Signal enhancement and efficient DTW-based comparison for wearable gait recognition

Danilo Avola [a], Luigi Cinque [a], Maria De Marsico [a], Alessio Fagioli [a], Gian Luca Foresti [b], Maurizio Mancini [a], Alessio Mecca [a],*

[a] *Sapienza, University of Rome, Department of Computer Science, Via Salaria 113, Rome, 00169, Italy*
[b] *University of Udine, Department of Mathematics, Computer Science, and Physics, Via delle Scienze 206, Udine, 33100, Italy*

ARTICLE INFO

ABSTRACT

The popularity of biometrics-based user identification has significantly increased over the last few years. User identification based on the face, fingerprints, and iris, usually achieves very high accuracy only in controlled setups and can be vulnerable to presentation attacks, spoofing, and forgeries. To overcome these issues, this work proposes a novel strategy based on a relatively less explored biometric trait, i.e., gait, collected by a smartphone accelerometer, which can be more robust to the attacks mentioned above. According to the wearable sensor-based gait recognition state-of-the-art, two main classes of approaches exist: 1) those based on machine and deep learning; 2) those exploiting hand-crafted features. While the former approaches can reach a higher accuracy, they suffer from problems, e.g., performing poorly outside the training data, i.e., lack of generalizability. This paper proposes an algorithm based on hand-crafted features for gait recognition that can outperform the existing machine and deep learning approaches. It leverages a modified Majority Voting scheme applied to Fast Window Dynamic Time Warping, a modified version of the Dynamic Time Warping (DTW) algorithm with relaxed constraints and majority voting, to recognize gait patterns. We tested our approach named MV-FWDTW on the ZJU-gaitacc, one of the most extensive datasets for the number of subjects, but especially for the number of walks per subject and walk lengths. Results set a new state-of-the-art gait recognition rate of 98.82% in a cross-session experimental setup. We also confirm the quality of the proposed method using a subset of the OU-ISIR dataset, another large state-of-the-art benchmark with more subjects but much shorter walk signals.

1. Introduction

Over the past decade, advancements in Computer Science have driven significant innovations across various research domains. For example, Sahu et al. (2020, 2023) have defined decision-making frameworks for building robust web applications and selecting the best renewable energy sources when faced with ambiguous or incomplete data. Different application areas have instead been explored by Gu et al. (2020) and Avola et al. (2022a, 2021), focusing on the explainability of deep learning approaches applied to medical imaging analysis, which is critical for handling sensitive data. To ensure the effectiveness of these advancements, addressing software reliability is essential in developing new solutions. Sahu et al. (2021) suggest models for predicting a system's reliability over time. Furthermore, the works by Sahu and Srivastava (2018, 2020, 2021) also analyze software reliability in industrial and medical scenarios, focusing on identifying bugs in large

datasets. Another pressing concern regarding data is security, as discussed by Attaallah et al. (2022), which can lead to potential breaches Almulihi et al. (2022). These issues are of particular importance in specific fields and application areas, such as person re-identification and monitoring Avola et al. (2022b); Shao et al. (2021), and biometrics Fan et al. (2017). As a matter of fact, the latter field has gained popularity in recent years, with applications in personal security. Indeed, modern smartphones can be unlocked by fingerprint images captured by dedicated sensors or facial recognition through embedded cameras, while banks often use the former to authorize monetary transactions.

Among biometrics, it is possible to distinguish between two families of traits, i.e., hard and soft. The former are generally related to physical characteristics and should present properties, i.e., universality, uniqueness, and permanence, desired for a robust and solid identification of a human being. Some examples are the iris Bowyer and Burge (2016); Nguyen et al. (2017), the already mentioned fingerprints Li

et al. (2021); Marasco and Ross (2014), and the face Guo and Zhang (2019); Wang and Deng (2021). However, such traits are prone to spoofing and forgery. Soft biometrics are instead related either to behavioral characteristics, which can be less reliable due to relative lack of permanence in the long term, or to some physical traits that may lack in one or more of the previously mentioned properties, especially uniqueness as in the case of hair color or face shape. Soft biometrics of the second group can be successfully exploited to identify a class of users, e.g., by hair color or gender, for supporting hard biometrics for identification in a multi-biometric setting Maity et al. (2021). On the other hand, behavioral traits are generally more challenging to forge or spoof.

Gait is a behavioral trait that, as the voice, is also influenced by the physical structure of the subject's body. Therefore, it presents some characteristics of hard biometrics, due to the high characterization of personal kinematic strategies, and some from the behavioral traits, such as the relative lack of permanence. It cannot be considered truly strong but still offers the advantages of soft traits, i.e., it is very robust against forgery and spoofing Muaaz and Mayrhofer (2017). Gait recognition has been studied following three approaches: 1) methods using floors equipped with pressure sensors Takeda et al. (2009), that offer a low feature resolution and are presently limited to gait analysis for diagnosis and rehabilitation; 2) methods processing video streams Singh et al. (2021), that are presently the most popular ones but suffer from typical problems of computer vision, such as intra-class variations due to different poses, illumination, occlusion, and perspective, and from trajectory crossings; 3) more recent methods using built-in sensors, usually accelerometers, of mobile devices such as smartphones and smartwatches Xu et al. (2017). This third category is less explored since the need for one or more worn devices entails the users' awareness and collaboration. However, this can be taken for granted when using gait for authentication.

This paper presents a strategy for gait recognition from a single wearable sensor, an accelerometer, available in any smartphone. The designed system is suited for a real-life scenario where a registered user can access a restricted area simply by walking toward it, as proposed in De Marsico et al. (2019); Mecca (2018). As for De Marsico et al. (2019), it is possible to configure a mobile application to capture the signal from two beacons that automatically trigger the start and stop of data acquisition along a hallway to a restricted area. Incidentally, it is possible to capture data with similar lengths and, therefore, easier to compare by a recognition server. The mobile application will send the acquired data to such a server to identify the user. The main advantages of this approach are that it does not require any specific user cooperation, except for installing the application, and that the system can be used as an independent module when considering a multi-biometric setting (e.g., using a camera at the end of the hallway to capture one or more video frames to perform face recognition or to extract other visual features useful to improve the overall recognition accuracy).

Currently, two main groups of methods exist for performing gait recognition from wearable sensor data: 1) those based on machine learning, also via deep architectures; 2) those exploiting hand-crafted features. Machine and deep learning are widely used to solve almost any computational problem. On the one hand, they can outperform algorithms using hand-crafted features without specific training, including those tackling gait recognition, as reported in Giorgi et al. (2017). On the other hand, these techniques suffer from typical problems, e.g., requiring a really huge amount of training data, and performing poorly outside the benchmark from which the training data originated O'Mahony et al. (2020). In the wearable-based gait recognition context, the lack of data is one of the main issues since there are really few datasets with a sufficient number of walk signals. Therefore, it still seems appropriate to investigate algorithms using hand-crafted features, possibly without training, to effectively solve problems in real-life scenarios, as in Van Gastel et al. (2015) and Ameur et al. (2019). Similarly, this paper shows how gait recognition can be improved by suitably devised improvements to the basic algorithms in the existing literature.

