Activity recognition based on gyroscope data

SILAS LENZ, Technical University Graz, Austria

1 INTRODUCTION

Human activity recognition is the task of classifying what a user is doing by analysing sensor data. This is most commonly used for fitness tracking, but also has other uses, such as turning off notifications when a user starts cycling or playing music when the user is running. The goal is to separate the following activities: `WALKING`, `RUNNING`, `BIKING` and `SITTING`.

2 DATA COLLECTION

Gyroscope data was collected from a mobile phone at a frequency of around 200Hz[1]. While running the phone was held in the left hand and for all other activities the phone was located in the left pocket. There are recordings from multiple sessions from all activities. In total 3870 seconds of usable data was collected, of which 595 was walking, 1140 running, 1130 cycling, 870 sitting and 135 walking down stairs.

3 PREPROCESSING

All data was split into sample windows of 5 seconds each with no overlaps. Examples of such 5-second chunks can be seen in Figure 1 representing only the rotation around the x-axis. This results in a total of 774 5 second samples.

[1]Using the app Physics Toolbox: https://play.google.com/store/apps/details?id=com.chrystianvieyra.physicstoolboxsuite

Author's address: Silas Lenz, Technical University Graz, Graz, Austria, slenz@student.tugraz.at.

2018. XXXX-XXXX/2018/6-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

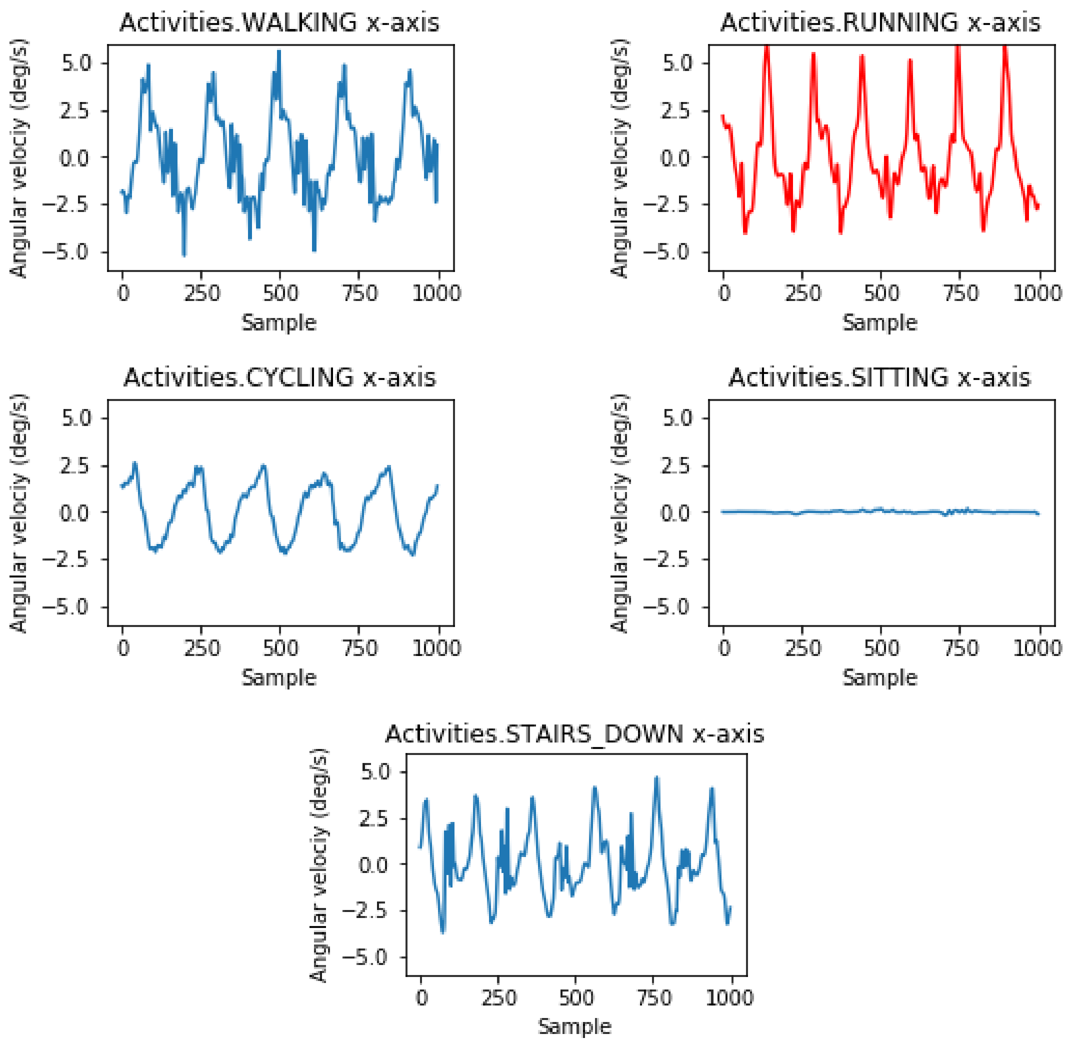

Fig. 1. Examples of gyroscope values around x-axis while doing the different activities

