

Article

Impact of Sliding Window Length in Indoor Human Motion Modes and Pose Pattern Recognition Based on Smartphone Sensors

Gaojing Wang ¹ , Qingquan Li ^{1,*} , Lei Wang ¹ , Wei Wang ², Mengqi Wu ² and Tao Liu ³ 

¹ State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; wanggaojing@whu.edu.cn

² School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China; wanngwei@whu.edu.cn (W.W.); mengqi.wu@whu.edu.cn (M.W.)

³ College of Resources and Environment, Henan University of Economics and Law, Zhengzhou 450002, China; liuzimo@whu.edu.cn

* Correspondence: liqq@szu.edu.cn (Q.L.); lei.wang@whu.edu.cn (L.W.)

Received: 7 May 2018; Accepted: 15 June 2018; Published: 18 June 2018



Abstract: Human activity recognition (HAR) is essential for understanding people's habits and behaviors, providing an important data source for precise marketing and research in psychology and sociology. Different approaches have been proposed and applied to HAR. Data segmentation using a sliding window is a basic step during the HAR procedure, wherein the window length directly affects recognition performance. However, the window length is generally randomly selected without systematic study. In this study, we examined the impact of window length on smartphone sensor-based human motion and pose pattern recognition. With data collected from smartphone sensors, we tested a range of window lengths on five popular machine-learning methods: decision tree, support vector machine, K-nearest neighbor, Gaussian naïve Bayesian, and adaptive boosting. From the results, we provide recommendations for choosing the appropriate window length. Results corroborate that the influence of window length on the recognition of motion modes is significant but largely limited to pose pattern recognition. For motion mode recognition, a window length between 2.5–3.5 s can provide an optimal tradeoff between recognition performance and speed. Adaptive boosting outperformed the other methods. For pose pattern recognition, 0.5 s was enough to obtain a satisfactory result. In addition, all of the tested methods performed well.

Keywords: human motion mode; human pose pattern; window length; machine-learning method; smartphone sensors

1. Introduction

Human activity recognition (HAR) has become a popular research topic. Analyzing human activities is an effective method for understanding the human context, living habits, and demands [1–14]. HAR can be used in many applications, such as precise marketing and human psychology. Scholars regard human motion, such as walking, being at rest, and riding an elevator, and posing, which includes activities such as calling and typing, as two highly interesting types of human activity [15,16]. These activities are particularly important for pedestrian navigation applications [17], because they support the robustness and accuracy of the navigation. Varying motion modes and pose patterns require different algorithms and constraints to obtain accurate positioning results [18]. For instance, when walking is detected, users' vertical locations should be fixed, whereas horizontal displacement and direction must be updated. When riding an elevator is detected, the horizontal location should be fixed, whereas the vertical location must be updated. When using an escalator is detected, horizontal

and vertical displacements should be updated. Moreover, the models of misalignment estimation (i.e., differentiating between pedestrian heading and smartphone orientation) differ for each motion and pose [19]. Therefore, awareness of user motion modes and pose patterns can determine the correct misalignment estimation model, and potentially improve positioning solutions. Additionally, the optimal type of sensor during positioning varies for each human pose [20]. For instance, a gyroscope is an optimal sensor for pedestrian dead reckoning (PDR) when users carry their smartphone in a trouser pocket. However, an accelerometer is an excellent option for phoning and typing recognition. HAR also provides guidance measures for patient treatment, and has thus attracted increased attention in the medical treatment field [21].

Considerable research has been conducted on human motion modes and pose pattern recognition. Researchers initially focused on the combinations of motion modes for motion mode recognition. Yang et al. [22] considered four motion modes: sitting, standing, walking, and running. Prasertsung et al. [23] focused on rising and falling modes, involving stairs. Choudhary et al. [24] proposed the vertical motion mode by presenting an elevator case; Bao et al. [25] focused on the complicated motion mode of riding escalators; and Elhoushi et al. [17] introduced the detection of walking transitioning to the escalator motion mode. However, much of the early research on motion modes focused on extracting motion information from wearable motion sensors that were attached to certain body parts [26–31]. This research field has leveraged the emergence of smartphones, because such devices are equipped with powerful micro-processing units and high-quality, versatile sensors. The smartphone is more acceptable for users than wearable sensors, because smartphones can operate as a multi-purpose personal assistant, whereas wearable sensors only meet specific demands [32,33]. Therefore, this study focused on smartphone sensor-based motion modes and pose pattern recognition.

Machine-learning methods are typically adopted as classifiers in the research on human motion and pose recognition [21–23]. When used to detect human motion modes and pose patterns, all sensor data should first be segmented using a windowing method and classified at every segmentation. Therefore, selecting the window length directly affects classification performance. A small window accelerates recognition but may negatively impact recognition performance. As such, the tradeoff between recognition performance and latency must be carefully considered during the algorithm design. User requirements should also be considered during this phase. Several applications aimed for excellent classification performance regardless of speed (e.g., counting the number of steps in a day [34,35]), whereas others aimed to create speed-critical applications (e.g., real-time positioning). Therefore, understanding the effect of window length on human motion modes and pose pattern recognition can help select the suitable window length during the algorithm design to meet specific user requirements.

In previous research, several techniques for data segmentation that divide sensor signals into usable small parts have been developed [36–41]. Among the known techniques, the sliding window approach is the most widely employed [42–45], being regarded as the best approach for research given its simplicity and stability, and a wide range of window lengths has been used in past studies. Windows as short as 0.5 s and 0.8 s were used to recognize walking, jogging, and going up or down the stairs [46], whereas a window of 1 s with a decision tree (DT) [47] was used to classify stationary, walking, running, and biking motion modes. Additionally, with a neural network [48], a window of 2 s was adopted to classify the motion modes of walking, upstairs and downstairs movement, running, and sitting with varying poses. An average accuracy of 93% was achieved. A window of 5 s was used to classify walking, standing, and climbing stairs using handheld smartphones, where multiple methods achieved a high accuracy score of 84% [49]. A window of 7.5 s was adopted in recognizing walking, stationary, running, and cycling for cases where the smartphone is inside the trouser pocket of users [50], and a classification accuracy score of 93.9% was achieved based on the K-nearest neighbor (KNN) machine learning.

In this paper, we initially examined the influence of window length on human motion modes and pose pattern recognition using five popular machine-learning algorithms. Subsequently, we examined

the performance of human motion modes and pose pattern recognition based on different window lengths and machine-learning algorithms using smartphone sensor data. Lastly, the suitable window length for human motion modes and pose pattern recognition is recommended.

The rest of this paper is constructed as follows. Section 2 outlines the methods used for activity classification methodologies. Section 3 describes the experimental setup. Section 4 analyzes the results, which are discussed in Section 5. Section 6 presents the limitations of the study and concludes this paper.

2. HAR Workflow

Generally, HAR includes four steps: data preprocessing, segmentation, feature extraction, and classification (Figure 1). Sensors can provide multiple data streams for use as data input, such as raw acceleration and air pressure.

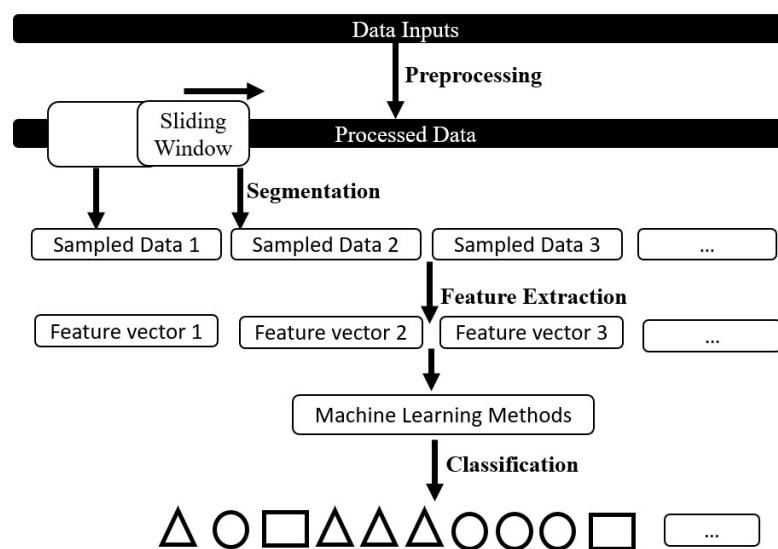


Figure 1. Overview of human activity recognition (HAR) workflow.

The subsequent feature extraction process determines useful features and distinguishes the activities. Feature extraction requires data segments for use as the input data. Thus, raw data streams should be cut into segments. The sliding window segmentation algorithm has been widely used to split sensor data and maximize data usage. Feature extraction is then performed on the data segments. Time-domain statistical and frequency-domain features [23,51,52] are conventionally used as feature input.

The key step in HAR is classification, which takes advantage of the extracted features. Machine-learning methods that can explore unique patterns for classification are popularly used in motion modes and pose pattern recognition. In this study, five machine-learning methods are examined: support vector machine (SVM), KNN, decision tree (DT), Gaussian naïve Bayesian (GNB), and adaptive boosting (Adaboost). The machine-learning methods used in this study are briefly introduced as follows.

The SVM theory was proposed by Vapnik and Chervonenkis [53]. The effectiveness of SVM has been proven to be effective at addressing many problems, such as handwritten digit recognition, face detection in images, and text categorization. SVM achieves high classification accuracy and is robust to noisy data and overfitting problems. Therefore, SVM is considered one of the top classifiers in terms of generalization, and is a popular machine-learning approach in HAR [54,55].

