Int. J. Appl. Math. Comput. Sci., 2025, Vol. 35, No. 4, 631–649
DOI: 10.61822/amcs-2025-0045

A QUATERNION–BASED AUGMENTING METHOD DEDICATED TO BIOMETRIC GAIT SYSTEMS

ALEKSANDER SAWICKI [a,*], KHALID SAEED [a,b]

[a]Faculty of Computer Science
Bialystok University of Technology
ul. Wiejska 45A, 15-351 Białystok, Poland

[b]Department of Computer Science and Electronics
Universidad de la Costa
Calle 58 55–66, Barranquilla, Colombia
e-mail: a.sawicki@pb.edu.pl

Collecting large amounts of real-world data can be expensive, time-consuming, or even impossible in some cases. This situation can be particularly acute for IMU-based gait-analysis data due to participant fatigue, changing walking surfaces, and instrumentation positioning. These factors can provoke changes in the characteristics (drift or distribution shift) of data collected over long durations or on different days, making analysis difficult. Ideally, one would prefer a larger data set whose distributional characteristics do not change. Data augmentation can help address this limitation by creating more training data from what is already available. It is a technique used to artificially increase the amount of data available for training machine learning models. It works by creating new data points from existing data through various modifications. We present a new data-augmentation technique based on modelling three types of common sensor disturbances: alignments, vibrations and drift. The new approach is compared to three state-of-the-art methods, and validated on three different data sets: two from publicly accessible repositories, and one from our own laboratory with 100 participants and 30 gait cycles per person. Improvements in classification accuracy are obtained for each data set: 35%, 90% and 21%, respectively. The experiments show that the use of data augmentation has a positive impact on the metrics of the gait biometric system. It enables increased efficiency in identifying subjects from a limited sample of data and does not require a problematic data acquisition process.

Keywords: biometrics, MEMS, CNN, augmentation.

1. Introduction

The aim of biometric systems is to verify a subject based on physiological and behavioural features or their combination. The development and implementation of biometrics applications are attracting the attention of both academic and commercial sectors (Zou *et al.*, 2020). Following the popularisation and a decrease in the price of based on a accelerometers microelectromechanical system (MEMS), an increase in the number of applications based on motion analysis has been observed in recent years. Applications related to gait-based verification of individuals are particularly promising. The intensification of work in this area is primarily motivated by the possibility of data acquisition using mobile devices such as smartphones or smartwatches (Sprager and Juric, 2015). Another advantage of these types of systems is the difficulty in deliberately mimicking the gait of other people (Giacomo *et al.*, 2017).

Moreover, the lack of active interaction of the participant with other devices (which occurs in the case of iris or finger sampling (Wan *et al.*, 2018) is worth noting. Unfortunately, despite the many advantages of behavioural systems based on gait analysis, they have a significant drawback, which is the distribution shift of gait samples collected during two separate days. The main factors influencing gait pattern include the subject's fatigue, different footwear, and the type of surface on which the walk is performed. When gait samples are

*Corresponding author

collected over a longer period, the gait pattern can also be influenced by weight change or muscular growth through sports exercise (Boyd and Little, 2005). The literature (Darwish, 2017) identifies carrying things while moving, and the participant's illness (fatigue, malaise, weakness) adds additional factors that may change the gait pattern. The aforementioned disadvantages of behavioural biometrics mean that they are sometimes combined with traditional physiological biometrics (Abate *et al.*, 2017). However, this approach requires active interaction between the subject and the device.

On the other hand, the use of sensors such as accelerometers or gyroscopes involves an additional inconvenience, i.e., the collected signals have a "fingerprint" related to the way the device is mounted. This is precisely related to the fact that data acquisition is carried out in the sensor's local reference system. This additional factor also affects the distribution shift between gait samples (for two separate days) and negatively affects the person verification process.

On the other hand, recent years have seen an increase in the number of deep learning (DL) solutions that allow obtaining very good classification results in areas such as natural language processing, speech or time series analysis. The effectiveness of the application depends on a large and diverse dataset to ensure good generalization properties of the classifiers (Delgado-Escano *et al.*, 2018). In many cases, acquiring a large number of learning samples is time-consuming, expensive, and, sometimes, impossible. Data augmentation is a set of mechanics and techniques for artificially generating additional data from an already existing training set. It is a kind of emergency solution when the acquisition of new data is impossible for various reasons (financial, organizational) (Steven Eyobu and Han, 2018).

It is worth noting that, in the field of image processing, data augmentation in the form of affine transformations is a widely used standard (Steven Eyobu and Han, 2018; Wieczorek *et al.*, 2024), as evidenced by the embedding of these transformations into deep learning frameworks. In the field of motion signal classification, this topic is not standardized and is more challenging. Data augmentation methods dedicated to image processing or recording audio cannot be directly transferred to the field of motion signals. The reason for this is the different nature of the sensors used as well as the (processing) characteristics (Kim and Jeong, 2021).

It should be added that a significant number of augmentation methods described in the scientific literature are used for biometric studies performed within a single motion tracking session (Abate *et al.*, 2017; Iso and Yamazaki, 2006; Um *et al.*, 2017). As mentioned earlier, gait cycles collected for a single session are very similar to each other, whilst samples collected over two days may differ significantly. Such studies are thus of questionable applicability from our perspective. This paper, therefore, focuses on the development of a data augmentation method for applications in biometric systems operating within two days.

2. Background and related works

In the literature it is possible to distinguish the main trends concerning the generation of motion signals: perturbation of the accelerometer and gyroscope signals using basic modification methods; generation of IMU signals using additional sources of information; and application of methods using a generative approach.

For the first group of solutions, signal perturbations are mainly based on the rotation of the signals in three-dimensional space. Scientific studies show that the rotation of the sensor is what influences their readings value (Ohashi *et al.*, 2017) whilst the translation has a marginal effect. The accelerometer signal at time t can be viewed as a vector in three-dimensional space. Simulating a different way of mounting the sensor (adding offset during mounting) can be realized by adding a rotation to the vector.

Iso and Yamazaki (2006) suggested distorting the original signals by modelling rotations around three axes in the range of 0–45° with a step of 15°. On the other hand, in the work of Um *et al.* (2017), a series of elementary transformations of the accelerometer signals recorded by the mobile were proposed for Parkinson's disease monitoring applications. The best indications were obtained for stand-alone rotation (v. I) or rotation and permutation (v. II) modification. In the proposed solution, a transformation matrix was created from a random angle and axis of rotation. The permutation consisted in dividing the signals into five equal parts and then randomly arranging the segments. The above-mentioned algorithms are of similar character, as they are based on signal rotation. First, the low computational complexity should be mentioned as an advantage of the solution. On the other hand, the disadvantages include the ability to generate signals that are not observable in real-world conditions. For example, the algorithm by Um *et al.* (2017) can generate 90° rotations, which will never be observed by a mobile phone located in the trousers. In addition, the application of these methods to gyroscope signals is debatable. These algorithms were only used to process accelerometer signals.

Simultaneous augmentation of the accelerometer and gyroscope measurement data was presented by Delgado-Escano *et al.* (2018) as a three-step mechanism. In the first step, the noise was added to each IMU channel. In the second one, the signal was scaled between 0.7 and 1.1, and lastly, non-uniformities of the signals sampling were modelled. The publication addresses the interesting topic of simultaneous perturbation of both accelerometer

and gyroscope signals. The proposed solution has the significant disadvantage that the signals generated for the accelerometer and the gyroscope are not "coherent". The proposed method of distorting the different sensor axes completely ignores the orientation dependence of both signals (described in detail in Section 3). This can be summarized by creating patterns that are not observable in the real world.

When it comes to generating signals using additional sources of information, three publications (i.e., Kwon *et al*., 2020; Pellatt *et al*., 2021; Xia *et al*., 2025) are particularly noteworthy. In the work of Kwon *et al*. (2020), the mechanics of artificially generating IMU measurement values from the YouTube platform videos are presented. The generation of gyroscope measurement values was done using only sensor orientation information. In the case of accelerometer signals, sensor displacement information from the video was considered. On the other hand, in the research by Pellatt *et al*. (2021) an analogous approach was taken, with the professional motion capture system Vicon being used to evaluate the displacement speed. In contrast, Xia *et al*. (2025) employed motion capture recordings of the entire skeleton as input, from which virtual IMU signal data were produced. The publications just mentioned (Kwon *et al*., 2020; Pellatt *et al*., 2021; Xia *et al*., 2025) address an issue similar to data augmentation, i.e., synthetic generation. The analytical transformations presented therein indicate that an angular velocity signal can be successfully generated from object orientation information. Generating a useful acceleration signal "from scratch", as it were, requires an additional source of displacement information. In addition, it should be noted that in the work of Xia *et al*. (2025) it was decided to use synthetic data as a basis for transfer learning of additional data from the actual IMU sensors. The authors emphasised that this was primarily motivated by the use of other coordinate systems by the virtual sensors. In addition to the typical augmentation mechanisms that have common features, a wide variety of experimental algorithms can be distinguished (Steven Eyobu and Han, 2018; Subramanian *et al*., 2015; Tran and Choi, 2020). In the innovative work of Steven Eyobu and Han (2018), instead of augmenting sensor signals, the focus was on perturbing feature vectors. In contrast, Subramanian *et al*. (2015) created additional signals using the Kabsh algorithm commonly employed in crystallography. On the other hand, Tran and Choi (2020) used an advanced gait cycle augmentation mechanism. In this case, signal values were modelled for each of the participants applying a normal distribution. In this approach, however, the accelerometer and gyroscope measurement values were not employed, and the signals were transferred to a new coordinate system (additional processing). The approaches by Steven Eyobu and Han (2018), Subramanian *et al*. (2015) as well as Tran and Choi (2020), despite their novelty, are not widely practiced.

The literature review indicates a field gap regarding the simultaneous augmentation of the accelerometer and gyroscope signals. There is also a lack of broader studies that focus more on the perturbation of orientation signals by adding additional perturbations, such as high-speed noise or slow drift.

It should be noted that, in the case of motion signals, work is also being carried out on the application of generative models. A two-step solution based on the use of variational autoencoders to generate the motion of the entire skeleton was developed by Maeda and Ukita (2022). In the first step, synthetic data were created using the learned encoder, and then, with analytical solutions, motion sequence correction was performed. The presented solutions applied the inverse kinematics equation and whole-skeleton data, making it impossible to be implemented when wearing a single motion sensor. In contrast, Carneros-Prado *et al*. (2024) used a combination of long short-term memory (LSTM) and dense networks to model gait, also across the whole motion skeleton. This solution is also not transferable to the described case study. Alzantot *et al*. (2017) employed combined LSTM-MDN (mixture density network) models for the generation of IMU movement sequences. In doing so, the authors pointed out the deterministic nature of the LSTM network as a necessity to include a probabilistic member. Finally, they emphasised that the paper did not perform a collation, i.e., a comparison of generated and orginal data. The authors did not evaluate the possibility of using the generated data for behavioural biometrics.

