Human Movement Science 87 (2023) 103042

Contents lists available at ScienceDirect

Human Movement Science

journal homepage: www.elsevier.com/locate/humov

Perspective on “in the wild” movement analysis using machine learning

Eva Dorschky [a], Valentina Camomilla [b], Jesse Davis [c], Peter Federolf [d], Jasper Reenalda [e,f], Anne D. Koelewijn [a,*]

[a] Machine Learning and Data Analytics (MaD) Lab, Department Artificial Intelligence in Biomedical Engineering (AIBE), Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[b] Department of Movement, Human and Health Sciences, University of Rome “Foro Italico”, Rome, Italy
[c] Department of Computer Science and Leuven.AI, KU Leuven, Leuven, Belgium
[d] Department of Sport Science, University of Innsbruck, Innsbruck, Austria
[e] Biomedical Signal and Systems group, University of Twente, Enschede, The Netherlands
[f] Roessingh Research and Development, Enschede, The Netherlands

ARTICLE INFO

Keywords:
Movement analysis
Sports
Machine learning
Wearable sensors
Free-living

ABSTRACT

Recent advances in wearable sensing and machine learning have created ample opportunities for “in the wild” movement analysis in sports, since the combination of both enables real-time feedback to be provided to athletes and coaches, as well as long-term monitoring of movements. The potential for real-time feedback is useful for performance enhancement or technique analysis, and can be achieved by training efficient models and implementing them on dedicated hardware. Long-term monitoring of movement can be used for injury prevention, among others. Such applications are often enabled by training a machine learned model from large datasets that have been collected using wearable sensors. Therefore, in this perspective paper, we provide an overview of approaches for studies that aim to analyze sports movement “in the wild” using wearable sensors and machine learning. First, we discuss how a measurement protocol can be set up by answering six questions. Then, we discuss the benefits and pitfalls and provide recommendations for effective training of machine learning models from movement data, focusing on data pre-processing, feature calculation, and model selection and tuning. Finally, we highlight two application domains where “in the wild” data recording was combined with machine learning for injury prevention and technique analysis, respectively.

1. Introduction

Measuring sports movement during training and competition allows monitoring athletes’ performance and their risk of injury (Camomilla, Bergamini, Fantozzi, & Vannozzi, 2018; Cust, Sweeting, Ball, & Robertson, 2019). Performance monitoring is relevant to assess motor capacity and physical demand, as well as to analyze technique and how technique impacts performance (Camomilla et al., 2018). In the complementary perspective of sport-related injuries, monitoring can be oriented to preventing, assessing, and informing the recovery from injuries (Preatoni et al., 2022).

* Corresponding author.
E-mail address: anne.koelewijn@fau.de (A.D. Koelewijn).

https://doi.org/10.1016/j.humov.2022.103042
Received 21 January 2022; Received in revised form 1 September 2022; Accepted 19 November 2022
Available online 6 December 2022
0167-9457/

The state-of-the-art approach to measure movement is to perform a biomechanical analysis from data recorded in a lab environment using optical motion capture (OMC) systems and force plates. Such an analysis includes calculating spatio-temporal variables, as well as joint angles, joint moments, and ground reaction forces. Musculoskeletal models can provide additional insights into muscle forces and activations, which cannot be measured directly. However, these state-of-the-art measurement techniques can only be used in a laboratory environment and are limited by high costs, a stationary setup, and a short duration.

To enable measurements in a natural environment, several new types of wearable sensors have been developed. Wearable sensors are sensors that are worn on or close to the body and measure the state of the human body, while being minimally invasive to the person who wears the sensor (Düking, Hotho, Holmberg, Fuss, & Sperlich, 2016). This permits performing measurements in any location with minimal change to the person's natural behaviour. In sports applications, these are used to measure movement and inertial forces during sport-related tasks (Aroganam, Manivannan, & Harrison, 2019), while they are also applied for medical purposes, such as classification of Parkinson's disease (Klucken et al., 2013). In sport applications such as performance enhancement (Camomilla et al., 2018) and injury risk mitigation (Sheerin, Reid, & Besier, 2019; Willy, 2018), measurements are typically performed using measurement units based on inertia (inertial measurement units, IMUs), often used in association with magnetic field sensors (magneto inertial measurement units, MIMUs). These new technologies allow for data collection outside of a lab setting, granting ecological validity while having a limited cost and improved portability with respect to lab-based equipment (Macadam, Cronin, Neville, & Diewald, 2019). Along with their ease of use, these characteristics are at the basis of a huge positive potential to provide actionable insights to sport practitioners in the field.

However, this positive potential is paralleled by several pitfalls that could lead to improper use (Seel, Kok, & McGinnis, 2020). The reliability of the measured quantities from wearable sensors may be affected by sensor noise and drift, calibration errors, movement artifacts, data transmission errors, or ferromagnetic disturbances (see Hughes et al., 2021 for a list of sport assessments more prone to them). In addition, movements "in the wild" must be evaluated based on incomplete measurements. Measurements are incomplete because on the one hand, not all types of sensors are available compared to a lab setup (e.g., ground reaction force sensors) and, on the other hand, researchers often aim to minimize the number of sensors to simplify measurements, increase mobility, and decrease costs. Therefore, to date, only a limited set of variables can be adequately computed from wearable sensor data such as acceleration peaks (e.g., peak tibial acceleration), temporal parameters (e.g. foot contact timings, stride, step or (swimming) stroke duration, and frequency), and vertical forces. However, other kinematic variables (e.g. step length, instantaneous velocity, displacements), absolute or relative body segment orientation (joint angles), and dynamic variables (e.g. horizontal force, momentum, stiffness) are thought to be more informative for performance and injury-specific monitoring (Camomilla et al., 2018; Willy, 2018). Computing these variables from inertial sensor data remains challenging due to the aforementioned reliability issues. Therefore, researchers are developing algorithms to fuse sensor signals (Kok, Hol, & Schön, 2017) and incorporate a priori knowledge about the measurement signals and specific movements (see Camomilla et al., 2018 for a list of examples).

