

An Intelligent Gait Data Processing Algorithm Based on Mobile Phone Accelerometers [†]

Nikolay Dorofeev ^{1,*}  and Anastasya Grecheneva ^{1,2} 

¹ Department of Management and Control in Technical Systems, Vladimir State University, Orlovskaya st. 23, 602264 Murom, Russia; grechenevaav@yandex.ru

² Department of Applied Informatics, Moscow Timiryazev Agricultural Academy, Timiryazevskaya st. 49, 127434 Moscow, Russia

* Correspondence: dorofeevnnv@yandex.ru; Tel.: +7-49234-77-236

[†] Presented at the 15th International Conference “Intelligent Systems” (INTELS’22), Moscow, Russia, 14–16 December 2022.

Abstract: This paper describes an algorithm for extracting human gait movements in data obtained from accelerometer sensors of a mobile phone, provided that the mobile phone is used in the usual mode. The algorithm also performs a classification of the selected movements based on a feed-forward neural network. The developed algorithm selects the best areas in the accelerometer data, which reflect individual steps, according to the optimality criterion. For the selected area, the optimality criterion is the maximum value of the correlation coefficient with all other data segments. The selected plots are used as templates. Changing the parameters of patterns over time is necessary to assess changes in the individual rate of the functioning of the musculoskeletal system. Due to the correction of tolerance limits at the segmentation stage, the algorithm adapts to the change in gait speed.

Keywords: gait; accelerometer; mobile phone; wearable device; automation; algorithm



Citation: Dorofeev, N.; Grecheneva, A. An Intelligent Gait Data Processing Algorithm Based on Mobile Phone Accelerometers. *Eng. Proc.* **2023**, *33*, 44. <https://doi.org/10.3390/engproc2023033044>

Academic Editors: Askhat Diveev, Ivan Zelinka, Arutun Avetisyan and Alexander Ilin

Published: 3 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Biometric information has long been an important component in control and security systems. Increasingly more attention is being paid to biometric data in medicine, statistics, and economics. Modern information technology tools provide the ability to quickly collect and process biometric data, including in a covert mode. Human gait parameters are an example of biometric data that are analyzed in a covert manner. From the point of view of biometrics, gait parameters have attracted researchers for a long time; many approaches and algorithms have been developed for assessing and analyzing gait parameters, but the relevance of this area is not decreasing. This is due to the sharp growth in mobile, personal, and miniature information and technical means in control and authentication; the development of a personalized approach to assessing human health; and innovations in the gaming world [1–3]. Current trends in the processing and analysis of human gait data are aimed at identifying changes in motor patterns and the factors that cause these changes [1,4–6]. Technical means that are used to record gait parameters can be divided into the following categories: stationary systems and systems using wearable sensors or sensors sewn into clothing, video recording systems, and systems based on wearable mobile devices [7–12]. The last category includes fitness bracelets, smart watches, smartphones, and—devices that include an acceleration or spatial orientation sensor. This category is the most interesting from the position of quasi-continuous measurements of gait parameters in the background due to prolonged contact with a person during the day. However, the apparent simplicity of measurements is complicated by the technical features of the wearable devices and the peculiarities of their use. Unlike stationary systems or systems based on wearable sensors, the measurement of gait parameters using wearable devices is possible only at one point, which can move relative to the human body during the

measurement process (for example, when moving a smartphone). In addition to the random movement of the smartphone, there are factors that affect measurements (various technical and design features of the device) and human gait in general (for example, physiological ones) [1,13]. Thus, the estimation of human movement (gait) parameters based on measurements at one point in real conditions leads to the need to solve an incorrect inverse problem, where, according to noisy data from one point, it is necessary to roughly estimate the values of parameters in a multi-parameter model of human gait, which evaluates the features of the functioning of the support motor apparatus and the reasons that change its work. All this indicates the need to develop methodological and algorithmic support for the processing and analysis of data on the parameters of a person's approach obtained using wearable devices. Increasing the reliability of the obtained estimates of the values of biometric gait parameters is important for the development of mobile information and analytical systems. The aim of this work is to increase the reliability of the estimated biometric gait parameters according to an acceleration sensor built into the smartphone by developing a neural network algorithm with preliminary correlation processing.

