Does Cycling Reveal Insights About You? Investigation of User and Environmental Characteristics During Cycling

Luca Hernández Acosta(✉), Sebastian Rahe, and Delphine Reinhardt

Georg-August-Universität, Göttingen 37077, Germany
{hernandez,reinhardt}@cs.uni-goettingen.de

Abstract. Smartwatches are increasingly being used as fitness and health trackers. To provide such a service, these devices have to collect and process movement data gathered by built-in accelerometers and gyroscopes. Based on these data, existing studies leveraging smartphones have shown that it is possible to distinguish users when they (1) walk, (2) perform different hand gestures, or (3) pick up their phone from the table. However, to the best of our knowledge, the case of cycling has not been addressed yet. The goal of this paper is to close this gap by investigating whether it is possible to infer information about users wearing a smartwatch coupled with their smartphone when cycling, their bike type, seat height, gear, and the terrain. In addition, we explore whether it is possible to distinguish individual users based on their movement patterns that may lead to their (re)identification. To this end, we conducted a user study with 17 participants, equipped with a smartphone and a smartwatch, who had to ride along a bike road for two km. Among others, our results show that it is possible to infer the four characteristics bike type, gear, seat height, and terrain with accuracies of 93.05%, 92.23%, 95.76%, 94.24% respectively and distinguish participants with a probability of 99.01%.

Keywords: Behavior Analysis · Activity Recognition · Bike Identification · User Recognition

1 Introduction

The use of smart devices, such as smartwatches and smartphones, in our professional and private life is steadily increasing [1–3]. In addition to support easy communication and quick access to information on the Internet, these devices are also equipped with various sensors allowing for a myriad of applications. For example, they are gaining popularity in the areas of fitness and health tracking [4]. In order to provide such services, the applications collect and analyze data of the devices' users. Often, movement data are collected by built-in sensors, such as the accelerometer and gyroscope, for these purposes. These data

S. Longfei and P. Bodhi (Eds.): MobiQuitous 2022, LNICST 492, pp. 172–190, 2023.
https://doi.org/10.1007/978-3-031-34776-4_10

are then used for various tasks, such as counting daily steps, analyzing individual walking behavior and deriving additional health-related information [5], or to detect severe falling incidents [6]. Related work has also shown that the movement data can be used to recognize and distinguish different activities, such as walking, running, or cycling [7–9]. Even non-sports activities such as eating, drinking, or writing on a keyboard can be determined from these data [10,11]. For walking, the data can even be used to distinguish and identify different users [12] by their gait. Kröger et al. give a detailed overview about potential inferences based on accelerometer data and show that besides activity and user recognition also personal information, such as the age or gender could be inferred [13]. However, cycling has not been considered yet. As a result, this raises the following research questions that we address in this paper:

- Which information about the cyclists can be derived from their movement data collected using their smartwatch and smartphone?
- Which information can be inferred about their bike (e.g., bike type, seat height, and gears)?
- Which information can be derived about the terrain?

To answer these questions, our contributions are as follows. First, we have built a data set consisting out of all combinations of the characteristics bike type, seat height, gear, and terrain. This data set is based on the sensor readings of one single person, resulting in a size of 1.1 gigabyte. Second, we conducted a user study with 17 participants (10 male, 7 female) between the ages of 19 and 64. In the study, they were equipped with a smartphone and a smartwatch and instructed to ride a predetermined bike road for two km. While the seat height could be adjusted depending on the participant, the type of bike remained the same for all participants. Inspired by existing studies on human gait recognition, we have processed the collected data and split it according to individual pedal rotations. The prepared data serves as basis for the training and testing of different machine learning models to explore whether characteristics about the participants, bikes, and/or terrain can be predicted and how unique cycling patterns are between participants. To this end, we have considered the following algorithms: Gaussian Naive Bayes (GNB), k-Nearest Neighbours (KNN), linear Support-Vector Classification (SVC), Decision Tree (DT), and Random Forest (RF). The results show that the best performances are obtained for the four characteristics bike type, gear, seat height, and terrain with accuracies of 93.05%, 92.23%, 95.76%, 94.24% respectively and distinguish participants with a probability of 99.01% all by utilizing the RF algorithm.

The remainder of this article is structured as follows. In Sect. 2, we discuss related studies that dealt with both activity and user recognition based on different activities and highlight our contributions. In Sect. 3, we are describing the methodology applied in our user study to collect our data. In Sect. 4, we describe the approach used for the analysis of the collected data, while we present our results in Sect. 5. In Sect. 6, we comment on the limitation of our study and future work, before making concluding remarks in Sect. 7.

2 Related Work

Related work can be split into the following two categories: (1) activity recognition and (2) user distinction.

In the field of activity recognition, both smartphones and smartwatches can be leveraged. Using smartphones and their accelerometers and gyroscopes, certain activities, such as walking, running, or cycling (e.g., [7–9,14–17]), can be recognized. Using smartwatches, additional activities such as eating, drinking, writing on a keyboard, and fitness activities can also be recognized (e.g., [10,11,18–20]).

Apart from activity recognition, movement data can also be used in order to identify users. Zou et al. show that using the collected accelerometer and gyroscope data from a smartphone in the field can be used to identify individual users based on their gait after training and evaluating the data with deep learning techniques [12]. The same applies for data that is collected by smartwatches. In a study performed by Andrew Johnston and Gary Weiss, the authors show that gait-based biometric identification is also possible with smartwatches [21]. Moreover, Häring et al. show that the pick up motion performed when picking up a smartphone from a desk could be used in order to support user authentication to the device, showing that the motion itself is highly depending on individual characteristics [22,23].

In contrast, we investigate to what extent it is possible to determine certain user characteristics based on data collected while cycling. Matkovic et al. show that is already possible to use the collected smartphone movement data in order to infer the respective bike type [24]. In a later study, they further include the detection of an e-scooter along with different bike types [25].