KNN groups feature vectors into clusters that represent different classes [56]. For KNN, the parameter k can be used to regulate underfitting and overfitting. Reducing the value of k

increases the sensitivity of the classifier to training data noise, but makes the classifier prone to overfit. Susi et al. [57] achieved accuracy rates ranging from 80% to 84% for upstairs and downstairs movement with $k = 1$. KNN has also been used in other studies [58,59].

A DT solves a classification problem through a series of cascading decision questions. A feature vector, which satisfies a specific set of questions, is assigned to a specific class. This method is represented graphically using a tree structure, where each internal node is a test on a feature compared with the threshold, and the remaining values refer to the decided classes. Its implementation is based on a loop of if/else conditions. Many types of DTs are generated by different algorithms. In our research, C4.5 was adopted. DTs have been used widely by researchers, many of whom agree that it provides highly accurate results [57].

The naïve Bayesian classifier determines the probability of an event that belongs to a certain class based on Bayesian theorem through using a naïve method [60], assuming that all of the input features are independent. When dealing with continuous data, a typical assumption is that the continuous values associated with each class are distributed according to Gaussian distribution. In such cases, naïve Bayesian (NB) is also called GNB. Compared with other algorithms, the NB classifier is profoundly easy to implement for training and evaluation algorithms. However, this simplicity leads to a much lower accuracy than that of many other classifiers. NB has obtained accuracy rates ranging from 68% to 72% for upstairs and downstairs motion modes, respectively, and 89% to 93% for walking and running, respectively [57].

The adaptive boosting (Adaboost) is a machine-learning meta-algorithm formulated by Yoav Freund and Robert Schapire [61]. It is a method that can be used with other machine-learning methods to improve recognition accuracy. Adaboost combines the outputs of plenty of “weak” classifiers into a weighted sum that represents the final output. AdaBoost is adaptive in the sense that subsequent weak learners are tweaked in favor of those instances misclassified by previous classifiers. The individual learners can be weak, but as long as the performance of each one is slightly better than random guessing, the final model can be proven to converge to a strong learner. It has been proved effective in HAR in previous researches [62,63]. In this paper, Adaboost emended with decision tree (C4.5) was adopted.

In our research, these methods were used to generate robust recognition results. In the training phase, training data streams were segmented with a fixed-length sliding window. Subsequently, features were extracted from data segmentations with equal lengths and fed into the classifiers. When the classifiers are trained and deployed for activity recognition, sensor data streams that must be identified should be sectioned into data segments with the same length as those used for the training data to ensure the effectiveness of the trained classifiers. Understanding the impact of window length on activity recognition can help determine the appropriate window length and present the classification result without bias.

3. Experiment Setup

An extensive experiment was performed in our research. The experiment is described in this section.

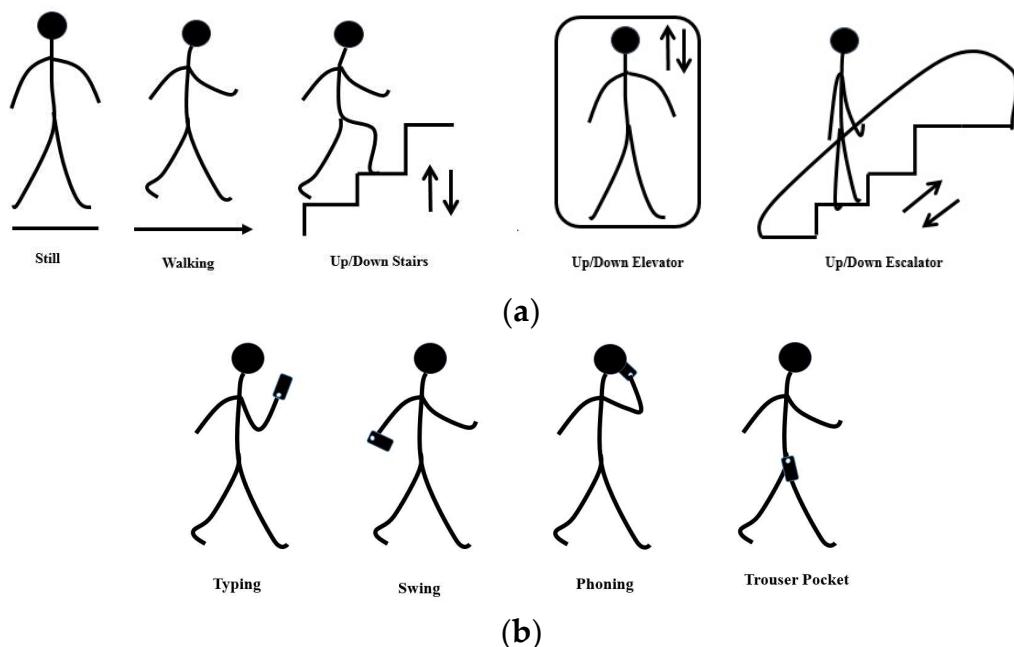
3.1. Data Acquisition

We collected a dataset with smartphones in a typical shopping center with nine floors equipped with escalators and elevators. We recruited 10 subjects for the data collection, whose heights ranged from 163 cm to 180 cm and weights ranged from 50 kg to 80 kg. Among the 10 subjects, seven were male and three (subjects 1–3) were female, aged 20 to 30 years old. To protect the privacy and personal information of the subjects, we only show the approximate range of their height and weight in Table 1.

Table 1. Subject information.

Weight (kg)	Height (cm)	[163,170]	[170,175]	[175,180]
[50,60)	Subject 1, 3			
[60,70)	Subject 8	Subject 2		Subject 4
[70,80)		Subject 9, 10	Subject 5, 6, 7	

In daily life, people may use their smartphone in different poses. To increase the robustness of our estimation, and in contrast to previous research, in which smartphones were fixed in one pose [23,51], we considered smartphone usage poses and motion modes when we collected the sensor data. Through observation, we considered eight common indoor pedestrian motion modes and four smartphone usage poses, as shown in Figure 2a,b. Based on these activities and previous works [18,64–66], accelerometer, barometer, and gravity sensors equipped in the smartphone were chosen in our research. The sensor data of every motion mode under each pose, and every pose under each motion mode, were collected. Notably, the sensor data of the dynamic motion modes (walking and going up and down stairs) were collected with three walking speeds: slow, normal, and fast. In addition, left-hand and right-hand usage was considered under each pose. Specifically, the left and right trouser pockets were considered in such poses. The data collection campaign lasted over one week, and 21 h of valid test data were collected. Figure 3 demonstrates the data collection scenarios. To protect the privacy, faces of the subjects in Figure 3 were covered with mosaics.

**Figure 2.** (a) Human motion modes and (b) pose patterns covered in this study.

In our experimental settings, walking distance was approximately 500 m, the escalator and stairs descended and ascended between the first and 10th floors, and the elevator covered from the first to the 26th floor. The subjects were blinded to the purpose of the experiment during data collection, and were thus allowed unrestricted smartphone usage to guarantee a natural performance [15,17,67–69].

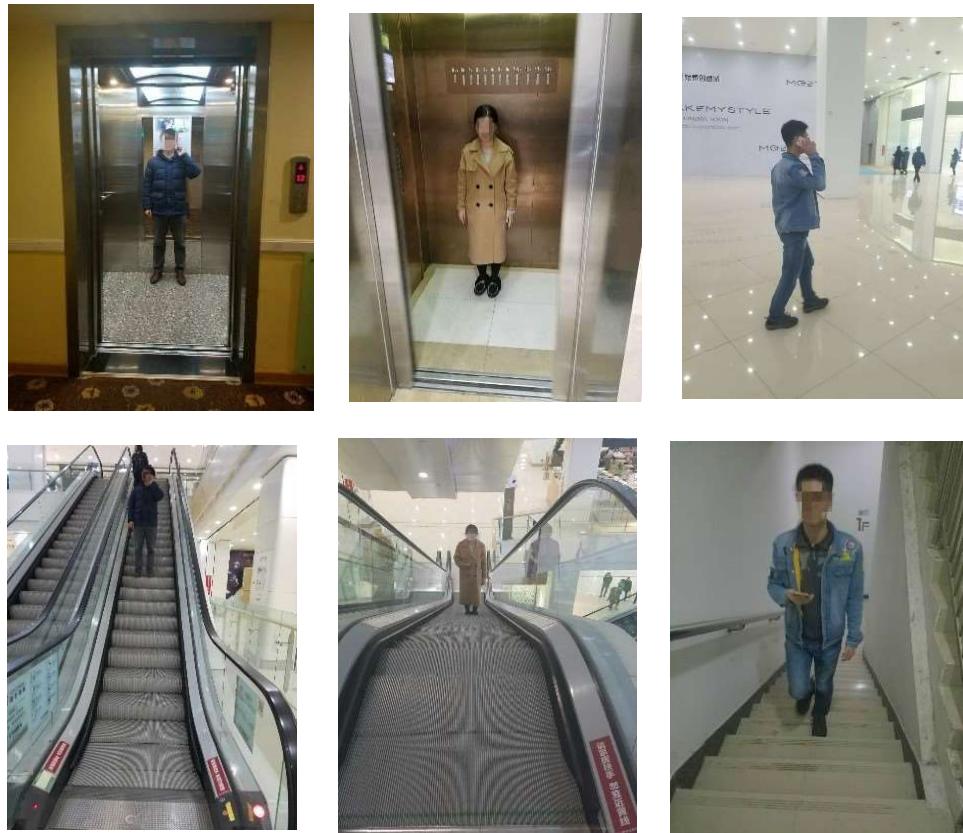


Figure 3. Typical data collection scene. **Upper row:** elevators 1 and 2 and walking. **Lower row:** escalators 1 and 2 and going upstairs (from left to right).

