When Virtual Reality Meets Internet of Things in the Gym: Enabling Immersive Interactive Machine Exercises

FAZLAY RABBI*, TAIWOO PARK*, BIYI FANG, MI ZHANG, Michigan State University
YOUNGKI LEE, Singapore Management University

With the advent of immersive virtual reality (VR) head-mounted displays (HMD), we envision that immersive VR will revolutionize the personal fitness experience in our daily lives. Toward this vision, we present JARVIS, a virtual exercise assistant that is able to provide an immersive and interactive gym exercise experience to a user. JARVIS is enabled by the synergy between Internet of Things (IoT) and immersive VR. JARVIS employs miniature IoT sensing devices removably attachable to exercise machines to track a multitude of exercise information including exercise types, repetition counts, and progress within each repetition in real time. Based on the tracked exercise information, JARVIS shows the user the proper way of doing the exercise in the virtual exercise environment, thereby helping the user to better focus on the target muscle group. We have conducted both in-lab experiments and a pilot user study to evaluate the performance and effectiveness of JARVIS, respectively. Our in-lab experiments with fifteen participants show that JARVIS is able to segment exercise repetitions with an average accuracy of 97.96% and recognize exercise types with an average accuracy of 99.08%. Our pilot user study with ten participants shows statistically significant improvements in perceived enjoyment, competence, and usefulness with JARVIS compared to a traditional machine exercise setting ($p < 0.05$). Finally, our surface electromyography (sEMG) signal analysis conducted during the pilot user study shows statistically significant improvement in terms of muscle activation ($p < 0.01$), indicating the potential of JARVIS in providing an engaging and effective guidance for machine exercises.

CCS Concepts: • **Human-centered computing** → *Ubiquitous and mobile computing systems and tools*;

Additional Key Words and Phrases: Virtual Assistant; Virtual Reality; Internet of Things; Activity Recognition; Human-Computer Interaction.

ACM Reference Format:
Fazlay Rabbi*, Taiwoo Park*, Biyi Fang, Mi Zhang and Youngki Lee. 2018. When Virtual Reality Meets Internet of Things in the Gym: Enabling Immersive Interactive Machine Exercises. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2, Article 78 (June 2018), 21 pages. https://doi.org/10.1145/3214281

1 INTRODUCTION

The promise of immersive virtual reality (VR) is starting to look very real with the emergence of head mounted displays (HMDs) such as Oculus Rift [2], Samsung Gear VR [3], and HTC VIVE [1]. Immersive VR soaks a user in a computer-generated simulated environment that naturally responds to the movements of the user. An engaged immersive virtual experience is thus realized by employing sensing technologies that capture the user's movements and using those information to update the sensory stimuli presented to the user via a HMD. As such,

* Authors contributed equally. Taiwoo Park is the corresponding author.

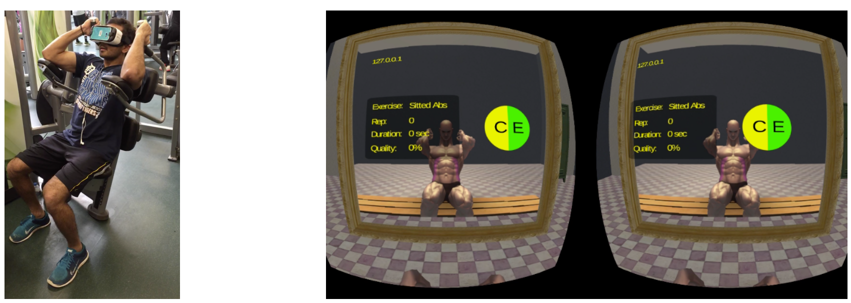
Authors' addresses: Fazlay Rabbi*, Taiwoo Park*, Biyi Fang, Mi Zhang, Michigan State University; Youngki Lee, Singapore Management University.

2474-9567/2018/6-ART78 $15.00
https://doi.org/10.1145/3214281

it creates an illusion of being immersed in a virtual environment in which the user can interact [39]. Given its unique capability of enabling such engaged virtual experience, immersive VR has been regarded as a technology that has the significant potential to revolutionize a wide range of industries such as entertainment, education, fitness and health care.

As one of its important application areas, we envision that immersive VR will help people improve their workout experiences in the gym environment. Today, workout in the gym has become one important part of people's modern life [40]. However, working out on the exercise machines in the gym could make exercisers feel easily bored [22, 24]. Moreover, without the feedback from professional trainers, it is challenging for novice exercisers to perform a certain exercise in an adequate pace with the right group of muscles employed. This prevents exercisers from making steady progress, and eventually makes exercisers lose their interests and motivation.

