A Preliminary Experimental Outline to Train Machine Learning Models for the Unobtrusive, Real-Time Detection of Acute Physiological Stress Levels during Training Exercises

André Jeworutzki
University of the West of Scotland
Hamburg University of Applied Sciences
Hamburg, Germany
andre.jeworutzki@haw-hamburg.de

Jan Schwarzer
Hamburg University of Applied Sciences
Hamburg, Germany
jan.schwarzer@haw-hamburg.de

Kai von Luck
Hamburg University of Applied Sciences
Hamburg, Germany
kai.vonluck@haw-hamburg.de

Susanne Draheim
Hamburg University of Applied Sciences
Hamburg, Germany
susanne.draheim@haw-hamburg.de

Qi Wang
University of the West of Scotland
Paisley, Scotland
qi.wang@uws.ac.uk

ABSTRACT

The automatic recognition of activities can assist people in keeping track of their health and in avoiding injuries. Nowadays, inertial measurement units have gained notable interest for such tasks due to being low-cost, small-sized, and easy-of-use. Inertial sensor technology in combination with physiological data allows to state holistic conclusions regarding, for example, an activity's quality. This research draws attention to the case of stress levels in sports, where researchers typically rely on obtrusive stress markers analyzed in laboratories (e.g., lactate and cortisol). While there are known stimuli for stress such as fatigue, existing knowledge is limited concerning methodological means and measurement standards for unobtrusively detecting stress in challenging contexts such as sports. In response, this work reports from our ongoing research, where we aim to develop the necessary means to unobtrusively detect stress levels in real-time based on machine learning algorithms. The main contribution of the present paper is a preliminary experimental outline. It illustrates the steps we intend to take to methodologically guide the data collection procedures and to train machine learning models towards this goal. In doing so, we hope contributing helpful insights to aid other researchers in designing stress-related studies in the sports context.

CCS CONCEPTS

• **Human-centered computing → Laboratory experiments**.

PETRA 2021, June 29-July 2, 2021, Corfu, Greece

ACM ISBN 978-1-4503-8792-7/21/06…$15.00
https://doi.org/10.1145/3453892.3461833

KEYWORDS

activity recognition, stress determination, physiological sensors, inertial measurement units, supervised machine

ACM Reference Format:
André Jeworutzki, Jan Schwarzer, Kai von Luck, Susanne Draheim, and Qi Wang. 2021. A Preliminary Experimental Outline to Train Machine Learning Models for the Unobtrusive, Real-Time Detection of Acute Physiological Stress Levels during Training Exercises. In *The 14th PErvasive Technologies Related to Assistive Environments Conference (PETRA 2021), June 29-July 2, 2021, Corfu, Greece.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3453892.3461833

1 INTRODUCTION

Based on information about user behavior, activity recognition enables computer systems to help users with their tasks [1]. The automatic recognition of physical activity has historically been of notable interest in computer vision research, while efforts to recognize activities beyond instrumented rooms led to a shift towards body-worn inertial measurements units (IMU) [12]. IMUs are now one of the dominant technical aids to assist in home-based exercise therapy [33], while being low-cost, small-sized, and easy-of-use devices [11]. Due to their characteristic of retrieving motion-related data such as acceleration and angular velocity, IMUs play an important role in determining human motion [47]. Fundamentally, IMU data allows to assess the execution quality of a training exercise in terms of technique and accuracy [20]. With the term quality, we lean towards a definition from the International Organization for Standardization, which understands quality "as the degree to which a set of inherent characteristic fulfills requirements" [35]. Feedback in terms of quality is an interesting field of research in sports and healthcare [48]. However, research indicates limitations of IMU sensors in terms of, for instance, detecting deviations from the ideal movement (e.g., improper timing of muscular activation) hence studies additionally incorporate biofeedback systems such as electromyographic (EMG) sensors to overcome such limitations [47].

The ability to simultaneously detect physical activity and stress levels can assist users in keeping track of their health, while IMU data in addition to physiological data builds a crucial foundation for this task [60]. Stress is considered to have a bidirectional relationship with injuries in activities such as training exercises. However, research is still controversial [56, 61]; as originally hypothesized by [58], stress increases muscle tension that can lead to a motor coordination disturbance as well as a reduction in flexibility and an increase in fatigue. A monitoring system that determines stress and fatigue during exercises could aid future research in examining this relationship. Since fitness-oriented exercises at home are often poorly or ineffectively performed [55], a monitoring system could also serve as an early warning system [28]. So far, there is a lack of studies investigating stress in challenging contexts such as training exercises [31, 43]. More research on methodological and measurement standards is needed [6]. Since there is no single stress marker that globally assesses an individual's stress response [6], we apply a multivariable approach that incorporates workload, lactate, and other unobtrusive biosignals such as heart rate to determine stress levels.

To the best of our knowledge, no existing study attempted to unobtrusively measure stress in the challenging context of training exercises. In our ongoing research, we examine methodological approaches and potential biosignal candidates. The present paper contributes an experimental outline that utilizes a supervised machine learning process to build a model that can determine stress levels without the need of invasive stress markers such as lactate and cortisol. Our research is part of a larger European funded research project. This project includes industry as well as university partners and targets the development of a product for the fitness market aiding people in their home-based training.

This paper is organized as follows: In Section 2, the relevant related literature is presented. Then, the research context is introduced in Section 3. Subsequently, Section 4 elaborates on the experimental setting, which is the foundation for our research. Section 5 then discusses our chosen approach, highlights research implications, and indicates research limitations. Finally, Section 6 concludes the present paper and provides recommendations for future research.

2 RELATED WORKS

Our literature review builds on the circumstance that physiological signals alone are not sufficient to determine stress levels and the use of IMU data to detect physical activities is encouraged to enrich investigations [60]. Therefore, the related work is elaborated on from two perspectives: firstly, it is concentrated on how studies approach the challenge of detecting human motion during training exercises; and, secondly, it is focused on the challenge of determining stress levels during training exercises, including fatigue as a stimulus for stress.

2.1 Detecting Human Activities Throughout Training Exercises

Traditionally, computer vision approaches have been at the forefront of recognizing human activities, while there is a shift towards body-worn inertial sensors observable (e.g., due to their capability to detect activities beyond instrumented rooms [12]). A commonly used general-purpose framework to design and evaluate activity recognition systems is the so-called *Activity Recognition Chain* (ARC) [48]. The ARC was introduced in [12] and prescribes different steps that transcend raw sensor data to classified pieces of information. In principle, any type of multimodal sensor data could be used with the ARC; which sensors are suitable depends on the application context [6, 41].

The initial step in the ARC is the task of preprocessing the incoming data to smooth and prepare the signal for subsequent processing and analysis. To this end, related studies leveraged different filters such as the Butterworth filter [9] and a moving average filter [47]. Such filters make it easier to divide the incoming signal into individual segments (e.g., where an activity begins and ends), since noise is largely filtered out. In the chain's next step, the filtered data is segmented into parts, each representing an individual repetition. However, the literature emphasizes that segmentation of time series data is difficult [11, 42]. After successfully segmenting the filtered stream of data, features are subsequently extracted on this basis. Features reduce the data to information that is discriminative for the corresponding activity that is performed [12]. Finally, there is the classification step in the chain. Here, feature vectors result in labeled decisions – e.g., an activity is labeled as correct or incorrect.