3. Accelerometer and gyroscope mathematical models

The measured values of the accelerometer can be described by a model related to the orientation of the sensor (Bieda *et al*., 2015; Sawicki *et al*., 2016),

$$a^S = {}^W_S\,\mathrm{R} \cdot (\tilde{a} + g), \tag{1}$$

where a^S is sensor acceleration (in the sensor local system), ${}^W_S\mathrm{R}$ is the rotation matrix from world coordinates to sensor coordinates, $\tilde{a}$ is acceleration of the movement (in the global coordinate system), and g is acceleration of the gravity (in the global coordinate system).

On the other hand, the measured values of the gyroscope are also described by a model related to the orientation of the sensor (Diebel, 2006),

$$\omega_t(q_t, \dot{q}_t) = 2 \cdot M(q_t) \cdot \dot{q}_t, \tag{2}$$

where ω_t denotes an angular velocity vector $(\omega x, \omega y, \omega z)$ in time t and M stands for a matrix mapping the

quaternion qt and its differential to angular velocities expressed in world coordinates.

The value of the matrix M is determined by that of the quaternion q at time t. This relationship is described by

$$M(q_t) = \begin{bmatrix} -q_x & q_w & -q_z & q_y \\ -q_y & q_z & q_w & -q_x \\ -q_z & -q_y & -q_x & q_w \end{bmatrix}, \quad (3)$$

where qw, qx, qy, qz are w, x, y, z components of the quaternion at time t. The differential of the quaternion is described by

$$\dot{q}_t = \frac{(q_{t+1} - q_t)}{\Delta T}, \quad (4)$$

where $q(t+1)$, qt denotes orientation in the form of a quaternion at time $t+1$ and t, while ΔT is the sampling period.

The accelerometer and gyroscope indications present signals in a local reference related to the sensor's orientation. For biometric applications, the sensor measurement values are affected by a kind of 'fingerprint' related to the initial sensor mounting way.

For biometric analysis of data from a single tracking session, "fingerprinting" has a positive effect. It can artificially increase the verification metric (F1-score) by up to 99%, whilst for data involving two days, the metric may remain insufficient. In such a case, data augmentation mechanisms that simulate different ways of sensor mounting are promising. Equations (3) and (4) indicate that the measurement values of the accelerometer and the gyroscope are connected to the sensor orientation. From our own observation, it makes sense to start the data augmentation process by augmenting the orientation signal.

4. Proposed data augmentation model

Previously (Sawicki, 2022), we proposed an elementary data augmentation mechanism in basic form. In the present study, we extend it with an additional "Gain accelerometer signal" module (Fig. 1). Further details are provided in Fig. 3, whilst the impact on identification results is illustrated in Fig. 8. The developed algorithm is comprised of two primary stages. In the initial stage, the sensor's orientation signal is subjected to perturbations, characterized by parameters such as $Offset$ (modelling different sensor initial alignments), $Noise$ (modelling fast-varying disturbances that simulate vibrations), and $Drift$ (modelling the slow tilt of the sensor during gait cycles). In the subsequent stage, the augmented orientation signal is employed to synthesize the angular velocity signal (gyroscope readings) and to model the perturbed accelerometer readings.

The augmentation module requires as input the accelerometer measurement signals and the sensor orientation in the form of a quaternion time series. In addition, it uses the $Offset$, $Noise$, $Drift$ and $Gain$ perturbation parameters. The outputs of the expansion module are three-channel acceleration and gyroscope (angular velocity) signals, which form the standard data set used by the classifiers. The process of orientation augmentation involves modelling varying disturbances as a result of setting three perturbation parameters. The first one ($Offset$) relates to a different way of mounting the sensor by the presence of a constant angle of rotation for the entire duration of the gait cycle. The second one ($Noise$) refers to the presence of short-term orientation distortions modelling oscillations. It is realised by modelling additional random rotation angles for each moment t of the gait duration. The last one ($Drift$) allows modelling the slow tilt of the sensor during the gait duration. It is realised by drawing two starting and ending rotation angles and filling in the entire gait cycle duration using Bezier interpolation. A type of uniform sampling was employed during the process of drawing values. This approach was influenced by the work of Giacomo *et al.* (2017). It must be acknowledged that the $Offset$ parameter maintains a constant value for the

Algorithm 1. Single channel orientation perturbation.

Require: $Offset, Noise, Drift$
1: O[1..3][1..128]{be a Offset array}
2: N[1..3][1..128]{be a Noise array}
3: D[1..3][1..128]{be a Drift array}
4: **for** ch:=1:3 **do**
5: $offset$:= random_uniform(-$Offset$, +$Offset$)
6: $driftstart$:= random_uniform(-$Drift$, +$Drift$)
7: $driftend$:= random_uniform(-$Drift$, +$Drift$)
8: **for** t:=1:128 **do**
9: O[ch][t]:= $offset$ value
10: N[ch][t]:= random_uniform($-Drift$, +$Drift$)
11: D[ch][t] assign Bezier interpolation between $driftstart$ and $driftend$
12: **end for**
13: **end for**
14: **return** O, N, D

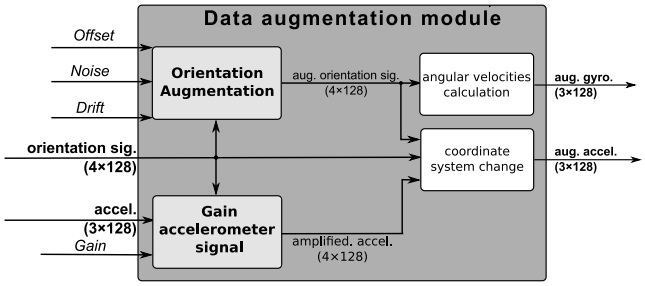

Fig. 1. Block diagram of the augmentation mechanism.

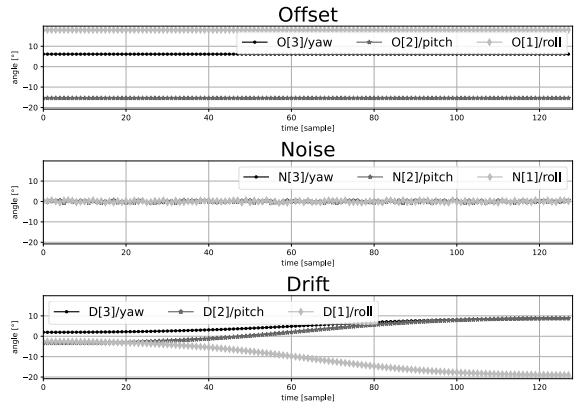

Fig. 2. Sample of the rotation angles used for the $Offset$, $Noise$ and $Drift$ perturbations.

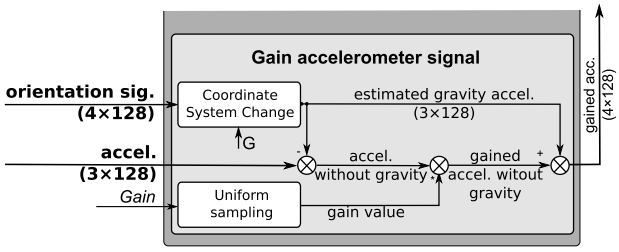

Fig. 3. Detail diagram of the accelerometer signal amplification module.

entirety of the gait cycle. Consequently, the parameter is randomized prior to the iteration of the individual gait cycle t frames. The presented method is performed independently for the three ch channels and results in arrays (O, N, D) modelling additional angles of rotation. The perturbation angle generation process is described in Algorithm 1. It should be noted that the values of the selected perturbations $Offset$, $Noise$, $Drift$ were chosen empirically. The effect of the perturbation settings on the classification values is investigated futher on in this section using the grid search method (results are shown in Fig. 9).

Since the augmentation process is concerned with producing different samples, the augmentation module does not start with a specific perturbation value, but with a numerical range. For a parameter, e.g., $Offset = 20$, a random sensor alignment in the range ⟨-20°, +20°⟩, relative to the original orientation, will be modelled. Figure 2 shows an example of additional rotation angles used for the perturbation of orientation signals. When using 3D space, quite typically the angles of rotation around the Z-axis are named yaw, around the Y-axis $pitch$, and around the X-axis $roll$. In general, the drawn angles of rotation yaw, pitch, and roll are transformed into quaternions (Kitowski *et al.*, 2023) representing the rotation in the respective axes of the coordinate system. A $q_{disturbances}$ quaternion is then formed to represent the rotation according to the equation

$$\begin{aligned} q_{yaw} &= cos\Big(\frac{yaw}{2}\Big) + 0i + 0j + sin\Big(\frac{yaw}{2}\Big)k, \\ q_{pitch} &= cos\Big(\frac{pitch}{2}\Big) + 0i + \Big(\frac{pitch}{2}\Big)j + 0k, \\ q_{roll} &= cos\Big(\frac{roll}{2}\Big) + \Big(\frac{roll}{2}\Big)i + 0j + 0k, \\ q_{disturbances} &= q_{yaw} \cdot q_{pitch} \cdot q_{roll}, \end{aligned} \tag{5}$$

where yaw, $pitch$, $roll$ are values of the rotation angles derived from the disturbances $Offset$, $Noise$, $Drift$, $q_{disturbances}$ is quaternion representing the modelled disturbances and $\cdot$ is a Hamilton operator implementing multiplication of quaternions (Kitowski *et al.*, 2023).

