Smart Health (2019)

Contents lists available at ScienceDirect

Smart Health

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

LiftRight: Quantifying strength training performance using a wearable sensor

Slobodan Milanko*, Shubham Jain*

Department of Computer Science, Old Dominion University, Norfolk, VA, United States

ARTICLE INFO

Keywords:
Strength training
Motion tracking
Exercise segmentation
Digital health
Exercise kinematics
Wearable sensor systems

ABSTRACT

Exercise is shown to affect the physical, psychological, and cognitive aspects of one's health. Despite the prevalence of fitness trackers for tracking heart-rate, steps, and other wellness metrics, very few tools exist to monitor strength training performance. This paper presents *LiftRight* which introduces a low-cost approach for quantifying an individual's performance during strength training sessions. Unique challenges include the lack of useful qualitative or quantitative metrics, and monitoring the complex motion of the exercises. We leverage an arm-mounted inertial sensor for tracking the user's arm and computing associated kinetics. We choose three upper-body strength training exercises, and segment the workout trace into constituent sets, repetitions, and phases. We also compute performance metrics such as velocity, range of motion, stability angle, and sticking points for eight candidates over a total period of 26 weeks. LiftRight achieves 96% accuracy in identifying sets and repetitions in a workout. In terms of lifting practices and its affects on performance, our quantitative and qualitative observations are found to be consistent with existing literature. We believe that the ability to accurately monitor weight training will lower the entrance barrier and help prevent injuries by helping users and trainers alike.

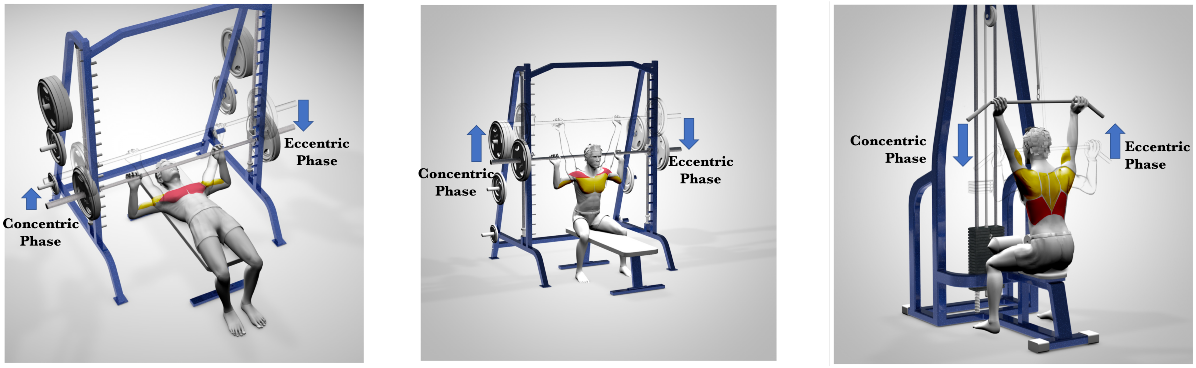
1. Introduction

Exercise can improve many aspects of one's life, such as their overall strength, sports performance, health and well-being, and even psychological stability (DHHS/PHS, 2000; Scully et al., 1998). Despite the proliferation of fitness tracking devices that monitor heart rate, number of steps, and even swimming (Fitbit, 2016; Garmin, 2016; Samsung, 2016; Apple, 2016), weight training is still a widely unexplored area. For fitness trackers to have a bigger impact on our health and lives, they need to expand beyond the limited assortment of cardio-based exercises. During the last few decades, the popularity of weight training exercises outside the professional domain has increased rapidly, primarily due to the amount of literature and blogs on the subject that have resulted in increased awareness on the benefits of weight training workouts. Some of these benefits include improved physical performance, improvement in cognitive abilities, prevention and management of type 2 diabetes, enhancing cardiovascular health, and promoting bone development, among others (Westcott; Winett & Carpinelli, 2001).

Despite the plethora of information hurled at the unsuspecting user, most of which may be contradictory, there exist limited tools for basic weight training utilities such as counting the number of sets and reps one performs. Many fitness trackers claim to track weight training workouts, but they require one to manually input the number of sets (PushBand, 2017) or even the number of repetitions for every session (Fitbod, 2016). Ding et al. (2015) integrate RFID technology in weight equipment towards exercise recognition. They focus on distinguishing between an array of weight training exercises, and require training data for exercise. In the past, researchers (Chang et al., 2007) have used accelerometer mounted on a glove to perform exercise detection and rep counting, but no performance was measured.

*Corresponding author.
e-mail: smila001@odu.edu (Slobodan Milanko), jain@cs.odu.edu (Shubham Jain)

http://dx.doi.org/10.1016/j.smhl.xxxx.xx.xxx
Available online 25 September 2019
2352-6483/



(a) Bench Press (b) Overhead Press (c) Lat Pulldown

Fig. 1: Our sample set of strength training exercises. The arrows denote the phases in each exercise.

The primary challenge in accurately monitoring weight training performance lies in the complex nature of various muscle involvement and body movement. Every exercise focuses on a different muscle group, which makes tracking movements harder because sensors would have to be placed at the most effective position for each exercise. This is in contrast to cardiovascular exercises where the entire body moves together and monitoring the performance becomes relatively easier. Moreover, lack of automated tools to measure form and performance have exposed users to more injuries while weight training (Matthews, 2014). Like any other sport, form or technique is of utmost importance. Lack of a fitness trainer or a spotter can expose one to the risk of continued bad form, or pushing through (what seems) temporary pain. Through this research work we are taking a step in the direction of continuous monitoring as a suggested preventive measure for these injuries.

Motivated by the void in performance monitoring for weight training, we present *LiftRight*. LiftRight is a wearable sensing system that performs fine-grained motion tracking for workout segmentation and quantifying weight training performance. It addresses the aforementioned challenges by utilizing an inertial measurement unit (IMU) that is worn around the upper arm to provide maximum coverage for a range of upper body exercises. For developing and validating LiftRight, we target three upper body weight lifting workouts that manipulate different muscles of the body: bench press, overhead press, and lat pulldowns.

LiftRight is a low-cost abstraction that captures upper body dynamics and computes several performance measures accurately. These measures can be directly consumed by the user or a fitness trainer to generate comprehensive performance analysis. We focus on segmenting the time-series workout trace from the IMU into sets which are further divided into constituent reps. For measuring user performance, LiftRight detects the range of motion and the velocity for each lifting/lowering cycle. LiftRight also determines recovery times and identifies other qualitative attributes- such as when fatigue starts setting in. This can help participants make informed decisions regarding meeting their goals. We focus on monitoring fundamental attributes for each workout. Capturing these statistics serves as a precursor to muscle fatigue detection, injury prevention, and overall health improvement.

In summary, the salient contributions of LiftRight are:

- **Fine-grained workout segmentation:** Design and implementation of a low-cost technique for computing body kinetics during weight training sessions using a single wearable sensor. LiftRight can achieve fine-grained detection in identifying the phases in each repetition of every set. Unlike other systems, LiftRight does not require any training data or samples of weight-training exercises being performed. It is a stand-alone system that does not rely on infrastructure.
- **Form and Performance Analysis:** Measure common performance metrics and validate them for three sample weight training exercises for 8 users, over a total period of 26 weeks of workout data. This amounts to a total of $4,000$ repetitions.

2. Background and Challenges

In this section, we present basic definitions and popular concepts in the field of weight training (Fleck & Kraemer, 2014; Bryant, 2013). Weight training is also known as strength training or resistance training. While there are subtle differences in the usage of these terms, for the purpose of this paper, we focus on resistance training with free weights or with a weight training machine, and use the above terms interchangeably.

