



OPEN

DATA DESCRIPTOR

Health & Gait: a dataset for gait-based analysis

Jorge Zafra-Palma^{1,2}✉, Nuria Marín-Jiménez^{3,4}, José Castro-Piñero^{3,4},
Magdalena Cuenca-García^{3,4}, Rafael Muñoz-Salinas^{1,2} & Manuel J. Marín-Jiménez^{1,2}✉

Acquiring gait metrics and anthropometric data is crucial for evaluating an individual's physical status. Automating this assessment process alleviates the burden on healthcare professionals and accelerates patient monitoring. Current automation techniques depend on specific, expensive systems such as OptoGait or MuscleLAB, which necessitate training and physical space. A more accessible alternative could be artificial vision systems that are operable via mobile devices. This article introduces Health&Gait, the first dataset for video-based gait analysis, comprising 398 participants and 1, 564 videos. The dataset provides information such as the participant's silhouette, semantic segmentation, optical flow, and human pose. Furthermore, each participant's data includes their sex, anthropometric measurements like height and weight, and gait parameters such as *step* or *stride length* and *gait speed*. The technical evaluation demonstrates the utility of the information extracted from the videos and the gait parameters in tackling tasks like sex classification and regression of weight and age. Health&Gait facilitates the progression of artificial vision algorithms for automated gait analysis.

Background & Summary

Gait capacity is an inherent ability of the human being, reflecting the primary means of mobility and daily functioning. Thus, many health conditions may impact gait patterns. The human gait comprises distinct gait cycles. A gait cycle begins when the heel of one foot makes contact with the ground and ends when the same foot, after taking a step, sets your foot on the ground. The gait cycle has two phases: Stance and Swing. Every cycle begins with a Stance phase, where the supporting foot goes from heel to toe, and continues with a Swing phase, where the foot is lifted and proceeds through the air. The Stance phase makes up 60% of the time, and the Swing phase makes up 40% of the time¹. A representation of the gait cycle is shown in Fig. 1 for better understanding.

Gait analysis is a systematic way of identifying variations in the gait pattern and finding the associated reasons and how they can affect the different human systems, mainly implying the locomotor system². The gait patterns are used for early disease diagnosis, monitoring and improvement of sports performance, and rehabilitation. Gait speed is a clinically significant and reliable metric for assessing functional capacity in a wide range of disease populations and is sometimes referred to as the sixth vital sign^{3–5}. For a complete understanding of the gait pattern, different spatiotemporal gait parameters must be taken into account, such as *step length*, *stride length* or *gait speed*⁶. Among them, *gait speed* is a key indicator of health and function in ageing and disease⁷.

In this sense, altered gait patterns have been observed in populations with neurodegenerative diseases, such as people with Parkinson's Disease⁸, Multiple Sclerosis⁹, or cognitive decline¹⁰. Likewise, alterations in gait patterns, such as gait asymmetry, are also documented in people with obesity¹¹. Therefore, human gait analysis is essential for health prevention and monitoring.

Gait performance methodologies include specific kinematic analysis through devices equipped with optical sensors, such as OptoGait system¹². The main limitations of these types of systems are high cost, space constraints, and the need to familiarise yourself with the software or wiring necessary for its correct installation and operation (requiring greater assembly and maintenance time). Consequently, alternatives that are simpler to set up, affordable, and with the same validity and reliability of measurement are necessary for this field. Thereby, methodologies for recognising an individual (i.e. biometric recognition) by their gait pattern, including the development of artificial vision systems that allow for the analysis of gait using low-cost devices such as a mobile phone are emerging^{13–15}.

¹University of Cordoba, Department of Computing and Numerical Analysis, Córdoba, 14071, Spain. ²Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC), Córdoba, 14004, Spain. ³GALENO Research Group, Department of Physical Education, Faculty of Education Sciences, University of Cadiz, Cádiz, Spain. ⁴Instituto de Investigación e Innovación Biomédica de Cádiz (INIBICA), Cádiz, Spain. ✉e-mail: jzafra@uco.es; mjmarin@uco.es

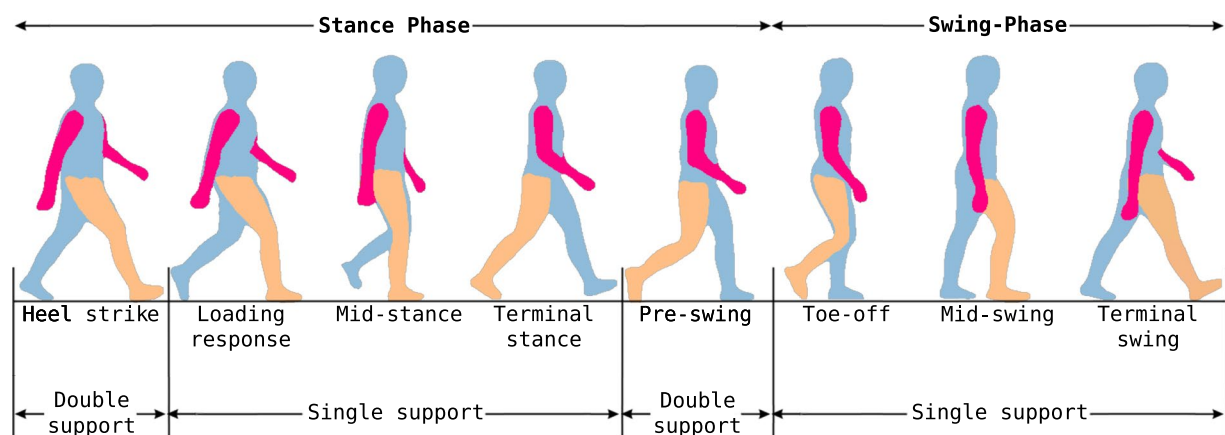


Fig. 1 Gait cycle. The leg considered for measuring the gait cycle (right) is highlighted, and the arms are highlighted to make visualisation easier. The different stages have been obtained from the semantic segmentation of one of the participants.

To date, there are numerous datasets focused on gait analysis. However, most of them use inertial sensors for feature extraction. For instance, the DUO-GAIT¹⁶ dataset contains gait information from the 6-minute walk test performing only the task of walking or combined with a cognitive task consisting of continuously subtracting seven from a random 4-digit number. This dual task seeks to assess how the simultaneous performance of cognitive activity may influence gait patterns. The number of participants is 16, healthy young adults wearing nine IMUs placed on different parts of their bodies. The same applies to the Diverse Gait¹⁷ dataset, which also uses inertial sensors for data collection, but in this case, they are placed only on the participants' feet. The dataset consists of 4,690 walking strides data and 19,093 gait labels. The dataset includes a total of 22 healthy participants (13 males and 9 females) with an average age of 32.5 ± 7.5 years. The authors use the dataset to develop an algorithm for detecting the different phases of the gait. The main drawback of these systems is that, despite being non-invasive from a clinical perspective, data collection may still involve physical contact or attachments that can be uncomfortable or even alter the usual gait pattern of participants. On the other hand, GAIT-IT¹⁸ is a dataset composed of 828 gait video sequences obtained from 21 participants, aimed at classifying five different types of gait pathologies. The authors develop a lightweight Convolutional Neural Network (CNN) architecture specifically designed to classify gait types, significantly reducing the number of trainable parameters, memory requirements and execution times while maintaining a classification accuracy comparable to the state-of-the-art. It also integrates with a remote web application, allowing users to upload videos and receive the gait-type classification, which is useful for healthcare professionals.

