ARTIFICIAL INTELLIGENCE FOR EGG/PPG SIGNAL PROCESSING ON MOBILE PLATFORMS

Galya Georgieva-Tsaneva
Institute of Robotics, Bulgarian Academy of Sciences, Bulgaria
galitsaneva@abv.bg

Krasimir Cheshmedzhiev
Institute of Robotics, Bulgarian Academy of Sciences, Bulgaria
cheshmedzhiev@gmail.com

ИЗКУСТВЕН ИНТЕЛЕКТ ЗА ОБРАБОТКА НА ЕКГ/ФПГ СИГНАЛИ НА МОБИЛНИ ПЛАТФОРМИ

Abstract

ECG/PPG signal processing is a cornerstone of modern cardiovascular diagnostics. While artificial intelligence has already enhanced ECG analysis through accurate detection, classification, and prediction of cardiac events, its integration into mobile platforms enables continuous, ubiquitous monitoring. This paper introduces a novel framework that couples state-of-the-art AI methodologies with the Digital Twin paradigm to create a personalized, real-time virtual replica of a patient's cardiac function. We survey deep learning and hybrid wavelet–neural approaches for QRS complex detection, arrhythmia classification, and heartbeat segmentation, and propose methods for incremental on-device learning to address data imbalance and inter-subject variability. Annotated datasets such as MIT-BIH are extended with synthetic augmentation to populate and calibrate the digital twin models, enabling generalization across heterogeneous populations. The proposed architecture emphasizes low-latency inference, energy-aware computation, and secure data flows suitable for mobile and wearable devices. By embedding interpretability layers and adaptive feedback loops, the system closes the gap between passive ECG monitoring and actionable, individualized cardiac care. Our results demonstrate that AI-driven ECG digital twins can significantly outperform traditional algorithms in accuracy and adaptability, filling a critical scientific gap and opening new pathways for predictive, preventive, and personalized cardiovascular healthcare.

Keywords: *Artificial Intelligence; ECG; PPG; Digital Twin; Mobile and Wearable Platforms; Deep Learning; Wavelet–Neural Networks; Incremental On-Device Learning; Interpretability; Personalized Cardiac Monitoring.*

INTRODUCTION

Electrocardiogram (ECG) and photoplethysmographic (PPG) signal processing remains the foundation of modern cardiovascular diagnostics, providing noninvasive insight into cardiac electrophysiology. Conventional algorithms for QRS complex detection, arrhythmia classification, and heart rhythm segmentation have achieved very good performance in offline conditions, but struggle to maintain accuracy, stability, and adaptability when implemented in real mobile or wearable environments.

In parallel, artificial intelligence (AI) – and in particular deep learning – is bringing a number of improvements to the analysis of complex biomedical signals. AI models can automatically extract discriminative features from raw ECG/EPG data, significantly better than classical technologies, and enable more accurate detection and prediction of cardiac events. However, challenges remain: high inter-individual variability, sensitivity to noise, limited annotated data, power constraints of peripheral devices, and limited interpretability of deep models. Recently, the concept of the "Digital Twin" has begun to emerge as a promising approach for personalized healthcare. A digital twin is a virtual, continuously updated representation of a given physical system (in this case, an individual's heart), powered by real-time sensor data and capable of simulating future states. Combining AI-enhanced cardiac analytics (ECG/FPG/HRV analysis) with a digital twin architecture on mobile platforms offers a path to predictive, preventive, and personalized cardiac care that goes beyond passive monitoring.

This paper proposes and evaluates a novel AI-driven ECG/FPG signal processing framework for mobile and wearable devices designed to build a user-specific digital twin of cardiac function. The aim of the paper is to review the fundamentals of the HRV digital twin concept, define it, and present a novel index for assessing fatigue in athletes.

HISTORY AND RELATED WORK

1. Traditional methods for processing ECG/FPG signals

Classical methods for ECG/FG analysis are based on filtering, morphological analysis, peaking and thresholding and manually defined features such as RR/PP interval length, QRS width, etc. These algorithms often work well in clean recordings, but lose accuracy in the presence of noise, moving artifacts, and intersubject variability.

Articles in the scientific literature include methods such as the Pan–Tompkins algorithm for QRS detection, adaptive thresholding, wavelet-based approaches, etc.

2. AI and Deep Learning in ECG Analysis

With the advent of machine learning and neural networks in cardiac diagnostics, numerous studies have emerged that use convolutional neural networks (CNN), recurrent neural networks (RNN), transformer architectures, and hybrid models for tasks such as arrhythmia classification, automated ECG waveform segmentation, and acute event prediction. In addition, the concept of a digital twin is increasingly integrated with artificial intelligence to build personalized, prognostic, and adaptive models.

A review by Chaparro-Cárdenas et al. examines the current state of digital twins and AI for personalized and predictive medicine, with a focus on the potential for individualized treatment [1]. Fuse et al. present a systematic review of the applications of large language models, foundation models, and digital twins for clinical analysis and allergology [2]. Kreuzer et al. focused on the use of artificial intelligence to build and simulate digital twins, systematizing methodological and application aspects [3]. In another large-scale study, Qian et al. demonstrated how machine learning allows the creation of populations of cardiac digital twins based on imaging and ECG data to assess electrophysiological processes such as conduction and repolarization [4]. Grandits et al. created an efficient digital twin of the ventricular conduction system using 12-lead ECG and visualization data, highlighting the possibilities for the identification of physiological parameters [5].

3. Concept and Implementation of Digital Twin in Healthcare

The concept of Digital Twin originates from industrial engineering and is widely used in manufacturing processes, monitoring and predictive control. However, in the context of healthcare, it has taken on a new meaning – the creation of a virtual copy of an individual's physiological state, which is fed with real-world data and used for diagnosis, monitoring, prediction and personalized treatment.

In recent years, a number of review and applied studies have presented different aspects of digital twins in cardiology and medicine.

Thangaraj et al. discuss how digital twins are combined with generative artificial intelligence to transform cardiovascular care through simulation and predictive models [6]. Coorey et al. review the interdisciplinary progress in creating health digital twins for cardiology applications, highlighting key challenges and prospects [7]. Sel et al. propose a comprehensive architecture for building a digital twin for cardiac health, including modeling, data synchronization, and clinical interpretation [8]. Kabir et al. examine the integration between digital twins and the Internet of Things (Healthcare IoT), analyzing the technological and security aspects of the implementation [9]. Bhagirath et al. emphasize the potential of digital twins to revolutionize cardiac electrophysiology through simulation models compatible with real ECG data [10].

Despite these significant contributions, most implementations remain conceptual or offline, lacking full-fledged real-time solutions based on data from wearable devices and with the ability for dynamic adaptation and feedback. This deficiency creates a need to develop a lightweight, fractal-adaptive, and AI-supported architecture capable of operating on edge devices with low latency and high predictive value.

