BJMB
Brazilian Journal of Motor Behavior

Research Article

Predicting Walkway Spatiotemporal Parameters Using a Markerless, Pixel-Based Machine Learning Approach

ARIANY K. TAHARA[1] | ABEL G. CHINAGLIA[1] | RAFAEL L. M. MONTEIRO[1] | BRUNO L. S. BEDO[2] | GUILHERME M. CESAR[3] | PAULO R. P. SANTIAGO[1,4]

[1] Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil
[2] School of Physical Education and Sport, University of São Paulo, São Paulo, SP, Brazil
[3] Department of Physical Therapy, Brooks College of Health, University of North Florida, FL, USA
[4] School of Physical Education and Sports of Ribeirão Preto, University of São Paulo, Ribeirão Preto, SP, Brazil

Correspondence to: Paulo Roberto Pereira Santiago
email: paulosantiago@usp.br
https://doi.org/10.20338/bjmb.v19i1.462

HIGHLIGHTS

- Developed a machine learning-based approach for gait analysis without commercial instrumented walkways.
- Predicted spatiotemporal gait parameters with markerless pose estimation and pixel-based data.
- Validated the proposed method against GAITRite™ measurements, demonstrating high accuracy in gait parameter estimation.
- Identified Gradient Boosting Regressor and Random Forest Regressor as the most effective models for spatial and temporal gait characteristics, respectively.
- Integrated the trained models into the open-source *vailá* Multimodal Toolbox, providing an accessible and cost-effective gait analysis solution, even for users without expertise in IT or programming.
- Provides a low-cost, practical alternative for clinical and research applications, reducing dependence on expensive motion capture systems.

ABBREVIATIONS

DCNN	Deep convolutional neural networks
EMG	Electromyography
EVS	Explained Variance Score
GBR	Gradient Boosting
GUI	Graphical user interface
ICC	Intraclass correlation coefficient
MAE	Mean Absolute Error
MaxE	Maximum Error
MCID	Minimal clinically important difference
MedAE	Median Absolute Error
RAE	Relative Absolute Error
RFR	Random Forest
RMSE	Root Mean Squared Error

PUBLICATION DATA

Received 31 01 2025
Accepted 04 06 2025
Published 10 07 2025

BACKGROUND: Traditional gait analysis relies on motion capture systems and sensor mats, which are precise but costly and constrained to specialized environments. Markerless pose estimation combined with machine learning offers a low-cost alternative for gait analysis.
AIM: This study aimed to develop and validate a markerless method for predicting spatiotemporal gait parameters using video-based pose detection and machine learning, leveraging the GAITRite™ system as a reference.
METHODS: Ten healthy adults walked barefoot on a GAITRite™ mat while their movements were recorded. Videos were processed using the *vailá* Multimodal Toolbox, which integrates MediaPipe for pose estimation and various machine learning algorithms. Gait parameters such as step length, stride velocity, and support time were predicted using regression models trained on pixel-based data.
RESULTS: The results demonstrate a clear division in model performance based on the target variable. For metrics related to spatial characteristics, such as Step Length, Step Width, Stride Length, Stride Width, and Stride Velocity, the Gradient Boosting (GBR) delivered the best results, exhibiting lower error metrics and higher R^2 scores. Only for the Support Base data the best model differed, with the Random Forest (RFR) outperforming the others. This model also displayed lower error values and higher R^2 and EVS score.
INTERPRETATION: The results indicate the feasibility of markerless pixel-based gait analysis as a low-cost alternative to traditional methods, broadening the accessibility of precise gait assessment for both clinical and research applications.

KEYWORDS: Markerless motion analysis | Gait parameters | Machine learning | Pixel-based data | Computer vision | Walkways

INTRODUCTION

Gait analysis is a fundamental tool in both clinical and research settings, providing important information about human movement control. Traditional methods often utilize instrumented walkways, such as the GAITRite™ and Zeno Walkway systems, which capture spatial and temporal gait parameters through pressure sensitive sensors embedded in mats. The GAITRite™ system has

demonstrated excellent inter-rater reliability (ICC = 0.97) and has been validated in studies comparing its gait measurements to reference tools such as 3D motion capture systems. Similarly, the Zeno Walkway system has shown strong accuracy in measuring step length and cadence, with high correlation (r = 0.98) to traditional motion capture methods [1–4]. However, despite their precision, these traditional systems present limitations, including high costs, specialized equipment necessity, and testing environment constraints. Recent advancements in computer vision and machine learning have introduced markerless motion capture systems as viable alternatives. These systems employ algorithms to estimate 2D and 3D joint positions from standard video footage without physical markers or specialized laboratories. Panconi et al. (2024) [5] demonstrated in a recent study the feasibility of using DeepLabCut — a markerless pose estimation framework—for clinical gait analysis. By comparing its output with that of the GAITRite™ system, the authors reported strong agreement for step length and stride time, with mean absolute errors below 3.5 cm and 0.03 s, respectively, and intraclass correlation coefficients (ICCs) above 0.90 for most parameters [5]. These results support the potential of markerless systems to achieve clinically relevant accuracy.

Markerless pose estimation based on deep learning promises to become an easy-to-access method to analyze human movements without using a high-cost motion capture system. This development in computer vision enables reliable analysis of human 2D joint position in images or videos with varying environments, including lighting and angle of view [6–8]. The algorithm outputs joint position (e.g., knee, ankle, foot, and hip) in the 2D image plane in each video frame. Furthermore, 3D motion reconstruction from the 2D pose in the image or video has been proposed without any markers or manual labeling [9–11].

Recent advances in markerless gait analysis have enhanced human motion tracking through a variety of pose estimation frameworks. MediaPipe has emerged as a robust option due to its low computational demands, ease deployment, and ongoing development, presenting an effective balance between performance and accessibility [8]. In contrast, OpenPose, despite its pioneering role in real-time multi-person tracking, requires substantial processing power and has received limited recent updates [6].

DeepLabCut offers high accuracy via custom labeling but relies heavily on manual annotation and GPU resources, which may restrict its scalability for real-time or large-scale applications [12,13]. Given these considerations, MediaPipe presents a lightweight, open-source alternative well suited for low-cost, real-time gait analysis on standard hardware [14].

Current approaches have successfully employed machine learning techniques to enhance gait analysis accuracy. Guffanti et al. (2024) [15] utilized supervised learning to improve the accuracy of gait parameter estimations from a robot-mounted 3D camera. By training artificial neural networks with data from a certified Vicon system, they achieved substantial improvements in detecting kinematic gait signals and descriptors across 207 gait sequences from 37 healthy participants.

Similarly, Nazari et al. (2022) [16] compared traditional machine learning models—such as Gaussian Naive Bayes, Decision Trees, and Random Forests—with deep convolutional neural networks (DCNNs) for gait phase detection using electromyography (EMG) data. Their findings indicated that DCNNs outperformed traditional models, achieving up to 89.5% accuracy in certain trials. These studies underscore the potential of integrating machine learning approaches to enhance the precision and applicability of gait analysis systems [15,16].

