



OPEN ACCESS

EDITED BY

Navrag B. Singh,
ETH Zürich, Switzerland

REVIEWED BY

Simone Tassani,
Pompeu Fabra University, Spain
Nicola Francesco Lopomo,
University of Brescia, Italy

*CORRESPONDENCE

Johannes Burdack,
✉ burdack@uni-mainz.de

RECEIVED 11 April 2023

ACCEPTED 25 July 2023

PUBLISHED 04 August 2023

CITATION

Burdack J, Giesselbach S, Simak ML,
Ndiaye ML, Marquardt C and
Schöllhorn WI (2023), Identifying
underlying individuality across running,
walking, and handwriting patterns with
conditional cycle-consistent generative
adversarial networks.
Front. Bioeng. Biotechnol. 11:1204115.
doi: 10.3389/fbioe.2023.1204115

COPYRIGHT

© 2023 Burdack, Giesselbach, Simak,
Ndiaye, Marquardt and Schöllhorn. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Identifying underlying individuality across running, walking, and handwriting patterns with conditional cycle-consistent generative adversarial networks

Johannes Burdack^{1*}, Sven Giesselbach^{2,3}, Marvin L. Simak¹,
Mamadou L. Ndiaye¹, Christian Marquardt⁴ and
Wolfgang I. Schöllhorn¹

¹Department of Training and Movement Science, Institute of Sport Science, Johannes Gutenberg-University, Mainz, Germany, ²Knowledge Discovery, Fraunhofer-Institute for Intelligent Analysis and Information Systems, Sankt Augustin, Germany, ³Lamarr Institute for Machine Learning and Artificial Intelligence, Sankt Augustin, Germany, ⁴Science&Motion GmbH, Munich, Germany

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 the 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. A total of 17 healthy adults (35.8 ± 11.1 years, eight women and nine men) each performed 627.9 ± 129.0 walking strides, 962.9 ± 182.0 running strides, and 59.25 ± 1.8 handwritings. Using the conditional cycle-consistent generative adversarial network (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, 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 was tested. 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.

KEYWORDS

cross-movement individuality, cross-signal individuality, CycleGAN, data augmentation, deep learning, generative adversarial network, movement pattern recognition, support vector machine

1 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 increasingly been asked for (Schöllhorn, 1993; Bates, 1996). Often, a mixture of colloquial (weak) and science-oriented (strong) understanding can be observed, which is even more confusingly equated with “personalized” (Ginsburg and Willard, 2009; Ng et al., 2009; Chan and Ginsburg, 2011; Buford et al., 2013). Colloquially, weak individuality often serves as an excuse for a lack of statistically significant group differences (Davids et al., 1999; Button et al., 2000; Nuzzo, 2014) or for not finding commonalities across individuals (Schöner et al., 1992; Button et al., 2000; Hecksteden et al., 2015; Barth et al., 2019). 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 (Jain et al., 2006; Kaye, 2010), 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 (Johansson, 1973; Cutting and Kozlowski, 1977; Troje, 2002), followed by biomechanical analyses of gait movements (Schöllhorn et al., 1999; Windle et al., 1999; Schöllhorn et al., 2002) and sporadically single sports movements (Bauer and Schöllhorn, 1997; Schöllhorn and Bauer, 1998b), analyses of a wide variety of movements have become increasingly popular. Besides walking (Schöllhorn et al., 2002a; Begg et al., 2005; Begg and Kamruzzaman, 2005), the individuality of movements has also been shown in the field of sports in sprinting (Schöllhorn et al., 2001), running (Schöllhorn and Bauer, 1998a; Maurer et al., 2012; Hoerzer et al., 2015; Hoitz et al., 2021), javelin (Schöllhorn and Bauer, 1998b) and discus throwing (Bauer and Schöllhorn, 1997), and horseback riding (Schöllhorn et al., 2006), as well as in the field of music when playing the flute (Albrecht et al., 2014). Similar features could be shown for team behavior in volleyball (Westphal and Schöllhorn, 2001; Jäger and Schöllhorn, 2007), soccer (Grunz et al., 2012; Rein and Memmert, 2016), or basketball (Schmidt, 2012; Kempe et al., 2014). Typical parameters investigated are biomechanical data from video recordings (Kaur et al., 2023), force plates (Horst et al., 2016; Horst et al., 2017a; Horst et al., 2017b), pressure insoles (Schöllhorn et al., 2002b; Jacques et al., 2022), EMG (Jaitner et al., 2001; Aeles et al., 2021), and brain signals (Liao et al., 2014; Lee et al., 2022). Besides these, wearable sensors are becoming increasingly popular (Hammad and El-Sankary, 2020). Situated “perturbations” such as emotions (Janssen et al., 2008), fatigue (Burdack et al., 2020a), or time alone (Schöllhorn and Bauer, 1998b; Horst et al., 2016; Horst et al., 2017a; Horst et al., 2017b) were not able to move the patterns out of the strong individual space. Consequently, robust evidence for an important role of individuality is provided in short-term adaptive behavior. What all studies listed so far have in common is 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 or learning the same movement by the same person never encounters comparable conditions again due to cognitive and body memory (Horst et al., 2020). To solve this problem, finding commonalities in learning 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, and javelin) with similar kinematic structure was provided (Horst et al., 2020). The aim of this study is to identify individual commonalities of movement forms with different kinematic structures.

The problem with this is that the classification models cannot transfer among 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 the potential to address this problem. Approaches from image generation offer solutions for analog problems. In image-to-image translation or style transfer, it is possible to learn a relationship between images from two domains, domain 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 landscape as it might look in summer or in winter.

In the area of data generation, Generative Adversarial Networks (GANs) (Goodfellow et al., 2014; Goodfellow et al., 2020) 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 the 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 (Zhu et al., 2017). 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 (Horst et al., 2019), finding matching pairs of movements between two movements can be problematic.

