

Activity Recognition and Intensity Estimation in Youth from Accelerometer Data Aided by Machine Learning

Xiang Ren · Wei Ding · Crouter, Scott E · Yang
Mu · Rong Xie

Received: date / Accepted: date

Abstract Physical activity monitoring for youth is an area of increasing scientific and public health interest due to the high prevalence of obesity and downward trend in physical activity. However, accurate assessment of such activity remains a challenging problem because of the complex nature in which certain activities are performed. In this study, we formulated the issue as a machine learning problem—using a diverse set of 19 physical activities commonly performed by youth—via two approaches: activity recognition and intensity estimation. With the aid of training data, we implemented a distance metric learning method called DML-KNN that utilizes time-frequency features and is capable of effectively classifying both continuous and intermittent movement in youth subjects. Four different time-frequency feature extraction methods were then systematically evaluated. Our results show that the DML-KNN method performed competitively, especially when using features extracted by the Tamura method for intensity estimation, and by the Square Coefficient method for activity recognition.

Keywords machine learning · feature extraction · physical activity classification

1 Introduction

The collection and analysis of human activity data via body-worn sensors has garnered considerable attention in recent years [18, 22, 34]. Studies have shown that accelerometer measurements during physical activity have a strong relationship to energy expenditure, which is widely accepted as the standard reference [9, 11, 12, 21, 28]. Advances in technology have allowed accelerometers with low power consumption, easy setup, and unobtrusive design to provide a promising tool for monitoring free-living physical activities [2, 16, 37].

Xiang Ren
Wuhan University International School of Software, 37 Luoyu Road, Hongshan District, Wuhan, Hubei Province, China
Tel.: +18043148424
E-mail: xiangren2000@gmail.com

Wei Ding
University of Massachusetts Boston 100 Morrissey Blvd. Boston, MA 02125-3393

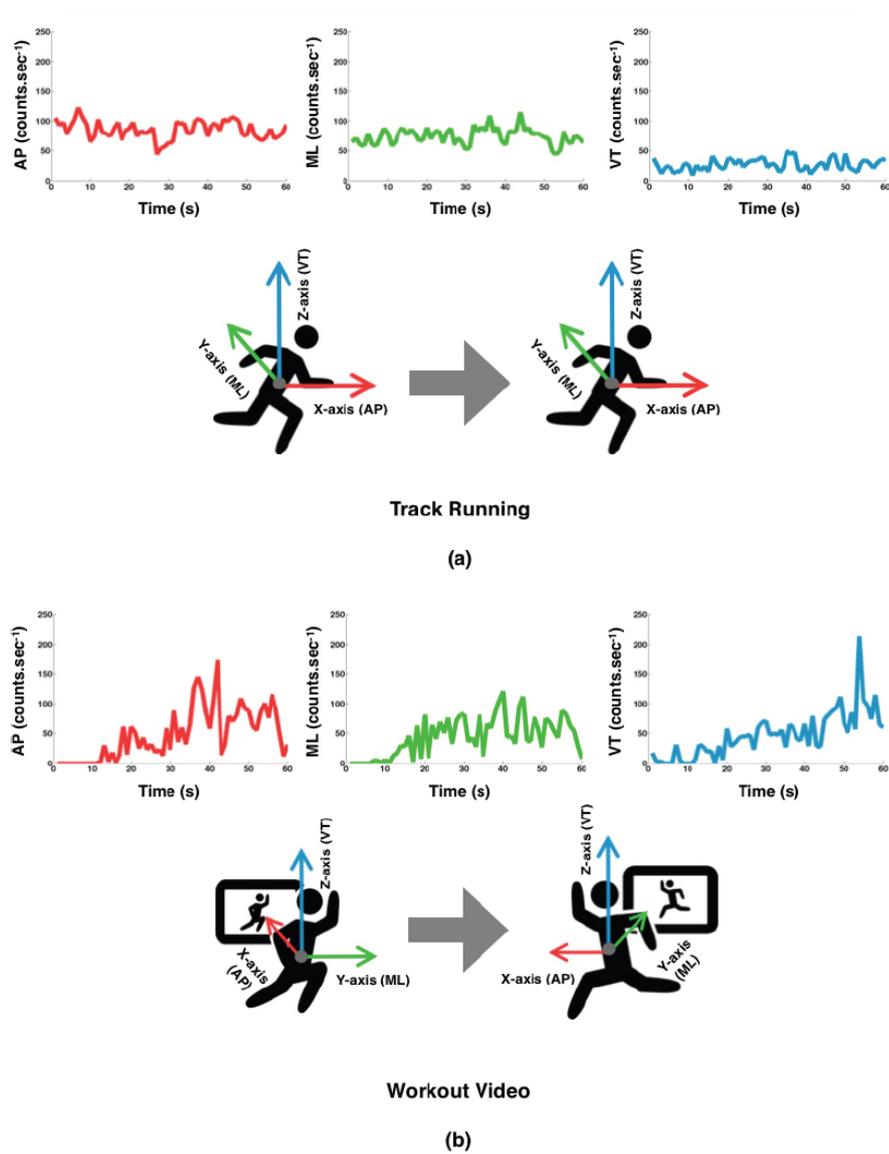


Fig. 1 Accelerometer measurements taken in one-second epochs of continuous (a, track running) and intermittent (b, workout video) movements in the anterior-posterior (AP), medio-lateral (ML), and vertical (VT) directions. Each set of graphs is accompanied by a pair of silhouette figures demonstrating some of the movements associated with each activity.

Physical activity recognition is well documented as an effective predictor of widespread chronic health problems among members of the public. The ability to differentiate between types of physical activities could provide valuable new insights in the fields of behavioral and clinical science [31]. Many previous studies have focused on monitoring and analyzing adults' physical activities (e.g., standing, walking, ascending and descending stairs, and traversing up and down gradient slopes[14, 20, 27, 28]), However, these works only analyzed simulated activities conducted in controlled environments by small numbers of subjects [3]. The breadth of research that uses machine learning algorithms to examine real-life youth physical activities is exceedingly limited[17, 32]. Importantly, the types of activities and how they are performed by youths tend to vary more greatly than for adults[8]. Thus, techniques and algorithms generated for adults from previous studies are not necessarily applicable to youth populations[32].

As the need for measurement precision has increased, techniques for monitoring and analyzing physical activity have resolved to divide the complexities of recorded behaviors into continuous movement (e.g. walking and running) and intermittent movement (e.g. vacuuming, basketball, etc.)[6, 7]. The consequences of this separation are most apparent for intermittent movement, which involve many postural transitions, and thus create sharp high-frequency transients in recorded data. Unlike continuous movement activities that produce cyclical motions and are less frequently misclassified, intermittent movements exhibit non-cyclical characteristics with more changes in frequency and acceleration that complicate classification. For example, Figure 1 shows two physical activities performed by youth in the anterior-posterior, medio-lateral, and vertical axes. From the figure, it is clear that track running produces movements more regularly distributed in both amplitude and frequency than the workout video activity, indicating that the subject is performing cyclical motions. In contrast, the workout video activity generates movements that contain postural changes in all three directions, as evidenced by sharp temporal spikes in each axis.

In order to address the sporadic behavioral patterns of youth, we introduced a combination of time-frequency feature extraction and a local Distance Metric Learning used on K Nearest Neighbors (DML-KNN) method to classify a diverse range of common activities. These included sedentary behaviors, household chores, locomotion, video games, and exercise and sports. Due to the presence of sharp high-frequency transients in intermittent movement data, information of interest is often buried within a combination of features, localized in the time and frequency domains. To obtain the desired data we first extracted time-frequency feature sets—henceforth referred to as just "features"—from accelerometer counts, then projected those features onto subspaces using a set of learned local distance metrics to make clear associations between similar activities and enhance discrimination between different activities. Finally, we employed K Nearest Neighbors (CkNN)[4, 8], a machine learning method that extends from standard kNN, for multiple instance classification. kNN is well-suited to handling noise and multimodal distributions, which are features common to data collected from accelerometers during physical activities. Comparison experiments with a SVM classifier revealed clear distinctions in accuracy when using time-frequency features to examine intermittent movement.

Notable features in this work include:

1. A new machine learning framework to generate time-frequency features from accelerometer data, which uses a local distance metric learning (DML-KNN) method for intensity

estimation and activity recognition, as well as an evaluation this method using a real-world examination of physical activities in youth.

2. Utilizing a larger sample size than most predecessors reported in literature.
3. The division of youth movement categories into continuous and intermittent, where intermittent movements contain sharp high-frequency changes in signal during transitions in motion—a feature not easily detected by previous methodologies.
4. Extensive comparison studies with a SVM method. The experimental results reveal the advantages of using our proposed DML-KNN classifier when examining time-frequency features.

2 Related work

Our method draws from two separate fields of study: feature extraction and activity classification using accelerometer data.

