Computational and Structural Biotechnology Journal 21 (2023) 3414–3423

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research article

Modeling biological individuality using machine learning: A study on human gait

Fabian Horst [a,*,1], Djordje Slijepcevic [b,1], Marvin Simak [a], Brian Horsak [c,d], Wolfgang Immanuel Schöllhorn [a], Matthias Zeppelzauer [b]

[a] Department of Training and Movement Science, Institute of Sport Science, Johannes Gutenberg-University Mainz, Mainz, Germany
[b] Institute of Creative Media Technologies, Department of Media & Digital Technologies, St. Pölten University of Applied Sciences, St. Pölten, Austria
[c] Center for Digital Health & Social Innovation, Department of Health Sciences, St. Pölten University of Applied Sciences, St. Pölten, Austria
[d] Institute of Health Sciences, St. Pölten University of Applied Sciences, St. Pölten, Austria

ARTICLE INFO

Article history:
Received 15 February 2023
Received in revised form 9 June 2023
Accepted 10 June 2023
Available online 13 June 2023

Keywords:
Human gait recognition
Biomechanics
Ground reaction forces
Explainable artificial intelligence
Layer-wise relevance propagation
Force-based gait recognition

ABSTRACT

Human gait is a complex and unique biological process that can offer valuable insights into an individual's health and well-being. In this work, we leverage a machine learning-based approach to model individual gait signatures and identify factors contributing to inter-individual variability in gait patterns. We provide a comprehensive analysis of gait individuality by (1) demonstrating the uniqueness of gait signatures in a large-scale dataset and (2) highlighting the gait characteristics that are most distinctive to each individual. We utilized the data from three publicly available datasets comprising 5368 bilateral ground reaction force recordings during level overground walking from 671 distinct healthy individuals. Our results show that individuals can be identified with a prediction accuracy of 99.3% by using the bilateral signals of all three ground reaction force components, with only 10 out of 1342 recordings in our test data being misclassified. This indicates that the combination of bilateral ground reaction force signals with all three components provides a more comprehensive and accurate representation of an individual's gait signature. The highest accuracy was achieved by (linear) Support Vector Machines (99.3%), followed by Random Forests (98.7%), Convolutional Neural Networks (95.8%), and Decision Trees (82.8%). The proposed approach provides a powerful tool to better understand biological individuality and has potential applications in personalized healthcare, clinical diagnosis, and therapeutic interventions.

1. Introduction

Classification and structuring the object of research is a fundamental aspect of biology. The process of classification involves identifying traits or characteristics unique to specific groups of organisms. Since Linne's fundamental work on the taxonomy in biology in 1735, various criteria for classification have been proposed, and advancements in technology and research methods have led to increasingly sophisticated classification systems over time.

In recent years, the development of classification systems in computational and structural biology has been influenced by the utilization of data-driven approaches, particularly (deep) machine learning (ML) methods. These methods are based on computational algorithms that enable the analysis and detection of patterns in the provided data. Employing ML methods for biological data could facilitate the exploration of novel criteria for taxonomies across different hierarchical levels (e.g., kingdoms, classes, families, and species) [1,2] and opened up the possibility for distinguishing and analyzing *individual organisms* within a species in more detail. Taking into account the diversity within species may lead to new insights and a more complete understanding of the biology and ecology of organisms and their interactions with the environment [3].

Using human gait as an example, we will demonstrate the application of ML methods to model biological individuality in biomechanics. The ability to recognize individual humans by their gait pattern, i.e., gait recognition, is a widespread phenomenon that has been studied in recent decades. This is indicated by studies showing

* Corresponding author.
E-mail address: horst@uni-mainz.de (F. Horst).
[1] Both authors contributed equally to this research.

https://doi.org/10.1016/j.csbj.2023.06.009

that individuals can identify friends and colleagues visually by their walking style [4], even at a distance and with limited visibility [5]. In recent years, biomechanical studies have corroborated these observations using ML-based classification. ML methods (e.g., artificial neural networks and support vector machines) were used to demonstrate that individuals can be accurately identified based on kinetic (e.g., 99.8% in 128 individuals [6]), kinematic (e.g., 100.0% in 57 individuals [7]), and electromyographic (e.g., 98.9% in 79 individuals [8]) data measured during walking. Further studies showed that biomechanical gait patterns allow for identifying individuals across different days [8,9], months [6], emotional states [10], and fatigue [11]. These findings suggest that human gait features satisfy even stringent biometric demands, such as possessing a high degree of uniqueness (i.e., a feature should not be identical for any two persons) and permanence (i.e., a feature should remain identifiable over time) [12]. In analogy to the handwritten signature, which is often used as proof of a person's identity in daily life, the notion of "gait signatures" emerged in recent years [8,13]. The gait signature describes the unique gait patterns of an individual person.

Beyond the classification of gait signatures, recent work has addressed the question of what makes a person's gait unique. One possibility to approach this question is to analyze which input features or portions of the input signal the ML models use for classification. To this end, explainability methods are used to characterize gait signatures by highlighting which gait features were used by the ML models to identify an individual person [7,14].

To date, biomechanical research on gait signatures has been conducted primarily on small-scale datasets. The growing number of publicly available ground reaction force (GRF) datasets in recent years [15–17] offers the possibility to investigate gait signatures on a larger scale. Following previous work on gait signatures, this work examines the individuality of GRF-based gait signatures by (1) proving their uniqueness in a large-scale dataset and (2) highlighting gait characteristics that are unique to each individual.

2. Related work

The idea that each individual has a distinct (unique) walking pattern (gait signature) has gained attention for various purposes, such as healthcare, authentication, and surveillance. To demonstrate the uniqueness of human gait patterns, a variety of sensor modalities have been utilized in the literature [18]. These include (i) vision-based approaches, which employ RGB, infrared, or depth cameras to extract gait features from images or videos, (ii) force- or pressure-based approaches, which use instrumented platforms or insoles to measure underfoot pressure distribution or GRFs during the stance phase, and (iii) accelerometry-based approaches, which record acceleration using an accelerometer or inertial measurement unit (IMU) in a mobile phone or wearable device.

Research on the uniqueness of gait signatures has primarily focused on the gait-based identification or re-identification of individuals (i.e., gait recognition) in surveillance scenarios [18]. Gait recognition in surveillance scenarios is often performed at a (large) distance *without* "user-interaction", i.e., no active participation of the monitored individual. To allow for remote recognition and monitoring, vision-based approaches are predominantly used. In addition to gait-related features, anthropometric characteristics (e.g., a person's height, and body proportions) are commonly utilized to compensate for variations in clothing, camera angles, and other variables that can bias gait recognition. Incorporating anthropometric attributes can enhance accuracy and reduce false positive error rates [19]. The emergence of deep learning methods has greatly advanced the field of image classification, enabling ML models to achieve performance in vision-based subject re-identification tasks with accuracies of 89.2% and 99.9% on large-scale datasets, such as the OU-ISIR [20] and the OU-MVLP [21] dataset with 4007 and 10,307 individuals, respectively [22].

