Journal of Biomechanics 48 (2015) 3975–3981

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Classification of team sport activities using a single wearable tracking device

Daniel W.T. Wundersitz [a,*], Casey Josman [b], Ritu Gupta [b], Kevin J. Netto [c], Paul B. Gastin [a], Sam Robertson [a,d]

[a] Centre for Exercise & Sports Science, School of Exercise & Nutrition Sciences, Deakin University, Melbourne, Victoria, Australia
[b] Department of Mathematics and Statistics, Curtin University, Perth, Western Australia, Australia
[c] School of Physiotherapy and Exercise Science, Curtin University, Perth, Western Australia, Australia
[d] Institute of Sport, Exercise and Active Living, Victoria University, Melbourne, Victoria, Australia

ARTICLE INFO

Article history:
Accepted 24 September 2015

Keywords:
Accelerometer
Gyroscope
Random Forest
Logistic Regression Tree
Support Vector Machine

ABSTRACT

Wearable tracking devices incorporating accelerometers and gyroscopes are increasingly being used for activity analysis in sports. However, minimal research exists relating to their ability to classify common activities. The purpose of this study was to determine whether data obtained from a single wearable tracking device can be used to classify team sport-related activities. Seventy-six non-elite sporting participants were tested during a simulated team sport circuit (involving stationary, walking, jogging, running, changing direction, counter-movement jumping, jumping for distance and tackling activities) in a laboratory setting. A MinimaxX S4 wearable tracking device was worn below the neck, in-line and dorsal to the first to fifth thoracic vertebrae of the spine, with tri-axial accelerometer and gyroscope data collected at 100 Hz. Multiple time domain, frequency domain and custom features were extracted from each sensor using 0.5, 1.0, and 1.5 s movement capture durations. Features were further screened using a combination of ANOVA and Lasso methods. Relevant features were used to classify the eight activities performed using the Random Forest (RF), Support Vector Machine (SVM) and Logistic Model Tree (LMT) algorithms. The LMT (79–92% classification accuracy) outperformed RF (32–43%) and SVM algorithms (27–40%), obtaining strongest performance using the full model (accelerometer and gyroscope inputs). Processing time can be reduced through feature selection methods (range 1.5–30.2%), however a trade-off exists between classification accuracy and processing time. Movement capture duration also had little impact on classification accuracy or processing time. In sporting scenarios where wearable tracking devices are employed, it is both possible and feasible to accurately classify team sport-related activities.

1. Introduction

Objective measurement of sports activities is essential for understanding the physical and technical demands related to sports performance (Aughey and Falloon, 2010). It is also important in evaluating the effectiveness of training programs designed to increase performance as well as those targeting both the prevention and rehabilitation of injury (Neville et al., 2010). Fundamental to furthering these understandings is the need to accurately collect specific information relating to the type, intensity and frequency of activities performed (Carling et al., 2009). Consequently, attempts to improve the techniques related to activity analysis in sports have been made in recent years.

At least partially responsible for these improvements are the considerable developments that have occurred in commercially available wearable tracking device technology. Wearable tracking devices typically integrate multiple sensors (e.g., global positioning system [GPS], accelerometer and gyroscope) into a single, versatile unit often worn on the upper back in a sports vest (Kelly et al., 2012). To date, the majority of research has focused on the GPS sensor contained within these devices to obtain basic descriptors of sports activities, such as speed, distance travelled, and the number of high-intensity efforts performed (Cummins et al., 2013). However, evidence suggests that more detailed analysis can be obtained using the accelerometer sensor (Ermes et al., 2008). Specifically, different activity types can be classified based on the features of the accelerometer signal.

* Correspondence to: Deakin University, School of Exercise and Nutrition Sciences, 221 Burwood Highway, Burwood, 3125 Victoria, Australia.
Tel.: +61 3 92417277; fax: +61 3 9244 6017.
E-mail address: dwunder@deakin.edu.au (D.W.T. Wundersitz).

http://dx.doi.org/10.1016/j.jbiomech.2015.09.015
0021-9290/

McNamara et al. (2015) developed a bowling detection algorithm for cricket. The researchers found that the algorithm was able to classify cricket bowling more effectively in training than game-play, with a maximum accuracy of 98.1% (training). Kelly et al. (2012) applied support vector machine (SVM) and hidden conditional random field algorithms to automatically detect tackling in rugby. The algorithm was able to consistently classify collisions, with a maximum accuracy of 95%. Similarly, Gastin et al. (2013) assessed the concurrent validity of a manufacturer-developed tackle detection algorithm (Catapult Sports), which was compared against video-replay and coded into three intensity categories. The researchers found a maximum classification accuracy of 78%, with tackled players more accurately detected than the players initiating the tackle. However, during game-play the algorithm was only able to correctly detect tackles 18% of the time. Although these findings are promising, more sophisticated and generalisable sport and activity specific algorithms are required (Gastin et al., 2014).

