Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 225 (2023) 374–383

Procedia Computer Science

www.elsevier.com/locate/procedia

27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2023)

Identification and authorization with single accelerometer data - implications from "Wearables in Arthritis" project

Pawel Mielnik[a,*], Anja Myhre Hjelle[a], Bjarte Pollen[b], Krzysztof Tokarz[c], Marcin Fojcik[b]

[a]*Helse Førde, Førde, Norway*
[b]*Western Norway University of Applied Sciences, Førde, Norway*
[c]*Silesian University of Technology, Gliwice, Poland*

Abstract

Introduction: Miniaturisation and development of mobile devices led to accelerometers being literally everywhere. The way we are moving seems intuitively very personal, resulting in increased interest in the possibility of accelerometer-based authorisation. We present results of an experiment based on data from the "Wearables in Arthritis" project.
Aims: The main aim of this work was to assess if it is possible to identify participants based on data from a single accelerometer.
Methods: Participants' accelerometer data were collected during clapping and walking. We trained models using k-nearest neighbours (KNN), naïve Bayes (NB), random forest (RF), extreme gradient boosting (xgboost – XGB), and single-hidden-layer neural network (NNET) algorithms. We analysed data separately from clapping and walking, and both together. The most effective algorithm was RF on clapping data, with an accuracy of 0.992. We examined the effect av "imposter" data from participants not used in training. The performance was lower but still acceptable, with an accuracy between 0.860-0.929, depending on the probability threshold.
Conclusions: The results show the potential of accelerometer data for biometric authorisation but also raise concerns about data privacy.

Keywords: accelerometry; data analyse; chronic arthritides;

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .
E-mail address: pawel.franciszek.mielnik@helse-forde.no

10.1016/j.procs.2023.10.022

1. Introduction

Accelerometry (ACC) is a widely accepted method for registering and analysing of human physical activity. Miniaturisation led to wide availability of inertial sensor modules (IMU), which can be found in many electronic devices such as mobile telephones, tablets, and smartwatches. IMUs are also used in robotics and aeronautics. In humans, they are used for classifying physical activity, such as walking and jogging [1,2]. IMU step counting analysis can estimate personal physical load, which can help evaluate health status [3]. For example, accelerometry can detect patient falls [4] and epileptic seizures [5] or be helpful in rehabilitation [6]. Accelerometry seems to be less risky in terms of personal data security than video and sound monitoring. However, it also raises concerns about identity security [7], which will be discussed later in the paper.

This work is a part of a larger project about accelerometry in diagnostics and follow-up of patients with chronic arthritides [8,9]. Chronic arthritides are a group of diseases characterised by prolonged inflammation in the joint, which can cause permanent joint destruction when not effectively treated [10]. Active joint inflammation results in pain, swelling, and reduced motion and accelerometry seems to be a potentially valuable instrument to detect it. Interim analysis of our data shows that accelerometry may be an effective method of detecting arthritis activity. However, we found some instability in classification models (unpublished data). Small number of participants can cause this instability, although influence of other parameters is also possible. In this study, we aimed to assess interpersonal difference between accelerometer data, focusing on the identification of individuals.

Biometry is an established method for personal identification and is widely utilised in modern appliances such as mobile telephones [11]. However, biometry methods mainly use fingerprints or face recognition. Accelerometer biometry was proposed previously [12], but the authors do not know about any comprehensive commercial implementation of the method.

The main aim of this work was to assess if it is possible to identify participants based on data from a single accelerometer. The secondary goal was to evaluate the machine learning methods for accelerometry classification, especially in the context of the whole project on accelerometry in chronic arthritides.

2. Relevant literature

Although accelerometer biometry is not broadly implemented, it is widely discussed in the literature. Most publications focus on gait analysis, the most natural human activity. Mäntyjärvi and colleagues proposed a portable device, identifying person by ACC analysis [13]. The authors compared subject reference to test signals using a correlation between samples, histograms, higher moment statistics and fast Fourier transform (FFT) analysis. It is not clearly stated, but it does not seem that the authors utilised machine learning. Despite that, they achieved good performance with the best effect for correlation with an equal error rate of 7%.

