ABSTRACT

Study Design: Cross-sectional study

Objectives: The main objective of this study was to develop and test classification algorithms based on machine learning, using accelerometers to identify the activity type performed by manual wheelchair users with SCI.

Setting: The study was conducted in the Physical Therapy department and the Physical Education and Sports department of the University of Valencia.

Methods: Twenty volunteers were asked to perform ten physical activities: lying down, body transfers, moving items, mopping, working on a computer, watching TV, arm-ergometer exercises, passive propulsion, slow propulsion and fast propulsion while fitted with four accelerometers placed on both wrists, chest and waist. The activities were grouped into five categories: sedentary, locomotion, housework, body transfers and moderate physical activity. Different machine learning algorithms were used to develop individual and group activity classifiers from the acceleration data for different combinations of number and position of the accelerometers.

Results: We found that although the accuracy of the classifiers for individual activities was moderate (55-72%), with higher values for a greater number of accelerometers, grouped-activities were correctly classified in a high percentage of cases (83.2 - 93.6%).
Conclusions: with only two accelerometers and the quadratic discriminant analysis algorithm we achieved a reasonably accurate group activity recognition system (> 90%). Such a system with the minimum of intervention would be a valuable tool for studying PA in persons with SCI.

Keywords: physical activity, machine learning, accelerometer, spinal cord injury
INTRODUCTION

Physical activity (PA) plays an important role in the health of persons with spinal cord injury (SCI). PA is a protective factor that reduces the risk of illnesses such as cardiovascular disease and Type 2 diabetes \(^1\)–\(^3\) and other common comorbidities in this population (e.g., pressure ulcers) \(^4\),\(^5\).

An appropriate method of quantifying PA levels in persons with SCI during their daily activities is essential for several reasons \(^6\). Firstly, these methods may be used in epidemiological studies to establish more precisely the effects of PA on their health. Secondly, it can be used to monitor the effectiveness of PA promotion programs in this population. Finally, with the appropriate hardware and software, those suffering from SCI may carry out continuous control of their energy expenditure and thereby adjust their physical and nutritional habits to achieve a healthy lifestyle.

Accelerometers are currently the devices most commonly used to measure PA although other methods, like heart rate \(^7\),\(^8\) and questionnaires \(^9\),\(^10\), have been validated for people with spinal cord injury. Early studies quantified PA by estimating energy expenditure. However recent works estimate not only energy expenditure, but also the type of activity being carried out, according to the acceleration pattern produced\(^11\)–\(^16\), which is important in studies on the SCI population. The performance of certain activities could either prevent or aggravate certain health problems (e.g., shoulder pain\(^17\),\(^18\)).
Although studies have been published that establish the necessary mathematical models for estimating types of physical activities\textsuperscript{11–16}, few of them have tackled this problem in subjects with SCI. Specifically, Postma et al.\textsuperscript{19} using a total of six accelerometers, were able to identify wheelchair propulsion from other activities (e.g., lying down, body transfer, doing dishes…). Their classifier achieved an accuracy of 92%. Later Hiremath et al.\textsuperscript{20} classified the type of activity performed by SCI subjects using accelerometry, galvanic skin response, skin temperature and near-body temperatures. They were able to distinguish between resting, propulsion, arm-ergometer and deskwork, with an accuracy of 96.2 \% using Quadratic Discriminant Analysis (QDA). Although 4 types of activities were included in this latter study, a broader study needed to be carried out in order to identify a wider range of activities. Therefore, the aims of the present work were:

1. To develop and test classification algorithms to identify a) 10 individual activities, b) 5 grouped-activities, performed by manual wheelchair users with SCI equipped with accelerometers.

2. To establish the minimum number of accelerometers needed for a given accuracy for each application.
MATERIAL AND METHODS

Participants

Twenty subjects took part in the study [40.03 (10.57) years, 75.8 (17.54) kg and 1.76 (0.09) m]. The researchers recruited participants from two different institutions: i. Hospital la Fe of Valencia and ii. Asociación Provincial de Lesionados Medulares y Grandes Discapacitados (ASPAYM). The subjects had suffered spinal damage between the T2 and L5 vertebrae, and had been diagnosed at least one year before the start of this study. The level and completeness of the SCI (Table 1) were determined by a complete neurological examination conducted by a medical specialist, using the American Spinal Injury Association Impairment Scale (AIS). Their independence status expressed as mean (SD) was 65.3 (7.61). This independence measurement was determined using Spinal Cord Independence Measure version III (SCIM III)\(^1\).

