Motion analysis based on accelerometric data

Tomáš Alexa[1]

[1]Faculty of Information Technology, CTU, Prague, Czech Republic

Abstract

This paper presents a machine learning pipeline for Human Activity Recognition (HAR) using wearable sensors, bridging clinical neurology and digital signal processing. Kinematic data from eight individuals were processed into a multidimensional feature space across the time, frequency, and wavelet domains. These extracted features were subsequently evaluated using a nested cross-validation framework and a 2,916-permutation grid search.

The optimal configuration – a Support Vector Machine (SVM) with a cubic kernel and a 10.24-second temporal window – achieved a peak validation accuracy of 98.12 %, yielding an overall mean test accuracy of 85.08 % during Leave-One-Subject-Out (LOSO) cross-validation. While the model demonstrated strong generalisation with 95 % to 100 % test accuracies for the majority of the participant pool, a localised test failure (48 % accuracy) in a single individual revealed an algorithmic vulnerability to static sensor tilt. Ultimately, this research validates the efficacy of optimised non-linear SVMs for continuous motion analysis while emphasising the necessity of dynamic gravity compensation for reliable, real-world clinical deployment.

Keywords

human activity recognition, gait analysis, time-frequency analysis, fast Fourier transform, Daubechies wavelet, accelerometer, wearable sensors, MATLAB Mobile

1. Introduction

Human Activity Recognition (HAR) has evolved into a cornerstone of pervasive healthcare, enabling continuous patient monitoring beyond traditional clinical settings. In clinical neurology, this continuous monitoring offers unprecedented opportunities for early diagnosis and disease tracking. However, to assess neurological health accurately in a free-living environment, a system must first reliably distinguish between baseline functional tasks.

This work serves as an initial proof-of-concept and baseline study, focusing on the classification of three fundamental mobility tasks: level-ground walking, stair ascent, and stair descent. Establishing a high-accuracy pipeline for these primary tasks under strict Leave-One-Subject-Out (LOSO) conditions is a prerequisite for deploying single-sensor architectures to detect subtle pathological gait anomalies.

The primary objective of the proposed system is to accurately distinguish these fundamental movements. Furthermore, this research investigates a critical question in modern medical engineering: whether highly accessible, consumer-grade hardware – such as standard smartphones – can serve as a viable alternative to expensive, professional diagnostic machinery.

The remainder of this paper is organised as follows: Section 2 describes the data acquisition protocol. Section 3 and Section 4 outline signal preprocessing and multi-domain feature engineering. Section 5 details the nested cross-validation and hyperparameter optimisation framework. Section 6 evaluates classification performance and discusses diagnostic anomalies, while Section 7 concludes the paper.

2. Data Acquisition

2.1. Hardware and Software Environment

Data acquisition was performed using a realme 8 smartphone (Model RMX3085). The device is equipped with a tri-axial MEMS accelerometer, capable of measuring linear acceleration and gravitational forces across the X, Y and Z axes.

ITAT'26: Information Technologies – Applications and Theory, September 25–29, 2026, Biele Karpaty, Slovakia

alexatom@fit.cvut.cz (T. Alexa)

0009-0008-7874-4743 (T. Alexa)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

published 2026-11-09

CEUR-WS.org/Vol-4323/paper12.pdf



Figure 1: Human activity recognition pipeline - data acquisition

To interface with the hardware, the MATLAB Mobile application was utilised. This setup enabled high-fidelity data logging directly within MATLAB. The sampling frequency was set to 100 Hz to ensure all walking frequencies were captured, yielding a temporal resolution of 10 ms between data points. The captured datasets represent raw acceleration, including the static component of Earth's gravity $g \approx 9.81\,\mathrm{m/s^2}$.

2.2. Experimental Protocol

Test subjects were eight healthy students (7 male and 1 female) from CTU, aged 22-23.

To capture the ascent and descent of the stairs, all eight test subjects were instructed to ascend and descend 14 floors of the stairs in the TH:A faculty building on the Dejvice campus. To obtain motion data on flat ground, each test subject walked for 3 minutes directly from the Faculty of Civil Engineering on a sidewalk straight without any stops. This protocol yielded a total of 24 distinct recordings, each lasting three to four minutes.

Throughout all trials, the smartphone was securely fastened to the lower back (lumbar region) of each test subject using a belt. This specific topological placement was selected to align directly with the human body's centre of mass. In clinical biomechanics and gait analysis, lumbar sensor placement is the gold standard used by commercial medical-grade systems during standardised clinical tests.