3.2. Adopted Sensors and Features

We used acceleration magnitude instead of vector to avoid the negative influence of smartphone orientation on motion mode recognition [23,70]. For pose pattern recognition, gravity sensor data were used to reduce the influence of various motion modes [71]. Actually, the gravity sensor is not a real sensor, but it is obtained by processing data provided by the accelerometer and gyroscope [72]. Hence, motion modes and pose pattern recognition were performed separately and simultaneously.

Table 2 presents a detailed description of the smartphone sensors and measurement types used in this study. In our experiment, we directly used the raw data stream collected at 50 Hz without any specific preprocessing to avoid relevant information loss. During segmentation, a window length ranging from 0.5 s to 7 s with an interval of 0.5 s was adopted for analysis, and the sliding overlap of 0.5 s was used for window sizes larger than 0.5 s. This range largely covers the window lengths used in previous research.

In this study, time-domain and frequency-domain features, which were normally adopted in previous studies [23,51,52], were used for classification (Table 3). These features were extracted for each segment after windowing. Notably, human actions belong to the low-frequency domain, and fast Fourier transformation (FFT) calculations are time-consuming. Thus, we adopted the second to ninth FFT coefficients as the frequency-domain features [52]. We forewent the first FFT coefficient because it represents a direct component, which is similar to the mean value of the sequence. Features were then extracted for every data stream for pattern recognition, as shown in Table 3.

Table 2. Description of adopted smartphone sensors.

Sensors	Purpose	Data Stream	Description	Manufacturer	Measuring Range	Measuring Accuracy
Gravity sensor	Pose pattern classification	G_x	Gravity force along x axis	Qualcomm	39.226593 m/s ²	0.00119 m/s ²
		G_y	Gravity force along y axis			
		G_z	Gravity force along z axis			
Accelerometer	Motion mode classification	$A = \sqrt{A_x^2 + A_y^2 + A_z^2}$		A_* is the specific force along * axis		
Barometer		P	Air pressure measurement	BOSCH	1100 hPa	0.00999 hPa

Table 3. Feature set.

No.#	FEATURE	DEFINITION
1	Mean	$mean(x) = \bar{x} = \frac{1}{N} \sum_{n=1}^N x[n]$
2	Absolute Mean	$mean(x) = \frac{ x }{ x }$
3	Variance	$var(x) = \sigma_x^2 = \frac{(x - \bar{x})^2}{N}$
4	Standard deviation	$std(x) = \sigma^x = \sqrt{var(x)}$
5	Mode	Values that appear most frequently in data set
6	Median	Middle value in a data set
7	Average Absolute Difference	$mean(x - \bar{x}) = x - \bar{x} $
8	75th Percentile	Value separating 25% higher data from 75% lower data in a data set.
9	Interquartile range	Difference between 75th and 25th percentile
10	Gradient (only for air pressure data)	The coefficient of first-order linear fitting
11	Coefficients of FFT (Fast Fourier Transform)	Energy of each frequency component

3.3. Performance Metric

The F1 score was used as a performance metric to gauge classification performance. It is a combination of precision and recall measures that can represent the detection result with less bias than the accuracy in multi-class classification problems, especially with disproportionate samples in each class [67]. Suppose that in classifying classes A and B, we obtained a confusion matrix (Table 4).

Table 4. Confusion matrix in classifying class A.

		Predicted Class	
		A	B
Actual Class	A	TP	FN
	B	FP	TN

In the matrix, true-positive (TP) is the number of observations that are positive and were predicted to be positive, false-negative (FN) is the number of observations that are positive but were predicted to be negative, true-negative (TN) is the number of observations that are negative and were predicted to be negative, and false-positive (FP) is the number of observations that are negative but were predicted to be positive. Precision, recall, and F1 score are defined as follows:

$$\begin{aligned}
 precision &= \frac{TP}{TP+FP} = \frac{\text{positive predicted correctly}}{\text{all positive predictions}} \\
 recall &= \frac{TP}{TP+FN} = \frac{\text{positive predicted correctly}}{\text{all positive observations}} \\
 F1 &= 2 \cdot \frac{precision \times recall}{precision + recall}
 \end{aligned} \tag{1}$$

A high F1 score indicates a high level of classification performance and agreement between the classification and true value.

3.4. Validation and Testing Strategy

Ten-fold cross-validation, leave-one-subject-out cross-validation (LOOCV), and bootstrapping strategies have been used in the literature. As previously summarized [73–75], LOOCV and bootstrapping are better for risk estimation, whereas 10-fold CV is the most accurate approach for model selection. Chen et al. [52] reported that, in contrast to rest-to-one mode in LOOCV, the all-to-one model better enhances robustness and is recommended for HAR. Therefore, in our study, 10-fold cross-validation was first used to select the machine-learning method parameters [73]. Parameters with higher average F1 scores in the cross-validation were selected. As for SVM, the linear kernel function and radial basis kernel function (RBF) were adopted in parameter selection for their popularity in HAR [52,53]. The parameter searching of k in KNN was performed in a wide range from 1 to 10. The searching range of the number of embedding decision trees in Adaboost was set from 10 to 100.

Finally, the linear kernel and a parameter of two were selected for SVM and KNN, respectively, and 20 decision trees were incorporated in Adaboost. Subsequently, a 100-time bootstrapping strategy was adopted to ensure statistical robustness and produce an asymptotic convergence to the correct estimation of system performance [73].

We list some of the bootstrapping distribution results in Figure 4 for brevity. These results were derived from the SVM motion mode classification on window sizes ranging from 0.5 s to 3 s. The figure shows the normal distribution fitting curves (black) based on the mean and SD of the bootstrapping results. The results are consistent with normal distribution, and the SD of the 100-time bootstrapping results was less than 1%. In our research, the maximum SD in motion modes and pose pattern recognition was 0.59%. Such a small variation denotes the reliability of our results. The mean values of 100-time bootstrapping results were used as the final results. Section 4 presents the results of motion mode and pose pattern classification.

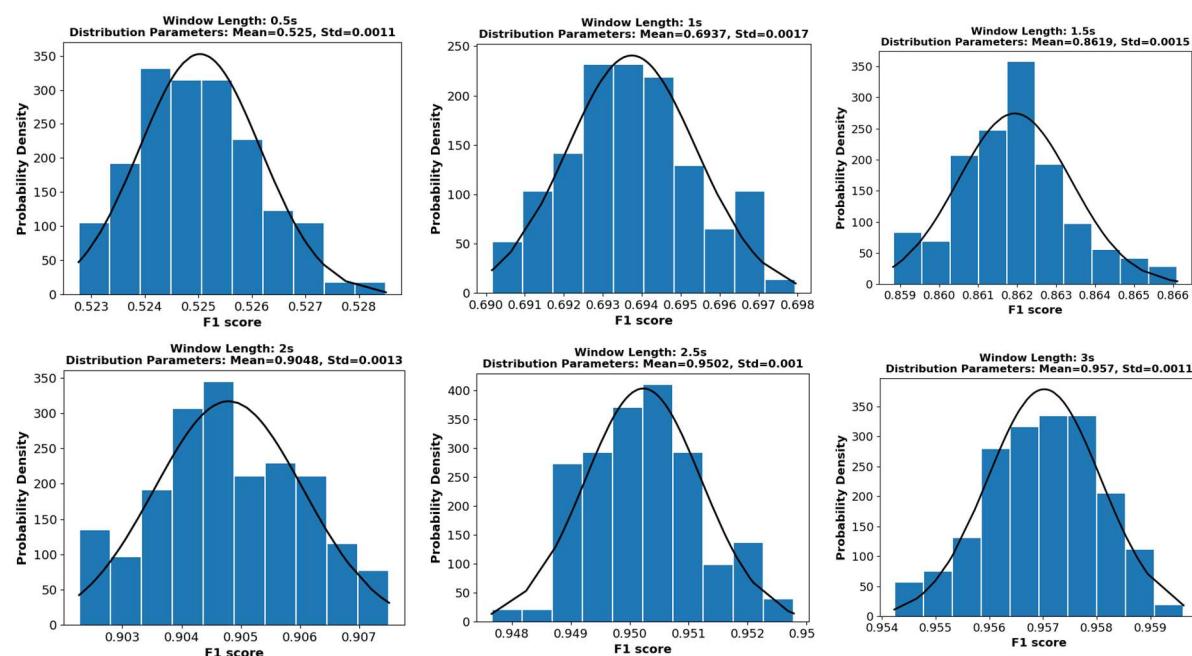


Figure 4. The 100-time bootstrapping results of motion mode recognition using support vector machine (SVM) and different window lengths (0.5–3 s).

4. Experiment Result and Analysis

4.1. Motion Mode Classification Result

First, we examined the influence of window length on the feature extraction of motion mode recognition. We visualized the compressed features extracted at different window lengths on a two-dimensional plane based on principal component analysis [76]. Figure 5 presents the results. The figure shows that data segments with long window lengths indicate feature separability. Points with the same color are increasingly concentrated with the increase in window length, and the boundaries between the features become evident. This notion is particularly true for walking (red) and going up and down stairs (black and green). When the window length is only one second, these point groups were lumped together. When the window length increased to three seconds, the boundaries between cases emerged. These results reveal the significant effect of window length on the classification of human motion modes. In addition, the linear boundaries among various point groups also prove the good performance of SVM with linear kernel.