2. Data Processing

The assessment of gait parameters using wearable devices, in particular a smartphone, is carried out in several stages, which in general can be divided into the following: primary processing of the initial data of the acceleration sensor; the step of determining segment repeating areas that may correspond to walking cycles in the analyzed data; evaluation of the parameters of the selected data sections; and comparison of the obtained results with an individual model that was obtained during previous measurements or based on the base model. Taken together, the steps described are computationally expensive, so some of them can be performed on remote servers. At the stage of primary processing of the initial data, noise and high-frequency noise are eliminated using a set of filters, which can be window functions or wavelet filters [14]. Similarly, negative effects are eliminated, for example, accidental movement of the smartphone, the appearance of bias errors, etc., while smoothing the useful signal and reducing the individual features of measurement and gait, so more complex data processing and analysis are required to eliminate interference and noise [1]. At this stage, the user's activity level is also detected to separate the gait phases from random movements and amplitude normalization is performed. Depending on the implementation and processing algorithms, some stages are combined, for example, the user activity assessment performed at the primary processing stage is combined with the stage of identifying repeating sections (the so-called data segmentation), or is performed after segmentation, which makes it possible to pre-separate the selected segments by movement type. Separation by types of movement can be carried out on the basis of threshold criteria (selection of action or inaction) on the basis of informative features using classifiers (including neural networks), as well as on the basis of the autocorrelation function. The selection in these areas corresponding to one of the types of cyclic movements (for example, a single or double step when walking) is performed by different algorithms, which are divided into two groups. The first group includes algorithms that are based on the cyclicity of movements and the search for extremums in the data, crossing zero, which correspond to the boundaries of the elements of the cycle (single or double step). This group also includes algorithms that are based on the analysis of the phase of movements, dynamic changes, and correlation analyses. The search for similar areas is carried out with time normalization, for example, using the DWT time warp method. The second group includes algorithms that allow estimating the boundaries of the segments on the "fly" within the time window, which negatively affects the accuracy of marking data sections when changing the temporal parameters of the gait and its individual phases. Selected data sections are analyzed for the presence of informative features and subjected to statistical evaluation of informative time and frequency parameters. Based on the results obtained, an individual model of a person's gait is built, which is used in the procedures for authentication and evaluation of the parameters of the functioning of the musculoskeletal

system. The input segments under study are compared with model templates using correlation and other statistical apparatus or using neural network technologies [15]. It is necessary to evaluate and take into account the influence of external and internal factors on the measurement results when analyzing the received data due to the peculiarities of measuring gait parameters using a smartphone. According to the works of S. Sprager and M.B. Juric [1], the measurement results are influenced by factors that can be classified as informative (useful) and factors that negatively affect the accuracy of the results and which must be compensated for or taken into account in data processing and analysis. Beneficial factors relate to human physiology—these are changes that occur in the human body during its life. These include uncontrolled changes (physiological and psychological) and controlled physiological and psychological changes [16–18]. Negative factors include the type of clothing and footwear, the parameters of the base on which the person walks, the uniformity of distribution of the additional load on the person or its absence, etc. Negative factors distort the measurement results, for example, when the phone makes random movements in the pocket of loose clothing or a bag. At the same time, some factors may also be useful (informative), corresponding to a certain style of gait in certain conditions. For example, the type of footwear is a negative factor and changes the natural gait of a person (transition from boots to shoes with long heels), but under these conditions, the gait still remains individual. For example, Figure 1 shows a graph of the change in the value of the correlation coefficient when comparing the waveform of the accelerometer of a smartphone (placed in the front pocket of a pair of trousers) when walking in trousers with different degrees of tightness. All measurement conditions, except for the different types of pants, were the same and the results were averaged. According to the results of the analysis of the graph in Figure 1, it follows that the type of pants when walking affects the shape of the accelerometer signal, distorting the signal by up to 30%. This can be misinterpreted and a decision may be made on the distortion of gait parameters, while the gait remains unchanged. Figure 2 also shows a graph of the change in the value of the correlation coefficient when comparing the accelerometer signals of a smartphone (placed in the front pocket of a pair of trousers) when walking in shoes of various types. Changing shoes distorts the shape of the accelerometer signal by up to 12%. Figures 1 and 2 show special cases from the data of one person. The degree of distortion of the accelerometer signal shape in other subjects under the same conditions is different, but the general trend in changes corresponds to the results presented. All this means that, depending on the tasks and scope, it is necessary to correctly classify factors and take this into account when processing and analyzing data [12]. Features of measuring gait parameters using a smartphone must be taken into account in the models and algorithms used, or classified in a template database. The data presented in Figures 1 and 2 were obtained under idealized conditions, when the number of interfering factors was minimized or their influence was the same.