The final augmented orientation was determined as a product of the original (q_t) and disturbed rotations ($q_{disturbances}$). In Eqn. (6), it is explicitly stated that the final augmented signal is comprised of perturbations from the $Offset$, $Noise$ and $Drift$ disturbances,

$$\begin{aligned} q_{augmented} &= q_t \cdot q_{disturbances}, \\ q_{augmented} &= q_t \cdot q_{offset} \cdot q_{noise} \cdot q_{drift}. \end{aligned} \tag{6}$$

Compared to the previous iteration of the augmentation algorithm (Sawicki, 2022), the modified version includes an acceleration gain module (Fig. 3). Accelerometer readings are affected by values caused by movement and by components of gravitational acceleration, which typically has a value of 1 g. By using information about the orientation of the sensor, it is possible to estimate the distribution of gravitational acceleration along each of the three axes of the sensor. As a result, it is possible to determine the values resulting from actual movement and modify it with a random gain parameter (Fig. 3). The action module is described in five steps in Algorithm 2. Two original signals are applied to the module: the orientation ($orientationsig$) and the accelerometer reading ($accelerationsig$), along with the gain parameter ($Gain$) and the known global acceleration $\langle 0; 0; 1\rangle$. The primary advantage of the presented approach is that it enables the perturbation of the acceleration resulting from the object's actual movement, thus allowing the gait speed to be modelled indirectly. This methodology may be especially pertinent in applications where gait cycles are normalized to a fixed duration (as is the case in the present experiments). In certain conditions, gait speed can be regarded as a biometric trait, and a significant modification may have a negative effect on verification results. Concerns have been raised regarding the complexity of the process involved in estimating the gravitational acceleration, particularly in the context of its subsequent integration into "Step

Algorithm 2. Advanced acceleration gain.

Require: $Offset, Noise, Drift$
Require: $orientation\ sig.$, $acceleration\ sig.$, $Gain$, G
1: $orientation\ \ sig.$[1..4][1..128]{be an orientation signal}
2: $accel.$[1..3][1..128]{be accelerometer readings}
3: $Gain$}{be a gain range}
4: G[1..3]{be a global gravity value}
5: $estimated\ gravity\ accel.$[1..3][1..128]{be esitmated gravity components}
6: **for** t:=1:128 **do**
7: $estimated\ \ gravity\ \ accel.$[1..3][t]:= transform_vector(G,$orientation\ sig.$[t]) {Convert global gravity vector G, to a local sensor system using the $orientation\ sig.$}
8: **end for**
9: $accel.\ \ without\ \ gravity$:=$accel.$-$estimated\ gravity\ accel.$ {Create movement only based signal}
10: $gain\ \ value$:= random_uniform(-$Gain$,+$Gain$) {draw gain indicator}
11: $gainedacc.\ \ signal\ \ without\ \ gravity$:= $(1 + gain\ value) \cdot accel.\ without\ gravity$ {model waling speed impact}
12: $gained\ acc.$:= $gained\ accel.\ without\ gravity\ +$ $estimated\,gravity\,accel.$
13: **return** $gained\,acc.$

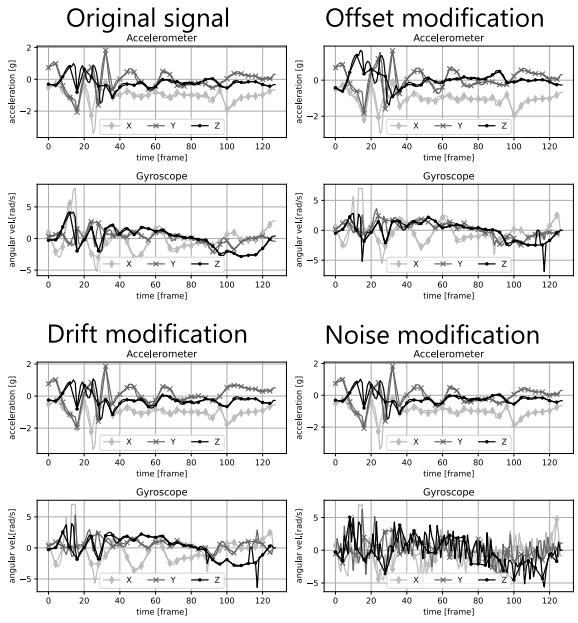

Fig. 4. Effect of perturbation parameters on accelerometer and gyroscope signals.

5". The values resulting from gravitational acceleration are present in the raw accelerometer measurement data and are therefore used by many deep learning approaches (Gadaleta and Rossi, 2018; Delgado-Escano *et al.*, 2018; Zou *et al.*, 2020).

Finally, the output of the augmentation module (Fig. 1) includes an acceleration signal transferred to a new coordinate system and a synthetically generated angular velocity signal. Calculation of the angular velocity is carried out on the basis of Eqn. (2) with the known $W(q_t)$ matrix (Eqn. (3)) and the differential of the modified quaternion $\dot{q}_t$ (Eqn. (4)).

Thanks to this approach, both the generated accelerometer and gyroscope signals are dependent on the sensor orientation. There is no situation where angular velocity and gravitational acceleration are modelled independently of each other. This approach, which seems to us to be incorrect, was adopted by Delgado-Escano *et al.* (2018). The fact that both angular velocity and acceleration depend on the (same) orientation was completely disregarded and independent perturbations of the individual measurement channels were made. Figure 4 presents examples of both original and augmented signals.

5. Description of databases

5.1. Signet dataset. *Participants*: The dataset was collected with 50 participants (SIGNET, 2016; Gadaleta and Rossi, 2018). However, only 31 people participated in two sessions. In addition, for two individuals, the available data had significant missing samples as a result of sampling problems. Therefore, the research study used a database limited to 29 participants. The sensor was embedded in a mobile phone and located in the right trouser pocket. Participants were asked to walk naturally for several minutes. Demographic information (age, gender, height, weight, etc.) was not made available. Also unavailable was information regarding instructions to participants and experimental procedures. Participants were asked to walk naturally for a period of five minutes with an Android mobile phone in their front trouser pocket (one of the Asus Zenfone 2, Samsung S3 Neo, Samsung S4, LG G2, LG G4 and Google Nexus 5 devices). Data acquisition was done wirelessly and the data was sent to the server using a proprietary application.

Instructions given to participants: Participants were asked to move freely for a period of about five minutes. At the same time, they were asked to keep their cell phone in their pants pocket. The way the phone was placed (e.g., phone camera lens at the top, etc.) was not consistent and could vary between acquisition sessions. The Signet dataset is available at `https://signet.dei.unipd.it/research/human-sensing/`.

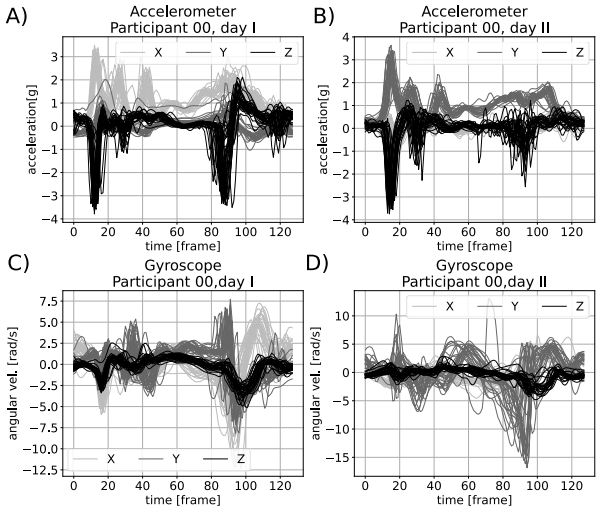

Fig. 5. Samples of Signet segmented gait cycles: accelerometer day I (a), accelerometer day II (b), gyroscope day I (c), gyroscope day II (d).

In the study, a gait cycle segmentation mechanism was implemented. Partitioning data blocks into gait and immobility fragments is a necessary process to increase the accuracy of biometric walking-based systems (Chi *et al.*, 2017). In the presented approach we implemented an algorithm that enables the division of a recorded time series into time fragments in which a single gait cycle was performed—the period of time from when one foot contacts the ground until the same foot contacts the ground again (Whittle, 2014). Therefore, 9584 gait cycles per participant were collected during the first day (M = 330.48, SD = 116.10). On the second day, 9678 gait cycles were recorded (M = 333.72, SD = 123.52).

The study participants were free to rotate their phone "upside down" within individual sessions. Consequently, when analysing the sample gait cycles for particular participants, "channel swapping" can be observed. Figures 5(a) and (b) refer to the accelerometer measurement values acquired on two separate days. Each colour represents a different sensor axis: light grey on the X axis, grey on the Y axis, and black on Z axis. It can be seen that the gait cycles have a characteristic pattern, with the X axis being "swapped" with the Y one due to the change in orientation of the phone during the experiment. Figures 5(c) and (d) refer to the gyroscope obtained on two separate days. In this case, the same colour labelling scheme was used. Also, the same channel relationship occurs for these signals—swapping the measurement axes X with Y. Finally, it should be noted that, despite the undoubted advantage of having two tracking sessions, the data was recorded within the same day for 28 of the 29 included participants. This may be concerning, because the participants had very similar measures of fatigue, emotion, etc. This makes the data collected somewhat different from the potential data used in real-life scenarios. By this we mean a situation where the biometric system is trained on a certain set of gait patterns, and its practical use (gait validation) occurs on a different day (where conditions may be different).

Advantages of the data corpus: Tests are conducted in quasi-real conditions; participants can have different types of footwear and clothing between motion capture sessions; gait samples are recorded over a period of approximately five minutes per trial; data recording covers two different sessions, or even more in the case of selected participants; data acquisition is done using a single smartphone device equipped with an accelerometer and a gyroscope.

Disadvantages of the data corpus: Data recording across two days and sensors located in the right thigh area for only 31 participants out of a large group of 50; interruptions in signal recording for two participants resulted in ultimately useful data only for 29 participants in the experiment; lack of information on footwear worn or demographic data; in 28 cases, both tracking sessions took place on the same day.

5.2. 200GaitData: A human gait database for normal walk collected by a smartphone accelerometer. *Participants*: A total of 93 participants took part in the study (Dehkordi and Farahmand, 2022; Vajdi *et al.*, 2019). However, manual inspection indicated that, for the sensor in the right pants pocket, only 55 participants had recordings of walking signals over two days. A very significant problem was encountered during data preprocessing—the segmentation of gait cycles for a large proportion of the subjects did not proceed correctly. Neither the ready-made implementation of the algorithm (Hoang *et al.*, 2015), the reimplementation of the method (Gadaleta and Rossi, 2018), nor the modification of our algorithm (Sawicki, 2021) were able to correctly process the available data.

Eventually, the segmentation process was carried out correctly for only 29 participants. The reason for this was walking in real-world conditions, including on soft ground such as grass. This meant that detecting the impact of a leg hitting the ground was done improperly. For many participants, there were numerous cases when the process went well up to a certain point, and then the algorithms stopped working properly. Despite the problems with the segmentation of the gait cycle, it was decided to use this database to show the difference from the Signet database, where other conditions of the experiment were required.

