

Identifying Underlying Individuality Across Running, Walking, and Handwriting Patterns with Conditional Cycle-Consistent Generative Adversarial Networks

Johannes Burdack ¹, Sven Giesselbach ², Marvin Simak ², Mamadou Ndiaye ², Christian Marquardt ², and Wolfgang Immanuel Schöllhorn ²

¹Department of Training and Movement Science

²Affiliation not available

October 30, 2023

Abstract

In recent years, the analysis of movement patterns has increasingly focused on the individuality of movements. After long speculations about weak individuality, strong individuality is now accepted, and the first situation-dependent fine structures within it are already identified. Methodologically, however, only signals of same movements have been compared so far. The goal of this work is to detect cross-movement commonalities of individual walking, running, and handwriting patterns using data augmentation. 17 healthy adults (35.8 ± 11.1 years, 8 females, 9 males) each performed 627.9 ± 129.0 walking strides, 962.9 ± 182.0 running strides, and 59.25 ± 1.8 handwritings. Using the conditional CycleGAN, conditioned on the participant's class, a pairwise transformation between the vertical ground reaction force during walking and running and the vertical pen pressure during handwriting was learned in the first step. In the second step, the original data of the respective movements were used to artificially generate the other movement data. In the third step, it was tested whether the artificially generated data could be correctly assigned to a person via classification using a support vector machine trained with original data of the movement. The classification F1-score ranged from 46.8% for handwriting data generated from walking data to 98.9% for walking data generated from running data. Thus, cross-movement individual patterns could be identified. Therefore, the methodology presented in this study may help to enable cross-movement analysis and the artificial generation of larger amounts of data.

Identifying Underlying Individuality Across Running, Walking, and Handwriting Patterns with Conditional Cycle–Consistent Generative Adversarial Networks

Johannes Burdack, Sven Giesselbach, Marvin Simak, Mamadou L. Ndiaye, Christian Marquardt, Wolfgang I. Schöllhorn

Abstract– In recent years, the analysis of movement patterns has increasingly focused on the individuality of movements. After long speculations about weak individuality, strong individuality is now accepted, and the first situation-dependent fine structures within it are already identified. Methodologically, however, only signals of same movements have been compared so far. The goal of this work is to detect cross-movement commonalities of individual walking, running, and handwriting patterns using data augmentation. 17 healthy adults (35.8 ± 11.1 years, 8 females, 9 males) each performed 627.9 ± 129.0 walking strides, 962.9 ± 182.0 running strides, and 59.25 ± 1.8 handwritings. Using the conditional CycleGAN, conditioned on the participant's class, a pairwise transformation between the vertical ground reaction force during walking and running and the vertical pen pressure during handwriting was learned in the first step. In the second step, the original data of the respective movements were used to artificially generate the other movement data. In the third step, it was tested whether the artificially generated data could be correctly assigned to a person via classification using a support vector machine trained with original data of the movement. The classification F_1 -score ranged from 46.8% for handwriting data generated from walking data to 98.9% for walking data generated from running data. Thus, cross-movement individual patterns could be identified. Therefore, the methodology presented in this study may help to enable cross-movement analysis and the artificial generation of larger amounts of data.

Cross-Movement Individuality, Cross-Signal Individuality, CycleGAN, Data Augmentation, Deep Learning, Generative Adversarial Network, Movement Pattern Recognition, Support Vector Machine

No external funding was provided to conduct this study.

J. Burdack, Department of Training and Movement Science, Institute of Sport Science, Johannes Gutenberg-University, Mainz, Germany (e-mail: burdack@uni-mainz.de),

S. Giesselbach, Knowledge Discovery, Fraunhofer-Institute for Intelligent Analysis and Information Systems, Sankt Augustin, Germany (e-mail: sven.giesselbach@iais.fraunhofer.de),

M. Simak, Department of Training and Movement Science, Institute of Sport Science, Johannes Gutenberg-University, Mainz, Germany (e-mail: masimak@uni-mainz.de),

I. INTRODUCTION

After weak individuality had been used only sporadically for quite some time and meanwhile acquired the status of a buzzword, especially in connection with learning, training and therapy, the investigation of strong individuality had been increasingly been asked for [1], [2]. Often a mixture of colloquial (weak) and science-oriented (strong) understanding can be observed, which is even more confusingly equated with “personalized” [3]–[6]. Colloquially, weak individuality often serves as an excuse for lack of statistically significant group differences [7]–[9] or for not finding commonalities across individuals [9]–[12]. In contrast, science on individuality is guided by the much stronger criteria of forensics, which must withstand legal proof for sentencing purposes. The two essential criteria are uniqueness and persistence [13], [14], for the proof of which, first, a larger amount of data is necessary and, second, a different statistical method than the average oriented in social sciences is required. Both conditions explain why it is only with the more recent development of appropriate methods and techniques that the study of the strong individuality of selected forms of movement has increased almost inflationary. Originating from the visual perception of walking individuals [15]–[17], followed by biomechanical analyses of gait movements [18]–[20] and sporadically single sports movements [21], [22] meanwhile analyses of a wide variety of movements have become increasingly popular. Besides walking [20], [23], [24], individuality of movements has also been shown in the field of sports in sprinting [25], running [26]–[29], javelin [21] and discus throwing [22], and horseback riding [30] and in the field of music for playing the flute [31]. Similar features could be shown for team behavior in

M. L. Ndiaye, Department of Training and Movement Science, Institute of Sport Science, Johannes Gutenberg-University, Mainz, Germany (e-mail: mandiaeye@uni-mainz.de),

Christian Marquardt, Science&Motion GmbH, Munich, Germany (e-mail: christian.marquardt@scienceandmotion.de)

W. I. Schöllhorn, Department of Training and Movement Science, Institute of Sport Science, Johannes Gutenberg-University, Mainz, Germany (e-mail: schoellw@uni-mainz.de).

volleyball [32], [33], soccer [34], [35], or basketball [36], [37]. Typical parameters investigated are biomechanical data from video recordings [38], force plates [39]–[41], pressure insoles [42], [43], EMG [44], [45], and brain signals [46], [47]. Besides these, wearable sensors are becoming increasingly popular [48]. Situated “perturbations” such as emotions [49], fatigue [50], or time alone [21], [39]–[41] were not able to move the patterns out of the strong individual space. In consequence, robust evidence for an important role of individuality is provided in short-term adaptive behavior. All studies listed so far have in common that they answer questions that could be carried out based on the comparison of a single movement technique. From this, the problem of individuality in longer-term learning, like in training or therapy, must be distinguished, especially with respect to the criterion of permanence. Repeating of learning of the same movement by the same person never encounters comparable conditions again due to cognitive and body memory [51]. To solve this problem, finding commonalities of the learning of different movements seems an appropriate approach, but requires the identification of movement independent individuality. In a first proposal, evidence for individual characteristics across three throwing techniques of the decathlon (final throwing phase of shot put, discus, javelin) with similar kinematic structure was provided [51]. The aim of this study is to identify individual commonalities of movement forms with different kinematic structure.

The problem with this is that the classification models cannot transfer between domains and thus only work on data from one domain. Consequently, it is not straightforward to train a classification model with walking data and test it on running data to identify any common underlying structures. However, new methods from the field of deep learning provide potential to address this problem. Approaches from image generation offer solutions for analogue problems. In image-to-image translation or style transfer, it is possible to learn a relationship between images from two domains A (e.g., horses) and domain B (e.g., zebras), so that realistic images of domain B can then be generated from images of domain A. This has been done without losing the image content. Images of horses can become pictures of zebras, or images of a landscape can become pictures of the same landscapes as it might look in summer or in winter.

In the area of data generation, the Generative Adversarial Networks (GANs) [52], [53] have proven to be extremely successful in generating new, previously unseen data that is somewhat similar to a given training data set. In order to solve “generative modeling” problems, the goal of a GAN is to learn a probability distribution of the data to be generated. Based on this probability distribution, the GAN then generates new data from this probability distribution. However, a major challenge with GANs is that they require a very large database of paired data to solve image-to-image translation [54]. This makes the GANs difficult to apply in the context of cross-movement studies for many problems. First, it is often not possible to generate the necessary large data sets, and second, due to the continuous change of movement patterns [55], finding matching pairs of movements between two movements can be problematic.