Figure 1. Distortion of the accelerometer waveform with different levels of trouser tightness.

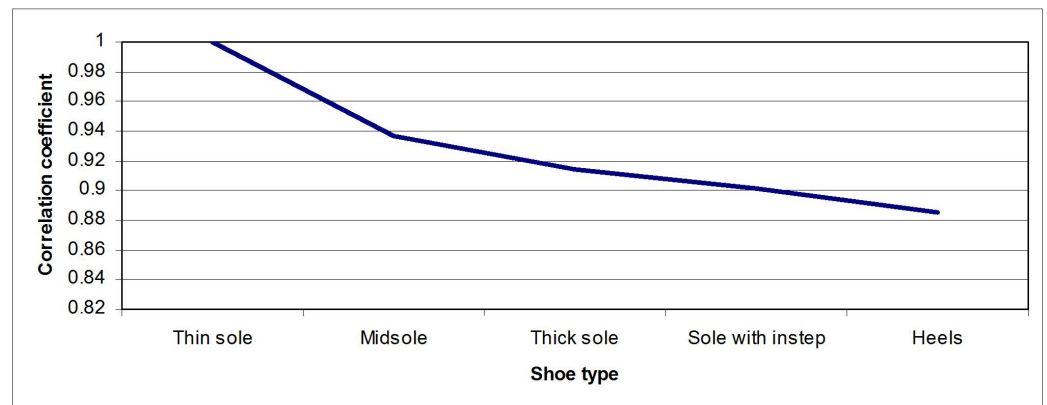


Figure 2. Distortion of the accelerometer waveform for different types of shoes.

In everyday life, such situations do not occur, so different factors can appear at random times and have a random degree of impact, i.e., in processing and analysis, it is difficult or impossible to accurately assess the components that affect the measurement results. From the position of automating the processing of measurement results, the main task is to select data sections corresponding to gait or other analyzed movements. Figure 3 shows a simple example of a time series of smartphone accelerometer data. The central section corresponds to the walking of a person, and in the extreme sections there are random, interfering movements. When segmenting this time series (for example, by extrema), interfering movements often fall into the category of informative signals, which correspond in duration to informative sections (in Figure 4 the boundaries of informative sections are marked in red).

