A Systematic Analysis of External Factors Affecting Gait Identification

Alia Saad
alia.saad@uni-due.de

Nick Wittig
nick.wittig@uni-due.de

Uwe Gruenefeld
uwe.gruenefeld@uni-due.de

Stefan Schneegass
stefan.schneegass@uni-due.de
University of Duisburg-Essen
Essen, Germany

Abstract

Inertial sensors integrated into smartphones provide a unique opportunity for implicitly identifying users through their gait. However, researchers identified different external factors influencing the user's gait and consequently impact gait-based user identification algorithms. While these previous studies provide important insights, a holistic comparison of external factors influencing identification algorithms is still missing. In this explorative work, we conducted a focus group with participants from biometrics research to collect and classify these factors. Next, we recorded the gait of 12 participants walking regularly and being influenced by eleven different external factors (e.g., shoes and floor types) in two separate sessions. We used a Deep Learning (DL) identification algorithm for analysis and validated the analysis results using within- and between- sessions data. We propose a categorization of gait covariates based on users' control levels. Floor types have the most significant impact on recognition accuracy. Finally, between-session analysis shows less accurate yet more robust results than within-session validation and testing.

1. Introduction

Smartphones are ubiquitously used, holding sensitive information that should be kept private. Traditional authentication methods such as Personal Identification Numbers (PIN) and passwords are vulnerable to observation and reconstruction attacks. Physiological biometrics systems, though less vulnerable, are not continuous and unlocked phones are in an imminent threat. Implicit continuous identification techniques can overcome these problems by depending on the ongoing recognition of distinct human behavior. In particular, identifying an individual based on the walking style, also known as gait identification, remains one of the most successful behavior-based implicit identification methods [29]. However, most gait studies are seldom conducted in the wild or consider other factors that might affect the identification robustness. In actual everyday situations, humans face external factors that might alter their typical walking patterns, and accordingly, the continuous implicit identification process [22]. Among these covariates are footwear, walking surfaces, objects carried, and other activities. Several studies showed that wearable sensors, such as accelerometers and gyroscopes, feasible and accurate for collecting gait data. In addition to the wearable sensors' non-contact and non-obtrusive nature, that also leads to higher robustness against spoofing attacks [16]. Since the vast majority of smartphones nowadays are embedded with these inertial sensors, more studies focused on gait identification using smartphones [15, 19].

Although the questions of what external factors can affect the walking pattern and what effects do these external factors impose concerning gait identification techniques were previously introduced [3, 26, 22], these factors were not profoundly categorized and investigated systematically, specifically for inertial sensor-based gait data [6]. It remains unclear how such factors would affect the performance of the gait classifiers, and how to develop countermeasures to mitigate the effects of such factors on the gait classification process. Focusing on deep learning methods, we base our classification on one of the most robust gait deep learning classification methods, reported by Zou et al. [36].

Using this implementation, we investigate different external factors that could affect the accuracy of gait-based identification. To categorize and summarize the external factors that would alter users' walking style, we conducted a focus group to identify different external factors influencing gait. We clustered them based on the level of control users have upon them: uncontrollable such as walking surfaces, semi- controllable, involving social situations involving another person, and controllable, like interacting with the phone, carrying an object, or wearing different shoes.

We collected data from twelve participants to assess the impact of the external possible covariates. The collection process was conducted twice. In our collected data, we considered the baseline to be the person walking alone on a hard floor, wearing their regular shoes, and not doing any activities. To compare with the baseline, we changed only one factor at a time from the baseline condition. The factors' order was counterbalanced using a Latin square design, including the baseline. Interestingly, our findings from validation between sessions showed that walking and talking achieved solid classification accuracy (73%), compared to the baseline (74%), unlike walking silently with another person, which resulted in low classification accuracy of 52.0%. Additionally, surface types are considered to be the factors with the lowest classification results.

The contributions of the paper are presented as follows:

1. Investigating and categorizing external gait affecting factors.
2. Understanding the impact of different factors upon the inertial sensor-based gait recognition.
3. Evaluating the performance of between- and within-session training upon the recognition accuracy.

2. Related Work

Biometric systems are mainly divided into *physiological* features, such as fingerprint and face recognition, and *behavioral*, that include keystroke dynamics [2, 31, 27], signature [5, 10], gaze [13, 24], and gait. Since the main interest of this work is gait identification, we focus on previous works that proposed gait recognition solutions and conducted studies on gait-affecting factors.