To this aim, it presents the MV-FWDTW algorithm. It joins a modified Majority Voting scheme with a novel version of the existing DTW algorithm, the Fast Window Dynamic Time Warping. As shown in the paper, MV-FWDTW achieves a recognition rate of 98.82% in a cross-session experimental setup. It reaches a 100% score at Rank #4 when considering the Cumulative Match Characteristic (CMC) curve, even with more than 150 identities, thus outperforming existing deep learning approaches for gait recognition.

The presented MV-FWDTW algorithm compares the accelerometer time series for gait recognition. It is specifically designed to account for relaxed constraints, providing a faster and more accurate version of the DTW procedure and exploiting majority voting to increase the recognition rate. It is worth noting that, unlike machine and deep learning-based approaches, MV-FWDTW does not require any training. The evaluation of the proposed method relied on extensive experiments on the ZJU-gaitacc benchmark dataset, whose results demonstrate the effectiveness of the proposed pipeline. Moreover, it has been tested on a subset of the OU-ISIR Ngo et al. (2014) dataset, a large state-of-the-art benchmark with more than 700 subjects but shorter walk signals, confirming its capability.

The paper contributions are as follows:

- a novel pre-processing strategy to de-noise and enhance signals from wearable sensors, exploiting state-of-the-art procedures and especially focusing on gait signals;
- a gait recognition module using a single wearable sensor suited for real indoor or outdoor scenarios, avoiding all the typical drawbacks of camera-based systems.
- a fast and accurate algorithm for gait recognition, MV-FWDTW, that does not require any training phase and proves its effectiveness on two large state-of-the-art datasets, namely ZJU-gaitacc and OU-ISIR.

In the rest of the paper, Section 2 reviews the main related work on wearable sensor-based gait recognition. Section 3 describes the proposed approach, focusing on both the pre-processing of signals and the comparison strategies. Section 4 details the datasets used to test the strategy, the experimental setup, and the achieved results. Finally, Section 5 draws some conclusions and possible future research directions.

2. Related work

Wearable sensor-based gait recognition mostly relies on data acquired from accelerometers (even if some proposals also use gyroscope data). A generic walk signal can be represented as $W = \{X, Y, Z\}$, where X, Y, and Z are time series containing data from the three accelerometer axes. Recognition strategies dealing with this kind of data can be divided into two categories: those relying on signal processing/comparison and those exploiting neural networks. The approaches in both categories can be further subdivided into those splitting the walks into single steps (the portion of a gait cycle that includes the movement of a single leg only) or cycles (sequence of right and left steps), and those using fixed-length frames De Marsico and Mecca (2019). The division into steps or cycles is more complex to automate. It requires fine-tuned strategies to be effective due to the difficulty of accurately identifying the actual beginning and ending of a step/cycle. However, this signal subdivision is more related to the intrinsic cycling nature of walk dynamics and usually produces better results, especially in the case of signal processing/comparison strategies and in presence of walk signals with a different number of steps. After segmentation, these parts (or relevant feature vectors of prevailing statistical nature extracted from them) are compared. Alternative methods use the entire signal without any kind of segmentation. The following subsections only consider few strategies dealing with the main dataset tested in our proposals, namely the ZJU-gaitacc Zhang et al. (2014). Interested read-

ers can refer to the surveys in literature De Marsico and Mecca (2019); Wan et al. (2018) for a more comprehensive overview.

2.1. Signal processing/comparison strategies

The proposal in Sun et al. (2018) presents a speed-adaptive authentication procedure based on Pearson Correlation Coefficient (PCC). Its first step is cycle segmentation. The 1-minute walk is processed through the Fast Fourier Transform (FFT) to estimate the cycle length during enrollment. The cycles are extracted starting from the first local minima and then looking for the following one with a displacement based on the previously estimated cycle length. During the testing phase, the input walk is divided into 8-second frames with an overlap of 4 seconds before applying the cycle segmentation. In both cases, the estimated cycles are normalized and interpolated to a fixed length to create a template. Recognition proceeds by comparing the templates from enrolled users and the probe (the unknown template to match) using the PCC.

The proposal in Mecca (2018) presents a strategy based on DTW that exploits the concept of gait stabilization. Such a phenomenon describes the natural stabilization of the human walk pattern after a certain number of steps. For this principle, the first and the last steps in a walk tend to be slightly different from the central ones because the starting (or ending) of the locomotor strategy always requires a voluntary action (see Fernandez-Lopez et al. (2017)). Therefore, the first phase of the procedure relies on step segmentation, which is based on the estimation of two parameters. The first one is the $10th$ value of the list, in descending order, of the relative maxima on the Y axis. The second parameter is computed as $\mu - \sigma$, where μ and σ are the average and standard deviation of the acceleration values on the same axis. The segmentation exploits these two parameters to find each step's start and end and avoid choosing spikes due to noise. After segmentation, the first k steps are discarded, with $k = 1, ..., 5$, as well as the last one, and the remaining part of the walk signal is used as a template for the user. The same procedure is applied to the incoming probes. The recognition relies on the basic formulation of the DTW algorithm. The best results are achieved discarding the highest number of steps.

2.2. Neural network-based strategies

The strategy proposed in Giorgi et al. (2017) is based on a Deep Convolutional Neural Network (CNN). The data processing phase comprises cycle extraction, filtering, and normalization. The first step aims at inferring the start and the end of cycles by looking for peaks in the magnitude vector mag, computed as $mag_i = \sqrt{x_i^2 + y_i^2 + z_i^2} \quad \forall i$. The second one enhances the signal quality using the low-pass Butterworth filter. The last step normalizes the cycles by linear interpolation. The obtained cycles are then used to train a Deep CNN. The proposal reports very high accuracy results; however, the network has been trained and tested with data from the same session. This is a more straightforward yet unrealistic scenario unless aiming at re-identification only (or short-term recognition).

The work in Nemes and Antal (2021) presents strategies based on extracting and testing different features. During the pre-processing, walking signals are normalized in the $[-1, 1]$ interval and then subdivided into cycles or time-fixed frames (different sizes in separate tests) by exploiting the dataset annotations. However, such annotations are handmade, so this is not suitable for new walks in a real-life scenario. The procedure then extracts 3 groups of features, namely: the raw data (concatenating X, Y, and Z axes), a set of 59 statistical features from the magnitude vector, and a set of 64 features extracted by the encoder part of two different convolutional autoencoder networks. The classification relies on the Random Forest algorithm. The results confirm that the performance achieved by using the cycles is higher than by using fixed-time frames. Moreover, as expected, using data from the same session for training and testing increases the performance, with the above-mentioned scope limitations.

