2015

Prediction of activity type in preschool children using machine learning techniques

M. Hagenbuchner
University of Wollongong, markus@uow.edu.au

Dylan P. Cliff
University of Wollongong, dylanc@uow.edu.au

Stewart Trost
Queensland University of Technology, s.trost@qut.edu.au

Van Tuc Nguyen
University of Wollongong, vtn966@uowmail.edu.au

Gregory E. Peoples
University of Wollongong, peoples@uow.edu.au

Publication Details

Hagenbuchner, M., Cliff, D. P., Trost, S. G., Van Tuc, N. & Peoples, G. E. (2015). Prediction of activity type in preschool children using machine learning techniques. Journal of Science and Medicine in Sport, 18 (4), 426-431.

Prediction of activity type in preschool children using machine learning techniques

Abstract

Objectives Recent research has shown that machine learning techniques can accurately predict activity classes from accelerometer data in adolescents and adults. The purpose of this study is to develop and test machine learning models for predicting activity type in preschool-aged children. **Design** Participants completed 12 standardised activity trials (TV, reading, tablet game, quiet play, art, treasure hunt, cleaning up, active game, obstacle course, bicycle riding) over two laboratory visits. Methods Eleven children aged 3-6 years (mean age = 4.8 ± 0.87; 55% girls) completed the activity trials while wearing an ActiGraph GT3X+ accelerometer on the right hip. Activities were categorised into five activity classes: sedentary activities, light activities, moderate to vigorous activities, walking, and running. A standard feed-forward Artificial Neural Network and a Deep Learning Ensemble Network were trained on features in the accelerometer data used in previous investigations (10th, 25th, 50th, 75th and 90th percentiles and the lag-one autocorrelation). **Results** Overall recognition accuracy for the standard feed forward Artificial Neural Network was 69.7%. Recognition accuracy for sedentary activities, light activities and games, moderate-to-vigorous activities, walking, and running was 82%, 79%, 64%, 36% and 46%, respectively. In comparison, overall recognition accuracy for the Deep Learning Ensemble Network was 82.6%. For sedentary activities, light activities and games, moderate-to-vigorous activities, walking, and running recognition accuracy was 84%, 91%, 79%, 73% and 73%, respectively. **Conclusions** Ensemble machine learning approaches such as Deep Learning Ensemble Network can accurately predict activity type from accelerometer data in preschool children.

Keywords

Physical activity, pattern recognition, accelerometry, neural networks, exercise, validity

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Hagenbuchner, M., Cliff, D. P., Trost, S. G., Van Tuc, N. & Peoples, G. E. (2015). Prediction of activity type in preschool children using machine learning techniques. Journal of Science and Medicine in Sport, 18 (4), 426-431.

Hagenbuchner M, et al. Prediction of activity type in preschool children using machine learning techniques. J Sci Med Sport (2014), http://dx.doi.org/10.1016/j.jsams.2014.06.003

Prediction of Activity Type in Preschool Children using Machine Learning Techniques

Markus Hagenbuchner[a], Dylan P. Cliff[b,*], Stewart G. Trost[c], Nguyen Van Tuc[a], Gregory E. Peoples[d]

a Faculty of Engineering and Information Science, University of Wollongong, Australia

b Faculty of Social Sciences, Early Start Research Institute, University of Wollongong, Australia

c Institute of Health and Biomedical Innovation, Queensland University of Technology, Australia

d School of Medicine, University of Wollongong, Australia

Abstract

Objectives: Recent research has shown that machine learning techniques can accurately predict activity classes from accelerometer data in adolescents and adults. The purpose of this study is to develop and test machine learning models for predicting activity type in preschool-aged children.

Design: Participants completed 12 standardised activity trials (TV, reading, tablet game, quiet play, art, treasure hunt, cleaning up, active game, obstacle course, bicycle riding) over two laboratory visits.

Methods: Eleven children aged 3-6 years (mean age = 4.8 ± 0.87; 55% girls) completed the activity trials while wearing an ActiGraph GT3X+ accelerometer on the right hip. Activities were categorised into five activity classes: sedentary activities, light activities, moderate to vigorous activities, walking, and running. A standard feed-forward Artificial Neural Network (ANN) and a Deep Learning Ensemble Network (DLEN) were trained on features in the accelerometer data used in previous investigations (10th, 25th, 50th, 75th and 90th percentiles and the lag-one autocorrelation).