Mitchell et al. (2013) recently proposed a method using a single accelerometer contained within a smartphone worn on the upper-back, with the aim of identifying seven different sporting activities (stationary, walking, jogging, sprinting, hitting a ball, standing tackle, dribbling a ball). An overall activity classification success rate of 75% was achieved using classification approaches that included SVM, Logistic Model Tree (LMT), and a range of Neural Network/Optimisation type classifiers. With the aim of achieving higher classification accuracy, multiple sensors (i.e., both accelerometer and gyroscope) have also been considered in the literature, rather than a single accelerometer sensor alone (e.g. (Leutheuser et al., 2013; Najafi et al., 2003)). Gyroscopes are insensitive to linear accelerations and gravity, and provide essential information pertaining to the rotational motions of the body during human activity (Kunze et al., 2010). As a gyroscope sensor is typically contained within most wearable tracking devices, this would appear to be a feasible approach to aid in the ability to classify of sporting activities.

Another important methodological consideration in the classification literature relates to the duration over which the activity is measured (movement capture duration) for a given classification algorithm (Trost et al., 2012). The optimal duration will ideally be long enough to capture the entire activity as it occurs, while also being short enough to not include any additional activities (Mitchell et al., 2013). Previous work classifying activity type has extracted features in accelerometer data from movement capture durations as short as 0.1 s (Bulling et al., 2014) or as long as 60 s (Trost et al., 2012). In team sports, however, most activities (sprinting, jumping, tackling etc.) can be performed over much shorter durations. For example, the lowest intensity movement (walking) occurs approximately 1.4–2.2 times per second (e.g. (Peacock et al., 2014)). Therefore, much shorter movement capture durations (e.g. 1.5 s or less) may be required to capture activities in team sports. Further, this may improve classification accuracy of these activities, as more periods are available for training (Mannini et al., 2013).

The aims of this study were threefold. First, to determine whether data obtained from wearable tracking device inputs (specifically, gyroscope and accelerometer sensors) alone or in combination can be used to classify team sport-related activities. Second, to determine the ability of three classification algorithms (LMT, RF, SVM) and movement capture durations (0.5, 1.0 and 1.5 s) for feature extraction to classify activities in team sports. Third, to consider the processing time and data collection burdens associated with these methods and identify the best option for practitioners.

2. Methods

2.1. Participants

Seventy-six recreationally active, healthy male participants (age 24.4 ± 3.3 years; height 181.8 ± 7.5 m; mass 77.4 ± 11.6 kg; mean $\pm$ SD) were recruited for participation in the study. All participants were regular competitors in one or more contact-based team sport events per week at the time of testing. The study protocol was approved by the relevant University Human Ethics Advisory Group (HEAG-H 135_2013); all procedures followed ethical guidelines for human research and participants provided written informed consent prior to participating.

The simulated team sport circuit involved a modified version of a circuit developed by Singh et al. (2010) and reported in Wundersitz et al. (2015b). Each circuit included three counter-movement jumps, an eight metre jog, an eight metre change of direction agility section (COD), two jumps for distance, a 10 m sprint, seven metres of walking, and a tackle bag to be taken to ground with maximum force. After the completion of each activity, the participant stood in a stationary position for approximately one second before commencing the next (i.e., three counter-movement jumps were performed in a row then a one second pause occurred). Stationary pauses ensure that there is no bias in data accumulation (e.g., influence of previous activity (Bayat et al., 2014)). Each individual circuit took approximately 45 s to complete, allowing the participant 15 s of rest before completing the next circuit (on 1 min). Each participant completed the circuit six times. During testing, each participant wore a single, wearable tracking device (MinimaxX S4, Catapult Innovations, Australia), which contained (among other sensors not utilised for this study) a 100 Hz tri-axial accelerometer and gyroscope. The device was worn in a tightly fitted manufacturer supplied sports vest and located below the neck (on the upper trunk), in-line and dorsal to the first to fifth thoracic vertebrae of the spine.

2.2. Data processing

The data collected comprised of accelerometer (X, Y, Z axes [$n_1=3$] and the resultant vector [$n_2=1$]) and gyroscope (X, Y, Z axes [$n_3=3$]) inputs for the duration of the circuit (seven total inputs; $n_1+n_2+n_3=7$). The resultant vector is defined as (Robertson et al., 2004):

$$\text{Resultant} = \sqrt{X^2+Y^2+Z^2}$$

where X, Y and Z are the medio-lateral, vertical and antero-posterior inputs of acceleration measured from the tri-axial accelerometer directions. For each of the eight activities of interest (counter movement jump, COD, jog, run and jump, sprint, stationary pause, tackle, and walk), the corresponding data was extracted and processed to generate features of interest.

Features were extracted from the data using three different movement capture durations of 0.5, 1.0, and 1.5 s respectively, each with a 50% overlap (Preece et al., 2009). These were chosen in such a way as to be long enough to capture the descriptive segment of each activity, but short enough to avoid overlap of the information. The feature set consisted of seven time domain features calculated for each of the seven inputs (minimum amplitude, maximum amplitude, mean amplitude, variance of amplitude, 25th percentile, 75th percentile and interquartile range; $m_1=49$). In addition, two frequency domain features were calculated with one spectral centroid for each of the seven inputs and a single bandwidth feature for the set of inputs as a whole (i.e., one bandwidth feature for the four accelerometer and three gyroscope inputs; ($m_2=8$)). Lastly, one custom energy feature was calculated for each sensor ($m_3=2$). The energy feature is defined as (Leutheuser et al., 2013):

$$E = \frac{\left(\frac{\left(\sum_{i=1}^{3} a_i\right)}{3}\right)}{p}$$

where a_i are the sum of the squared values for axes i ($i=X, Y, Z$) corresponding to either the accelerometer or gyroscope and p is the number of observations per axis. Thus, for each activity a total of 59 accelerometer and gyroscope features were calculated ($m_1+m_2+m_3=59$). The amplitude and percentile features provided important descriptors of the time domain of each input, while the percentiles also deliver a more representative measure of statistical dispersion (Rafiuddin et al., 2011). The spectral density and bandwidth features provided important descriptors relating to central mass and the frequency domain (obtained via fast Fourier transformation). The energy feature can be used to distinguish between sedentary and high intensity movements (Sugimoto et al., 1997).

