# Mobile Motion Tracking for Disease Prevention and Rehabilitation Using Apple ARKit

Lara Marie REIMER[a,1], Severin WEIGEL[a], Florian EHRENSTORFER[a], Malintha ADIKARI[a], Wolfgang BIRKLE and Stephan JONAS[a]

[a] *Technical University of Munich, Munich, Germany*

**Abstract.** Background: Physical activity helps improve the overall quality of life. The correct execution of physical activity is crucial both in sports as well as disease prevention and rehabilitation. Little to no automated commodity solutions for automated analysis and feedback exist. Objectives: Validation of the Apple ARKit framework as a solution for automatic body tracking in daily physical exercises using the smartphones' built-in camera. Methods: We deliver insights into ARKit's body tracking accuracy through a lab experiment against the VICON system as Gold Standard. We provide further insights through case studies using apps built on ARKit. Results: ARKit exposes significant limitations in tracking the full range of motion in joints but accurately tracks the movement itself. Case studies show that applying it to measure the quantity of execution of exercises is possible. Conclusion: ARKit is a light-weight commodity solution for quantitative assessment of physical activity. Its limitations and possibilities in qualitative assessment need to be investigated further.

**Keywords.** mHealth, Mobile Applications, Fitness Trackers, Augmented Reality

## 1. Introduction

Regular physical exercise is known to be beneficial for the overall quality of life. It improves functional capacity and reduces long-term risks for diseases like Diabetes mellitus Type 2 and Alzheimer's while improving overall health, health-related aspects, and mood [1, 2, 3]. Physical exercise can also positively influence hemic and oncological diseases' treatment and rehabilitation by preventing muscular atrophy and improving the patients' mood [4, 5, 6]. However, the correct execution of physical exercise is essential. Wrong exercise execution can lead to biochemical stress, injuries, and osteoarthritis in the respective joints [7]. To avoid wrong movements, regular supervision of the exercising person by experienced personnel is crucial. Various systems have been developed and evaluated to allow a more profound analysis of human motion and detect problematic movements in exercises. These systems include optical, magnetic, inertial, or mechanic sensors to detect and measure different kinds of metrics about human body motion [8, 9]. Modern motion capture systems track indoor and outdoor activities with different accuracy based on the underlying technologies [10]. However, none of those

---

[1] Corresponding Author: Lara Marie Reimer, Technical University of Munich, Germany, E-Mail: reimer@tum.de

systems currently allows daily usage as they require custom hardware for tracking and often are connected to high costs.

Many people track their daily lives with mobile devices like smartphones, -watches, and sleep or fitness trackers [11]. These devices are equipped with more advanced sensors such as camera systems, gyroscopes, accelerometers, and optical sensors. Also, the devices themselves are getting more powerful and allow resource-intensive applications, e.g., for on-device machine learning. These advances include powerful software applications to perform automatic detection of objects, among them the human body. Software frameworks for mobile devices like Apple ARKit[2] or Vision[3] or Google's Tensorflow Pose Estimation[4] automatically detect the joints' positions through image recognition techniques. These positions can be used for further analysis and enable various new applications, e.g., in mobile games and healthcare.

Using motion detection applications on mobile devices could add to the user's health, allowing quantitative analysis of physical activity. The application area is broad and includes prevention of injuries and sickness, progress, and aggravation tracking, e.g., in rehabilitation, during treatment, or as a motivation to be more active and exercise in general [11]. Mobile devices might be able to deliver a low-cost, commodity alternative to established solutions. Boulos et al. state that while many mobile applications for tracking physical activity have been developed, most of them rely on few sensor data. Used data is often limited to GPS and heart rate. Especially in the area of resistance exercises, personalized coaching applications are missing [12]. However, systems based on wearable sensors can recognize and count exercises [13], even without prior, exercise-specific training [14]. Current research shows that approaches using Apple ARKit can track the lower extremities' motion [15]. A more detailed evaluation of the applicability of ARKit, including the upper extremity, is missing. The aim of this work is to evaluate the suitability and applicability of on-device motion tracking using Apple ARKit and give an outlook to their application in three different use cases: goalkeeper training, physiotherapy and resistance training, and golf.

