This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

IMU-based Pelvic Rotation Detection: A Novel Dataset, Benchmark Classifiers, and Sensor Placement Optimization

Meixing Liao, *Graduate Student Member, IEEE*, Eva Libeert, Joeri Van Cauwelaert, Chunzhuo Wang, *Member, IEEE*, Simon Brumagne, Jean-Marie Aerts, *Member, IEEE*, and Bart Vanrumste, *Senior Member, IEEE*

***Abstract*— Prolonged sedentary behavior has been a major public health concern strongly associated with low back pain (LBP), the leading cause of disability worldwide. Monitoring pelvic rotations (PRs), a key factor in postural control, is crucial for LBP prevention and management. However, PR detection remains underexplored, and publicly available datasets are lacking. This study comprehensively investigates the feasibility of using inertial measurement units (IMUs) and machine learning (ML) to recognize PRs. We introduce a novel public dataset comprising over 15 hours of high-resolution IMU data from 39 sedentary participants, specifically designed for PR detection across multiple activities, including sitting, standing, sit-to-stand-to-sit (STSTS) transitions, walking, and targeted anterior/posterior PRs during sitting and standing. We benchmark six classical ML classifiers and identify XGBoost as the most effective model, achieving weighted F1 scores of 0.988 and 0.895 in 2-class and 8-class configurations, respectively. We further conduct a systematic sensor placement optimization, demonstrating that a two-sensor configuration (pelvis and left upper leg) reached performance comparable to that of the full-body 17-sensor system. In the comprehensive 8-class classification task (sit, stand, STSTS, walk, and targeted anterior/posterior PR), this optimal combination achieved a weighted F1 score of 0.872, closely matching the result obtained using all sensors. Our work provides the community with a unique dataset, a robust baseline for PR recognition, and a practical minimal-sensor solution, thereby paving the way for real-world wearable systems aimed at reducing prolonged static postures and improving LBP management in daily life.**

***Index Terms*— Inertial sensors, Machine Learning, Human Activity Recognition, Pelvic Rotation, Wearable Health Monitoring**

I. Introduction

SPENDING approximately one-third of the day sitting has become the epitome of life for many individuals [1]. However, the increasing prevalence of prolonged sitting is accompanied by potential health problems, among which low back pain (LBP) is one of the most common [2]. This chronic condition is now recognized as the leading cause of years of healthy life lost due to disability, affecting more than 600 million people worldwide [3].

One of the key contributors to LBP is poor posture, which may cause excessive stress on the spine during prolonged static postures [4], defined as remaining in the same position with minimal movement for extended durations [5]. A critical factor influencing posture control and spinal alignment is pelvic rotation (PR), which plays a fundamental role in adjusting spinal curvature and displacing the center of mass to maintain balance, redistribute load [6], and enable transitions between postures, such as sitting, standing, and walking [7], [8]. Despite its clinical importance, PR detection remains underexplored in current research, representing a gap that this study aims to address.

There is a growing trend toward extending the applications of sensory technologies beyond activity and posture recognition [9], [10], [11]. In healthcare, such systems are increasingly integrated into telehealth applications [12], such as elderly care [13], freezing of gait assessment [14], and spine motion analysis [15], allowing for timely interventions and personalized care. Despite advances in sensor-based health monitoring, studies explicitly focusing on LBP and PR detection using sensors remain limited.

In particular, in the field of sedentary behavior monitoring, widely used sensors include inertial measurement units (IMUs)

This work was supported in part by the China Scholarship Council (CSC) under Grant No.202208090040 and in part by Flanders AI (Corresponding author: M. Liao).

Meixing Liao is with the e-Media Research Lab; the ESAT-STADIUS Division; the M3-BIORES Research Group, Department of Biosystems; and the Research Group for Musculoskeletal Rehabilitation, Department of Rehabilitation Sciences, KU Leuven, 3000 Leuven, Belgium (email: meixing.liao@kuleuven.be)

Eva Libeert was with the M3-BIORES Research Group, Department of Biosystems, KU Leuven, 3001 Leuven, Belgium (email: eva.libeert@kuleuven.be)

Joeri Van Cauwelaert was with the e-Media Research Lab, KU Leuven, 3001 Leuven, Belgium (email: joeri.vancauwelaert@student.kuleuven.be)

Chunzhuo Wang is with the e-Media Research Lab and the ESAT-STADIUS Division, KU Leuven, 3000 Leuven, Belgium (email: chunzhuo.wang@kuleuven.be)

Simon Brumagne is with the Research Group for Musculoskeletal Rehabilitation, Department of Rehabilitation Sciences, KU Leuven, 3001 Leuven, Belgium (email: simon.brumagne@kuleuven.be)

Jean-Marie Aerts is with the M3-BIORES Research Group, Department of Biosystems, KU Leuven, 3001 Leuven, Belgium (email: jean-marie.aerts@kuleuven.be)

Bart Vanrumste is with the e-Media Research Lab and the ESAT-STADIUS Division, KU Leuven, 3000 Leuven, Belgium (email: bart.vanrumste@kuleuven.be)

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

[16], pressure sensors [17], and cameras [18]. Regarding their placement, these sensors are typically embedded in chairs [19], seat cushions [16], shoe insoles [20], or worn on the body [21], [22]. Although existing approaches can estimate sitting duration to indicate sedentary behavior, they typically provide limited insight into postural patterns, offering minimal guidance to users to reduce prolonged static postures.

Beyond sedentary monitoring, recent studies have advanced posture assessment using different modalities. Tanthuwapathom et al. validated the reliability of an IMU-based system for sitting posture assessment against physical therapist video evaluations using the RULA method, showing strong correlations but noting limitations related to calibration and controlled laboratory settings [23]. Xing et al. designed and validated a depth camera–based posture assessment system that can detect common anomalies such as uneven shoulders, pelvic tilt, and scoliosis with high accuracy compared to optical motion capture, though its applicability was limited to static screening contexts [24], while Lee et al. used human pose estimation to assess sit-to-stand balance across younger and older adults, identifying compensatory arm and trunk movements as critical features to distinguish groups [25]. The same potential and limitations are found in the reviews. For example, Simpson et al. systematically reviewed wearable systems for spinal posture analysis, concluding that IMU-based devices show promise but require further validation for long-term and real-world applications [26]. More recently, Figueira et al. reviewed wearables for posture monitoring and feedback in occupational settings, reporting that IMU-based systems with vibrotactile feedback are the most common, but still face challenges in comfort, battery life, and user acceptance [27]. In addition, García-Jaén et al. systematically reviewed lumbo-pelvic rhythm monitoring using wearable technology with sensory biofeedback, highlighting the potential of IMU-based devices while also noting major challenges related to standardization, usability, and long-term effectiveness [28].

To fill the gaps and take a first step toward the broader goal of providing personalized and detailed feedback for reducing prolonged postures, this study explores the feasibility of using commercially available wearable sensors and machine learning (ML) classifiers to recognize PRs during activities of daily life. We focus on detecting two types of PRs: 1) subconscious PRs, which naturally occur during transitions, such as walking or sit-to-stand-to-sit (STSTS) movements; 2) targeted PRs, which can be performed purposely during static postures like sitting or standing. We use the term "targeted PRs" rather than "intentional PRs" to emphasize that these rotations can be guided or encouraged, for example, through posture feedback systems, or exercises prescribed by physical therapists for LBP rehabilitation. This distinction is especially relevant in real-world settings where standing up and moving may not be feasible, such as during lessons, meetings, or driving. In these cases, PRs offer a practical strategy to compensate for prolonged static postures and support low back health without leaving the seat. The main contributions of this study are as follows:

1) Dataset Collection: We collected a novel IMU-based dataset[1] from forty-five subjects focused on PRs using the Xsens Awinda Motion Capture System (Xsens).
2) Classifier Benchmarking: We evaluated the performance of six classic ML classifiers, including Decision Tree (DT), K-Nearest Neighbors (KNN), Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and XGBoost (XGB) to identify the most effective model for recognizing PRs.
3) Sensor Placement Optimization: We investigated the impact of sensor placement on classification performance, demonstrating the benefits of combining multiple sensor units to improve classification accuracy.

With the contributions mentioned above, this study provides valuable information on the use of wearable sensors and ML for posture recognition, and paves the way for future systems aimed at reducing prolonged postures in daily life.

II. Methods

A. Data Collection

The study was approved by the KU Leuven Social and Societal Ethics Committee (protocol number: G-2023-6913-R2 (MIN), date of approval: December 11, 2023). In total, forty-five sedentary adults were recruited, all without self-reported discomfort at the time of participation.

Measurements were conducted in an empty experimental room with consistent lighting and minimal ambient noise to reduce external distractions. Each session started by familiarizing participants with the protocol and collecting their signed informed consent. The sensor calibration process followed the standard Xsens MVN software-guided workflow, including the N-pose initialization recommended by the manufacturer and 10 seconds of walking around. Body dimensions were measured for high-quality sensor calibration, after which participants changed into tight-fit shirts and leggings. Throughout the session, they were instructed to perform a series of activities (Table I), including static postures (sitting and standing), subconscious PRs, such as STSTS and walking, and targeted PRs during sitting and standing in both anterior and posterior directions. During all measurements, participants wore the Xsens (Xsens MVN, Enschede, The Netherlands)[2], which incorporates seventeen wearable IMUs with an output sampling frequency of 60 Hz, providing high-resolution IMU data for accurate motion tracking. The sensor placement is shown in Fig. 1.

