Machine Learning Techniques for Gait Analysis in Skiing

Machine Learning Techniques for Gait Analysis in Skiing

Savya Sachi Gupta[1], Moa Johansson[1], Dan Kuylenstierna[1,2], David Larsson[1], Julia Ortheden[1], Markus Pettersson[1]

[1] Chalmers University of Technology, Gothenburg, Sweden
[2] Skisens AB, Gothenburg, Sweden
savyasachi92@gmail.com, moa.johansson@chalmers.se

Abstract. We investigate the use of supervised machine learning on data from ski-poles equipped with force sensors, with the goal of automatically identifying which sub-technique the skier is using. Our first contribution is a demonstration that sub-technique identification can be done with high accuracy using only sensors in the pole. Secondly, we also compare different machine learning algorithms (LSTM neural networks and random forests) and highlight their respective strengths and weaknesses, providing practitioners working with sports data some guidance for choice of machine learning algorithms.

Keywords: cross-country skiing, gait analysis, machine learning

1 Introduction

In cross-country skiing, the proportion of power generated through the poles depends on which sub-technique the skier is using, which can be measured through sensor-equipped ski poles. In double poling, all force is applied through the poles, whereas in the skating sub-techniques (referred to as “gears 1-5”, following the notation in [5]) an increasing proportion is generated by the legs. To estimate the total work, it is therefore of interest to automatically identify which sub-technique has been used. Furthermore, the most effective sub-technique will depend on features of the terrain, the snow conditions and the individual strengths of the skier. As such, sub-technique identification is of interest for athletes and coaches for several reasons: Firstly, it can directly be used to estimate power to steer intensity in interval training. This is already used in practice by some long-distance specialists, who tend to use only the double poling technique, where all force is applied via the poles. With accurate technique classification, this can also be adapted to conventional cross-country skiers, using other sub-techniques. Secondly, it will also allow to analyse the proportion of time a skier spends in each gear over the course of a race or training session, even without the use of cameras covering the entire course.

Previous work on gait analysis in skiing typically use multiple sensors (accelerometers, IMUs, gyroscopes) attached to the body and/or skis [4, 9, 7, 8, 6, 2]. A more light-weight and flexible system with sensors built into the poles was investigated in the small pilot study by Johansson et al. [3], showing high accuracy in classifying skating sub-techniques using an LSTM neural network model. We use the same ski-pole sensors, provided to us by Skisens AB[3], and extend previous work by training machine learning models on larger datasets, including more individuals and also covering experiments with both classical and skating techniques. Furthermore, we train, evaluate and contrast both neural network models as well as random forest models.

2 Methods and Data

We conduct two separate case studies: one on freestyle skating techniques, and one on classical sub-techniques. In both settings we pre-process the data to identify force peaks representing individual strokes, each labelled with its associated sub-technique. We considered two machine learning models with different properties: a LSTM neural network model, which takes the time-series of sensor readings representing a stroke directly as input, and a random forest classifier, which uses derived features as inputs to learn from (e.g., stroke length, frequency and peak power).These techniques were chosen as they represent two of the main families of machine learning algorithms.

Random forests are ensembles of *decision trees*, a simple idea working well for tabular data, looking at one feature at the time to decide the classification of a datapoint. However, for data collected as a time-series of sensor readings, such features must first be computed. This *feature engineering* step relies on domain expertise, it is by no means always obvious what the best features to use for separating the classes are. If the wrong ones are chosen, the model may not perform as well as one would have hoped. Neural networks on the other hand, work directly on the data (time series in our case). During training, the network adapts and updates its internal parameters to maximise accuracy on the classification task, without the need for a feature engineering step. The network learns whatever features that best helps it to solve the task at hand, however these features are not in general human interpretable, so we may not know *why* a certain label was picked. Still, neural networks have outperformed earlier methods on a wide range of tasks such as natural language processing and computer vision. The implementations of the machine learning models use standard Python libraries and are described in more detail in [1].

We used two datasets described below, each collected using Skisens handles sampled at 100Hz. At each time-step, the force in both the left and right pole was recorded. Unlike earlier work this new version of the Skisens handles only record force, and does not include angular data as in [3]. For the LSTM models, the time series was split into separate pole-pushes and zero padded to the same length.

Dataset 1: Freestyle techniques. The dataset for freestyle skiing was provided by Skisens, and contain data collected from ten junior elite skiers (7 male, 3

[3] www.skisens.se

female) skiing on an indoor treadmill with sub-technique, speed and slope varied according to a set protocol. The sub-techniques considered were double poling (DP) and skating gears 2, 3 and 4[4] (G2, G2, G4). In total this dataset contains 16 007 pole pushes, however not evenly distributed between the gears, with DP and G3 being twice as frequent as G2 and G4.

Dataset 2: Classical techniques Unlike the Dataset 1, the protocol for data collection of classical techniques was designed by the authors, specifically with machine learning in mind to ensure roughly the same amount of time in each sub-technique (double poling (DP), step double poling (SDP) and diagonal stride (DS)). Data was again collected from junior elite skiers (9 male, 5 female) on an indoor treadmill. In addition, some data was also collected from outdoor roller skiing (3 male subjects). In total the dataset contains 33 519 pole pushes roughly evenly distributed between the three classical sub-techniques.