Besides these physics-based processing methods, data-driven, or machine learning, methods are also explored as a potential option to estimate the variables of interest. Machine learning is a powerful tool where the recorded data are used to make inferences about the data themselves (Bishop, 2006). Supervised learning, unsupervised learning, and reinforcement learning are increasingly investigated in human movement analysis, for example, to classify normal and pathological gait, to map IMU data to biomechanical variables, to discover or cluster movement patterns, or to learn controllers that drive biomechanical models (Cust et al., 2019; Ferber, Osis, Hicks, &

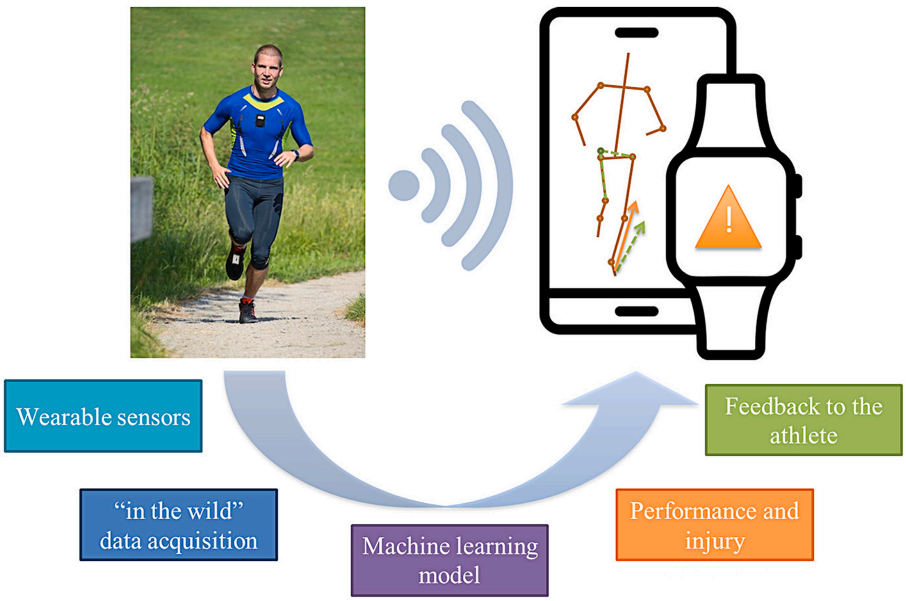


Fig. 1. Overview of "in the wild" movement recording and analysis using machine learning.

Delp, 2016; Gurchiek, Cheney, & McGinnis, 2019; De Vree & Carloni, 2021). Machine learning models have been developed mainly for image analysis and speech recognition, and their potential in mobile motion analysis is far from being fully exploited (Ferber et al., 2016; Gurchiek et al., 2019). However, similar to wearable sensors, the positive potential of machine learning is paralleled by pitfalls in its implementation that could lead to improper model outcomes. Therefore, both for wearable sensors and machine learning algorithms, a full awareness of both sides of the coin is at the foundation of obtaining actionable insights and exploiting these sensors' and methodology's potential within the boundaries set by their limitations (Hughes et al., 2021).

Therefore, in this perspective paper, we aim to provide an overview of current methods and challenges regarding "in the wild" movement analyses using wearable sensors and machine learning (Fig. 1). This perspective paper summarizes and extends upon the symposium "Smart Technologies and Wearable Sensors" which was held during the sixth International Conference on Complex Systems in Sports (ICCSS). First, we discuss six questions that should be answered when designing a protocol for any movement recording study, and suggest procedures to answer these questions. Then, we discuss benefits and important considerations related to the use of machine learning models for temporal movement data, and provide a checklist to aid researchers developing machine learning models. We conclude by discussing concepts relevant to two specific application domains for machine learning applied to "in the wild" movement recording and analysis.

2. Designing an "in the wild" movement analysis study

To successfully monitor sports movements, six questions should be answered: where?, who?, why?, what?, when? and how?. Table 1 provides an overview of possible answers to these questions. So far, the first three questions have been delimited: the ideal "where" for movement monitoring is the field of sport performance, coaches and athletes are the main actors "who" monitor it, and "why" can be split up into movement quality assessment, e.g., to monitor performance, injury prevention, and movement classification, e.g. to assess the effect of a training program. The appropriate choice of data analytics and modeling methods is then driven by the preliminary identification of "what" variables provide the best trade off between being informative and reliably measured "in the wild", "when" to assess them according to the aim of the evaluation, and "how" the retrieved data can be built into actionable insights, also in the light of the recent increase of using artificial intelligence towards this aim (Araújo, Couceiro, Seifert, Sarmento, & Davids, 2021).

The answer to the question of "what?", requires selecting the key variables on the basis of a well focused research question that is rooted in the knowledge of the performance model specific to the sport under analysis. These key variables should be presented to the user, who could be a sports scientist, an athlete, or their coach, in an understandable way. Here, a participatory design could aid successful exchange of the identified key variables. The role of these variables in describing performance or predicting injury-related information can be proved using machine learning approaches (Claudino et al., 2019). Furthermore, if these key variables are unknown, unsupervised learning and causal discovery of long-term data monitoring can reveal "what" variables can provide insight into the research question as well (Reenalda, Maartens, Buurke, & Gruber, 2019).

The answer to the question "when?", or the timing of the test, as well as the measurement frequency, depends on the aim of testing. A general monitoring of an athlete's performance or their capacity profile can benefit from a sparse sampling at the beginning, middle and end of a season. A weekly test, or one scheduled according to the time of adaptation to training stimuli could be more adequate to monitor the efficacy of training strategies, especially in adaptation to huge changes in external factors (e.g. injuries or team changes). Continuous monitoring could support injury forecasting systems. These systems attempt to estimate future states of a complex system, which are the result of the interaction among many factors of a constantly changing environment. These interactions and factors all influence the macroscopic dynamic behavior of the athlete (Fonseca et al., 2020). Besides the aim of testing, the measurement frequency also depends on the identified key variables. Depending on these variables, the speed of the movement of interest, and the used sensors, an appropriate measurement frequency should be chosen. To monitor training efficacy, a high frequency measurement of the trained motion is required, while for performance monitoring, e.g., recording a task duration or a maximal force during the task might be sufficient.

The answer to the question "how?" involves transforming wearable-based measures of movement into actionable insights. The type

Table 1
Checklist for designing an "in the wild" movement analysis study.

Question	Answer
Where are the movements being monitored?	"in the wild"
Who is monitoring the movements?	Athletes and coaches
Why is movement monitored?	Movement quality assessment (performance monitoring)
	Injury prevention
	Movement classification (training intervention)
What measurements and models are required?	If output parameters are known, find measurements and models that are informative and reliable.
	If output parameters are not known, use unsupervised learning or causal discovery.
When are measurements taking place?	Depends on **why?**
	Performance monitoring: sparsely throughout season
	Injury prevention: continuously
	Training intervention: before and after training program
How are the measurements being used?	Create actionable insights
	Find desired and undesired behavior

of measurement, as well as the number of different measurements should be defined, which are necessary to achieve the actionable insight. If multiple sensor are used, their synchronization is important to successfully fuse their signals. Furthermore, a context specific evaluation framework is required that allows identifying the correct form of specific movements and the common deviations according to an agreed clinical consensus (e.g. retreats on anterior cruciate ligament injuries (Shultz et al., 2019), or patellofemoral pain (Collins et al., 2018), or current evidence synthesized with expert opinion for gait retraining (Barton et al., 2016)). To build this framework, (1) sensors must be exposed to the specific movement characteristics of correct and incorrect performance, classified through expert supervision and based on clinical consensus or scientific evidence; (2) a data query pathway must be developed to identify if the task is in its correct form or not (i.e., if selected key parameters stay within safe limits, as defined in the former step, possibly tailored to the analysed subject) and, if incorrect, to identify specific deviation; (3) an actuation strategy, to be built using expert-supervision, is then required to, if correct, send a positive feedback or, if incorrect, send specific corrective feedback, either in real time or right after the movement execution (Whelan et al., 2016).