2.3. Data analysis

In classification problems, it is common to compare the performance of different algorithms in order to determine which addresses the relevant problem most effectively. Classification algorithms are now typically preferred to traditional analysis approaches when assessing potentially non-linear data, due to their often improved overall classification performance (Witten and Frank, 2005; Robertson et al., 2015). To classify the eight activities of interest, three classification algorithms (LMT, RF and SVM) were employed. These were chosen for implementation based on their prevalence in the previous literature investigating similar problems. The LMT is a commonly used classification algorithm, which performs competitively with other machine learning classifiers and has the additional advantage of being easy to interpret (Landwehr et al., 2005). It combines the two complementary classification techniques of tree induction and linear regression (Hornik et al., 2009). Random Forest is another classification algorithm, which in its application grows multiple classification trees and builds upon them until each tree is at its largest (Breiman, 2001). The mean classification performance of the trees is then taken, which further assists in protecting against model overfitting (Breiman, 2001). Overfitting refers to the development of a model so specific to a particular training set that the findings are not generalisable when validated on a new test set of data (Robertson et al., 2015; Morgan et al., 2013). Additionally, RF is considered to have various useful features including high efficiency with large data sets and built-in ensemble classifiers (Breiman, 2001). Support Vector Machine differs slightly from the previous two algorithms, in that it attempts to find the best separating vector between two groups within a set of descriptors (Bennett and Bredensteiner, 2000). In this study, a radial kernel was used and both gamma ($\gamma \in [10^{-6}, 10^{-1}]$) and cost ($c \in [0.1, 10]$) values were tuned (Meyer et al., 2014). For classification of data with more than two groups (as seen here), the original problem is split into multiple binary problems which are then classified and compared. The problem receiving the most votes per instance is then assigned as the classifier. Readers interested in a more detailed explanation of these and other classifiers are directed towards the work of Zaki and Meira (2014).

As both classification accuracy and processing time was assessed, the analysis was conducted in two phases. First, the aim of phase one was to ascertain the data collection burden to achieve the desired classification accuracy. Specifically, the accuracy of each classification algorithm was investigated in four different ways for