3. Signal Preprocessing

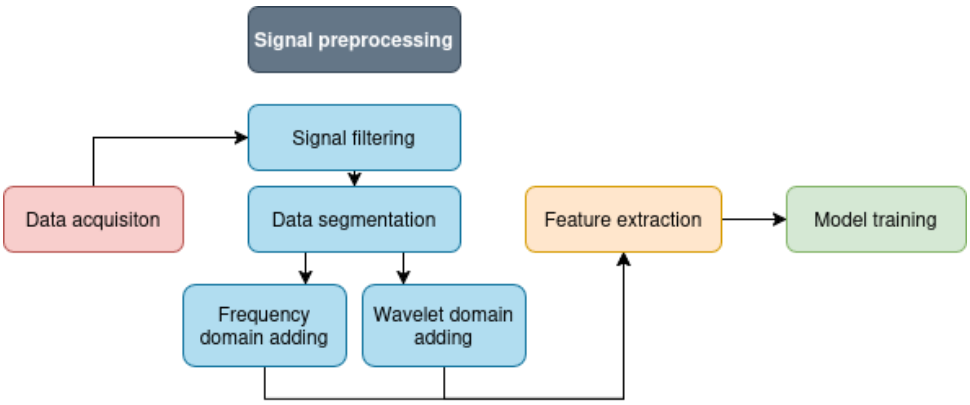
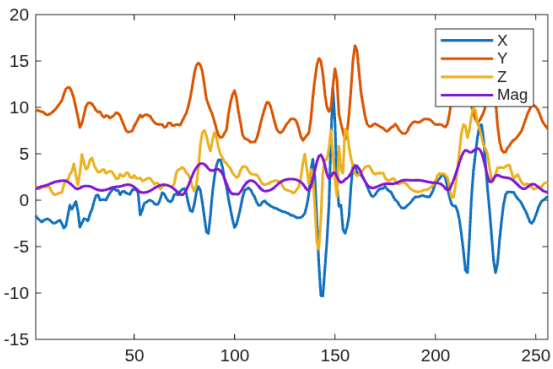


Figure 2: Human activity recognition pipeline - signal preprocessing

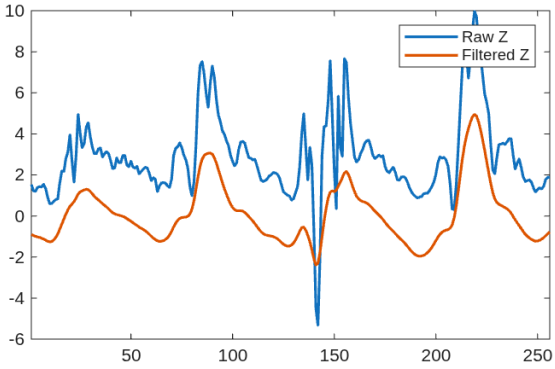
Prior to feature extraction, raw sensor data required preprocessing to correct hardware inconsistencies. Smartphone inertial sensors frequently exhibit slight fluctuations in their sampling rates, a phenomenon known as jitter. To establish a strictly uniform temporal grid, raw time-series values were resampled using linear interpolation. This step ensured a precise spacing of 10 ms (equivalent to a 100 Hz sampling rate) between all discrete data points and was implemented utilising the MATLAB `retime` [1] function.

3.1. Signal Filtering

Our data still contains captured noise and the signal's 0 Hz (DC) component representing the constant pull of Earth's gravity. To solve that, we used a first-order Butterworth passband filter with upper



(a) X, Y, Z axes and Magnitude



(b) Raw vs filtered Z axis

Figure 3: Axes after preprocessing

cut-off frequencies of 5 Hz and 20 Hz, and a lower cut-off frequency of 0.5 Hz. This approach resulted in two parallel datasets of each observation:

Narrowband filtered dataset Processed with a passband of 0.5 Hz to 5 Hz to isolate only the fundamental frequencies corresponding to standard walking and stepping cadence

Wideband filtered dataset Processed with a passband of 0.5 Hz to 20 Hz to preserve higher-frequency transient impacts and harmonics.

Evaluating both bandwidths in parallel enabled the downstream hyperparameter search and feature selection algorithms to empirically determine which fundamental cadence (narrowband) or transient impact dynamics (wideband) yielded superior subject-independent generalisation without a priori assumptions.

Filtering was performed using MATLAB's `butter` [1] function to compute the transfer function coefficients, followed by `filtfilt` [1]. The latter performs zero-phase digital filtering by processing the data in both the forward and reverse directions, effectively preventing temporal phase distortion.

Additionally, the Euclidean signal magnitude was computed in the time domain to provide a rotation-invariant metric:

$$\text{Mag} = \sqrt{X^2 + Y^2 + Z^2}$$

Signal Magnitude (Mag) is treated as a 4th independent signal channel alongside raw X, Y, and Z axes. All downstream feature extraction steps were computed on Mag as well as the individual axes. Although Mag is mathematically derived from the three spatial axes, its inclusion significantly improves classification robustness.

3.2. Data Segmentation

Continuous recordings lasting three to four minutes yielded approximately 20,000 time steps per trial, which is computationally inefficient and unsegmented for discrete classification. Furthermore, to account for transient states – specifically the delay between initiating the recording and the subject achieving a steady-state walking rhythm – the first and last 5 seconds of each recording were cropped. This ensured the retained kinematic data strictly matched its assigned target label.

