Performance Monitoring for Exercise Movements using Mobile Cameras

Morteza Mohammady Gharasuie
Old Dominion University

Noah Jennings
Old Dominion University

Shubham Jain
Stony Brook University

ABSTRACT

Despite numerous devices targeted to fitness tracking, the strength training domain has often been overlooked and understudied. In this paper, we propose a smartphone camera based approach to track users' strength training workouts, as well as metrics pertaining to their form and performance. Our goal is to detect the repetitions in a workout without requiring user intervention or any training data from the user. Unlike many existing systems, our proposed system is scalable, low-cost, and widely accessible. We gather data from two sources for 5 exercises across 25 subjects. We compute performance metrics such as range of motion, velocity, and duration from each repetition with median errors less than 10%. These results demonstrate that commercial off the shelf smartphone cameras can be used to accurately detect and count repetitions in user movements, as well as to compute rep-by-rep user performance.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in ubiquitous and mobile computing**;

KEYWORDS

Fitness monitoring, Mobile camera, Early event detection

ACM Reference Format:
Morteza Mohammady Gharasuie, Noah Jennings, and Shubham Jain. 2021. Performance Monitoring for Exercise Movements using Mobile Cameras. In *ACM Workshop on Body Centric Computing Systems (BodySys 21), June 24, 2021, Virtual, WI, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3469260.3469665

1 INTRODUCTION

Pose estimation is an active area of computer vision research that focuses on understanding posture and movement in images and videos. In addition to applications in action recognition, gaming industry, and virtual environments [19], it has been used for analyzing player movements [9, 31] in team sports and personal workout sessions. Tracking of posture during exercise and physical workouts has the potential to monitor and improve one's overall performance [2, 6, 21, 25]. Unfortunately, the number of injuries in the fitness domain, particularly in weight training, have been rising [1, 16]. We believe that these injuries can be alleviated with improved surveillance and timely feedback. Typically, coaches can monitor a user's movements during exercise and provide corrective feedback. However, instructors are neither affordable nor accessible for everyone. Existing approaches for automated performance monitoring require additional equipment like wearable inertial sensors [2, 11, 21, 25, 26]. However, these additional sensors are expensive, and can add to the user's cognitive load by requiring them to carry and charge an additional device. Techniques using wireless signals [12] are often affected by the environment.

In this paper, we implement and validate state of the art pose estimation techniques in the domain of physical activity and exercise monitoring in uncontrolled environments. Our goal is to build scalable solutions to make the proposed technology widely accessible and to allow smart device users to monitor their movement-specific progress. However, exercise involves targeted movements and tracking small body movements in real-world environment poses significant challenges. Most existing attempts at monitoring fine-grained movement have focused on using multi-camera motion capture (MoCap) systems [31]. Such infrastructure based approaches are not scalable and do not allow average users to receive metrics pertaining to their form and performance. To address this challenge, we focus on using cameras on users' smartphones and laptops. The large-scale availability of these devices makes our approach accessible to a wider population. We target the strength training space to design and evaluate a repetition detection system. The ability to detect temporal events as they happen in a real-time video stream is the cornerstone of bridging the gap between theoretical vision techniques and real world applicability.

We propose a performance monitoring technique that relies on pose estimation to track user movements during weight training. A video stream is captured using commercial off-the-shelf (COTS) device with a camera (laptop or smartphone). The correctness of athletic movements is often assessed by the athlete themselves, or those around them. However, there is little agreement on the criteria to determine correctness. There are, however, widely accepted practices that are followed in the weight training community. Therefore, *rather than determining correctness of form, we focus on computing movement related metrics that can then be used to determine variations in a user's form over time.* We detect repetitions as a user exercises, and for each repetition we compute the range of motion, duration, and velocity. Additionally, we analyze the variation in rest times as an indirect measure of fatigue. In building this system, *we design a low-cost scalable approach that is capable of allowing users to track their performance.* Our proposed technique can also be enhanced to provide feedback to the user in real-time.