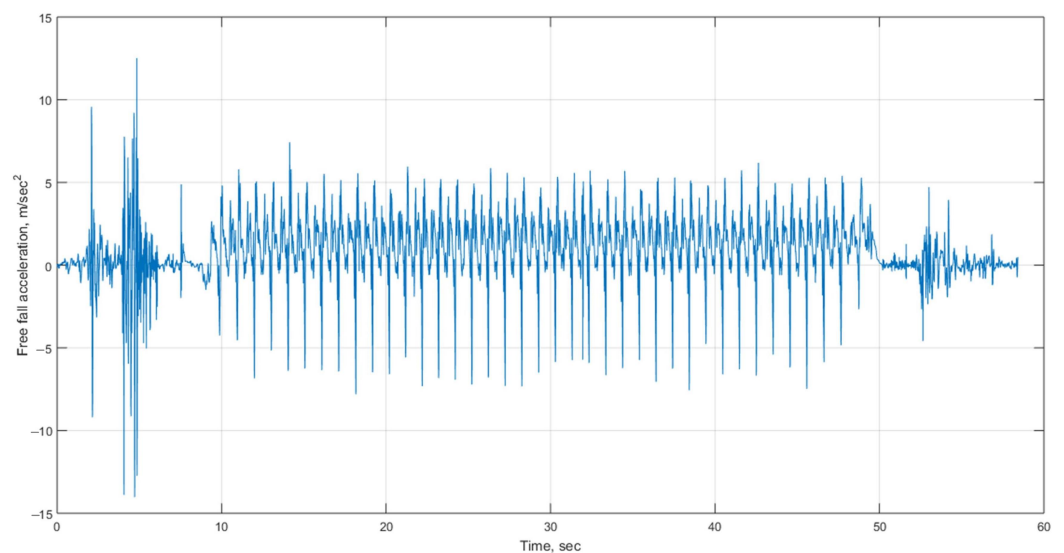


Figure 3. Time series example.

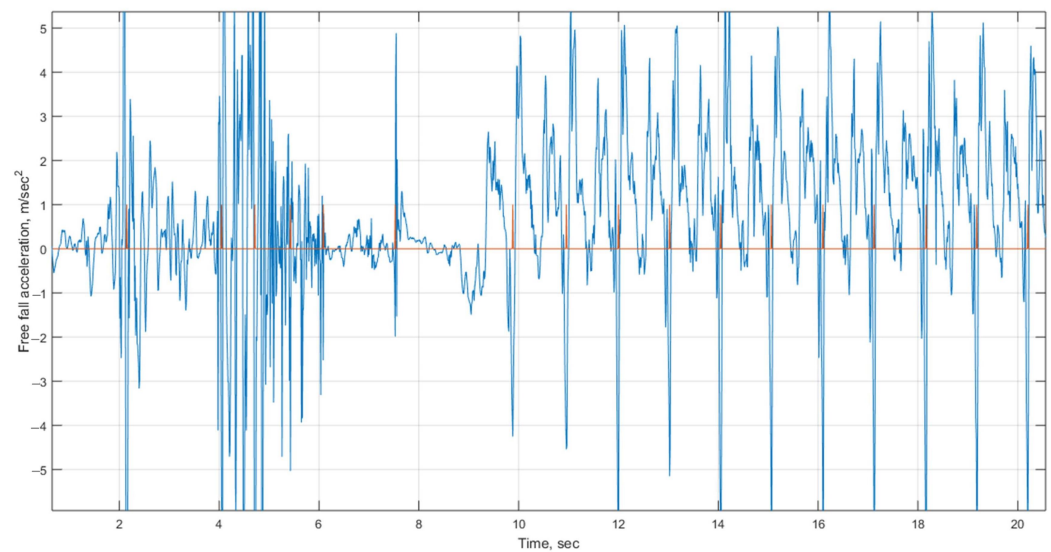


Figure 4. An example of non-informative areas getting into the selection results.

3. Processing Algorithm

In addition to this, when processing a time series with a window function during segmentation, situations arise when one movement corresponds to two time windows with its subsequent division into two movements. An example can be seen in Figure 5 (flat peak of the red mark), where the central section corresponding to one movement falls into two time windows due to the edge of the time window hitting the extremum. To reduce the problems described above, the authors propose an algorithm (Figure 6) for processing smartphone accelerometer measurements. For the high-quality operation of the algorithm, it is necessary to accumulate measurements from the smartphone accelerometer over time T . In authentication systems and personalized medicine, in terms of gait parameters, even a few minutes is not critical, since time series and statistical indicators of gait parameters are analyzed. As specified earlier, in such systems, more attention is paid to the reliability of the estimates obtained, which means a decrease in the number of negative factors in the processing results. The developed algorithm selects areas with the best quality for analysis and accumulation in the template database. The quality criterion is the maximum value of the correlation coefficient. The data section, which in the time domain is most similar to others, carries the least distortion and reflects the constancy in gait over the analyzed period. The quality of the selected areas simplifies the correction of the individual model and improves the accuracy of clustering and classification. In the first steps, the algorithm performs data preprocessing within the analyzed time T ; the standard deviation STD is found after subtracting the constant component, the period is determined, and the local time window $Temp$ is formed based on the amplitude spectrum. The local window $Temp$, the doubled period Tsr of the harmonic Fsr , which has the maximum amplitude in the spectrum, is taken:

$$Temp =]2 \times Tsr \times Fd[, \quad (1)$$

where Fd is the measurement frequency.