In the context of healthcare scenarios, biomechanical research on gait signatures diverges from surveillance research by analysing intrinsic properties of the musculoskeletal locomotor system rather than focusing on the task of re-identification. A goal in biomechanics is to model the individual's gait by identifying subject-specific characteristics (e.g., for tailoring interventions to subject-specific needs). Biomechanical gait analysis is performed at rather small distance *with* "user-interaction", i.e., markers and electrodes are placed on the person's body, persons are familiarised to the experimental protocol and follow instructions from the examiner. Three-dimensional gait analysis (3DGA) using high-resolution recording devices (e.g., infrared cameras and force platforms) in standardized laboratory conditions is an established standard in the biomechanics community. 3DGA focuses on the quantitative description and biomechanical analysis of human gait from a kinematic (i.e., joint angles), kinetic (i.e., ground reaction forces and joint moments), and muscular (i.e., electromyographic activity) point of view [23,24]. The extensive amount of effort required to collect all these data modalities for each individual in 3DGA results in a considerably smaller pool of available datasets compared to the vision-based approaches in surveillance scenario. GRF data represent a trade-off between the time- and resource-consuming to capture but accurate 3DGA data and the rather easy to capture but less accurate vision-based data from RGB cameras. GRF data provide a high accuracy making it suitable for biomechanical applications. However, they provide less differentiated information on the gait process as the motion information of all joints gets superimposed in the GRF data. In the absence of body segment- and joint-specific motion information, classification of gait signatures based on GRFs can be considered a challenging task.

GRF measurements provide high-resolution data that are recorded with force or pressure platforms embedded in the ground. For the analysis using GRFs, often either geometric features, spectral features, or a combination of both are employed. Geometric features are mostly based on key events in the GRF profile, such as local peaks and valleys of the characteristic M-shaped waveform [25]. Spectral features can be obtained by signal transformations, such as the Fast Fourier Transform (FFT) for power spectral densities [26]), or the Wavelet Packet Decomposition (WPD) [27,28]. In contrast to these feature-oriented approaches, studies used the entire waveform as a *holistic* description of the GRFs.

In total, we identified ten studies that used a time-continuous (holistic) representation of GRF data to investigate the uniqueness of gait signatures. Table 1 provides an overview of the current state of research on the classification of gait signatures using a time-continuous representation of GRF data. We included only studies that examined GRFs during walking at self-selected walking speed without additional interventions (e.g., carrying loads). All identified studies were limited to healthy individuals.

The identified studies varied in several aspects, including the number of individuals studied (N), the number of gait recordings per individual, the footwear condition, the dimensionality and laterality of GRFs collected (only the vertical GRF (GRF_V) component or all three GRF components (GRF_{3D}); unilateral or bilateral), the features used to characterize the gait, and the ML method applied. Some studies had as few as 9 individuals, and others involved up to 200. Similarly, the number of gait recordings per individual ranged from 6 to over 100. Both parameters are central for contextualizing the obtained results in terms of their generalizability. It is also important to note that most individuals wore shoes, but several studies only provide vague information on the type of worn shoes. The studies used different ML methods for classification, including hidden Markov models (HMMs), k-nearest neighbor (k-NN) with dynamic

Table 1
A review of studies that classified gait signatures based on a time-continuous representation of GRF data from 1997 to 2023, including the number of individuals studied (N), recordings per individual, footwear conditions (barefoot or shod), type of GRF components used (GRF_V for vertical only or GRF_{3D} for all three GRF components), laterality used (bilateral or unilateral), classification method applied, and classification performance in terms of Correct Classification Rate (CCR), or Equal Error Rate (EER). The studies are presented in chronological order. Data not provided in the studies are marked as not available (na). For the sake of clarity, only the results of the best approach of a study are shown in the table.

Reference	N	Recordings	Footwear	Dimensionality	Laterality	ML method	Performance
Addlesee et al. [29]	15	20	na	GRF_V	Unilateral	HMM	CCR: 91%
Vera-Rodriguez et al. [30]	17	~ 170	Shod	GRF_V	Unilateral	PCA+SVM	EER: 12.5%
Janssen et al. [10]	38	9–12	Barefoot	GRF_{3D}	Unilateral	SVM	CCR: 98.5%
Janssen et al. [11]	9	6	Barefoot	GRF_{3D}	Unilateral	SVM	CCR: 100%
Derlatka [31]	132	14–20	Shod	GRF_{3D}	Bilateral	k-NN (DTW)	EER: ~ 2.8%
Derlatka and Bogdan [32]	200	14–22	Shod	GRF_{3D}	Bilateral	k-NN (DTW)	CCR: 97.4%
Marcin [33]	81	28–40	Shod	GRF_{3D}	Bilateral	k-NN (DTW)	CCR: 99.0%
Horst et al. [6]	128	10	Barefoot	GRF_{3D}	Bilateral	SVM	CCR: 99.8%
Horst et al. [7]	57	20	Barefoot	GRF_{3D}	Bilateral	SVM	CCR: 100.0%
Duncanson et al. [34]	118	8–10	Shod	GRF_{3D}	Bilateral	ConvRNN	CCR: 96.0%

time warping (DTW), support vector machines (SVMs), and convolutional recurrent neural networks (ConvRNNs).

The studies show that holistic GRF analysis can be used to accurately identify individual gait signatures, with prediction accuracies over 90%. High accuracies were achieved by analyzing the GRF_{3D} using SVMs. Both bilateral [6,7], as well as unilateral data [10,11] performed well with correct classification rates (CCR) of 98.5% (unilateral) up to 100.0% (bilateral). Other studies using k-NN (DTW) [32,33] and ConvRNN [34] classifiers achieved similar levels of accuracy, with CCRs ranging from 96.0% to 99.0%. Using only the unilateral GRF_V Addlesee et al. [29] achieved a slightly worse outcome of 91%. Two studies did not report the CCR but the Equal Error Rate (EER) - an often used metric in authentication. The EER is where the percentage of false acceptances and false rejections is the same. While Derlatka [31] achieved a low EER of ~ 2.8%, again with bilateral GRF_{3D}, Vera-Rodriguez et al. [30] got a higher EER with unilateral GRF_V and an approach based on dimensionality reduction by principal component analysis (PCA).

Overall, these studies suggest that the identification of gait signatures using a holistic (time-continuous) representation of GRFs is a promising approach with high accuracy rates achieved. The studies suggest that using bilateral GRF_{3D} data can improve prediction accuracy compared to using unilateral GRF_V. Additionally, the studies found that using shoes during gait analysis did not significantly affect classification accuracy.

3. Methods

3.1. Datasets

For our classification experiments, we utilized subsets of the AIST Gait Database [15], the GaitRec dataset [16], and the Gutenberg Gait Database [17], which contain bilateral GRF recordings during level overground walking. The number of gait analysis sessions and the number of gait recordings per session varied among participants and datasets. Our classification experiments were standardized by including only data from participants who walked barefoot, were injury-free, and walked at a self-selected (preferred) speed. Moreover, we used data from the initial gait analysis session of participants who had at least eight gait recordings available. For participants who had more than eight gait recordings available, we randomly selected eight recordings. In total, we used 5368 gait recordings from 671 distinct participants. Table 2 summarizes demographic details for the different datasets and provides an overview of the data used for our classification experiments.

3.2. Data recording

Bilateral analog force plate signals were recorded by asking participants to walk at their self-selected (preferred) speed on a level and approximately 10 m long walkway. The analog force plate signals of successive stance phases of the right and left foot were recorded by two or more force plates. For more details on the hardware configurations and settings used, refer to the descriptions of the underlying datasets (i.e., AIST Gait Database [15], GaitRec dataset [16], and Gutenberg Gait Database [17]).

3.3. Data processing

The three-dimensional GRF components – anterior-posterior (GRF_{AP}), medio-lateral (GRF_{ML}), and vertical (GRF_V) – were calculated in a uniform way based on the analog force plate signals. To ensure consistency and comparability across all datasets, we standardized the data processing steps. The analog force plate signals were downsampled to 250 Hz and the GRF signals were filtered using a second-order Butterworth bidirectional low-pass filter with a cut-off frequency of 20 Hz. The stance phase was determined based on the filtered GRF signals using a GRF_V threshold of 25 N. Each GRF signal was time-normalized to 101 data points, corresponding to 100% stance phase, and normalized to the body weight in units of $N/(kg{*}9.81m/s^2)$. The GRF_{AP} and GRF_{ML} signals were unified to ensure consistency regardless of the walking direction. For this purpose, the medial and anterior forces were transformed into positive values, while the lateral and posterior forces were transformed into negative values. Overall, this standardized data processing approach ensured that the gait recordings were directly comparable across all datasets.