In all three use cases, complex motion needs to be tracked. If executed incorrectly, these exercises can overstress several joints and promote injuries. The usage of mobile motion capture to assess motion and detect incorrect movements could balance training, prevent injuries, and provide progress supervision. Existing approaches in goalkeeper training build on custom sensors to analyze motion [16]. In our approach, we rely on ARKit as an alternative approach using a commodity device (Case Study 1). Resistance training plays a substantial role in improving and maintaining physical strength and fitness, either during the prevention of diseases and injuries or physiotherapy. Research proves that active physical exercise is among the crucial factors in physiotherapy of common conditions, e.g., in chronic low back pain [17]. We design an algorithm to capture and assess body-weight exercises and provide individual feedback in a mobile coaching application in resistance training (Case Study 2). Golf consists of complex motion sequences (intra- and intermuscular) that require a high amount of coordination, flexibility, and strength, which needs to be practiced continuously to generate high performance. Even for professional golf trainers, an individual assessment of the golf swing and related exercises and tracking progress is challenging. Several solutions have been developed to support golf training. Existing solutions use different sensors attached

---

[2] https://developer.apple.com/augmented-reality/arkit/
[3] https://developer.apple.com/documentation/vision
[4] https://www.tensorflow.org/lite/models/pose_estimation/overview

to the body and the golf club to measure motion [18, 19] or focus on camera-based techniques, like the Coach's Eye application[5], which allows calculating joint angles in a single, selected video frame manually. We propose a more light-weight, automated approach by combining an optical sensor and a single Inertial Measurement Unit (IMU) sensor for qualitatively assessing motion in golf-related exercises (Case Study 3).

## 2. Methods

Within the scope of this paper, we aim at investigating two research questions:

- **RQ 1**: Which accuracy does Apple ARKit provide in contrast to the VICON system?
- **RQ 2**: What are potential use cases for an ARKit-based system?

To answer **RQ 1**, we performed a lab experiment in which we compared motion data generated by Apple ARKit against the VICON system's motion data. To answer **RQ 2**, we are conducting several case studies in different health and exercise science areas, each of them consisting of a mobile prototype application running the ARKit framework.

### 2.1. Suitability of ARKit for Motion Recognition and Tracking

A total of 12 subjects participated in the experiment, 5 females and 7 males, all of good health without physical impairments. The participants' height ranged from 1.56m to 1.96m and their weight from 52.2kg to 97.5kg. In the lab experiment, subjects had to perform 9 different exercises focusing on both the upper and lower extremities, including running on a treadmill at 3 different speeds, passing and catching a ball, jumping jacks, cuttings, and squats. All joint angles are calculated using Euler angles in 3 dimensions: x referring to flexion/extension, y referring to inversion/eversion, and z referring to the rotation. Both systems measured shoulder, elbow, neck, knee, and ankle angles. We calculated minimum, maximum, mean, standard deviation (SD), and range of motion (ROM) and compared the ARKit values against the VICON angles for all exercises.

For the study setup, we used a 10-camera VICON setup with the full-body Plug-in Gait model provided by VICON Nexus[6]. For the recording with ARKit, we placed an iPad Pro 11" 2020 with LiDAR sensor on a tripod in front of the subject with a distance of 3m.