Integrating machine learning techniques with markerless pose estimation enhances the analysis of gait patterns. Machine learning models can process pixel-based data to predict various gait metrics, offering a low-cost and accessible solution for motion analysis. This approach facilitates gait assessment in diverse environments without the constraints associated with traditional motion capture systems. Supervised machine learning methods can be used for predicting gait analysis. The goal of this kind of model is to estimate a function f(x) that transforms the input x into a desired output y as accurately as possible. To determine this function, the supervised learning algorithm is provided with labeled examples (the training set), i.e., for a known input x, the corresponding expected output y. After the training phase, the supervised learning model tries to predict the output for new input data [17].

This study aims to develop and validate a method that uses machine learning algorithms to predict gait measures traditionally obtained from systems such as GAITRite™, using only pixel-based data derived from a computer vision pose detection model. We hypothesize that this method will provide accurate gait assessments, thereby broadening the accessibility and applicability of gait analysis in clinical and research contexts.

METHODS

Participants, procedures and instruments

Seventeen healthy adults, able to walk free of auxiliary devices (8 women and 9 men, aged 18–40 years), participated in the present study. Participants were instructed to execute ten trials walking at a self-selected pace, barefoot, on a GAITRite™ mat (CIR Systems Inc., Clifton, NJ, USA). Each trial began with the participant standing on the edge of the GAITRite™ mat, initiating the walk with their preferred leg. During the walking trials, movements were recorded using a GoPro Hero 10 Black Edition camera (GoPro, Inc., USA), set to a resolution of 1920×1080 pixels and 60 Hz frame rate. The camera was positioned 1 meter from the edge of the GAITRite™ mat, providing a frontal view of the walking.

The GAITRite™ system consists of a mat equipped with six sensor pads, covering an active area of 3.66 × 0.61 meters, capable of detecting foot contact. These sensors are activated by mechanical pressure, allowing the software to compute the coordinates (x: mediolateral and y: anteroposterior) of foot contacts in real-time. Temporal events are recorded directly by the sensors without

reliance on derived formulas. The extracted variables included step length, step time, and step width, which characterize individual steps, stride length, stride time, and stride width, representing complete gait cycles. Additional parameters, such as stride velocity and support base width, provided insights into walking stability and speed. The system also documented single and double support times, crucial for understanding dynamic balance during walking.

At the same time, we employed a markerless video-based system to capture additional gait data. Although this system requires fixed cameras and controlled environments, such constraints are often suitable for laboratory and clinical settings where standardization is crucial. Compared to other methods, such as inertial sensors or instrumented mats, markerless systems offer distinct advantages, including lower cost, non-invasiveness, and the ability to reprocess previously recorded video data [18,19]. Additionally, recent advancements in pose estimation algorithms have significantly enhanced the accuracy and robustness of markerless systems, making them suitable for use in structured environments outside of highly controlled settings [14]. Consequently, we chose to use a fixed-camera setup in this study as a methodological approach to provide a non-invasive, cost-effective solution for gait analysis, aligned with the goal of enhancing accessibility through open-source tools.

Data processing

To evaluate kinematic variables for each trial, videos were processed using the *vailá* Multimodal Toolbox [20], a multimodal platform developed in Python 3.12.8. The graphical user interface (GUI) of *vailá* is shown in Figure 1. This toolbox integrates MediaPipe [8], an advanced 2D pose estimation framework that tracks and overlays anatomical landmarks on video frames. MediaPipe enables the detection of joints and anatomical points, providing pixel coordinates of the identified landmarks through its skeleton detection algorithm.

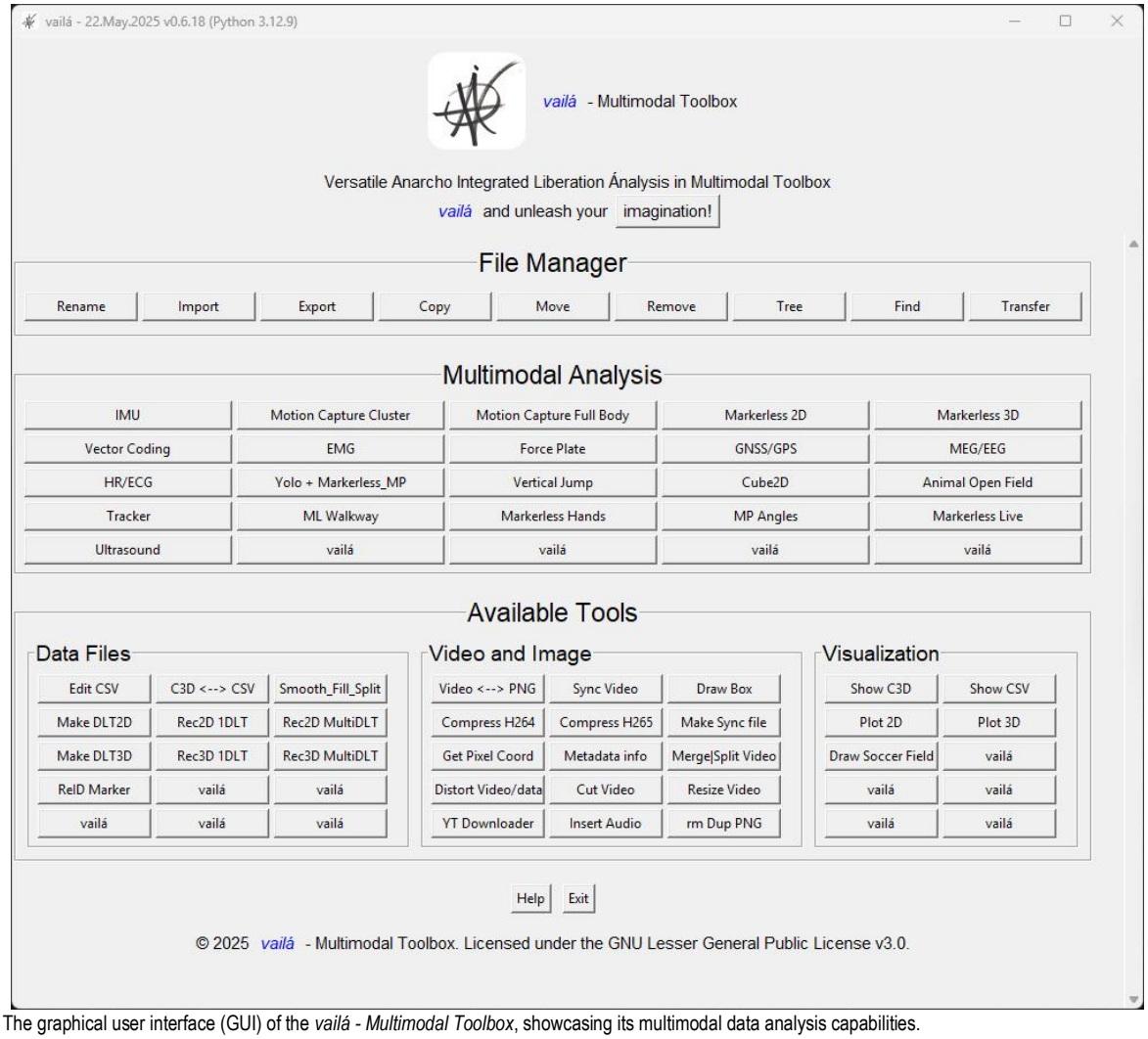

Figure 1. The graphical user interface (GUI) of the *vailá - Multimodal Toolbox*, showcasing its multimodal data analysis capabilities.

