Expert Systems With Applications 147 (2020) 113195

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Gender recognition using motion data from multiple smart devices

Jianmin Dong, Youtian Du, Zhongmin Cai*

MOE KLINNS Lab, Xi'an Jiaotong University, No. 28, Xianning West Road, Xi'an, Shaanxi 710049, PR China

A R T I C L E I N F O

Article history:
Received 4 April 2019
Revised 7 January 2020
Accepted 7 January 2020
Available online 16 January 2020

Keywords:
Gender recognition
Motion sensor
Multiple smart devices
Performance evaluation
Walking behavior

A B S T R A C T

Using multiple smart devices, such as smartphone and smartwatch simultaneously, is becoming a popular life style with the popularity of wearables. This multiple-sensor setting provides new opportunities for enhanced user trait analysis via multiple data fusion. In this study, we explore the task of gender recognition by using motion data collected from multiple smart devices. Specifically, motion data are collected from smartphone and smart band simultaneously. Motion features are extracted from the collected motion data according to three aspects: time, frequency, and wavelet domains. We present a feature selection method considering the redundancies between motion features. Gender recognition is performed using four supervised learning methods. Experimental results demonstrate that using motion data collected from multiple smart devices can significantly improve the accuracy of gender recognition. Evaluation of our method on a dataset of 56 subjects shows that it can reach an accuracy of 98.7% compared with the accuracies of 93.7% and 88.2% when using smartphone and smart band individually.

1. Introduction

The increasing integration of sensors into smartphones has made sensor data analysis for the study of user behavior and traits an interesting new research direction for many researchers in the community of expert systems (Jain & Kanhangad, 2018; Kwon, Kang, & Bae, 2014; Ronao & Cho, 2016). This trend is recently being further fostered by wearables, such as smartwatches and smart bands, which are quickly becoming popular complements to smartphones for notifications and activity tracking. Various wearables and their applications are developed and widely used in consumer markets (Yoon, Park, & Lee, 2016). CCS Insight (Insight, 2018) reported that consumer awareness of smartwatch has risen by 10%, and the market for wearable devices will be worth $25 billion by 2019. Using multiple smart devices, such as a smartphone and a smart band (watch) together, is becoming a popular life style. Given that wearables are worn by users nearly everywhere they go, they provide an ideal platform for the study of user behaviors and characteristics.

One interesting problem in sensor data analysis on smart devices is gender recognition. Gender information has many applications in security, recommendation systems, and market research. It is used to improve accuracies of biometric identification systems in conjunction with other information or traits. Jain et al. showed that gender information introduced to a fingerprint recognition system improves accuracy by 5% (Jain, Dass, & Nandakumar, 2004). Gender information also plays an important role in recommendation systems, online shopping, and market surveys (Bologna et al., 2013; Nunes, 2008). It is useful to advertisers and vendors because it provides meaningful cues to make accurate predictions or support customized services. Many works have been done for gender recognition by using various data sources such as face (Golomb, Lawrence, & Sejnowski, 1990; Nguyen, Cho, Shin, Bang, & Park, 2014b), electroencephalogram (EEG) (Hu, 2018), voice (Harb & Chen, 2003), gait (Li, Maybank, Yan, Tao, & Xu, 2008; Yu, Tan, Huang, Jia, & Wu, 2009), keyboard strokes, and mouse behaviors (Shen, Cai, Maxion, & Guan, 2013). Motion data collected by smart devices can be used to analyze user traits, such as gender (Jain & Kanhangad, 2016; 2018; Weiss & Lockhart, 2011). Better accuracy may be achieved for gender recognition with data collected simultaneously from multiple smart devices than using data only from a single device.

In this study, we thoroughly investigate the problem of whether having behavioral data sensed from multiple devices will truly improve the performance of user trait prediction. In particular, we use two commonly used smart devices, namely, a smartphone and a smart band, to collect motion data and perform gender recognition. The major contributions of this study are summarized as follows.

- We thoroughly investigate the value of behavioral data collected from multiple smart devices and perform extensive ex-

* Corresponding author.
E-mail addresses: jianmind23@stu.xjtu.edu.cn (J. Dong), duyt@xjtu.edu.cn (Y. Du), zmcai@sei.xjtu.edu.cn (Z. Cai).

https://doi.org/10.1016/j.eswa.2020.113195
0957-4174/

periments to demonstrate the achieved improvements for gender recognition.
- We propose a methodological framework for a complete analysis of behavioral data collected from multiple smart devices including data collection, preprocessing, feature extraction and selection, and design of the classifiers and their evaluations.
- We extract motion features from time, frequency and wavelet domains for the behavioral data collected by smart devices. We select the most effective features to represent the characteristics of gender by evaluating the redundancies between motion features.
- Experimental results demonstrate that using motion data collected from multiple smart devices can significantly improve the accuracy of gender recognition. Evaluation of our method on a dataset of 56 subjects shows that it can reach an accuracy of 98.7% compared with the accuracies of 93.7% and 88.2% when using motion data only from an individual smartphone and smart band, respectively.
- We establish a motion dataset of 56 subjects for gender recognition via a carefully designed procedure using a smartphone and a smart band simultaneously. This dataset will be useful for subsequent studies in behavioral data collected from multiple smart devices.

The rest of the paper is organized as follows. Related works are discussed in Section 2. Data collection is described in Section 3. Details of feature analysis are presented in Section 4. Classifiers for gender recognition are introduced in Section 5. The experimental results and analysis are shown in Section 6. Deployment issues are shown in Section 7. Limitations and future work are provided in Section 8. The conclusion is presented in Section 9.

2. Related works

Gender recognition is an interesting topic in pattern recognition and machine learning. Traditional approaches use several different types of data sources, such as face (Golomb et al., 1990; Nguyen et al., 2014b), EEG (Hu, 2018), voice (Harb & Chen, 2003), gait (Li et al., 2008; Yu et al., 2009), keyboard strokes and mouse movements (Shen et al., 2013), and text (Deitrick et al., 2012; Figueroa, 2017; H & Omman, 2014; Jianle Chen, Tianqi Xiao, Jie Sheng, & Teredesai, 2017; Nguyen et al., 2014a; Schwartz et al., 2013), to predict gender. Specifically, Golomb et al. (1990) presented a method of predicting users' gender using face images and achieved an average accuracy of 91.9%. Hu (Hu, 2018) proposed an approach of using EEG signals collected from 28 subjects to recognize gender and achieved an accuracy of 99.8%. Harb and Chen (2003) introduced a gender identification approach using people's voice and achieved an accuracy of 92.0%. Yu et al. (2009) investigated gender classification using gait information and achieved average accuracies of 67.8% using the lower body silhouettes, 94.3% using the upper body silhouettes, and 95.4% using the whole-body silhouettes. Shen et al. (2013) applied the data of keyboard strokes and mouse movements to predict users' gender. They achieved the accuracies of 83.0% using data of keyboard strokes and 85.0% using data of mouse movements.

Many researchers have used text information for gender recognition. Nguyen et al. (2014a) studied the prediction of an author's gender and age from text. They showed that more than 10% of the Twitter users do not employ languages that associate with their biological sex. Schwartz et al. (2013) analyzed 700 million words, phrases, and topic instances collected from Facebook messages of 75,000 volunteers and found striking variations in language with personality, gender, and age. Figueroa (2017) investigated a large-scale corpus automatically constructed from the integration of Yahoo! Search and Yahoo! Answer profiles to determine the gender of a questioner. Their best results reported an accuracy of 74.5%. Jianle Chen et al. (2017) proposed an approach that combines the latent semantic indexing method with K-nearest neighbor (KNN) to predict the gender based on a real-life collection of posts for blog pages. H and Omman (2014) studied gender identification by using blog data and achieved a best accuracy of 78.8%. Deitrick et al. (2012) examined the gender recognition problem on the popular Enron email dataset. They achieved accuracies of 88% with the stylometric features and around 95% with the word based features.