The different steps of the ARC are more thoroughly discussed in Section 4, while we focus on providing a brief introduction to ARC in this context. It is worth highlighting that there are a variety of different solutions for each step in this chain. For example, the task of data segmentation can be solved by supervised or unsupervised machine learning approaches. Furthermore, algorithms exist that allow for light-weight online segmentation and algorithms for more heavy-weight offline segmentation. Lin et al. discuss these various approaches of data segmentation in more detail [42]. Similarly, the definition of appropriate features and their calculation vary notably in related studies. Whereas some studies rely on statistical features (e.g., mean, median, and variance) [47], other leverage dynamics features (e.g., energy and energy ratio) as well [9]. Finally, there is a notable number of classifiers to be utilized for the classification process. To name a few, support-vector machines (SVM), decision trees, and k-nearest neighbor (KNN) algorithms are three of the potential candidates that a researcher can choose from [12].

2.2 Detecting Stress in Training Exercises

Attention is now drawn to the field of stress in the context of sports, its relation to fatigue, and how studies have attempted to measure it during training exercises.

There is a long debate across multidisciplinary fields about the concept of stress [15]. Since each discipline has its own concepts on stress, a common definition is unlikely [6, 17]. Stress can be classified as acute or chronic [15, 29]. While chronic stress is pathological and psychological in nature, acute stress is the immediate response of the body to a stimulus (stressor) [29, 49]. The acute response triggers alertness, energy release, physiological regulation, and immunological activation to compensate for the effects of the stressor [29]. During training exercises, the body experiences an acute stress response in which more oxygen and energy are required. The heart rate increases so that more blood is pumped through the body and thus oxygen is transported to improve cardiorespiratory function [7]. Stress could be understood as a response to a disturbance of homeostatic balance by events or conditions (stressors) [49]. For example, untrained people suffer from more stress due to higher demand for oxygen and energy, while trained people become accustomed to use less oxygen; their body will eventually feel the stress over a longer period [7, 14]. The physiological reactions are summarized as follows:

- Sympathoadrenal system (SAM axis): Sympathetic activation and parasympathetic withdrawal cause increased heart rate and respiratory rate, bronchial and pupil dilation, sweaty skin, and other symptoms. The body is rapidly prepared for a physical “fight or flight” stress response [7].
- Hypothalamic-pituitary-adrenal axis (HPA axis): Slowly activated by the secretion of cortisol leading to increased catabolism, anabolism inhibition, and depression of the immune system. Typically activated by mental tasks [22].

In addition, stress is highly subjective and individual in all aspects [19, 21]. There is a lack of research on methodological and measurement standards to determine stress during challenging contexts such as training exercises [2, 6, 19, 41, 50], for which stress is a natural physiological response [4, 7, 26, 53]. In principle, there are countless stimuli that are associated with stress [21, 46]. One of them is the performed quality of a training exercise [3, 31]. Fatigue is another stimulus for stress [38, 40]. Physical activity could be viewed as providing stimuli that promote specific and varied adaptations of the body depending on the type, intensity, and duration of exercise performed [14, 26, 27]. Chronic exercise training does not eliminate the acute exercise response, but it can attenuate the overall effect of the response as the body adapts to the training stimulus in a positive way. An excessive intensity and/or volume of training may lead to maladaptation [26, 27]. Hence, a stress response is dependent on the athlete and the exercise. An unfamiliar exercise is likely to elicit a higher metabolic stress response than a familiar, routine exercise, e.g., a long-distance runner will probably have a different stress response profile for a given exercise than a weightlifter. Exercises represent an effective methodological tool to study the body’s response to metabolic stress, and from a clinical perspective, offers an alternative treatment choice to drug intervention strategies [7].

Thus far, only a few studies have attempted to investigate physiological stress during training exercises. For example, Magiera et al. focused on the effect of physical and mental stress on the heart rate as well as cortisol and lactate concentrations [43]. They found that the heart rate is most sensitive to physical and mental stress. Hong et al. investigated the influence of physical activity on stress recognition with physiological responses [31]. The authors used different stressors to induce stress and found that, among others, stress models for each physical activity should be built due to variations in physiological changes caused by physical activity. Alamudun et al. introduced two multivariate signal processing algorithms to cope with the differences in physiology between participants and changes in physical activity [3]. They found that these two algorithms can bring noticeable improvements for the process of stress prediction. Wong et al. used IMU data to distinguish stress and high intensity activity in daily life [60].

Based on our literature work, we created a tabular overview of existing stress markers (see Table **??**). Stress (and fatigue) markers can be classified as subjective or objective depending on the measurement technique [24]. Subjective stress markers, on the one hand, are traditionally used by psychologists in the form of questionnaires, interviews, or self-reports, which are usually conducted retrospectively. Subjective markers are not suitable to continuously monitor stress during training exercises but can be used to determine stress levels before and after an exercise. Objective stress markers, on the other hand, are quantifiable and cover physiological, physical, behavioral responses, and other contextual data. They can reduce the possibility of self-deception, falsification, fabrication, attention, or recall bias, which is usually present in subjective markers [51]. Objective markers are measured either obtrusively or unobtrusively [6]. Biomedical researchers rely on obtrusive biochemical markers, typically hormones, to measure stress [6, 53]. One of these hormones is cortisol, which is commonly used in studies on stress [19]. Another less expensive marker is lactate [13, 30] which was once incorrectly attributed to muscle fatigue [13]. Such obtrusive biochemical markers provide accurate quantitative data [6]. However, they are not suitable for real-time monitoring systems due to their inherent nature and that they, at times, necessitate analyzing data in a laboratory. Unobtrusive stress markers, such as heart rate or muscle activity, are measured by sensors that are attached to the body. They provide continuous data in real-time and do not require analysis in a laboratory [19]. Yet, unobtrusive markers are susceptible to noise or artifacts due to individual’s body parts movements or activities [19]; however, studies show that they can provide relevant indicators to determine stress [2, 6, 19, 21, 24, 32, 50, 60, 62].

Regardless of the stress marker, Arza et al. state that a single stress marker cannot globally assess an individual’s stress response, because stress causes different physiological reactions, and a multivariable approach is therefore suggested [6]. Due to the multifaceted characteristics of stress, determining a ground truth is a difficult process [19]. Some studies use subjective measures of perceived stress. Other studies rely on biosignals or biomarkers that they consider reliable for determining stress. In many studies, ground truth is established by placing a subject in a neutral and in a stressful situation to label the collected data accordingly. Others use the amount of workload and cognitive demand that is being applied as the stressor [6, 19, 28].

In summary, stress cannot be objectively and unobtrusively monitored in real-time [21]. Determining stress is challenging because of the subjectivity and individual nature of stress [21]. Moreover, the start, the duration, and the intensity of a stress event is often not clearly identifiable [21]. There is also no commonly agreed methodological or measurement standard for unobtrusive markers [6, 19]. The relationship between the body’s activation of biochemical stress markers and the intensity of the stress perceived is both complex and understudied [6]. However, it has been shown that unobtrusive stress markers can be used to approximate stress (and implicitly fatigue) in real-time [6, 19].

