiments demonstrated high accuracy in activity recognition (macro F1-score of 90.73%) and robust HR estimation even in the presence of motion artifacts (MAE below 10 bpm). The personalized assessment revealed that participant age substantially influences physiological-load patterns, confirming that demographic-aware normalization is essential for accurate workload interpretation. The main factors influencing system performance have been identified, and directions for improving the models across diverse user groups and limited-signal-quality conditions are discussed.

Keywords: photoplethysmography; human activity recognition; heart rate estimation; wearable sensors; machine learning; accelerometer signals; motion artifact removal; KID-PPG; physiological monitoring

1. Introduction

In the rapidly evolving world of modern technologies, Human Activity Recognition (HAR) is becoming increasingly widespread. It is an important and complex technology that can serve as the foundation for numerous new applications—in smart homes, industrial environments, office worker assistance systems, surveillance systems, and interactive games, as well as interfaces for home healthcare and medical systems. HAR is a multidisciplinary research field closely related to machine learning (ML), artificial intelligence, machine perception, and human-computer interaction, as well as psychology and sociology.

With the advent of miniaturized sensors that can be attached to the body, it has become possible to collect and store data regarding various aspects of human activity during daily life. This technology has great potential for use in automated systems that monitor daily activities, which continuously record body movements over extended periods. Such monitoring systems rely on recognition algorithms that effectively interpret sensor readings and identify different types of activities.

 doi:10.20944/preprints202603.0657.v1

There are three well-known approaches to equipping a recognition system with sensors: (1) video-based systems, (2) systems using sensors embedded in the environment, and (3) systems based on body-worn sensors. Although video systems provide detailed observation, they are expensive, environment-dependent, and raise privacy concerns. Systems with environmental sensors also depend on external factors. In contrast, body-worn sensors are independent of environmental conditions, relatively inexpensive, and ensure privacy, making them the most convenient and effective option for long-term physical activity monitoring. They are well-suited for collecting data on daily physical activity patterns over extended periods. Such systems can be integrated into clothing, attached to the body, embedded in jewelry, or exist as other wearable devices that a person can carry.

HAR systems based on wearable sensor data can provide high-precision, real-time monitoring of motor behavior. The use of a triaxial accelerometer enables reliable differentiation among various activities by analyzing dynamic changes in acceleration across different planes. The photoplethysmogram (PPG) signal contains information about the pulse wave, allowing indirect estimation of heart rate (HR) under dynamic conditions [1]. However, the signal is susceptible to motion artifacts (MA), which significantly complicate the reliable extraction of the pulse component during physical activity.

The main objective of this work is to propose a methodology for automated analysis of worker physiological load based on accelerometer and PPG data, using modern ML and signal processing techniques. This includes investigating the most informative features of accelerometers and PPG signals for HAR, applying modern deep learning models for HR estimation, and integrating these results to analyze workers' physiological states in relation to their individual characteristics. The methodology should enable noninvasive, continuous monitoring and an objective assessment of workload by analyzing current physiological parameters against typical group-specific values.

In this study, automated classification of worker activities is performed using features derived from the temporal and frequency characteristics of accelerometer and PPG signals. The main objective at this stage is to apply established signal preprocessing and feature extraction techniques and to select an optimal classification algorithm that is robust to noise and MA while achieving high accuracy in recognizing different types of motor activity. The task of HR estimation involves applying methods for PPG denoising and MA suppression, particularly via adaptive filtering based on synchronized accelerometer data, and applying modern peak-detection methods to determine the inter-pulse interval (IPI). Special attention is paid to implementing deep knowledge-driven models that account for physiological regularities and measurement uncertainty [2]. An important stage of the research is the validation of the obtained HR estimates by comparing them with reference ECG data. A comprehensive analysis of physiological state is performed by interpreting the obtained HR data in the context of the activities performed and the worker's individual characteristics, such as age, sex, and fitness level. The main goal at this stage is to develop an approach to workload interpretation that accounts for both normative and personalized indicators of cardiac activity and performed activity. The methodology should ensure segmentation of the physiological workload range into activity-intensity levels and the identification of overexertion or underload states.

This work makes several contributions to the field of wearable sensor-based physiological monitoring:

1. A complete pipeline is developed and validated, integrating discrete wavelet transform (DWT), amplitude demodulation, and a knowledge-informed deep learning (KID-PPG) model for robust HR estimation from PPG signals under MA conditions. The framework achieves a mean absolute error below 10 bpm, even during dynamic activities, and is validated against reference ECG data.
2. The optimal subset of 164 features was identified through a systematic analysis of 341 time-, frequency-, and wavelet-domain features extracted from accelerometer and PPG signals, which enables activity classification with a 90% F1-score using Random Forest (RF). We demonstrate that while accelerometer features dominate for activity recognition, specific PPG spectral features provide complementary discriminative information.

 doi:10.20944/preprints202603.0657.v1

3. A statistical approach to workload evaluation was introduced that accounts for individual characteristics (age, sex, fitness level) and activity intensity by comparing effort indices against demographic-specific distributions. This enables objective identification of normal load, fatigue, and overload states without requiring individualized baseline calibration.
4. The proposed modular system combines activity classification, adaptive PPG filtering, probabilistic HR estimation, and physiological state assessment in a unified framework. The architecture supports continuous, noninvasive monitoring using only wrist-worn accelerometers and PPG sensors.
5. All system components were validated on the PPG-DaLiA dataset containing data of 15 participants with varying age (21–55 years), sex, body mass index (BMI) (20.2–27.6), and fitness levels (1–6), across eight daily activities. The leave-subject-out evaluation demonstrates strong generalization capability across demographic groups.

The proposed approach can be applied to workplace health monitoring systems, occupational safety platforms, sports performance analysis, and rehabilitation programs requiring continuous physiological state assessment from wearable sensors. It will enable timely detection of potential signs of overload or insufficient activity—essential for ensuring safety, productivity, and fatigue prevention of personnel.

2. Related Works

HAR relies on effective transformation of raw sensor data into activity labels [3]. The ML approaches here are categorized into traditional and deep learning-based [4]. The traditional approach comprises feature engineering (e.g., statistical and spectral characteristics: mean, standard deviation (SD), correlation) and classification using algorithms such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), Decision Trees, RF, and XGBoost [5–8]. These methods perform well for basic activities but require expert feature engineering and are less effective for complex cases [9]. Deep learning approaches automatically learn features from minimally processed data [10]. Common architectures include CNNs—effective for local patterns [2], recurrent networks (LSTM, GRU) [11], hybrid models (CNN–LSTM, CNN–GRU) [4,12,13], autoencoders [14], and transformers [15]. Deep models achieve higher accuracy, especially for complex or hierarchical activities, but require substantial computational resources and large datasets. Modern ML methods are also applied in HAR, including ensemble learning (combining RF, XGBoost, and stacking for complex activities) [16], transfer learning (fine-tuning CNN or DenseNet121 models on specialized datasets such as KU-HAR) [17,18], and federated learning (training on decentralized data without data transfer, preserving privacy and system architecture constraints) [19]. The choice of an optimal solution depends on the type of sensors, activities, data volume and quality, and requirements for accuracy and computational resources [3].

Triaxial accelerometer data are a key source of information for wearable-sensor-based HAR systems. The accelerometer measures linear acceleration along the X, Y, and Z axes, enabling determination of motion, body posture, and device orientation. In addition to activity recognition, accelerometer data are essential for MA compensation in biosignals, particularly PPG. In [2], acceleration data are utilized in an adaptive filtering model to remove motion components from PPG and improve HR estimation accuracy. In [20], accelerometer and PPG spectra are fed into a deep neural network for HR estimation, with recognized activity used as an additional feature to improve performance. Raw accelerometer signals often contain noise and MA; thus, preprocessing is performed as an initial step, including noise filtering, smoothing (moving average, exponential smoothing, Savitzky–Golay filter) [21], normalization, and data calibration.

Processing the PPG signal is a fundamental step in obtaining clinically meaningful information, particularly for monitoring cardiac activity and assessing physiological state. PPG signals acquired from wearable devices are affected by noise and, most notably, MA, which significantly complicate the estimation of HR and IPI [22]. The quality and morphology of the PPG signal are influenced by age, health condition, sex, and skin tone: arterial stiffness increases with age, changing the pulse

 doi:10.20944/preprints202603.0657.v1

wave shape [23]; in oncology patients and males, signal quality tends to be lower [24]; and darker skin tones reduce signal acquisition quality due to higher melanin content [25]. To reduce noise, preprocessing techniques are employed, with digital filtering as the primary method [26]. A common approach is to apply band-pass filtering to extract the useful frequency range of the PPG signal, typically the cardiac activity band, for example, 0.67–8 Hz [22] or 0.8–2.5 Hz [1]. To mitigate the effects of MA, digital filtering—particularly band-pass filtering—is used, with the Butterworth filter standing out for its smooth response [27]. DWT with inverse reconstruction allows efficient removal of both low- and high-frequency interferences [28,29]; for example, the Coiflet-1 (coif1) wavelet with four decomposition levels fits well with the morphology of PPG signals [1]. Amplitude demodulation, including Hilbert transform-based approaches, reduces the impact of amplitude artifacts and is helpful for the determination of IPI [1,30]. Different algorithms are used to detect heartbeats in PPG signals: simple thresholding, the AMPD method for noisy signals [31], and the sum-of-slopes function (SSF) for detecting the onset of systole [32]. Particular attention is given to removing MA—the KID-PPG approach [2] employs an adaptive linear filter based on a two-layer convolutional neural network (CNN) that utilizes accelerometer data to subtract motion components from the PPG signal. The choice of an optimal algorithm depends on the trade-off among accuracy, complexity, and noise robustness. The quality of PPG-based HR estimation is validated by comparison with reference ECG signals [32,33]. This validation approach is used in studies that employ the PPG-DaLiA dataset [34,35], where ECG provides the ground-truth HR values for evaluating PPG- and accelerometer-based models [2,20].

The next step is feature extraction for ML models. Statistical time-domain features include mean, SD, variance, maximum, skewness, kurtosis, and correlation coefficients. Specific metrics include ENMO [36], MAD [37], MIMS [38], and activity intensity [39]. Frequency-domain analysis employs the Fast Fourier Transform (FFT) to extract spectral features such as the dominant frequency, subband energy, and spectral entropy; in [20], a time–frequency spectrum (0–4 Hz) is used for HR estimation. Wavelet-based methods are applied to obtain scalograms from accelerometer signals, thereby enabling time-series classification using CNNs (e.g., DenseNet121) [17,40]. FFT-based spectral representations are also used as inputs to CNNs. While [2] primarily focuses on spectral and temporal representations for deep learning, [20] also considers traditional feature extraction from 8-second windows for simpler models such as XGBoost.