2.1. Gait Recognition Systems

In prior surveys, authors summarized the process of recognizing individuals' gaits in five main steps [33, 28]. The process starts with *data acquisition*, where a set of temporal-spatial raw data is collected, and then become subject to *preprocessing*. The features (e.g., gait cycles) are then extracted, a step denoted as *feature representation*. Then, *dimension selection*, where the most significant features are extracted. Lastly, *classification* takes place.

Mainly, there are three gait recognition approaches: floor sensors, machine vision, and inertial sensors [12, 8, 9]. Floor-based recognition studies are restricted by certain locations and conditions, and harder to be conducted in-the-wild. Vision-based setups require additional equipment to be ubiquitous. Moreover, they are vulnerable to occlusions and inconsistent lighting conditions [28]. For these reasons, we focus on the inertial sensors-based gait recognition method. Commonly known as wearable sensors, this approach uses attached accelerometers, gyroscopes and other sensors to retrieve gait data from the person wearing the sensors [29, 17]. The data could be collected from sensors attached to the person's wrist, hip, leg, shoe sole, or embedded in a smartphone that the person uses [35, 20, 7]. Accelerometers and gyroscopes capture movement dynamics in the X, Y and Z directions, which can accurately describe user's movements patterns.

Smartphone-based Gait Recognition Smartphone-based gait recognition solutions are non-obtrusive and do not require a dedicated setup or location. Furthermore, they are showing high recognition accuracy results. Recently, researchers investigated inertial sensors performance changes, by comparing the gait recognition performance with different phone placements, a bag (backpack and hand carry bag), and when held in hand (left and right), all collected with different walking speeds (slow, normal and fast) [22]. Their investigations show that the pocket position produces the best recognition accuracy performance. Zou et al. [36] used CNN to segment gait data into active walking segments, based their identification upon the LSTM and CNN techniques, independently and combined. To overcome the influence of the phone placement in the pocket, they eliminated the phone's orientation factor.

Gait Classification Methods Most of the existing solutions relied on different Machine Learning (ML) algorithms, varying between Support Vector Machines (SVM) [14], Random Forest, K-Nearest Neighbor (KNN) [4]. More recent solutions used Deep Learning methods, such as Long Short-Term Memory (LSTM) to recognize gait patterns collected from Inertial Measurements Units (IMU) [11, 32, 36, 25]. The wearable inertial sensors showed promising results and are more ubiquitous than the other previous approaches. However, keeping one or more sensors attached to the body is not usable and could be inconsistent for extended durations.

2.2. Gait-Affecting Factors

Different researchers focused on understanding different factors affecting gait identification, namely covariate factors [34, 21]. One of the earliest attempts was the gait identification challenge problem [23], in which Philips et al. used the machine vision approach with two cameras to evaluate and compare the effects of floor (grass and concrete), and different shoes (flat and high heels). Focusing on inertial sensors, Subramanian et al. proposed an orientation invariant solution for phone-embedded inertial sensors for gait matching [30]. They considered phone placement (pocket and holster), and activities as talking, texting and only walking, in addition to time, where collection took place twice with several days between the collection sessions, as independent variables. Their results are based on combining factors rather than observing the effect of each factor independently. Results showed that EER is significantly lower with their orientation independent solution, which we considered in our solution. Their findings show that slower walking does not have significance upon gait recognition, compared to faster walking speeds. The phone placement in the pocket has shown the highest overall accuracy results (up to 96.65%). However, they only evaluated the factors combined, making it harder to assess the influence of individual factors.

2.3. Summary

Some factors have more impact upon gait recognition performance than others. While some works already addressed the external factors affecting gait identification, a comprehensive classification and investigation of factors affecting gait identification remains unexplored. In previous studies, data collected in several time intervals was combined and analyzed, which might compromise the classification accuracy. Additionally, traditional machine learning approaches are widely investigated, in comparison to novel deep learning approaches.

3. General Approach

In this work, we investigate the robustness of gait-based identification systems. We choose smartphone inertial sensors-based approaches, and we examine the effects of different external factors on the accuracy of user identification. We define our main research questions as: **RQ1:** How to comprehensively investigate and categorize different external gait covariates? **RQ2:** What are the effects of different factors upon gait classification? **RQ3:** How would the gait recognition performance change based on between- and within-session training? First, we conducted a focus group with experts in the field to collect and classify external factors influencing gait. Then, to understand the impact of different external factors affecting gait, we collected a dataset consisting of the walking data of twelve subjects, walking in twelve variations of the generated external factors. Our dataset was collected in two different sessions, within a time interval of 7-14 days. We selected and reimplemented Zou et al. [36] deep learning solution that uses CNN and LSTM network for gait recognition. We used the collected baseline data for between- and within-session training and compared their impact on the robustness of the recognition accuracy. Additionally, we examined the impacts of the external factors upon the recognition process, also between and within two sessions.