3. Method

Let $W_i^{id} = \{X, Y, Z\}$ be the *i-th* walk belonging to the subject id, where $X = \{x_1, x_2, ..., x_n\}$, $Y = \{y_1, y_2, ..., y_n\}$, and $Z = \{z_1, z_2, ..., z_n\}$ are three one-dimensional time series with the same length n. In this representation, the generic x_i, y_i, and z_i are the values collected by the accelerometer over the three space dimensions X, Y, and Z, respectively, in the time instant i. Since each person walks at a different pace, the walk length n to cover a similar distance can differ from subject to subject, but it can also be different for walks of the same person considering, e.g., the walking speed. The length difference can be a practical problem in the case of neural network-based approaches, where an identical length is often, if not always, required. Cutting the walks to have the same length is possible, but some helpful information and patterns could be lost. The proposed methodology uses a different strategy to overcome this limitation. The procedure entails two independent tasks carried out by the pre-processing module and the comparison module.

3.1. Pre-processing module

The pre-processing module first identifies the relevant portion of the signal. In fact, when a human starts walking from a still stance, the first portion of the signal usually contains a lot of noisy data. Afterward, the module enhances the signal data quality by removing the overall noise to improve the results provided by the comparison module.

3.1.1. Retrieval of the relevant signal portion

This phase is based on the idea proposed in De Marsico and Mecca (2015, 2017). The procedure chooses the start and the end points of the relevant portion of the walk by exploiting a threshold T computed over the values on the Y axis only. In fact, due to the intrinsic nature of the human walk, the vertical oscillations (in this case, those captured by the Y accelerometer axis) are empirically demonstrated to be more discriminative for gait recognition. T is computed independently for each walk W to account for both intra-class and inter-class variability as: $T = \mu + \sigma$, where $\mu = \frac{\sum_{i=1}^{n} y_i}{n}$, and $\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \mu)^2}$, i.e., the mean and standard deviation over the Y axis for a walk W containing n samples. The start s of the useful portion of W is then set to the point corresponding to the first relative maximum greater than T. The ending point e is computed similarly, looking for the last relative maximum greater than T. The s and e points are then projected onto the other two axes for the same sequence length. Given a walk W (for sake of clarity of notation, we omit the subject's ID and the walk index), this produces a reduced walk $W_s^e = \{X_s^e, Y_s^e, Z_s^e\}$, where X_s^e, Y_s^e, and Z_s^e are the reduced intervals on the accelerometer axes, while s and e are the starting and the ending points of the relevant sequence, respectively. A graphical example is illustrated in Fig. 1. This procedure improves the accuracy of the comparison module and provides a computational speed up since it discards a portion of non-meaningful samples De Marsico and Mecca (2017).

3.1.2. Noise removal and signal quality enhancement

This phase entails three steps. The former relies on the Z-score normalization. Given a reduced walk W_s^e, in the normalized reduced walk $W_{norm} = \{X_{norm}, Y_{norm}, Z_{norm}\}$ the i-th element of each normalized time series is computed along the corresponding axis as $x_i' = \frac{x_i - \mu_X}{\sigma_X}$, $y_i' = \frac{y_i - \mu_Y}{\sigma_Y}$, and $z_i' = \frac{z_i - \mu_Z}{\sigma_Z}$, respectively, with $x_i \in X_s^e$, $y_i \in Y_s^e$, $z_i \in Z_s^e$ $\forall i = 1, ..., n$. Normalization also enhances signal quality by highlighting the peaks and better delineating the wave shape. The second step removes noise by exploiting a moving average with window size k. For a normalized walk W_{norm} it computes a walk with averaged values $W_{avg} = \{X_{avg}, Y_{avg}, Z_{avg}\}$, where the following equations give the new i-th element of each time series:
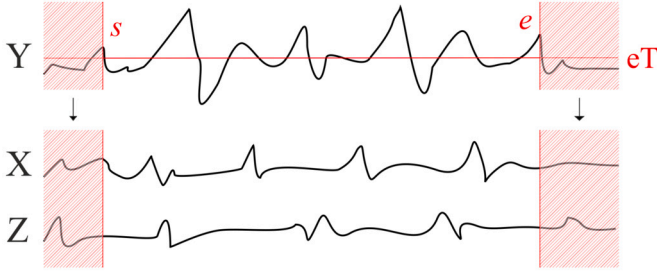
Fig. 1. A graphical example of the relevant signal portion retrieval strategy. The first and the last relative maxima greater than T are searched on the Y axis, identifying the start (s) and the end (e) points, respectively. These are then projected onto the X and Z axes to maintain the same length.

$$\bar{x}_i = \begin{cases} \frac{\sum_{i-\lfloor k/2 \rfloor}^{i+\lfloor k/2 \rfloor} x_i'}{k}, & \text{if } \lfloor k/2 \rfloor \leq i \leq n - \lfloor k/2 \rfloor; \\ x_i', & \text{otherwise}, \end{cases} \quad (1)$$

$$\bar{y}_i = \begin{cases} \frac{\sum_{i-\lfloor k/2 \rfloor}^{i+\lfloor k/2 \rfloor} y_i'}{k}, & \text{if } \lfloor k/2 \rfloor \leq i \leq n - \lfloor k/2 \rfloor; \\ y_i', & \text{otherwise}, \end{cases} \quad (2)$$

$$\bar{z}_i = \begin{cases} \frac{\sum_{i-\lfloor k/2 \rfloor}^{i+\lfloor k/2 \rfloor} z_i'}{k}, & \text{if } \lfloor k/2 \rfloor \leq i \leq n - \lfloor k/2 \rfloor; \\ z_i', & \text{otherwise}, \end{cases} \quad (3)$$

where $x_i' \in X_{norm}$, $y_i' \in Y_{norm}$, and $z_i' \in Z_{norm}$, $\forall i = 1, \ldots, n$, are the normalized elements in W_{norm}, and $\bar{x}_i$, $\bar{y}_i$, and $\bar{z}_i$, $\forall i = 1, \ldots, n$, are the new values computed via moving average. The third step is inspired by De Marsico and Mecca (2018), where a Gaussian convolution allows further increasing the signal quality. The empirical choice to apply a Gaussian kernel with $\sigma = 4$ provides the best trade-off between computational time and performance. Given an averaged walk W_{avg} this step produces the corresponding enhanced walk $W_{enh} = \{X_{enh}, Y_{enh}, Z_{enh}\}$ where the i-th element of each time series is obtained as:

$$\hat{x}_i = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\bar{x}_i^2}{2\sigma^2}}, \quad (4)$$

$$\hat{y}_i = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\bar{y}_i^2}{2\sigma^2}}, \quad (5)$$

$$\hat{z}_i = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\bar{z}_i^2}{2\sigma^2}}, \quad (6)$$

where $\bar{x}_i \in X_{avg}$, $\bar{y}_i \in Y_{avg}$, and $\bar{z}_i \in Z_{avg}$, $\forall i = 1, \ldots, n$, are the elements in W_{avg}, while $\hat{x}_i$, $\hat{y}_i$, and $\hat{z}_i$, $\forall i = 1, \ldots, n$, are the elements of the final W_{enh} time series after the pre-processing module has completed its task.