The continuous data was subsequently segmented into discrete temporal windows. To optimise computational efficiency for the FFT algorithm, window lengths were restricted to powers of two: 2.56, 5.12, and 10.24 seconds (corresponding to 256, 512, and 1024 time steps, respectively). To evaluate the impact of temporal continuity, windows were generated with overlap configurations of 0 %, 25 %, and 50 %. Applying this combination of filter bandwidths, window lengths, and overlap percentages resulted in 18 distinct dataset configurations. Through this segmentation process, the initial 24 continuous recordings were partitioned into thousands of discrete samples.

3.3. Domains Adding

So far, we have only the time domain. To provide a broader range of features, we added two additional domains: frequency and wavelet domains. After segmenting our data, we computed the frequency-domain spectra for each filtered and non-filtered axis *X*, *Y*, and *Z*, as well as for the signal magnitude, and the wavelet-domain spectra for each filtered axis and the signal magnitude. For the frequency-domain calculation, we used the `fft` [1] function, which computes the DFT of the input signal using the FFT algorithm. For the wavelet transform, we used the MATLAB function `wavedec` [1] that returns the wavelet decomposition of the input signal. The function takes three parameters: the first is the input data, the second is the level of decomposition, and the third is the analysing wavelet. As an analysis wavelet, we used the Daubechies 4 (`db4`) mother wavelet at the level of decomposition 3.

4. Feature Extraction

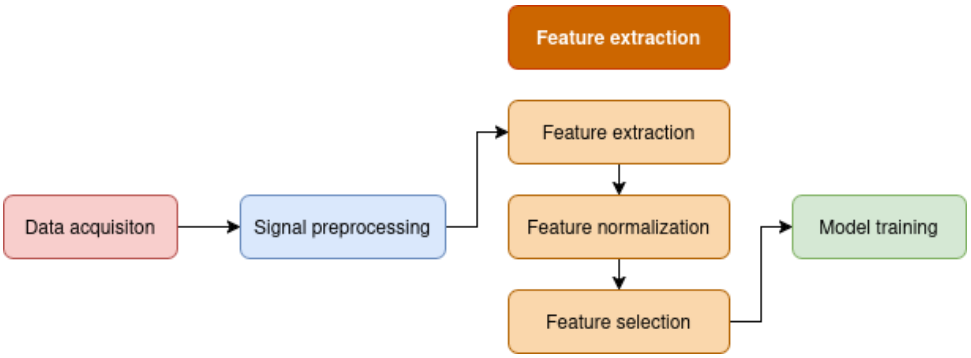


Figure 4: Human activity recognition pipeline - feature extraction

Extracting descriptive statistical features from raw inertial time-series data is essential for capturing stable kinematic patterns across subjects [2, 3]. Building upon established feature extraction approaches in human activity recognition [4, 5], an exhaustive initial feature space was generated across all filtered and unfiltered axes, as well as the signal magnitude. Redundant metrics, such as the time between peaks, were omitted because frequency-domain transformations inherently capture these properties. This comprehensive multi-domain dataset was subsequently optimised using automated feature selection algorithms.

4.1. Time Domain Features

Nine independent time-domain features (mean, median, standard deviation, variance, skewness, kurtosis, minimum, maximum, and root mean square) were extracted for each of the seven data streams. Additionally, 42 pairwise correlation coefficients were computed, yielding a total of 105 time-domain features.

Specifically, the feature duplication occurred during automated feature extraction, where pairwise cross-correlations between accelerometer axes were computed symmetrically in both directions. Because Pearson correlation is commutative, this resulted in identical feature vectors for all axis pairs.

4.2. Frequency Domain Features

Dominant frequency and spectral energy were extracted across all seven data streams, yielding 14 frequency-domain features.

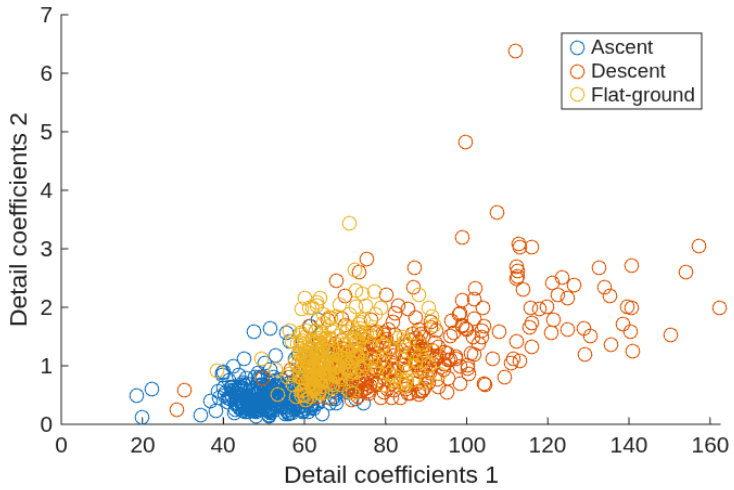


Figure 5: Wavelet Daubechies coefficients across decomposition levels.