3. Machine learning models for “in the wild” movement analysis

The combination of machine learning and “in the wild” recordings can greatly improve the transformation of movement measures into actionable insights, since machine learning models can potentially be used for real-time analysis, while analyses can also be performed on large datasets. Machine learning models have been implemented for a range of applications, including classification, regression, and clustering. Classification of technique (Rindal, Seeberg, Tjønnås, Haugnes, & Sandbakk, 2017) or specific events (Giles, Kovalchik, & Reid, 2020) during the movement can be done more efficiently with machine learning models than based on expert annotation. Furthermore, machine learning models for regression have recently been shown similar or even better performance than physics-based models (Dorschky et al., 2020; Hernandez, Dadkhah, Babakeshizadeh, & Kulić, 2021). Machine learning models have also been used for clustering, to derive features from data that cannot easily be identified by experts, such as differences and similarities between groups of athletes (Clermont, Benson, Edwards, Hettinga, & Ferber, 2019) or important features of technique (Federolf, Reid, Gilgien, Haugen, & Smith, 2014; Witte, Ganter, Baumgart, & Peham, 2010).

Once they have been trained, machine learning models can be implemented in efficient and dedicated hardware to make real-time predictions on the incoming data stream. The real-time potential can be used to enhance the performance of athletes during training

Table 2
Checklist for developing a machine learning model.

Data pre-processing		
Extrapolation	Does my training data include the application?	Environmental conditions Participant characteristics Movement types Sensor placement
Cross-validation	Is the validation data independent of training data?	Leave (at least) one subject out Respect temporal dependencies
Data augmentation	How can we generate more training data?	Perturb input data Simulate training data
Feature calculation		
Feature computation	Do I know what features to use to generate the desired output?	Yes, hand-craft features No, automatic extraction
Feature selection	How to pick the best features?	Filter-based methods Wrapper methods Embedded methods
Feature learning	Can I avoid manual feature computation and selection?	Deep learning Representation learning
Model selection and tuning		
Type	Do I have labels which I want to predict?	Yes, supervised learning No, unsupervised learning
Model type	What model type fits my application best?	Traditional machine learning models Models with uncertainty Deep neural networks
Model fit	Is the training and validation error high? Is there a big gap between training and validation error?	Underfitting Overfitting
Little data	Is my model over-parameterized?	Explicit regularization Implicit regularization
	Do I have any prior knowledge of the underlying physics?	Data augmentation Physics-based constraints Physics-guided neural networks
	Are related data/ models available?	Transfer learning
Carbon footprint	Does a simple algorithm already provide good results?	Efficiency vs. accuracy

and competition, specifically by providing real-time feedback to athletes and coaches, by developing tools for technique correction (Camomilla et al., 2018), or by measuring injury predictive factors (Claudino et al., 2019). Its advantages for large datasets can be exploited when performing long-term monitoring for injury forecasting systems (Fonseca et al., 2020), especially for athletes with disabilities, where personalized systems can be developed that monitor to anticipate injuries specific to their disability (Rum et al., 2021).

Another advantage of machine learning models is that they can learn invariant representations of the data, eliminating the need for extensive and error-prone calibration (Hernandez et al., 2021). In contrast, physics-based approaches often require models to be calibrated, e.g. scaled to the subject's anthropometrics, while sensor-to-segment alignments need to be estimated (Stanev et al., 2021). This is a limiting factor when analyzing movement "in the wild" as the data collection is unsupervised and calibration may be performed incorrectly.

4. Developing an appropriate pipeline to train machine learning models

While machine learning can be a powerful technique to learn predictive models from movement data, these data pose several challenges that researchers must take into consideration during their analysis. Therefore, the success of machine learning is highly dependent on the availability of training data, data processing, and model selection. Several considerations arise while pre-processing the data, while calculating features, and during model selection and tuning, as shown in Table 2.

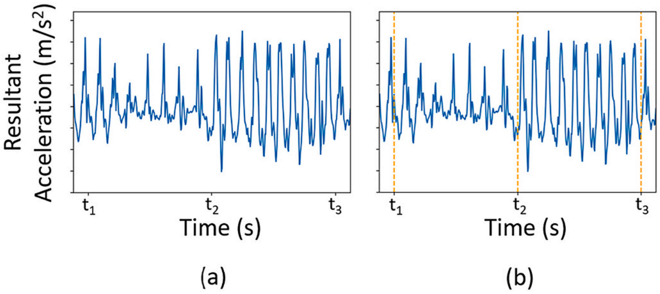
4.1. Data pre-processing

Data pre-processing is performed to ensure that the data is of sufficient quality to develop a model of good quality. In this step, the data is split up into a training, validation, and test set. Furthermore, it is ensured that the data is of good quality and appropriate for the model application. The training data is used to learn the model and its performance is evaluated on the validation set. The test set cannot play any role in constructing the model and serves as a final evaluation after model development and tuning. However, the difficulties of collecting data and the nature of movement data itself pose issues for which we provide recommendations that can be followed when designing the evaluation setup.

First, machine-learned models often struggle to extrapolate from the data used to train the model (Sambu Seo, Wallat, Graepel, & Obermayer, 2000). Current machine learning models have been trained and evaluated on laboratory data. Therefore, it is questionable how well they perform "in the wild", since movement patterns may change in outdoor environments due to uneven surfaces, slopes, or weather conditions (Milner, Hawkins, & Aubol, 2020). Furthermore, not only external conditions matter, also the characteristics of the study participants should be sufficiently similar. For example, suppose the goal is to train a system to recognize activities such as going up and down stairs, walking, or running. Moreover, the model is trained on data collected from healthy 20 and 30 year olds. However, if you try to deploy this learned model on elderly people suffering from osteoarthritis, then the model will perform very poorly (i.e., will incorrectly detect many activities) because this population of individuals exhibits radically different movement patterns than the younger, healthier population used to train the model. A machine learned model must be trained on data that is similar to the data it will encounter "in the wild". This requirement is important to consider when designing a data collection protocol.