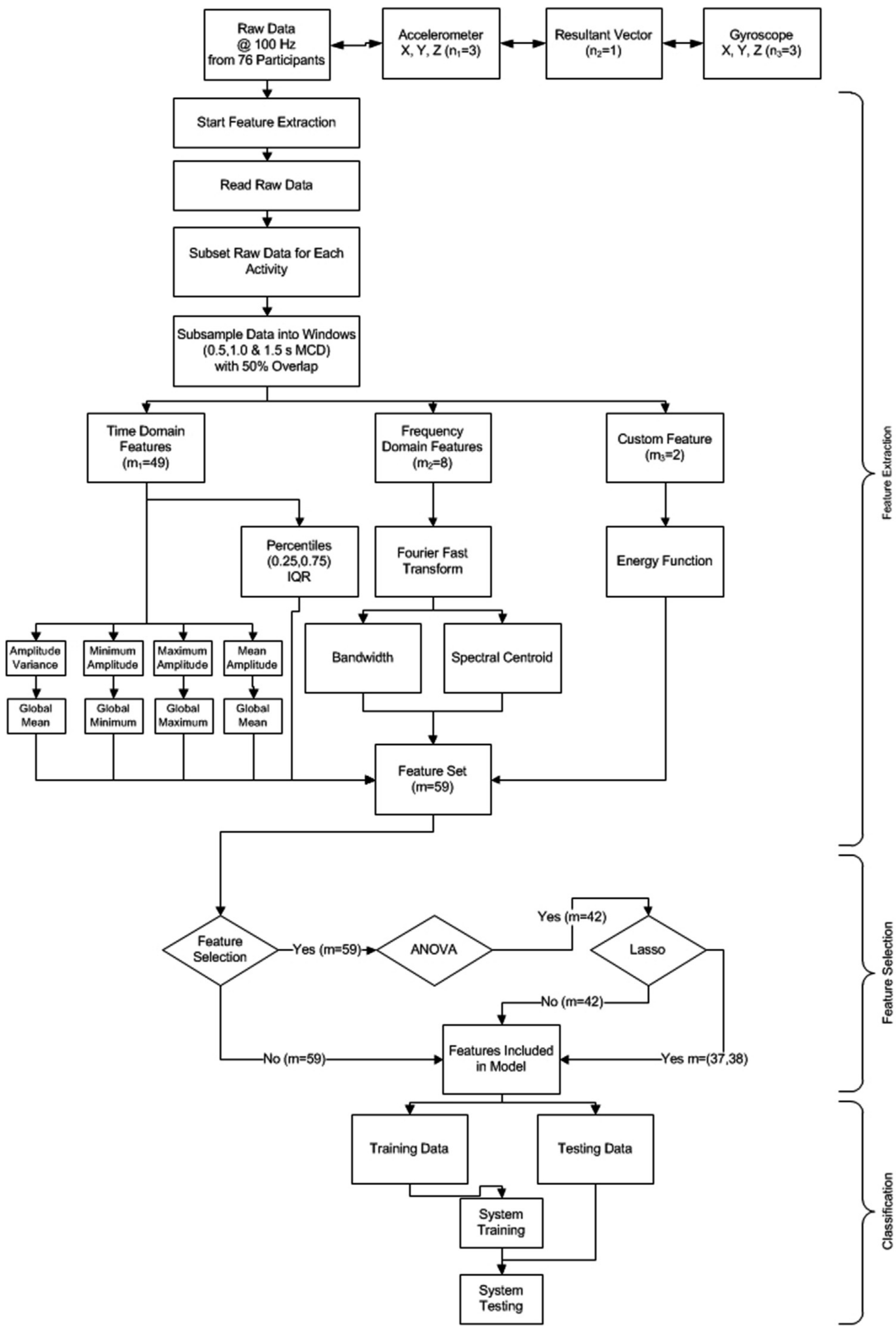

Fig. 1. Overview of the feature extraction and classification process. IQR, interquartile range; MCD, movement capture duration.

each of the three moment capture durations. These were, (i) all 59 features from accelerometer, resultant vector and gyroscope (ii) only the accelerometer and resultant vector features (m=33), (iii) only the accelerometer features (m=26), and (iv) only the gyroscope features (m=26).

Phase two aimed to investigate novel feature selection methodology in three different ways for the full set of inputs ($n_1+n_2+n_3=7$). This was undertaken as a large number of features do not always assist in obtaining better classification accuracy. Irrelevant features may introduce 'noise' leading to a loss in accuracy and over-fitting. Furthermore, the resulting model may take longer to implement and be more difficult to interpret (Liu and Motoda, 2012). In addition to processing time, comparisons of accuracies across each feature selection method, movement capture duration and classifier were also investigated. These were, (i) all 59 features were considered (no feature selection used), (ii) only features with significant results ($p<0.05$) for one-way analysis of variance (ANOVA) across classification groups were selected (Rabiei et al., 2012), with each feature examined for possible significance individually (Srivastava and Khatri, 1979), and (iii) all features that were selected though ANOVA (m=42) were passed for screening under Lasso regression simultaneously (Hastie et al., 2007) and further reduced (m=37 [0.5 s] and 38 [1.0 and 1.5 s]). Under this screening a feature was retained if it was contained in all of the three the feature sets produced by Lasso model based on the criterion of Mallow's C_p ($\min_{p=1,2,\ldots,m}(C_p-p)$) (Mallows, 1995), residual sum of squares ($\min_{p=1,2,\ldots,m}(\mathrm{RSS})$), and coefficient of determination (Walpole et al., 1993). In this instance, m was the number of features included in the model as determined by the feature selection methodology (no feature selection, ANOVA feature selection, ANOVA and lasso feature selection).

The computed set of feature data were split into a training and testing data set in accordance with a leave-one-out cross-validation methodology. The classification model was developed on the training set and its accuracy was ascertained on the testing data set. A single activity was randomly chosen (with equivalent probability) for each participant and assigned to the validation set (76 activities). From this validation set, a training set of 75 activities and a testing set of 1 activity not included in the training set are assigned. The leave-one-out cross-validation was repeated 76 times validating on each possible training and testing set combination. This process was repeated a further 10 times with the validation set being randomly re-sampled. The classification accuracy, defined as percent of correctly classified cases, was computed for each repeat, resulting in a hybrid 10-fold leave-one-out cross validation (Witten and Frank, 2005). The three algorithms were compared for mean classification accuracy across 10-fold cross validation. Fig. 1 below gives an overview of both the feature extraction and classification process.

All analyses were conducted on a 64-bit Windows operating system computer with Intel® Core™ i7-2670QM CPU and 8 GB RAM. All statistical analyses were conducted using R (version 3.0.1, R Core Team, Australia), which makes use of the following packages: e1071 (Meyer et al., 2014), lars (Hastie and Efron, 2013), RF (Liaw and Wiener, 2002) and RWeka (Hornik et al., 2009).

3. Results

The accuracy of classifiers per input variation and movement capture duration are presented in Table 1. Throughout all classification iterations LMT greatly outperformed both RF and SVM classifiers, obtaining classification rates 79% or above. As the number of input variables increased from three to seven, the accuracy of the classifiers generally remained the same or increased. The SVM and RF classifiers generally exhibited the strongest accuracy with a 1.5 s movement capture duration, while the LMT was generally strongest with a 1.0 s movement capture duration.

Table 1
Accuracy (Mean ± SD %) of classifiers and input variations for the movement capture durations after 10-fold leave-one-out cross-validation.