In comparison to the studies performed by Matkovic et al., we do not only explore bike type identification but also other characteristics, such as the gear, seat height, and the terrain, which to the best of our knowledge have not been investigated before. Moreover, we do not only rely on data collected by a smartphone like in other studies, but also consider data from smartwatches and investigate the usefulness of these data to identify the different characteristics. Last but not least, we also investigate whether it is possible to use the collected data to distinguish users based on their individual cycling patterns.

3 Data Collection

In this section, we describe the methodology used in our user study. We outline the devices used for data collection, where the devices are placed with our participants, how they are recruited, and the steps taken to comply with data protection.

3.1 Used Devices

Firstly, we have decided to use smartphone and smartwatch data to collect the movement data. Our decision is motivated by the fact that both devices are able to collect data at different positions on cyclists. While smartphones are often located in trousers' front/back pockets, a smartwatch is usually worn at the wrist. The smartwatch thus allows us to gather additional knowledge about users' movements and increases the probability to successfully infer additional characteristics about them and their context. The data collected by the smartwatch is transferred to the smartphone via a Bluetooth connection, where it is stored. We used a Google Pixel 4 and a Samsung Galaxy Watch (SM-R805F). Both configured to collect the considered sensor data 400 Hz 100 Hz, respectively. Pseudonyms were used to avoid the linking of the collected data to the participants.

3.2 Data Protection and Ethical Aspects

Before starting the study, we have distributed a consent form to the participants in order to inform them about both data collection and processing modalities following the *General Data Protection Regulation* (GDPR). Note that our study was submitted to the Data Protection Officer of our institution. A verification by the ethical board of our institution is however not mandatory in our field. Nevertheless, we have limited the efforts for the participants to the minimum. They have been informed that they could opt out at any time and that their data would be removed. On average, each participant took about 30 min to complete the study.

3.3 Recruitment

The participants were recruited within our social circle. Our recruitment strategy, however, does not impact our results due to absence of subjective questions in our study. The study took place between 14th of September 2021 and 14th of December 2021. Every participant executed the task once.

Table 1. Observed characteristics

Dimensions	Selected alternatives
Bike type	Road bike, Mountain bike, Gravel bike, E-bike
Gear	15th, 21st
Seat height	Low, Normal, High
Terrain	Dirt road, Bike road, Stone road, Asphalt road

3.4 Smartphone and Smartwatch Placement

In order to have the same conditions for all participants, the smartphone was located in the right front pocket of their trousers, upside down with the charging port at the top and the screen facing away from the body. Furthermore, the smartwatch was worn on their right wrist.

3.5 Scenario, Dimensions, and Parameter Variations

To answer our research questions, we aim at exploring whether the collected data can reveal information about five different dimensions: (1) the bike type, (2) seat height, (3) gear, (4) terrain, and (5) the participant. For the first four dimensions, we have selected different alternatives to define our ground truth as summarized in Table 1. For the bike type, this means a road bike, a mountain bike, a gravel bike, and an e-bike as displayed in Fig. 1. For all bikes, the seat height can be adjusted to low, normal, and high. We set the different seat heights by measuring the distance between the centre of the chain set and the top of the seat resulting in 60 cm for low, 85 cm for normal, and 95 cm for high. We further set the gears to the 15th or 21st gear for the different runs. We finally considered different terrains as depicted in Fig. 2: dirt road, bike road, stone road, and asphalt road. For analysis of the fifth dimension, i.e., possible differences between participants, all participants have ridden the same bike on the same road for two km, only the seat height has been adjusted depending on the participant. We have chosen these dimensions because we presume that they have a direct impact on the collected sensor data. Indeed, while bike type and seat height could affect the posture of the cyclist, gear could affect pedaling speed, and different terrains could create vibrations affecting the collected movement data.

(a) Road bike (b) Mountain bike (c) Gravel bike (d) E-bike

Fig. 1. Bike types

4 Data Processing

In this section, we describe our approach to process the data collected according to the settings described in Sect. 3 in order to explore the different dimensions of interest.

(a) Dirt road (b) Bike road (c) Stone road (d) Asphalt road

Fig. 2. Terrains

4.1 Preprocessing

We have first applied the following preprocessing step. Due to the high sampling rate 400 Hz, different readings share the same timestamp. We have hence first distributed all timestamps evenly across all recordings following this function:

Let the set $\{T_1, T_2, T_3, \ldots, T_n\}$ be the unrepaired timestamps. Then

$$T_{new}(x) = T_1 + x\frac{T_n - T_1}{n}$$

for $x = 1, 2, \ldots, n$ and where T_x is the xth recorded timestamp. $x \in \mathbb{N}$

4.2 Identification of Pedal Rotations

Like in gait recognition (see Sect. 2), we aim at first identifying repetitive patterns. While such patterns are defined by two consecutive steps in the case of gait recognition, we consider a single pedal rotation as a representation for the periodic repetitions during pedaling. In order to detect such a pedal rotation, we use the gyroscope data to identify the circular movement of the legs and therefore the identification of one complete pedal rotation.

In a first attempt to detect the bounds symbolizing the start and the end of a single pedal rotation, three options shown in Fig. 3 were possible: searching for (1) local maxima (see Fig. 3a), (2) minima (see Fig. 3b), or (3) turning points around the zero value whenever the value turns from negative to positive (see Fig. 3c). Among these options, we have selected the zero point approach, shown in Fig. 3c, because it turned out to be the most reliable approach to detect the bounds of a single pedal rotation.