For each static sitting and standing task, participants hold their usual postures with arms folded across the chest for half a minute for a clearer recording view of the low back. In the targeted PR tasks, each repetition started in the sitting/standing position, followed by an anterior PR, a posterior PR, and returned to the initial position. To increase data diversity, participants performed PRs at three self-controlled amplitudes:

[1]The dataset is published on KU Leuven RDR under restricted access in accordance with GDPR. Interested researchers may request and obtain full access after signing a data usage agreement with KU Leuven. DOI: https://doi.org/10.48804/SHJX2D.

[2]Xsens product details are available at: https://www.movella.com/products/motion-capture/xsens-mvn-awinda.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

TABLE I
ACTIVITIES DURING DATA COLLECTION

Task	Duration / repetition
Sit	30 seconds
Targeted PR during sitting • Full RoM • Half of RoM • Small fraction of RoM	5 repetitions for each amplitude
STSTS	5 times as fast as possible
Stand	30 seconds
Targeted PR during standing • Full RoM • Half of RoM • Small fraction of RoM	5 repetitions for each amplitude
Walk forward 3 steps and • return by walking 3 steps backwards • turn 180° clockwise (right side) and walk 3 steps back to the start.	5 repetitions for each

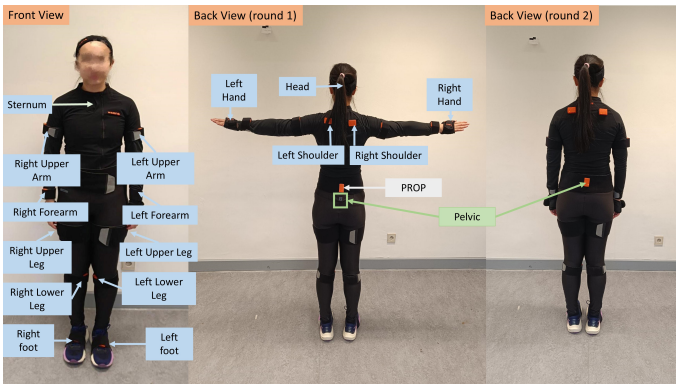


Fig. 1. The placements of the Xsens MVN Awinda motion capture sensors.

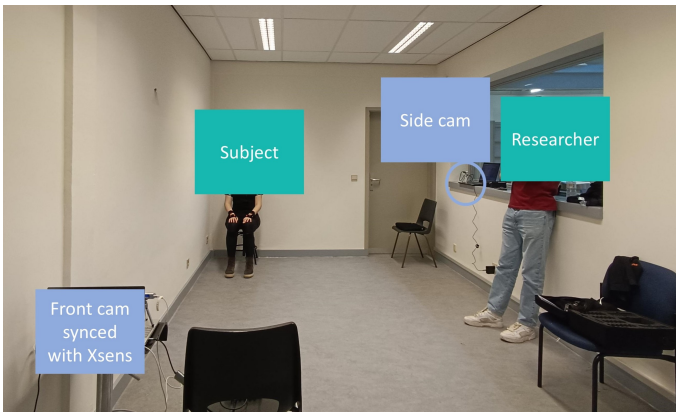


Fig. 2. The measurement environment.

their full range of motion (RoM), approximately half RoM, and a small fraction of RoM. Each amplitude level was repeated five times. All repetitions were supervised in real time by the researchers, with synchronized video recordings and the Xsens Avatar available for both on-site and post-hoc validation. The three amplitude levels were not defined by exact angles, but were used to mimic real-world variability in PRs, where movements naturally occur at varying magnitudes. This design also provides potential for future personalized modeling, as RoM can vary substantially among individuals. As the foundational study, this paper focuses solely on benchmarking PR recognition using generalized models, and PRs in the same direction and posture were considered as one class. Thus, minor deviations in executing PRs did not affect labeling in this study. During the walking task with the turn, participants were instructed to always turn clockwise (to the right side), ensuring that their back was consistently captured by the side camera placed on the left. This standardized protocol guaranteed clear and reliable recordings of pelvic motion.

To evaluate the robustness of pelvic sensor placement, participants completed tasks twice using different positions (Fig. 1). In the first round, the pelvic sensor ("PELV") was placed in the back pocket of the leggings on the sacrum, while the additional sensor, named "PROP" by the manufacturer of Xsens, was activated and positioned on a stretchable belt just above the "PELV" sensor. In the second round, the "PROP" sensor was disabled, and the "PELV" sensor was relocated to the stretchable belt above the legging pocket to ensure the same placement across both rounds. The remaining sixteen sensor units were placed according to the official Xsens MVN placement guidelines [29], using the manufacturer-provided shirts with velcro patches and elastic straps to ensure secure and standardized attachment. Fig. 2 demonstrates the measurement environment: the entire session was recorded using a laptop camera (HP 5MP camera, 30 fps, 2560 * 1920), synchronized with Xsens, together with two Intel RealSense D435 cameras (30 fps, 1280 * 720). One RealSense camera was placed next to the Xsens Awinda station (receiver), facing the participant from the front, and the other was three meters away on the left side. Except during walking tasks, the front camera and the Xsens hubs were 4.6 meters away from the participant, while for the walking tasks, the distance was increased to 5.9 meters. The synchronized video recordings served as the ground truth for data segmentation.

B. Data Preprocessing and Processing

As each task was conducted and logged separately, Xsens data annotations were made concurrently with the measurement sessions. To prepare the dataset for ML classifiers, four main preprocessing steps were performed: 1) Segment extraction; 2) Quaternion rotation; 3) Windowing; and 4) Feature extraction and processing.

1) Segment extraction: Xsens recordings were parsed and manually segmented using Xsens MVN software (version 2024.0), with synchronized camera recordings serving as reference. Each segment corresponded to a continuous time interval during which a specific activity (sitting, standing, targeted anterior/posterior PR during sitting/standing, STSTS, and walking) was performed. Complete and correctly recorded

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

task repetitions were identified for each participant and individually exported as Excel files. Repetitions with parsing issues or low-quality recordings in certain participants (e.g., signal dropouts or misaligned segments) were excluded. These manually verified segments served as ground truth labels for model training and evaluation.

2) Quaternion rotation: While raw accelerometer signals were available from the Xsens MVN export, raw gyroscope data were not. Therefore, quaternion data provided by Xsens were used to represent orientation for feature extraction. To account for inter-subject calibration differences, after calibration, all orientations were realigned using quaternion-based rotation into the local reference frame, a fixed right-handed Cartesian frame defined by Xsens with its origin at the right heel of the avatar, the X-axis pointing forward, the Y-axis to the left, and the Z-axis upward [29], [30]. Quaternions were preferred over Euler angles for their robustness in handling 3D rotations without gimbal lock [31].

3) Windowing: A sliding window approach was applied to segment the continuous IMU data for feature extraction and classification. We initially considered multiple window lengths (0.1, 0.5, 1.0, and 2.0 seconds) and stride lengths (stride/window ratio = 0.1, 0.2, 0.5, and 0.8). Shorter windows (e.g., 0.1 seconds) produced substantially more samples. As static activities and walking typically lasted more than 30 seconds, they generated many more windows than short activities, such as targeted PRs and STSTS, which generally lasted only a second per direction. Conversely, larger windows (1–2 seconds) frequently included null segments within targeted PR windows, reducing class-specific information. Considering these trade-offs, observed segment durations (Supplementary Table VIII), and class-level window balance analysis (Supplementary Table IX), we selected a 0.5-second window with a 0.1-second stride. This configuration provided the best compromise between capturing short-duration characteristics of targeted PRs and maintaining manageable sample counts and acceptable class balance, and was therefore retained for all further analyses.

4) Feature extraction and processing: To characterize the windows, thirteen time-domain and three frequency-domain features were hand-crafted to capture relevant motion patterns [32]. These features were derived from the ten degrees of freedom available per sensor: 3-axis acceleration, 3-axis magnetic field, and 4-axis quaternion orientation. The inclusion of magnetic field signals enabled the capture of heading-related (yaw) information that contributes to PR detection. Although magnetic signals can be affected by environmental disturbances, all recordings were conducted in a controlled laboratory environment with short durations, minimizing noise. Given that IMUs were placed at seventeen distinct body locations, with an additional sensor ("PROP") used in the first round, features were calculated for each sensor for both rounds, resulting in 2880 features in the first round and 2720 in the second. Table II lists a summary of hand-crafted features, sensor locations, and degrees of freedom.

Before classification, we applied a feature selection process to reduce computational complexity. First, features with no variability across the dataset were removed. Next, highly correlated features (Pearson correlation coefficient > 0.95) were excluded to reduce redundancy. Finally, the remaining features were normalized using Min-Max scaling to the range $[0, 1]$ to ensure consistent feature weighting and improve classifier performance [33]. The feature selection process was conducted once on the entire dataset prior to model training and cross-validation for an objective comparison across folds and sensor configurations. The unsupervised, label-independent selection process minimized the risk of look-ahead bias.

