



# Gait-Based Biometrics System

Aleksander Sawicki<sup>1</sup>(✉)  and Khalid Saeed<sup>1,2</sup> 

<sup>1</sup> Faculty of Computer Science, Bialystok University of Technology, Bialystok, Poland  
a.sawicki@pb.edu.pl

<sup>2</sup> Department of Computer Science and Electronics, Universidad de La Costa, Barranquilla, Colombia

**Abstract.** Behavioral biometrics is the field of study related to person identification based on the way an activity is performed. Despite the difficulties of implementation and achieving high recognition metrics, this field has advantages that attract the interest of the scientific community. In the case of gait analysis, active interaction between the user and the acquisition device is not required. This enables signals to be safely assessed remotely, that is important in the times of pandemic. Furthermore, it is not easy to deliberately mimic a person's gait. The work concerns on the development of system that enables identifying individuals based on gait with the use of wearable sensors such as accelerometers or gyroscopes. The work describes the data preprocessing pipeline and the innovative data augmentation mechanism performed with the use of generative models. The validation of the system is carried out using three different datasets collected under laboratory, semi-laboratory and field conditions. This article focuses on presenting a comprehensive solution, with a special authors' aspect of data augmentation.

**Keywords:** Biometrics · CNN · IMU

## 1 Introduction

In this work, we aim to present a complex gait-based personal identification system that use wearable sensors. The work presents the construction of a biometric system using the data from a triaxial accelerometer and a triaxial gyroscope. Sensors of this type can be embedded in smartphone or smartwatch devices. The demonstration of the system execution can be found at: <https://vimeo.com/818295324>. Currently, the gait analysis systems are not reserved for the worlds created by Hollywood directors, but are an actual core of research and implementation. In order to provide credence to the interest, two references can be mentioned [1, 2]. One of them is the Polish research and development project “BIOMETRICS - a biometric system for identifying persons using their gait” [1]. The project is involved in the construction of a corridor in which data acquisition and analysis of human gait could be carried out. It is based on a vision-based manner and maintaining a constant type of substrate and lighting. One of the potential users of the product was declared to be the internal use of the Polish Border Guard.

On the other hand, it is worth mentioning the opening of the European competition “Beyond the state-of-the-art biometrics on the move for border checks” [2]. This European grant mainly concerns non-contact gait analysis of persons for the purpose of border controls. However, without granting the official award or even without the opening of the competition, it is difficult to detail the actual use-case. These two references [1, 2] indicate that behavioral biometric systems are of interest not only to academia, but to important institutions at national/international level as well as to private capital.

Classical physiological biometrics systems that for example are based on a fingerprints or an iris photos, require active user interaction - the user must walk up to the data acquisition device. It is not possible to identify a person remotely. In addition, such systems could be susceptible to various types of attacks in which artificial samples are provided instead of the original data. In previous years, researchers have demonstrated that the so-called MasterPrints attack [3] has successfully frauded biometrics systems.

Under such circumstances, it may seem that the use of the additional keys, in the authentication process - in the form of RFID cards, is a more reasonable solution. However, such approach is not ideal. When an unwanted person obtains the key through theft/robbery, it will automatically gain unauthorized access. It should be emphasized that the theft of a carrier such as an RFID card is a relatively simple operation, but the intentional imitation of its owner’s gait is much more difficult [4].

The target users of this system may be institutions that aim to increase the security access. The entities that are concerned, for example, about the theft of the access keys and would like to upgrade security systems might use an additional behavioral biometrics module that is resistant to device theft. The final point is that nowadays, solutions based on behavioral biometrics in practice do not exist on their own. They complement the systems based on biometrics of physiological features or key-based authorization systems.

## 1.1 Dataset Description

The evaluation of the performance of the proposed biometrics system was conducted using three data corpora, in each of which participants took part in two acquisition sessions during two separate days. This approach is quite relative to a real-world scenario in which the biometric system is trained once and then evaluated (even under changed conditions, e.g., footwear, ground surface). The validation of the system was carried out with three datasets: (I) Białystok University of Technology; laboratory conditions with constants shoes types and ceramic walking surface; 100 Participants, acquisition performed using Perception Neuron Inertial Motion Capture system [5]. Signals were collected using proprietary software “Axis Neuron”; (II) Signet (University of Padova [6, 7]); semi-laboratory conditions with varying clothes and constant walking surface; 28 Participants, acquisition performed using Android mobile phone. Data acquisition was carried out using the developed custom application; (III) Boston dataset [8]; field conditions with varying clothes and walking surfaces; 29 Participants, acquisition performed using iPhone. Official “SensorLog” app was used for data acquisition.