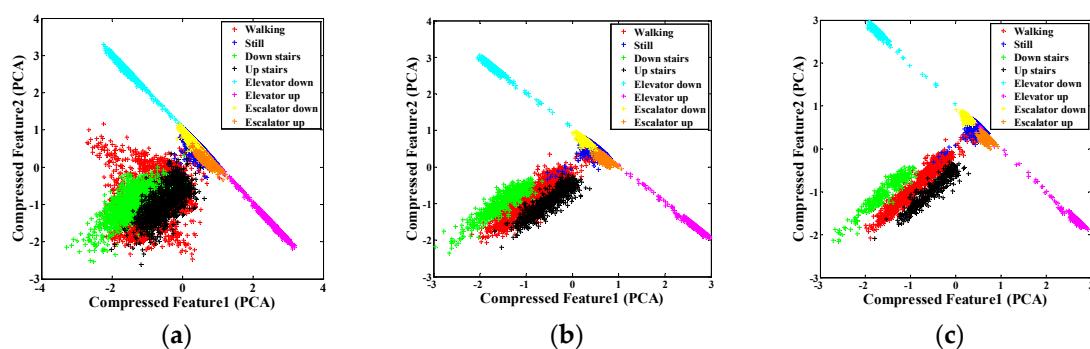


Figure 5. Distribution of compressed features of human motion modes with various window lengths: (a) 1 s; (b) 2 s; and (c) 3 s.

4.1.1. Global Evaluation

We then evaluated the influence of window length on motion mode recognition using the classification results. Figure 6 presents the results. The figure shows the average F1 score of the eight motion modes with window lengths varying from 0.5 s to 7 s using the five classification methods. Initially, we observed that the F1 score considerably increased with the expansion in window length. The SVM F1 score increased significantly from 52.5% to 98% when the window length increased from 0.5 s to 3.5 s. The F1 score improvement using other classification methods was also significant at 66.37% to 98% for DT, 70.14% to 98.12% for KNN, 56.34% to 93.18% for GNB, and 73% to 98.49% for Adaboost (“ABOOST”). These results prove that a notable improvement in classification performance using an increased window length occurs across all of the adopted methods.

Despite the evident benefit of expanding the window length, blindly increasing the window size to improve performance is unreasonable, because the additional benefit of expanding the window length is evidently reduced with the increased in window size. For SVM, the F1 scores increased by less than 1.5% after 3.5 s. To a lesser extent, this result applies to the DT, KNN, GNB, and Adaboost models. In addition, F1 scores decreased when the window length exceeds six seconds.

In a real application, a large window length leads to large recognition latency. However, when we moderately reduced our performance requirement, recognition latency evidently decreased with a minimal tradeoff with recognition performance. If we require an F1 score above 99%, then a sliding window larger than four seconds with SVM, KNN, DT, or Adaboost is satisfactory. However, if the required F1 score decreased to 95%, the window length can be reduced to 2.5 s. If we further lower the required F1 score to 90%, then a window length of two seconds is satisfactory.

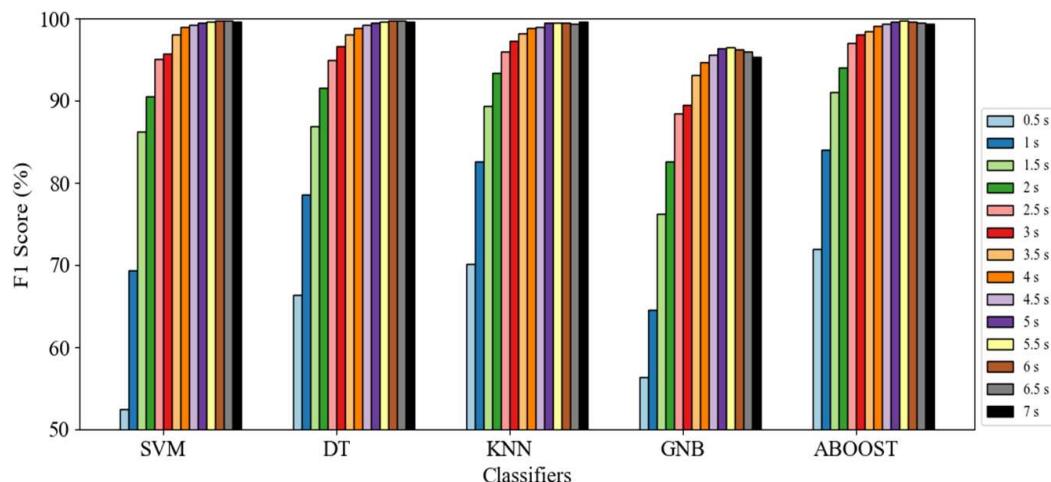


Figure 6. Average F1 score of motion mode classification using different machine-learning methods and window lengths.

In summary, the motion mode classification results show that the impact of the sliding window length is obvious, with the difference between the F1 scores based on different window lengths being larger than 40%, regardless of the adopted machine-learning method. Performance generally improved with greater window lengths. However, the improvement became increasingly smaller and a cut-off window size emerged, after which the improvement was negligible. Based on the result, a window length between 2.5–3.5 s was proven to be the optimum value, given the tradeoff between recognition performance and speed, so this length is recommended for real-time applications with low latency requirements. As for applications that emphasize recognition performance, a window of six seconds is recommended.

4.1.2. Motion Mode-Specific Analysis

In addition to multiple motion mode recognition, specific motion mode recognition may also be required. Therefore, the impact of sliding window length on the specific motion mode was also examined. Figure 7 depicts the recognition results of a specific motion mode with different window lengths.

Figure 7 shows the impact of window length on every motion mode. The F1 score increased by expanding the window size by approximately 20% for still and walking detection, 5% for up and down elevator detection, 40% for up and down stairs detection, and notably, 60% for up and down escalator detection. The enhancement occurred with all of the methods.

In addition, the improvement caused by expanding the window size for specific motion mode detection became less distinct with the increase in window length, which is similar to the result obtained from the overall performance result analysis. However, different cut-off points for enhancement were observed for different motion modes. For instance, the main benefit of expanding the window size in up and down elevator detection occurred at a window size of 0.5 s to 1.5 s, with improvements of less than 0.5% after 1.5 s. However, the same was true after 3.5 s for still and working detection. Therefore, the suitable windows differ according to motion modes for users who are mainly concerned about specific motion mode recognition. Based on Figure 6, Table 5 summarizes the recommended window length for specific motion mode recognition.

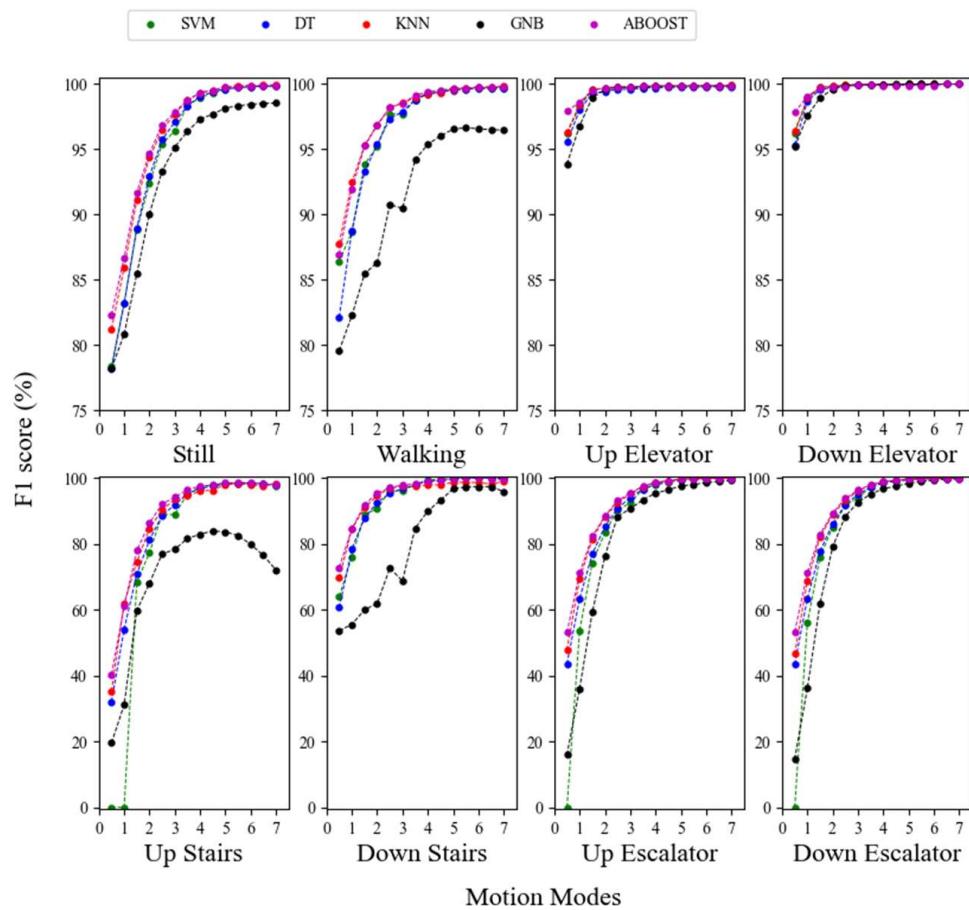


Figure 7. Relationship between F1 score and window length (horizontal axis) for different motion modes.

Table 5. Recommended window size for specific motion mode classification.