3.2. Comparison module

This module is based on a modified version of the well-known DTW algorithm Senin (2008). It compares two signals s and t (with possible different lengths) and provides a distance score d. Differently from other signal comparison techniques, such as Manhattan and Euclidean distances, the computed d is more related to the overall shape of the signals than to the single values. DTW has a recursive formulation, but it is generally impractical. For this reason, it is usually implemented as a dynamic programming algorithm. The basic formulation of this algorithm has a computational complexity of $O(n^2)$ but with a high coefficient, which is considered one of its main drawbacks. However, thanks to the matrix representation, it is possible to introduce some useful tricks to speed up the algorithm and increase its accuracy significantly. The plain text and comments in Algorithm 1 show the basic formulation of DTW. The final value in DTW[n][m] is the distance score between two generic

time series, representing the *warping* effort to align them. From the complete DTW table, it is possible to extrapolate the DTW path P_{DTW} (or *warping path*), which contains all matching pairs taken into account to compute the DTW distance. Formally, the warping path is defined as:

$$P_{DTW} = [(1,1), \ldots, (i,j), \ldots, (n,m)] \quad (7)$$

where i and j are matched indexes of s and t, respectively, and (i,j) is included in P_{DTW} if and only if it is a matching required to obtain the minimum cost in the creation of the DTW table. It is also worth noticing that P_{DTW} always starts with $(1,1)$ and ends with (n,m), i.e., matches the pairs of first and last elements of the two time series. Furthermore, a meaningful path must respect the monotonicity property, i.e., if $j > i$ are indices from the first sequence, then there must not be two indices $l > k$ in the other sequence, such that index i is matched with index l and index j is matched with index k, and vice versa. In other words, it is impossible to go back in time by cross-matching past samples in the other sequence.

3.2.1. Reduced warping window

The original DTW neglects an additional constraint holding for DTW in most applications, namely the *warping* constraint that limits the time difference between two matched points in the s and t series. Projecting this on the DTW matrix, the algorithm is forced to compute the values in a region close to the matrix main diagonal, i.e., to match samples with similar indexes in the two series. This can be achieved by limiting the DTW search space to a window w_r centered on the diagonal. According to the experiments in Ratanamahatana and Keogh (2005), this constraint drastically limits the number of operations and can also improve the final comparison results, e.g., for time series with few possible values such as binary time series approximated from electric signals. In the application to gait recognition, as reported in Section 4, it is possible to consider tiny warping windows without loss of accuracy, encouraging the use of a reduced search space.

3.2.2. Relaxed endpoints constraint

Another limitation of the basic DTW algorithm is related to the fixed choice of the matching for the starting and ending points. As shown in Equation (7), a DTW path starts by matching the first points of the two sequences and ends by matching the last ones. As suggested in Silva et al. (2016), it is possible to relax this constraint to let the DTW algorithm choose a starting (ending) point in a broader range to produce an improved path. To this end, as constrained by the warping window, our proposal takes as possible starting (ending) point one of those along the corners of the DTW table. Two values w_n and w_m proportional to the n and m lengths are computed according to the desired reduced warping window size w_r, acting as the *relaxation factor* parameter of Silva et al. (2016):

$$w_n = \lfloor n * (1 - w_r/100) \rfloor$$
$$w_m = \lfloor m * (1 - w_r/100) \rfloor \quad (8)$$

The first and last points in the path are then chosen as follows:

$$P_{DTW}[FIRST] = min(\{DTW(0,i) \; \forall i = 1, \ldots, w_n,$$
$$\cup DTW(j,0) \; \forall j = 1, \ldots, w_m\}), \quad (9)$$

$$P_{DTW}[LAST] = min(\{DTW(n, m-j) \; \forall j = 1, \ldots, w_m,$$
$$\cup DTW(n-i, m) \; \forall i = 1, \ldots, w_n\}), \quad (10)$$

Fig. 2 graphically illustrates the reduced warping window and the relaxed endpoints constraint strategies.

3.2.3. Fast window DTW (FWDTW) comparison strategy

By modifying the DTW algorithm using the two presented strategies, we define the Fast Window DTW (FWDTW) procedure that can generate an improved DTW path compared to the classical DTW. The pseudocode
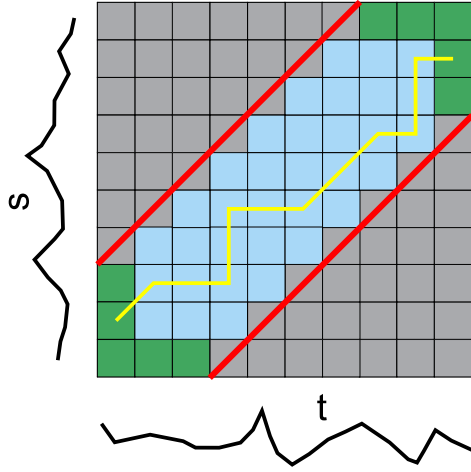
Fig. 2. A graphical example of the reduced warping window (the oblique red lines) and the relaxed endpoints constraints (the green squares). The best warping path starts from one of the bottom-left corner green squares and ends in one of the upper-right green ones. The yellow line represents a possible warping path. The light blue squares indicate candidate elements of the best warping path, while the gray ones are not considered since they require matching points from s and t outside the reduced warping window. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

for the proposed FWDTW approach is reported in Algorithm 1. By leveraging this algorithm to separately compute the distances between two walks along the X, Y, and Z axes, it is possible to define a comparison strategy for gait recognition. Given two generic walk sequences $W_{enh}^1 = \{X_{enh}^1, Y_{enh}^1, Z_{enh}^1\}$ and $W_{enh}^2 = \{X_{enh}^2, Y_{enh}^2, Z_{enh}^2\}$, pre-processed as described in Section 3.1, a distance score d is obtained as follows:

$$
\begin{aligned}
d = \ & w_x * FWDTW(X_{enh}^1, X_{enh}^2) + \\
& w_y * FWDTW(Y_{enh}^1, Y_{enh}^2) + \\
& w_z * FWDTW(Z_{enh}^1, Z_{enh}^2),
\end{aligned}
\tag{11}
$$

where w_x, w_y, and w_z, are weights for the X, Y, and Z axis, respectively. These weighting factors are used to handle the different entropy of the three axes to increase the recognition capabilities of the system.