4.3. Wavelet Domain Features

Wavelet energy E was calculated from detail coefficients across three decomposition levels for four data streams (filtered X, Y, Z axes, and signal magnitude), resulting in 12 wavelet-domain features. Detail coefficient mapping exhibits distinct task clustering, as shown in Figure 5.

$$E = \sqrt{\sum_{i=1}^{n} a_i^2}$$

4.4. Feature Normalization

Z-score normalisation was applied to standardise features to a mean of 0 and a standard deviation of 1 using the MATLAB `normalize` [1] function. To prevent data leakage, normalisation parameters were derived exclusively from the training subsets. To optimise computational efficiency during nested cross-validation, normalised datasets were precalculated for all 6-, 7-, and 8-subject training permutations.

4.5. Feature Selection

The concatenated multi-domain pipeline produced 131 independent features. Initial screening identified ReliefF ($k = 10$) as the most discriminative selection algorithm over MRMR and Chi-square. Feature ranking computations were restricted strictly to the training subsets to preserve the integrity of the unseen test data.

Selecting $k = 10$ provides an optimal trade-off: it suppresses sensitivity to transient motion spikes and sensor jitter while preserving the algorithm's capability to detect non-linear feature interactions, avoiding the over-smoothed global estimates associated with larger k values.

5. Model Training



Figure 6: Human activity recognition pipeline - model training

Randomly splitting time-series kinematic data across subjects introduces subject-level data leakage. To ensure strict generalisation to unseen individuals and prevent artificially inflated metrics, a nested Leave-One-Subject-Out (LOSO-CV) cross-validation architecture was implemented.

5.1. Nested Cross-Validation and LOSO

The nested cross-validation framework comprises two distinct evaluation loops: an inner loop dedicated to hyperparameter optimisation and an outer loop reserved strictly for unbiased performance testing.

The outer loop iterates eight times, withholding one subject for unbiased testing while passing the remaining seven to the inner loop. Within the inner loop, a secondary LOSO-CV (six training subjects, one validation subject) identifies the optimal hyperparameter configuration. This model is then retrained on all seven training subjects and evaluated against the unseen eighth test subject.

The final reported test accuracy represents the arithmetic mean across the eight outer folds. Lastly, standard cross-validation across all subjects determines the optimal configuration for training the final deployment model on the full dataset.

5.2. Training Pipeline

Each nested iteration followed a strict four-step execution workflow: subject-wise data splitting, feature normalisation using parameters (μ, σ) derived solely from the training fold, classifier training, and fold evaluation on unseen subjects. This execution sequence guaranteed complete isolation of unseen test distributions prior to final evaluation.

While standard Z-score normalisation is effective for well-bounded time-series windows, it remains sensitive to extreme statistical outliers in the training distribution. To mitigate this risk, transient motion artefacts were eliminated prior to feature calculation by trimming recording boundaries and applying band-pass filtering.

5.3. Classification Algorithms and Optimisation Space

To construct an optimal model, a grid search was conducted over six dimensions: band-pass cut-off frequencies, window size, window overlap, classification algorithms, feature selection methods, and the number of retained features.

The classifiers comprised Support Vector Machines (quadratic, cubic, and Gaussian kernels), k-Nearest Neighbours, and Random Forest. Across all algorithm variants, 162 distinct classifier configurations were evaluated (126 SVM: 54 quadratic, 36 cubic, 36 Gaussian; 24 kNN; 12 Random Forest). Combined with 18 signal preprocessing combinations (varying filters, window lengths, and overlaps), the total optimisation space encompassed 2,916 unique training permutations (162×18).

5.3.1. Signal Preprocessing Parameters

Preprocessing hyperparameters were evaluated across three dimensions: upper cut-off frequency (5 Hz narrowband vs 20 Hz wideband, with lower cut-off fixed at 0.5 Hz for gravity removal), window length (2.56, 5.12, and 10.24 s to balance stability against transient kinematic noise), and overlap fraction (0%, 25%, and 50% to mitigate step truncation across window boundaries).

5.3.2. Algorithm Parameters

Classifier hyperparameters were tuned across their core intrinsic dimensions: Support Vector Machines over box constraint C (polynomial: 1, 3, 10, 30; gaussian: 1, 10, 30, 100) and kernel scale γ (3, 10, 30, 100); k-Nearest Neighbours over neighbour count k (1, 5, 10, 30) and distance weighting (equal, inverse); and Random Forest over tree count (50, 100) and minimum leaf size (1, 5, 20).

We acknowledge the asymmetry between the SVM search space (126 configurations) and ensemble methods (12 configurations). This restriction was imposed by computational constraints in the nested

Table 1
Overall test accuracies per isolated subject (rounded to two decimal places).

Test Subject	Accuracy (%)
Subject 1	95.76
Subject 2	78.26
Subject 3	100.00
Subject 4	48.44
Subject 5	64.56
Subject 6	100.00
Subject 7	98.28
Subject 8	95.38
Average	**85.08**

cross-validation architecture, prioritising a fine-grained grid search for SVMs given the high sensitivity of soft-margin and kernel-scale parameters.