### 2.2. Case Study 1: Recognition of Exercises in Goalkeeper Training

In our first case study, we wanted to identify patterns specific for individual exercises, which is the baseline for further analysis of the motion. For this, we recorded 15 different exercises with ARKit and an additional camera, which served as the basis for labeling the data afterward. The exercises consisted of 6 dive variations, 5 catch variations, 2 throw variations, and 2 kick variations. 4 goalkeepers were recorded. Their football trainer was supervising the recording sessions to guarantee the correct execution of the exercises. Through 10 session recordings in total, we created a dataset of 1050 single exercise executions mapped on the 15 exercises. We used machine learning techniques to train a classifier predicting the matching. We reduced the 6 different dive classes to 2

---

[5] https://www.coachseye.com
[6] https://docs.vicon.com/display/Nexus26/Full+body+modeling+with+Plug-in+Gait

classes due to similarities in the execution as a preparation. Based on research by Ronao et al. [20], we applied an approach based on Convolutional Neural Networks (CNN).

## 2.3. Case Study 2: Recognition of Body-Weight Exercises in Physiotherapy and Resistance Training using a Pose-based Approach

In case study 2, we developed an algorithm to recognize and assess body-weight exercises using ARKit. The algorithm allows app users to record new exercises, store them in a database, detect correct exercise executions of stored exercises, detect incorrect repetitions, and provide individual feedback to the user on improving the motion. Due to the COVID-19 pandemic, an evaluation of the prototype application remains open.

## 2.4. Case Study 3: Classification of Golf-Related Exercises

As part of case study 3, we developed a system consisting of an ARKit-based mobile application prototype and an IMU sensor to measure motion, rotation, and acceleration in an exercise related to the golf swing. The camera is placed in front of the player. The IMU sensor is attached to the hip. The system automatically assesses the exercise execution compared to a reference recording. The automatic assessment is performed by a Dynamic Time Warping (DTW) algorithm, which compares the exercise data against the reference data and evaluates the distance between the compared time series. Through the DTW approach, we classify the executions and provide a grading. We recorded around 300 exercise executions, and a golf trainer assessed them to generate training data. A prototype was implemented, but the automatic assessment could not be qualitatively validated due to the ongoing COVID-19 pandemic. Through an online questionnaire, we gained insights into how golf trainers assessed the prototype's usefulness.

## 3. Results

### 3.1. Suitability of ARKit for Motion Recognition and Tracking

The comparison of the Euler angles calculated by VICON against the Euler angles calculated based on the ARKit data exposed several deviations between the two systems across all tracked joints (see Table 1 for one joint). Across all joints and dimensions, the ARKit measurements revealed smaller SD and ROM values than the VICON measurements. Additionally, ARKit and VICON use different reference points to calculate the Euler angles, which led to mirrored values in the analysis (Figure 1b).

**Table 1.** Results of the VICON and ARKit measurements in the squat exercise. The values represent the knee joint angles. All values are calculated in degrees.

| System | Angle | Dimension | Min | Max | Mean | SD | ROM |
|--------|-------|-----------|-----|-----|------|----|----|
| VICON | Knee | x | 0.57 | 119.61 | 46.35 | 42.45 | 119.04 |
| | Right | y | -3.60 | 16.52 | 5.86 | 6.73 | 20.12 |
| | | z | -22.86 | 36.30 | 1.53 | 20.97 | 59.16 |
| | Knee | x | 0.00 | 117.98 | 46.58 | 42.22 | 117.97 |
| | Left | y | -3.50 | 13.77 | 4.46 | 5.76 | 17.27 |
| | | z | -17.14 | 41.08 | 7.69 | 20.29 | 58.22 |
| ARKit | Knee | x | 23.15 | 54.37 | 29.44 | 6.39 | 31.21 |
| | Right | y | -9.07 | 1.28 | -2.81 | 2.13 | 10.34 |
| | | z | -2.54 | 18.12 | 4.66 | 4.45 | 20.66 |
| | Knee | x | 22.77 | 55.88 | 28.79 | 7.11 | 33.11 |
| | Left | y | -2.08 | 5.79 | 1.66 | 1.48 | 7.87 |
| | | z | -19.85 | 3.00 | -5.20 | 4.31 | 22.85 |