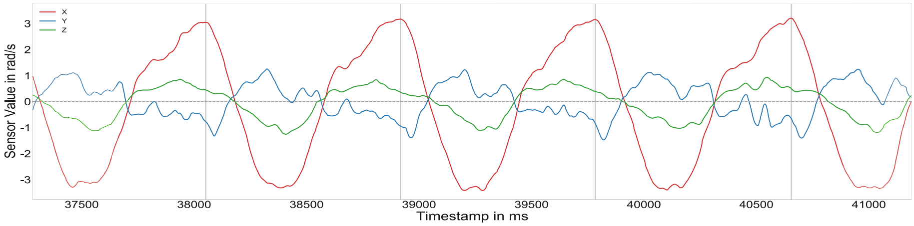
One of the main problems that occurred in the process of identifying a pedal rotation was (1) that a non-pedal rotation is sometimes identified as a real pedal rotation (type I error; false positive) and on the other hand (2) some pedal rotations were not identified at all (type II error; false negative). When examining our participants' data, we notice that for some participants the turning point approach did not work as expected because the X-values oscillated above 0, leading to false positives. To make our approach more reliable in correctly identifying pedal rotations, we have further considered the Z values of the gyroscope data, which have a similar pattern as the X values. By doing so, the identification of false pedal rotations due to the oscillation of the X values around zero can be avoided. The principle behind this idea is that a new bound of a pedal rotation is only detected if both X and Z values are positive, if at least one negative value between X and Z was detected in the previous timestamp. The rotation is considered as complete when at least one timestamp exists where both X and Y values are negative.

Following this adjusted approach, we have eliminated all false positive pedal rotations and obtained fewer false negative detected pedal rotations leading to approximately 250 to 400 detected pedal rotations in rounds of 4 min cycling sessions. Even though we still have a few false negative pedal rotations this scenario is highly preferred as false positive pedal rotations might have a bad impact on the later classification of the characteristics.

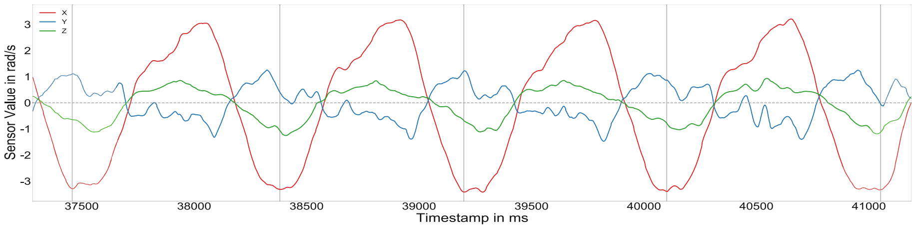
4.3 Feature Extraction

After the successful identification of the individual pedal rotations, we now focus on feature extraction. Since all pedal rotations can differ in both length and number of sensor readings, we need to find an approach that makes all pedal rotations comparable. To solve this issue, we bin the sensor readings in a pedal rotation by separating the readings in fixed intervals. In more details, we use the start and end timestamps of a detected pedal rotation and calculate the respective timestamp for every sensor reading in the pedal rotation that matches our fixed interval following this function, where T_s is the first timestamp of a pedal rotation and T_e is the last timestamp of a pedal rotation:

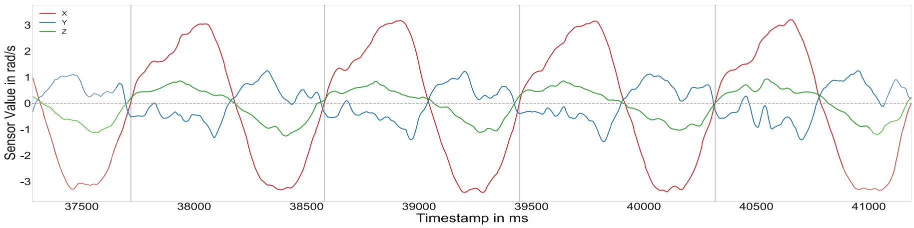
$$T_x = T_s + x * \frac{T_e - T_s}{i - 1}$$



(a) Smartphone gyroscope data - pedal rotation bounds by maxima



(b) Smartphone gyroscope data - pedal rotation bounds by minima



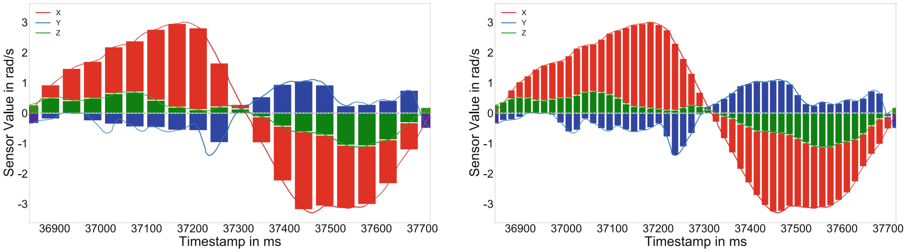
(c) Smartphone gyroscope data - pedal rotation bounds by zero point

Fig. 3. Pedal rotation bounds detection methods

We next determine the sensor value for each calculated timestamp. If a sensor reading exists for this timestamp, we obviously use its value as the new value for our bin. Otherwise, we apply the following linear interpolation with T_a the timestamp before, T_b the timestamp after, V_a as the value at T_a and V_b as the value at T_b:

$$V_x = V_a + (V_b - V_a) * \frac{x - T_a}{T_a - T_b}$$

Figure 4 illustrates how such a binning process looks like for different intervals of 20 bins (see Fig. 4a) and 50 bins (see Fig. 4b). We hence follow this approach in order to normalise the values for each individual pedal rotation with varying lengths and number of readings.