4. Categorization of External Factors on Gait

Excluding medical studies, existing works investigating factors impact on gait recognition mainly focused on surfaces, clothing, including shoes, and carrying objects. However, categorizing these factors and investigating their impact comparatively is unaccomplished. Therefore, we conducted a focus group to formulate a detailed structure for the external gait affecting factors.

4.1. Focus Group

In correspondence to the first question on identifying and structuring different external factors affecting gait, we conducted a focus group. We invited four participants (female=2, male=2), aged between 28 and 33 years (mean=29.75, SD=2.22), all are researchers knowledgeable in biometrics and HCI. After filling the participation consent and demographics forms, the session started by introducing the gait recognition topic, followed asking the participants to mention all the gait affecting factors of their knowledge. We used a collaborative online tool[1], enabling participants to see others' responses. We used the card sorting technique to cluster these factors. Next, we defined the *external* gait affecting factors as causes for gait alterations, that do not intentionally originate from the subjects themselves, create the need for adaptation, and irrelevant to the subjects' physical or mental states (e.g., stress). Otherwise, the factors would be considered as *internal*. Both internal and external factors are categorized in Table 1. In this work, we are primarily investigating the external factors.

The participants continued with listing more aspects that might influence the regular walking styles. Some of these factors were environmental such as weather conditions, surface conditions, or activities such as dog walking, walking with another person, or receiving a notification or a call. In a second iteration, the group proceeded with refining the structure by clustering the factors based on the level of control the affected person has upon these factors, resulting in categorizing them into *uncontrollable*, *semi-Controllable*, and *controllable*. Participants defined factors as uncontrollable when the user does not have any control upon the factor, such as surface conditions, or obstacle avoidance. While interacting with others or dog walking were considered reasons to change the regular walk, but the user might have some control over it. Most of the activities were considered controllable, as the user has full control or choice over these factors. Participants categorized the controllable factors to carrying objects, interaction with their phones, and different clothing, such as trousers and shoes.

[1] Miro. https://miro.com, last accessed 20.04.2022.

Table 1. Overview of the internal and external gait affecting factors evaluated. The internal factors are sorted into their types, whereas the external factors are sorted into their level of control.

Type of Factor	Category	Factor
Mental	Emotion	Stress
		Fear
	Cognition	Task Difficulty
		Goal Planning
Physical	Temporary	Injuries
		Muscle Development
	Permanent	Balance Problems
Combined	State	Medication
		Caffeine, Alcohol
		Exhaustion

Level of Control	Category	Factor
Uncontrollable	Surface Condition	Grass
		Gravel
Semi-controllable	Interacting /w People	Silent Company
		Talking Company
Controllable	Phone Usage	Texting
		Phone Call
	Carrying Objects	Moving Box
		Backpack
		Shopping Bag
	Shoe Types	Flip-Flops
		Winter Boots

4.2. Factors Selection

First, we defined the *baseline* condition in each subcategory. For surfaces, we selected hard floor, a flat surface with no bumps or obstacles. For shoe types, we selected every day shoes, such as sneaker shoes. Since carrying objects and interactions with people or phone are considered to affect the gait recognition, we eliminated all activities in the baseline condition. We wanted to investigate at least two from each category in our study. Here, we based our selection on the most commonly encountered factors, and the ones discussed in previous studies. Therefore, we selected grass and gravel from the uncontrollable category, and silent and talking semi-controllable interactions with people. For the controllable aspects, starting with the phone usage, we decided on texting and talking over the phone. Second, carrying objects, we were interested to investigate the three types of weight loading: front, back and side. Accordingly, we selected carrying a moving box, a backpack, and a shopping bag, respectively. Since the previously selected factors would need to be studied in public areas, changing clothes was difficult to realize. Thus, we limited the clothing category to flip-flops and winter boots as shoe types.

5. Explorative Pilot Study

In our study, we collected gait data in two different sessions, and used deep learning algorithms to test the recognition accuracy performance and the effects of the external factors upon the recognition process.