3.4. Data analysis

The GRF recordings were primarily classified with a linear Support Vector Machine (SVM), due to its runtime efficiency [35]. In addition, we also investigated other ML methods, such as SVMs with polynomial and radial basis kernel functions, Convolutional Neural Networks (CNNs), Random Forests (RFs), and Decision Trees (DTs). For the linear SVM, we utilized the L2-regularized L2-loss and a linear kernel function implemented in the Liblinear Toolbox (version 1.4.1) [36]. To determine the optimal hyperparameters for each ML method (except for DTs), we performed a grid search using a two-fold cross-validation on the training data. For the SVM, we investigated the regularization parameter C, performing a grid search within the range of $C = \{10^{-1}, 10^{-0.5}, \ldots, 10^{3.5}, 10^{4}\}$. In the case of the CNNs, we explored different configurations including the number of convolutional layers and the number of filters {{32, 32}, {32, 32, 32, 32}, {32, 32, 32, 32, 32, 32}, {32, 64}, {32, 32, 64, 64}}, and the size of

Table 2
Demographic details of the utilized dataset (including details on individual dataset level).

Dataset	N	Sex (m/f)	Age (yrs.) Mean (SD)	Body Mass (kg) Mean (SD)	Body Height (m) Mean (SD)	Recordings
AIST [15]	266	123/143	49.2 (19.2)	59.2 (10.3)	1.63 (0.08)	2128
GaitRec [16]	56	25/31	35.9 (10.9)	72.9 (15.2)	1.72 (0.08)	448
Gutenberg [17]	349	205/142	24.2 (7.0)	70.9 (12.0)	1.76 (0.09)	2792
Total	**671**	**353/316**	**35.6 (18.1)**	**66.4 (13.1)**	**1.70 (0.11)**	**5368**

The information presented in the table refers only to the data included in our classification experiments and not to the complete underlying datasets.

the dense layer {256, 512}, while certain parameters remained constant, i.e., dropout rate of 0.25, batch size of 256, filter size of three and learning rate of 10^{-3}. Due to the sensitivity of RFs to the number of individual trees, we conducted a grid search over this hyperparameter using the values {100, 200, 300}.

The GRF signals were normalized to the absolute maximum value determined in the training data. Specifically, each signal was divided by the absolute maximum value in the training data, resulting in signals with a value range between − 1 and + 1. Furthermore, we concatenated the normalized GRF signals into an input vector before providing it to the ML models. The size of the input vector depended on the lengths (and number) of the signals used, e.g., using bilateral data of all three GRF components resulted in a vector size of 1×606.

3.4.1. Performance evaluation

The 5368 gait recordings were divided into two parts for the classification experiments. The larger part, containing 4026 recordings (75%), was used to train the ML models. The remaining 1342 recordings (25%) served as test data for evaluating the performance of the ML models. To maintain a consistent distribution of each participant's data, six gait recordings from each participant were randomly selected to serve as training data and the two remaining recordings were used as test data. We used the same training and test split in all classification experiments.

3.4.2. Explainability evaluation

The explainability results were derived from the ML model that demonstrated the highest performance for classifying bilateral GRF_{3D} data, i.e., the linear SVM (Table 3). To analyze which input features are used by the SVM model, we utilized Layer-wise relevance propagation (LRP) [37]. LRP is a popular explainability method designed to reveal how ML models work and how they ground their predictions. To this end, LRP decomposed the predictions of the trained SVM models into relevance scores for each value i in the input vector. The relevance scores R_i were calculated based on the product of each value x_i of the input vector x and the weight w_i of the weight vector w of the trained SVM models:

$$R_i = x_i{}^* w_i \tag{1}$$

Relevance scores indicate which information (and to what extent) was used by the SVM model for its prediction. Positive scores represent input features supporting the classification, while negative scores represent input features that are against a given classification. For this work, the ground truth class labels (i.e., participant labels) were decomposed, and only positive input relevance scores were analyzed. Subsequently, positive relevance scores were normalized to their respective maximum. To obtain an indicator of the overall relevance of the input features for this classification task, we calculated the total relevance [38] for the trained models, which is the absolute sum of the relevance values obtained for all recordings from the test data. The total relevance was then aggregated at the GRF component and stance phase level and visualized using the code provided by Hoitz et al. [39]. All data processing and analysis were performed within the MATLAB 2021b (MathWorks, USA) framework.

Table 3
Performance evaluation on the test data across different ML methods, i.e., Support Vector Machines (SVMs), Convolutional Neural Networks (CNNs), Random Forests (RFs) and Decision Trees (DTs). The ML models were trained on bilateral GRF_{3D}.

ML method	Accuracy	Precision	Recall	F1 Score
SVM (linear)	**99.3%**	**99.2%**	**99.3%**	**99.1%**
SVM (polynomial)	96.7%	97.1%	96.7%	96.3%
SVM (radial basis)	97.3%	97.6%	97.3%	97.0%
Convolutional Neural Network	95.8%	96.4%	95.8%	95.3%
Random Forest	98.7%	99.1%	98.7%	98.6%
Decision Tree	82.8%	85.0%	82.8%	81.5%

4. Results

This section first presents the quantitative classification results for automated identification of gait signatures and subsequently provides explanations for the predictions. As a fixed train-test split was used, we report the classification performance and explainability results obtained on the test data.

4.1. Classification results

Table 3 presents the classification performance results of different ML methods for classifying individuals based on the bilateral GRF_{3D} signals during walking. The highest prediction accuracy, i.e., 99.3%, was achieved by the linear SVM model. The linear SVM correctly classified 1332 out of 1342 gait recordings to their respective individuals, with only 10 misclassifications. The RF ranked second with an accuracy of 98.7%, followed by the two other SVM variants (97.3% and 96.7%) and the CNN (95.8%). The DT performed significantly worse than the other methods, with an accuracy of 82.8%.

For the ML method with the highest performance in classifying bilateral GRF_{3D} data, i.e., the linear SVM (Table 3), we conducted a separate post-hoc analysis of the classification results. First, we formed groups based on metadata information such as age, sex, and the originating dataset. Subsequently, we calculated and visualized the amount of correctly and incorrectly classified gait recordings for these groups. Fig. 1 shows the results of the post-hoc analysis. In terms of age, all but one misclassification were observed in the group of young individuals, while no misclassification was found among older individuals. With respect to sex, the misclassifications were evenly distributed among male and female individuals. Regarding the datasets, GaitRec had no misclassifications, while AIST and Gutenberg had four and six misclassifications, respectively. Due to the very high overall classification accuracy of 99.3% (Table 3), the number of misclassifications is very low, amounting to only 10 cases. However, the post-hoc analysis reveals that the SVM did not struggle to model certain sub-groups in the data, as the classification accuracies in the post-hoc analysis are above 99.0% for all groups.