In [41], a system for non-invasive blood pressure estimation based on pulse wave analysis from a flexible piezoelectric sensor is proposed. The authors extract 32 time-, frequency-, and wavelet-domain features, consider physiological parameters (BMI, body temperature, sex, age), and employ an RF regression model to estimate systolic and diastolic pressure. Although the sensor type (piezoelectric) and main goal (blood pressure estimation) differ from this study, the ML-based approach leveraging multi-domain features and individual characteristics is relevant. The work [42] focuses on classifying physical load (fatigue) levels from ECG signals during exercise. The authors develop a robust R-peak detection method using a hybrid BiLSTM–CNN model with adaptive post-processing, which is crucial for accurate IPI extraction. An SVM classifier trained on a combination of IPI and other features (wavelets, local binary patterns, higher-order statistics, CWT-based features) is used to identify five fatigue states. This study is significant in the context of cardiac signal analysis under motion and noise, and although it employs ECG rather than PPG, both share the goal of robust HR estimation and ML-based state classification. In [43], HAR based on a hip-placed accelerometer is combined with PPG-derived HR monitoring to compute physical workload in real time using Frimat's scoring criteria. The system links recognized activities to qualitative workload categories and demonstrates body adaptation tracking over time. However, the workload assessment relies on fixed universal thresholds (Frimat's coefficients) and the generic Fox–Haskell formula (2), without accounting for sex-specific HR models, fitness level, or demographic-group normalization.

Article [44] systematizes modern approaches to non-contact physiological monitoring using remote PPG extracted from video data and deep learning. Various architectures—CNNs, RNNs, GANs, and transformers—are used for remote extraction of PPG signals and estimation of HR, respiration

 doi:10.20944/preprints202603.0657.v1

rate, blood oxygen saturation, IPI, and blood pressure. Although this paper focuses on camera-based methods, whereas the current work utilizes wearable sensors, it highlights key challenges in PPG-like signal processing (MA, lighting variation, skin tone) and the role of deep learning, providing a valuable theoretical background. In [20], HR estimation is improved by integrating predicted user activity (obtained from a HAR model based on PPG and accelerometer data) into a deep CNN architecture, reducing HR estimation error. In [2], a KID-PPG approach is proposed, where accelerometer data are explicitly used to model and remove MA from PPG via CNN-based adaptive filtering, alongside a probabilistic HR estimation mechanism accounting for uncertainty. Both studies, though focused on HR estimation, emphasize the importance of accelerometer data and activity context for PPG analysis, which is directly relevant to this work.

Analysis of the reviewed works reveals active use of ML for physiological signal processing and human state classification. The use of multi-domain features, the consideration of individual parameters, and the enhancement of robustness are key research directions. While some works have combined activity recognition with HR-based workload estimation [43], such approaches typically rely on fixed, universally applied workload categories that do not differentiate between demographic subgroups. Other studies incorporate individual parameters (age, BMI, sex) as input features to ML models [41] but do not use them for group-level normalization of physiological responses. In contrast, this study proposes an integrated two-stage approach that combines PPG and accelerometer data from wearable sensors for personalized physiological load assessment: estimated HR is contextualized against demographic-specific statistical distributions defined by sex, age, and activity intensity, enabling automated identification of normal load, fatigue risk, and overload states without prior individual calibration or fixed universal thresholds.

3. Materials and Methods

3.1. Employed datasets

This study utilizes the publicly available PPG-DaLiA dataset [34,35], designed to support research in PPG-based HR estimation, HAR, and physiological-state analysis under realistic daily-life conditions. The dataset's main advantage lies in its multimodal, synchronized recordings, including triaxial accelerometer, PPG, and ECG data, as well as activity labels and participant physiological parameters. This makes it highly suitable for the comprehensive analysis and testing of modern ML algorithms.

Fifteen volunteers (seven males and eight females, mean age 30.6 years) participated in data collection. For each participant, height, weight, Fitzpatrick skin type, and fitness level were recorded. The data collection protocol included eight daily-life activities: sitting at rest, stair ascent and descent, table soccer, cycling, driving, lunch break, walking, and desk work. The total recording duration per participant was approximately 2.5 hours.

Two synchronized devices were used for data acquisition: the chest-worn *RespiBAN Professional* (recording ECG, accelerometer, respiration; sampling rate 700 Hz) and the wrist-worn *Empatica E4* (triaxial accelerometer 32 Hz, PPG 64 Hz, electrodermal activity and skin temperature 4 Hz). In this work, we use accelerometer and PPG data from *Empatica E4* as the primary input, while reference HR values derived from ECG serve to validate performance.

Signal synchronization between devices was achieved via double-tap markers at the start and end of each session, ensuring precise time alignment and drift correction. The data were segmented into 8-second windows with a 2-second step size. Reference HR values were computed as the mean instantaneous HR within each window. The dataset structure for each participant includes: activity labels (4 Hz), reference HR values (from ECG), physiological parameters, R-peak indices, accelerometer and PPG signals from multiple locations, and a unique participant ID. For this research, the key components are segmented PPG and accelerometer signals from the Empatica E4, along with the corresponding reference HR. An additional advantage of this dataset is its representation of activities across different load intensities, which is essential for accurately assessing algorithm performance. In total, 64,697 eight-second signal windows are stored in this dataset, of which 47,182 correspond to

 doi:10.20944/preprints202603.0657.v1

the eight labeled activities; the remaining windows represent transition periods between activities. A circular diagram (Figure 1) visualizes the distribution of activity windows across all subjects.

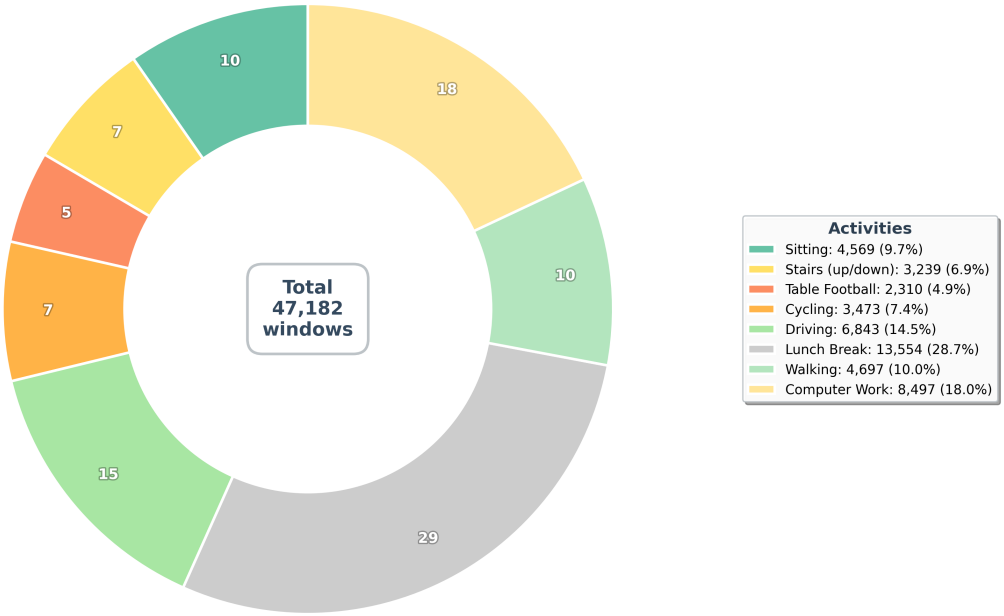


Figure 1. Circular diagram showing activity-window distribution in the PPG-DaLiA dataset.

Thanks to its multimodality and high annotation quality, the PPG-DaLiA dataset enables validation of various HAR, HR-estimation, and physiological-analysis approaches under conditions that are close to real-world wearable-sensor use.

3.2. Signal preprocessing

The investigated HAR methods involve joint processing of triaxial accelerometer and PPG signals to improve input data quality and ensure reliable classification. The relevance of including PPG features in HAR tasks is confirmed by recent studies [20], which demonstrate that PPG-derived characteristics also carry discriminative information for activity recognition. Therefore, in this study, both accelerometer and PPG parameters are incorporated into the feature vector.

At the preprocessing stage, raw accelerometer and PPG signals undergo noise reduction. For accelerometer data, sequential median filtering is applied to suppress impulsive noise, followed by a third-order Butterworth filter with a cutoff frequency of 15 Hz to attenuate high-frequency components while preserving motion-related content. A low-pass filter with a 0.3 Hz cutoff then separates gravitational and body-motion components, enabling independent analysis of posture and movement.

PPG signals undergo four-level discrete wavelet decomposition to remove baseline drift and noise while retaining cardiovascular-relevant frequency components. Amplitude demodulation is subsequently performed using the Hilbert transform [45], improving feature stability and HR-estimation accuracy.

3.3. Physical Activity Classification

Figure 2 presents the block diagram of the feature engineering process for automated classification of physical activity types from accelerometer and PPG signals. The algorithm involves sequential data processing, feature formation and selection, and evaluation of different ML models.

 doi:10.20944/preprints202603.0657.v1

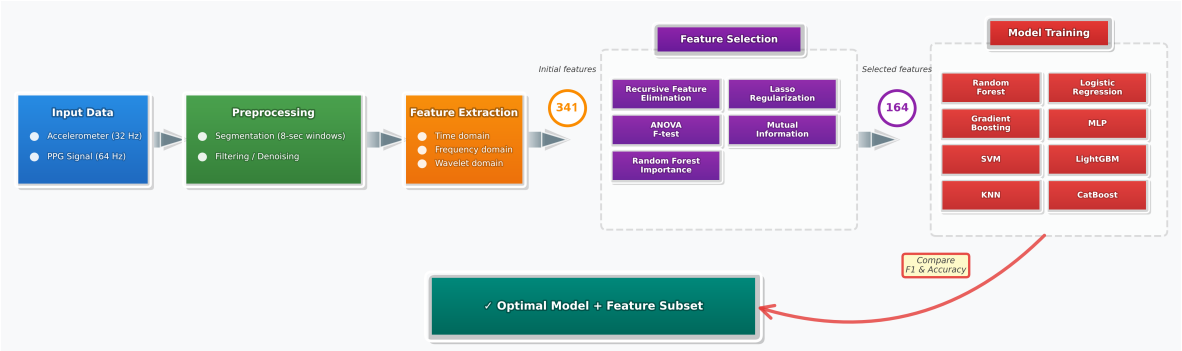


Figure 2. Block diagram of the feature engineering process for physical activity classification.

The initial stage consists of acquiring accelerometer signals sampled at 32 Hz and PPG signals sampled at 64 Hz. After preprocessing, the accelerometer and PPG signals are segmented into overlapping 8-second time windows to standardize analysis and facilitate feature formation. The next stage includes computing an initial feature set.