5.3.3. Feature Selection Parameters

To balance resolution and computational cost, feature selection was executed using the ReliefF algorithm ($k = 10$). The quantity of top-ranked features retained was optimized dynamically per classifier family: SVM Quadratic (20, 30, 40, 50, 100, 131), SVM Cubic (20, 30, 40, 50), SVM Gaussian (40, 50, 100, 131), kNN (30, 100, 131), and Random Forest (50, 131).

Evaluating a polynomial cubic kernel over a 131-dimensional feature space scales poorly in terms of memory and floating-point operations. Within our nested Leave-One-Subject-Out (LOSO) cross-validation architecture – where every configuration is iteratively trained and validated across inner and outer loops – running the full 131-feature set on cubic SVMs caused a significant computational bottleneck.

6. Results and Visualisation

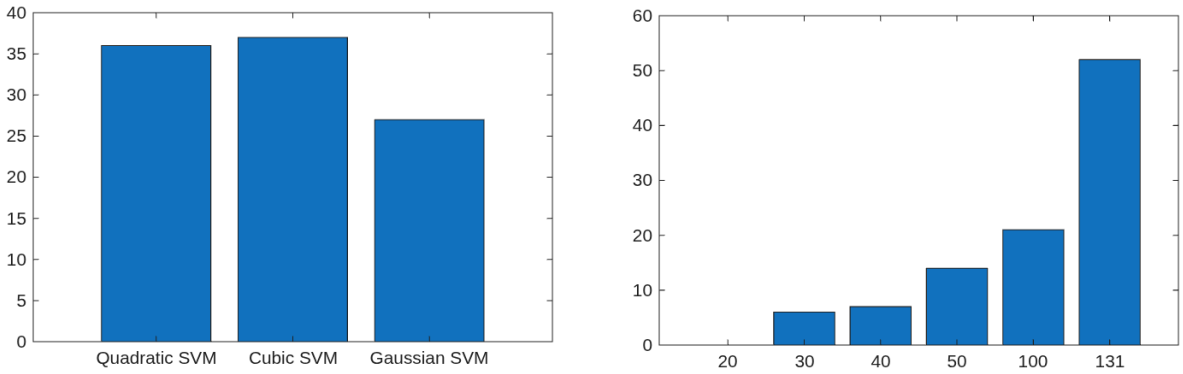
6.1. Overview of Classification Performance

The evaluation of the proposed classification pipeline yielded an overall test accuracy of **85.08%** across the LOSO cross-validation architecture. A detailed breakdown of individual subject accuracies is provided in Table 1.

An analysis of the 2,916 hyperparameter configurations evaluated during the nested cross-validation revealed a profound dominance of Support Vector Machine (SVM) architectures over ensemble methods. Remarkably, SVM models accounted for the top 371 performing configurations, with the highest-ranking Random Forest model appearing at position 372.

Within the SVM architectures, non-linear kernels demonstrated superior capability in mapping the complex kinematic boundaries of the mobility tasks. The highest-performing model overall was a cubic SVM, closely followed by quadratic and Gaussian variants. An analysis of the top 100 configurations further illustrates this trend: cubic SVMs appeared 37 times, quadratic SVMs 36 times, and Gaussian SVMs 27 times. Furthermore, the top 10 absolute configurations were entirely composed of these SVM models (four Gaussian, three cubic, and three quadratic).

An evaluation of the feature subset sizes across the top 100 configurations reveals a clear trend toward higher feature dimensionality, as illustrated in Figure 7b. Of these 100 top-performing pipelines, 20 features did not appear at all, while 30 features appeared 6 times, 40 appeared 7 times, and 50 appeared 14 times. In contrast, the largest feature subsets heavily dominated the distribution: a subset size of 100 features appeared 21 times, and the full feature set of 131 features appeared 52 times, accounting for over half of the top 100 positions.



(a) Distribution of algorithms

(b) Feature subset dimensionality

Figure 7: Results among top 100 performing models

6.1.1. Algorithm Performance Comparison

To systematically evaluate the efficacy of different machine learning paradigms, the peak classification accuracies of the five core algorithm variants were compared, along with their absolute rankings within the 2,916-permutation hyperparameter space.

The empirical results reveal a profound architectural divide. Support Vector Machine configurations exclusively claimed the top three positions in the entire grid search. The absolute champion was the Cubic SVM, achieving a peak accuracy of **98.12 %** (Rank 1), closely followed by the Quadratic SVM at **97.95 %** (Rank 2) and the Gaussian RBF SVM at **97.92 %** (Rank 3).

In stark contrast, the ensemble and distance-based classifiers experienced severe performance degradation when forced to generalise. The highest-performing Random Forest model achieved an accuracy of **93.07 %**, ranking 372nd. The K-Nearest Neighbours pipeline performed worst, achieving a peak accuracy of **92.41 %** and ranking 455th.

Several factors contribute to this disparity. One possibility is that the finer hyperparameter grid evaluated for SVMs offered more optimisation opportunities than the coarser search spaces of RF and kNN.

6.2. Optimal Pipeline Configuration

The highest-performing configuration across the entire hyperparameter search space was an SVM with a cubic kernel (degree 3), achieving a peak validation accuracy of **98.12 %**. The optimal hyperparameters for this architecture included a box constraint (C) of 10 and a kernel scale (γ) of 30, coupled with an optimised subset of exactly 40 features.