In a stress test, we investigated how the amount of samples (per individual) in the training data affects the generalization ability of the linear SVM. In this side experiment, the test data remained unchanged (containing the same two samples per individual as in the previous experiments). We examined two dimensions by using different subsets of the training data, i.e., by varying the total number of individuals and by varying the number of samples per individual. First, we randomly divided the training data into 10 folds of equal

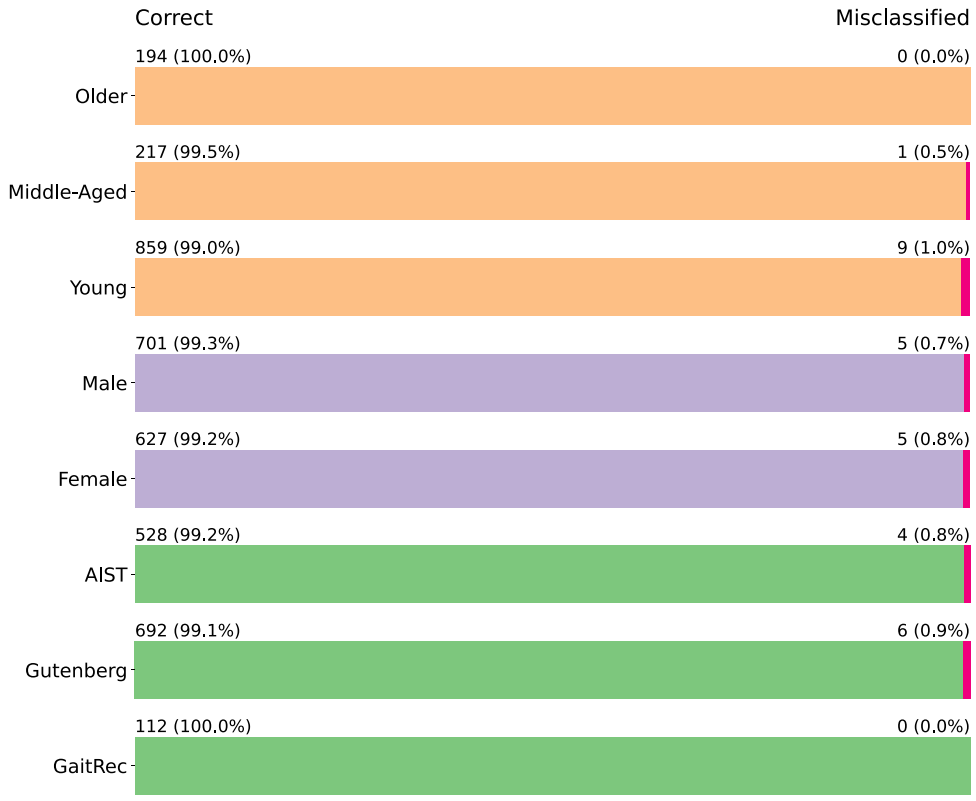

Fig. 1. Post-hoc analysis of the prediction accuracy on the test data with respect to metadata information (i.e., age, sex, and underlying dataset). Participants were divided into three age groups: young (1–39 years), middle-aged (40–64 years), and older (65–99 years) adults. The analyzed SVM model was trained on bilateral GRF_{3D}.

Table 4
Evaluation of the prediction accuracy for (linear) SVMs with respect to the number of considered individuals (ranging from 67 to 671 individuals over 10 folds) and the number of training samples per individual (ranging from one to six samples per individual). The SVMs were trained on bilateral GRF_{3D}.

N	1 Sample	2 Samples	3 Samples	4 Samples	5 Samples	6 Samples
67 (1/10)	91.0%	97.8%	99.3%	100.0%	100.0%	100.0%
134 (2/10)	94.0%	98.1%	99.3%	99.3%	99.6%	99.6%
201 (3/10)	91.5%	97.8%	98.8%	99.3%	99.8%	99.8%
268 (4/10)	89.7%	96.3%	98.1%	98.7%	99.1%	99.4%
336 (5/10)	89.7%	96.3%	98.1%	98.2%	99.4%	99.4%
403 (6/10)	89.8%	96.4%	98.3%	98.9%	99.4%	99.5%
470 (7/10)	89.4%	96.8%	98.3%	98.3%	99.0%	99.5%
537 (8/10)	88.4%	96.5%	97.7%	98.5%	98.8%	99.2%
604 (9/10)	88.2%	96.5%	98.1%	98.3%	98.8%	99.3%
671 (10/10)	88.5%	96.5%	97.8%	98.3%	98.8%	99.2%

Table 5
Evaluation of the prediction accuracy for (linear) SVMs with respect to different GRF dimensionalities and lateralities.

Input	GRF_{AP}	GRF_{ML}	GRF_{V}	GRF_{3D}
Unilateral (right)	82.5%	78.0%	87.7%	97.9%
Unilateral (left)	82.9%	79.0%	86.5%	98.3%
Bilateral (right+left)	91.7%	91.7%	91.6%	**99.3**%

size of individuals (N = 67) and trained a linear SVM on one fold. Then, the number of folds used for training was gradually increased until all 10 folds were used. With regard to the second dimension, we modified the training data to contain either one, two, three, four, five, or six samples per individual. Table 4 presents the results of the stress test for the linear SVM. In the experiment, we initially trained/tested an SVM model on 67 individuals. As the number of individuals in the training/test data increased, the performance of the SVMs decreased slightly and reached its worst performance with the entire training data (671 individuals). This drop in performance was particularly evident with a smaller number of samples per individual. For instance, a 2.5% drop in performance was observed with 1 sample, whereas with 6 samples, the drop was only 0.8%. As the number of samples per individual in the training data increases, the classification performance increases. Using only one sample per individual yields a 10.7% difference in performance compared to using six samples per individual for the maximum number of individuals (671). A similar pattern can be observed when SVMs are trained with fewer folds, although in this case the difference decreases slightly (e.g., for 67 individuals it is 9.0%). This underlines the importance of a sufficient number of samples per individual in the training data to achieve optimal performance of the model.

Finally, we trained linear SVMs based on different combinations of GRF dimensionality and laterality (Table 5). The use of bilateral instead of unilateral signals resulted in a substantial improvement. For the individual GRF components, GRF_{AP} showed an improvement of 9.2% (right) and 8.8% (left), GRF_{ML} showed the largest improvement of 13.7% (right) and 12.7% (left), and GRF_{V} showed the smallest improvement of 3.9% (right) and 5.1% (left). The combination of all GRF components GRF_{3D} showed an improvement of 1.4% (right) and 1.0% (left). Considering the individual GRF components, regardless of