A further development of the GAN that circumvents the pairwise data problem that respectively offers a possible solution and can be successfully applied on relatively small data sets is the cycle-consistent GAN (CycleGAN) (Zhu et al., 2017). 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 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

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.

use CycleGAN to identify common individual patterns across different movements. Specifically, we aim to find individual commonalities underlying the 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 movement A. Specifically, we generate the other two movements from the walking, running, and handwriting movements (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 (Bauer and Schöllhorn, 1997; Schöllhorn et al., 2002a; Schöllhorn et al., 2002b; Schöllhorn et al., 2006; Janssen et al., 2008; Janssen et al., 2011; Schmidt, 2012; Albrecht et al., 2014; Horst et al., 2016; Horst et al., 2017a; Horst et al., 2017b; Horst et al., 2019; Burdack et al., 2020a), 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?

2 Materials and methods

2.1 Participants and ethics statement

The study participants were 17 athletically active, healthy adults (eight women and nine men; 1 left-handed and 16 right-handed) who regularly handwrote and ran 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.

2.2 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 (Dal et al., 2010). The PWS is a speed that the participants prefer in their leisure time, for example, when going for a walk, and the PRS is 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.

As presented in Figure 1, the participants performed six sets of 4 min of running or walking. Each was followed by a 4-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.

2.3 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).



FIGURE 1

Experimental sequence with the 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.

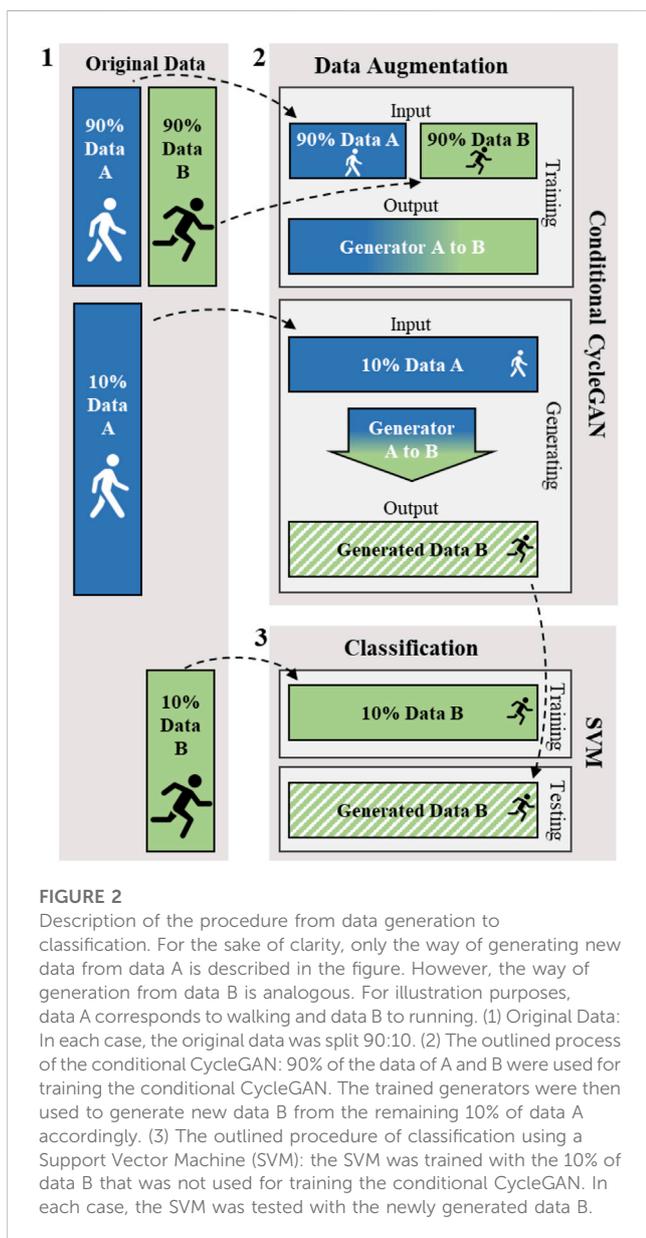


FIGURE 2

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.

2.4 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. (2020b) 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] (Chau, 2001a; Chau, 2001b; Hsu et al., 2003). 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 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 zero 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 (Savitzky and Golay, 1964), which smoothens on the least squares method while maintaining the shape and height of the waveform peaks (Schafer, 2011). 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.

2.5 Data analysis

2.5.1 Data analysis procedure

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.

2.5.2 Data generating–conditional CycleGAN

With the CycleGAN (Zhu et al., 2017), 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.

2.5.2.1 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 $\{y_j\}_{j=1}^M$ where $y_j \in Y$, and the class embedding $\{z_k\}_{k=1}^P$ where $z_k \in Z$. For simplicity, the indices i, j , and k 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 for D_Y to distinguish between $\{y\}$ and $\{G(x, z)\}$. In the following, the terms adversarial loss (Goodfellow et al., 2014; Goodfellow et al., 2020), cycle-consistency loss (Zhu et al., 2017), and identity-mapping loss (Taigman et al., 2017), which are elementary for the conditional CycleGAN, are described and finally summarized in the objective function.

Adversarial Loss (Goodfellow et al., 2014; Goodfellow et al., 2020): To both functions, adversarial losses are applied. For the mapping function $G: (X, Z) \rightarrow Y$ and its discriminator D_Y , the adversarial loss is:

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

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 (Zhu et al., 2017): 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: $x \rightarrow G(x, z) \rightarrow F(G(x, z), z) \approx x$, which 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:

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

Identity-mapping Loss (Taigman et al., 2017): To promote the successful reproduction of the input, an identity-mapping loss is formulated:

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

Full Objective: To summarize, the complete objective is:

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

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

The goal is to solve the following equation:

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

2.5.2.2 Architecture and training details

The basic architecture and training details are based on the CycleGAN architecture of Zhu et al. (Zhu et al., 2017) 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. (Johnson et al., 2016), and the discriminator architecture is based on PatchGANs (Li and Wand, 2016; Isola et al., 2017; Ledig et al., 2017). We implemented the class conditioning according to the conditional GAN model of Isola and colleagues (Isola et al., 2017). To implement the code, we used Tensorflow 2.9.2 (Abadi et al., 2015). In the following, we point out all differences in 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 one-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 the same padding.