Data acquisition took place using two iPhone 6's, attached to the left waist and right thigh of the participants. The final corpus studied included gait cycles collected under real-world conditions on the Boston University

campus. The data acquisition process took place both on hard ground (e.g., concrete pavements) and on soft surfaces such as lawns. The acquisition involved walking over a distance of 200 miles (320 meters). Participants were asked to walk at their comfortable speed. First, they gave verbal consent after learning the details of the experiment and answering demographic questions. In the subsequent session, each subject traversed a distance of 320 metres between locations A and B in both forward and reverse directions, thus completing a total of 640 metres. The experiment was conducted at the same location with the same zero sea level for all subjects. There were no changes in slope along the way.

Instructions given to participants: Each participant had one smartphone fixed around his waist and the other around his right thigh. The participant in the experiment started walking at point A and then walked a distance of about 320 meters. At point B, the person turned around and waited for five seconds. The participant then walked back another 320 m to the starting point A. Upon reaching it, the smartphones used for acquisition were removed. Participants were asked to walk at their own pace and there was no time limit.

200GaitData dataset (Dehkordi and Farahmand, 2022): The experiments used a limited database set of 29 people, including 26 men and three women. The experimental participants were (M = 23.34, SD = 5.97) years old, weighed (M = 75.14, SD = 17.66) kg and were (M = 172.03, SD = 8.44) cm tall. A total of 25,569 gait samples were collected: 13,035 samples during the first day and 12,534 during the second. For each participant on the first day, an average of (M = 449.48, SD = 36.30) samples was acquired , and on the second day (M = 432.21, SD = 33.95.) The collected data closely resembled the signals in Fig. 5 and has the same problem of the sensor's dependence on its mounting method. Therefore, it was decided to skip its graphical representation. Finally, it should be added that (as in the case of the Signet dataset), for 10 people, two sessions were conducted within the same day. This makes the advantages of the supposedly collected data actually reflecting everyday life scenarios debatable. Additionally, despite the presence of two mobile phones, it is not possible to synchronize the signals between them (the timestamp is related to the time when the device was started).

Advantages of the data corpus: Tests conducted under field conditions; gait samples recorded over a distance of approximately 320 m in one direction and then 320 m in the opposite one; data acquisition using two iPhone devices equipped with an accelerometer and a gyroscope.

Disadvantages of the data corpus: Data recording across two days and sensors located in the right thigh area are available for only 55 participants out of a large group of 93; lack of ability to segment gait cycles for a large proportion of participants (29 participant signals were eventually used), as a result of data acquisition for the "soft" surface; data recording using different mobile phones, which can affect the values measured by the accelerometer and the gyroscope (related to sensor calibration); in 10 cases, both tracking sessions took place on the same day; inability to synchronize signals for two mobile phones.

5.3. Bialystok dataset. *Participants*. 100 students and university personnel (47 men, 53 women) were selected in a convenience sample from the Bialystok academic community.

Demographics: age in years (M = 32.18, SD = 8.73), height in meters (M = 1.73, SD = 0.11) and weight in kilograms (M = 75, SD = 19). Exclusion criteria: healthy. Participants do not claim any health case.

Data-acquisition instruments: (a) Perception Neuron device 32 (Noitom (Noitom, 2019)), (b) Microsoft Kinect v2.0 depth camera (Microsoft, 2015; Guzsvinecz *et al*., 2019). The Perception Neuron device included a dedicated body suit with 17 sensors, each comprising an accelerometer, a gyroscope, and a magnetometer. With the built-in data fusion filter (an algorithm that estimates the orientation of the sensor-based triaxial accelerometer, gyroscope and magnetometer signals (NXP, 2016)), an additional sensor-orientation signal was available and, as a result of "T-pose" calibration in which the participant stands straight with arms held horizontally to the side, additional tilt signals for individual limbs were available. The calibration process calculates the difference between the orientation of the sensor and the limb, and requires temporary assumption of a strictly defined pose.

Instructions given to participants: Participants were asked to perform the gait naturally, with no imposed speed or choice of starting leg. However, they were asked not to hold anything in their hands during the gait and to use their natural walk. A single gait trial consisted of the participant standing in front of the starting line and, upon hearing a voice command, starting to walk. The walk ended after a distance of about 3 m. Upon hearing the return command, the participant would return to the starting location. The return of the participant was not recorded.

Procedure: Data acquisition took place in a laboratory environment. Participants performed 20 gait trials on a hard surface over a distance of 3 m, from a start-point to an end-point (distance determined by the depth camera). A gait trial is defined as a period of time in which the participant covers a distance of 3 m, with the number of steps taken varying between participants.

The gait-cycle segmentation algorithm is able to detect the moment the foot strikes the ground, through

a motion sensor located in the thigh area. Due to the different heights of the participants, as well as different walking styles for the straight, fixed-length gait path, a varying number of gait cycle samples was observed. During the first day, a total of 3376 gait cycles samples were collected, i.e., (M = 33.8, SD = 6.1) gait cycles for each participant. A similar number of samples of 3321 gait cycles were collected during the second day: (M = 33.2, SD = 6.6) gait cycles for each participant.

The development of the first author's corpus of data (BUT ALEK) was due to the limitations of other publicly available datasets, especially the size of the collection, data acquisition within two days, and the presence of orientation recordings. In constructing the corpus, particular attention was directed towards ensuring a balanced representation of male and female participants. Participants with injuries, who could introduce additional bias and artificially distort the results, were excluded from the experiment. To ensure the motivation of participants, they were informed about the experiment and given a small gift at the conclusion. It is our contention that every effort was made to minimize additional conditions affecting the outcome of the identification. For the BUT ALEK dataset please contact `aleksander.sawicki.work@gmail.com`.

Figure 6 shows an example of the signals recorded for the selected participant. In this case, as in Fig. 5, each sensor channel is the colour-coded light grey X-axis, grey Y-axis, and black Z-axis. When using the Perception Neuron system, we will not observe the "swapping axis" phenomenon resulting from a different rotation of the sensor. However, it is not possible to wear the body suit identically between tracking sessions. The sensor will have a different tilt. This causes the accelerometer to notice a change in average values. In Fig. 6, it can be observed that the Z-axis black signal for the second session has a constant positive offset.

Advantages of the data corpus: The number of study participants was significantly higher than in the competing databases. For each participant, each tracking session was carried out on a separate day. Data acquisition was performed using the Perception Neuron inertial motion capture system, as well as the Microsoft Kinect v2.0 depth camera. The inertial system allowed for simultaneous acquisition from 17 motion sensors positioned on different parts of the body.

Disadvantages of the data corpus: Data recording takes place at a distance of approximately 3 m, due to the requirement to record MS Kinect, which resulted in a limited number of recorded gait cycles; wireless transmission errors in data acquisition were present. They were supplemented with previous values by the Perception Neuron system—constant values can be observed in Fig. 5.

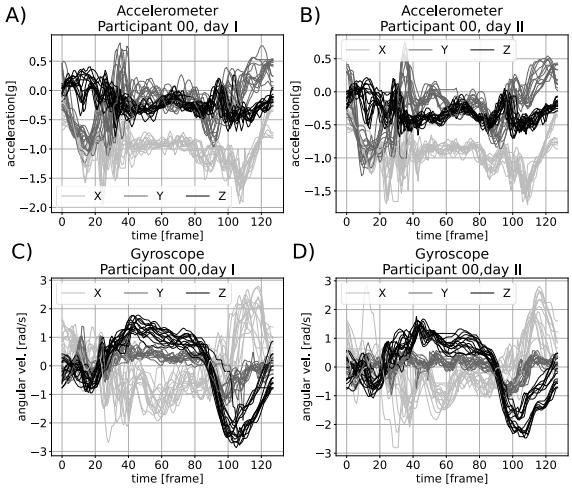

Fig. 6. Samples of BUT segmented gait cycles: accelerometer day I (a), accelerometer day II (b), gyroscope day I (c), gyroscope day II (d).

Table 1 provides a brief summary of the most important factors of each database.

6. Evaluation

The objective of this study is to make a comparison between the novel augmentation technique and its competitors by utilizing three publicly accessible databases. The classification process was executed by employing neural network architectures specifically designed for the processing of IMU signals. In the initial experiment, a Vanilla CNN network (Zou *et al.*, 2020) was employed, comprising alternating convolutional layers and max pooling. In the subsequent experiment, a multi-input network (Delgado-Escano *et al.*, 2018) was utilized, in which accelerometer and gyroscope signals were applied separately to the network input. In this case, a late data fusion architecture was employed.

In the final experiment, a CNN network with an additional attention mechanism was applied (Huang *et al.*, 2021).

Tables 2, 3, and 4 provide a detailed summary of the individual architectures. It is worth noting that, in the case of the baseline model, defined as "Vanilla CNN" and "CNN with attention mechanism", the input data had a length of 128 steps/pixel. For the "Multi-input CNN" architecture, the length was 100 steps/pixel. In addition, for each of the datasets used, due to the changing number of participants in the experiment, modifications were made to the number of neurons in the last layer. Nevertheless, for our 100-person database for the three successive CNN architectures, the number of trained parameters was, respectively, 334,564, 339,246 and 1,759,180. In the case of the multi-input CNN

Table 1. Dataset summary.

Name	Signet Research Group	200GaitData	BUT ALEK
University	Padova, Italy	Boston, US	Bialystok University of Technology, Poland
Environment	Quasi-real	Field	Laboratory
Number of available/used participants	31/29	55/29	100/100
Task	Walk for 5 min	Walk 320 m between locations and then return	Twenty repetitions of walking for a dist. of 3 m
Sensor number	1	2	17
Data acquisition tool	Android mobile phone	iPhone mobile phones	Perception Neuron inertial tracking system
Constant gait ground type	Yes	No	Yes
Constant type of footwear	No	Yes	Yes
Number of gait cycles in Session I	9584	13035	3376
Number of gait cycles in Session II	9678	12534	3321

Table 2. Vanilla CNN architecture.

Layer	Type	Details
1	convolution_1	in=1, out=32, ks=[1,9],str=[2]
2	max_pooling_1	ks=[1,2], str=[2]
3	convolution_2	in=32, out=64, ks=[1,3],str=[1]
4	convolution_3	in=64, out=128, ks=[1,3], str=[1]
5	max_pooling_3	ks=[1,2], str=[2]
6	convolution_4	in=128, out=128, ks=[6,1], str=[1]
7	dense	in=2048, out=100
8	softmax	

architecture, that number was about five times higher than for the previous networks. This indicates a much higher input requirement, but also a much longer training time for this type of model.