2.1 Feature extraction

Time-domain features are most widely used for activity recognition[9–11, 23, 32, 33]. For example in[30] data features are extracted from each one minute window of accelerometer data and categorized them as follows:

1. 10th, 25th, 50th, 75th, and 90th percentile values for every one-second interval; sixty intervals in each one-minute window.
2. Lag-1 to lag-8 autocorrelations, representing temporal relationships.

While the presence of unavoidable artificial error effects can produce noise in data with one-second epochs, the 10th, 25th, 50th, 75th, and 90th percentile values provide workable representations of the distribution present in each one-minute window, and serve as equivalent dimensions for each instance.

Lag-1 to lag-8 autocorrelation captures the correlation between neighboring seconds (1-8 seconds) that cannot be discerned by examining percentiles alone. Given a time-series dataset $Y = \{Y_1, Y_2, \dots, Y_N\}$ at time $X = \{X_1, X_2, \dots, X_N\}$, the lag-k (with k from 1 to 8) autocorrelation is defined as the function:

$$r_k = \frac{\sum_{i=1}^{N-k} (Y_i - \bar{Y})(Y_{i+k} - \bar{Y})}{\sum_{i=1}^N (Y_i - \bar{Y})^2} \quad (1)$$

Where \bar{Y} is the mean of time-series dataset Y. The larger the value of lag-k, the greater the correlation.

Previous works have postulated that applying a combination of time-domain and frequency-domain features can generate more accurate results than using either one separately[8]. These include statistical time-frequency features[21, 26, 36]; a combination of time-domain features and Fourier Transforms to extract the mean, entropy, energy, and correlation in the frequency domain[4, 5, 12]; wavelet packet decompositions to extract time-frequency features between 0.25HZ and 17HZ, where most useful information is contained[14], and

applying discrete wavelet transforms to analyze healthy activities in adults[15, 20, 25, 27].

Inspired by the aforementioned works, we combined wavelet time-frequency features and time-domain features to further improve activity recognition. This enhancement to our analytical arsenal was crucial, as we were better equipped to identify the movement patterns of our 112 youth participants, who were measured performing 19 daily physical activities.

2.2 Activity classification from accelerometer data

Existing classification methods vary depending on the application: decision trees being one such method that have performed admirably for activity recognition in uncontrolled out-of-laboratory environments [4, 10, 29]. Notably, Hong et al. used a hierarchical structure that divided the activity recognition task into body motion and hand motion with object identification, then applied a decision tree classifier that was sensitive enough to recognize instrumental activity (I-ADL) [12]. However, decision trees can also be applied toward intensity estimation, with the data collected from accelerometers and gyro sensors placed on the ankle, hip and wrist [21]. Some researchers focus on estimation of energy expenditure and prediction of activity type using either artificial neural networks (ANN) [11, 32], or the Bipart method [17]. Another option is to first use a random forest classifier to predict physical activity types, then use a random forest of regression trees to estimate energy expenditure [9]. Still others have used support vector machines (SVM) to perform activity recognition based on a more practical, long-term single sensor system [15]. Cleland et al. showed that optimal accelerometer placement for adults was on the hip, and that SVM achieved the best classification results when differentiating between activities exhibiting significantly different accelerometer readings, such as sitting and walking [5]. Bao et al. focused on differentiating activities performed without researcher supervision, finding decision tree classifiers to be most accurate, while k-Nearest neighbor was the second most accurate algorithm [4]. Other investigators have used meta-level classifiers to produce results with remarkable accuracy [8, 26].

In this study, we first used DML to project generated time-frequency features of youth physical activities onto a more discriminative subspace, maximizing the distance between data points from different movement groups, while minimizing the distance between points from the same group. We then used the k-Nearest neighbor classifier method for physical activity recognition. We begin by recognizing that both intensity estimation and activity recognition play key roles in classification. For intensity estimation, we classified activities into sedentary behaviors (< 1.5 metabolic equivalents (MET)), light physical activity (1.5-2.99 METs), moderate physical activity (3.00-5.99 METS), or vigorous physical activity (≥ 6.0 METs) categories; for activity recognition, we aim to correctly identify motions associated with a wide range of activity types[26].

3 Time-frequency features for activity recognition and intensity estimation

Ultimately, both activity recognition and intensity estimation are classification problems. As the subjects of this study were youths, our methods placed particular emphasis on distinguishing between the continuous and intermittent movements, as youths are prone to

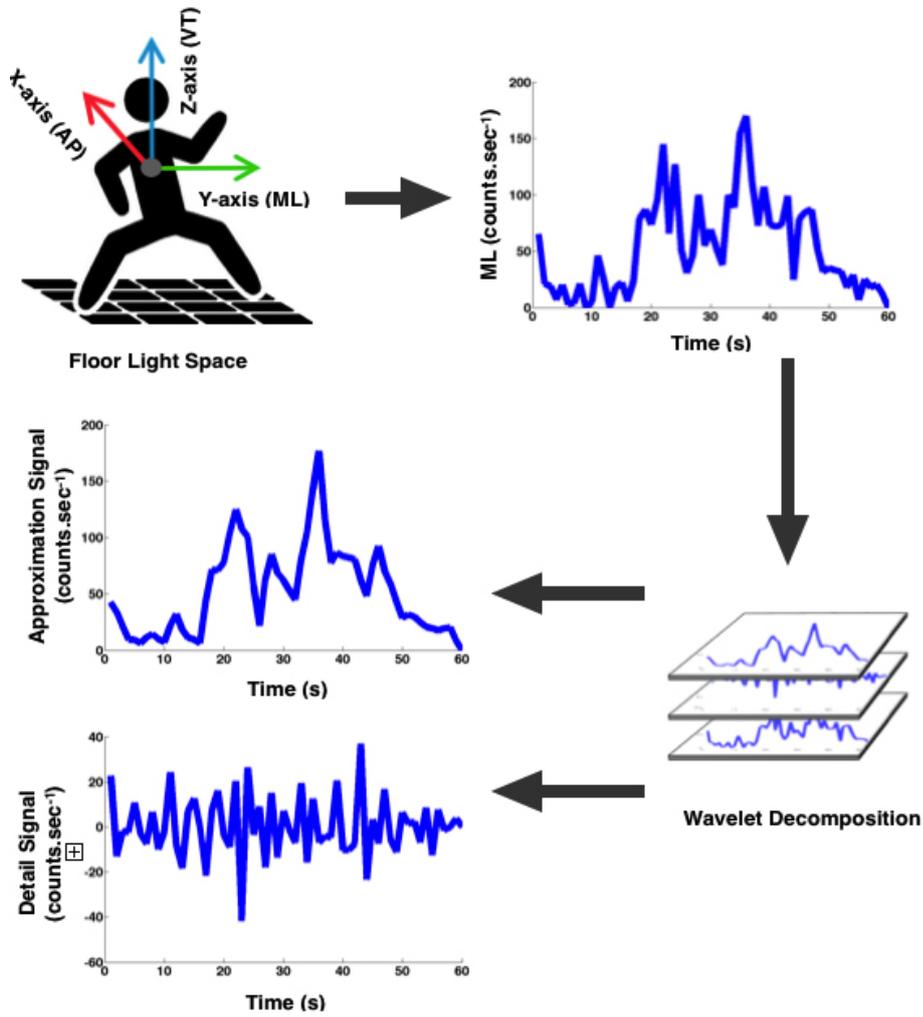


Fig. 2 An example of signal wavelet decomposition using data from the floor Light Space activity.

performing continuous movement activities in a manner that analyses could easily misinterpret as intermittent when examining accelerometer data. To this end, we first extracted time-frequency domain features from accelerometer data, then used a local distance metric learning method to create a projection matrix. This matrix mapped the time-frequency domain features onto a lower dimension, placing youth activity points from identical classes close to each other, while positioning points from different classes at a greater distance. Finally, taking computational efficiency into account, we applied a simple KNN classifier.

3.1 Wavelet features generation

We employ wavelet feature generation to extract additional frequency information not readily available from the accelerometry data, specifically accelerometer counts with sharp frequency changes—indicating posture transients, which cannot be extracted from time-domain features. The decision to use wavelets derived from their suitability for interpreting intermittent movements containing numerous frequency changes. Also, unlike Fourier transforms, which can only extract frequency information, wavelets have the capacity to generate both time and frequency characteristics, which better capture frequency fluctuations on short time scales.

In a discrete youth activity wavelet transform, an accelerometer signal $x(n)$, where x is the accelerometer signal and n is the length, is split into an approximation coefficient $CA1$, using a low-pass filter g , and a detail signal $CD1$, using a high-pass filter h . Using $CA1$ and $CD1$, we can reconstruct the original accelerometer signal into approximation signal $A1$ and detail signal $D1$. Approximation signal $A1$ contains the overall structure of the youth activity, while detail signal $D1$ contains detailed information, like postural transients and frequency changes performed. Fig 2 provides an example using data gathered from youth playing floor Light Space, an interactive game activity that requires players to touch lighted floor panels using their feet. After decomposing the medio-lateral activity signal into approximation and detail components, we can observe that the approximation signal follows the overall trend of the original, while the detail signal fluctuates wildly in the course of the subject's medio-lateral movements, though on a smaller scale.