C. Classification

1) Classifiers: To recognize PRs using motion capture data, six common ML algorithms, including RF [34], SVM [35], KNN [36], DT [37], LR [38], and XGB [39] were implemented using the scikit-learn library (v1.2.2) [40] in Jupyter Notebook (Python 3.10.12). To fairly compare the model performance, we trained the baseline models without hyperparameter tuning. The full list of hyperparameters for each classifier is summarized in Table III.

2) Class configurations: Since the primary focus of this study is PR recognition, we designed a hierarchical classification scheme. Starting with 2-class classification (PR vs. static postures), the scheme further distinguished between targeted and subconscious PR, directional differences (anterior/posterior), posture-specific contexts (sit vs. stand), and finally, to posture-specific and complete activity-based categorization (8 classes). The full class configurations are illustrated in Fig. 7 in the supplementary.

3) Sensor placement optimization: To establish a baseline for PR recognition, we first evaluated classifier performance using the full-body sensor setup in all classification configurations. Given that wearing seventeen sensors is impractical for daily use, we investigated whether a single-sensor setup could achieve comparable performance. Based on individual sensor results, we further explored combinations of two sensors to improve accuracy while maintaining minimal wearability.

4) Grouped Leave-One-Subject-Out cross-validation: To fairly assess the robustness of the classifiers, a grouped Leave-One-Subject-Out cross-validation was applied. Thirty-nine eligible participants were divided into five groups, four with eight participants each and one with seven participants [41]. In each fold, the model was trained on four groups and tested on the remaining group, and the process was repeated five times so that each group served as the test set once. This validation strategy simulated real-world conditions in which the models encounter data from previously unseen users. All models were trained and tested on data from (1) Round 1, (2) Round 2, and (3) the combined dataset.

D. Evaluation

To evaluate model performance under class imbalance, we used the weighted F1 score as one of the evaluation metrics, providing an overall measure that accounts for class distribution. In addition, as PRs are the key focus of this study, we report class-specific F1 scores. Confusion matrices were also plotted to illustrate misclassifications among classes. This multi-metric approach captures both overall performance and

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

TABLE II
OVERVIEW OF HAND-CRAFTED FEATURES, SENSOR LOCATIONS, AND DEGREES OF FREEDOM

Category	Items
Time-domain features (13)	Mean, variance, standard deviation, root mean square, median, skewness, kurtosis, zero crossing, slope sign changes, mean absolute value, interquartile range, median absolute deviation, peak to peak.
Frequency-domain features (3)	Spectral peak frequency, spectral peak magnitude, FFT energy.
Sensor locations (17 + 1*)	Pelvic, sternum, head, right shoulder, right upper arm, right forearm, right hand, left shoulder, left upper arm, left forearm, left hand, right upper leg, right lower leg, right foot, left upper leg, left lower leg, left foot, PROP*.
Degrees of Freedom (10)	Free acceleration (x, y, z), magnetic field (x, y, z), orientation (q0, q1, q2, q3).

*PROP refers to the additional sensor used in the first round to optimize the placement of the pelvic sensor.

TABLE III
HYPERPARAMETERS USED FOR BASELINE CLASSIFIERS

Model	Hyperparameters
DT	criterion = gini; max_depth = None; min_samples_split = 2; min_samples_leaf = 1; max_features = None.
KNN	n_neighbors = 5; weights = uniform; metric = minkowski; p = 2.
LR	penalty = l2; C = 1.0; solver = lbfgs; max_iter = 1000; multi_class = auto.
RF	n_estimators = 100; criterion = gini; max_depth = None; min_samples_split = 2; min_samples_leaf = 1; bootstrap = True; max_features = sqrt.
SVM	kernel = rbf; C = 1.0; gamma = scale; degree = 3; tol = 10^{-3}.
XGB	n_estimators=100; learning_rate = 0.1; max_depth = 6; subsample = 1.0; colsample_bytree = 1.0; gamma = 0; objective = multi:softprob.

the ability to detect the specific subtle PRs from the common activities of daily life.

The weight F1 score is calculated as follows:

$$\text{Weighted F1 score} = \sum_{i=1}^{k} \frac{\#\text{ True Instances}_i}{\#\text{ Total Instances}} \times \text{F1 score}_i \quad (1)$$

where k is the total number of classes, True Instances$_i$ is the number of true instances for class i, and F1 score$_i$ is the F1 score for class i, defined as:

$$\begin{aligned} \text{F1 score}_{\text{class } i} &= 2 \times \frac{\text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \\ &= \frac{2\text{TP}_i}{2\text{TP}_i + \text{FP}_i + \text{FN}_i} \end{aligned} \quad (2)$$

where TP$_i$, FP$_i$, FN$_i$ represent the number of true positives, false positives, and false negatives for class i, respectively [42].

E. Statistical Analysis

1) Model comparison: To assess whether significant differences existed among the models for each class configuration (2, 3, 4, 6, and 8) across both recording rounds, we employed the Friedman test [43]. This non-parametric test is suitable for repeated measures and helps determine whether differences among multiple related groups are statistically significant. When the Friedman test indicated a significant result (p-value < 0.05), we followed up with the Nemenyi post-hoc test to identify which specific models differed from each other [44].

2) Sensor comparison: Several paired t-tests were performed to determine whether there was a statistically significant difference between the performance of various sensor units. The findings contribute to the optimization of sensor placement.

3) Round comparison: Similarly, we performed a paired t-test to ensure that there were no statistically significant differences between the two recording rounds, allowing data augmentation by combining the data from both rounds.

4) Pelvic rotation direction comparison: To examine potential differences in model performance between targeted posterior and anterior PRs, we used the Wilcoxon signed rank test to compare class-specific F1 scores between both the two PR directions [45].

III. RESULTS

Our evaluation includes extensive analytical comparisons involving multiple classifiers, sensor and class configurations, supported by statistical significance testing.

A. Dataset

As the first objective, we collected a novel dataset from 45 subjects, specifically designed to detect PRs across multiple activities and postures. Measurements took place in a controlled experimental setting and repeated over two rounds following the same protocol. Table IV presents the demographic statistics of the participants, including age, weight, height, and body mass index (BMI).

After data processing using the Xsens MVN Analyze software, recordings of six participants were excluded due to unparsable or low-quality files. The final published dataset consists of more than fifteen hours of Xsens recordings, resulting in 8251 valid activity segments from thirty-nine individuals.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

TABLE IV
PARTICIPANT DEMOGRAPHICS

Index	All (n = 45)		Female (n = 23)		Male (n = 22)	
	Mean	Std	Mean	Std	Mean	Std
Age	29.6	10.1	29.9	10.1	29.2	10.4
Weight	69.1	14.0	60.2	7.7	78.5	13.1
Height	176.3	10.7	168.6	7.4	184.4	7.1
BMI	22.1	2.7	21.4	2.1	22.7	3.1

TABLE V
OVERALL WEIGHTED AVERAGE F1 SCORES FOR RECORDING ROUND 1 USING THE FULL SET OF SENSORS (18 IMUS)

Models	2-class	3-class	4-class	6-class	8-class
XGB	**0.988**	**0.976**	**0.906**	**0.900**	**0.895**
RF	0.967	0.969	0.877	0.868	0.879
SVM	0.957	0.955	0.861	0.857	0.861
LR	0.916	0.871	0.845	0.849	0.858
KNN	0.911	0.904	0.782	0.767	0.747
DT	0.820	0.725	0.734	0.706	0.715

A paired t-test comparing the two recording rounds showed no statistically significant difference (p-value = 0.391). Consequently, the two rounds were merged to augment the data for future analyses.

B. Classifier Selection

Table V summarizes the weighted F1 scores for Round 1. As the results demonstrated, XGB consistently achieved the highest F1 scores, ranging from 0.988 (2-class) to 0.895 (8-class). RF followed closely, while SVM and LR performed reasonably well, especially in simpler configurations. In contrast, KNN and DT exhibited lower performance, particularly as the number of classes increased. Results from Round 2 followed the same pattern: XGB again led across all configurations, reaching 0.992 (2-class) and 0.909 (8-class), followed by RF.

Activity-specific analysis (Fig. 3 as an example) further confirmed the feasibility of detecting PRs from activities of daily life, while highlighting the difficulty of specifying the direction of targeted PRs. XGB consistently performed the best across all PR-related classes, followed by RF, reflecting the trends observed in the overall F1 scores.

In the multi-class configuration, XGB achieved significantly higher performance than KNN, DT, and LR, with p-values ranging from 0.003 to 0.008 (Table VI). Although RF performed similarly, with no statistically significant difference from XGB, the latter demonstrated greater consistency across both rounds and class configurations. Therefore, XGB was selected for further analyses.

C. Sensor Location Optimization

This sensor location optimization constitutes a systematic ablation study, isolating the effect of each sensor and sensor combination on classification performance.