BodySys 21, June 24, 2021, Virtual, WI, USA

ACM ISBN 978-1-4503-8600-5/21/06…$15.00
https://doi.org/10.1145/3469260.3469665

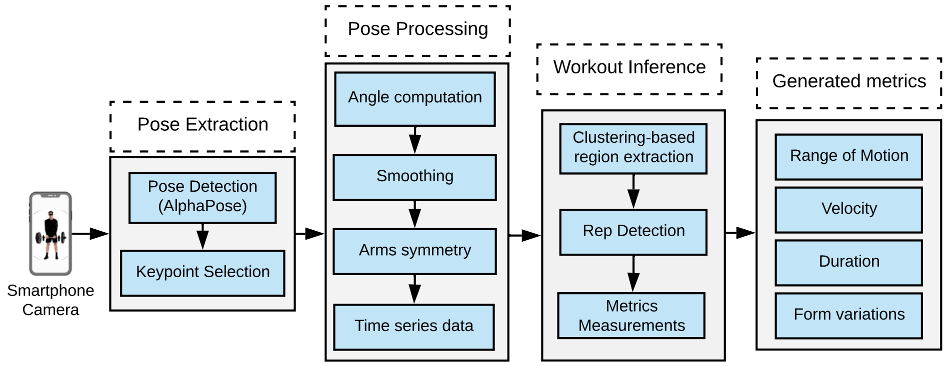


Figure 1: System Overview

2 BACKGROUND

We present terminology used in the athletic community, prior research [4, 20, 21, 28], and throughout this paper.

Strength training workouts. Strength training, also known as weight training or resistance training, is a popular form of physical activity designed to target specific muscle groups by working against resistance stemming from free weights, weight machines, or body weight. Controlled body mechanics and correct posture are an important part of these workouts. If performed incorrectly, these movements can lead to injuries over time [16, 28].

Repetitions. In the context of athletic workouts, a repetition or a rep, is one complete motion of an exercise and consists of a concentric phase and an eccentric phase. Multiple repetitions performed consecutively is known as a *set*, multiple sets of the same type of motion is an *exercise*, and multiple exercises form a *workout session*.

Concentric Phase. It is the positive phase of a repetition where the athlete's major muscles contract [24]. This portion of many exercises tends to be in the opposite direction of gravity, and therefore we refer to it as *lifting* phase in the rest of this paper.

Eccentric Phase. It is the negative phase of a repetition where the athlete's muscles lengthen or relax. This portion of many exercises tends to be in the same direction as gravity, and therefore referred to as the *lowering* phase in this paper.

Pose Estimation. Human pose estimation is a popular computer vision technique to identify key points (often the human joints) in a person's body captured in an image. Typically, these algorithms can track the shoulders, elbows, wrists, knees, and ankles. Pose estimation is a useful tool in analyzing the posture and movements of a user, enabling us to compute quantitative metrics.

3 SYSTEM DESIGN

The overall system design is presented in Figure 1. It consists of three components: Pose Extraction, Pose Processing, and Workout Inference. The details regarding each component are discussed in the following subsections.

3.1 Pose Extraction

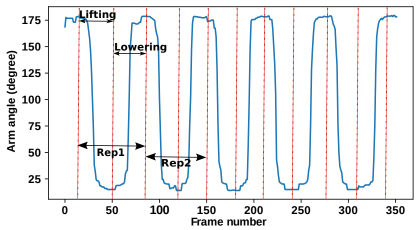
The incoming video stream from camera, smartphone or webcam is provided to the pose extraction module. Open source pose estimation models such as AlphaPose [8], OpenPose [5], and DeepPose [30] were tested, wherein AlphaPose was found to be most accurate and was thus used in our research. Alphapose [8] is a state of the art pose estimation technique that processes each incoming frame from the video to localize human joints, also known as keypoints. The keypoints returned by the pose estimation algorithm are nose, ears, eyes, shoulders, elbows, wrists, hips, knees, and ankle joints. Every keypoint is represented as a point (x, y) in the frame's 2D coordinate system. The keypoints are filtered to retain the shoulder, elbow, wrist and hip joints, in keeping with our goal of processing upper body movement mechanics. These keypoints are processed in the next step to get a series of spatio-temporal values for user arm motion in each frame.

3.2 Pose Processing

After extracting the keypoints in the previous module, we compute the angle between the upper arm and the lower arm for all exercises except lateral raise. Due to inherent noise in pose estimation algorithms, there are jitters in keypoint detection for consecutive frames, often caused by lighting or clothing. Even when the user is not moving, wrist detection, for example, in consecutive frames may be displaced by a few pixels. To ensure that our algorithm is robust to this jitter, we discretize the raw angles and smooth our observations by using a median filter. We analyze the variations of this angle in both arms to draw inferences regarding the user's movements.