Results: Overall recognition accuracy for the standard feed forward ANN was 69.7%. Recognition accuracy for sedentary activities, light activities and games, moderate-to-vigorous activities, walking, and running was 82%, 79%, 64%, 36% and 46%, respectively. In comparison, overall recognition accuracy for the DLEN was 82.6%. For sedentary activities, light activities and games, moderate-to-vigorous activities, walking, and running recognition accuracy was 84%, 91%, 79%, 73% and 73%, respectively.

Conclusion: Ensemble machine learning approaches such as DLEN can accurately predict activity type from accelerometer data in preschool children.

Keywords: Physical activity; Pattern recognition; Accelerometry; Neural networks; Exercise; Validity.

Introduction

Due to the limitations of self-reports and pedometers, as well as the intermittent activity patterns of children, accelerometry has become the ’best-practice methodology’ for assessing physical activity (PA) and sedentary behaviour in pre-schoolers, school-aged children and adolescents[1,2]. To interpret accelerometry count data, researchers have typically used cut-points developed from regression or receiver operating characteristic curve analyses to estimate time spent in sedentary behaviour, and light, moderate and vigorous intensity PA. However, conventional regression-based approaches are limited in their ability to accurately predict energy expenditure across a wide range of activities[3,4,5], because the relationship between accelerometer counts and energy expenditure (EE) differs according to the type of activity performed. Not surprisingly, cut-point methods exhibit 28%-45% misclassification of PA intensity in children and adolescents[3,5,6]. As accelerometry use is widespread, this level of misclassification has significant implications for understanding and promoting PA among children and adolescents internationally.

Innovative data processing methodologies such as those utilising machine learning approaches, provide PA researchers with the potential to substantially improve the accuracy of PA measurement. Machine learning is an area of research concerned with the design and development of algorithms that allow computers to “learn” from data. The ability to recognise complex patterns and make intelligent decisions based on data is the main focus of machine learning research. An important class of machine learning algorithms is Artificial Neural Networks (ANN). ANNs are typically applied to applications where the complexity of the data or the task makes the design of alternative approaches impractical.

To date, just two studies have employed ANNs to predict activity type in children and adolescents. Trost and colleagues[6] developed and tested an ANN to classify PA type from second-by-second hip-worn ActiGraph data in 5 to 15 year-olds. Participants completed 12 activity trials that were categorised into 5 activity types: sedentary, walking, running, light intensity house-hold activities or games, and moderate-to-vigorous games or sports. Mean accuracy for activity type ranged from 81.3% to 88.4%. De Vries et al. trained an ANN to predict 9-12 year old children’s PA type from accelerometers worn on the hip and ankle[7]. The overall classification accuracy across the seven activity types evaluated ranged from 57.2% (GT1M/ankle placement) to 76.8% (GT3X/hip placement).

Although the aforementioned studies indicate that machine learning approaches are feasible and offer enhanced accuracy for accelerometry-based assessments of PA in school-aged children and adolescents, the validity of neural networks developed in preschool-aged children has not been investigated. Due to developmental, biomechanical, and behavioural factors, such as differences in motor proficiency[8], and PA types and patterns[1,9], models developed in older children might not be generalizable to young children. To our knowledge, machine learning based accelerometry data modeling approaches are yet to be evaluated in pre-school children. Furthermore, previous models developed in school-aged children and adolescents have been trained and tested using conventional feed-forward ANNs with a single hidden layer, also known as Multi-Layer Perceptron Networks (MLP). Therefore, this study aimed to examine and compare the accuracy of MLP as well as more advanced models, such as a deep-learning-inspired neural network, for predicting PA type in preschool children.

Methods

Eleven children aged 3-6 years (mean age = 4.8 ± 0.87; 55% girls; mean BMI = 15.9 ± 1.0 kg/m^2, 9.1% overweight[10]) were recruited to participate in the study via University staff email lists and word-of-mouth. Parent consent was obtained prior to participation. The study was approved by the University of Wollongong Human Research Ethics Committee.