4. Problems and Limitations in Mobile/edge Deployment

Implementing Digital Twins and AI models in mobile or edge devices is a key step towards realizing personalized and continuous healthcare. However, the transfer from cloud or server environments to on-premises devices brings with it a number of technical and ethical challenges:

- Limited resources – processing power, RAM, battery capacity and cooling are critical factors in edge implementation.
- Low latency and real-time requirements – the digital twin must process and respond to physiological changes almost instantaneously.
- Data privacy and security – the need to protect sensitive medical information (e.g. ECG, PPG, RR intervals).
- Inter-subject variability and domain adaptation – systems must be personalized to the physiology of each individual patient.
- Interpretability and trust – healthcare professionals require explainable predictions that can be clinically substantiated.

Johnson and Saikia explore how wearable devices such as watches and bracelets can serve as entry points for digital twins, highlighting challenges with data processing and real-time synchronization [11]. Volkov et al. conduct an extensive review of existing platforms that unify IoT, Digital Twin, and mobile medicine, identifying the lack of integrated standards and modular architectures as a major obstacle to mass adoption [12]. Chen and colleagues present the concept of a Human Digital Twin powered by Mobile AIGC (AI-generated Content) and demonstrate how generative models can be used to make personalized predictions on resource-constrained devices [13].

These developments (Table 1) show significant potential for personalized edge solutions, but emphasize the need for optimized models, input stream compression, and prioritized security, especially in the context of continuous monitoring and real-time alarm mechanisms.

Table 1. Overview of what has been done to date on the VSC digital twin concept

Area / Subtask	Examples of existing developments/research	Main contribution	Limitations/gaps (in the context of Digital Twin)
Peak detection / QRS / wave segmentation	Grandits et al. (2025) [5]	High sensitivity and accuracy in QRS (≈99.6%)	Works with recorded data; not real-time on wearables
Noise/artifact reduction	Edder (2025) [14]; Gaoudam (2025) [15]	Transformer-based noise removal; preserved morphology	Offline processing; lack of optimization for edge devices
HRV analysis / state differentiation	Johnson et al. (2024) [11]; Gaoudam (2025) [15]	AI-mining SDNN, RMSSD and state classification	Lack of a dynamic digital twin framework with these metrics
Cardio data prediction/modeling	Grandits et al. (2025) [5]; Qian et al. (2025) [4]; Bhagirath et al. (2024) [10]	Creating digital twins from ECG and imaging diagnostics	Not integrated with edge/mobility solutions
Calibration / uncertainty analysis	Fuse et al. (2025) [2]	Assessment of anatomical variations and uncertainties in models	Lack of real dynamic adaptation of the twin in real time
IoT / connectivity / data protection	Kabir et al. (2025) [9]; Volkov et al. (2021) [12]; Chen et al. (2024) [13]	Demonstration of IoT platforms with cloud analytics and AI for wearable devices	Lack of two-way synchronization and adaptive learning

Methodology of the Proposed Conceptual Framework of a Digital Twin of a VCH, Managed by Artificial Intelligence

1. System Overview

The proposed architecture integrates a mobile or wearable device for ECG/FPG acquisition, on-device preprocessing, AI-based feature extraction and classification, and a cloud-synchronized digital twin of the heart. Each patient has a virtual replica that is continuously updated with incoming ECG/FPG streams and model outputs. The mobile device offers low-latency inference and transmits only essential parameters and/or anonymized features to the digital twin, thereby reducing bandwidth and preserving privacy.

2. Data Collection and Synchronization from Mobile/wearable Devices

ECG/EPG signals are collected using single or multi-input wearable devices (smartwatches, chest straps, custom IoT patches, or purpose-built cardio devices) at a manufacturer-defined sampling rate. The device's pre-processing modules perform noise reduction (baseline removal, bandpass filtering), determination of normal cardiac intervals (RR/PP), and segmentation of the time series before feeding the data to the AI model.

3. Cardiology Data Modeling

Modeling cardiac data using statistical approaches is often based on the use of Gaussian (normal) distributions, which describe the variation of physiological parameters around their mean value. In the analysis of heart rate variability (HRV), Gaussian models allow for a quantitative assessment of the dispersion of RR intervals, the dynamics of the mean heart rate, and parameters such as standard deviation or confidence intervals. In Bayesian probabilistic modeling, normal distributions are often used to calculate conditional probabilities $(X \mid C\ i\ )$, which describe the belonging of the observed data to a given physiological state (rest, fatigue, stress). More complex approaches include the use of Gaussian Mixture Models (GMM), which allow for the description of multimodal distributions characteristic of cardiac data with the presence of arrhythmias, noise, or transitions between states. Although Gaussian models provide a good statistical basis and allow reliable hypothesis derivation, their limitation is manifested in data with nonlinear or chaotic dynamics, where supplementation with attractor and nonlinear analysis methods is necessary.

4. Hybrid Wavelet-neural Networks for Noise-resistant ECG Feature Extraction

To improve robustness against real-world noise and motion artifacts, the framework uses hybrid architectures combining discrete wavelet transforms (DWT) with convolutional and recurrent neural networks. The wavelet stage decomposes ECG signals into multi-resolution components, which are then fed to CNN or CNN-LSTM feature training blocks. This approach captures both time-frequency structure and higher-level patterns, improving QRS complex detection, arrhythmia classification, and heart rate segmentation on mobile hardware.

5. Personalization Training

A key innovation of the framework is on-device learning, which allows the AI model to adapt to individual patient characteristics without the need for complete retraining in the cloud. Fine-tuning or transfer learning modules update the model parameters as new labeled events occur, compensating for inter-subject variability and class imbalance, while respecting the energy and storage constraints of the mobile device.

6. Layers for Interpretation and User Feedback

To build trust and support critical decision-making, the digital twin includes interpretable AI layers (e.g., attention heat maps, relevance scores) and generates textual explanations of the model's results. Users (including study subjects, athlete coaches, physicians) can review detected anomalies or predicted risks, validate or correct them through a secure dashboard, and these corrections provide feedback to improve the personalized model. This closed feedback loop transforms passive observation into actionable, adaptive care.

7. Security and Privacy Considerations in Edge Computing

Given the sensitivity of cardiac data, the framework implements end-to-end encryption, anonymizes transmitted features, and provides secure device authentication. Federated or split learning can be used to train global models without sharing raw cardiac signals. Data retention policies meet GDPR/HIPAA guidelines, and lightweight cryptographic protocols minimize latency and energy costs on mobile platforms.

Dataset Preparation and Model Calibration

1. Using Annotated Datasets

To train and evaluate the proposed AI-driven digital twin framework, annotated ECG/FPG databases, such as the MIT-BIH Arrhythmia Database, the PTB Diagnostic ECG Database, and the PhysioNet Challenge datasets, as well as our own datasets collected by the author team for the needs of implementing the tasks of projects at the Bulgarian National Science Foundation, can be used. These repositories provide beat- and rhythm-level annotations necessary for supervised training of peak detectors, arrhythmia classifiers, and heart rate segmentation models.