3.3 Workout Inference

In the third module, the time series data is available as a sequence of angles that represent the orientation and the position of the arm in each camera frame. Our first task is to detect repetitions from this sequence of arm angles. Because exercises contain repetitive movements, performed one after another, the angles exhibit a pattern in this time series data. This is illustrated in Figure 2a for the bicep curls exercise (shown in Figure 2b). The dashed red lines in Figure 2a show the beginning and end of each phase. For each movement during bicep curls, the user's repetition starts with the arm almost straight down by their side, and they start curling it up, bending at the elbow. At the highest point, the wrist is close to the shoulder and the arm angle measurement now drops to a few degrees. This is the end of the lifting phase. When the user brings their arm back to the starting position, we mark a repetition complete.



(a) Arm angle from one set of bicep curl exercise.

(b) Angle measurement.

Figure 2: Pattern in the range of angle in one set of the bicep curl exercise. The angle is measured using shoulder, elbow and wrist points. The red dashed lines show the beginning and end of lowering and lifting phases.

To detect these repetitions, we focus on identifying phase transitions in Figure 2a. These can be seen as the positive and negative peaks in the figure. A pair of phases constitutes one repetition. A clustering method [18] is used to identify phase transition regions. We define an empirically derived threshold for the arm angle, that is different for each exercise, since each has a different movement and range of motion. When the angle values exceed the threshold, the data samples between the last pair of threshold crossings are passed to the k-means clustering algorithm. Rather than clustering the arm angle, we use the y-coordinate value of the wrist locations. This is designed to capture the spatial aspect of the user's motion. We define the number of initial centers k as 3. Having three clusters separates the angle values in three regions. The cluster that has a center with the biggest or smallest y-axis value is detected as the region where a phase transition occurred. Each alternate transition marks the end of the previous rep and beginning of the next one. For rapid movements, these transitions or peaks are observed to be somewhat sharp. However, for slow movements where a user spends some time at the top or bottom of each phase, the detected curve shows somewhat flat peaks. Each phase is therefore represented by a start time and an end time. The detected reps are then used to compute relevant metrics.

3.4 Workout Metrics

Based on previous research and literature survey [3, 4, 13, 21], we select three quantifiable metrics to be computed for each detected repetition: range of motion (R), duration, and velocity. These metrics are used to evaluate the algorithm's ability in measuring the metrics, and monitoring variations in the metrics' values during a set or a session. R is defined as the distance moved by the weight during each repetition. To maintain scalability of our solution by avoiding camera calibration, we use pixel based distance for R. The velocity V is calculated as: $V = \frac{R}{t}$, where t denotes the time taken to complete one movement or phase. Duration is another metric that is affected by the movement velocity. As a user fatigues, the lifting is often longer (slower) and lowering is shorter (faster). Therefore, detecting these metrics play a significant role in informing users about their motion, allowing them to adjust their movement in real-time.

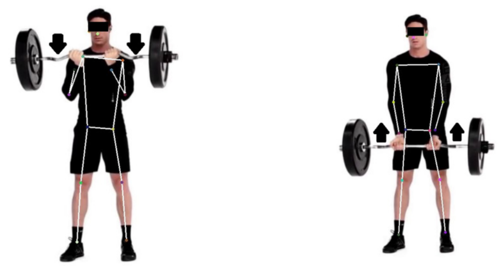


Figure 3: Demostration of the bicep curls exercise. The images show the user's pose at the end of lifting phase (left) and end of lowering phase (right).

4 EVALUATION

We evaluate our system by answering the following questions:

- How well can the system detect and count repetitions?
- How accurately can the system measure the range of motion, duration, and velocity?
- How early can the system detect a repetition before it ends?
- Is the system capable of capturing variations in form?

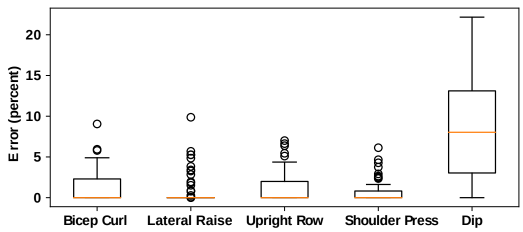
Data Collection. In order to study the performance of the proposed system and answer the above questions, we collected data for various upper body exercises: bicep curls, shoulder press, tricep dips, upright rows, and lateral raise. One of these exercises is shown in Figure 3 with the pose estimation output superimposed on them. We collected data by recording front view of participants, using a mobile camera, when they were exercising at their gym. This allowed us to take the real environment into account when designing our algorithm. We also compiled additional data by gathering videos from the internet, particularly YouTube. These are videos posted by athletes as part of their training. Overall, we gathered data for a total of 25 subjects. Different exercise equipment were used in these videos. The exact number of sets and repetitions in each exercise is shown in Table 1. The ground truth for each video was labeled manually, by annotating keypoints of interest in subsequent frames.