Similarly, Gafurov et al. used absolute (Manhattan) distance, correlation, histogram analysis, and higher-order moments in person identification based on ACC walking analysis [14]. The authors reported similar effectiveness as in the previous works. However, they used a different method of source data extraction. They used six consecutive walking cycles, excluding the first two, which were more unstable.

The problem of gait segmentation was discussed elsewhere by Derawi [15]. The author addressed the need for walking segmentation based on the gait cycle selection. It is because the walking patterns can change dynamically not only in real life but also in laboratory conditions.

User recognition by ACC gait data obtained from Android-based smartphones is the subject of work by Juefei-Xu and colleagues [16]. The authors used standard Android devices with an ACC registering application instead of laboratory devices. It caused problems with ACC pacing due to the operating system scheduler, which is not real-time designed. Despite it, the authors achieved a reasonable 99.4% subjects verification rate (VR) and 0.1% false accept rate (FAR). After extracting features from ACC signal, they used supported vector machines (SVM) as the analysis tool.

The other approach of gait-based person identification is presented in a work by Cola et al. [17]. The author proposed a real-time system for ACC-based authorisation. It consists of a gait detection module and an authentication module. The system starts to register gait samples after detecting eight consecutive steps. As in the previous publication, mathematical features were extracted from the ACC signal before proper analysis. Again, a similar performance to Juefei-Xu was found.

The ACC signal can be sensitive to disturbances, as shown by Zali et al. [18]. In this experiment, the researchers showed that ACC data analysis could detect if the participant had drunk a small amount of alcohol.

A different approach was presented by Carlson et al. [19]. The researchers used both a linear accelerometer and a gyroscope to identify users. They assessed data from mobile phone sensors under answering, during the call, and during hanging out. Experiments showed high accuracy in subject classification.

The rise of smartwatches in recent years has drawn researchers' attention to gesture recognition as an authentication tool. Buriro et al. analysed the publicly available ACC and gyroscope dataset from a smartwatch [20]. They used data collected under clapping and found them capable of identifying the subjects. The authors reported that the k-nearest neighbours (KNN) was the most and random forest (RF) the least effective algorithm. Experiments showed high accuracy in subject classification, although small data samples were used for training. The authors did not use extern samples, not included in the training set, as a potentially confusing factor.

3. Methods

We used accelerometry data and basic anthropomorphic data from participants included in the main study until 28th February 2023. All data analysis was performed with R v. 4.2.2 [21]. We used caret package for R for model training and testing [22]. Graphics were made with R library ggplot [23].

3.1. Ethics and approval

The study has been approved by the local Ethics Committee and accepted by the personal data protection officer. All participants have signed informed consent form (patient consent form).

3.2. Material

We analysed data from 28 participants, therein 14 healthy controls and 14 arthritis patients. Two participants were male, and 25 were female. All participants had to fulfil inclusion/exclusion criteria per the protocol. Controls were volunteers that had no arthritis or other exclusion disorders. All participants were asked to perform physical activities such as walking (fast and slowly), clapping, going downstairs/upstairs, moving arms to the sides and forward, sitting in a chair, standing up from the couch, and standing still. Basic demographic and anthropomorphic data such as age, gender, height, and weight was also collected. Patients' data was supplemented with clinical data about arthritis to evaluate arthritis activity status. Clinical data was not used in this manuscript.

This manuscript presents only accelerometer signals from fast walking and clapping. The choice of walking was because it is the most natural human movement. Clapping is a simple repetitive activity involving only the upper limbs and reducing the influence of gait disturbances due to diseases.

3.3. Data collection

Accelerometer data was collected using Arduino Nano 33 developing board equipped with 9-axis IMU LSM9DS1 from STMicroelectronics and nRF52840 microcontroller from Nordic Semiconductors. We used homebrewed software both for the board and recorder. For the board, we used Arduino IDE. The standard PC was the recorder with recording software developed in Python 3.10.6. ACC data was saved as plain text in CSV format. We registered three linear axes: X, Y, and Z. Participants wore the device on the left forearm, about 5 cm above the wrist, similar to a watch.