2.1. Basic Definitions

We define some basic terms here that are used throughout the paper. Figure 1 demonstrates our sample exercises.

1. **Repetition:** One complete motion of an exercise, known as a *rep*. It consists of two phases - concentric and eccentric.
2. **Set:** A group of repetitions performed continuously without stopping or resting. A set typically ranges from 1 to 15 repetitions.
3. **Concentric phase:** Also known as *positive phase*, is defined as the phase of the rep in which the major muscles contract.

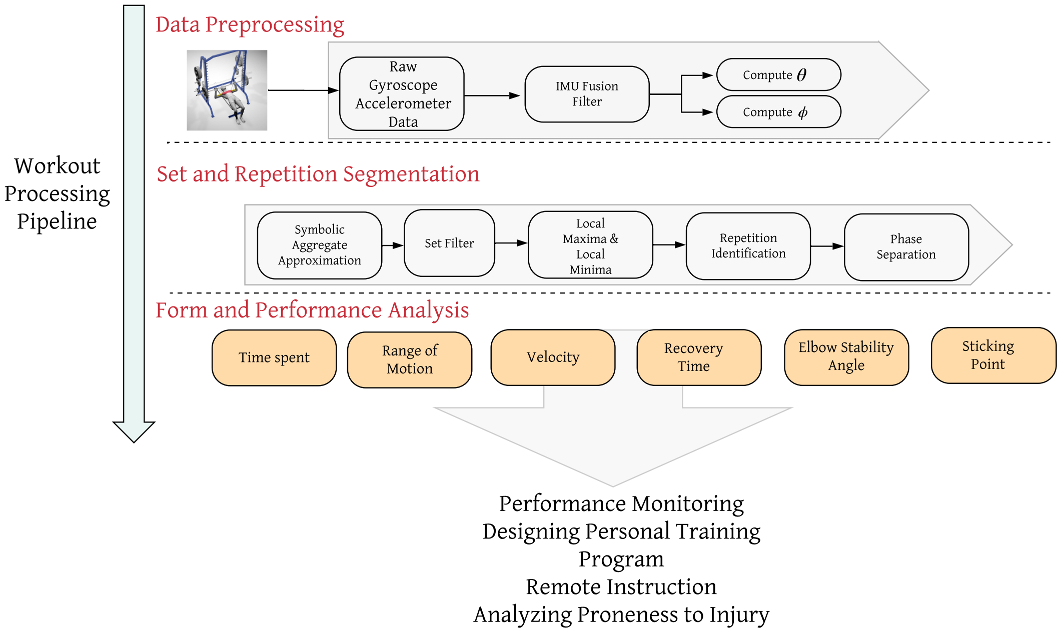


Fig. 2: LiftRight Overview.

4. **Eccentric phase:** Also known as *negative phase*, is defined as the phase of the rep in which the involved muscles lengthen or relax.

5. **Recovery periods:** Rest time in between sets in a training session.

6. **Training Program:** Is determined by the resistance/weight used, the number of sets and reps, and the duration of the recovery/rest periods.

7. **Sticking Point:** A moment in time where the applied force cannot overcome the applied resistance, either leading to rep failure or lockup.

2.2. Challenges

In order to understand upper body dynamics and compute performance measures for strength training sessions, we face the following challenges:

Lack of standard guidelines for form and motion. Aside from power lifting standards (IPF, 2016), there are no universally accepted guidelines for the right way to perform a bench press, for example. Lack of such guidelines render it difficult to define what is considered correct or incorrect technique for weight lifting. LiftRight addresses this challenge by focusing only on computing underlying metrics and accurately capturing its variation over time rather than determining the 'correctness of form'. These metrics can then be used by an expert, human or virtual, to determine form efficiency.

Lack of standard performance measures. Often people measure their performance in terms of the number of sets and repetitions in each set, while also factoring in the load. However, these measures alone are misleading. Previous research (Padulo et al., 2012) has shown that the form of the motion, which includes factors such as velocity, force, and range are equally important in determining the quality of a workout session. LiftRight computes velocity, range of motion, and elbow stability angle in addition to the number of sets and reps in each workout.

Diverse individual style of weight lifting. The ergonomic style of lifting weights varies largely among people. For example, for a workout such as bench press, different people hold the bar at a different grip, which affects their form and motion during the exercise. To overcome this challenge, LiftRight employs fundamental signal processing techniques that do not require training data, and hence are generalizable over subjects.

Limited and subtle body motion. Often only a specific part of the body is performing the action of load lifting, depending on the muscles being worked out. This makes it challenging to capture meaningful data, and sensor placement is important. For example, a shoe mounted sensor may not capture upper body movements during a bench press. LiftRight's arm mounted sensor covers a wide range of upper body exercises.

3. System Overview

In this section we discuss the design goals and present an overview of the proposed system.

3.1. Overview

Here, we introduce the main architectural components of LiftRight and the high-level work flow of the system, as shown in Figure 2. LiftRight is a motion tracking system that focuses on weight lifting workouts by capturing upper body dynamics. The prototype is an inertial measurement unit (IMU) which is mounted on the lifter's upper arm. LiftRight collects accelerometer and gyroscope data from the IMU and passes it to the *Workout Processing Pipeline*, which is composed of three modules. The first is *Data Preprocessing*, which computes the arm's orientation along two planes from the IMU readings. The second is *Set Segmentation and Rep Identification*, which processes the time-series arm orientation data to partition the sets. Once the sets are segmented, this module identifies patterns signifying a repetition. The last unit conducts *Form and Performance Analysis* by computing primary and derived metrics for each rep in each set. Particularly, we compute the delay of rep detection, range of motion, and velocity.

3.2. Performance Metric Selection

Numerous metrics have been proposed over time to assess form and performance, but measuring all of them is outside the scope of this work. We present the metrics LiftRight computes and tracks over time, by monitoring upper body dynamics using a single arm-mounted inertial sensor. Note that training programs are largely based on personal goals. Despite their motives, they tend to overlap and rely upon several common metrics to analyze performance.

A rich body of existing literature (Fleck & Kraemer, 2014; Matthews, 2012; Bryant, 2013; Brzycki, 1993; Matthews, 2014; DHHS/PHS, 2000; Hatfield, 1993) and our local fitness trainers concur that a general goal of weight lifting exercises is to improve *power* or *strength*. Power can be improved by lifting the same weight the same distance in shorter time (Fleck & Kraemer, 2014). Thus implying that *speed of movement* is an important factor in improving power. Recovery times are an equally important measure since their primary goal is to rebuild the energy needed for the next set (Hatfield, 1993; Bryant, 2013; Hatfield, 1993).

The most controversial similarity between all training programs is form. While there are many different characteristics that can define form, it typically relates to the efficiency of performing an exercise correctly. This efficiency can be described with various characteristics, of which we choose range of motion and stability. Range of motion accounts for the total distance traveled by different muscle groups. Full range of motion, especially for beginners, helps improve strength and accumulate muscular mass (Pinto et al., 2012; Matthews, 2014; Hatfield, 1993). Poor form can be identified when partial transitions are practiced (Bryant, 2013). Stability is another crucial factor in determining form, referring to the ability to safely control the applied resistance levels (Hatfield, 1993). An underdeveloped stability can promote a multitude of incorrect exercising patterns and potentially cause injury, depending on load size.

Many other factors affect performance during weight training, for example the body mass, genetics, day or time of the week, diet, general fitness level, psychological state, or even how much sleep one has been getting. We, however, focus on metrics that can be derived from upper body dynamics and can help determine changes from one session to another.