1) Single sensor performance: The performance of individual sensors was evaluated across all class configurations and both recording rounds. Fig. 4 illustrates the F1 scores of different sensor units in the 8-class configuration in Round 1, with similar patterns found in other configurations.

The chart indicates that the "walk" class was the most easily recognized by all sensors. Most sensors also identified STSTS transitions (sit-to-stand/stand-to-sit), showing the feasibility of recognizing subconscious PRs using the tested classifiers. Focusing on particular sensors, the "PELV" sensor achieved the highest F1 scores for most tasks. The proximity of this sensor to the pelvis makes it more effective in detecting PRs. Similarly, upper leg sensors performed well, particularly in detecting subconscious PRs (STSTS and walking), where their placement captures notable lower limb movement involved in these tasks. Besides, the "PROP" sensor demonstrated better performance in recognizing PR-related tasks compared to the rest of the sensors. In contrast, sensors positioned at other locations, such as on the head, forearms, and feet, showed much lower F1 scores across nearly all tasks. These sensors minimally contributed to classification, with F1 scores below 0.5 in most cases, making them less critical for PR detection.

Focusing on the pelvic sensors, the F1 scores for the "PELV" sensor in Round 1 ("PELV" R1), placed in the legging pocket, were consistently higher than those of the other two sensor placements ("PELV" R2 and "PROP" R1). Paired t-tests comparing the three pelvic sensor positions confirmed statistically significant differences: "PELV" R1 vs "PROP" (p-value = 3.49e-07), and "PELV" R1 vs "PELV" R2 (p-value = 0.00034), indicating that even small differences in placement can affect model performance. The comparison between "PELV" R2 and "PROP" R1 was not significant (p-value = 0.071), suggesting that these two belt placements performed similarly.

2) Sensor combination performance: The results obtained from individual sensor units led to the hypothesis that performance can be improved by integrating pelvic sensors with those of the upper leg. Consequently, in the first round, four combinations were evaluated: "PELV + left upper leg", "PELV + right upper leg", "PROP + left upper leg", and "PROP + right upper leg". In the second round, two combinations, "PELV + left upper leg" and "PELV + right upper leg", were evaluated in all class configurations.

The weighted F1 scores for the first round are visualized in Fig. 5 as a heatmap, further highlighting the consistent performance of the combinations in various class configurations. In both 2-class and multi-class configurations, the use of only two sensors resulted in nearly equivalent F1 scores to the entire sensor set. For 2-class scenarios, the overall weighted F1 score was 0.988, with the combination of two sensors reaching an F1 score that was only 1.7% lower. Regarding 3-class scenarios, the scores were 0.976 and 0.969, respectively, showing the effectiveness in recognizing PRs. The classification of targeted anterior and posterior PRs becomes more challenging when the complexity of classification increases, particularly when models are trained on reduced sensor combinations. Compared to using the full sensor set, these configurations show a notable drop in F1 scores for the targeted PR classes.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

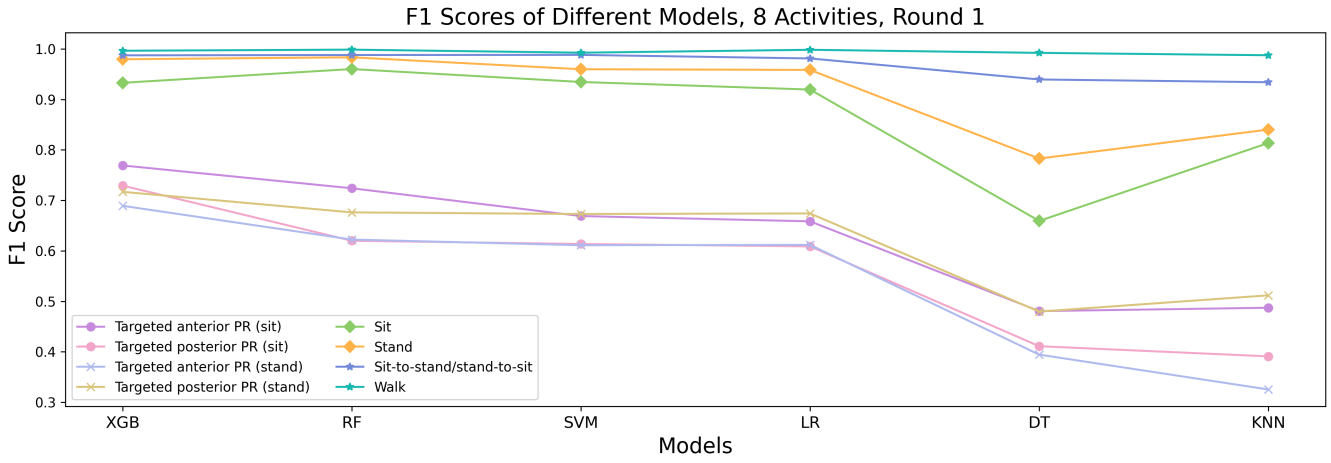


Fig. 3. Activity-specific F1 scores of models across 8 activities (8-class classification) in Round 1. Sorted in descending order by the average F1 scores of the targeted PR classes.

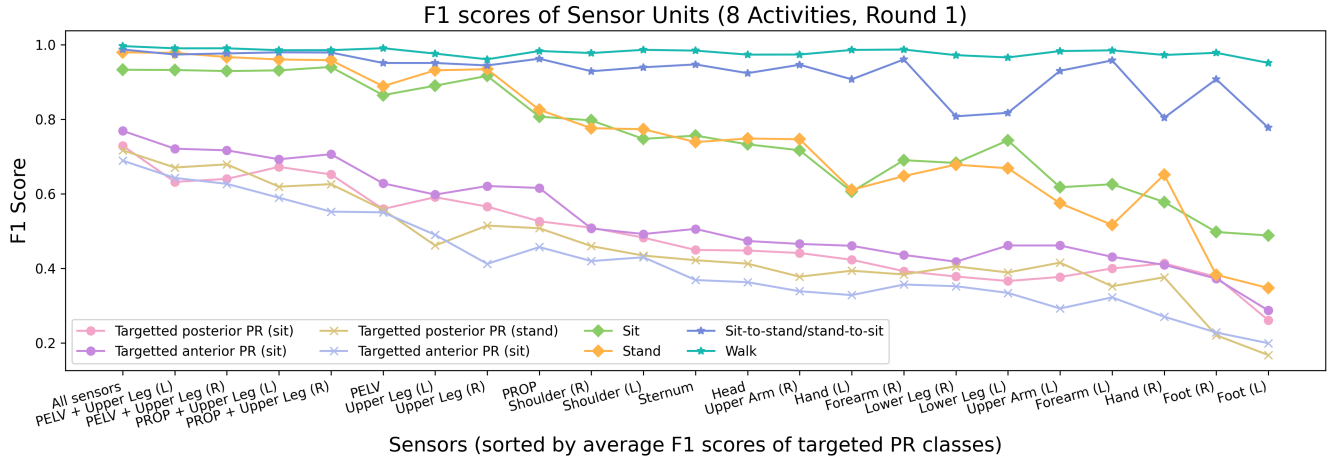


Fig. 4. F1 scores for single sensor units (8 classes, Round 1). Sorted in descending order by the average F1 scores of the targeted PR classes. (L) and (R) represent left and right sides.

In 8-class configurations, although the "PELV + left upper leg" combination maintains an overall weighted F1 score of 0.872, the class-specific F1 scores for the targeted PR classes are 0.721 (anterior PR during sitting), 0.632 (posterior PR during sitting), 0.643 (anterior PR during standing), and 0.671 (posterior PR during standing), highlighting the challenges of PR direction detection.

Among the combinations evaluated, "PELV + left upper leg" achieved the highest F1 scores in nearly all scenarios, closely followed by "PELV + right upper leg". This observation was supported by a paired t-test, as there was no statistically significant difference between these sensor combinations in both rounds. The slight advantage of the "PELV + left upper leg" combination can be attributed to the protocol design, as participants were required to turn clockwise during the walking tasks, causing a larger movement on the left upper leg than on the right upper leg (Table I). Taking into account the average performance, the combination "PELV + left upper leg" was chosen as the optimal combination.

For a more detailed comparison of three configurations (the "PELV" sensor alone, the "PELV + left upper leg" combination, and all sensors), we plotted the confusion matrices (Fig. 6), which illustrates the impact of different sensor setups on classification performance.

We observed in Fig. 6[a] that while the "PELV" sensor alone performs well for static postures and subconscious PR tasks, it struggles to accurately classify targeted PRs, particularly when distinguishing between anterior and posterior directions. This is further reflected in the confusion with nearby dynamic movements, for example, the model misclassified 2724 instances of "targeted anterior PR during sitting" as another class.