A further development of the GAN, that circumvents the pairwise data problem respectively offers a possible solution

and can be successfully applied on relatively small data sets, is the cycle-consistent GAN (CycleGAN) [54]. Here, images from domain A can be translated into domain B, but the basic content from domain A is preserved. A well-known example that makes use of this method is the FaceApp (FaceApp Technology Limited, Cyprus). Given images of faces, the app allows for a transformation that makes someone’s face laugh, look older, look younger, or appear in the style of the opposite sex.

While the CycleGANs work quite well on images, it has not been applied to movement measurements. Therefore, in this work, we will use the CycleGAN to identify common individual patterns across different movements. Specifically, we aim to find individual commonalities underlying walking, running, and handwriting patterns of the same person.

With the CycleGAN, we generate artificial movement data of movement B from the original data of a movement A. Specifically, we generate the other two movements from the movements walking, running, and handwriting (i.e., walking to running, walking to handwriting, running to walking, running to handwriting, handwriting to walking, and handwriting to running). The movements were chosen from the point of view that with walking and running we have two related movements and with handwriting one very different from them. Based on former studies on individuality [20], [22], [50], [55], [56], [30], [31], [37], [39]–[41], [43], [49], we assume that individuals can be distinguished by their walking, running, and handwriting patterns. From this, we derive the following research questions:

Can CycleGANs artificially generate pairwise data between walking, running, and handwriting movements, and can this artificial data be assigned to the correct individuals?

II. MATERIAL AND METHODS

A. Participants and Ethics Statement

The study participants were 17 athletically active, healthy adults (8 females, 9 males; 1 left-handed, 16 right-handed) who regularly handwrite and run for health reasons (the group characteristics are shown in Table 1). Before participating in the study, the participants signed informed consent forms. The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of the Johannes Gutenberg-University Mainz (2022/05; 5/23/2022). Each participant visited the biomechanics laboratory once, where all measurements took place.

TABLE 1
Participant Characteristics

	M	SD
Age (years)	35.8	11.1
Height (cm)	172.1	6.2
Body mass (kg)	68.0	9.3
Preferred walking speed (km/h)	4.2	0.5
Preferred running speed (km/h)	8.4	1.4

Data are presented as mean (M) and standard deviation (SD), preferred walking and running speed was determined while walking/running on the treadmill.

B. Experimental Protocol

At the beginning of the study, the preferred walking speed (PWS) and preferred running speed (PRS) on the treadmill were determined for each participant [57]. The PWS is a speed that the participants prefer in their leisure time, for example when going for a walk, and the PRS a speed at which they “feel comfortable” and “can keep going for a very long time”. At the same time, the determination of PWS and PRS also served as a habituation to the treadmill. This was followed by familiarization with writing on the digitizing tablet with a pressure-sensitive pen, with everyone writing the sentence that was also written in the data collection (see below) five times.



Fig. 1. Experimental sequence with chronological sequence of the three walking, three running and six handwriting conditions. The walking and running conditions were always alternated with handwriting. In addition, walking and running were always alternated for stress control reasons. The starting condition (walking or running) was randomized between participants.

As presented in Figure 1, the participants performed 6 sets of four minutes running or walking. Each was followed by a four-minute break during which they performed 10 handwriting trials. To achieve greater variation within participants (for more robust training of the Deep Learning models, we increased the variance), we varied the speed in each of the three walks and runs slightly from slow 85% PWS/PRS, to normal: 100% PWS/PRS, and fast 115% PWS/PRS. To avoid sequence effects, we randomly shuffled the order of the walking and running conditions across all participants, with the only restriction being that walking and running must always alternate due to load control. Between each walk and run, the phrase “Wellen folgen den Bewegungen” [English: “waves follow the movements”] was handwritten 10 times. The sentence was chosen because it was as neutral and as contentless as possible in terms of meaning. Again, to provoke greater variation in the data, a new instruction was given for each handwriting set, which was implemented at the discretion of the participants. The instructions included writing “normal” or “bigger”, “smaller”, “faster”, “slower”, and “more beautiful” as usual. Again, we randomized the order of each instruction.

C. Data Acquisition

The movements investigated in this study are walking, running, and handwriting. Walking and running were performed on a treadmill (cos12148, h/p/cosmos, Leipzig, Germany) and recorded with pressure soles (pedar, novel, Munich, Germany) at a frequency of 100 Hz. The handwriting was performed using a pressure sensitive pen (Wacom Pro Pen 3D, Wacom, Düsseldorf, Germany) with 4096 pressure levels on a digitizing tablet on which a paper was adjusted (Wacom Intuos Pro Paper Edition L, Wacom, Düsseldorf, Germany) with a recording frequency of 200 Hz and recorded with the software CSWin (CSWin 2016, MedCom Verlag, Munich, Germany).

D. Data Processing

For the locomotion tasks, the vertical ground reaction force (GRF) was calculated from the pressure data using Pedar Mobile Expert software (version 8.2). The stance phase from heel strike to toe-off of the left and right foot was determined using a vertical GRF threshold of 50N. Each ground contact with one foot was time normalized to 128 values. Burdack et al. [58] showed that the exact vector length plays a minor role as long as the curve shape is preserved. In addition, the data were normalized by body weight and scaled to the range [0, 1] [59]–[61]. After scaling, the step pairs from the left and right ground contact were combined into one vector of 256 values (128 data points left foot + 128 data points right foot). Each vector begins with a left ground contact and ends with the corresponding right one. If the data of a ground contact was incorrect during recording, this and the corresponding ground contact of the other foot were deleted from the recording.

For handwriting, only the vertical pen pressure data where the pressure on the pen was greater than 0 was considered. Furthermore, we considered only the first letter W for the handwriting analysis. If the W was not written in one piece, the test was discarded. In the case that the e was written from the W without settling, the point with the least pressure between the W and the e determined the end of the W. In addition, the handwriting data were filtered with a 1st-order Savitzky-Golay filter with a window size of 13 [62], which smoothes on the least squares method while maintaining the shape and height of the waveform peaks [63]. Data were also time-normalized to 256 data points, to have the same length as the GRF data, z-standardized, and scaled to the range [0, 1]. The reduction of the entire sentence to the letter W had several reasons. First, preliminary measurements showed that compressing the signal of the entire sentence to 256 data points meant that the handwriting could no longer be generated sufficiently well. Derived from this, we wanted to obtain a signal that was similarly complicated and on a similar time scale in execution as that of the locomotion movements.

E. Data Analysis

1) Data Analysis Procedure

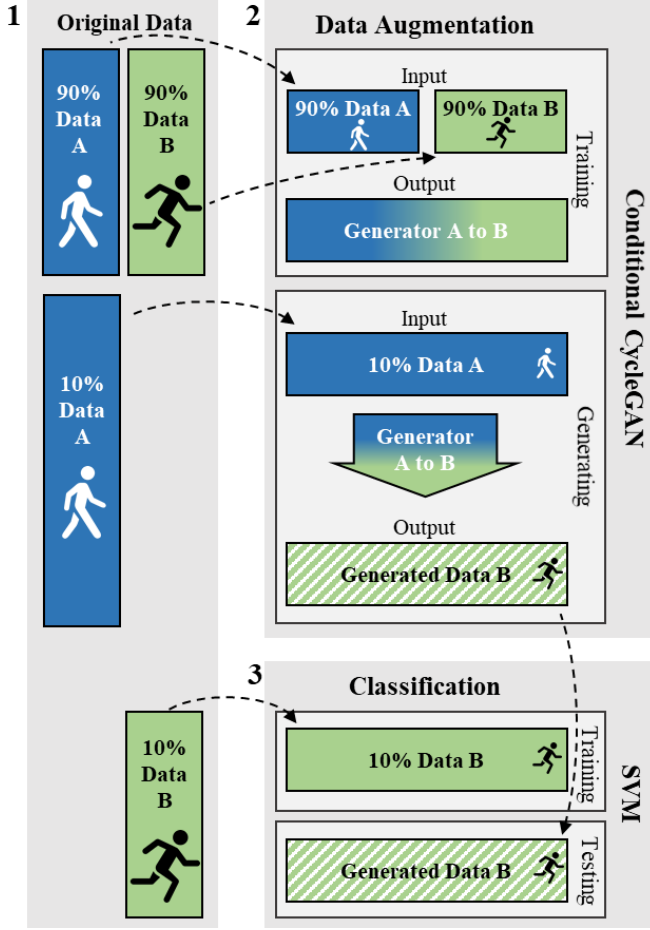


Fig. 2. Description of the procedure from data generation to classification. For the sake of clarity, only the way of generating new data from data A is described in the figure. However, the way of generation from data B is analogous. For illustration purposes, data A corresponds to walking and data B to running. (1) Original Data: In each case, the original data was split 90:10. (2) The outlined process of the conditional CycleGAN: 90% of the data of A and B were used for training the conditional CycleGAN. The trained generators were then used to generate new data B from the remaining 10% of data A accordingly. (3) The outlined procedure of classification using a Support Vector Machine (SVM): the SVM was trained with the 10% of data B that was not used for training the conditional CycleGAN. In each case, the SVM was tested with the newly generated data B.

