

THE APPLICATION OF AN ENSEMBLE OF CONVOLUTIONAL NEURAL NETWORKS FOR HUMAN RECOGNITION BASED ON THE GROUND REACTION FORCES

Marcin DERLATKA*

*Institute of Biomedical Engineering, Faculty of Mechanical Engineering,
Bialystok University of Technology, Wiejska 45C, 15-351 Bialystok, Poland

m.derlatka@pb.edu.pl

received 13 August 2025, revised 07 November 2025, accepted 07 November 2025

Abstract: One of the biometric methods is human recognition based on ground reaction forces (GRFs) generated during a person's gait. Conventional methodologies for gait recognition have relied on the manual extraction of features from measured GRF signals, in conjunction with the utilization of machine learning algorithms. Recently, convolutional neural networks (CNNs) have become increasingly popular due to their ability to automatically extract features from signal data. However, the CNNs don't always produce optimal results for human recognition. In this study, we emphasize a novel aspect of the approach: the use of an ensemble of homogeneous CNN classifiers, all sharing the same architecture but trained on different combinations of GRF components. This strategy leverages diversity originating purely from data representation rather than architectural variation, demonstrating that even identical CNNs can complement each other when exposed to distinct training data. The objective of this paper is to design a biometric system that recognizes humans based on GRFs and an ensemble of classifiers, in which the base classifiers will be CNNs. The study utilized a dataset for a total of 5,980 gait cycles from 322 individuals. The architecture of the base classifiers was consistent, and all possible combinations of GRF components were used to train the base CNNs. The optimal results were obtained when all six GRF components were used for CNN training, achieving a recognition rate of 96.57%. Combining seventeen base classifiers into a homogeneous ensemble further improved the performance, yielding a 99.57% correct recognition rate. This demonstrates the effectiveness of ensemble learning with identical CNN architectures in enhancing gait-based biometric recognition.

Key words: Human gait recognition, Biometrics, Ground Reaction Forces, Convolutional Neural Networks, Ensemble Classifiers

1. INTRODUCTION

Biometrics, as a technique for the automatic recognition of people based on their unique features, has gained significance in recent years in the context of security and access control. The most frequently utilized biometric methods include those utilizing fingerprint [1], face [2], iris [3], voice [4], or hand vein analysis [1,5]. Another promising, although still relatively uncommonly used, biometric method is the identification of a person by the way they walk [6]. Technologies gauging gait encompasses video analysis methods [7] as well as those making use of such devices as accelerometers [8] or force plates [9]. Measurements performed with the employment of force plates register loads, so-called Ground Reaction Forces (GRFs), exerted by the foot on the surface during the stance phase. These forces are an expression of the dynamic attributes of gait biomechanics and generate a unique biomechanical signature based on a person's body mass, anatomic structure, movement patterns, and motor skills. In contrast to solutions based on computer vision, GRF analysis is not affected by changes in light levels, clothing, or angle of observation, with measurements done in real time without the necessity for silhouette recognition. It does, however, require that the test subject cleanly step onto the force plate. The GRF signature is exclusive to any given person, which makes it a valuable source of information for identification systems [10, 11].

Traditional approaches to gait recognition are based on the manual extraction of characteristics from obtained GRF signals and the use of classical machine learning algorithms. Through the

utilization of continuous wavelet transform and the SVM classifier, the authors of [12] were able to attain a high effectiveness of recognition even with varying walking speeds or additional body loading. Based on GRF signals, Michałowska determined characteristics, separately for the left and the right leg, that were time-dependent (such as time of gait cycle) and force-dependent (including maxima of loading response phase for vertical component of GRF) [13]. In the work of [14], in turn, after the division of the GRF signal into individual components corresponding to human gait phases, utilization of the Dynamic Time Warping (DTW) algorithm and the k-nearest neighbours (kNN) classifier yielded over 97% correct recognition for a sample of 200 people.

Recently, increasing popularity has been achieved by deep learning methods, including ones employing convolutional neural networks (CNNs) that, during learning, can automatically identify features from data expressed as a time series, eliminating the necessity of manual selection of traits that contribute to the greatest extent to the differentiation between classes. Moreover, the application of CNNs often leads to superior classification performance compared with classical algorithms. One example of the use of CNNs is the work of [15], where a simple one-dimensional convolutional network (1D-CNN) was proposed to classify GRF patterns to distinguish between healthy and impaired human gaits. The work of [16] also introduces a 1D-CNN, GaitRec-Net, which on a sample containing data concerning over 2,000 people had a task to automatically differentiate between patients exhibiting impaired gait patterns (such as people with hip, knee or foot injuries) and healthy individuals, was able to achieve 91.62% correct

classifications besting such classical machine learning methods as support vector machine (SVM), kNN and Naive Bayes.