Our dataset, Health&Gait, consists of 1,564 videos of 398 participants walking in a controlled closed environment, where each video has associated information such as silhouette, semantic segmentation, optical flow, and participant's pose. In addition, each participant has associated anthropometric information such as weight, height, body mass index (BMI), etc.; gait parameters obtained from OptoGait and MuscleLAB; and gait parameters estimated from the pose information by Computer Vision. To the best of our knowledge, no previous datasets collect these anthropometric and kinematic characteristics. As demonstrated in the technical evaluation, our dataset is useful for inferring the sex, weight, and age of the subject, using solely the gait information extracted from the videos or combining it with anthropometric and kinematic characteristics provided by other sensors. It is an ideal dataset to use as the starting point for developing more complex gait analysis algorithms. Therefore, it could be implemented as a simple and quick clinical screening technique for health prevention.

Methods

Experimental design. The proposed dataset, termed Health&Gait, has been recorded over eight months, from 2020-10 to 2021-05. A total of 398 participants were recorded walking indoors, with 1,564 video sequences. All participants belong to the ADULT-FIT Project (DEP2017-88043-R) sample and were recruited from the cities of Cádiz Bay, Spain, through leaflets, local newspapers, and social media. The inclusion criteria for this study were: (i) age: adults (18-64 years old); (ii) not having a physical or mental illness that prevents you from doing physical activity; (iii) intention to carry out all the tests that make up the study; and, (iv) able to read and understand the informed consent as well as the object of the study. The exclusion criteria for this study were: (i) acute or terminal illness; (ii) myocardial infarction three months before starting the study; (iii) unstable cardiovascular disease; (iv) medical prescription that prevents the performance of the test; and, (v) injury or circumstance that makes it impossible to carry out the test correctly. They were informed of the protocol and signed an informed consent. In particular, the participants in the study, once informed, agreed in the written consent form to share their data in an anonymized form that does not allow for their identification. This consent was obtained after ensuring that all individuals were fully aware of the extent and nature of the data being collected, as well as the measures taken to protect their privacy and anonymity throughout the research process. The main goal of this project was to develop a field-based physical fitness test battery related to health based on their criterion-validity,

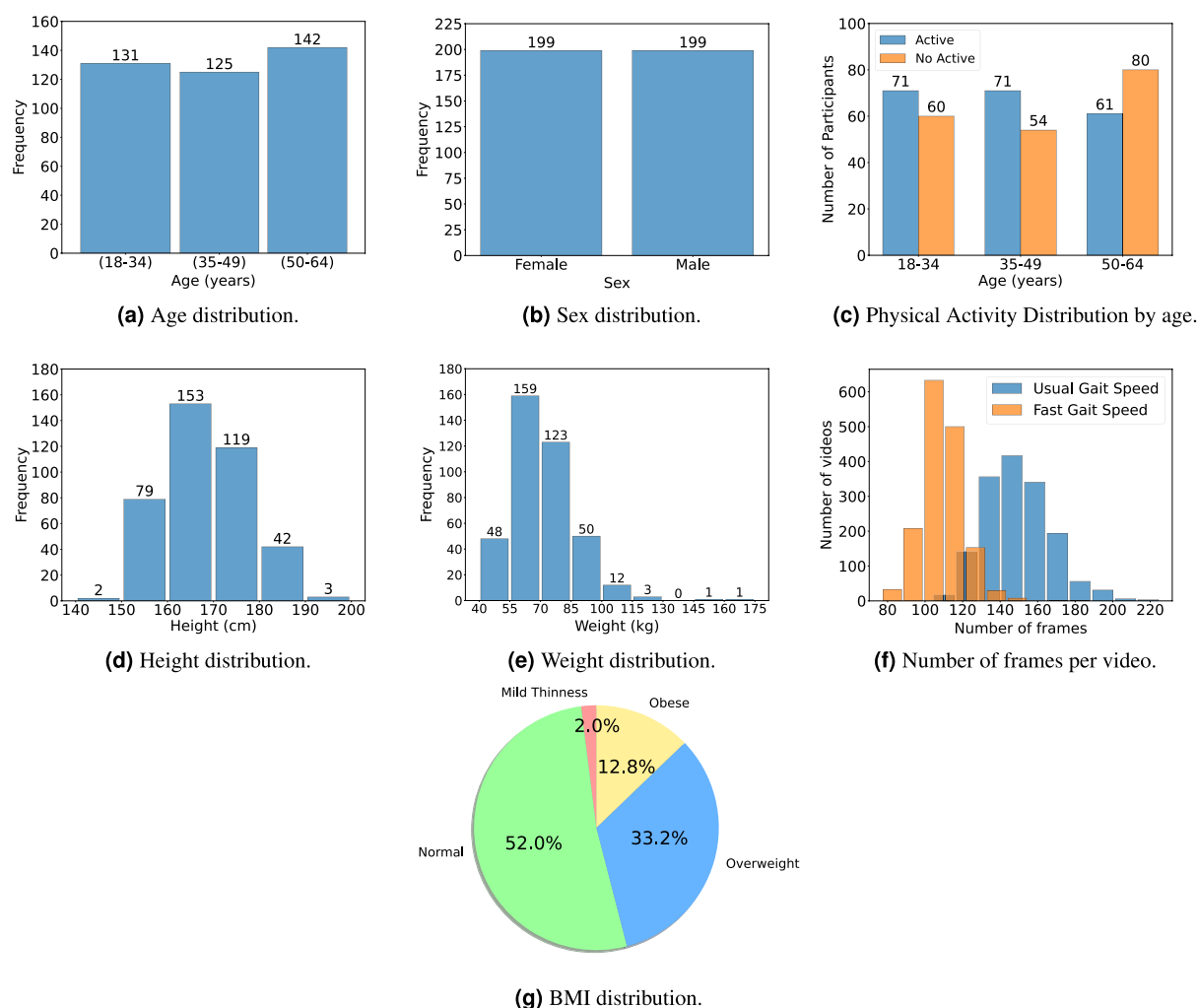


Fig. 2 General dataset statistics (Best viewed in digital format).

predictive validity, reliability, feasibility, safety and responsiveness for use in adults. The project was approved by the Review Committee for Research of Cadiz, Spain, and The Declaration of Helsinki was rigorously followed.

The participants were instructed to walk along a 10-metre corridor in a controlled environment (laboratory setting) with the following instructions:

i) “Walk along this corridor, from this cone to the cone in front of you, at a normal pace, as you usually walk. Do it back and forth”; ii) “Repeat the same path as fast as you can, without running”.

Additionally, whenever possible, the participant was asked to perform these actions with and without wearing their jacket to add more variability to the data, being valuable, for example, to verify if this alters gait parameters or complicates the inference of information from the data. Therefore, the four scenarios present in the dataset were as follows:

- Usual gait speed (‘UGS’) without jacket (WoJ).
- Fast gait speed (‘FGS’) without a jacket.
- Usual gait speed with jacket (‘WJ’).
- Fast gait speed with jacket.