Table 1 here

The exclusion criteria were: i) history of depressive or cognitive disorders; ii) posttraumatic cervical myelopathy, motor or sensory impairment of the upper extremities, ischemic heart disorder, or recent osteoporotic fractures; iii) Presence of tracheotomy or iv) sacrotuberous ulcers or hypertension.

All the subjects gave written consent to participate in the study, which was previously approved by the university’s ethical committee. We certify that
all applicable institutional and governmental regulations concerning the
ethical use of human volunteers were followed during the course of this
research.

*Data collection*

The subjects were asked to perform ten physical activities (using their own
wheelchair): lying down, body transfers, moving items, mopping, working
on a computer, watching TV, arm-ergometer exercise, passive propulsion,
slow propulsion and fast propulsion. A detailed description of each activity
can be found in a previous study. Each activity was carried out for 10
minutes with 1-2 minutes’ rest between activities, with only one exception
in the case of body transfers, in which the activity took place for one minute
followed by one minute’s rest for a total of ten minutes to avoid overloading
the shoulder musculoskeletal system. All these measurements have been
supervised by the same researcher to ensure the successful completion of
these activities.

During these activities body forces were monitored by four accelerometers
(Actigraph model GT3X, Actigraph, Pensacola, FL, USA) being the
sampling frequency 30 Hz. A bandpass digital filter between 0.25 and 2.5
Hz was implemented in order to reduce the influence of the static
acceleration and the higher frequency components (manufacturer hardware
characteristic). Then, the accelerations (expressed in counts) were rectified
and integrated in 1-second epochs. The accelerometers were placed one on
each wrist, one on the non-dominant waist and on the non-dominant side of the chest (Figure 1). Elastic belts were used in order to minimize movements of the accelerometers; and the spatial orientation were similar in all the subjects.

Figure 1 here

Signal processing

The Matlab R2012a (Mathworks Inc, Natick, USA) was used for signals processing. We worked out fourteen variables for each axis (i.e. X, Y, Z and resultant vector) at minutes: four, five, six and seven for each activity.

The standard deviation, variance and the $10^{th}$, $25^{th}$, $50^{th}$, $75^{th}$ and $90^{th}$ percentiles, interquartile range and the range between the $10^{th}$ and $90^{th}$ percentiles were calculated in the time domain. The lag-one correlation of each minute was also worked out as a measure of temporal dynamics. The acceleration signal was analyzed using the two-level wavelet transform, the mother wavelet being Daubechies 2. We calculated the Euclidean norm of the detail coefficients of the first and second levels of resolution and the approximation coefficients of the second level (i.e. ND$_1$, ND$_2$, NA$_2$). The sample entropy was computed for each axis (tolerance=0.3 SD; pattern length=2). Finally, we computed the cross-correlation between the three orthogonal axes (i.e., X-Y, Y-Z and X-Z cross-correlations). The total
number of variables was 59 for each accelerometer (i.e., 14 variables for the four axes and three variables for the correlation between axes).

Data Analysis

Classifiers were designed for individual-activities and grouped-activities; those for individual-activities had ten possible categories (i.e. each activity performed) and grouped-activities had five (Table 2), established according to the activity’s objective or function.

Table 2 here

In order to determine the required number of accelerometers to properly identify the activities or groups of activities, the data from several accelerometers were combined. The configurations tested were: i) dominant wrist accelerometer, ii) non-dominant wrist accelerometer, iii) both wrist accelerometers and iv) all four accelerometers

The first step was to split the database (800 data = 20 subjects*10 PAs*4 min/PA) into two data sets (figure 2). One was used to train and validate the classifiers (n=640) and the other to test them (n=160). We checked that there were no statistically significant differences in the computed variables between data sets by means of the Wilcoxon rank sum test (p>0.05) and that the percentage of cases of each activity was the same in both data sets.

A principal component analysis was then applied to reduce the dimensions of the data matrix parameters. This analysis was applied to the training set
of the above-cited four combinations of accelerometers. These databases
were reduced from 59, 59, 118 and 236 variables respectively to 22, 22, 41
and 78 principal components (99% of the variance was maintained). The
coefficients of this analysis of the training set were applied to the test set, so
as to obtain the principal components of these data. The principal
components of the two data sets were used as inputs in the subsequent
analysis.