3.2.4. Majority voting with FWDTW

Works focused on the recognition task generally create a user template by extracting some aggregate features into a feature vector, which is then used to recognize a person by retrieving the most similar user identity from the available gallery. The proposed recognition strategy puts more attention on the constancy of the walking pattern of each user. To this aim, it considers every walk $W_{enh,i}^{id} = \{X_{enh,i}^{id}, Y_{enh,i}^{id}, Z_{enh,i}^{id}\}$ in the gallery G as a template where id indicates a specific identity, while i corresponds to the i-th walk stored in G for that identity. In closed-set identification (all probes belong to an enrolled identity), the identification procedure usually arranges the distances in increasing order. The returned identity is the one in the first position. The proposed system instead returns the identity with the highest number of occurrences in the first best b entries. In case of a tie, the system prioritizes the best comparison score, similar to classical recognition approaches. This scheme can be considered as a special case of Majority Voting, which is normally used for multi-classifier fusion. In this case, it is applied to a single classifier's ranking when the gallery includes multiple templates per identity (with a single template per identity it reduces to the usual scheme). We define the algorithm obtained by applying this scheme to the results of FWDTW as MV-FWDTW. It achieves better accuracy by preventing the wrong recognition of outlier walks.

Algorithm 1 Pseudocode of the FWDTW algorithm. In black are the lines and comments related to the basic DTW version. In red (light gray) underlined text and comments from lines 3 to 12 and line 15, the updates used to handle the reduced warping window (Section 3.2.1). Specifically, in line 15, the j starting/ending values are conditioned by i to follow the diagonal window. In blue (dark gray) underlined text from lines 20 to 25, the additional steps to manage the relaxed endpoints constraint (Section 3.2.2).

```
 1: function FWDTW(s[1, …, n], t[1, …, m], w_r)
    Input: s, t = accelerometer time series to be compared
            w_r = warping window size
    Output: DTW[n][m] = temporal distance matrix

 2:     DTW[0, …, n][0, …, m]
 3:     w_n = ⌊n * (1 − w_r/100)⌋
 4:     w_m = ⌊m * (1 − w_r/100)⌋
 5:     ▷ start of matrix initialization for reduced warping
 6:     for i ← 0, …, n do
 7:         for j ← 0, …, m do
 8:             if (i == 0 ∧ j < w_m) ∨ (j == 0 ∧ i < w_n) then
 9:                 DTW[i][j] ← 0
10:             else
11:                 DTW[i][j] ← ∞
12:     ▷ end of matrix initialization for reduced warping
13:     ▷ in basic DTW all the matrix values are initialized
            to ∞ except for DTW[0][0] = 0
14:     for i ← 1, …, n do
15:         for j ← max(1, i − w_n), …, min(n, i + w_m) do
16:             ▷ in basic DTW j runs from 1 to m.
17:             cost ← |s[i] − t[j]|
18:             DTW[i][j] ← cost + min(DTW[i − 1][j],
                        DTW[i][j − 1], DTW[i − 1][j − 1])
19:     ▷ here, the basic DTW returns DTW[n][m]
20:     dist ← ∞
21:     for i ← n − w_n, …, n do
22:         dist ← min(dist, DTW[i][m]),
23:     for j ← m − w_m, …, m do
24:         dist ← min(dist, DTW[n][j]),
25:     return dist
```

---

4. Experiments

Extensive experiments were conducted to evaluate the proposed FWDTW algorithm and the final MV-FWDTW.

4.1. Dataset

It is worth pointing out that the proposed method and presented experiments are to be considered from the perspective of using gait recognition for medium-long term biometric applications. On the one hand, being a behavioral trait, gait may lack sufficient permanence for those applications. On the other hand, the fact that it is possible for humans to recognize a well-known person from the gait pattern encourages one to search for some effective invariant feature. With this in mind, experiments need to rely on data acquired at different times. The dataset chosen for the experiments is the ZJU-gaitacc Zhang et al. (2014), one of the largest benchmarks for gait recognition systems based on wearable sensors. This dataset, though a little bit outdated, is, unfortunately, the only available one that provides data usable in a realistic context, i.e., presenting data collected into at least two time-separated sessions and a sufficient number of multiple walks per session per subject, to also account for slight changes in the short term. It contains 3D signals from five body locations, i.e., right wrist, left upper arm, right pelvis, left thigh, and right ankle, of 175 subjects walking along a 20-meter hallway. As already mentioned, to increase data variability, signals from 153 of these identities were captured again in a different session after at least one week. Each person was recorded six times in each session, resulting in 1836 sequences. To simulate a realistic everyday scenario, our experiments only use the pelvis location. This is a reasonable position

Table 1

Ablation study for the moving average window k. Baseline DTW performance without this operation is $k = $ '-'. The best results are in bold.

Method	k	Sensors	Rank #1	Rank #5	Rank #10
DTW	-	Pelvis	96.21%	98.34%	99.10%
DTW	5	Pelvis	97.17%	**99.10%**	**100.00%**
DTW	7	Pelvis	**97.21%**	99.05%	**100.00%**
DTW	9	Pelvis	97.19%	99.07%	**100.00%**

for an acquisition device such as a smartphone or a dedicated instrument. In addition, among the available body locations, the hip is the closest to the body's center of gravity, resulting in a stable place to acquire gait data De Marsico and Mecca (2019). We split the dataset into two separately recorded sessions, to obtain D_1 and D_2, each containing each user's walks from the same session. The experiments used these two splits as either probe set P or gallery G in turn. The 22 subjects with recordings from a single session were discarded to set up a cross-session evaluation suitable for closed-set identification. The session-based partition into probe and gallery represents a more realistic scenario where template walks of a person are captured on a given day for enrolling, but recognition exploits walks captured on a different day. As reported in De Marsico and Palermo (2022), using only walks from one session or mixing up the two sessions in the gallery and probe sets would result in heavily increased recognition rates. However, these can only be realistic in a scenario of short-term re-identification. To further assess the quality of the presented strategy, it has also been tested on a subset of 40 subjects of the OU-ISIR dataset Ngo et al. (2014), which consists of walk signals from 744 subjects. It is worth pointing out that the signals are very short (taken along a about 3-meter-long walk) and only two templates for each subject are provided, manually extracted from a single walk that also contains two slope-dependent portions. This limits the applicability only to short-term re-identification. Since the DTW algorithm usually requires longer sequences to provide accurate results, in the experimental setup we tested on both the original shorter signals and on an augmented and longer version of them. In the latter case, all the walking signals are doubled.

4.2. Implementation details

The proposed MV-FWDTW recognition was implemented in Java, and all experiments were performed on an Intel Core i7-6700HQ @2.6GHz with 16 GB of RAM and an nVIDIA GTX1060M GPU with 6 GB of RAM.