Participants completed 12 structured activity trials (see Supplementary Table for a description of each activity) over two laboratory visits scheduled within a 3-wk period. Participants undertook the following six trials at visit 1: watching TV (TV), sitting on floor being read to (reading), standing making a collage on a wall (art), walking (walking), playing an active game against an instructor (active game), and completing an obstacle course (obstacle course). The remaining six trials were completed at visit 2: sitting on a chair playing a computer tablet game (tablet), sitting on floor playing quietly with toys (quiet play), treasure hunt (treasure hunt), cleaning up toys (clean-up), bicycle riding (bicycle), and running (running). Each trial was completed for 4-5 min. These 12 activities were then grouped into five activity classes: sedentary activities (TV, reading, tablet, and quiet play), light activities and games (art, treasure hunt, and clean-up), moderate to vigorous activities (active game, obstacle course, and bicycle), walking, and running.

Participants were fitted with an ActiGraph GT3X+ (ActiGraph, Pensacola, FL) on the mid-axillary line at the iliac crest. The GT3X+ records time varying accelerations ranging in magnitude from ±6g. The acceleration output is digitised by a 12-bit analog-to-digital converter at a user-specified rate (30-100 Hz). A sampling frequency of 100 Hz was used in this study.

For each activity trial, 1s count data between minutes 2 and 4 was used for analyses. Since each of the eleven participants performed 12 different activity trials, there were a total of 120s *11 subjects *12 trials = 15,840 instances of data available for the experiments. The 120s segment was divided into non-overlapping time windows. Window sizes of 10s, 15s, 20s, 30s, and **60s** were evaluated (Parameters in bold font indicate the optimal configuration). For each window, features were extracted from those data instances. For ease of comparisons we utilised the same features used by Trost and colleagues[6]. These included the 10th, 25th, 50th, 75th and 90th percentiles and the lag-one autocorrelation values.

Three different ANNs were evaluated in this study: the standard feed-forward Multi-Layer Perceptron Network (MLP), the Self-Organizing Map (SOM), and the Deep Learning Ensemble Network (DLEN). The MLP is a supervised learning model and commonly consists of three layers: input, hidden and output layers[11]. Neurons in those layers are fully connected by a set of adjustable parameters called “weights”. These weights are updated by a learning function which requires an input (training) set consisting of numeric features and associated target values. Consequently, the number of neurons in the input and output layer must match the dimension of input samples and the dimension of class labels respectively. The dimension of the hidden layer can be adjusted freely. The schematic of the MLP is shown in Figure 1(a).

< insert Figure 1 here >

The SOM is an unsupervised learning model that is popularly applied to tasks requiring dimension reduction or clustering[12]. The SOM is computationally very efficient which makes it particularly useful for data mining[12]. Figure 1(b) depicts the schematic of the SOM. Both MLP and SOM take in inputs in the form of vectors. If those vectors are long in size, it refers to the high dimensional input/data space. The SOM can project its input vectors to a 2-dimensional grid referred to as the "activation map", such that each input vector is then represented by a 2-dimensional vector or low dimensional data.

Because the MLP tends to perform poorly when dealing with limited number of samples and high dimensional input space, it makes sense to combine the SOM with MLP since they have complementary properties. The SOM has advantages over the MLP in that the algorithm is trained unsupervised. The resulting model is much less sensitive to "noise" or variability in the data. The MLP on the other hand is trained supervised, and has good generalisation properties. Therefore, adopting concepts from Deep Learning[13], we evaluated the performance of the ensemble model DLEN consisting of a SOM as a first layer, followed by an MLP as a second layer. Both layers were trained on the same set of data with the second layer receiving the output of the first layer as an additional input.

The MLP and SOM models were implemented in plain C programming language. The SOM's parameters including the learning rate was selected from 0.6, 0.8, **1.0**, 1.2 and the radius in 12, 15, 20, **25**. The SOM activation map sizes tried were 19x17, **20x19**, 23x20 and 25x22. A number of MLP configurations were decided by assigning the size of the hidden layer to 3, 8, 13, 17 or **25** and the learning rate to 0.001, **0.01** or 0.5. For each validation round, the MLP and SOM were evaluated 10 times using different random initial conditions. The trained models providing the best performance on the training set was selected to produce the result for the test set. Both the MLP and SOM were trained for 10,000 iterations.