(a) Smartphone gyroscope data - single pedal rotation with bins marked (20)

(b) Smartphone gyroscope data - single pedal rotation with bins marked (50)

Fig. 4. Smartphone gyroscope data - single pedal rotation with bins marked

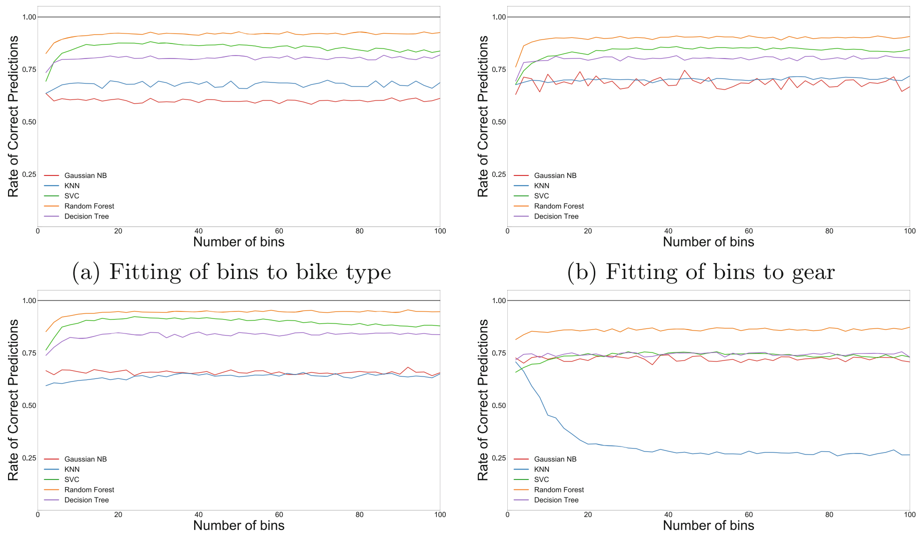
4.4 Classification

For the classification of the pedal rotations identified according to our approach detailed in Sect. 4 according to the specific characteristics given in Table 1, we explore and compare the following five machine learning algorithms: *Gaussian Naive Bayes* (GNB), *k-Nearest Neighbours* (KNN), *linear Support-Vector Classification* (linear SVC), *Decision Tree* (DT) and *Random Forest* (RF).

To avoid under- or overfitting our models, we have tested different configurations of our features. Since the binned sensor values are our most important features, we have investigated how the performance of the machine learning models changes depending on the size of our bins. In Fig. 5, we see that performance for all machine learning models increases as the number of bins increases, up to 25 bins. Just for the classification of the terrain and only for KNN we observe a dramatic decrease in the rate of correct classification, while the number of bins increases. Moreover, we notice that for the linear SVC performance degradation occurs once the number of bins exceeds 40. Based on these results, we have decided to select a number of 40 bins for further analysis.

Except for GNB, all other algorithms have hyper parameters that need to be tuned in order to reach the best performance. We have therefore run grid tests with cross-validation and obtained the following set of hyper parameters:

- KNN: number of neighbours = 6
- linear SVC: regularisation parameter = 1
- RF: maximum depth = 16; number of trees = 20
- DT: maximum depth = 5



(a) Fitting of bins to bike type

(b) Fitting of bins to gear

(c) Fitting of bins to seat height

(d) Fitting of bins to terrain

Fig. 5. Fitting of bins to characteristics

Since the grid search for the optimal hyper parameters is quite computationally expensive, we do not exclude that additional fitting of these hyper parameters can lead to even more accurate models and thus to a classification with even a higher accuracy.

5 Results

To explore the best combination of features, we have run through all possibilities with smartphone and/or smartwatch data available. This means 24 possible combinations of the features minima/maxima, variance, and standard deviation. We discuss the corresponding results in what follows.

5.1 Bike Type

For the detection of the correct bike type, we find that the best classification can be made when considering the RF classifier, closely followed by linear SVC as shown in Fig. 6. A more detailed insight shows that the overall best performance can be achieved with the combination of all features maxima/minima, variance, and standard deviation. In this case, we achieve an accuracy of 93.05%, shown in Fig. 6a. In comparison, with the smartphone data only, the highest accuracy is equal to 91.48%, as depicted in Fig. 6b. Moreover, we observe that the road bike and the e-bike both with a precision of to 94% are the best classified types, while the gravel bike is worst with 90%. The mountain bike and the road bike are the two most likely bike types to be confused as shown in Fig. 7. Reasons for the F1-Score being higher than the accuracy is most probably due to the imbalanced nature of our dataset.

Features				GNB	KNN	SVC	RF	DT
S	M	V	S	66.99%	73.97%	91.47%	93.05%	75.28%
S	M	V	-	64.61%	72.26%	92.53%	92.93%	76.44%
S	M	-	S	62.88%	71.78%	92.26%	92.62%	75.62%
S	M	-	-	65.99%	69.22%	91.25%	92.11%	78.15%
S	-	V	S	66.47%	72.20%	89.64%	92.50%	77.05%
S	-	V	-	64.37%	73.33%	91.28%	92.59%	77.54%
S	-	-	S	66.47%	65.04%	91.77%	92.62%	77.08%
S	-	-	-	65.50%	65.41%	89.61%	90.52%	76.68%

(a) Performance with smartwatch data

Features				GNB	KNN	SVC	RF	DT
S	M	V	S	56.70%	88.14%	83.41%	90.36%	70.86%
S	M	V	-	55.43%	87.36%	84.66%	89.63%	71.68%
S	M	-	S	54.73%	86.48%	85.51%	91.48%	73.65%
S	M	-	-	54.40%	86.14%	78.47%	89.75%	73.23%
S	-	V	S	55.22%	86.66%	83.32%	89.96%	72.41%
S	-	V	-	57.06%	87.11%	78.53%	90.05%	74.41%
S	-	-	S	54.79%	84.32%	87.33%	89.36%	72.20%
S	-	-	-	54.12%	84.87%	75.50%	88.36%	70.86%

(b) Performance without smartwatch data

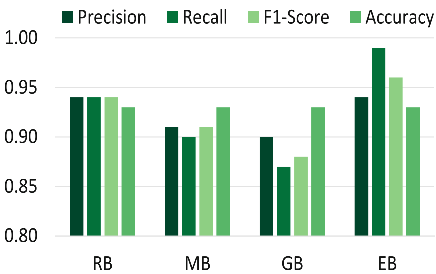
Fig. 6. Performance to infer bike type of all classifiers based on different combinations of features sensor data (S), maxima/minima (M), variance (V), and standard deviation (S)

5.2 Gear

The classifier with the highest accuracy to infer the 15th and 21st gear is the RF classifier, as shown in Fig. 8. The combination of features that results in the best accuracy is represented by both the smartphone and smartwatch data in addition with the maxima/minima and variance giving an accuracy of 92.23%, shown in Fig. 8a. In comparison with the smartphone data only, we only achieve an accuracy of 90.60%, depicted in Fig. 8b. In both cases the other observed classifiers are not able to achieve accuracies that are higher than 90%. Moreover, we observe that the precision and the recall are higher for the 15th gear, meaning that the RF classifier is slightly skewed towards predicting this gear as shown in Fig. 9.