Other specific settings not shown in Figure 3 for the discriminator and generator are noted below. For the discriminator, we used 70×1 PatchGANs according to Isola et al. (2017). The convolutional layers had a stride of two for all layers except the output layer, where the stride is one. In addition, the slope of the leaky ReLU layers was set to $\alpha = 0.2$. The generator model according to Johnson et al. (2016) 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 one, while it is two 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

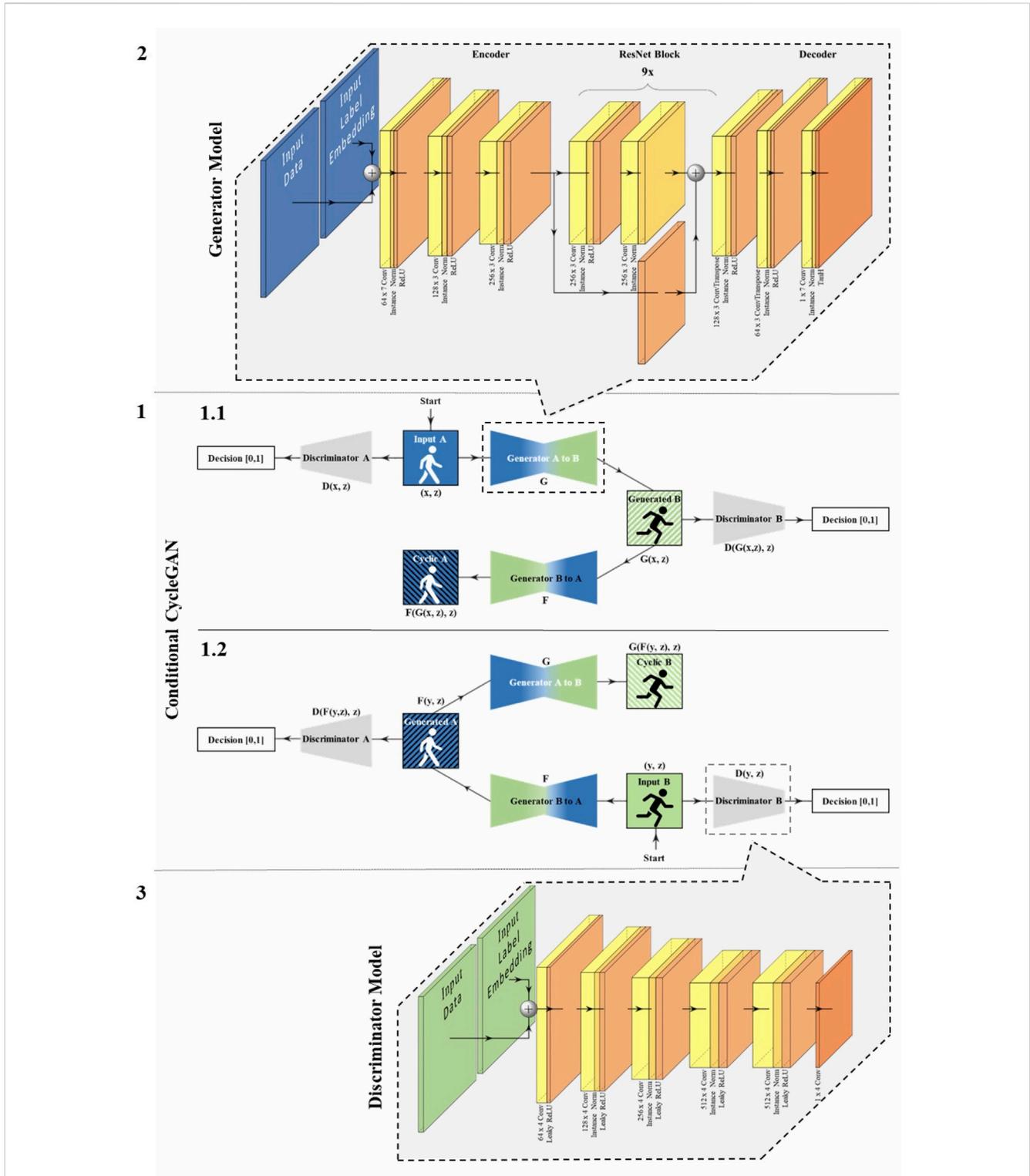


FIGURE 3 Schematic structure of the conditional CycleGAN, the generator, and the discriminator model. (1) The conditional CycleGAN with walking exemplary for Data A and running for Data B: In 1.1, the path is shown starting from Data A, and in 1.2 starting from Data B. It is important to emphasize that 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 × 7 Conv = convolutional Layer with 64 filters and a kernel size of seven). (3) PatchGAN discriminator model: presented are the layers including filter and kernel size.

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 ± 110.7) [905 (53.3 ± 1.5)]	1,067 (62.9 ± 12.3) [101 (5.9 ± 0.3)]
Running	14,722 (866.6 ± 163.9) [905 (53.3 ± 1.5)]	1,636 (96.3 ± 18.1) [101 (5.9 ± 0.3)]
Writing	905 (53.3 ± 1.5)	101 (5.9 ± 0.3)

An SVM with five-fold cross-validation was applied. The table shows the total number of trials for the training and test data from all 17 participants. The mean amount and standard deviations of the trials of each participant are shown in round brackets. In square brackets, the number of trials of the walking and running data adjusted for the number of trials of the writing data is also shown.

according to the pair of data. For the conditional CycleGAN between walking and running data, as suggested in the study by Zhu et al. (2017), 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} was 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.