The leave-one-subject-out cross validation approach was used for model assessment. Thus, the model was trained on all input samples except for the data of one participant as the test set. After training, the model was then tested on the left-out data. The experiment was repeated until each participant was considered exactly once for testing. For comparison purposes the MLP results served as the baseline. The confusion matrix and overall accuracy (ACC) were reported.

Results

Table 1 presents confusion matrices for the three ANNs with window sizes of 10s, 30s, and 60s. The average recognition accuracy for 60s windows for the MLP, SOM, and DLEN was 69.7%, 53.8%, and 82.6%, respectively. With 10 s windows, recognition accuracy decreased marginally to 60.6%, 51.5%, and 72.0%, respectively. The performance improvement of the DLEN was largely derived from an increase in the ability to predict walking and running. In particular, relative to MLP, the accuracy of the DLEN improved from 45.5% to 72.7% for running, and from 36.4% to 72.7% for walking. Similar improvements in walking recognition were observed when compared to the SOM; however the SOM failed to recognise any running windows. The confusion matrices show that the MLP and SOM commonly confused walking and running with more generic classes such as light activities and games and moderate activities.

< Insert Table 1 here >

Accelerometry data is available from an earlier study evaluating the accuracy of a standard MLP in 5–15 year-old school-aged children and adolescents (n = 100) [6]. As in the current study, activity trials were categorised into five activity classes; sedentary activities, light house-hold activities or games, moderate-to-vigorous games and sports, walking, and running. We hypothesised that the MLP model would provide similar recognition accuracy to that reported by Trost et al.[6] and that the DLEN would provide higher recognition accuracy than the standard MLP or SOM. The results are summarised in Table 2. In agreement with the results of Trost et al.[6] recognition accuracy for the MLP was 88.4%. Recognition accuracy for the SOM and DLEN was higher than that observed for preschool children at 75.1% and 89.7%.

< Insert Table 2 here >

Discussion

To our knowledge, this is the first study to develop, test and compare neural networks to classify activity type from accelerometer data in preschool-aged children. The findings indicate that a standard feed-forward MLP using the feature set described by Staudenmayer and colleagues[14] and tested in children and adolescents by Trost et al.[6] exhibited fair to poor recognition accuracy (69.7%) for classifying PA type in young children. However, through the application of a deep-learning-inspired ensemble network, substantial improvements in recognition accuracy were achieved (82.6%). The recognition of walking and running increased most substantially from 36.4% and 45.5% to 72.7% in both cases. When the DLEN was tested in a sample of school-aged children and adolescents, recognition of PA type from processed 1Hz accelerometer data (89.7%) was higher than for young children, although minimal gains were achieved relative to MLP (88.4%).

When compared to MLP, DLEN provided substantially improved recognition accuracy in preschool-aged children. Improvements were most prominent for walking (+36%), running (+27%), moderate-to-vigorous activities (+15%), and light activities and games (+12%). The SOM as a pre-training module in DLEN brings about two benefits. The first benefit is that SOM reduces dimensionality of the problem by reducing the number of potential solutions[13], assisting the MLP to find an optimal solution during the training procedure. The second benefit is that SOM is flexible in the mapping size[13], which can assist in identifying distinctions between the activity classes by "stretching out" the data, allowing it to handle potential heterogeneity in the classes.

Overall recognition accuracy for the standard feed-forward MLP was lower among preschool-aged children than previously found among school-aged children and adolescents (88.4%).[6] Trost and colleagues found that recognition accuracy for sedentary activities and walking exceeded 90%, whereas running trials were correctly recognised 79% of the time[6]. Among preschool-aged children, the MLP correctly recognised sedentary activities 82% of the time, however, recognition accuracy for walking and running were considerably lower at 36% and 46%, respectively. For walking, 9%, 27%, and 27% of 60s windows were misclassified as sedentary, light activities and games, and running, respectively. Likewise, 27% and 27% of running trial windows were misclassified as light activities and games and walking, respectively. One explanation for these contrasting findings is that there may have been more variability in the data for preschool children, possibly because the “hybrid classes”, such as light activities and games, may have been more heterogeneous (e.g., treasure hunt and clean-up may have included some walking and running) compared to the data for children and adolescents. Two of the activities in Trost et al.’s[6] light-intensity household activities or games category (floor sweep and laundry task) had significant periods of walking, just as the moderate-to-vigorous intensity games and sports trial of basketball included significant periods of walking and running. Thus, both the present study and the study by Trost and colleagues included heterogeneous activity classes that were distinct from continuous walking or running in isolation. Nevertheless, because preschool children performed different types of activities and exhibited greater variability in performing them, we believe the greater improvement in performance provided by DLEN over the standard MLP in preschoolers was a function of DLEN’s ability to handle learning problems with limited number of samples and to accommodate more complex movement patterns. Likewise, the larger sample size for children and adolescents may have contributed to the higher recognition accuracy for MLP in that age group, because this provided the model with a greater number of correct solutions for each activity class during training.