3.3. Design Goals

We compute the following quantitative and qualitative metrics that denote various aspects of form and performance, as discussed in Section 3.2.

3.3.1. Quantitative Metrics

Quantitative metrics represent values that provide basic kinetics information associated with exercise motion. They can be optimized for any strength training program with a specific goal in mind, e.g. bodybuilding or weight loss. We choose these metrics as they are widely acknowledged measures of exercise performance (Matthews, 2014; Fleck & Kraemer, 2014; Bryant, 2013; Brzycki, 1993; Hatfield, 1993).

- Number of sets and repetitions in each workout.
- Time spent in each phase: measured in $milliseconds$.
- Range of motion: measured in $degrees$.
- Velocity of movement: measured in $degrees/second$.

3.3.2. Qualitative Metrics

We also focus on more subjective measures that are hard to quantify, such as recovery time, elbow stability angle, and fatigue. Traditionally, fatigue is measured via questionnaires and blood lactate measures, or by observation. We identify sticking points in the participant's workout trace to identify when fatigue is setting in. This can be particularly useful for preventing injuries.

4. Workout Processing Pipeline

In this section, we describe the details of how LiftRight processes IMU data to compute performance measures.

4.1. Data Preprocessing

Once the IMU is mounted on the upper arm, accelerometer and gyroscope data is collected via Bluetooth onto a laptop. In order to measure and track upper body movements, we define a *shoulder coordinate system* where the shoulder is considered as the origin. In this system, the sensor is mounted on the left arm and when the user is standing, the X axis is pointed vertically toward their head, the Z axis is emanating out from the left arm (to their left), parallel to the ground, and the Y axis is pointing perpendicularly out of their torso. This coordinate system can be used to map all upper body movements that involve the torso or the shoulder. The tri-axial accelerometer and tri-axial gyroscope data is mapped to this coordinate system as shown in Figure 3, corresponding to sensor orientation shown in Figure 5a.

Mapping into the shoulder coordinate system allows us to identify two key rotations around the shoulder joint. The first is the angular rotation of the arm during the concentric and eccentric phases, referred to as θ, which is calculated along the $x - y$ plane. The second is the elbow stability angle, which is the angular distance between the upper arm and the user's torso, referred to as ϕ. We use a Kalman Filter implementation to smooth the raw data and calculate these values. The benefit of this approach is that jitters caused by rapid movements and minute vibrations of the upper arm are reduced. Lastly, we do not assume that the inertial sensor is perfectly aligned with gravity before a user commences an exercise. To address arbitrary sensor orientation, LiftRight captures the initial orientation of the sensor, and applies it as an offset, o, to the initial readings. The calculated ϕ and θ values can be obtained using Equation 1, given a quaternion input of q_0,q_1,q_2,q_3 representing the w, i, j, and k components. Note, one would have to react to singularity as needed (Blanco, 2010).

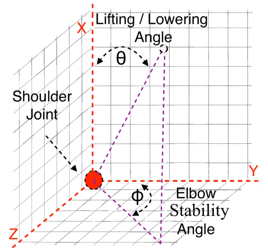


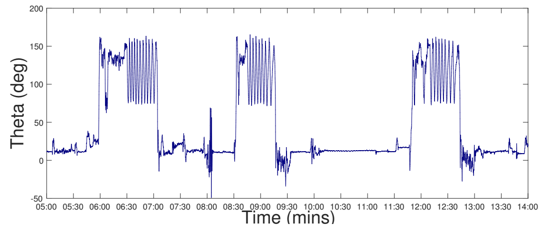
Fig. 3: Tracking motion through the shoulder coordinate system.

$$\begin{bmatrix}\phi\\ \theta\end{bmatrix} = \begin{bmatrix}\text{atan2}(2(q_0q_1 + q_2q_3), 1 - 2(q_1^2 + q_2^2))\\ \text{asin}(2(q_0q_2 - q_3q_1))\end{bmatrix} \tag{1}$$

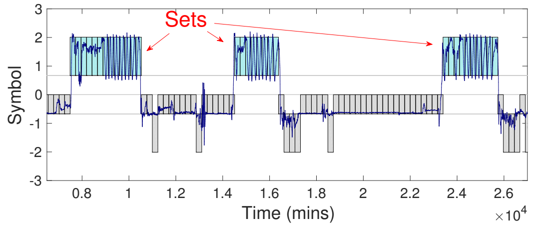
4.2. *Set segmentation*

After preprocessing the data, we obtain a filtered version of the lifting/lowering angles, θ, as shown in Figure 4a. Next, we segment this filtered time series data to obtain the sets. What makes this particularly challenging is the aperiodic nature of when sets occur, usually separated by a random recovery time chosen by the candidate. However, we observe that commencing and concluding an exercise causes a significant deviation in θ measurements compared to the data observed during the rest periods.

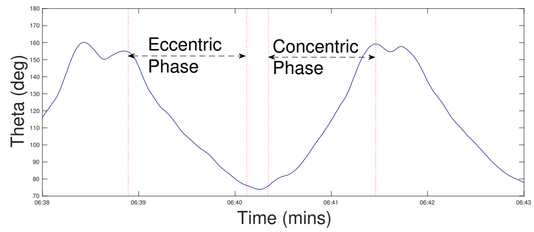
To exploit these findings, we use Symbolic Aggregate approXimation, or SAX (Lin et al., 2007, 2003). SAX provides symbolic representations for time-series data by offering dimensionality-reduction. It extends Piecewise Aggregate Approximation, PAA, which discretizes the time-series signal X of length n into a vector X' of length m, where $m << n$. SAX takes these discrete bins and produces symbols corresponding to the time-series features with equal probability, using an alphabet size $\alpha > 2$. The z-normalized values of the time-series follow a Normal distribution. Using this property, SAX picks equal-sized areas under the Normal curve using lookup tables at the boundaries of segments. Figure 4b shows the outcome of processing θ through SAX, with an empirically derived alphabet size of 4 and PAA segmentation of 256. The symbol corresponding to workout sets in our dataset are highlighted in blue, as belonging to symbol 4.



(a) Low-pass filtered lifting/lowering angle.



(b) Symbol generation.



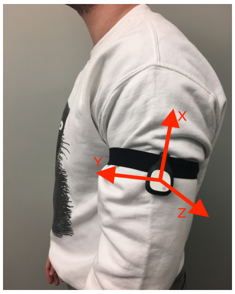
(c) Detected rep with phases.

Fig. 4: Workout processing pipeline summary. The low-pass filtered angles in (a) are segmented (b). (c) shows a sample rep detected from a set to extract exercise phases.

4.3. *Repetition Extraction*

Once sets are identified, individual reps are parsed, along with their eccentric and concentric phases. A closer look at each rep demonstrates a V-like pattern, as shown in Figure 4c. Hence, to extract reps, LiftRight finds all the local maxima within each set. Once two or more peaks are identified using local maxima, p_1 and p_2, a local minima is found in between, v_1. The three points potentially represent one rep. However, since the V-like pattern commonly occurs in noisy data, we further filter potential reps by matching them to a rep template. The rep template is a single rep extracted using the ground truth labels for a random workout. It does not matter which person, day, or exercise is chosen for this one sample rep to be used as a template, since all weight lifting exercises have concentric and eccentric phases. The template is matched against potential reps using Dynamic Time Warping (DTW). Despite the variation in the range of motion for each exercise, distance thresholds are found to be similar due to the similarity in motion patterns. The threshold is obtained from the ROC curve in Figure 6. The rep template is the only known sample LiftRight requires. It does not need any training data or exercise templates.