It is worth noting that deep learning networks do not always produce the best human recognition results. In the work of [10], it has been shown that, using GRFs data, the SVM classifier achieved 99.3% effectiveness in the identification of 671 people, while CNN reached 95.8%. In the paper of [17], in turn, it has been demonstrated that traditional algorithms like Scale-Invariant Feature Transform (SIFT) may attain better results than CNNs in situations where data is limited or in scenarios in which the network has not been trained using test classifications such as Open-World mode. It is necessary to point out, however, that such studies are rather the exception, and usually CNNs allow the realization of better results.

Literature concerning machine learning has repeatedly shown that ensemble learning allows the achievement of better classification results than the utilization of a single classifier [18, 19, 20]. Studies in gait biometrics, including the author's previous work, similarly indicate that assembling even simple classifiers into ensembles often yields substantial improvements in accuracy and robustness. Many previous studies employing ensemble of classifiers combine models of different types, but such a strategy increases design complexity and computational cost [21]. In contrast, homogeneous ensembles that generate diversity at the data level offer a simpler and more practical solution that preserves implementation uniformity while still benefiting from ensemble effects. The present work aims to address gaps in literature related to the subject of human gait recognition. Its main objective is to present of a method for recognizing individuals based on their manner of walking, using Ground Reaction Forces (GRFs) and a homogenous ensemble of base classifiers where each base classifier is a convolutional neural network.

The main contributions of this work are specified below:

- Empirical demonstration of the effectiveness of a homogeneous ensemble classifier composed of convolutional neural networks (CNNs) for human recognition based on ground reaction forces (GRFs).
- Comprehensive testing and comparison of models trained on all relevant combinations of the six GRF components, leading to the identification of configurations that yield the highest recognition performance.
- Evaluation of the impact of both the number and recognition accuracy of the base classifiers on the overall performance of the ensemble, providing insights into the optimal ensemble structure.
- Validation of the proposed human recognition algorithm on a large dataset collected by the author, which represents one of the most extensive GRF-based databases described in the literature.

2. MATERIAL AND METHODS

Data: The present work utilized a set of data presented in [22]. It contains GRF components for both feet for 5,980 gait cycles gathered from 322 people, including 139 women and 183 men. The measurements were made at the Institute of Biomedical Engineering of the Bialystok University of Technology. During the testing, the participants were asked to walk at their own pace through a testing path concealing two 60 cm x 40 cm Kistler force plates registering data with a frequency of 960 Hz. Movement was initiated at a signal

from the person conducting the measurement. If the walker did not cleanly step on either platform, the transition was not recorded, and the starting point was slightly adjusted. Each person traversed the testing path several times wearing their own sports shoes. To avoid fatigue, a one to two-minute rest was observed after every ten trials. GRFs obtained by individual force plates included three components: medial/lateral, anterior/posterior, and vertical (Figure 1).