Closer inspection of the misclassifications which occurred between the sedentary and moderate-to-vigorous PA classes revealed that these misclassifications only occurred for tablet and quiet play, during which two participants exhibited a significant degree of body movement. Similarly, instances of bicycle riding (three cases), active game (two cases) and obstacle course (one case) were misclassified as sedentary. The hip remains largely inactive during cycling, possibly explaining these misclassifications. The other cases were misclassified because the participants had a rest period during the trial (i.e. stopping and standing still).

Model accuracy was optimised using 60s windows, however, shorter windows may be required to characterise the pulsatile and sporadic nature of preschoolers' free-living PA.[1] Importantly, other than for walking, reducing the window size to 10s had a limited impact on the DLEN's accuracy. As walking was consistently misclassified as running, the accuracy for a single "active locomotion" (walk and run combined) category, which is justified given that recommendations for preschoolers focus on total (light, moderate and vigorous) PA[15], would be 72.7%, 77.3%, and 81.2% for windows of 10s, 30s, and 60s, respectively, based on these findings. Thus, the DLEN using a 10s window may be a viable option for field-based studies.

This study had a number of strengths. It is the first to evaluate machine learning approaches to accelerometry data analysis in preschool-aged children. The activity protocol adhered to best practice recommendations[16]; it included a wide variety of common developmentally-appropriate activities ranging in intensity from sedentary to vigorous, and including both ambulatory and free-living tasks. An innovative modelling approach that has not yet been explored in PA research, involving a deep-learning-inspired ensemble neural network, was examined. This alternative model was compared to a standard ANN that has shown promising results in adults[14], and children and youth[6]. Further, the capability of DLEN for recognising PA type was confirmed when tested in a large sample of school-aged children. The experiments have shown that the use of a more suitable classifier can improve the accuracy more substantially than would be obtainable from an increased sample size; DLEN improved accuracy from 69.7% to 82.6% whereas a 10-fold sample size increase improved the performance from 82.6% to just 89.7%.

Some limitations should also be considered when interpreting the findings. Although the number of available data points was sufficient to evaluate and compare different machine learning models, the relatively small number of preschool-aged participants might influence the generalisability of the findings. Likewise, the activity trials were completed in a controlled laboratory environment that might not reflect the free-living behaviours of young children. Similarly, the inclusion of frequency domain features could potentially improve the discrimination between pattern classes. Therefore, larger studies using simulated or entirely free-living activity protocols are required to test the accuracy of machine learning approaches against the criterion measure of direct observation for activity type recognition in preschool children. To accommodate the intermittent activity patterns of young children, future studies should use high frequency raw acceleration signal and extract features over shorter time windows (< 5s).

Conclusion

Neural networks can be used to predict activity type using a single waist-mounted accelerometer in preschool-aged children. Compared to a standard feed-forward MLP, a deep-learning-inspired ensemble neural network provided enhanced accuracy among preschool children, and comparable accuracy in school-aged children. These results contribute to an emerging body of evidence supporting the application of pattern recognition approaches to accelerometry data analysis in children and youth.

Practical Implications

- Neural networks can be used to accurately predict activity type from waist-mounted accelerometry data in preschool-aged children.
- Compared to the accuracy of a standard feed-forward ANN (MLP) for recognising activity type, a deep-learning-inspired ensemble neural network provided the best accuracy among preschool children.
- The standard MLP and DLEN provided comparable accuracy for predicting activity type in school-aged children and adolescents.

- Improved accuracy for measuring physical activity in preschool children can assist in understanding and promoting physical activity in young children, and can contribute to addressing important questions such as: i) how much and which types of activity are important for health?, ii) how active are preschool children?, iii) what are the key determinants of physical activity, and iv) which strategies are most effective for promoting physical activity in young children?