2.5.2.3 Data classification

Since there are no theoretical or practical empirical values for the time series-based method used in this study in terms of the minimum amount of data, we roughly followed the sizes of the image datasets used in the CycleGAN article (Zhu et al., 2017). 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 walking or running stride and handwriting included 256 data points (walking/running: 256 data points = 128 data points of left food contact + 128 data points of the right food contact).

Participant classifications were based on support vector machine (SVM) (Cortes and Vapnik, 1995; Boser et al., 1996; Müller et al., 2001; Scholkopf and Smola, 2002) 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, where each participant represented one class. As shown in Table 2 for the original data (and in more detail for the generated data in the Supplementary Table S1), the amount of training and testing data used for the classifications varied in size due to the different data sets. In addition, to relate the results of the participant 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.

To evaluate the results of the multi-class classifications, the performance indicators accuracy, F_1 -score, precision, and recall were calculated after five-fold cross-validation. The number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) define these metrics:

$$\begin{aligned} \text{Accuracy} &= \frac{TP + TN}{TP + TN + FP + FN} \\ \text{Precision} &= \frac{TP}{TP + FP} \\ \text{Recall} &= \frac{TP}{TP + FN} \\ F_1\text{-score} &= 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned}$$

The baseline reference is the zero-rule baseline (ZRB), which results from the theoretical accuracy when the classifier always predicts the most frequented class of the training set. Since our dataset is unbalanced, the ZRB is calculated by dividing the number of training trials of the most frequent class by the number of all training trials in the corresponding classification task:

$$\text{ZRB} = \frac{\text{Training Trials Most Frequent Class}}{\text{All Training Trials}}$$

For a detailed overview of the calculated ZRB for each classification task, see Supplementary Table S2.

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 (Pedregosa et al., 2011).

3 Results

3.1 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, an 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.

3.2 Qualitative analysis of generated data

Figure 4 shows an example of the data of the participants p3 (Figure 4.1) and p5 (Figure 4.2). Looking at the shape of the curves, it is noticeable that the generated running (genRunning), generated

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 [98.0]	99.7 [98.0]	99.7 [98.4]	99.7 [98.0]	7.5 [6.0]	9,594 [905]	1,067 [101]
Running	99.6 [99.0]	99.6 [99.0]	99.6 [99.1]	99.6 [99.0]	6.7 [6.0]	14,722 [905]	1,636 [101]
Writing	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; F₁, F₁-Score; Prec, Precision; and Rec, Recall.

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., Figure 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 fewer smooth curve components (Figure 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 (Figure 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 (Figure 4.2.3) is somewhat wavier than that of p3 (Figure 4.1.3), which is also reflected in the corresponding genWriting data.

3.3 Classification with generated test data

Table 4 presents the results of the person classifications with generated test data. The classification F₁-score of the genRunnings generated from walking data and the genWalkings from running data was 92.5% and 98.9%, respectively. The classification F₁-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 F₁-score was 46.8% and 50.0%, respectively. Thus, the results are 7.7–14.0 times better than the ZRB guess probability.

4 Discussion

4.1 Person classification based on original data

To test the underlying assumption of the 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 (F₁-score: running: 99.6%, walking: 99.7%, and writing: 99.0%). These results are therefore within the range of previous studies on the individuality of movement patterns (Bauer and Schöllhorn, 1997; Schöllhorn et al., 2002a; Schöllhorn et al., 2002b; Schöllhorn et al., 2006; Janssen et al., 2008; 2011; Schmidt, 2012; Albrecht et al., 2014; Horst et al., 2016; Horst et al., 2017a; Horst et al., 2017b; Horst et al., 2019; Burdack et al., 2020a).

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%–98.0% F₁-score) and running (99.6%–99.0% F₁-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 (Burdack et al., 2020a). Similar to the study by Schöllhorn et al. (2002a), 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 explaining 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 (Burdack et al., 2020b).