2. Data Preprocessing and Segmentation

All raw ECG/FPG recordings undergo standardized preprocessing steps, including baseline removal, bandpass filtering (0.5–40 Hz), and resampling to a predefined frequency. Signals are segmented into fixed-length windows or individual heartbeats using annotation markers. This unified processing ensures comparability across heterogeneous sources.

3. Synthetic Data Extension for Digital Twin Calibration

To address class imbalance and further enrich the diversity of training data, synthetic ECG/FPG/HR signals and noise are generated. Techniques include waveform morphing, noise addition (to test the robustness of algorithms on noisy data), simulated motion artifacts, and GAN-based synthesis of rare arrhythmias. These augmented signals are used to calibrate and personalize digital twin models, improving generalization to unseen patient profiles.

4. Working with Intersubject Variability

The methodology includes cross-subject validation and incremental fine-tuning of the device to adapt global models to individual users. This approach allows the mobile AI module to learn patient-specific morphology without retraining from scratch, while leveraging a large shared cardiac dataset.

5. Metrics for Model Calibration and Evaluation

Calibration curves, reliability plots, and Brier scores can be used to evaluate the probabilistic outputs of the classifiers. Power consumption and latency metrics are measured directly on the target mobile hardware to ensure the feasibility of real-time implementation. Model interpretability is assessed through attention map visualization and feature relevance scores.

The Heart Rate Variability Digital Twin Concept (HRV Digital Twin)

1. Definition

The HRV Digital Twin is a personalized, dynamically updated virtual model that reflects the dynamics of heart rate variability of a specific individual in real time, which is trained and operated in real time and can be implemented on mobile and wearable platforms. It uses a continuous stream of RR intervals, ECG/PPG-derived cardio intervals, HRV metrics (SDNN, RMSSD, LF/HF, DFA α, entropy measures, etc.) and additional activity and sleep data from mobile wearable devices. Unlike classical cardiology digital twins that focus on anatomy, electrophysiology and hemodynamics, this twin has autonomic regulation, expressed through HRV, as its core. The proposed framework defines HRV as a basic state variable and develops indices reflecting the dynamics of fatigue, stress and recovery.

2. Architecture

Main components:

- Data from wearable devices (ECG patches, smartwatches, PPG sensors) – RR intervals, HRV indicators, pulse signal amplitude, activity, sleep.
- Personal baseline – individual HRV values (SDNN, RMSSD, LF/HF, fractal/entropy indicators), determined at rest and under different loads.
- Modeling and prediction module – hybrid AI model (wavelet + CNN-LSTM or other interpretable model) that learns HRV dynamics and predicts future states.
- State indices – Fatigue Digital Twin Index (FDTI) for real-time fatigue/stress assessment and Recovery Digital Twin Index (RDTI) for assessment of recovery rate.
- Personalization – automatic adaptation of model parameters to new data from a specific person (few-shot / on-device learning).
- Visualization and interface – graphs and alarms for athletes, doctors or users.

3. Functions and Application

- Instant assessment – shows the current autonomic workload (FDTI).
- Recovery prediction – predicts when HRV will return to baseline values (RDTI).
- "What if" simulations – the twin can predict the effect of workload, stress or intervention on HRV.
- Sports and rehabilitation – supports planning of training cycles, monitoring of overtraining, cardiac rehabilitation.
- Research tool – provides a basis for analysis of individual and population models of autonomic regulation.

4. Scientific Novelty and Contribution

For the first time, HRV is defined as the central foundation of a digital twin, rather than a peripheral biomarker. The current framework integrates a new index (FDTI) for real-time fatigue/stress assessment and provides a methodological basis for developing a second index (RDTI) for recovery. The HRV Digital Twin is designed for mobile and wearable platforms, with personalization and interpretability, and can be extended to predictive modeling and recommendations, which creates a new research line and contributes to personalized medicine and sports science.

The diagram in Figure 1 presents the data flow and main modules of the proposed heart rate variability digital twin (HRV Digital Twin). On the left are the data sources – mobile and wearable devices (ECG patches, smartwatches, PPG sensors) that continuously collect RR intervals, HRV metrics, pulse signal amplitude, activity, and sleep. The data goes through a pre-processing stage (filtering, normalization, synchronization) and a hybrid module that extracts noise-resistant time-frequency and dynamic characteristics.

The next block is the digital twin module, which stores the individual's personal baseline (individual HRV values at rest and during exercise), adapts through on-device learning and generates states and predictions. In this module, the FDTI (Fatigue Digital Twin Index) indices are calculated – for a momentary assessment of fatigue/stress – and the RDTI (Recovery Digital Twin Index) – for an assessment of the recovery rate. The resulting indices and predictions are visualized in an interface for the user/physician (right part of the diagram) and can be used for a momentary assessment, personalized exercise planning and "what if" simulations.

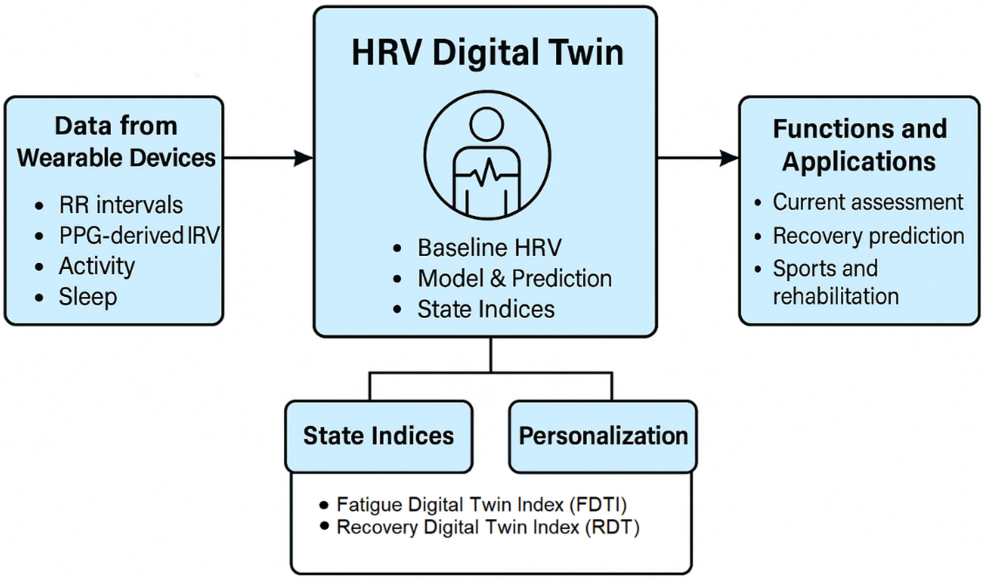


Figure 1. Architecture of the proposed digital twin of heart rate variability (HRV Digital Twin)

A general presentation of the two new indices is given in Table 2.