In this paper, we propose a *virtual exercise assistant*, named JARVIS, that is able to provide an *immersive* and *interactive* gym exercise experience to a user, which we envision has the potential to fundamentally change the way how people work out in the gym. JARVIS is enabled by the synergy between two emerging technologies: Internet of Things (IoT) and immersive VR. At the backend, JARVIS uses a miniature IoT sensing device removably attachable to exercise machines to continuously and automatically track a multitude of exercise information including exercise types, repetition counts, and progress within each repetition in real time. At the frontend, JARVIS leverages an immersive VR HMD to create a virtual exercise environment along with a virtual body of the user within the environment. Based on the extracted exercise information, JARVIS shows the user the proper way of doing the exercise, instructs the user to adjust exercise pace, and guides the user to focus on the particular muscle group that the exercise targets. In this way, JARVIS is creating an immersive and interactive gym exercise experience that was not previously available. Figure 1 illustrates how JARVIS is being used during machine exercise and the immersive and interactive VR gym exercise experience it provides.



(a) (b)

Fig. 1. (a) a user using JARVIS during machine exercise; (b) stereoscopic screen images of the virtual exercise environment along with a virtual body of the user within the environment.

Motivating Scenario. Steve has been seeking an effective way to adhere to regular machine exercise schedules. Personal trainers could not be with him all the time, and most of the existing mobile fitness apps require him to manually log his exercise progress, which is burdensome. Steve decided to try JARVIS virtual exercise assistant. He brings a VR HMD and a miniature IoT sensing device to a gym. He attaches the sensing device onto a trunk curl machine and wears the HMD. In a few repetitions, JARVIS automatically detects the current exercise type and immediately displays a virtual exercise environment for the trunk curl exercise. JARVIS shows a "virtual Steve" in the virtual exercise environment that follows Steve's exercise movements. JARVIS also shows the real-time exercise progress information, and highlights the muscle group on which he needs to focus. This helps Steve easily give more attention on the target muscle group. Once Steve finishes trunk curl, JARVIS turns off the VR exercise scene and switches to outside pass-through camera to let Steve navigate to the next exercise machine.

Challenges. The design of JARVIS presents several challenges. First, the foundation of JARVIS is to recognize the types of machine exercises and to track exercise progress automatically. This requires a sensing mechanism to capture movements during machine exercises. Unfortunately, most exercise machines in the gyms are not yet instrumented with sensing capabilities to capture exercise movements. Second, high accuracy and low latency are two key requirements of highly interactive systems to achieve high perceived usability and satisfaction [13, 17, 30]. As a highly interactive VR application, JARVIS needs to meet those strict performance requirements. Therefore, designing lightweight sensor data processing algorithms that incur minimal latency while achieving high accuracy of exercise information tracking represents a unique challenge for JARVIS. Third, a VR HMD that employs a smartphone as a host processing and display device is prone to the overheat problem due to continuous high-quality VR scene rendering. Therefore, reducing the computation load of the VR rendering while satisfying the needed visual quality represents another unique challenge for JARVIS.

Approach. To address the first challenge, JARVIS employs a miniature IoT sensing device that can be easily attached to exercise machines and track machine exercises by capturing machine movements during exercises. This *machine-attachable* approach not only equips exercise machines with sensing capabilities without being instrumented but also turns JARVIS into a mobile system that allows a user to enjoy immersive VR exercise experience anywhere.

The proposed machine-attachable approach also helps achieve higher exercise information tracking accuracy. The most commonly used approach for tracking machine exercises in the literature is using a smartphone or a wearable sensing device such as a wristband [11, 27, 43]. Compared to this *human-wearable* approach, the proposed machine-attachable approach has two key advantages for tracking machine exercises with high accuracy. First, the machine-attachable approach is able to easily capture machine exercises which target abdominal and lower limb muscles that the human-wearable approach has difficulty to capture. Second, since the sensing device is attached to the exercise machine, the machine-attachable approach only captures machine movements caused by machine exercises. In contrast, the human-wearable approach captures not only body movements caused by machine exercises but also non-exercise body movements between exercise sessions, which requires significant signal processing efforts to filter out [34]. JARVIS leverages these advantages and employs lightweight sensor data processing algorithms to accurately track machine exercises in real time.

Finally, to address the overheat problem, JARVIS employs a mixed visual quality technique that effectively reduces the overall computation workload of onboard CPU/GPU by highlighting the muscle group targeted by the machine exercise being performed using high-quality rendering while drawing the remaining body parts using low-quality rendering. Highlighting the muscle group also has an effect of attracting users' attention during exercises, which will lead to greater mind-muscle connection (MMC) [42] and muscle activation [10].

Research Questions. With the design of JARVIS, we aim to answer two research questions at the highest level:

- **RQ1**. Is the machine-attachable approach incorporated by JARVIS able to accurately recognize the types of machine exercises and track exercise progress in real time?
- **RQ2**. Does the virtual exercise environment provided by JARVIS have the potential of leading to more effective, engaging, and enjoyable machine exercises over traditional machine exercises without VR?