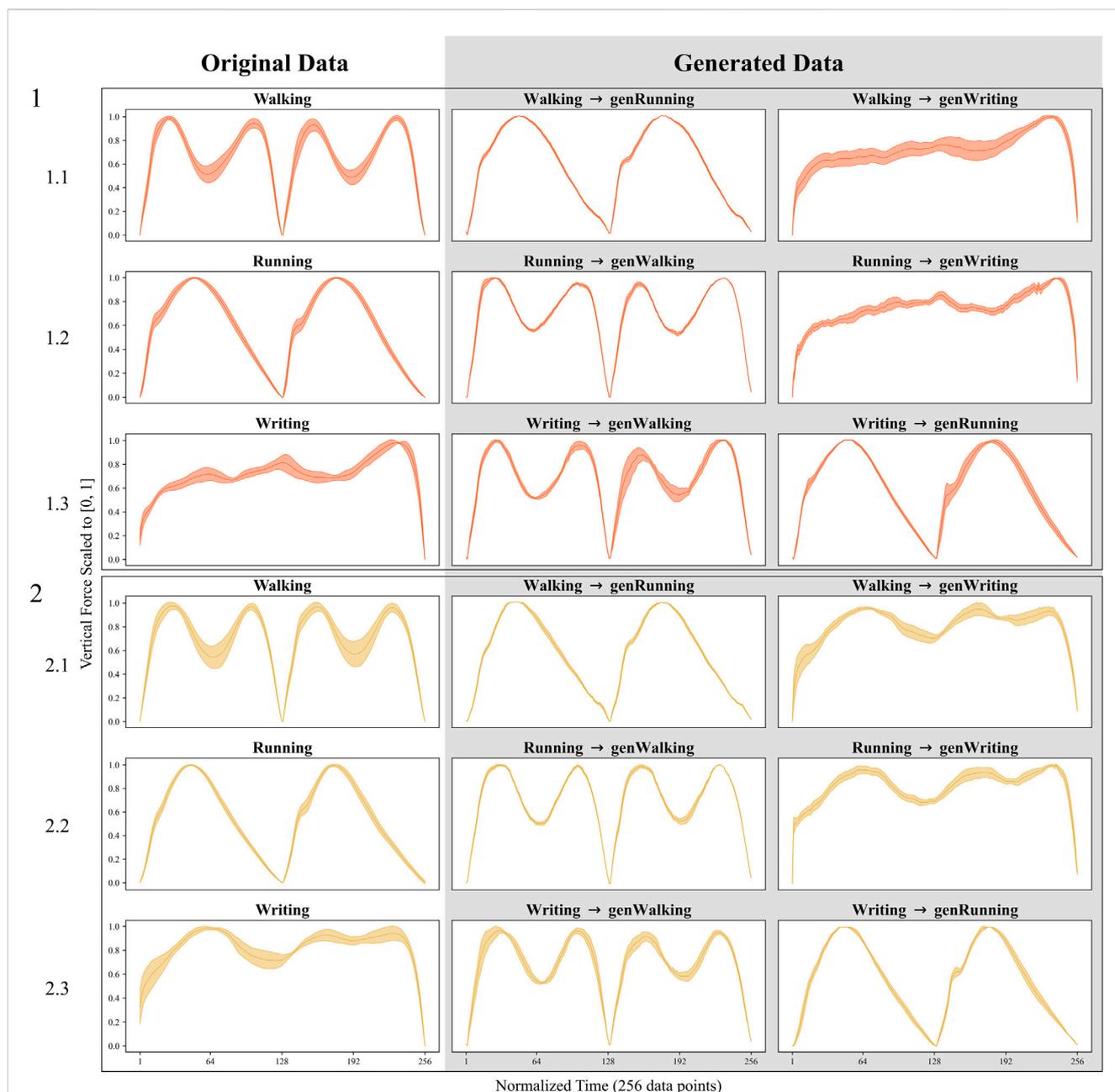


FIGURE 4

Original and CycleGAN-generated Data. Shown is 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; and genWriting, generated handwriting data.

4.2 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 (Zhu et al., 2017) 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% F1-score, corresponding to almost 8 times and up to 14 times the guess probability, respectively. Consequently, this provided the first

TABLE 4 Classification results of person recognition using SVM with generated test data.

Classification problem		Classification score [%]					Number of trials	
Test Data	Generated from	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; F₁, F₁-score; Prec, Precision; Rec, Recall; genWalking, Generated walking data; genRunning, Generated running data; and genWriting, Generated handwriting data.

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% F₁-score) and genWalking (98.9% F₁-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 (Figure 4) 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 (Goodfellow et al., 2014; Goodfellow et al., 2020). 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 guessing 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 78.7% and 78.4% F₁-score and the

genWriting data from running (50.0% F₁-score) and walking (46.1% F₁-score) data are at least 7.7 times better than the ZRB. Thus, the results are worse than those of the genRunning and genWriting data generated from the walking and running data. The figures (Figure 4) 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 the 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 a 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 the generation, but this 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 remain a subject of future research. The extent to which analyzing the first letter of the sentence while handwriting, and a possible altered variation in movement associated with this, may also have influenced the results remains a subject for future investigation. 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.

4.3 Identification of underlying individuality 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 (Horst et al., 2020). 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 (Bauer and Schöllhorn, 1997; Schöllhorn et al., 2002a; Schöllhorn et al., 2002b; Schöllhorn et al., 2006; Janssen et al., 2008; Janssen et al., 2011; Schmidt, 2012; Albrecht et al., 2014; Horst et al., 2016; Horst et al., 2017a; Horst et al., 2017b; Horst et al., 2019; Burdack et al., 2020a).

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 the 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. It would be of interest to see how additional movement components such as three-dimensional GRF, which has been shown to be better than vertical GRF in classifying individuals (Horst et al., 2023), could convey additional information about the individual. While a transfer from kinematic data to GRF data during walking could be shown using GANs (Bicer et al., 2022), there is still great potential to be exploited cross-movement wise. For example, it is important to find out the extent to which the large thematic overlap of locomotion movements might affect inherent biases related to movement character or motor and neurological control compared to handwriting. It is therefore even more encouraging that more different movement signals, such as that of handwriting, could be adequately generated. This gives room for an experimental extension to other signals as well as to linkages with other signals such as audio (e.g., voice), ECG, or EEG (John et al., 2022), which, however, remain 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, etc.

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 the 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 (Buss and Plomin, 1975; Funder and Colvin, 1991; Sherman et al., 2010; Marcus and Roy, 2017). 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.

4.4 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 (Saxena and Cao, 2022) 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 seems to provide the potential for domain-specific optimization.

The particular potential of the conditional CycleGAN lies in the possibility of 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 and 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.

5 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 the 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.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems*. arXiv. doi:10.5281/zenodo.5043456
- Aeles, J., Horst, F., Lopuschkin, S., Lacourpaille, L., and Hug, F. (2021). Revealing the unique features of each individual's muscle activation signatures. *J. R. Soc. Interface* 18 (174), 20200770. doi:10.1098/RSIF.2020.0770
- Albrecht, S., Janssen, D., Quarz, E., Newell, K. M., and Schöllhorn, W. I. (2014). Individuality of movements in music - finger and body movements during playing of the flute. *Hum. Mov. Sci.* 35, 131–144. doi:10.1016/j.humov.2014.03.010
- Barth, V., Käsbauer, H., Ferrauti, A., Kellmann, M., Pfeiffer, M., Hecksteden, A., et al. (2019). Individualized monitoring of muscle recovery in elite badminton. *Front. Physiology* 10, 778. doi:10.3389/fphys.2019.00778
- Bates, B. T. (1996). Single-subject methodology: An alternative approach. *Med. Sci. Sports Exerc.* 28 (5), 631–638. doi:10.1097/00005768-199605000-00016
- Bauer, H. U., and Schöllhorn, W. I. (1997). Self-organizing maps for the analysis of complex movement patterns. *Neural Process. Lett.* 5, 193–199. doi:10.1023/a:1009646811510