Table 2. Tabular summary of the two indices.

Index	Time scale	Main entrances	Expected direction
FDTI (Fatigue)	0–2 h after loading (windows 60–120 s)	↓SDNN, ↓RMSSD, ↑LF/HF, ↑DFA α_1 (or specific change), ↓SampEn, ↑ΔHR, ↓PPG amplitude, ↑PPG variability, ↓PTT	Higher FDTI ⇒ higher acute fatigue
RDTI (Recovery)	2–24 h (units of 5–30 min + trajectories)	Trend towards baseline: ↑SDNN, ↑RMSSD, LF/HF → rest zone, SampEn/MSE ↑, PPG amplitude stabilization, PTT → baseline; night RMSSD ↑	Higher RDTI ⇒ more complete recovery

Definition of the Fatigue Digital Twin Index

To provide a quantitative and personalized assessment of fatigue and stress in real time, the Fatigue Digital Twin Index (FDTI) has been developed within the proposed HRV Digital Twin. The index is based on a combination of selected features of the heart rate variability and pulse signal, extracted after the hybrid wavelet-neural processing stage.

Equation: FDTI is defined as a weighted linear combination of K features f_k (t), which are standardized against the individual baseline of the respective individual:

$$\mathrm{FDTI(t)} = = \sum_{k=1}^{K} w_k \frac{f_k(t)-\mu_k}{\sigma_k}, \quad (1)$$

where:

– f_k (t) are the time series of the selected HRV features (SDNN, RMSSD, LF/HF, DFA α, entropy metrics, wavelet coefficient energies, etc.);
– μ_k and σ_k are the mean and standard deviation of the feature for the respective individual at rest (personal baseline);
– w_k are the weights determined by training on a reference cohort (e.g. a group of athletes with labels "rest", "fatigue", "stress") and subsequently adapted to the specific individual through on-device incremental learning.

Interpretation: A higher FDTI value reflects increased autonomic load (fatigue/stress), and a lower value reflects a closer physiological state. The index can be calculated beat-by-beat on the mobile hardware, updating with each new measurement of RR intervals/PPG.

Advantages: This formulation allows integration of classic HRV metrics and new wave/entropy features into a single metric; personalization through individual normalization and weight adaptation; use for alarms, visualization, and prediction within the HRV Digital Twin.

JUSTIFICATION FOR THE DETERMINATION OF THE NEW INDEX

1. Selecting the parameters

The following parameters were selected for inclusion in the post-exercise fatigue index: SDNN, RMSSD, LF/HF, SD1, SD2/SD1, DFA α_1, SampEn. The rationale for the selection is presented in Table 3.

Table 3. Parameter and its physiological significance

Parameter	Physiological description
RMSSD	Sharp decline → acutely suppressed parasympathetic
SD1	Related to RMSSD, also falls with fatigue
LF/HF	Sharply increases (↑ LF, ↓ HF) → sympathetic dominance
SampEn	Declines → reduced regulatory complexity
DFA α1	Deviates above 1.2 → "loss" of fractal structure
HR (bpm)	Increases → tachycardia
SD2/SD1	Increases during acute stress

2. Orienting the signs so that their increase corresponds to the accumulation of fatigue

The used form of inclusion of the features through the variables x_i is shown in formulas (2) to (8) and Table 4.

- $x_1 = \frac{1}{SDNN}$ (2)
- $x_2 = \frac{1}{RMSSD}$ (3)
- $x_3 = LF/HF$ (4)

- $x_4 = \frac{1}{SD1}$ (5)
- $x_5 = SD2/SD1$ (6)
- $x_6 = DFA_{\alpha 1}$ (7)
- $x_7 = \frac{1}{SampEn}$ (8)

Table 4. How to enable parameters

k	Character	Transformation	Physiological significance at ↑
x_1	SDNN	1/SDNN	↓ variability
x_2	RMSSD	1/RMSSD	↓ parasympathetic tone
x_3	LF/HF	LF/HF	↑ sympathetic dominance
x_4	SD1	1/SD1	↓ rapid variability
x_5	SD2/SD1	SD2/SD1	↑ instability
x_6	DFA α_1	DFA α_1	loss of fractal structure
x_7	SampEn	1/SampEn	↓ complexity

3. Normalization (Pre-training)

For each metric xk and state s (Post or +2h):

$$z_{k,s} = \frac{x_{k,s}}{x_{k,Pre}} - 1 \quad (9)$$

This is the relative change from "before training" (0 means no change; positive – more fatigue).

4. FDTI Index

With equal weight of the individual features, we add up the normalized features and take the average:

$$\text{FDTI(t)} = \sum_{k=1}^{K} \frac{1}{\text{K}} \frac{f_k(t) - \mu_k}{\sigma_k}, \qquad \text{K - number of parameters} \quad (10)$$

The weights wk can be trained (e.g. with logistic regression or Linear Discriminant Analysis), the formula becomes:

$$\text{FDTI(t)} = \sum_{k=1}^{K} w_k \cdot z_{k,s} \text{ ,} \sum w_k = 1. \quad (11)$$

Personalization of the FDTI Formula in the Context of a Digital Twin

1. Individual baseline

- Normalization itself $\frac{x_{k,s}}{x_{k,Pre}} - 1$ makes the index personalized: each athlete has their own "zero" line.

- With regular use, the digital twin can update baseline values dynamically (periodic training, different conditions).

2. Adaptive weights w_k

- Instead of the same w_k for everyone, individual weights can be trained on the historical data of the particular athlete.
- Example: for one athlete RMSSD is very sensitive to fatigue, for another – LF/HF; the system automatically increases the weight of the most informative indicators for him.
- Mathematical form:

$$FDTI^{\,i}\,(t) = \sum_{k=1}^{K} w_k^{\,i} \,.\, z_{k,s}^{\,i} \qquad (12)$$

Where w_k^i are personal weights of *i-th* individual and $z_{(k,\ s)}^i$ – normalized signs of *i–th* individual.

3. Incremental (on-device) learning

- The digital twin can use new records to update w_k^i – using online learning.
- Thus, the index gradually "learns" the reaction of the specific person.

4. Calibration with external markers

- If subjective fatigue scales (RPE) or biochemical markers (lactate) are created, the system can use these "labels" for further training and relate the HRV profile to real fatigue.

5. Adaptive alarm thresholds

- In addition to the formula, the thresholds for "fatigue" can also be individual. For each athlete, the digital twin calculates a personal "green/yellow/red zone" of FDTI.

In practice, this means that the formula is not "static", but a living part of the digital twin, which learns from each user's data and after a certain number of iterations adapts to their specific individuality.