Given the popularity of smartphones, the motion data collected by smart devices provide a new data source for gender recognition. Two groups of researchers have performed gender recognition using motion data collected from a smartphone (Jain & Kanhangad, 2016; 2018; Weiss & Lockhart, 2011). Weiss and Lockhart (2011) used accelerometer data recorded by a smartphone in a walking user's pants pocket to predict gender. Jain and Kanhangad (2016, 2018) proposed a method of combining motion data from accelerometer and gyroscope of a smartphone to improve the performance of gender recognition. Weiss and Jain achieved best accuracies of 71.2% (Weiss & Lockhart, 2011) and 94.4% (Jain & Kanhangad, 2018), respectively. Their works revealed strong correlations between motion data and gender. However, they only investigated motion data from a single device, that is, the smartphone, for gender recognition.

Given the emergence of wearables, such as smart band and smartwatch, using multiple smart devices, such as smartphone and smart band simultaneously, in our daily lives has become a frequent phenomenon. In this study, we develop a new approach to recognize gender using motion data from multiple smart devices simultaneously.

3. Data collection

In this part, we introduce our approach of motion data collection from two smart devices simultaneously. We use an Android mobile phone as our smartphone and a Microsoft Band as smart band. The two devices provide a set of open Software Development Kits for the development of data collection applications. We collect walking data as our motion data for gender recognition because walking is the most common activity in our daily lives.

3.1. Data collection system

The smartphone used in our data collection is Samsung Galaxy S4, and the smart band is Microsoft Band 1. For conciseness, we refer to them as S4 and MSBand in later parts of this paper. In the data collection system, an application for the control and data retrieval of MSBand and S4 is developed on a third Android smartphone. The application is connected to MSBand and S4 via Bluetooth. The application issues commands to the MSBand and S4 to simultaneously start and stop the collection of walking data. When the data are being generated, the application will continuously retrieve the motion data from MSBand and S4 and deposit them in files on the control smartphone.

Two motion sensors of accelerometer and gyroscope from MSBand and S4 are utilized to collect walking data. The sampling rates of accelerometer and gyroscope in the MSBand are 62 and 31Hz, respectively, and the sampling rates of accelerometer and gyroscope in S4 are 90Hz.

3.2. Data collection

In this study, we recruit 56 volunteers, record their user traits, and collect their walking data with the MSBand and S4 while they are walking in natural modes.

All 56 volunteers are college students who are healthy and have no injuries. We record the gender information of each subject. We recruit 28 male and 28 female subjects. The ages of the subjects are within the range from 20 years to 30 years.

All subjects are required to wear MSBand on their left wrist and with S4 in the right-front pocket of their trousers, because the two positions are commonly used in the previous studies (Kwapisz, Weiss, & Moore, 2011; Primo, Phoha, Kumar, & Serwadda, 2014). Ten seconds of motion data collected in one round of walking is called one sample. Each subject is asked to participate in the data collection process for five times on different days. Each time, 10 rounds of walking along a long and straight corridor are performed. The data collection lasted for about 4 weeks. We obtain 2800 left wrist samples and 2800 right leg samples from the 56 subjects. For each subject, we have 50 samples of walking data sensed by the MSBand on the left wrist and another 50 samples from the S4 in right trouser pocket. Every data sample of S4 or MSBand consists of two files: one is a 3D acceleration data file, and the other is a 3D angular velocity data file. We use their gender information to generate the corresponding labels or ground truths for the collected walking data samples. The dataset will be published with this paper[1].

3.3. Comparison withexisting datasets

To date, two groups of researchers have performed gender recognition using motion data collected from a smartphone (Jain & Kanhangad, 2016; 2018; Weiss & Lockhart, 2011).

Weiss and Lockhart (2011) used accelerometer data recorded by a smartphone in a walking user's pants pockets to predict gender. The study recruited 66 subjects consisting of 38 (57.6%) men and 28 (42.4%) women. The ages of subjects were in the range from 18 years to 24 years. A best accuracy of 71.2% was reported (Weiss & Lockhart, 2011). Jain and Kanhangad (2016) proposed a method of combining motion data from accelerometer and gyroscope of a smartphone to improve the performance of gender recognition. They collected 42 subjects with 25 males and 17 females.

No previous datasets provide motion data collected from multiple smart devices simultaneously. We have to establish the dataset to explore the new settings for simultaneous usage of multiple devices.

4. Feature extraction and analysis

Feature extraction is an important step for gender recognition using motion sensor data. In our approach, features extracted from motion data are called motion features. We conduct an empirical analysis to search for effective and discriminative motion features for gender recognition.

4.1. Data preprocessing

4.1.1. Filtering and resampling

Mean filter algorithm is used with five-point approximation to reduce the effect of noise (Mostayed, Kim, Mazumder, & Park, 2008). Given that the sampling rates of accelerometer and gyroscope in MSBand and S4 are different, we resample all the data samples at the re-sampling rate of 90Hz.

4.1.2. Cycle detection and segmentation

Given that walking is a periodic activity, one original sample of 10s of walking consists of several walking cycles. A walking cycle corresponds to the complete process of one step of walking, that is, two consecutive movements performed by the two legs. We utilize the (peak, 0, valley, 0, peak) pattern in one cycle of the motion signal to cut the original data sequence of 10s into smaller segments corresponding to individual walking cycles. Feature calculation and extraction is performed on these smaller segments, and the results for the same sample are averaged to generate its feature vector.

4.2. Feature extraction

Motion features extracted in our study can be generally classified into three categories: time, frequency, and wavelet domain features.

- Time domain features: Twelve features are extracted from time domain, namely, standard derivation (STD), mean, min, max, peak, kurtosis, energy, root mean square errors (RMS), median, mean crossing rate (rate of times signal crossing mean value), gait cycle, and amplitude.
- Frequency domain features: We use fast Fourier transformation technique to transform signals from time domain into frequency domain. In frequency domain, we extract five features: mean, standard derivation, kurtosis, skewness, and energy of shape and amplitude as the frequency domain features.
- Wavelet domain features: Wang, Ambikairajah, Lovell, and Celler (2007a) identified that low-frequency signals are more salient in characterizing human activities. In our work, we apply a five-level wavelet decomposition to extract low-frequency information from the three dimensions of accelerometer and gyroscope signals for smartphone and smart band. A fourth-order Daubechies wavelet is applied to extract wavelet coefficients of the low-frequency part at the 4th and 5th levels for the calculation of the wavelet energies.

We select features from time, frequency, and wavelet domains to characterize motion sensor data that may relate to differences in gender. For example Yu et al. (2009) pointed out that females prefer to sway their hips while males tend to sway their shoulders during walking. Time domain features, namely, mean, max, and min are selected to describe the information of strength, speed, and amplitude of swinging arms and legs. Frequency domain features are selected to describe distributions of frequency components that provide an overall characterization of walking paces. Wavelet domain features are selected to characterize the low-frequency part of the behavior and generate a comprehensive characterization from a multi-scale point of view.

All features from time, frequency, and wavelet domains are listed in Table 1.

Table 1 shows 264 features extracted from one sample of walking data. We extracted 12, 5, and 1 features for every dimension of the acceleration and angular velocity signals in time, frequency, and wavelet domains, respectively. Given that we record a left-hand sample and a right-leg sample, we have 264 motion features in total characterizing one round of walking for one subject.