Input description	Classifier	Movement capture duration (s)		
		0.5	1.0	1.5
Accelerometer, resultant vector and gyroscope ($n_1+n_2+n_3=7$, m=59)	RF	0.39 ± 0.14	0.42 ± 0.12	0.43 ± 0.12
	LMT	0.92 ± 0.04	0.92 ± 0.03	0.92 ± 0.04
	SVM	0.37 ± 0.16	0.40 ± 0.15	0.35 ± 0.18
Accelerometer and resultant vector ($n_1+n_2=4$, m=33)	RF	0.32 ± 0.12	0.32 ± 0.12	0.33 ± 0.13
	LMT	0.89 ± 0.04	0.90 ± 0.04	0.89 ± 0.03
	SVM	0.33 ± 0.15	0.29 ± 0.14	0.36 ± 0.16
Accelerometer only (n_1=3, m=26)	RF	0.39 ± 0.16	0.37 ± 0.17	0.37 ± 0.16
	LMT	0.88 ± 0.05	0.89 ± 0.04	0.88 ± 0.04
	SVM	0.29 ± 0.13	0.28 ± 0.13	0.27 ± 0.14
Gyroscope only (n_3=3, m=26)	RF	0.41 ± 0.10	0.41 ± 0.11	0.42 ± 0.08
	LMT	0.80 ± 0.05	0.79 ± 0.07	0.80 ± 0.08
	SVM	0.32 ± 0.10	0.29 ± 0.10	0.27 ± 0.09

LMT, Logistic Model Tree; m, total number of observations; n_{1-3}, number of inputs; RF, Random Forest; SD, standard deviation; SVM, Support Vector Machine.

Table 2 presents the processing times for the accelerometer X, Y, Z, resultant vector and gyroscope inputs for all classifiers and movement duration combinations. The processing time (for both extraction and classification) was reduced (range 1.5–30.2%) using ANOVA or ANOVA and Lasso feature selection methods. The ANOVA and Lasso feature selection method was generally slower than the pure ANOVA method. The reduction in processing time had little effect on LMT (0–3%) and RF (0–5%) classifier accuracy, and a larger effect on SVM classifier accuracy (0–15%).

Fig. 2 presents the full models (n=7, m=59) activity-specific classification accuracies on the basis of 0.5, 1.0 and 1.5 s movement capture durations. For all three movement capture durations, classification accuracy exceeded 86%. Walking (98–99%) and stationary (95–98%) were best classified, whereas tackling (86–91%) and run and jump (89–90%) showed lower classification rates in the 1.0 (tackling) and 1.5 (tackling, run and jump) movement capture durations. Differences in movement capture duration classification accuracy ranged from 0% (COD/ jog [0.5 versus 1.5 s] and sprint/walk [0.5 versus 1.0 s]) to 8% (run and jump [1.0 versus 1.5 s]).

4. Discussion

The results of this study demonstrate that accurate activity classification using accelerometer and gyroscope inputs is achievable in a team-sport simulated circuit. Specifically, results showed that the highest performing algorithm for this purpose was LMT with an overall mean classification rate ranging from 79% to 92%. Further, the highest classification rate was achieved by combining all seven inputs from the accelerometer and gyroscope. Notably, the classification rate was substantially lower in the RF and SVM than those obtained using the LMT approach.

The findings of the current study are somewhat comparable to previous accelerometer input classification work (Mitchell et al., 2013). Mitchell et al. (2013) found that LMT (74%) outperformed SVM (55%) for activity classification in football (soccer), however, no differences were noted between classifiers when field hockey-specific activities were performed. Similar classification performance with the current study was found when multiple classifiers were combined. The stronger individual classifier results in the current study may be due to differences in experimental methodology. Specifically, Mitchell et al. (2013) did not assess gyroscope inputs, used lower frequency sampled accelerometer data (16–25 Hz), and also assessed activities such as dribbling (soccer) and hitting (field hockey) the ball (not assessed in this study). As the higher sample rate in the current study (100 Hz) may have contributed to the increased comparative classification performance, it may be that further increases in sample rate (> 100 Hz) could aid classification performance. However, this may have a negative effect on processing time.

When all seven accelerometer and gyroscope inputs were combined, highest rates of activity classification were achieved (e.g., mean classification accuracy of 92%). This was not surprising however, given that there was more information available for algorithm training. For example, previous research has shown that by combining both accelerometer and gyroscope inputs the classification rate during daily living and tennis-specific activities can be improved by as much as 14% (Bulling et al., 2014). Interestingly, classification performance was only improved by 2–3% when the three gyroscope inputs were included with accelerometer inputs

Table 2
Accuracy (Mean ± SD%) and processing time (s) of classifiers and model selection variations for the movement capture durations after 10-fold leave-one-out cross-validation. All comparisons are from accelerometer, resultant vector and gyroscope inputs ($n_1+n_2+n_3=7$, $m=59$) of differing feature selection methodology.