In closed-set identification, the system returns a list of the gallery templates identities sorted by increasing distance from the probe, i.e., the input walk to recognize. The system evaluation relies on standard metrics, i.e., the Cumulative Match Score (CMS) at Rank #1, also referred to as Recognition Rate (RR), and the Cumulative Match Characteristic (CMC) curve. The former is the percentage of the tests in which the correct identity of a probe is returned in the first position of the ordered list of the gallery identities. The CMC curve plots the CMSs at ranks (up to the gallery size), where CMS(k) is the percentage of the total number of tests in which the correct probe identity is returned within the first k positions of the list (the curve is monotonically increasing). The curve provides an insight into the overall ability of the system to quickly return the correct identity, though not in the first position. Other popular measures used in the literature are CMS(5) and CMS(10), and also the rank k where CMS(k) = 100%. The lower such k, the better.

4.3. Performance evaluation

Several ablation studies assessed the FWDTW algorithm to fully evaluate it. They identified the best parameter combination of FWDTW to join the derived implementation with MV-FWDTW. Both the best

Table 2

Ablation study on warping window size w_r. DTW with $w_r = 100$ is a baseline with no window restriction. Avg. Time indicates the average computation time required to compare two walks. The best accuracy results are in bold.

Method	w_r	Rank #1	Rank #5	Rank #10	Avg. Time
DTW	100	97.21%	99.05%	**100.00%**	32 ms
FWDTW	70	97.55%	99.12%	**100.00%**	23 ms
FWDTW	50	97.67%	99.12%	**100.00%**	17 ms
FWDTW	30	97.89%	99.27%	**100.00%**	11 ms
FWDTW	20	**98.04%**	**99.32%**	**100.00%**	7 ms
FWDTW	10	98.03%	99.30%	**100.00%**	4 ms

FWTDW implementation and the derived MV-FWDTW were finally compared with other state-of-the-art works tackling gait recognition from wearable sensors and using the ZJU-gaitacc benchmark. To avoid partition-based bias, the reported cross-session results are computed by averaging the performances obtained using D_1 as probe set and D_2 as gallery and then swapping their role.

The first batch of experiments evaluated the influence of the size k of the moving average window described in Section 3.1.2. Table 1 reports the results obtained using the classic DTW algorithm. These show that the baseline model, i.e., without the moving average denoising (first line in Table 1), already achieves remarkable performances. This confirms that signals associated with the pelvis sensor carry meaningful information, allowing accurate gait recognition via the DTW algorithm. Nevertheless, by applying the de-noising procedure, the improved quality of signals spurs a direct performance increase across all recognition ranks independently from the chosen window size, with all dimensions gaining ~1% at Rank #1, Rank #5, and Rank #10 compared to the baseline. This outcome can be associated with the constancy of individual gait, which can be better captured using smoothed signals. This is also confirmed by the 100% RR achieved at Rank #10, even though more than 150 identities are available.

The second ablation study examined the effects of different sizes w_r of the warping window. Table 2 reports the results obtained using DTW without any warping window reduction (i.e., $w_r = 100$) as a baseline and the proposed FWDTW algorithm with different w_r dimensions. As can be observed, restricting the warping window size to 70%, 50%, 30%, 20%, and 10% of the original time series, drastically decreases the average computational time due to a smaller search space, requiring 87.5% less time compared to the baseline when using a 10% window size. More interestingly, such a decrease in time does not affect identification accuracy but rather slightly improves by $\sim 0.80\%$ with the smaller windows, i.e., 20% and 10%. This suggests that by moving away from the DTW table diagonal, i.e., allowing larger time differences in the matched series points, optimal local choices might result in an overall worse warped path, thus demonstrating the validity of the warping window restriction as it improves both execution time and RR.

The last batch of experiments aims at selecting the best axes weights in the weighted distance between two time series computed in Equation (11). In all the different combinations, the Y axis gets a weight of at least 0.5 as in accordance with De Marsico and Mecca (2017), it contains the highest amount of information when performing user recognition via gait. The empirical results show that the best combination for the proposed system is $w_x = 0.2$, $w_y = 0.7$, and $w_z = 0.1$, indicating that all axes capture some walk's characteristics, but the Y one still contains stronger information, in accordance with prior studies.

We compared the proposed FWDTW algorithm and the final MV-FWDTW strategy with other state-of-the-art proposals. For the latter, we use the first $b = 10$ scores to select identities for the majority voting strategy since the ZJU-gaitacc dataset has six walks per user and it is a reasonable number of templates for the application context. Table 3 reports the results obtained using a moving average window of dimension $k = 7$, and a warping window size $w_r = 20$ (the parameters achieving the best accuracy according to Tables 1 and 2). The available proposals in literature use different protocols that employ either a same-session

Table 3

State-of-the-art comparison on the ZJU-gaitacc dataset for works employing either a same-session or a (more complex) cross-session testing procedure.

Method	Sensors	Rank #1	Test Protocol
Giorgi et al. (2017)*	All	95.00%	same-session
Mecca (2018)	Pelvis	96.49%	same-session
Sun et al. (2018)	All	96.90%	same-session
Nemes and Antal (2021)*	All	98.67%	same-session
Nemes and Antal (2021)*	All	59.88%	cross-session
Zhang et al. (2014)	Pelvis	73.40%	cross-session
Zhang et al. (2014)	All	95.80%	cross-session
FWDTW	Pelvis	98.04%	cross-session
MV-FWDTW	Pelvis	**98.82%**	cross-session

*marked methods employ deep learning models.

Table 4

Results achieved on a 40-subject subset of the OU-ISIR dataset with the FWDTW.

Data	Rank #1	Rank #5	Rank #10
Original	71.45%	80.22%	86.22%
Doubled	98.18%	100.00%	100.00%

or cross-session testing scenario, i.e., mix or do not mix walks recorded in different sessions in the probe set P and gallery G, or even use walks from a single session for both sets. These strategies can all seem sound, but avoiding cross-session testing, though providing better results, is not suitable in real-life scenarios except for short-term re-identification. Experiments in Nemes and Antal (2021) show a considerable performance drop when using the more complex and realistic cross-session protocol, though using a neural architecture for feature extraction. A further difference lies in the number of exploited sensors. As described in Section 4.1, the dataset contains signals captured by sensors in five distinct body locations. While leveraging the entire data naturally results in higher performances, as noted in Zhang et al. (2014), using the single pelvis sensor corresponds to a more feasible real-case scenario. Moreover, placing five acquisition devices on the user is way more intrusive than using a single one. Also, the latter might be easily embedded in the user's smartphone or into a microchip mounted, for example, on a belt. All considered, the proposed MV-FWDTW consistently outperforms the existing literature by using signals associated exclusively with the pelvis sensor in a cross-session setup, even when compared with methods using deep learning models. This behavior is also confirmed when analyzing the CMC curves. Fig. 3 shows that FWDTW and MV-FWDTW achieve higher recognition rates across all rankings. Moreover, the MV-FWDTW reaches a 100% rate at a Rank #4, contrary to the best literature approach that achieves a 99.33% recognition rate at Rank #10. This fully highlights the effectiveness of the proposed signal pre-processing and comparison strategy via relaxed endpoints and reduced search space.