5.1. Study Design

To evaluate the performance of the framework and the external gait affecting factors, we conducted a within-subject study. Our study was conducted into two different sessions, separated by a time ranging between 7 and 14 days. The split is a crucial procedure in biometric investigations to ensure acquiring realistic data validating the biometric traits of each individual. External factors (11) deduced from the focus group are the independent variables. In addition to the *baseline* consisting of walking alone, in regular shoes on a hard floor, the total number of gait recordings is 12 times per session. The sequence of is counterbalanced using a Latin square design. The dependent variable is the classifier accuracy.

Android Application We developed and used an Android application. The application collected inertial data from the accelerometer and gyroscope of the smartphone, with a frequency of 100Hz. We recorded the tri-axial data from both sensors, along with a timestamp for each point of collection. The collection is triggered by a start button, and locking the device is enabled to avoid accidental display touches during the process. All data is stored in CSV files, anonymized[2].

5.2. Procedure and Apparatus

Before the collection, participants received a list of items to bring on both days of the process. The items were listed as personal flip-flops, winter boots, and their smartphones, in addition to wearing trousers with frontal pockets. Experimenters brought a moving box, a shopping bag, and a backpack, and four weighted plates of a total weight of 5 kilograms. The weights are similar to the ones used by Ming et al. [18].We invited participants to an open area, that consists of 50 meters long straight pathways of different types (hard, grass and gravel). We define the *baseline walk* as a normal walk on a hard pedestrian path, each participant wears their regular every day shoes, and is not performing any tasks. In the first session, participants signed a consent form, and confirmed of not having any internal influence upon their gait, such as an injury or under any substance influence. Upon signing the consent form, participants were given a smartphone with our developed application, and asked to walk steadily and straightforwardly in one of the predefined paths. We used one phone among all participants to control the experiment and reduce the noise. Each walk is recorded when the participant presses that start button in the phone application, locks the phone, and starts walking the 50 me-

[2]Our dataset is publicly avaiable here:https://www.hci.wiwi.uni-due.de/en/research

Table 2. Validation Datasets Summary

Stage	Dataset	Subjects	Accuracy		
			Observed	Reported	Difference
Cycle Extraction	Dataset #8 by Zou et al. [36]	118	84.23%	85.57%	1.34%
Identification Accuracy	Dataset #1 by Zou et al. [36]	118	92.27%	92.51%	0.24%

ters from a defined point. At the end of the path, each subject then stops, turns around and walks towards the starting point, the total distance is therefore 100 meters. There were no obstacles in any of the paths. Before each walk, participants were asked to change shoes, walk in a predefined path, or carry a different object, separately. The baseline and the factors are counterbalanced in a Latin square design. The second session takes place within a duration of 7 to 14 days after the first collection. Each participant repeats the 12 walks, following the same order of the first day. Each session took 30-45 minutes per person. We understand that the phone placement in the pocket is not practical in real-life situations. However, we opted for a single phone placement among all participants as a dependent variable to investigate the impact of the factors upon gait recognition.

5.3. Replication of Gait Identification System

To systematically evaluate the effects of different factors, we reimplemented the Deep Learning (DL) based gait recognition solution by Zou et al. [36]. Here, we considered the implementation shared by the authors on GitHub[3]. The algorithm uses a gait extraction neural network, trained to find the active walking segments of a recorded data sample. The next step is the cycle segmentation, where these walking segments are then split into gait cycles, that are later used in the training of the gait identification network.

Validation with Preexisting Datasets We validate our implementation using two datasets (#1 and #8) from Zou et al. [36], consisting of 118 participants. Thereby, we can verify that we correctly implemented and trained the described algorithms for which source code was not available. The used datasets were collected for 98 subjects in one day and the 20 subjects in 2 different days.[36]. They used dataset #8 for cycle extraction and dataset#1 for testing the identification accuracy. Here, dataset #8 is a manually annotated dataset provided specifically for cycle extraction. As seen in Table 2, our implementation results compared to theirs showed differences of 1.34% and 0.24% for the cycle extraction and identification accuracy stages, respectively.

5.4. Participants

We recruited 12 participants (2 female), aged between 23 and 81 years (M = 33.5, SD = 17.2), with no foot injuries or substance influence. Participants were reimbursed with 20 euros upon the completion of both sessions.