Feature selection	Classifier	Movement capture duration and [processing time] (s)		
		0.5 [236.3]	1.0 [231.0]	1.5 [228.9]
Full model ($m=59$)	RF	0.39 ± 0.14 [7.5]	0.42 ± 0.12 [7.6]	0.43 ± 0.12 [7.7]
	LMT	0.92 ± 0.04 [53.5]	0.92 ± 0.03 [53.7]	0.92 ± 0.04 [56.7]
	SVM	0.37 ± 0.16 [104.6]	0.40 ± 0.15 [102.8]	0.35 ± 0.18 [106.0]
ANOVA ($m=42$)	RF	0.36 ± 0.13 [6.6]	0.39 ± 0.14 [7.5]	0.39 ± 0.15 [6.9]
	LMT	0.91 ± 0.04 [42.5]	0.91 ± 0.04 [44.9]	0.92 ± 0.04 [44.4]
	SVM	0.23 ± 0.07 [77.1]	0.30 ± 0.16 [80.6]	0.27 ± 0.13 [81.5]
ANOVA and Lasso ($m=37/38$)	RF	0.36 ± 0.16 [7.0]	0.38 ± 0.16 [6.8]	0.38 ± 0.18 [7.1]
	LMT	0.91 ± 0.05 [44.2]	0.91 ± 0.03 [45.2]	0.89 ± 0.05 [46.7]
	SVM	0.24 ± 0.06 [77.6]	0.25 ± 0.11 [76.7]	0.26 ± 0.12 [83.2]

LMT, Logistic Model Tree; m, total number of observations; n_{1-3}, number of inputs; RF, Random Forest; SD, standard deviation; SVM, Support Vector Machines. Note that features were reduced to 37 (0.5 s movement capture duration) and 38 (1.0 and 1.5 s movement capture duration) using a combination of ANOVA and lasso regression. Note that the processing times reported in the column sub-heading refers to the amount of time it took to extract the relevant features, and the processing times reported in the table refer to the amount of time it took to classify all activities.

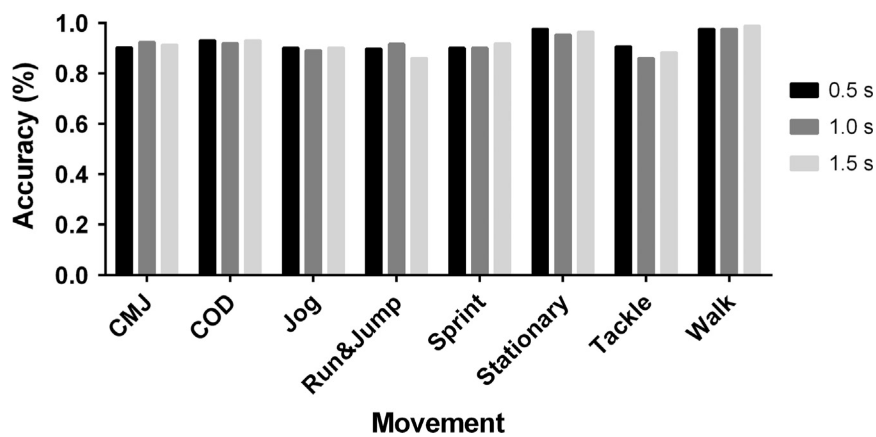

Fig. 2. Full model ($n=7$, $m=59$) activity classification accuracies for each movement capture duration. CMJ, countermovement jump; COD, change of direction.

in the current study, meaning this sensor contributed less to activity classification in these contexts (e.g. 8–10% decrease in classification accuracy compared to accelerometer inputs). This study made one of the first attempts to evaluate the effect of gyroscope inputs alone in classifying sporting activities. Generally lower rates of activity classification for the gyroscope inputs may be due to the upper back being predominantly exposed to linear motions, as compared to rotational motion that a gyroscope measures (Aminian and Najafi, 2004). Gyroscopes placed on the limbs (e.g. wrist and ankle) may be better able to aid classification, as limb motion is essentially a rotation around the corresponding joint (Kunze et al., 2010). Therefore, consideration of the location of the device and the activities performed may be important in deciding on the number of inputs included in future classification assessments. Furthermore, no study has assessed the validity and reliability of the gyroscope contained within the wearable tracking device, whereas a number of studies have been published in regards to the accelerometer (e.g. (Wundersitz et al., 2015a; Boyd et al., 2011)).

This investigation also analysed how movement capture durations of 0.5, 1.0 and 1.5 s affect the classification accuracy across different input and activity variations. There was no clear influence of movement capture duration on classification accuracy. Comparatively, Bulling et al. (2014) assessed durations ranging from 0.1 to 9.0 s during daily living and tennis-specific activities. The researchers found classification accuracy peaked at 1.0 s and dramatically decreased thereafter. Mitchell et al. (2013) also assessed durations from 1.0 to 9.0 s and found classification accuracy was maintained for all movement capture durations during field hockey, however soccer-based activity classification accuracy decreased past 3.0 s. Larger movement capture durations have also been used in lifestyle activity classification settings (e.g. (Mannini et al., 2013; Pärkkä et al., 2006; Trost et al., 2012)), however, the frequency of sporting activities may result in two or more activities occurring in longer movement capture durations (e.g., greater than 1.0 s), dramatically increasing classification difficulty (Mitchell et al., 2013; Mannini et al., 2013). Team sport activities, therefore, may benefit from shorter movement capture durations than are typically employed in the lifestyle activity classification literature. It should also be acknowledged that inter- and intra-participant variations and set movement capture durations as used in the current study may have contributed to the lower than expected classification rates. Future work may consider alternative approaches, such as a sliding window approach, where a capture duration with a time length of T is slid across the data and if an activity is detected within the window it is flagged.