3 RESEARCH CONTEXT

This study is part of a European funded, interdisciplinary research project. The project aims at developing a smart training shirt for the

Table 1: An overview of commonly used stress markers.

Subjective Stress Markers	Objective Stress Markers	
	Obtrusive	Unobtrusive (real-time)
E.g., interviews, self-reports, and questionnaires.	Salvia, hair, and blood samples (e.g., cortisol or lactate).	Wearables (e.g., heart rate), contextual (e.g., air quality), video-based (e.g., thermal imaging), behavioral (e.g., physical activity).

home-based fitness market to assess a person's movements during repetitive training exercises in real-time and to provide additional health information (e.g., repetition counter). For this project, we decided to initially focus on repetitive exercises, as such exercises can be more easily assessed (e.g., in terms of the labeling process and data segmentation), but we plan to expand to non-repetitive exercises in the future. The collaborating university and industrial project partners have expertise in computer science, medical engineering, sports science, and embedded systems. The prototype in development consists of four body-worn sensors integrated into the textile of the shirt. By utilizing their mobile devices, users will be provided with immediate visual and acoustic feedback about the quality of the exercise performed as well as will be able to track their training progress over time. The current prototype integrates four IMUs, one on each shoulder, the chest, and the abdomen. A custom embedded system is used for data collection and data processing. Additional physiological sensors for biosignals (electrocardiogram, electromyography, electrodermal activity, respiration, and pulse) will be added in the future. Based on this data, supervised machine learning algorithms are trained to assess the quality of movement and detect, for example, muscle fatigue. First tests (see Figure 1) were conducted in our laboratory called Creative Space for Technical Innovations[1] to create sets of labeled training data and to implement first parts of the outlined ARC variant in software (see Section 4). Building on this context, the present study's goal is to allow for unobtrusively detecting stress levels during training exercises in the future.

Figure 1: Preliminary tests in our laboratory – Left: IMU data was leveraged to detect muscle fatigue; Right: Sets of push-up exercises were utilized to train first machine learning models based on IMU data.

4 EXPERIMENTAL OUTLINE

The experimental outline consists of a supervised machine learning training process that is built on a custom software implementation of the ARC. While we implemented the latter, we are currently working on the realization of the former.

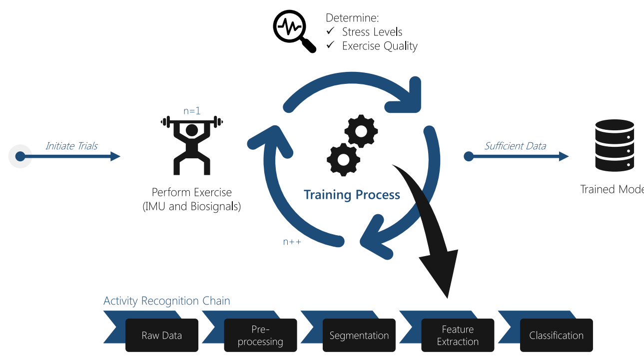

Figure 2: An overview of the training process and its individual elements.

4.1 Training Process

Figure 2 shows the experimental procedure to train a model that can classify the quality of performed repetitions and the corresponding stress levels. We chose supervised machine learning to guide the training process in this study. The training process begins with one participant who is initially wired with sensors and then performs a repetitive exercise. The unobtrusive senor data is collected during the exercise. Further data is collected separately: firstly, the quality of the performed repetitions is assessed (labeled) by sports experts; secondly, the subject fills in a questionnaire for perceived exertion; thirdly, a blood sample is taken for the lactate values at the index finger. This procedure is repeated with the same subject for different exercises. After all exercises have been completed by the subject, the training is repeated with a new participant until sufficient data has been collected. By using the word sufficient we gear towards a machine learning model that shows good results in evaluating its accuracy, precision, recall, and f-measures. We also aim at collecting a balanced set of training data to avoid too stark class imbalances that would deteriorate the classification performance [54]. Related studies recruited a varying number of (healthy) participants to reach satisfactory results. For instance, while Seiffert et al. recruited as few as two participants [48], Morris et al. gathered data from as much as 114 people (i.e., in both the training and evaluation phase) [45]. A notable number of studies recruited no more than 20 participants (e.g., [8, 25, 42, 47]) hence we initially target a similar cohort in our research. However, Morris et al. [45] note that the variation in the form inevitably affects the recognition accuracy and, consequently,

[1] https://csti.haw-hamburg.de/

see the necessity to conduct large-scale trainings. While we intent to recruit first subjects from within our research group, we therefore also plan to prospectively gather data from more people. Yet, we are unaware of any general rule of thumb as regard to a minimal required number of subjects to reach satisfactory classification performance at scale. We expect that the recognition accuracy will incrementally improve on a subject-to-subject basis.

Features are calculated for each repetition to train a model. For this purpose, all sensor data passes through each stage of the ARC (see Section 4.2). The IMU sensor data is used as the basis for finding individual repetitions and for segmenting all other sensor data. The labels are added to each corresponding feature set (for each repetition). The trained model is eventually able to classify individual repetitions by using only unobtrusive sensor data without any labels. Since our first prototype has only IMU sensors, we could only create features for the performed quality.

Figure 3 presents the training process in detail. It is divided into two phases: a baseline and a stress phase. This structure is based on [6] and was modified to reduce the time required per participant who originally had to be available on two days. In addition, the stress phase now has a variable duration. All subjects should be evaluated in the morning to avoid differing physiological responses due to circadian changes [6, 16, 52, 53].

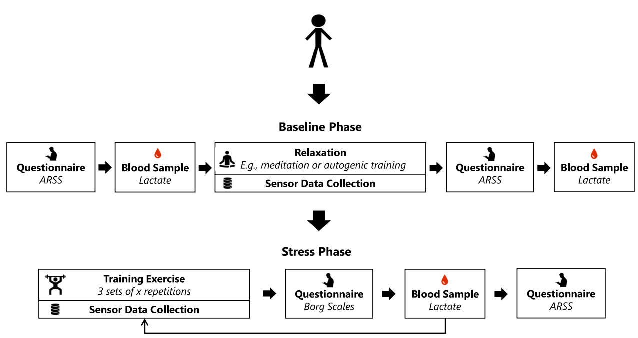

Figure 3: The rationale of the stress determining procedure throughout the training process.

In the baseline phase, individual stress levels are determined by placing each participant in a rested, neutral state for protocol calibration and the determination of normal conditions [19]. The participant first receives a short briefing regarding the procedure. After all sensors are firmly attached and in the correct position on the body, it is assured that each sensor transmits data. The subject fills in the Acute Recovery and Stress Scale (ARSS) [37]. A blood sample is taken from an index finger and analyzed for lactate levels with a handheld meter. The subject then performs a relaxation exercise (e.g., mediation or autogenic training) for ten minutes. During the relaxation exercise, data is recorded by the sensors. Upon completion of the relaxation exercise, the subject fills in the ARSS again and a second blood sample is taken to end the baseline phase.