Age, sex, and physical activity levels (active or non-active) were balanced, as can be seen in Fig. 2a-c. Figure 2d,e show the distributions of height and weight, which resemble a normal distribution. The BMI distribution of the participants is shown in Fig. 2g. Most of the participants have a normal or overweight BMI, with a minority of those showing obesity or mid-thinness. Finally, Fig. 2f shows the distribution of the number of frames of the recorded videos.

For each participant, we collected their age and sex, anthropometric measurements, and gait parameters. The anthropometric measurements are height, weight, body mass index, hip, waist and neck circumferences, percentage of body fat mass, and lean mass. The gait parameters measured are *step length*, *stride length*, *cadence*,

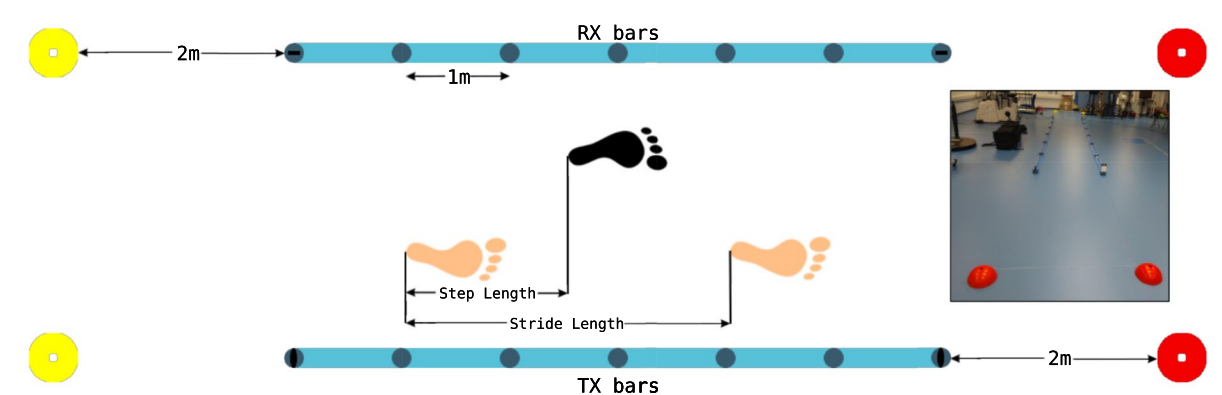


Fig. 3 Recording setup and step length, stride length visualisation.

single-limb support, *bipedal support*, and *gait speed*. Moreover, each participant was asked the number of days they exercised.

Anthropometric measurements were always obtained by the same trained evaluator (to avoid intra-evaluator variability), who was of the same sex as the participant. All measurements were collected barefoot, in light sports clothing, and with a 3-hour fast.

Finally, for each video, information about the participants' silhouette, semantic segmentation, optical flow, and pose is estimated using Computer Vision techniques. Using the pose information, an estimation of the gait parameters, *step length*, *stride length*, *cadence*, and *gait speed* is done for subsequent comparisons with the parameters obtained directly from the OptoGait sensors. For the neck circumference, the protocol established by the Center for Disease Control and Prevention was followed¹⁹.

Data acquisition. The dataset was recorded at the University of Cádiz, under the ADULT-FIT project, by a research group of specialists in Sports Sciences and physical fitness assessment.

A CASIO EX-ZR1000 camera was used to record the videos at a resolution of $1,920 \times 1,080$ pixels with a frame rate of 30Hz, using the H.264 compression format.

Anthropometry. Height was measured using a TANITA HR001 portable height rod (Tanita®, Illinois, USA; sensitivity, 1 mm). Weight was measured using an OMRON BF-400 electronic scale (Omron Healthcare Europe BV, Hoofddorp, The Netherlands; sensitivity, 100 g). Hip, waist and neck circumferences were assessed using the tape measure using SECA 201 (Seca Int, Hamburg, Germany; range, 0–205 cm; sensitivity, 0.1 cm). Height, weight, hip and waist circumferences were measured using the protocol described by the International Society for the Advancement of Kinanthropometry (ISAK)²⁰.

BMI was calculated as weight (kg) divided by squared height (m^2). The percentage of body fat mass (%BF) and lean mass (kg) were determined by bioimpedance Tanita MC 780-P MA (Tanita Co., Guangzhou, China), according to the protocol described by the National Institute of Health²¹. For its correct evaluation, the participants were asked about their level of hydration²².

Walking test. The OptoGait system and the Optical Timing Gates (Photo-cells) from MuscleLAB were used to obtain the gait parameters and the participants' speed.

OptoGait is an advanced optical detection system that operates using two optical bars: one that emits and another that receives. Our system setup consists of 12 bars, six transmitters and six receivers forming a 6-metre-long corridor, as each bar is one metre long. Each one contains 96 LEDs communicating on an infrared (visible) frequency with the same number of LEDs on the opposite bar. The system is designed to be placed on the ground or on a treadmill, where it can identify interruptions in LED communication caused by a person's movements. These interruptions allow the system to calculate the duration and position. During the execution of a running, gait or series of jumps test, the contact and flight times can be measured with an accuracy of one-thousandth of a second and the position of the interrupted LEDs with a space resolution of 1.041 cm. This raw data is then processed in real-time by specialised software, which derives movement analysis metrics. The original system includes a pair of cameras discarded due to their lack of quality.

The Optical Timing Gates (Photo-cells) from MuscleLAB are wireless and designed for timing and assessing performance during sprint and agility tests. The system consists of two photocells and their corresponding receivers. Although it is a wireless system, the synchronisation of both photocells has been carried out using a data synchronisation unit (DSU) from MuscleLAB. The DSU synchronises all connected devices, wired or wireless, to ± 1 ms. This system is placed at the beginning and end of the corridor delimited by the OptoGait to measure walking speeds over six metres.

The recording setup is shown in Fig. 3. A pair of cones are placed two metres from the OptoGait at the beginning and end to mark an acceleration area before entering the OptoGait. The Optical Timing Gates (Photo-cells) from MuscleLAB are placed at the beginning and end of the OptoGait, right above the first and last sensors.



Fig. 4 Two examples of the different data types from the dataset for two participants (a) and (b).

Data processing. The original videos were scaled to a resolution of 960×540 pixels to ease the training with deep learning techniques and to reduce the total dataset size. Sound was also removed. Then, as each video contained two walking directions, they were split into two video clips, one per direction.

The processing of the clips to obtain the different types of data is detailed below. Some examples of these are shown in Fig. 4, where each row corresponds to a participant and the different modalities have been extracted at the same instant of time.

Optical flow computation. We obtained the optical flow using two methods: Dual-TVL1²³ (traditional) and GMFlow²⁴ (learning-based), to compare which optical flow representation achieves better results. We store the optical flow as an RGB image with a resolution of 480×270 pixels to reduce file size.

For the Dual-TVL1 method, the HSV (Hue, Saturation, Value) colour model is used, where the Hue channel represents angle values bounded between 0 and 179, and the Value channel represents the magnitude of the optical flow between 0 and 255.

On the other hand, GMFlow utilises the Middlebury colour code²⁵ for optical flow representation. In the case of GMFlow, a modification has been made to the provided code such that lighter colours represent regions with higher optical flow magnitude, while regions with optical flow magnitude close to zero are associated with the colour black. This modification aims to facilitate the association of the value zero with the absence of optical flow for learning-based methods.