A further practical consideration relates to the effect different model selection variations had on processing time and classification accuracy. The number of features included in the analysis (range 37 to 59) tended to improve classification accuracy, but this was generally at the cost of increased processing time. The classifier chosen also appeared to affect this relationship. For example, the processing burden was shortest for RF (6.6 to 7.7 s) and longest for SVM (76.7 to 106.0 s), with the LMT falling in the middle (42.5 to 56.7 s). However, movement capture duration had minimal impact on processing time in the current study.

Removing low-level contributing features from the training process can have a positive effect on classification performance, with ANOVA feature selection reducing processing time by approximately 15%. For reduced processing time, using ANOVA or ANOVA and Lasso feature selection, the classification accuracy only decreased by 5% and 4% respectively. Such a model would require approximately 230 s of time for feature extraction to occur and a further 45–54 s for classification of the activities performed using LMT. Comparable results are reported in literature with much larger volumes of data accumulation and smaller number of classification groups. Nathan et al. (2012) using accelerometer and GPS data, gathered over 750,000 measurements and achieved accuracy of over 84% for RF and SVM classifiers. Leutheuser et al. (2013) also using a large dataset and pre-clustering of the activities, achieved an accuracy of 87% for SMV method. Mitchell et al. (2013) reported a similar trend in classification accuracy with a LMT method, on a much smaller dataset. To this end, the similar between classification rates to previous studies using a reduced number of measurement features is encouraging. This is especially important where real-time classification (e.g., during training and gameplay) may be a future aim. Based on these findings, presently the LMT algorithm combined with accelerometer inputs alone, provided the best trade-off between classification accuracy and processing time for use in this context.

In conclusion, a variety of classification algorithms, movement capture durations and feature selection methods were compared to determine the most parsimonious approach to classify multiple simulated team sport-related activities. The LMT was shown to be highly accurate using data obtained from a single accelerometer and gyroscope sensor contained within wearable tracking device technology. Consequently, in sporting scenarios where wearable tracking devices are employed, it is both possible and feasible to use accelerometer and gyroscope data to accurately classify sporting activities. Further, the processing time can be reduced through feature selection models, however a trade-off exists between classification accuracy and processing time. With this in mind, accelerometer inputs alone appear to be the most parsimonious approach from this location. Further development and validation of algorithms in sports is required. Once developed, the ability of these algorithms to classify team sport activities during game-play should be performed. Further exploration of accelerometer and gyroscope features, and feature reduction is needed to provide real-time classification in the future.

Conflict of interest statement

None of the authors have financial or other conflicts of interest in regards to this research.

Acknowledgements

None.