4.3. Featurenormalization

Motion features extracted from different aspects have different measurements. We perform feature normalization to rescale all motion features into the same scale of N(0, 1) for eliminating the imbalance caused by different scales of feature measurements. We implement the feature normalization for each feature based on Eq. (1).

$$x_{norm} = \frac{x - \mu}{\sigma} \tag{1}$$

where x is the original feature value, μ is the mean value of the feature values, σ is the STD of the feature values, and x_{norm} is the normalized feature value.

[1] https://github.com/Jianmindong2019/WalkingDataSet.

Table 1
Motion features descriptions.

Features	# of T.	# of F.	# of W.
mean	12	24	
STD	12	24	
min	12		
max	12		
peak	12		
kurtosis	12	24	
mean crossing rate	12		
energy	12	12	12
RMS	12		
median	12		
skewness		24	
gait cycles	12		
amplitude	12		
	144	108	12
Total:		264	

T.: time domain features; F.: frequency domain features; W.: wavelet domain features.

4.4. Feature selection

We conduct an empirical feature analysis approach using a method proposed in Estévez, Tesmer, Perez, and Zurada (2009) to reduce the computational load, reserve discriminative information, and eliminate redundant features. This feature selection method includes two steps: 1) choosing the features with a close relationship with the corresponding labels as the "Feature_Label" set, and 2) reducing a few features with high redundancy and constructing the selected feature set. Additional details are given as follows.

Step 1 : Construction of "Feature_Label" set.

We use correlation coefficient (CC) (Mukaka, 2012) to analyze two relationships: 1) the relationships between motion features and their corresponding gender labels and 2) the relationships between each pair of 264 motion features. Under the situation of calculating CC between a continuously measured variable and a binary variable, CC is equivalent to point biserial CC, which is specially proposed for linear CC calculation when one variable is continuous and the other one is dichotomous (Cohen, West, & Aiken, 2014) (Appendix). The CC is calculated using the following Eq. (2).

$$CC_j = \frac{\sum_{i=1}^{N}(f_{ij}-\bar{f}_j)(T_i-\bar{T})}{\sqrt{\sum_{i=1}^{N}(f_{ij}-\bar{f}_j)^2}\sqrt{\sum_{i=1}^{N}(T_i-\bar{T})^2}}, j=1,\cdots,M; \qquad (2)$$

Where f_{ij} stands for the j^{th} feature of i^{th} sample, T_i stands for the label of i^{th} sample, M and N stand for the total number of features and samples, respectively.

We have 56 independent subjects; thus, we have 54 degrees of freedom (DOF). With this DOF, if the coefficient is over 0.26, then the feature is significantly correlated with its corresponding label at the significance level 0.05 (Mukaka, 2012). The results of $|CC|$ are shown in Fig. 1. A total of 118 top-performing features with a coefficient of over 0.26 are selected as the "Feature_Label" set. We observe time domain features have higher $|CC|$ on average than frequency and wavelet domain features. This condition proves that time domain features can describe the motion data better.

Step 2 : Construction of the selected feature set.

Fig. 2 shows the $|CC|$ between each pair of 118 features in the "Feature_Label" set. The warm color denotes a higher dependence, and the cold color denotes a lower dependence. The figure shows some warm colors exist, which means pairwise features have significant associations. Specifically, Fig. 2 shows that some features from time, frequency, and wavelet domains have high correlations with each other, whereas others have low correlations. These results shed light into some interesting facts. For example, time domain features extracted from 3D angular velocity of smartphone have high correlations with each other. This relationship reveals that the fact that the movement patterns of the upper leg, where the smartphone is placed, is relatively simple, and the angular movements only exhibit approximately one DOF and can be characterized roughly by one scalar variable. However, the motion features extracted from accelerations of smart band have low correlations with each other. This result is likely due to that motion patterns of the wrist are complex and have at least several DOFs. These motion patterns require additional independent variables for characterization.

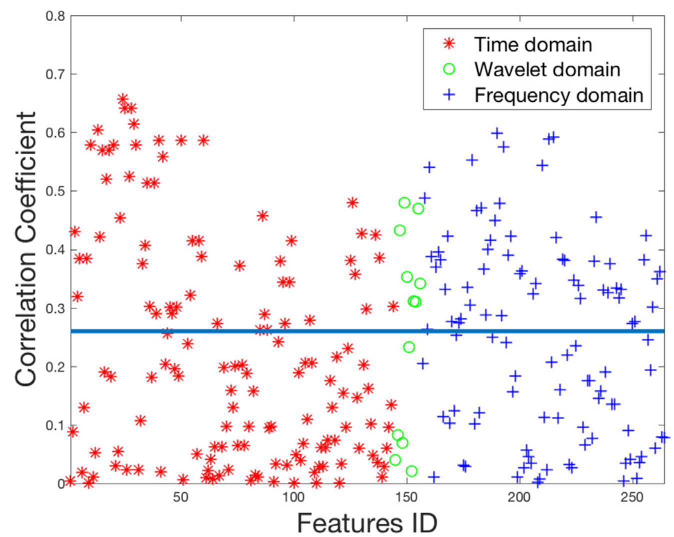

Fig. 1. The $|CC|$ between motion features and gender.

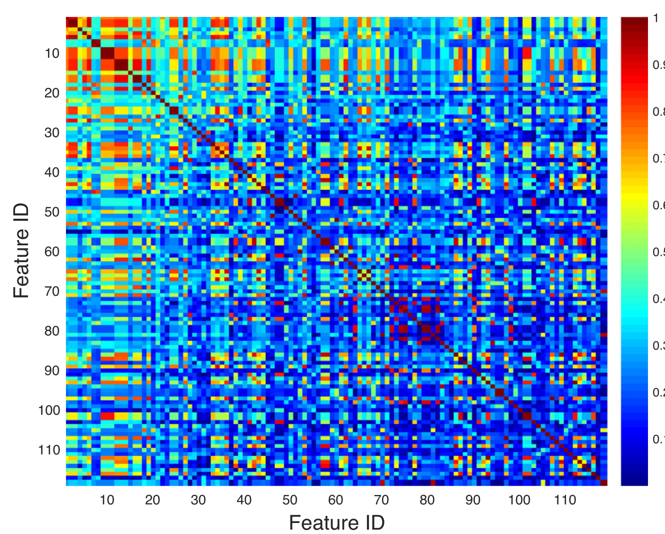

Fig. 2. The $|CC|$ between pairwise features in Feature_Label set.

Therefore, we compute the $|CC(f_a, f_b)|$ between every two features f_a and f_b in the "Feature_Label" set and then choose the pair features with $|CC|$ larger than 0.9 (Mukaka, 2012) (given that 0.9 means very high positive correlation). We delete one of the pair features with smaller $|CC|$ with its corresponding label from the "Feature_Label" set and finally obtain the selected feature set. A total of 96 features are selected, and the selected feature set is shown in Table 2.

In Table 2, only about one-third of all features (96 out of 264) are selected as the selected feature set, which illustrates that only a small part of our raw feature set is efficient for gender recog-

Table 2
Motion feature descriptions.

Features	# of T.	# of F.	# of W.	# of Total
Raw feature set	144	108	12	264
"Feature_Label" set	56	57	5	118
Selected feature set	46	45	5	96

T.: time domain features; F.: frequency domain features; W.: wavelet domain features.

nition. This selected feature set includes 46 motion features from time domain, 5 motion features from wavelet domain, and 45 motion features from frequency domain. Therefore, some motion features from time, frequency, and wavelet domains provide useful information for gender recognition.

We also explore the effect of the mean filter preprocessing on frequency-related features. Mean filter preprocessing has double-sided effects on the perspective of selected features. It removes not only noise, but also high-frequency components. We evaluate the settings for with and without mean filter preprocessing and find that it is useful for gender recognition as a whole.