Human pose estimation. For human pose estimation, Alpha-Pose²⁶ was used, employing its FastPose model with a ResNet50 backbone, YOLOv3 detector, and an input size of 256×192 pixels. The model was pre-trained on the COCO keypoint train 2017 dataset. The model estimates a total of 17 keypoints for each human pose.

Body silhouette extraction. To obtain the silhouettes, YOLOv8²⁷ was used, specifically the pre-trained YOLOv8x-seg model on the COCO dataset, with an input size of 640×640 pixels.

Body semantic segmentation. For the semantic estimation of the person, a chart-based DensePose Estimation method²⁸ is used to find the correspondence of each pixel in the image with a 3D object mesh. The 3D object is divided into charts, and each chart is associated with a body part. Specifically, the human body is divided into 24 parts.

Our dataset uses 14 body parts since we do not differentiate between the front and back, as in²⁹. The result is stored as an RGB image with a resolution of 960×540 pixels, where a different colour represents each part. The correspondence between each part and its RGB colour is indicated in Table 1.

Gait parameters estimation. The following paragraphs detail the procedure for estimating gait parameters using only the visual information obtained from the camera to approximate the gait parameters obtained with OptoGait. It should be noted that when estimating gait parameters, a distinction is made between UGS and FGS cases. However, no distinction is made between WoJ and WJ classes, nor in the direction of movement.

To ensure that the measurements are equivalent to those obtained by OptoGait, the working area is restricted by the first and last OptoGait sensors in the hallway so that the sequence starts and ends when the person crosses it. While the OptoGait sensors were in the same place during all the recording sessions, the position of the camera changed because it was on a tripod, which was manually placed and removed in every session. Thus, Segment Anything Model (SAM)³⁰ has been used to obtain the bounding boxes of the OptoGait sensors. As the variation of the position of the sensors between the different recordings is not very high, in order to facilitate the segmentation task, SAM is provided with information of the bounding boxes where these sensors are expected to be.

The start and end of the measurements are determined by the instant at which the participant passes in front of the sensors. To determine this, semantic segmentation has been used. Although silhouette and pose could have been potentially used, to achieve more precision in the measurement, semantic segmentation was chosen to determine exactly when the top of the foot entered the measurement zone. With the binary silhouette, we do














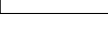
Part of the Body	Colour	Hex Code
Torso		#FF0000
Right Hand		#0000FF
Left Hand		#00FF00
Right Foot		#800000
Left Foot		#FFFF00
Upper Leg Right		#808000
Upper Leg Left		#008080
Lower Leg Right		#800080
Lower Leg Left		#808080
Upper Arm Left		#FF8000
Upper Arm Right		#FF0080
Lower Arm Left		#FF80FF
Lower Arm Right		#80FF80
Head		#FFFFFF

Table 1. Correspondence between each body part and its RGB colour in the semantic segmentation representation. Best viewed in colour.

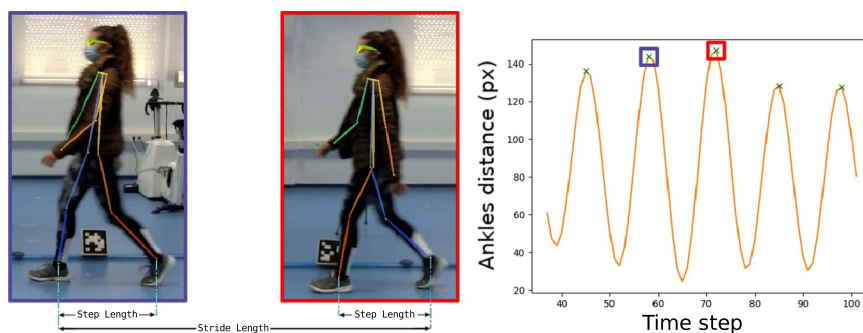


Fig. 5 Visualisation of the step and stride estimation process from the distance between ankles using pose estimation.

not have such body information directly, and it could happen that the hands, due to swinging, would enter the recording zone before the foot. In the case of the body pose, only information about the ankle is available, but in this situation, it would have been the same. Using the cross product between the lines formed by the sensors and the point of the segmentation that is furthest to the right or left, depending on whether the participant is entering or exiting the measurement area and the direction of movement, the exact moments when enters and leaves the measurement area are obtained.

On the other hand, in order to make the two bars of the OptoGait sensors parallel to each other, a rectification process is carried out. This procedure involves obtaining the homography matrix that transforms the points of the measurement area so that both lines of sensors of the OptoGait are parallel.

Lastly, it is necessary to determine the scale to convert distances from pixels to metres, which can be easily obtained from the images considering that the distance between the first and last OptoGait sensor is four metres.

We can estimate the gait parameters once we have all the above information. To estimate the *step length*, we use the participant's pose, specifically, the ankle coordinates. We calculate the distance between the ankles for each frame within the measurement area. We detect when a participant takes a step as the moment when the distance between the ankles is greatest. To facilitate the automatic process of finding these peaks, the distance vector is filtered using the Fast Fourier Transform, eliminating the highest frequencies that add noise. As shown in Fig. 5, once we have identified the peaks, we calculate the distance between the ankles at these points and obtain the average. The previously calculated peak values are used to estimate the *stride*. In this case, the distance between the ankles of the same leg between two consecutive steps is measured using the pose information. For the *speed*, as we know the moments when the participant enters and leaves the measurement area, the video frame rate of 29.97, and the scale, it can be easily estimated. Finally, to estimate the *cadence*, the number of steps is needed, which are the number of peak values calculated and the time the participant spends in the measurement area, previously calculated.

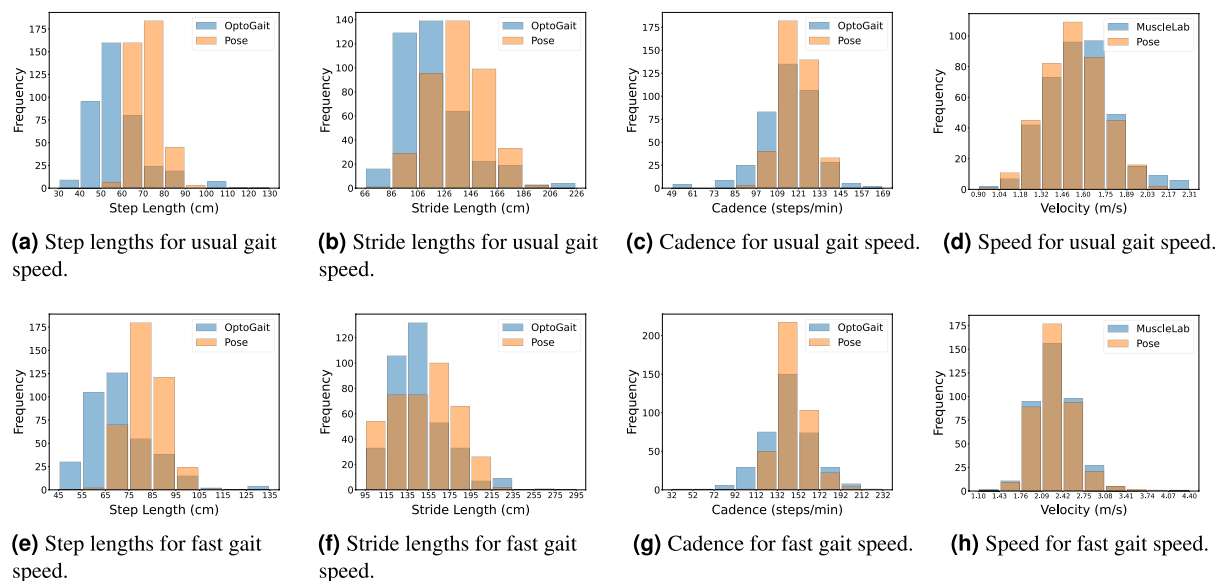


Fig. 6 Comparison between gait parameters obtained with OptoGait or MuscleLAB and the estimations using human pose.