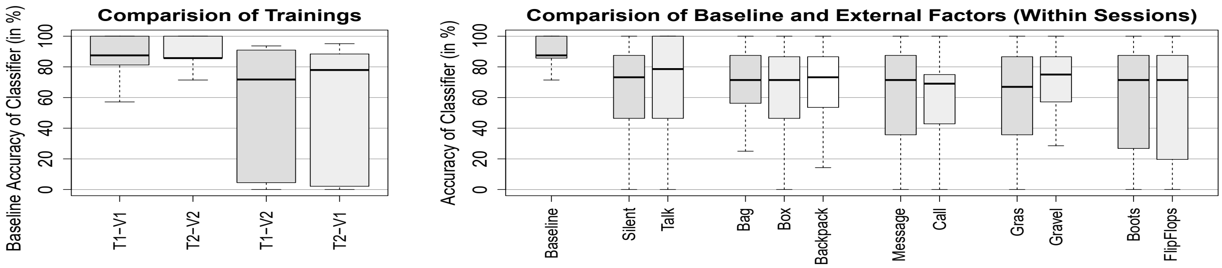
[3] Implementation from Zou et al. on GitHub. https://github.com/qinnzou/Gait-Recognition-Using-Smartphones, last accessed 20.04.2022.

6. Results

We discuss our classification accuracies we observed during training/validation within/between sessions and how the classification accuracy is influenced by external factors. For descriptive statistics, we report mean (M), median (Md), and interquartile-range (IQR). Effect sizes of performed statistic tests are reported with r (r=0.1 small effect, r=0.3 medium effect, and r=0.5 large effect).

Baseline Classification Accuracies First, we report the classifications accuracies that we achieved when training and validating with the baseline of one session. Here, we applied an 80/20 split for training and validating, respectively. For the first session, we reached a median classification accuracy of 87.5% (IQR=15.6%) and for the second, we observed 85.7% (IQR=14.3%). Second, we report the classification accuracies from training with the baseline of one session and validating with the baseline of the other session. For training with the first and validating on the second session, we achieved an accuracy of 71.8% (IQR=84.8%) while vice versa we observed an accuracy of 78.0% (IQR=86.1%). All classification accuracies are reported in Figure 6. Combining the results from training and validating within sessions results in a median accuracy of 87.5%, while combining the results from training and validating between sessions results in a median accuracy of 74.4%. As we do not assume normality and compare two matched groups within subjects, we performed a Wilcoxon Signed-rank test. Here we found a significant difference between training/validating within and between sessions (W=241, Z=2.60, p=0.008, r=0.38). We can conclude that training and validating between session significantly reduces the classification accuracy compared to training/validating within sessions.

Influence of External Factors within Sessions We report the classification accuracies, resulting from our training and validating within sessions. Hence, we consider the effect of *condition* (baseline vs. external factors) on classification accuracy. For each condition, we consider two accuracies (first session, second session) per participant (n=12). The median (interquartile-range) accuracies for each condition are (in descending



(a) Comparison of training approaches. (b) Comparison of factor influence on the algorithm from Zou et al. [36].

Figure 1. We report the classification accuracies observed. In a) we compare the training within and between sessions (T1=training with first session, T2=training with second session, V1=validating with first session, and V2=validating with second session). in b) we compare the influence of the external factors in relation to the baseline within sessions.

order): *baseline*=87.5% (IQR=14.3%), *talk*=78.6% (IQR=51.8%), *gravel*=75.0% (IQR=29.0%), *backpack*=73.2% (IQR=30.8%), *silent*=73.2% (IQR=39.3%), *bag*=71.4% (IQR=28.1%), *box*=71.4% (IQR=37.9%), *message*=71.4% (IQR=48.2%), *boots*=71.4% (IQR=59.8%), *flip-flops*=71.4% (IQR=65.2%), *call*=69.1% (IQR=25.0%), and *grass*=67.0% (IQR=46.9%). The classification accuracies for each condition are plotted in Figure 6. Since we do not assume normality, we performed a Friedman test that revealed a significant effect of condition on classification accuracy (χ^2(11)=27.01, p=0.005, N=12). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences between some conditions (see Table 3). For the classification accuracy within sessions, we can conclude: *baseline* > *silent*, *box*, *message*, *call*, *grass*, *boots*, and *flip-flops*.

Table 3. Pairwise comparisons baseline and significant external factors for training and validating within sessions.