For activity classification, a feature vector is constructed for each time window. This vector must include quantitative characteristics that best describe the activity within that window. Each window yields an extended set of 341 features covering time, frequency, and wavelet domains. Key features include statistical characteristics (mean, SD, median, skewness, kurtosis), signal energy and entropy, peak and trough characteristics, autocorrelation, cross-correlation, FFT-based spectral parameters, and wavelet-packet energy features reflecting non-stationary motion and cardiac patterns. Additional features include body acceleration statistics, inter-axis correlations, and directly computed PPG spectral features.

Time-domain features

These are calculated directly from time series data. According to [46,47], the most common and informative are mean value, SD, variance, median, median absolute deviation, skewness, kurtosis, root mean square, energy: $E = \sum_{i=1}^{N} x_i^2$; area under the absolute signal: $A = \sum_{i=1}^{N} |x_i|$; as well as extremes and range, interquartile range, signal entropy, correlation coefficients between axes, and quantiles.

Frequency-domain features

These are derived from the signal spectrum obtained using FFT. They describe how signal energy is distributed across frequencies:

- Frequency of the component with maximum amplitude.
- Mean weighted frequency.
- Spectral energy within predefined frequency bands.
- Skewness and kurtosis of the spectral distribution.

Wavelet-domain features

As an alternative or complement to FFT, the wavelet transform is effective for analyzing non-stationary HAR signals. Wavelet packet decomposition provides a more detailed time–frequency representation [47]. For each terminal node j of the decomposition, the normalized energy is computed as (1):

$$E_j = \frac{\sum_k |W_{j,k}|^2}{\sum_m \sum_k |W_{m,k}|^2} \tag{1}$$

In HAR tasks, the main signal energy is often concentrated in low-frequency wavelet components, and the Haar wavelet has demonstrated high effectiveness for constructing energy-based features [47].

In total, the 341 features are distributed as follows: 197 accelerometer time-domain features (14 statistical measures computed for 12 base signals, inter-axis correlation coefficients for three axial pairs, two angle calculations, and four autoregressive coefficients for six axis vectors), 80 accelerometer

 doi:10.20944/preprints202603.0657.v1

frequency-domain features (10 FFT-based metrics for eight frequency-domain signals), 48 accelerometer wavelet-domain features (16 wavelet packet energy bands for each of three axes), and 16 PPG features (10 time-domain statistical measures, two single-frequency metrics, and four band energy metrics).

To avoid redundancy, reduce feature-space dimensionality, and prevent overfitting, feature selection is applied using recursive feature elimination, RF importance, ANOVA F-score, Lasso coefficients, and mutual information. Low-variance and highly correlated ($|r| > 0.85$) features are removed. Feature selection was performed exclusively on the training data (13 subjects) to prevent information leakage; the two held-out test subjects were not involved in any feature selection or model training step.

The next step is to compare the accuracy and F1-score for various feature subsets across all models. Each segmented 8-second window is represented by a feature vector and assigned one of eight activity labels. For multiclass classification, several ML models are trained in parallel and tested: RF, SVM, Gradient Boosting, KNN, Multilayer Perceptron (MLP), and modern LightGBM and CatBoost. Training and testing follow a Group K-Fold cross-validation scheme (grouped by subject ID, ensuring that data from the same participant never appear in both training and validation folds), and performance is evaluated using F1-score, precision, recall, and accuracy.

The final stage is the selection of the optimal model and identification of the most informative features for the activity classification task. As a result, a subset of 164 informative features is formed. The analysis revealed that time- and frequency-domain body-acceleration features are most informative for activity recognition, whereas frequency and certain wavelet-domain PPG features significantly improve the accuracy of physiological-state and HR estimation.

Thus, the combination of robust signal preprocessing, extended feature engineering from accelerometer and PPG signals, and effective feature selection provides the foundation for a reliable, automated HAR system that accounts for individual physiological differences.

3.4. HR Estimation

The PPG signal, especially under dynamic conditions, often contains substantial noise and MA, which complicate accurate HR estimation. To improve quality, a multi-stage preprocessing pipeline is required. Figure 3 presents the block diagram of the algorithm for HR estimation based on PPG and accelerometer signal processing. The algorithm performs stepwise data processing with quality control at each stage, ensuring the reliability of the obtained results.

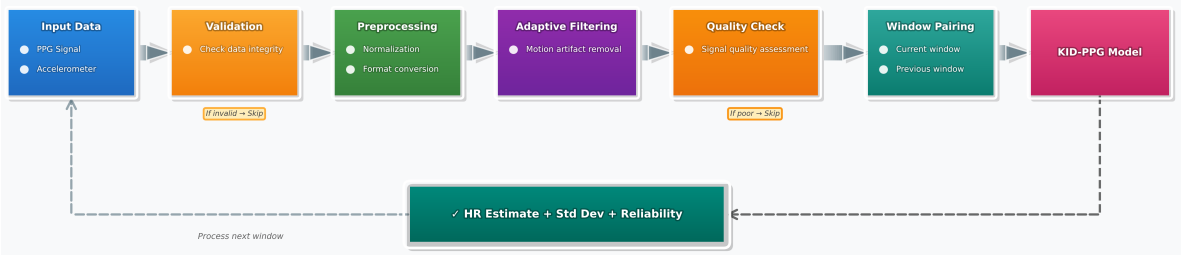


Figure 3. Block diagram of the feature engineering process for HR estimation.

The algorithm starts by receiving segments of PPG and accelerometer signals. In the first step, a validation is performed: if invalid or malformed signal values are detected, processing for that window is terminated, and the system proceeds to the next segment. For valid input data, preprocessing is applied, including normalization and format conversion to unify data representation. Then, adaptive filtering is performed to remove MA and reduce noise in the PPG signal. Following [1], this stage applies DWT with subsequent signal reconstruction to suppress low-frequency baseline drift and high-frequency noise. In this study, the Coiflet-1 wavelet with four decomposition levels is used, as its morphology aligns well with the PPG pulse waveform. Only components carrying cardiac-rhythm content are retained, while noise-dominant components are attenuated. The next step performs

 doi:10.20944/preprints202603.0657.v1

amplitude demodulation via the Hilbert transform to obtain the signal envelope [45]. Normalization at this stage mitigates amplitude artifacts and stabilizes subsequent peak detection.

Next, a signal-quality check is performed: if the signal remains unreliable (anomalies or invalid values persist), processing of that window is stopped. Otherwise, a pair of the current and previous cleaned signal windows is formed. This enables more robust HR estimation via inter-window analysis.

For HR estimation, several systolic-peak detectors are considered, including classical thresholding, the AMPD method, and SSF. Each approach has advantages depending on the signal-to-noise ratio and waveform complexity. The accuracy of HR estimation is assessed by comparison with reference values obtained from the synchronously recorded ECG signal. However, even with careful preprocessing, classical algorithms do not always provide reliable HR estimates in the presence of substantial MA. To overcome these limitations, this work adopts the KID-PPG model [2]. KID-PPG is a knowledge-informed deep learning model that processes data in 8-second windows with a 2-second overlap. The inputs consist of a PPG signal ($N \times 1$) and a triaxial accelerometer signal ($N \times 3$), where $N = 8 \times f_s$ and f_s is the sensor sampling frequency. The model comprises two stages. The first stage is a linear adaptive filter—a two-layer convolutional network that processes the three-channel accelerometer signal alongside the PPG signal, applying spatio-temporal filtering to estimate and subtract MA and produce a cleaned blood volume pulse signal. This filter is adapted using stochastic gradient descent (SGD) for 1000 epochs, with the loss computed as the mean squared error of the FFT of the inputs. The second stage is a temporal attention model—a deep network consisting of three convolutional blocks (32, 48, and 64 filters), an average pooling layer, a multi-head temporal attention layer with a residual connection, layer normalization, and fully connected layers. The output layer parameterizes a normal distribution by its mean and SD, enabling probabilistic HR estimation with built-in uncertainty quantification. This model is trained using the Adam optimizer (learning rate $= 0.0005$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$). In this study, the pre-trained model weights provided by the original authors were used directly, while only the adaptive filtering stage was fine-tuned per window. Owing to these properties, KID-PPG is selected as the primary approach in this study, offering superior robustness to MA and high HR-estimation accuracy on real multimodal data compared with classical methods.

After processing one window, the system proceeds to the next signal segment. The process repeats for all available windows, enabling continuous analysis and data processing.

3.5. Personalized Load Assessment

Assessment of workers' physiological state from biometric data is central to health-monitoring systems. The main indicators are HR, activity level, BMI, age, sex, and fitness level. Figure 4 illustrates the block diagram of the proposed algorithm for the comprehensive analysis of a worker's physiological state, incorporating HR estimation, effort index (EI) calculation, demographic group, and activity-intensity classification.

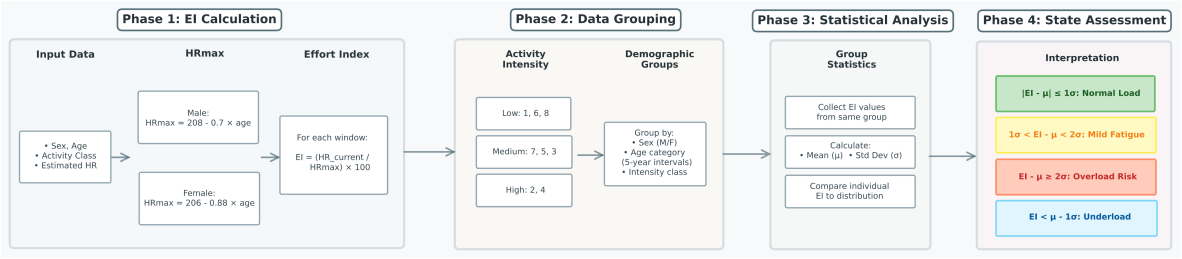


Figure 4. Block diagram of the algorithm for analyzing the worker's physiological state.

The algorithm begins by obtaining personal data, including sex, age, recognized activity class, and estimated HR for each signal window. The first step verifies the completeness and correctness of the input (sex, age, HR, activity ID). If missing or invalid data are found, the corresponding segment is skipped.

 doi:10.20944/preprints202603.0657.v1

Next, the individual's maximum HR is determined. The algorithmic approach involves normalizing the worker's HR relative to their individual maximum HR, computed using validated formulas, such as the widely used Fox–Haskell formula [48] (2):

$$HR_{max} = 220 - \text{age}. \tag{2}$$

However, this formula tends to overestimate HR for older individuals. For men, the Tanaka equation [49] provides a more accurate estimate (3):

$$HR_{max} = 208 - 0.7 \times \text{age}. \tag{3}$$

For women, the Gulati formula [50] is recommended (4):

$$HR_{max} = 206 - 0.88 \times \text{age}. \tag{4}$$

Age is the primary factor influencing HR_{max}. Sex also plays a role, as the HR decline curve with age differs between men and women. BMI and fitness level mainly affect resting HR but have little influence on HR_{max}. Athletes generally exhibit a lower resting HR and faster recovery after exertion, although HR_{max} differs only slightly from that of untrained individuals of the same age [51]. Hence, in the proposed approach, the individual's maximum HR is determined by sex: the formula (3) is used for males and (4) for females.