In the subsequent stress phase, one or more training exercises are performed. Each training exercise consists of three sets. A set consists of repetitions (see Section 4.2.3). Each subject is instructed to perform as many repetitions as possible. In our preliminary experiments, this usually took less than two minutes, though this depends highly on the exercise and individual's fitness level. During each set, data is recorded by the sensors. After each set, there is a break of ten minutes. Meanwhile, the subject fills in the Borg Scales [10] on perceived exertion and another blood sample is taken after the seventh minute, as peak values can be observed three to eight minutes after exercise [23]. Once all exercises are completed, the subject fills in the ARSS one more time. This ends the stress phase and another baseline phase can begin with the next participant.

The described training process builds on the premise that sufficient training data can be collected for each stress level. Furthermore, it is presumed that workload [28], heart rate [43], and lactate values [30] correlate with stress. Magiera et al. state that the lactate level depends on recovery periods, while no effect of fatigue, when recovery periods were greater than 20 minutes, were experienced [43]. According to Kop and Kupper, there is a bidirectional relationship between fatigue and stress, so recovery periods should be considered [40]. Moreover, lactate accumulates only when the training intensity is above the anaerobic threshold (through short, intense exercises). Heart rate is affected by an increase in fatigue after a short recovery time [43]. Since our stress phase tends to be short and intense, we opted for 10-minute recovery periods, which also keeps the total time per participant low.

4.2 Implementation of the Activity Recognition Chain

As part of the aforesaid research project, we already implemented a custom variant of the ARC in MATLAB (see Figure 4). This implementation constitutes the software foundation for the experimental outline summarized in Figure 2. In the following, we introduce this custom chain and the different choices we made throughout the implementation. The central element of this implementation is the Online Classification Thread, which continually executes the different stages of the chain in real-time. Below, we present the individual stages that are visualized in Figure 4 and indicate how data passes through the chain to transform raw signals to classified stress levels. It is noted that we have implemented the chain based on IMU data and are planning to do the same with data stemming from physiological sensors in the future. Similar to Guo et al., our implementation lays a focus on processing the acceleration (IMU) data to create segments [25].

4.2.1 Raw Data Stage. Sensor data is constantly received and processed by the Receiver Thread. This thread handles all data connections to the sensors and parses the raw data to value objects which are used internally for representation. Based on related studies (e.g., [25]), we decided for working with the Euclidean norm to combine the x, y, and z axes for each IMU sensor into one signal. In doing so, it is no longer necessary to determine which of the three axes has the highest signal strength. Additionally, an exact orientation of the attached IMU sensors is no longer necessary. This is due to the reason that the gravity that accelerometers measure spreads across the three axes and the Euclidean norm summarizes the magnitudes in one signal.

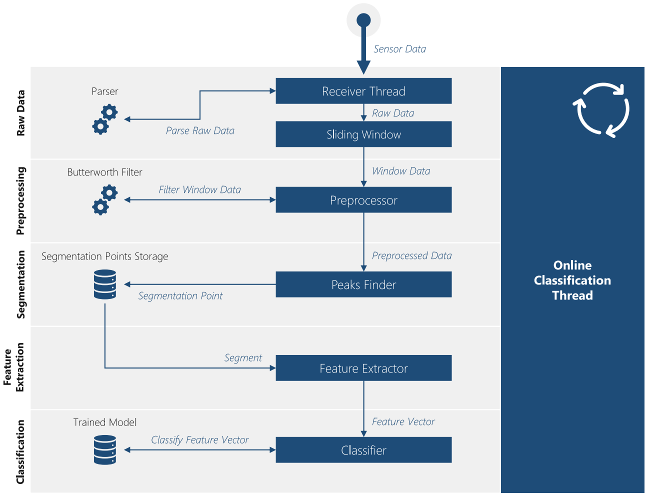

Figure 4: The different components of our custom implementation of the ARC.

Because we went with calculating the Euclidean norm, we were required to work with a sliding window approach [34]. This approach is embodied in the Sliding Window component, which limits the amount of data to be processed, primarily because the chain is intended to run on an embedded system with limited computation and storage capacity. The Sliding Window stores a total of 1000 milliseconds worth of sensor data. Currently, the prototype works with a frequency of 200 Hz but we intend to lower this frequency to 50–100 Hz as suggested by Trimpop et al. [57]. A reasonably chosen window size is crucial – a very large window would delay the real-time feedback and a small window would potentially result in detecting too many irrelevant data points in a time series. The Sliding Window component is essentially a data structure that follows the first in, first out (FIFO) principle. The window continually moves across the incoming signals. With each movement step, a new data point is added to the window and, at the same time, the oldest data point is removed. The Sliding Window is transferred through each of the next three stages.

4.2.2 Preprocessing Stage. The Preprocessor component uses the data from the Sliding Window for interpolation (i.e., if data is missing due to network losses) and for filtering the incoming sensor data. Related studies leveraged different filter techniques to smooth incoming sensor signals; reducing additive noise is the key concern at this stage due to sensor variability and limited digitization processes [19, 59]. Utilized candidates were, among others, the Butterworth filter [9] and the moving average filter [47]. In our preliminary experiments, we tested different configurations of the two mentioned filters (e.g., different window sizes for the moving average filter). Finally, we decided to use the Butterworth filter because it is efficient for real-time filtering and produces smooth signals (see Figure 5), which are beneficial for our segmentation approach (see Section 4.2.3). The following parameters were determined manually by experimental test runs: an order of 3, a cutoff frequency of 0.8, and a sample rate of 200 Hz – analogous to the sample rate of the IMU sensor. The goal was to produce a smooth signal with as little oscillation as possible and without attenuating the signal beyond recognition. We also noticed a significant shift of the signal to the right on the time axis (about 450 ms) when low cutoff frequencies (like 0.8) were used (the shift has been corrected in Figure 5). In summary, we observed that the filter settings have a notable impact on the resulting signal (i.e., especially the total number of maxima and minima).

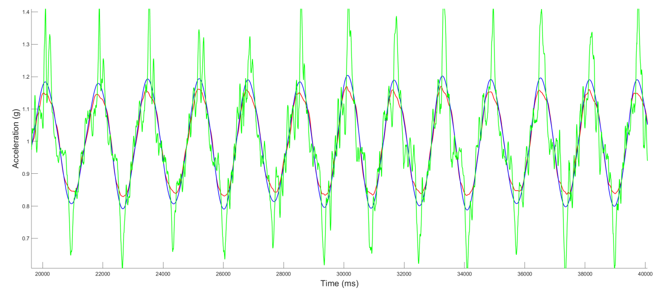

Figure 5: Comparison of filtered IMU signals (push-ups): raw data (green), moving average filter (red), and Butterworth filter (blue).