We empirically compare our feature selection method and other methods, such as Fisher score (Wang, Liu, & Zheng, 2007b), mutual information (Pohjalainen, Räsänen, & Kadioglu, 2015), KS test (Chen & Shen, 2017), and PCA. Our method performs the best on our dataset.

5. Classifiers for gender recognition

In our study, we utilize four classifiers, namely, support vector machine (SVM) (RBF kernel function), KNN (Euclidean distance), back propagation (BP) neural network and random forest (RF), to perform gender recognition. All hyperparameters applied in these classifiers are provided by running experiments many times to search for the optimal values.

5.1. SVM

SVM was introduced for novelty detection by Schölkopf, Williamson, Smola, Shawe-Taylor, and Platt (2000). It maps the training data into a high-dimensional space and tries to find an optimal hyperplane with a large margin. For the nonlinear separable training data, other kernel functions can be implemented to map them into a new high-dimensional feature space where the points are linear separable. In our study, we adopt LIBSVM (Chang & Lin, 2011) to implement RBF kernel SVM. For the RBF kernel, we set the parameter $g = 0.01$ and $nu = 0.1$.

5.2. KNN

KNN is a type of instance-based learning algorithm (Duda & Hart, 1973). The basic idea of this algorithm is to classify the testing sample to the label that is most frequent among the k training samples nearest to that query point. In our study, we use Euclidean distance and standard implementation provided in Matlab to evaluate the performance. Here, we set $k = 6$.

5.3. BP neural network

BP is a prevalent classification method of identifying patterns (Duda & Hart, 1973) and currently one of the most widely used neural network model (Shen, Zhang, Guan, & Maxion, 2016). It is a multilayer feedforward network training by error backward propagation algorithm. Its goal is to minimize the sum of square error of the network by adjusting the network weights and threshold constantly using BP. Here, we use a two-hidden-layer network. We use the standard implementation provided in Matlab. Sigmoid function is used as activation function and the prediction outputs is [-1,1].

5.4. RF

RF is an ensemble method with multiple decision trees (Breiman, 2001). A separate bootstrap sample strategy from the training dataset is used to obtain the trees. Each tree classifies the data, and the final result is obtained through a majority vote among the trees. During the training stage, training samples of each decision tree (800 decision tress in the forest) are randomly selected from the entire dataset. Every node of the decision tree selects the best weak classifier. All decision trees are combined to form the RF classifier. During the testing stage, each decision tree outputs two confidence coefficients when a sample is tested, and the final judgment is based on the average of all decision tree results. The standard implementation provided in Matlab is used.

5.5. Evaluation Metric

In this study, we use the *accuracy* as our metric to evaluate the performance of gender recognition. Accuracy is calculated as described as Eq. (3).

$$Accuracy = \frac{N_{Correct}}{N_{Total}} \tag{3}$$

where $N_{Correct}$ and N_{Total} represent the number of correctly recognized testing samples and the number of total testing samples, respectively.

6. Experiments and analysis

In this part, we evaluate the performance of gender recognition using motion data collected from multiple smart devices. Five experiments are conducted: 1) the evaluation of multiple and single device scenarios, 2) the effect of feature selection and dimension reduction, 3) the effect of sample sizes, 4) the effect of subject sizes and 5) the effect of different sensors. A 10-fold cross validation is used to partition the training and testing datasets. Each 10-fold experiment is repeated for 10 times. Each time, we randomly divide the 28 male and 28 female subjects into 10 partitions of roughly equal sizes. Motion data from subjects in 9 partitions are used for training, and the rest is used for testing. We apply feature selection only on the training dataset to prevent the leakage of testing information. In partitioning the training and testing datasets, motion data from the same human subject are either used for training or testing but not for both. Thus, the information leakage due to the authentication effect is also avoided.

6.1. Gender recognition using motion data from multiple devices

We design three scenarios to evaluate the effect of gender recognition using multiple sensor data sources: 1) multiple devices: using motion data from the smartphone and the smart band simultaneously; 2) smartphone only: using motion data from the smartphone device only; 3) smart band only: using motion data from the smart band device only. A 10-fold cross validation is used to partition the training and testing datasets. Each 10-fold experiment is repeated for 10 times. Each time, we randomly divide the 28 male and 28 female subjects into 10 partitions of roughly equal sizes. Motion data from subjects in 9 partitions are used for training, and the rest is used for testing.

Experiments are conducted with respect to different feature subsets and classifiers to investigate the effects on gender recognition. In each scenario, we use four feature subsets to evaluate the performance of gender recognition.

Table 3
The accuracies (%) of gender recognition using multiple devices (with STD in parentheses).

Classifiers	T.		F.		W.		T. & F. & W.	
	All	Sel.	All	Sel.	All	Sel.	All	Sel.
Multiple devices								
SVM	71.9 (2.2)	95.6 (2.1)	76.1 (4.1)	**93.4** (3.6)	71.9 (5.2)	77.6 (4.2)	82.4 (2.8)	97.5 (3.4)
KNN	85.4 (6.0)	89.4 (4.3)	78.5 (6.7)	85.3 (4.9)	73.1 (5.0)	74.2 (5.1)	86.7 (6.5)	99.5 (4.6)
BP	88.7 (1.6)	**95.6** (2.0)	83.8 (3.7)	90.1 (2.4)	79.4 (4.6)	**84.7** (4.1)	**92.1** (2.1)	97.4 (1.8)
RF	86.8 (4.8)	94.1 (2.8)	83.7 (5.4)	88.8 (3.2)	73.4 (6.4)	80.2 (4.7)	88.0 (5.9)	**98.7** (2.9)
Smartphone only								
SVM	71.8 (4.8)	95.4 (2.9)	85.1 (4.3)	91.6 (4.2)	80.1 (5.1)	82.1 (4.8)	79.9 (2.2)	93.4 (3.4)
KNN	73.7 (7.5)	85.9 (3.7)	82.1 (4.9)	85.3 (3.0)	74.9 (5.0)	70.4 (5.9)	86.5 (6.3)	88.7 (5.7)
BP	89.4 (3.4)	91.7 (2.2)	85.7 (2.7)	86.2 (2.6)	**81.4** (5.4)	84.5 (4.1)	91.8 (3.2)	92.7 (4.3)
RF	**90.3** (5.2)	91.6 (3.9)	**87.6** (5.6)	90.2 (4.3)	74.5 (6.3)	79.2 (5.7)	89.4 (4.6)	93.7 (2.7)
Smart band only								
SVM	82.9 (5.5)	76.5 (2.7)	84.2 (4.2)	82.3 (3.1)	76.4 (6.6)	74.2 (3.3)	86.5 (4.4)	85.4 (3.1)
KNN	67.3 (6.7)	74.3 (4.3)	80.1 (4.2)	74.1 (4.5)	65.7 (6.3)	73.5 (4.1)	83.7 (5.3)	86.3 (4.2)
BP	76.0 (6.2)	75.2 (3.1)	83.2 (2.4)	80.1 (3.5)	76.3 (4.6)	75.9 (4.3)	86.9 (3.1)	87.5 (2.1)
RF	75.6 (7.3)	86.8 (3.2)	80.6 (6.3)	83.1 (2.6)	66.7 (5.0)	75.4 (2.1)	81.6 (5.0)	88.2 (4.0)

T.: time domain features; F.: frequency domain features; W.: wavelet domain features; All: all features; Sel.: selected feature set.

- T.: features from time domain;
- F.: features from frequency domain;
- W.: features from wavelet domain;
- T.&F.&W.: features from all domains of time, frequency, and wavelet.