Figure 2 shows a schematic example of the data analysis flow for a portion of the data (Data A). It is important to emphasize that the conditional CycleGAN training data is separate from the generation data and from the SVM training data. While the conditional CycleGAN training for data A and B occurs simultaneously, the paths of data A and B are strictly separated from the time of data generation. The details of data generation and classification are described below.

2) Data Generating – Conditional CycleGAN

The CycleGAN is a deep learning approach originally developed for unpaired image-to-image translation [54]. Image-to-image translation involves generating a new synthetic version of a given image with a specific modification, e.g., translating a photograph into a Monet-style painting while preserving the original content. Training a model for image-to-image translation usually requires a large data set of paired data. Creating such data sets can be difficult, expensive, or even impossible, such as with photographs of paintings by long-dead artists, or pairs of walking and running data. The CycleGAN is a technique that involves automatic training of image-to-image translation models without paired examples. The models are trained in an unsupervised manner using a collection of data from the source and target domains, which need not be connected in any way.

In our case, we want to translate the movements of walking, running and handwriting, keeping individual features in each case. Again, an unpaired approach is particularly appropriate because we want to translate data from one movement domain (e.g., walking) into data from another movement domain (e.g., running) without knowing a possible mapping of the different movements of each domain to each other.

In doing so, our approach deviates somewhat from the original CycleGAN formulation. Since we want to preserve the individual component via the movement data transfer as well, we have added a class condition to the conventional CycleGAN in the following.

a) Formulation of the conditional CycleGAN

The goal of the conditional CycleGAN, conditioned on the participant's class label, is to learn mapping functions between two domains X and Y under the condition of the classes Z given the training samples $\{x_i\}_{i=1}^N$ where $x_i \in X$ and $\{y_i\}_{i=1}^N$ where $y_i \in Y$, and the class embedding $\{z_i\}_{i=1}^N$ where $z_i \in Z$. For simplicity, the indices i and j are omitted in the following. The data distribution is denoted as $x \sim p_{data}(x)$ and $y \sim p_{data}(y)$, for the input of the original data, $z \sim p_{class}(z)$ for the class embedding, and $x, z \sim p_{data}(x, z)$ and $y, z \sim p_{data}(y, z)$ for the data under the class condition. The conditional CycleGAN includes two mappings $G: (X, Z) \rightarrow Y$ and $F: (Y, Z) \rightarrow X$. Furthermore, there are two adversarial discriminators D_X and D_Y , where the aim of D_X is to distinguish between data $\{x\}$ and translated data $\{F(y, z)\}$ and correspondingly of D_Y to distinguish between $\{y\}$ and $\{G(x, z)\}$. In the following, the terms adversarial loss [52], [53], cycle-consistency loss [54], and identity-mapping loss [64], which are elementary for the conditional CycleGAN, are described and finally summarized in the objective function.

Adversarial Loss [52], [53]: To both function adversarial losses are applied. For the mapping function $G: (X, Z) \rightarrow Y$ and its discriminator D_Y , the adversarial loss is:

$$\begin{aligned} \mathcal{L}_{GAN}(G, D_Y, X, Y, Z) &= \mathbb{E}_{y, z \sim p_{data}(y, z)} [\log D_Y(y, z)] \\ &+ \mathbb{E}_{x, z \sim p_{data}(x, z), z \sim p_{class}(z)} [\log(1 \\ &- D_Y(G(x, z), z))] \end{aligned}$$

Equation 1

where G aims to generate data $G(x, z)$ that look similar to data from domain Y , while D_Y tries to distinguish them from the real samples y . G tries to minimize this goal against an adversary D , which in turn tries to maximize it: $\min_G \max_{D_Y} \mathcal{L}_{GAN}(G, D_Y, X, Y, Z)$. The adversarial loss for the mapping function $F: (Y, Z) \rightarrow X$ and its discriminator D_X is formulated accordingly.

Cycle-Consistency Loss [54]: In addition, to reduce the space of possible mapping functions G and F , the mapping functions should be cycle-consistent. For each data x from domain X , the data translation cycle should be able to return x to the original data: $x \rightarrow G(x, z) \rightarrow F(G(x, z), z) \approx x$, that is called *forward cycle consistency*. The *backwards cycle consistency* applies accordingly to y from domain Y . To achieve cycle consistency, the following cycle-consistency loss is expressed:

$$\begin{aligned} \mathcal{L}_{CYC}(G, F) &= \mathbb{E}_{x, z \sim p_{data}(x, z), z \sim p_{data}(z), x \sim p_{data}(x)} [\|F(G(x, z), z) - x\|_1] \\ &+ \mathbb{E}_{y, z \sim p_{data}(y, z), z \sim p_{data}(z), y \sim p_{data}(y)} [\|G(F(y, z), z) - y\|_1] \end{aligned}$$

Equation 2

Identity-mapping Loss [64]: To promote the successful reproduction of the input, an identity mapping loss is formulated:

$$\begin{aligned} \mathcal{L}_{ID}(G, F) &= \mathbb{E}_{y, z \sim p_{data}(y, z), y \sim p_{data}(y)} [\|G(y, z) - y\|_1] \\ &+ \mathbb{E}_{x, z \sim p_{data}(x, z), x \sim p_{data}(x)} [\|F(x, z) - x\|_1] \end{aligned}$$

Equation 3

Full Objective: To summarize, the complete objective is:

$$\begin{aligned} \mathcal{L}_{CYC}(G, F, D_X, D_Y) &= \mathcal{L}_{GAN}(G, D_Y, X, Y) + \mathcal{L}_{GAN}(F, D_X, Y, X) \\ &+ \lambda_{CYC} \mathcal{L}_{CYC}(G, F) + \lambda_{ID} \mathcal{L}_{ID}(G, F) \end{aligned}$$

Equation 4

where λ_{CYC} and λ_{ID} control the relative importance of the two objectives, respectively.

The goal is to solve the following equation:

$$G^*, F^* = \arg \min_{G, F} \max_{D_X, D_Y} \mathcal{L}(G, F, D_X, D_Y)$$

Equation 5

b) Architecture and Training Details

The basic architecture and training details are based on the CycleGAN architecture of [54] and its implementation on GitHub (<https://github.com/junyanz/CycleGAN>). Thereby, the generator architecture is in turn based on the GAN architecture of Johnson et al. [65] and the discriminator architecture is based on PatchGANs [66]–[68]. We implemented the class conditioning according to the conditional GAN model of Isola and colleagues [66]. To implement the code, we used Tensorflow 2.9.2 [69]. In the following, we point out all differences to specifications from the originally proposed constructs and training parameters.

The specific layers used for the generator and discriminator models, including their filter and kernel sizes, are shown in Figure 3. We fitted all layers to a 1-dimensional input. Moreover, for the convolutional layers shown, we initialized the model weights with a random Gaussian with a mean of 0.00

and a standard deviation of 0.02. In addition, we used same padding.

Other specific settings not shown in Figure 3 for the discriminator and generator are noted below. For the discriminator, we used 70x1 PatchGANs according to Isola et al. [66]. The convolutional layers had a stride of 2 for all layers except the output layer, where the stride is 1. In addition, the slope of the leaky ReLU layers was set to $\alpha = 0.2$. The generator model according to Johnson et al. [65] consists of an encoder, nine consecutive residual networks (ResNet) for transformation, and a decoder. The stride of the first and last convolutional layers and ResNet convolutions is 1, while it is 2 in every other convolutional layer. For all parameters of the discriminator and generator not mentioned, we used the default values of Tensorflow (version 2.9.2).