Second, a subtler version of the aforementioned problem can arise from the fact that (1) movement models are trained on data from a small number of subjects because data collection is expensive and time consuming and (2) each individual has a unique movement pattern. In this situation, we recommend to perform cross validation by dividing the data on a subject level (De Brabandere et al., 2020). That is, all the data belonging to one subject should either appear only in the training set or only in the testing set. This makes the setup more realistic as you will never have training data about all subjects that the model will be deployed on. This setup evaluates the model's ability to make accurate predictions on unseen subjects.

Third, there are temporal dependencies in the data that must be considered and respected. For example, suppose you have collected data by asking someone to perform a set of multiple consecutive repetitions of an activity (e.g., a physical therapy exercise). The



Min (m/s²)	Max (m/s²)	Mean (m/s²)	StDev (m/s²)	Number of Peaks	Label
5.74	17.90	9.70	2.02	8	Walking
2.43	18.74	9.96	4.17	10	Descend Stairs

(c)

Fig. 2. (a) Shows the resultant of the acceleration recorded by a triaxial accelerometer collected from someone performing daily activities. (b) Shows the signal divided into two five second windows. (c) A simple feature-based representation of each window. The features are the minimum, maximum, mean, standard deviation and number of peaks within the window. The last column shows the activity performed during (the majority of) each window.

temporal nature of the data means that the data from consecutive repetitions will be highly correlated. Therefore, all data about one set of repetitions should either be in the training set or the test set (Decroos, Schütte, De Beéck, Vanwanseele, & Davis, 2018; Singh et al., 2021). It should not appear in both. Similarly, when analyzing data collected from athletes over the course of the season, it can also be important to use a temporal split of the data (Jaspers et al., 2018; Rossi et al., 2018). In this setup, data from the first half of the season is used to train the model and the model's predictions are evaluated on the second half of the season. In both cases, failing to respect the temporal ordering and correlations can yield over optimistic estimates of a model's predictive performance.

4.2. Feature calculation

When applying machine learning algorithms to data produced "in the wild", these algorithms can in general not operate directly on the data produced by sensors, which are time series (Fig. 2a). Standard machine learning algorithms require that the data is represented in a spreadsheet-like table (Fig. 2c). The typical approach is to convert the data into a tabular format by partitioning the series into a number of windows (Fig. 2b) and and defining a number of features about the signals within a window such as simple statistics (min, max, etc.), advanced statistics (entropy, kurtosis, etc.), time series features (number of peaks, autocorrelation), functions (linear trends, FFT, etc.), and domain-specific features (e.g., stride regularity for running). Sometimes it is even necessary to derive new signals. For example, computing and analyzing the resultant of the acceleration can help make the analysis more robust to different sensor attachments (e.g., account for different ways the sensor's axes are oriented) (Op De Beéck, Meert, Schütte, Vanwanseele, & Davis, 2018; De Brabandere, Op De Beéck, Hendrickx, Meert, & Davis, 2021; Sheerin et al., 2019). Similarly, it can also be useful to analyze the derivative of the acceleration (Robberechts et al., 2021). Unfortunately, trying to hand-craft the most appropriate features for a given problem is an arduous and time consuming task. Luckily, there has been significant progress in producing open-source software for partially automating this task for time series data (Christ, Braun, Neuffer, & Kempa-Liehr, 2018; De Brabandere et al., 2021; Fulcher & Jones, 2017). Advanced machine learning models, such as convolutional neural networks, can automatically learn features from the data, replacing manual feature engineering.

4.3. Model selection and tuning

Model selection is a challenging component of design good machine learning models, because of the large number of possible models, the fact that there is no single best-performing model according to the "no free lunch" theorem (Wolpert & Macready, 1997), and the different tasks (classification, regression, and clustering) that can be performed. Classification and regression models are trained using supervised learning, and can be used to distinguish between normal and abnormal movement patterns based on the derived features (classification) or one can learn an equation that maps the feature values to a measure of performance (regression). Models typically used for these tasks can be both traditional machine learning models and deep neural networks. Traditional machine learning models, such as support vector machines, random forests, decision trees, and smaller neural networks, can be trained using smaller amounts of data. Deep neural networks, large neural networks with many layers and tuneable parameters, can be applied to very large datasets. Besides model type, a distinction can also be made in the type of output, which can be exact or it can include an uncertainty distribution, such as in Bayesian models, where a distribution is calculated. Including uncertainty can be important when risks should be avoided as much as possible, for example for tasks related to injury prevention. When the model has an exact output, the prediction uncertainty can be estimated using post-processing, for example by calculating the distance to the decision boundary in support vector machines (Kremer, Steenstrup Pedersen, & Igel, 2014). On the other hand, unsupervised machine learning can reveal patterns from unlabeled data, and is usually applied for clustering, for example to automatically identify groups with similar technique, and dimensionality reduction, for example to better understand similarities between movements of individuals. Typical models are k-means clustering for clustering or principal component analysis for dimensionality reduction.

Recently, deep neural networks have become popular for classification and regression tasks. These models are neural networks with many processing layers, which outperform traditional machine learning models in many tasks. When using deep neural networks, convolutional layers, which act like a filter, can be added for automatic feature extraction. Convolutional layers are effective for time-series sensor data, since there is a correlation over time and sensor axes. Therefore, we recommend that the ordering of sensor data is chosen to take advantage of these correlations between sensor axes. In addition, recurrent layers can model the dynamic behavior over time, and allow for real-time calculation of variables based on a short-term history. A disadvantage of deep neural networks is that a sufficient amount of training data is required, and training consumes a tremendous amount of energy (Schwartz, Dodge, Smith, & Etzioni, 2020). Furthermore, complex neural networks tend to be less reliable, as a different locally optimal model is found each time training is performed. Therefore, we recommend to re-initialize and perform training more than once and apply the model with the lowest error. To avoid the disadvantages of deep learning, we recommend to first test less complex models, such as traditional machine learning models, since these often already deliver a sufficient accuracy, while being more efficient, reliable and comprehensive. To improve performance of traditional models, gradient boosting can be applied. Gradient boosting involves a set of weak models which just have to outperform random guessing. The models are trained in a iterative, stage-wise manner to fit the (pseudo) residuals of the model's prediction (Friedman, 2001). Another approach would be to make use of existing deep neural networks and apply transfer learning (Johnson et al., 2020). Transfer learning could be especially important when models should be highly personal, for example for top-level athletes or athletes with a disability. To provide good starting points for such models, data and model sharing is fundamental to advance research in "in the wild" human movement analysis.