3.4. Data preparation

We applied a high-pass Butterworth filter to all ACC axes to remove gravidity impact. The ACC signal was divided into smaller 3 seconds batches. We extracted the following features from every data chunk: mean, standard deviation, median absolute deviation, variance, signal magnitude vector, signal magnitude area, skewness, kurtosis, furrier frequencies, autocorrelation coefficients, zero-crossing count, the correlation between axes, FFT energy, and total entropy. The features were described elsewhere [9]. Variables were scaled before analysis.

The data was balanced with the undersampling. It concerned only eight patients with more than one visit (3 visits in four patients and 2 for the rest). Undersampling was done randomly, maintaining an equal proportion of samples according to visits. Data were subsequently randomly divided into train and test bulks in a ratio of 4/1.

3.5. Data analysis

All training models, model comparisons, and performance evaluations were made using R's caret library. Caret library is a software for rapid machine learning model design and testing. It employs over 200 different algorithms containing many tuning parameters. Process details differ between algorithms and are more precisely described in library documentation. "Train()" caret function resamples data with K-fold cross-validation to find the most efficient model. We built models using k-nearest neighbours (KNN), naïve Bayes (NB), random forest (RF), extreme gradient boosting (xgboost – XGB), and single-hidden-layer neural network (NNET) algorithms. We used the default tuning parameter for methods that can be applied. The models were built separately for clapping, walking ACC data and joined activities. As walking and clapping were not performed simultaneously, we connected ACC data chunks consecutively.

For KNN analysis, it was used a built-in R function. For the rest of the algorithms, the train function engages other libraries: "naivebayes" library for NB [24], "randomForest" for RF [25], and xgboost and NNET libraries with the same names as the names of the methods [26,27].

Accuracy was the primary performance outcome. We calculated accuracy and kappa for predictions on test data for all models.

We investigated performance of the training models by analysing accuracy distribution for all cycles of training cross-validation.

We analysed together data from patients and controls. We focused on subject identification by ACC signal in this work. The arthritis status is a potential noise source. To avoid subject recognition by arthritis activity, we used patient data from visits when high activity was observed and when the disease was in remission in equal proportions.

In the second part of the study, we performed an experiment simulating a real authentication problem. First, we randomly chose eight subjects, four from each group, whose ACC clapping data was used as "imposters". We generated in this way 705 "imposter" samples. Then, we trained the model with the rest of the clapping data using the RF algorithm. Finally, we compiled "imposter" with other subjects' test data and calculated classification probabilities for every subject. Those probabilities were used to classify data with thresholds ranging from 0.5 to 0.676, the maximum probability observed in "imposters" data. We calculated the false acceptance rate (FAR), false rejection rate (FRR), true acceptance rate (TAR) and accuracy.

FAR was defined as falsely accepted (false positive) to all "imposter" samples number. FRR is a ratio of false rejected subjects (false negative) to all subject samples. TAR is a per cent of accepted subjects' samples.

4. Results

Table 1 shows group characteristics. The mean age was 53, with a relatively narrow 95% CI. It is because most real-life arthritis patients are over 50. Our cohort does not differ from the usual chronic arthritis patient population.

Table 1. Participants characteristics

Variable	mean	95pr CI upper	95pr CI lower
Age	53.54	48.52	58.55
Height	165.39	162.10	168.69
Weight	71.41	66.02	76.81

We have achieved decent results of classification for all classification results and all activity combinations. It means that ADD analysis is an effective method for person identification. However, controlling using the test data reveals that RF was the most effective classification algorithm for all activities, with the highest accuracy for clapping (0.992, 95% CI 0.978-0.998). It corresponds to the same value of sensitivity and 0.999 specificity.

Table 2. Test classification results for both clapping, walking and combined data.