The approximation coefficient $CA1$ can be further split into a second level approximation coefficient $CA2$ and detail coefficient $CD2$. If the frequency of the original signal is f_{max} , the band-pass of the first approximation signal $A1$ is $[0, f_{max}/2]$, and the band-pass of the first detail signal $D1$ is $[f_{max}/2, f_{max}]$. This process can be repeated, splitting the signal into yet more coefficients, as shown in Fig 3. The maximum level of decomposition depends on the number of data points. Specifically, if N is the number of the data points in one observation, the number of levels cannot be greater than j , where $N = 2^j$. As the data sets used contained 60 data points for each axis, the maximum decomposition level was 5.

The coefficients can be derived from the dilation function ϕ and the wavelet function ψ , respectively. These functions are defined as:

$$\phi_{j,k}(n) = 2^{-j/2} \phi(2^{-j}n - k) \quad (2)$$

$$\psi_{j,k}(n) = 2^{-j/2} \psi(2^{-j}n - k) \quad (3)$$

Where $j \in Z$ represents the decomposition level of the dilation or wavelet function, and $k \in Z$ represents an index for translation in time.

Detail coefficients CD produced by decomposition are typically small, consisting mainly of high-frequency noise, while approximate coefficients CA are less noisy and contain information about broader shifts in the data values. Thus, the process of wavelet feature generation involves two components: decomposing a sample to obtain the wavelet coefficients, and extracting useful information from these coefficients.

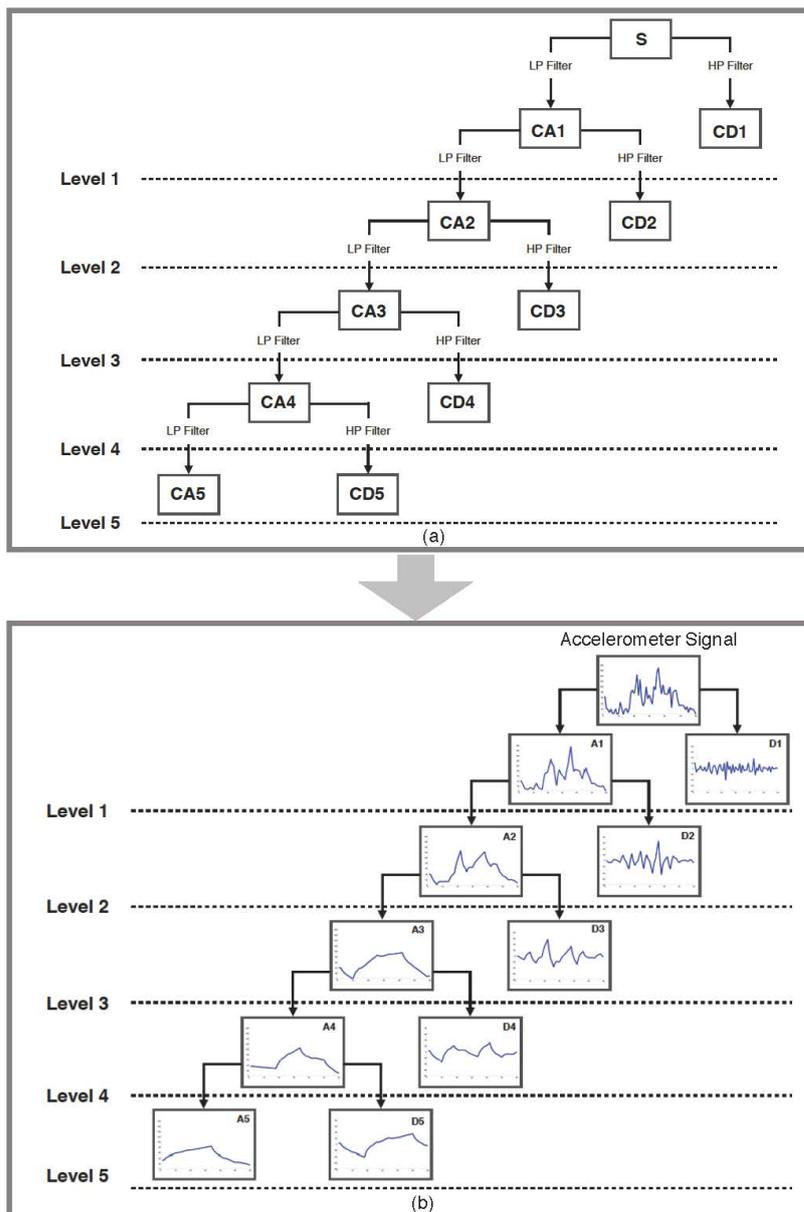


Fig. 3 (a) Illustration of the process of wavelet decomposition, where S is the original signal, CA represents approximation coefficients, and CD represents detail coefficients. LP filter and HP filter are low-pass filters and high-pass filters, respectively. By using CA and CD, we can reconstruct the approximation and detail signals from each level. (b) Applying this process to data from the floor Light Space activity.

3.2 Proposed time-frequency features

The features we choose to extract include both time-domain features and frequency-domain features, with the latter being able to reveal information from intermittent movements that the former cannot. In all, we utilized four different methods for wavelet feature generation to study accelerometer data: Tamura, Nyan, Square coefficient, and Magnitude coefficient.

The Tamura method decomposes accelerometer data into five levels using a wavelet transform, with ‘Daubechies 3’ as the wavelet mother. Daubechies wavelets are a family of orthogonal wavelets defining a discrete wavelet transform where the number characterizes the length of the wavelet. Features are calculated as the sum of the squared detail coefficients at levels four and five [24]. The Nyan method, previously used by Nyan et al., [20] establishes wavelet features in much the same way [24]. However, rather than treating the scales separately, the summations at levels four and five are added together. The squared coefficient method decomposes each component of the tri-axial acceleration signal into five levels using a ‘Daubechies 2’ wavelet mother, and calculates the sum of the square detail coefficients at levels one through five for each axis. The magnitude coefficient method is the final approach used, and operates similarly to the squared coefficient method, but instead generates the absolute values of each decomposed component of the tri-axial acceleration signal.

A number of time-domain features, including several not examined in previous studies, are analyzed:

1. Maximum and minimum values—1st and 100th percentile: Maximum and minimum values can aid in distinguishing activities from different intensity categories.
2. Values at percentiles from the 5th to the 100th, in 5-percentile increments: A greater number of percentiles are employed than previous works that recorded activity in adults to counteract the increased variability in movements observed in youth.
3. Lag-1 to lag-8 autocorrelations, representing temporal relationships: Much like in previous works, autocorrelations are used to detect and eliminate dependencies in the data.
4. Number of counts with a value of zero, in each axis: sedentary activities—such as Supine rest, reading, watching TV, and searching the internet—contained a large number of zero-valued data readings; to better differentiate these activities, we counted the number of zero values observed in each axis.

3.3 Semi-supervised learning for activity classification

Though methods exist for classifying accelerometer data from generated features—including SVM [15], ANN [13, 14], GMM [19, 35], decision trees, and Nave Bayes [1, 4]—we opted to devise a semi-supervised learning method to utilize our proposed time-frequency features, and was designed to yield competitive results in the physical activity domains of our youth subjects, especially for the classification of intermittent movements involving sharp frequency changes and postural transients.

For time series classification we treated adjacent samples as a block, given that samples adjacent in time predominantly possess similar structures and information. Using these blocks we were able to independently identify a local distance metric from training and test sets that minimized the distance between samples from the same class, while simultaneously maximizing the distance between samples from different classes. Finally we used a kNN

classifier to implement classification.

Given a set with time-frequency domain features $F = \{f_1, f_2, \dots, f_N\} \in \mathbb{R}^{m \times N}$, containing N samples, where each sample has m time-frequency domain attributes, we can use the equation for Mahalanobis distance to define the distance between f_i and f_j as :

$$d_A(f_i, f_j) = \sqrt{(f_i - f_j)^T A (f_i - f_j)} \quad (4)$$

where A is positive and definite. If we replace A in equation (4) with WW^T using Cholesky decomposition, we obtain:

$$\begin{aligned} d_A(f_i, f_j) &= \sqrt{(f_i - f_j)^T WW^T (f_i - f_j)} \\ &= \sqrt{(W^T (f_i - f_j))^T (W^T (f_i - f_j))} \\ &= \| W^T (f_i - f_j) \| \end{aligned} \quad (5)$$

Our goal is to obtain a projection matrix W that can map all the activity points in the time-frequency domain to lower dimensions, where activity points from the same class are located as close to each other as possible, while points from different classes are as far from each other as possible.