4 A VERY SIMPLE APPROACH

When looking at the plots there are two obvious differences. One is the frequency of movements and one is its amplitude. So an easy classification approach would be to do a fast Fourier transform and look at the main frequencies and their strength. Calculating those values for all chunks results in the following stats.

Value	Walking	Running	Cycling	Sitting	Stairs (down)
Mean main frequency	1.11	1.41	0.86	2.65	1.34
Variance	0.52	0.04	0.23	4.59	0.32
Average main amplitude	2.04	2.55	0.83	0.01	1.16

Table 1

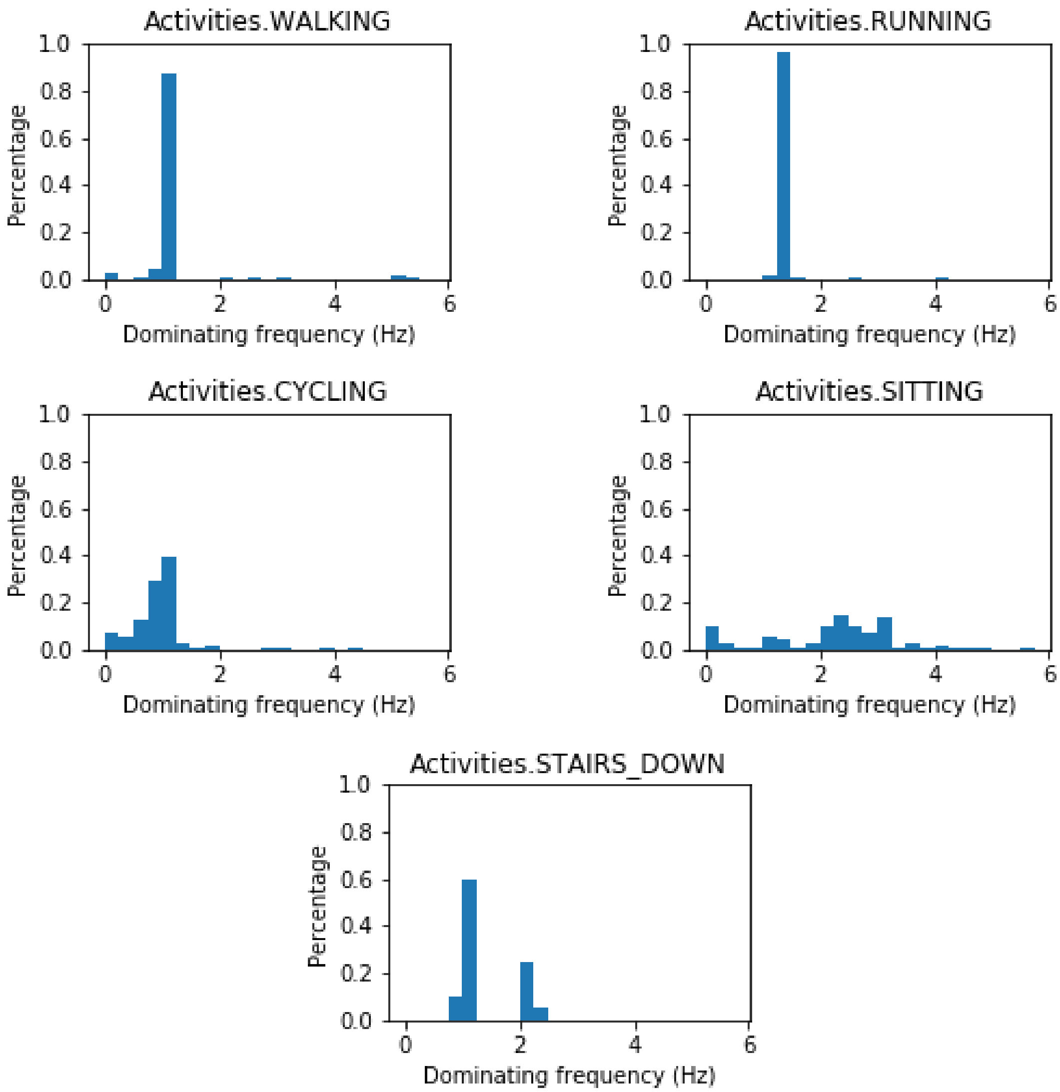

Fig. 2. Histogram on dominating frequency around x-axis per chunk

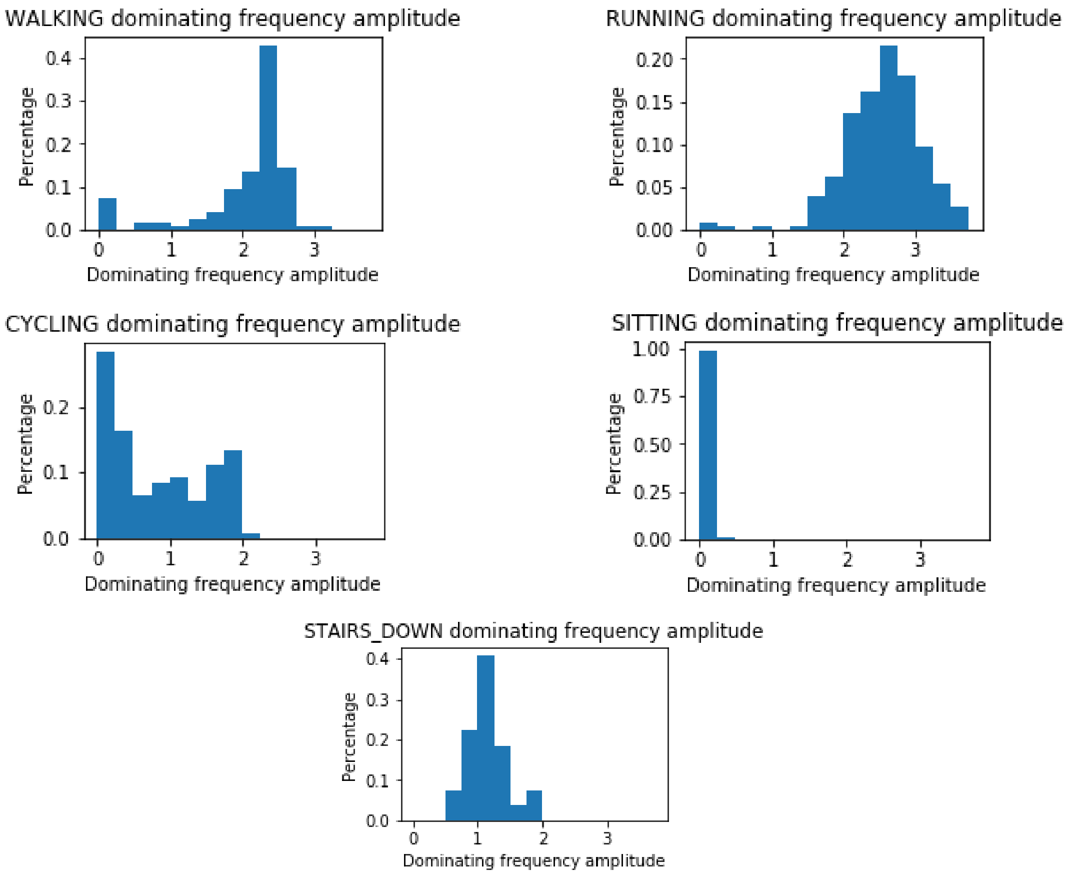

Fig. 3. Histogram on the dominating frequency amplitude around x-axis per chunk

Looking at Figure 2 it is clearly visible that walking and running usually has a higher frequency than cycling, while sitting is all over the place. However when looking at Table 1 and Figure 3 we can see that the average amplitude is close to zero while sitting, so that could be classified as noise. Now we know that running has quick and strong movements, while cycling is slow and "soft", walking is in the middle on both and sitting is very weak but with almost uniform distribution over the frequency space from 0-5Hz.

Since all activities except stairs have clear distinguishing features even when just using the main frequency and amplitude it seems like it would be possible generate a good decision tree. To simplify the model only the rotation around x-axis is used, because it shows the most significant movements with little noise. Using sklearns DecisionTreeClassifier and a 75/25 random split training/test results in the confusion matrix in figure 4. A 10-fold cross validation confirms the reasonably good results with an accuracy of 85.70% (+/- 3.74%). This is now the baseline to beat using more advanced methods.