In Fig. 7, the results of person identification for the BUT ALEK database are presented. The figure displays box plots for the three investigated classifiers. Due to the non-deterministic mechanism of neural networks (random setting of the initial network weights), each experiment

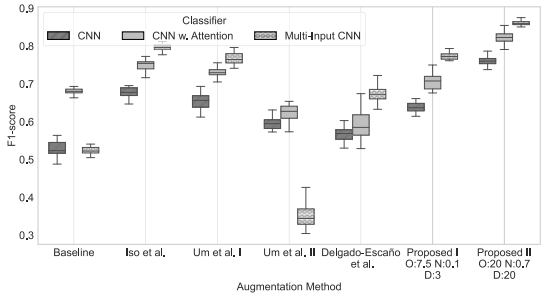

Fig. 7. Comparison of augmentation influence for three selected classifiers.

was repeated multiple times. Simple validation repeated 50 times was used to evaluate the learned classifiers. In the evaluation process, the segmented gait cycle from the first day was utilized as the training set, and the cycle from the second day was employed as the validation set. The horizontal axis shows a summary of our approach and selected other augmentation methods. From the conducted experiments, it can be observed that, for the baseline case, the identification metric ranges from approximately 0.5 to 0.7. The choice of the classifier has an obvious impact on the accuracy of the biometric system. The CNN with the attention mechanism (Table 4) allowed obtaining better average verification rates than the Vanilla CNN classifier (Table 2) in each case studied. It can be noted that the use of multi-input CNNs (Table 3) in combination with the developed method allowed obtaining very good results. At the same time, this network architecture is characterized by high instability in regards to other augmentation methods, such as no augmentation—baseline or Um *et al.* II. Therefore, a classifier with an attention mechanism was used in the

Table 3. Multi-input CNN architecture.

Layer	Type	Details
1	convolution_1	in=1, out=240, ks=[1,10],p=VAL
2	batch_norm_1	n_features=240
3	max_pooling_1	ks=[1,2], str=[1,2]
4	convolution_2	in=240, out=300, ks=[1,7],p=VAL
5	batch_norm_2	n_features=300
6	max_pooling_2	ks=[1,2], str=[1,2]
7	convolution_3	in=300, out=360, ks=[1,5],p=VAL
8	batch_norm_3	n_features=360
9	max_pooling_3	ks=[1,2], str=[1,2]
10	convolution_3	in=360, out=420, ks=[1,3],p=VAL
11	batch_norm_4	n_features=420
12	aver_pooling_4	ks=[1,5], str=[1,2]
13	dropout	P=0.5
14	dense	in=2520, out=100
15	softmax	

Table 4. CNN witch the attention mechanism architecture.

Layer	Type	Details
1	convolution_1	in=1, out=32, ks=[1,9],str=[1,2]
2	max_pooling_1	ks=[1,2], str=[1,2]
3	batch_norm_1	n_features=32
4	convolution_2	in=32, out=64, ks=[1,3],str=[1]
5	convolution_3	in=64, out=128, ks=[1,3], str=[1,3]
6	max_pooling_3	ks=[1,2], str=[2]
7	batch_norm_2	n_features=128
8	convolution_4	in=128, out=128, ks=[6,1],p=VAL
9	batch_norm_4	n_features=128
10	attention	chann.=128,red.=8
11	dense	in=2048, out=100
8	softmax	

current research. In Fig. 7, it can be observed that our proposed augmentation mechanism allows obtaining different metric rates depending on the parameters set. In the case of the proposed configuration II, the observed rates are the highest in all the trials performed. The results are presented in tabular form in Table 5.

In that table, the mean F1-score for the 50-fold simple validation is presented, with results for the various neural network architectures that were examined. The individual rows show results in the absence and with augmentation. Our mechanism is presented for three different settings: the $Offset$, $Noise$ and $Drift$ parameters. The notation used is abbreviated as O: $Offset$, N: $Noise$ and D: $Drift$.

The highest verification efficiency of 0.86 was observed for the multi-input CNN classifier in cooperation with the "Proposed II" augmentation methods. Conversely, the use of other configurations of perturbation parameters resulted in lower metrics. This finding underscores the necessity of meticulous parameter selection. Results presented in Fig. 6 and Table 5 were obtained under the condition that the gain parameter (see Fig. 1) was set to zero.

In a subsequent study we examined the effect of the number of trials generated when the $Gain$ parameter was 0.0 or 0.1. The graph was drawn for convolutional neural networks with an attention mechanism. The parameter 0.1 represents the multiplication of the accelerometer signal by a random rate in the range 0.9 to 1.1, simulating an increase or decrease in walking speed of up to 10%. Figure 8 shows a box plot where the X-axis represents

Table 5. F1-score metrics for the BUT dataset.

Approach	CNN	M-I CNN	CNN w. attention
Baseline	0.53	0.53	0.68
Iso *et al.*	0.67	0.8	0.75
Delgado-Escano *et al.*	0.56	0.67	0.59
Proposed I O:7.5 N:0.1 D:3	0.64	0.78	0.70
Proposed II O:20 N:0.7 D:20	0.76	0.86	0.82

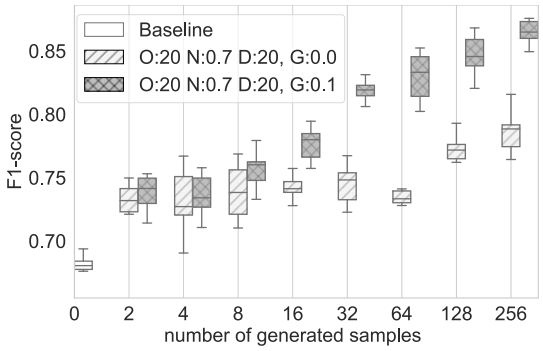

Fig. 8. Influence of the gain parameter on F1-score.

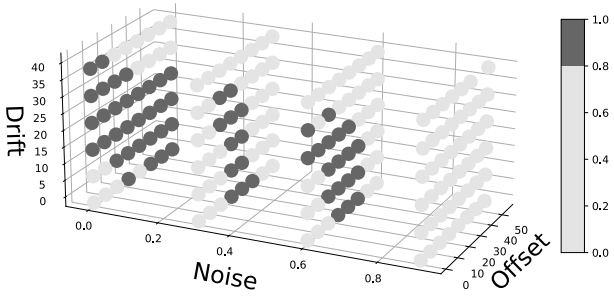

Fig. 9. Impact of $Offset$, $Noise$ and $Drift$ parameters on biometric system metrics.

the number of generated samples and the Y-axis—the F1 value.

As shown in Fig. 8, the value of the $Gain$ parameter has no significant effect on the classification results when the number of samples generated is low. On the other hand, for a higher number of generated data, information about speed of the subject is desirable and has a positive impact on the system performance. Although the gait cycles are interpolated to a fixed length in the preprocessing stage, some of the motion velocity information is retained in the accelerometer measurement signals. By modeling the gait speed within a 10% range, the accuracy of the biometric system could be increased. The Kruskal–Wallis test was used to statistically compare the results. Statistical significance of the augmentation results was obtained for each of the cases of the number of samples generated. The achieved p-values were $7.14 \cdot 10^{-5}$ (2 samples), $7.87 \cdot 10^{-5}$ (4 samples), $3.17 \cdot 10^{-5}$ (8 samples), $7.16 \cdot 10^{-6}$ (16 samples), $2.49 \cdot 10^{-6}$ (32 samples), $2.49 \cdot 10^{-6}$ (64 samples), $2.49 \cdot 10^{-6}$ (128 samples), and $2.49 \cdot 10^{-6}$ (256 samples).

In the ensuing research, the emphasis was placed on ascertaining which of the listed augmentation mechanism parameters were of greater significance, or what influence the advantage of the "Proposed II" configuration had over others. To this end, a grid search hyperparameter testing approach was employed. In each of the examined tests, a CNN with an attention mechanism was utilised, with the caveat that the experiments were limited to 10 repetitions due to the time-consuming nature of classifier training (Figure 9).

Each of the three axes is associated with the $Offset$, $Noise$ and $Drift$ parameters. The $Offset$ parameter was tested for the values {0.8, 16, 24, 32, 40, 48}, the $Noise$ parameter for the values {0.0, 0.3, 0.6, 0.9} and the $Drift$ parameter for the values {0.8, 16, 24, 32, 40}. The graph shows a scatterplot where the mean values above a 0.8 F1-score are indicated in grey.

Based on the presented results, it can be observed that, for a parameter $Noise$ equal to 0.9, an F1-score of 0.8 is not achieved. A high value of this parameter makes the signals of the augmented angular velocity very noisy (Fig. 4), which has a negative impact on the performance of the classifier. On the other hand, with a zero setting of the $Noise$ parameter, metrics above a 0.8 F1-score can be achieved. The $Noise$ parameter, like $Gain$, should take on negligible values and be used to add small perturbations.

Analysing the first "left-side layer" (results for $Noise$ equal to zero), an interesting relationship can be observed. It is possible to achieve very good identification metrics with a zero value for the $Offset$ parameter, provided that the $Drift$ parameter is at least 32. Conversely, it is possible to achieve very good results with a zero value for the $Drift$ parameter, provided that the $Offset$ parameter is at least 40. The tests carried out indicate the these parameters are critical. Their values should not be increased to exaggerated ones, and, for example, for the $Drift$ parameter, it is not recommended to exceed a value of 32.

A critical examination of the augmentation mechanisms reveals a salient question: *Does the number of generated samples influence the accuracy of the biometric system?* Within the domain literature, the number of newly generated samples exhibits considerable variation. For instance, Iso and Yamazaki (2006) report a

sample generation of 64, whereas Giacomo *et al*. (2017) limit it to a maximum of two.

Figure 10 shows a summary of the impact of the number of samples generated on the F1-score metrics. The graph was derived using 10 time repeated simple validation for a CNN with an attention mechanism. The horizontal axis presents the number of newly generated samples (ratio per the original gait number), and the vertical axis is the F1-score metrics. Using a box plot, a comparison of four data augmentation mechanisms is presented.

The results of the proposed method have a typical pattern. As the number of generated samples increases, the efficiency of the system grows. At the same time, a limitation can be observed: as the horizontal axis moves, the rate of change decreases. The Um method (Um *et al*., 2017), in its basic version, has a constant pattern for all tested cases. However, for cases $\{2, 4, 8, 16\}$, the obtained results are better than those for the proposed method. The method of Um *et al.* II, on the other hand, gives very low score results while worsening the baseline result. The addition of a permutation mechanism has a negative effect on data augmentation. The Delgado-Escaño method (Delgado-Escano *et al*., 2018) achieves generally good results, but, for the cases $\{32\}$ and $\{256\}$, significant instability in the results was observed. After meticulously examining the base case and generating 256 samples, the Kruskal–Wallis test yielded a p-value of $3.36 \cdot 10{-}5$. This indicates the discrepancy between the baseline and augmented results. The application of the Conover post-hoc tests resulted in p-values of less than 0.001 in all examined scenarios, with the exception of the Delgado-Escano *et al.* and Um *et al.* (version I) method pair (p-value < 0.01).