	RB	MB	GB	EB
RB	1058	56	14	0
MB	47	886	38	15
GB	15	28	481	27
EB	0	2	4	610

(a) Bike type confusion matrix



(b) Bike type classification results

Fig. 7. Confusion matrix and classification report for the bike types of road bike (RB), mountain bike (MB), gravel bike (GB), and e-bike (EB)

Features				GNB	KNN	SVC	RF	DT
S	M	V	S	61.41%	78.76%	82.69%	91.65%	83.05%
S	M	V	-	61.44%	78.48%	66.60%	92.23%	82.96%
S	M	-	S	62.33%	72.60%	83.42%	91.59%	83.79%
S	M	-	-	62.91%	73.03%	83.97%	90.46%	81.99%
S	-	V	S	62.42%	78.36%	59.92%	91.44%	82.05%
S	-	V	-	63.88%	78.67%	80.92%	91.34%	83.42%
S	-	-	S	62.33%	71.87%	83.75%	91.31%	83.66%
S	-	-	-	62.97%	72.23%	80.89%	90.12%	80.80%

(a) Performance with smartwatch data

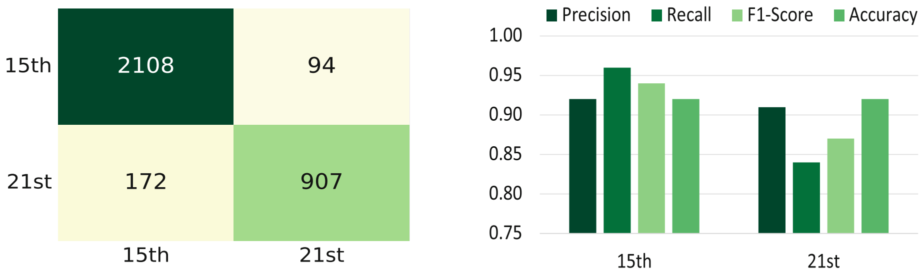
Features				GNB	KNN	SVC	RF	DT
S	M	V	S	77.26%	86.66%	83.78%	89.78%	84.23%
S	M	V	-	75.59%	86.96%	79.14%	89.69%	82.23%
S	M	-	S	75.44%	84.87%	81.23%	90.60%	83.38%
S	M	-	-	74.59%	86.39%	82.08%	89.51%	80.87%
S	-	V	S	76.11%	85.60%	88.87%	89.66%	82.72%
S	-	V	-	74.62%	85.08%	82.02%	90.02%	82.87%
S	-	-	S	75.74%	83.78%	85.08%	90.48%	83.14%
S	-	-	-	73.32%	84.41%	65.16%	88.14%	80.56%

(b) Performance without smartwatch data

Fig. 8. Performance to infer gear of all classifiers based on different combinations of features sensor data (S), maxima/minima (M), variance (V), and standard deviation (S)

5.3 Seat Height

When observing the performance of all classifiers to infer the seat height, we see that the best results are achieved by the RF classifier, closely followed by linear SVC as shown in Fig. 10. We achieve a classification performance of 95.76% for the seat height by using the RF classifier with the feature combination of both smartphone and smartwatch data, maxima/minima, and standard deviation, as shown in Fig. 10a. By excluding the smartwatch data we notice that the classification results are still high and for some classifiers such as KNN and DT even higher than using smartphone and smartwatch data in combination, depicted in Fig. 10b. This shows, that the smartphone data are sufficient to classify the seat height, while the smartwatch data in some cases even reduced the classification performance. Moreover, we observe that normal and high seat settings sometimes are confused with each other, while low and high almost never get confused as shown in Fig. 11.



(a) Gear confusion matrix (b) Gear classification results

Fig. 9. Confusion matrix and classification results for the gears of 15 and 21

Features				GNB	KNN	SVC	RF	DT
S	M	V	S	72.75%	70.86%	93.20%	95.03%	82.78%
S	M	V	-	70.77%	69.86%	92.38%	94.85%	83.42%
S	M	-	S	71.65%	67.02%	90.03%	95.76%	83.72%
S	M	-	-	72.63%	68.61%	92.47%	94.67%	83.69%
S	-	V	S	71.35%	70.89%	91.95%	94.57%	82.69%
S	-	V	-	72.48%	71.17%	93.45%	94.64%	82.41%
S	-	-	S	70.98%	65.53%	92.81%	95.00%	82.57%
S	-	-	-	72.75%	64.40%	93.17%	94.67%	81.93%

(a) Performance with smartwatch data

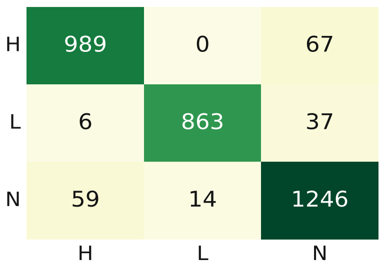
Features				GNB	KNN	SVC	RF	DT
S	M	V	S	69.13%	94.85%	91.78%	94.24%	82.38%
S	M	V	-	69.07%	94.24%	93.03%	95.21%	82.72%
S	M	-	S	69.41%	94.06%	94.33%	95.51%	83.51%
S	M	-	-	69.41%	94.03%	92.87%	95.15%	84.08%
S	-	V	S	69.44%	94.51%	94.00%	94.97%	82.87%
S	-	V	-	70.41%	94.00%	94.12%	95.33%	82.20%
S	-	-	S	68.28%	94.15%	93.30%	94.88%	82.38%
S	-	-	-	70.32%	94.39%	93.97%	94.51%	81.99%