Comparison			W	Z	p	r
baseline	vs.	*silent*	134	2.99	0.019	0.43
baseline	vs.	*box*	139	3.13	0.010	0.45
baseline	vs.	*message*	178	3.32	0.004	0.48
baseline	vs.	*call*	200.5	3.66	<0.001	0.52
baseline	vs.	*grass*	164	3.65	<0.001	0.54
baseline	vs.	*boots*	201.5	3.02	0.002	0.44
baseline	vs.	*flip-flops*	194	3.35	0.004	0.48

Influence of External Factors between Sessions We report the classification accuracies, resulting from our cross-validation between sessions. Here, we consider the effect of *condition* (baseline vs. external factors) on classification accuracy. For each condition, we consider two accuracies (cross-validation of two sessions) per participant (n=12). The median (interquartile-range) accuracies for each condition are (in descending order): *baseline*=74.0% (IQR=86.7%), *talk*=73.0% (IQR=82.0%), *message*=71.8% (IQR=78.2%), *boots*=71.8% (IQR=81.9%), *flip-flops*=71.8% (IQR=78.1%), *bag*=69.5% (IQR=65.1%),

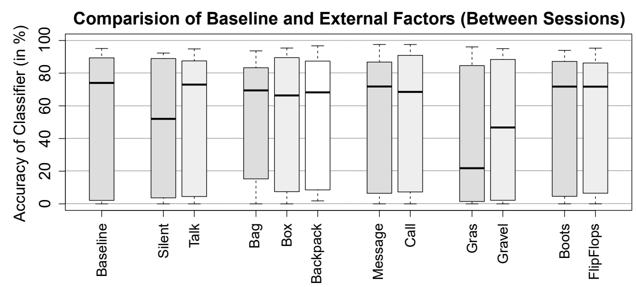


Figure 2. Comparison of factor influence on the Zou et al. [36] algorithm for training and testing between sessions (cross-validation).

call=68.5% (IQR=82.2%), *backpack*=68.3% (IQR=78.2%), *box*=66.4% (IQR=81.0%), *silent*=52.0% (IQR=84.8%), *gravel*=46.8% (IQR=85.7%), and *grass*=21.8% (IQR=82.5%). The classification accuracies for each condition are plotted in Figure 2. Since we do not assume normality, we performed a Friedman test that revealed no significant effect of condition on classification accuracy (χ^2(11)=16.85, p=0.112, N=12).

7. Discussion

Understanding of different Gait Covariates In **RQ1**, we collect and classify the different covariates influencing gait. The covariates generated by the focus group are broad and covering different environmental, and personal activities or attire. In a controlled user study, only specific representatives of each influencing factor can be investigated. Different surface condition, for example, can influence the gait in a specific way. Additionally, the factors might influence each other. Thus, interaction effects between the factors need to be considered. Given the number of factors we found in the focus group, a systematic assessment might not be feasible anymore. One approach to tackle this challenge would be a real world deployment of a gait identification system that records user data in context. By means such as contextual inquiry, the data could be annotated and analyzed. This could lead to a more holistic understanding of the different factors.

Investigating Factors Influence In **RQ2**, we wanted to understand the impact of external factors on the classification performance. Results show that accuracy significantly changes with different factors. We recognize the factors with the lowest classification accuracies are silent company 52.0%, and the other two floor types: gravel and grass, with classification accuracies of 46.8% and 21.8%, respectively. We start with silently walking along another person. Normally, people talk while walking with their company, as reflected in the results where the talking with another person is the external factor with the highest classification accuracy. We believe that silent walks would be strange and awkward, and result in uncomfortable or different walks, as reflected in the results. We observe the different surface types showing the worst classification accuracy results. In our design, hard regular floor is associated to the baseline condition. Except for the surface types, all other conditions are collected with people walking on the hard floor, and it is justifiable to find the floor types result in lower accuracies.

Within- and Between- Sessions We explored how the gait recognition performance would change based on between- and within-session training(**RQ3**). Accordingly, we analyzed the collected data in within- and between-sessions designs. Our results show that the validation accuracy is significantly lower for training and validating between sessions. The average training accuracy has dropped from 86.6% to 74.9%, compared to with-session training. Behavioral biometric features are showing robust results. However, there are many other aspects that contribute to change in behavior, such as fatigue or stress. Many studies used data collected from one rather than multiple sessions. Although they yield high accuracies, such designs entirely neglect influencing factors that occur between days for participants, such as fatigue or stress. We argue that reporting accuracies from between sessions comparisons are more robust when presenting novel identification approaches.

Counter-Strategies As we concluded that external factors affect the performance of gait recognition systems, we suggest several counter-strategies to minimize their effect. Deep learning gait recognition systems require large amounts of data to perform well, increasing the data used for training can improve their performance. We propose training the system continuously with the data recorded during runtime. Considering the in-the-wild use, this counter-strategy improve the system's robustness against external factors. Another possible strategy could be to determine external factors representing entire groups of factors. For example, the data created for the factors *silent* and *talk* might be similar enough that one of these factors is enough to train the system to recognize both. Using this idea, we could drastically reduce the required factors for training the system while potentially achieving high accuracy.