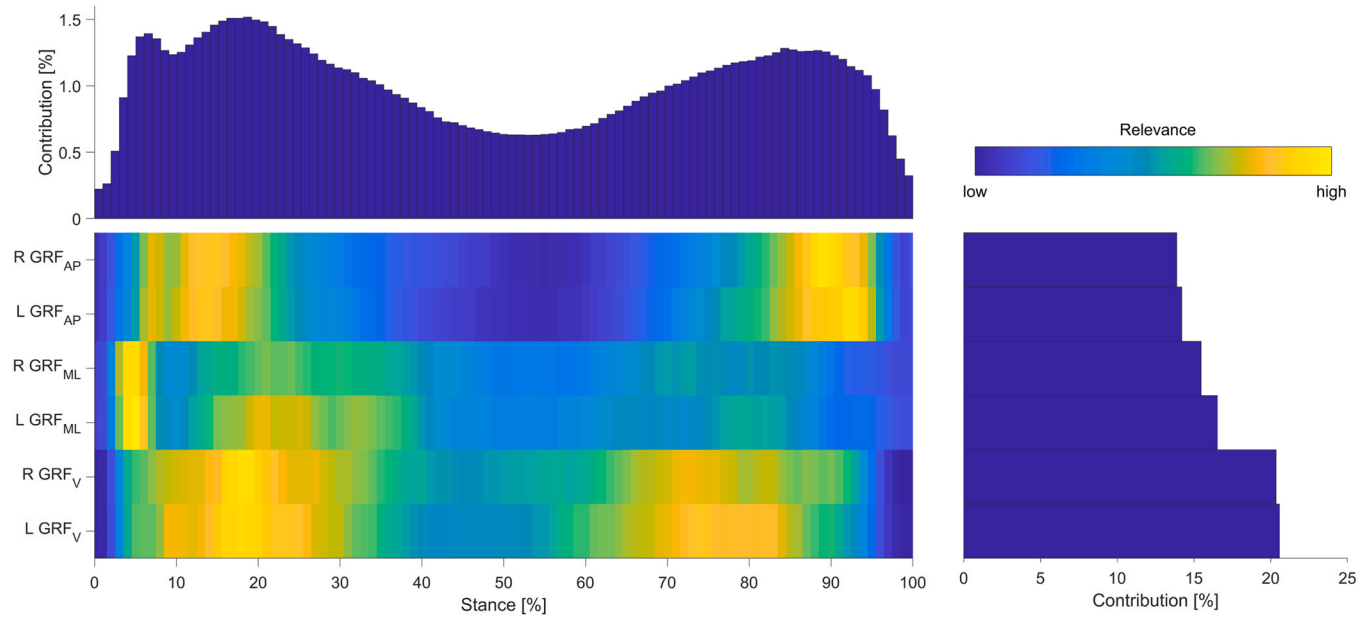

Fig. 2. Model explanation of the SVM trained on bilateral GRF_{3D} data. Input relevance scores obtained by LRP. The top part on the left shows the summed contribution of the relevance scores for each of the 101 time points of the stance phase. In the bottom part on the left, lighter colors indicate variables of high relevance, while darker colors indicate variables of low relevance. The bottom right part highlights the summed contribution of relevance scores of each of the GRFs, namely medio-lateral (GRF_{ML}), anterior-posterior (GRF_{AP}), and vertical (GRF_V).

the leg, GRF_V had the highest discriminative power for classifying gait signatures, with an advantage of 5.2% (right) and 3.6% (left) over the second most informative component, GRF_{AP}. The component with the lowest discriminative power was GRF_{ML} with a performance of 78.0% (right) and 79.0% (left) in the unilateral case. However, in the bilateral case the performance between GRF_{AP}, GRF_{ML}, and GRF_V was nearly identical. Using GRF_{3D} results in a substantial improvement of 10.2% (right), 11.8% (left), and 7.7% (right+left) compared to the strongest single GRF component.

4.2. Explainability results

The explainability results for the linear SVM trained on bilateral GRF_{3D} data are presented in Fig. 2, which shows the aggregated relevance scores for all test samples. The upper left part shows the overall contribution of relevance scores for each of the 101 time points during the (time-normalized) stance phase. In the lower left part, brighter colors indicate high relevance, while darker colors indicate low relevance for the input features. The lower right part highlights the summed relevance scores for each of the GRF dimensions.

The SVM utilized regions identified as relevant by LRP from all bilateral GRF components. These relevant regions are mainly located around the peaks of each GRF component. The highest relevance scores are found in GRF_V, which is consistent with the classification results of GRF_V that showed the highest performance of each individual GRF component (Table 5). Lower relevance scores are observed for GRF_{AP} and GRF_{ML}. However, the latter has a slightly higher relevance compared to GRF_{AP}, which is not consistent with the results in Table 5, where GRF_{AP} seems to have a higher discriminative power.

Figure 3 provides a more detailed examination of the individual LRP explanations at the decision level. The first subfigure shows the average GRF_{3D} signals (along with the standard deviation), with the color coding reflecting the total relevance, also shown in Fig. 3. Each of the following subfigures shows two correctly classified recordings (from the test data) for a particular individual, with the color coding reflecting the relevance values for the decision. The explanations confirm that for each individual relevant regions are found in close proximity to the peaks, but not necessarily in each GRF dimension. Furthermore, the relevant regions for each individual are locally consistent across different recordings, whereas different regions are often relevant between different individuals.

5. Discussion

In this study, we have shown how ML methods together with explainability methods can enrich the assessment of biological individuality. By using the example of human gait, we provide evidence that unique features for each moving individual can be identified in a large-scale and heterogeneous dataset. The identification and characterization of gait signatures may inspire the classification process in computational and structural biology and is a crucial step towards a better understanding of human behavior in the context of research, but also for well-being and healthcare.

Our results demonstrate that ML models can learn gait signatures, providing evidence for unique characteristics of individual gait patterns and distinct differences between individuals on a previously unprecedented scale of nearly 700 individuals. The comparison between the obtained prediction accuracy of 99.3% (bilateral GRF_{3D}) and the zero-rule baseline of 0.0015% (= 1/671) illustrates the significance of the present results (Table 3). Our results confirm the findings of previous studies (Section 2) and demonstrate a high degree of uniqueness of gait signatures even in a large-scale dataset comprising data collected independently across three different gait laboratories in Japan and Central Europe. The outstanding performance of the linear SVM, surpassing other employed ML methods with an accuracy of 99.3% (Table 3), indicates the linear separability of gait signatures and underscores the unique nature of individual gait data.

Our study achieved a comparable performance level to previous studies that utilized a time-continuous (holistic) data representation from bilateral GRF_{3D} as input for the identification of gait signatures (i.e., 99.3% in 671 individuals in our classification experiments compared to 100.0% in 57 individuals [7] or 99.8% in 128 individuals [6]). However, the number of considered individuals affects the