After determining HR_{max} for each worker, the current HR is normalized using the EI as the ratio of the estimated HR to the individual HR_{max}, expressed as a percentage (5):

$$EI = \frac{HR_{current}}{HR_{max}} \times 100\%. \tag{5}$$

This approach ensures that age and sex are directly incorporated into workload estimation, providing a personalized and physiologically justified assessment.

Then, according to the activity identifier, each window is assigned to an intensity class: low (1-sitting, 6-lunch break, 8-desk work), medium (7-walking, 5-driving a car, 3-table soccer), or high (2-stairs, 4-cycling). Segments with undefined activity IDs are skipped. The activity labeled "break" was excluded from further analysis because the PPG-DaLiA dataset lacks a clear description of the participants' actions during this activity. This prevents unambiguous classification in terms of physiological load level and therefore excludes it from grouping.

Approaches to state analysis can be categorized into algorithmic (traditional) and statistical. For **algorithmic** analysis of physiological workload, the EI is usually divided into standard intensity zones based on the percentage of HR_{max}, as recommended by established exercise prescription guidelines [52]. These zones typically range from very light ($< 57\%$ HR_{max}) through moderate (64–76% HR_{max}) to near-maximal ($\geq 96\%$ HR_{max}) intensity. Each zone allows relating the worker's current physiological load to the nature of the performed activity. If the effort index is below the zone threshold, it may indicate underactivation; matching the zone corresponds to the optimal load; exceeding it signals tension or overload. The main drawback of this traditional approach lies in its reliance on unified ranges that neglect individual variability and may lead to inaccurate assessments. A related approach is the Karvonen method [53], which defines target intensity as a percentage of heart rate reserve ($HR_{max} - HR_{rest}$) rather than HR_{max} alone, offering improved precision for individuals whose resting HR deviates substantially from population norms. However, it requires a reliable measurement of each individual's resting HR, introducing an additional per-person calibration step that limits applicability in continuous monitoring scenarios; moreover, publicly available wearable-sensor datasets rarely include standardized resting HR measurements.

The **statistical** approach proposed in this work provides a personalized evaluation: for each group, the EI distribution is analyzed by sex, age, and activity, and the mean and SD are reported. The worker's current index is compared to this distribution—typical values indicate a normal state,

 doi:10.20944/preprints202603.0657.v1

significant excess indicates a risk of fatigue, and lower values suggest insufficient activity. Thus, the system accounts for both individual and group-level variability. To form the reference group, each worker is categorized by sex and age (in 5-year intervals). Next, the algorithm checks whether sufficient reference data are available within the same demographic group and intensity class. If insufficient reference data exist, analysis of the current segment is not possible. When sufficient reference data are available, EI values of other workers in the same demographic and intensity group are collected. The group mean and SD of EI are then computed. For each window, the worker's effort index is compared to the group mean and SD: below (mean − SD), within one SD, between one and two SDs, or above two SDs. Interpretation of these categories follows established criteria (Table 1).

Table 1. Evaluation and interpretation of physiological load.

Deviation from Expected HR	Physiological Interpretation	Fatigue Level	Health/Load Impact	Description
Within 1σ	Normal cardiovascular response	Normal / Balanced	Normal load	Expected effort maintained over several hours.
1σ–2σ	Moderate increase in sympathetic activity	Growing fatigue	Mild fatigue. Short-term load is acceptable,	A break may be required after prolonged activity.
$\geq 2\sigma$	High cardiovascular load	Near exhaustion	Risk of overload, dehydration, or overheating	Rest recommended; signs of physiological stress present.
Below expected	Insufficient activation or high training level	Recovery / Underload	Low workload	Worker under-engaged or not sufficiently active.

Table 1 summarizes the interpretation criteria used to classify each signal window into one of four physiological states based on the worker's deviation from the demographic-group mean EI. Windows falling within one SD represent a normal cardiovascular response appropriate for the given activity and demographic profile. Deviations of one to two SDs indicate growing sympathetic activation and suggest that a break may be needed if the elevated load persists. Values exceeding two SDs signal a risk of physiological overload, warranting rest or workload reduction. Conversely, EI values below the expected range may reflect insufficient activation or, alternatively, a high level of cardiovascular fitness. This statistical framework avoids the limitations of fixed HR-zone thresholds [52] by anchoring the assessment to group-specific distributions, thereby accommodating inter-individual variability arising from differences in age, sex, and activity type.

3.6. Validation and Experimental Setup

Evaluation of the proposed methodology was conducted in two main directions: evaluation of the quality of physical activity classification and verification of HR estimation accuracy. To objectively assess the models' generalization capability, data from two randomly selected subjects of opposite sexes were used. Their recordings were not included in the model training phase.

In the first testing stage, an RF activity classification model was trained using 164 selected features. Model hyperparameters were tuned via randomized search with Group K-Fold cross-validation (grouped by subject ID) to maximize classification accuracy. After training was completed, the model's performance was evaluated on the two designated test subjects.

The second direction of the test involved applying the KID-PPG model to HR estimation. The main focus was on tuning the adaptive filtering parameters, since the model—which employs knowledge-

 doi:10.20944/preprints202603.0657.v1

informed, physics-driven training—was used in a pre-trained form, as recommended by the authors, and did not require additional retraining on the dataset. This approach was chosen because the model had already undergone multi-stage training and testing across various datasets, including PPG-DaLiA, demonstrating strong generalization. Adaptive filtering parameters were varied by changing the optimizer, learning rate, and number of epochs.

Within the comprehensive system evaluation, a full procedure was performed for each signal window of the test subjects: activity classification, PPG filtering using accelerometer data, HR estimation via KID-PPG, and statistical analysis of the physical-load intensity distribution for each subject, accounting for demographic group affiliation. Detailed testing results, including illustrative model visualizations, are presented in the following sections.

4. Results and Discussion

4.1. Feature Selection

Figure A1 displays the aggregated feature ranking across multiple selection methods (Recursive Feature Elimination, RF importance, ANOVA F-score, Lasso, and Mutual Information). The histogram reveals that body acceleration features (BodyAcc) dominate the top rankings, with statistical measures (mean, standard deviation, correlation coefficients) and wavelet energy features appearing most frequently among high-priority features. Notably, accelerometer-derived features across the X, Y, and Z axes constitute the majority of the top 164 selected features. PPG-related features, particularly those from frequency-domain analysis (wavelet energy bands) and spectral characteristics, fall within the mid-to-low priority range but remain essential for comprehensive activity characterization. The gradient from dark purple (highest rank) to yellow (lowest rank) illustrates a progressive decrease in discriminative power, with the majority of high-impact features concentrated in the temporal and wavelet domains of body acceleration signals.

The analysis of feature ranking reveals that body acceleration features dominate the top-ranked positions, particularly time-domain statistical measures (mean, standard deviation, root mean square) computed across the three accelerometer axes. These features effectively capture the distinctive motion patterns associated with different activities. Notably, frequency-domain features from both accelerometer and PPG signals also demonstrate substantial discriminative power, especially wavelet packet energy coefficients that capture non-stationary patterns inherent in human movement and cardiac activity. Among PPG-derived features, spectral characteristics in the 0.67–8 Hz band show significant importance for activity classification, complementing the motion-based features. This finding supports recent research demonstrating that cardiovascular responses vary systematically across activity types, making PPG features valuable for HAR tasks beyond their primary role in HR estimation.

Highly correlated features ($|r| > 0.85$) were systematically removed to minimize redundancy. The resulting feature subset balances discriminative power with practical considerations for real-time implementation on resource-constrained wearable devices.

Figure 5 presents the Kendall correlation matrix, revealing strong positive correlations (0.73-0.94) across all activity pairs, indicating substantial overlap in feature distributions between different activities. The highest correlations occur between Stairs-Walking (0.94), Stairs-Break (0.90), and Desk Work-Sitting (0.91), suggesting these activity pairs share similar accelerometer and PPG characteristics. Activities involving sedentary behavior (Sitting, Desk Work, Lunch) show moderate-to-high intercorrelations (0.87-0.91), whereas dynamic activities (Cycling, Table Soccer, Driving) exhibit correlation coefficients ranging from 0.72 to 0.83. This high baseline correlation across all activities underscores the challenge of discriminative classification and motivates comprehensive feature engineering to capture subtle distinguishing characteristics.

 doi:10.20944/preprints202603.0657.v1

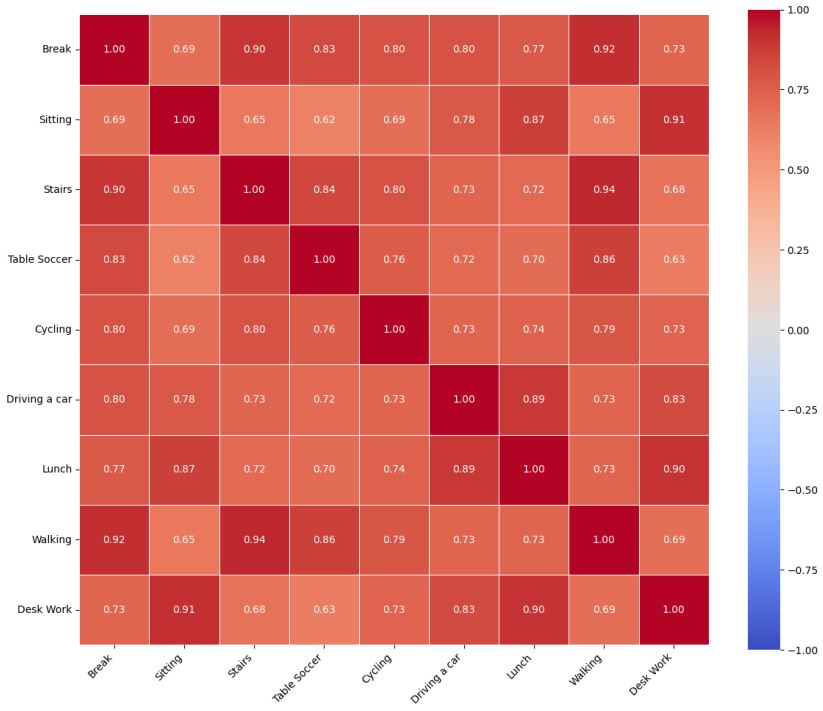
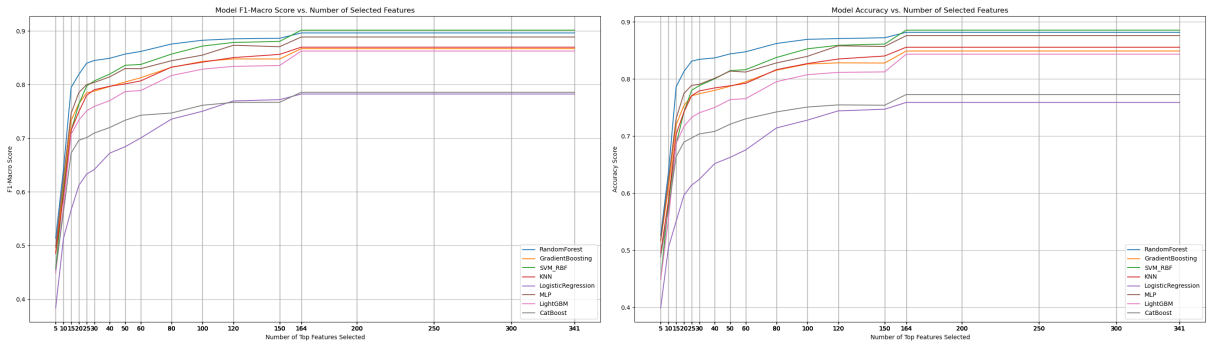


Figure 5. Kendall correlation matrix between activity types.