For each feature subset, we evaluate the effect of feature selection using the method proposed in Section 4.4. We have "All" for all the features in a feature subset and "Sel." for the selected features in a feature subset. For each feature subset, four classifiers are used to perform gender recognition. We use the measurement of *accuracy* according to Eq. 3 to evaluate the performance. The average accuracies and the corresponding STDs are shown in Table 3.

Table 3 shows the performance of using samples from multiple smart devices simultaneously, smartphone only, and smart band only for gender recognition. The results show that all the accuracies are above 64.2%, which are better than random guesses. This finding illustrates the motion data collected from smart devices contains useful information for gender recognition. Furthermore, the best accuracy 98.7% is achieved in the scenario of multiple devices for selected all domain features using the RF classifier. The best accuracies for using smartphone or smart band individually are 93.7% and 88.2%, respectively. The scenario of multiple devices also achieves the best result for five out of eight investigated feature subsets. All these results demonstrate that using the data from the smartphone and smart band simultaneously significantly improves the accuracy of gender recognition when compared with using smartphone or smart band individually.

Table 3 shows that the four selected models of the multiple-device scenarios are either the best model of all or the best model for each single channel of T., F., and W. We also explore the feature portfolio of the full space model (T.&F.&W.) compared with that of the selected full space model (selected feature set) in the scenario of multiple devices. In the selected full space model, the number of features incorporated is 96 compared with the total number of 264 for the full space model. About 1/3 of T. features, 2/5 of F. features, and 1/2 of W. features remain after feature selection. This result shows that all the three channels provide important information for gender recognition. We also observe that, although both devices provide a significant number of the selected features, the smart band provides less selected features than the smartphone. Therefore, the motion pattern of hip provides more gender-specific information than that of wrist.

We also perform experiments to explore the feature constituents of the best model for each individual channel. The number of selected features from one channel is the same as the result when selection is performed on the full space. This result may suggest that selected features from different channels are roughly independent and can complement each other in predicting the gender trait of a user.

Comparing the four classifiers shows that the RF performs the best. It achieves the best performance for all three scenarios when using the selected and combined feature subset.

We compare our approaches with the existing methods that have been used for gender recognition in Weiss' and Jain's works (Jain & Kanhangad, 2016; 2018; Weiss & Lockhart, 2011). We implement the previous methods and apply them to our dataset. We achieve the best accuracy of 93.7% using RF compared with 92.7% and 94.4% achieved by using Weiss' and Jain's methods, respectively, for data captured by the smartphone. We achieve the best accuracy of 88.2% using RF compared with 87.5% and 85.4% achieved by using Weiss' and Jain's methods, respectively, for data captured by the smart band. We achieve the best accuracy of 98.7% using RF compared with 97.4% and 97.5% achieved by using Weiss' and Jain's methods, respectively, for data from multiple devices.

On the basis of the results from Table 3, we use the selected features from the three domains (T.&F.&W.) as the feature set and RF as the classifier to investigate the effect on gender recognition with respect to the factors of feature selection and dimension reduction, training sample sizes, training subject sizes, and different sensors.

6.2. The effect of feature selection and dimension reduction

We add the comparison experiments of one-step PCA, two-step PCA, one-step CC, and our approach. One-step PCA uses only PCA for feature reduction. Two-step PCA uses CC for feature selection in the first step and PCA for feature reduction in the second step. One-step CC approach uses only CC to select features. In one-step PCA, we set the number of transformed features from 1(73.2%) to 100(100%) in a step of 9 initially and 10 afterward to investigate the effect of feature sizes on gender recognition. The best threshold is 40(100%). In two-step PCA, we also set the number of transformed features from 1(97.1%) to 100(100%) in a step of 9 initially and 10 afterward to investigate the effect of feature sizes on gender recognition. The best threshold is 20(100%). All the ma-

Table 4
The accuracies (%) of gender recognition using feature selection and dimension reduction.

Classifiers	all features	one-step CC	one-step PCA	two-step PCA	our approach
KNN	86.7	96.1	72.9	76.9	95.9
SVM	82.4	97.8	78.3	90.8	97.5
BP	92.1	96.8	85.8	88.4	97.4
RF	88.0	98.5	80.5	81.1	98.7

nipulations of features are performed on the training set without any knowledge of the testing set. A 10-fold cross validation is performed 10 times on the training datasets. The average accuracies are shown in Table 4.

Table 4 shows that one-step CC outperforms PCA. The relationships with classification labels provide more valuable information to select important features than the relationships within the features. The better performance using two-step PCA than using one-step PCA also verifies this point. Moreover, the results demonstrate that one-step CC can be improved with the second step of feature reduction (dimension reduction) in the proposed approach. Therefore, removal of redundant features will help improve the final results. The results also show our approach outperforms two-step PCA. Thus, our approach of using CC for feature reduction is more appropriate.

6.3. Effect ofsample sizes

In the previous experiment, all 50 samples of motion data from each subject are used for training. In this experiment, we use less than 50 samples from each subject for training and investigate the effect of using different sample sizes for gender recognition. The sample size is related to the difficulty in training a well-performed model.

We set the training sample sizes from 1 to 50 for each subject in a step of 4 initially and 5 afterward to investigate the effect of sample size on gender recognition. A total of 11 scenarios of different training sample sizes are evaluated. A 10-fold cross validation is used to partition the training and testing datasets. Each 10-fold experiment is repeated for 10 times. Each time, we randomly divide the 28 male and 28 female subjects into 10 partitions of roughly equal sizes. Motion data from subjects in 9 partitions are used for training, and the rest is used for testing.

Three scenarios of using motion data from multiple devices, smartphone only, and smart band only are compared. We use the selected features from the three domains (T.&F.&W.) as the feature set and RF as the classifier in the experiment. We use the measurement of *accuracy* according to Eq. (3) to evaluate the performance. The results are averaged for the 100 runs in each of the 11 scenarios, and the average accuracies are shown in Fig. 3.

Fig. 3 shows that the sample size has an evident effect on the performance of gender recognition. The results show that the accuracies increase from above 70% to close to 100% when the training sample sizes change from 1 to 50. Steeper rises of accuracies are observed when the sample sizes are small, but the trends flatten when the sample sizes are over 15 where the accuracies reach above 90%. Thus, we only need a small sample size with 15 samples of motion data to achieve an accuracy of 90% for gender recognition. We can achieve higher accuracies with more samples. Given that a user is taking a smartphone or wearing a smart band for lengthy periods of time, the requirement of sample size is easy to satisfy in real-life applications.

The results also show that, for the three configurations of using motion data from multiple smart devices simultaneously, smartphone only, and smart band only, the multiple-device scenario achieves the best performance for nearly all sample sizes (except for 1). This result suggests the two-device-combined approach

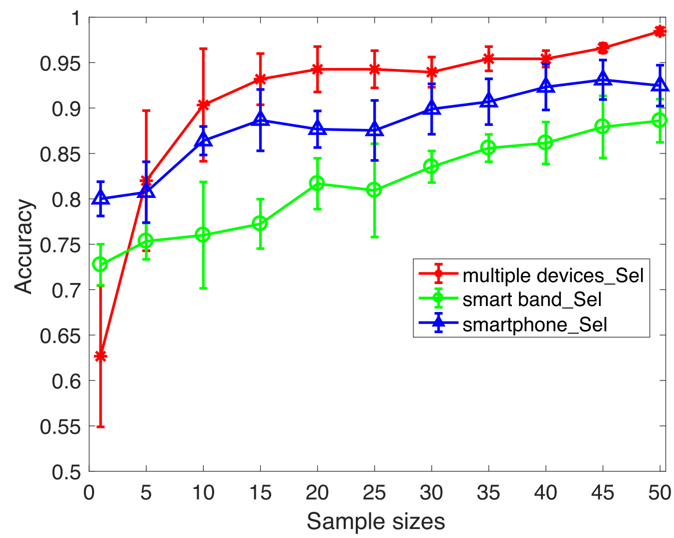

Fig. 3. The performance of gender recognition with different sample sizes.

can retrieve more complete information from the behavior and achieves better accuracies for gender recognition. For example, it can achieve an accuracy of approximately 92% using only 20 training samples.