4.2.3 Segmentation Stage. The preprocessed data is forwarded to the Peaks Finder component, which detects individual repetitions of an exercise. The literature indicates different means to accomplish repetition detection such as minima and maxima searches [42], also known as Zero-Velocity Crossing [11]. We decided for a search of maxima. Every time a new data point is added to the Sliding Window, the window is checked if a new maximum can be found. For this purpose, the value in the center of the window is used as reference. An algorithm checks if all values to the left and to the right of the central reference point are smaller. If this is the case, a maximum is found and the maximum is stored in a separate storage. The window then continues to move forward until another maximum is found. All sensor data from the first to the second maximum is considered as one segment (i.e., one repetition of the current exercise). However, some exercises may consist of multiple maxima per repetition (e.g., squats), in which case a segment is created only after every x maximum. It is assumed that the exercise to be performed is known in advance.

After a segment is found, further analysis is applied to discard unwanted segments. Since Zero-Velocity Crossing algorithms tend to over-segmentation [11], some segments do not represent a valid repetition. For example, if the duration of a segment exceeds or falls below a certain threshold or if the variance (amplitude) of the segment is very low. A more advanced analysis could be the calculation of prominence [44]. However, finding the right thresholds to filter segments manually is a challenging task [11]. Another approach would be to not discard any segments and let the trained model decide whether the segment is a valid repetition.

Finally, each found segment is based on the data of the IMU sensors and will be used in the future as reference for segmenting all the other sensor data such as heart rate, which is then used to determine the stress level per repetition.

4.2.4 Feature Extraction Stage. The next stage receives and processes the latest segment found and is referred to as the Feature Extractor component in Figure 4. The foundation for this stage was a literature review to find suitable features that can characterize individual repetitions most accurately. We found a diversity of types of features for this task such as dynamic [9], statistical [25], or frequency-based features [5]. We finally went with a set of statistical features used by Guo et al. [25], who exclusively built their study on such features. However, we also limited our selection to statistical features in this early stage of our research, because we were able to readily calculate them with built-in MATLAB functions. In the end, our preliminary set of features consisted of the measures skewness, kurtosis, std, var, mode, median, range, trimmean, and mean. The selection of appropriate features is critical in influencing the accuracy of the trained model to successfully detect repetitions [33]. During this stage, the input vector of the size of n dimensions (i.e., the number of data points in a segment) is transformed to a vector, the feature vector, which has the size of one dimension. The feature vector consists of a summary of calculated values (e.g., mean and variation) that are finally passed to the next stage. In our study, the aforesaid features were calculated for the IMU used (i.e., its accelerometer and gyroscope sensors) as well as each axis of these sensors. As a result, the feature vector had a total of 54 unique measures per IMU sensor. In the future, we will enrich this IMU feature vector with data stemming from biosensors.

4.2.5 Classification Stage. The Classifier component incorporates a trained model that can determine either the quality of the performed repetition of an exercise or, in the future, the stress level (i.e., in the form of the feature vector). Like the previous stage, we initially conducted a literature review to find suitable classifier candidates. Examples are SVMs [47], decision trees [33], random forests [9], and Naive Bayes classifiers [8], whereas most of the mentioned studies leveraged and compared a set of different classifiers. However, for exploration purposes, we leveraged MATLAB's own Classification Learner App[2] to test and find suitable classifiers. In our test runs, we utilized a 5-fold cross validation and found that especially SVM classifier variants, a tree algorithm, as well as KNN algorithms showed a comparably high accuracy with our dataset. The dataset consisted of 304 push-up exercises from six different subjects. Each subject performed three sets of push-ups to the point of exhaustion. Subsequently, these push-ups were labelled by our collaborating sports science partners as correct or incorrect. Overall, 202 push-ups were labelled as correct, while 102 push-ups were identified as incorrectly executed. Table 2 shows an overview of the most accurate classifiers that were trained in MATLAB. Interestingly, while visually exploring the dataset in a scatter plot, we found that both the standard deviation and the variance of the gyroscope's x and y axis were particularly suitable to separate both classes of data (i.e., correctly and incorrectly executed push-ups). The results presented in Table 2 are based on these two measures from the feature vector.

The outcome of the classification stage are labels such as correct and incorrect that reflect the internally determined decision of the trained model. It is noted that in our preliminary experiments, we concentrated on a binary classification problem, while we are planning to incorporate more classes to distinguish correctly from incorrectly performed repetitions (i.e., multi-class classification). We would also like to emphasize that, although related work shows similar good results [47], the findings presented Table 2 illustrate notable high accuracies and we therefore plan to conduct further investigations in the future (e.g., as regard to overfitting). We will also incorporate more data from other subjects.

[2]https://www.mathworks.com/help/stats/classificationlearner-app.html

Table 2: An overview of the most accurate classifiers.

Classifiers	**Accuracy**
Cubic SVM	98.0 %
Medium KNN	98.0 %
Coarse Tree, Quadratic SVM, Fine Gaussian SVM, Coarse Gaussian SVM, and Cubic KNN	All 97.7 %

5 DISCUSSION

The main contribution of this paper is a preliminary experimental outline, how we plan to unobtrusively detect stress levels during training exercises. Our work's novelty primarily arises from the circumstance that there exist methodological and measurement challenges when it comes to determining stress in this demanding context [2, 6, 41, 50]. To the best of our knowledge, the present study is one of only a few studies that attempts to investigate stress during trainings exercises [31, 43]. However, we experienced several challenges throughout our ongoing research. For example, as indicated in the literature [11, 42], we similarly experienced the segmentation process as difficult. While different filters were examined in preparation for the segmentation stage, various settings for the peak-finding algorithms were tested to reduce over-segmentation. Our preliminary results concur with the related literature as regard to the classification process. Like Bevilacqua et al. [9] and Guo et al. [25], we also found that SVM classifiers (i.e., variants of this classifier) are particularly accurate in correctly determining the quality of an exercise. In that regard, our study stands in some contrast to [8] as we did not find as strong the support for the Naive Bayes classifier in our test runs. This classifier, according to Baumbach and Dengel, shows similar accuracy performance to more complex classifiers [8]. It is worth highlighting that other studies also took entirely different avenues to the segmentation problem and used, for instance, machine learning algorithms such as clustering to unveil similarity in the data [8]. Again, Lin et al. provide a thorough overview in that regard [42]. Our contribution is, however, also underlined by the non-academic part of the research project – i.e., a product for the fitness market is to be developed. Segmentation approaches are rarely applied beyond academic contexts [18] and more of such approaches are warranted that operate in real-time, produce accurate segments, and are computationally inexpensive [11].

We see the following implications for research and practice. Researchers, on the one hand, profit the most from our elaborations in terms of the literature work and the design decisions we made. Future studies can leverage this knowledge to design their own studies and to make profound contributions to the field. Practitioners (e.g., fitness studio personnel and health tool developers), on the other hand, benefit most notably from the fact that we envision ways to unobtrusively detect stress levels. Prospectively, studies such as ours may lead to less expensive equipment and labor-intensive work (e.g., analyzing blood samples in laboratories) to determine stress levels during training exercises. Likewise, the feedback for the trainee is enriched and potential injuries may be avoided. In sports and healthcare, feedback regarding an execution's quality is an interesting aspect to consider [48].

Our work is not without limitations. Firstly, most of the hardware is provided to us by the cooperating company developing the embedded system. Hence, whether there are more accurate sensors on the market to more accurately detect repetitions, is an issue beyond the scope of this study and may affect the overall results (e.g., the segmentation procedure).