Even though the results exhibited a considerably smaller range of motion, the motion itself was detected in every repetition of the exercises (Figure 1). The ROM was consistently smaller throughout all repetitions in all exercises, supported by the smaller
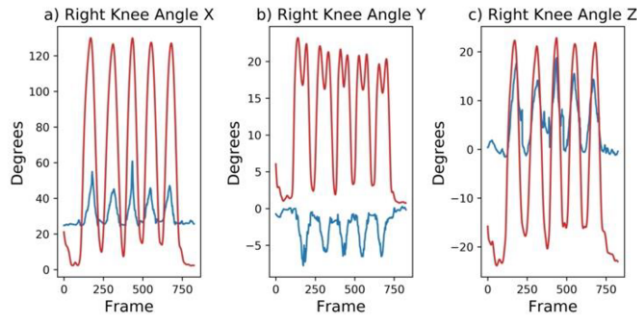


**Figure 1.** Knee joint angles during squat by ARKit (blue) and VICON (red).

SD values. We calculated the Pearson correlation-coefficients for shoulder, elbow, knee, and ankle angle for all participants for the squat exercise for further analysis. We calculated the mean and SD values across all participants, dimensions, and left and right sides based on the coefficients. We used the correlation coefficients' absolute values for the correct detection of the motion, as due to the different reference points of VICON and ARKit, the values of the y dimension of ARKit were mirrored for the joints on the right side of the body. The mean correlation-coefficients and SD for the shoulder, elbow, knee, and ankle angles were $0.406 \pm 0.236$, $0.085 \pm 0.117$, $0.705 \pm 0.144$, and $0.654 \pm 0.184$, respectively.

**Table 2.** Pearson correlation-coefficients of the VICON and ARKit measurements for all 12 participants in the Squat exercise.

| | Shoulder | | Elbow | | Knee | | Ankle | |
|------|------|------|------|------|------|------|------|------|
| | L | R | L | R | L | R | L | R |
| Min | 0.017 | 0.011 | 0.000 | 0.000 | 0.297 | 0.410 | 0.222 | 0.207 |
| Max | 0.801 | 0.825 | 0.658 | 0.447 | 0.901 | 0.898 | 0.909 | 0.911 |
| Mean | 0.367 | 0.439 | 0.084 | 0.085 | 0.674 | 0.736 | 0.614 | 0.693 |
| SD | 0.227 | 0.241 | 0.128 | 0.104 | 0.147 | 0.134 | 0.173 | 0.186 |

### 3.2. Case Study 1: Recognition of Exercises in Goalkeeper Training

Figure 2 visualizes the output of the 3-dimensional positional data of a session with recordings of 5 exercises for the right hand and right foot. The different patterns are identifiable, as well as the repetitions of each exercise.
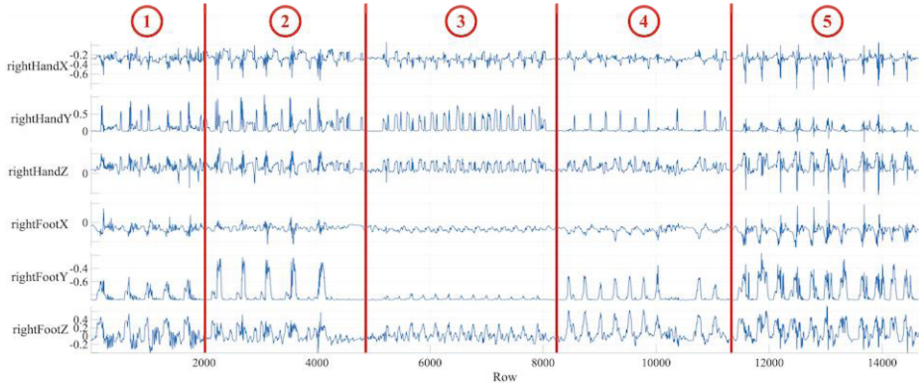


**Figure 2.** Dataset of a session including 5 exercises: Dive High Right (1), Dive High Left (2), Catch Hand (3), Catch Body (4), and Catch Ground (5).