Adding the left upper leg sensor resulted in a marked improvement, as shown in Fig. 6[b]. The class "targeted anterior PR during sitting" shows enhanced recognition, with correct classifications increasing from 9784 to 10752, and the class "targeted posterior PR during sitting" improves from 6660 to 7963. It is also noticeable that the confusion between sitting tasks and standing tasks is reduced, as seen in the decrease in misclassifications across the matrix. Although the complete

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

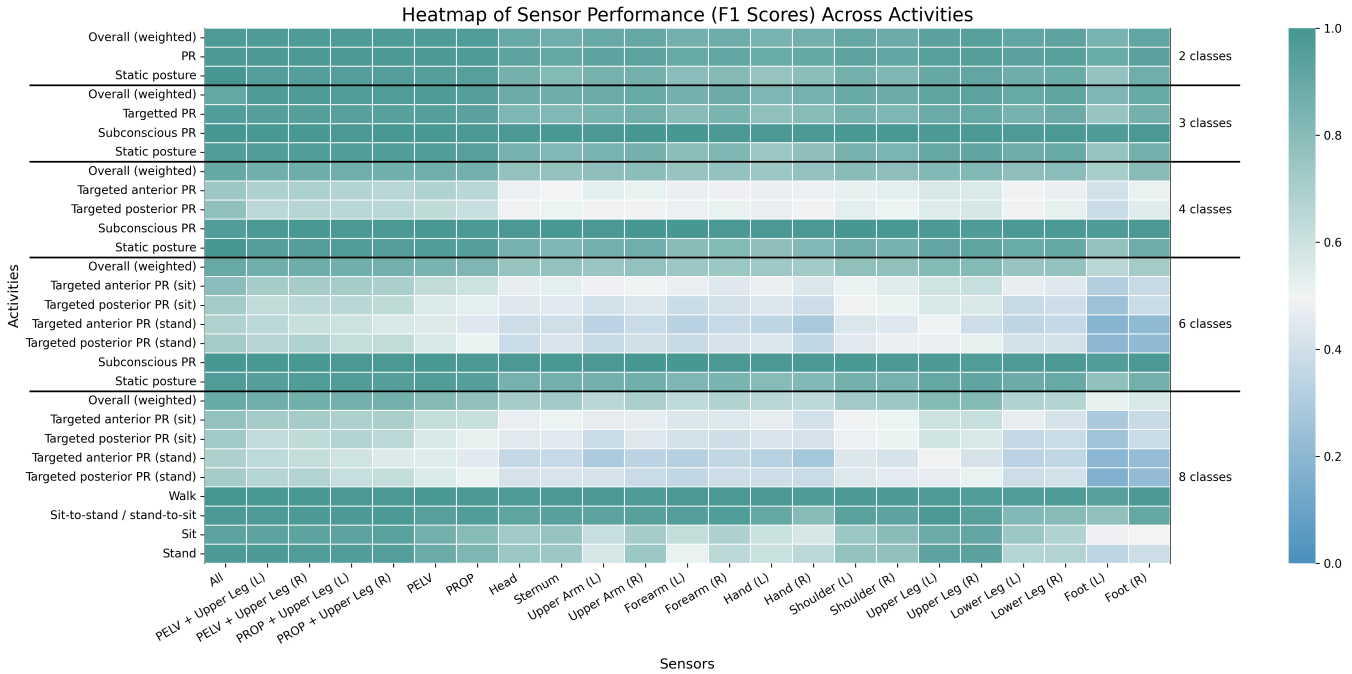


Fig. 5. Heatmap of F1-scores for different sensor configurations across activity classes. Sensors are shown along the x-axis and activity categories along the y-axis. (L) and (R) indicate the left and right sides, respectively.

TABLE VI
FRIEDMAN AND NEMENYI TEST RESULTS FOR COMPARING MODELS

Configuration	Round 1		Round 2	
	P-values (Friedman)	**Significant Models (Nemenyi)**	**P-values (Friedman)**	**Significant Models (Nemenyi)**
2-class	0.427	No	0.170	No
3-class	0.044	XGB > DT (p-value = 0.030)	0.037	XGB > DT (p-value = 0.030)
4-class	0.001	XGB > LR (p-value = 0.040); XGB > KNN (p-value = 0.002); XGB > DT (p-value = 0.005)	0.002	XGB > KNN (p-value = 0.001); XGB > DT (p-value = 0.009)
6-class	0.002	XGB > KNN (p-value = 0.049); XGB > DT (p-value = 0.013); SVM > DT (p-value = 0.021)	0.003	XGB > KNN (p-value = 0.008); RF > KNN (p-value = 0.032)
8-class	2.527e-05	XGB > KNN (p-value = 0.009); XGB > DT (p-value = 0.013); RF > KNN (p-value = 0.001); RF > DT (p-value = 0.001)	8.138e-05	XGB > KNN (p-value = 0.004); XGB > DT (p-value = 0.013); RF > KNN (p-value = 0.002); RF > DT (p-value = 0.008)

Note: Model A > Model B: Model A achieves a higher F1 score than Model B.

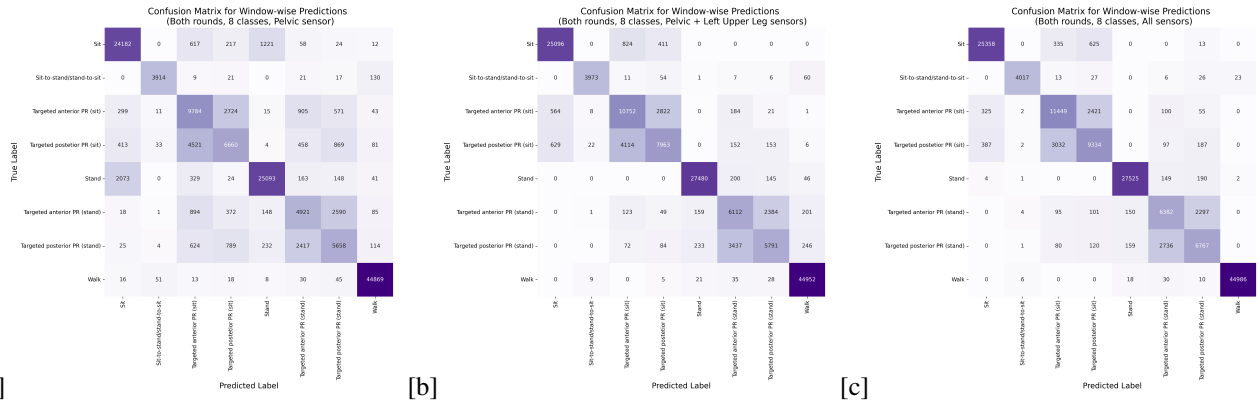
The p-values in brackets indicate the statistical significance of these comparisons based on the Nemenyi post-hoc test.

sensor setup shows a slight improvement over the two-sensor combination in specific tasks, such as differentiating between posterior and anterior PRs, the difference in performance is negligible. For static postures and subconscious PR tasks, such as sitting, standing, and walking, the combination of two sensors achieves near identical results to the combination of all sensors, with both configurations accurately classifying over 90%.

Friedman and post-hoc Nemenyi tests were performed to statistically evaluate differences between the sensor configurations ("All sensors", "PELV + left upper leg", and "PELV"). In the 2-class scenario, the differences were not significant. However, for multiple classes, the pairwise comparison results showed the following p-values: 1) "All sensors" vs "PELV + left upper leg": 0.0222; 2) "PELV + left upper leg" vs "PELV": 0.0223; 3) "All sensors" vs "PELV": 0.0010. Significant p-values indicate notable differences in performance between using all sensors and a reduced number of sensors. These results suggest that the "PELV + left upper leg" combination improved the model performance compared to using only the "PELV" sensor. The Wilcoxon signed rank test further supported these observations, showing a significant difference between the two directions (p-value = 1.907e-06), with anterior PRs consistently achieving higher F1 scores.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765



[a] [b] [c]

Fig. 6. Confusion matrices for 8-class classification using merged data of both rounds: (a) "PELV" only, (b) "PELV + left upper leg" combination, and (c) all sensors. True and predicted labels follow the same order: sit, sit-to-stand/stand-to-sit, targeted anterior PR (sit), targeted posterior PR (sit), Stand, targeted anterior PR (stand), targeted posterior PR (stand), and Walk.

IV. Discussion

A. Dataset

This study introduces a novel dataset that focuses on detecting PRs through IMU data. This dataset not only captured a wide range of static postures and dynamic movements, but also presented multiple class configurations, ranging from binary to 8-class scenarios. These multi-class configurations allow the evaluation of more complex movements involving PRs, such as STSTS and targeted anterior/posterior PRs.

Compared to other publicly available MoCap or IMU-based datasets (Table VII), such as the CMU Motion Capture Database [46], TotalCapture [47], or MoVi [48], they primarily focus on large amplitude movements, like activities of daily life, sports movements, and general locomotion. Although PRs naturally occur during walking and STSTS, these are typically not emphasized or labeled in existing datasets. By explicitly capturing and labeling these targeted PRs in both sitting and standing, the design of this unique dataset fills a gap in the discipline, where PRs remain underexplored.

Furthermore, the use of Xsens in this dataset opens up possibilities for data augmentation. Since it is a commercially and widely used system in the discipline of human activity recognition (HAR), techniques such as synthetic data generation or combining with other datasets based on Xsens can be applied to data augmentation [49], [50]. The combination of static postures and subconscious PRs also provides opportunities to explore how PRs influence transitions between common daily activities, such as sitting, standing, and walking. Insights from these analyses can support the development of more effective wearable technologies for posture recognition and targeted interventions for posture-related health problems, particularly those linked to sedentary behavior and LBP.