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of the Johannes Gutenberg-University Mainz. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JB and MN recorded the data. JB, SG, and WS conceived the presented idea. JB, SG, and MS performed the data analysis and designed the figures. JB and WS wrote the manuscript. JB, SG, MS, MN, CM, and WS reviewed and approved the final manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

CM was employed by the Company Science&Motion GmbH. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbioe.2023.1204115/full#supplementary-material>

- Begg, R. K., and Kamruzzaman, J. (2005). A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data. *J. Biomechanics* 38 (3), 401–408. doi:10.1016/j.jbiomech.2004.05.002
- Begg, R. K., Palaniswami, M., and Owen, B. (2005). Support vector machines for automated gait classification. *IEEE Trans. Biomed. Eng.* 52 (5), 828–838. doi:10.1109/TBME.2005.845241
- Bicer, M., Phillips, A. T. M., Melis, A., McGregor, A. H., and Modenese, L. (2022). Generative deep learning applied to biomechanics: A new augmentation technique for motion capture datasets. *J. Biomechanics* 144, 111301. doi:10.1016/j.jbiomech.2022.111301
- Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1996). “A training algorithm for optimal margin classifier,” in Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory, Pittsburgh, Pennsylvania, July 27–29, 1992. doi:10.1145/130385.130401
- Buford, T. W., Roberts, M. D., and Church, T. S. (2013). Toward exercise as personalized medicine. *Sports Med.* 43 (3), 157–165. doi:10.1007/s40279-013-0018-0
- Burdack, J., Horst, F., Aragonés, D., Eekhoff, A., and Schöllhorn, W. I. (2020a). Fatigue-related and timescale-dependent changes in individual movement patterns identified using support vector machine. *Front. Psychol.* 11, 551548. doi:10.3389/fpsyg.2020.551548
- Burdack, J., Horst, F., Giesselbach, S., Hassan, I., Daffner, S., and Schöllhorn, W. I. (2020b). Systematic comparison of the influence of different data preprocessing methods on the performance of gait classifications using machine learning. *Front. Bioeng. Biotechnol.* 8, 260. doi:10.3389/fbioe.2020.00260
- Buss, A. H., and Plomin, R. (1975). *A temperament theory of personality development*. New Jersey, United States: Wiley-Interscience.
- Button, C., Davids, K., Bennett, S., and Taylor, M. (2000). Mechanical perturbation of the wrist during one-handed catching. *Acta Psychol.* 105 (1), 9–30. doi:10.1016/S0001-6918(00)00044-5
- Chan, I. S., and Ginsburg, G. S. (2011). Personalized medicine: Progress and promise. *Annu. Rev. Genomics Hum. Genet.* 12 (1), 217–244. doi:10.1146/annurev-genom-082410-101446
- Chau, T. (2001a). A review of analytical techniques for gait data. Part 1: Fuzzy, statistical and fractal methods. *Gait Posture* 13 (1), 49–66. doi:10.1016/S0966-6362(00)00094-1
- Chau, T. (2001b). A review of analytical techniques for gait data. Part 2: Neural network and wavelet methods. *Gait Posture* 13 (2), 102–120. doi:10.1016/S0966-6362(00)00095-3
- Cortes, C., and Vapnik, V. N. (1995). Support-vector networks. *Mach. Learn.* 20(3), 273–297. doi:10.1023/A:1022627411411
- Cutting, J. E., and Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bull. Psychonomic Soc.* 9 (5), 353–356. doi:10.3758/BF03337021
- Dal, U., Erdogan, T., Resitoglu, B., and Beydagi, H. (2010). Determination of preferred walking speed on treadmill may lead to high oxygen cost on treadmill walking. *Gait Posture* 31 (3), 366–369. doi:10.1016/j.gaitpost.2010.01.006
- Davids, K., Button, C., and Bennett, S. (1999). Modeling human motor systems in nonlinear dynamics: Intentionality and discrete movement behaviors. *Nonlinear Dyn. Psychol. Life Sci.* 3 (1), 3–30. doi:10.1023/A:1022429522099
- Funder, D. C., and Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *J. Personality Soc. Psychol.* 60 (5), 773–794. doi:10.1037/0022-3514.60.5.773
- Ginsburg, G. S., and Willard, H. F. (2009). Genomic and personalized medicine: Foundations and applications. *Transl. Res.* 154 (6), 277–287. doi:10.1016/j.trsl.2009.09.005
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). “Generative adversarial nets,” in *Advances in neural information processing systems*. Editors Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (New York: Curran Associates, Inc), 2672–2680.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2020). Generative adversarial networks. *Commun. ACM* 63 (11), 139–144. doi:10.1145/3422622
- Grunz, A., Memmert, D., and Perl, J. (2012). Tactical pattern recognition in soccer games by means of special self-organizing maps. *Hum. Mov. Sci.* 31 (2), 334–343. doi:10.1016/j.humov.2011.02.008
- Hammad, I., and El-Sankary, K. (2020). “Using machine learning for person identification through physical activities,” in Proceedings - IEEE International Symposium on Circuits and Systems, Sevilla, Spain, October 10–21, 2020. doi:10.1109/ISCAS45731.2020.9181231/VIDEO
- Hecksteden, A., Kraushaar, J., Scharhag-Rosenberger, F., Theisen, D., Senn, S., and Meyer, T. (2015). Individual response to exercise training - a statistical perspective. *J. Appl. Physiology* 118 (12), 1450–1459. doi:10.1152/jappphysiol.00714.2014
- Hoerzer, S., von Tscherner, V., Jacob, C., and Nigg, B. M. (2015). Defining functional groups based on running kinematics using Self-Organizing Maps and Support Vector Machines. *J. Biomechanics* 48 (10), 2072–2079. doi:10.1016/j.jbiomech.2015.03.017