Furthermore, 200 epochs were trained with a batch size of 64. The discriminator used the Adam solver with a constant learning rate of 0.0002. While for the generator, we adjusted the learning rates according to the pair of data. For the conditional CycleGAN between walking and running data, as suggested in the study by Zhu et al. [54], the learning rate was set to 0.0002 for the first 100 epochs and then linearly decreased to 0 for the next 100 epochs. For the conditional CycleGAN between walking and handwriting data or running and handwriting data, in the first 50 epochs the learning rate was 0.0128 ($= 0.0002 * 64$ (batch size)) in the next 50 epochs it was 0.0016 ($0.0002 * 8$ (square root of the batch size)) and in the last 100 epochs it was linearly decreasing to 0. In addition, λ_{CYC} were set to 10 and λ_{ID} to 5 in Equation 4. The trained generator with the lowest loss value over each of the 200 epochs was selected for data generation.

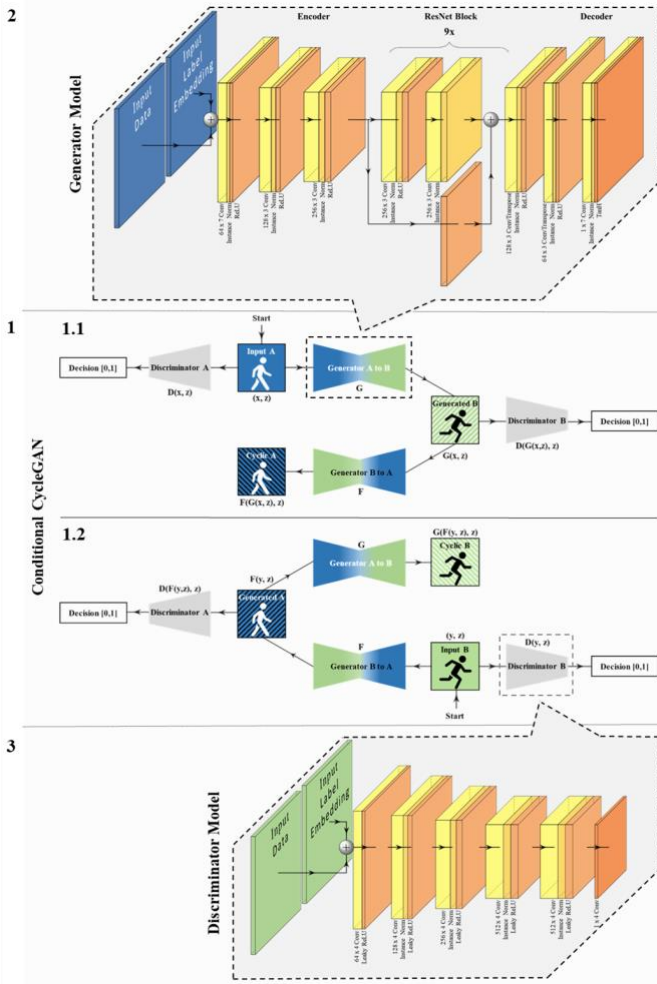


Fig. 3. Schematic structure of the conditional CycleGAN, the generator and the discriminator model. (1) The conditional CycleGAN with walking exemplary for Data A and running for Data B: In 1.1 the path is shown starting from Data A and in 1.2 starting from Data B. It is important to emphasize that the identically named generators and discriminators are the same in each case and are trained from both directions. Furthermore, the basic generator and discriminator models are the same. (2) Generator model: presented are the layers including filter and kernel size (e.g., 64 x 7 Conv = convolutional Layer with 64 filters and a kernel size of 7). (3) PatchGAN discriminator model: presented are the layers including filter and kernel size.

c) Data Classification

We collected 10,661 walking strides (627.9 ± 129.0 per person), 16,358 running strides (962.9 ± 182.0 per person), and 1,067 handwritings (59.25 ± 1.8 per person). Each vector of a walking or running stride and handwriting included 256 data points (walking/ running: 256 data points = 128 data points of left foot contact + 128 data points of the right foot contact).

Participant classifications were based on support vector machine (SVM) [70]–[73] with an extensive hyperparameter search in terms of kernel (linear, radial basis function, sigmoid, and polynomial) and cost parameter ($C = 2^{-5}, 2^{-4.75}, \dots, 2^{15}$).

We examined the performance to discriminate walking, running, and handwriting patterns between participants using a multi-class classification with 17 classes. As shown in Table 2

for the original data (and in more detail for the generated data in the appendix in Table A1), the matrices used for the classifications varied in size due to the different data sets. Therefore, prediction accuracy, F_1 -score, precision, and recall were calculated using 5-fold cross-validation. In addition, to relate the results of the person classification based on the original data, we exactly matched the training and test splits of the walking and running classification to those of the writing classification. As a baseline reference, we also calculated the zero-rule baseline (ZRB) for the respective classification problem. The ZRB reflects the theoretical accuracy obtained when the classifier always predicts the most frequent class in the training set.

The classification was performed within Python version 3.9.12 (Python Software Foundation, Wilmington, DE, United States) using the scikit-learn toolbox version 1.1.3 [74].

TABLE 2
Description of the Data of the Baseline Participant Classification by Means of SVM Based on the Original Walking, Running, and Handwriting Data.

Original Data	SVM	
	Number of Training Data	Number of Test Data
Walking	9,594 (565.1 $\pm$ 110.7) [905 (53.3 $\pm$ 1.5)]	1,067 (62.9 $\pm$ 12.3) [101 (5.9 $\pm$ 0.3)]
Running	14,722 (866.6 $\pm$ 163.9) [905 (53.3 $\pm$ 1.5)]	1,636 (96.3 $\pm$ 18.1) [101 (5.9 $\pm$ 0.3)]
Writing	905 (53.3 $\pm$ 1.5)	101 (5.9 $\pm$ 0.3)

A SVM with 5-fold cross-validation was applied. The table presents the total number of trials for both training and test sets across 17 participants, with the mean and standard deviation values shown in round parentheses. In square brackets are additionally the number of trials of the walking and running data adjusted to the number of trials of the writing data.

III. RESULTS

A. Participant Classification on Original Data

The basic assumption of this study is the existence of the distinguishability of the individual movement patterns between the persons. To test this assumption, a SVM was used to classify individuals based on walking, running, or handwriting data. As presented in Table 3, it is possible to distinguish the participants from each other in both the walking- and running data, and the handwriting data with more than 98.0% classification F_1 -score each.

TABLE 3
Classification Scores of Person Recognition Using SVM Based on Original Data.

Classification Problem	Classification Score [%]					Number of Trials	
	Acc	F_1	Prec	Rec	ZRB	Train	Test
Walking	99.7	99.7	99.7	99.7	7.5	9,594	1,067

	[98.0]	[98.0]	[98.4]	[98.0]	[6.0]	[905]	[101]
Running	99.6	99.6	99.6	99.6	6.7	14,722	1,636
Writing	99.0	99.0	99.1	99.0	[6.0]	[905]	[101]
	99.0	99.0	99.2	99.0	6.0	905	101

In squared parentheses are the results of the original walking and running classification, where the number of trials was exactly matched to that of the writing classification. ZRB = Zero-Rule Baseline, Acc = Accuracy, $F_1 = F_1$ -Score, Prec = Precision, Rec = Recall.

B. Qualitative Analysis of Generated Data

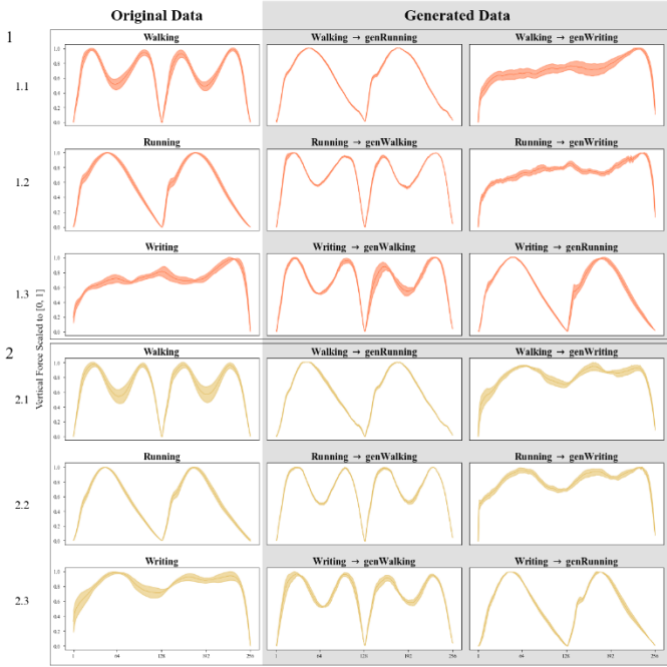


Fig. 4. Original and CycleGAN generated Data. Exemplary shown are the test set data of participant 3 (1) and participant 5 (2). Panels 1.1 and 2.1 show the original walking data and the genRunning and genHandwriting data generated from them. Accordingly, panels 1.2 and 2.2 refer to the original running data, and panels 1.3 and 2.3 to the original handwriting data each with the data generated from them. The mean values and standard deviation over the respective curves are shown in each case. On the x-axis, the respective courses over the 256 time points are shown. The y-axis shows the vertical force scaled to the interval [0, 1]. For the walking and running data, the values 1 to 128 represent the ground contact of the left foot and the values 129 to 256 of the right foot. genRunning = generated running data, genWalking = generated walking data, genWriting = generated handwriting data.