Acknowledgments

We thank Anja Grossek, Melidna Smith, Christel van Loo and Rachel Jones for assistance with data collection assistance and protocol development.

References

1. Cliff DP, Reilly JJ, Okely AD. Methodological considerations in using accelerometers to assess habitual physical activity in children aged 0-5 years. *J Sci Med Sport* 2009;**12**(5):557–567.

2. Trost S. State of the art reviews: measurement of physical activity in children and adolescents. *Am J Lifestyle Med* 2007;**1**(4):299–314.

3. Janssen X, Cliff DP, Reilly JJ, et al. Predictive validity and classification accuracy of ActiGraph energy expenditure equations and cut-points in young children. *PLoS ONE*. 2013;**8**(11):e79124.

4. Trost SG, Loprinzi PD, Moore R, et al. Comparison of accelerometer cut-points for predicting activity intensity in youth. *Med Sci Sports Exerc* 2011;**43**(7):1360–1368.

5. Trost SG, Way R, Okely AD. Predictive validity of three Actigraph energy expenditure equations for children. *Med Sci Sports Exerc* 2006;**38**(2):380–387.

6. Trost SG, Wong W-K, Pfeiffer KA, et al. Artificial neural networks to predict activity type and energy expenditure in youth. *Med Sci Sports Exerc* 2012;**44**(9):1801–1809.

7. De Vries SI, Engels M, Galindo Garre F. Identification of children's activity type with accelerometer-based neural networks. *Med Sci Sports Exerc* 2011;**43**(10):1994–1999.

8. Gallahue DL, Ozmun JC. Goodway JD. Understanding Motor Development: Infants, Children, Adolescents, Adults. 7th ed. New York: McGraw-Hill 2012.

9. Obeid J, Nguyen T, Gabel L, et al. Physical activity in Ontario preschoolers: prevalence and measurement issues. *Appl Physiol Nutr Metab* 2011;**36**:291**–**297.

10. Cole TJ, Bellizzi MC, Flegal KM, et al. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000; **320**: 1-6.

11. Haykin S. Neural Networks: A Comprehensive Foundation. 2nd ed. Prentice Hall 1999.

12. Kohonen T. Self-Organizing Maps. In: Geva S, Kamps J, Trotman A, eds. Springer Series in Information Sciences. Berlin, Heidelberg: Springer 1995.

13. Schmidhuber J. Learning complex, extended sequences using the principle of history compression. *Neural Comput* 1992;**4**(2):234–242.

14. Staudenmayer J, Pober D, Crouter S, et al. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *J Appl Physiol* 2009;**107**:1300–1307.

15. Australian Government Department of Health and Ageing. Australia's physical activity and sedentary behaviour guidelines. http://www.health.gov.au/internet/main/publishing.nsf/content/health-pubhlth-strateg-phys-act-guidelines (accessed May 2014).

16. Bassett Jr DR, Rowlands A, Trost SG. Calibration and validation of wearable monitors. *Med Sci Sports Exerc* 2012;**44**(1S): S32-S38.

Table 1: Leave-one-out performance of ANN models among preschool children (n = 11).