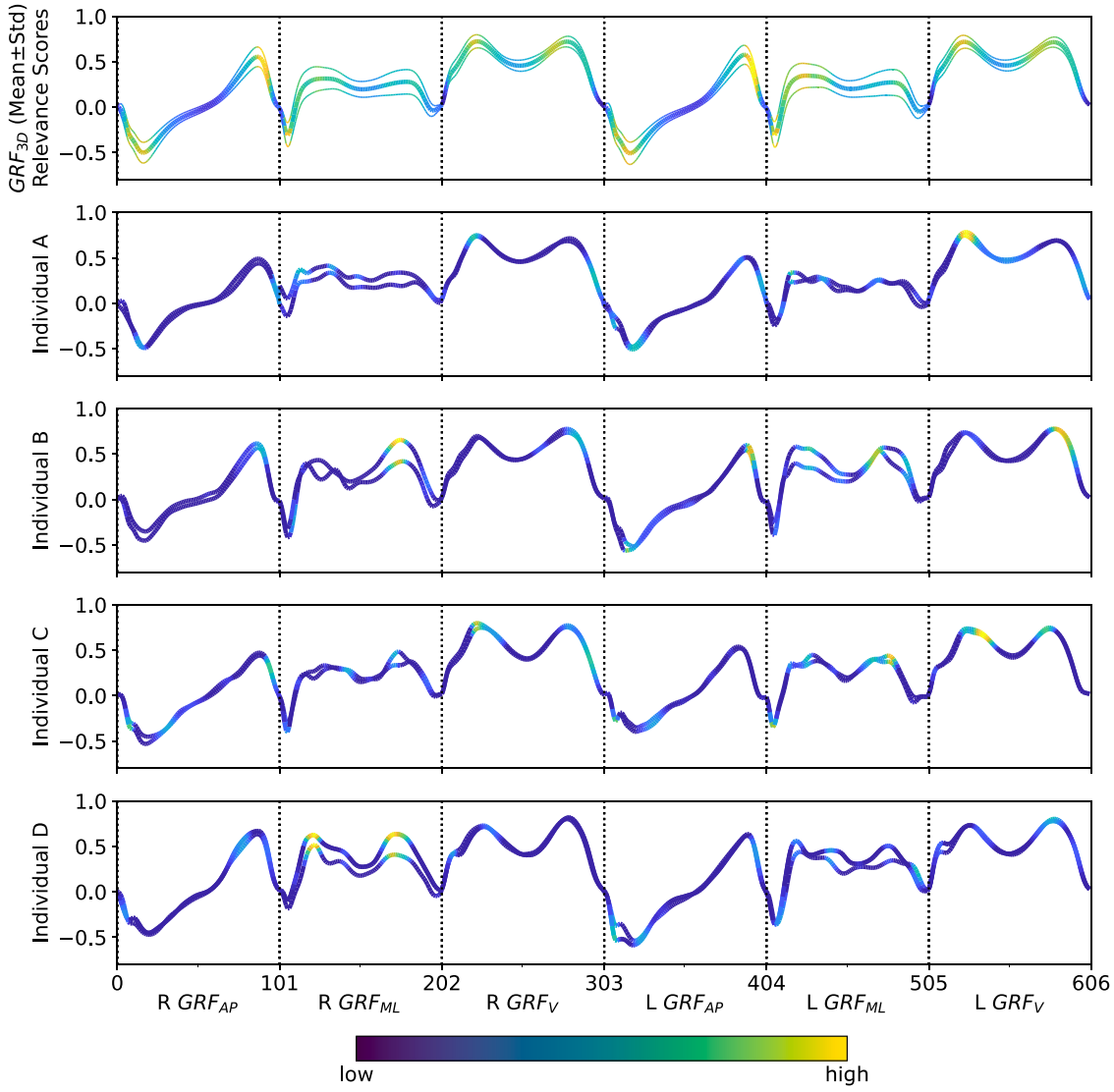

Fig. 3. Decision explanations of the SVM trained on bilateral GRF_{3D} data. Input relevance scores obtained by LRP. The first subfigure displays the mean GRF_{3D} signals (with standard deviation) color-coded according to the total relevance, while each subsequent subfigure shows two recordings for an individual, with color-coded decision relevance scores. The explainability results confirm that gait signatures are unique across individuals and consistent for the same individual.

classification performance in this task. This is indicated by the decrease in prediction accuracy when the number of individuals increases in our stress-test experiment. As shown in Table 4, the increase in the number of individuals is accompanied by a slight decrease in the prediction accuracy, i.e., from 100.0% (67 individuals) to 99.2% (671 individuals), when six training samples per individual are used. This confirms that accurately classifying gait signatures becomes more difficult with a larger number of individuals. Nonetheless, the degree of uniqueness of GRF-based gait signatures appears to be higher than previously expected [12]. Although GRFs are an aggregated variable that provides less differentiated information about body segment and joint movements compared to vision-based approaches, GRF-based gait signatures demonstrate a high level of uniqueness (by exceeding an accuracy of 99%). This indicates that a comparable level of uniqueness to vision-based gait signatures can be achieved [22].

Our results show that all bilateral GRF components (GRF_{AP}, GRF_{ML}, GRF_{V}) contain complementary information that are relevant for the classification and description of gait signatures. This is highlighted by the highest prediction accuracy for GRF_{3D} (Table 5) and the considerable amount (> 10%) of the total relevance for all GRF components (Fig. 2). However, the results also demonstrate that gait signatures can be accurately modeled even when some components of the data are not available, e.g., when only a single GRF dimension is used. For example, the SVM achieved an accuracy of 86.5–87.7% when using only unilateral GRF_{V}, a commonly used component in the literature. Furthermore, the results revealed that even with only unilateral GRF_{3D} data the prediction accuracy was only 1.2–1.7% worse than the best performing configuration with bilateral GRF_{3D}. This is relevant for research and application scenarios where only limited or incomplete data is available. In addition, the results of our stress test suggest that the amount of samples per individual used for training the models is an influencing factor on the classification performance (Table 4). The results demonstrate that for a high prediction accuracy of over 99.0%, at least six training samples per individual were required for a dataset size of 671 healthy individuals. However, even with just a single sample per individual for training, a prediction accuracy of 88.5% was achieved when using an ML model on the dataset of 671 individuals.

The model explanation for visualizing overall relevance in Fig. 2 shows that the SVM uses relevant regions in all three GRF dimensions. By examining the decision explanations in Fig. 3, we can observe that these regions can occur in all GRF components of an individual, but often only appear in a subset of them. These relevant

regions are mainly located around the peaks of each GRF component, where the variability of the data is also the highest. This is reasonable because unique features are likely to be found in these regions due to higher variability [40]. Furthermore, the aggregated explanation for each GRF dimension shows that the vertical GRF component (GRF_V) has the highest discriminatory power, followed by GRF_{AP} and GRF_{ML}. At the decision level, LRP explanations confirm that relevant regions are located in close proximity to the peaks, but also in other local minima and maxima. The decision explanations reveal that the relevance scores for the records of an individual are highly spatially correlated, e.g., for individual A in Fig. 3, highly relevant regions are identified for both records in the first peak in GRF_V. This finding is reinforcing the concept of distinct gait signatures. Another result supporting the uniqueness of gait signatures is that significant regions tend to manifest in different locations for different individuals.

The discovery that individual gait signatures can be detected even in well-standardized experimental conditions, i.e., healthy participants walked barefoot at a self-selected speed, underlines the significance of considering individual characteristics in biological research. Current research strategies usually prioritize the analysis of average behavior or uniform responses to treatments within groups of individuals. This approach, which often relies on estimates of central tendencies like mean values, places less emphasis on studying individual subjects [41,42]. To date, the impact of individual gait signatures on the outcomes of group-based evaluations of treatment effects has not been well documented and merits further research [42]. While demographic and anthropometric characteristics are often considered as potentially confounding variables when researching human gait, the present results highlight the importance of considering individual gait signatures in future research more strongly. Leveraging explainability methods to discover which input features or sections of the input signals contribute most to individual gait signatures can help researchers to tailor their analyses. In particular, LRP enables to analyze sample/class explanations (Fig. 3) that provide information about the specific gait signature of an individual. In addition, model explanation (Fig. 2) can provide general information about gait signatures and inform which signal regions appear to be particularly vulnerable to subject-specific behavior.

The uniqueness of GRF-based gait signatures identified in this study is also of interest for authentication and surveillance scenarios. Like other biometric features (e.g., fingerprints or voice), human gait seems to be a distinctive feature that can be used for identification purposes. The use of GRF data for gait recognition in authentication and surveillance scenarios is still in an early stage [18]. By increasing the number of individuals by a factor of approximately four compared to previous studies (Section 2), our results represent a significant advance in demonstrating the potential of GRF-based gait signatures in serving as a viable alternative or complement to vision-based approaches in some real-world authentication and surveillance scenarios.

In healthcare scenarios, our findings emphasize the importance to consider individual gait signatures in the interpretation of gait data, as advocated by previous studies [6,41]. Utilizing explainability methods to determine which characteristics contribute most to individual gait signatures may support clinicians in finding and tailoring their assessments and interventions to subject-specific needs. Determining the most suitable intervention for a given individual is an ongoing challenge for clinicians. However, further work is needed to develop methodological frameworks for personalized treatments [43]. In this context, caution is warranted when directly inferring individual adaptation processes through interventions solely based on the finding of individual gait signatures [44]. Thus, an interesting research direction would be the exploration of the extent to which the knowledge gained from gait signatures can significantly impact rehabilitation, clinical diagnostics, and personalized treatment approaches.

6. Future research

This study utilized data from healthy individuals who walked barefoot at their self-selected speed during a single session in a well-standardized laboratory setting. The aim was to capture natural gait patterns while minimizing the influence of potentially confounding variables such as footwear, multiple testing sessions, and externally imposed speed constraints. This approach allows for a baseline assessment of the uniqueness of gait signatures within a typical laboratory setting used in biomechanical research. However, it may not capture the complexity and variations present in different populations. To enhance the applicability and generalizability of gait signatures, future studies should explore the uniqueness of individual gait signatures in diverse populations, including individuals with injuries, orthopaedic and neurological disorders, who often exhibit higher (intra-subject) variability in their gait patterns. From an ML perspective, the utilization of zero- and few-shot learning approaches [45,46] could be promising directions for investigating uniqueness of gait signatures in open-set tasks.