Figure 4 shows an example of the data of the participants p3 (Fig 4.1) and p5 (Fig 4.2). Looking at the shape of the curves, it is noticeable that the generated running (genRunning), generated walking (genWalking) and generated handwriting (genWriting) data correspond to the curve of the respective original data. However, based on the examples, it is noticeable that genWalking and genRunning data generated from walking or running data show a distribution with lower variance. In addition, the examples of genRunning data from walking data show a small wave at the end of the left and right ground contact, which does not occur in the original running data. In

the genWalking and genRunning data, which are based on the handwritten data, it can also be seen that there are sections within the curves that vary noticeably in variance (e.g., Fig 4.1.3: the right ground contact in each case for genWalking and genRunning). A more detailed look at the genWriting data shows that they deviate somewhat more from the original data than genRunning data and genWalking data. In addition, for example, there are also less smooth curve components (Fig 4.1.2). Furthermore, it is noticeable that the first value of the generated data tends to be too small (i.e., by 0.0) and does not match the original data (i.e., by 0.2).

Looking at the data at the individual participant level, we see that individual characteristics of the original walking, running, and handwriting data were carried over into the genWalking, genRunning, and genWriting data, respectively. For example, in both p3 and p5, the somewhat stronger impact peak of the right ground contact during running is also transferred accordingly in the respective genRunning data. In addition, in the genWalking data at p3 (Fig 4.1.2), the left ground contact shows a higher loading peak than the terminal stance peak, which is exactly the opposite in the right ground contact but corresponds to the original data in both cases. Also, in handwriting, for example, it can be seen that the pattern of p5 (Fig 4.2.3) is somewhat wavier than that of p3 (Fig 4.1.3), which is also reflected in the corresponding genWriting data.

C. Classification with Generated Test Data

TABLE 4
Classification Results of Person Recognition Using SVM with Generated Test Data.

Test Data	Generated from	Classification Score [%]					Number of Trials	
		Acc	F1	Prec	Rec	ZRB	Train	Test
genRunning	← Walking	93.6	92.5	95.9	93.6	6.7	1,636	1,067
genWriting	← Walking	52.2	46.8	49.7	52.2	6.0	101	1,067
genWalking	← Running	98.9	98.9	98.9	98.9	7.5	1,067	1,636
genWriting	← Running	53.9	50.0	50.1	53.9	6.0	101	1,636
genRunning	← Writing	71.3	78.7	75.5	86.7	6.7	1,636	101
genWalking	← Writing	73.3	78.4	73.4	88.1	7.5	1,067	101

ZRB = Zero-Rule Baseline, Acc = Accuracy, $F_1 = F_1$ -score, Prec = Precision, Rec = Recall, genWalking = Generated walking data, genRunning = Generated running data, genWriting = Generated handwriting data.

Table 4 presents the results of the person classifications with generated test data. The classification F_1 -score of the genRunnings generated from walking data and the genWalkings from running data was 92.5% and 98.9%, respectively. The classification F_1 -score of the genRunnings and genWalkings from handwriting data was 78.7% and 78.4%, respectively. For the genWriting data generated from the walking and running data, the F_1 -score was 46.8% and 50.0%, respectively. Thus, the results are 7.7 to 14.0 times better than the ZRB guess probability.

IV. DISCUSSION

A. Person classification Based on Original Data

To test the underlying assumption of individuality of the collected walking, running, and handwriting patterns in the

present study, person classifications were performed. The results of the person classifications based on the original data confirm our assumption that both vertical GRF in walking or running and vertical pen pressure in handwriting clearly differ between participants in each case (F_1 -score: running 99.6%, walking 99.7%, writing 99.0%). These results are therefore within the range of previous studies on individuality of movement patterns [20], [22], [50], [55], [56], [30], [31], [37], [39]–[41], [43], [49].

In order to be able to relate the results of the person classifications between the movements based on the original data, the data sets of the walking and running data were adjusted to the size of the data set of the writing data. Due to the reduction of the data set, the classification results of the original walking (99.7% to 98.0% F_1 -score) and running (99.6% to 99.0% F_1 -score) data slightly decreased. When comparing the classification results between the vertical GRF of walking and running and the vertical pen pressure patterns in handwriting, no differences were shown between the signals of running and those of writing. However, there is a difference, albeit very slight, between the results based on the signals of walking and those of walking or writing. A possible explanation for the somewhat more individual patterns in running compared to walking could be due to, for example, a different frequency spectrum or higher applied forces [50]. Similar to the study by Schöllhorn et al. [20], where the most extreme heel heights of shoes resulted in the highest recognition rate, running can be seen as a more extreme movement that forces the participants to show their individuality. Whereas in walking speed there are many more possibilities for compensation. However, for larger data sets, the above points seem to play a minor role in influencing person classification. One approach to explain the slightly better handwriting results compared to those of walking could be the different localization of the movement control in the central nervous system. The time normalization and accompanying possible interpolation of the vectors should have only a subordinate influence on the results of the classification [58].

B. Data Generation and Person Classification Based on Generated Data

In the first step of data analysis, transformations between the vertical GRF data of walking and running movements and the vertical pen pressure of handwriting were learned using the deep learning method CycleGAN [54] with a participant's class conditioning. Then, based on the learned transformations, the data of each of the other two movements was artificially generated from the third movement. This generated data was then tested in a person classification trained with the original data.

The results ranged from 46.8% to 98.9% F_1 -score, correspond to almost 8 times and up to 14 times the guess probability, respectively. Consequently, this provided first evidence that it is possible to learn pairwise transformations between the respective movement data on the one hand, and preservation of individual structures on the other hand.

The generation of the genRunning (92.5% F_1 -score) and genWalking (98.9% F_1 -score) data from the original walking

and running data worked particularly well. This impression can be confirmed by looking at the figures (Figure 4) of the generated data. The generated genWalking and genRunning data from the original walking and running data are not only very similar in their general shape to the original data, but also reflect the respective individual characteristics such as impact peak or time course features in their curves. However, the figures also show differences between generated and original data. For example, the variance of the generated data was shown to be significantly lower than that of the original data. This observation could be attributed to the fact that GANs aim to learn probability distributions that accurately represent the underlying data generation process [52], [53]. Thus, the generator might have tried to stay as close as possible to the learned probability distribution. In addition, curve features were also uncovered that did not appear to have a substantial effect on classification but did not match the original data. On the one hand, this provides a reason for caution for the use of the generated data, and on the other hand, it reveals optimization potential that should be addressed in future research.

To be able to explain the approximated 6% difference in the classification results of the genRunning and genWalking data, several explanatory approaches come into question. One possible approach could be due to the highly non-convex optimization process of the conditional CycleGAN, so that the generator was stuck in a local minimum, for example, or the initially learned weights negatively influenced the learning process of the genRunnings. Alternatively, it could be that the running data contains more or different information than the walking data, and it is therefore easier for the conditional CycleGAN to generate genWalking data from the running data than genRunning data from the walking data. However, if we put the classifications in relation to the guess probability, we notice that the results with 13.2 and 13.8 times the ZRB roughly correspond.