Each video was preprocessed to enhance the visibility and detectability of the participant's movements by the pose estimation model. First, a region of interest was defined around the participant, and all pixels outside this window were set to black, minimizing

visual noise from background elements and improving joint localization. Next, the video was duplicated and horizontally mirrored to simulate the participant walking from the opposite side of the frame. This mirroring ensured that, in at least one version of the video, the participant appeared closer to the camera during the walking phase, which improved the resolution of visual features and facilitated more accurate landmark detection.

From these two versions, only the second half of the mirrored video was retained for analysis, as it typically contained the most stable and natural gait cycles. These steps collectively enhanced the reliability of the joint detection algorithm by reducing occlusion, increasing contrast between the participant and background, and maximizing the visual prominence of lower-limb joints during the stance and swing phases of gait. An example of the detection process is illustrated in Figure 2.

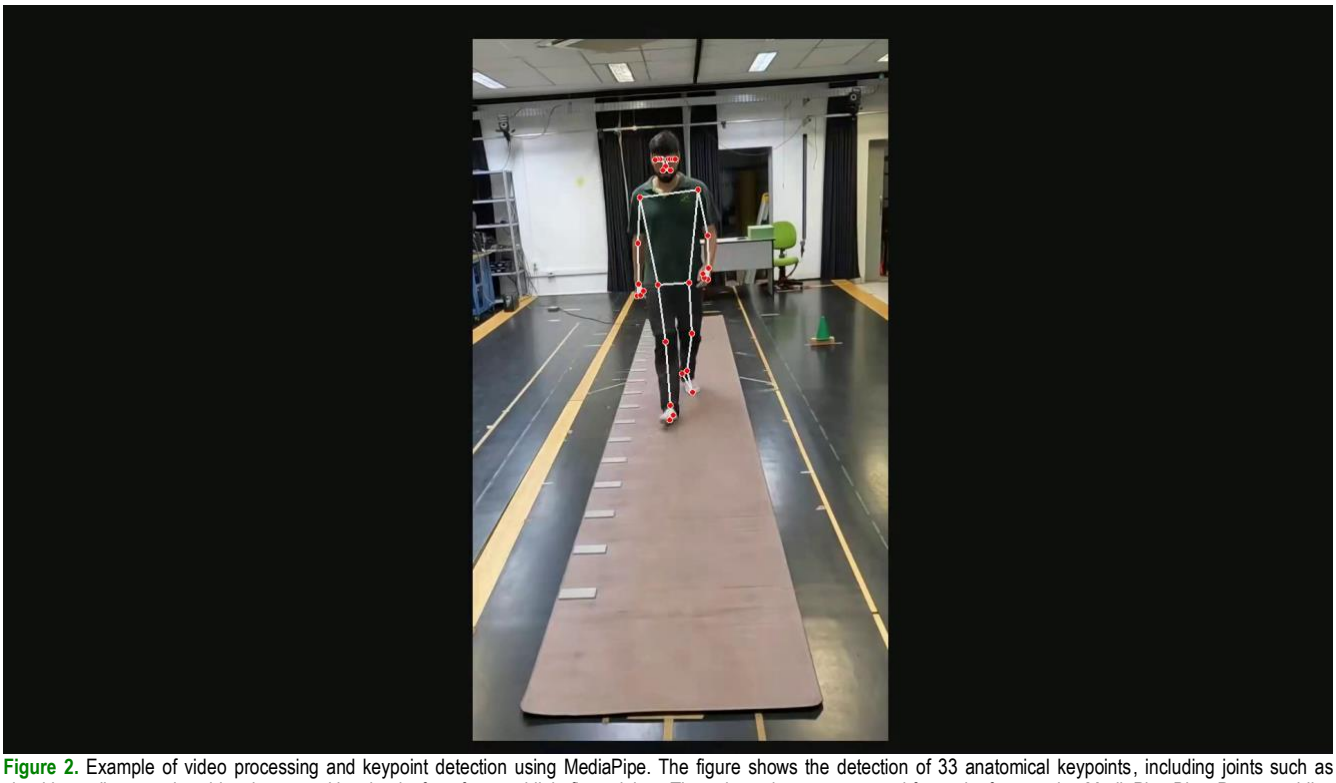

Figure 2. Example of video processing and keypoint detection using MediaPipe. The figure shows the detection of 33 anatomical keypoints, including joints such as shoulders, elbows, wrists, hips, knees, ankles, heels, feet, face and little finger joints. These keypoints are extracted frame-by-frame using MediaPipe BlazePose, enabling the generation of spatiotemporal data for gait analysis. The video was recorded in a controlled environment, with the participant walking on an instrumented walkway (GAITRite™), which served as a reference for validating the machine learning model predictions. For more details on MediaPipe BlazePose and its capabilities, refer to the official documentation: MediaPipe

These spatial and temporal parameters served as reference metrics for evaluating the accuracy of gait data derived from video analysis. By integrating video processing with the GAITRite™ system, a comprehensive dataset was created to support machine learning-based prediction models. This dataset was developed using the *vailá* ML Walkway feature, accessible via the ML Walkway button in the main interface (see Figure 3). The sub-interface, as shown, enabled the training and validation of ML models, as well as the processing of gait features and running of predictions, streamlining the analysis workflow.

As a standard preprocessing step, all stride-related values equal to zero provided by the GAITRite™ system were excluded prior to model training. This preliminary filtering ensured that only valid and complete measurements were used, minimizing the impact of sensor artifacts and enhancing the reliability of the reference data. These zero values exclusion resulted in the removal of the first and last gait cycles. These cycles are also considered the moments of acceleration and deceleration of the participants. Thus, only the middle steps of the participants were used.

The screen coordinates in pixels of the MediaPipe detections corresponding to the right and left feet were selected for analysis. The pixel data from each trial was divided equally by the number of steps taken in each trial. Using a Python-based framework, spatial and temporal features were extracted from the gait analysis data for each step. The extracted features included mean, variance, range, speed, and step length, computed for each participant and trial data.