Figure 2 here

We used three different machine-learning algorithms to design the
classifiers: linear discriminant analysis (LDA), quadratic discriminant
analysis (QDA) and support vector machines (SVM). The classifiers were
designed and validated using a 10-fold stratified cross-validation, which was
performed twenty times to reduce the randomization effect. The optimal
combination of variables was determined using a forward sequential feature
selection algorithm that included only those variables that significantly
improved classifier accuracy. The feature selection algorithm stopped when
the addition of any new variable did not improve classifier accuracy by
0.5%. Once the classifiers were designed with the training set, we applied
them to the test set and computed the classification accuracy:

\[
\text{Accuracy} = \frac{\text{TruePositives} + \text{TrueNegatives}}{\text{TruePositives} + \text{TrueNegatives} + \text{FalsePositives} + \text{FalseNegatives}}
\]
RESULTS

Table 3 shows the accuracy of the different classifiers implemented in the test set, using the information from the different accelerometer configurations to distinguish each of the 10 individual activity types. As expected, it can be seen that in general the accuracy of the classifier improves as the number of accelerometers increases. However, the accuracy obtained is always less than 75%, regardless of the number/position of the accelerometer and the classification algorithm used.

*Table 3 here*

Figure 3 shows the accuracy of the individual activity classifiers in each of the 10 categories. It can be observed that in many activities accuracy values near or above 90% are achieved, particularly when two or four accelerometers are used. However, some activities (e.g. PC work or passive propulsion), which could be confused with each other, have particularly low accuracy values, giving a slightly low overall accuracy value for the classifier.

*Figure 3 here*

On the other hand, the grouped-activity classifiers showed good accuracy in all cases (between 83.2% and 93.6%) (Table 4). Again, it can be seen that in general, the higher the number of accelerometers, the higher the classification accuracy. In contrast, the classification algorithm does not
seem to significantly influence the prediction capability. It is noteworthy that there are three classifiers with accuracy values above 90%: i) two wrists with QDA, ii) all with QDA and iii) all with SVM.

Table 4 here

The accuracy of the classifiers for each category is shown in Figure 4. It can be observed that those with the lowest values are body transfers and locomotion. It is also noteworthy that the accuracies of the body transfer and housework categories seem to be the most dependent on the number of accelerometers used, whereas the accuracy of the other three categories is fairly stable, regardless of the number of accelerometers and algorithms used.

Figure 4 here

Finally, Table 5 shows the confusion matrix of the QDA classifier for grouped-activities, which uses information from the accelerometers on both wrists. As shown, the rate of properly classified sedentary activities is very high (93.75-100%) and only 6.25% of the cases of working with computers or passive propulsion are misclassified. The classification error in the locomotion category is mainly due to the fact that the slow propulsion activity is misclassified as housework in 39.34% of cases. In the housework category, high accuracy values are observed for both activities. 90.56% of the moving items cases and 85.94% of the mopping cases were properly
classified. 14.41% of the transferring activity cases were misclassified as housework. Finally, it is noteworthy that 100% accuracy is reached in moderate physical activity.

Table 5 here

**DISCUSSION**

In the present work we designed and implemented several classifiers using only recordings from accelerometers in SCI patients to distinguish a) 10 individual activities and b) 5 categories of grouped-activities according to the activity's aim or function. None of the classifiers obtained an overall accuracy over 73% in identifying the 10 activities, regardless of the number of accelerometers and the algorithm used. The relatively low values are most likely due to the fact that some of the activities shared similar patterns, e.g. watching television, working with a PC or passive propulsion. Additional information would be needed to overcome this limitation.

When the activities were grouped by their aim or function, promising results were obtained. In general it has been observed that the more accelerometers used, the higher the classifier accuracy. Three classifiers were obtained with an average accuracy above 90%: i) two wrists with QDA, ii) all with QDA and iii) all with SVM. In configurations ii) and iii), the use of four accelerometers did not provide a significant increase in the accuracy of the classifier using the QDA algorithm. Compared with configuration iii),
classifier i) has the advantage that the QDA algorithm is computationally much more efficient and could be easily implemented in a real-time system. Moreover, using only two accelerometers greatly simplifies the recording protocol and also improves patient comfort during recording. This suggests that the optimal setting of the classifier to distinguish the 5 categories of SCI activities tested was obtained with the QDA algorithm and the accelerometers on both wrists.

Sedentary activities and moderately intensive physical activities obtained good rates of correct classification (always above 93.75%). These results are comparable with those of other authors, who obtained 92% accuracy in distinguishing different activities in SCI patients\(^1^9\). However, in this latter study six accelerometers were used and only two categories were classified: two types of wheelchair propulsion versus other activities: lying down, body transfer, doing dishes\(^1^9\). The accuracy values obtained in the present work are similar to those obtained by other authors\(^2^0\): obtained 96.2% in identifying 4 types of activities (rest, deskwork, arm-ergometer and propulsion). Unlike other authors, who used input variables of acceleration, galvanic skin response, skin temperature and near body\(^2^0\), in the present work only acceleration data (from the two wrists) was used.