Figure 6 demonstrates the relationship between the number of selected features and classification performance across eight ML models. Both F1-score and accuracy improve rapidly within the initial feature range (5-30 features), reaching near-optimal performance at approximately 164 features. Beyond this threshold, performance gains become marginal across all models, indicating feature saturation.



(**a**) (**b**)

Figure 6. Models' F1-score (a) and accuracy (b) depending on the number of features.

The ensemble methods (RF, Gradient Boosting, SVM_RBF) converge to F1-scores of 0.88-0.91 and accuracies of 0.87-0.89 with the optimal 164-feature subset. KNN and Logistic Regression show more gradual improvement curves and lower asymptotic performance (F1-score ≈ 0.78-0.85), suggesting sensitivity to feature dimensionality. MLP demonstrates intermediate performance but exhibits greater variability. The behavior of the plateau beyond 164 features supports the decision to retain this subset, balancing classification accuracy with computational efficiency while avoiding redundancy and potential overfitting from the entire 341-feature set.

The feature selection process identified 164 optimal features from an initial set of 341 characteristics extracted from accelerometer and PPG signals. The resulting dimensionality was reduced by 52% while maintaining classification performance, thereby improving computational efficiency and reducing the risk of overfitting.

 doi:10.20944/preprints202603.0657.v1

4.2. Activity Classification

The RF classifier trained on 164 selected features achieved strong overall performance on the test set, as shown in Table 2. The model demonstrated robust generalization, with a macro F1-score of 90.73%, indicating balanced performance across activity classes despite dataset imbalance.

Table 2. Overall classification performance of the RF model.

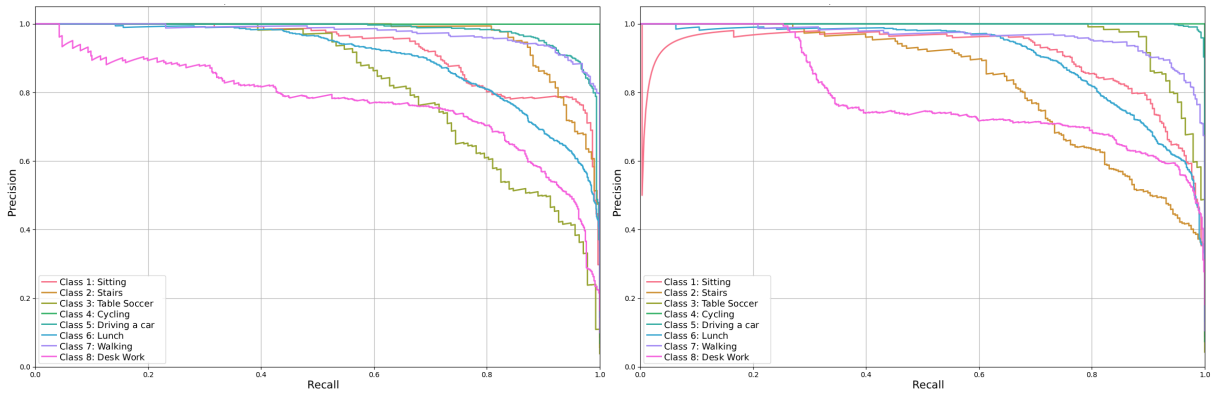
Accuracy	F1-macro	F1-weighted
0.8936	0.9073	0.8940

To evaluate the generalization capability of the model across diverse demographic profiles, two independent test subjects with contrasting characteristics were selected:

1. S7: Female, age 21, BMI 20.5, fitness level 2.
2. S12: Male, age 43, BMI 27.6, fitness level 5.

These subjects represent the opposite ends of the demographic spectrum in the dataset, enabling assessment of model robustness across variations in age, sex, body composition, and fitness level.

Figure 7 shows the precision–recall (PR) curve of the RF model for each activity, enabling the assessment of the precision and recall relationships for different activity classes for the first (F1-score 90.2%) and second (F1-score 86%) users.



(**a**) (**b**)

Figure 7. PR curve for the RF activity classification model for the first (a) and second (b) test users.

As can be observed, the PR curves indicate a strong ability to correctly classify most activity types. Particularly high precision and recall are observed for the classes "Table soccer", "Sitting", "Stairs", and "Lunch break", where precision values approach 1 even at high recall. The model best distinguished the "Table soccer" and "Lunch break" classes due to their characteristic features. At the same time, for the "Desk Work" and "Cycling" classes, a notable drop in precision is observed as recall increases, indicating a considerable number of false positives for these activities. This may be related to both contextual similarity between activities and class imbalance in the training set.

Figure 8 presents the confusion matrix, which illustrates the probability of assigning signals to each activity class in the test data and serves as an additional indicator of the model's generalization quality for the first and second test users.

 doi:10.20944/preprints202603.0657.v1

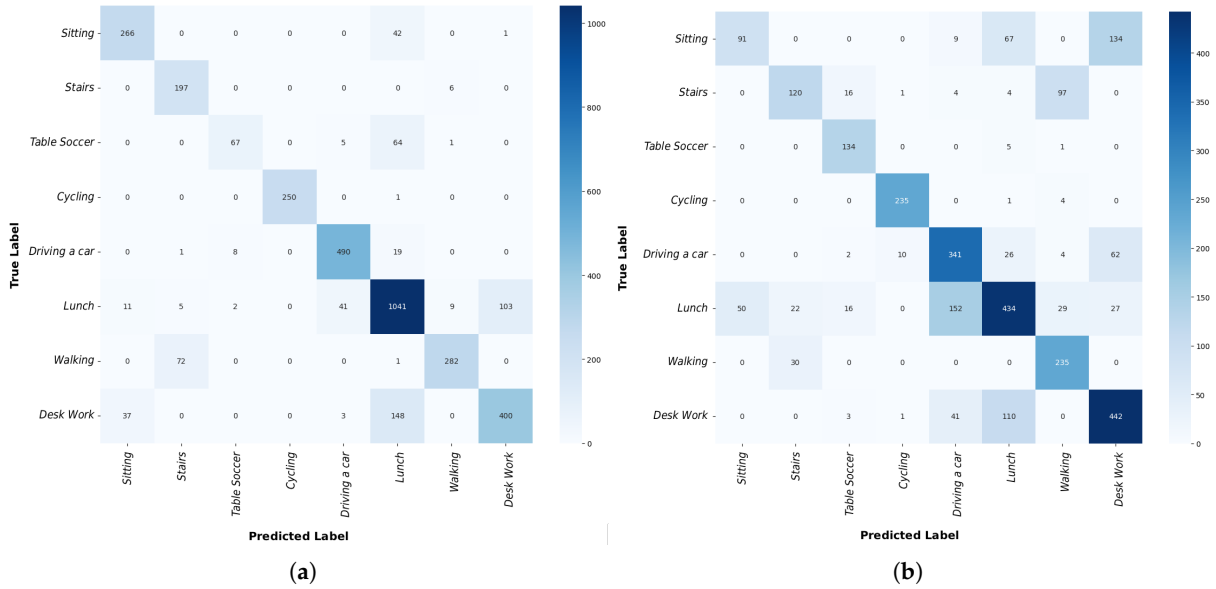


Figure 8. Confusion matrix for the classification results of the first (a) and second (b) test users.

The most frequent confusions occur between "Desk Work" and "Lunch break" and between "Sitting" and "Desk Work". Conversely, for activities associated with active movement—namely, "Table soccer", "Cycling", and "Walking"—the model generally maintains high accuracy and minimizes misclassifications, which is important for physical-activity monitoring systems.

The impact of class imbalance is also evident: activities with more samples exhibit more stable metrics, whereas less frequent classes show lower performance and greater error variability, highlighting the need for further dataset balancing or class-weight optimization during training. To address these issues, forming *intensity groups* proved useful: grouping activities with similar features or outcomes simplifies error interpretation and improves the system's practical relevance.

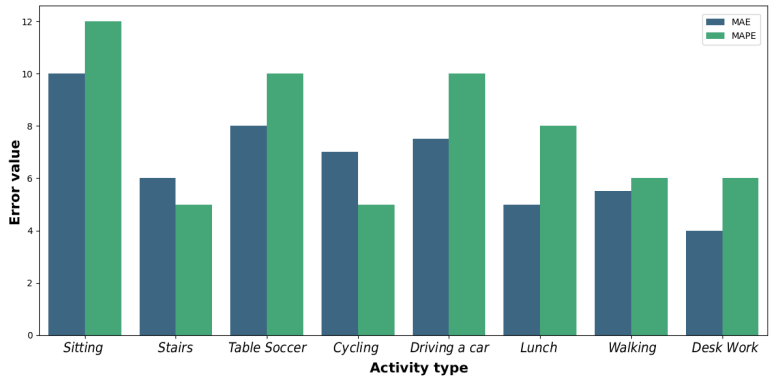
4.3. HR Estimation

Various HR estimation methods were analyzed in detail, including classical peak detection algorithms (derivative- and threshold-based), AMPD, and modern wavelet filters employing adaptive algorithms. The comparative analysis confirmed that classical methods are sufficiently accurate only for high-quality signals; under more complex conditions (with noise and MA), the deep learning KID-PPG model achieves the best accuracy and robustness in the final estimation.