Data acquisition for each case proceeded differently. For dataset (I), participants completed 20 repetitions along a straight distance of about 3 m. For the base (II), subjects completed a standard walk for about 5 min, and for base (III), participants covered a distance of approximately 640 m. The implemented gait cycle segmentation algorithm

recognized a total of 6,697 gait cycles for the base (*I*), 19,262 for base (*II*) and 25,569 for base (*III*). Small intervals between training and testing sessions characterize all selected databases. The case where the period is much longer is much more demanding and is omitted in this study. According to the literature, a period of 9 months is the time when the gait does not change significantly [9], after which it is recommended to collect new samples.

### 1.2 Preprocessing

Preprocessing was carried out in three steps. The first was detecting the moment when a person’s right leg touches the ground. Despite the fact that the IMU sensor was placed in the trouser pocket/right thigh area, it was possible to indicate this moment in time. In the segmentation process, only the accelerometer measurement data were used. For this purpose, the algorithm described in our previous work [10] was applied. The segmentation process made it possible to extract the so-called gait cycles, i.e. (the gait cycle is defined as the time inversion between two successive occurrences of the foot contacting the ground [11]). The second preprocessing step involved assembling the data into a coherent dataset by removing outliers and performing frequency filtering. At the same time, it should be noted that the data processed in this step concerned the IMU measurement values directly read by the sensor. Therefore, they were in the so-called local reference system depending on the mounting method. In Fig. 1 (middle step), it can be seen that there are significant differences between the two days for the X and Y axes. The data in their original form depend on the way the sensor is placed in the trouser pocket.

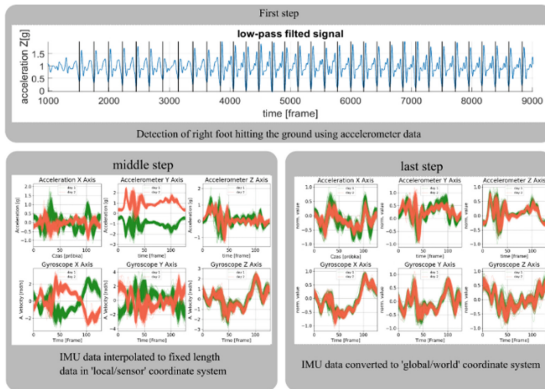


Fig. 1. IMU preprocessing flow diagram

The last step involved converting the data from a local reference system to a global world system. This conversion requires additional information about the phone’s orientation at the measurement time (The application uses time series of quaternions). The result of this step is the independence of the measurement data from the way the sensor is mounted.

### 1.3 Data Augmentation

In our research, we used the technique of data augmentation, which involves the use of distributional models. In the development of the model, we made use of studies such as [12, 13]. In each of these publications, the output of the neural networks was not expected to be a single value, but instead modeled the mean value with variance. For this purpose, a architecture consisting of an LSTM member and a Dense layer was used (Fig. 2). The idea behind this form of augmentation is to use historical samples to train a person's typical gait cycles, which can have a positive effect on identification results. Moreover, a model of this type has an interesting property. With a small computational cost, the variance of the generated time series can be easily multiplied/amplified. This approach is in opposition to the use of VAE-based models (such as RH-VAE [14] aka pyraug [15]). In which the so-called bottleneck connecting the encoder and decoder is modeled by normal distribution parameters.

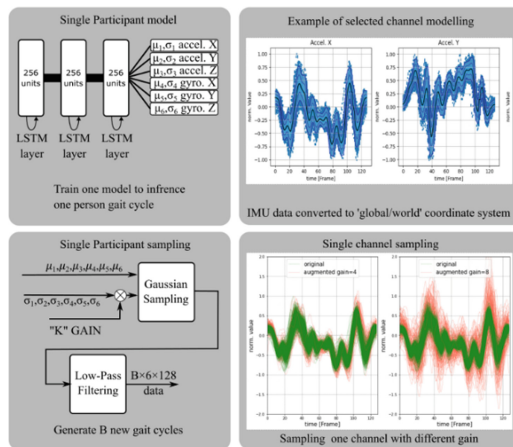


Fig. 2. Data augmentation flow diagram

### 1.4 Classification Results

The classification was carried out as follows. The data collected during the first day constituted the training set, while the gait cycles collected during the second day entirely constituted the test set. A CNN network with an attentional mechanism [16] was used as a classifier. For augmentation, 50% of the synthetic gait samples were added for each participant. Table 1 shows results in the form of an F1-score metrics.