Activity	Method	Accuracy	Kappa	Accuracy Lower CI	Accuracy Upper CI
Clapping	KNN	0.980	0.979	0.960	0.991
Clapping	NB	0.959	0.958	0.935	0.976
Clapping	NNET	0.875	0.870	0.838	0.906
Clapping	RF	0.992	0.992	0.978	0.998
Clapping	XGB	0.980	0.979	0.960	0.991
Walking	KNN	0.739	0.730	0.698	0.778
Walking	NB	0.735	0.725	0.693	0.774
Walking	NNET	0.670	0.658	0.626	0.712
Walking	RF	0.857	0.852	0.822	0.887
Walking	XGB	0.845	0.839	0.809	0.876
Both	KNN	0.967	0.966	0.944	0.982
Both	NB	0.903	0.899	0.869	0.930
Both	NNET	0.763	0.754	0.717	0.804
Both	RF	0.985	0.984	0.967	0.994
Both	XGB	0.982	0.981	0.964	0.993

All models were most efficient in the classification of clapping and worst in gait, with accuracy under 0.9 in all cases. Astonishingly, adding walking to clapping data did not improve the recognition of subjects compared to clapping data analysis alone, despite providing much better performance than walking. The NNET algorithm provided the worst results thorough all activities. Aggregated performance results are shown in Table 2. Figure 1 shows the heatmap diagrams for all algorithms. It visualises confusion matrixes for test data against prediction, coded with colour tone. One can see no particular pattern in misclassifications which could have suggested a significant bias.

A closer look at the trained model shows similar algorithm effectiveness as from test data (Figure 2). The NNET had the lowest mean accuracy of all resampling epochs and the worst precision.

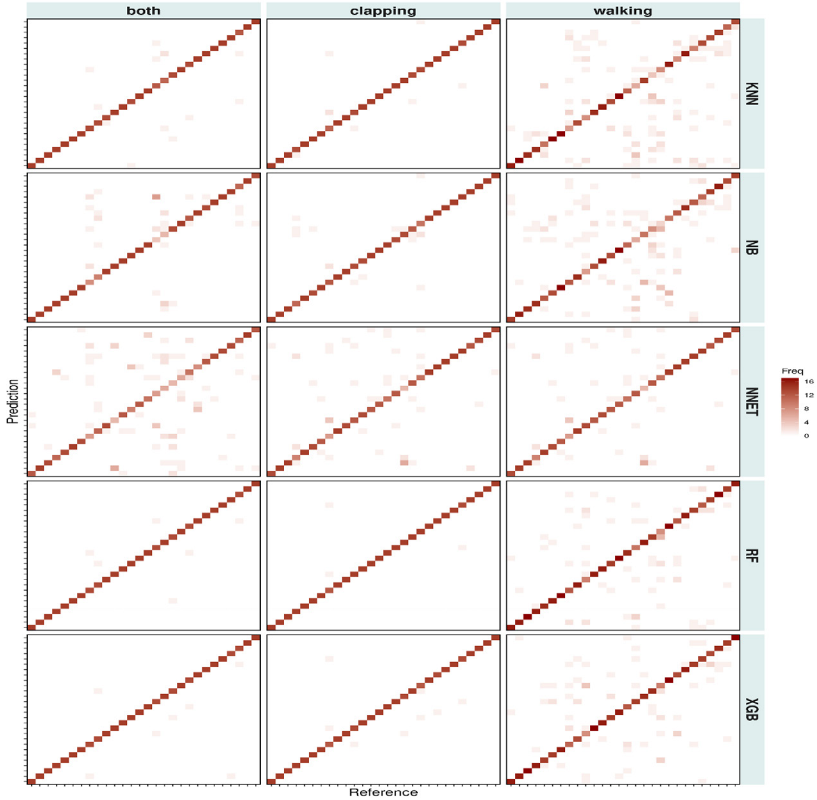

Fig. 1. Heatmap diagram presenting classification results for all three activities and all algorithms.