The results of the genRunning and genWalking data from handwriting data are with 78.7% and 78.4% F_1 -score and the genWriting data from running (50.0% F_1 -score) and walking (46.1% F_1 -score) data also each at least 7.7 times better than the guess probability. Thus, the results are worse than those of the genRunning and genWriting data from the walking and running data. The figures provide a first explanation in this respect. At first glance, the genRunning and genWalking data from the handwriting data and the genWriting data from the running and walking data correspond quite closely to the original data, including the adoption of curve-specific individual characteristics. The genRunning and genWalking data from the handwriting data, for example, show a relatively strong unequal distribution of the variance in certain curve segments, which does not occur in the original data. However, this could also be due to the relatively small amount of handwriting data generated, so that individual trials with a greater deviation from the mean are more significant and are shown relatively overrepresented. The genWriting data also shows a lower variance than the original handwriting data and represents the original curve in part only in rough form. A further potential explanation for the lower classification results compared to the data generated between walking and running, as well as the discrepancy between the genRunning and

genWalking data and the genWriting data derived from walking and running data, could be attributed to the quantity of handwriting data. The limited training data available for the conditional CycleGAN could result in the generator learning a probability distribution that does not accurately represent the original data. One more possible explanation is that generating handwriting is inherently more challenging than generating walking or running. Additionally, the bipedal nature of walking and running may provide additional information about the relationship between the left and right steps, which could facilitate generation, but is not present in the case of handwriting. To what extent the different central nerves control locomotion movements such as walking or running and arm or wrist movements such as writing influences the identification of individual patterns remains a subject of future research. A further possible explanation could lie in the architecture and training parameters of the conditional CycleGAN, so that parameter tuning or optimization of the architecture of the generator or discriminator (details in chapter IV.D) can achieve domain-specific improvements.

C. Identification of Individual Commonalities Across Movements

The results provide further evidence for the possibility of automatic recognition of movement patterns across movements. So far, this could only be done for very similar movements using the joint angle curves in shot put, discus and javelin throwing, taking into account all kinematic variables except that of the throwing arm [51]. With the proposed method, we can extend this approach considerably. As shown in this work, transferable individual movement features can also be found in movement data that differ significantly in their time course and originate from very different movements.

Other previous studies on the automatic identification of individual movement patterns investigated these in each case only based on single movement signals [20], [22], [50], [55], [56], [30], [31], [37], [39]–[41], [43], [49].

Even though we can provide the first cross-movement approaches in this work, we are only at the beginning of cross-movement research. We were able to provide first evidence in this study, using vertical forces and pressure data respectively, that it is possible to learn transformations between these data to generate artificial data that still preserve latent patterns of the original data. Thus, we provide a "proof of concept" of the presented method, which has the potential to represent a starting point for further research in this context.

In this respect, an extension of the study to other data signals as well as other biological data would be useful. While a transfer from kinematic data to GRF data during walking could be shown using GANs [75], there is still great potential to be exploited cross-movement wise. Whether there are also connections to other signals such as audio (e.g. voice), ECG or EEG [76] remains the subject of future research.

Furthermore, while we have found individual cross-movement commonalities, what these explicitly look like and what characteristics they exhibit should be addressed in future research. In addition to the individual movement component, it might also be possible to find further latent patterns or subcomponents across movements. Other movement

components could be the movement technique or situational adaptations to fatigue, emotions, environment, or the like.

The individual patterns across movements also suggest that there might be characteristic features of a person that are at least reflected in several movements. Whether a link can be established between movement and psychological characteristics of a person could also be the subject of future work. In addition, it remains to be investigated whether individual movement patterns behave similarly and are consistent with patterns found in behavioral research [77]–[80]. This could be of specific interest for economizing training or therapy. Whether changing the gait by training or therapy, which is sometimes observed after psychotherapy, and whether this has an effect on the handwriting or vice versa could be one area of a more holistic approach to future practical applications.

D. Data Generation and Cross-Movement Analysis by Conditional CycleGANs

The methodology presented in this paper has overcome the problem of mapping two biomechanical time series signals to each other while transferring the individual component. It was thus possible to learn the transfer of movement A to movement B while obtaining latent patterns of movement A. This could provide fundamental new opportunities in future experiments where one is looking for latent structures between movements or movement-signals. In the following, we discuss optimization potentials and application possibilities of conditional CycleGANs.

First, it must be emphasized restrictively, that in the context of this study, the person condition of the CycleGAN was necessary to learn the transformation including the preservation of the individual component from movement A to movement B. Specifically, this means that it was not possible, for example, to generate a person's individual handwriting data from the walking data without prior knowledge of that person's handwriting.

For the conditional CycleGANs, the amount of walking (approximately 10,000 steps), running (approximately 14,000 steps), and handwriting data (approximately 900 trials) appears to be sufficient to learn transformations between movements based on vertical ground reaction forces or vertical pen pressure. However, the results suggest that the quality of the learned transformation depends largely on the amount of data. Such that possibly the 900 handwritten trials were too few to produce deceptively real trials. While the amount of walking and running data provided a good basis to perform person classification with very high recognition rates, the generated curves also showed "errors". In order to use data for generation, domain-specific optimizations would need to be made [81] so that the generated data are not only indistinguishable using a classifier but make biological sense and are indistinguishable from original data by experts. Apart from a larger data set, we see the greatest optimization potential in adjusting the learning rate of the generator's optimizer (especially in relation to the learning rate of the discriminator). In addition, the number of epochs, or the selected batch size also seem to provide potential for domain-specific optimization.

Particular potential of the conditional CycleGAN lies in the possibility cross-movement analysis due to the learnability of transformations from one movement signal to another

movement signal. In addition, as in the setting of the present work, the conditional CycleGAN could be used to identify further latent patterns across movements. In doing so, it would also be possible to search for cross-person abstract patterns of for example fatigue, emotion, or illness. Another potential of the conditional CycleGAN could be especially in an area where it is difficult to collect large amounts of data. There, artificial augmentation of small data sets could open up the possibility of using data-intensive methods (e.g., deep learning approaches, machine-learning classification). The advantage of the conditional CycleGAN in this context is that this algorithm requires relatively little data, and that the training data need not be paired, or the construction of artificial pairs is obsolete.

V. CONCLUSION

In recent years, the analysis of movement patterns has increasingly focused on the individuality of movements, revealing individual patterns with situation-dependent fine structures. However, previous research methods only allowed the comparison of very similar movement signals. In this study, we were able to identify similarities between individual walking, running, and handwriting patterns across different movements through data augmentation, revealing individual patterns across movements. This further extends the understanding of strong individuality.

Based on the results of movement science studies that use machine learning methods to investigate the uniqueness of individual movement patterns, and the findings presented in this study, it can be inferred that our understanding of the individuality of human movement and the influence of individuality on targeted development, improvement, or recovery is still in its beginning stages. Understanding individual cross-movement commonalities in movement may offer insights into the underlying more general individuality of central nervous system physiology and structure. Future applications of this approach have the potential to investigate the extent to which the central nervous system or muscle physiology can be altered beyond the individual domain.

In addition, this study provides proof of concept that it is possible to use the conditional CycleGAN to artificially generate cross-movement data with latent movement characteristics of the original movement without relying on paired data. In summary, the methodology presented in this study helps to enable cross-movement analysis and artificially generate larger data sets.

APPENDIX

TABLE A1

Description of the Data Sizes in the Whole Analysis Process Depending on the Different Classification Tasks.

Dependent on the Different Classification Task					
Classification Problem		Cond. CycleGAN		SVM	
Generated Test Data	Generated from Original Data	Number of Training Data	Number of Trials		
			Train (Original Data)	Test (Generated Data)	
genWalking	Running	Walking	9,594	1,067	1,636
		Running	14,722	(62.9 ± 12.3)	(96.3 ± 18.1)

genRunning	Walking	Running	14,722	1,636	1,067
		Walking	9,594	(96.3 ± 18.1)	(62.9 ± 12.3)
genWalking	Writing	Walking	9,594	1,067	101
		Writing	905	(62.9 ± 12.3)	(5.9 ± 0.25)
genWriting	Walking	Writing	905	101	1,067
		Walking	9,594	(5.9 ± 0.25)	(62.9 ± 12.3)
genRunning	Writing	Running	14,722	1,636	101
		Writing	905	(96.3 ± 18.1)	(5.9 ± 0.25)
genWriting	Running	Writing	905	101	1,636
		Running	14,722	(5.9 ± 0.25)	(96.3 ± 18.1)

A SVM with 10-fold Cross-Validation for original walking and running data and 5-fold for original writing data was applied. The table presents the total number of trials for both the conditional CycleGAN training and the SVM training and test sets across 17 participants, with the mean and standard deviation values shown in parentheses.

genRunning = generated running data, genWalking = generated walking data, genWriting = generated handwriting data.