Once the reps are filtered, they are divided into constituent phases. This allows us to determine when concentric or eccentric phase ended or started, and the small duration spent in the transition to the next phase. To achieve this we use linear regression. As the local minima v_1 approaches p_1 or p_2, a line of best fit is modeled to determine the slope of change. The point at which this slope exhibits significant change, i.e. $\pm 0.5°$, it is marked as either the end of the eccentric or start of the concentric phase. From this, LiftRight derives the start and stop time for each phase.



(a) Sensor Orientation

(b) Angle of interest

Fig. 5: The orientation of the sensor is shown in (a). This orientation is used to capture both Lifting and Lowering Angles, shown in (b), and Elbow Stability Angle.

4.4. Form and Performance Analysis

Once concentric and eccentric phases are identified in each rep, we obtain the time spent in each phase. Δt_c denotes the time spent in the concentric phase, and Δt_e is the time spent in eccentric phase. For each phase, we compute the range of motion, given by $\Re$, with a subscript that indicates the phase:

$$\Re_c = \theta_c^{stop} - \theta_c^{start} \tag{2}$$

Here, θ_c denotes the angles in the concentric phase of the rep, with θ_c^{stop} and θ_c^{start} representing final and initial angles, respectively. The difference between the two represents range of motion $\Re$, the angular distance traveled by the arm during a phase, defined in degrees. To determine velocity for each phase in deg/sec, we define it as the angle ($\Re$) moved by the arm, over time Δt. Thus, the velocity for the concentric phase, υ_c is given by:

$$\upsilon_c = \frac{\Re_c}{\Delta t_c} \tag{3}$$

Another metric for form is the elbow stability angle, denoted by ϕ, as shown in Equation 1. The change in elbow stability angle in the concentric phase is given by: $\mathcal{E}_c = \phi_c^{stop} - \phi_c^{start}$. In bench press for example, lifting the weight while maintaining the form will require ϕ to exhibit little or no change. We discuss in Section 6 how this angle changes over a workout session. Lastly, we obtain set recovery times by observing the elapsed time between consecutive sets.

5. Implementation and Experiment Setup

Prototype Implementation. In our current implementation, inertial readings were recorded using the MetaMotionR platform (MbientLab, 2016). We capture this data through Bluetooth Low Energy, at $50Hz$, on a MAC OSX Swift application that we developed. The sensor is power efficient and we were able to record multiple 30 to 45 minute workout sessions per charge. To ensure that movements during an exercise do not cause sensor displacement, it was placed in a cover on an adjustable elastic band, that was snug around the volunteer's arm without being too tight or slipping. The sensor mounting position is shown in Figure 5a.

Participants. We conducted our experiments with 8 volunteers, 2 female and 6 male. They were all in the 27-40 years range. We focused on a diverse population for this feasibility study. Our participants are generally fit, and three of them work out at the gym regularly. 5 participants performed the exercises for 4 weeks each, while 3 performed them for 2 weeks each.

Experiment Setup. We did not provide the participants any specific instructions on how to perform the exercise for our data collection. If at any time the participant was unable to lift the bar, a spotter helped lift the bar back to a safe position. The participants were allowed to choose the time to rest in-between sets, the recovery time. They were not aware of the goal of the study or the metrics captured from their data, so as not to bias them. For each session, we fitted them with the sensor in the beginning, and measurements were recorded for the entire workout. Participants only performed one exercise in each session. Each session was recorded with a Logitec C922 camera. This recording was later used as ground truth for identifying when lifting and lowering stages commenced and ended for each session.

Training Plan. To measure performance metrics, the volunteers participated in the following three exercises: Bench press, Overhead press, and Lat pulldown. For bench press, volunteers used either the conventional bench or the Smith machine for completing the exercise. Overhead press was completed solely via the Smith machine and Lat Pulldowns via a Lat Pulldown machine. It is worth mentioning that both the conventional bench and the Lat Pull Down machines allow for more degrees of freedom during exercise, comparable to free weights. Each volunteer was allowed to choose a weight they felt comfortable with, and they were to retain it for the total duration of the experiment. The only exclusion to this was one volunteer who found the initial training weight low during the bench press, and increased it in the following session. The sessions were scheduled to account for the recommended recovery times that each muscle group needed to ensure that no overtraining was initiated, and each contained up to 8 sets and 10 repetitions per set (Bryant, 2013; Hatfield, 1993).

6. Evaluation

In this section, we aim to evaluate LiftRight by answering the following questions:

- How well can LiftRight identify sets and repetitions from a weight lifting workout session?
- How accurately can LiftRight detect the beginning and end of each phase of a rep?
- Can LiftRight accurately measure primary metrics such as range of motion and velocity of the arm?
- Can LiftRight determine events of qualitative value, such as when a lifter is starting to get tired?
- How does accuracy and usability of LiftRight compare to a commercially available tracker?

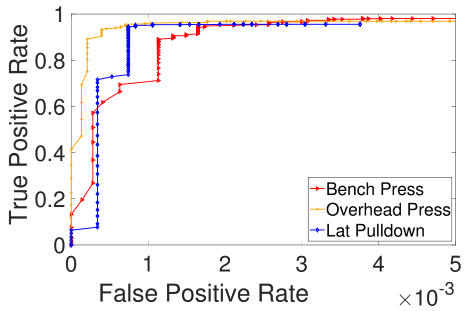


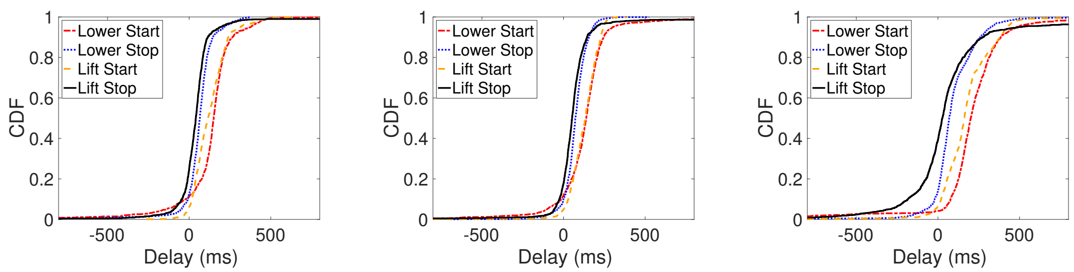
Fig. 6: ROC curve for rep identification.

6.1. *Identifying sets and repetitions*

One of LiftRight's strongest characteristics is its ability to divide a workout motion signal into constituent sets, and further subdivide each set into reps, without the need for training data.

Baseline. To record the ground truth for rep counting, we manually counted the number of sets and reps in each workout session from the captured video data.

Result. LiftRight achieved 100% set detection accuracy for all three exercises, for all participants. Figure 6 demonstrates LiftRight's performance in identifying the number of repetitions. Every correct rep detection is a true positive. In a collection of 4,000 reps, LiftRight can correctly identify 96% reps with 0.3% false positive for Lat Pulldowns; for Overhead Press it can correctly detect 97% reps with 0.2% false positives; and for Bench Press, it can detect 98% reps with 0.4% false positives. Overall, our rep detection rate is higher than 96% for all three exercises, with less than 1% false positives. In comparison to other work (Chang et al., 2007), our false positive rate is much lower. It is worth pointing out that the same algorithm was used for all the exercises to generate this curve.



(a) Overhead Press (b) Bench Press (c) Lat Pulldown

Fig. 7: Distribution of exercise phase detection delay. Negative values indicate early detection by LiftRight.

6.2. Delay in exercise timing

Beyond rep identification, LiftRight accurately detects phase timings from each rep.