Secondly, lactate has limited use in measuring stress due to the anaerobic threshold and recovery periods [43]. Lactate has only been used in recent studies on stress [13, 30]; most studies on stress use cortisol as obtrusive marker, because cortisol secretion is directly associated with activation of the hypothalamic-pituitary-adrenal axis [6, 19]. However, cortisol is more expensive to analyze per sample.

Thirdly, further research is needed to identify which set of stress markers is best suited to determine stress levels [6, 19]. We will initially focus on heart rate variability, lactate, and perceived stress scales, but other markers such as electrodermal activity or body temperature show promising results as well [60]. Furthermore, the use of subjective stress scales can be prone to self-deception, fabrication, and attention bias [51].

Fourthly, according to Giannakakis et al. [19] and Morris et al. [45], it is important to take the natural variability into account when measuring stress in experiments. A laboratory environment that does not aesthetically resemble a gym may result in different data as if the experiments had been carried out in a real environment. Giannakakis et al. suggest keeping the experimental environmental conditions constant [19], which is challenging due to the variety of contextual stressors such as the duration of the experiment, rest periods, noise factor, temperature, lightning, or air quality [19, 36, 39, 53].

Finally, the segmentation approach we have chosen is known to be efficient, to require little computational effort, and to allocate a comparably small amount of system memory. However, the approach also has some drawbacks. First, it can lead to over-segmentation of the data [11], and second, it does not generalize well across primitives and subjects [42]. Moreover, the approach is very sensitive to the chosen size of the sliding window (see Section 4.2.1) and how the data is preprocessed and filtered (see Section 4.2.2).

6 CONCLUSION AND FUTURE WORK

This study responses to recent developments regarding the lack of methodological and measurement standards to detect stress during challenging contexts such as training exercises [2, 6, 41, 50]. It introduces a preliminary experimental outline that illustrates, how we plan to unobtrusively detect stress in our experiments. To this end, we elaborated on both the rationale behind the training process to be developed and the specific variant of the ARC that we already implemented. We intend to conduct test runs with participants in the near future and we are currently preparing for these studies by, among others, evaluating different technical means to measure blood lactate levels and assessing suitable exercise candidates.

In summary, there exists no commonly accepted definition of stress [6], whereas stress is caused by various stimuli such as fatigue [38, 40]. Due to the limited number of studies scrutinizing stress levels during trainings exercises, we conclude that this field encompasses promising avenues for future research. In outlining our preliminary implementation and design decisions, we hope that other researchers will find helpful assistance in preparing their own studies as they embark on similar endeavors including injury prevention and rehabilitation training.

ACKNOWLEDGMENTS

This study is part of a research project funded by the European Regional Development Fund and the Hamburgische Investitions- und Förderbank.

REFERENCES

[1] G. D. Abowd, A. K. Dey, R. Orr, and J. Brotherton. 1997. Context-awareness in wearable and ubiquitous computing. In *Digest of Papers. First International Symposium on Wearable Computers.* 179–180. https://doi.org/10.1109/ISWC.1997.629943

[2] Jordi Aguiló, Pau Ferrer-Salvans, Antonio García-Rozo, Antonio Armario, Ángel Corbí, Francisco J. Cambra, Raquel Bailón, Ana González-Marcos, Gerardo Caja, Sira Aguiló, Raúl López-Antón, Adriana Arza-Valdés, and Jorge M. Garzón-Rey. 2015. Project ES3: attempting to quantify and measure the level of stress. *Revista de neurologia* 61, 9 (01 Nov 2015), 405–415. https://pubmed.ncbi.nlm.nih.gov/26503316 26503316[pmid].

[3] F. Alamudun, J. Choi, R. Gutierrez-Osuna, H. Khan, and B. Ahmed. 2012. Removal of subject-dependent and activity-dependent variation in physiological measures of stress. In *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops.* 115–122. https://doi.org/10.4108/icst.pervasivehealth.2012.248722

[4] A. Angeli, M. Minetto, A. Dovio, and P. Paccotti. 2004. The overtraining syndrome in athletes: A stress-related disorder. *Journal of Endocrinological Investigation* 27, 6 (June 2004), 603–612. https://doi.org/10.1007/bf03347487

[5] D. Anguita, A. Ghio, L. Oneto, X. Parra, and Jorge Luis Reyes-Ortiz. 2013. A Public Domain Dataset for Human Activity Recognition using Smartphones. In *ESANN.*

[6] Adriana Arza, Jorge Mario Garzón-Rey, Jesús Lázaro, Eduardo Gil, Raul Lopez-Anton, Conchita de la Camara, Pablo Laguna, Raquel Bailon, and Jordi Aguiló. 2018. Measuring acute stress response through physiological signals: towards a quantitative assessment of stress. *Medical & Biological Engineering & Computing* 57, 1 (Aug. 2018), 271–287. https://doi.org/10.1007/s11517-018-1879-z

[7] Derek Ball. 2014. Metabolic and endocrine response to exercise: sympathoadrenal integration with skeletal muscle. *Journal of Endocrinology* 224, 2 (Nov. 2014), R79–R95. https://doi.org/10.1530/joe-14-0408

[8] Sebastian Baumbach and Andreas Dengel. 2017. Measuring the Performance of Push-ups - Qualitative Sport Activity Recognition. In *Proceedings of the 9th International Conference on Agents and Artificial Intelligence.* SCITEPRESS - Science and Technology Publications. https://doi.org/10.5220/0006114503740381

[9] A. Bevilacqua, B. Huang, R. Argent, B. Caulfield, and T. Kechadi. 2018. Automatic classification of knee rehabilitation exercises using a single inertial sensor: A case study. In *2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN).* 21–24. https://doi.org/10.1109/BSN.2018.8329649

[10] G. A. Borg. 1982. Psychophysical bases of perceived exertion. *Medicine and science in sports and exercise* 14, 5 (1982), 377–381. https://pubmed.ncbi.nlm.nih.gov/7154893 7154893[pmid].

[11] Louise Brennan, Antonio Bevilacqua, Tahar Kechadi, and Brian Caulfield. 2020. Segmentation of shoulder rehabilitation exercises for single and multiple inertial sensor systems. *Journal of Rehabilitation and Assistive Technologies Engineering* 7 (Jan. 2020), 205566832091537. https://doi.org/10.1177/2055668320915377

[12] Andreas Bulling, Ulf Blanke, and Bernt Schiele. 2014. A Tutorial on Human Activity Recognition Using Body-Worn Inertial Sensors. *ACM Comput. Surv.* 46, 3, Article 33 (Jan. 2014), 33 pages. https://doi.org/10.1145/2499621

[13] Marinella Coco, Andrea Buscemi, Tiziana Ramaci, Matej Tusak, Donatella Di Corrado, Vincenzo Perciavalle, Grazia Maugeri, Valentina Perciavalle, and Giuseppe Musumeci. 2020. Influences of Blood Lactate Levels on Cognitive Domains and Physical Health during a Sports Stress. Brief Review. *International Journal*

of Environmental Research and Public Health 17, 23 (Dec. 2020), 9043. https://doi.org/10.3390/ijerph17239043