To conclude the performance evaluation, experiments tested the proposed strategy on a subset of 40 subjects of the OU-ISIR dataset. It is worth pointing out that, in this context, since only one walk per subject is available in the gallery, it is not possible to use the MV-FWDTW, so the FWDTW strategy is used instead. As it is possible to notice from Table 4, due to the much shorter walk signals and the presence of only one gallery template (which also reduces intra-class variability), the results achieved by the proposed strategy drastically drop when using the original signals. However, with a data augmentation step consisting of a simple duplication of each walk signal, the results become comparable with those achieved on the ZJU-gaitacc benchmark, though relating to walks too close in time. This confirms the strategy's applicability when long enough signals are provided, as it happens in the realist context mentioned in Section 2.
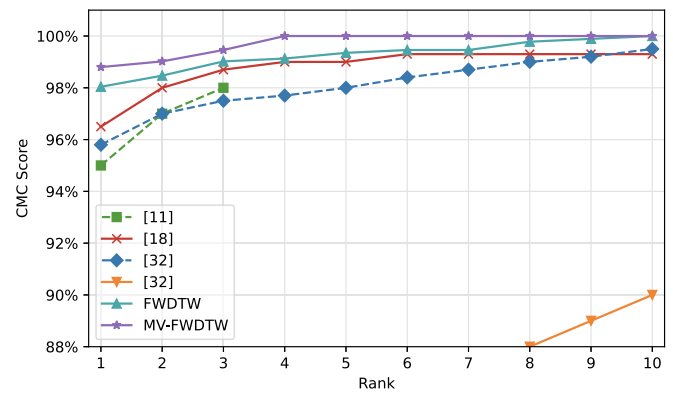

Fig. 3. Comparison of state-of-the-art CMCs. Solid lines are for methods using only the pelvis sensor; dashed ones are for those using all five sensors.

5. Conclusion

This paper presents a novel gait recognition strategy based on signals captured by a single smartphone accelerometer, suitable for real-life identification scenarios due to the gait robustness to forgery and spoofing. Instead of being based on training on large amounts of data, like in a machine or deep learning scenario, our strategy leverages handcrafted features in an algorithm that does not require any training, so that it can be more robust on unknown data. The proposed approach retrieves the informative part of a signal and applies several enhancing operations, i.e., Z-score normalization, moving average, and Gaussian convolution. The enhanced signals are then processed using a modified version of the Dynamic Time Warping algorithm, the FWDTW, which uses speed-up and refinement strategies to increase the overall robustness and reduce the computational demand, i.e., a reduced warping window and relaxed endpoints. The majority voting strategy is then adapted to obtain the final MV-FWDTW recognition. To evaluate the presented proposal, experiments were performed on ZJU-gaitacc, a benchmark dataset containing many users (153) and walks (12 per user). Results demonstrate the effectiveness of the proposed approach that outperforms existing works using either a cross-session or (the easier) same-session testing protocol, one or all wearable sensors available in the dataset, and neural network models. The quality of the strategy has been further tested on a subset of the OU-ISIR benchmark, which contains shorter walk signals collected in a single session.

In future work, we plan to collect an extensive dataset containing, for each walk acquisition, a frontal video and a synchronized set of wearable sensor data from several body locations. In fact, to the best of our knowledge, no dataset contains synchronized video and wearable gait acquisition, and, in general, frontal video gait recognition is a realistic yet not much-explored topic. This will allow testing the viability of the proposed recognition setup in a multi-biometrics setting. In this frontal gait video context, we will exploit previous findings on skeleton data retrieved from video sequences Avola et al. (2019, 2020a,b) and explore further fusion approaches to advance gait recognition systems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors are unable or have chosen not to specify which data has been used.

References

Almulihi, A.H., Alassery, F., Khan, A.I., Shukla, S., Gupta, B.K., Kumar, R., 2022. Analyzing the implications of healthcare data breaches through computational technique. Intell. Autom. Soft Comput. 32.

Ameur, B., Belahcene, M., Masmoudi, S., Ben Hamida, A., 2019. Hybrid descriptors and weighted PCA-EFMNet for face verification in the wild. Int. J. Multimed. Inf. Retr. 8, 143–154.

Attaallah, A., Alsuhabi, H., Shukla, S., Kumar, R., Gupta, B.K., Khan, R.A., 2022. Analyzing the big data security through a unified decision-making approach. Intell. Autom. Soft Comput. 32.

Avola, D., Bacciu, A., Cinque, L., Fagioli, A., Marini, M.R., Taiello, R., 2022a. Study on transfer learning capabilities for pneumonia classification in chest-x-rays images. Comput. Methods Programs Biomed. 221, 106833.

Avola, D., Cascio, M., Cinque, L., Fagioli, A., Petrioli, C., 2022b. Person re-identification through Wi-Fi extracted radio biometric signatures. IEEE Trans. Inf. Forensics Secur. 17, 1145–1158.

Avola, D., Cascio, M., Cinque, L., Foresti, G.L., Massaroni, C., Rodolà, E., 2019. 2-D skeleton-based action recognition via two-branch stacked LSTM-RNNs. IEEE Trans. Multimed. 22, 2481–2496.

Avola, D., Cinque, L., Fagioli, A., Foresti, G., Mecca, A., 2021. Ultrasound medical imaging techniques: a survey. ACM Comput. Surv. 54, 1–38.

Avola, D., Cinque, L., Fagioli, A., Foresti, G.L., Massaroni, C., 2020a. Deep temporal analysis for non-acted body affect recognition. IEEE Trans. Affect. Comput.

Avola, D., Cinque, L., Fagioli, A., Foresti, G.L., Pannone, D., Piciarelli, C., 2020b. Bodyprint—a meta-feature based LSTM hashing model for person re-identification. Sensors 20, 5365.

Bowyer, K.W., Burge, M.J., 2016. Handbook of Iris Recognition. Springer.

De Marsico, M., Mecca, A., 2015. Biometric walk recognizer. In: International Conference on Image Analysis and Processing (ICIAP), pp. 19–26.

De Marsico, M., Mecca, A., 2017. Biometric walk recognizer. Multimed. Tools Appl. 76, 4713–4745.

De Marsico, M., Mecca, A., 2018. Benefits of Gaussian convolution in gait recognition. In: 2018 International Conference of the Biometrics Special Interest Group (BIOSIG), pp. 1–15.