Crucially, this optimal machine learning configuration was explicitly tied to the most extensive temporal window evaluated: 10.24 seconds with a 0% overlap. Biomechanically, this extended window size is highly logical; rather than analysing fragmented kinematics, a 10.24-second frame captures multiple complete, continuous gait cycles, providing the classifier with a stable and highly periodic signal pattern.

Furthermore, the optimal signal preprocessing for this configuration utilised a band-pass filter with cut-off frequencies of 0.5 Hz and 5 Hz. This specific frequency band perfectly isolates the fundamental dynamics of human locomotion – which typically oscillates between 1.5 Hz and 2.5 Hz – while successfully eliminating 0 Hz gravitational artefacts (such as static sensor tilt) and high-frequency mechanical noise.

6.2.1. Feature Importance

To understand the underlying physical mechanisms driving the optimal cubic SVM's performance, the 40-feature subset selected by the ReliefF algorithm ($k = 10$) was evaluated based on individual feature

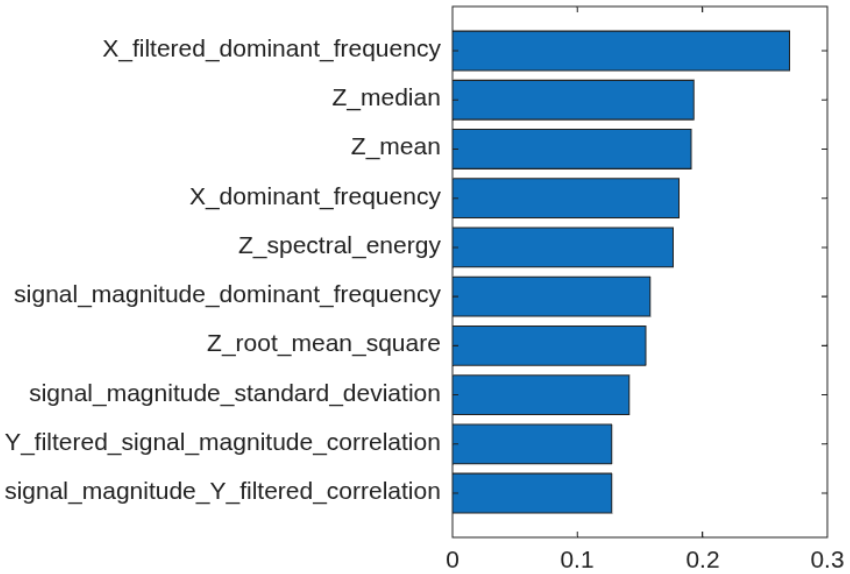


Figure 8: Relative importance scores of the top 10 features selected by the ReliefF algorithm, revealing a heavy reliance on frequency-domain dynamics and raw static metrics.

importance weights. Plotting all 40 features simultaneously introduces visual redundancy; therefore, the top 10 highest-ranked kinematic features are isolated and illustrated in Figure 8.

The feature distribution profile reveals a strategic blend of frequency-domain, time-domain, and raw statistical metrics. The single most dominant feature across the entire grid search space was the *X-axis Filtered Dominant Frequency*.

However, the structural composition of this top tier exposes a critical architectural vulnerability. The second and third-most important features selected by the algorithm are the raw, unfiltered *Z-axis Median* and *Z-axis Mean*. In a standard, upright subject, these features effectively capture the stable orientation of the pelvis relative to the gravitational vector during specific tasks.

6.3. Individual Subject Performance

While aggregate test accuracy provides a broad metric of a system's capability, understanding the pipeline's real-world generalisation requires analysing individual participants' results. Table 1 details the classification accuracy achieved when each of the eight participants was independently isolated as the unseen test set during the LOSO cross-validation.

The evaluation across all eight participants reveals three distinct performance tiers:

- **High-Accuracy Group (Subjects 1, 3, 6, 7, and 8):** The optimal cubic SVM pipeline demonstrated high classification accuracy across five of the eight participants, achieving test accuracies ranging from 95.38% to 100.00% (Subjects 3 and 6). This suggests strong feature consistency within this young adult cohort; however, these results represent an initial feasibility assessment, and broader generalisability across wider demographic profiles remains to be verified.
- **Moderate-Accuracy Group (Subjects 2 and 5):** Performance degraded moderately for Subject 2 (78.26%) and Subject 5 (64.56%). Localised error inspection revealed that misclassifications in this group occurred primarily between stair ascent and stair descent, driven by individual variations in stepping cadence and lower impact energy during vertical transitions.
- **Outlier Group (Subject 4):** A severe performance drop was observed exclusively in Subject 4 (48.44%), where level walking and stair descent were systematically misclassified.