Despite the comparably small dataset, the classifier achieved a validation accuracy of around 75% using the CNN in the test set (Figure 3). The dives and the two versions of the jump catch achieved high rates of correct predictions with around 70-80%. For comparison, we tested the same approach with Long Short-Term Memory Networks (LSTM) and Deep Neural Networks (DNN). Due to the small dataset, both alternatives
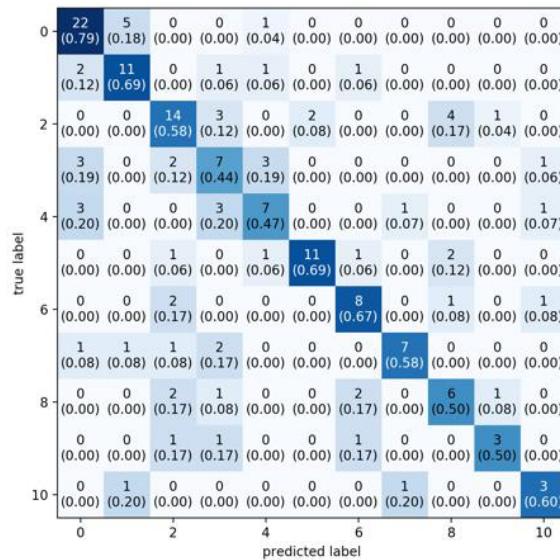


**Figure 3.** Confusion matrix with precision values of the CNN prediction on 11 classes: (0) Dive Right (1) Dive Left (2) Catch Hand (3) Catch Body (4) Catch Ground (5) Jump Catch (6) Jump Catch Run (7) Throw Low (8) Throw High (9) Side Kick (10) Side Kick Ground.

showed overfitting tendencies and did not reach the accuracy of the CNN with an accuracy of 55% and 65%, respectively, for all classes.

The first case study shows that the data generated by ARKit is accurate enough to distinguish between goalkeeper exercises. Even with a small dataset, a machine learning classifier can classify around 75% of the exercises correctly.

## 3.3. Case Study 2: Recognition of Body-Weight Exercises in Physiotherapy and Resistance Training using a Pose-based Approach

To build a mobile, ARKit-based application enabling tracking various dynamic body-weight exercises, we designed a workflow for app users to add new exercises to the app. The prerequisite for registering a new exercise is that it has to consist of clearly different poses. An exercise needs to expose motion in at least one trackable joint in three poses as observable in a squat with full extension of both knees, flexion of at least 90deg in both knees, and a full extension of both knees again (Figure 4). The system captures the three poses, including the measured angles. For each pose, a set of essential joints can be defined, e.g., knee, ankle, and hip joint in the squat. This setup should be done by qualified personnel, e.g., a physiotherapist, to ensure the exercise's correctness.



**Figure 4.** Exercise repetition tracking using the pose-based approach. ARKit's recognition is shown by the robot overlay. ARKit tracks the poses in a, b, and c, and updates the repetition count after pose c is matched.

For exercise recognition, we convert the 3-dimensional positional joint data provided by ARKit into angles. Our proposed algorithm observes every motion detected by ARKit and matches the motion's progression against the curve expected by the exercise. Once all poses have been reached in the correct order, the algorithm considers the repetition as completed and increases the count, as shown in Figure 4c. The algorithm is flexible enough to recognize a variety of body-weight exercises without training a specific machine learning model through this approach.

Considering the initial lab experiment results, approaches on qualitatively assessing exercise repetitions need to be investigated. Using the proposed algorithm allows the creation and tracking of dynamic body-weight exercises based on joint angle calculations.