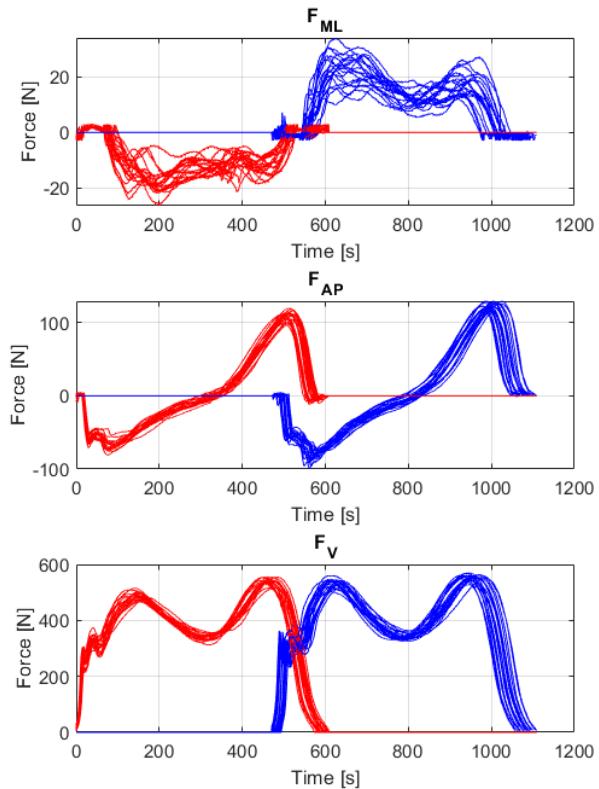


Fig. 1. Components of GRF in: medial/lateral— F_{ML} ; vertical— F_V ; anterior/posterior— F_{AP} direction of the left lower limb (blue line) and of the right one (red line) in sport shoes. The graph shows a dozen steps of a woman aged 21 years with a weight of 48.8kg and height of 164.6cm

Registered GRFs were presented as time series x_1, x_2, \dots, x_N , where N is the number of samples. Generally, the duration time of the support phase of a person's gait depends on several factors and varies so N is variable. To facilitate the comparison of two differing gait cycles, the number of the longest gait cycle samples was established, with the remaining, shorter cycles filled in with 0. Thanks to that, a data set with an even number of samples with $N=1,643$ was attained. Subsequently, these data vectors were used in the study without normalizing the obtained GRFs.

2.1. Base classifiers

The principal part of any biometric system is the module that assigns the considered biometric signature to a particular person represented within the database. This designation is realized through the utilization of classifiers. As mentioned above, the present work employed an ensemble of classifiers that used CNNs as base classifiers. Each CNN possessed the same general architecture, presented in Figure 2 and Table 1. A certain

difference between the utilized CNNs was constituted in the number of channels (e.g. time series) representing components of GRFs fed into the CNNs' input. The character of data used, of course, caused the number of channels to fall within the range of 1 ÷ 6. The number of classes (Fig. 2 – person ID) corresponds to the number of individuals included in the dataset, that is, 322.

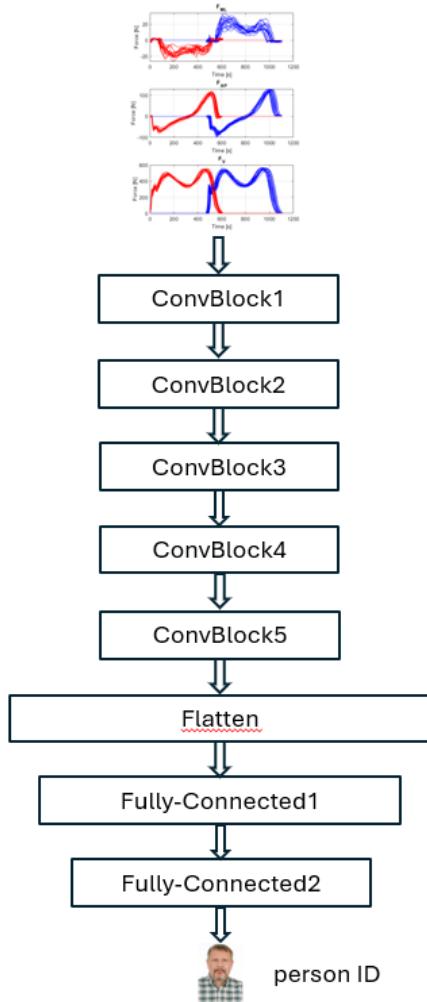


Fig. 2. Architecture and signal processing of CNN base classifier

Tab. 1. The summary of architecture of convolution neural network

No of layer	No. of Conv Block	Type of layer	Kernel size	No of kernels	Output size
1	-	Input	-	-	1643 x channels
2	1	Conv1D	5	64	1639 x 64
4		Max Pooling	2	-	819 x 64
5	2	Conv1D	3	128	817x128
7		Max Pooling	2	-	408x128
8	3	Conv1D	3	256	406x256
10		Max Pooling	2	-	203x256
11	4	Conv1D	3	512	201x512
13		Max Pooling	2	-	100x512
14	5	Conv1D	3	1024	98x1024
15		Max Pooling	2	-	49x1024
16	-	Flatten	-	-	50 176

17	-	Fully-Connected1	-	1000 neurons	1000
18	-	Fully-Connected2	-	700 neurons	700
19	-	Output	-	322 neurons	322

CNNs are well known for their effectiveness in applications connected to the classification of images [23]. This work employs architecture of a CNN to identify characteristics within time series describing GRFs. A CNN consists of several layers, and each one of those has a strictly defined task.