Personalization of wk weights for different groups of athletes. The wk weights in the FDTI index are initially calibrated on a reference training cohort (e.g. athletes from a specific sport), optimizing the ability to distinguish between rest/fatigue/stress states. When applying the index to a new group of athletes, the global weights serve as an initial value, but the system allows for customization by quickly recalculating or fine-tuning wk with a small number of labeled data from the respective group. This approach preserves the general structure of the index but increases its accuracy in the presence of inter-subject differences, thus FDTI adapts to the specific physiological profiles of different populations and supports the construction of individualized digital twins.

RESULTS

This study used the weights presented in Table 5, obtained through correlation analysis and Table 6 shows the calculated FDTI indices for each of the studied wrestlers immediately after training and 2 hours later.

Table 5. Parameter weights in the index

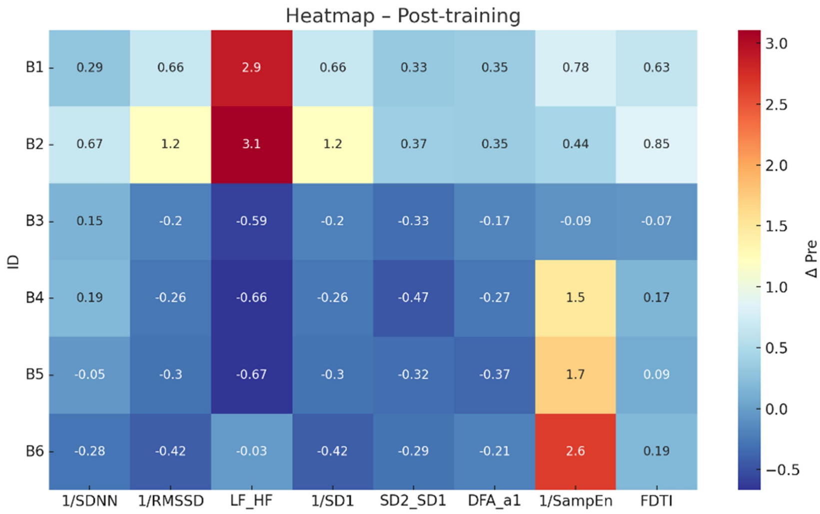
Parameter	Rationale for weight assignment
SDNN	Decreases with fatigue → use 1/SDNN ↑ → moderate weight (0.14)
RMSSD	Decreases with fatigue → use 1/RMSSD ↑ → moderate weight (0.14)
LF/HF	Increases under stress → use LF/HF directly → moderate weight (0.14)
SD1	Decreases with fatigue → use 1/SD1 ↑ → moderate weight (0.14)
SD2/SD1	Reflects balance/imbalance → moderate weight (0.14)
DFA α_1	Decreases with fatigue, indicates nonlinear structure → higher weight (0.15)
SampEn	Clearly drops with fatigue → use 1/SampEn ↑ → higher weight (0.15)

Table 6. FDTI index (post-workout and 2 hrs after)

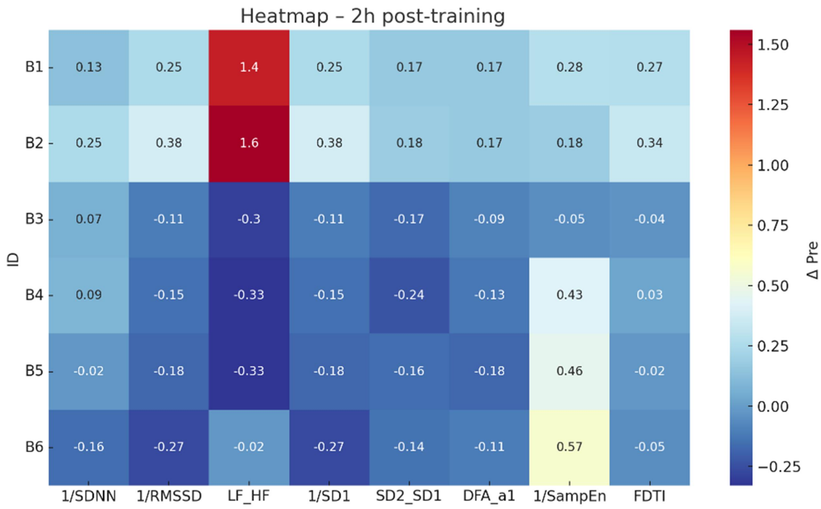
ID	FDTI_post	FDTI_2h
B1	0.92	0.43
B2	1.02	0.49
B3	0.22	0.11
B4	0.30	0.05
B5	0.21	0.02
B6	0.27	0.09

Figure 2 presents heatmaps of the relative changes (relative to pre-training values) of the seven HRV parameters included in the fatigue index formula (1/SDNN, 1/RMSSD, LF/HF, 1/SD1, SD2/SD1, DFA α_1, 1/SampEn), and the combined FDTI index in wrestlers B1–B6. The upper image presents the results immediately after training, and the lower one – two hours after training. Red colors indicate an increase in the corresponding parameter/index (fatigue indicator), and blue ones – a decrease or recovery from baseline. Analysis of the selected HRV indicators shows distinct changes in all studied athletes immediately after the load. The heatmap (Figure 2, upper image) demonstrates a dominant increase in 1/SDNN and 1/SampEn, respectively reflecting reduced variability and regulatory complexity. In most participants, an increase in LF/HF and SD2/SD1 was also observed, which is an indicator of sympathetic dominance and rhythm instability. These trends were most pronounced in B1 and B2, which is confirmed by the combined FDTI index (>0.9–1.0), indicating significant fatigue immediately after training.

Two hours after the load (Figure 2, bottom image), the values of the parameters and the combined FDTI index decreased in all athletes, indicating a partial recovery of autonomic regulation. The index remained positive, but lower (e.g. 0.4–0.5 in B1 and B2), while in the remaining participants it approached the baseline. This model demonstrates the applicability of the new index for tracking the dynamics of fatigue and recovery in an individual and group plan.



A) Immediately after training



B) Two hours after training

Figure 2. Heatmap of relative changes

Summarizing Correlation Analysis

The new composite index FDTI demonstrated a very high positive correlation with the main HRV parameters reflecting sympathetic dominance and a decrease in variability (1/RMSSD, 1/SD1, LF/HF, SD2/SD1; $r=0.95–0.98$, $p<0.01$) immediately after training (Table 7). The strong relationship indicates that the index reliably captures the physiological changes associated with fatigue and is mainly driven by the components reflecting rapid variability and geometric dispersion. Two hours after exercise (Table 8), the correlations weaken, which corresponds to a partial recovery of autonomic regulation and confirms the sensitivity of the index to the dynamics of fatigue and recovery.