If we employ two distance metrics to replace W with $W_1 W_2$, where W_1 and W_2 are respectively the distance metrics learned from test data and training data, the distance between f_i and f_j becomes:

$$d_A(f_i, f_j) = \| W_2^T W_1^T (f_i - f_j) \| \quad (6)$$

Equation (6) is equivalent to projecting all time-frequency features onto the space defined by projection matrices W_1 , learned from the test set, and W_2 , learned from the training set. For the next step, we will use test data distance metric W_1 to illustrate the procedure, though W_2 operates identically. For the time-frequency domain features, given a test feature sample f_i , we defined features with the same physical activity class as $F_i^s = \{f_{i_1}^s, f_{i_2}^s, \dots, f_{i_{k_1}}^s\}$, and features with different physical activity classes as $F_i^d = \{f_{i_1}^d, f_{i_2}^d, \dots, f_{i_{k_2}}^d\}$. For each sample f_i , we constructed a block structure that contained the values for the sample itself, as well as the values of the nearest samples from the same physical activity class and the nearest samples from different activity classes. Each block was defined as:

$$\begin{aligned} F_i &= \{f_i, F_i^s, F_i^d\} \\ &= \{f_i, f_{i_1}^s, f_{i_2}^s, \dots, f_{i_{k_1}}^s, f_{i_1}^d, f_{i_2}^d, \dots, f_{i_{k_2}}^d\} \end{aligned} \quad (7)$$

Using f_i as an example, we minimized the distance between features of the same physical activity class using the equation:

$$\begin{aligned} & \arg \min_{A_1} \sum_{j=1}^{k_1} d_{A_1}^2(f_i, f_{ij}^s) \\ &= \arg \min_{W_1} \sum_{j=1}^{k_1} \| W_1^T (f_i - f_{ij}^s) \|^2 \end{aligned} \quad (8)$$

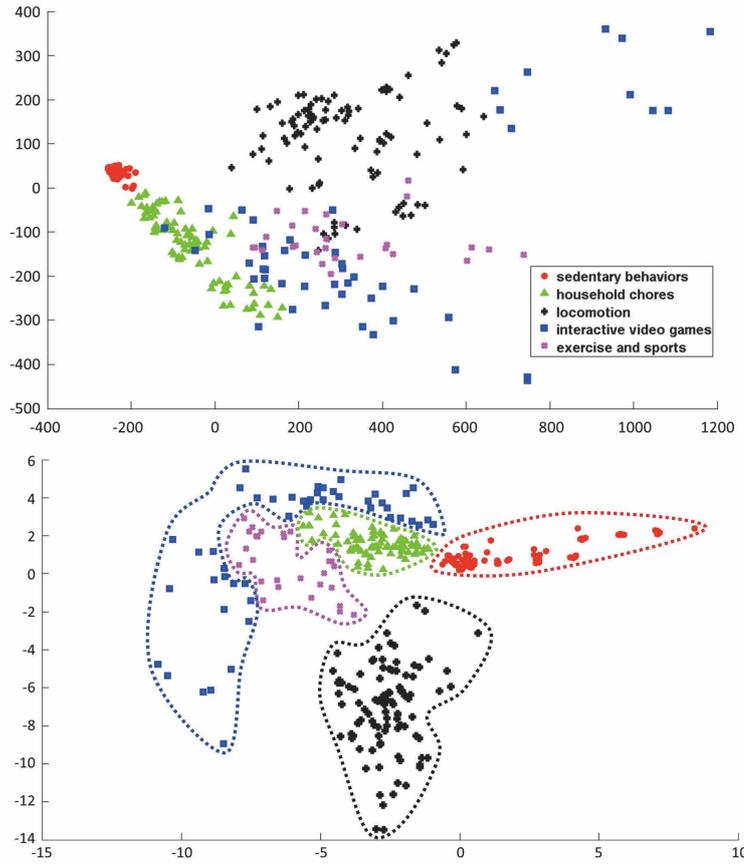


Fig. 4 Anterior-posterior (AP), Medio-Lateral (ML), and Vertical (VT) axis accelerometer data with 1-second epochs was used to derive a 2-dimensional PCA projection (top), after feature extraction. Features were then projected onto the subspace (bottom).

Meanwhile, we maximized the distance between features from different activity classes, as illustrated in Fig 4, using the equation:

$$\begin{aligned}
 & \arg \max_{A_1} \sum_{q=1}^{k_2} d_{A_1}^2(f_i, f_{i_q}^d) \\
 & = \arg \max_{W_1} \sum_{q=1}^{k_2} \|W_1^T(f_i - f_{i_q}^d)\|^2
 \end{aligned} \tag{9}$$

where $A_1 = W_1^T W_1$ is the distance metric learned from the test set.

Combining equations (8) and (9) produces:

$$\arg \min_{w_1} \left(\sum_{j=1}^{k_1} \|W_1^T(f_i - f_{i_j}^s)\|^2 - \beta \sum_{q=1}^{k_2} \|W_1^T(f_i - f_{i_q}^d)\|^2 \right) \tag{10}$$

where $\beta \in [0, 1]$ is a scaling parameter that unifies the different measures of within-class distance and between-class distance.

If we define the coefficient vector as:

$$\omega_i = \left[\underbrace{1, \dots, 1}_{k_1}, \underbrace{-\beta, \dots, -\beta}_{k_2} \right] \quad (11)$$

we can reduce equation (10) to:

$$\begin{aligned} & \arg \min_{W_1} \left(\sum_{j=1}^{k_1} \|W_1^T (f_i - f_{ij}^s)\|^2 - \beta \sum_{q=1}^{k_2} \|W_1^T (f_i - f_{iq}^d)\|^2 \right) \\ &= \arg \min_{W_1} \left(\sum_{j=1}^{k_1+k_2} \|W_1^T (F_i\{1\} - F_i\{j+1\})\|^2 (\omega_i)_j \right) \\ &= \arg \min_{W_1} \text{tr}(W_1^T F_i L_i F_i^T W_1) \end{aligned} \quad (12)$$

We can then create an expression L_i , which encapsulates both the local geometry and the discriminative information:

$$L_i = \begin{bmatrix} \sum_{j=1}^{k_1+k_2} (\omega_i)_j & -\omega_i^T \\ -\omega_i & \text{diag}(\omega_i) \end{bmatrix} \quad (13)$$

From there, $F_i = \{f_i, F_i^s, F_i^d\} \in R^{k_1+k_2+1}$ is selected from the global coordinate $F = \{f_1, f_2, \dots, f_N\} \in R^{m \times N}$, such that:

$$F_i = F S_i \quad (14)$$

where $S_i \in R^{N \times (k_1+k_2+1)}$ is a selection matrix containing elements defined as:

$$(S_i)_{pq} = \begin{cases} 1 & \text{if } p = D_i\{q\} \\ 0 & \text{else} \end{cases} \quad (15)$$

where $D_i = [i, i_1, \dots, i_{k_1+k_2}]$ is the index set for F_i . In order to make the projection matrix W_1 linear and orthogonal, we iterate for all N samples according to the procedure:

$$\begin{aligned} & \arg \min_{W_1} \sum_{i=1}^N \text{tr}(W_1^T F_i L_i F_i^T W_1) \\ &= \arg \min_{W_1} \text{tr} \left(W_1^T F \sum_{i=1}^N (S_i L_i S_i^T) F^T W_1 \right) \\ &= \arg \min_{W_1} \text{tr}(W_1^T F L F^T W_1) \quad \text{s.t. } W_1^T W_1 = I_d \end{aligned} \quad (16)$$

where the alignment matrix $L = \sum_{i=1}^N S_i L_i S_i^T \in R^{N \times N}$. We then use standard Eigen-decomposition to obtain the solution to equation (16):

$$F L F^T u = \lambda u \quad (17)$$

where the column vectors u_1, u_2, \dots, u_d are the solutions of equation (17), ordered according to the eigenvalues $\lambda_1 < \lambda_2 < \dots < \lambda_d$. Therefore, the optimal projection matrix is $W_1 = [u_1, u_2, \dots, u_d]$. Following the same process for W_2 as W_1 , we can obtain the final projection matrix $W = W_1 W_2$. For more detailed technical discussion, please refer to our previous work[17].

4 Experiments

4.1 Data collection

112 children (58 male, 54 female), aged eight to twelve, participated in this study. An ActiGraph GT3X tri-axial accelerometer positioned on the right hip of each participant was used for physical activity measurement during all trials. This was done by a trained research assistant and checked periodically during testing to ensure the accelerometer was in the correct location. When using the ActiGraph, the raw acceleration data first goes through a low pass filter to remove unwanted noise. It next goes through a band-pass filter when converting to epochs, which are what is used for the analysis. Both ActiGraph filters are proprietary. In converting to epochs the data are full-wave rectified and for the time of interest (e.g. 5 sec) the area under the curve is calculated, which gives the epoch value. Numerous studies have shown these devices are reliable and consistent between participants. However they are proprietary by nature so only the manufacturer knows exactly how the output data is generated. As for their influence on human activities, the ActiGraph is battery powered and data is downloaded to a computer after use, thus there are no cables that would restrict the participants' movements.