A one-dimensional convolutional layer (Conv1D) detects local patterns along a time series. The convolutional operation is defined as:

$$(f * g)[n] = \sum_{k=-\infty}^{k=\infty} f[n] \cdot g[n - k] \quad (1)$$

where: f represents the input time series and g is the convolutional filter or kernel.

Every filter moves along the input sequence and „learns” to identify characteristic features such as edges or trend changes. The number of filters determines how many different patterns a layer can recognize simultaneously. The kernel size resolves the range of an individual filter. In subsequent layers, the number of filters grows, allowing the model to recognize more complex and abstract features.

An activation function is applied after each convolution layer, most often (including the present work), it is a Rectified Linear Unit (ReLU) defined as:

$$\text{ReLU}(x) = \max(0, x) \quad (2)$$

where x is an activation of neuron.

One-dimensional MaxPooling reduces sequence length by selecting the greatest value within a window (in this paper, it is every two elements). It is a form of downsampling that reduces the size of data and the number of calculations in subsequent layers. At the same time, it enhances the most relevant signal traits because it retains the strongest activations. Pooling also adds a slight translational invariance, so minor signal shifts do not change the result significantly.

Additionally, after every convolutional layer in CNN, Batch Normalization (BN) is used. This equalizes output values from previous layers so that their average is zero while their variance is one. This stabilizes the learning process and allows for the use of larger learning coefficients. Thanks to that, the gradients are better distributed, preventing the vanishing or explosion of gradients in deep learning neural networks. During learning, the values of two other parameters, scale and shift, are added, allowing the restoration of the right range of values if it is beneficial. In practice, BN often accelerates learning and improves precision.

The flatten layer modifies multidimensional data ([sequence, filters]) to a one-dimensional vector. Thanks to this, convolutional layers can be joined with fully-connected (Dense) layers. In a fully-connected layer, every neuron is connected to every neuron of the previous layer. It allows the model to learn the global relationships between all input characteristics. Dense layers are often utilized to connect and interpret complex representations extracted by previous convolutional layers. The activation function (ReLU)

determines how the neurons react to a signal.

The last Dense layer with the softmax activation function creates a probability distribution over classes, and the number of neurons in this layer is equal to the number of recognized people.

The model was compiled using the well-known Adam optimizer and CNNs were trained using categorical cross-entropy as the loss function [24]. Accuracy was used as the primary evaluation metrics. Learning was mostly done with parameters with standard values, with the number of epochs set to 50 and the batch size set to 64. Additionally, during learning, in order to minimize the possibility of overfitting, a dropout equal to 0.1 has been utilized.

2.2. Ensemble Decision Aggregation

Within the presented solution the weighted vote with weight based on rank order technique for combining classifier decisions was utilized. The author is aware of the existence of several other methods for the combining of base classifier decisions; however, a choice to use this relatively simple method which most likely will result in underestimated results of classification has been made.

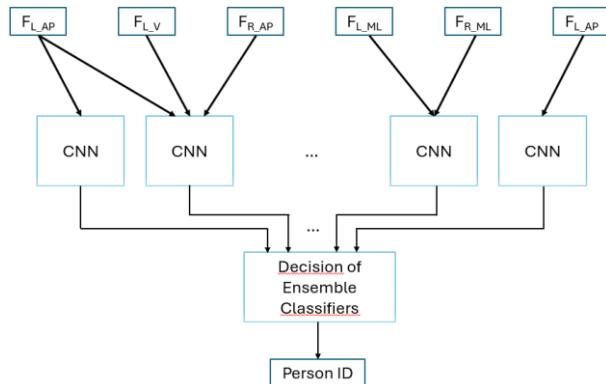


Fig. 3. Flowchart of the proposed method

In this case the weighted value connected to every label depends on rank R , which has been determined based on the accuracy of all base classifiers. The final decision was the class label with the largest total of weights:

$$\text{Person_ID} = \arg \max \left(\sum_{j=1}^k w_j \cdot d_{j,i} \right) \quad (3)$$

where: Person_ID - class label; k - the number of base classifiers, $d_{j,i}$ - decision (class) of the j -th classifier, $d_{j,i} \in \{0, 1\}$, if j -th classifier chooses class i then $d_{j,i} = 1$ otherwise $d_{j,i} = 0$, $w_j = [w_1, \dots, W_R, \dots, w_k]$ - weights, which are calculated from the following formula:

$$w_R = \frac{k+1-R}{k} \quad (4)$$

where: R - indicates the rank for j -th classifier, $R = \{1, 2, \dots, k\}$. In the event of a draw, the class indicated by a greater number of base classifiers was chosen. A schematic of the entire process has been shown in Figure 3.