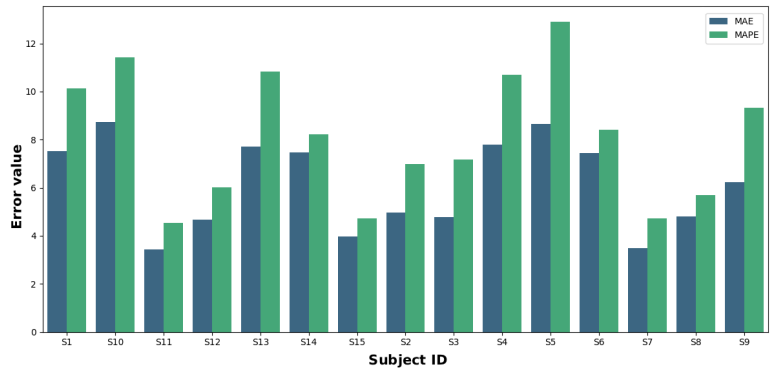
To evaluate the performance of the KID-PPG model, a series of experiments was conducted with varying adaptive filtering parameters (different optimizers and learning rates). Figure 9 presents histograms of MAE and MAPE for each activity type and user, allowing for the evaluation of model quality by activity category.

The MAE and MAPE evaluations for each activity and user (Figure 9) indicate variability in prediction accuracy across scenarios. MAE ranges from 4 to 10 bpm across activity types, while MAPE ranges from 5% to 13%, depending on the user. Typically, better accuracy is observed for static activities (e.g., desk work or sitting), whereas dynamic and complex movements (e.g., table soccer and driving) exhibit higher errors. It is worth noting that the KID-PPG model not only provides a quantitative HR estimate but also outputs a numerical measure of result reliability—the SD of the estimate—which depends on the input signal quality. Based on this parameter, the adaptive filtering parameters for PPG recordings are adjusted using accelerometer data: for noisier signals, a larger number of epochs can be selected.

 doi:10.20944/preprints202603.0657.v1



(**a**)



(**b**)

Figure 9. MAE and MAPE of the KID-PPG model for each activity type (a) and test users (b).

4.4. Physiological State Assessment

The personalized load assessment methodology described in Section 3.5 was applied to all 15 participants. For each subject, the effort index was computed for every eight-second window, assigned to an intensity class (low, medium, or high), and compared against the corresponding demographic-group distribution. Table A1 presents the resulting physiological load distributions, broken down by intensity class and subject.

Across the full cohort, the largest share of signal windows falls within one SD of the group mean (median across subjects: 40.2% for low, 43.5% for medium, and 48.1% for high intensity), corresponding to a normal cardiovascular response. Overload states ($\geq 2\sigma$) occur most frequently in the low-intensity group (median 12.5%), followed by medium intensity (median 11.3%), while high-intensity activities exhibit the lowest overload rate (median 6.7%). This pattern reflects the system's design: because the effort index is evaluated relative to activity-appropriate reference ranges, elevated HR during vigorous exercise is expected and therefore does not trigger overload classification, whereas unexpectedly high HR during sedentary or moderate tasks indicates a genuinely atypical physiological response.

Analysis of physiological-load distributions for individual users (Table A1) reveals several patterns linked to demographic and fitness characteristics. The youngest participants with low BMI (S7, S11, and S13; ages 21–24, BMI 20.5–22.0) exhibit the most homogeneous load profiles: their average within-1σ rates exceed 48%, and overload ($\geq 2\sigma$) rates are near zero (0.0% for S7 and S13, 2.0% for S11), indicating highly predictable cardiovascular responses regardless of fitness level (which ranges from 2 to 6 in this subgroup).

A distinct age-related pattern emerges among the three oldest participants (S8, S10, and S12; ages 43–55). While their high-intensity overload rates remain moderate (5.8–7.5%), the medium-intensity group reveals pronounced overload: 37.5% for S8, 51.0% for S10, and 28.7% for S12. This suggests that moderate activities—which are expected to produce only modest cardiovascular responses—elicit disproportionately elevated HR in older participants, potentially reflecting age-related reductions in

 doi:10.20944/preprints202603.0657.v1

cardiovascular efficiency. Notably, S12 (BMI 27.6, the highest in the cohort) shows consistently lower within-1σ rates (average 36.3%), suggesting that elevated BMI compounds the age effect.

User S6 (fitness level 1, the lowest in the dataset) presents a unique pattern: in the low-intensity group, the overload rate reaches 42.0% with only 15.1% of windows within 1σ, whereas medium- and high-intensity distributions are unremarkable (10.9% and 8.9% overload, respectively). This anomaly is likely attributable to data quality issues rather than physiology alone: the PPG-DaLiA dataset authors reported difficulties in recording this participant's signals, with some segments removed, which may have distorted the low-intensity reference distribution.

These findings indicate a substantial influence of participants' individual characteristics—particularly age—on physiological-load distributions and underscore the importance of both demographic-aware normalization and rigorous data quality control in the interpretation of automated monitoring results.

4.5. Limitations and Future Work

Several limitations should be acknowledged. First, the PPG-DaLiA dataset, while multimodal and well-annotated, comprises only 15 participants with a limited age range (21–55 years) and does not include participants with chronic conditions or occupational populations performing real industrial tasks. The demographic-group normalization is therefore based on relatively small reference groups, and its robustness should be verified on larger and more diverse cohorts. Second, the KID-PPG computation time of approximately 5 seconds per eight-second window (at 1000 iterations) is suitable for delayed near-real-time processing but not yet sufficient for true real-time analysis. Given the model's computational complexity, deployment on low-power wearable devices would require offloading processing to a remote server using a modular architecture, which in turn raises requirements for robust data protection and anonymization of personal health information. Third, certain activity pairs—particularly sedentary classes such as "desk work" and "sitting"—exhibit high feature-space similarity (Kendall correlation 0.91), leading to elevated misclassification rates. Class imbalance in the dataset further contributes to performance variability across less-represented activities. Fourth, the current approach groups activities into three broad intensity levels (low, medium, high). A finer-grained intensity classification or continuous intensity estimation could improve the sensitivity of the physiological load assessment, particularly for activities near category boundaries. Future work should address these limitations by: (1) validating the methodology on larger occupational datasets collected in real workplace environments; (2) optimizing the KID-PPG pipeline for edge-device deployment, potentially through model distillation or quantization; (3) expanding the range of analyzed physiological parameters beyond HR (e.g., HR variability, respiratory rate); (4) investigating adaptive demographic grouping strategies that account for evolving participant baselines over time; and (5) implementing multi-level personal data protection mechanisms suitable for continuous occupational monitoring.

5. Conclusions

This study presented a methodology for personalized assessment of workers' physiological load by integrating accelerometer-based activity recognition with PPG-derived HR estimation from wrist-worn wearable sensors. The methodology was validated on the PPG-DaLiA dataset comprising 15 participants with diverse demographic profiles (ages 21–55, BMI 20.2–27.6, fitness levels 1–6). The RF classifier trained on an optimized subset of 164 time-, frequency-, and wavelet-domain features achieved a macro F1-score of 90.73% for eight-class activity recognition. The KID-PPG deep learning model provided robust HR estimation with MAE below 10 bpm across activity types, including dynamic conditions with substantial motion artifacts. The proposed personalized load assessment framework, which normalizes effort indices against demographic-group-specific distributions, revealed that participant characteristics—particularly age—substantially influence physiological-load patterns. Notably, older participants (ages 43–55) exhibited elevated overload rates during medium-intensity activities (up to 51.0%), whereas the youngest participants (ages 21–24) demonstrated highly

 doi:10.20944/preprints202603.0657.v1

predictable cardiovascular responses with near-zero overload rates. These findings confirm that demographic-aware normalization is essential for accurate physiological state interpretation and that generic HR thresholds are insufficient for diverse worker populations. The proposed approach can be applied to workplace health monitoring systems, occupational safety platforms, sports performance analysis, and rehabilitation programs. Future work should focus on optimizing computational efficiency for real-time deployment on edge devices, expanding the range of analyzed physiological parameters, and validating the methodology on larger and more diverse occupational datasets.

Author Contributions: Conceptualization, O.L., M.M., and O.P.; methodology, O.L., M.M., and O.P.; software, O.L.; validation, O.L.; formal analysis, O.L.; investigation, O.L.; resources, O.L., M.M., and O.P.; data curation, O.L.; writing—original draft preparation, O.L., and M.M.; writing—review and editing, O.L., M.M., and O.P.; visualization, O.L. and M.M.; supervision, M.M., and O.P.; project administration, M.M., and O.P.; funding acquisition, M.M., and O.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was co-funded by the European Union HORIZON TMA MSCA Doctoral Networks / HORIZON-MSCA-2023-DN-01 / project TUAI / grant agreement N° 101168344. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. Co-funded by the Statutory Research funds of the Department of Distributed Systems and Informatic Devices, Silesian University of Technology, Gliwice, Poland (Grants BK-244/RAu8/2025 02/110/BK_25/1036).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the PPG-DaLiA dataset as reported by the original dataset authors [34,35]. No additional human subjects were involved in this secondary analysis study.

Data Availability Statement: The PPG-DaLiA dataset used in this study is publicly available at [35]. Code and additional analysis scripts developed for this work can be made available upon reasonable request to the corresponding author.

Acknowledgments: During the preparation of this manuscript, the authors used Claude Opus 4.6 for the purposes of proofreading and partial generation of block diagrams and Google NotebookLM for generating some visual elements for the graphical abstract. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AMPD	Automatic Multiscale-based Peak Detection
BMI	Body Mass Index
CNN	Convolutional Neural Network
DWT	Discrete Wavelet Transform
ECG	Electrocardiography
EI	Effort Index
FFT	Fast Fourier Transform
HAR	Human Activity Recognition
HR	Heart Rate
IPI	Inter-Pulse Interval
KNN	k-Nearest Neighbors
MA	Motion Artifacts
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MLP	Multilayer Perceptron
PPG	Photoplethysmography
RF	Random Forest

 doi:10.20944/preprints202603.0657.v1

RMS	Root Mean Square
SD	Standard Deviation
SSF	Sum-of-Slopes Function
SVM	Support Vector Machine