All participants completed supine rest, and then were randomly assigned to complete six out of the other 18 structured physical activities. Thus, each participant completed a total of seven activities. These 19 activities were then labeled and sorted into five categories:

1. Sedentary behaviors: supine rest, reading, watching TV, and searching the internet.
2. Household chores: sweeping and vacuuming.
3. Locomotion: slow track walking, brisk track walking, walking with a 10-lb backpack, and track running.
4. Interactive video games: Nintendo Wii, floor Light Space, wall Light Space, Dance Dance Revolution (DDR), and Trazer.
5. Exercise and sports: playing catch, soccer around cones, Sport Wall, and workout video.

In order to decrease bias, we implemented each experiment 10 times to establish an average accuracy.

4.2 Feature generation

In accordance with previous studies, bouts were divided into a sequence of non-overlapping windows, each one minute in duration. This was done for ease of comparison, as activities lasted for different lengths of time (all activities were performed for 7-min, except supine rest which lasted for 30-min).

Each 1-minute window was further divided into 1-second intervals. Feature extraction was performed five times for each interval: once for the original features, and once using each of the four wavelet methods for the improved features. DML-KNN and SVM classifiers were then employed to classify and predict labels from the features generated.

4.3 Intensity estimation and activity recognition

Intensity estimation was performed to divide the 19 recorded activities into five categories (sedentary behaviors, and light, moderate, and vigorous physical activity). Activity recognition involved distinguishing the 19 activities from each other. We used the DML-KNN to classify features generated by the time-domain method and also the four different time-frequency classification methods.

4.4 Training validation and precision measure

Three types of training validation were performed: 10-fold cross validation (CV), leave-one-person-out (LOPO) and random splitting (RS).

1. For CV, the raw dataset was divided into ten distinct segments. Labels were predicted for each segment by using the other nine segments as training data. Once this had been done for each of the ten segments, voting was performed in blocks. CV validation is normally used to combat over-fitting.
2. The LOPO method removes one sample from the data set and uses it for validation. The process is repeated for each sample in the set until all have been used for validation. LOPO validation is the closest simulation on how the method operates in practice.
3. The process for RS was first performed by applying a label to 10% of the raw data, with the remaining 90% used for training (RS_90). RS was then repeated eight times, each trial using 10% less of the raw data for training (RS_80 through RS_10). RS Validation is used to evaluate robustness and flexibility.

We evaluated the performance of each method against four different metrics: overall accuracy, precision, recall (or sensitivity), and F-score. Overall accuracy was taken as the proportion of true examples of all activity types that were correct,

$$Accuracy = (TP + TN) / (TP + FP + FN + TN) \quad (18)$$

Where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives. Precision was defined as the proportion of predicted examples of each activity type that were correct, meaning

$$Precision = TP / (TP + FP) \quad (19)$$

Recall was defined as the proportion of true examples of an activity type that were correctly identified, and can therefore be calculated as:

$$Recall = TP / (TP + FN) \quad (20)$$

Precision and recall always conflict such that in cases where precision is high, recall is consistently low, and vice versa. However, if, for example, we erroneously classified all samples as supine rest, the recall result of supine rest would be 100%, but precision would be atrocious. To counteract such anomalies, we use the F-score, a combined measurement of precision and recall that is calculated as:

$$F - score = 2PR / (P + R) \quad (21)$$

Where P is precision, and R is recall.

Table 1 Overall accuracy obtained from combining five different feature extraction methods with SVM and DML-KNN classifiers.

Classifier	Feature Extraction Method					
		Time-domain	Tamura	Nyan	Square	Magnitude
	DML-KNN	86.72	91.37	91.14	88.91	89.86
	SVM	69.34	84.52	83.31	83.98	86.07

5 Results

Using the time-domain features we examined in previous works as a basis for comparison (see Section 2.1), we constructed features using the Tamura, Nyan, Square Coefficient and Magnitude Coefficient methods previously described. These features were classified using DML-KNN, and results were compared to classification performed using a SVM. Finally, we employed a confusion matrix to discern the effects of classification.

5.1 Intensity estimation

Table 1 shows that by combining our proposed time-frequency features with a DML-KNN classifier, we were able to achieve significantly greater classification accuracy than previously utilized methods. Figure 5 displays a comparison of F-score measurements between SVM and DML-KNN. From the figure and table, we can see that for continuous movement DML-KNN averaged only about 3% better than SVM, though SVM did benefit significantly from using time-frequency features over time-domain features. However, for intermittent movement DML-KNN outperformed SVM by an average of 24%, demonstrating DML-KNNs superiority when dealing with sharp high-frequency, and transient information. The time-frequency features produced better results than time-domain features for both classifiers.

From the classification accuracies shown in Table 2, we observe that our proposed time-frequency method can produce more accurate results than previous time-domain classification attempts for youth activities. For cross-validation, the Nyan and Square coefficient methods achieved the highest accuracies, with 92.2% and 92.3%, respectively. For LOPO validation, the Tamura and Nyan methods achieved the best accuracies, at 91.7% and 91.1%, respectively. For RS validation, all four time-frequency methods outperformed time-domain methods, with no dramatic drop in accuracy as the proportion of data used for training was decreased from 90% to 10%—even for RS_10, where the highest accuracy seen was 80.2%.

After predicting category types, voting was performed for individual blocks, which improved accuracy in some cases. Notably, some samples selected for RS validation, unlike for LOPO and CV validation, were not adjacent to any other samples in the same block. Thus, the highest overall accuracy seen for RS was only 88.3%, when 90% of the data was used for training.

As physical activities performed by youth are different from adults, and the volume of accelerometer data is large, we designed an experiment named "shotgun" that focused on accelerometer data with a tendency to being misclassified. In "shotgun" trials, we implement each experiment 10 times, and in each instance we randomly selected 12% of the entire dataset. In Tables 3 through 7, it can be observed that all methods correctly classify

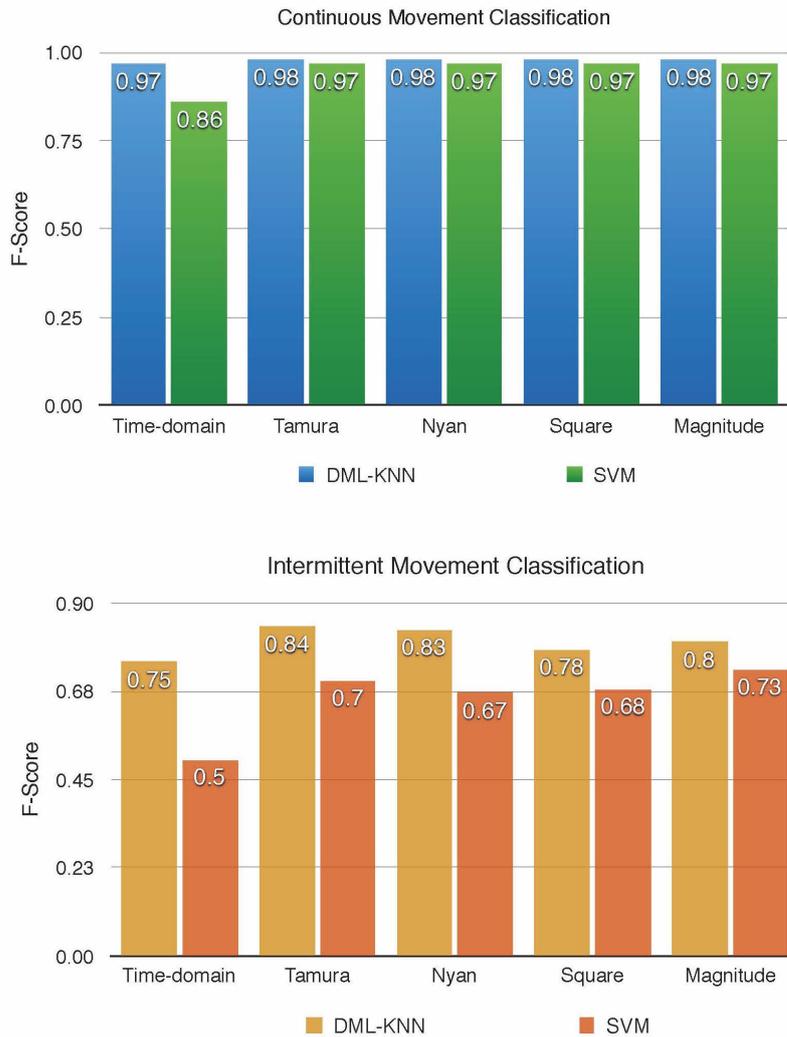


Fig. 5 DML-KNN and SVM classifier F-score measurements for intensity estimation of continuous movement (top) and intermittent movement (bottom) using time-domain features and four time-frequency features.

sedentary behaviors in almost all cases. This can be attributed to the abundance of zero-valued accelerometer measurements associated with sedentary behaviors when compared to activities in other categories. We also found that the four new methods are competitive at identifying activities in all categories, especially in the household chores. Some misclassification was observed for interactive video games and exercise and sports, but the newly implemented feature generation methods still exhibited better accuracy in the field of youth activity.