Table 7. Correlation Matrix (Post-training) Pearson Correlation Tables (Post & 2h)

	1/SDNN	1/RMSSD	LF/HF	1/SD1	SD2/SD1	DFA α_1	1/SampEn	FDTI
1/SDNN	1.000	0.864	0.658	0.864	0.641	0.712	-0.622	0.802
1/RMSSD	0.864	1.000	0.944	1.000	0.940	0.951	-0.441	0.980
LF/HF	0.658	0.944	1.000	0.944	0.989	0.981	-0.230	0.967
1/SD1	0.864	1.000	0.944	1.000	0.940	0.951	-0.441	0.980
SD2/SD1	0.641	0.940	0.989	0.940	1.000	0.967	-0.275	0.946
DFA α_1	0.712	0.951	0.981	0.951	0.967	1.000	-0.388	0.943
1/SampEn	-0.622	-0.441	-0.230	-0.441	-0.275	-0.388	1.000	-0.395
FDTI	0.802	0.980	0.967	0.980	0.946	0.943	-0.395	1.000

Table 8. Correlation matrix (2 h Post-training)

	1/SDNN	1/RMSSD	LF/HF	1/SD1	SD2/SD1	DFA α_1	1/SampEn	FDTI
1/SDNN	1.000	0.812	0.603	0.812	0.605	0.677	-0.588	0.765
1/RMSSD	0.812	1.000	0.911	1.000	0.922	0.941	-0.409	0.960
LF/HF	0.603	0.911	1.000	0.911	0.978	0.964	-0.220	0.950
1/SD1	0.812	1.000	0.911	1.000	0.922	0.941	-0.409	0.960
SD2/SD1	0.605	0.922	0.978	0.922	1.000	0.959	-0.260	0.934
DFA α_1	0.677	0.941	0.964	0.941	0.959	1.000	-0.345	0.932
1/SampEn	-0.588	-0.409	-0.220	-0.409	-0.260	-0.345	1.000	-0.360
FDTI	0.765	0.960	0.950	0.960	0.934	0.932	-0.360	1.000

Table 9 shows the calculated Pearson correlation [16] between HRV parameters and the FDTI index. After exercise, the index has the strongest positive correlation with 1/RMSSD, 1/SD1 and LF/HF ($r \approx 0.96$–0.98). Two hours after exercise, all correlations weaken but remain high, indicating that the index continues to reflect key changes in variability.

Table 9. Pearson correlation between HRV parameters and FDTI index

Parameter	r (Post-training), p	r (2 h Post-training), p
1/SDNN	$r = 0.802$, $p \approx 0.055$	$r = 0.765$, $p \approx 0.077$
1/RMSSD	$r = 0.980$, $p \approx 0.001$	$r = 0.960$, $p \approx 0.003$
LF/HF	$r = 0.967$, $p \approx 0.002$	$r = 0.950$, $p \approx 0.004$
1/SD1	$r = 0.980$, $p \approx 0.001$	$r = 0.960$, $p \approx 0.003$
SD2/SD1	$r = 0.946$, $p \approx 0.006$	$r = 0.934$, $p \approx 0.009$
DFA α_1	$r = 0.943$, $p \approx 0.006$	$r = 0.932$, $p \approx 0.010$
1/SampEn	$r = -0.395$, $p \approx 0.438$	$r = -0.360$, $p \approx 0.483$

Figure 3 shows a boxplot of the calculated FDTI values for the six wrestlers. The values are normalized to the individual "pre-training" baseline (Pre = 0). An increase in the index is seen immediately after loading and a partial recovery after 2 hours:

– Post (immediately after loading) – the median is above zero, with two clearly higher individuals.

– +2h (two hours after) – the values decrease, but remain slightly positive in some athletes.

Since the sample is small, these values serve only as a demonstration of the concept; with a larger cohort, you will be able to train the weights w_k more precisely and the FDTI will become even more informative.

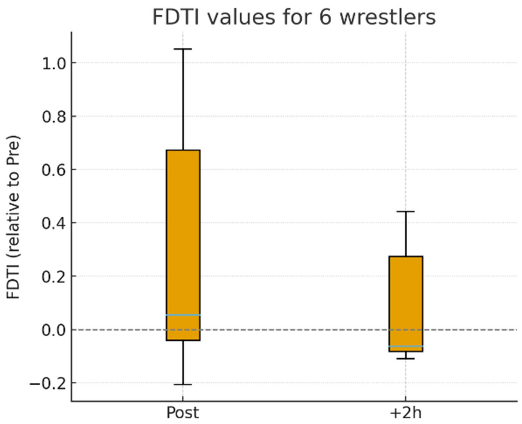


Figure 3. Boxplot of the calculated FDTI values for the six wrestlers

Figure 4 presents the individual FDTI values for six wrestlers (B1–B6) in two states – immediately after training (Post) and two hours after training (+2h). The index was calculated as a weighted combination of HRV indicators (SDNN, RMSSD, LF/HF, SD1, SD2/SD1, DFA α1, SampEn) and normalized to the individual baseline (Pre=0). Each panel of the figure shows a boxplot with the distributions of FDTI, calculated using the hybrid model of HRV indicators and normalized to the individual baseline (Pre=0). The differences in the fatigue response are visible: in B1 and B2 the index increases significantly after exercise and decreases after two hours, while in the other wrestlers the changes are less pronounced. In B1 and B2, a distinct increase in FDTI after exercise (on average ~0.85–1.05) is observed, which partially normalizes after two hours (~0.38–0.45). This reflects the classic autonomic response “acute fatigue → partial recovery”. At B3–B6, FDTI values are around or below zero and indicate a less pronounced response, which may be the result of individual characteristics or different workload.

These results demonstrate that FDTI captures the personalized response of the body to training and allows for real-time assessment of fatigue. Despite the small sample size (N=6), the effect at B1 and B2 is clearly pronounced, which supports the applicability of the approach. It is planned to expand the study with a larger number of athletes and introduce a recovery index (RDTI) to more fully describe the dynamics of recovery.

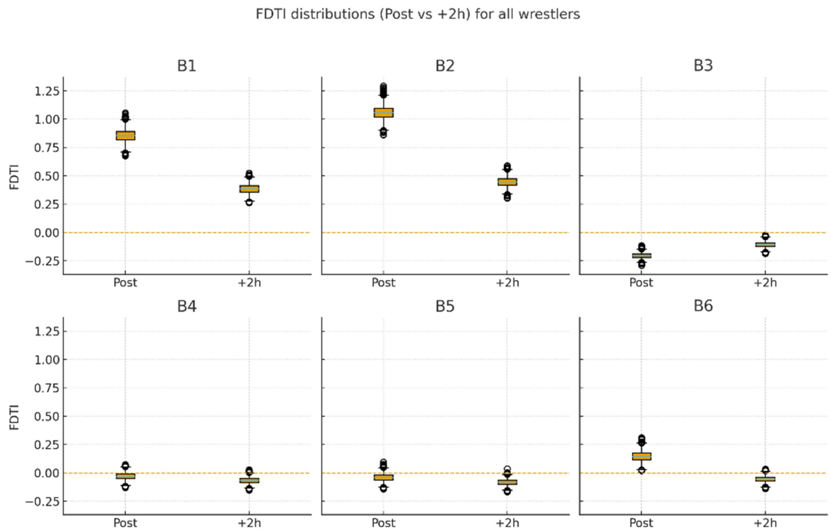


Figure 4. Individual values of the Fatigue Digital Twin Index (FDTI) in six wrestlers (B1–B6) for the states “immediately after training” (Post) and “two hours after training” (+2h).