6.4. Effect of subject sizes

In previous experiments, motion data from 25 male and 25 female subjects are used for training (approximately 90% of 56 subjects), and the remaining 3 male and 3 female subjects are used for testing (nearly 10% of 56 subjects). We design seven scenarios of different subject dataset to further explore the effect of training subject sizes on gender recognition. We change the training subject sizes from 2 (1 M & 1 F) to 54 (27 M & 27 F) in a step of 8 initially and 10 afterward to investigate the effect of subject sizes on gender recognition. For each scenario, M/F is always 50/50, and train+test=56. In each scenario, the subjects used for training are selected randomly from all 56 subjects (28 males and 28 females), and the remaining subjects are used for testing. For each subject, all 50 samples of motion data are used. Three scenarios of using motion data from multiple devices, smartphone only, and smart band only are compared. We use the selected features from the three domains (T.&F.&W.) as the feature set and RF as the classifier in the experiment. We use the measurement of *accuracy* according to Eq. 3 to evaluate the performance. The experiment is repeated for 100 times, and the results are shown in Fig. 4.

Fig. 4 shows the performance of gender recognition using seven different training subject sizes in three scenarios of multiple smart devices, smartphone only, and smart band only. The results show that the training subject size is also an important factor for gender recognition. For all three scenarios, the accuracy of gender recognition increases from above 55% to close to 100% when the training subject size increasings from 2 to 54. These results illustrate that more subjects provide more useful information from motion behavior for gender recognition. We also observe that the accura-

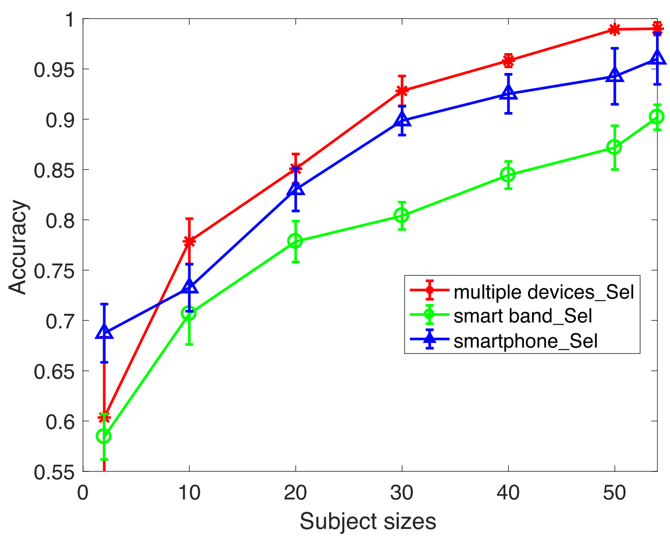

Fig. 4. The performance of gender recognition with different subject sizes.

cies of gender recognition increase fast when the subject sizes are small, but the speed of improvement slows down when the subject size becomes increasingly larger. The results also show the accuracy can reach above 70% only with 10 subjects (5 M & 5 F) for training. Although this accuracy is not very high, it clearly demonstrates significant similarities between behaviors of the same gender. The results show the best accuracy of 98.7% for combined devices using 50 subjects for training. Although the results may be better with more subjects for training, the accuracy of 98.7% is sufficiently good for some real applications.

For the compared approaches of using motion data from multiple devices, smartphone only and smart band only, we notice the approach of using motion data from multiple devices is significantly better than the two other approaches, while the approach of using motion data from smart band only performs the worst. The reason may be largely related to the quality of information retrieved from the motions of different body parts for gender recognition. The smart band retrieves information from the wrist. Wrist movements may be less discriminative between the two genders, and the motion data collected from wrist may be noisier because wrist movements are subject to a larger DOF and closely related to more complex behavioral contexts. By contrast, motion data collected from the pant pocket are more related to the movements of the hip, which are not so flexible and restricted by some physical constraints related to gender characteristics. If we can extract and combine useful information from the two data sources, as we did in the approach of using motion data from multiple devices, then we can reduce noise and generate a complete description of gender differences. This way can lead to a better performance as shown in this experiment.

We also observe that the accuracy of 90% with 30 training subjects (15 M & 15 F) is satisfactory. This result illustrates the proposed method can achieve satisfactory performance without requiring a large amount of subjects for training.

6.5. Effect of different sensors

To evaluate the effect of gender recognition using motion data collected from two sensors, namely, accelerometer and gyroscope, three scenarios are designed: 1) accelerometer: using motion data from accelerometer only; 2) gyroscope: using motion data from gyroscope only; 3) accelerometer and gyroscope: using motion data from accelerometer and gyroscope.

A 10-fold cross validation is used to partition the training and testing datasets. Each 10-fold experiment is repeated for 10 times. Each time, we randomly divide the 28 male (50 samples each) and 28 female subjects (50 samples each) into 10 partitions of roughly equal sizes. Motion data from subjects in 9 partitions are used for training, and the rest is used for testing. We use two feature subsets of selected features (”Sel.”) and all features (”All”) from the three domains (T.&F.&W.) as the feature sets and RF as the classifier in the experiment. Three scenarios of using motion data from multiple devices, smartphone only, and smart band only are compared. We use the measurement of *accuracy* according to Eq. (3) to evaluate the performance. The average accuracies are shown in Table 5.

Table 5 shows the accuracies of using motion data from different sensors of accelerometer, gyroscope, and combined accelerometer and gyroscope: all results are above 76%. These results illustrate that motion data collected from accelerometer and gyroscope can provide useful information for gender recognition.

We also observe that combining motion data from accelerometer and gyroscope achieves the best performance of 98.7% compared with 94.7% and 90.4% using motion data from accelerometer only and gyroscope only, respectively, in the scenario of using multiple devices. In the scenario of using smartphone only, combining motion data from accelerometer and gyroscope achieves the best performance of 93.7% compared with 90.7% and 87.1% using motion data from accelerometer only and gyroscope only, respectively. In the scenario of using smart band only, combining motion data from accelerometer and gyroscope achieves the best performance of 88.2% compared with 86.5% and 83.9% using motion data from accelerometer only and gyroscope only, respectively. These results illustrate that combining motion data collected from multiple sensors provides more comprehensive descriptions of motion behavior than that from one sensor. Motion data collected from accelerometer perform better than those from gyroscope, which means acceleration presents more information for gender recognition than angular velocity during walking. Acceleration describes the changes of walking speed, while angular velocity explains the changes in walking direction during walking.

The performance of using selected feature set is better than using all features in three scenarios of multiple smart devices, smartphone only, and smart band only. This result proves that feature selection technique is a good choice for better performance with low-dimension feature subset. The results also show that using motion data collected from multiple smart devices of smartphone and smart band achieves the best performance Therefore, information from multiple smart devices provides different descriptions about walking behavior that can be combined to predict gender.

The best accuracy of 98.7% is achieved in the scenario of multiple devices by using motion data of accelerometer and gyroscope under the selected feature set.

7. Deployment issues

7.1. Availability of motion data

The availability of the motion data is an important issue in using the proposed solution. We list the following three situations in which the required motion data will be available.