When tuning a model, it is important to ensure that the model is not under- or overfitted to the data. Under- and overfitting is also known as the bias-variance trade-off. Underfitting occurs when the machine learning model is unable to represent the training data and the training error is high, while overfitting occurs when the model represents the training data too well (e.g., modeling random noise in the data) and the gap between training error and validation error, which is also called the generalization error, is high. Deep neural networks are especially prone to overfitting, since these models have many tuneable parameters. Overfitting can be avoided by increasing the size of the training dataset, however this is very challenging for datasets of “in the wild” movements. To ensure that the model does not overfit, while also avoiding underfitting, we recommend to choose the model with the smallest training and validation error, while the gap between the validation error and training error should be small and decrease over time during training.

Different techniques, such as regularization schemes (Goodfellow, Bengio, & Courville, 2016), data augmentation, and including prior knowledge into models, can be applied when models are prone to overfitting. Explicit regularization schemes either aim to remove nodes (and thus learning parameters) from the network (dropout) or add a function to the loss function to minimize network weights, thereby forcing the weights to zero, effectively removing the node as well. Implicit regularization schemes, such as stochastic gradient descent and data augmentation, have shown to be effective. On the other hand, data augmentation and including prior knowledge do not only reduce the change of overfitting, but also improve model generalizability, meaning that the model’s ability to estimate outputs for unseen data is improved. Data augmentation is a technique that increases the size of the training dataset with the aim to make this dataset more representative. For example, additional synthetic training data can be generated by perturbing the input data. Artificially inducing measurement errors or virtually rotating the sensors can ensure that the model is insensitive to noise, calibration errors, and inaccurate sensor placement (Weber, Gühmann, & Seel, 2021). Another option is to generate new movement patterns by simulating biomechanical models (Dorschky et al., 2020; Mundt et al., 2020; Sharifi Renani, Eustace, Myers, & Clary, 2021). Dorschky et al. (2020) showed that the accuracy of gait kinematics can be improved by up to 27% when adding simulated IMU data to the training dataset. However, simulated data can only improve generalization if the discrepancy between simulated and measured data (reality gap) is small. Including prior knowledge in the model design can also improve model generalizability. To achieve this, hybrid models, also referred to as physics-informed machine learning, are being increasingly investigated (Karniadakis et al., 2021). For example, hybrid models include physics-guided neural network architectures, and physical constraints in the cost function (Willard, Jia, Xu, Steinbach, & Kumar, 2020; Karniadakis et al., 2021).

5. Further considerations for specific application domains

This section presents application considerations for two different example domains where “in the wild” movement data recording was combined with machine learning to answer new questions that could not be answered otherwise. First, Section 5.1 presents considerations regarding approaches related to long-term monitoring, specifically to estimate tibial loading as a means to prevent injuries due to fatigue. Then Section 5.2 presents considerations for providing athletes with actionable insights related to their technique based on data recorded during training. We also exemplify how to increase the relevance of the data to the athlete, providing them with actionable insights they can actually benefit from.

5.1. Application: injury prevention

Neuromuscular fatigue is a main risk factor for running related injuries, suggesting that the body in a fatigued state is less able to attenuate impact forces sufficiently while landing on the ground. Inertial sensor technology, combined with force sensing and modeling techniques, provides a base to understand changes in kinetic and kinematic parameters and the shock attenuation strategies the body uses to counteract the negative effects of fatigue during running. These methods can be used outside the lab to assess shock attenuation strategies in a real-world setting (Reenalda, Maartens, Homan, & Buurke, 2016; Wouda et al., 2018). The load on the knee can be estimated using inertial sensors data combined with a neural network (Stetter, Krafft, Ringhof, Stein, & Sell, 2020). The role of the knee in attenuating the shock due to the foot hitting the ground can be modeled using wavelet transform analysis. Both show a change in knee loading and shock attenuation from active to passive strategies. This shift in strategy was also observed in fatiguing experiments on the athletic track as well as during an actual marathon. Shock attenuation derived from accelerations at the tibia, sacrum and sternum showed that runners adopt a more passive strategy to attenuate the shock when fatigued, which might result in a higher risk of injuries (Reenalda et al., 2019). This passive mechanism means that instead of using muscle contractions to actively control the ankle, knee and hip as shock attenuators, the body relies more on elastic and bony structures to dissipate the impulse. Therefore, providing feedback, either directly or indirectly via a trainer or coach, about the shock attenuation strategy is important to prevent injury, since the runner could then attempt to improve their strategy, even though adjustments are challenging for an athlete who is already fatigued.

The load on the tibia can be modeled using kinetic and kinematic variables. This has already been demonstrated using optical motion capture systems (Matijevich, Branscombe, Scott, & Zelik, 2019) but inertial sensor technology gives the opportunity to estimate tibial bone loading due to fatigue “in the wild”. To identify this fatigue, machine learning approaches showed that statistical features of the data provide additional information over traditional biomechanics. It allows for sensor reduction to unobtrusively and accurately classify fatigue in runners (Marotta, Buurke, van Beijnum, & Reenalda, 2021).

Novel approaches in inertial sensor technology together with new data analysis and modeling techniques provide valuable insights into the load on the body and the strategies used to cope with this load. Identified parameters and mechanisms can eventually be used to directly or indirectly provide feedback to runners to assist injury prevention without compromising performance.

5.2. Application: technique analysis in sports

In many sports, as athletes acquire and improve their skill, they converge to specific, individual movement patterns, often called the individual technique of the athlete. The individual technique characteristics can be a decisive factor for success in sports (Lees, 2002). Consequently, there is a high demand for methods facilitating comparison of techniques between athletes or groups of athletes. Instructors/coaches who are not only able to assess individual athletes' techniques, but can also advise on exercises for improving technique, are sought after (Faber, Koopmann, Büsch, & Schorer, 2021). However, in many sports this is quite an advanced skill and even the technique assessments and given recommendations of the best coaches remain highly subjective. With recent developments in sensor and app technologies, it has become a vision that technique analysis could be automatized and machine learning approaches together with biomechanical/motor control modeling may lead to devices and apps that support athletes and coaches through providing more objective analysis for more evidence-based training recommendations.

While in recent years research and technological developments have made good progress towards realizing the outlined vision, there are technical and conceptual challenges that still have to be addressed. One such challenge arises from the fact that human movement is the behavior of a complex system and an individual movement pattern might be seen as an emergent feature of this system (Federolf et al., 2021). An emergent feature of a complex system is a property that cannot be predicted from only the properties of the components of this system. For quantifying and comparing technique between individuals, this suggests that it may be insufficient to rely on only a few variables or sensors (Federolf et al., 2021). In a best case, a representation of the full body, i.e., of the movements of all main body segments, should be attempted; whereas relying on single-sensor technologies is inherently limited. Single-sensor systems may allow a classification of technique styles, but not a true comparison of individual techniques between different athletes.