4.1 Repetition Detection Performance

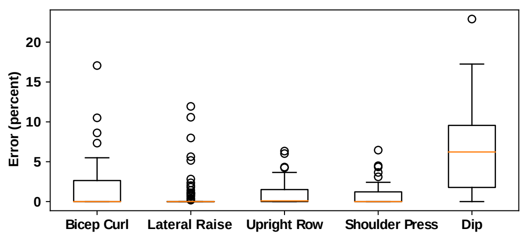
To answer the first question, we evaluate the repetition detection accuracy of the proposed system. The accuracy of system is shown in Table 1. Our algorithm detects 100% repetitions for 4 exercises. A slightly lower accuracy of 80.3% is observed for dips due to the higher diversity among people while performing the exercise movement. Observations from our data show that some people have very small range of movement in dips, making it more challenging for the algorithm.

Table 1: Detection accuracy of the proposed system

Exercise	Number of Sets	Number of Reps	Detected Reps	Detection Rate
Bicep Curl	10	54	54	100%
Lateral Raise	10	113	113	100%
Shoulder Press	10	73	73	100%
Upright Row	10	53	53	100%
Dip	8	61	49	80.3%

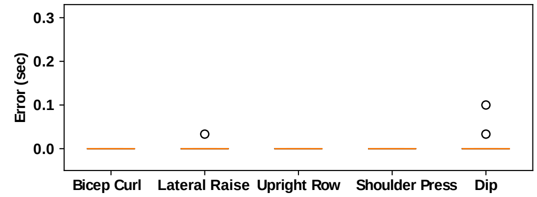


(a) Lifting Motion

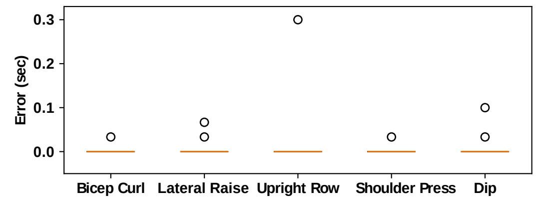


(b) Lowering Motion

Figure 4: Error in range of motion.

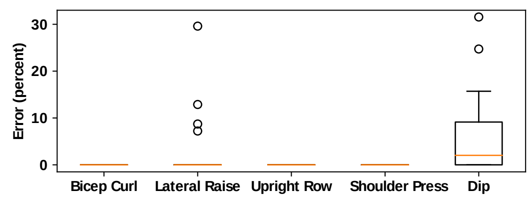


(a) Lifting Motion

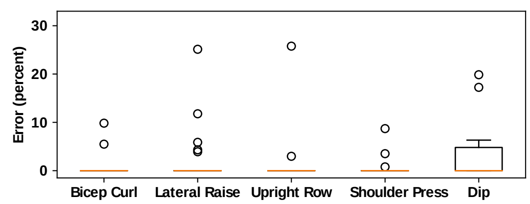


(b) Lowering Motion

Figure 5: Error in duration.



(a) Lifting Motion



(b) Lowering Motion

Figure 6: Error in velocity.

4.2 Metrics Evaluation

To evaluate the proposed system, we first focus on assessing the accuracy for each metric computed by the system. We identify repetitions and compute duration, range of motion (R), and velocity for each phase (lifting and lowering).

Range of Motion. The system performance in computing user's R is shown in Figure 4. While the R is computed in pixel, the error is shown as a percentage of the ground truth. This figure shows the percent error during the lifting and lowering phases of each exercise. We can see that for bicep curls, lateral raise, upright rows, as well as shoulder press, we observe errors between 0%-3% for lifting and lowering movements. We observe degraded performance in dips due to the increased variability in the motion, where median errors in R were observed to be around 9%.

Duration. The duration for each phase in each repetition is measured in number of frames, which is then converted to seconds based on the frame rate. From Figure 5 we can see that for both lifting and lowering motion, the duration measurements exhibit almost 0 seconds of error in most cases. We observe occasional outliers that exhibit errors around 0.1 seconds. This also indicates that our system was able to segment the data into repetitions and constituent phases precisely.