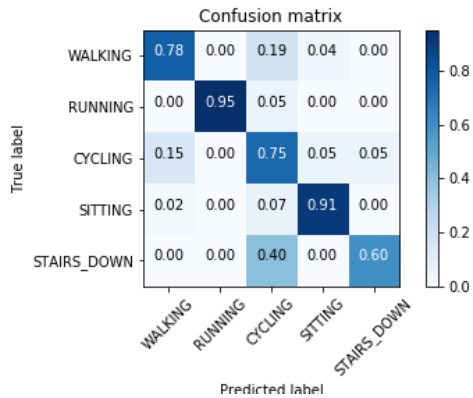

Fig. 4. Confusion matrix with decision tree

5 A MORE ADVANCED APPROACH

Since the sensor data is a time series a reasonable approach would be an LSTM classifier. Here the full dataset is used as the LSTM classifier seems to be able to extract a little bit of additional information from the other axes. Classes are weighted while training to offset the differing amount of data. This results in slightly improved results compared to the simple model. A sample confusion matrix that is quite representative can be seen in figure 5, although these vary much more than using a decision tree. 10-fold cross validation gives 90.43% (+/- 2.69%) accuracy.

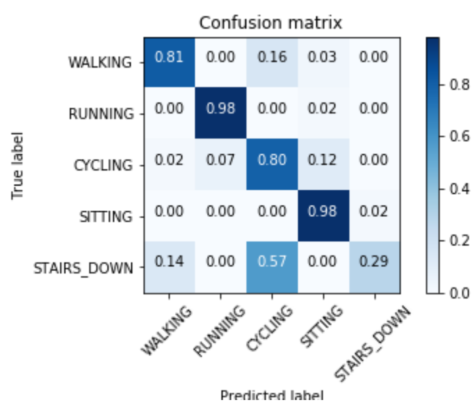

Fig. 5. Confusion matrix with LSTM classifier

6 EVALUATION

Some simple frequency analysis and a decision tree already resulted in good results. Walking down stairs is hard to classify for multiple reasons. One is that I can't be arsed to walk down stairs for 15minutes, so the amount of data available is much lower. The other is that it looks quite similar to cycling due to the lower amplitudes and in about 70% of the cases the frequency is also in the same range. This means that some samples will end up classified as cycling, which is very visible in the confusion matrices 4 and 5, but doesn't affect the accuracy much as it represents just 3.5% of the dataset. There are also sections of flat floor between the stair sections in the data.

Running and sitting is very easy to classify for both classifiers as they look significantly different from the others. Cycling on the other hand has frequencies from 0Hz to 1.5Hz and equally varying amplitudes, which is why portions of walking and walking down stairs are classified as cycling. I would also argue that some of the $\sim$ 10% that the LSTM usually classifies as sitting is accurate but incorrectly labelled and represents the downtime at stoplights.

7 CONCLUSIONS

Both methods have a decent accuracy and for most classes would be "good enough" for fitness tracker purposes. On the other hand there are some low hanging fruit that could potentially improve the results a few more percentage points. Another conclusion is that classical signal processing and simple classification methods should not be rejected in favour of complex machine learning without trying. I expected much worse results for the first model.

8 POTENTIAL IMPROVEMENTS

The simple model uses only one axis of rotation, which should be good enough for most cases but misses the rotations around the y and z axis. While they are usually smaller and mostly consist of noisy data from shaking in the pocket it probably also differs between activities and could be used to further improve accuracy.

Additionally using heart rate measurements was attempted but deemed unrealistic as data exports from the sensor where only available once a month (GDPR requests are the only method of export). Other studies suggest relatively minor improvements[1].

Another sensor that could have been used was the accelerometer. That was initially attempted, but quickly ruled out as it showed much noisier data compared to the gyroscope. Recording both simultaneously also resulted in a frequency of ~25Hz instead of 200Hz. Might have been helpful for the stairs class (different amount of movement up vs. down).

Using a majority vote on three or more chunks in a row could also have resulted in better results and act as a filter for inaccurate source data.

REFERENCES

[1] Saeed Mehrang, Julia Pietila, Johanna Tolonen, Elina Helander, Holly Jimison, Misha Pavel, and Ilkka Korhonen. 2017. Human Activity Recognition Using A Single Optical Heart Rate Monitoring Wristband Equipped with Triaxial Accelerometer. *EMBEC NBC 2017 IFMBE Proceedings* (2017), 587âĂŞ590. https://doi.org/10.1007/978-981-10-5122-7_147