7. Evalutation of biometric systems involving mobile devices

In the next step, it was decided to jointly present the results for the Signet/University of Padova and 200GaitData/Boston University databases, which were created in the case of acquisition via an Android and an iPhone smartphone. First, it should be noted that, due to the different ways of positioning the sensor in the trouser pocket, it was necessary to model rotations in the range of $\pm 180°$,$\pm 90°$,$\pm 180°$ for the $Offset$ parameter. The preliminary trials conducted to ascertain the efficacy of the parameter settings derived directly from the BUT ALEK database yielded unsatisfactory results. A $Noise$ parameter of 0.1 was used in each of the analysed cases, and different gait speeds were modelled within a range of 5% ($Gain = 0.05$). The decision was made to undertake a comparison of two values of the parameter $Drift$, namely, 16 and 32.

Figure 11 demonstrates the results for the Signet database. The horizontal presents the number of augmented results and the vertical one axis shows the F1-score metric. Some common features with previous experiments can be observed. First, with a small number of generated samples, the Um method provides the highest results. An increase in the number of generated samples facilitates growth of the proposed augmentation method metric. Furthermore, the application of the Um method in version II typically yields suboptimal results, while the Delegado-Escaño approach occupies a median position in the ranking. The results of the proposed method were analysed by comparing two different settings. In the first one, a gait cycle was modelled with a $Drift$ disturbance of a large value of 50, and in the second setting—of 16. It was found that a parameter equal to 16 allowed better results to be achieved.

As illustrated in Table 6, the validation results are presented in a numerical format. Each row of the table corresponds to the mean F1-score calculated for the 10-time repeated cross-day validation. The columns represent the division by the argumentation method. The bottom of the table displays the results achieved in the base case without data augmentation.

The maximum score attained was 0.62 for the proposed augmentation method, while the baseline score was 0.46. Despite a substantial decline in the number of participants in the experiment, as compared to the BUT ALEK database (100 in contrast to 29), a discernible deterioration in the classification metrics is evident.

In the case of the Kruskal–Wallis test for the 256-sample generation scenario, a p-value of $2.0 \cdot 10^{-8}$ was reached when comparing the absence of augmentation and the occurrence of artificial samples. For the Conover post-hoc tests, no statistically significant difference was observed between the proposed metod with Drift:50 and the method of Um *et al.* (version I), wheras for the Drift:16 variant the difference was present.

In contrast, Fig. 12 presents the outcomes for the most challenging dataset, i.e., 200GaitDataset (Boston University). Contrary to previous experiments, the Um *et al.* method II (rotation and permutation) attains substantially superior results in comparison to the Delegado-Escaño approach. Conversely, there are no statistically significant differences between the proposed method and the Um *et al.* I one. For this particular database, no performance enhancement over competing methods was achieved for the proposed approach. In contrast to the previous experiment (Signet dataset), altering the $Drift$ parameter between the values {16, 50} does not have a significant impact on verification outcomes. For the generation of 128 samples (and the base case), the Kruskall–Wallis test indicated a p-value of 2.04×10-8, suggesting a statistical difference. Conversely, the Conover post hoc test revealed a statistical discrepancy between the proposed methods and the one

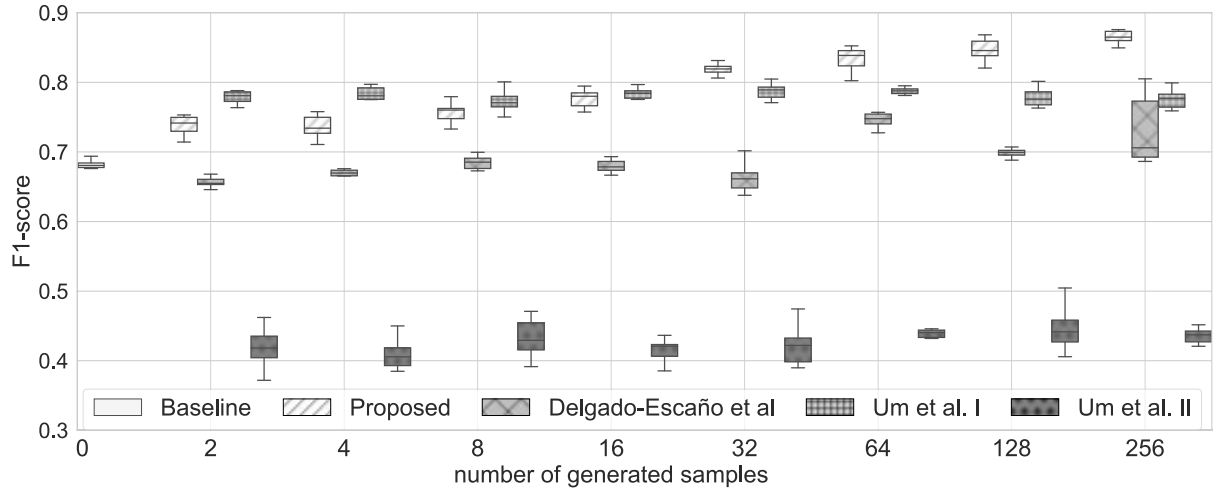

Fig. 10. Effect of the number of samples generated on the classification results of the BUT dataset.

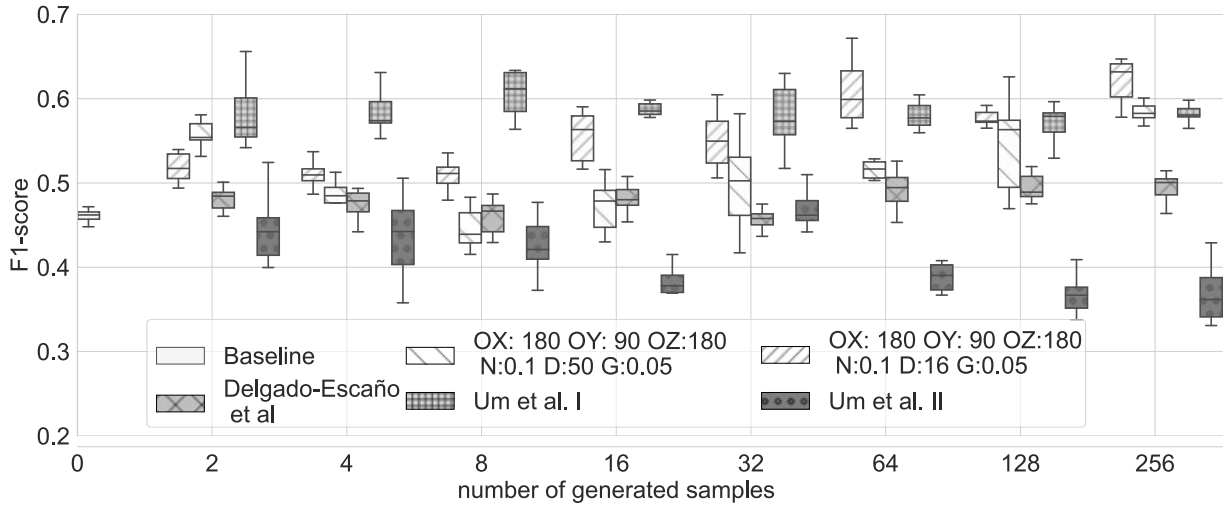

Fig. 11. Effect of the number of samples generated on the classification results of the Signet dataset.

that attained the highest score. However, no statistically significant variation was observed among the variants of the method proposed in our paper.

8. Implementation aspects

8.1. Quaternion algebra. It is important to acknowledge the utility of quaternion algebra, which stems from the effectiveness of this representation. The implementation of orientation perturbation methods (Fig. 1), signal transformations between reference systems, and synthetic generation of angular velocities was facilitated by the use of quaternions. While alternative approaches, such as matrix representations, could achieve similar functionalities, they would demand greater computational resources. The use of this form of representation required additional normalisation that was performed at the input of the Coordinate system change block (Fig. 1) in order to preserve orthonormality. This ensures that the signals are correctly transformed to the new coordinate system. However, some doubt about the creation of the quaternion differential $\dot{q}_t$ (Eqn. (4)) needs to be dispelled. In practice, a signal is created whose magnitude depends on the sampling period. Note that the determined signal $\dot{q}_t$ is used to create the angular velocity signal ω_t (Eqn. (2)). In this case, it is not advisable to normalize the dq signal Eqn. (4), since this would artificially limit the angular velocity values (Eqn. (2)) to the (-2 rad/s

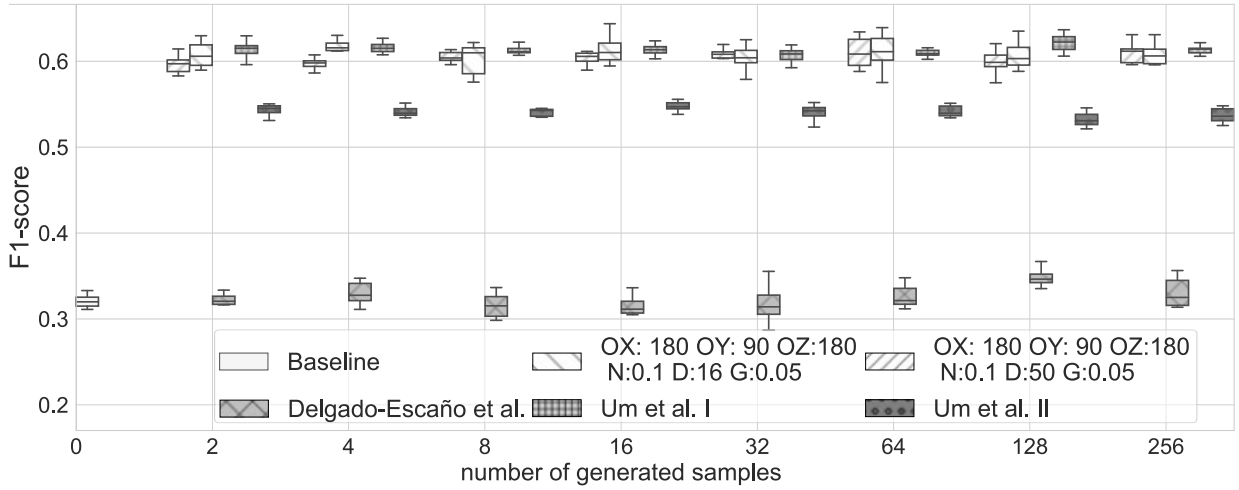

Fig. 12. Effect of the number of generated samples on the classification results of the Boston dataset.