Limitations and Future Work We recognize the following limitations in our work. The participants number, whether in the focus group or the data collection, is limited. We acknowledge that a larger number of participants is preferable to increase statistical power for data analysis. Hence, we consider this study to be more explorative in nature, paving the way for hypothesis-driven follow-up studies. We also know that only employing one deep learning-based algorithm lacks comparisons to other approaches, and accordingly, missing insights to develop a robust gait recognition system. We used one phone among participants, yet we expect that different devices usage would affect the results. Finally, there are many external and internal gait affecting factors that we did not consider in our design. However, the design was limited to these twelve walks per session by time and exhaustion constraints.

We believe these limitations can be surpassed in future studies. Extending the number of participants and the gait affecting factors, should be considered in the training process. We also consider training with various factors as a viable, robust solution to mitigate consequential inaccuracies. We also consider additional approaches, including machine learning methods such as Siamese Networks [1], to reach extended and potent results.

8. Conclusion

Smartphone inertial sensor-based gait recognition methods are gaining attention for their robustness and ubiquitous properties. However, accurate user identification is vulnerable to factors impacting users while walking. In this paper, we investigated and categorized possible factors that impact the gait identification process. We conducted a focus group and collected data from twelve participants with smartphone inertial sensors to investigate these factors. We used the state-of-the-art deep learning-based implementation of smartphone-based gait identification to validate our results. Our main findings show that different factors affect classification robustness, and between-session analysis is crucial for recognition robustness. Accordingly, we believe that considering these findings would lead to more accurate and potent gait recognition solutions in the future.

Acknowledgement

This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 425869382 and is part of Priority Program SPP2199 Scalable Interaction Paradigms for Pervasive Computing Environments.

References

[1] O. Adel, M. Soliman, and W. Gomaa. Inertial gait-based person authentication using siamese networks. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE, 2021. 7

[2] S. P. Banerjee and D. L. Woodard. Biometric authentication and identification using keystroke dynamics: A survey. *Journal of Pattern Recognition Research*, 7(1):116–139, 2012. 2

[3] J. E. Boyd and J. J. Little. Biometric gait recognition. In *Advanced Studies in Biometrics*, pages 19–42. Springer, 2005. 1

[4] S. Choi, I.-H. Youn, R. LeMay, S. Burns, and J.-H. Youn. Biometric gait recognition based on wireless acceleration sensor using k-nearest neighbor classification. In *2014 International Conference on Computing, Networking and Communications (ICNC)*, pages 1091–1095. IEEE, 2014. 2

[5] N. L. Clarke and A. Mekala. The application of signature recognition to transparent handwriting verification for mobile devices. *Information management & computer security*, 2007. 2

[6] P. Connor and A. Ross. Biometric recognition by gait: A survey of modalities and features. *Computer Vision and Image Understanding*, 167:1–27, 2018. 1

[7] M. De Marsico, D. De Pasquale, and A. Mecca. Embedded accelerometer signal normalization for cross-device gait recognition. In *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–5. IEEE, 2016. 2

[8] M. O. Derawi. Accelerometer-based gait analysis, a survey. *Norwegian Information Security Conference*, 1:33–44, 2010. 2

[9] R. Divya and R. Lavanya. A systematic review on gait based authentication system. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pages 505–509. IEEE, 2020. 2

[10] R. Doroz and P. Porwik. Handwritten signature recognition with adaptive selection of behavioral features. In *Computer Information Systems–Analysis and Technologies*, pages 128–136. Springer, 2011. 2

[11] C. Filipi Gonçalves dos Santos, D. d. S. Oliveira, L. A. Passos, R. Gonçalves Pires, D. Felipe Silva Santos, L. Pascotti Valem, T. P. Moreira, M. Cleison S. Santana, M. Roder, J. Paulo Papa, et al. Gait recognition based on deep learning: a survey. *ACM Computing Surveys (CSUR)*, 55(2):1–34, 2022. 2

[12] D. Gafurov. A survey of biometric gait recognition: Approaches, security and challenges. In *Annual Norwegian computer science conference*, pages 19–21. Annual Norwegian Computer Science Conference Norway, 2007. 2

[13] C. Galdi, M. Nappi, D. Riccio, V. Cantoni, and M. Porta. A new gaze analysis based soft-biometric. In *Mexican Conference on Pattern Recognition*, pages 136–144. Springer, 2013. 2