Another important direction for future research is to explore the persistence of individual gait signatures. Limited studies have been conducted on how gait signatures are influenced by various factors such as time, circadian rhythms, menstrual cycle, fatigue, footwear, gait speed, mood, and injury [18]. Longitudinal data collections should be incorporated, capturing gait patterns over multiple sessions and under different conditions. This would provide insights into the temporal dynamics and contextual variations of gait signatures. Understanding the impact of these confounding variables is crucial for developing reliable applications for real-world scenarios.

Expanding the scope of the study to encompass more realistic settings and diverse populations will contribute to a more comprehensive understanding of gait signatures and their practical applications in various research and real-world scenarios.

7. Conclusion

The present study examined the individuality of GRF-based gait signatures on a large scale. The utilized dataset comprises GRF data from nearly 700 healthy individuals of different sex, age, physical constitution, and was collected in three gait laboratories located in Japan and Central Europe. Our work demonstrated that modelling the biological individuality of organisms using ML offers great potential for advancing research in biology. By exploring not only extraordinary characteristics of biological structures (e.g., fingerprint and face) but also identifying unique behavior of individual organisms (e.g., gait), researchers can gain additional insights into the functioning of biological systems, their interaction with the environment, and their ability to adapt over time [3,47]. One way to study the biological individuality of organisms in more detail is through the use of ML driven classification approaches in combination with explainability methods. The latter can identify the most relevant characteristics used for classification. For example, explainability analysis can inform about age- [48] and sex- [49] related behavioral characteristics in human gait. In our study, we demonstrated how an explainability methods can be utilized to identify gait signatures. According to the model explanations, the trained SVM models used input features in all three GRF components that correspond to local minima and maxima in order to achieve a prediction accuracy of 99.3%. The intriguing aspect of this observation lies in the fact that, within the field of biomechanics, the primary emphasis is placed on GRF_V and GRF_{AP}, while GRF_{ML} receives comparatively little attention. In addition, the LRP explanations at the sample level showed a high degree of agreement between the regions used

to classify an individual. This provides further evidence for the concept of a unique gait signature. The proposed ML-based approach employed to classify individuals based on their distinctive gait patterns and to identify gait signatures can serve as a template for demonstrating how an automated classification of behavior can effectively complement existing classification systems for diverse taxa [50,51].

CRediT authorship contribution statement

Fabian Horst: Conceptualization, Formal analysis, Visualization, Data curation, Software, Writing - original draft. **Djordje Slijepcevic:** Conceptualization, Formal analysis, Data curation, Visualization, Software, Writing - original draft. **Marvin Simak:** Data curation, Writing – original draft. **Brian Horsak:** Conceptualization, Writing – review & editing, Funding acquisition. **Wolfgang Immanuel Schöllhorn:** Conceptualization, Writing – review & editing. **Matthias Zeppelzauer:** Conceptualization, Writing – review & editing, Funding acquisition.

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

This work was partly funded by the Austrian Research Promotion Agency of Lower Austria (GFF, Gesellschaft für Forschungsförderung NÖ) within the Endowed Professorship for Applied Biomechanics and Rehabilitation Research (SP19-004) and the IntelliGait3D project (FTI17-014).