The individual accuracies reveal a highly polarised performance distribution. The optimal cubic SVM pipeline demonstrated exceptional generalisation across most test subjects, achieving near-perfect accuracies (95 to 100 %) for five of the eight participants, including a flawless 100 % classification rate

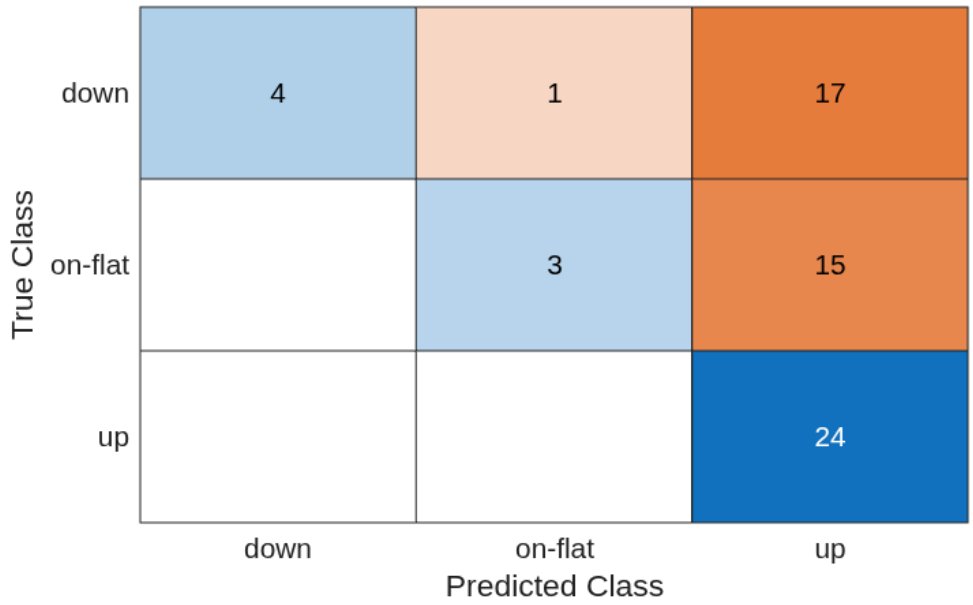


Figure 9: Confusion matrix detailing the classification bias for Subject 4

for Subjects 3 and 6. However, this high baseline was offset by a severe performance degradation in Subject 4 (48.44 %). This distinct inter-subject variability highlights the inherent challenges of universal gait classification and necessitates a deeper diagnostic analysis of the outlier.

This inter-subject variability illustrates that while the pipeline performs exceptionally well for most subjects, individual kinematic styles and shifts in hardware placement necessitate a deeper diagnostic analysis of localised failures.

6.4. The Subject 4 Anomaly

To identify the underlying failure mechanism responsible for the 48% classification accuracy observed in Subject 4, a localised confusion matrix was constructed and evaluated (Figure 9). Rather than exhibiting stochastic or random misclassifications across the three mobility domains, the optimal cubic SVM model demonstrated a profound, systematic bias toward a single class: *Stair Ascent* (Up).

Out of 64 total evaluation windows, the pipeline predicted the *Stair Ascent* class a disproportionate 56 times (87.5 % of all instances). While this systematic bias resulted in a flawless 100% classification rate (24/24) for true stair ascent, it severely compromised the accuracy for flat-ground walking (16.6 %) and stair descent (18.1 %).

Biomechanically, stair ascent is uniquely characterised by two distinct inertial features captured by a lumbar-mounted IMU: a pronounced forward trunk flexion (bending forward) and high vertical acceleration peaks during weight acceptance. The fact that level walking and stair descent were overwhelmingly misclassified as stair ascent strongly implies that Subject 4's raw kinematic data possessed a permanent forward tilt. A physical sensor misalignment could have caused this – such as the smartphone shifting or tilting forward within the belt during the data collection session – or by a unique personal gait adaptation wherein the subject naturally walked with an exaggerated forward posture.

This physical anomaly interacted catastrophically with the pipeline's feature selection phase. Because the inner loop utilised the ReliefF algorithm (k = 10) exclusively on the training subjects, the algorithm heavily weighted raw gravitational features, which served as excellent class separators for the normal, upright subjects. Consequently, when presented with the statistically tilted data from Subject 4, the cubic SVM prioritised these heavily weighted raw features over the band-pass-filtered signals. The model interpreted the constant gravitational offset as a continuous forward trunk lean, leading it to classify almost all movements as stair ascent.

This diagnostic underscores a critical vulnerability in Leave-One-Subject-Out architectures. When raw features are exposed to the classifier, a localised hardware shift or individual postural anomaly can completely mask the dynamic, filtered frequency responses of the underlying physical activity.

7. Conclusion and Future Research

7.1. Summary of the Project

The objective of this work was to differentiate fundamental movements accurately. The core contribution of this work involved the design, execution, and evaluation of a robust machine learning pipeline. An empirical dataset was constructed.

The raw kinematic data underwent systematic preprocessing, band-pass filtering, and temporal segmentation. To capture a highly descriptive biomechanical profile, a comprehensive feature space was generated spanning the time, frequency, and wavelet domains. To mitigate the curse of dimensionality, the dataset was normalised, and the most critical predictive variables were isolated using feature selection algorithms.