Motion Mode	Recommended Window Size			
	F1 Score			
	85%	90%	95%	99%
Stationary	1.5 s	2 s	3 s	4.5 s
Walking	1 s	1.5 s	3 s	4 s
Up elevator	0.5 s	0.5 s	0.5 s	1.5 s
Down elevator	0.5 s	0.5 s	0.5 s	1.5 s
Up stairs	2 s	3 s	3.5 s	5 s
Down stairs	1.5 s	2 s	2.5 s	4 s
Up escalator	2 s	2.5 s	3.5 s	4.5 s
Down escalator	2 s	2.5 s	3 s	4.5 s

As is shown in Figures 6 and 7, a subtle reduction on performance requirement can generally allow us to evidently shorten the needed window length. This will be important for applications that require rapid detection, such as fall detection or indoor positioning [77]. Moreover, there are also other applications that further emphasize recognition performance such as an analysis of people's movement in a whole day or counting the number of step in a day. Based on these different application needs, we listed the recommended window sizes that can guarantee different performance requirements (F1 score of 85% to 99%) in Table 5.

Table 5 also shows that taking elevators is an easily distinguishable motion mode. They can be recognized with a F1 score close to 94% with an interval of 0.5 s. To achieve similar performance, much larger windows are needed for the other motion modes, mainly because the high operating speed of elevators causes evident variation in air pressure, so that the classifiers can distinguish this mode from the others.

Conversely, the low operating speed of the escalator results in a much longer window to capture sufficient signal variation to achieve the same classification performance as with elevator classification. In elevator classification, any classifier with a 0.5-s window can operate with an F1 score of 94%. Nevertheless, the window length must be at least 2.5 s to obtain a similar performance for up and down escalator movement. Furthermore, the up and down elevator evaluations in Figure 6 had similar patterns, because their corresponding sensor signals have opposite signals and are approximately equal in magnitude. The same was true for the up and down escalator case.

To summarize, we explored the impact of sliding window length on specific motion mode recognition, and found that expanding the window evidently improved the recognition performance of each motion mode, regardless of the method adopted. The enhancement of the F1 score was over 50%. However, improvement by lengthening the sliding window becomes increasingly less with the expansion in window size, which renders blindly increasing window length for better performance unreasonable. Different enhancement cut-off window sizes exist for different motion modes. Based on the results, suitable window sizes are recommended according to the motion mode to be recognized and varying application needs.

4.2. Pose Classification Result

In this section, we analyze the effect of window length on human pose pattern classification and explore a suitable window length. First, we analyzed the effect of window length on the compressed pose pattern feature distribution. Figure 8 presents the results. The effect of window length on pose-specific features differed from that of motion mode. A change in boundaries among pose points was not evident as window length expanded. In contrast to motion mode classification, the effect of window length on human pose classification was limited.

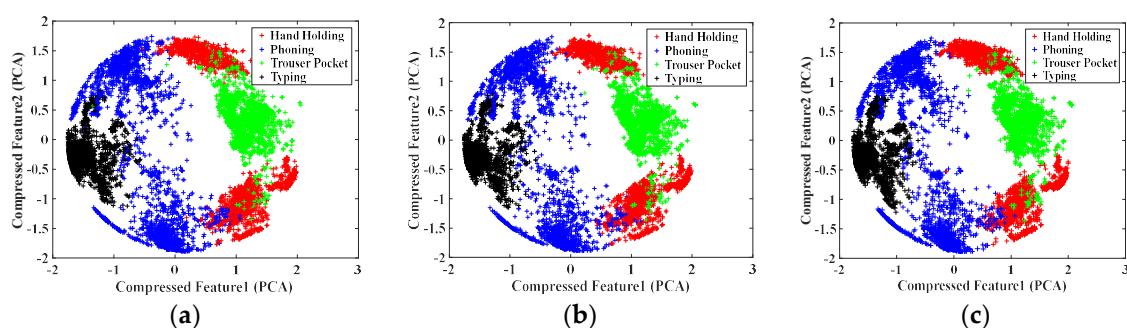


Figure 8. Distribution of compressed features of human poses with different window lengths: (a) 1 s; (b) 2 s; and (c) 3 s.

4.2.1. Global Evaluation

In Figure 9, we compare the pose classification performance of each methodology with different window lengths. Each bar cluster represents different machine-learning methods, and bars in each cluster represent the average F1 score of poses based on window lengths of 0.5 s to 7 s. Among the classifiers, the performances of SVM, DT, KNN, and Adaboost were close to 99% based on a window of only 0.5 s. Although GNB performed the worst, it still achieved a score of 97%. The enhancement from expanding the window size was not evident (less than 0.8%), and a clear increasing trend in the F1 score was not apparent with the increase in window length.

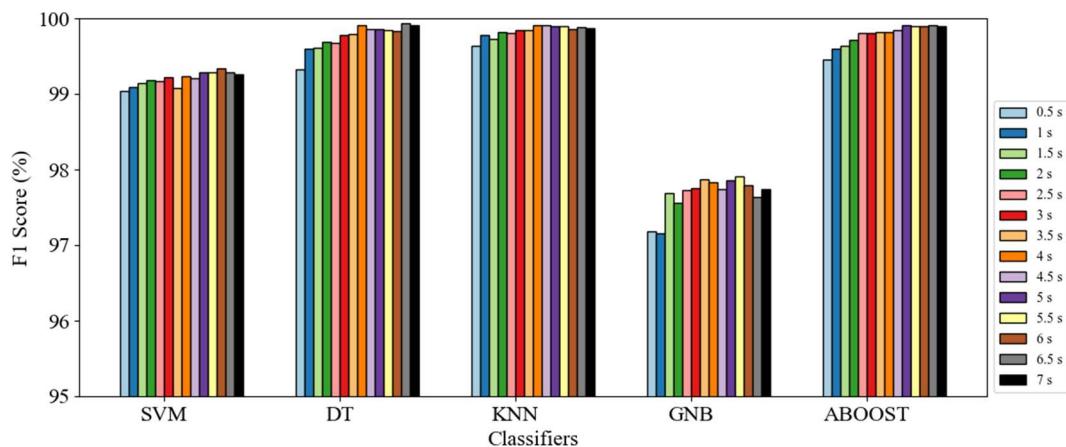


Figure 9. Average F1 score of pose classification using different machine-learning methods and window lengths.

4.2.2. Pose-Specific Analysis

Figure 10 depicts the details of classification performance for specific poses, with the F1 score for the poses using varying classification methods and window lengths. Initially, the classification F1 scores of ‘swing’ and ‘trouser pocket’ were similar, which also applied to calling and typing. We rationalize that when users are typing or calling, their pose patterns are unique, so they are easily distinguishable. Therefore, classifiers can recognize the pose patterns with an F1 score close to 100%. For ‘swing’, we found that people often held the smartphone in a similar manner as the smartphone being carried in the trouser pocket. In this case, although ‘swing’ and the trouser pocket can be distinguished from each other under dynamic motion mode, identifying the difference when the user was static was difficult. The confused samples between swing and trouser pocket when the user was static resulted in relatively poor performance for the swing and trouser pocket poses compared with typing and calling.

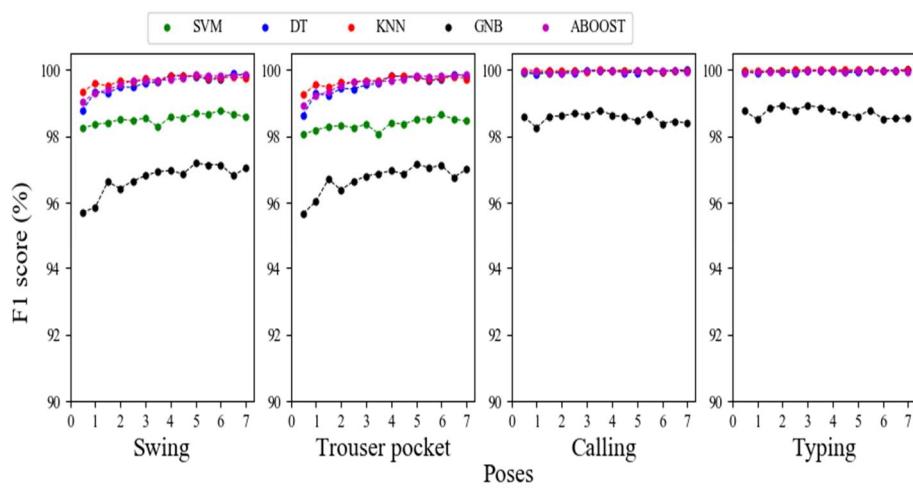


Figure 10. Relationship between F1 score and window length for pose classification.

As depicted in Figure 10, an F1 score close to 99% was achieved based on a window of only 0.5 s for SVM, KNN, DT, and Adaboost in swing and trouser pocket classification. GNB performed the worst. However, GNB also received an F1 score higher than 95% with a 0.5-s window. For typing and calling classification, the results were even better. In these cases, every classifier performed impressively, and even the worst performance exceeded 98%. Therefore, for pose classification, a window of 0.5 s is

sufficient, and using a longer window is unnecessary because the improvement is negligible and does not equal the sacrificed recognition speed. An F1 score beyond 95% was achieved based on a 0.5-s window with all of the classifiers.

In summary, in contrast to motion mode recognition, the influence of sliding window length on pose pattern recognition is not evident. Information extracted from gravity sensor data changes was sufficient to accurately classify these poses. Results show that a sliding window as short as 0.5 s can guarantee an F1 score higher than 95% for all of the pose patterns and machine-learning methods.