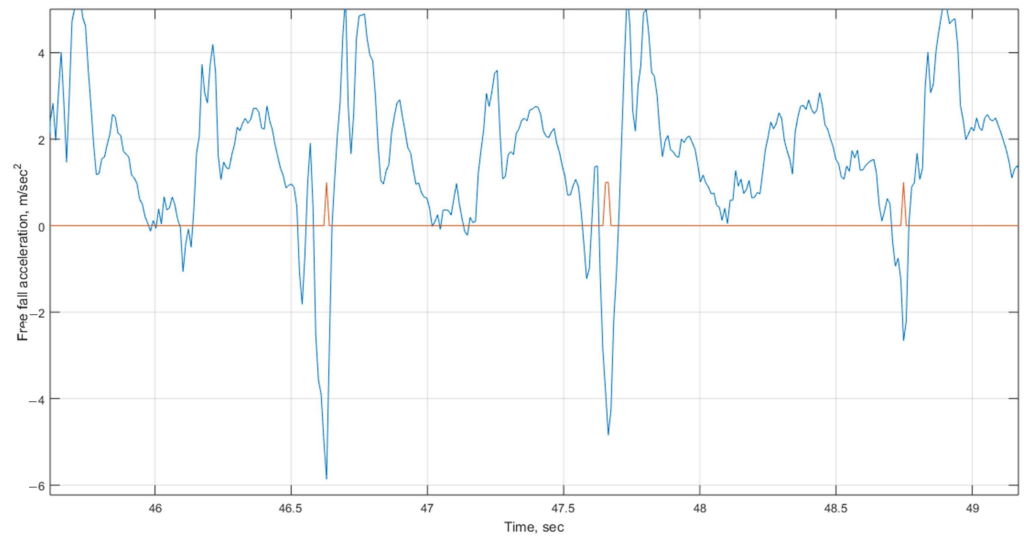


Figure 5. Movement separation example.

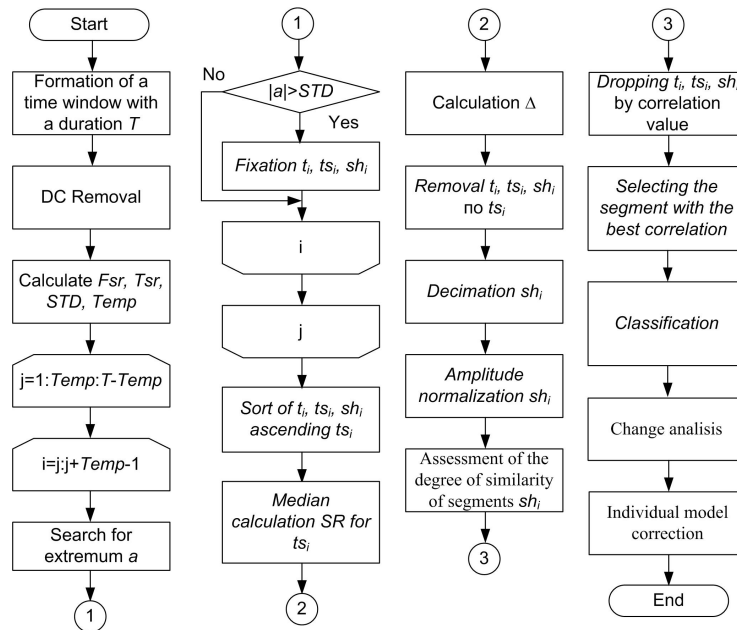


Figure 6. Processing algorithm.