[14] Edward F Coyle. 2000. Physical activity as a metabolic stressor. *The American Journal of Clinical Nutrition* 72, 2 (Aug. 2000), 512S–520S. https://doi.org/10.1093/ajcn/72.2.512s

[15] Elissa S. Epel, Alexandra D. Crosswell, Stefanie E. Mayer, Aric A. Prather, George M. Slavich, Eli Puterman, and Wendy Berry Mendes. 2018. More than a feeling: A unified view of stress measurement for population science. *Frontiers in Neuroendocrinology* 49 (April 2018), 146–169. https://doi.org/10.1016/j.yfrne.2018.03.001

[16] J. J. Forsyth and T. Reilly. 2004. Circadian rhythms in blood lactate concentration during incremental ergometer rowing. *European Journal of Applied Physiology* 92, 1-2 (June 2004), 69–74. https://doi.org/10.1007/s00421-004-1059-8

[17] Reinhard Fuchs and Markus Gerber (Eds.). 2018. *Handbuch Stressregulation und Sport.* Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49322-9

[18] S. Gharghabi, Y. Ding, C. M. Yeh, K. Kamgar, L. Ulanova, and E. Keogh. 2017. Matrix Profile VIII: Domain Agnostic Online Semantic Segmentation at Superhuman Performance Levels. In *2017 IEEE International Conference on Data Mining (ICDM).* 117–126. https://doi.org/10.1109/ICDM.2017.21

[19] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis. 2019. Review on psychological stress detection using biosignals. *IEEE Transactions on Affective Computing* (2019), 1–1. https://doi.org/10.1109/TAFFC.2019.2927337

[20] Oonagh M Giggins, Kevin T Sweeney, and Brian Caulfield. 2014. Rehabilitation exercise assessment using inertial sensors: a cross-sectional analytical study. *Journal of NeuroEngineering and Rehabilitation* 11, 1 (2014), 158. https://doi.org/10.1186/1743-0003-11-158

[21] Martin Gjoreski, Mitja Luštrek, Matjaž Gams, and Hristijan Gjoreski. 2017. Monitoring stress with a wrist device using context. *Journal of Biomedical Informatics* 73 (Sept. 2017), 159–170. https://doi.org/10.1016/j.jbi.2017.08.006

[22] Lívea Dornela Godoy, Matheus Teixeira Rossignoli, Polianna Delfino-Pereira, Norberto Garcia-Cairasco, and Eduardo Henrique de Lima Umeoka. 2018. A Comprehensive Overview on Stress Neurobiology: Basic Concepts and Clinical Implications. *Frontiers in Behavioral Neuroscience* 12 (July 2018). https://doi.org/10.3389/fnbeh.2018.00127

[23] Matthew L. Goodwin, James E. Harris, Andrés Hernández, and L. Bruce Gladden. 2007. Blood Lactate Measurements and Analysis during Exercise: A Guide for Clinicians. *Journal of Diabetes Science and Technology* 1, 4 (July 2007), 558–569. https://doi.org/10.1177/193229680700100414

[24] Aishwarya Goyal, Shailendra Singh, Dharam Vir, and Dwarka Pershad. 2016. Automation of Stress Recognition Using Subjective or Objective Measures. *Psychological Studies* 61, 4 (Nov. 2016), 348–364. https://doi.org/10.1007/s12646-016-0379-1

[25] Xiaonan Guo, Jian Liu, and Yingying Chen. 2020. When your wearables become your fitness mate. *Smart Health* 16 (May 2020), 100114. https://doi.org/10.1016/j.smhl.2020.100114

[26] Anthony C Hackney. 2006. Stress and the neuroendocrine system: the role of exercise as a stressor and modifier of stress. *Expert Review of Endocrinology & Metabolism* 1, 6 (Nov. 2006), 783–792. https://doi.org/10.1586/17446651.1.6.783

[27] Anthony C. Hackney and Amy R. Lane. 2015. Exercise and the Regulation of Endocrine Hormones. In *Progress in Molecular Biology and Translational Science.* Elsevier, 293–311. https://doi.org/10.1016/bs.pmbts.2015.07.001

[28] Michael John Hamlin, Danielle Wilkes, Catherine A. Elliot, Catherine A. Lizamore, and Yaso Kathiravel. 2019. Monitoring Training Loads and Perceived Stress in Young Elite University Athletes. *Frontiers in Physiology* 10 (Jan. 2019). https://doi.org/10.3389/fphys.2019.00034

[29] D.H. Hellhammer, A.A. Stone, J. Hellhammer, and J. Broderick. 2010. Measuring Stress. In *Encyclopedia of Behavioral Neuroscience.* Elsevier, 186–191. https://doi.org/10.1016/b978-0-08-045396-5.00188-3

[30] Robin Hermann, Daniel Lay, Patrick Wahl, Walton T. Roth, and Katja Petrowski. 2019. Effects of psychosocial and physical stress on lactate and anxiety levels. *Stress* 22, 6 (May 2019), 664–669. https://doi.org/10.1080/10253890.2019.1610743

[31] Jin-Hyuk Hong, Julian Ramos, and Anind K. Dey. 2012. Understanding Physiological Responses to Stressors during Physical Activity. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (Pittsburgh, Pennsylvania) *(UbiComp '12).* Association for Computing Machinery, New York, NY, USA, 270–279. https://doi.org/10.1145/2370216.2370260

[32] Karen Hovsepian, Mustafa al'Absi, Emre Ertin, Thomas Kamarck, Motohiro Nakajima, and Santosh Kumar. 2015. cStress. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15.* ACM Press. https://doi.org/10.1145/2750858.2807526

[33] B. Huang, O. Giggins, T. Kechadi, and B. Caulfield. 2016. The limb movement analysis of rehabilitation exercises using wearable inertial sensors. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* 4686–4689. https://doi.org/10.1109/EMBC.2016.7591773

[34] Shima Imani, Frank Madrid, Wei Ding, Scott E. Crouter, and Eamonn Keogh. 2020. Introducing time series snippets: a new primitive for summarizing long time series. *Data Mining and Knowledge Discovery* 34, 6 (July 2020), 1713–1743. https://doi.org/10.1007/s10618-020-00702-y

[35] ISO. 2015. I.O.S., "ISO 9000:2015 Quality management systems - Fundamentals and vocabulary," International Organization for Standardization, 2015. https://www.iso.org/. Accessed: 2021-04-17.

[36] Johanna Kallio, Elena Vildjiounaite, Jani Koivusaari, Pauli Räsänen, Heidi Similä, Vesa Kyllönen, Salla Muuraiskangas, Jussi Ronkainen, Jari Rehu, and Kaisa Vehmas. 2020. Assessment of perceived indoor environmental quality, stress and productivity based on environmental sensor data and personality categorization. *Building and Environment* 175 (May 2020), 106787. https://doi.org/10.1016/j.buildenv.2020.106787

[37] Michael Kellmann and Sarah Kölling. 2019. The Acute Recovery and Stress Scale. In *Recovery and Stress in Sport.* Routledge, 16–38. https://doi.org/10.4324/9780429423857-4

[38] Rüya D Kocalevent, Andreas Hinz, Elmar Brähler, and Burghard F Klapp. 2011. Determinants of fatigue and stress. *BMC Research Notes* 4, 1 (July 2011). https://doi.org/10.1186/1756-0500-4-238

[39] C.E. Koch, B. Leinweber, B.C. Drengberg, C. Blaum, and H. Oster. 2017. Interaction between circadian rhythms and stress. *Neurobiology of Stress* 6 (Feb. 2017), 57–67. https://doi.org/10.1016/j.ynstr.2016.09.001

[40] W.J. Kop and N. Kupper. 2016. *Fatigue and stress.* Academic Press, 345–350.