Table 6. F1-score metrics for the Signet dataset.

	Um *et al.* I	Um *et al.* II	Delgado-Escaño	Proposed Drift:16	Proposed Drift:50
2	**0.58**	0.44	0.48	0.52	0.52
4	**0.58**	0.44	0.47	0.51	0.51
8	**0.61**	0.42	0.46	0.51	0.51
16	**0.59**	0.38	0.48	0.56	0.56
32	**0.59**	0.47	0.46	0.55	0.55
64	0.58	0.38	0.49	**0.61**	0.51
128	**0.57**	0.36	0.50	**0.57**	0.55
256	0.58	0.37	0.49	**0.62**	0.58
0 (Base.)	0.46				

Table 7. F1-score metrics for 200GaitData.

	Um *et al.* I	Um *et al.* II	Delgado–Escaño	Proposed Drift:16	Proposed Drift:50
2	**0.61**	0.54	0.32	0.60	0.60
4	**0.62**	0.54	0.33	0.61	0.60
8	**0.61**	0.54	0.32	0.60	0.60
16	**0.61**	0.55	0.32	**0.61**	0.60
32	**0.61**	0.52	0.32	**0.61**	0.59
64	**0.61**	0.53	0.33	**0.61**	**0.61**
128	**0.62**	0.53	0.35	0.61	0.60
256	**0.61**	0.54	0.32	**0.61**	**0.61**
0 (Base.)	0.32				

to 2 rad/s) range. It should be noted that, in certain publicly available implementations (including (Stanley, 2016), $\dot{q}_t$ normalization is not performed (see line 12 of the differentiate.py file). Furthermore, the formula corresponding to (Eqn. (4)) describing the relationship between the orientation q_t and the quaternion differential $\dot{q}_t$ is presented by (Kou and Xia, 2017, Section 1.1.2), whlist emphasizing the need for a constant motion speed during the sampling period.

8.2. Limitations on the number of participants in the datasets. The necessity of recordings being made over two days and the availability of orientation signals constituted significant limitations. Consequently, the utilization of numerous publicly accessible databases was impeded. To address these challenges, a decision was made to construct an authoritative corpus of data, thereby enabling the execution of an evaluation of the developed augmentation method.

In the context of public collections utilizing mobile phones, concerns regarding the transferability of the solution have been raised (due to the number of approximately 30 participants). It is acknowledged that the implementation of additional subjects would enhance the efficacy of the tests. In further research, we anticipate using databases with a larger number of participants. It should be noted, however, that the number of labels used in the literature is typically on the order of 20–30,

as illustrated by Lee *et al.* (2022) with 20 labels and Luo *et al.* (2020) with 30. There are publications in the VR behavioural biometrics literature examining the effect of the number of participants on the accuracy of a biometric system (Pfeuffer *et al.*, 2019). In this study, the number of participants was varied from two to 18, with marginal degradation in accuracy for a 16 to 18 change. An analogous experiment is planned for behavioral gait biometrics for our gait dataset.

At this point, it is worth noting again that the laboratory approach yielded baseline rates of a 0.68 F1-score (Table 5) for 100 participant, and a 0.46 F1-score for 29 participants in the quasi-real-life conditions (Table 6). The fact that lower results were achieved despite three times fewer participants under quasi-real-life conditions highlights the problems associated with applicability. A broader analysis was not made because of the footwear used for individual participants (available description for our base). Such an approach would have reduced the number of labels and artificially increased the rates of biometric systems.

9. Conclusions

The research presented in this paper proposed a novel method of augmenting the gait cycle that has applications in gait-based biometric identification systems. The paper detailed a dedicated augmentation mechanism for motion sensor applications that is based on a set of an accelerometer and a gyroscope. The conducted experiments present an augmentation mechanism that uses the modeling of three types of perturbation: different initial sensor alignment, vibration, and slow tilt of the sensor during the motion. The novel approach proposed here involves a two-part process. Firstly, the orientation signal is disturbed. Secondly, using the perturbed orientation signal, accelerometer measurement values are modelled, and angular velocity measurements (gyroscope measurement values) are synthetically generated. An additional component is added to model the effect of gait speed on accelerometer readings.

The augmentation mechanism was evaluated by contrasting the results with those obtained from state-of-the-art methods. The validation of the solution was carried out for three data corpora: the BUT ALEK/Bialystok University of Technology lab-like condition dataset (100 subjects, approximately 30 gait cycles per person), a quasi-real Signet/Padova University set (29 subjects, approximately 300 gait cycles per participant), and a real condition one, i.e., 200GaitDataset/Boston University (29 subjects, approximately 400 gait cycles per participant).

Depending on the adopted method of data acquisition, varied identification metrics were achieved, summarizing the case for a CNN classifier with an attention mechanism. In cases without data augmentation, the results of F1-scores for the laboratory, quasi-real, and real conditions were 0.68, 0.46, and 0.32, respectively. The use of data augmentation increased the verification measures to 0.82, 0.62, and 0.62 F1-scores.

Data augmentation allowed the largest increase in verification metrics for our corpus. The dataset was characterized by a relatively large group of participants and a relatively small number of gait cycles per person. Augmentation increased the verification performance from a baseline of 0.53 to 0.86 for the multi-input CNN classifier, and 0.68 to 0.82 for the CNN with an attention mechanism (Table 5).

For the datasets created with mobile phones, the increase of F1-score metrics was not as significant as in laboratory scenarios. In the case of the Signet/University of Padova dataset, the F1-score increased from 0.46 to 0.62. In the case of the 200GaitDtaset/Boston University database, the change ranged from 0.32 to 0.61.

The proposed data augmentation mechanism achieved the highest F1-score identification for the ALEK/Bialystok University of Technology dataset and the Signet/University of Padova corpus. On the other hand, in the case of real conditions, it provided us with results similar to the method of Um *et al* (2017). In the conducted experiments, it was observed that it is the modelling of the initial alignment ($Offset$) and the displacement of the sensor during movement ($Drift$) that have a decisive influence on the verification rates (Fig. 9), with the parameter modelling the interference (fast-variable rotation). Even though the gait cycle in the experiments carried out was normalized to a fixed length (so as to constitute the input of the classifiers), the information on gait speed is implicitly preserved in the accelerometer measurement signals. Thanks to the modelling of the value of the ground acceleration and the possibility of extracting the acceleration resulting from motion, it was possible to amplify it. The experiments (Fig. 8) show that walking speed can be an indirect biometric trait. We recommend modelling gait speed within small limits, e.g., 5%. Modeling gait speed within 20% negatively affects the properties of the biometric system.

The process of data augmentation should be carried out thoughtfully so that the new samples generated reflect observable data. This coincides with the literature requirements of data augmentation methods (Kim and Jeong, 2021). In the case of mobile phone datasets, it is important that the sensor can be located in a trouser pocket in several different ways. Generating data in an identical way as in the Bialystok University of Technology dataset covers observable cases poorly. Therefore, in the case of these datasets, it was necessary to model the $Offset$ parameter with much larger values.

It should also be noted that, in the case of the ALEK

and Signet datasets, increasing the number of generated samples has a positive impact on verification rates (Figs. 10 and 11). The highest rates in the experiments were achieved by generating 256 artificial samples per original. This is due to the assignment of numerical ranges for the augmentation mechanism, rather than a specific numerical value. Setting the $Offset$ parameter to 180° will draw a number in the range $\pm 180^\circ$. As the number of generated samples grows, the possibility of generating boundary samples increases.

It is important to note the different results for acquisitions involving mobile phones—for the quasi-real conditions (Signet/University of Padova) and real-life conditions (200GaitDataset/Boston University) datasets. The application of the Um *et al.* II method, which has an additional signal permutation component, scientifically changed the original signals. This transformation in the quasi-real dataset was shown to have a detrimental effect on verification rates. Conversely, for real-life condition datasets, it was demonstrated that this transformation can result in the achievement of relatively high metrics. In contrast, the Delgado-Escaño method (which does not introduce drastic signal changes) in real-life conditions datasets allows only the lowest verification rates to be obtained. In addition, changing the Drift parameter between the $\{16, 50\}$ values does not significantly affect verification rates in that set.

These experiments indicate that, in the case of fully real conditions (varying shoe type and varying ground type), the recorded signals have a significant distribution shift. The validation set samples are so dissimilar from the training set that even a very large distortion parameter is not able to decrease the verification rates.

Finally, it should be noted that the 200GaitData database, in which both the type of substrate and the type of footwear were changed, features a rather low application potential. In a real application that can potentially be implemented in real-life scenarios, it is expected that dedicated corridors with fixed length and fixed substrate type can be built. In our view, developing solutions in which individuals provide gait samples within a single session (and within that session they are validated) also misses the mark for potential real-world solutions. The subsequent research will be conducted within two primary domains. First, work will be done on enhancing the verification indicators within the Signet dataset. Concurrently, we intend to undertake pilot studies to assess the efficacy of the developed methodologies in real-time data processing scenarios.

Acknowledgment

This work was supported by the grant WZ/WI-IIT/5/2023 of the Bialystok University of Technology and funded with resources for research by the Ministry of Science and Higher Education in Poland, as well as the grant 2021/41/N/ST6/02505 of the Bialystok University of Technology and funded with resources for research by the National Science Centre, Poland. For the purpose of Open Access, the author has applied a CC BY-NC-ND public copyright license to any Author Accepted Manuscript (AAM) version arising from this submission.

References

Abate, A.F., Nappi, M. and Ricciardi, S. (2017). I-Am: implicitly authenticate me—Person authentication on mobile devices through EAR shape and ARM gesture, *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **49**(3): 469–481.

Alzantot, M., Chakraborty, S. and Srivastava, M. (2017). Sensegen: A deep learning architecture for synthetic sensor data generation, *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Kona, USA*, pp. 188–193.

Bieda, R., Grygiel, R. and Galuszka, A. (2015). Naive Kalman filtering for estimation of spatial object orientation, *2015 20th International Conference on Methods and Models in Automation and Robotics (MMAR), Międzyzdroje, Poland*, pp. 955–960.

Boyd, J.E. and Little, J.J. (2005). Biometric gait recognition, *Advanced Studies in Biometrics: Summer School on Biometrics, Alghero, Italy,* pp. 19–42, DOI: 10.1007/114936482.