Appendix A

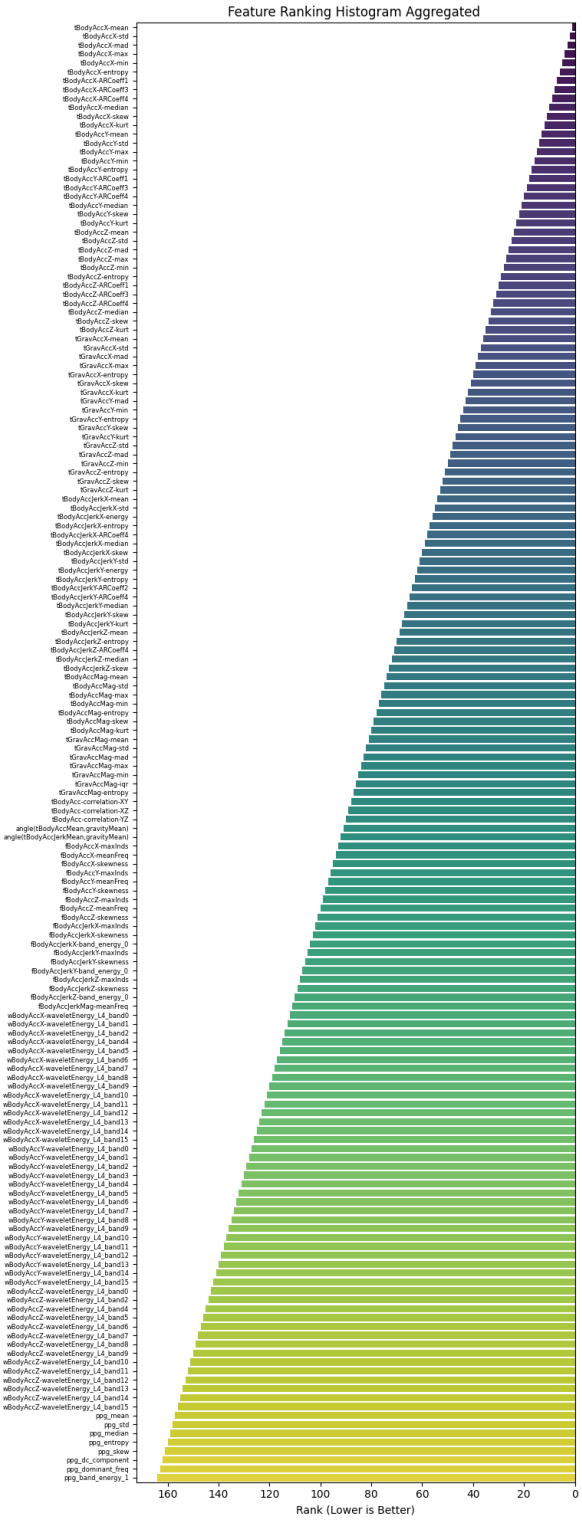


Figure A1. Aggregated result of feature importance determination for activity classification.

 doi:10.20944/preprints202603.0657.v1

Appendix B

Table A1. Physiological load distribution by intensity class per subject.

Subject ID	Gender	Age	BMI	Physical Fitness	Intensity Class	Below expected (%)	Within 1σ (%)	Between 1σ–2σ (%)	$\geq 2\sigma$ (%)
S1	Male	34	23.5	6	Low	18.0	32.0	33.8	16.2
					Medium	15.4	45.2	28.1	11.3
					High	14.3	49.6	28.1	8.0
S2	Male	28	22.4	5	Low	14.5	48.5	32.6	4.4
					Medium	16.4	46.0	28.8	8.8
					High	15.3	47.2	30.1	7.4
S3	Male	25	20.8	5	Low	12.8	35.2	33.3	18.7
					Medium	8.3	39.4	31.9	20.4
					High	12.6	52.7	26.6	8.1
S4	Male	25	20.2	5	Low	8.4	40.2	36.0	15.4
					Medium	12.6	45.3	28.7	13.5
					High	18.6	48.1	28.6	4.8
S5	Female	21	21.6	4	Low	0.0	36.0	51.5	12.5
					Medium	0.0	43.5	43.0	13.5
					High	0.0	44.0	44.0	12.0
S6	Female	37	22.6	1	Low	25.0	15.1	17.9	42.0
					Medium	16.2	43.2	29.7	10.9
					High	13.8	43.3	33.9	8.9
S7	Female	21	20.5	2	Low	24.1	54.2	21.7	0.0
					Medium	28.4	47.3	24.2	0.0
					High	28.6	46.9	24.4	0.0
S8	Male	43	21.8	5	Low	12.3	36.5	34.2	16.9
					Medium	17.1	22.7	22.7	37.5
					High	14.7	42.0	36.6	6.7
S9	Female	28	21.5	5	Low	15.9	44.5	32.2	7.5
					Medium	16.7	46.9	22.8	13.6
					High	15.6	48.9	31.6	3.9
S10	Female	55	20.8	5	Low	17.4	38.8	30.8	12.9
					Medium	26.0	11.5	11.5	51.0
					High	14.3	50.4	29.5	5.8
S11	Female	24	22	5	Low	13.2	61.4	24.6	0.9
					Medium	6.1	64.6	26.9	2.4
					High	12.0	61.3	24.0	2.7
S12	Male	43	27.6	5	Low	20.2	40.8	29.8	9.2
					Medium	22.2	24.5	24.5	28.7
					High	16.3	43.6	32.6	7.5
S13	Female	21	21.8	6	Low	25.8	52.0	22.2	0.0
					Medium	30.8	43.1	26.2	0.0
					High	25.4	48.8	25.8	0.0
S14	Female	26	23.2	4	Low	15.5	38.5	30.5	15.5
					Medium	12.3	55.3	27.2	5.3
					High	12.0	53.8	28.4	5.8
S15	Male	28	23.6	5	Low	25.9	46.0	25.9	2.1
					Medium	25.1	37.7	33.5	3.8
					High	8.4	46.7	30.8	14.0

References

1. Polak, A.G.; Klich, B.; Saganowski, S.; Prucnal, M.A.; Kazienko, P. Processing Photoplethysmograms Recorded by Smartwatches to Improve the Quality of Derived Pulse Rate Variability. *Sensors* **2022**, *22*. https://doi.org/10.3390/s22187047.
2. Kechris, C.; Dan, J.; Miranda, J.; Atienza, D. KID-PPG: Knowledge Informed Deep Learning for Extracting Heart Rate From a Smartwatch. *IEEE Transactions on Biomedical Engineering* **2025**, *72*, 870–877. https://doi.org/10.1109/tbme.2024.3477275.

 doi:10.20944/preprints202603.0657.v1

3. Kulsoom, F.; Narejo, S.; Mehmood, Z.; Chaudhry, H.N.; Butt, A.; Bashir, A.K. A review of machine learning-based human activity recognition for diverse applications. *Neural Computing and Applications* **2022**, *34*, 18289–18324. https://doi.org/10.1007/s00521-022-07665-9.
4. Pavliuk, O.; Mishchuk, M. Smartwatch-Based Human Staff Activity Classification: A Use-Case Study in Internal Logistics Systems Utilizing AGVs. In Proceedings of the 2024 IEEE International Conference on Big Data (BigData), 2024, pp. 8198–8207. https://doi.org/10.1109/BigData62323.2024.10825909.
5. Bennasar, M.; Price, B.A.; Gooch, D.; Bandara, A.K.; Nuseibeh, B. Significant Features for Human Activity Recognition Using Tri-Axial Accelerometers. *Sensors* **2022**, *22*. https://doi.org/10.3390/s22197482.
6. Attal, F.; Mohammed, S.; Dedabrishvili, M.; Chamroukhi, F.; Oukhellou, L.; Amirat, Y. Physical Human Activity Recognition Using Wearable Sensors. *Sensors* **2015**, *15*, 31314–31338. https://doi.org/10.3390/s151229858.
7. Nematallah, H.; Rajan, S.; Cretu, A.M. Logistic Model Tree for Human Activity Recognition Using Smartphone-Based Inertial Sensors. In Proceedings of the 2019 IEEE SENSORS, 2019, pp. 1–4. https://doi.org/10.1109/SENSORS43011.2019.8956951.
8. Hsu, Y.L.; Lin, S.L.; Chou, P.H.; Lai, H.C.; Chang, H.C.; Yang, S.C. Application of nonparametric weighted feature extraction for an inertial-signal-based human activity recognition system. In Proceedings of the 2017 International Conference on Applied System Innovation (ICASI), 2017, pp. 1718–1720. https://doi.org/10.1109/ICASI.2017.7988270.
9. Zhang, W.; Zhao, X.; Li, Z. A Comprehensive Study of Smartphone-Based Indoor Activity Recognition via Xgboost. *IEEE Access* **2019**, *7*, 80027–80042. https://doi.org/10.1109/ACCESS.2019.2922974.
10. Pavliuk, O.; Mishchuk, M. A novel Deep-Learning model for Human Activity Recognition based on Continuous Wavelet Transform. In Proceedings of the 5th International Conference on Informatics & Data-Driven Medicine (IDDM), 2022, Vol. 3302, *CEUR Workshop Proceedings*, p. 236–245. Available online: https://ceur-ws.org/Vol-3302/paper14.pdf.
11. Zhang, S.; Li, Y.; Zhang, S.; Shahabi, F.; Xia, S.; Deng, Y.; Alshurafa, N. Deep Learning in Human Activity Recognition with Wearable Sensors: A Review on Advances. *Sensors* **2022**, *22*. https://doi.org/10.3390/s22041476.
12. Lu, X.; Ling, Y.; Liu, S. Temporal Convolutional Network with Wavelet Transform for Fall Detection. *Journal of Sensors* **2022**, *2022*, 7267099. https://doi.org/10.1155/2022/7267099.
13. Abdel-Basset, M.; Hawash, H.; Chakrabortty, R.K.; Ryan, M.; Elhoseny, M.; Song, H. ST-DeepHAR: Deep Learning Model for Human Activity Recognition in IoHT Applications. *IEEE Internet of Things Journal* **2021**, *8*, 4969–4979. https://doi.org/10.1109/JIOT.2020.3033430.
14. Yuan, H.; Chan, S.; Creagh, A.P.; Tong, C.; Acquah, A.; Clifton, D.A.; Doherty, A. Self-supervised learning for human activity recognition using 700,000 person-days of wearable data. *npj Digital Medicine* **2024**, *7*, 91. https://doi.org/10.1038/s41746-024-01062-3.
15. Dirgová Luptáková, I.; Kubovčík, M.; Pospíchal, J. Wearable Sensor-Based Human Activity Recognition with Transformer Model. *Sensors* **2022**, *22*. https://doi.org/10.3390/s22051911.
16. Batool, S.; Khan, M.H.; Farid, M.S. An ensemble deep learning model for human activity analysis using wearable sensory data. *Applied Soft Computing* **2024**, *159*, 111599. https://doi.org/10.1016/j.asoc.2024.111599.
17. Pavliuk, O.; Mishchuk, M.; Strauss, C. Transfer Learning Approach for Human Activity Recognition Based on Continuous Wavelet Transform. *Algorithms* **2023**, *16*. https://doi.org/10.3390/a16020077.
18. Jia, Q.; Guo, J.; Yang, P.; Yang, Y. A holistic multi-source transfer learning approach using wearable sensors for personalized daily activity recognition. *Complex & Intelligent Systems* **2024**, *10*, 1459–1471. https://doi.org/10.1007/s40747-023-01218-w.
19. Shiri, F.M.; Perumal, T.; Mustapha, N.; Mohamed, R. Deep Learning and Federated Learning in Human Activity Recognition with Sensor Data: A Comprehensive Review. *Computer Modeling in Engineering & Sciences* **2025**, *145*, 1389–1485. https://doi.org/10.32604/cmes.2025.071858.
20. Veluguri, S.P. Deep PPG: Improving Heart Rate Estimates with Activity Prediction. In Proceedings of the 2025 1st International Conference on AIML-Applications for Engineering & Technology (ICAET), 2025, pp. 1–6. https://doi.org/10.1109/ICAET63349.2025.10932185.
21. Ortiz, B.L.; Gupta, V.; Kumar, R.; Jalin, A.; Cao, X.; Ziegenbein, C.; Singhal, A.; Tewari, M.; Choi, S.W. Data Preprocessing Techniques for AI and Machine Learning Readiness: Scoping Review of Wearable Sensor Data in Cancer Care. *JMIR Mhealth Uhealth* **2024**, *12*, e59587. https://doi.org/10.2196/59587.