[41] Kalliopi Kyriakou, Bernd Resch, Günther Sagl, Andreas Petutschnig, Christian Werner, David Niederseer, Michael Liedlgruber, Frank Wilhelm, Tess Osborne, and Jessica Pykett. 2019. Detecting Moments of Stress from Measurements of Wearable Physiological Sensors. *Sensors* 19, 17 (Sept. 2019), 3805. https://doi.org/10.3390/s19173805

[42] J. F. Lin, M. Karg, and D. Kulić. 2016. Movement Primitive Segmentation for Human Motion Modeling: A Framework for Analysis. *IEEE Transactions on Human-Machine Systems* 46, 3 (2016), 325–339. https://doi.org/10.1109/THMS.2015.2493536

[43] Artur Magiera, Robert Roczniok, Ewa Sadowska-Krępa, Katarzyna Kempa, Oskar Placek, and Aleksandra Mostowik. 2018. The Effect of Physical And Mental Stress on the Heart Rate, Cortisol and Lactate Concentrations in Rock Climbers. *Journal of Human Kinetics* 65, 1 (Dec. 2018), 111–123. https://doi.org/10.2478/hukin-2018-0024

[44] MathWorks. 2021. Prominence. https://de.mathworks.com/help/signal/ug/prominence.html. Accessed: 2021-04-17.

[45] Dan Morris, T. Scott Saponas, Andrew Guillory, and Ilya Kelner. 2014. RecoFit. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM. https://doi.org/10.1145/2556288.2557116

[46] Barry S. Oken, Irina Chamine, and Wayne Wakeland. 2015. A systems approach to stress, stressors and resilience in humans. *Behavioural Brain Research* 282 (April 2015), 144–154. https://doi.org/10.1016/j.bbr.2014.12.047

[47] Ana Pereira, Duarte Folgado, Ricardo Cotrim, and Inês Sousa. 2019. Physiotherapy Exercises Evaluation using a Combined Approach based on sEMG and Wearable Inertial Sensors. In *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies.* SCITEPRESS - Science and Technology Publications. https://doi.org/10.5220/0007391300730082

[48] M. Seiffert, F. Holstein, R. Schlosser, and J. Schiller. 2017. Next Generation Cooperative Wearables: Generalized Activity Assessment Computed Fully Distributed Within a Wireless Body Area Network. *IEEE Access* 5 (2017), 16793–16807. https://doi.org/10.1109/ACCESS.2017.2749005

[49] Hans Selye. 1956. *The stress of life.* McGraw-Hill.

[50] Dhruv R. Seshadri, Ryan T. Li, James E. Voos, James R. Rowbottom, Celeste M. Alfes, Christian A. Zorman, and Colin K. Drummond. 2019. Wearable sensors for monitoring the physiological and biochemical profile of the athlete. *npj Digital Medicine* 2, 1 (July 2019). https://doi.org/10.1038/s41746-019-0150-9

[51] Saul Shiffman, Arthur A. Stone, and Michael R. Hufford. 2008. Ecological Momentary Assessment. *Annual Review of Clinical Psychology* 4, 1 (April 2008), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

[52] Rodrigo P Simões, Renata G Mendes, Viviane Castello, Heloísa G Machado, Larissa B Almeida, Vilmar Baldissera, Aparecida M Catai, Ross Arena, and Audrey Borghi-Silva. 2010. Heart-Rate Variability and Blood-Lactate Threshold Interaction During Progressive Resistance Exercise in Healthy Older Men. *Journal of Strength and Conditioning Research* 24, 5 (May 2010), 1313–1320. https://doi.org/10.1519/jsc.0b013e3181d2c0fe

[53] Matthew A. Stults-Kolehmainen and Rajita Sinha. 2013. The Effects of Stress on Physical Activity and Exercise. *Sports Medicine* 44, 1 (Sept. 2013), 81–121. https://doi.org/10.1007/s40279-013-0090-5

[54] Yanmin Sun, Andrew K. C. Wong, and Mohamed S. Kamel. 2009. Classification of imbalanced data: a review. *International Journal of Pattern Recognition and Artificial Intelligence* 23, 04 (June 2009), 687–719. https://doi.org/10.1142/s0218001409007326

[55] Portia E Taylor, Gustavo J M Almeida, Takeo Kanade, and Jessica K Hodgins. 2010. Classifying human motion quality for knee osteoarthritis using accelerometers. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology.* IEEE. https://doi.org/10.1109/iembs.2010.5627665

[56] Ulrika Tranæus, Andreas Ivarsson, and Urban Johnson. 2017. Stress and Injuries in Elite Sport. In *Handbuch Stressregulation und Sport.* Springer Berlin Heidelberg,

451–466. https://doi.org/10.1007/978-3-662-49322-9_22

[57] John Trimpop, Hannes Schenk, Gerald Bieber, Friedrich Lämmel, and Paul Burggraf. 2017. Smartwatch Based Respiratory Rate and Breathing Pattern Recognition in an End-Consumer Environment. In *Proceedings of the 4th International Workshop on Sensor-Based Activity Recognition and Interaction* (Rostock, Germany) *(iWOAR '17)*. Association for Computing Machinery, New York, NY, USA, Article 4, 5 pages. https://doi.org/10.1145/3134230.3134235

[58] Jean M. Williams and Mark B. Andersen. 1998. Psychosocial antecedents of sport injury: Review and critique of the stress and injury model'. *Journal of Applied Sport Psychology* 10, 1 (March 1998), 5–25. https://doi.org/10.1080/10413209808406375

[59] David A. Winter. 2009. *Biomechanics and Motor Control of Human Movement*. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470549148

[60] Johnny Chun Yiu Wong, Jun Wang, Eugene Yujun Fu, Hong Va Leong, and Grace Ngai. 2019. Activity Recognition and Stress Detection via Wristband. In *Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia* (Munich, Germany) *(MoMM2019)*. Association for Computing Machinery, New York, NY, USA, 102–106. https://doi.org/10.1145/3365921.3365950

[61] Aurelio Zafra, Víctor Rubio, and Enrique Ortega. 2015. *In: Sports Injuries PREDICTING AND PREVENTING SPORT INJURIES: THE ROLE OF STRESS.* 87–104.

[62] Abdulaziz Zamkah, Terence Hui, Simon Andrews, Nilanjan Dey, Fuqian Shi, and R. Simon Sherratt. 2020. Identification of Suitable Biomarkers for Stress and Emotion Detection for Future Personal Affective Wearable Sensors. *Biosensors* 10, 4 (April 2020), 40. https://doi.org/10.3390/bios10040040