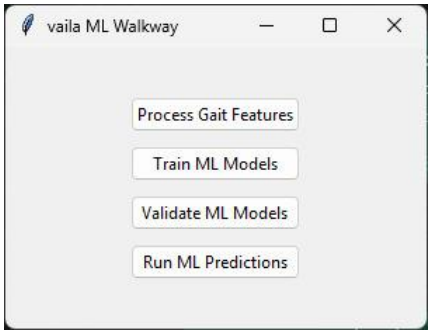

Figure 3. The *vailá* ML Walkway sub-interface opens upon selecting the ML Walkway button. It provides options for Process Gait Features, Train ML Models, Validate ML Models, and Run ML Predictions, facilitating machine learning-based gait analysis.

In addition to the processing options we employed, *vailá* also supports various other video processing and landmark detection methods. The additional video adjustment features include selecting video metadata, modifying video codecs, synchronizing multiple videos, and correcting lens distortion. Figure 4 shows a flow chart with the possibilities for adjusting videos and detecting the landmarks in *vailá*.

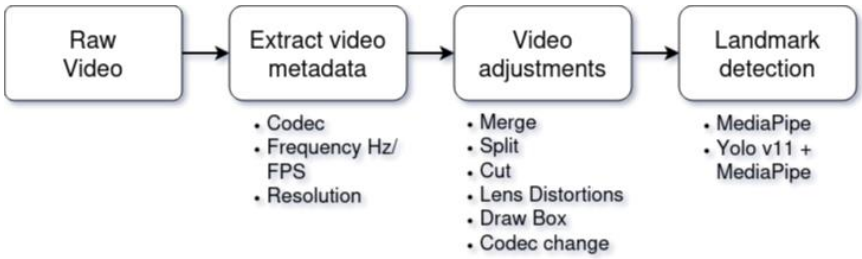

Figure 4. Flowchart illustrating the video processing and landmark detection workflow using *vailá*: (1) Raw Video: Input of unprocessed video data recorded during the gait analysis; (2) Extract video metadata: Extraction of technical details such as codec type, frame rate (Hz/FPS), and resolution for standardization and quality control; (3) Video adjustments: Operations to prepare the videos, including merging, splitting, cutting, correcting lens distortions, drawing bounding boxes, and changing codecs to ensure compatibility with processing tools; (4) Landmark detection: Application of pose estimation algorithms, such as MediaPipe or Yolo v11 combined with MediaPipe, to identify and track anatomical landmarks for further analysis.

Machine Learning algorithms

This dataset, containing gait metrics derived from the pixel coordinates of each participant, was used as input features for training machine learning algorithms. The target variables for these algorithms were (01) Step Length, (02) Step Time, (03) Step Width, (04) Stride Length, (05) Stride Time, (06) Stride Width, (07) Stride Velocity, (08) Support Base Size, (09) Support Time Single, (10) Support Time Double. These metrics are some of the gait metrics exported by the GAITRite™ system. Seven regression Machine Learning algorithms were trained separately for each target variable, resulting in distinct training processes for each of the ten metrics exported by the GAITRite™ system. The algorithms included XGBoost, Gradient Boosting Regressor, K-Nearest Neighbors, Linear Regression, Multilayer Perceptron, Random Forest Regressor, and Support Vector Regression. Each algorithm offers a unique approach to regression tasks [23,24] as shown in Table 1.

Table 1. The definitions of Machine Learning Models.

Model	Definition
XGBoost	An optimized implementation of gradient boosting that iteratively combines weak predictive models, such as decision trees, to improve prediction accuracy.
Gradient Boosting Regressor (GBR)	Applies a similar boosting methodology, focusing on minimizing prediction errors by correcting residuals over multiple iterations.
K-Nearest Neighbors (KNN)	A non-parametric method that predicts a target value by averaging the values of the closest k neighbors in the feature space.
Linear Regression (LR)	Models the relationship between input features and target variables by fitting a linear equation to the data, assuming linearity and independence between predictors.
Multilayer Perceptron (MLP)	A type of neural network that learns complex patterns in data through layers of interconnected neurons, optimizing weights using backpropagation.
Random Forest Regressor (RFR)	Builds multiple decision trees during training and averages their predictions to enhance accuracy and reduce overfitting.
Support Vector Regression (SVR)	Constructs a hyperplane in high-dimensional space to predict values while maintaining a margin of tolerance (epsilon) around the actual targets.

All models were implemented using Python libraries such as *scikit-learn* and *xgboost* for both algorithm development and

performance evaluation. Two dataset-splitting strategies were employed for model training to assess their impact on model performance: the traditional train-test split and k-fold cross-validation. In the first approach, the dataset was divided into 75% for training and 25% for testing, using a fixed random seed to ensure reproducibility. The k-fold cross-validation technique with 10 splits was applied in the second approach, involving data shuffling and a fixed random seed to guarantee reproducibility. Both strategies were compared to determine which approach provided better generalization and accuracy.

To ensure robust performance evaluation, data from two participants were used exclusively as an unseen test set in both approaches after the training performance analysis. The training set consisted of 854 data points, each of which represents a step. The test set included 213 data points; this dataset allowed an unbiased evaluation of the model's predictive capabilities. The performance of models was assessed using the evaluation metrics outlined in Table 2.

Table 2. Evaluation Metrics.

Metric	Definition
Root Mean Squared Error (RMSE)	Calculates the square root of the average squared differences between predicted and actual values, providing error estimates in the same unit as the original data. Lower RMSE values indicate better model performance.
Mean Absolute Error (MAE)	Computes the average of absolute differences between predicted and actual values, offering an intuitive measure of prediction accuracy. Lower MAE values indicate smaller average errors.
Median Absolute Error (MedAE)	Measures the median of absolute prediction errors, offering robustness to outliers.
Relative Absolute Error (RAE)	Compares the absolute prediction error of the model to a naïve mean-based model. Lower RAE values reflect better model performance.
Maximum Error (MaxE)	Identifies the maximum absolute prediction error, helping to detect extreme prediction deviations.
R2 Score	Quantifies the proportion of variance explained by the model. Scores closer to 1 represent better model fits.
Explained Variance Score (EVS)	Evaluates the proportion of variance in the data captured by the model, with higher scores indicating better fits.

The metrics were considered to select the model with the best performance. As the first criteria, the R^2 score, followed by the EVS and RAE. Prediction errors as in MAE and MaxE composed third and fourth criteria of analysis of the values, anterior of MedAE. The last metric analyzed was the RMSE. The machine learning algorithm that achieved the best performance across the evaluation metrics was selected for each target variable. Figure 5 shows a flow chart of the machine learning process.

The chosen models were then integrated into the *vailá* toolbox, enabling their application for predicting gait metrics in new datasets. This integration aims to facilitate broader accessibility to accurate and efficient gait analysis, leveraging the predictive capabilities of machine learning within a user-friendly framework. Additionally, for those interested in training their own models or enhancing the existing ones with new data, the corresponding codes and resources are made available within the *vailá* toolbox for further use and development.