 doi:10.20944/preprints202603.0657.v1

22. Charlton, P.H.; Kotzen, K.; Mejía-Mejía, E.; Aston, P.J.; Budidha, K.; Mant, J.; Pettit, C.; Behar, J.A.; Kyriacou, P.A. Detecting beats in the photoplethysmogram: benchmarking open-source algorithms. *Physiological Measurement* **2022**, *43*. https://doi.org/10.1088/1361-6579/ac826d.
23. Allen, J.; Murray, A. Age-related changes in the characteristics of the photoplethysmographic pulse shape at various body sites. *Physiological Measurement* **2003**, *24*, 297–307. https://doi.org/10.1088/0967-3334/24/2/306.
24. Moscato, S.; Palmerini, L.; Palumbo, P.; Chiari, L. Quality Assessment and Morphological Analysis of Photoplethysmography in Daily Life. *Frontiers in Digital Health* **2022**, *4*. https://doi.org/10.3389/fdgth.2022.912353.
25. Puranen, A.; Halkola, T.; Kirkeby, O.; Vehkaoja, A. Effect of skin tone and activity on the performance of wrist-worn optical beat-to-beat heart rate monitoring. In Proceedings of the 2020 IEEE SENSORS, 2020, pp. 1–4. https://doi.org/10.1109/SENSORS47125.2020.9278523.
26. Liang, Y.; Elgendi, M.; Chen, Z.; Ward, R. An optimal filter for short photoplethysmogram signals. *Scientific Data* **2018**, *5*, 180076. https://doi.org/10.1038/sdata.2018.76.
27. Mejía-Mejía, E.; Allen, J.; Budidha, K.; El-Hajj, C.; Kyriacou, P.A.; Charlton, P.H. Photoplethysmography signal processing and synthesis. In *Photoplethysmography*; Academic Press, 2022; pp. 69–146. https://doi.org/10.1016/B978-0-12-823374-0.00015-3.
28. Shuvo, S.B.; Alam, S.S.; Ayman, S.U.; Chakma, A.; Salvi, M.; Seoni, S.; Barua, P.D.; Molinari, F.; Acharya, U.R. Application of Wavelet Transformation and Artificial Intelligence Techniques in Healthcare: A Systemic Review. *WIREs Data Mining and Knowledge Discovery* **2025**, *15*, e70007. https://doi.org/10.1002/widm.70007.
29. Merino-Monge, M.; Castro-García, J.A.; Lebrato-Vázquez, C.; Gómez-González, I.M.; Molina-Cantero, A.J. Heartbeat detector from ECG and PPG signals based on wavelet transform and upper envelopes. *Physical and Engineering Sciences in Medicine* **2023**, *46*, 597–608. https://doi.org/10.1007/s13246-023-01235-6.
30. Sepúlveda-Cano, L.M.; Gil, E.; Laguna, P.; Castellanos-Dominguez, G. Sleep apnoea detection in children using PPG envelope-based dynamic features. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2011, Vol. 2011, pp. 1483–1486. https://doi.org/10.1109/IEMBS.2011.6090362.
31. Zong, W.; Heldt, T.; Moody, G.; Mark, R. An open-source algorithm to detect onset of arterial blood pressure pulses. In Proceedings of the Computers in Cardiology, 2003, 2003, pp. 259–262. https://doi.org/10.1109/CIC.2003.1291140.
32. Mejía-Mejía, E.; May, J.M.; Torres, R.; Kyriacou, P.A. Pulse rate variability in cardiovascular health: a review on its applications and relationship with heart rate variability. *Physiological Measurement* **2020**, *41*, 07TR01. https://doi.org/10.1088/1361-6579/ab998c.
33. Charlton, P.H.; Allen, J.; Bailón, R.; Baker, S.; Behar, J.A.; Chen, F.; Clifford, G.D.; Clifton, D.A.; Davies, H.J.; Ding, C.; et al. The 2023 wearable photoplethysmography roadmap. *Physiological Measurement* **2023**, *44*, 111001. https://doi.org/10.1088/1361-6579/acead2.
34. Reiss, A.; Indlekofer, I.; Schmidt, P.; Van Laerhoven, K. Deep PPG: Large-Scale Heart Rate Estimation with Convolutional Neural Networks. *Sensors* **2019**, *19*. https://doi.org/10.3390/s19143079.
35. Reiss, A.; Indlekofer, I.; Schmidt, P. PPG-DaLiA, 2019. https://doi.org/10.24432/C53890.
36. van Hees, V.T.; Gorzelniak, L.; Dean León, E.C.; Eder, M.; Pias, M.; Taherian, S.; Ekelund, U.; Renström, F.; Franks, P.W.; Horsch, A.; et al. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. *PLOS ONE* **2013**, *8*, e61691. https://doi.org/10.1371/journal.pone.0061691.
37. Vähä-Ypyä, H.; Vasankari, T.; Husu, P.; Suni, J.; Sievänen, H. A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. *Clinical Physiology and Functional Imaging* **2015**, *35*, 64–70. https://doi.org/10.1111/cpf.12127.
38. John, D.; Tang, Q.; Albinali, F.; Intille, S. An Open-Source Monitor-Independent Movement Summary for Accelerometer Data Processing. *Journal for the Measurement of Physical Behaviour* **2019**, *2*, 268–281. https://doi.org/10.1123/jmpb.2018-0068.
39. Bai, J.; He, B.; Shou, H.; Zipunnikov, V.; Glass, T.A.; Crainiceanu, C.M. Normalization and extraction of interpretable metrics from raw accelerometry data. *Biostatistics* **2014**, *15*, 102–116. https://doi.org/10.1093/biostatistics/kxt029.
40. Karas, M.; Muschelli, J.; Leroux, A.; Urbanek, J.K.; Wanigatunga, A.A.; Bai, J.; Crainiceanu, C.M.; Schrack, J.A. Comparison of Accelerometry-Based Measures of Physical Activity: Retrospective Observational Data Analysis Study. *JMIR mHealth and uHealth* **2022**, *10*, e38077. https://doi.org/10.2196/38077.

 doi:10.20944/preprints202603.0657.v1

41. Jiang, E.; Nie, B.; Cao, Z.; Yu, Z.; Li, S.; Lu, Y.; Yu, C.; Yin, N. Non-Invasive Blood Pressure Estimation Using Multi-Domain Pulse Wave Features and Random Forest Regression. *Electronics* **2025**, *14*. https://doi.org/10.3390/electronics14071409.
42. Su, X.; Wang, X.; Ge, H. Exercise ECG Classification Based on Novel R-Peak Detection Using BILSTM-CNN and Multi-Feature Fusion Method. *Electronics* **2025**, *14*. https://doi.org/10.3390/electronics14020281.
43. Manjarres, J.; Narvaez, P.; Gasser, K.; Percybrooks, W.; Pardo, M. Physical Workload Tracking Using Human Activity Recognition with Wearable Devices. *Sensors* **2020**, *20*. https://doi.org/10.3390/s20010039.
44. Chen, W.; Yi, Z.; Lim, L.J.R.; Lim, R.Q.R.; Zhang, A.; Qian, Z.; Huang, J.; He, J.; Liu, B. Deep learning and remote photoplethysmography powered advancements in contactless physiological measurement. *Frontiers in Bioengineering and Biotechnology* **2024**, *12*. https://doi.org/10.3389/fbioe.2024.1420100.
45. Laval, X.; Mailhes, C.; Martin, N.; Bellemain, P.; Pachaud, C. Amplitude and phase interaction in Hilbert demodulation of vibration signals: Natural gear wear modeling and time tracking for condition monitoring. *Mechanical Systems and Signal Processing* **2021**, *150*, 107321. https://doi.org/10.1016/j.ymssp.2020.107321.
46. Tian, Y.; Wang, X.; Yang, P.; Wang, J.; Zhang, J. A Single Accelerometer-based Robust Human Activity Recognition via Wavelet Features and Ensemble Feature Selection. In Proceedings of the 2018 24th International Conference on Automation and Computing (ICAC), 2018, pp. 1–6. https://doi.org/10.23919/IConAC.2018.8749005.
47. Abid, M.H.; Nahid, A.A.; Islam, M.R.; Parvez Mahmud, M.A. Human Activity Recognition Based on Wavelet-Based Features along with Feature Prioritization. In Proceedings of the 2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA), 2021, pp. 933–939. https://doi.org/10.1109/ICCCA52192.2021.9666294.
48. Fox, S.M.; Naughton, J.P. Physical activity and the prevention of coronary heart disease. *Preventive Medicine* **1972**, *1*, 92–120. https://doi.org/10.1016/0091-7435(72)90079-5.
49. Tanaka, H.; Monahan, K.D.; Seals, D.R. Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology* **2001**, *37*, 153–156. https://doi.org/10.1016/s0735-1097(00)01054-8.
50. Gulati, M.; Shaw, L.J.; Thisted, R.A.; Black, H.R.; Bairey Merz, C.N.; Arnsdorf, M.F. Heart rate response to exercise stress testing in asymptomatic women: the St. James women take heart project. *Circulation* **2010**, *122*, 130–137. https://doi.org/10.1161/CIRCULATIONAHA.110.939249.
51. Lach, J.; Wiecha, S.; Śliż, D.; Price, S.; Zaborski, M.; Cieśliński, I.; Postuła, M.; Knechtle, B.; Mamcarz, A. HR Max Prediction Based on Age, Body Composition, Fitness Level, Testing Modality and Sex in Physically Active Population. *Frontiers in Physiology* **2021**, *12*, 695950. https://doi.org/10.3389/fphys.2021.695950.
52. Garber, C.E.; Blissmer, B.; Deschenes, M.R.; Franklin, B.A.; Lamonte, M.J.; Lee, I.M.; Nieman, D.C.; Swain, D.P. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise. *Medicine & Science in Sports & Exercise* **2011**, *43*, 1334–1359. https://doi.org/10.1249/MSS.0b013e318213fefb.
53. Karvonen, M.J.; Kentala, E.; Mustala, O. The effects of training on heart rate; a longitudinal study. *Annales medicinae experimentalis et biologiae Fenniae* **1957**, *35*, 307–315. PMID: 13470504.