Baseline. To obtain the exact times for rep start/stop, we manually labeled each phase of each rep.

To get accurate ground truth, the camera was set to capture video at 120 frames per second. The lift and lower start/stop times were labeled down to the resolution of one-eighth of a second by iterating through every frame. The video data as well as data from the sensor was recorded on the same device, and was therefore synchronized in time.

Result. Figure 7 shows the delay distribution in detecting the start and stop times of lifting and lowering phase of each rep, for all participants for each exercise. Negative values indicate that the algorithm performed early detection. LiftRight detected the lift/lower start/stop as early as 250 milliseconds. Early detections were often caused by the participant gripping the bar to start lifting.

6.3. Range of Motion.

Range of motion, $\mathfrak{R}$, is an indicator of performance and form that changes during the session.

Baseline. To address the challenge of obtaining ground truth for the range of motion of the user's arm, we devised a tool to obtain the lifting/lowering angle from the captured video. The camera was placed at the same height as the user's torso, close to their head. In this position, the arms were parallel to the camera plane, as shown in Figure 5b. We randomly selected 46 reps across 5 people. In each of the 46 reps, we identified 10 uniformly distributed frames, and obtained image coordinates for a marker on the shoulder and the elbow in each frame. We used these coordinates to compute the Euler angle of the arm in the image plane, and use the values as ground truth. To ensure that our tool was accurate, we validated it against a 10 camera motion capture system (Vicon). Over all, we were able to acquire 460 angle measurements for ground truth, over 46 reps. The samples were selected for both concentric and eccentric phases which varied between first and last set/rep.

Result. We compute two errors for validating our range of motion measurements.
The first is the error in the measurement of θ, treating each ground truth sample as an independent measurement in time. For this we compute the error between ground truth and LiftRight's measurement for θ at each of the 460 points. Figure 8 shows the CDF for this individual error. We see that despite the fast movement of the human arm, most values are within sub 7° of error. Similarly, we compute the error in lifting angle by comparing the range of motion computed from the ground truth and that returned by LiftRight. These are the first and last θ measurements from each of the 46 reps. Figure 9 shows the error for range of motion measurements, where most of the values fall within 9°. These results are comparable with those reported in prior work by Shen et al. (2016), where the user's hand with the inertial sensor was undergoing slower movements.

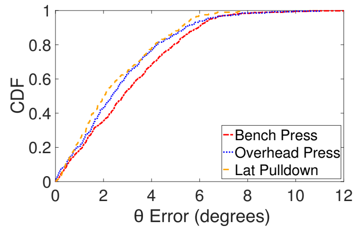


Fig. 8: Error in computing θ.

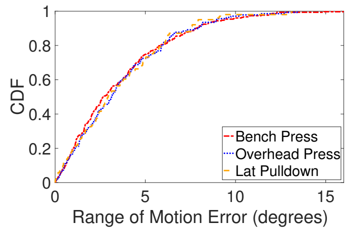


Fig. 9: Error in computing range of motion.

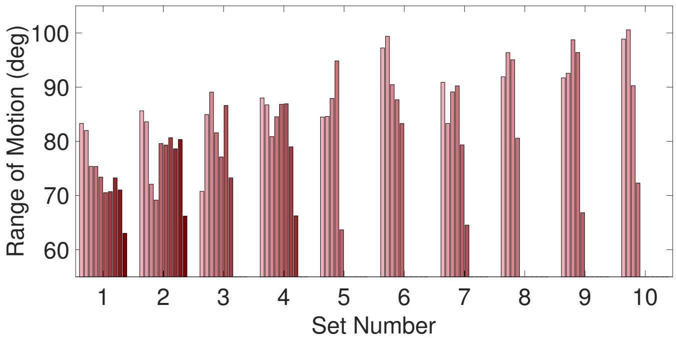


Fig. 10: Range of motion of a single participant for one session of bench press. We see that $\mathfrak{R}$ deteriorates towards the end of each set as the fatigue sets in. Shades of red represent reps in each set.

Observations. $\mathfrak{R}$ exhibits a small variation from set to set. Figure 10 shows how a single volunteer's range of motion reduces within a set as they get tired, leading to a decrease over ten sets. This is an indication of poor form over time, and can be used by a human trainer or a sophisticated software to provide feedback to the user.

6.4. Velocity

Baseline. We use the baseline measurements obtained for range of motion and delay timing, as discussed in the earlier subsection, to obtain ground truth values for velocity.

Result. Figure 11 shows LiftRight's performance in measuring the lifting/lowering velocity. Median velocity errors resemble errors obtained in range of motion. Measuring velocity during quick rapid movements with sudden stops is a challenging problem, and not many previous work have addressed it. Since the arm covers a range of about 100° in each rep, we consider a median error of 5° – 10° reasonable.

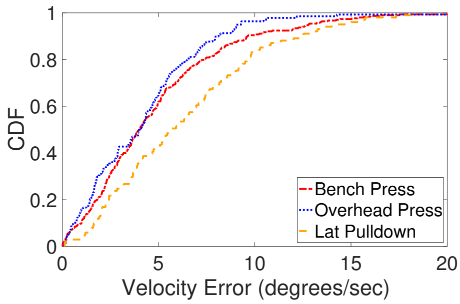
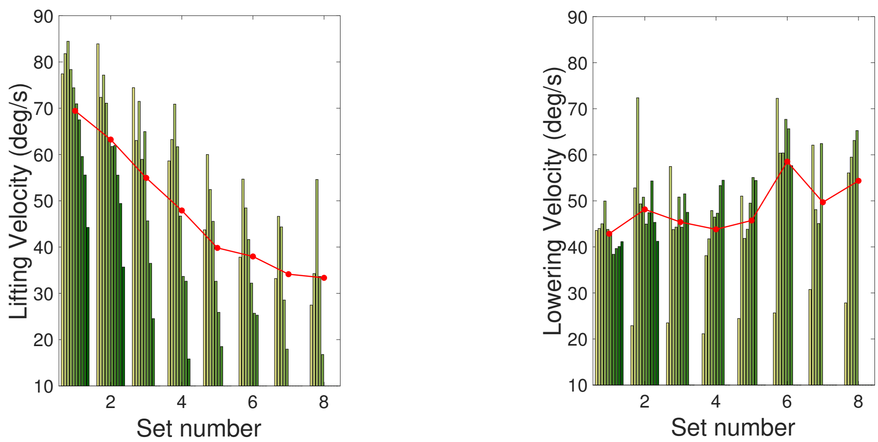


Fig. 11: LiftRight's performance for lifting/lowering velocity measurement.

Observations. Intuitively, as participants proceed through a set and perform reps, the same amount of effort will cause a decline in velocity, possibly due to fatigue. Figure 12a and 12b display the velocity in each set for a single bench press session. Light to dark shades of green represent reps in a set. We can see in Figure 12a that in each set, the lifting velocity decreases as more reps are performed, leading to a total decrease by 52% in the mean velocity over the workout. Similarly, Figure 12b shows that lowering velocity increases over sets by 23.76%. As more energy is spent, participants are more likely to let gravity assist them in lowering the weights.



(a) Lifting Velocity

(b) Lowering Velocity

Fig. 12: Velocity observed in a single bench press workout session. Shades of green represent reps in each set. The red line marks mean velocity in each set. Mean lifting velocity for each set is seen to decline during a session, while mean lowering velocity rises.

Figure 13 aggregates the information from all participants for bench press data. We can see that for each volunteer the average lifting velocity for each workout session improves over the duration of our data collection, which is consistent with literature (Bryant, 2013). Note that Candidate 2 increased the resistance level after day one, leading to a decrease of 38% in lifting velocity during day two, which was eventually gained back by the completion of the training program.