Gait Type	Gait parameter	Mean Error + std	Pearson correlation
UGS	Step Length	16.47 ± 8.20	0.39
	Stride Length	29.05 ± 18.56	0.30
	Cadence	10.48 ± 11.59	0.41
	Speed	0.16 ± 0.23	0.39
FGS	Step Length	15.48 ± 9.23	0.26
	Stride Length	30.68 ± 23.35	0.11
	Cadence	15.14 ± 14.45	0.56
	Speed	0.17 ± 0.15	0.80

Table 2. Mean error between the values obtained by OptoGait or MuscleLAB and those estimated using human pose, and Pearson correlation coefficient. The mean error of Step Length and Stride Length are measured in cm, Cadence in steps per minute, and Speed in meters per second.

In Fig. 6, the gait values obtained with the OptoGait system and MuscleLAB are compared with the estimated ones. The top row shows the estimates obtained for a usual gait speed, and the bottom row shows the estimates obtained from a fast gait speed. It can be seen that, in the case of the *step length* estimation, the estimated values tend to be higher than the parameters obtained with OptoGait. Something similar occurs in the case of the *stride length*, although the overlap is greater. Finally, for the cases of *cadence* and *speed*, the values obtained are similar to those obtained from the sensors. Table 2 shows the mean error between the values obtained with OptoGait and MuscleLAB and the values estimated from human pose. These errors are possibly due to the error in pose estimation models, the necessary rectification of the camera viewpoint, and the shortening of the measurement area, as not all of the area was covered by the camera. Additionally, both feet do not necessarily have to be on the same plane, which can also lead to measurement variations. The Pearson correlation coefficient is also shown, where a slight positive correlation is observed. In the case of Speed FGS, it is observed that it is the variable with the highest correlation, showing a strong correlation with a value of 0.80, while Stride Length is the variable with the lowest correlation, exhibiting a weak correlation with a value of 0.11.

Data Records

The dataset is available at Zenodo³¹. For each participant, the following data has been recorded: 2D pose estimation of their joints, semantic segmentation images, optical flow images, silhouettes, anthropometric measurements, and spatio-temporal gait parameters. Figure 4 shows the pose estimated, semantic segmentation, silhouette, and optical flow obtained for two participants in two of the frames of our dataset. In Fig. 7, a schematic of the database is shown, which is comprised of four main directories corresponding to the information of the pose, optical flow, semantic segmentation, and silhouette, as well as three .csv files, *participants_measures.csv*, *gait_parameters.csv*, and *gait_parameters_estimation.csv* which store anthropometric metrics, spatio-temporal gait parameters obtained with OptoGait and MuscleLAB, and spatio-temporal parameters estimated from the

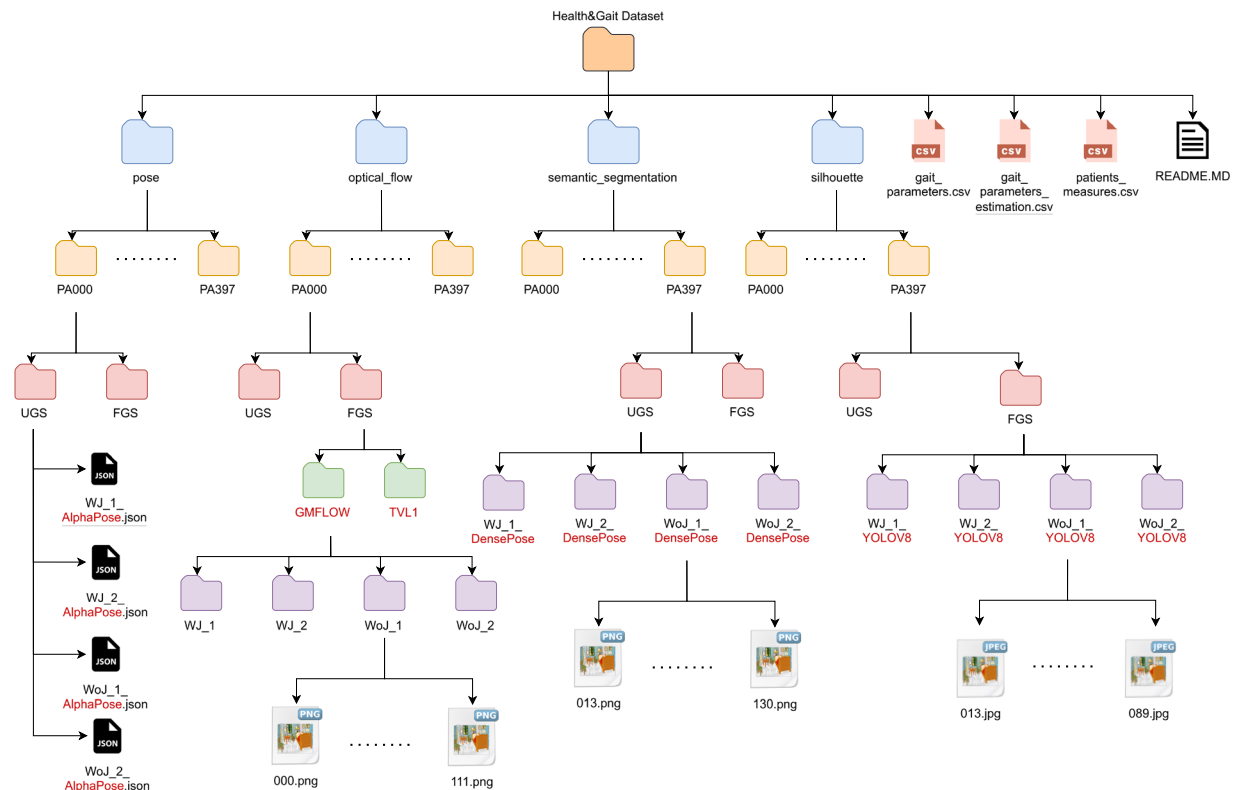


Fig. 7 Directory and file scheme of the Health&Gait database.

pose, respectively. To relate the different records to each other, a unique participant identifier is used in the following format: PAXXX, where X is a digit ranging from zero to nine, so that there is a directory for each participant and in the CSV files there is a column with the participant identifier. For the pose, optical flow, semantic segmentation, and silhouette, each participant's data is divided into usual gait speed and fast gait speed, and for each of these, there is further differentiation between information about the class without a jacket and with a jacket. It should be noted that in the case of optical flow, there are a couple of extra directories for the two selected methods, GMFlow and TVL1. In the final directories, a '1' or '2' in the directory name indicates the direction of the walk, with '1' corresponding to movement from right to left and '2' from left to right. The information contained in each data record is detailed below, as well as the format used.