5. Discussion

Based on the findings, we propose some useful inferences, providing suggestions for future work. We also summarize the limitations of this work to improve our work in the future.

Suggestions. Even though it is easy to see from the standard workflow of HAR that sliding window length influences the HAR result directly, few researches have been done that reveal such impact in detail. Plenty of recent works [17,23,51,52] still keep selecting the window size intuitively. However, determining the window length based on experience usually create bias in the results. For instance, the recognition performance of a similar motion mode group (walking, stationary, and going up and down stairs) with similar feature sets (time-domain and frequency-domain) was analyzed including the same method (random forest) [52,78]. However, the result presented in Qian et al. [78] exceeded the F1 score of 95%, whereas that in Yufei et al. [52] was less than 85%. Qian et al. attributed their good performance to a new strategy that they introduced in classification. However, they neglected that they used a sliding window length of 5 s, whereas Yufei et al. [52] used a short window of 1 s. Based on our study, the main reason for the improvement in Qian et al. [78] may have been caused by the much longer sliding window rather than the proposed strategy.

A detailed analysis in wearable sensor-based HAR has been presented by Banos et al. [67], but no studies have been conducted yet in smartphone sensor-based HAR. Smartphone-based HAR has different application contexts with wearable sensors [9,67,79]. Wearable sensors that attach to different parts of the human body permit measuring the motion experienced by each body limb and trunk, thus better capturing the human body dynamics. This guarantees the ability of wearable sensors in complex HAR such as sports activities [67,80], but the smartphone can only obtain the dynamic information of a certain part of the body, which makes it relatively weak. However, the smartphone has been extensively used in people's daily life as the most popular device, and is more acceptable than wearable sensors for people to carry every day. Therefore, smartphone-based HAR is becoming increasingly popular in recognizing people's daily activities [51,52,78,81]. As the results in [67] cannot apply to systems using smartphones, a comprehensive analysis is of concern in smartphone sensor-based HAR. Our research provides a comprehensive analysis in this field and makes up this gap.

Our study proves that the motion mode recognition result is influenced heavily by window size, which is independent of classification methods, and the influence on F1 score could be larger than 40%. For the users who largely require recognition performance, a longer window generally results in much better performance. However, blindly increasing the window size is also unreasonable because the improvement after a cut-off window length may be too small to be considered and not worth the sacrificed recognition speed. The improvement cut-off length is proved to be 6 s with an F1 score beyond 99%. As for users who largely focus on reducing recognition latency, they can shorten window sizes by reducing the required accuracy. In this case, a window between 2.5–3.5 s with an F1 score around 95% is recommended. In addition, the improvement cut-off points and the trade-off between performance and window length have been proven to differ according to motion modes. Therefore, window sizes fulfilling various accuracy requirements for specific motion mode recognition are listed in Table 5 for reference.

Our study of pose pattern recognition shows that the impact of window size on pose pattern recognition is limited, and gravity information proved effective in pose recognition even under various motion modes. Based on the variation in the gravity components on each axis of a smartphone

local coordinate system, pose patterns can be classified accurately using a short window size of 0.5 s. This result provides a good reference for researchers who are interested in pose pattern recognition, especially in the field of indoor positioning.

In addition, results corroborate that Adaboost and KNN are more effective than other methods in motion mode and pose pattern recognition. GNB is not recommended based on the bad performance due to its simple principle and assumption that all of the input features are independent, so it cannot extract sufficient useful variation in the features to distinguish the activities. Based on this notion, Adaboost and KNN are recommended for use.

Study generalization. Regarding generalization, the tested recognition systems correspond with those that are widely used in related works. Furthermore, simplicity and comprehensiveness were key elements considered during our study, which enabled us to focus on the potential impact of segmentation on recognition. Thus, in this paper, the data directly captured through the sensors were used, thereby avoiding filtering or preprocessing. These procedures typically remove certain parts of raw signals, which potentially lead to a change in the signal space and limit the applicability of these results to other designs [67]. Moreover, time and frequency domain features were considered in our study to generalize our results, because these features were typically used in previous research [23,51,52]. The feature set extracted from different window sizes were kept constant to eliminate the potential bias from different feature sets on recognition to objectively present the impact of window size.

The motion modes and pose patterns considered in our research are common in people's daily lives in indoor environments. We selected these motion modes and poses by observing human activities in a classical supermarket, which were also hot topics in previous studies [21,23,51,52].

As for the sensors used in our paper, accelerometers are widely used in HAR and have been proven to be effective in recognizing stationary and dynamic modes [64]. A barometer proved to be effective in recognizing vertical moving modes, such as the use of escalators and elevators. As for the elevator and escalator use cases, the subjects move in constant speed such that their acceleration is zero. However, the change in air pressure acts as an effective factor for classification. Therefore, the barometer is usually adopted in height-change motion mode recognition [64–66]. Thus, we used the accelerometer and barometer based on previous research suggestions to consider motion modes [64]. As for pose recognition, recognition based on gravity data has become increasingly popular because the variation in the gravity component on each axis of the smartphone under different poses is easily distinguishable [71]. The magnitude of gravity is constant, so that pose recognition will not be influenced by different motion modes. Therefore, the gravity sensor was used in our paper.

Sampling rate. Our results may also be influenced by the sampling rate. For this reason, we opted to define the window range in terms of time rather than sample amount. Maurer et al. [82] evaluated the effect of sampling rate on recognition accuracy, and found that evident gain did not exist for a sampling frequency over 20 Hz. Therefore, the results obtained could be, in principle, applied to other monitoring systems with sampling rates over 20 Hz.

Performance metric. In many studies, the recognition results of the system are normally measured in terms of accuracy or precision. Despite the extensive use of these metrics in many fields, they are biased in presenting the results, especially in the presence of imbalance issues in the experiment samples. Therefore, we adopted the F1 score in our work, which does not have this limitation [83]. Consequently, the results obtained could be generalized for each activity independent of the number of available instances for each target activity.

Limitations. Our work aimed to conduct a systematic evaluation of the impact of sliding window length on human motion mode and pose pattern recognition using smartphone motion sensors. However, we acknowledge that certain limitations are evident in our work.

1. Firstly, for motion mode recognition, neither a gyroscope nor a magnetometer were used. Although experiments have proven that the barometer and accelerometer are effective and sufficient, the current trends show that using additional sensors could help improve the

recognition performance and system robustness. Therefore, an analysis using other smartphone sensors could be of interest and will be explored in future work.

2. Secondly, the dataset is relatively impoverished, because the data collection was taxing for researchers and subjects. Sufficient amounts of data could hardly be acquired over a short time period. In the future, we will recruit additional subjects so that our data will cover a wider range of ages, heights, and weights of the subjects, and so on. We also aim to establish a comprehensive human motion mode and pose pattern dataset for public use.
3. Finally, in this study, we mainly focused on revealing the impact of segmentation on HAR and manually tuning the window size. However, testing different window sizes before designing the system is time-consuming and inefficient. Advanced methods that could automatically tune the segmentation parameters based on the characteristics of the human activities to be distinguished would be considerably useful. Our future study will also focus on this aspect.

6. Conclusions

In human motion modes and pose pattern recognitions, windowing is a basic step used by the majority of scholars [64]. However, research largely relies on randomly selected values without careful analysis.

In this paper, we presented a comprehensive study that analyzed the influence of window length on human motion mode and pose pattern recognition. We evaluated the effect of window length on motion mode and pose pattern using five well-accepted classification methods. The results demonstrated that the window length affects motion mode recognition, but does not affect pose pattern recognition.

For motion mode classification, recognition performance generally improved by increasing the window length. However, the improvement became increasingly obscure, and a cut-off point was found to exist, after which improvement was negligible and not worth the sacrificed recognition speed. This result affirms that a window length between 2.5–3.5 s provides the best tradeoff between performance and latency. Adaboost performs the best in this window length range. Additionally, we proposed the recommended window lengths for use with varying motion mode classification requirements. In terms of pose classification, the effect of window length is limited, and the benefit of increasing the window length was less than 1% on the F1 score. All of the classification methods with 0.5-s windows achieved satisfactory results. In addition to the analysis on motion modes and pose cluster classification, the classification performance of specific motion modes and poses was analyzed. The suitable window length and technique can be determined through the use of experiments. The results provide a comprehensive understanding of the effect of window length on different classification methods, motion modes, and pose patterns, which subsequently determine the suitable window length and algorithm for motion modes and pose pattern recognition.

Author Contributions: Q.L. and L.W. conceived and designed the experiments; G.W. and W.W. performed the experiments; G.W., M.W. and T.L. analyzed the data; G.W. and L.W. wrote the paper; and all authors proof-read the paper.

Funding: This work was supported in part by the National Key Research Development Program of China (2016YFB0502203); by the National Natural Science Foundation of China (41371377, 91546106, 41401444, 41671387); and by the Shenzhen Future Industry Development Funding Program (201507211219247860).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. European Commission (EC). *Horizon 2020—The Framework Programme for Research and Innovation*; Technical Report; European Commission: Brussels, Belgium, 2013.
2. Lin, J.J.; Mamykina, L.; Lindtner, S.; Delajoux, G.; Strub, H.B. Fish'N'Steps: Encouraging physical activity with an interactive computer game. In Proceedings of the 8th International Conference on Ubiquitous Computing, Orange County, CA, USA, 17–21 September 2006; pp. 261–278.