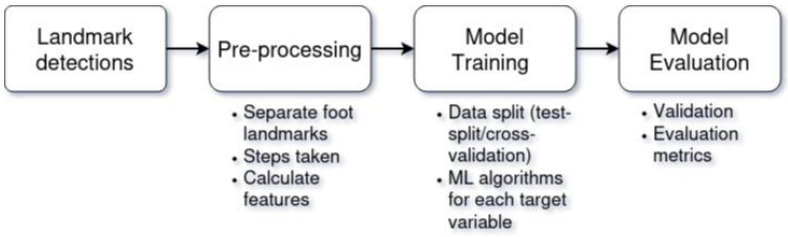

Figure 5. Flowchart illustrating the workflow for machine learning-based gait parameter prediction using the ML Walkway button from vailá: (1) Landmark detections: Identification of anatomical landmarks from video data using pose estimation tools such as MediaPipe; (2) Pre-processing: Extraction of relevant features, including separation of foot landmarks, detection of number of steps valid, calculation of step based metrics, and generation of input data; (3) Model training: Division of data into training and testing sets (e.g., cross-validation), and application of machine learning algorithms tailored for each target variable (e.g., step length, stride velocity); (4) Model evaluation: Validation of model performance using appropriate metrics (e.g., R^2, MAE, RMSE), ensuring accurate predictions for spatiotemporal gait parameters.

RESULTS

Only the performance metrics of the best models, selected based on their highest performance during validation, are presented among the evaluated results. In all validation cases, the best results were obtained using cross-validation. The unseen test results for each of the variables can be found in Tables 3 to 12, with the highlighted rows in the tables corresponding to the models selected as the best performing.

Regarding step length - Table 3, XGBoost showed the highest R^2 (0.7089) and EVS (0.7124), explaining a significant portion of data variance. It also achieved the lowest RAE (0.4925), MAE (0.028), MedAE (0.021), and RMSE (0.038), indicating high accuracy and robustness. For step time - Table 4, XGBoost again led with the highest R^2 (0.631) and EVS (0.634), and the lowest RAE (0.562), MAE (0.018), MedAE (0.013), and RMSE (0.025), reflecting precise predictions and low error. In step width - Table 5, XGBoost stood out with the highest R^2 (0.716) and EVS (0.720), and the lowest RAE (0.481), MAE (0.027), MedAE (0.021), and RMSE (0.037), showing reduced prediction errors.

Table 3. Step Length Results: Performance Evaluation Across Machine Learning Models.

Model	R^2	EVS	RAE	MAE (m)	MaxE (m)	MedAE (m)	RMSE (m)
XGBoost*	0.709	0.712	0.493	0.028	0.119	0.021	0.038
KNN	0.622	0.625	0.553	0.032	0.141	0.024	0.043
MLP	0.378	0.39	0.73	0.042	0.182	0.033	0.055
SVR	0.316	0.336	0.741	0.043	0.185	0.032	0.058
RFR	0.683	0.689	0.513	0.03	0.141	0.023	0.04
GBR	0.646	0.652	0.562	0.032	0.117	0.026	0.042
LR	0.401	0.41	0.737	0.042	0.142	0.033	0.052

*XGBoost had the superior performance in some metrics makes it the most consistent choice for step length, even the GBR having a slightly lower MaxE.

Table 4. Step Time Results: Performance Evaluation Across Machine Learning Models.

Model	R^2	EVS	RAE	MAE (s)	MaxE (s)	MedAE (s)	RMSE (s)
XGBoost*	0.631	0.634	0.562	0.018	0.088	0.013	0.025
KNN	0.604	0.61	0.581	0.019	0.091	0.014	0.026
MLP	-2.902	-2.81	1.925	0.062	0.273	0.052	0.08
SVR	-0.205	0.146	1.201	0.039	0.1	0.038	0.045
RFR	0.602	0.603	0.578	0.019	0.096	0.014	0.026
GBR	0.611	0.612	0.589	0.019	0.094	0.014	0.025
LR	0.521	0.525	0.675	0.022	0.091	0.018	0.028

*XGBoost demonstrated superior overall performance, making it the best option for step time, although LR presented a slightly lower MaxE.

Table 5. Step Width Results: Performance Evaluation Across Machine Learning Models.

Model	R^2	EVS	RAE	MAE (m)	MaxE (m)	MedAE (m)	RMSE (m)
XGBoost*	0.716	0.72	0.481	0.028	0.121	0.021	0.037
KNN	0.643	0.646	0.534	0.031	0.14	0.023	0.042
MLP	0.383	0.394	0.729	0.042	0.189	0.032	0.055
SVR	0.333	0.354	0.73	0.042	0.184	0.031	0.057
RFR	0.693	0.699	0.504	0.029	0.138	0.022	0.039
GBR	0.666	0.672	0.539	0.031	0.12	0.024	0.04
LR	0.416	0.424	0.732	0.042	0.141	0.034	0.054

*XGBoost outperformed in most metrics, justifying its selection for step width, although GBR had a slightly lower MaxE.

For stride length - Table 6, GBR achieved the highest R^2 (0.864) and EVS (0.867), with the lowest RMSE (0.200), indicating a strong fit and low errors. While RFR had lower MAE (0.112 vs. 0.118), MedAE (0.049 vs. 0.064), and MaxE (0.888 vs. 0.892), and XGBoost a lower RAE (0.253 vs. 0.281), GBR's balance of explained variance and RMSE made it preferable. For stride time - Table 7, GBR led with highest R^2 (0.848) and EVS (0.851), and lowest RMSE (0.161), reflecting excellent fit and low error dispersion. Although RFR showed lower MAE (0.084 vs. 0.095), MedAE (0.028 vs. 0.050), and MaxE (0.717 vs. 0.746), and XGBoost lower RAE (0.242 vs. 0.294), GBR's overall balance was superior.

Table 6. Stride Length Results: Performance Evaluation Across Machine Learning Models.

Model	R^2	EVS	RAE	MAE (m)	MaxE (m)	MedAE (m)	RMSE (m)
XGBoost	0.828	0.832	0.254	0.109	1.013	0.042	0.22
KNN	0.725	0.728	0.372	0.157	1.132	0.058	0.281
MLP	0.639	0.645	0.567	0.238	0.991	0.167	0.325
SVR	0.039	0.177	0.726	0.315	1.322	0.104	0.541
RFR	0.859	0.862	0.262	0.112	0.888	0.049	0.204
GBR*	0.864	0.867	0.281	0.118	0.892	0.065	0.198
LR	0.408	0.426	0.838	0.35	0.939	0.344	0.416

*GBR was preferred for its balance between high explained variance and low RMSE, which are essential for stride length.

Table 7. Stride Time Results: Performance Evaluation Across Machine Learning Models.

Model	R^2	EVS	RAE	MAE (s)	MaxE (s)	MedAE (s)	RMSE (s)

XGBoost	0.809	0.811	0.243	0.077	0.937	0.029	0.176
KNN	0.684	0.688	0.374	0.12	0.97	0.031	0.232
MLP	0.66	0.668	0.542	0.173	0.814	0.116	0.241
SVR	0.578	0.587	0.555	0.178	0.883	0.099	0.271
RFR	0.84	0.844	0.259	0.084	0.746	0.029	0.166
GBR*	0.848	0.851	0.294	0.095	0.717	0.05	0.161
LR	0.335	0.354	0.89	0.282	0.722	0.253	0.339

*GBR was chosen for its superior balance, especially in the explained variance and RMSE metrics, which are crucial for Stride Time.