(b) Performance without smartwatch data

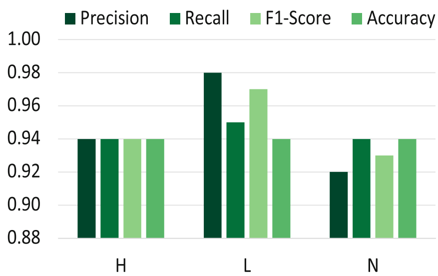
Fig. 10. Performance to infer seat height of all classifiers based on different combinations of features sensor data (S), maxima/minima (M), variance (V), and standard deviation (S)

5.4 Terrain

For terrain detection, the RF classifier performed best, followed closely by the linear SVC, as shown in Fig. 12. To achieve the highest accuracy of 94.24%, a combination of all features: smartphone and smartwatch in addition with minima/maxima, variance, and standard deviation is needed, as depicted in Fig. 12a. By exploring the performances of all classifiers without the smartwatch data in Fig. 12b, we see that the accuracies dropped more significantly in contrast to the other dimensions such as the bike type, gear, and seat height when excluding the smartwatch data. This confirms that different terrains especially affect the steering behavior as well as the shaking of the handle bars. As expected, we also observe that bike road and asphalt terrains can get confused for each other as shown in Fig. 13.



(a) Seat height confusion matrix



(b) Seat height classification results

Fig. 11. Confusion matrix and classification results for the seat heights of high (H), low (L), and normal (N)

Features				GNB	KNN	SVC	RF	DT
S	M	V	S	78.36%	88.05%	91.95%	94.24%	91.80%
S	M	V	-	78.27%	87.32%	89.91%	93.78%	90.98%
S	M	-	S	77.63%	49.80%	82.11%	93.90%	90.73%
S	M	-	-	78.67%	48.13%	84.36%	92.08%	87.56%
S	-	V	S	78.94%	88.08%	90.49%	93.75%	91.28%
S	-	V	-	78.21%	86.96%	85.71%	92.81%	90.61%
S	-	-	S	78.63%	38.86%	85.89%	92.81%	90.83%
S	-	-	-	76.96%	36.21%	70.13%	86.96%	72.42%

(a) Performance with smartwatch data

Features				GNB	KNN	SVC	RF	DT
S	M	V	S	46.51%	74.20%	81.96%	89.48%	80.99%
S	M	V	-	47.45%	74.80%	76.02%	88.48%	80.90%
S	M	-	S	48.48%	73.32%	81.78%	89.36%	81.20%
S	M	-	-	45.45%	72.47%	71.10%	85.81%	77.35%
S	-	V	S	44.09%	70.71%	80.23%	88.36%	80.29%
S	-	V	-	46.36%	71.98%	74.95%	86.11%	81.08%
S	-	-	S	45.09%	69.92%	54.67%	86.60%	81.20%
S	-	-	-	43.69%	68.37%	70.92%	81.84%	70.77%

(b) Performance without smartwatch data

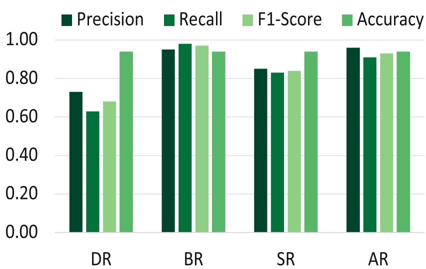
Fig. 12. Performance to infer terrain of all classifiers based on different combinations of features sensor data (S), maxima/minima (M), variance (V), and standard deviation (S)

5.5 User Distinction

The classifier that worked out best to distinguish the participants of our user study is the RF classifier, closely followed by the linear SVC and GNB, as shown in Fig. 14. The highest accuracy of 99.01% resulted in a combination of the features smartphone and smartwatch data in addition with variance and standard deviation, depicted in Fig. 14a. When comparing those results with the performances that are achieved without the smartwatch data, shown in Fig. 14b, we see that the exclusion of the smartwatch data positively affected the KNN and linear SVC classifiers while it only slightly reduced the performance of the RF classifier. Therefore, we want to state that the smartwatch data does not drastically affect the performance to distinguish our participants and that the smartphone data alone is sufficient to distinguish our participants. The confusion matrix and the classification report respectively depicted in Fig. 15a and Fig. 15b show that there are no big confusions between the different participants.

	DR	BR	SR	AR
DR	121	44	27	0
BR	12	2194	4	24
SR	31	5	182	0
AR	2	57	0	578

(a) Terrain confusion matrix



(b) Terrain classification results

Fig. 13. Confusion matrix and classification results for the terrains of dirt road (DR), bike road (BR), stone road (SR), and asphalt road (AR)

Features				GNB	KNN	SVC	RF	DT
S	M	V	S	96.31%	82.84%	95.32%	97.21%	79.15%
S	M	V	-	96.22%	79.80%	96.22%	98.11%	74.22%
S	M	-	S	96.06%	81.77%	96.63%	98.44%	80.13%
S	M	-	-	97.13%	81.61%	97.04%	98.60%	73.15%
S	-	V	S	95.40%	80.79%	95.48%	99.01%	64.86%
S	-	V	-	95.16%	81.61%	96.14%	98.60%	75.21%
S	-	-	S	96.72%	78.57%	95.73%	98.03%	66.34%
S	-	-	-	95.89%	80.62%	95.48%	98.52%	71.43%