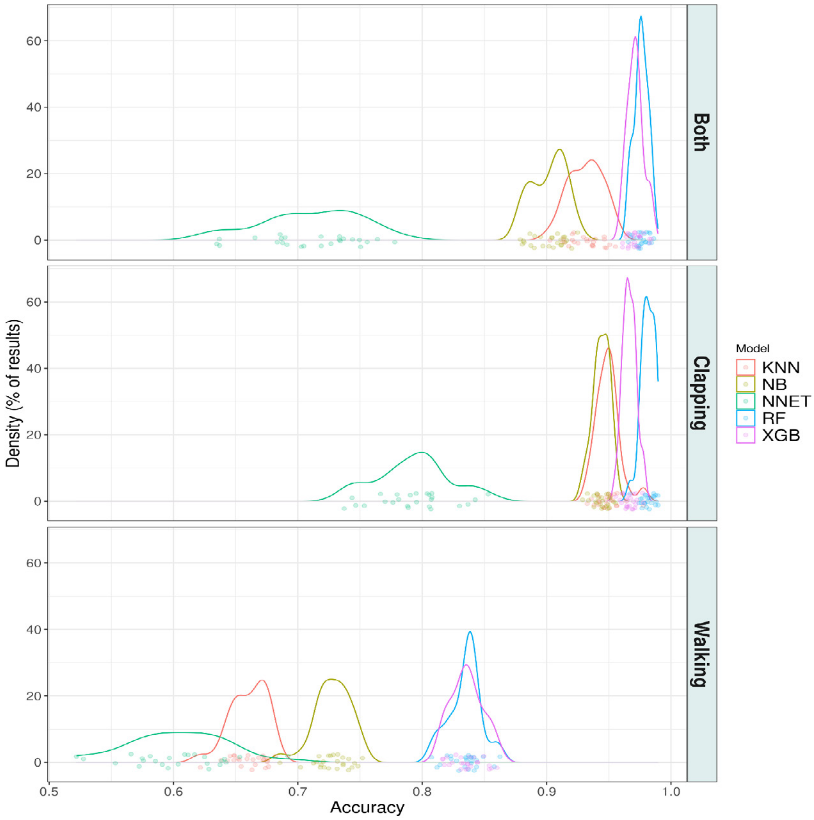

Fig. 2. Density diagram showing performance and precision of training models. It is not based on test data. Every dot represents the result of training validation.

Experimenting with "imposters" data showed that the RF algorithm could effectively classify ACC samples despite the noise from foreign data. Furthermore, the ROC curve showed good classification performance despite the threshold of probability value (Figure 3). AUC was 0.998 with a 95% CI of 0.993-1.0. Most of the wrong classification was made for one subject, as seen in the diagram.

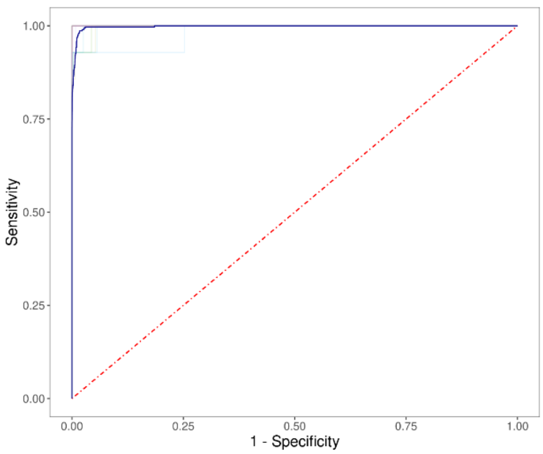

Fig. 3. ROC curve for RF classification in the model with "imposters". The blue line shows averaged ROC curve for all subjects. Bright lines represent curves for each subject. As it can be seen, one subject has a poorer response.

FAR and FRR fluctuated naturally with the used threshold. FAR was zero at the point of 0.676, but both FRR and TAR were reduced to 27.3 and 72.8%, respectively (Table 3). Figure 3 shows estimated FAR and FRR ratios for different threshold values.

Table 3. Results of the experiment with "imposters."