Such a size of the local time window is guaranteed to capture the entire area corresponding to the movement. By moving the local *Temp* window along the analyzed sequence *T* with a step equal to *Temp*, the search for extremums (minimums or maxima) within the local window is performed. The found extremum and its time position *ti* are saved if its value is greater than the standard deviation *STD*, which makes it possible to discard noise components, including non-informative movements. Two adjacent selected extrema form the boundaries of the *shi* segment with a duration of *tsi*. This approach to selecting segments does not distort the data, which is typical for filtering algorithms based on moving averages. The resulting durations *tsi* of the selected sections are arranged in ascending order with sorting of the corresponding values of *ti*, *shi*, *tsi*. According to the found median of the sorted set of durations *tsi*, and taking into account the possible deviation in the duration of the segments by Δ , segments are rejected based on the expression $SR - \Delta < tsi < SR + \Delta$:

$$\pm \Delta = SR \pm \left[\frac{SR \times Int}{100} \right], \quad (2)$$

where Int is the allowable percentage deviation in the duration of the found segments from the obtained median in percent. This percentage adjusts to the person's gait and at the initial stage can take a value of 10%.

The remaining sections are compressed by the decimation procedure to the size $\min(ts_i)$ of the shortest section sh_i . Then, by calculating the value of the correlation coefficient, a search is made for the area most similar to the others, and the areas that have a correlation coefficient of less than a given threshold are removed (in this work, the value of the correlation coefficient equal to 0.9 was taken as a threshold). As a result, the sections of the time series and their parameters (the value of the correlation coefficient, the mean and standard deviation) which remain are fed to the neural network classifier. In this work, the neural network classifier was built on the basis of a feed-forward network. At the output of the classifier, the number of the movement is obtained, which corresponds to a particular section, as well as the degree of correspondence of the movement to a particular person. In the case of the expected outcome, the individual model is adjusted and the trend in individual gait measures is determined.

4. Results and Conclusions

The application of the developed algorithm increased the reliability of the obtained estimates and the quality of intermediate data at the stage of primary processing (Figure 7). At the stage of primary processing, data sections were selected that relate only to useful data (analyzed movements during walking) and no duplication of sections was found. As a result of the correlation processing of the selected sections, heavily noisy sections were removed that corresponded to random movements not related to the analyzed ones, as well as sections that had distortions, reflecting stumbling while walking, stopping and changing trajectory, the type of movement, etc. When constructing the neural classifier, feed-forward networks were used; their training was carried out by the gradient descent method with cross-entropy as an optimization criterion. The activation function for the hidden layers was the sigmoid and for the output layer it was the normalized exponential function. Testing a neural network classifier with at least 100 hidden layers on a test set showed a correct decision probability of 0.95. Reducing the number of hidden words to 10 led to a decrease in the probability of correct decision making to 0.9 (Figure 8). The classification of the subject's own movements was 100%.

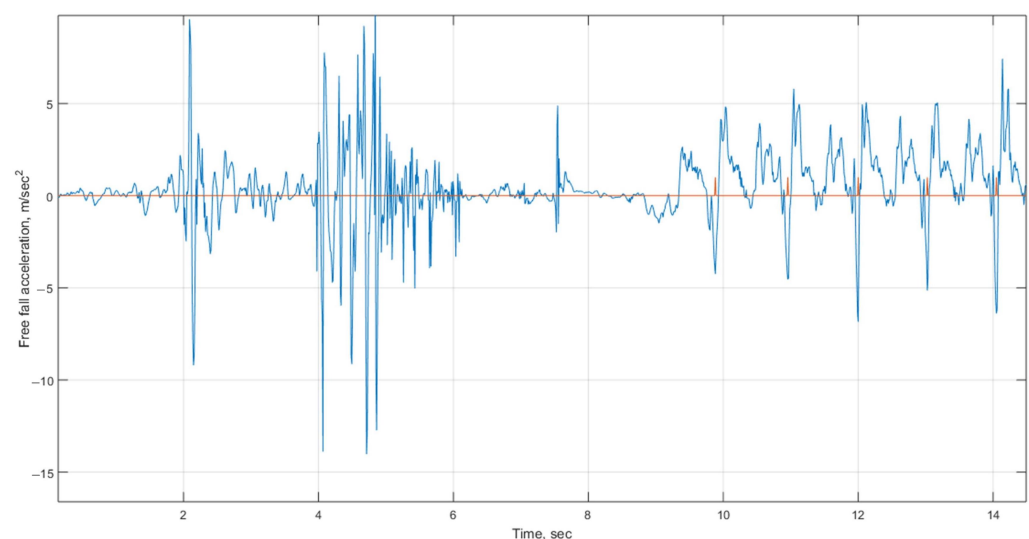


Figure 7. The result of primary processing of the time series.