[14] M. R. Hestbek, C. Nickel, and C. Busch. Biometric gait recognition for mobile devices using wavelet transform and support vector machines. In *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 205–210. IEEE, 2012. 2

[15] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Cell phone-based biometric identification. In *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–7. IEEE, 2010. 1

[16] A. Mahfouz, T. M. Mahmoud, and A. S. Eldin. A survey on behavioral biometric authentication on smartphones. *Journal of information security and applications*, 37:28–37, 2017. 1

[17] M. D. Marsico and A. Mecca. A survey on gait recognition via wearable sensors. *ACM Computing Surveys (CSUR)*, 52(4):1–39, 2019. 2

[18] D. Ming, Z. Xue, L. Meng, B. Wan, Y. Hu, and K. D. Luk. Identification of humans using infrared gait recognition. In *2009 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurements Systems*, pages 319–322. IEEE, 2009. 4

[19] M. Muaaz and R. Mayrhofer. Orientation independent cell phone based gait authentication. In *Proceedings of the 12th International Conference on Advances in Mobile Computing and Multimedia*, pages 161–164, 2014. 1

[20] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi. The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication. *Pattern Recognition*, 47(1):228–237, 2014. 2

[21] M. S. Nixon, T. Tan, and R. Chellappa. *Human identification based on gait*, volume 4. Springer Science & Business Media, 2010. 2

[22] J. Permatasari, T. Connie, and T. S. Ong. Inertial sensor fusion for gait recognition with symmetric positive definite gaussian kernels analysis. *Multimedia Tools and Applications*, 79(43):32665–32692, 2020. 1, 2

[23] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer. The gait identification challenge problem: Data sets and baseline algorithm. In *Object recognition supported by user interaction for service robots*, volume 1, pages 385–388. IEEE, 2002. 2

[24] M. Porta and A. Barboni. Strengthening security in industrial settings: A study on gaze-based biometrics through free observation of static images. In *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1273–1277. IEEE, 2019. 2

[25] S. Potluri, S. Ravuri, C. Diedrich, and L. Schega. Deep learning based gait abnormality detection using wearable sensor system. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3613–3619. IEEE, 2019. 2

[26] A. Primo, V. V. Phoha, R. Kumar, and A. Serwadda. Context-aware active authentication using smartphone accelerometer measurements. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 98–105, 2014. 1

[27] B. S. Saini, N. Kaur, and K. S. Bhatia. Keystroke dynamics for mobile phones: A survey. *Indian Journal of Science and Technology*, 9(6):1–8, 2016. 2

[28] J. P. Singh, S. Jain, S. Arora, and U. P. Singh. Vision-based gait recognition: A survey. *IEEE Access*, 6:70497–70527, 2018. 2

[29] S. Sprager and M. B. Juric. Inertial sensor-based gait recognition: A review. *Sensors*, 15(9):22089–22127, 2015. 1, 2

[30] R. Subramanian, S. Sarkar, M. Labrador, K. Contino, C. Eggert, O. Javed, J. Zhu, and H. Cheng. Orientation invariant gait matching algorithm based on the kabsch alignment. In *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2015)*, pages 1–8. IEEE, 2015. 2

[31] P. S. Teh, A. B. J. Teoh, and S. Yue. A survey of keystroke dynamics biometrics. *The Scientific World Journal*, 2013, 2013. 2

[32] L. Tran, T. Hoang, T. Nguyen, H. Kim, and D. Choi. Multi-model long short-term memory network for gait recognition using window-based data segment. *IEEE Access*, 9:23826–23839, 2021. 2

[33] C. Wan, L. Wang, and V. V. Phoha. A survey on gait recognition. *ACM Computing Surveys (CSUR)*, 51(5):1–35, 2018. 2

[34] S. Yu, D. Tan, and T. Tan. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 4, pages 441–444. IEEE, 2006. 2

[35] Y. Zhang, G. Pan, K. Jia, M. Lu, Y. Wang, and Z. Wu. Accelerometer-based gait recognition by sparse representation of signature points with clusters. *IEEE transactions on cybernetics*, 45(9):1864–1875, 2014. 2

[36] Q. Zou, Y. Wang, Q. Wang, Y. Zhao, and Q. Li. Deep learning-based gait recognition using smartphones in the wild. *IEEE Transactions on Information Forensics and Security*, 15:3197–3212, 2020. 1, 2, 3, 5, 6