- Hoitz, F., von Tscherner, V., Baltich, J., and Nigg, B. M. (2021). Individuality decoded by running patterns: Movement characteristics that determine the uniqueness of human running. *Plos One* 16 (4), e0249657. doi:10.1371/JOURNAL.PONE.0249657
- Horst, F., Eekhoff, A., Newell, K. M., and Schöllhorn, W. I. (2017a). Intra-individual gait patterns across different time-scales as revealed by means of a supervised learning model using kernel-based discriminant regression. *PLoS ONE* 12 (6), e0179738. doi:10.1371/journal.pone.0179738
- Horst, F., Mildner, M., and Schöllhorn, W. I. (2017b). One-year persistence of individual gait patterns identified in a follow-up study – a call for individualised diagnose and therapy. *Gait Posture* 58, 476–480. doi:10.1016/j.gaitpost.2017.09.003
- Horst, F., Janssen, D., Beckmann, H., and Schöllhorn, W. I. (2020). Can individual movement characteristics across different throwing disciplines be identified in high-performance decathletes? *Front. Psychol.* 11, 2262. doi:10.3389/fpsyg.2020.02262
- Horst, F., Kramer, F., Schäfer, B., Eekhoff, A., Hegen, P., Nigg, B. M., et al. (2016). Daily changes of individual gait patterns identified by means of support vector machines. *Gait Posture* 49, 309–314. doi:10.1016/j.gaitpost.2016.07.073
- Horst, F., Lapuschkin, S., Samek, W., Müller, K. R., and Schöllhorn, W. I. (2019). Explaining the unique nature of individual gait patterns with deep learning. *Sci. Rep.* 9 (1), 2391. doi:10.1038/s41598-019-38748-8
- Horst, F., Slijepcevic, D., Simak, M., Horsak, B., Schöllhorn, W. I., and Zeppelzauer, M. (2023). Modeling biological individuality using machine learning: A study on human gait. *Comput. Struct. Biotechnol. J.* 21, 3414–3423. doi:10.1016/j.csbj.2023.06.009
- Hsu, C., Chang, C., and Lin, C. (2003). *A practical guide to support vector classification*. Available at: <http://www.csie.ntu.edu.tw/~cjlin>.
- Isola, P., Zhu, J.-Y., Zhou, T., Efros, A. A., and Research, B. A. (2017). “Image-to-image translation with conditional adversarial networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, Hawaii, USA, 21–26 July 2017.
- Jacques, T., Bini, R., and Arndt, A. (2022). Inter-limb differences in *in-vivo* tendon behavior, kinematics, kinetics and muscle activation during running. *J. Biomechanics* 141, 111209. doi:10.1016/j.jbiomech.2022.111209
- Jäger, J. M., and Schöllhorn, W. I. (2007). Situation-orientated recognition of tactical patterns in volleyball. *J. Sports Sci.* 25(12), 1345–1353. doi:10.1080/02640410701287230
- Jain, A. K., Bolle, R. M., and Pankani, S. (2006). *Biometrics: Personal identification in network society*. Berlin, Germany: Springer.
- Jaitner, T., Mendoza, L., and Schöllhorn, W. I. (2001). Analysis of the long jump technique in the transition from approach to takeoff based on time-continuous kinematic data. *Eur. J. Sport Sci.* 1 (5), 1–12. doi:10.1080/17461390100071506
- Janssen, D., Schöllhorn, W. I., Lubienietzki, J., Fölling, K., Kokenge, H., and Davids, K. (2008). Recognition of emotions in gait patterns by means of artificial neural nets. *J. Nonverbal Behav.* 32 (2), 79–92. doi:10.1007/s10919-007-0045-3
- Janssen, D., Schöllhorn, W. I., Newell, K. M., Jäger, J. M., Rost, F., and Vehof, K. (2011). Diagnosing fatigue in gait patterns by support vector machines and self-organizing maps. *Hum. Mov. Sci.* 30 (5), 966–975. doi:10.1016/j.humov.2010.08.010
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14 (2), 201–211. doi:10.3758/bf03212378
- John, A. T., Barthel, A., Wind, J., Rizzi, N., and Schöllhorn, W. I. (2022). 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.* 16, 816334. doi:10.3389/fnbeh.2022.816334
- Johnson, J., Alahi, A., and Fei-Fei, L. (2016). “Perceptual losses for real-time style transfer and super-resolution,” in 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, 694–711. doi:10.1007/978-3-319-46475-6_43
- Kaur, R., Motl, R. W., Sowers, R., and Hernandez, M. E. (2023). A vision-based framework for predicting multiple sclerosis and Parkinson’s disease gait dysfunctions—a deep learning approach. *IEEE J. Biomed. Health Inf.* 27 (1), 190–201. doi:10.1109/JBHI.2022.3208077
- Kaye, D. H. (2010). Probability, individualization, and uniqueness in forensic science evidence. *Brooklyn Law Rev.* 75 (4), 1163–1185.
- Kempe, M., Grunz, A., and Memmert, D. (2014). Detecting tactical patterns in basketball: Comparison of merge self-organizing maps and dynamic controlled neural networks. *Eur. J. Sport Sci.* 15 (4), 249–255. doi:10.1080/17461391.2014.933882
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., et al. (2017). “Photo-realistic single image super-resolution using a generative adversarial network,” in Computer Vision and Pattern Recognition Conference, Honolulu, HI, USA, July 21 2017 to July 26 2017.
- Lee, H. S., Schreiner, L., Jo, S.-H., Sieghartsleitner, S., Jordan, M., Pretl, H., et al. (2022). Individual finger movement decoding using a novel ultra-high-density electroencephalography-based brain-computer interface system. *Front. Neurosci.* 16, 1009878. doi:10.3389/fnins.2022.1009878
- Li, C., and Wand, M. (2016). “Precomputed real-time texture synthesis with markovian generative adversarial networks,” in *Computer vision - ECCV 2016*. ECCV 2016. Editors B. Leibe, J. Matas, N. Sebe, and M. Welling (Cham: Springer), 702–716. doi:10.1007/978-3-319-46487-9_43