Regarding stride width - Table 8, MLP outperformed all models with the highest R^2 (0.918) and EVS (0.919), and lowest RAE (0.269), MAE (0.715), MaxE (0.033), MedAE (0.602), and RMSE (0.941), indicating minimal prediction errors. For stride velocity - Table 9, GBR achieved the highest R^2 (0.8565) and EVS (0.8585), along with the lowest RMSE (19.9245), demonstrating superior fit and low error spread. Though RFR had lower MAE (12.5205 vs. 12.9127), MedAE (6.7749 vs. 8.3038), and MaxE (83.1067 vs. 87.3396), and XGBoost had a lower RAE (0.2771 vs. 0.3108).

Table 8. Stride Width Results: Performance Evaluation Across Machine Learning Models.

Model	R^2	EVS	RAE	MAE (m)	MaxE (m)	MedAE (m)	RMSE (m)
XGBoost	0.715	0.718	0.53	0.014	0.05	0.012	0.018
KNN	0.420	0.425	0.727	0.019	0.080	0.015	0.025
MLP*	0.918	0.919	0.269	0.007	0.033	0.006	0.009
SVR	0.445	0.451	0.722	0.019	0.074	0.016	0.025
RFR	0.663	0.666	0.562	0.015	0.054	0.012	0.019
GBR	0.634	0.637	0.590	0.016	0.056	0.013	0.020
LR	0.859	0.861	0.356	0.009	0.041	0.008	0.012

*MLP showed a strong fit and minimal prediction errors, with consistently superior performance across all metrics.

Table 9. Stride Velocity Results: Performance Evaluation Across Machine Learning Models.

Model	R^2	EVS	RAE	MAE (m/s)	MaxE (m/s)	MedAE (m/s)	RMSE (m/s)
XGBoost	0.834	0.837	0.277	0.115	0.913	0.056	0.21
KNN	0.736	0.738	0.381	0.159	1.096	0.077	0.268
MLP	0.678	0.682	0.537	0.222	0.918	0.169	0.299
SVR	0.115	0.227	0.739	0.314	1.256	0.13	0.504
RFR	0.849	0.85	0.299	0.125	0.831	0.068	0.204
GBR*	0.856	0.858	0.311	0.129	0.873	0.083	0.199
LR	0.449	0.465	0.8	0.329	0.92	0.324	0.39

*GBR was selected for its strong performance in variance and RMSE metrics, important for stride velocity.

For base of support - Table 10, MLP showed the highest R^2 (0.898) and EVS (0.899), with the lowest RAE (0.311), MAE (0.820), MaxE (0.036), MedAE (0.643), and RMSE (0.011). In single and double support time - Tables 11–12, RFR achieved the highest R^2 (0.495) and EVS (0.500) in single support, with the lowest MAE (0.015), MaxE (0.091), MedAE (0.012), and RMSE (0.021), showing precise and stable predictions. Although KNN had a slightly lower RAE (0.689 vs. 0.695), RFR outperformed it in other metrics, supporting its selection.

Table 10. Support Base Results: Performance Evaluation Across Machine Learning Models.

Model	R^2	EVS	RAE	MAE (m)	MaxE (m)	MedAE (m)	RMSE (m)
XGBoost	0.662	0.665	0.591	0.016	0.057	0.013	0.02
KNN	0.417	0.424	0.746	0.02	0.084	0.015	0.026
MLP*	0.898	0.899	0.311	0.008	0.036	0.006	0.011
SVR	0.381	0.392	0.772	0.02	0.087	0.016	0.027
RFR	0.633	0.637	0.611	0.016	0.061	0.013	0.02
GBR	0.607	0.611	0.631	0.017	0.061	0.013	0.021
LR	0.833	0.834	0.407	0.011	0.042	0.009	0.014

*MLP showed a reliable fit and minimal prediction errors, with consistent superiority across all metrics.

For double support time, RFR also led with highest R^2 (0.548) and EVS (0.554), and lowest MAE (0.024), MaxE (0.148), MedAE (0.018), and RMSE (0.034). Despite XGBoost's marginally lower RAE (0.566 vs. 0.575), RFR's superior overall performance justified its choice.

The step-by-step of *vailá* Multimodal Toolbox process is available in the project's README file, the ReadTheDocs documentation, and the YouTube tutorial video in the link bellow https://github.com/vaila-multimodaltoolbox/vaila. These resources provide guidance on setting up the environment, running the models, and interpreting the results.

Table 11. Support Time Single Results: Performance Evaluation Across Machine Learning Models.

Model	R^2	EVS	RAE	MAE (s)	MaxE (s)	MedAE (s)	RMSE (s)
XGBoost	0.399	0.407	0.718	0.016	0.112	0.012	0.023
KNN	0.476	0.48	0.695	0.015	0.093	0.011	0.022
MLP	-6.304	-6.098	2.684	0.058	0.269	0.048	0.077
SVR	-0.752	-0.538	1.342	0.029	0.13	0.025	0.038
RFR*	0.495	0.5	0.689	0.015	0.091	0.012	0.021
GBR	0.418	0.425	0.738	0.016	0.098	0.013	0.023
LR	0.406	0.414	0.77	0.017	0.091	0.014	0.023

*RFR surpassed it in the other indicators, supporting its selection for the support time single results.

Table 12. Support Time Double Results: Performance Evaluation Across Machine Learning Models..

Model	R^2	EVS	RAE	MAE (s)	MaxE (s)	MedAE (s)	RMSE (s)
XGBoost	0.536	0.54	0.566	0.024	0.151	0.017	0.034
KNN	0.45	0.463	0.62	0.026	0.162	0.019	0.037
MLP	-1.228	-1.157	1.36	0.057	0.255	0.046	0.075
SVR	-0.309	-0.11	1.185	0.05	0.168	0.049	0.057
RFR*	0.548	0.554	0.575	0.024	0.148	0.018	0.034
GBR	0.471	0.476	0.621	0.026	0.154	0.02	0.037
LR	0.468	0.473	0.648	0.027	0.145	0.02	0.037

*RFR demonstrated better performance across the remaining indicators, validating its selection.

DISCUSSION

Kinematic gait analysis involves measuring a range of spatiotemporal variables, such as step time, stride time, step length, stride length, speed, and others to quantitatively evaluate the walking [23]. This analysis has been done mostly with marker-based methods, such as motion tracking systems that measure joint kinematic parameters by directly attaching markers to anatomical landmarks on the human body [24]. Although these motion tracking systems can track body motion with great precision and accuracy, they have several limitations, such as being time-consuming for experiment setup and data post-processing, requiring expertise for correct marker placement, and potentially altering one's natural body movement patterns [25].