Every base classifier was trained using 10-fold cross-validation. Each time the same division of data into folds was utilized thanks to which results obtained by different classifiers were comparable. The quality of every classifier was determined by its accuracy. This represents the proportion of true positive results (both true positive as well as true negative) in the selected population:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \cdot 100\% \quad (5)$$

where TP, TN, FP and FN denote: true positive, true negative, false positive and false negative.

3. RESULTS AND DISCUSSION

Accuracy of selected base classifiers has been presented in Table 2. The results indicate that the accuracy of human recognition for CNNs working based on 1 element of GRF oscillates between 85% to 94.3%. It can be noticed that signals recorded for the left leg (the first force plate) have greater accuracy. This difference may be a consequence of a slight variance in the type of the employed Kistler's force plates, with the first force plate being a model 9286AA and the second model 9286AA-A. Disappointing are the recognition results of the base classifier ID6, which worked using the vertical GRF component of the right leg, since it allowed the achievement of only 88.09% of correct identification. This result for the left leg is consistent with information presented earlier in the literature, where it has been ascertained that the vertical component exhibits the greatest potential in differentiating between individual people [10, 11]. The use of both force plates, analogous to the configuration applied for the left lower limb, should therefore provide higher human recognition accuracy than that presented in this study.

Tab. 2. Mean accuracy of person identification depending on the number of channels and types of signals used

ID of base classifier	Components of GRFs used for learning	Accuracy \pm SD[%]
1	F_{L_ML}	$89,4314 \pm 2.4037$
2	F_{L_AP}	$91,8562 \pm 2.8214$
3	F_{L_V}	$94,3144 \pm 1.5766$
4	F_{R_ML}	$85,0836 \pm 2.7157$
5	F_{R_AP}	$90,0502 \pm 1.9168$
6	F_{R_V}	$88,0936 \pm 3.2901$
7	F_{L_ML}, F_{R_ML}	$93,1271 \pm 1.3641$
8	F_{L_AP}, F_{R_AP}	$91,3712 \pm 2.8435$
9	F_{L_V}, F_{R_V}	$94,6488 \pm 1.3025$
10	F_{L_AP}, F_{L_V}	$95,4849 \pm 2.0797$
11	F_{R_AP}, F_{R_V}	$92,6756 \pm 2.8579$
12	$F_{L_ML}, F_{L_AP}, F_{L_V}$	$96,2876 \pm 1.2359$
13	$F_{R_ML}, F_{R_AP}, F_{R_V}$	$94,2642 \pm 1.1316$
14	$F_{L_ML}, F_{L_AP}, F_{R_ML}, F_{R_AP}$	$95,5686 \pm 1.2014$
15	$F_{L_AP}, F_{L_V}, F_{R_AP}, F_{R_V}$	$96,2709 \pm 1.4558$
16	$F_{L_ML}, F_{L_AP}, F_{R_AP}, F_{R_V}$	$96,3378 \pm 1.2819$
17	All	$96,5719 \pm 1.1403$

Base classifiers that learned through the use of data containing time series describing two components of GRFs (ID of base classifiers 7-11) most often attained higher recognition rates than CNNs working on one channel. The sole exception is the ID8 classifier, which, despite the fact that it worked utilizing the same signals as

classifiers ID2 and ID5, achieved only an accuracy of 91.3712% while the results of classifier ID2 reached a level of 91.8562% of correct recognitions. These kinds of exceptions are not seen in CNNs, which have a greater number of channels (from 3 to 6). It is also clear that as the number of channels grows, so does the accuracy of recognition, reaching 96.57% correct identification for a convolutional network employing all GRF elements. This signifies that when it comes to recognizing a person, full information of a phenomenon provides greater possibilities for differentiating between particular people.

Similar results have been presented in [10], where a linear classifier SVM achieved greater precision for signals representing GRF elements of both legs than for a single lower extremity. Additionally, Horst et. al presented classification values with the use of all GRF elements for both legs, which is the same as that for classifier ID17 from Table 2. The CNN classifier achieved 95.8% of correct classifications while the SVM classifier attained a precision on the level of 99.3%, greatly exceeding the results of that study.