Carneros-Prado, D., Dobrescu, C.C., Cabañero, L., Villa, L., Altamirano-Flores, Y.V., Lopez-Nava, I.H., González, I., Fontecha, J. and Hervás, R. (2024). Synthetic 3D full-body skeletal motion from 2D paths using RNN with LSTM cells and linear networks, *Computers in Biology and Medicine* **180**: 108943, DOI: 10.1016/j.compbiomed.2024.108943.

Chi, W., Wang, J. and Meng, M. Q.-H. (2017). A gait recognition method for human following in service robots, *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **48**(9): 1429–1440.

Darwish, S.M. (2017). Design of adaptive biometric gait recognition algorithm with free walking directions, *IET Biometrics* **6**(2): 53–60.

Dehkordi, N.R. and Farahmand, S. (2022). Gait Database, `https://figshare.com/articles/dataset/Gait_Database/20346852`.

Delgado-Escano, R., Castro, F.M., Cozar, J.R., Marin-Jimenez, M.J. and Guil, N. (2018). An end-to-end multi-task and fusion CNN for inertial-based gait recognition, *IEEE Access* **7**: 1897–1908, DOI: 10.1109/ACCESS.2018.2886899.

Diebel, J. (2006). Representing attitude: Euler angles, unit quaternions, and rotation vectors, *Matrix* **58**(15–16): 1–35.

Gadaleta, M. and Rossi, M. (2018). IDNet: Smartphone-based gait recognition with convolutional neural networks, *Pattern Recognition* **74**: 25–37, DOI: 10.1016/j.patcog.2017.09.005.

Giacomo, G., Martinelli, F., Saracino, A. and Alishahi, M. (2017). Try walking in my shoes, if you can: Accurate gait recognition through deep learning, *Proceedings of the International Conference on Computer Safety, Reliability, and Security, Trento, Italy*, pp. 12–15.

Guzsvinecz, T., Szucs, V. and Sik-Lanyi, C. (2019). Suitability of the kinect sensor and leap motion controller—A literature review, *Sensors* **19**(5): 1072.

Hoang, T., Choi, D. and Nguyen, T. (2015). On the instability of sensor orientation in gait verification on mobile phone, *SECRYPT 2015: Proceedings of the 12th International Conference on Security and Cryptography, Colmar, France*, pp. 148–159.

Huang, H., Zhou, P., Li, Y. and Sun, F. (2021). A lightweight attention-based CNN model for efficient gait recognition with wearable IMU sensors, *Sensors* **21**(8): 2866.

Iso, T. and Yamazaki, K. (2006). Gait analyzer based on a cell phone with a single three-axis accelerometer, *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services, Helsinki, Finland*, pp. 141–144.

Kim, M. and Jeong, C.Y. (2021). Label-preserving data augmentation for mobile sensor data, *Multidimensional Systems and Signal Processing* **32**(1): 115–129.

Kitowski, Z., Piskur, P. and Orłowski, M. (2023). Dual quaternions for the kinematic description of a fish-like propulsion system, *International Journal of Applied Mathematics and Computer Science* **33**(2): 171–181, DOI: 10.34768/amcs-2023-0013.

Kou, K.I. and Xia, Y.-H. (2017). Linear quaternion differential equations: Basic theory and fundamental results, *arxiv* 1510.02224.

Kwon, H., Tong, C., Haresamudram, H., Gao, Y., Abowd, G.D., Lane, N.D. and Ploetz, T. (2020). IMUTube: Automatic extraction of virtual on-body accelerometry from video for human activity recognition, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **4**(3): 1–29.

Lee, S., Lee, S., Park, E., Lee, J. and Kim, I.Y. (2022). Gait-based continuous authentication using a novel sensor compensation algorithm and geometric features extracted from wearable sensors, *IEEE Access* **10**: 120122–120135, DOI: 10.1109/ACCESS.2022.3221813.

Luo, Y., Coppola, S., Dixon, P., Li, S., Dennerlein, J. and Hu, B. (2020). A database of human gait performance on irregular and uneven surfaces collected by wearable sensors, *Scientific Data* **7**: 219, DOI: 10.1038/s41597-020-0563-y.

Maeda, T. and Ukita, N. (2022). Motionaug: Augmentation with physical correction for human motion prediction, *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, USA,* pp. 6417–6426.

Microsoft (2015). *Kinect for Windows v2 Sensor, Microsoft, Redmond,* https://web.archive.org/web/20150522142846/http://www.microsoft.com/en-us/kinectforwindows/meetkinect/default.aspx.

NXP (2016). NXP Sensor Fusion Library for Kinetis MCUs, Rev. 0.8, NXP, Eidhoven, https://www.nxp.com/docs/en/data-sheet/NSFK_DS.pdf.

Ohashi, H., Al-Nasser, M., Ahmed, S., Akiyama, T., Sato, T., Nguyen, P., Nakamura, K. and Dengel, A. (2017). Augmenting wearable sensor data with physical constraint for DNN-based human-action recognition, *ICML 2017 Times Series Workshop, Sydney, Australia*, pp. 6–11.

Pellatt, L., Dewar, A., Philippides, A. and Roggen, D. (2021). Mapping vicon motion tracking to 6-axis IMU data for wearable activity recognition, *in* M.A.R. Ahad *et al.* (Eds), *Activity and Behavior Computing*, Springer, Singapore, pp. 3–20.

Noitom (2019). *Perceptrion Neuron 32 Official Website*, Noitom International Miami, https://web.archive.org/web/20190116201850/https://neuronmocap.com/products/perception_neuron

Pfeuffer, K., Geiger, M., Prange, S., Mecke, L., Buschek, D. and Alt, F. (2019). Behavioural biometrics in VR: Identifying people from body motion and relations in virtual reality, *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI'19), Glasgow, UK*, pp. 1–12, DOI: 10.1145/3290605.3300340.

Sawicki, A. (2022). Augmentation of accelerometer and gyroscope signals in biometric gait systems, *International Conference on Computer Information Systems and Industrial Management, Barranquilla, Columbia*, pp. 32–45.

Sawicki, A., Walendziuk, W. and Idzkowski, A. (2016). The gravitational acceleration components elimination from the accelerometer measurement data, *Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2016, Wilga, Poland*, pp. 485–494.

Sawicki, A.and Saeed, K. (2021). Application of LSTM networks for human gait-based identification, *in* W. Zamojski *et al.* (Eds), *Theory and Engineering of Dependable Computer Systems and Networks*, Springer International Publishing, Cham, pp. 402–412.

SIGNET (2016). *SIGNET Smartphone Inertial Signals for User Identification—DATASET*, University of Padova, Padova, https://signet.dei.unipd.it/research/human-sensing.

Sprager, S. and Juric, M.B. (2015). Inertial sensor-based gait recognition: A review, *Sensors* **15**(9): 22089–22127.

Steven Eyobu, O. and Han, D.S. (2018). Feature representation and data augmentation for human activity classification based on wearable IMU sensor data using a deep LSTM neural network, *Sensors* **18**(9): 2892.

Subramanian, R., Sarkar, S., Labrador, M., Contino, K., Eggert, C., Javed, O., Zhu, J. and Cheng, H. (2015). Orientation invariant gait matching algorithm based on the Kabsch alignment, *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2015), Hong Kong, China*, pp. 1–8.

Stanley, M. (2016). *Trajectory and Sensor Simulation Toolkit, Rev. 1.4*, https://github.com/memsindustrygroup/TSim.

Tran, L. and Choi, D. (2020). Data augmentation for inertial sensor-based gait deep neural network, *IEEE Access* **8**: 12364–12378, DOI: 10.1109/ACCESS.2020.2966142.

Um, T.T., Pfister, F.M., Pichler, D., Endo, S., Lang, M., Hirche, S., Fietzek, U. and Kulic (2017). Data augmentation of wearable sensor data for Parkinson's disease monitoring using convolutional neural networks, *Proceedings of the 19th ACM International Conference on Multimodal Interaction, Glasgow, UK*, pp. 216–220.

Vajdi, A., Zaghian, M.R., Dehkordi, N.R., Rastegari, E., Maroofi, K., Farahmand, S., Jia, S., Pomplun, M., Haspel, N. and Bayat, A. (2019). Human gait database for normal walk collected by smartphone accelerometer, *arXiv* 1905.03109.

Wan, C., Wang, L. and Phoha, V.V. (2018). A survey on gait recognition, *ACM Computing Surveys* **51**(5): 1–35.

Whittle, M.W. (2014). *Gait Analysis: An Introduction*, Butterworth-Heinemann, Oxford.

Wieczorek, M., Wojtas, N., Wituła, R., Krawczyk, A. and Rycerz, K. (2024). A custom deep learning architecture with image augmentation for intelligent gastrointestinal tract tissue classification, *International Journal of Applied Mathematics and Computer Science* **34**(4): 597–616, DOI: 10.61822/amcs-2024-0040.

Xia, C., Chen, X., Huang, L., Min, T. and Zhang, D. (2025). Open-source virtual IMU sensor platform for developing a customized wearable human activity recognition system, *IEEE Transactions on Instrumentation and Measurement* **74**: 1–14, Article no. 9508514, DOI: 10.1109/TIM.2025.3548063.

Zou, Q., Wang, Y., Wang, Q., Zhao, Y. and Li, Q. (2020). Deep learning-based gait recognition using smartphones in the wild, *IEEE Transactions on Information Forensics and Security* **15**: 3197–3212, DOI: 10.1109/TIFS.2020.2985628.

Aleksander Sawicki is a researcher at the Department of Digital Media and Computer Graphics of the Faculty of Computer Science, Bialystok University of Technology, Poland. His research interests cover IMU signal processing and generative models. His work in this area particularly concerns behavioural biometrics, with a special focus on gait recognition using motion sensors.

Khalid Saeed is a full professor of computer science at the Bialystok University of Technology and a half-time visiting professor at Universidad de La Costa, Barranquilla, Colombia. He was with the Warsaw University of Technology in 2014–2019 and with the AGH University of Krakow in 2008–2014. He received his BSc degree from Baghdad University in 1976, and his MSc and PhD degrees from the Wroclaw University of Technology in Poland in 1978 and 1981, respectively. He obtained his DSc degree in computer science from the Polish Academy of Sciences in Warsaw in 2007. He was granted the professorial title by the President of Poland in 2014. He has authored more than 300 publications.

Received: 18 January 2025
Revised: 24 April 2025
Re-revised: 25 June 2025
Accepted: 25 July 2025