Table 2 Overall accuracy obtained from five different experimental methods for classifying individual activities, with voting. Each method was evaluated using three types of training validation: 10-fold cross validation (CV), leave-one-person-out (LOPO), and random splitting (RS_90 through RS_10).

	Time-domain	Tamura	Nyan	Square	Magnitude
CV	89.05	91.77	92.21	92.28	91.87
LOPO	86.72	91.73	91.14	88.91	89.86
RS_90	81.98	85.23	85.82	86.88	88.25
RS_80	82.24	85.57	85.34	86.28	86.10
RS_70	84.00	85.01	85.81	86.08	85.55
RS_60	83.49	85.71	86.52	85.94	85.04
RS_50	83.87	86.00	86.55	86.58	85.78
RS_40	83.84	84.92	85.07	85.64	86.50
RS_30	83.04	84.72	84.82	85.29	85.85
RS_20	81.43	82.43	81.96	83.44	82.70
RS_10	77.91	78.78	79.51	79.06	80.15

Table 3 Confusion matrix for categorized data with the time-domain feature generation method. Only LOPO experiments were considered. Each trial was repeated using ten different randomly selected data subsets. Values shown are averages rounded to the nearest integer.

	Sedentary behaviors	Household chores	Locomotion	Interactive video games	Exercise and sports
Sedentary behaviors	245	0	0	0	1
Household chores	1	97	0	7	2
Locomotion	0	1	101	1	0
Interactive video games	8	19	1	73	8
Exercise and sports	8	13	2	18	62

Table 4 Confusion matrix for categorized data with the Tamura feature generation method. Only LOPO experiments were considered. Each trial was repeated using ten different randomly selected data subsets. Values shown are averages rounded to the nearest integer.

	Sedentary behaviors	Household chores	Locomotion	Interactive video games	Exercise and sports
Sedentary behaviors	245	0	0	1	0
Household chores	1	103	0	1	3
Locomotion	0	1	102	0	1
Interactive video games	6	12	0	82	8
Exercise and sports	3	9	1	10	80

Table 5 Confusion matrix for categorized data with the Nyan feature generation method. Only LOPO experiments were considered. Each trial was repeated using ten different randomly selected data subsets. Values shown are averages rounded to the nearest integer.

	Sedentary behaviors	Household chores	Locomotion	Interactive video games	Exercise and sports
Sedentary behaviors	245	0	0	0	1
Household chores	0	103	0	2	2
Locomotion	0	2	101	0	0
Interactive video games	5	13	0	84	7
Exercise and sports	3	12	2	12	75

Table 6 Confusion matrix for categorized data with the Square coefficient feature generation method. Only LOPO experiments were considered. Each trial was repeated using ten different randomly selected data subsets. Values shown are averages rounded to the nearest integer.

	Sedentary behaviors	Household chores	Locomotion	Interactive video games	Exercise and sports
Sedentary behaviors	245	0	0	1	0
Household chores	0	104	0	1	2
Locomotion	0	1	102	0	0
Interactive video games	7	15	0	80	7
Exercise and sports	4	14	2	21	62

Table 7 Confusion matrix for categorized data with the Magnitude coefficient feature generation method. Only LOPO experiments were considered. Each trial was repeated using ten different randomly selected data subsets. Values shown are averages rounded to the nearest integer.

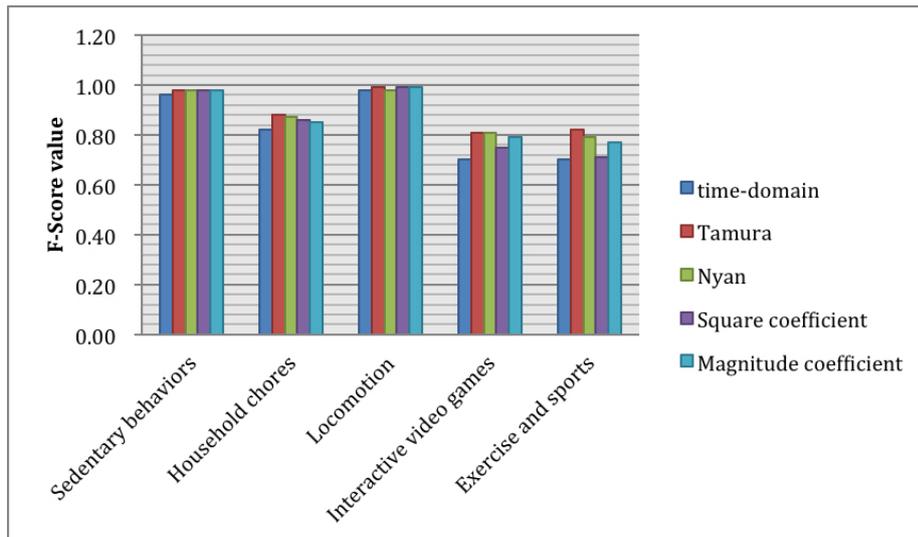
	Sedentary behaviors	Household chores	Locomotion	Interactive video games	Exercise and sports
Sedentary behaviors	245	0	0	0	0
Household chores	0	100	1	4	3
Locomotion	0	0	101	1	0
Interactive video games	6	14	0	82	6
Exercise and sports	6	14	1	13	70

Table 8 shows the intensity estimation performance of each feature generation method, as measured by recall, precision, and F-score. Fig 6 shows a comparison of F-scores achieved by each method, with the Tamura method performing most competitively in youth activity classification. Additionally, the Tamura, Square coefficient and Magnitude coefficient methods performed better in the continuous movement categories, including sedentary behaviors and locomotion. While for the intermittent movement categories, the Tamura and Nyan methods obtained the better F-scores.

Overall, classification of household chores was more successful than interactive video games and exercise and sports. This is likely due to the smaller diversity of postural transitions inherent to household chores, when compared to interactive video games and exercise and sports. The latter produces more sharp frequency changes. Furthermore, the postural transitions observed in household chores occur predominantly when a subject changes direction or posture, but then proceeds to perform the same movements as before. In contrast, postural transitions in interactive video games and exercise and sports tended to indicate a complete shift in the types of movements the subject performed. In the voting phase, different labels would sometimes earn the same number of votes, and ties were broken by selecting the first applicable label from among those with equal votes. Thus, while the correct label may have received as many votes as an incorrect label, it was eliminated due to its placement in the order of activity categories.

Table 8 Recall, precision and F-score measurements of time-frequency features from four wavelet feature generation methods: Tamura, Nyan, Square coefficient, and Magnitude coefficient.

		Recall	Precision	F-score
Sedentary behaviors	Time-domain	0.94	1.00	0.96
	Tamura	0.96	1.00	0.98
	Nyan	0.97	1.00	0.98
	Square coefficient	0.96	1.00	0.98
	Magnitude coefficient	0.95	1.00	0.98
Household chores	Time-domain	0.75	0.91	0.82
	Tamura	0.82	0.95	0.88
	Nyan	0.79	0.96	0.87
	Square coefficient	0.78	0.97	0.86
	Magnitude coefficient	0.78	0.93	0.85
Locomotion	Time-domain	0.97	0.98	0.98
	Tamura	0.99	0.98	0.99
	Nyan	0.98	0.98	0.98
	Square coefficient	0.98	0.99	0.99
	Magnitude coefficient	0.98	0.99	0.99
Interactive video game	Time-domain	0.74	0.67	0.70
	Tamura	0.87	0.76	0.81
	Nyan	0.86	0.77	0.81
	Square coefficient	0.78	0.73	0.75
	Magnitude coefficient	0.82	0.76	0.79
Exercise and sports	Time-domain	0.85	0.60	0.70
	Tamura	0.87	0.78	0.82
	Nyan	0.88	0.72	0.79
	Square coefficient	0.87	0.60	0.71
	Magnitude coefficient	0.89	0.67	0.77

**Fig. 6** F-score results of time-domain, Tamura, Nyan, Square coefficient, and Magnitude coefficient methods.

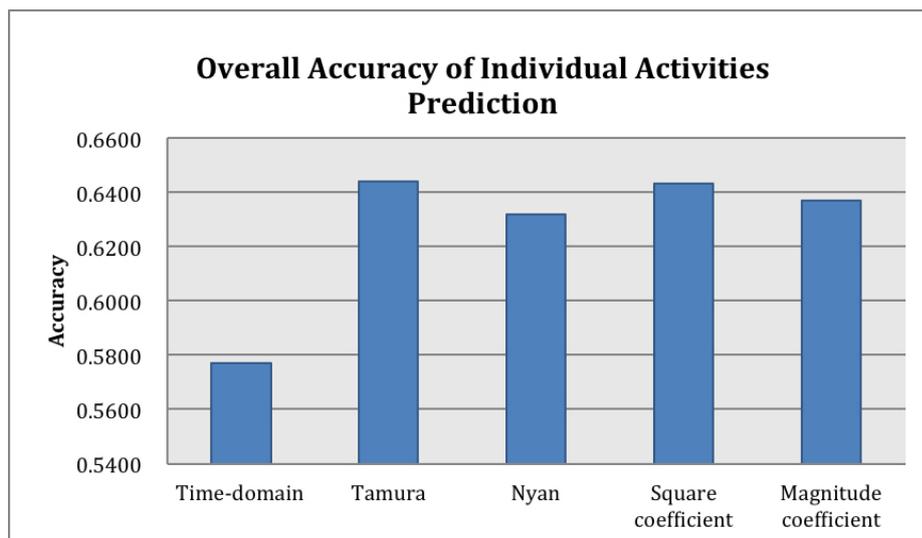


Fig. 7 Overall accuracy of activity recognition results using time-domain, Tamura, Nyan, Square coefficient and Magnitude coefficient methods. Only LOPO experiments were considered.

