Logistic regression with elastic net regularisation can successfully identify groups of runners based on their technique using consumer-oriented wearable sensor data

Adrian R. Rivadulla[1, 2], Xi Chen[2] Dario Cazzola[1], Grant Trewartha[3], Ezio Preatoni[1]
[1]Department for Health, University of Bath, Bath, UK
[2]Institute for the Augmented Human, Department of Computer Science, University of Bath, Bath, UK
[3]School of Health & Life Sciences, Teesside University, Middlesbrough, UK

Email: arr43@bath.ac.uk

Summary

We previously identified that runners can be clustered in two main groups based on their technique: the "neutral" and "tilted pelvis" clusters. Cluster-specific interventions may enhance training success, but identifying what cluster an athlete belongs to currently requires a lab-based optical motion capture system. Here, we develop and validate a regularised logistic regression model that uses data from consumer-oriented wearable technologies to allocate runners to one of the two running technique classes. Using 2-3 convenient sensor locations was enough to achieve testing scores ≥ 0.82, enabling reasonably confident allocation of new runners to the neutral and tilted pelvis techniques. This method facilitates large scale cluster-specific training development and provides real-world solutions for the wider running community.

Introduction

Technique-based training in running is influenced by inter-individual variability, requiring personalised training approaches over "one size fits all" methods [1]. However, individualised training may not be accessible to every runner. Previously, we identified two clusters of runners: the "neutral pelvis" and the "tilted pelvis" clusters [2]. The tilted pelvis cluster exhibited smaller duty factor and greater anterior pelvic tilt, trunk to pelvis extension and hip flexion throughout the stride. These differences in technique could inform targeted training to enhance coaching interventions.

The original cluster identification relied on lab-based motion capture, which is usually restricted to research or elite sport facilities. Wearable sensors are widely used by runners and present a more practical alternative to enable the broader implementation of cluster-specific training. This study aimed to develop a classification algorithm to allocate runners to the neutral and tilted pelvis groups using wearable sensor data.

Methods

The same 84 runners previously used for clustering [2] were included in this study. Participants ran at 12 km/h on a treadmill for 4 minutes whilst wearing pressure insoles (NURVV, London, UK; 1000 Hz) which computed spatio-temporal running parameters, and six inertial measurement units (IMUs) (Delsys, Massachusetts, US; 370.7 Hz) placed on popular consumer-based locations (foot, tibia, sacrum, trunk, upper arm, wrist).

Input features for the classifier were generated using domain knowledge and comprehensive time series feature extraction via the "tsfresh" algorithm [3]. A logistic regression model was trained to allocate runners to the neutral and tilted pelvis classes using elastic net regularisation to reduce overfitting and recursive feature elimination to improve model interpretability and efficiency. Participants were randomly allocated to the development (70%) or test (30%) sets for model training and final testing preserving class balance. Hyperparameter optimisation for elastic net regularisation and recursive feature elimination were performed on the development set following a 5-fold cross-validation scheme.

Results and Discussion

The top three performing models required only 2-3 sensors and achieved testing scores ≥ 0.82 (Table 1).

Table 1. Testing scores for top three sensor networks named by their location (IMUs) or type (insole).

Model	Accuracy	Precision	Recall	F1
Foot, insole	0.85	0.80	0.85	0.84
Tibia, sacrum	0.83	0.84	0.83	0.83
Upper arm, wrist, insole	0.82	0.83	0.82	0.82

These results suggest that the neutral and tilted pelvis running techniques can be identified reasonably accurately without the need for an optical motion capture system. The identified wearable sensor networks and models can allocate new unseen runners to the neutral and tilted pelvis groups, enabling the design and testing of large-scale cluster-specific training interventions. These sensor networks were sparse and used locations of consumer-based wearables, integrating seamlessly with the standard kit of a runner and facilitating the integration of these models on existing technologies.

Conclusions

The algorithms presented in this study enable runners to use widely adopted wearable sensors to identify which cluster they belong to and to plan more personalised coaching interventions.

References

[1] van Hooren, B et al. (2020). *J Sports Sci*, **38**: 214-230.
[2] Rivadulla et al. (2024) *Sports Biomech*: 1-24.
[3] Christ et al. (2018) *Neurocomputing*, **307**: 72-77.