The 2D pose information is stored using the JSON format file, where for each participant video, the frame number, x and y coordinates of the 17 joints (*null* if not possible to estimate) using the COCO 17 keypoints representation³², and the bounding box are stored. In Fig. 8, you can see a skeleton with the estimated joints. The semantic segmentation information is stored using PNG images, where each segmented part is associated with a colour as indicated in Table 1. The total number of parts is 14. The image resolution is 960×540 pixels. Storing the optical flow of each sequence in binary files would significantly increase the size of the dataset. Therefore, a representation in PNG images with a resolution of 480×270 pixels has been chosen, where the meaning of each pixel value is indicated in the 'Data processing' section. The silhouette information is stored using JPEG images, with a resolution of 960×540 pixels, where the white pixels indicate where the participant is in the image, and the black pixels indicate the absence of the participant.

In Table 3, a summary of the information provided in the file *participants_measures.csv* is displayed along with the measurement units used for each attribute. The file contains an entry for each participant in the dataset; however, the variables *PA_level*, *Weight*, *BMI*, *Percentage_fat_mass*, *Lean_mass*, and *HR_Final*, contain missing values, representing a total of 1.08% of all data, due to measurement errors or because these values were not taken for that participant.

Table 4 describes the metrics obtained with the OptoGait system. It should be noted that in the file *gait_parameters.csv*, there is a field for normal and fast gait speed, indicated by a UGS or FGS, respectively. The variables *Step_UGS*, *Stride_UGS*, *Cadence_UGS*, *MonoSP_UGS*, *BiSP_UGS*, *Velocity_UGS*, *Step_FGS*, *Stride_FGS*, *Cadence_FGS*, *MonoSP_FGS*, and *BiSP_FGS* contain missing values, representing a total of 4.19% of all data. In the file name *gait_parameters_estimation.csv*, the gait parameter values estimated from each participant's pose are stored. As previously discussed, the estimated parameters include *step*, *stride*, *cadence*, and *speed* for both normal and fast walking paces. The labels used are identical to those employed in the *gait_parameters.csv*.

Figure 9 shows the relationship between speed and body mass index and the distribution of participants according to age and sex. It can be seen that the participants with the highest speed are in the Normal BMI category, while the maximum speed of patients with obesity has a much lower value. The fact that the average speed of an obese person is more or less in the normal range may be because the distance over which speed is

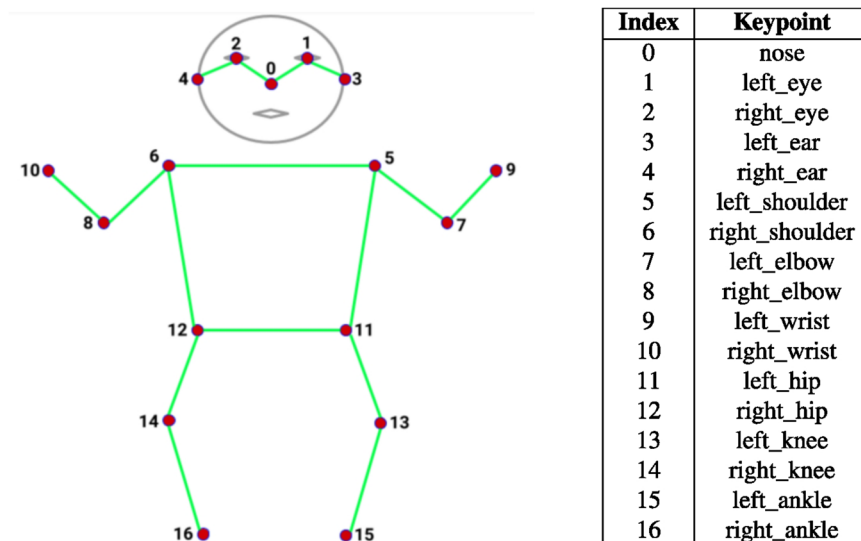


Fig. 8 COCO Body Format 17 keypoints (<https://github.com/tensorflow/tfjs-models/blob/master/pose-detection/README.md>).

Attributes	Description	Unit
Sex	Participant sex	0:female, 1:male
Age	Participant Age	Years
PA_level	Level of physical activity	>=3 days: Active, <3 days: Non active
Height	Participant height	cm
Weight	Participant weight	kg
BMI	Body Mass Index	kg/m ²
WaistC	Waist circumference	cm
HipC	Hip circumference	cm
NeckC	Neck circumference	cm
Percentage fat mass	The total mass of fat divided by total body mass	%
Lean mass	The difference between total body weight and body fat weight	kg

Table 3. Attributes in the file *participants_measures.csv*.

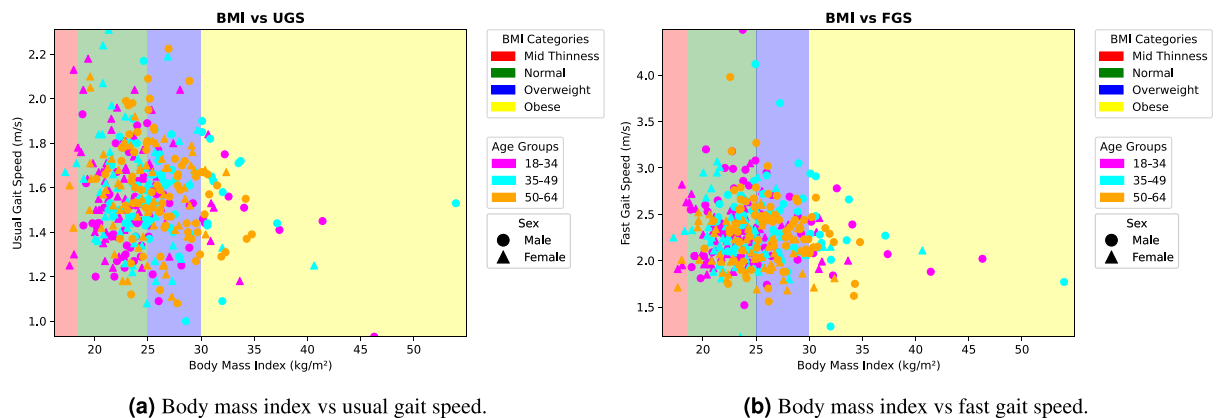
Attributes	Description	Unit
Step_UGS / Step_FGS	The distance between the two toes or heels of the feet in sequence for usual/fast gait speed.	cm
Stride_UGS / Stride_FGS	The distance between the two toes or heels of sequential strides of the same foot for usual/fast gait speed.	cm
Cadence_UGS / Cadence_FGS	The number of steps taken per unit of time for usual/fast gait speed.	Steps/min
MonoSP_UGS / MonoSP_FGS	Time in the swing phase where only one limb is in contact with the ground for usual/fast gait speed.	sec
BiSP_UGS / BiSP_FGS	Time that both feet are on the ground for usual/fast gait speed.	sec
Speed_UGS / Speed_FGS	Participant velocity for usual/fast gait speed.	m/s

Table 4. Attributes in the file *gait_parameters.csv*. There is an attribute for each type of gait speed.

measured is not large enough to begin to notice very noticeable differences in speed in participants with obesity. Finally, you can observe how the majority of participants with obesity tend to be male while, on the contrary, the majority of participants with mid-thinness are female.