Different wearable sensor systems allow for full-body movement representation (Fig. 3). When using such a system, the next question is, which variables should be used for quantifying the movements. For example, extracting joint angles from the measurement system as variables has the advantage that these angles can be compared between athletes with no need for further data normalization. However, communicating outcomes back to an athlete or coach can be ineffective if they involve only relatively minute differences in joint angles. A better approach is to visualize technique differences using, e.g., an avatar or stick-figure movements (Gløersen, Myklebust, Hallén, & Federolf, 2018). Such visualizations are easier when body segment and joint positions rather than joint angles are used as variables. However, the disadvantage of this choice is that comparison of athletes of different size and stature then requires additional normalization procedures (Federolf et al., 2014; Gløersen et al., 2018).

Other challenges when using full-body motion capture systems arise from the tasks of retaining an overview of the multitude of variables and data points and of systematizing the extraction of technique features of interest. One example of a suitable methodology to address these tasks is the application of principal component analysis (PCA). Since body segment movements are highly correlated within human movements, a relatively low number of principal components covers, even for complex movement sequences, a large proportion of the variance in the data (Daffertshofer, Lamoth, Meijer, & Beek, 2004). Projecting full-body movement data onto the resultant principal component vectors provides a coordinate system tailored to the movement of interest, in which techniques or other movement features can be compared efficiently (Federolf et al., 2014; Witte et al., 2010). In numerous sports, PCA has already been applied successfully to study technique features or to compare techniques between groups of athletes, for example, in running (Mohr, Pieper, Löffler, Schmidt, & Federolf, 2021), walking (Malloggi et al., 2021), diving (Young & Reinkensmeyer, 2014), gymnastics (Busquets, Ferrer-Uris, Angulo-Barroso, & Federolf, 2021), juggling (Zago et al., 2017), alpine (Federolf et al., 2014) or cross-country skiing (Gløersen et al., 2018; Pellegrini, Zoppirolli, Boccia, Bortolan, & Schena, 2018). These solutions can easily be applied for "in the wild" measurements and many would be suitable for real-time variable extraction and feedback systems.

6. Conclusion

In this perspective paper, we have discussed different opportunities and challenges regarding analysis of "in the wild" movement

Fig. 3. Left: skier equipped with a full-body wearable sensor system (Xsens™) for a technique analysis study. Right: the avatar created for this movement by wearing a full-body Xsens™-software. Note: The two images represent the same time point of the skiing motion, however, the observer perspective differs slightly.

data using machine learning. We have highlighted the importance of careful data recordings. This is important when recording movement in general, to collect the correct data at the correct point in time, but becomes even more important when machine learning is used for processing, to ensure that the trained model works for the data it should be analyzing, and that the correct analysis is performed. We have discussed several considerations regarding both data recording in general and specific to training of machine learning models, to ensure that in future researchers can properly address these in experiments. Furthermore, we highlighted how "in the wild" movement data can be used in two application domains, specifically monitoring injury risk and technique analysis.

By combining machine learning with "in the wild" recordings, the main advantages are the potential for real-time feedback, while analyses can also be performed on large datasets. The real-time potential can be used to enhance the performance of athletes during training and competition. Its advantages for large datasets can be exploited when performing long-term monitoring to forecast injuries, especially by developing personalized models, while statistical features could be found that provide more insight than traditional parameters as well. The combination of real-time feedback with analysis on large datasets allows for computationally efficient calculation of many variables, e.g. joint angles of a kinematic model, which can then be used to create avatars of athlete's movements to provide actionable insights. In conclusion, this perspective paper provides researchers with guidance and directions for future development of machine learning models for movement analysis "in the wild".

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

AK and ED were partly supported by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) under Grant SFB 1483 – Project-ID 442419336. AK was also supported by a faculty endowment by adidas AG. JD was partially supported by the interuniversity BOF fund (IBOF/21/075) and the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" program.

References

Araújo, D., Couceiro, M., Seifert, L., Sarmento, H., & Davids, K. (2021). *Artificial Intelligence in Sport Performance Analysis*. Routledge.

Aroganam, G., Manivannan, N., & Harrison, D. (2019). Review on wearable technology sensors used in consumer sport applications. *Sensors, 19*(9), 1983.

Barton, C. J., Bonanno, D., Carr, J., Neal, B., Malliaras, P., Franklyn-Miller, A., et al. (2016). Running retraining to treat lower limb injuries: a mixed-methods study of current evidence synthesised with expert opinion. *British Journal of Sports Medicine, 50*(9), 513–526.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag.

Busquets, A., Ferrer-Uris, B., Angulo-Barroso, R., & Federolf, P. (2021). Gymnastics experience enhances the development of bipedal-stance multi-segmental coordination and control during proprioceptive reweighting. *Frontiers in Psychology, 12*.

Camomilla, V., Bergamini, E., Fantozzi, S., & Vannozzi, G. (2018). Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: A systematic review. *Sensors, 18*(3), 873.

Christ, M., Braun, N., Neuffer, J., & Kempa-Liehr, A. W. (2018). Time series feature extraction on basis of scalable hypothesis tests (tsfresh–a python package). *Neurocomputing, 307*, 72–77.

Claudino, J. G., de Oliveira Capanema, D., de Souza, T. V., Serrão, J. C., Pereira, A. C. M., & Nassis, G. P. (2019). Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: a systematic review. *Sports Medicine-Open, 5*(1), 1–12.

Clermont, C. A., Benson, L. C., Edwards, W. B., Hettinga, B. A., & Ferber, R. (2019). New considerations for wearable technology data: Changes in running biomechanics during a marathon. *Journal of Applied Biomechanics, 35*(6), 401–409.

Collins, N. J., Barton, C. J., Van Middelkoop, M., Callaghan, M. J., Rathleff, M. S., Vicenzino, B. T., et al. (2018). 2018 consensus statement on exercise therapy and physical interventions (orthoses, taping and manual therapy) to treat patellofemoral pain: recommendations from the 5th international patellofemoral pain research retreat, gold coast, australia, 2017. *British Journal of Sports Medicine, 52*(18), 1170–1178.

Cust, E. E., Sweeting, A. J., Ball, K., & Robertson, S. (2019). Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance. *Journal of Sports Sciences, 37*(5), 568–600.

Daffertshofer, A., Lamoth, C. J., Meijer, O. G., & Beek, P. J. (2004). Pca in studying coordination and variability: A tutorial. *Clinical Biomechanics, 19*(4), 415–428.

De Brabandere, A., Emmerzaal, J., Timmermans, A., Jonkers, I., Vanwanseele, B., & Davis, J. (2020). A machine learning approach to estimate hip and knee joint loading using a mobile phone-embedded imu. *Frontiers in Bioengineering and Biotechnology, 8*, 320.