On the other hand, the accuracy values obtained for the activity recognition systems in SCI patients compare favorably with those published regarding the able-bodied population. Trost et al.\(^1^6\) obtained 88.4% accuracy in
classifying activities clustered into the following categories: sedentary, light household activities and games, household activities and moderate-to-high-intensity sports, walking and running. Also in this context Khan et al.\textsuperscript{12} reached 97.9\% of properly classified recording time in the following activities: lying, standing, walking and running. Liu et al.\textsuperscript{13} combined several sensors (two accelerometers and a flow meter) and achieved 84.7\% correct classification in 13 different activities. Therefore, the activity recognition systems proposed in the present study show similar accuracy to those in other populations when considering groups of similar activities.

It is remarkable that the grouped-activities classifier, employing the recordings from 2 accelerometers with the QDA algorithm, often identified some locomotion activities, such as housework. In spite of the fact that rapid propulsion was correctly distinguished from other household chores, probably due to the greater magnitude of the accelerations, slow propulsion was misclassified as housework in 39.34\% of cases. This may be because while performing household tasks (mopping or moving objects) the subjects had to propel the wheelchair at a slow speed (similar to slow propulsion). The inclusion of additional parameters that take into account the temporal structure of the data or the variation of the spectral parameters over time could help to improve accuracy in these cases.

Finally, this study has some limitations. Firstly, it would be advisable to expand the database in terms of the numbers of both subjects and activities.
Secondly, although some extent of variability has been included in the data used to design the classifiers since participants used their own wheelchair, which could have different dynamic responses for each of the movements, the physical activities were carried out in a controlled environment, following the instructions of a supervisor, with a break between activities so as to minimize fatigue. Future studies should confirm the good results obtained in this work in conditions closer to everyday life. In such conditions events such as transitions between activities, the type or inclination of the surfaces, etc. could worsen classification accuracy. In summary, we believe that this work provides the basis for a minimally intrusive expert system that would monitor daily physical activity in SCI subjects, for whom monitoring is of great significance.

In short, the highest accuracy values (83.2 - 93.6%) were those obtained on activities grouped according to objective or function. Classifiers of individual activities showed lower classification accuracy (55 – 72.5%). The best performance was obtained from four accelerometers and QDA or SVM algorithms. However, an activity recognition system with good accuracy (> 90%) was also achieved with only two accelerometers and the QDA algorithm. Due to the fact that 2 accelerometers are less stressful for the subject, it would be useful to implement this system in future studies to identify activities in subjects with spinal cord injuries.
X García-Massó gratefully acknowledge the support of the University of Valencia under project UV-INV-PRECOMP13-115364
CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.
TITLES AND LEGENDS TO FIGURES

Figure 1. Location of the accelerometers
Figure 2. Schematic overview of the process to obtain the individual activity classifiers. The process is the same for individual and grouped-activity classifiers.
Figure 3. Accuracy of the classifiers for individual activities with the algorithms: Top- linear discriminant analysis, Middle- quadratic discriminant analysis and bottom-support vector machines.
Figure 4. Accuracy of the classifiers for grouped activities with the algorithms: Top- linear discriminant analysis, Middle- quadratic discriminant analysis and bottom-support vector machines.
REFERENCES