	Activity Type	Window	ANN classification of Activity Type 1	2	3	4	5	Total ACC
SOM	1.Sedentary	10s	**0.750[33]**	0.068[3]	0.159[7]	0.000[0]	0.023[1]	
		30s	**0.727[32]**	0.159[7]	0.091[4]	0.000[0]	0.023[1]	
		60s	**0.591[26]**	0.318[14]	0.068[3]	0.023[1]	0.000[0]	
	2.Light activities and games	10s	0.091[3]	**0.515[17]**	0.182[6]	0.121[4]	0.091[3]	
		30s	0.212[7]	**0.515[17]**	0.212[7]	0.061[2]	0.000[0]	
		60s	0.000[0]	**0.576[19]**	0.303[10]	0.000[0]	0.121[4]	
	3.Moderate-to-vigorous activities	10s	0.303[10]	0.091[3]	**0.545[18]**	0.000[0]	0.061[2]	
		30s	0.212[7]	0.152[5]	**0.636[21]**	0.000[0]	0.000[0]	
		60s	0.152[5]	0.152[5]	**0.667[22]**	0.030[1]	0.000[0]	
	4.Walking	10s	0.182[2]	0.818[9]	0.000[0]	**0.000[0]**	0.000[0]	
		30s	0.091[1]	0.727[8]	0.000[0]	**0.000[0]**	0.182[2]	
		60s	0.091[1]	0.545[6]	0.000[0]	**0.364[4]**	0.000[0]	
	5.Running	10s	0.091[1]	0.909[10]	0.000[0]	0.000[0]	**0.000[0]**	**0.515**
		30s	0.182[2]	0.364[4]	0.182[2]	0.273[3]	**0.000[0]**	**0.530**
		60s	0.000[0]	0.727[8]	0.273[3]	0.000[0]	**0.000[0]**	**0.538**
MLP	1.Sedentary	10s	**0.750[33]**	0.136[6]	0.068[3]	0.000[0]	0.045[2]	
		30s	**0.750[33]**	0.114[5]	0.068[3]	0.000[0]	0.068[3]	
		60s	**0.818[36]**	0.068[3]	0.068[3]	0.023[1]	0.023[1]	
	2.Light activities and games	10s	0.061[2]	**0.788[26]**	0.091[3]	0.000[0]	0.061[2]	
		30s	0.000[0]	**0.879[29]**	0.030[1]	0.000[0]	0.091[3]	
		60s	0.030[1]	**0.788[26]**	0.061[2]	0.091[3]	0.030[1]	
	3.Moderate-to-vigorous activities	10s	0.182[6]	0.182[6]	**0.606[20]**	0.000[0]	0.030[1]	
		30s	0.182[6]	0.182[6]	**0.606[20]**	0.000[0]	0.030[1]	
		60s	0.212[7]	0.121[4]	**0.636[21]**	0.000[0]	0.030[1]	
	4.Walking	10s	0.091[1]	0.727[8]	0.000[0]	**0.000[0]**	0.182[2]	
		30s	0.000[0]	0.727[8]	0.000[0]	**0.000[0]**	0.273[3]	
		60s	0.091[1]	0.273[3]	0.000[0]	**0.364[4]**	0.273[3]	
	5.Running	10s	0.000[0]	0.727[8]	0.091[1]	0.091[1]	**0.091[1]**	**0.606**
		30s	0.000[0]	0.636[7]	0.091[1]	0.091[1]	**0.182[2]**	**0.636**
		60s	0.000[0]	0.273[3]	0.000[0]	0.273[3]	**0.455[5]**	**0.697**

DLEN	1.Sedentary	10s	**0.750[33]**	0.182[8]	0.068[3]	0.000[0]	0.000[0]	
		30s	**0.795[35]**	0.136[6]	0.068[3]	0.000[0]	0.000[0]	
		60s	**0.841[37]**	0.045[2]	0.091[4]	0.023[1]	0.000[0]	
	2.Light activities and games	10s	0.000[0]	**0.818[27]**	0.121[4]	0.030[1]	0.030[1]	
		30s	0.000[0]	**0.879[29]**	0.061[2]	0.061[2]	0.000[0]	
		60s	0.030[1]	**0.909[30]**	0.030[1]	0.030[1]	0.000[0]	
	3.Moderate-to-vigorous activities	10s	0.121[4]	0.091[3]	**0.727[24]**	0.000[0]	0.061[2]	
		30s	0.121[4]	0.061[2]	**0.758[25]**	0.000[0]	0.061[2]	
		60s	0.182[6]	0.030[1]	**0.788[26]**	0.000[0]	0.000[0]	
	4.Walking	10s	0.000[0]	0.182[2]	0.091[1]	**0.364[4]**	0.364[4]	
		30s	0.000[0]	0.091[1]	0.091[1]	**0.364[4]**	0.455[5]	
		60s	0.000[0]	0.182[2]	0.000[0]	**0.727[8]**	0.091[1]	
	5.Running	10s	0.000[0]	0.273[3]	0.000[0]	0.091[1]	**0.636[7]**	**0.720**
		30s	0.000[0]	0.273[3]	0.000[0]	0.091[1]	**0.636[7]**	**0.758**
		60s	0.000[0]	0.182[2]	0.000[0]	0.091[1]	**0.727[8]**	**0.826**

Values in boldface indicate the proportion [absolute value] of time segments correctly classified.