Finally, a rigorous nested cross-validation framework was deployed to train and evaluate multiple classification architectures. By systematically optimising the hyperparameters across a 2,916-permutation grid search, the study conclusively demonstrated the superiority of non-linear classification models. The pipeline ultimately revealed that the optimal configuration – a Support Vector Machine utilising a cubic kernel, a 10.24-second temporal window, and an optimised 40-feature subset – achieved a peak validation accuracy of 98.12 %. Crucially, when evaluated on completely unseen hold-out data, this architecture yielded a robust overall test accuracy of 85.08 %.

7.2. System Limitations and Feature Redundancy

Following completion of model training and validation, a retrospective analysis of the feature engineering pipeline identified an artefact in the time-domain extraction phase. Specifically, a subset of correlation coefficients was inadvertently duplicated within the final 131-dimensional feature space prior to algorithm selection.

Because the ReliefF algorithm was utilised as the primary feature selection mechanism, these redundant features were not filtered out. By design, ReliefF evaluates the predictive weight of features individually based on nearest-neighbour margin maximisation; it does not include a penalisation mechanism for inter-feature redundancy, unlike algorithms such as Minimum Redundancy Maximum Relevance (MRMR). Consequently, highly correlated redundant features received identical, elevated importance scores and were subsequently propagated into the final SVM training sets.

While this redundancy introduces minor computational inefficiencies, it does not invalidate the predictive integrity of the underlying SVM architectures. SVMs, particularly those employing non-linear kernels with robust structural risk minimisation (regularisation), are inherently resilient to collinearity. The redundant vectors serve only as scalar multipliers in the high-dimensional feature space. The nested K-fold cross-validation results empirically confirm the validity of the models; the high test-fold accuracies demonstrate that the SVM successfully established generalisable decision boundaries without catastrophic overfitting, despite the presence of redundant correlation metrics. Future iterations of this pipeline should remove redundant correlation features or implement MRMR prior to classification to enforce strict feature orthogonality.

7.3. Future Work

While this study successfully establishes an accurate baseline for kinematic classification, several avenues remain to expand the pipeline's robustness and clinical applicability.

Firstly, the experimental group should be significantly expanded. While the existing dataset of eight healthy subjects provided sufficient variance to validate the machine learning architecture, a larger,

more demographically diverse sample size is required to ensure universal generalisation.

Secondly, the biomechanical taxonomy of the system should be broadened. The current pipeline works for classifying fundamental dynamic locomotion; however, future iterations should incorporate a wider spectrum of human activity. This includes high-impact kinematics, such as running or jumping, as well as static postural transitions, such as standing, prolonged sitting, or transitioning from a seated to a standing position.

Thirdly, future research should transition this offline methodology into a real-time, continuous monitoring application. By leveraging the computational efficiency of the Fast Fourier Transform, the classification pipeline can, in theory, be deployed directly on low-power wearable microprocessors or mobile architectures, enabling real-time neurological motion analysis outside controlled laboratory settings.

Fourthly, evaluating end-to-end deep learning time-series architectures represents a compelling avenue for future investigation. While our current framework relies on manual multi-domain feature engineering, sequence-based models such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and 1D Temporal Convolutional Networks (1D-TCNs) can directly capture complex temporal dependencies from raw accelerometric streams without explicit domain transformations. Although deep architectures are prone to overfitting on small cohorts due to their high parameter density, testing lightweight or pre-trained deep models on larger participant datasets will provide key insights into whether automated feature representation learning can surpass highly optimised non-linear SVMs.

Acknowledgments

I would like to express my sincere thanks to Mgr. Irena Šindelářová, Ph.D., for valuable guidance, insightful feedback, and support throughout the development of this paper.

Declaration on Generative AI

During the preparation of this work, the author(s) used Grammarly in order to: Grammar and spelling check. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] The MathWorks, Inc, Matlab Documentation, 2026. URL: https://www.mathworks.com/help/releases/R2026a/, accessed: May 15, 2026.

[2] O. D. Lara, M. A. Labrador, A survey on human activity recognition using wearable sensors, IEEE Communications Surveys & Tutorials 15 (2013) 1192–1209. doi:10.1109/surv.2012.110112.00192.

[3] G. Ogbuabor, R. La, Human activity recognition for healthcare using smartphones, in: Proceedings of the 2018 10th International Conference on Machine Learning and Computing, ICMLC '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 41–46. URL: https://doi.org/10.1145/3195106.3195157. doi:10.1145/3195106.3195157.

[4] J. R. Kwapisz, G. M. Weiss, S. A. Moore, Activity recognition using cell phone accelerometers, ACM SIGKDD Explorations Newsletter 12 (2011) 74–82. doi:10.1145/1964897.1964918.

[5] S. Jamian, T. S. Gunawan, M. Kartiwi, R. Ahmad, K. Abdul Kadir, M. N. Nordin, Human activity and posture classification using smartphone sensors and matlab mobile, in: 2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), IEEE, 2022, pp. 1–6. doi:10.1109/i2mtc48687.2022.9806551.