Velocity. Velocity is measured by the system in pixels/second. Figure 6 shows the percent error for velocity measurement in each exercise. Our system can compute velocity with less than 2% error for 4 exercises from our dataset. Due to the error in measuring range of motion during dips, we see higher errors, up to 15%, in velocity measurements.

4.3 Early repetition detection

The goal of early detection is to identify repetitions before they finish. By detecting repetitions as they happen, the system can potentially generate alerts and warnings in real-time, much like a human would. Figure 7 shows the CDF for early repetition detection. We can detect repetitions up to 1.5 seconds before they end. For users that perform movements slowly and pause in between

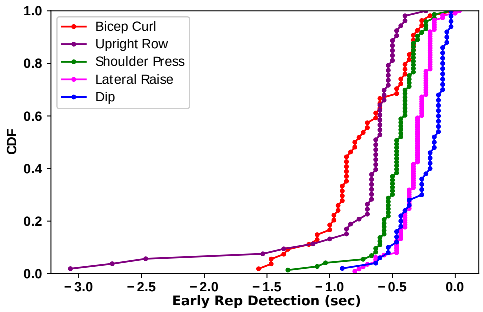


Figure 7: Early repetition detection by the proposed system. 0 on the x-axis indicates the time a rep ends. Early detection can support live feedback about a rep to allow a user to make adjustments to the next rep.

repetitions, our system can detect repetitions even 3 seconds before they end. Bicep curls and upright rows can be detected earlier compared to faster or smaller movements, such as those during lateral raise or dips. All repetitions are detected by the time they finish. This allows for computing and displaying the metrics for each repetition in real-time.

4.4 Observing user form and fatigue

The time in between two repetitions is referred to as rest or recovery time. We hypothesize that as a set progresses, users get fatigued, which may result in an increase in the duration of rest times. Figure 8 shows the rest time for 2 different users during 2 consecutive sets of lateral raise. For both, subjects 1 and 2, we can see that the rest times increase across each set. Subject 1 exhibits higher rest times than Subject 2. Indicators of fatigue, such as an increase in the rest times in between repetitions or even in between sets can be used to determine user performance over time. As fatigue sets in, user form is seen to deteriorate [21]. Strength training workouts emphasize on a controlled motion. Therefore, velocity, range of motion, and rest times combined can enable detecting when a user is getting fatigued, thus generating warnings that could potentially prevent injuries caused by deteriorated form.

Further, we use pose to determine the changes in user's form and posture over time. Figure 9 shows the positions of wrist and elbow keypoints for a user completing the lifting phase of a bicep curl. The output of the pose detection system is shown in red and yellow colors, for wrist and elbow respectively. Green and grey colors show the ground truth labels for the wrist and elbow positions. The ground truth is obtained by annotating the keypoints in each frame manually. We draw two key insights from this image. First, the detected keypoints overlap with the ground truth labels, thus showing that our pose detection algorithm works well. Second, we can see the variability of joint positions during an exercise. This is extremely important since higher variability can indicate incorrect form. As shown earlier [21] the stability of the elbow is an important factor in determining user form. We can incorporate graphical cues into our system to encourage users to keep certain joints stable - for example the elbow when performing bicep curls.

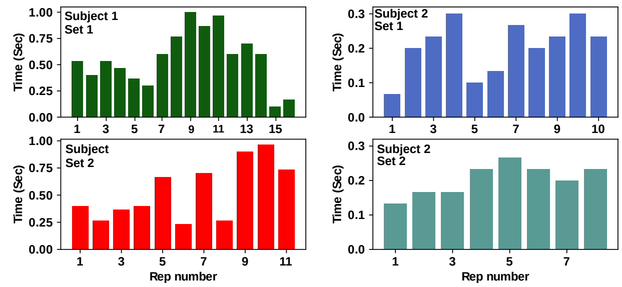


Figure 8: Measured rest times between repetitions during lateral raise exercise for 2 subjects. The second half of the graph shows increase in rest times - indicating the onset of fatigue.

Figure 9: Analyzing variation in wrist and elbow keypoints at the end of lifting phase for a bicep curl. Yellow and red points are output from pose detection algorithm, and grey and green colors are ground truth. Variations in joint positions can indicate user form over time.