6.5. *Qualitative Metrics*

The metrics presented in this section are hard to quantify and existing tools cannot be used to accurately measure or validate them. Thus, we present our observations of these metrics as qualitative metrics, that can be combined with other parameters to generate feedback on user performance.

6.5.1. *Elbow stability angle*

Figure 14a and 14b show the change in elbow stability angle, $\mathcal{E}$, for two separate bench press sessions for the same candidate. The solid blue line indicates the mean $\mathcal{E}$ for each set. In bench press one can emphasize their pecs more if the elbows are away from the sides (Hatfield, 1993). Since $\mathcal{E}$ measures $\Delta\phi$, higher values of $\mathcal{E}$ represent larger movement of the elbows inward while lifting. While small movement is acceptable, higher values indicate bad/inconsistent form through poor stabilization. For the same resistance load, Figure 14a shows higher values of $\mathcal{E}$, averaging to 15° compared to values measured two sessions later, shown in Figure 14b, averaging to 8°. In session 4 for the same candidate, the $\mathcal{E}$ values are lower and exhibit less variance. We posit that this indicates that, for this participant, stability improves over time.

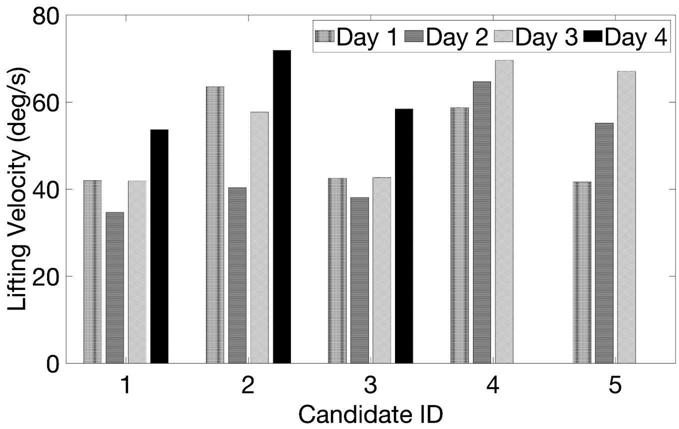
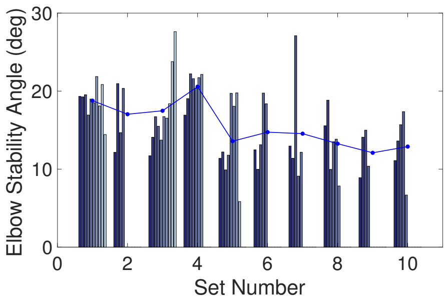
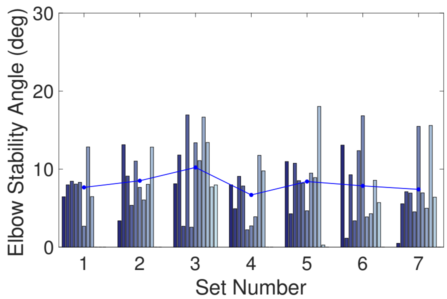


Fig. 13: Lifting velocity for all days of bench press data collection. For most volunteers, lifting velocity improves over the duration of data collection. Candidate 2 started with a lighter load on Day 1, but increased the load for subsequent sessions.



(a) 2^{nd} bench press session.



(b) 4^{th} bench press session.

Fig. 14: Change in elbow stability angle between the 2^{nd} and 4^{th} bench session of participant 2. Shades of blue represent reps in each set.

6.5.2. Occurrence of sticking points

Through observations, trainers identify an athlete's sticking points. In our system, sticking points and other points of struggle can be identified from the Lifting/Lowering Angle, θ, as shown in Figure 15. LiftRight identifies sticking points by computing the first-order gradient of the trace, and looking for deceleration during the lift phase of each rep. Unfortunately, there are no existing measures for verifying the occurrence of a sticking point. The most popular way of determining the occurrence of sticking points is through manual observation. Therefore, the ground truth was obtained via manual observation of the video. Compared to our observation, our algorithm could detect 85% of all sticking points.

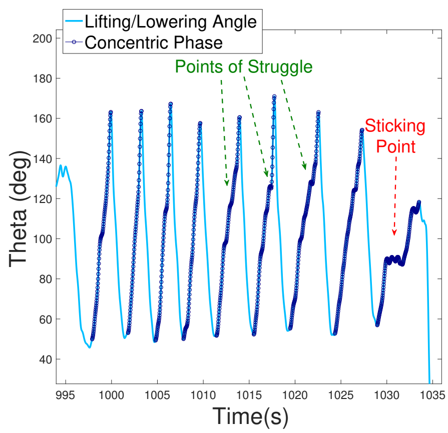


Fig. 15: Sticking point and points of struggle over a set of Bench Press.

6.6. Sensor Positioning Rationale

LiftRight focuses on tracking motion via a single inertial sensor placed on the body, that allows it to monitor common upper body exercises. Choosing this placement, however, can be a challenging task, as quantitative data that can be captured is location specific. In the scope of this work, capturing common upper body exercises, we find that the following two parts of the limbs are the most involved: upper arm and forearm. To examine the efficacy of each mounting location, we attached a similar sensor to the user's upper arm and forearm. We computed the range of motion captured via each sensor, and compared it to our ground truth. We found that the forearm mounting position does not capture the complete range of motion for our exercises.

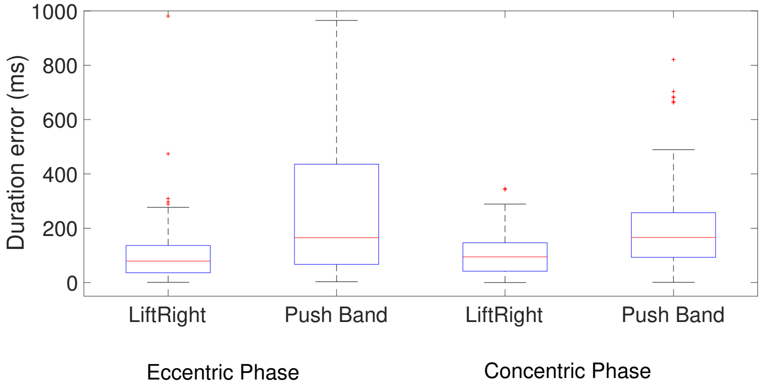


Fig. 16: Eccentric and Concentric phase duration error between LiftRight and the Push Band

Moreover, an upper arm sensor mounting position will allow us to monitor and understand popular exercises like the bench press, bent over rows, flye, over head presses, pull ups and downs, and chin ups; allowing us to quantify performance when working out the deltoids, biceps, triceps, pectorals, upper and middle back. A similar list might not be as easily created when the sensor positioning changes to the forearm, hands, or chest regions, given that complex, full body range workouts are not captured.

6.7. Comparison with Push Band

We compare LiftRight with a commercial state-of-the-art tracker that monitors performance for strength training - the PushBand (2017). One of the biggest advantages LiftRight has over Push Band is that LiftRight automatically segments user's sets and reps. The Push Band, on the other hand, needs the user to indicate the start and stop of each set manually. Forgetting to indicate it even once can lead to incorrect estimations of workout statistics. Additionally, the Push Band does not compute other metrics that LiftRight offers, such as the range of motion, elbow stability, recovery times, and sticking points. There are, however, several similarities between LiftRight and Push Band. First, both systems rely on the user to state what exercise they are doing. Second, they both provide a way to capture average concentric velocity and exercise phase duration. For the former, our system captures the results in deg/s while the Push Band captures in m/s. The latter however, is measured in $milliseconds$ by both. Therefore, we compare actual differences in duration of each phase based on ground truth from our videos.