3. Consolvo, S.; McDonald, D.W.; Toscos, T.; Chen, M.Y.; Froehlich, J.; Harrison, B.; Klasnja, P.; LaMarca, A.; LeGrand, L.; Libby, R.; et al. Activity sensing in the wild: A field trial of ubikit garden. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008; pp. 1797–1806.
4. Lillo, I.; Niebles, J.C.; Soto, A. Sparse composition of body poses and atomic actions for human activity recognition in RGB-D videos. *Image Vis. Comput.* **2017**, *59*, 63–75. [[CrossRef](#)]
5. Jalal, A.; Kim, Y.H.; Kim, Y.J.; Kamal, S.; Kim, D. Robust human activity recognition from depth video using spatiotemporal multi-fused features. *Pattern Recognit.* **2017**, *61*, 295–308. [[CrossRef](#)]
6. Lu, Y.; Wei, Y.; Liu, L.; Zhong, J.; Sun, L.; Liu, Y. Towards unsupervised physical activity recognition using smartphone accelerometers. *Multimed. Tools Appl.* **2017**, *76*, 10701–10719. [[CrossRef](#)]
7. Chen, C.; Jafari, R.; Kehtarnavaz, N. A survey of depth and inertial sensor fusion for human action recognition. *Multimed. Tools Appl.* **2017**, *76*, 4405–4425. [[CrossRef](#)]
8. Gravina, R.; Ma, C.; Pace, P.; Aloisio, G.; Russo, W.; Li, W.; Fortino, G. Cloud-based Activity-as-a-Service cyber-physical framework for human activity monitoring in mobility. *Future Gener. Comput. Syst.* **2017**, *75*, 158–171. [[CrossRef](#)]
9. Wannenburgh, J.; Malekian, R. Physical activity recognition from smartphone accelerometer data for user context awareness sensing. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *47*, 3142–3149. [[CrossRef](#)]
10. Theodoridis, S.; Koutroumbas, K. *Pattern Recognition*, 4th ed.; Elsevier: London, UK, 2009.
11. Tunca, C.; Alemdar, H.; Ertan, H.; Incel, O.D.; Ersoy, C. Multimodal wireless sensor network-based ambient assisted living in real homes with multiple residents. *Sensors* **2014**, *14*, 9692–9719. [[CrossRef](#)] [[PubMed](#)]
12. Kunze, K.S.; Lukowicz, P.; Junker, H.; Troster, G. Where am I: Recognizing On-body Positions of Wearable Sensors. In Proceedings of the LoCA 2005: Location- and Context-Awareness, Oberpfaffenhofen, Germany, 12–13 May 2005; pp. 264–275.
13. Zeng, M.; Nguyen, L.T.; Yu, B.; Mengshoel, O.J.; Zhu, J.; Wu, P.; Zhang, J. Convolutional neural networks for human activity recognition using mobile sensors. In Proceedings of the 2014 6th International Conference on Mobile Computing, Applications and Services (MobiCASE), Austin, TX, USA, 6–7 November 2014.
14. Ordóñez, F.J.; Roggen, D. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* **2016**, *16*, 115. [[CrossRef](#)] [[PubMed](#)]
15. Yang, J. Toward physical activity diary: Motion recognition using simple acceleration features with mobile phones. In Proceedings of the 1st International Workshop on Interactive Multimedia for Consumer Electronics, Beijing, China, 19–23 October 2009; pp. 1–10.
16. Foerster, F.; Smeja, M.; Fahrenberg, J. Detection of posture and motion by accelerometry: A validation study in ambulatory monitoring. *Comput. Hum. Behav.* **1999**, *15*, 571–583. [[CrossRef](#)]
17. Elhoushi, M.; Georgy, J.; Noureldin, A.; Korenberg, M.J. Motion mode recognition for indoor pedestrian navigation using portable devices. *IEEE Trans. Instrum. Meas.* **2015**, *65*, 208–221. [[CrossRef](#)]
18. Frank, K.; Nadales, M.; Robertson, P. Reliable real-time recognition of motion related human activities using MEMS inertial sensors. In Proceedings of the 23rd International Technical Meeting Satellite Division Institute of Navigation (ION GNSS), Portland, OR, USA, 21–24 September 2010; pp. 2919–2932.
19. Ali, A.S.; Georgy, J.; Wright, D.B. Estimation of heading misalignment between a pedestrian and a wearable device. In Proceedings of the International Conference on Localization and GNSS 2014 (ICL-GNSS 2014), Helsinki, Finland, 24–26 June 2014; pp. 1–6.
20. Xiao, Z.; Wen, H.; Markham, A.; Trigoni, N. Robust pedestrian dead reckoning (R-PDR) for arbitrary mobile device placement. In Proceedings of the 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Busan, Korea, 27–30 October 2014; pp. 187–196.
21. Bussmann, J.B.J.; Martens, W.L.J.; Tulen, J.H.M.; Schasfoort, F.C.; Van Den Berg-Emans, H.J.G.; Stam, H.J. Measuring daily behavior using ambulatory accelerometry: The activity monitor. *Behav. Res. Methods Instrum. Comput.* **2001**, *33*, 349–356. [[CrossRef](#)] [[PubMed](#)]
22. Yang, J.-Y.; Chen, Y.-P.; Lee, G.-Y.; Liou, S.-N.; Wang, J.-S. Activity recognition using one triaxial accelerometer: A neuro-fuzzy classifier with feature reduction. In Proceedings of the 6th ICEC 2007: Entertainment Computing—ICEC 2007, Shanghai, China, 15–17 September 2007; pp. 395–400.
23. Prasertsung, P.; Horanont, T. A classification of accelerometer data to differentiate pedestrian state. In Proceedings of the 20th International Computer Science and Engineering Conference: Smart Ubiquitous Computing and Knowledge, ICSEC 2016, Chiang Mai, Thailand, 14–17 December 2016.

24. Choudhury, T.; Consolvo, S.; Harrison, B.; Hightower, J.; LaMarca, A.; LeGrand, L.; Rahimi, A.; Rea, A.; Bordello, G.; Hemingway, B.; et al. The mobile sensing platform: An embedded activity recognition system. *IEEE Pervasive Comput.* **2008**, *7*, 32–41. [[CrossRef](#)]
25. Bao, L.; Intille, S.S. Activity recognition from user-annotated acceleration data. In Proceedings of the Pervasive 2004: Pervasive Computing, Vienna, Austria, 21–23 April 2004; pp. 1–17.
26. Frank, K.; Nadales, V.; Robertson, P.; Angermann, M. Reliable realtime recognition of motion related human activities using MEMS inertial sensors. In Proceedings of the 23rd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2010), Portland, OR, USA, 21–24 September 2010.
27. Chen, Y.-P.; Yang, J.-Y.; Liou, S.-N.; Lee, G.-Y.; Wang, J.-S. Online classifier construction algorithm for human activity detection using a tri-axial accelerometer. *Appl. Math. Comput.* **2008**, *205*, 849–860. [[CrossRef](#)]
28. Jin, G.H.; Lee, S.B.; Lee, T.S. Context awareness of human motion states using accelerometer. *J. Med. Syst.* **2007**, *32*, 93–100. [[CrossRef](#)]
29. Janidarmian, M.; Roshan Fekr, A.; Radecka, K.; Zilic, Z. A comprehensive analysis on wearable acceleration sensors in human activity recognition. *Sensors* **2017**, *17*, 529. [[CrossRef](#)] [[PubMed](#)]
30. Ertuğrul, Ö.F.; Kaya, Y. Determining the optimal number of body-worn sensors for human activity recognition. *Soft Comput.* **2017**, *21*, 5053–5060. [[CrossRef](#)]
31. Cornacchia, M.; Ozcan, K.; Zheng, Y.; Velipasalar, S. A survey on activity detection and classification using wearable sensors. *IEEE Sens. J.* **2017**, *17*, 386–403. [[CrossRef](#)]
32. Diraco, G.; Leone, A.; Siciliano, P. An active vision system for fall detection and posture recognition in elderly healthcare. In Proceedings of the 2010 Design, Automation & Test in Europe Conference & Exhibition (DATE 2010), Dresden, Germany, 8–12 March 2010; pp. 1536–1541.
33. Brusey, J.; Rednic, R.; Gaura, E.I.; Kemp, J.; Poole, N. Postural activity monitoring for increasing safety in bomb disposal missions. *Meas. Sci. Technol.* **2009**, *20*, 075204. [[CrossRef](#)]
34. Zhang, H.; Yuan, W.; Shen, Q.; Li, T.; Chang, H. A handheld inertial pedestrian navigation system with accurate step modes and device poses recognition. *IEEE Sens. J.* **2015**, *15*, 1421–1429. [[CrossRef](#)]
35. Pan, M.S.; Lin, H.W. A step counting algorithm for smartphone users: Design and implementation. *IEEE Sens. J.* **2015**, *15*, 2296–2305. [[CrossRef](#)]
36. Sekine, M.; Tamura, T.; Togawa, T.; Fukui, Y. Classification of waist-acceleration signals in a continuous walking record. *Med. Eng. Phys.* **2000**, *22*, 285–291. [[CrossRef](#)]
37. Yoshizawa, M.; Takasaki, W.; Ohmura, R. Parameter exploration for response time reduction in accelerometer-based activity recognition. In Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing, Zurich, Switzerland, 8–12 September 2013; pp. 653–664.
38. Aminian, K.; Rezakhanlou, K.; de Andres, E.; Fritsch, C.; Leyvraz, P.F.; Robert, P. Temporal feature estimation during walking using miniature accelerometers: An analysis of gait improvement after hip arthroplasty. *Med. Biol. Eng. Comput.* **1999**, *37*, 686–691. [[CrossRef](#)] [[PubMed](#)]
39. Aminian, K.; Najafi, B.; Bla, C.; Leyvraz, P.F.; Robert, P. Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes. *J. Biomech.* **2002**, *35*, 689–699. [[CrossRef](#)]
40. Wan, J.; O’Grady, M.J.; O’Hare, G.M.P. Dynamic sensor event segmentation for real-time activity recognition in a smart home context. *Pers. Ubiquitous Comput.* **2015**, *19*, 287–301. [[CrossRef](#)]
41. Mortazavi, B.; Lee, S.I.; Sarrafzadeh, M. User-centric exergaming with fine-grain activity recognition: A dynamic optimization approach. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, Seattle, WA, USA, 13–17 September 2014.
42. Atallah, L.; Lo, B.; King, R.; Yang, G.Z. Sensor positioning for activity recognition using wearable accelerometers. *IEEE Trans. Biomed. Circuits Syst.* **2011**, *5*, 320–329. [[CrossRef](#)] [[PubMed](#)]
43. Gjoreski, H.; Gams, M. Accelerometer data preparation for activity recognition. In Proceedings of the International Multiconference Information Society, Ljubljana, Slovenia, 10–14 October 2011.
44. Jiang, M.; Shang, H.; Wang, Z.; Li, H.; Wang, Y. A method to deal with installation errors of wearable accelerometers for human activity recognition. *Physiol. Meas.* **2011**, *32*, 347–358. [[CrossRef](#)] [[PubMed](#)]
45. Kwapisz, J.R.; Weiss, G.M.; Moore, S.A. Activity recognition using cell phone accelerometers. In Proceedings of the 17th Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 21–24 August 2011; Volume 12, pp. 74–82.