De Brabandere, A., Op De Beéck, T., Hendrickx, K., Meert, W., & Davis, J. (2021). Tsfuse: Automated feature construction for multiple time series data. *Machine Learning*.

De Vree, L., & Carloni, R. (2021). Deep reinforcement learning for physics-based musculoskeletal simulations of healthy subjects and transfemoral prostheses' users during normal walking. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 29*, 607–618.

Decroos, T., Schütte, K., De Beéck, T. O., Vanwanseele, B., & Davis, J. (2018). Amie: Automatic monitoring of indoor exercises. In *Proceedings 2018 Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 424–439).

Dorschky, E., Nitschke, M., Martindale, C. F., van den Bogert, A. J., Koelewijn, A. D., & Eskofier, B. M. (2020). CNN-based estimation of sagittal plane walking and running biomechanics from measured and simulated inertial sensor data. *Frontiers in Bioengineering and Biotechnology, 8*. https://doi.org/10.3389/fbioe.2020.00604

Düking, P., Hotho, A., Holmberg, H.-C., Fuss, F. K., & Sperlich, B. (2016). Comparison of non-invasive individual monitoring of the training and health of athletes with commercially available wearable technologies. *Frontiers in Physiology, 7*, 71.

Faber, I. R., Koopmann, T., Büsch, D., & Schorer, J. (2021). Developing a tool to assess technical skills in talented youth table tennis players—A multi-method approach combining professional and scientific literature and coaches' perspectives. *Sports Medicine-Open, 7*(1), 1–24.

Federolf, P., Angulo-Barroso, R. M., Busquets, A., Ferrer-Uris, B., Gløersen, Ø., Mohr, M., et al. (2021). Letter to the editor regarding the assessment of center of mass and center of pressure during quiet stance: Current applications and future directions. *Journal of Biomechanics, 128*, Article 110729.

Federolf, P., Reid, R., Gilgien, M., Haugen, P., & Smith, G. (2014). The application of principal component analysis to quantify technique in sports. *Scandinavian Journal of Medicine & Science in Sports, 24*(3), 491–499.

Ferber, R., Osis, S. T., Hicks, J. L., & Delp, S. L. (2016). Gait biomechanics in the era of data science. *Journal of Biomechanics, 49*(16), 3759–3761.

Fonseca, S. T., Souza, T. R., Verhagen, E., Van Emmerik, R., Bittencourt, N. F., & Mendonça, L. D. (2020). Sports injury forecasting and complexity: A synergetic approach. *Sports Medicine*, 1–14.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 1189–1232.

Fulcher, B. D., & Jones, N. S. (2017). hctsa: A computational framework for automated time-series phenotyping using massive feature extraction. *Cell Systems, 5*(5), 527–531.

Giles, B., Kovalchik, S., & Reid, M. (2020). A machine learning approach for automatic detection and classification of changes of direction from player tracking data in professional tennis. *Journal of Sports Sciences, 38*(1), 106–113.

Gløersen, Ø., Myklebust, H., Hallén, J., & Federolf, P. (2018). Technique analysis in elite athletes using principal component analysis. *Journal of Sports Sciences, 36*(2), 229–237.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. http://www.deeplearningbook.org.

Gurchiek, R. D., Cheney, N., & McGinnis, R. S. (2019). Estimating biomechanical time-series with wearable sensors: A systematic review of machine learning techniques. *Sensors, 19*(23), 5227.

Hernandez, V., Dadkhah, D., Babakeshizadeh, V., & Kulić, D. (2021). Lower body kinematics estimation from wearable sensors for walking and running: A deep learning approach. *Gait and Posture, 83*, 185–193.

Hughes, G. T., Camomilla, V., Vanwanseele, B., Harrison, A. J., Fong, D. T., & Bradshaw, E. J. (2021). Novel technology in sports biomechanics: Some words of caution. *Sports Biomechanics*, 1–9.

Jaspers, A., Op De Beéck, T., Brink, M. S., Frencken, W. G., Staes, F., Davis, J., et al. (2018). Relationships between the external and internal training load in professional soccer: What can we learn from machine learning? *International Journal of Sports Physiology and Performance, 13*(5), 625–630.

Johnson, W. R., Mian, A., Robinson, M. A., Verheul, J., Lloyd, D. G., & Alderson, J. A. (2020). Multidimensional ground reaction forces and moments from wearable sensor accelerations via deep learning. *IEEE Transactions on Biomedical Engineering, 68*(1), 289–297.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics, 3*(6), 422–440.

Klucken, J., Barth, J., Kugler, P., Schlachetzki, J., Henze, T., Marxreiter, F., et al. (2013). Unbiased and mobile gait analysis detects motor impairment in parkinson's disease. *PloS One, 8*(2), Article e56956.

Kok, M., Hol, J. D., Schön, T. B. (2017). Using inertial sensors for position and orientation estimation. arXiv preprint arXiv:1704.06053.

Kremer, J., Steenstrup Pedersen, K., & Igel, C. (2014). Active learning with support vector machines. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 4*(4), 313–326.

Lees, A. (2002). Technique analysis in sports: A critical review. *Journal of Sports Sciences, 20*(10), 813–828.

Macadam, P., Cronin, J., Neville, J., & Diewald, S. (2019). Quantification of the validity and reliability of sprint performance metrics computed using inertial sensors: A systematic review. *Gait & posture, 73*, 26–38.

Malloggi, C., Zago, M., Galli, M., Sforza, C., Scarano, S., & Tesio, L. (2021). Kinematic patterns during walking in children: Application of principal component analysis. *Human Movement Science, 80*, Article 102892.

Marotta, L., Buurke, J. H., van Beijnum, B.-J. F., & Reenalda, J. (2021). Towards machine learning-based detection of running-induced fatigue in real-world scenarios: Evaluation of imu sensor configurations to reduce intrusiveness. *Sensors, 21*(10), 3451.

Matijevich, E. S., Branscombe, L. M., Scott, L. R., & Zelik, K. E. (2019). Ground reaction force metrics are not strongly correlated with tibial bone load when running across speeds and slopes: Implications for science, sport and wearable tech. *PloS one, 14*(1), Article e0210000.

Milner, C. E., Hawkins, J. L., & Aubol, K. G. (2020). Tibial acceleration during running is higher in field testing than indoor testing. *Medicine and Science in Sports and Exercise, 52*(6), 1361–1366.

Mohr, M., Pieper, R., Löffler, S., Schmidt, A. R., & Federolf, P. A. (2021). Sex-specific hip movement is correlated with pelvis and upper body rotation during running. *Frontiers in Bioengineering and Biotechnology, 9*, 521.