ACKNOWLEDGMENT

None

- [1] B. T. Bates, "Single-subject methodology: An alternative approach," *Med. Sci. Sports Exerc.*, vol. 28, no. 5, pp. 631–638, 1996.
- [2] W. I. Schöllhorn, *Biomechanische Einzelfallanalyse im Diskuswurf: Prozess und Produktorientierte Technikanalyse mechanischer Energieformen [Biomechanical single case analysis in discus throwing: process and product oriented technique analysis of mechanical energy forms]*. Frankfurt am Main: Harri Deutsch, 1993.
- [3] I. S. Chan and G. S. Ginsburg, "Personalized Medicine: Progress and Promise," *Annu. Rev. Genomics Hum. Genet.*, vol. 12, no. 1, pp. 217–244, Sep. 2011.
- [4] G. S. Ginsburg and H. F. Willard, "Genomic and personalized medicine: foundations and applications," *Transl. Res.*, vol. 154, no. 6, pp. 277–287, Dec. 2009.
- [5] P. C. Ng, S. S. Murray, S. Levy, and J. C. Venter, "An agenda for personalized medicine," *Nature*, vol. 461, no. 7265, pp. 724–726, Oct. 2009.
- [6] T. W. Buford, M. D. Roberts, and T. S. Church, "Toward Exercise as Personalized Medicine," *Sport. Med.*, vol. 43, no. 3, pp. 157–165, Mar. 2013.
- [7] R. Nuzzo, "Scientific method: Statistical errors," *Nature*, vol. 506, no. 7487, pp. 150–152, Feb. 2014.
- [8] K. Davids, C. Button, and S. Bennett, "Modeling Human Motor Systems in Nonlinear Dynamics: Intentionality and Discrete Movement Behaviors," *Nonlinear Dyn. Psychol. Life Sci.* 1999 31, vol. 3, no. 1, pp. 3–30, 1999.
- [9] C. Button, K. Davids, S. . Bennett, and M. . Taylor, "Mechanical perturbation of the wrist during one-handed catching," *Acta Psychol. (Amst.)*, vol. 105, no. 1, pp. 9–30, Sep. 2000.
- [10] G. Schöner, P. G. Zanone, and J. A. S. Kelso, "Learning as Change of Coordination Dynamics: Theory and Experiment," *J. Mot. Behav.*, vol. 24, no. 1, pp. 29–48, Mar. 1992.
- [11] V. Barth et al., "Individualized Monitoring of Muscle Recovery in Elite Badminton," *Front. Physiol.*, vol. 10, Jun. 2019.
- [12] A. Hecksteden, J. Kraushaar, F. Scharhag-Rosenberger, D. Theisen, S. Senn, and T. Meyer, "Individual response to exercise training - a statistical perspective," *J. Appl. Physiol.*, vol. 118, no. 12, pp. 1450–1459, Jun. 2015.
- [13] A. K. Jain, R. M. Bolle, and S. Pankani, *Biometrics: Personal identification in network society*. New York, NY: Springer, 2006.
- [14] D. H. Kaye, "Probability, individualization, and uniqueness in forensic science evidence," *Brooklyn Law Rev.*, vol. 75, no. 4, pp. 1163–1185, 2010.
- [15] J. E. Cutting and L. T. Kozlowski, "Recognizing friends by their walk: Gait perception without familiarity cues," *Bull. Psychon. Soc.*, vol. 9, no. 5, pp. 353–356, May 1977.
- [16] N. F. Troje, "Decomposing biological motion: A framework for analysis and synthesis of human gait patterns," *J. Vis.*, vol. 2, no. 5, pp. 2–2, Sep. 2002.
- [17] G. Johansson, "Visual perception of biological motion and a model for its analysis," *Percept. Psychophys.*, vol. 14, no. 2, pp. 201–211, Jun. 1973.
- [18] C. M. Windle, S. M. Gregory, and S. J. Dixon, "The shock attenuation characteristics of four different insoles when worn in a military boot during running and marching," *Gait Posture*, vol. 9, no. 1, pp. 31–37, Mar. 1999.
- [19] W. I. Schöllhorn, D. J. Stefanyshyn, B. M. Nigg, and W. Liu, "Recognition of individual walking patterns by means of artificial neural nets," *Gait Posture*, vol. 10, no. 1, pp. 85–86, Sep. 1999.
- [20] W. I. Schöllhorn, B. M. Nigg, D. J. Stefanyshyn, and W. Liu, "Identification of individual walking patterns using time discrete and time continuous data sets," *Gait Posture*, vol. 15, no. 2, pp. 180–186, 2002.
- [21] W. I. Schöllhorn and H. U. Bauer, "Identifying individual movement styles in high performance sports by means of self-organizing Kohonen maps," in *XVI Annual Conference of the International Society for Biomechanics in Sport*, 1998, pp. 574–577.
- [22] H. U. Bauer and W. I. Schöllhorn, "Self-organizing maps for the analysis of complex movement patterns," *Neural Process. Lett.*, vol. 5, pp. 193–199, 1997.
- [23] R. K. Begg and J. Kamruzzaman, "A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data," *J. Biomech.*, vol. 38, no. 3, pp. 401–408, Mar. 2005.
- [24] R. K. Begg, M. Palaniswami, and B. Owen, "Support Vector Machines for Automated Gait Classification," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 5, pp. 828–838, May 2005.
- [25] W. I. Schöllhorn, F. Röber, T. Jaitner, W. Hellstern, and W. Käubler, "Discrete and continuous effects of traditional and differential training in sprint running," in *Perspectives and Profiles 6th European College on Sports Science Congress*, 2001, p. 331.
- [26] F. Hoitz, V. von Tscharner, J. Baltich, and B. M. Nigg, "Individuality decoded by running patterns: Movement characteristics that determine the uniqueness of human running," *PLoS One*, vol. 16, no. 4, pp. e0249657–e0249657, Apr. 2021.
- [27] C. Maurer, P. Federolf, V. von Tscharner, L. Stirling, and B. M. Nigg, "Discrimination of gender-, speed-, and shoe-dependent movement patterns in runners using full-body kinematics," *Gait Posture*, vol. 36, no. 1, pp. 40–45, May 2012.
- [28] S. Hoerzer, V. von Tscharner, C. Jacob, and B. M. Nigg, "Defining functional groups based on running kinematics using Self-Organizing Maps and Support Vector Machines," *J. Biomech.*, vol. 48, no. 10, pp. 2072–2079, Jul. 2015.
- [29] W. I. Schöllhorn and H. U. Bauer, "Assessment of running patterns using neural networks," in *Sport und Informatik*, J. Mester and J. Perl, Eds. Cologne: Sport & Buch Strauss, 1998, pp. 208–309.
- [30] W. I. Schöllhorn, C. Peham, T. Licka, and M. Scheidl, "A pattern recognition approach for the quantification of horse and rider interactions," *Equine Vet. J.*, vol. 38, no. S36, pp. 400–405, Aug. 2006.
- [31] S. Albrecht, D. Janssen, E. Quarz, K. M. Newell, and