The evolution of computational capacity and the advent of machine learning have facilitated the evaluation of kinematic data, primarily through human pose detection networks that apply computer vision techniques. Markerless pose estimation algorithms can be applied to new or old videos, provided sufficient image resolution, and while marker-based methods are limited by the marker set used during data collection, old markerless video data could be reprocessed with new pose estimation algorithms to improve accuracy or extract more in-depth measures [26]. Stenum et al. [14] validated the accuracy of spatiotemporal gait parameters estimated from pose detection against 3-dimensional motion capture and found errors within 0.02 s and 0.05 m for temporal and spatial parameters, respectively. In this context, the present study aimed to develop and validate a method that applies machine learning algorithms to predict gait measures traditionally obtained from systems like GAITRite™, using only pixel-based data derived from a computer vision pose detection model.

A study using DeepLabCut for retraining computer vision models for gait analysis reported for step length and stride time, mean absolute errors below 3.5 cm and 0.03 s, when comparing its output with that of the GAITRite™ system [5]. Our results show a MAE of 2.8 cm for step length and 0.077 s for stride time. This indicates that the combination of markerless pose estimation and machine learning in this study improved step length predictions but faced greater difficulty with stride time compared to previous research.

The results demonstrate a clear division in model performance based on the target variable. For metrics related to spatial characteristics, such as Step Length, Step Width, Stride Length, Stride Width, and Stride Velocity, the GBR consistently delivered the best results, exhibiting lower error metrics and higher R^2 scores. Only for the Support Base data, the chosen model differed, with the RFR outperforming the other models. This model also displayed low error values and high R^2 and EVS scores.

For metrics associated with temporal characteristics, such as Step Time, Stride Time, Single Support Time, and Double Support Time, the RFR outperformed other models. This suggests that the data distribution and relationships among temporal features might align more effectively with the ensemble structure of the Random Forest algorithm.

The variables that exhibited strong model performance included Step Length, Step Width, Stride Time, Stride Width, Stride Length, Stride Velocity, and Support Base. These models achieved high values of explained variance, indicating a substantial proportion of the variance in the data was captured, alongside consistently low prediction errors, reflecting accurate and reliable estimations.

Conversely, the models predicting Step Time, Single Support Time, and Double Support Time demonstrated comparatively lower explained variance values, marginally exceeding 50%. Despite this, these models maintained low error metrics, suggesting that while the overall variance explained was moderate, the predictions remained precise within the scope of the data.

Even with the R^2 considered a strong performance in machine learning regression tasks — particularly in the context of human

motion, where noise from pose estimation, inter-individual variability, and recording conditions naturally introduce some irreducible variance, it is also known that R^2 does not directly quantify the clinical impact of prediction errors. Clinical utility depends not only on the proportion of variance explained but also on whether the absolute error falls within clinically acceptable thresholds, such as the minimal clinically important difference — MCID. In this proof-of-concept study, the aim was to demonstrate the feasibility of using markerless video-based methods to estimate spatiotemporal gait parameters with reasonably high accuracy.

Despite these advancements, the application of machine learning algorithms and markerless methods requires knowledge of computer vision and advanced programming skills, which can limit accessibility and application. One of the goals of this study was to address this limitation by facilitating access to these analyses through the inclusion of the best-performing models in the *vailá* Multimodal Toolbox [20]. Furthermore, as an open-source tool, *vailá* is freely available to the community, encouraging collaboration, continuous improvement, and the integration of new applications. This openness not only enhances the tool's accessibility but also fosters innovation and broader adoption in the field.

Although the present study employed MediaPipe for landmark detection due to its accessibility and ease of integration, alternative approaches such as the YOLOv11 network [27] may offer advantages in scenarios requiring multi-person tracking or greater robustness under challenging video conditions. YOLO-based models are known for their high detection accuracy and real-time performance, but they demand higher computational resources and more complex implementation workflows. While YOLOv11 was not used in this study, its potential integration into the *vailá* framework could be explored in future work to enhance flexibility and adaptability across diverse gait analysis contexts.

This study has several limitations. One notable limitation is related to the GAITRite™ equipment, which occasionally provides zero values for stride variables. These zero values can significantly impact the performance of certain machine learning algorithms, necessitating adjustments during data processing. Additionally, the data used for training were derived from images captured by a camera with specific parameters, including image resolution, video codec, acquisition frequency, and the distance between the GaitRite™ and the camera. Consequently, we cannot rule out the possibility that variations in these parameters may have influenced the predictive performance of the models. Another limitation is the potential challenge machine learning algorithms face in generalizing to populations that differ from the training dataset, particularly when analyzing gait parameters in clinical populations or across diverse age groups. Furthermore, algorithms such as Gradient Boosting can be computationally intensive, and their feasibility may be limited when working with large datasets on devices with restricted computational capacity.

CONCLUSION

This study validates the application of machine learning algorithms to predict gait metrics traditionally obtained from GAITRite™ systems, using only pixel-based data from markerless pose estimation. Among the tested models, the XGBoost demonstrated superior performance in step parameters, while Gradient Boosting Regressor excelled in stride length, time, and velocity. Multi-Layer Perceptron outperformed other models in stride width, and Random Forest Regressor showed the best results for single and double support times.

These findings demonstrate the feasibility of using low-cost, accessible tools for precise gait analysis, broadening its applicability in diverse settings, including clinical and research environments. This work contributes to democratizing gait analysis by reducing reliance on specialized hardware, ensuring broader accessibility without compromising measurement reliability.

REFERENCES

1. Parati M, Ambrosini E, De Maria B, Gallotta M, Dalla Vecchia LA, Ferriero G, et al. The reliability of gait parameters captured via instrumented walkways: a systematic review and meta-analysis. European Journal of Physical and Rehabilitation Medicine. 2022;58(3):363-77. doi: 10.23736/S1973-9087.22.07037-X
2. McDonough AL, Batavia M, Chen FC, Kwon S, Ziai J. The validity and reliability of the GAITRite system's measurements: A preliminary evaluation. Archives of physical medicine and rehabilitation. 2001;82(3):419-25. doi: 10.1053/apmr.2001.19778
3. Van Uden CJ, Besser MP. Test-retest reliability of temporal and spatial gait characteristics measured with an instrumented walkway system (GAITRite®). BMC musculoskeletal disorders. 2004;5:1-4. doi: 10.1186/1471-2474-5-13.
4. Vallabhajosula S, Humphrey SK, Cook AJ, Freund JE. Concurrent validity of the Zeno walkway for measuring spatiotemporal gait parameters in older adults. Journal of geriatric physical therapy. 2019;42(3):E42-50. doi: 10.1519/JPT.0000000000000168
5. Panconi G, Grasso S, Guarducci S, Mucchi L, Minciacchi D, Bravi R. Deep-Learning-Based Marker less Pose Estimation Systems in Gait Analysis: DeepLabCut Custom Training and the Refinement Function. arXiv preprint arXiv:240710590. 2024. doi: 10.48550/arXiv.2407.10590
6. Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. IEEE transactions on pattern analysis and machine intelligence. 2019;43(1):172-86. doi: 10.1109/TPAMI.2019.2929257
7. Pishchulin L, Insafutdinov E, Tang S, Andres B, Andriluka M, Gehler PV, et al. Deepcut: Joint subset partition and labeling for multi person pose estimation. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 4929-37. doi:10.48550/arXiv.1511.06645