Technical Validation

To demonstrate the quality control of the dataset, only state-of-the-art methods, which underwent a peer-review process, have been used to obtain the different provided data representations, as previously indicated in Sec. Methods (Subsec. Data processing). In addition, we have used the OptoGait system to measure gait parameters, which operates at a frequency of 1000 Hz and has an accuracy of 1 cm, as stated in its user manual



(a) Body mass index vs usual gait speed.

(b) Body mass index vs fast gait speed.

Fig. 9 Visualisation of the influence of body mass index on participants' speed.

(<https://medical.microgate.it/sites/default/files/manuali/optogait/Manual-EN.pdf>). The Optical Timing Gates (Photo-cells) from MuscleLAB, the system used to measure speed, have a resolution value of 2 ms, allowing for precise measurement of displacement times, as obtained from the MuscleLAB website (<https://www.musclelab-system.com/products/>). Furthermore, all measurements have been carried out under the supervision of sports science specialists.

Additionally, a series of computational experiments have been conducted to demonstrate the quality control of the dataset. For this, the experiments have been conducted to determine whether the biometric information of participants is encoded in the representations used for data. For this purpose, the task of sex classification and regression of participants' weight and age are used. For each task, the different data representations from the dataset videos, namely, silhouette, semantic segmentation, and optical flow, were used. Additionally, the gait parameters derived from the OptoGait and those estimated from the videos using pose information have been used to verify the utility of gait parameters in these tasks.

The tasks of sex classification and regression of age and weight are approached, on the one hand, using the Mobile Video Networks (MoViNets)³³ video classification architecture, training the model with information from the silhouette, semantic segmentation, and optical flow. On the other hand, a multilayer perceptron (MLP) and XGBoost³⁴ are trained using the gait parameters, which are *step length*, *stride length*, *cadence* and *speed*.

The MoViNets architecture has been selected as it is a Convolutional Neural Network (CNN) designed to achieve high classification rates while reducing computational and memory requirements, making it ideal for use on lower-performance devices, such as mobile devices. On the other hand, MLP and XGBoost have been chosen as they are state-of-the-art models in a wide variety of applications for classification and regression.

For the MoViNets architecture, the images need to be resized to 224×224 pixels and normalised. For the MLP and XGBoost models, an exhaustive hyperparameter tuning was conducted using the Bayesian Optimization algorithm³⁵. The search space is the same for each of the experiments conducted. Below, the hyperparameter search space for multilayer perceptron is detailed.

- Number of layers: takes values in the range [1, 5].
- Number of units per layer: takes values in the range [32, 512] with a step of 32.
- Activation function: Activation function used in the hidden layers. Takes values *ReLU* or *tanh*.
- Dropout: Boolean variable that indicates whether a dropout of 25% of the weights is performed or not.
- Learning rate: takes values in the range $[1 \times 10^{-4}, 1 \times 10^{-2}]$ using a logarithmic sampling.

Below, the hyperparameter search space for XGBoost is detailed.

- Max Depth: the maximum depth of a tree. Used to control over-fitting. Takes values in the range [3, 10], and the values 12 or 15
- Min child weight: defines the minimum sum of weights of observations required in a child. Takes values in the range [1, 8].
- Gamma: specifies the minimum loss reduction required to make a split. Takes values in the range [0, 0.5] with a step of 0.1, and the values 1.0, 1.5 and 2.0.
- Subsample: denotes the fraction of observations to be random samples for each tree. Takes values in the range [0.5, 1.0] with a step of 0.1.
- Colsample bytree: denotes the fraction of columns to be random samples for a tree. Takes values in the range [0.5, 1.0] with a step of 0.1.
- Learning rate: takes values in the range [0.01, 0.3] with a step of 0.05.
- Number of estimators: takes the value 50 or values in the range [100, 1000] with a step of 100.
- Alpha: L1 regularization. Takes the values 0, 0.1, 3 or values in the range [0.5, 2] with a step of 0.5.
- Lambda: L2 regularization. Takes the values in the range [0.5, 2.0] with a step of 0.5, and the values 3, 4.5, 5, 6, 7 and 8.

To create the training and test sets, a 4-fold stratified split is applied based on the participants in the dataset, ensuring that there are no samples from the same participant in both the training and test sets. Additionally, 10% of the participants in the training set are selected for the validation set. The results obtained were averaged across each of the folds.

For the sex classification and age estimation tasks, participants are selected stratified according to their age and sex, ensuring a balanced representation of men and women in each age group. For the weight estimation task, participants are also selected in a stratified way, organised to ensure equal representation of men and women across the four body mass index (BMI) categories.

Moreover, all random number generators involved in obtaining the results have been seeded with a value of 27.

For classification, the average accuracy and F-Score across the four partitions were used, while for regression problems, the average mean absolute error along with the standard deviation.

To view more detailed results and discussions of the experiments conducted, please visit the GitHub repository (<https://github.com/AVAuco/healthgait>)

Usage Notes

The reader is provided with all the code necessary to replicate the results of the computational experimentation performed, including details on the partitions of participants used for training, validation, and testing. The main functionalities of the code include:

- Preprocessing of RAW videos for pose estimation, silhouette, semantic segmentation, and optical flow.
- Creation of participant partitions used for training, validation, and testing.
- Estimation of gait parameters using pose information and scene information for camera perspective correction.
- Training of models for sex classification, using either visual information or gait parameters estimated or obtained from OptoGait.
- Training of models for weight and age regression, using either visual information or gait parameters estimated or obtained from OptoGait.
- Hyperparameter tuning for MLP and XGBoost models.
- Evaluation of trained models to obtain classification or regression metrics.
- Generation of confusion matrices and error histograms.

Other potential dataset uses might include attempts to estimate other patient attributes, such as physical activity level or body mass percentage. Additionally, one could leverage the latest advancements in classification or regression algorithms, for example, by using architectures based on Transformers. Finally, multimodal architectures could be employed to determine whether the dataset's information is complementary for addressing different problems.

The main limitation of the dataset is that raw video recording are not provided, which, combined with the challenging scenario where only a single camera with a side view is used, may result in some post-processed outcomes being unreliable.

Limitations

The limitation in the experimental design is that it was not possible to synchronise the video recordings with the measured gait parameters, i.e. the measurements were obtained first, and then the video recordings were made. Despite this issue, the fluctuation in the data is not very high since a person's walking pattern will not change radically from one moment to another.

While it was initially overlooked due to resource limitations, incorporating a multi-camera recording system alongside an advanced motion capture setup would be valuable in future data collection efforts, enabling comprehensive 3D motion extraction.

Code availability

In addition to releasing the data, we also make the code used in the experiments available. All code is available on GitHub at <https://github.com/AVAuco/healthgait>.