- Liao, K., Xiao, R., Gonzalez, J., and Ding, L. (2014). Decoding individual finger movements from one hand using human EEG signals. *PLoS ONE* 9 (1), e85192. doi:10.1371/journal.pone.0085192
- Marcus, J., and Roy, J. (2017). Values, personality, and sustainability behaviour: An integrative analysis. *Acad. Manag. Proc.* 2017 (1), 13105. doi:10.5465/AMBPP.2017.13105abstract
- Maurer, C., Federolf, P., von Tscharnar, V., Stirling, L., and Nigg, B. M. (2012). Discrimination of gender-speed-and shoe-dependent movement patterns in runners using full-body kinematics. *Gait Posture* 36 (1), 40–45. doi:10.1016/J.GAITPOST.2011.12.023
- Müller, K. R., Mika, S., Rätsch, G., Tsuda, K., and Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *Handb. Neural Netw. Signal Process.* 12 (2), 181–201. doi:10.1109/72.914517
- Ng, P. C., Murray, S. S., Levy, S., and Venter, J. C. (2009). An agenda for personalized medicine. *Nature* 461 (7265), 724–726. doi:10.1038/461724a
- Nuzzo, R. (2014). Scientific method: Statistical errors. *Nature* 506 (7487), 150–152. doi:10.1038/506150a
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Rein, R., and Memmert, D. (2016). Big data and tactical analysis in elite soccer: Future challenges and opportunities for sports science. *SpringerPlus* 5 (1), 1410. doi:10.1186/S40064-016-3108-2
- Savitzky, A., and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36 (8), 1627–1639. doi:10.1021/AC60214A047/ASSET/AC60214A047.FP.PNG_V03
- Saxena, D., and Cao, J. (2022). Generative adversarial networks (GANs). *ACM Comput. Surv.* 54 (3), 1–42. doi:10.1145/3446374
- Schafer, R. W. (2011). What is a savitzky-golay filter? *IEEE Signal Process. Mag.* 28 (4), 111–117. doi:10.1109/MSP.2011.941097
- Schmidt, A. (2012). Movement pattern recognition in basketball free-throw shooting. *Hum. Mov. Sci.* 31 (2), 360–382. doi:10.1016/j.humov.2011.01.003
- Schölkopf, B., and Smola, A. J. (2002). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. United States: MIT Press.
- Schöllhorn, W. I., and Bauer, H. U. (1998a). “Assessment of running patterns using neural networks,” in *Sport und Informatik*. Editors J. Mester and J. Perl (Germany: Sport and Buch Strauss), 208.
- Schöllhorn, W. I., and Bauer, H. U. (1998b). “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*. Editors H. J. Riehle and M. Vieten (USA: International Society for Biomechanics in Sport), 574–577.
- Schöllhorn, W. I. (1993). 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]. *Harri Dtsch.* 19.
- Schöllhorn, W. I., Nigg, B. M., Stefanyshyn, D. J., and Liu, W. (2002a). Identification of individual walking patterns using time discrete and time continuous data sets. *Gait Posture* 15 (2), 180–186. doi:10.1016/S0966-6362(01)00193-X
- Schöllhorn, W. I., Schaper, H., Kimmeskamp, S., and Milani, T. (2002b). Inter- and intra-individual differentiation of dynamic foot pressure patterns by means of artificial neural nets. *Gait Posture* 16 (1), 172–173.
- Schöllhorn, W. I., Peham, C., Licka, T., and Scheidl, M. (2006). A pattern recognition approach for the quantification of horse and rider interactions. *Equine Veterinary J.* 38 (36), 400–405. doi:10.1111/j.2042-3306.2006.tb05576.x
- Schöllhorn, W. I., Röber, F., Jaitner, T., Hellstern, W., and Käubler, W. (2001). “Discrete and continuous effects of traditional and differential training in sprint running,” in *Perspectives and profiles 6th European college on sports science congress*. Editors J. Mester, G. King, H. K. Strüder, E. Tsolakidis, and A. Osterburg (Germany: Sport and Buch Strauss), 331.
- Schöllhorn, W. I., Stefanyshyn, D. J., Nigg, B. M., and Liu, W. (1999). Recognition of individual walking patterns by means of artificial neural nets. *Gait Posture* 10 (1), 85–86. doi:10.1016/S0966-6362(99)90454-X
- Schöner, G., Zanone, P. G., and Kelso, J. A. S. (1992). Learning as change of coordination dynamics: Theory and experiment. *J. Mot. Behav.* 24 (1), 29–48. doi:10.1080/00222895.1992.9941599
- Sherman, R. A., Nave, C. S., and Funder, D. C. (2010). Situational similarity and personality predict behavioral consistency. *J. Personality Soc. Psychol.* 99 (2), 330–343. doi:10.1037/A0019796
- Taigman, Y., Polyak, A., and Wolf, L. (2017). “Unsupervised cross-domain image generation,” in 2017 International Conference on Learning Representations, Toulon, France, April 24–26, 2017.
- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *J. Vis.* 2 (5), 2. doi:10.1167/2.5.2
- Westphal, N., and Schöllhorn, W. I. (2001). “Identifying Volleyballteams by their tactical moves,” in 6th annual congress of the European College of Sport Science, Cologne 2001 - Germany, 24–28 July, 55.
- Windle, C. M., Gregory, S. M., and Dixon, S. J. (1999). The shock attenuation characteristics of four different insoles when worn in a military boot during running and marching. *Gait Posture* 9 (1), 31–37. doi:10.1016/S0966-6362(99)00002-8
- Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. (2017). “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, Oct. 22 2017 to Oct. 29 2017, 2242–2251. doi:10.1109/ICCV.2017.244