5.2 Activity Recognition

The four previously mentioned time-frequency methods—Tamura, Nyan, Square Coefficient and Magnitude Coefficient—along with a time-domain method included for comparison, were used to identify the 19 different physical activities performed by our participants. From Figure 7, we can see that the Tamura and the Square Coefficient methods achieved the best overall accuracies in youth activity recognition, with 64.4% and 64.3%, respectively. The Nyan and Magnitude Coefficient methods achieved accuracies comparable to each other, while the time-domain method performed unsatisfactorily with an overall accuracy of 57.7%.

Tables 9 through 11 provide a clearer picture of each method's performance. For household chores, the Squared coefficient and Tamura methods achieved preferable F-scores. For locomotion activities, the Tamura and Nyan methods performed better—achieving the highest F-scores—especially for walking with a 10-lb backpack and track running, which proved troublesome for the other methods. The Square coefficient method achieved the highest F-scores in the interactive video games category, though all methods struggled to classify the wall light space activity, resulting in F-scores between 0.40 and 0.49. The time-domain method was unsatisfactory at recognizing the Nintendo Wii activity than all other methods, with an F-score of only 0.34. Of the exercise and sports activities, soccer around cones was the most consistently recognized by all methods. The Square coefficient method was superior at identifying the workout video activity, with an F-score of 0.73, compared to F-scores of 0.51 to 0.65 for the other methods.

Figure 8 compares the activity recognition performance of the five time-domain and time-frequency features when used in conjunction with the DML-KNN classifier. The four time-frequency features all achieved better F-Scores than the time-domain features for both

Table 9 Recall measurement of activity recognition result.

		Time domain	Tamura	Nyan	Square	Magnitude
Sedentary behaviors	Supine rest	0.55	0.55	0.55	0.55	0.55
	Reading	0.19	0.26	0.25	0.30	0.27
	Watching TV	0.04	0.15	0.08	0.12	0.04
	Searching the internet	0.44	0.41	0.44	0.42	0.43
Household chores	Sweeping	0.55	0.86	0.76	0.86	0.75
	Vacuuming	0.72	0.75	0.71	0.78	0.64
Locomotion	Slow track walking	0.80	0.93	0.93	0.90	0.90
	Brisk track walking	0.74	0.78	0.81	0.81	0.85
	Walking with a 10-lb backpack	0.81	0.89	0.88	0.89	0.93
	Track running	0.92	0.92	0.92	0.80	0.92
Interactive video games	Nintendo wii	0.31	0.67	0.56	0.66	0.59
	Floor light space	0.50	0.62	0.67	0.63	0.67
	Wall light space	0.52	0.34	0.38	0.39	0.43
	Dance dance revolution	0.46	0.59	0.66	0.61	0.66
	Trazer	0.79	0.70	0.74	0.64	0.69
Exercise and sports	Playing catch	0.59	0.59	0.54	0.63	0.57
	Soccer around cones	0.66	0.74	0.74	0.81	0.74
	Sport wall	0.54	0.65	0.63	0.52	0.50
	Workout video	0.43	0.56	0.52	0.68	0.67
Average		0.56	0.63	0.62	0.63	0.62

continuous and intermittent movement classification. Though the time-frequency features performed comparably when analyzing continuous movement, the Square Coefficient method proved best when dealing with intermittent movement, followed by the Tamura and Magnitude Coefficient methods.

6 Conclusions

Our primary objective was to examine common physical activities performed by youth using two approaches: intensity estimation and activity recognition. Data was gathered via a single hip-mounted accelerometer on each of the 112 youth participants. We proposed four different time-frequency wavelet methods, and one previously used time-domain method, to extract features from accelerometer counts, then developed and implemented a semi-supervised learning method for activity classification.

Overall, continuous activities were more accurately classified than intermittent activities, likely owing to postural transitions inherent in the latter group that negatively impact accuracy. Our results show that time-frequency methods are considerably more capable at classifying intermittent movements, which contain sharp-frequency and posture transients—Tamura and Nyan are better at youth activity intensity estimation, while Tamura and Square Coefficient methods are best suited for youth activity recognition. However, due to the more

Table 10 Precisions of one time-domain and four time-frequency methods.

		Time domain	Tamura	Nyan	Square	Magnitude
Sedentary behaviors	Supine rest	0.56	0.63	0.63	0.64	0.62
	Reading	0.17	0.25	0.25	0.23	0.23
	Watching TV	0.14	0.67	0.50	0.60	0.17
	Searching the internet	0.18	0.18	0.19	0.18	0.20
Household chores	Sweeping	0.47	0.62	0.59	0.58	0.60
	Vacuuming	0.54	0.64	0.61	0.70	0.60
Locomotion	Slow track walking	0.80	0.90	0.82	0.87	0.87
	Brisk track walking	0.83	0.91	0.96	0.88	0.92
	Walking with a 10-lb backpack	0.81	0.89	0.92	0.86	0.89
	Track running	0.82	0.88	0.88	0.91	0.88
Interactive video games	Nintendo wii	0.38	0.57	0.62	0.66	0.61
	Floor light space	0.50	0.56	0.57	0.56	0.59
	Wall light space	0.47	0.59	0.42	0.65	0.46
	Dance dance revolution	0.43	0.40	0.51	0.44	0.53
	Trazer	0.92	0.88	0.71	0.90	0.83
Exercise and sports	Playing catch	0.73	0.80	0.78	0.85	0.76
	Soccer around cones	0.90	0.80	0.77	0.78	0.83
	Sport wall	0.82	0.77	0.89	0.74	0.78
	Workout video	0.63	0.75	0.68	0.79	0.67
Average		0.59	0.67	0.65	0.67	0.63

strenuous computational demands of the Square coefficient method, the Tamura method is still preferable for cases where the two methods perform comparably.

Acknowledgements The investigators wish to express their sincerest thanks to National Institutes of Health (R21HL093407) for providing the youth accelerometer data examined in this study.

References

1. Allen FR, Ambikairajah E, Lovell NH, Celler BG (2006) Classification of a known sequence of motions and postures from accelerometry data using adapted gaussian mixture models. *Physiological Measurement* 27(10):935
2. Aminian K, Najafi B (2004) Capturing human motion using body-fixed sensors: outdoor measurement and clinical applications. *Computer Animation and Virtual Worlds* 15(2):79–94
3. Awais M, Mellone S, Chiari L (2015) Physical activity classification meets daily life: Review on existing methodologies and open challenges. In: *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, IEEE*, pp 5050–5053
4. Bao L, Intille SS (2004) Activity recognition from user-annotated acceleration data. In: *Pervasive computing, Springer*, pp 1–17

Table 11 Activity recognition F-scores of one time-domain and four time-frequency methods.