Thresholds	Accuracy	Accuracy Lower CI	Accuracy Upper CI	FAR	FRR	TAR
0.500	0.860	0.837	0.881	15.177	11.224	88.776
0.525	0.881	0.859	0.900	11.915	11.905	88.095
0.550	0.896	0.875	0.914	8.511	14.966	85.034
0.575	0.917	0.898	0.933	4.681	17.007	82.993
0.600	0.929	0.911	0.944	2.270	18.707	81.293
0.625	0.925	0.907	0.940	1.135	22.789	77.211
0.650	0.923	0.905	0.939	0.426	25.170	74.830
0.676	0.920	0.901	0.936	0.000	27.211	72.789

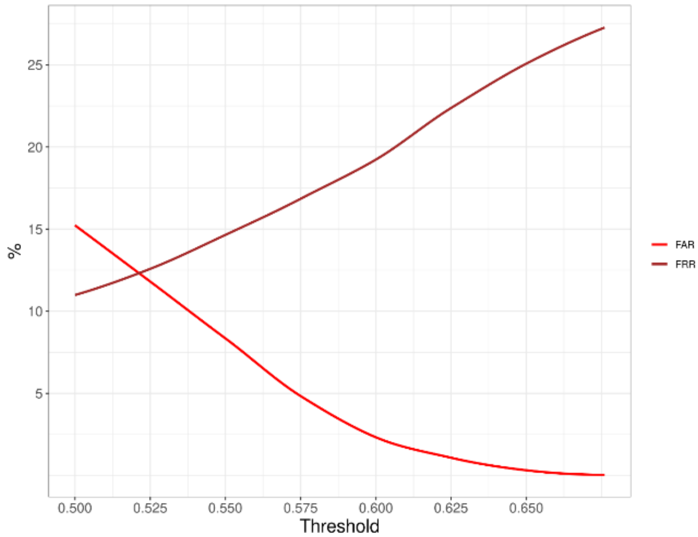

Fig. 4. Estimated FAR and FRR to classification threshold. The cross point is at threshold 0.524.

5. Discussion

The results are similar to those previously reported for both physical activity types [17,20]. In addition, our classification models performed better on ACC clapping than walking data, which was not investigated earlier. The best effective algorithm in the study was the random forest. Buriro et al. demonstrated that KNN was more efficient for clapping-based classification [20]. The RF classification had an accuracy of only 86, whereas KNN was 90.7. We had a very high performance for almost all engaged algorithms on clapping (Table 2). Only NNET accuracy was below 90.

The experiment with «imposter» data has shown that the algorithm can perform well despite confusing factors. The algorithm produced lower efficacy in this situation, but it was still high. Proper maneuvering with a classification probability threshold can reduce the risk of failed authorisation. Mantyjarvi et al. used more “imposter” samples than in this work [13]. Although the authors used different classification methods without using machine learning, the results are similar. The authors concluded that ADD is promising biometrical authorization method.

On the other hand, our results and earlier research reveal that ACC data is strictly personal and should be protected along with passwords, images, signatures, fingerprints etc. [14, 15, 16] The same data protection rules should then apply to ACC data as to the others. It encompasses obtaining consent for collecting and processing data, reporting data leaks etc. Our study participant group is unique due to consists of both healthy volunteers and patients with active arthritis, which can significantly disturb movement patterns. The decent classification outcomes of subjects in our cohort imply that ADD can be invulnerable to additional influence. The primary study limitation is the relatively small size of the study group. The group is also relatively homogenous because it originated from one region's population.

Different biometric technics are employed in personal identification and authorisation nowadays. ADD should be further investigated as an option, especially in addition to face, voice or fingerprint recognition. Parallel use of different data sources can increase the viability of biometric authorisation.

6. Conclusions

We found that ADD data from a single sensor is highly specific for an individual. Therefore, it can be possible to detect precisely to whom the sample belongs—those findings have many potential implications for biometry authorisation development and personal data safety.

The results also have implications for the main project. First, it showed that strong individual signal dependence could negatively inflict the possibility for evaluation of clinical status on a cohort basis. However, it should not affect assessments on an individual basis.