ANN: Artificial Neural Network; SOM: Self-Organizing Map; MLP: Multi-layer Perceptron network;

DLEN: Deep Learning Ensemble Network.

Table 2: Performance of ANN models among school-aged children and adolescents (n = 100) using a 60s window.

		ANN classification of Activity Type					**Total**
	Activity Type	**1**	**2**	**3**	**4**	**5**	**ACC**
SOM	1.Sedentary	**0.960[95]**	0.040[4]	0.000[0]	0.000[0]	0.000[0]	
	2.Light HH and games	0.242[24]	**0.727[72]**	0.020[2]	0.000[0]	0.010[1]	
	3. Moderate-to-vigorous	0.045[3]	0.364[24]	**0.561[37]**	0.000[0]	0.030[2]	
	4.Walking	0.010[1]	0.152[15]	0.141[14]	**0.606[60]**	0.091[9]	
	5.Running	0.000[0]	0.030[1]	0.212[7]	0.121[4]	**0.636[21]**	**0.719**
MLP	1.Sedentary	**0.939[93]**	0.051[5]	0.000[0]	0.010[1]	0.000[0]	
	2.Light HH and games	0.172[17]	**0.808[80]**	0.010[1]	0.010[1]	0.000[0]	
	3.Moderate-to-vigorous	0.015[1]	0.061[4]	**0.909[60]**	0.015[1]	0.000[0]	
	4.Walking	0.010[1]	0.000[0]	0.030[3]	**0.929[91]**	0.030[3]	
	5.Running	0.000[0]	0.000[0]	0.091[3]	0.152[5]	**0.758[25]**	**0.884**
DLEN	1.Sedentary	**0.949[94]**	0.051[5]	0.000[0]	0.000[0]	0.000[0]	
	2.Light HH and games	0.172[17]	**0.818[81]**	0.010[1]	0.000[0]	0.000[0]	
	3.Moderate-to-vigorous	0.015[1]	0.031[2]	**0.939[62]**	0.015[1]	0.000[0]	
	4.Walking	0.010[1]	0.000[0]	0.051[5]	**0.929[92]**	0.010[1]	
	5.Running	0.000[0]	0.000[0]	0.091[3]	0.121[4]	**0.788[26]**	**0.897**

Values in boldface indicate the proportion [absolute value] of time segments correctly classified.

ANN: Artificial Neural Network; SOM: Self-Organizing Map; HH: house-hold; MLP: Multi-layer Perceptron network; DLEN: Deep Learning Ensemble Network.

Figure Legends

Figure 1: (a) the schematic of an MLP. Shown is an MLP that takes the mappings of a SOM as additional input, which is referred to as the DLEN, and (b) the schematic of a two-dimensional SOM of size 3x3.

Supplementary Table: Activity Trials

Activity Class	Activity Trial	Description
Sedentary	Watching TV	Sit in a comfortable chair watching TV.
	Tablet computer activity	Sit in a chair at a table completing a developmentally-appropriate puzzle activity on a computer tablet.
	Reading	Sit on the floor on a cushion and listen to a story-book.
	Quiet play	Sit on the floor playing with toys/blocks/puzzles/dolls.
Light activities and games	Cleaning up toys	Collect toys and equipment and return them to appropriate boxes.
	Standing art	Create a collage on a whiteboard by sticking art materials onto contact paper.
	Treasure hunt	Walk through the activity room (20m x 10m) and search for and collect hidden sea creatures.
Moderate-to-vigorous activities	Bicycle riding	Ride a bicycle around the activity room (one lap = 45m), with or without training wheels, as selected by parent/child.
	Obstacle course	Move through an obstacle course involving jumping through hoops, crawling through a tunnel, hopping, climbing up foam stairs and jumping down.
	Active game	*Clean up your backyard* - Try to keep your playing area (4m x 3m) “clean” by throwing all bean-bags onto the instructors playing area. The instructor will do the same. Game ends when your playing area is clean (Based on child’s ability, instructor increases/decreases difficulty by playing faster/slower).
Walking	Walking	Walk with instructor at a self-selected comfortable speed around the marked perimeter of the activity room (one lap = 45m)
Running	Running	Run with instructor at a self-selected speed around the marked perimeter of the activity room (one lap = 45m)