References

[1] Borba VH, Martin C, Machado-Silva JR, Xavier SC, de Mello FL, Iñiguez AM. Machine learning approach to support taxonomic species discrimination based on helminth collections data. Parasit Vectors 2021;14:230. https://doi.org/10.1186/s13071-021-04721-6
[2] Badirli S, Picard CJ, Mohler G, Richert F, Akata Z, Dundar M. Classifying the unknown: insect identification with deep hierarchical bayesian learning. Methods Ecol Evol 2023;14:1515–30. https://doi.org/10.1111/2041-210X.14104
[3] Baedke J. What Is a Biological Individual? Cham: Springer International Publishing; 2019. p. 269–84. https://doi.org/10.1007/978-3-030-18202-1_13
[4] Cutting JE, Kozlowski LT. Recognizing friends by their walk: gait perception without familiarity cues. Bull Psychon Soc 1977;9:353–6. https://doi.org/10.3758/BF03337021
[5] Stevenage SV, Nixon MS, Vince K. Visual analysis of gait as a cue to identity. Appl Cogn Psychol 1999;13:513–26. https://doi.org/10.1002/(SICI)1099-0720(199912)13:6<513::AID-ACP616>3.0.CO;2-8
[6] Horst F, Mildner M, Schöllhorn WI. One-year persistence of individual gait patterns identified in a follow-up study - a call for individualised diagnose and therapy. -10 Gait Posture 2017;58:476–80. https://doi.org/10.1016/j.gaitpost.2017.09.003
[7] Horst F, Lapuschkin S, Samek W, Müller K-R, Schöllhorn WI. Explaining the unique nature of individual gait patterns with deep learning. Sci Rep 2019;9:2391. https://doi.org/10.1038/s41598-019-38748-8
[8] Hug F, Vogel C, Tucker K, Dorel S, Deschamps T, Le Carpentier É, et al. Individuals have unique muscle activation signatures as revealed during gait and pedaling. J Appl Physiol 2019;127:1165–74. https://doi.org/10.1152/japplphysiol.01101.2018
[9] Horst F, Kramer F, Schäfer B, Eekhoff A, Hegen P, Nigg BM, et al. Daily changes of individual gait patterns identified by means of support vector machines. Gait Posture 2016;49:309–14. https://doi.org/10.1016/j.gaitpost.2016.07.073
[10] Janssen D, Schöllhorn WI, Lubienetzki J, Fölling K, Kokenge H, Davids K. Recognition of emotions in gait patterns by means of artificial neural nets. J Nonverbal Behav 2008;32:79–92. https://doi.org/10.1007/s10919-007-0045-3
[11] Janssen D, Schöllhorn WI, Newell KM, Jäger JM, Rost F, Vehof K. Diagnosing fatigue in gait patterns by support vector machines and self-organizing maps. Hum Mov Sci 2011;30:966–75. https://doi.org/10.1016/j.humov.2010.08.010
[12] Jain AK, Bolle R, Pankanti S. Biometrics: Personal identification in networked society (Eds.). New York, NY: Springer; 2006.
[13] Gandevia SC, Muceli S, Héroux M. Signing up to motor signatures: a unique link to action. J Appl Physiol 2019;127:1163–4. https://doi.org/10.1152/japplphysiol.00643.2019
[14] Aeles J, Horst F, Lapuschkin S, Lacourpaille L, Hug F. Revealing the unique features of each individual's muscle activation signatures. J R Soc Interface 2021;18:20200770. https://doi.org/10.1098/rsif.2020.0770
[15] Kobayashi Y, Hida N, Nakajima K, Fujimoto M, Mochimaru M. AIST Gait Database 2019, ⟨https://unit.aist.go.jp/harc/ExPART/GDB2019_e.html⟩, 2019. Online; accessed 12 April 2022.
[16] Horsak B, Slijepcevic D, Raberger A-M, Schwab C, Worisch M, Zeppelzauer M. GaitRec, a large-scale ground reaction force dataset of healthy and impaired gait. Sci Data 2020;7:143. https://doi.org/10.1038/s41597-020-0481-z
[17] Horst F, Slijepcevic D, Simak M, Schöllhorn WI. Gutenberg gait database, a ground reaction force database of level overground walking in healthy individuals. Sci Data 2021;8:232. https://doi.org/10.1038/s41597-021-01014-6
[18] Connor P, Ross A. Biometric recognition by gait: a survey of modalities and features. Comput Vis Image Underst 2018;167:1–27. https://doi.org/10.1016/j.cviu.2018.01.007
[19] G. Veres L. Gordon J. Carter M. Nixon What image information is important in silhouette-based gait recognition? Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2004 doi: 10.1109/CVPR.2004.1315243.
[20] Iwama H, Okumura M, Makihara Y, Yagi Y. The ou-isir gait database comprising the large population dataset and performance evaluation of gait recognition. IEEE Trans Inf Forensics Secur 2012;7:1511–21. https://doi.org/10.1109/TIFS.2012.2204253
[21] Takemura N, Makihara Y, Muramatsu D, Echigo T, Yagi Y. Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition. IPSJ Trans Comput Vis Appl 2018;10:4. https://doi.org/10.1186/s41074-018-0039-6
[22] Sepas-Moghaddam A, Etemad A. Deep gait recognition: a survey. IEEE Trans Pattern Anal Mach Intell PP 2022. https://doi.org/10.1109/TPAMI.2022.3151865
[23] Baker R. Measuring Walking: A Handbook of Clinical Gait Analysis. London: Mac Keith Press; 2013.
[24] Winter DA. Biomechanics and Motor Control of Human Movement. Hoboken, NJ: Wiley; 2009.
[25] Orr RJ, Abowd GD. The smart floor: a mechanism for natural user identification and tracking. CHI '00 Extended Abstracts on Human Factors in Computing Systems ACM; 2000. p. 275–6. https://doi.org/10.1145/633292.633453
[26] J. Suutala J. Roning. Towards the adaptive identification of walkers: Automated feature selection of footsteps using distinction sensitive LVQ 2004), 2004. PSIPS1415.
[27] Moustakidis SP, Theocharis JB, Giakas G. Subject recognition based on ground reaction force measurements of gait signals. IEEE transactions on systems, man, and cybernetics. Part B, Cyber: a Publ IEEE Syst, Man, Cyber Soc 2008;38:1476–85. https://doi.org/10.1109/TSMCB.2008.927722
[28] Moustakidis SP, Theocharis JB, Giakas G. Feature extraction based on a fuzzy complementary criterion for gait recognition using GRF signals. 17th Mediterranean Conference on Control and Automation. IEEE; 2009. p. 1456–61. https://doi.org/10.1109/MED.2009.5164752
[29] Addlesee M, Jones A, Livesey F, Samaria F. The ORL active floor. IEEE Pers Commun 1997;4:35–41. https://doi.org/10.1109/98.626980
[30] Vera-Rodriguez R, Evans N, Lewis R, Fauve B, Mason J. An experimental study on the feasibility of footsteps as a biometric. European Signal Process Conference (EUSIPCO) 2007:748–52.
[31] Derlatka M. Modified kNN algorithm for improved recognition accuracy of biometrics system based on gait. In: Saeed K, Chaki R, Cortesi A, Wierzchon S, editors. Computer Information Systems and Industrial Management, volume 8104 of Lecture Notes in Computer Science Berlin Heidelberg: Springer; 2013. p. 59–66. https://doi.org/10.1007/978-3-642-40925-7_6
[32] Derlatka M, Bogdan M. Ensemble kNN classifiers for human gait recognition based on ground reaction forces. 8th International Conference on Human System Interaction (HSI). IEEE; 2015. p. 88–93. https://doi.org/10.1109/HSI.2015.7170648
[33] Marcin D. Human gait recognition based on ground reaction forces in case of sport shoes and high heels. 2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA). IEEE; 2017. p. 247–52. https://doi.org/10.1109/INISTA.2017.8001165
[34] Duncanson K, Thwaites S, Booth D, Abbasnejad E, Robertson W, Thewlis D. The most discriminant components of force platform data for gait based person re-identification, 2021. doi: 10.36227/techrxiv.16683229.v1.
[35] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97. https://doi.org/10.1007/BF00994018
[36] Fan RE, Chang K-W, Hsieh C-J, Wang X-R, Lin C-J. Liblinear: a library for large linear classification. J Mach Learn Res 2008;9:1871–4.
[37] Bach S, Binder A, Montavon G, Klauschen F, Müller K-R, Samek W. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PloS One 2015;10:e0130140. https://doi.org/10.1371/journal.pone.0130140
[38] Slijepcevic D, Horst F, Lapuschkin S, Horsak B, Raberger A-M, Kranzl A, et al. Explaining machine learning models for clinical gait analysis. ACM Trans Comput Healthc (HEALTH) 2021;3:1–27. https://doi.org/10.1145/3474121
[39] Hoitz F, vonTscharner V, Baltich J, Nigg BM. Individuality decoded by running patterns: movement characteristics that determine the uniqueness of human running. PloS One 2021;16:e0249657. https://doi.org/10.1371/journal.pone.0249657
[40] Giakas G, Baltzopoulos V. Time and frequency domain analysis of ground reaction forces during walking: an investigation of variability and symmetry. Gait Posture 1997;5:189–97. https://doi.org/10.1016/S0966-6362(96)01083-1

[41] Schöllhorn WI, Nigg BM, Stefanyshyn D, Liu W. Identification of individual walking patterns using time discrete and time continuous data sets. Gait Posture 2002;15:180–6. https://doi.org/10.1016/S0966-6362(01)00193-X
[42] Fisher AJ, Medaglia JD, Jeronimus BF. Lack of group-to-individual generalizability is a threat to human subjects research. Proc Natl Acad Sci 2018;115:E6106–15. https://doi.org/10.1073/pnas.1711978115
[43] Schöllhorn WI, Rizzi N, Slapšinskaitė-Dackevičienė A, Leite N. Always pay attention to which model of motor learning you are using. Int J Environ Res Public Health 2022;19:711. https://doi.org/10.3390/ijerph19020711
[44] Horst F, Janssen D, Beckmann H, Schöllhorn WI. Can individual movement characteristics across different throwing disciplines be identified in high-performance decathletes? Front Psychol 2020;11:2262. https://doi.org/10.3389/fpsyg.2020.02262
[45] Krondorfer P, Slijepčević D, Unglaube F, Kranzl A, Breiteneder C, Zeppelzauer M, et al. Deep learning-based similarity retrieval in clinical 3d gait analysis. Gait Posture 2021;90:127–8. https://doi.org/10.1016/j.gaitpost.2021.09.066
[46] Duncanson KA, Thwaites S, Booth D, Hanly G, Robertson WSP, Abbasnejad E, et al. Deep metric learning for scalable gait-based person re-identification using force platform data. Sensors 2023;23:3392. https://doi.org/10.3390/s23073392
[47] Montévil M. Measurement in biology is methodized by theory. Biol Philos 2019;34:35. https://doi.org/10.1007/s10539-019-9687-x
[48] Slijepcevic D, Horst F, Simak M, Lapuschkin S, Raberger A-M, Samek W, et al. Explaining machine learning models for age classification in human gait analysis. Gait Posture 2022;97:252–3. https://doi.org/10.1016/j.gaitpost.2022.07.153
[49] Horst F, Slijepcevic D, Zeppelzauer M, Raberger A-M, Lapuschkin S, Samek W, et al. Explaining automated gender classification of human gait. Gait Posture 2020;81:159–60. https://doi.org/10.1016/j.gaitpost.2020.07.114
[50] Gibb AC, Amplo H, Struble M, Kawano SM. A step forward: functional diversity and emerging themes of slow-speed locomotion in vertebrates. Integr Comp Biol 2022;62:1235–45. https://doi.org/10.1093/icb/icac139
[51] Struble MK, Gibb AC. Do we all walk the walk? A comparison of walking behaviors across tetrapods. Integr Comp Biol 2022;62:1246–80. https://doi.org/10.1093/icb/icac125