B. Classifier Performance and Model Selection

Through benchmarking six baseline classifiers (Table V), XGB achieved the best performance in all configurations, followed by RF. This result aligns with previous HAR studies, showing the consistent and robust performance of XGB [51], [52], [53]. The advancement of tree-based models can be attributed to the high-dimensional, nonlinear nature of the hand-crafted feature space, where tree-based ensembles effectively capture feature interactions and perform implicit feature selection [34], [39].

Although rule-based approaches have been used in some motion analysis tasks, they often rely on manually defined thresholds and struggle to generalize across different subjects or activity contexts [9]. Given the complexity of PRs and the variability in populations, we opted for data-driven ML models that are more adaptable and consistent with current HAR practices.

To maintain objectivity and accountability, all classifiers were deployed with default hyperparameters to avoid biasing results toward particular models or configurations through fine-tuning. Although this approach ensured a fair baseline comparison and was in line with our objectives, previous work has shown that tailored hyperparameter tuning can substantially improve performance [54]. Thus, the results presented should be considered as baseline performance that could be enhanced.

In general, XGB remains a practical and interpretable model for HAR. However, it requires time-consuming window-based feature extraction and struggles with direction-specific classification of subtle movements such as anterior versus posterior PRs. This limitation likely arises because tree-based models, such as XGB, operate on aggregated features from each window and are therefore limited in their temporal modeling capabilities, which are essential to distinguish fine-grained dynamics between the two PR directions [9]. This inherent limitation of tree-based models in capturing temporal dependencies contrasts with sequence-aware architectures, which are better suited for such temporal dynamics. Moreover, within short, fixed-length windows, PRs may resemble either static postures or even the opposite direction, especially when the amplitude is small or the motion is executed unevenly. In addition, the biomechanical difference between anterior and posterior PRs lies primarily in subtle shifts in pelvic orientation rather than large, easily distinguishable movements, which have been the primary focus in most previous HAR tasks. This challenge highlights the need for more temporally aware architectures in future work.

Recent advances in deep learning for IMU-based HAR, such as temporal convolutional networks and transformer-based

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

TABLE VII
A COMPARISON WITH PUBLICLY AVAILABLE MOCAP OR IMU-BASED DATASETS

Dataset	Size (Participants/Recordings)	Modality	Movement Types
CMU Motion Capture [46]	Over 2600 trials of motions in six categories and twenty-three subcategories by one hundred forty-four human subjects.	Marker-based motion capture	A broad range of human movements, including walking, running, jumping, dancing, and various daily activities.
TotalCapture [47]	Five subjects of each recorded twelve sequences.	IMUs, marker-based motion capture	Full-body movements cover daily activities, sports, and performance-related motions.
MoVi [48]	Nine hours of optical motion capture data, seventeen hours of video data from four different points of view recorded by stationary and hand-held cameras, and 6.6 hours of IMU data recorded from sixty female and thirty male actors performing a collection of twenty-one everyday actions and sports movements.	Marker-based motion capture, videos, and IMUs	Walking, jogging, running in place, side gallop, crawling, vertical jumping, jumping jacks, kicking, stretching, cross arms, sitting down on a chair, crossing legs while sitting, pointing, clapping hands, scratching one's head, throwing and catching, waving, pretending to take a picture, pretending to talk on the phone, pretending to check one's watch.
Our Dataset	Eight hundred fifty-eight recordings from thirty-nine subjects (six of the forty-five recruited participants were excluded due to unparseable or low-quality data).	IMUs	Sitting, standing, STSTS, walking, posterior and anterior PR during sitting and standing.

approaches, as well as hybrid CNN–LSTM (Convolutional Neural Networks - Long Short-Term Memory) models, have demonstrated the strong effectiveness of sequentially aware architectures [13], [55], [56]. Additionally, transfer learning and physics-informed self-supervised pre-training strategies have shown promising improvement in generalizability in HAR [57], [58]. Building on these state-of-the-art models, we see a clear path forward to distinguish subtle sequential differences.

C. Sensor Location Optimization

While the "PELV" sensor demonstrated reasonably satisfactory performance on its own, incorporating the left upper leg sensor greatly enhanced the detection accuracy of the direction of targeted PR (Fig. 6). The combination of "PELV + right upper leg" also shows satisfactory results, only slightly lower than "PELV + left upper leg" and higher than all the other configurations. This asymmetry between the upper legs likely reflects both the task design and the natural biomechanics, as most of the participants were right-leg dominant and our protocol required subjects to turn clockwise (Table I), which led the left leg to contribute more to balance and stabilization with a larger movement. These findings suggest that the added upper leg sensor provides valuable complementary information about lower body dynamics, which the "PELV" sensor alone could not fully capture. Statistical analyses confirmed significant differences between sensor configurations. Practically, the optimized combination prioritizes both accuracy and minimal amount of sensor use without a significant drop in performance compared to using the full set of sensors. It also ensures user comfort and wearability, as the sacrum-mounted sensor can be integrated into a legging or underwear pocket, while the upper leg sensor can be discreetly placed in the front trouser pocket using Velcro. This balance of usability and accuracy makes the setup feasible for daily posture monitoring outside the laboratory.

Although we are, to our best knowledge, the first study focusing on PR detection, our findings align with the other researchers who investigated the optimization of sensor placement for motion capture systems. For instance, the work conducted by Phan et al. demonstrated the possibility of classifying ten exercise groups using a single IMU mounted on the pelvis, receiving 89% accuracy compared to 96% accuracy using a set of ten IMUs [59].

Furthermore, the analysis of three different placements of pelvic sensors indicated that placing the sensor in the legging pocket on the sacrum ("PELV" R1) reached the best performance. The optimal placement of the pelvic sensor coincided with placing the sensor on the sacrum to analyze pelvic movement [60]. The differences among the three pelvic sensor placements tested underline that even small positional changes, such as alternative placements mounted on a belt, can degrade signal quality, especially for subtle movements such as PRs. These findings emphasize the importance of precise sensor placements in wearable systems and support the legging pocket as an optimal location, which offers a balance of signal stability, user comfort, and suitability for daily monitoring.

D. Limitations and Future Work

In the following, we discuss the limitations faced in this study, together with proposed future work to address them.

1) Controlled environment and sensor placement: The data were collected in a controlled environment with a pre-defined protocol under the supervision of the researchers, which may

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

not fully reflect the variability of real-world scenarios. Although three pelvic sensor placements were compared and the legging pocket on the sacrum ("PELV" R1) was optimal, this measurement was conducted during short (< 2 h), controlled sessions with consistent, tight-fitting clothing. While tight clothing helps maintain stable sensor placement, looser clothing is more commonly worn in daily life and may allow sensor movement and introduce noise. Our ongoing work is validating these findings in semi-controlled environments where PRs occur spontaneously, and sensor placement varies, as an intermediate step toward the real world. Future designs should therefore prioritize secure, tight-fit attachment solutions, such as integrated pockets in leggings or underwear, or include adaptive fastening using belts with quick calibration methods to ensure consistent data quality across clothing conditions and users.

2) Population diversity: Even though the sample size of 39 participants is sufficient for initial analysis and included both genders, the subjects were recruited through the personal network, resulting in a relatively homogeneous group in terms of age, BMI, and university-affiliated populations. This lack of population diversity may limit the generalizability of the findings to broader populations. Future work will include individuals with LBP and other populations to ensure greater inclusion.

3) Baseline models and temporal dynamics: This study employed classic ML classifiers with default hyperparameters. Although this approach avoided bias toward any particular model, it also reserves the space to improve the classification through tailored hyperparameter tuning and the use of advanced deep learning models. In particular, classifiers struggled with direction-specific PRs, where subtle sequential differences between anterior and posterior PRs were difficult to capture with static fixed-length windows. The use of a single fixed window size, although suitable for the targeted activities in this study, may not generalize well to other movements with different temporal dynamics. To address these challenges, our ongoing work is exploring state-of-the-art sequence-to-sequence models, including Bidirectional Long Short-Term Memory Network (Bi-LSTM) and its combinations with Convolutional Neural Networks (CNNs) [61], [62], and Multi-Stage Temporal Convolutional Network (MS-TCN++) [63], to improve performance and robustness in PR direction recognition. Future work will focus on transitioning the study to real-world scenarios and exploring state-of-the-art ML techniques. Moreover, future research could investigate adaptive windowing strategies, which dynamically adjust the window length based on the movement duration. Such approaches could better accommodate both short transitions and longer static activities, reducing class imbalance and potentially improving PR direction recognition [64]. In addition to high temporal resolution, future models should also incorporate explainability mechanisms, as it will be critical when transitioning PR feedback systems toward clinical or ergonomic applications, where model transparency and trustworthiness are essential [65]. Last but not least, although this study grouped all PRs in the same direction as one class regardless of amplitude, future work may investigate amplitude-specific PR recognition to enable semi-personalized or personalized modeling.

By addressing these critical areas, we aim to translate this research into practical applications that advance daily posture monitoring and prevent musculoskeletal problems, particularly LBP.