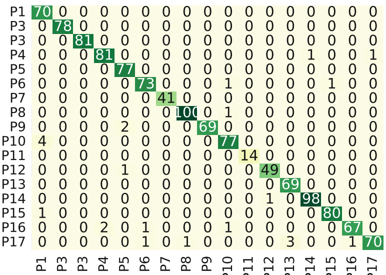
(a) Performance with smartwatch data

Features				GNB	KNN	SVC	RF	DT
S	M	V	S	92.35%	97.31%	98.21%	98.05%	69.49%
S	M	V	-	92.03%	97.48%	98.05%	97.72%	66.48%
S	M	-	S	91.38%	98.54%	98.86%	98.05%	73.80%
S	M	-	-	91.46%	98.05%	98.05%	98.05%	70.06%
S	-	V	S	91.94%	98.37%	98.05%	98.37%	64.77%
S	-	V	-	90.40%	98.13%	98.45%	97.97%	69.73%
S	-	-	S	91.46%	97.48%	98.13%	97.64%	67.78%
S	-	-	-	91.29%	97.07%	98.45%	98.37%	73.80%

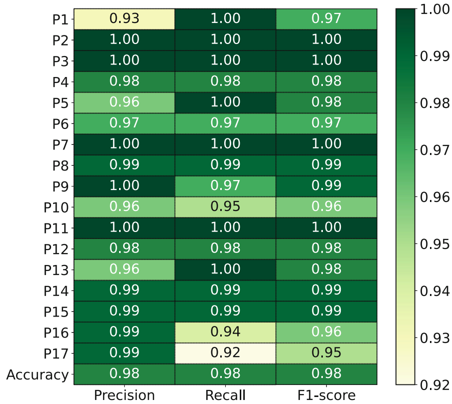
(b) Performance without smartwatch data

Fig. 14. Performance of all classifiers based on different combinations of features sensor data (S), maxima/minima (M), variance (V), and standard deviation (S)

In summary, we have shown that it is possible to infer characteristics such as the bike type, seat height, gear, and terrain when cycling by utilizing machine learning classifiers like RF and linear SVC. Besides the inference of those characteristics, we further show that our participants can be distinguished from each other by taking advantage of the data collected about our participants and training a RF classifier.



(a) User study confusion matrix



(b) User study classification report

Fig. 15. Confusion matrix and classification report for the user study and our 17 participants

6 Limitations and Outlook

Although we are able to correctly identify the characteristics discussed and also show that we are able to identify our participants based on the collected movement data, our results are based on a limited number of participants. As the number of people increases, the accuracy in identifying individual users might decrease. Moreover, our sample is not representative of the population. As a result, it would be interesting in the future to explore the feasibility to infer information, such as the age or the gender based on the unique cycling movements.

Moreover, it would be interesting to investigate the impact of the placement of the smartphone from the front pocket of the trouser to the back pocket, or other usual locations, such as a backpack [26].

In our study, we have also only explored different terrains without major decline or incline, which could be another interesting feature to observe in future studies.

Furthermore, we would like to note that we have not considered other factors, such as the weather or the way participants change directions with their bikes, which could be other interesting aspects for future studies.

While we have focused on showing the feasibility of inferring different types of information about our participants, their bike, and the terrain from movement data in this paper, we have left open, which solutions could be applied to prevent such inferences, while still allowing the underlying fitness applications to function.

An interesting direction to follow is how to communicate the resulting risks to their privacy and make transparent which data are collected. Works in this direction exist for smartwatches [2,27] or fitness trackers in general [28,29], but they are dedicated to other contexts and do not cover specifically our biking scenario. Similarly, the impact of applied privacy-preserving solutions on their data could be communicated to the cyclists like it is the case for the application of differential privacy on health data [30–32]

7 Conclusion

In this study, we have explored the potential of the movement data, generated by a smartphone and a smartwatch, to classify user as well as environmental characteristics when riding a bike. The four characteristics we observe are the bike type, gear, seat height, and terrain. Additionally, we examine on the potential to distinguish/identify users from each other solely on their unique cycling behavior. To this end, we first created a dataset based on movement data collected from a single person, representing all different combinations of the characteristics. Then, we conducted a user study with 17 participants who were equipped with a smartphone and a smartwatch and collected data while cycling for two km on the same road for all participants.

The results show that the RF classifier performed best among all classifiers that we have explored. Moreover, we were able to show that the smartwatch data can increase the performance accuracy for most characteristics. Especially for the detection of the terrain the smartwatch data increased the accuracy compared to the smartphone data alone. Our results show that we achieve prediction accuracies for our four characteristics bike type, gear, seat height, and terrain of 93.05%, 92.23%, 95.76%, 94.24% respectively. Also in the user study we were able to recognize a person from a crowd of 17 participants with a probability of 99.01%.

Overall, our results shed light on the performance of sensor readings collected by devices we use every day to predict user-specific and environmental characteristics. As a result, applications making use of these data can create fine-grained user profiles that users are often not aware of. We therefore recommend to design new user-friendly solutions, increase the transparency of the collected data, and allow users to make informed decisions about their privacy.