We compare against the Push Band for 9 workout sessions for the three mentioned exercises. For a total of 339 reps, both LiftRight and the Push Band miscounted 2 repetitions. Without relying on any set start/stop information provided by the user, this leads LiftRight to less than 0.06% error compared to manually segmented data. It's worth noting that during real time feedback, the Push Band over counted by 2 reps, which were later removed by the software upon set completion. While recall and precision are very similar between the two systems, eccentric and concentric rep duration are not. The results are shown in Figure 16. The mean duration error for LiftRight in concentric and eccentric phases was 94 ms and 79 ms, respectively. This is nearly doubled by the Push Band, whose average error is 166 ms and 165 ms. The reason this difference is important to highlight is because most metrics rely on accurate rep segmentation. Incorrect estimates of rep/phase duration can render other metrics useless. We also compared the trend for velocity estimates provided by LiftRight and Push, and found them to be highly correlated. Based on the results, it is clear that LiftRight outperforms the Push Band.

7. Related Work

Previous research has shown that a popular way of quantifying performance in strength training is through modified gym environments and equipment. One example is shown by Drinkwater et al. (2007), who use an optical encoder to measure the range of motion exerted on the bench press machine. The aim is to monitor the distance and the speed at which the bar travels to characterize how someone is performing the eccentric and concentric phases. An alternative to this is explained by Elliott et al. (1989), where cinematography captures alike metrics. The underlying difference between the two solutions being the needed, yet expensive, modification to the gym to support the implementation. Both however, require a significant number of manual steps to process and capture exercise data. Infrastructure changes have also triggered additional work, such as exercise detection and classification. One example is demonstrated by Ding et al. (2015) who use RFID technology as a mechanism to extrapolate wireless signals towards exercise identification and segmentation. While the general idea in solving the problem is fascinating, it is unclear if the benefits of exercise feature extraction match the high overhead of infrastructure support.

Wearable solutions also exist, dating back to 2013, where two wrist worn inertial measurement units are used for tracking motion of an exercise (Ochoa, 2013). Aside from counting repetitions, they monitor upper limb motion symmetry, over extension, and speed. The ultimate benefits of using such a system are apparent, both providing the user with a way to monitor improvement and potentially retard injury. However, the segmentation strategy for identifying when sets and repetitions occurred is missing, alike the accuracy of the aforementioned metrics. Work by Chang et al. (2007) can be viewed as a partial improvement on the former pitfall, by introducing a windowed segmentation strategy. However, the byproduct of this solution yields a limited assortment of useful metrics, being repetition counting and identification of performed exercises. Alone, these results do not provide adequate amount of information in determining how a user can improve. Other work also introduces alternative tactics, as found in Velloso et al. (2013) and Pernek et al. (2012), by showcasing the power of pattern matching, either originating from the user (self training the system) or an expert. Based on our observations, we find that such solutions are easily corrupted by behavior occurring outside of the exercise set window. In other words, only when the person remains idle while at rest can high efficiency be reached. However, such conditions are uncommon outside of lab environments. Rontu et al. (2010) address the second pitfall, accurate metrics, by measuring the relationship between inertial acceleration and performance. Specifically, they used an IMU to capture the correlation between linear acceleration and sub-maximal weight lifts. This of course being a fine grained metric that classifies the quality of the workout further than repetition counting. Regardless of the proposed solution, most wearable options tend to lack automation, form detection, and ease of implementation to have a justifiable use case for the general public. In addition, very few provide an analysis as to how accurate and efficient their system is.

The commercial side of research also contributes to mobile sensing, usually embedded within proprietary hardware and software packages. Some examples include the Fitbod (Fitbod, 2016), BioStrap (BioStrap, 2016), and the Push Band (PushBand, 2017). Out of the three, Fitbod has the most amount of manual engagement with the user, in which it requires data entry for which exercise was performed, along with the number of sets and reps completed. This leaves the automated footprint very small as validation between user notes and body kinematics during exercise is non existent. To some extent, BioStrap and the Push Band shorten the manual boundaries of tools like Fitbod through the use of machine learning algorithms. Aside from exercise detection, their goal is to provide many metrics, including but not limited to: lifting velocity, energy expenditure, and set / rep counting. The goal being to require little to no human guidance. On the contrary, we find that solutions that follow this problem solving pattern tend to seek advice from the user to know when the set has commenced in an attempt to decrease segmentation recall error. In addition we find that most provide a limited assortment of metrics, and come with additional requirements such as: subscription fees, multi sensor chaining for fine grained analysis, and reliance on clean training data for accurate exercise detection.

Lastly, the research community actively promotes fatigue detection and injury prevention solutions, such as the work by Mokaya et al. (2016) and Manero et al. (2016). The former monitors muscular vibrations via inertial sensors and the latter, muscular activity via EKG to estimate when exercise should stop. Considering that weight training has been shown to cause injuries, monitoring and capturing this estimation is proven vital by Malina (2006). While we agree with the authors that exercise injuries should be monitored, we believe that there are many unexplored features within each rep that can contribute to more accurate injury prevention. Other research focuses on injury prevention in outdoor scenarios, such as pedestrian safety (Jain & Gruteser, 2018; Jain et al., 2015, 2014).

In summary, most work related to our problem space is solved through infrastructure modification, wearable devices that rely on machine learning, or simply the use of manual data entry tools. In our view, the first solution lacks scalability and application within the general public due to environmental and financial requirements. Machine learning approaches, while a great solution to a multitude of domains, rely on clean classification data to function correctly. Considering most tools tend to rely on this approach when all else fails, we believe that there might be other potential ways to solve the problem, without relying on any knowledge that defines what is right. Lastly, while manual data entry tools are a great way to introduce the concept of workout journals in the digital space, their lack of user interaction and performance analysis leave it as a last resort in today's automated world. Current solutions provide limited accuracy specification and rely on the user to help with exercise data segmentation.

8. Discussion

Strength training is a vast field, and this work only covers fundamental metrics for performance and form analysis. We discuss some of the limitations of LiftRight and potential ideas for future work.

Scaling to other upper body exercises: The upper arm mounting position of the sensor can be easily extended to other upper body exercises such as upright rows, chin-ups, inclined flye, barbell high pull, lateral rise, among many others. It does, however, have its limitations in detecting few upper body exercises, such as the bicep curls, due to limited or no motion of the upper arm.

Value to users and trainers. Many factors affect exercise performance. In this work, we have quantified the core factors that find use in the design of training programs. These values can help both, users and trainers, to make informed decisions regarding their performance, form, limitations, and overall fitness. They can be plugged into existing training plans to determine improvement or change over time and provide feedback.

Scaling to different equipment and free-weights: In the current scope of the project we have limited ourselves to the exercise equipment described in Section 5. We believe that LiftRight can be extended to other equipment or free-weights since the underlying motion doesn't change. This was confirmed with our use of the conventional bench and Lat Pullown machine. We will validate the scalability to different equipment in the future.

Future work: In the future, we can provide a platform to track how those needing physical therapy are recovering, i.e. their form, exercise patterns, intensity of exercise, etc. LiftRight could also be extended by implementing existing exercise activity detection algorithms to eliminate the need to ask the user which exercise is being performed. Furthermore, in the future, we aim to understand the relationship and dependencies between various quantitative and qualitative metrics, and their joint effect on injuries.