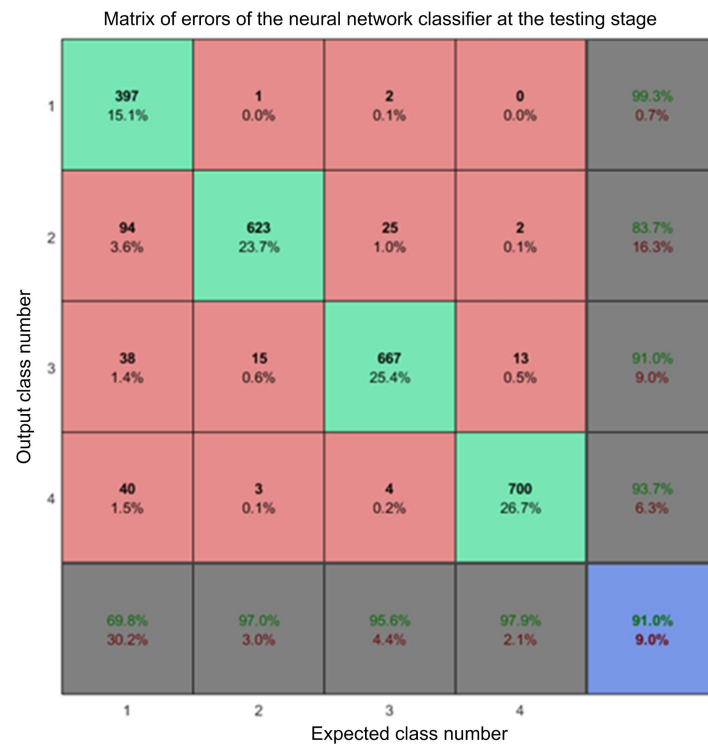


Figure 8. Classification results into four groups.

The application of the developed algorithm in practice made it possible to automate the process of selecting the best informative areas and forming a database of human movement patterns. Automatic adjustment of the allowable limit of the size of the data section allows one to track the change in gait speed and adjust the size of the local time window when selecting sections. The neural network classifier determines the type of movement and it belonging to a particular person. The developed algorithm is not inferior to similar solutions in terms of the quality of selection and discrimination of movements in the accelerometer data of a mobile phone.

Author Contributions: Conceptualization, N.D. and A.G.; methodology, N.D.; software, N.D.; validation, N.D. and A.G.; formal analysis, A.G.; investigation, N.D.; resources, N.D.; data curation, A.G.; writing—original draft preparation, A.G.; writing—review and editing, N.D.; visualization, N.D.; supervision, A.G.; project administration, A.G.; funding acquisition, A.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the grant of the President of the Russian Federation No. MK-1558.2021.1.6.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sprager, S.; Juric, M.B. Inertial Sensor-Based Gait Recognition: A Review. *Sensors* **2015**, *15*, 22089–22127. [[CrossRef](#)] [[PubMed](#)]
2. Connor, P.; Ross, A. Biometric recognition by gait: A survey of modalities and features. *Comput. Vis. Image Underst.* **2018**, *167*, 1–27. [[CrossRef](#)]
3. Guelta, B.; Tlemsani, R.; Chouraqui, S.; Benouis, M. An Improved Behavioral Biometric System based on Gait and ECG signals. *Int. J. Intell. Eng. Syst.* **2019**, *12*, 147–156. [[CrossRef](#)]
4. Ren, Y.; Chuah, M.C. User Verification Leveraging Gait Recognition for Smartphone Enabled Mobile Healthcare Systems. *IEEE Trans. Mobile Comput.* **2015**, *14*, 1961–1974. [[CrossRef](#)]