Comparing conclusions from Table 2 with the results of the study [19], where, among others, seven varying ensembles of characteristics were analyzed, it can be seen that only one set of parameters allowed a base classifier to reach a better result (99.46%). As mentioned before, both studies work with the same data set.

Tab. 3. Mean accuracy of person identification depending on the base classifiers used

ID	ID of base classifiers	Accuracy \pm SD [%]
EC_1	1+2+3+4+5+6	98.7793 \pm 0.4321
EC_2	1+2+3+4+5+6+17	99.2140 \pm 0.2736
EC_3	7+8+9	97.8930 \pm 0.8162
EC_4	7+8+9+17	98.7625 \pm 0.6217
EC_5	1+2+3+4+5+6+7+8+9	99.1973 \pm 0.3326
EC_6	1+2+3+4+5+6+7+8+9+17	99.4147 \pm 0.2398
EC_7	7+8+9+10+11	98.8127 \pm 0.5196
EC_8	7+8+9+10+11+17	99.1639 \pm 0.3053
EC_9	12+13+17	98.4114 \pm 0.6884
EC_10	1+2+3+4+5+6+7+8+9+10+11	99.3645 \pm 0.2708
EC_11	1+2+3+4+5+6+7+8+9+10+11+17	99.4482 \pm 0.2848
EC_12	1+2+3+4+5+6+7+8+9+10+11+12+13	99.3478 \pm 0.2423
EC_13	1+2+3+4+5+6+7+8+9+10+11+12+13+17	99.4314 \pm 0.2518
EC_14	14+15+16	98.1940 \pm 0.4303
EC_15	14+15+16+17	98.7625 \pm 0.3629
EC_16	1+2+3+4+5+6+7+8+9+10+11+12+13+14+15+16	99.5317 \pm 0.2238
EC_17	All base classifiers	99.5652 \pm 0.2257
EC_18	2+3+5+7+8+9+10+11+12+13+14+15+16+17	99.4816 \pm 0.2548
EC_19	10+12+14+15+16+17	99.0635 \pm 0.3797

Table 3 presents the results of the work of the ensemble classifiers. In this event, the main premise was that this set of classifiers would consist of a minimum of 3 base classifiers, and the maximum number of classifiers would be 17. Additionally, it was assumed that one of the combinations tested would contain base classifiers whose accuracy reached over 90% (EC_18) or over 95% (EC_19). The analysis of results presented in Table 3 shows that the least accurate ensemble classifier correctly recognizes a much greater

number of gait cycles than the best base classifier (97.893% vs 96.5719%).

Figure 4 shows the accuracy of the ensemble classifier depending on the base classifiers. When testing more than one combination of base classifiers whose the same number (e.g., EC_3 and EC_9), the average value is marked. This graph shows that accuracy increases with the number of base classifiers used. According to the conclusions of other authors [25], this rise is at first relatively large, but the addition of another base classifier only slightly improves the quality of recognition. In situations where a base classifier of a lower quality than those used thus far, it may even lead to a slight decrease in accuracy. This type of occurrence can be seen in the case of a classifier consisting of 10 base classifiers (EC_6) where the removal of the base classifier ID17 from the set and the addition of classifiers ID10 and ID11 instead(EC_10), despite the rise in the number of base classifiers, resulted in the reduction in classification quality from 99.4147% to 99.3645%.

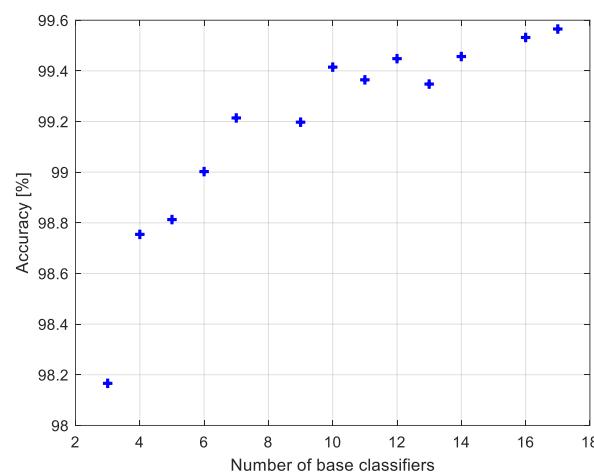


Fig. 4. The average accuracy of the ensemble classifier depends on the number of base classifiers

It is also worth drawing attention to the fact that the quality of a set of classifiers depends on the quality of base classifiers. Thus, the utilization of the best base classifier (ID17) always improves the accuracy of classification, e.g. the accuracy of EC_1 is equal to 98.7793% whereas the accuracy of EC_2 is 99.214%. The combination of these two factors is significant since the use of only the best base classifiers (EC_18 and EC_19) results in a lower quality of a classifier ensemble than EC_17 with their smaller number.