5 RELATED WORK

Earlier approaches for tracking and monitoring exercises include wearable devices and computer vision, discussed here.

Wearable Devices. Sensors worn by the user such as an Inertial Measurement Units (IMU) or smart watches have been commonly used to collect and process the signal pertaining to user motion. Milanko et al. [21, 22] propose LiftRight for segmenting workout sessions and providing feedback on performance and form during exercise. They use an arm-mounted inertial sensor for tracking arm movements and computing associated metrics. Kwapisz et al. [15] use an accelerometer to recognize when the user starts a particular exercise and accurately classify patterns in body movements. Shen et al. [27] processed IMU data from a smart watch to classify arm and hand posture. They were able to estimate the location of the wrist within a few centimeters. Today, many commercial products such as [25], [2], and [10] have gained popularity in the domain of personal fitness. They combine the use of wearable sensors like smart watches and sensor processing to help the user plan, track, assess, and improve their athletic ability in real-time. Despite the advancements, wearable devices require a user to carry an additional device, making it inconvenient and expensive.

Computer Vision. Recent advancements in computer vision, particularly in deep learning, have resulted in a wide range of applications including fitness tracking. Levy et al. [17] propose a repetition detection system that uses convolution neural network to estimate the length of repetition in video. Recently, human pose estimation methods [5, 8, 30] have been used to detect joints in human body, which can be used to analysis motion during exercise. Pose trainer [7] is an end-to-end algorithm that uses OpenPose [5] human pose estimation framework to provide feedback on fitness exercise form. Gymcam [14] uses camera to track repetitive motions in a gym. It detects, identifies and tracks partially visible exercises. More recently, home gym systems [23, 29] use a combination of vision and human instructors to provide guidance to users. However, these systems are very expensive and not scalable. To the best of our knowledge, none of these works detect repetitions early or quantify the performance of each repetition. Our research fills this gap and can potentially enable real-time feedback to the user.

6 CONCLUSION AND FUTURE WORK

We have proposed a cost-effective and scalable solution that relies on commercial off-the-shelf mobile cameras and open source libraries to allow users to compute workout metrics in real-time. We demonstrate that we can accurately count the number of repetitions for 5 different exercises for data captured across 25 subjects. Our system can also detect repetitions early, before they are over, thus enabling potential feedback to the user in real-time. By providing the user access to various metrics related to their physical training, we hope our system can reinforce fitness tracking, prevent injuries, and aid in physical development at a lower cost.

Currently, the system does not detect the exercise being performed. Moreover, the system has not been evaluated with multiple users in the camera's view. In the future, we aim to increase the range of workouts to include lower body and full body exercises. We plan to broaden the set of metrics we learn from the user's workout such as the rate of force development. Our system will eventually incorporate audio feedback that is intended to supplement the user's performance, form, motivation, and experience, much like a personal trainer. This will further our efforts to create a more efficient weight training environment and correct form related issues in real-time.

Our system could eventually integrate wearable sensors for more extensive and accurate metric creation; while still housing the current methodology. A Bluetooth connection from the camera device to the wearable sensor could give us insights into the user's heart rate, energy consumption, etc. over time. We envision application being used not only by athletes, but also in physical therapy. Improving such an application could help accelerate a patient's recovery time and potentially prevent exercise-related injuries.

ACKNOWLEDGMENTS

This material is based in part upon work supported by the National Science Foundation under Award No. 2110193 and the National Institute of General Medical Sciences of the National Institutes of Health under Award Number T34GM118259. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Matching funds from Old Dominion University are also acknowledged.

REFERENCES

[1] 2018. CPSC NEISS. https://www.cpsc.gov/cgibin/NEISSQuery/home.aspx

[2] beast. 2016. Lift. Track. Improve. https://www.thisisbeast.com/en

[3] J. Bryant. 2013. *Bench Press: The Science*. BookBaby. https://books.google.com/books?id=UvJWDQAAQBAJ

[4] Matt Brzycki. 1993. Strength testing—predicting a one-rep max from reps-to-fatigue. *Journal of Physical Education, Recreation & Dance* (1993).

[5] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. 2019. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019).

[6] Keng-Hao Chang, Mike Y Chen, and John Canny. 2007. Tracking free-weight exercises. In *International Conference on Ubiquitous Computing*. Springer, 19–37.

[7] Steven Chen and Richard R Yang. 2020. Pose Trainer: correcting exercise posture using pose estimation. *arXiv preprint arXiv:2006.11718* (2020).