		Time domain	Tamura	Nyan	Square	Magnitude
Sedentary behaviors	Supine rest	0.56	0.59	0.59	0.59	0.59
	Reading	0.18	0.25	0.25	0.26	0.25
	Watching TV	0.06	0.24	0.14	0.19	0.16
	Searching the internet	0.26	0.25	0.27	0.26	0.26
Household chores	Sweeping	0.51	0.72	0.67	0.69	0.68
	Vacuuming	0.62	0.69	0.66	0.74	0.69
Locomotion	Slow track walking	0.80	0.91	0.87	0.88	0.88
	Brisk track walking	0.78	0.84	0.88	0.85	0.86
	Walking with a 10-lb backpack	0.81	0.89	0.90	0.87	0.89
	Track running	0.87	0.90	0.90	0.85	0.88
Interactive video games	Nintendo wii	0.34	0.62	0.59	0.66	0.62
	Floor light space	0.50	0.59	0.62	0.59	0.60
	Wall light space	0.49	0.43	0.40	0.49	0.44
	Dance dance revolution	0.45	0.47	0.58	0.51	0.54
	Trazer	0.85	0.78	0.73	0.75	0.74
Exercise and sports	Playing catch	0.65	0.68	0.64	0.72	0.68
	Soccer around cones	0.76	0.77	0.75	0.79	0.77
	Sport wall	0.65	0.71	0.74	0.61	0.67
	Workout video	0.51	0.64	0.59	0.73	0.65
Average		0.56	0.63	0.62	0.63	0.62

5. Cleland I, Kikhia B, Nugent C, Boytsov A, Hallberg J, Synnes K, McClean S, Finlay D (2013) Optimal placement of accelerometers for the detection of everyday activities. *Sensors* 13(7):9183–9200
6. Crouter SE, Clowers KG, Bassett DR (2006) A novel method for using accelerometer data to predict energy expenditure. *Journal of applied physiology* 100(4):1324–1331
7. Crouter SE, Horton M, Bassett Jr DR (2012) Use of a 2-regression model for estimating energy expenditure in children. *Medicine and science in sports and exercise* 44(6):1177
8. Dalton A, O’Laighin G (2013) Comparing supervised learning techniques on the task of physical activity recognition. *Biomedical and Health Informatics, IEEE Journal of* 17(1):46–52
9. Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D, Marshall S (2014) A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiological measurement* 35(11):2191
10. Ermes M, Parkka J, Mantyjarvi J, Korhonen I (2008) Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *Information Technology in Biomedicine, IEEE Transactions on* 12(1):20–26
11. Freedson PS, Lyden K, Kozey-Keadle S, Staudenmayer J (2011) Evaluation of artificial neural network algorithms for predicting mets and activity type from accelerometer data: validation on an independent sample. *Journal of Applied Physiology* 111(6):1804–1812
12. Hong YJ, Kim IJ, Ahn SC, Kim HG (2010) Mobile health monitoring system based on activity recognition using accelerometer. *Simulation Modelling Practice and Theory*

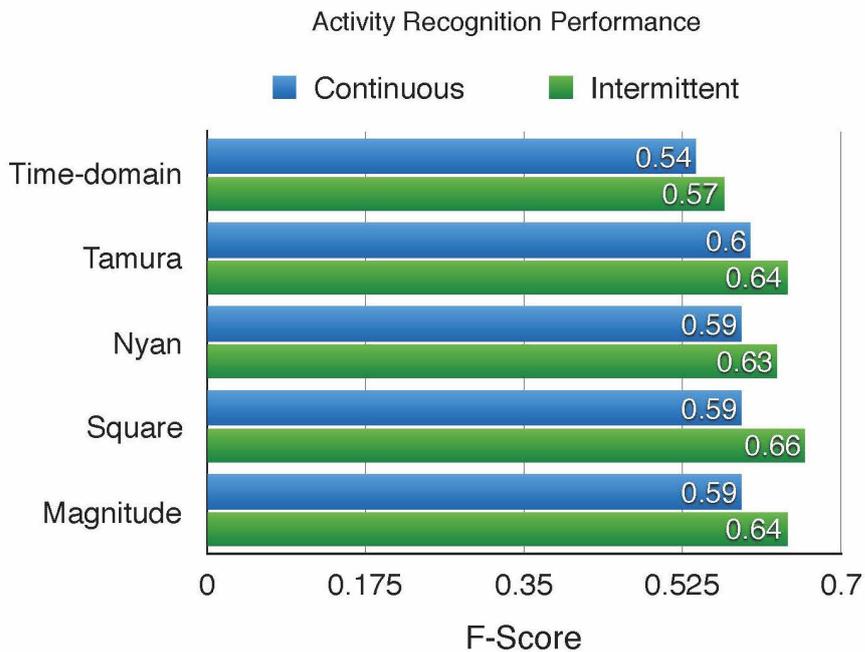


Fig. 8 DML-KNN classifier F-score measurements for activity recognition of continuous movement and intermittent movement using time-domain features and four time-frequency features.

18(4):446–455

13. Khan AM, Lee YK, Lee SY, Kim TS (2010) A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *Information Technology in Biomedicine, IEEE Transactions on* 14(5):1166–1172
14. Lovell N, Wang N, Ambikairajah E, Celler BG (2007) Accelerometry based classification of walking patterns using time-frequency analysis. In: *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE, IEEE*, pp 4899–4902
15. Mannini A, Intille SS, Rosenberger M, Sabatini AM, Haskell W (2013) Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and science in sports and exercise* 45(11):2193–2203
16. Mathie M, Celler BG, Lovell NH, Coster A (2004) Classification of basic daily movements using a triaxial accelerometer. *Medical and Biological Engineering and Computing* 42(5):679–687
17. Mu Y, Lo HZ, Ding W, Amaral K, Crouter SE (2014) Bipart: Learning block structure for activity detection. *Knowledge and Data Engineering, IEEE Transactions on* 26(10):2397–2409
18. Najafi B, Aminian K, Loew F, Blanc Y, Robert P, et al (2002) Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly. *Biomedical Engineering, IEEE Transactions on* 49(8):843–851
19. Nguyen TT, Liew AWC, Tran MT, Nguyen MP (2014) Combining multi classifiers based on a genetic algorithm—a gaussian mixture model framework. In: *Intelligent Com-*

- puting Methodologies, Springer, pp 56–67
20. Nyan M, Tay F, Seah K, Sitoh Y (2006) Classification of gait patterns in the time–frequency domain. *Journal of biomechanics* 39(14):2647–2656
 21. Pärkkä J, Ermes M, Korpipää P, Mäntyjärvi J, Peltola J, Korhonen I (2006) Activity classification using realistic data from wearable sensors. *Information Technology in Biomedicine, IEEE Transactions on* 10(1):119–128
 22. Parkka J, Ermes M, Antila K, van Gils M, Manttari A, Nieminen H (2007) Estimating intensity of physical activity: a comparison of wearable accelerometer and gyro sensors and 3 sensor locations. In: *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE, IEEE*, pp 1511–1514
 23. Pober DM, Staudenmayer J, Raphael C, Freedson PS, et al (2006) Development of novel techniques to classify physical activity mode using accelerometers. *Medicine and science in sports and exercise* 38(9):1626
 24. Preece SJ, Goulermas JY, Kenney LP, Howard D (2009) A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. *Biomedical Engineering, IEEE Transactions on* 56(3):871–879
 25. Preece SJ, Goulermas JY, Kenney LP, Howard D, Meijer K, Crompton R (2009) Activity identification using body-mounted sensors—a review of classification techniques. *Physiological measurement* 30(4):R1
 26. Reiss A, Stricker D (2011) Introducing a modular activity monitoring system. In: *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, IEEE*, pp 5621–5624
 27. Sekine M, Tamura T, Togawa T, Fukui Y (2000) Classification of waist-acceleration signals in a continuous walking record. *Medical engineering & physics* 22(4):285–291
 28. Sekine M, Tamura T, Akay M, Fujimoto T, Togawa T, Fukui Y (2002) Discrimination of walking patterns using wavelet-based fractal analysis. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* 10(3):188–196
 29. Skotte J, Korshøj M, Kristiansen J, Hanisch C, Holtermann A (2014) Detection of physical activity types using triaxial accelerometers. *J Phys Act Health* 11(1):76–84
 30. Staudenmayer J, Pober D, Crouter S, Bassett D, Freedson P (2009) An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *Journal of Applied Physiology* 107(4):1300–1307
 31. Troiano RP, McClain JJ, Brychta RJ, Chen KY (2014) Evolution of accelerometer methods for physical activity research. *British journal of sports medicine* 48(13):1019–1023
 32. Trost SG, Wong WK, Pfeiffer KA, Zheng Y (2012) Artificial neural networks to predict activity type and energy expenditure in youth. *Medicine and science in sports and exercise* 44(9):1801
 33. Trost SG, Zheng Y, Wong WK (2014) Machine learning for activity recognition: hip versus wrist data. *Physiological measurement* 35(11):2183
 34. Veltink PH, Bussmann HB, De Vries W, Martens WL, Van Lummel RC (1996) Detection of static and dynamic activities using uniaxial accelerometers. *Rehabilitation Engineering, IEEE Transactions on* 4(4):375–385
 35. Yi C, Jeong S, Han KS, Lee H (2013) Age-group classification for family members using multi-layered bayesian classifier with gaussian mixture model. In: *Multimedia and Ubiquitous Engineering, Springer*, pp 1153–1159
 36. Zhou Y, Cheng Z, Jing L, Hasegawa T (2015) Towards unobtrusive detection and realistic attribute analysis of daily activity sequences using a finger-worn device. *Applied Intelligence* pp 1–11

-
37. Zijlstra W, Aminian K (2007) Mobility assessment in older people: new possibilities and challenges. *European Journal of Ageing* 4(1):3–12