5. Khabir, K.M.; Siraj, M.S.; Ahmed, M.; Ahmed, M.U. Prediction of gender and age from inertial sensor-based gait dataset. In Proceedings of the Joint 2019 8th International Conference on Informatics, Electronics & Vision (ICIEV) & 3rd International Conference on Imaging, Vision & Pattern Recognition (IVPR, Spokane, WA, USA, 30 May–2 June 2019; pp. 371–376.
6. Shema-Shiratzky, S.; Beer, Y.; Mor, A.; Elbaz, A. Smartphone-based inertial sensors technology—Validation of a new application to measure spatiotemporal gait metrics. *Gait Posture* **2022**, *93*, 102–106. [[CrossRef](#)] [[PubMed](#)]
7. Garufov, D.; Bours P. User authentication based on foot motion. *SIViP* **2011**, *5*, 457–467.
8. Reyes, O.C.; Vera-Rodriguez, R.; Scully, P.J.; Ozanyan K.B. Analysis of Spatio-Temporal Representations for Robust Footstep Recognition with Deep Residual Neural Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *4*, 1–13.
9. Sokolova, A.I.; Konushin, A.S. Methods of human identification by gait in video. *Proc. ISP RAS* **2019**, *1*, 1–19.
10. Helvas, A.V.; Belyaikina, N.G.; Gilya-Zetinov, A.A.; Chernikova, D.D.; Shabunin, V.M.; Yapyrintsev E.O. Gesture recognition using a neural network and the application of this approach to create a new generation of gaming gadgets. *Proc. Mosc. Inst. Phys. Technol.* **2017**, *2*, 1–7.
11. Teh, P.S.; Zhang, N.; Tan, S.-Y.; Shi, Q.; Khoh, W.H.; Nawaz, R. Strengthen user authentication on mobile devices by using user's touch dynamics pattern. *J. Ambient Intell. Humaniz. Comput.* **2020**, *11*, 4019–4039. [[CrossRef](#)]
12. Zhang, X.; Yao, L.; Huang, C.; Gu, T.; Yang, Z.; Liu Y. DeepKey: A Multimodal Biometric Authentication System via Deep Decoding Gaits and Brainwaves. *ACM Trans. Intell. Syst. Technol.* **2020**, *11*, 3393619. [[CrossRef](#)]
13. Grecheneva, A.V.; Dorofeev, N.V.; Goryachev M.S. Estimation of human biomechanics during registration with a wearable device. *J. Phys. Conf. Ser.* **2021**, *2096*, 012117. [[CrossRef](#)]
14. Lu, H.; Huang, J.; Saha, T.; Nachman L. Unobtrusive gait verification for mobile phones. In Proceedings of the 2014 ACM International Symposium on Wearable Computers, Seattle, WA, USA, 13–17 September 2014; pp. 91–98.
15. Oguz, A.; Ertugrul O.F. Human identification based on accelerometer sensors obtained by mobile phone data. *Biomed. Signal Process. Control* **2022**, *77*, 103847. [[CrossRef](#)]
16. Tandon, R.; Javid, P.; Giulio I.D. Mobile phone use is detrimental for gait stability in young adults. *Gait Posture* **2021**, *88*, 37–41. [[CrossRef](#)] [[PubMed](#)]
17. Pierce, A.; Ignasiak, N.K.; Eiteman-Pang, W.K.; Rakovsky C. Mobile phone sensors can discern medication-related gait quality changes in Parkinson's patients in the home environment. *Comput. Methods Programs Biomed. Update* **2021**, *1*, 100028. [[CrossRef](#)]
18. Lunardini, F.; Malavolti, M.; Pedrocchi, A.L.G.; Borghese, N.A.; Ferrante S. A mobile app to transparently distinguish single-from dual-task walking for the ecological monitoring of age-related changes in daily-life gait. *Gait Posture* **2021**, *86*, 27–32. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.