Received: 11 March 2024; Accepted: 18 December 2024;

Published online: 10 January 2025

References

1. Webster, J. B. & Darter, B. J. 4 - principles of normal and pathologic gait. In *Atlas of Orthoses and Assistive Devices (Fifth Edition)*, 49–62.e1 (Elsevier, 2019).
2. Perry, J. & Burnfield, J. Gait analysis: Normal and pathological function, second edition. *Journal of Sports Science and Medicine* **9**, 353 (2010).
3. Czech, M. D. *et al.* Age and environment-related differences in gait in healthy adults using wearables. *npj Digital Medicine* **3**, 127 (2020).
4. Studenski, S. *et al.* Gait Speed and Survival in Older Adults. *JAMA* **305**, 50–58 (2011).
5. Abellan Van Kan, G. *et al.* Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people: an international academy on nutrition and aging (iana) task force. *The Journal of nutrition, health and aging* **13**, 881–889 (2009).
6. Sethi, D., Bharti, S. & Prakash, C. A comprehensive survey on gait analysis: History, parameters, approaches, pose estimation, and future work. *Artificial Intelligence in Medicine* **129**, 102314 (2022).
7. Stuck, A. K., Bachmann, M., Fullemann, P., Josephson, K. R. & Stuck, A. E. Effect of testing procedures on gait speed measurement: A systematic review. *PLOS ONE* **15**, e0234200 (2020).

8. Carvajal-Castano, H. A., Lemos-Duque, J. D. & Orozco-Arroyave, J. R. Effective detection of abnormal gait patterns in parkinson's disease patients using kinematics, nonlinear, and stability gait features. *Human Movement Science* **81**, 102891 (2022).
9. Coca-Tapia, M., Cuesta-Gomez, A., Molina-Rueda, F. & Carratala-Tejada, M. Gait pattern in people with multiple sclerosis: A systematic review. *Diagnostics* **11**, 584 (2021).
10. Masse, F. A. A. *et al.* Progression of gait changes in older adults with mild cognitive impairment: A systematic review. *Journal of Geriatric Physical Therapy* **44**, 119–124 (2021).
11. Pau, M. *et al.* Kinematics adaptation and inter-limb symmetry during gait in obese adults. *Sensors* **21**, 5980 (2021).
12. Microgate. Optogait product details (2024). <https://medical.microgate.it/en/products/optogait>. [Accessed on: September 27, 2023].
13. Zeng, X., Zhang, X., Yang, S., Shi, Z. & Chi, C. Gait-based implicit authentication using edge computing and deep learning for mobile devices. *Sensors* **21**, 4592 (2021).
14. Connor, P. & Ross, A. Biometric recognition by gait: A survey of modalities and features. *Computer Vision and Image Understanding* **167**, 1–27 (2018).
15. Sepas-Moghaddam, A. & Etemad, A. Deep gait recognition: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45**, 264–284 (2023).
16. Zhou, L., Fischer, E., Brahms, C. M., Granacher, U. & Arnrich, B. Duo-gait: A gait dataset for walking under dual-task and fatigue conditions with inertial measurement units. *Scientific Data* **10**, 543 (2023).
17. Huang, C., Zhang, F., Xu, Z. & Wei, J. The diverse gait dataset: Gait segmentation using inertial sensors for pedestrian localization with different genders, heights and walking speeds. *Sensors* **22**, 1678 (2022).
18. Albuquerque, P., Machado, J. P., Verlekar, T. T., Correia, P. L. & Soares, L. D. Remote gait type classification system using markerless 2d video. *Diagnostics* **11**, 1824 (2021).
19. Control CfD, Prevention., Atlanta, GA, USA: CDC. *National health and nutrition examination survey: Anthropometry procedures manual* (2007).
20. Marfell-Jones, M., Olds, T., Stewart, A. & Carter, L. *ISAK Accreditation Handbook*. ISAK Accreditation Handbook. (International Society for the Advancement of Kinanthropometry (ISAK), 2006).
21. Bioelectrical impedance analysis in body composition measurement: National institutes of health technology assessment conference statement. *The American Journal of Clinical Nutrition* **64**, 524S–532S (1996).
22. Armstrong, L. *et al.* Urinary indices during dehydration, exercise, and rehydration. *International Journal of Sport Nutrition* **8**, 345–355 (1998).
23. Zach, C., Pock, T. & Bischof, H. A duality based approach for realtime tv-l1 optical flow. In *Pattern Recognition, Proceedings*, vol. 4713, 214+ (2007).
24. Xu, H., Zhang, J., Cai, J., Rezatofighi, H. & Tao, D. Gmflow: Learning optical flow via global matching. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8111–8120 (2022).
25. Baker, S. *et al.* A database and evaluation methodology for optical flow. *International Journal of Computer Vision* **92**, 1–31 (2011).
26. Fang, H.-S. *et al.* Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45**, 7157–7173 (2023).
27. Jocher, G., Chaurasia, A. & Qiu, J. Ultralytics YOLO <https://github.com/ultralytics/ultralytics> (2023).
28. Guler, R. A., Neverova, N. & Kokkinos, L. Densepose: Dense human pose estimation in the wild. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 7297–7306 (2018).
29. Kirillov, A., Girshick, R., He, K. & Dollar, P. Panoptic feature pyramid networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019)*, 6392–6401 (2019).
30. Kirillov, A. *et al.* Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 4015–4026 (2023).
31. Zafra-Palma, J. *et al.* Health&Gait: a video dataset for gait-based analysis <https://doi.org/10.5281/zenodo.14039922> (2024).
32. Lin, T.-Y. *et al.* Microsoft COCO: Common Objects in Context. In Fleet, D., Pajdla, T., Schiele, B. & Tuytelaars, T. (eds.) *European Conference on Computer Vision (ECCV)*, vol. 8693, 740–755 (2014).
33. Kondratyuk, D. *et al.* MoViNets: Mobile video networks for efficient video recognition. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2021*, 16015–16025 (2021).
34. Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794 (2016).
35. Head, T., Kumar, M., Nahrstaedt, H., Louppe, G. & Shcherbatyi, I. scikit-optimize/scikit-optimize <https://doi.org/10.5281/zenodo.4014775> (2020).

Acknowledgements

This project was partially supported by the project TED2021-129151B-I00/AEI/10.13039/501100011033/ European Union NextGenerationEU/PRTR; the Ministry of Economy, Industry and Competitiveness in the 2017 call for R&D Projects of the State Program for Research, Development and Innovation Targeting the Challenges of the Company; the Regional Government of Andalusia and University of Cadiz: Research and Knowledge Transfer Fund (PPIT-FPI19-GJ4F-10); and project PID2023-147296NB-I00 of the Spanish Ministry of Economy, Industry and Competitiveness.

Author contributions

M.J.M.-J., J.Z.-P. and N.M.-J. conceived the dataset design; N.M.-J., J.C.-P. and M.C.-G. collected the subjects' videos and data; M.J.M.-J. and J.Z.-P. conceived the experiments; J.Z.-P. and M.J.M.-J. conducted the experiments; M.C.-G. and R.M.-S. analysed the results; J.Z.-P. and N.M.-J. wrote the first draft of the manuscript; and J.Z.-P. and R.M.-S. prepared the data and code release. All authors reviewed the manuscript and contributed to its current version.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.Z.-P. or M.J.M.-J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025