**Table 1.** Classification Results

Augmentation Method	Dataset: BUT	Dataset: Signet	Dataset: Boston
BASELINE (NO AUG.)	0.921 ± 0.017	0.815 ± 0.020	0.712 ± 0.019
PURAU[14]	0.914 ± 0.010	0.697 ± 0.020	0.699 ± 0.014
PROPOSED (NO GAIN)	0.812 ± 0.010	0.812 ± 0.010	0.748 ± 0.015
PROPOSED (GAIN 16)	<b>0.932 ± 0.007</b>	<b>0.831 ± 0.019</b>	<b>0.754 ± 0.009</b>

## 2 Conclusions

This paper described the authors' biometric system based on the measurement values of sensors such as accelerometers and gyroscope. The paper proposed the use of distributional models in the data augmentation process. The system was validated using three data corpora with different degrees of representation of real-world conditions.

**Acknowledgment.** This work was supported by grant 2021/41/N/ST6/02505 from Białystok University of Technology and funded with resources for research by National Science Centre, Poland. For the purpose of Open Access, the author has applied a CC-BY public copyright license to any Author Accepted Manuscript (AAM) version arising from this submission.

## References

1. "Biometrics - a biometric system for identifying persons using their gait, POIR.01.01.01-00-1942/20. Accessed 16 Mar 2023
2. "Beyond the state-of-the-art biometrics on the move for border checks" project page. <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/horizon-cl3-2023-bm-01-03>. Accessed 16 Mar 2023
3. Roy, A., Memon, N., Ross, A.: MasterPrint attack resistance: a maximum cover based approach for automatic fingerprint template selection. In: IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS), pp. 1–9 (2019)
4. Wan, C., Wang, L., Phoha, V.V.: A survey on gait recognition. *ACM Comput. Surv.* **51**, 5, Article 89 (2019). <https://doi.org/10.1145/3230633>
5. Perception Neuron 32 official website. [https://neuronmocap.com/products/perception\\_neuron/](https://neuronmocap.com/products/perception_neuron/). Accessed 16 Jan 2019
6. SIGNET research, Smartphone gait signals dataset, Department of Information Engineering, University of Padova, Italy. <https://signet.dei.unipd.it/research/> Accessed 02 April 2023
7. Gadaleta, M., Rossi, M.: IDNet: Smartphone-based gait recognition with convolutional neural networks. *Pattern Recogn.* **74**, 25–37 (2018)
8. Vajdi, A., et al.: Human gait database for normal walk collected by smart phone accelerometer. arXiv preprint [arXiv:1905.03109](https://arxiv.org/abs/1905.03109) (2019)
9. Matovski, D., et al.: The effect of time on gait recognition performance. *Trans. Inf. Forensics Secur.* **7**(2), 543–552 (2011)
10. Sawicki, A., Saeed, K.: Application of LSTM networks for human gait-based identification. In: Zamojski, W., Mazurkiewicz, J., Sugier, J., Walkowiak, T., Kacprzyk, J. (eds.) *Theory and Engineering of Dependable Computer Systems and Networks. DepCoS-RELCOMEX 2021.*

- Advances in Intelligent Systems and Computing, vol. 1389, pp. 402–412. Springer, Cham (2021). [https://doi.org/10.1007/978-3-030-76773-0\\_39](https://doi.org/10.1007/978-3-030-76773-0_39)
11. Whittle, M.: *Gait Analysis: an Introduction*. Butterworth-Heinemann, Oxford (1991)
  12. Graves A.: Generating sequences with recurrent neural networks. [arXiv:1308.0850](https://arxiv.org/abs/1308.0850) (2013)
  13. Alzantot, M., Chakraborty, S., Srivastava, M.: SenseGen: a deep training architecture for synthetic sensor data generation. In: 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops) (2017)
  14. Chadebec, C., Thibeau-Sutre, E., Burgos, N., et al.: Data augmentation in high dimensional low sample size setting using a geometry-based variational Autoencoder. [arXiv:2105.00026](https://arxiv.org/abs/2105.00026) (2021)
  15. Pyraug augmentation framework official repository. <https://github.com/clementchadebec/pyraug>. Accessed 16 Mar 2023
  16. Huang, H., Zhou, P., Li, Y., Sun, F.: A Lightweight attention-based CNN model for efficient gait recognition with wearable IMU sensors. *Sensors* 2021, **21**, 2866 (2021)