[8] Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, and Cewu Lu. 2017. RMPE: Regional Multi-person Pose Estimation. In *ICCV*.

[9] Mykyta Fastovets, Jean-Yves Guillemaut, and Adrian Hilton. 2013. Athlete Pose Estimation from Monocular TV Sports Footage. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

[10] Fitbod. 2021. https://www.Fitbod.me

[11] Xiaonan Guo, Jian Liu, and Yingying Chen. 2017. FitCoach: Virtual fitness coach empowered by wearable mobile devices. In *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*. 1–9. https://doi.org/10.1109/INFOCOM.2017.8057208

[12] Xiaonan Guo, Jian Liu, Cong Shi, Hongbo Liu, Yingying Chen, and Mooi Choo Chuah. 2018. Device-Free Personalized Fitness Assistant Using WiFi. 2, 4, Article 165 (Dec. 2018), 23 pages. https://doi.org/10.1145/3287043

[13] Frederick Hatfield. 1993. *Hardcore bodybuilding: a scientific approach*. McGraw-Hill.

[14] Rushil Khurana, Karan Ahuja, Zac Yu, Jennifer Mankoff, Chris Harrison, and Mayank Goel. 2018. GymCam: Detecting, recognizing and tracking simultaneous exercises in unconstrained scenes. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2018).

[15] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. 2011. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter* 12, 2 (2011), 74–82.

[16] Mark E Lavallee and Tucker Balam. 2010. An overview of strength training injuries: acute and chronic. *Current sports medicine reports* 9, 5 (2010), 307–313.

[17] Ofir Levy and Lior Wolf. 2015. Live repetition counting. In *Proceedings of the IEEE international conference on computer vision*.

[18] Stuart Lloyd. 1982. Least squares quantization in PCM. *IEEE transactions on information theory* 28, 2 (1982), 129–137.

[19] E. Marchand, H. Uchiyama, and F. Spindler. 2016. Pose Estimation for Augmented Reality: A Hands-On Survey. *IEEE Transactions on Visualization and Computer Graphics* 22, 12 (2016), 2633–2651.

[20] S. Milanko and S. Jain. 2019. Demo Abstract: LiftRight: Quantifying Training Performance using a Wearable Sensor. In *2019 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. 1–2.

[21] Slobodan Milanko and Shubham Jain. 2020. LiftRight: Quantifying strength training performance using a wearable sensor. *Smart Health* (2020), 100115.

[22] Slobodan Milanko, Alexander Launi, and Shubham Jain. 2020. Designing Just-in-Time Detection for Gamified Fitness Frameworks. *arXiv preprint arXiv:2005.08834* (2020).

[23] Mirror. 2018. The Nearly Invisible Interactive Home Gym. https://www.mirror.com

[24] Johnny Padulo, Guillaume Laffaye, Luca Paolo Ardigò, and Karim Chamari. 2013. Concentric and Eccentric: Muscle Contraction or Exercise? *Journal of Human Kinetics* 37, 1 (2013), 5 – 6.

[25] Push. 2013. PUSH Band 2.0. https://www.trainwithpush.com/

[26] Meera Radhakrishnan, Darshana Rathnayake, Ong Koon Han, Inseok Hwang, and Archan Misra. 2020. ERICA: Enabling Real-Time Mistake Detection amp; Corrective Feedback for Free-Weights Exercises. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems* (Virtual Event, Japan) *(SenSys '20)*. Association for Computing Machinery, New York, NY, USA, 558–571. https://doi.org/10.1145/3384419.3430732

[27] Sheng Shen, He Wang, and Romit Roy Choudhury. 2016. I Am a Smartwatch and I Can Track My User's Arm. Association for Computing Machinery, New York, NY, USA.

[28] Andrea Soro, Gino Brunner, Simon Tanner, and Roger Wattenhofer. 2019. Recognition and Repetition Counting for Complex Physical Exercises with Deep Learning. *Sensors* 19 (02 2019), 714. https://doi.org/10.3390/s19030714

[29] Tempo. 2019. https://tempo.fit

[30] Alexander Toshev and Christian Szegedy. 2014. DeepPose: Human Pose Estimation via Deep Neural Networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[31] Vicon. 2019. Biomechanics and Sport: Our worldwide user base pushes our research performance to the gold standard. https://www.vicon.com/motion-capture/biomechanics-and-sport