<table>
<thead>
<tr>
<th>Subject</th>
<th>Neurological level</th>
<th>AIS Score</th>
<th>Time of injury</th>
<th>Aetiology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T4</td>
<td>B</td>
<td>229</td>
<td>Trauma</td>
</tr>
<tr>
<td>2</td>
<td>T11-12</td>
<td>A</td>
<td>264</td>
<td>Trauma</td>
</tr>
<tr>
<td>3</td>
<td>T4</td>
<td>A</td>
<td>88</td>
<td>Trauma</td>
</tr>
<tr>
<td>4</td>
<td>T7</td>
<td>A</td>
<td>81</td>
<td>Trauma</td>
</tr>
<tr>
<td>5</td>
<td>T5</td>
<td>A</td>
<td>24</td>
<td>Trauma</td>
</tr>
<tr>
<td>6</td>
<td>T4</td>
<td>A</td>
<td>236</td>
<td>Tumour</td>
</tr>
<tr>
<td>7</td>
<td>T4</td>
<td>A</td>
<td>34</td>
<td>Trauma</td>
</tr>
<tr>
<td>8</td>
<td>L5-S1</td>
<td>B</td>
<td>59</td>
<td>Surgery</td>
</tr>
<tr>
<td>9</td>
<td>T10-11</td>
<td>A</td>
<td>233</td>
<td>Trauma</td>
</tr>
<tr>
<td>10</td>
<td>T5</td>
<td>A</td>
<td>359</td>
<td>Trauma</td>
</tr>
<tr>
<td>11</td>
<td>T4-5</td>
<td>A</td>
<td>153</td>
<td>Trauma</td>
</tr>
<tr>
<td>12</td>
<td>T12</td>
<td>A</td>
<td>401</td>
<td>Congenital sclerosis</td>
</tr>
<tr>
<td>13</td>
<td>T4</td>
<td>A</td>
<td>90</td>
<td>Trauma</td>
</tr>
<tr>
<td>14</td>
<td>T5</td>
<td>A</td>
<td>290</td>
<td>Trauma</td>
</tr>
<tr>
<td>15</td>
<td>T5</td>
<td>A</td>
<td>122</td>
<td>Trauma</td>
</tr>
<tr>
<td>16</td>
<td>T5-6</td>
<td>A</td>
<td>79</td>
<td>Tumour</td>
</tr>
<tr>
<td>17</td>
<td>T7</td>
<td>A</td>
<td>67</td>
<td>Trauma</td>
</tr>
<tr>
<td>18</td>
<td>T12</td>
<td>A</td>
<td>19</td>
<td>Multiple sclerosis</td>
</tr>
<tr>
<td>19</td>
<td>T12-L1</td>
<td>B</td>
<td>435</td>
<td>Trauma</td>
</tr>
<tr>
<td>20</td>
<td>T5</td>
<td>A</td>
<td>193</td>
<td>Trauma</td>
</tr>
</tbody>
</table>

Time of injury is expressed in months. AIS = American Spinal Injury Association Impairment Scale.
### Table 2. Accuracy of the individual-activities classifiers

<table>
<thead>
<tr>
<th></th>
<th>Dominant</th>
<th>Non-Dominant</th>
<th>Two wrists</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>LDA</em></td>
<td>61.4</td>
<td>63.3</td>
<td>62.9</td>
<td>69.3</td>
</tr>
<tr>
<td><em>QDA</em></td>
<td>55</td>
<td>63</td>
<td>67.8</td>
<td>72.5</td>
</tr>
<tr>
<td><em>SVM</em></td>
<td>59.1</td>
<td>61.5</td>
<td>68.9</td>
<td>65.9</td>
</tr>
</tbody>
</table>

Data are expressed as a percentage of total cases that belong to that category. *LDA* = Linear Discriminant Analysis; *QDA* = Quadratic Discriminant Analysis; *SVM* = Support Vector Machines.

### Table 3. Accuracy of the grouped-activities classifiers

<table>
<thead>
<tr>
<th></th>
<th>Dominant</th>
<th>Non-Dominant</th>
<th>Two wrists</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>LDA</em></td>
<td>85.9</td>
<td>83.9</td>
<td>87.1</td>
<td>89.4</td>
</tr>
<tr>
<td><em>QDA</em></td>
<td>84.5</td>
<td>86.7</td>
<td>90.4</td>
<td>90.7</td>
</tr>
<tr>
<td><em>SVM</em></td>
<td>83.2</td>
<td>87</td>
<td>86.8</td>
<td>93.6</td>
</tr>
</tbody>
</table>

Data are expressed as a percentage of total cases that belong to that category. *LDA* = Linear Discriminant Analysis; *QDA* = Quadratic Discriminant Analysis; *SVM* = Support Vector Machines.
Table 4. Confusion matrix of the QDA classifier, implemented using information from two accelerometers placed in both wrists, for grouped-activities.

<table>
<thead>
<tr>
<th>Real type of activity</th>
<th>QDA grouped-activities classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sedentary</td>
</tr>
<tr>
<td>Lying down</td>
<td>100</td>
</tr>
<tr>
<td>PC work</td>
<td>93.75</td>
</tr>
<tr>
<td>Watching TV</td>
<td>100</td>
</tr>
<tr>
<td>Passive propulsion</td>
<td>93.75</td>
</tr>
<tr>
<td>Slow propulsion</td>
<td>0</td>
</tr>
<tr>
<td>Fast propulsion</td>
<td>0</td>
</tr>
<tr>
<td>Moving items</td>
<td>0</td>
</tr>
<tr>
<td>Mooping</td>
<td>0</td>
</tr>
<tr>
<td>Transferring</td>
<td>0</td>
</tr>
<tr>
<td>Arm-ergometer</td>
<td>0</td>
</tr>
</tbody>
</table>

Data are expressed as a percentage of total cases that belong to that category. MPA = Moderate Physical Activity, QDA = Quadratic Discriminant Analysis