References

1. Richter, A., Kühtreiber, P., Reinhardt, D.: On the Impact of Information provided to Employees on their Intention to disclose Data collected by Smart Watches to their Employers. Association for Information Systems (AIS) (2022)
2. Richter, A., Kühtreiber, P., Reinhardt, D.: Enhanced privacy in smart workplaces: employees' preferences for transparency indicators and control interactions in the case of data collection with smart watches. In: Proceedings of the 37th International Conference on Information Security and Privacy Protection (IFIP SEC) (2022)
3. Hernández Acosta, L., Reinhardt, D.: A survey on privacy issues and solutions for voice-controlled digital assistants. Pervasive Mob. Comput. (PMC) **80**, 101523 (2021)
4. Siepmann, C., Kowalczuk, P.: Understanding continued smartwatch usage: the role of emotional as well as health and fitness factors. Electr. Markets (EM) **31**(4), 795–809 (2021)
5. Apple: Measuring Walking Quality Through iPhone Mobility Metrics (2021). https://www.apple.com/in/healthcare/docs/site/Measuring_Walking_Quality_Through_iPhone_Mobility_Metrics.pdf. Accessed 05 July 2022
6. Leiding, B., Bochem, A., Hernández Acosta, L.: Automated sensor-fusion based emergency rescue for remote and extreme sport activities. In: Proceedings of the 15th IEEE International Wireless Communications and Mobile Computing Conference (IWCMC) (2019)
7. Anjum, A., Ilyas, M.U.: Activity recognition using smartphone sensors. In: Proceedings of the 10th IEEE Consumer Communications and Networking Conference (CCNC) (2013)
8. Voicu, R.A., Dobre, C., Bajenaru, L., Ciobanu, R.I.: Human physical activity recognition using smartphone sensors. Sensors **19**(3), 458 (2019)
9. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. ACM SIGKDD Explor. Newsl. **12**(2), 74–82 (2011)
10. Weiss, G.M., Timko, J.L., Gallagher, C.M., Yoneda, K., Schreiber, A.J.: Smartwatch-based activity recognition: a machine learning approach. In: Proceedings of the 13th IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)
11. Balli, S., Sağbaş, E.A., Peker, M.: Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm. Meas. Control **52**(1–2), 37–45 (2019)
12. Zou, Q., Wang, Y., Wang, Q., Zhao, Y., Li, Q.: Deep learning-based gait recognition using smartphones in the wild. IEEE Trans. Inf. Forensics Secur. **15**, 3197–3212 (2020)

13. Kröger, J.L., Raschke, P., Bhuiyan, T.R.: Privacy implications of accelerometer data: a review of possible inferences. In: Proceedings of the 3rd International Conference on Cryptography, Security and Privacy, pp. 81–87 (2019)
14. Straczkiewicz, M., James, P., Onnela, J.P.: A systematic review of smartphone-based human activity recognition methods for health research. NPJ Digit. Med. **4**, 148 (2021)
15. Ramanujam, E., Perumal, T., Padmavathi, S.: Human activity recognition with smartphone and wearable sensors using deep learning techniques: a review. IEEE Sens. J. **21**(12), 13029–13040 (2021)
16. Mekruksavanich, S., Jitpattanakul, A.: LSTM networks using smartphone data for sensor-based human activity recognition in smart homes. Sensors **21**(5), 1636 (2021)
17. Sena, J., Barreto, J., Caetano, C., Cramer, G., Schwartz, W.R.: Human activity recognition based on smartphone and wearable sensors using multiscale DCNN ensemble. Neurocomputing **444**, 226–243 (2021)
18. Zimbelman, E.G., Keefe, R.F.: Development and validation of smartwatch-based activity recognition models for rigging crew workers on cable logging operations. PLoS ONE **16**(5), e0250624 (2021)
19. Mallol-Ragolta, A., Semertzidou, A., Pateraki, M., Schuller, B.: harAGE: a novel multimodal smartwatch-based dataset for human activity recognition. In: Proceedings of the 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG) (2021)
20. Fatima, S.: Activity recognition in older adults with training data from younger adults: preliminary results on in vivo smartwatch sensor data. In: Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (2021)
21. Johnston, A.H., Weiss, G.M.: Smartwatch-based biometric gait recognition. In: Proceedings of the 7th IEEE International Conference on Biometrics Theory, Applications and Systems (BTAS) (2015)
22. Lee, W.H., Liu, X., Shen, Y., Jin, H., Lee, R.B.: Secure pick up: implicit authentication when you start using the smartphone. In: Proceedings of the 22ND ACM on Symposium on Access Control Models and Technologies (2017)
23. Haring, M., Reinhardt, D., Omlor, Y.: Pick me up and i will tell you who you are: analyzing pick-up motions to authenticate users. In: Proceedings of the 16th IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops) (2018)
24. Matkovic, V., Waltereit, M., Zdankin, P., Uphoff, M., Weis, T.: Bike type identification using smartphone sensors. In: Proceedings of the 19th ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers (2019)
25. Matkovic, V., Waltereit, M., Zdankin, P., Weis, T.: Towards bike type and e-scooter classification with smartphone sensors. In: Proceedings of the 17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous) (2020)
26. Aboo, A.K.: Survey on human activity recognition using smartphone. AL-Rafidain J. Comput. Sci. Math. **15** (2021)
27. Shaw, P., Mikusz, M., Davies, N., Clinch, S.: Using smartwatches for privacy awareness in pervasive environments. In: Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications (HotMobile) (2017)

28. Murmann, P., Beckerle, M., Fischer-Hübner, S., Reinhardt, D.: Reconciling the what, when and how of privacy notifications in fitness tracking scenarios. Pervasive Mob. Comput. **77**, 101480 (2021)
29. Velykoivanenko, L., Niksirat, K.S., Zufferey, N., Humbert, M., Huguenin, K., Cherubini, M.: Are those steps worth your privacy? fitness-tracker users' perceptions of privacy and utility. Proc. ACM Interact. Mob. Wearable Ubiquit. Technol. **5**(4), 1–41 (2021)
30. Murmann, P., Reinhardt, D., Fischer-Hübner, S.: To be, or not to be notified. In: Proceedings of the 34th IFIP International Conference on ICT Systems Security and Privacy Protection (2019)
31. Kühtreiber, P., Pak, V., Reinhardt, D.: Replication: the effect of differential privacy communication on german users' comprehension and data sharing attitudes. In: Proceedings of the 18th Symposium on Usable Privacy and Security (SOUPS) (2022)
32. Saifuzzaman, M., Ananna, T.N., Chowdhury, M.J.M., Ferdous, M.S., Chowdhury, F.: A systematic literature review on wearable health data publishing under differential privacy. Int. J. Inf. Secur. **21**(4), 847–872 (2022)