- Applications having legitimate access to the motion data. Few applications, such as fitness and navigation applications, have functions that directly rely on the analysis of the motion data. These applications have ready access to the motion data, such as walking, and can use the proposed solutions to retrieve gender information directly.
- Motion sensor data as a control signal. Sensor data from accelerometer and gyroscope are often used in many games or

Table 5
The accuracies (%) of gender recognition using different sensor data (with STD in parentheses).

Scenarios	Devices	Features	Acc. (%)	Gyro. (%)	Acc. & Gyro. (%)
Multiple devices	smartphone & smart band	Sel.	**94.7** (1.8)	**90.4** (3.1)	**98.7** (2.9)
		All	91.4 (3.2)	80.1 (6.3)	88.0 (5.9)
Single device	smartphone only	Sel.	90.7 (4.4)	87.1 (5.8)	93.7 (2.7)
		All	87.5 (5.3)	76.5 (5.9)	89.4 (4.6)
	smart band only	Sel.	86.5 (4.1)	83.9 (3.2)	88.2 (4.0)
		All	84.2 (4.5)	78.5 (6.0)	81.6 (5.0)

Acc.: Accelerometer; Gyro.: Gyroscope; Sel.: selected feature set; All: all features.

daily applications as a control signal to flip the screen, move an object, or other actions that respond to the position and angle changes of smart devices. A remote website can also use JavaScript to obtain local motion data in the name of control signals. Thus, the technology proposed is applicable if a user opens an application or visits a website when walking.

- Motion data shared to a third party. Sharing health and fitness information is popular in social media. New technologies and services, such as smartwatch and tele-health services (Filipowicz & Keller, 2018; Wikipedia, 2019), promote real-time sharing of sensor data, such as heartbeats and motion data.

7.2. Application scenarios

Gender information is useful for recommendation and adaptive service and can be used as a soft biometric to improve robustness and accuracy of authentication techniques. Thus, these areas are potential application scenarios for the proposed approach. In implementation, smart devices are shared among multiple persons in some cases. Given that the analysis of gender only uses 10s of walking data and requires no other information of the user, our approach can deal with the situation seamlessly. For applications in shops or exhibitions, users are commonly invited to use a special application, which can provide recommendation and navigation services when entering the facility. Thus, motion data will be legitimately accessed by these applications with user consent.

8. Limitations and future work

8.1. Dataset

A limitation of our work is that the subject group does not cover age groups other than from 18 years to 34 years. We choose subjects of young people from our campus, and they are considered convenience samples in some cases (Ross, 1978). However, young people are currently a typical group of people who are highly likely to use smart devices, such as smartphone and smart band(watch) (Chaffey, 2018). Thus, they are justified as an important group of the target users to whom the technology investigated may be applied. In the future study, a more diversified subject pool and dataset will be established and will consist of more people of various ages and backgrounds. We will validate our method on the new dataset and examine the representation of age and other background information on the motion sensor signals.

8.2. Requirement formultiple devices

This study investigates gender recognition in a new scenario where a user wears a smart band and has a smartphone simultaneously. In a practical setting, the two devices may not be on the user's body at the same time. However, the scenario investigated in this study will become increasingly common with the increasing popularity of smart wearables. Moreover, even though a user only uses a single device, our results still show accuracies of 93.7% and 88.2% for smartphone and smart band individually.

8.3. Advanced machine learning approaches and activity scenarios

This study only uses simple classifiers and basic machine learning techniques to explore the problem of gender recognition under a common scenario of walking. The reason is that the goal of this work is to demonstrate that the new setting of multiple smart devices will lead to better performance of gender recognition than previous single-device setting. In the future work, we will explore more advanced machine learning techniques, such as semi-supervised learning and end-to-end learning approaches, to improve the generalization of this work.

Another interesting aspect is the activity scenario. Other than walking, the usage of multiple smart devices is versatile. In our future work, we will investigate the difference degree between genders in the usage of the smart devices and other activities, such as jogging and jumping.

9. Conclusions

In this study, we thoroughly investigate the problem of whether having behavioral data sensed from multiple devices will truly improve the performance of user trait prediction using gender recognition as the specific scenario. We demonstrate that using data from multiple devices will improve gender recognition accuracy by 4%. We achieve the best accuracy of 98.7% in the scenario of using motion data from multiple devices compared with the accuracies of 93.7% and 88.2% using the smartphone and smart band individually, by using motion features extracted from time, frequency, and wavelet domains. These results also illustrate that motion data have important information for gender recognition, and using motion data from multiple devices achieves better performance than that of smartphone only and smart band only. In the experiments, we evaluate the effect of feature selection, sample sizes, subject sizes, and sensors on the performance of gender recognition. The results with feature selection achieve higher accuracies than using all features, which demonstrates the effectiveness of feature selection. Our experiments also show that the sample size has an evident effect on the performance of gender recognition. The accuracies increase from above 75% to 98.7% when the training sample size changes from 1 to 50. The results illustrate that more samples can lead to higher accuracies. The training subject size is an important factor for gender recognition. The accuracy of gender recognition increases from above 70% to 98.7% as it increases from 2 to 54. The results demonstrate that training sample size and subject size largely affect the performance of gender recognition because a larger size can lead to a better description for the uncertainty of human motions. We also show that the combination of motion data from accelerometer and gyroscope can improve the performance of gender recognition.

In the future work, we will collect additional motion dataset from different scenarios, such as different ages and backgrounds, and use additional learning algorithms to obtain more accurate and stable models for gender recognition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Credit authorship contribution statement

Jianmin Dong: Methodology, Writing - original draft. **Youtian Du:** Methodology, Writing - review & editing. **Zhongmin Cai:** Writing - review & editing.

Acknowledgments

This work was supported in part by the National Key R&D Program of China [grant number 2018AAA0101501], National Natural Science Foundation of China [grant number 61772415], and China Scholarship Council [grant number 201706280220].

Appendix A

Lemma 1. *Given a continuously measured variable X and a binary variable Y, the point biserial correlation* r_{pb} *is mathematically equivalent to the Pearson correlation* r_{XY}*, that is,* $r_{pb} = r_{XY}$.

Proof. Assume that the binary variable Y has the two values -1 and 1. If the data set X is divided into two groups: X_1 labeled the value 1 on Y and X_0 labeled the value -1 on Y, we can represent the point-biserial correlation coefficient as

$$r_{pb} = \frac{\bar{X}_1 - \bar{X}_0}{\sigma_X}\sqrt{\frac{n_1 n_0}{n^2}} \tag{4}$$

where $\bar{X}_1$ is the mean value on all data points X_1, $\bar{X}_0$ the mean value on all data points X_0, n_1 the number of data points in X_1, n_0 the number of data points in X_0, n the total sample size of X (obviously, $n = n_0 + n_1$) and σ_X the standard deviation of X. We know the identical equation as

$$n\mathrm{E}[X] = n_0\bar{X}_0 + n_1\bar{X}_1 \tag{5}$$

where $\mathrm{E}[X]$ is the mean value on the continuous variable X.