The best classification result was achieved for an ensemble of classifiers that employed all base classifiers (EC_17). The attained accuracy (99.5652%) is one of the highest that has so far been presented in literature. A better result has only been produced in a study [19] where recognition of people on the basis of GRF signals generated during walking was correct with respect to 99.65% of strides. It must be highlighted that this greater accuracy [19] was reached through the use of an ensemble of heterogeneous classifiers, while the result utilizing an ensemble of homogeneous classifiers was a bit lower than the one from the present study (99.55%).

4. CONCLUSION

The paper presents the working of a biometric system for the recognition of a person on the basis of GRFs recorded during walking. The realization of this task was achieved through the use of an

ensemble of homogeneous base classifiers, being convolutional neural networks. Generally, the attained results for the recognition of people are very good and confirm the considerable possibilities for the application of gait as a biometric. The analysis of the outcomes confirmed that the quality of the ensemble of classifiers improves along with increase in the number of base classifiers, as well as with greater accuracy of recognition of individual base classifiers. The utilization, in turn, of the optimal set of features allows for the achievement of better classification results than with respect to the employment of CNNs, where the selection of significant attributes occurs during learning.

Further work in this area can be carried out in three directions. First, other algorithms based on deep learning for the constructing of base classifiers should be tested, and their impact on the result should be evaluated. Second, the robustness of the ensemble of classifiers to changes in the patterns of movement of people subjected to the procedure of recognition, caused, for example, by varying types of footwear or asymmetrical loading, should be scrutinized. Third, the resilience of such a system to deliberate attempts at impersonation should be investigated.

