We also focus on a set of exercise activities that are part of the Otago exercise program(16 movement).

The purpose of this paper is to assess the classiffcation performance of different groups of IMU sensors for different activities.

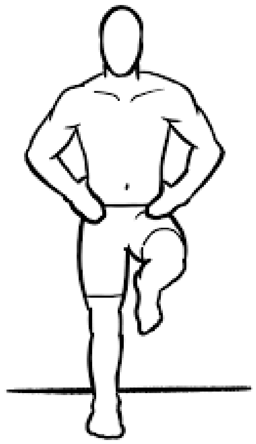
| | |
|---|---|
| We group the target activities into four categories | identifying the best placement for the inertial units, as well as the most effcient combination of sensors |

we grouped 16 different activities into four categories : walk ,walking balance ,standing balance, and strength.

***Walk***: it is composed of backwards walking (bawk), sideways walking (sdwk),walking and turning around (wktrn). these three activities have all wide and diverse range of movements, especially for the foot sensors.

***Walking Balance***: it is composed of heel to toe walking backwards (het-owkbk), heel walking (hewk), tandem walking (tdwk), toe walking (towk).These activities are based on similar ranges of movements.

***Standing Balance***: it is composed of single leg stance (sls), and tandem stance(tdst). The signals sampled from these two activities are mainly flat, as they require the subject to move only once in order to change the standing leg from left to right.



***Strength:*** it is composed of knee extension (knex), knee flextion (knfx), hip abduction (hpabd), calf raise (cars), toe raise (tors), knee bend (knbn), and sit to stand (std). As some of these activities are performed by using each individual leg separately, all the sensor configurations involving both right and left sides are not applicable for this group.

The group of 19 participants consists of 7 males and 12 females. Participants have a mean age of 22.94±2.39, a mean height of 164.34±7.07 cm, and a mean weight of 66.78±11.92 kg.

## Dataset

Once the signal is acquired from the activity, it is segmented into small overlap-ping windows of 204 points, corresponding to roughly 2 seconds of movements ,with a stride of 5 points.

2sec movement-> 5point -> Signal -> 204 points (overlap-ping windows)

reduced size of the windows is generally associated with a better classification performance.

each window comes in the form of a matrix of values, of shape 6N×204, where N is the number of sensors used to sample the window.

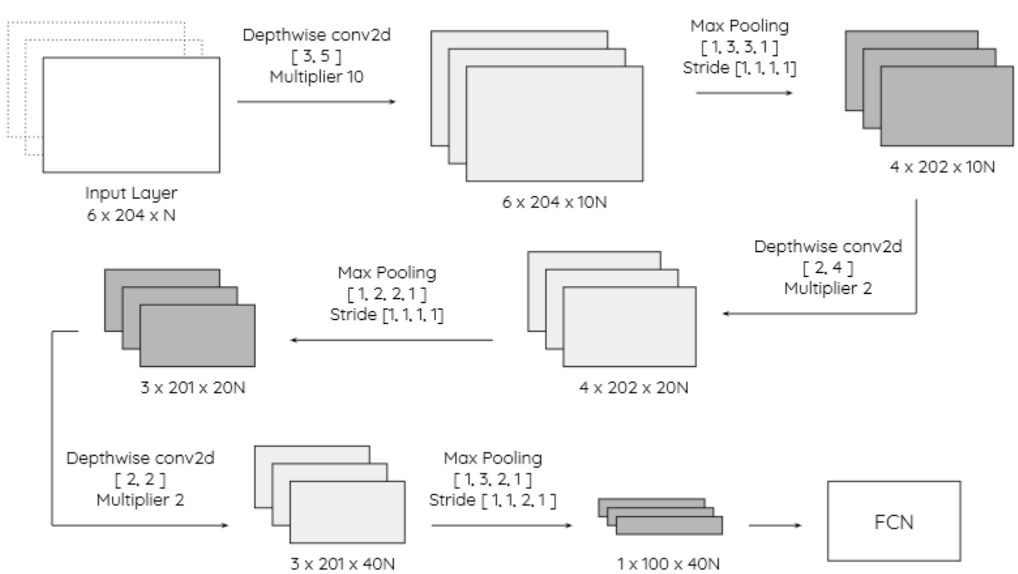
As the activities have different execution times, and different subjects may execute the same activity at different speed, the resulting dataset is not balanced. We partition the available datasets based on the subjects rather than on the windows.

the input of the network has shape 6(dimensions)×204×N, where N is the number of channels and is equal to the number of sensors used for the sampling.

After the input layer, three convolutional layers interleave with three max pooling layers.

The depthwise convolution operation generates multiple feature maps for every input(window:6*204*N) channel, with kernels of size 3×5, 2×4 and 2×2 in the first, second and third convolutional layer respectively.

The input of every convolutional layer is properly padded so that no loss of resolution is determined from the convolution operation.



A fully connected network follows, composed by three dense layers of 500, 250 and 125 units. The dense layers are regularized with dropout during the training  phase .

ReLU function -> softmax function

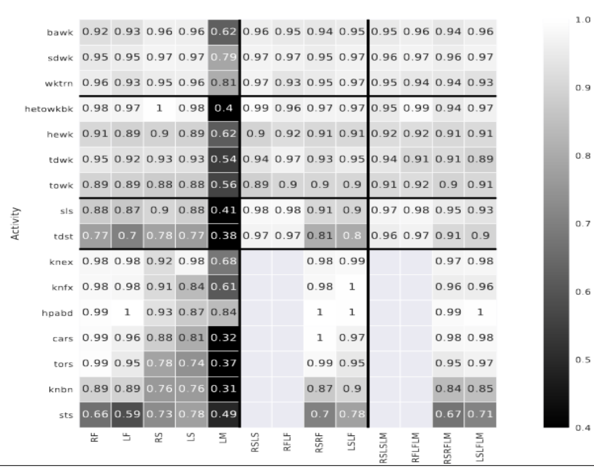
We select a set of hyperparameters that are kept constant for all the activity groups and sensor configurations, based on literature best practices [4] and empirical observations.

The number of training epochs varies from 150 to up to 300, according to the behavior of individual configurations.

## Experimental results

In order to evaluate our classifier, we collect individual precision and recall scores for every combination of sensors and activities, and we then compute the F-scores

Each tile in the picture contains the F-score obtained for the corresponding activity and sensor configuration.




In most cases, combinations of two or three sensors lead to better results compared to the adoption of single inertial units.