Mundt, M., Koeppe, A., David, S., Witter, T., Bamer, F., Potthast, W., et al. (2020). Estimation of Gait Mechanics Based on Simulated and Measured IMU Data Using an Artificial Neural Network. *Frontiers in Bioengineering and Biotechnology, 8*(February), 1–16.

Op De Beéck, T., Meert, W., Schütte, K., Vanwanseele, B., & Davis, J. (2018). Fatigue prediction in outdoor runners via machine learning and sensor fusion. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 606–615).

Pellegrini, B., Zoppirolli, C., Boccia, G., Bortolan, L., & Schena, F. (2018). Cross-country skiing movement factorization to explore relationships between skiing economy and athletes' skills. *Scandinavian Journal of Medicine & Science in Sports, 28*(2), 565–574.

Preatoni, E., Bergamini, E., Fantozzi, S., Giraud, L. I., Orejel Bustos, A. S., Vannozzi, G., et al. (2022). The use of wearable sensors for preventing, assessing, and informing recovery from sport-related musculoskeletal injuries: A systematic scoping review. *Sensors, 22*(9), 3225.

Reenalda, J., Maartens, E., Buurke, J. H., & Gruber, A. H. (2019). Kinematics and shock attenuation during a prolonged run on the athletic track as measured with inertial magnetic measurement units. *Gait & posture, 68*, 155–160.

Reenalda, J., Maartens, E., Homan, L., & Buurke, J. J. (2016). Continuous three dimensional analysis of running mechanics during a marathon by means of inertial magnetic measurement units to objectify changes in running mechanics. *Journal of Biomechanics, 49*(14), 3362–3367.

Rindal, O. M. H., Seeberg, T. M., Tjønnås, J., Haugnes, P., & Sandbakk, Ø. (2017). Automatic classification of sub-techniques in classical cross-country skiing using a machine learning algorithm on micro-sensor data. *Sensors, 18*(1), 75.

Robberechts, P., Derie, R., Van den Berghe, P., Gerlo, J., De Clercq, D., Segers, V., et al. (2021). Predicting gait events from tibial acceleration in rearfoot running: A structured machine learning approach. *Gait & Posture, 84*, 87–92.

Rossi, A., Pappalardo, L., Cintia, P., Iaia, F. M., Fernández, J., & Medina, D. (2018). Effective injury forecasting in soccer with gps training data and machine learning. *PloS One, 13*(7), Article e0201264.

Rum, L., Sten, O., Vendrame, E., Belluscio, V., Camomilla, V., Vannozzi, G., et al. (2021). Wearable sensors in sports for persons with disability: A systematic review. *Sensors, 21*(5), 1858.

Sambu Seo, Wallat, M., Graepel, T., & Obermayer, K. (2000). Gaussian process regression: Active data selection and test point rejection. In , *Neural Computing: New Challenges and Perspectives for the New Millennium: Vol. 3. Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000* (pp. 241–246). IEEE. URL:http://ieeexplore.ieee.org/document/861310/.

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green ai. *Communications of the ACM, 63*(12), 54–63.

Seel, T., Kok, M., & McGinnis, R. S. (2020). Inertial sensors—Applications and challenges in a nutshell. *Sensors, 20*(21), 6221.

Sharifi Renani, M., Eustace, A. M., Myers, C. A., & Clary, C. W. (2021). The use of synthetic imu signals in the training of deep learning models significantly improves the accuracy of joint kinematic predictions. *Sensors, 21*(17), 5876.

Sheerin, K. R., Reid, D., & Besier, T. F. (2019). The measurement of tibial acceleration in runners—A review of the factors that can affect tibial acceleration during running and evidence-based guidelines for its use. *Gait & Posture, 67*, 12–24.

Shultz, S. J., Schmitz, R. J., Cameron, K. L., Ford, K. R., Grooms, D. R., Lepley, L. K., et al. (2019). Anterior cruciate ligament research retreat viii summary statement: an update on injury risk identification and prevention across the anterior cruciate ligament injury continuum, march 14–16, 2019, greensboro, nc. *Journal of Athletic Training, 54*(9), 970–984.

Singh, A., Le, B. T., Nguyen, T. L., Whelan, D., O'Reilly, M., Caulfield, B., et al. (2021). Interpretable classification of human exercise videos through pose estimation and multivariate time series analysis. In *International Workshop on Health Intelligence* (pp. 181–199). Springer.

Stanev, D., Filip, K., Bitzas, D., Zouras, S., Giarmatzis, G., Tsaopoulos, D., et al. (2021). Real-time musculoskeletal kinematics and dynamics analysis using marker-and imu-based solutions in rehabilitation. *Sensors, 21*(5), 1804.

Stetter, B. J., Krafft, F. C., Ringhof, S., Stein, T., & Sell, S. (2020). A Machine Learning and Wearable Sensor Based Approach to Estimate External Knee Flexion and Adduction Moments During Various Locomotion Tasks. *Frontiers in Bioengineering and Biotechnology, 8*(January).

Weber, D., Gühmann, C., & Seel, T. (2021). RIANN—A Robust Neural Network Outperforms Attitude Estimation Filters. *Ai, 2*(3), 444–463.

Whelan, D., O'Reilly, M., Huang, B., Giggins, O., Kechadi, T., & Caulfield, B. (2016). Leveraging imu data for accurate exercise performance classification and musculoskeletal injury risk screening. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 659–662). IEEE.

Willard, J. D., Jia, X., Xu, S., Steinbach, M., & Kumar, V. (2020). Integrating physics-based modeling with machine learning: A survey. *arXiv, 1*(1), 1–34.

Willy, R. W. (2018). Innovations and pitfalls in the use of wearable devices in the prevention and rehabilitation of running related injuries. *Physical Therapy in Sport, 29*, 26–33.

Witte, K., Ganter, N., Baumgart, C., & Peham, C. (2010). Applying a principal component analysis to movement coordination in sport. *Mathematical and Computer Modelling of Dynamical Systems, 16*(5), 477–488.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation, 1*(1), 67–82.

Wouda, F. J., Giuberti, M., Bellusci, G., Maartens, E., Reenalda, J., Van Beijnum, B.-J. F., et al. (2018). Estimation of vertical ground reaction forces and sagittal knee kinematics during running using three inertial sensors. *Frontiers in Physiology, 9*, 218.

Young, C., & Reinkensmeyer, D. J. (2014). Judging complex movement performances for excellence: A principal components analysis-based technique applied to competitive diving. *Human Movement Science, 36*, 107–122.

Zago, M., Pacifici, I., Lovecchio, N., Galli, M., Federolf, P. A., & Sforza, C. (2017). Multi-segmental movement patterns reflect juggling complexity and skill level. *Human Movement Science, 54*, 144–153.