8. Lugaresi C, Tang J, Nash H, McClanahan C, Uboweja E, Hays M, et al.. MediaPipe: A Framework for Building Perception Pipelines; 2019. Available from: https://arxiv.org/abs/1906.08172. doi: 10.48550/arXiv.1906.08172
9. Kanazawa A, Zhang JY, Felsen P,Malik J. Learning 3d human dynamics from video. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2019. p. 5614-23. doi: 10.48550/arXiv.1812.01601
10. Pavllo D, Feichtenhofer C, Grangier D, Auli M. 3d human pose estimation in video with temporal convolutions and semi-supervised training. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2019. p. 7753-62. doi:10.1109/CVPR.2019.00794
11. Kocabas M, Athanasiou 325 N, Black MJ. Vibe: Video inference for human body pose and shape estimation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2020. p. 5253-63. doi:10.48550/arXiv.1912.05656
12. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nature neuroscience. 2018;21(9):1281-9. doi: 10.1038/s41593-018-0209-y
13. Mathis MW, Mathis A. Deep learning tools for the measurement of animal behavior in neuroscience. Current opinion in neurobiology. 2020;60:1-11. doi: 10.1016/j.conb.2019.10.008
14. Stenum J, Rossi C, Roemmich RT. Two-dimensional video-based analysis of human gait using pose estimation. PLoS computational biology. 2021;17(4):e1008935. doi: 10.1371/journal.pcbi.1008935
15. Guffanti D, Brunete A, Hernando M, A´lvarez D, Rueda J, Navarro E. Supervised learning for improving the accuracy of robot-mounted 3D camera applied to human gait analysis. Heliyon. 2024;10(4). doi:10.48550/arXiv.2207.01002
16. Nazari F, Mohajer N, Nahavandi D, Khosravi A. Comparison of gait phase detection using traditional machine learning and deep learning techniques. In: 2022 IEEE international conference on systems, man, and cybernetics (SMC). IEEE; 2022. p. 403-8. doi:10.1109/SMC53654.2022.9945397
17. Vandermeeren S, Bruneel H, Steendam H. Feature selection for machine learning based step length estimation algorithms. Sensors. 2020;20(3):778. doi:10.3390/s20030778
18. Uhlrich SD, Falisse A, Kidzin´ski L , Muccini J, Ko M, Chaudhari AS, et al. OpenCap: Human movement dynamics from smartphone videos. PLoS computational biology. 2023;19(10):e1011462. doi I: 10.1371/journal.pcbi.1011462
19. Needham L, Evans M, Cosker DP, Wade L, McGuigan PM, Bilzon JL, et al. The accuracy of several pose estimation methods for 3D joint centre localisation. Scientific reports. 2021;11(1):20673. doi: 10.1038/s41598-021-00212-x
20. Santiago PRP, Chinaglia AG, Flanagan K, Bedo BLS, Mochida LY, Aceros J, et al.. vail´a: Versatile Anarcho Integrated Liberation A´nalysis in Multimodal Toolbox; 2024. Available from: https://arxiv.org/abs/2410.07238. doi: https://doi.org/10.48550/arXiv.2410.07238
21. Raschka S, Mirjalili V. Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2. Packt publishing ltd; 2019.
22. Bishop CM, Nasrabadi NM. Pattern recognition and machine learning. vol. 4. Springer; 2006. doi:10.1117/1.2819119
23. Chen W, Xu Y, Wang J, Zhang J. Kinematic analysis of human gait based on wearable sensor 360 system for gait rehabilitation. Journal of Medical and Biological Engineering. 2016;36:843-56. doi:10.1007/s40846-016-0179-z
24. Reissner L, Fischer G, List R, Giovanoli P, Calcagni M. Assessment of hand function during activities of daily living using motion tracking cameras: A systematic review. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. 2019;233(8):764-83. doi: 10.1177/0954411919851302
25. Carse B, Meadows B, Bowers R, Rowe P. Affordable clinical gait analysis: An assessment of the marker tracking accuracy of a new low-cost optical 3D motion analysis system. Physiotherapy. 2013;99(4):347-51. doi: 10.1016/j.physio.2013.03.001
26. Wade L, Needham L, McGuigan P, Bilzon J. Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. PeerJ. 2022;10:e12995. doi: 10.7717/peerj.12995
27. Jocher G, Qiu J. Ultralytics YOLO11; 2024. Available from: https://github.com/ultralytics/ultralytics.

ACKNOWLEDGEMENTS

We want to express our gratitude to the School of Physical Education and Sports of Ribeirão Preto, University of São Paulo; the Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil; and the School of Physical Education and Sport, University of São Paulo, São Paulo, Brazil, for providing support during the research period.

Availability of data and materials

The *vailá* Multimodal Toolbox is available at our GitHub repository: https://github.com/vaila-multimodaltoolbox/vaila. All code is open-source, and the repository includes detailed documentation, example datasets, and installation scripts for Linux, macOS, and Windows. Additional documentation, including information about the *vailá* project and instructions for executing the ML Walkway, can be found on ReadTheDocs at https://vaila.readthedocs.io/en/latest/.

Citation: Tahara AK, Chinaglia AG, Monteiro RLM, Bedo BLS, Cesar GM, Santiago PRP. (2025). Differences in delayed onset muscle soreness behavior and joint position sense in men and women. *Brazilian Journal of Motor Behavior*, 19(1):e462.
Editor-in-chief: Dr Fabio Augusto Barbieri - São Paulo State University (UNESP), Bauru, SP, Brazil.
Associate editors: Dr José Angelo Barela - São Paulo State University (UNESP), Rio Claro, SP, Brazil; Dr Natalia Madalena Rinaldi - Federal University of Espírito Santo (UFES), Vitória, ES, Brazil; Dr Renato de Moraes – University of São Paulo (USP), Ribeirão Preto, SP, Brazil.

Funding: This study was partially funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. This study was financed, in part, by the São Paulo Research Foundation (FAPESP), Brazil. Process Numbers #2019/17729-0, #2021/15134-9 #2024/15658-6 and #2024/17521-8. It also received funding from the National Council for Scientific and Technological Development (CNPq—Brasil) (Process Numbers #312597/2021-5 and #432259/2018-0). This research was funded by the 2023 Pediatric Physical Therapy Research Grant from the Foundation for Physical Therapy Research. Additionally, financial support was provided by the Dean's Office for Research and Innovation of the University of São Paulo through the Support Program for New Professors (BLSB).
Competing interests: The authors have declared that no competing interests exist.
DOI: https://doi.org/10.20338/bjmb.v19i1.462