V. Conclusion

This study investigated the feasibility of applying IMU-based motion capture systems and ML classifiers to recognize PRs. First, we collected a novel dataset that focused on PRs, addressing the lack of datasets in this domain. Second, we illustrated the potential of using IMUs and ML models to detect PR patterns and discovered the advancements and limitations of XGB classifiers. Third, our sensor placement analysis shows that fixing the sensor in the legging back pocket on the sacrum is optimal, while adding a left upper leg sensor primarily improves the distinction between targeted anterior and posterior PR directions and reduces confusion between sitting and standing tasks. By taking the first step of applying wearable sensors to PR detection, this research further paved the way for self-care for posture-related issues, such as LBP, which affects millions of people worldwide.

Acknowledgment

The authors would like to thank all the participants for their time and kind support.

References

[1] A. Loyen, H. P. van der Ploeg, A. Bauman, J. Brug, and J. Lakerveld, "European sitting championship: Prevalence and correlates of self-reported sitting time in the 28 european union member states," *PLOS ONE*, vol. 11, no. 3, pp. 1–17, Mar. 2016. DOI: 10.1371/journal.pone.0149320.

[2] S.-M. Chen, M.-F. Liu, J. Cook, S. Bass, and S. K. Lo, "Sedentary lifestyle as a risk factor for low back pain: A systematic review," *International archives of occupational and environmental health*, vol. 82, pp. 797–806, 2009. DOI: 10.1007/s00420-009-0410-0.

[3] M. L. Ferreira et al., "Global, regional, and national burden of low back pain, 1990–2020, its attributable risk factors, and projections to 2050: A systematic analysis of the global burden of disease study 2021," *The Lancet Rheumatology*, vol. 5, no. 6, e316–e329, 2023. DOI: 10.1016/S2665-9913(23)00098-X.

[4] É. A. Csuhai, A. C. Nagy, Z. Váradi, and I. Veres-Balajti, "Functional analysis of the spine with the idiag spinalmouse system among sedentary workers affected by non-specific low back pain," *International Journal of Environmental Research and Public Health*, vol. 17, no. 24, p. 9259, 2020. DOI: 10.3390/ijerph17249259.

[5] F. B. Benatti and M. Ried-Larsen, "The effects of breaking up prolonged sitting time: A review of experimental studies," *Medicine & Science in Sports & Exercise*, vol. 47, no. 10, pp. 2053–2061, 2015. DOI: 10.1249/MSS.0000000000000654.

[6] P. W. Hodges, V. S. Gurfinkel, S. Brumagne, T. Smith, and P. Cordo, "Coexistence of stability and mobility in postural control: Evidence from postural compensation for respiration," *Experimental brain research*, vol. 144, pp. 293–302, 2002. DOI: 10.1007/s00221-002-1040-x.

[7] P. Cordo, V. Gurfinkel, T. Smith, P. Hodges, S. Verschueren, and S. Brumagne, "The sit-up: Complex kinematics and muscle activity in voluntary axial movement," *Journal of Electromyography and Kinesiology*, vol. 13, no. 3, pp. 239–252, 2003. DOI: 10.1016/s1050-6411(03)00023-3.

[8] K. Claeys, W. Dankaerts, L. Janssens, and S. Brumagne, "Altered preparatory pelvic control during the sit-to-stance-to-sit movement in people with non-specific low back pain," *Journal of Electromyography and Kinesiology*, vol. 22, no. 6, pp. 821–828, 2012. DOI: 10.1016/j.jelekin.2012.04.007.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

[9] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Computing Surveys (CSUR)*, vol. 46, no. 3, pp. 1–33, 2014. DOI: 10.1145/2499621.

[10] V. Bianchi, M. Bassoli, G. Lombardo, P. Fornacciari, M. Mordonini, and I. De Munari, "Iot wearable sensor and deep learning: An integrated approach for personalized human activity recognition in a smart home environment," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8553–8562, 2019. DOI: 10.1109/JIOT.2019.2920283.

[11] A. Y. A. Amer, J.-m. Aerts, B. Vanrumste, and S. Luca, "A localized learning approach applied to human activity recognition," *IEEE Intelligent Systems*, vol. 36, no. 3, pp. 58–71, 2021. DOI: 10.1109/MIS.2020.2964738.

[12] E. Ramanujam, T. Perumal, and S. Padmavathi, "Human activity recognition with smartphone and wearable sensors using deep learning techniques: A review," *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13 029–13 040, 2021. DOI: 10.1109/JSEN.2021.3069927.

[13] C. Wang et al., "Detection of eating gestures in older persons using imu sensors with multistage temporal convolutional network," *IEEE Sensors Journal*, vol. 24, no. 21, pp. 35 231–35 244, 2024. DOI: 10.1109/JSEN.2024.3460651.

[14] P.-K. Yang et al., "Freezing of gait assessment with inertial measurement units and deep learning: Effect of tasks, medication states, and stops," *Journal of NeuroEngineering and Rehabilitation*, vol. 21, no. 1, p. 24, 2024. DOI: 10.1186/s12984-024-01320-1.

[15] W. Wang, Y. Peng, Y. Sun, J. Wang, and G. Li, "Towards wearable and portable spine motion analysis through dynamic optimization of smartphone videos and imu data," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 10, pp. 5929–5940, 2024. DOI: 10.1109/JBHI.2024.3419591.

[16] C. Ma, W. Li, R. Gravina, J. Cao, Q. Li, and G. Fortino, "Activity level assessment using a smart cushion for people with a sedentary lifestyle," *Sensors*, vol. 17, no. 10, p. 2269, 2017. DOI: 10.3390/s17102269.

[17] X. Ran, C. Wang, Y. Xiao, X. Gao, Z. Zhu, and B. Chen, "A portable sitting posture monitoring system based on a pressure sensor array and machine learning," *Sensors and Actuators A: Physical*, vol. 331, p. 112 900, 2021. DOI: 10.1016/j.sna.2021.112900.

[18] J. M. Hänggi, S. Spinnler, E. Christodoulides, E. Gramespacher, W. Taube, and A. Doherty, "Sedentary behavior in children by wearable cameras: Development of an annotation protocol," *American journal of preventive medicine*, vol. 59, no. 6, pp. 880–886, 2020. DOI: 10.1016/j.amepre.2020.06.033.

[19] R. Kumar, A. Bayliff, D. De, A. Evans, S. K. Das, and M. Makos, "Care-chair: Sedentary activities and behavior assessment with smart sensing on chair backrest," in *2016 IEEE International Conference on Smart Computing (SMARTCOMP)*, 2016, pp. 1–8. DOI: 10.1109/SMARTCOMP.2016.7501682.

[20] A. Piau et al., "A smart shoe insole to monitor frail older adults' walking speed: Results of two evaluation phases completed in a living lab and through a 12-week pilot study," *JMIR mHealth and uHealth*, vol. 9, no. 7, e15641, 2021. DOI: 10.2196/15641.

[21] S. Kozey-Keadle, A. Libertine, K. Lyden, J. Staudenmayer, and P. S. Freedson, "Validation of wearable monitors for assessing sedentary behavior," *Medicine & Science in Sports & Exercise*, vol. 43, no. 8, pp. 1561–1567, 2011. DOI: 10.1249/MSS.0b013e31820ce174.

[22] S. Mekruksavanich, N. Hnoohom, and A. Jitpattanakul, "Smartwatch-based sitting detection with human activity recognition for office workers syndrome," in *2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON)*, 2018, pp. 160–164. DOI: 10.1109/ECTI-NCON.2018.8378302.

[23] R. Tanthuwapathom, U. Manupibul, W. Jarumethitanont, W. Limroongreungrat, S. Ongwattanakul, and W. Charoensuk, "Reliability of sitting posture between physical therapist video-based evaluation and smart imu system using rapid upper limb assessment (rula)," *Scientific Reports*, vol. 15, no. 1, p. 1441, 2025. DOI: 10.1038/s41598-025-85159-z.

[24] Q. Xing, R. Hong, Y. Shen, and Y. Shen, "Design and validation of depth camera-based static posture assessment system," *Iscience*, vol. 26, no. 10, 2023. DOI: 10.1016/j.isci.2023.107974.

[25] C.-H. Lee, T. Mendoza, C.-H. Huang, and T.-L. Sun, "Vision-based postural balance assessment of sit-to-stand transitions performed by younger and older adults," *Gait & Posture*, vol. 117, pp. 245–253, 2025. DOI: 10.1016/j.gaitpost.2025.01.001.

[26] L. Simpson, M. M. Maharaj, and R. J. Mobbs, "The role of wearables in spinal posture analysis: A systematic review," *BMC musculoskeletal disorders*, vol. 20, no. 1, p. 55, 2019. DOI: 10.1186/s12891-019-2430-6.

[27] V. Figueira et al., "Wearables for monitoring and postural feedback in the work context: A scoping review," *Sensors*, vol. 24, no. 4, p. 1341, 2024. DOI: 10.3390/s24041341.