We can obtain the mean value on the binary variable Y as

$$\mathrm{E}[Y] = \frac{n_1 - n_0}{n}. \tag{6}$$

Further, we have

$$\mathrm{E}[Y^2] = 1. \tag{7}$$

Hence, we calculate the standard deviation of Y as

$$\sigma_Y = \sqrt{\mathrm{E}[Y^2] - (\mathrm{E}[Y])^2} = 2\sqrt{\frac{n_1 n_0}{n^2}}. \tag{8}$$

Then, we calculate the covariance of X and Y as

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])] \\ &= \frac{2n_1 n_0}{n^2}(\bar{X}_1 - \bar{X}_0).\end{aligned} \tag{9}$$

For the aforementioned X and Y, we can write the general form of the Pearson correlation coefficient as

$$r_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}. \tag{10}$$

Here, combining 8,9 and 10, we have

$$r_{XY} = \frac{\bar{X}_1 - \bar{X}_0}{\sigma_X}\sqrt{\frac{n_1 n_0}{n^2}} = r_{pb} \tag{11}$$

Therefore, for the binary classification problem on the data set X labeled by the two-values Y, the point biserial correlation is mathematically equal to the pearson correlation between X and Y. □

References

Bologna, L., Gotti, E., Da Roit, F., Intermesoli, T., Rambaldi, A., Introna, M., & Golay, J. (2013). Ofatumumab is more efficient than rituximab in lysing B chronic lymphocytic leukemia cells in whole blood and in combination with chemotherapy. *The Journal of Immunology, 190*(1), 231–239.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Chaffey, D. (2018). Mobile marketing statistics compilation. https://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/.

Chang, C.-C., & Lin, C.-J. (2011). Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST), 2*(3), 27.

Chen, Y., & Shen, C. (2017). Performance analysis of smartphone-sensor behavior for human activity recognition. *IEEE Access, 5*, 3095–3110.

Cohen, P., West, S. G., & Aiken, L. S. (2014). *Applied multiple regression/correlation analysis for the behavioral sciences*. Psychology Press.

Deitrick, W., Miller, Z., Valyou, B., Dickinson, B., Munson, T., & Hu, W. (2012). Author gender prediction in an email stream using neural networks. *Journal of Intelligent Learning Systems and Applications, 4*, 169–175. doi:10.4236/jilsa.2012.43017.

Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. A Wiley-Interscience Publication, New York: Wiley, 1973.

Estévez, P. A., Tesmer, M., Perez, C. A., & Zurada, J. M. (2009). Normalized mutual information feature selection. *IEEE Transactions on Neural Networks, 20*(2), 189–201.

Figueroa, A. (2017). Male or female: What traits characterize questions prompted by each gender in community question answering? *Expert Systems with Applications, 90*, 405–413. doi:10.1016/j.eswa.2017.08.037.

Filipowicz, L., & Keller, J. (2018). How to send someone your heartbeat with apple watch or iphone. https://www.imore.com/how-send-someone-your-heartbeat-apple-watch-iphone.

Golomb, B. A., Lawrence, D. T., & Sejnowski, T. J. (1990). Sexnet: A neural network identifies sex from human faces.. In *Proceedings of the Nips: 1* (p. 2).

H, S. P., & Omman, B. (2014). Feature selection techniques for gender prediction from blogs. In *Proceedings of the first international conference on networks soft computing (ICNSC2014)* (pp. 355–359). doi:10.1109/CNSC.2014.6906657.

Harb, H., & Chen, L. (2003). Gender identification using a general audio classifier. In *Proceedings of the international conference on multimedia and expo, ICME: 2* (pp. II–733). IEEE.

Hu, J. (2018). An approach to eeg-based gender recognition using entropy measurement methods. *Knowledge-Based Systems, 140*, 134–141.

Jain, A., & Kanhangad, V. (2016). Investigating gender recognition in smart-phones using accelerometer and gyroscope sensor readings. In *Proceedings of the international conference on computational techniques in information and communication technologies (icctict)* (pp. 597–602).

Jain, A., & Kanhangad, V. (2018). Gender classification in smartphones using gait information. *Expert Systems with Applications, 93*, 257–266.

Jain, A. K., Dass, S. C., & Nandakumar, K. (2004). Soft biometric traits for personal recognition systems. In *Biometric authentication* (pp. 731–738). Springer.

Jianle Chen, Tianqi Xiao, Jie Sheng, & Teredesai, A. (2017). Gender prediction on a real life blog data set using LSI and KNN. In *Proceedings of the IEEE 7th annual computing and communication workshop and conference (CCWC)* (pp. 1–6). doi:10.1109/CCWC.2017.7868410.

Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter, 12*(2), 74–82.

Kwon, Y., Kang, K., & Bae, C. (2014). Unsupervised learning for human activity recognition using smartphone sensors. *Expert Systems with Applications, 41*(14), 6067–6074.

Li, X., Maybank, S. J., Yan, S., Tao, D., & Xu, D. (2008). Gait components and their application to gender recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C, 38*(2), 145–155.

Mostayed, A., Kim, S., Mazumder, M. M. G., & Park, S. J. (2008). Foot step based person identification using histogram similarity and wavelet decomposition. In *Proceedings of the international conference on information security and assurance, ISA* (pp. 307–311). IEEE.

Mukaka, M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal, 24*(3), 69–71.

Nguyen, D.-P., Trieschnigg, R., Dogruoz, A., Gravel, R., Theune, M., Meder, T., & de Jong, F. (2014a). Why gender and age prediction from tweets is hard: Lessons from a crowdsourcing experiment. In *Proceedings of the 25th international conference on computational linguistics, coling* (pp. 1950–1961). United States: Association for Computational Linguistics (ACL).

Nguyen, D. T., Cho, S. R., Shin, K. Y., Bang, J. W., & Park, K. R. (2014b). Comparative study of human age estimation with or without preclassification of gender and facial expression. *The Scientific World Journal, 2014*, 1–15.

Nunes, M. A. S. N. (2008). *Recommender systems based on personality traits*. Université Montpellier II-Sciences et Techniques du Languedoc Ph.D. thesis..

Pohjalainen, J., Räsänen, O., & Kadioglu, S. (2015). Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits. *Computer Speech & Language, 29*(1), 145–171.

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

Ronao, C. A., & Cho, S.-B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems with Applications, 59*, 235–244.

Ross, K. N. (1978). *Sample design for educational survey research*. Pergamon Press Oxford.

Schölkopf, B., Williamson, R. C., Smola, A. J., Shawe-Taylor, J., & Platt, J. C. (2000). Support vector method for novelty detection. In *Advances in neural information processing systems* (pp. 582–588).

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., … Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLOS ONE, 8*(9), 1–16. doi:10.1371/journal.pone.0073791.

Shen, C., Cai, Z., Maxion, R. A., & Guan, X. (2013). On user interaction behavior as evidence for computer forensic analysis. In *Proceedings of the international workshop on digital watermarking* (pp. 221–231). Springer.

Shen, C., Zhang, Y., Guan, X., & Maxion, R. A. (2016). Performance analysis of touch-interaction behavior for active smartphone authentication. *IEEE Transactions on Information Forensics and Security, 11*(3), 498–513.

Wang, N., Ambikairajah, E., Lovell, N. H., & Celler, B. G. (2007a). Accelerometry based classification of walking patterns using time-frequency analysis. In *Proceedings of the 29th annual international conference of the IEEE engineering in medicine and biology society* (pp. 4899–4902). IEEE.

Wang, S., Liu, C.-L., & Zheng, L. (2007b). Feature selection by combining fisher criterion and principal feature analysis. In *Proceedings of the international conference on machine learning and cybernetics: 2* (pp. 1149–1154). IEEE.

Weiss, G. M., & Lockhart, J. W. (2011). Identifying user traits by mining smart phone accelerometer data. In *Proceedings of the fifth international workshop on knowledge discovery from sensor data* (pp. 61–69). ACM.

Wikipedia (2019). Telehealth. https://en.wikipedia.org/wiki/Telehealth.

Yoon, H., Park, S.-H., & Lee, K.-T. (2016). Lightful user interaction on smart wearables. *Personal and Ubiquitous Computing, 20*(6), 973–984.

Yu, S., Tan, T., Huang, K., Jia, K., & Wu, X. (2009). A study on gait-based gender classification. *IEEE Transactions on Image Processing, 18*(8), 1905–1910.