3 Results

Case Study 1: Freestyle techniques. Both models were evaluated using 10-fold cross validation with seven skiers used for training the models and data from three skiers held out as a test-set. Smaller classes (G2 and G4) were over-sampled in an attempt to compensate for the class imbalance in the dataset. The random forest model reached an average accuracy of 78% (DP 69 %, G2 87%, G3 70%, G4 63%), while the LSTM model reached only 63% (DP 60%, G2 52%, G3 65%, G4 86%). While the LSTM did seemingly well for gear 4, it often misclassified gears 2 and 3 into this category (20% and 10% respectively). The Random Forest on the other hand, most often mistook gear 4 for double poling (23% of samples). We noticed that there were some data-issues where the time series from the left and right handle were not always in sync. This, in combination with the unbalanced classes of sub-techniques might have contributed to the comparatively low performance for the LSTM model, while the Random Forest was less affected.

Case Study 2: Classical techniques. Evaluation of the models trained for classical techniques were again evaluated using 10-fold cross validation. The Random Forest model achieved an average accuracy of 74% (DP 72,5%, DS 86%, SDP 69%), while the LSTM model performed considerably better with an average accuracy of 86% (DP 83%, DS 92.5%, SDP 84%). The most common mistake in both Random Forest and LSTM models was between double poling and step double poling (Random Forest: 15% and 19% respectively, LSTM: 13% and 10%), which was not surprising, as hand movements are very similar.

4 Discussion and Conclusion

We have shown that using only data from force sensors in the skipole handles it is possible to accurately classify cross-country skiing sub-techniques, both in

[4] Gears 1 and 5 were not included as gear 1 is almost never used in practice, and gear 5 only uses the legs.

the classical and freestyle cases. The Random Forest model appears to be less sensitive to the size and quality of the dataset, as it shows similar performance in both case-studies. We hypothesise that the relatively poor performance of the LSTM model on freestyle techniques was primarily due to data quality issues, as better results have been achieved in a previous smaller study [3]. In Case Study 2, on classical techniques, the LSTM model outperforms the Random Forest. We believe this is because the training dataset is larger, of better quality and class balanced. Note that for DP, which appear in both case studies, the LSTM accuracy increases from 60% to 83%, which would support this hypothesis.

The Random Forest model can be sensitive to which features it is passed, in an earlier experiment where additional features were present, these proved decremental to performance (see [1] for details). Choosing and computing the right features for learning is both important and highly non-trivial, and should ideally also involve a domain expert. On the other hand, an attractive aspect of neural networks, such as LSTMs, is that they do not need to be explicitly told what features to look for in the raw time-series data for the pole push. Given enough labelled examples the neural network will adapt and discover some distinguishing features of the sub-techniques itself. However, as learned information is encoded in the network's many trainable parameters (the weights), it is not obvious to a human user what these distinguishing features are.

References

1. S. S. Gupta. Predicting cross-country skiing techniques using machine learning. Master's thesis, Chalmers University of Technology, 2021.
2. J. Jang, A. Ankit, J. Kim, Y. Jang, H. Kim, J. Kim, and S. Ziong. A unified deep-learning model for classifying the cross-country skiing techniques using wearable gyroscope sensors. *Sensors*, 18(11), 2018.
3. M. Johansson, M. Korneliusson, and N. L. Lawrence. Identifying cross country skiing techniques using power meters in ski poles. In *Nordic Artificial Intelligence Research and Development. NAIS 2019*, 2019.
4. F. Marshland, K. Lyons, J. Anson, G. Waddington, C. Macintosh, and D. Chapman. Identification of cross-country skiing movement patterns using micro-sensors. *Sensors*, 12(4), 2012.
5. J. Nilsson, P. Tveit, and O. Eikrehagen. Effects of speed on temporal patterns in classical style and freestyle cross-country skiing. *Sports Biomechanics*, 3(1), 2004.
6. O. Rindal, T. Seeberg, J. Tjønnås, P. Haugnes, and Ø. Sandbakk. Automatic classification of sub-techniques in classical cross-country skiing using a machine learning algorithm on micro-sensor data. *Sensors*, 18(2), 2017.
7. Y. Sakurai, F. Zenya, and Y. Ishige. Automated identification and evaluation of subtechniques in classical-style roller skiing. *Journal of Sports Science and Medicine*, 13, 2014.
8. Y. Sakurai, F. Zenya, and Y. Ishige. Automatic identification of subtechniques in skating-style roller skiing using inertial sensors. *Sensors*, 16, 2016.
9. T. Stöggl, A. Holst, A. Jonasson, E. Andersson, T. Wunch, C. Norström, and H.-C. Holmberg. Automatic classification of the sub-techniques (gears) used in cross-country ski skating employing a mobile phone. *Sensors*, 14, 2014.