[28] M. García-Jaén, S. Sebastia-Amat, G. Sanchis-Soler, and J. M. Cortell-Tormo, "Lumbo-pelvic rhythm monitoring using wearable technology with sensory biofeedback: A systematic review," *Healthcare*, vol. 12, no. 7, 2024, ISSN: 2227-9032. DOI: 10.3390/healthcare12070758.

[29] M. T. B.V., *Xsens mvn user manual*, English, version Revision Z, Jan. 13, 2023, 160 pp. [Online]. Available: https://www.movella.com/hubfs/MVN_User_Manual.pdf.

[30] K. Shoemake, "Animating rotation with quaternion curves," 3, vol. 19, New York, NY, USA: Association for Computing Machinery, Jul. 1985, pp. 245–254. DOI: 10.1145/325165.325242.

[31] J. H. Challis, "Quaternions as a solution to determining the angular kinematics of human movement," *BMC Biomedical Engineering*, vol. 2, no. 1, p. 5, 2020. DOI: 10.1186/s42490-020-00039-z.

[32] Y. Wang, S. Cang, and H. Yu, "A survey on wearable sensor modality centred human activity recognition in health care," *Expert Systems with Applications*, vol. 137, pp. 167–190, 2019. DOI: 10.1016/j.eswa.2019.04.057.

[33] S. G. K. Patro and K. K. Sahu, "Normalization: A preprocessing stage," *CoRR*, vol. abs/1503.06462, 2015. DOI: 10.48550/arXiv.1503.06462. arXiv: 1503.06462.

[34] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001. DOI: 10.1023/A:1010933404324.

[35] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, ISSN: 0885-6125. DOI: 10.1023/A:1022627411411.

[36] E. Fix and J. L. Hodges, *Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties*. [Wiley, International Statistical Institute (ISI)], 1989, vol. 57, pp. 238–247. DOI: 10.2307/1403797. Accessed: Aug. 27, 2025.

[37] W.-Y. Loh, "Classification and regression trees," *Wiley interdisciplinary reviews: data mining and knowledge discovery*, vol. 1, no. 1, pp. 14–23, 2011. DOI: 10.1201/9781315139470.

[38] D. R. Cox, "The regression analysis of binary sequences," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 20, no. 2, pp. 215–232, 1958. [Online]. Available: https://www.jstor.org/stable/2983890.

[39] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16, San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 785–794, ISBN: 9781450342322. DOI: 10.1145/2939672.2939785.

[40] F. Pedregosa et al., "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011. [Online]. Available: https://www.jmlr.org/papers/v12/pedregosa11a.html.

[41] T. Hastie, J. Friedman, and R. Tibshirani, "Model assessment and selection," in *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer New York, 2001, pp. 193–224, ISBN: 978-0-387-21606-5. DOI: 10.1007/978-0-387-21606-5_7.

[42] A. A. Taha and A. Hanbury, "Metrics for evaluating 3d medical image segmentation: Analysis, selection, and tool," *BMC medical imaging*, vol. 15, pp. 1–28, 2015. DOI: 10.1186/s12880-015-0068-x.

[43] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *Journal of the american statistical association*, vol. 32, no. 200, pp. 675–701, 1937. DOI: 10.2307/2279372.

[44] P. B. Nemenyi, *Distribution-free multiple comparisons*. Princeton University, 1963. [Online]. Available: https://www.proquest.com/openview/c1f3e8829e8351e9c2a1c5e51778c6cf/1.

[45] R. F. Woolson, "Wilcoxon signed-rank test," *Wiley encyclopedia of clinical trials*, pp. 1–3, 2008. DOI: 10.1002/9780471462422.eoct979.

[46] W. Dewaele, B. Hahne, C. Graphics Lab, and A. A. Diwan, *Fbx conversion of the cmu graphics lab motion capture database*, en, 2024. DOI: 10.4121/0448AAB2-3332-449F-A8E2-D208CB58C7DF.V1.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2026.3665765

[47] M. Trumble, A. Gilbert, C. Malleson, A. Hilton, and J. Collomosse, "Total capture: 3d human pose estimation fusing video and inertial sensors," in *2017 British Machine Vision Conference (BMVC)*, 2017. DOI: 10.5244/C.31.14.

[48] S. Ghorbani et al., "Movi: A large multi-purpose human motion and video dataset," *PLOS ONE*, vol. 16, no. 6, pp. 1–15, Jun. 2021. DOI: 10.1371/journal.pone.0253157.

[49] J. DelPreto et al., "Actionsense: A multimodal dataset and recording framework for human activities using wearable sensors in a kitchen environment," in *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, Eds., vol. 35, Curran Associates, Inc., 2022, pp. 13 800–13 813. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2022/file/5985e81d65605827ac35401999aea22a-Paper-Datasets_and_Benchmarks.pdf.

[50] M. Palermo, S. M. Cerqueira, J. André, A. Pereira, and C. P. Santos, "From raw measurements to human pose-a dataset with low-cost and high-end inertial-magnetic sensor data," *Scientific Data*, vol. 9, no. 1, p. 591, 2022. DOI: 10.1038/s41597-022-01690-y.

[51] M. A. Shafique and S. S. Marchán, "Investigating the impact of information sharing in human activity recognition," *Sensors*, vol. 22, no. 6, p. 2280, 2022. DOI: 10.3390/s22062280.

[52] S. Abbas et al., "Advancing healthcare and elderly activity recognition: Active machine and deep learning for fine- grained heterogeneity activity recognition," *IEEE Access*, vol. 12, pp. 44 949–44 959, 2024. DOI: 10.1109/ACCESS.2024.3380432.

[53] L. S. Ambati and O. El-Gayar, "Human activity recognition: A comparison of machine learning approaches," *Journal of the Midwest Association for Information Systems (JMWAIS)*, vol. 2021, no. 1, p. 4, 2021. [Online]. Available: https://aisel.aisnet.org/jmwais/vol2021/iss1/4.

[54] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, no. 10, pp. 281–305, 2012. [Online]. Available: http://jmlr.org/papers/v13/bergstra12a.html.

[55] Y. Shavit and I. Klein, "Boosting inertial-based human activity recognition with transformers," *IEEE Access*, vol. 9, pp. 53 540–53 547, 2021. DOI: 10.1109/ACCESS.2021.3070646.

[56] P. Tokas, V. B. Semwal, and S. Verma, "A real-time deployable attention-driven cnn–lstm framework for human activity recognition using wearable sensor," *IEEE Sensors Journal*, vol. 25, no. 21, pp. 40 461–40 471, 2025. DOI: 10.1109/JSEN.2025.3610667.

[57] S. Ashry, S. Das, M. Rafiei, J. Baumbach, and L. Baumbach, "Transfer learning of human activities based on imu sensors: A review," *IEEE Sensors Journal*, vol. 25, no. 3, pp. 4115–4126, 2025. DOI: 10.1109/JSEN.2024.3510097.

[58] H. Yuan et al., "Self-supervised learning for human activity recognition using 700,000 person-days of wearable data," *NPJ digital medicine*, vol. 7, no. 1, p. 91, 2024. DOI: 10.1038/s41746-024-01062-3.

[59] V. Phan, K. Song, R. S. Silva, K. G. Silbernagel, J. R. Baxter, and E. Halilaj, "Seven things to know about exercise classification with inertial sensing wearables," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 6, pp. 3411–3421, 2024. DOI: 10.1109/JBHI.2024.3368042.

[60] S. Brumagne, R. Lysens, and A. Spaepen, "Lumbosacral position sense during pelvic tilting in men and women without low back pain: Test development and reliability assessment," *Journal of Orthopaedic & Sports Physical Therapy*, vol. 29, no. 6, pp. 345–351, 1999. DOI: 10.2519/jospt.1999.29.6.345.

[61] S. Ashry, T. Ogawa, and W. Gomaa, "Charm-deep: Continuous human activity recognition model based on deep neural network using imu sensors of smartwatch," *IEEE Sensors Journal*, vol. 20, no. 15, pp. 8757–8770, 2020. DOI: 10.1109/JSEN.2020.2985374.

[62] X. Chen et al., "A novel cnn-bilstm ensemble model with attention mechanism for sit-to-stand phase identification using wearable inertial sensors," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 1068–1077, 2024. DOI: 10.1109/TNSRE.2024.3366907.

[63] S. Li, Y. A. Farha, Y. Liu, M.-M. Cheng, and J. Gall, "Ms-tcn++: Multi-stage temporal convolutional network for action segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 6, pp. 6647–6658, 2023. DOI: 10.1109/TPAMI.2020.3021756.

[64] C. Ma, W. Li, J. Cao, J. Du, Q. Li, and R. Gravina, "Adaptive sliding window based activity recognition for assisted livings," *Information Fusion*, vol. 53, pp. 55–65, 2020, ISSN: 1566-2535. DOI: 10.1016/j.inffus.2019.06.013.

[65] P. Tokas, V. B. Semwal, and S. Jain, "A lightweight and explainable hybrid deep learning model for wearable sensor-based human activity recognition," *IEEE Sensors Journal*, vol. 25, no. 12, pp. 22 618–22 628, 2025. DOI: 10.1109/JSEN.2025.3564045.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on February 25,2026 at 09:33:46 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.