- W. I. Schöllhorn, "Individuality of movements in music - Finger and body movements during playing of the flute," *Hum. Mov. Sci.*, vol. 35, pp. 131–144, Jun. 2014.
- [32] N. Westphal and W. I. Schöllhorn, "Identifying Volleyballteams by their tactical moves," in *6th annual congress of the European College of Sport Science*, 2001, p. 55.
- [33] J. M. Jäger and W. I. Schöllhorn, "Situation-orientated recognition of tactical patterns in volleyball," <https://doi.org/10.1080/02640410701287230>, vol. 25, no. 12, pp. 1345–1353, Oct. 2007.
- [34] R. Rein and D. Memmert, "Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science," *Springerplus*, vol. 5, no. 1, Dec. 2016.
- [35] A. Grunz, D. Memmert, and J. Perl, "Tactical pattern recognition in soccer games by means of special self-organizing maps," *Hum. Mov. Sci.*, vol. 31, no. 2, pp. 334–343, Apr. 2012.
- [36] M. Kempe, A. Grunz, and D. Memmert, "Detecting tactical patterns in basketball: Comparison of merge self-organising maps and dynamic controlled neural networks," <http://dx.doi.org/10.1080/17461391.2014.933882>, vol. 15, no. 4, pp. 249–255, May 2014.
- [37] A. Schmidt, "Movement pattern recognition in basketball free-throw shooting," *Hum. Mov. Sci.*, vol. 31, no. 2, pp. 360–382, Apr. 2012.
- [38] R. Kaur, R. W. Motl, R. Sowers, and M. E. Hernandez, "A Vision-Based Framework for Predicting Multiple Sclerosis and Parkinson's Disease Gait Dysfunctions—A Deep Learning Approach," *IEEE J. Biomed. Heal. Informatics*, vol. 27, no. 1, pp. 190–201, Jan. 2023.
- [39] F. Horst et al., "Daily changes of individual gait patterns identified by means of support vector machines," *Gait Posture*, vol. 49, pp. 309–314, Sep. 2016.
- [40] F. Horst, A. Eekhoff, K. M. Newell, and W. I. Schöllhorn, "Intra-individual gait patterns across different time-scales as revealed by means of a supervised learning model using kernel-based discriminant regression," *PLoS One*, vol. 12, no. 6, 2017.
- [41] F. Horst, M. Mildner, and W. I. Schöllhorn, "One-year persistence of individual gait patterns identified in a follow-up study – A call for individualised diagnose and therapy," *Gait Posture*, vol. 58, pp. 476–480, 2017.
- [42] T. Jacques, R. Bini, and A. Arndt, "Inter-limb differences in in-vivo tendon behavior, kinematics, kinetics and muscle activation during running," *J. Biomech.*, vol. 141, p. 111209, Aug. 2022.
- [43] W. I. Schöllhorn, H. Schaper, S. Kimmeskamp, and T. Milani, "Inter- and intra-individual differentiation of dynamic foot pressure patterns by means of artificial neural nets," *Gait Posture*, vol. 16, no. Supplement 1, pp. 172–173, 2002.
- [44] T. Jaitner, L. Mendoza, and W. I. Schöllhorn, "Analysis of the long jump technique in the transition from approach to takeoff based on time-continuous kinematic data," *Eur. J. Sport Sci.*, vol. 1, no. 5, pp. 1–12, 2001.
- [45] J. Aeles, F. Horst, S. Lapuschkin, L. Lacourpaille, and F. Hug, "Revealing the unique features of each individual's muscle activation signatures," *J. R. Soc. Interface*, vol. 18, no. 174, Jan. 2021.
- [46] H. S. Lee et al., "Individual finger movement decoding using a novel ultra-high-density electroencephalography-based brain-computer interface system," *Front. Neurosci.*, vol. 16, Oct. 2022.
- [47] K. Liao, R. Xiao, J. Gonzalez, and L. Ding, "Decoding Individual Finger Movements from One Hand Using Human EEG Signals," *PLoS One*, vol. 9, no. 1, p. e85192, Jan. 2014.
- [48] I. Hammad and K. El-Sankary, "Using machine learning for person identification through physical activities," *Proc. - IEEE Int. Symp. Circuits Syst.*, vol. 2020-October, 2020.
- [49] D. Janssen, W. I. Schöllhorn, J. Lubienetzki, K. Fölling, H. Kokenge, and K. Davids, "Recognition of emotions in gait patterns by means of artificial neural nets," *J. Nonverbal Behav.*, vol. 32, no. 2, pp. 79–92, 2008.
- [50] J. Burdack, F. Horst, D. Aragonés, A. Eekhoff, and W. I. Schöllhorn, "Fatigue-Related and Timescale-Dependent Changes in Individual Movement Patterns Identified Using Support Vector Machine," *Front. Psychol.*, vol. 11, p. 2273, Sep. 2020.
- [51] F. Horst, D. Janssen, H. Beckmann, and W. I. Schöllhorn, "Can Individual Movement Characteristics Across Different Throwing Disciplines Be Identified in High-Performance Decathletes?," *Front. Psychol.*, vol. 11, p. 2262, Sep. 2020.
- [52] I. Goodfellow et al., "Generative adversarial nets," in *Advances in Neural Information Processing Systems* 27, 2014, pp. 2672–2680.
- [53] I. Goodfellow et al., "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020.
- [54] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2242–2251.
- [55] F. Horst, S. Lapuschkin, W. Samek, K. R. Müller, and W. I. Schöllhorn, "Explaining the unique nature of individual gait patterns with deep learning," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019.
- [56] D. Janssen, W. I. Schöllhorn, K. M. Newell, J. M. Jäger, F. Rost, and K. Vehof, "Diagnosing fatigue in gait patterns by support vector machines and self-organizing maps," *Hum. Mov. Sci.*, vol. 30, no. 5, pp. 966–975, Oct. 2011.
- [57] U. Dal, T. Erdogan, B. Resitoglu, and H. Beydagi, "Determination of preferred walking speed on treadmill may lead to high oxygen cost on treadmill walking," *Gait Posture*, vol. 31, no. 3, pp. 366–369,

- Mar. 2010.
- [58] J. Burdack, F. Horst, S. Giesselbach, I. Hassan, S. Daffner, and W. I. Schöllhorn, “Systematic Comparison of the Influence of Different Data Preprocessing Methods on the Performance of Gait Classifications Using Machine Learning,” *Front. Bioeng. Biotechnol.*, vol. 8, p. 260, Apr. 2020.
- [59] T. Chau, “A review of analytical techniques for gait data. Part 1: fuzzy, statistical and fractal methods,” *Gait Posture*, vol. 13, no. 1, pp. 49–66, Feb. 2001.
- [60] T. Chau, “A review of analytical techniques for gait data. Part 2: neural network and wavelet methods,” *Gait Posture*, vol. 13, no. 2, pp. 102–120, Apr. 2001.
- [61] C. Hsu, C. Chang, and C. Lin, “A practical guide to support vector classification,” 2003. [Online]. Available: <http://www.csie.ntu.edu.tw/~cjlin>. [Accessed: 15-Jan-2019].
- [62] A. Savitzky and M. J. E. Golay, “Smoothing and Differentiation of Data by Simplified Least Squares Procedures,” *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964.
- [63] R. W. Schafer, “What is a savitzky-golay filter?,” *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111–117, 2011.
- [64] Y. Taigman, A. Polyak, and L. Wolf, “Unsupervised cross-domain image generation,” in *In ICLR*, 2017.
- [65] J. Johnson, A. Alahi, and L. Fei-Fei, “Perceptual Losses for Real-Time Style Transfer and Super-Resolution,” 2016, pp. 694–711.
- [66] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, and B. A. Research, “Image-To-Image Translation With Conditional Adversarial Networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1125–1134.
- [67] C. Li and M. Wand, “Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks,” in *Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science()*, vol 9907, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Springer, Cham., 2016, pp. 702–716.
- [68] C. Ledig et al., “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network,” in *In CVPR*, 2017.
- [69] M. Abadi et al., “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015. [Online]. Available: <https://www.tensorflow.org/>.
- [70] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A Training Algorithm for Optimal Margin Classifier,” *Proc. Fifth Annu. ACM Work. Comput. Learn. Theory*, vol. 5, 1996.
- [71] C. Cortes and V. N. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [72] K. R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf, “An introduction to kernel-based learning algorithms,” in *Handbook of Neural Network Signal Processing*, vol. 12, no. 2, 2001, pp. 181–201.
- [73] B. Schölkopf and A. J. Smola, *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT Press, 2002.
- [74] F. Pedregosa et al., “Scikit-learn: Machine learning in python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
- [75] M. Bicer, A. T. M. Phillips, A. Melis, A. H. McGregor, and L. Modenese, “Generative deep learning applied to biomechanics: A new augmentation technique for motion capture datasets,” *J. Biomech.*, vol. 144, p. 111301, Nov. 2022.
- [76] A. T. John, A. Barthel, J. Wind, N. Rizzi, and W. I. Schöllhorn, “Acute Effects of Various Movement Noise in Differential Learning of Rope Skipping on Brain and Heart Recovery Analyzed by Means of Multiscale Fuzzy Measure Entropy,” *Front. Behav. Neurosci.*, vol. 16, p. 20, Feb. 2022.
- [77] A. H. Buss and R. Plomin, *A temperament theory of personality development*. Wiley-Interscience, 1975.
- [78] R. A. Sherman, C. S. Nave, and D. C. Funder, “Situational similarity and personality predict behavioral consistency,” *J. Pers. Soc. Psychol.*, vol. 99, no. 2, pp. 330–343, Aug. 2010.
- [79] D. C. Funder and C. R. Colvin, “Explorations in behavioral consistency: Properties of persons, situations, and behaviors,” *J. Pers. Soc. Psychol.*, vol. 60, no. 5, pp. 773–794, 1991.
- [80] J. Marcus and J. Roy, “Values, Personality, and Sustainability Behaviour: An Integrative Analysis,” *Acad. Manag. Proc.*, vol. 2017, no. 1, p. 13105, Aug. 2017.
- [81] D. Saxena and J. Cao, “Generative Adversarial Networks (GANs),” *ACM Comput. Surv.*, vol. 54, no. 3, pp. 1–42, Apr. 2022.