References

1. Bårdstu HB. (2016) Detection of physical activity types with accelerometers in adolescents during semi- structured free-living. BEV3901 Master's Thesis Hum Mov Sci Trondheim.
2. Bayat A, Pomplun M, Tran DA. (2014) A Study on Human Activity Recognition Using Accelerometer Data from Smartphones. Procedia Comput Sci.**34:** 450–457.
3. Lee I-M, Shiroma EJ. (2014) Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. Br J Sports Med. **48:** 197–201.
4. Chen K-H, Yang J-J, Jaw F-S. (2016) Accelerometer-based fall detection using feature extraction and support vector machine algorithms. Instrum Sci Technol. **44:** 333–342.
5. Jallon P, Bonnet S, Antonakios M, Guillemaud R. (2009) Detection system of motor epileptic seizures through motion analysis with 3D accelerometers. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Minneapolis, MN: IEEE 2466–2469.
6. Iqbal MH, Aydin A, Brunckhorst O, Dasgupta P, Ahmed K. (2016) A review of wearable technology in medicine. J R Soc Med. **109:** 372–380.
7. Middya AI, Roy S, Mandal S, Talukdar R. (2021) Privacy protected user identification using deep learning for smartphone-based participatory sensing applications. Neural Comput Appl. **33:** 17303–17313.
8. Mielnik P, Tokarz K, Mrozek D, Czekalski P, Fojcik M, Hjelle AM. (2019) Monitoring of Chronic Arthritis Patients with Wearables - A Report from the Concept Phase. In: Nguyen NT, Chbeir R, Exposito E, Aniorté P, Trawiński B, (eds) Computational Collective Intelligence. Springer International Publishing.
9. Mielnik P, Fojcik M, Tokarz K, Rodak Z, Pollen B. (2021) Detecting of minimal changes in physical activity using one accelerometer sensor. Advances in Computational Collective Intelligence: 13th International Conference, ICCCI 2021
10. Smolen JS, Landewé RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, (2019) EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis. 2020. **79:** 685–699.
11. Dijmărescu I, Iatagan M, Hurloiu I, Geamănu M, Rusescu C, Dijmărescu A. (2022) Neuromanagement decision making in facial recognition biometric authentication as a mobile payment technology in retail, restaurant, and hotel business models. Oeconomia Copernic **13:** 225–250.
12. Accelerometer Biometric Competition. https://kaggle.com/competitions/accelerometer-biometric-competition
13. Mantyjarvi J, Lindholm M, Vildjiounaite E, Makela S, Ailisto H. (2005) Identifying Users of Portable Devices from Gait Pattern with Accelerometers. Proceedings (ICASSP' 05) IEEE International Conference on Acoustics, Speech, and Signal Processing

14. Gafurov D, Snekkenes E, Bours P. (2007) Gait Authentication and Identification Using Wearable Accelerometer Sensor. IEEE Workshop on Automatic Identification Advanced Technologies.
15. Derawi MO. Accelerometer-Based Gait Analysis, A survey. (2010) Norwegian Information Security Conference : NISK 2010 .
16. Juefei-Xu F, Bhagavatula C, Jaech A, Prasad U, Savvides M. (2012) Gait-ID on the move: Pace independent human identification using cell phone accelerometer dynamics. 2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS).
17. Cola G, Avvenuti M, Vecchio A. (2017) Real-Time Identification Using Gait Pattern Analysis on a Standalone Wearable Accelerometer. Comput J. **60(8):** 1173–1186
18. Zaki THM, Sahrim M, Jamaludin J, Balakrishnan SR, Asbulah LH, Hussin FS. (2020) The Study of Drunken Abnormal Human Gait Recognition using Accelerometer and Gyroscope Sensors in Mobile Application. 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)
19. Carlson C, Chen T, Cruz J, Maghsoudi J, Zhao H, Monaco JV. (2015) User Authentication with Android Accelerometer and Gyroscope Sensors.
20. Buriro A, Ricci F. (2023) ClapAuth: A Gesture-Based User-Friendly Authentication Scheme to Access a Secure Infrastructure. In: Saracino A, Mori P, (eds). Emerging Technologies for Authorisation and Authentication. Cham: Springer Nature Switzerland
21. R Core Team. (2021) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
22. Kuhn M. (2020) Caret: Classification and Regression Training.
23. Wickham H. (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York
24. Majka M. (2019) Naivebayes: High Performance Implementation of the Naive Bayes Algorithm in R.
25. Breiman L. (2001) Random Forests. Mach Learn **45:** 5–32.
26. Chen T, Guestrin C. (2016) XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
27. Venables WN, Ripley BD. (2002) Modern Applied Statistics with S. : Springer; New York