The graph in Figure 5 shows the dynamics of the FDTI (Fatigue Digital Twin Index) for each of the six wrestlers (B1–B6) at three time points: before training (Pre), immediately after training (Post), and two hours later (2h). Each line corresponds to one athlete and tracks how his/her fatigue index (FDTI) changes over time. The graph shows that all wrestlers have a sharp increase in the index immediately after training; after two hours, most athletes show partial recovery (decrease in FDTI). B4 and B6 have the lowest FDTI in the post-training state, which may be a sign of good adaptation or insufficient intensity. This visualization highlights individual autonomic responses to load and recovery, which is the essence of the HRV digital twin concept.

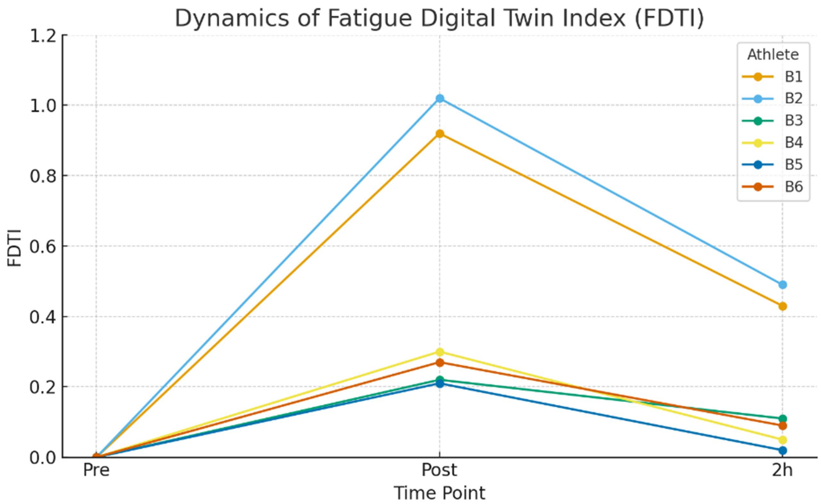


Figure 5. FDTI dynamics

Discussion

The presented results show that the Fatigue Digital Twin (FDTI) index, constructed as an integrated metric of classical, nonlinear and time-frequency HRV indicators, can capture acute changes in autonomic regulation after training load. In two of the six athletes (B1 and B2), a clear “peak” of FDTI was observed immediately after load and partial normalization after two hours, which is consistent with physiological expectations of sympathetic dominance and subsequent recovery. Less pronounced responses in the remaining athletes highlight the need for personalized weights and larger cohorts for training the model.

These initial data demonstrate the potential of the concept of a “digital twin” of HRV, in which individual biosignals are analyzed in real time and converted into a personalized fatigue index. This opens up the possibility not only for monitoring, but also for predicting the risk of overtraining and optimizing recovery in sports practice.

The direct benefit of the created index is providing an assessment of HRV through a single index, instead of through several, which is beneficial for users, who will not have to consider the interaction of several HRV parameters and do not need to understand their essence in depth.

HRV signals can be registered through the wearable device developed in our section, based on PPG (MAX30102) and ECG AFE (MAX30003), with integrated temperature and inertial sensors. The device has an MCU STM32U5A5 and a Bluetooth Low Energy module for data transmission to a mobile application. The system stores and transmits in real time key HRV indicators (SDNN, RMSSD, LF/HF, etc.) to a smartphone/computer/cloud for further processing.

Justification for the need for such a device in the HRV digital twin

- Continuous measurement: The digital twin needs frequent and reliable input data to update its model.
- Multisensor data synchronization: in addition to HRV, the device collects movement and temperature, which allows for artifact correction and load context.
- On-device processing: some of the algorithms (filtering, peaks) are also calculated on the device, which reduces latency and saves battery.
- Security and personalization: data can be encrypted and linked to the specific athlete profile, creating a true "twin".
- Integration with FDTI: the index can be calculated in real time and visualized on the mobile application, to alert in case of excessive fatigue, to adapt training or recovery.

Table 10 compares the main characteristics of the developed wearable device with typical PPG bracelets/watches and shows its advantages for collecting high-quality HRV data required for the digital twin concept.

Table 10. Comparative characteristics of the developed wearable device with typical PPG bracelets/watches

Characteristic	Developed Device (Authors' Prototype)	Typical PPG Wristband/Smartwatch
Sensors	PPG (MAX30102) + ECG (MAX30003) + Temperature (MAX30205) + Accelerometer (LSM6DSL)	PPG only; in some cases, accelerometer
Sampling Rate	100–250 Hz (raw signals)	25–50 Hz (filtered or aggregated data)
Data Synchronization	Yes (PPG, ECG, motion, and temperature in one synchronized stream)	Limited, no ECG
Data Processing	Local: filtering, peak detection, HR and HRV calculations	Usually only heart rate; HRV approximated
Data Transmission	Bluetooth Low Energy, encrypted	Bluetooth or proprietary protocol, often unencrypted
Personalization	Individual baseline profiles and adaptive FDTI weights	Typically absent; generic algorithms
On-device Analysis	Real-time FDTI computation possible	None or very limited
Artifact Correction	Uses accelerometer and temperature for motion/noise reduction	Usually absent or limited
Application	Research platform for HRV digital twin modeling	Consumer use – fitness or wellness tracking

As can be seen from Table 10, the prototype we developed significantly outperforms standard PPG bracelets/watches in terms of sampling rate, data synchronization, local processing, and the ability to custom calculate FDTI, making it a suitable hardware basis for implementing the HRV digital twin.

Future Work

The present study presents a framework for mobile AI-based processing of ECG signals, but the concept of a Digital Twin is at this stage mostly in a conceptual phase. In future work, the actual construction and integration of a cloud-based digital twin of the heart with the mobile application, which would be synchronized in real time with data from portable devices, is envisaged. It is planned to expand the model with multimodal inputs (ECG, PPG, blood pressure, temperature, gyroscope data, etc.) and implement methods for federated and confidential training on large user populations (athletes, individuals in fatigue and stress conditions, patients). The impact of personalized digital twins on clinical practice will be further investigated through long-term case studies and assessment of their predictive and preventive value.