## 3.4. Case Study 3: Classification of Golf-Related Exercises

We evaluated the Golf Coach app's usability in an online questionnaire. In total, 22 golf trainers. The trainers' age ranged between 19 to 66 years, with a mean of 38.5 years and a median of 32 years. Their experience as a golf trainer ranged from 1 year to 40 years, with a mean of 13.14 years and a median of 7.5 years. 95.5% stated that they would use

an app to track their trainees' progress, and 70% had used mobile coaching apps before. After a guided tour through the app, the participants were asked whether they would use it. 59% stated that they would use it, 9% would use it from time to time. 9% would prefer to test it first. 5% said that they would not use it. 18% chose not to answer. 55% stated that progress tracking would be the most important feature to them.

## 4. Discussion & Future Work

The lab experiment and the case studies provide evidence that mobile applications based on ARKit can track joint motion. Table 1 shows that the recognition is not as accurate as the reference values detected by VICON. ARKit does not seem to capture a full extension of the knee, as the minimum value provided is 22.77deg, compared to 0.00deg seen by VICON. Flexion is recognized, but to a much smaller extent than by the VICON system, with 119.61deg maximum in VICON and 54.37deg maximum in ARKit in the same joint. ARKit exposes smaller mean, SD, and ROM values. Similar observations apply to the inversion/eversion and rotation dimensions. The Pearson correlation-coefficients analysis shows that the ARKit and VICON measurements are strongly correlated, especially in the lower extremities. Even smaller changes in inversion/eversion are tracked in the squat motion's turning points, as shown in Figure 1b. Interestingly, the upper extremities' joint angle motion exhibit a considerably low correlation.

The ROM detected by ARKit is comparatively but reliably smaller than the ROM of VICON. This poses the question of whether an algorithm can be developed, which can approximate the ARKit values to the correct values provided by the VICON system. Enabling this would allow performing not only a quantitative motion analysis but also a qualitative analysis. A qualitative analysis would allow additional possibilities in motion analysis, mainly regarding feedback to prevent wrong motion.

Motion tracking using ARKit, especially regarding the lower extremity, seems to be a promising, light-weight approach. Even though the lab experiment shows that a quantitative assessment of exercise executions, e.g., for repetition counting, is feasible, the experiment was performed with a relatively limited number of participants and different exercises. A more extensive experiment is needed to gain further insights, which includes a more diverse set of participants and exercises. Through this, the limitations of the ARKit-based tracking need to be further investigated.

In case study 1, we showed that we can use ARKit data as reference values to recognize patterns in recordings. It remains open to which extent we can use reference data provided by ARKit to allow qualitative analysis of the motion, as proposed in case studies 2 and 3. The case studies served to explore the possibilities of ARKit-based motion tracking in health-related fields. Even though the first results show a high interest of potential app users and the lab experiment and data analysis expose such approaches' potential, the case studies need to be tested and validated in more extensive studies.

## 5. Conclusion

In this paper, we were able to show that even though ARKit exposes major inaccuracies in tracking the ROM, it is reliable in tracking the motion itself. Therefore, the ARKit framework can be used to assess physical exercise, recognize exercises, and count repetitions. ARKit enables various use cases for mobile applications, especially in the

prevention and rehabilitation of diseases and injuries in humans. In this paper, we presented 3 different case studies. In case study 1, we classified exercises in goalkeeper training, which shows that we can identify patterns specific to exercises in the data provided by ARKit. In case study 2, we presented an algorithm that enables the creation and automatic tracking of dynamic body-weight exercises. In case study 3, we combined ARKit with an IMU sensor to enable a qualitative assessment of an exercise related to the golf swing.

ARKit seems to be a promising, light-weight alternative to well-established motion tracking systems. Its limitations and possibilities need to be further investigated.

## References

[1]  W. B. Kannel, P. Sorlie, Some Health Benefits of Physical Activity: The Framingham Study. *Arch Intern Med*, **139**(8) (1979), 857–861.

[2]  A.P.A. Penedo and F. Ja, Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Curr Opin in Psychiatr*, **18** (2005), 189 – 193.

[3]  D.E.R. Warburton, C.W. Nicol, S.S.D. Bredin, Health benefits of physical activity: the evidence. *CMAJ*, **174**(6) (2006), 801-809.