De Marsico, M., Mecca, A., 2019. A survey on gait recognition via wearable sensors. ACM Comput. Surv. 52, 1–39.

De Marsico, M., Mecca, A., Barra, S., 2019. Walking in a smart city: investigating the gait stabilization effect for biometric recognition via wearable sensors. Comput. Electr. Eng. 80, 106501.

De Marsico, M., Palermo, A., 2022. Towards the suitability of gait wearable signal processing for long term recognition. In: 2022 IEEE International Joint Conference on Biometrics (IJCB). IEEE, pp. 1–9.

Fan, Y.-Y., Liu, S., Li, B., Guo, Z., Samal, A., Wan, J., Li, S.Z., 2017. Label distribution-based facial attractiveness computation by deep residual learning. IEEE Trans. Multimed. 20, 2196–2208.

Fernandez-Lopez, P., Sanchez-Casanova, J., Tirado-Martín, P., Liu-Jimenez, J., 2017. Optimizing resources on smartphone gait recognition. In: International Joint Conference on Biometrics (IJCB), pp. 31–36.

Giorgi, G., Martinelli, F., Saracino, A., Sheikhalishahi, M., 2017. Try walking in my shoes, if you can: accurate gait recognition through deep learning. In: International Conference on Computer Safety, Reliability, and Security, pp. 384–395.

Gu, D., Li, Y., Jiang, F., Wen, Z., Liu, S., Shi, W., Lu, G., Zhou, C., 2020. VINet: a visually interpretable image diagnosis network. IEEE Trans. Multimed. 22, 1720–1729.

Guo, G., Zhang, N., 2019. A survey on deep learning based face recognition. Comput. Vis. Image Underst. 189, 102805.

Li, S., Zhang, B., Fei, L., Zhao, S., Zhou, Y., 2021. Learning sparse and discriminative multimodal feature codes for finger recognition. IEEE Trans. Multimed., Early Access.

Maity, S., Abdel-Mottaleb, M., Asfour, S.S., 2021. Multimodal low resolution face and frontal gait recognition from surveillance video. Electronics 10, 1013.

Marasco, E., Ross, A., 2014. A survey on antispoofing schemes for fingerprint recognition systems. ACM Comput. Surv. 47, 1–36.

Mecca, A., 2018. Impact of gait stabilization: a study on how to exploit it for user recognition. In: 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), pp. 553–560.

Muaaz, M., Mayrhofer, R., 2017. Smartphone-based gait recognition: from authentication to imitation. IEEE Trans. Mob. Comput. 16, 3209–3221.

Nemes, S., Antal, M., 2021. Feature learning for accelerometer based gait recognition. In: 15th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI). IEEE, pp. 479–484.

Ngo, T.T., Makihara, Y., Nagahara, H., Mukaigawa, Y., Yagi, Y., 2014. The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication. Pattern Recognit. 47, 228–237.

Nguyen, K., Fookes, C., Jillela, R., Sridharan, S., Ross, A., 2017. Long range iris recognition: a survey. Pattern Recognit. 72, 123–143.

O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G.V., Krpalkova, L., Riordan, D., Walsh, J., 2020. Deep learning vs. traditional computer vision. In: Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), vol. 1. Springer, pp. 128–144.

Ratanamahatana, C.A., Keogh, E., 2005. Three myths about dynamic time warping data mining. In: International Conference on Data Mining (ICDM), pp. 506–510.

Sahu, K., Alzahrani, F.A., Srivastava, R., Kumar, R., 2020. Hesitant fuzzy sets based symmetrical model of decision-making for estimating the durability of web application. Symmetry 12, 1770.

Sahu, K., Alzahrani, F.A., Srivastava, R., Kumar, R., 2021. Evaluating the impact of prediction techniques: software reliability perspective. Comput. Mater. Continua 67.

Sahu, K., Srivastava, R., 2018. Soft computing approach for prediction of software reliability. Neural Netw. 17, 19.

Sahu, K., Srivastava, R., 2020. Needs and importance of reliability prediction: an industrial perspective. Inf. Sci. Lett. 9, 33–37.

Sahu, K., Srivastava, R., 2021. Predicting software bugs of newly and large datasets through a unified neuro-fuzzy approach: reliability perspective. Adv. Math. Sci. J. 10, 543–555.

Sahu, K., Srivastava, R., Kumar, S., Saxena, M., Gupta, B.K., Verma, R.P., 2023. Integrated hesitant fuzzy-based decision-making framework for evaluating sustainable and renewable energy. Int. J. Data Sci. Anal. 16, 371–390.

Senin, P., 2008. Dynamic Time Warping Algorithm Review. Information and Computer Science, vol. 855. Department University of Hawaii at Manoa, Honolulu, USA, p. 40.

Shao, Z., Cheng, G., Ma, J., Wang, Z., Wang, J., Li, D., 2021. Real-time and accurate UAV pedestrian detection for social distancing monitoring in Covid-19 pandemic. IEEE Trans. Multimed., Early Access, 2069–2083.

Silva, D.F., Batista, G., Keogh, E., et al., 2016. On the effect of endpoints on dynamic time warping. SIGKDD MiLeTS 16, 10.

Singh, J.P., Jain, S., Arora, S., Singh, U.P., 2021. A survey of behavioral biometric gait recognition: current success and future perspectives. Arch. Comput. Methods Eng. 28, 107–148.

Sun, F., Mao, C., Fan, X., Li, Y., 2018. Accelerometer-based speed-adaptive gait authentication method for wearable IoT devices. IEEE Int. Things J. 6, 820–830.

Takeda, T., Taniguchi, K., Asari, K., Kuramoto, K., Kobashi, S., Hata, Y., 2009. Biometric personal authentication by one step foot pressure distribution change by load distribution sensor. In: IEEE International Conference on Fuzzy Systems, pp. 906–910.

Van Gastel, M., Stuijk, S., de Haan, G., 2015. Motion robust remote-PPG in infrared. IEEE Trans. Biomed. Eng. 62, 1425–1433.

Wan, C., Wang, L., Phoha, V.V., 2018. A survey on gait recognition. ACM Comput. Surv. 51, 1–35.

Wang, M., Deng, W., 2021. Deep face recognition: a survey. Neurocomputing 429, 215–244.

Xu, W., Shen, Y., Zhang, Y., Bergmann, N., Hu, W., 2017. Gait-watch: a context-aware authentication system for smart watch based on gait recognition. In: Proceedings of the Second International Conference on Internet-of-Things Design and Implementation, pp. 59–70.

Zhang, Y., Pan, G., Jia, K., Lu, M., Wang, Y., Wu, Z., 2014. Accelerometer-based gait recognition by sparse representation of signature points with clusters. IEEE Trans. Cybern. 45, 1864–1875.