CONCLUSION

In this study, the concept of a Digital Twin for HRV is defined and a new index – Fatigue Digital Twin Index (FDTI) – is presented for the assessment of fatigue and stress in athletes based on an integrated analysis of HRV indicators. The index was calculated on real-world data from six wrestlers in three states (before, immediately after and two hours after training) and demonstrated its ability to capture the individual autonomic response to exercise. In some athletes, a clear peak of FDTI after training and partial recovery after two hours were observed, confirming the physiological validity of the approach.

These initial results support the concept of a digital twin of HRV, which allows for continuous, personalized and interpretable monitoring of fatigue in real time. Future studies are planned to expand the cohort, optimize the index weights and develop additional indices – for example, for recovery (RDTI) and for predicting the risk of overtraining – as the next step towards a comprehensive digital twin system in sports.

ACKNOWLEDGMENTS

This research was funded by the National Science Fund of Bulgaria (scientific project "Modeling and creation of a sensor system for research and analysis of the body's health"), Grant Number KP-06-M67/5, 13.12.2022.

REFERENCES

1. Chaparro-Cárdenas, S.L.; Ramirez-Bautista, J.-A.; Terven, J.; Córdova-Esparza, D.-M.; Romero-Gonzalez, J.-A.; Ramírez-Pedraza, A.; Chavez-Urbiola, E.A. A Technological Review of Digital Twins and Artificial Intelligence for Personalized and Predictive Healthcare. Healthcare 2025, 13, 1763. DOI: https://doi.org/10.3390/healthcare13141763
2. Y. Fuse, S. N. Murphy, H. Ikari, A. Takahashi, K. Fuse, and E. Kawakami, Artificial intelligence in clinical data analysis: A review of large language models, foundation models, digital twins, and allergy applications, Allergology International, 2025. DOI: https://doi.org/10.1016/j.alit.2025.06.005
3. T. Kreuzer, P. Papapetrou, and J. Zdravkovic, Artificial intelligence in digital twins—A systematic literature review, Data & Knowledge Engineering, vol. 151, p. 102304, 2024. DOI: https://doi.org/10.1016/j.datak.2024.102304
4. Qian S, Ugurlu D, Fairweather E, Toso LD, Deng Y, Strocchi M, Cicci L, Jones RE, Zaidi H, Prasad S, Halliday BP, Hammersley D, Liu X, Plank G, Vigmond E, Razavi R, Young A, Lamata P, Bishop M, Niederer S. Developing cardiac digital twin populations powered by machine learning provides electrophysiological insights in conduction and repolarization. Nat Cardiovasc Res. 2025 May;4(5):624-636. DOI: https://doi.org/10.1038/s44161-025-00650-0
5. T. Grandits, K. Gillette, G. Plank, and S. Pezzuto, Accurate and efficient cardiac digital twin from surface ECGs: Insights into identifiability of ventricular conduction system, Medical Image Analysis, vol. 105, p. 103641, 2025. DOI: https://doi.org/10.1016/j.media.2025.103641

6. Thangaraj PM, Benson SH, Oikonomou EK, Asselbergs FW, Khera R. Cardiovascular care with digital twin technology in the era of generative artificial intelligence. Eur Heart J. 2024 Dec 1;45(45):4808-4821. DOI: https://doi.org/10.1093/eurheartj/ehae619
7. Coorey G, Figtree GA, Fletcher DF, Snelson VJ, Vernon ST, Winlaw D, Grieve SM, McEwan A, Yang JYH, Qian P, O'Brien K, Orchard J, Kim J, Patel S, Redfern J. The health digital twin to tackle cardiovascular disease-a review of an emerging interdisciplinary field. NPJ Digit Med. 2022 Aug 26;5(1):126. DOI: https://doi.org/10.1038/s41746-022-00640-7
8. K. Sel, D. Osman, F. Zare, S. M. Shahrbabak, L. Brattain, J.-O. Hahn, O. T. Inan, and R. Jafari, Building Digital Twins for Cardiovascular Health: From Principles to Clinical Impact, Journal of the American Heart Association, vol. 13, no. 19, 2024. DOI: https://doi.org/10.1161/JAHA.123.031981
9. M. R. Kabir, F. S. Shishir, S. Shomaji, and S. Ray, Digital twins in healthcare IoT: A systematic review, High-Confidence Computing, vol. 5, no. 3, 2025, Art. no. 100340. DOI: https://doi.org/10.1016/j.hcc.2025.100340
10. Bhagirath P, Strocchi M, Bishop MJ, Boyle PM, Plank G. From bits to bedside: entering the age of digital twins in cardiac electrophysiology. Europace. 2024 Dec 3;26(12):euae295. DOI: https://doi.org/10.1093/europace/euae295
11. Z. Johnson and M. J. Saikia, Digital Twins for Healthcare Using Wearables, Bioengineering (Basel), vol. 11, no. 6, p. 606, Jun. 2024. DOI: https://doi.org/10.3390/bioengineering11060606
12. I. Volkov, G. Radchenko, and A. Tchernykh, Digital Twins, Internet of Things and Mobile Medicine: A Review of Current Platforms to Support Smart Healthcare, Programming and Computer Software, vol. 47, pp. 578–590, 2021. DOI: https://doi.org/10.1134/S0361768821080284
13. J. Chen, C. Yi, H. Du, D. Niyato, J. Kang, J. Cai, and X. Shen, A Revolution of Personalized Healthcare: Enabling Human Digital Twin With Mobile AIGC, IEEE Network, vol. 38, no. 6, pp. 234–242, 2024. DOI: https://doi.org/10.1109/MNET.2024.3366560
14. Edder, A.; Ben-Bouazza, F.-E.; Tafala, I.; Manchadi, O.; Jioudi, B. Self Attention-Driven ECG Denoising: A Transformer-Based Approach for Robust Cardiac Signal Enhancement. Signals 2025, 6, 26. DOI: https://doi.org/10.3390/signals6020026
15. Gaoudam N, Sakhamudi SK, Kamal B, Addla N, Reddy EP, Ambala M, Lavanya K, Palaparthi EC, Bhattam A, Periasamy P, Sayana SB, Medabala T. Wearable Devices and AI-Driven Remote Monitoring in Cardiovascular Medicine: A Narrative Review. Cureus. 2025 Aug 16;17(8):e90208. DOI: https://doi.org/10.7759/cureus.90208
16. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. Anesth Analg. 2018 May;126(5):1763-1768. DOI: https://doi.org/10.1213/ANE.0000000000002864

Received: 28-10-2025 Accepted: 22-12-2025 Published: 29-12-2025

Cite as:
Georgieva-Tsaneva, G., Cheshmedzhiev, K. (2025). "Artificial Intelligence for EGG/PPG Signal Processing on Mobile Platforms", Science Series "Innovative STEM Education", volume 07, ISSN: 2683-1333, pp. 272-290, 2025. DOI: https://doi.org/10.55630/STEM.2025.0725