[4]  J. Midtgaard, M. T. Baadsgaard, T. Møller, B. Rasmussen, M. Quist, C. Andersen, M. Rørth, and L. Adamsen. Self-reported physical activity behaviour; exercise motivation and information among Danish adult cancer patients undergoing chemotherapy. *Eur J of Oncol Nurs*, **13**(2) (2009), 116 – 121.

[5]  L. E. Stephenson, D. G. Bebb, R. A. Reimer, S. N. Culos-Reed, Physical activity and diet behaviour in colorectal cancer patients receiving chemotherapy: associations with quality of life. *BMC Gastroenterol*, **9**(1) (2009).

[6]  C. Visovsky. Muscle Strength, Body Composition, and Physical Activity in Women Receiving Chemotherapy for Breast Cancer. *Integr Cancer Ther*, **5**(3) (2006), 183–191.

[7]  R.H. Sandmeier, Osteoarthritis and Exercise: Does Increased Activity Wear Out Joints?. *Perm J*, **4**(4) (2000), 26-28.

[8]  M. Yahya, J.A. Shah, K.A. Kadir, Z.M. Yusof, S. Khan, A. Warsi, ""Motion capture sensing techniques used in human upper limb motion: a review"", *Sensor Rev*, **39**(4) (2019), 504-511.

[9]  M. Field, Z. Pan, D. Stirling, F. Naghdy, Human motion capture sensors and analysis in robotics, *Ind Robot*, **38**(2) (2011), 163-171.

[10]  E. van der Kruk, M.M. Reijne, Accuracy of human motion capture systems for sport applications; state-of-the-art review, *Eur J Sport Sci*, **18**(6) (2018), 806-819.

[11]  O.D. Lara, M.A. Labrador. A Survey on Human Activity Recognition using Wearable Sensors. *IEEE Commun Surv Tut*, **15**(3) (2013), 1192-1209.

[12]  M. N. Kamel Boulos, S.P. Yang. Mobile physical activity planning and tracking: a brief overview of current options and desiderata for future solutions. *mHealth*, **7**(13).

[13]  S. Ishii, A. Yokokubo, M. Luimula, G. Lopez. ExerSense: Physical Exercise Recognition and Counting Algorithm from Wearables Robust to Positioning. *Sensors,* **21**(1), 91.

[14]  S. Zelman, M. Dow, T. Tabashum, T. Xiao, M.V. Albert, Accelerometer-Based Automated Counting of Ten Exercises without Exercise-Specific Training or Tuning. *J Healthc Eng*, *2020*, 8869134.

[15]  S. Basiratzadeh, E. D. Lemaire, N. Baddour, Augmented Reality Approach for Marker-based Posture Measurement on Smartphones, *Annu Int Conf IEEE Eng Med Biol Soc*, 2020, 4612-4615.

[16]  J. Haladjian, D. Schlabbers, S. Taheri, M. Tharr, and B. Bruegge, Sensor-based Detection and Classification of Soccer Goalkeeper Training Exercises. *ACM Trans. Internet Things*, **1**(2) (2020).

[17]  M. van Middelkoop, S.M. Rubinstein, T. Kuijpers, A.P. Verhagen, R. Ostelo, B.W. Koes, M.W. van Tulder, A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non‐specific low back pain. *Eur Spine J*, **20** (2011), 19–39.

[18]  H. Ghasemzadeh, V. Loseu, R. Jafari. Wearable coach for sport training: A quantitative model to evaluate wrist-rotation in golf. *J Ambient Intell Smart Environ*, **1**(2) (2009), 173–184.

[19]  U. Jensen., M. Schmidt, M. Hennig, et al., An IMU-based mobile system for golf putt analysis. *Sports Eng*, **18** (2015), 123–133.

[20]  C.A. Ronao, S.B. Cho, Human activity recognition with smartphone sensors using deep learning neural networks, *Expert Syst Appl*, **59** (2016), 235-244.