References

Aminian, K., Najafi, B., 2004. Capturing human motion using body-fixed sensors: outdoor measurement and clinical applications. Comput. Animat. Virtual Worlds 15, 79–94.
Aughey, R.J., Falloon, C., 2010. Real-time versus post-game GPS data in team sports. J. Sci. Med. Sport 13, 348–349.
Bayat, A., Pomplun, M., Tran, D.A., 2014. A study on human activity recognition using accelerometer data from smartphones. Procedia Comput. Sci. 34, 450–457.
Bennett, K.P., Bredensteiner, E.J., 2000. Duality and geometry in SVM classifiers. In: Proceedings of the 17th International Conference on Machine learning. Stanford University, California, USA.
Boyd, L., Ball, K., Aughey, R.J., 2011. The reliability of MinimaxX accelerometers for measuring physical activity in Australian football. Int. J. Sports Physiol. Perform. 6, 311–321.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Bulling, A., Blanke, U., Schiele, B., 2014. A tutorial on human activity recognition using body-worn inertial sensors. ACM Comput. Surv. 46, 1–33.
Carling, C., Reilly, T., Williams, A.M., 2009. Performance Assessment for Field Sports. Routledge, New York.
Cummins, C., Orr, R., O'Connor, H., West, C., 2013. Global positioning systems (GPS) and microtechnology sensors in team sports: a systematic review. Sports Med. 43, 1025–1042.
Ermes, M., Parkka, J., Mantyjarvi, J., Korhonen, I., 2008. Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. IEEE Trans. Inf. Technol. Biomed. 12, 20–26.
Gastin, P., McLean, O., Breed, R., Spittle, M., 2014. Tackle and impact detection in elite Australian football using wearable microsensor technology. J. Sports Sci. 32, 947–953.
Gastin, P., McLean, O., Spittle, M., Breed, R., 2013. Quantification of tackling demands in professional Australian football using integrated wearable athlete tracking technology. J. Sci. Med. Sport. 16, 589–593.
Hastie, T., Efron, B., 2013. Least angle regression, lasso and forward stagewise. R package version 1.2.
Hastie, T., Taylor, J., Tibshirani, R., Walther, G., 2007. Forward stagewise regression and the monotone lasso. Electron. J. Stat. 1, 1–29.
Hornik, K., Buchta, C., Zeileis, A., 2009. Open-source machine learning: R meets Weka. Comput. Stat. 24, 225–232.
Kelly, D., Coughlan, G., Green, B., Caulfield, B., 2012. Automatic detection of collisions in elite level rugby union using a wearable sensing device. Sports Eng. 15, 81–92.
Kunze, K., Bahle, G., Lukowicz, P., Partridge, K., 2010. Can magnetic field sensors replace gyroscopes in wearable sensing applications? In: Proceedings of the 14th IEEE International Symposium on Wearable Computing. Seol, Korea.
Landwehr, N., Hall, M., Frank, E., 2005. Logistic Model Trees. Mach. Learn. 59, 161–205.
Leutheuser, H., Schuldhaus, D., Eskofier, B.M., 2013. Hierarchical, multi-sensor based classification of daily life activities: comparison with state-of-the-art algorithms using a benchmark dataset. PLoS ONE 8, e75196.
Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.
Liu, H., Motoda, H., 2012. Feature selection for knowledge discovery and data mining, Springer Science & Business Media. Kluwer Academic Publishers, Norwell, MA.
Mallows, C.L., 1995. More comments on Cp. Technometrics 37, 362–372.
Mannini, A., Intille, S.S., Rosenberger, M., Sabatini, A.M., Haskell, W., 2013. Activity recognition using a single accelerometer placed at the wrist or ankle. Med. Sci. Sports Exerc. 45, 2193–2203.
McNamara, D., Gabbett, T., Chapman, P., Naughton, G., Farhart, P., 2015. The validity of microsensors to automatically detect bowling events and counts in cricket fast bowlers. Int. J. Sports Physiol. Perform. 10, 71–75.
Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2014. e1071: Misc Functions of the Department of Statistics (e1071). Comprehensive R Archive Network, TU Wien.
Mitchell, E., Monaghan, D., Connor, N., 2013. Classification of sporting activities using smartphone accelerometers. Sensors 13, 5317–5337.
Morgan, S., Williams, M.D., Barnes, C., 2013. Applying decision tree induction for identification of important attributes in one-versus-one player interactions: a hockey exemplar. J. Sports Sci. 31, 1031–1037.
Najafi, B., Aminian, K., Paraschiv-Ionescu, A., Loew, F., Bula, C.J., Robert, P., 2003. Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. IEEE Trans. Biomed. Eng. 50, 711–723.
Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., Getz, W.M., 2012. Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. J. Exp. Biol. 215, 986–996.
Neville, J., Wixted, A., Rowlands, D., James, D., 2010. Accelerometers: an underutilized resource in sports monitoring. In: Proceedings of the 6th International Conference on Intelligent Sensors, Sensor Networks and Information Processing. Queensland Univeristy of Technology, Brisbane, Australia.
Pärkkä, J., Ermes, M., Korpipää, P., Mäntyjärvi, J., Peltola, J., Korhonen, I., 2006. Activity classification using realistic data from wearable sensors. IEEE Trans. Inf. Technol. Biomed. 10, 119–128.
Peacock, L., Hewitt, A., Rowe, D.A., Sutherland, R., 2014. Stride rate and walking intensity in healthy older adults. J. Aging Phys. Act. 22, 276–283.
Preece, S.J., Goulermas, J., Kenney, L.P.J., Howard, D., 2009. A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. IEEE Trans. Biomed. Eng. 56, 871–879.
Rabiei, M., Gupta, R., Cheong, Y.P., Sanchez, G.A., 2012. A novel approach in extracting predictive information from water-oil ratio for enhanced water production mechanism diagnosis. APPEA J. 50, 567–579.
Rafiuddin, N., Uzzaman Khan, Y., Farooq, O., 2011. Feature extraction and classification of EEG for automatic seizure detection. In: Proceedings of the International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT). Aligarh, Uttar Pradesh, India.
Robertson, D., Caldwell, G., Hamill, J., Kamen, G., Whittlesey, S., 2004. Research Methods in Biomechanics. Human Kinetics, South Australia.
Robertson, S., Back, N., Bartlett, J.D., 2015. Explaining match outcome in elite Australian Rules football using team performance indicators. J. Sports Sci., 1–8.

Singh, T.K., Guelfi, K.J., Landers, G., Dawson, B., Bishop, D., 2010. Reliability of a contact and non-contact simulated team game circuit. J. Sports Sci. Med. 9, 638–642.
Srivastava, S., Khatri, C.G., 1979. An introduction to Multivariate Statistics. North-Holland, New York.
Sugimoto, A., Hara, Y., Findley, T.W., Yoncmoto, K., 1997. A useful method for measuring daily physical activity by a three-direction monitor. Scandanavian J. Rehabil. Med. 29, 37–42.
Trost, S.G., Wong, W.K., Pfeiffer, K.A., Zheng, Y., 2012. Artificial neural networks to predict activity type and energy expenditure in youth. Med. Sci. Sports Exerc. 44, 1801–1809.
Walpole, R.E., Myers, R.H., Myers, S.L., Ye, K., 1993. Probability and Statistics for Engineers and Scientists. Macmillan, New York.
Witten, I.H., Frank, E., 2005. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, Sydney.
Wundersitz, D., Gastin, P., Richter, C., Robertson, S., Netto, K., 2015a. Validity of a trunk-mounted accelerometer to assess peak accelerations during walking, jogging and running. Eur. J. Sport Sci. 15, 382–390.
Wundersitz, D., Gastin, P., Robertson, S., Davey, P., Netto, K., 2015b. Validation of a Trunk-mounted Accelerometer to Measure Peak Impacts during Team Sport Movements. Int. J. Sports Med. 36, 742–746.
Zaki, M.J., Meira Jr, W., 2014. Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press, New York.