9. Conclusion

In this paper, we presented the design, implementation, and evaluation of LiftRight, wherein we demonstrate the feasibility of performing fine-grained exercise segmentation and monitoring using a single wearable sensor. We validated our techniques on real-world data from 8 participants for three weight lifting exercises. In building this system we have developed the underlying tools for automatically verifying years of observations made by power lifters and fitness trainers. To the best of our knowledge, LiftRight is the only system that performs automated workout session segmentation down to the granularity of the exercise phase, with 96% accuracy in rep identification and little to no delay in detecting each exercise phase. We believe LiftRight is an important step in monitoring not only exercise form and performance, but towards safer weight lifting, injury prevention, and general well being.

References

Apple (2016). Apple watch series 3. https://www.apple.com/watch/.

BioStrap (2016). Biostrap. http://bit.ly/2KbwUsG.

Blanco, J.-L. (2010). A tutorial on se (3) transformation parameterizations and on-manifold optimization. *University of Malaga, Tech. Rep*, *3*.

Bryant, J. (2013). *Bench Press: The Science*. BookBaby. URL: https://books.google.com/books?id=UvJWDQAAQBAJ.

Brzycki, M. (1993). Strength testingâĂŤpredicting a one-rep max from reps-to-fatigue. *Journal of Physical Education, Recreation & Dance*, .

Chang, K.-H., Chen, M. Y., & Canny, J. (2007). Tracking free-weight exercises. In *Proceedings of the 9th International Conference on Ubiquitous Computing* UbiComp '07. Berlin, Heidelberg: Springer-Verlag.

DHHS/PHS (2000). *Promoting Better Health for Young People Through Physical Activity and Sports. a Report to the President from the Secretary of Health and Human Services and the Secretary of Education. Fall 2000.* CCC Healthy Youth.

Ding, H., Shangguan, L., Yang, Z., Han, J., Zhou, Z., Yang, P., Xi, W., & Zhao, J. (2015). Femo: A platform for free-weight exercise monitoring with rfids. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems* (pp. 141–154). ACM.

Drinkwater, E. J., Galna, B., McKenna, M. J., Hunt, P. H., & Pyne, D. B. (2007). Validation of an optical encoder during free weight resistance movements and analysis of bench press sticking point power during fatigue. *Journal of Strength and Conditioning Research*, .

Elliott, B. C., Wilson, G. J., & Kerr, G. K. (1989). A biomechanical analysis of the sticking region in the bench press. *Med Sci Sports Exerc*, .

Fitbit (2016). Fitbit. https://www.fitbit.com/home.

Fitbod (2016). Fitbod. https://www.fitbod.me.

Fleck, S. J., & Kraemer, W. J. (2014). *Designing Resistance Trainin Programs*. Human Kinetics.

Garmin (2016). Garmin. http://bit.ly/2Ar2vXC.

Hatfield, F. (1993). *Hardcore bodybuilding: a scientific approach*. McGraw-Hill.

IPF (2016). *International Powerlifting Federation Technical Rules Book*. IPF.

Jain, S., Borgiattino, C., Ren, Y., Gruteser, M., & Chen, Y. (2014). On the limits of positioning-based pedestrian risk awareness. In *MARS 2014* MARS '14 (pp. 23–28). New York, NY, USA: ACM. URL: http://doi.acm.org/10.1145/2609829.2609834. doi:10.1145/2609829.2609834.

Jain, S., Borgiattino, C., Ren, Y., Gruteser, M., Chen, Y., & Chiasserini, C. F. (2015). Lookup: Enabling pedestrian safety services via shoe sensing. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services* MobiSys '15 (pp. 257–271). New York, NY, USA: ACM. URL: http://doi.acm.org/10.1145/2742647.2742669. doi:10.1145/2742647.2742669.

Jain, S., & Gruteser, M. (2018). Recognizing textures with mobile cameras for pedestrian safety applications. *IEEE Transactions on Mobile Computing*, (pp. 1–1). doi:10.1109/TMC.2018.2868659.

Lin, J., Keogh, E., Lonardi, S., & Chiu, B. (2003). A symbolic representation of time series, with implications for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery* (pp. 2–11). ACM.

Lin, J., Keogh, E., Wei, L., & Lonardi, S. (2007). Experiencing sax: a novel symbolic representation of time series.

Malina, R. M. (2006). Weight training in youth-growth, maturation, and safety: an evidence-based review. *Clinical Journal of Sport Medicine*, .

Manero, R. R., Shafti, A., Michael, B., Grewal, J., Fernández, J. L. R., Althoefer, K., & Howard, M. J. (2016). Wearable embroidered muscle activity sensing device for the human upper leg. In *IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC),*. IEEE.

Matthews, M. (2012). *CARDIO SUCKS: The Simple Science of Losing Fat Fast... Not Muscle*. Waterbury Publishers.

Matthews, M. (2014). *Beyond Bigger Leaner Stronger: The Advanced Guide to Building Muscle, Staying Lean, and Getting Strong*. Oculus Publishers.

MbientLab (2016). Mbientlab - smart wireless sensors. https://mbientlab.com/.

Mokaya, F., Lucas, R., Noh, H. Y., & Zhang, P. (2016). Burnout: a wearable system for unobtrusive skeletal muscle fatigue estimation. In *Proceedings of the 15th International Conference on Information Processing in Sensor Networks* (p. 8). IEEE Press.

Ochoa, G. G. (2013). Development of a wearable free-weight exercise assistant.

Padulo, J., Mignogna, P., Mignardi, S., Tonni, F., & D'Ottavio, S. (2012). Effect of different pushing speeds on bench press. *Int J Sports Med*, . doi:10.1055/s-0031-1299702.

Pernek, I., Hummel, K. A., & Kokol, P. (2012). Exercise repetition detection for resistance training based on smartphones. *Personal and Ubiquitous Computing*, *17*, 771–782.

Pinto, R. S., Gomes, N., Radaelli, R., Botton, C. E., Brown, L. E., & Bottaro, M. (2012). Effect of range of motion on muscle strength and thickness. *The Journal of Strength & Conditioning Research*, .

PushBand (2017). The push band. https://www.trainwithpush.com/push-band.

Rontu, J.-P., Hannula, M. I., Leskinen, S., Linnamo, V., & Salmi, J. A. (2010). One-repetition maximum bench press performance estimated with a new accelerometer method. *The Journal of Strength & Conditioning Research*, .

Samsung (2016). Samsung gear fit2. http://bit.ly/2vtyTmX.

Scully, D., Kremer, J., Meade, M. M., Graham, R., & Dudgeon, K. (1998). Physical exercise and psychological well being: a critical review. *British journal of sports medicine*, *32*, 111–120.

Shen, S., Wang, H., & Roy Choudhury, R. (2016). I am a smartwatch and i can track my user's arm. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services* MobiSys '16 (pp. 85–96). New York, NY, USA: ACM. URL: http://doi.acm.org/10.1145/2906388.2906407. doi:10.1145/2906388.2906407.

Velloso, E., Bulling, A., & Gellersen, H.-W. (2013). Motionma: motion modelling and analysis by demonstration. In *CHI*.

Vicon (). Biomechanics and sport: Our worldwide user base pushes our research performance to the gold standard. Available at https://www.vicon.com/motion-capture/biomechanics-and-sport (2019/04/11).

Westcott, W. L. (). Resistance training is medicine: effects of strength training on health. *Current sports medicine reports*, .

Winett, R. A., & Carpinelli, R. N. (2001). Potential health-related benefits of resistance training. *Preventive Medicine*, *33*, 503 – 513. URL: http://bit.ly/2M1FN9T. doi:https://doi.org/10.1006/pmed.2001.0909.