46. Wang, J.H.; Ding, J.J.; Chen, Y.; Chen, H.H. Real time accelerometer-based gait recognition using adaptive windowed wavelet transforms. In Proceedings of the IEEE Asia Pacific Conference on Circuits and Systems, Kaohsiung, Taiwan, 2–5 December 2012; pp. 591–594.
47. Sun, L.; Zhang, D.; Li, B.; Guo, B.; Li, S. Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. In Proceedings of the 7th International Conference on Ubiquitous Intelligence and Computing, Xi'an, China, 26–29 October 2010; pp. 548–562.
48. Khan, A.M.; Lee, Y.K.; Lee, S.; Kim, T.S. Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis. In Proceedings of the 5th International Conference on Future Information Technology, Busan, Korea, 21–23 May 2010; pp. 1–6.
49. Lee, Y.S.; Cho, S.B. Activity recognition using hierarchical hidden markov models on a smartphone with 3D accelerometer. In Proceedings of the 6th International Conference on Hybrid Artificial Intelligent Systems, Wroclaw, Poland, 23–25 May 2011; pp. 460–467.
50. Siirtola, P.; Röning, J. User-independent human activity recognition using a mobile phone: Offline recognition vs. In real-time on device recognition. In Proceedings of the 9th International Conference on Distributed Computing and Artificial Intelligence, Salamanca, Spain, 28–30 March 2012; pp. 617–627.
51. Li, P.; Wang, Y.; Tian, Y.; Zhou, T.S.; Li, J.S. An Automatic User-Adapted Physical Activity Classification Method Using Smartphones. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 706–714. [CrossRef] [PubMed]
52. Chen, Y.; Shen, C. Performance analysis of smartphone-sensor behavior for human activity recognition. *IEEE Access* **2017**, *5*, 3095–3110. [CrossRef]
53. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
54. Cheng, W.-C.; Jhan, D.-M. Triaxial accelerometer-based fall detection method using a self-constructing cascade-AdaBoost-SVM classifier. *IEEE J. Biomed. Health Inform.* **2013**, *17*, 411–419. [CrossRef] [PubMed]
55. Yin, J.; Yang, Q.; Pan, J.J. Sensor-based abnormal human activity detection. *IEEE Trans. Knowl. Data Eng.* **2008**, *20*, 1082–1090. [CrossRef]
56. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*; Morgan Kaufmann: San Francisco, CA, USA, 2005.
57. Susi, M.; Borio, D.; Lachapelle, G. Accelerometer signal features and classification algorithms for positioning applications. In Proceedings of the International Technical Meeting of The Institute of Navigation 2011, San Diego, CA, USA, 24–26 January 2011; pp. 158–169.
58. Kwapisz, J.R.; Weiss, G.M.; Moore, S.A. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explor. Newslett.* **2010**, *12*, 74–82. [CrossRef]
59. Ravi, N.; Dandekar, N.; Mysore, P.; Littman, M.L. Activity recognition from accelerometer data. In Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence, Pittsburgh, PA, USA, 9–13 July 2005; Volume 3, pp. 1541–1546.
60. Hand, D.J.; Yu, K. I diot's Bayes—not so stupid after all? *Int. Stat. Rev.* **2001**, *69*, 385–398.
61. Joglekar, S. Adaboost—Sachin Joglekar's Blog. Available online: <https://codesachin.wordpress.com/tag/adaboost/> (accessed on 3 August 2016).
62. Daghstani, T.; Alshammari, R. Improving Accelerometer-Based Activity Recognition by Using Ensemble of Classifiers. *Int. J. Adv. Comput. Sci. Appl.* **2016**, *7*, 128–133. [CrossRef]
63. Ponce, H.; Miralles-Pechuán, L.; Martínez-Villaseñor, M.D.L. A flexible approach for human activity recognition using artificial hydrocarbon networks. *Sensors* **2016**, *16*, 1715. [CrossRef] [PubMed]
64. Elhoushi, M.; Georgy, J.; Noureldin, A.; Korenberg, M.J. A survey on approaches of motion mode recognition using sensors. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 1662–1686. [CrossRef]
65. Frank, K.; Diaz, E.M.; Robertson, P.; Sanchez, F.J.F. Bayesian recognition of safety relevant motion activities with inertial sensors and barometer. In Proceedings of the 2014 IEEE/ION Position, Location and Navigation Symposium—PLANS 2014, Monterey, CA, USA, 5–8 May 2014; pp. 174–184.
66. Zhao, X.; Saeedi, S.; El-Sheimy, N.; Syed, Z.; Goodall, C. Towards arbitrary placement of multi-sensors assisted mobile navigation system. In Proceedings of the 23rd International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2010), Portland, OR, USA, 21–24 September 2010; pp. 556–564.
67. Banos, O.; Galvez, J.M.; Damas, M.; Pomares, H.; Rojas, I. Window size impact in human activity recognition. *Sensors* **2014**, *14*, 6474–6499. [CrossRef] [PubMed]

68. Khan, A.M.; Lee, Y.K.; Lee, S.; Kim, T.S. Accelerometer’s Position Independent Physical Activity Recognition System for Long-term Activity Monitoring in the Elderly. *Med. Biol. Eng. Comput.* **2010**, *48*, 1271–1279. [CrossRef] [PubMed]
69. Kawahara, Y.; Kurasawa, H.; Morikawa, H. Recognizing User Context Using Mobile Handsets with Acceleration Sensors. In Proceedings of the 2007 IEEE International Conference on Portable Information Devices, Orlando, FL, USA, 25–29 May 2007.
70. Lee, S.M.; Yoon, S.M.; Cho, H. Human activity recognition from accelerometer data using Convolutional Neural Network. In Proceedings of the 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju, Korea, 13–16 February 2017.
71. Park, J.G.; Patel, A.; Curtis, D.; Teller, S.; Ledlie, J. Online pose classification and walking speed estimation using handheld devices. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, Pittsburgh, PA, USA, 5–8 September 2012.
72. Android developers/motion sensors. Available online: https://developer.android.com/guide/topics/sensors/sensors_motion#sensors-motion-grav (accessed on 24 April 2018).
73. Stone, M. Asymptotics for and against cross-validation. *Biometrika* **1977**, *64*, 29–35. [CrossRef]
74. Varian, H. Bootstrap Tutorial. *Math. J.* **2005**, *9*, 768–775.
75. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. [CrossRef]
76. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **1933**, *24*, 417. [CrossRef]
77. Guo, S.; Xiong, H.; Zheng, X.; Zhou, Y. Activity Recognition and Semantic Description for Indoor Mobile Localization. *Sensors* **2017**, *17*, 649. [CrossRef] [PubMed]
78. Guo, Q.; Liu, B.; Chen, C. two-layer and multi-strategy framework for human activity recognition using smartphone. In Proceedings of the 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 22–27 May 2016.
79. Tahavori, F.; Stack, E.; Agarwal, V.; Burnett, M.; Ashburn, A.; Hoseinitabatabaei, S.A.; Harwin, W. Physical activity recognition of elderly people and people with parkinson’s (PwP) during standard mobility tests using wearable sensors. In Proceedings of the 2017 International Smart Cities Conference (ISC2), Wuxi, China, 14–17 September 2017.
80. Barshan, B.; Yüksek, M.C. Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units. *Comput. J.* **2013**, *57*, 1649–1667. [CrossRef]
81. Ronao, C.A.; Cho, S.B. Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Syst. Appl.* **2016**, *59*, 235–244. [CrossRef]
82. Maurer, U.; Smailagic, A.; Siewiorek, D.P.; Deisher, M. Activity recognition and monitoring using multiple sensors on different body positions. In Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks (BSN’06), Cambridge, MA, USA, 3–5 April 2006; pp. 113–116.
83. Beitzel, S.M. *On Understanding and Classifying Web Queries*; Illinois Institute of Technology: Chicago, IL, USA, 2006.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).