REFERENCES

- Ren H, Sun L, Guo J, Han C. A dataset and benchmark for multi-modal biometric recognition based on fingerprint and finger vein. *IEEE Trans Inf Forensics Secur.* 2022; 17: 2030-2043. <https://doi.org/10.1109/TIFS.2022.3175599>
- Moriuchi E. An empirical study of consumers' intention to use biometric facial recognition as a payment method. *Psychol Mark.* 2021; 38(10): 1741-1765. <https://doi.org/10.1002/mar.21495>
- Yin Y, He S, Zhang R, Chang H, Zhang J. Deep learning for iris recognition: a review. *Neural Comput Appl.* 2025; 37:11125-11173. <https://doi.org/10.1007/s00521-025-11109-5>
- Shakil S, Arora D, Zaidi T. Feature based classification of voice based biometric data through Machine learning algorithm. *Mater Today Proc.* 2022; 51, 240-247. <https://doi.org/10.1016/j.matpr.2021.05.261>
- Jia W, Xia W, Zhang B, Zhao Y, Fei L, Kang W, Huang D, Guo G. A survey on dorsal hand vein biometrics. *Patt Recog.* 2021; 120, 108122. <https://doi.org/10.1016/j.patcog.2021.108122>
- Shen C, Yu S, Wang J, Huang GQ, Wang L. A Comprehensive Survey on Deep Gait Recognition: Algorithms, Datasets, and Challenges. *IEEE Trans Biom Behav Identity Sci.* 2024; 7(2), 270-292 <https://doi.org/10.1109/TBIOM.2024.3486345>
- Asif M, Tiwana MI, Khan US, Ahmad MW, Qureshi WS, Iqbal J. Human gait recognition subject to different covariate factors in a multi-view environment. *Results Eng.* 2022; 15: 100556. <https://doi.org/10.1016/j.rineng.2022.100556>
- Webber M, Rojas RF. Human activity recognition with accelerometer and gyroscope: A data fusion approach. *IEEE Sens J.* 2021; 21(15): 16979-16989. <https://doi.org/10.1109/JSEN.2021.3079883>
- Terrier P. Gait recognition via deep learning of the center-of-pressure trajectory. *Appl Sci.* 2020; 10(3): 774. <https://doi.org/10.3390/app10030774>
- Horst F, Slijepcevic D, Simak M, Horsak B, Schöllhorn WI, Zepelzauer M. Modeling biological individuality using machine learning: A study on human gait. *Comput Struct Biotechnol J.* 2023; 21: 3414-3423. <https://doi.org/10.1016/j.csbj.2023.06.009>
- Duncanson KA, Thwaites S, Booth D, Hanly G, Robertson WS, Abbasnejad E, Thewlis D. Deep metric learning for scalable gait-based person re-identification using force platform data. *Sensors.* 2023; 23(7): 3392. <https://doi.org/10.3390/s23073392>
- Moustakidis SP, Theocharis JB, Giakas G. Subject recognition based on ground reaction force measurements of gait signals. *IEEE Trans Syst Man Cybern B Cybern.* 2008; 38(6): 1476-1485. <https://doi.org/10.1109/TSMCB.2008.927722>
- Michałowska M, Walczak T, Grabski JK, Cieślak M. People identification based on dynamic determinants of human gait. *Vibrations in Physical Systems.* 2018; 29: 2018012.
- Derlatka M, Bogdan M. Ensemble kNN classifiers for human gait recognition based on ground reaction forces. *Proc - 8th Conf Human Syst Interact 2015.* 2015: 88-93. <https://doi.org/10.1109/HSI.2015.7170648>
- Shafew A, Kim D, Kim D. Simple 1D CNN Model for Accurate Classification of Gait Patterns Using GRF Data. *J Infor Commun Converg Eng.* 2024; 22(4), 303-309. <https://doi.org/10.56977/jicce.2024.22.4.303>
- Pandey C, Roy DS, Poonia RC, Altameem A, Nayak SR, Verma A, Saudagar AKJ. GaitRec-Net: A Deep Neural Network for Gait Disorder Detection Using Ground Reaction Force. *PPAR Res.* 2022; 1: 9355015. <https://doi.org/10.1155/2022/9355015>
- Hernandez-Diaz K, Alonso-Fernandez F, Bigun J. One-shot learning for periocular recognition: Exploring the effect of domain adaptation and data bias on deep representations. *IEEE Access.* 2023; 11: 100396-100413. <https://doi.org/10.1109/ACCESS.2023.3315234>
- Zhou X, He J, Yang C. An ensemble learning method based on deep neural network and group decision making. *Knowl Based Syst* 2022; 239:107801. <https://doi.org/10.1016/j.knosys.2021.107801>
- Derlatka M, Borowska M. Ensemble of heterogeneous base classifiers for human gait recognition. *Sensors.* 2023; 23(1): 508. <https://doi.org/10.3390/s23010508>
- Mohammed A, Kora R. A comprehensive review on ensemble deep learning: Opportunities and challenges. *J King Saud Univ Comput Inf Sci.* 2023; 35(2): 757-774. <https://doi.org/10.1016/j.jksuci.2023.01.014>
- Mogan JN, Lee CP, Lim KM, Ali M, Alqahtani A. Gait-CNN-ViT: Multi-model gait recognition with convolutional neural networks and vision transformer. *Sensors.* 2023; 23(8): 3809. <https://doi.org/10.3390/s23083809>
- Derlatka M, Parfeniuk M. Real-world measurements of ground reaction forces of normal gait of young adults wearing various footwear. *Sci Data.* 2023; 10(1): 60. <https://doi.org/10.1038/s41597-023-01964-z>
- Saleh AM, Hamoud T. Analysis and best parameters selection for person recognition based on gait model using CNN algorithm and image augmentation. *J Big Data.* 2021; 8(1): 1. <https://doi.org/10.1186/s40537-020-00387-6>
- Kingma DP. Adam: A method for stochastic optimization. *arXiv preprint.* 2014; arXiv:1412.6980. <https://doi.org/10.48550/arXiv.1412.6980>
- Nakach FZ, Idris A, Zerouaoui H. Deep Hybrid Bagging Ensembles for Classifying Histopathological Breast Cancer Images. *Proc - 15th Int Conf Agents Artif Intell – ICAART 2023.* 2023;2:289-300. <https://doi.org/10.5220/0011704200003393>

The work was supported by the Bialystok University of Technology under Grant W/WM-IIB/2/2024

Marcin Derlatka:  <https://orcid.org/0000-0002-6736-5106>



This work is licensed under the Creative Commons BY-NC-ND 4.0 license.