To answer the above research questions, we have conducted both in-lab experiments to evaluate the technical components of JARVIS as well as a pilot user study focusing on the *Seated Abs* machine exercise to evaluate its usability. In our preliminary in-lab experiments with fifteen participants, we show that JARVIS is able to segment repetitions and classify exercise in real time. The mixed visual quality technique effectively reduces the computational workload of the VR HMD device and prevents it from being overheated. Our pilot user study with ten participants indicates the potential of effective guidance and increased enjoyment brought by the immersive VR exercise environment. Compared to a traditional machine exercise setting, JARVIS shows significant differences in three subcategories of Intrinsic Motivation Inventory (IMI) [37, 41]: perceived enjoyment, competence, and usefulness ($p < 0.05$ for each subcategory). The participants' interview results confirm that the participants found the VR user interface and the virtual self representation effective in helping them perform machine exercises. Finally, we quantitatively analyze the perceived user engagement and muscle activities of the users using high-fidelity clinical surface Electromyography (sEMG) sensors. Our results show that JARVIS provides a statistically significant difference in muscle activation level ($p < 0.01$, $p < 0.05$ for external and internal oblique abdominis respectively) compared to traditional machine exercises without JARVIS.

2 DESIGN GOALS AND RELATED WORK

2.1 Portable and Machine-Attachable Sensing Device

JARVIS is designed to be *portable* and can be easily *attached to* and *removed from* exercise machines. Most of previous works employed motion sensors that are worn on different human body locations such as armbands [11, 27, 29, 48], wristband [28, 36], glove [48], waist/chest belt [11, 48], or several of these body locations [43]. In the domain of machine exercises, our approach of using miniature sensing devices temporarily attached to exercise machines is favored over using wearable devices because it lowers the user's burden and cumbersomeness on putting sensing devices on different parts of their bodies during exercises. More importantly, wearable devices capture not only movements related to machine exercises but also non-exercise body movements between exercise sessions. In contrast, by attaching the sensing device onto exercise machines, the collected sensor data is only related to machine exercises. This significantly simplifies the sensor data processing effort and leads to higher exercise information tracking accuracy.

2.2 Universal Sensing Platform

JARVIS is designed to provide a *universal sensing platform* that can be eventually used for any exercise machine in a *plug and play* manner. Pioneer works in powering exercise machines with sensing capabilities have explored customized instrumentation of exercise machines by integrating a variety of types of sensors into different exercise machines [14, 33]. This approach requires considerable efforts from machine manufacturers in modifying exercise machines, leading to significant increase of the costs of exercise machines. In contrast, JARVIS aims to reduce the burden of machine manufacturers by developing a uniform sensing device that can provide sensing capability to any exercise machine without customized modification. We envision that in the future, every exercise machine will have a standardized slot/interface for the uniform sensing device to plug in. The exerciser can plug out the device from a machine after finishing the exercise and move on to another machine. As such, JARVIS acts as a personal device that tracks an individual's exercises on any machine.

2.3 Online Machine Exercise Information Tracking

JARVIS is designed to provide highly accurate machine exercise information with low latency in an *online* manner. A large body of previous works did not target highly interactive applications but focused on automatically generating workout summary during or after exercise sessions [27]. As such, there has not been a dire need of online sensor data processing. In contrast, JARVIS aims to continuously track a multitude of exercise information including exercise types, repetition counts, and progress within each repetition in real time. This is essential for highly interactive VR applications like virtual exercise assistant that JARVIS targets.

2.4 Immersive Movement and Fitness Training Experience

Finally, JARVIS is designed to provide a truly *immersive* VR gym exercise experience. There has been a series of prior works that provide an immersive movement training experience using motion cameras and augmented

Table 1. Comparison between existing works and JARVIS.

Author and Year	Repetition Segmentation Accuracy	Type Recognition Accuracy	Target Exercises	Target Application
[11] Chang et al., 2007	94.1%	85%	9 dumbbell exercises	Not specified
[14] Ding et al., 2015	94%	91%	10 dumbbell exercises	Offline summary
[27] Morris et al., 2014	86.90%	93.8%	13 various types of exercises at a time	Online tracking and offline summary
[28] Mortazavi et al., 2014	89.9%	N/A	5 free and body weight exercises	Offline summary
[29] Muehlbauer et al., 2011	85.1%	91%	10 machine weight exercises	Not specified
[36] Pernek et al., 2013	0.994 (F1 score)	N/A	9 machine weight exercises	Not specified
[43] Seeger et al., 2011	97.58%	95%	5 cardio and 11 weight lifting	Online summary
[48] Velloso et al., 2014	N/A	78.2%	5 different styles of biceps curl	Not specified
JARVIS	**97.96%**	**99.08%**	12 machine weight exercises	Online interactive VR

images such as YouMove [7] and MotionMA [47]. These works exemplified the potential of AR/VR applications to enable effective movement and fitness training. Inspired by these works, JARVIS aims to deliver truly immersive fitness training to anywhere, starting from gyms with exercise machines, by employing IoT sensing devices and VR HMD devices. To achieve this goal, JARVIS provides the user with an immersive VR machine exercise environment and visualizes a virtual body of the user that follows the user's exercise movements. More importantly, JARVIS aims to naturally guide the user to give effective focus on specific muscle group. Inspired by several instruction methods from the fields of kinesiology and sports physical therapy, we leverage the flexibility of visual presentation in VR to color specific muscle group where the user needs to give focus on to achieve more effective muscle activation including verbal and manual cues [6, 21, 38]. Table 1 summarizes these previous works and compares them with JARVIS proposed in this paper.

3 OVERVIEW OF JARVIS

Figure 2 illustrates the system architecture of JARVIS. As shown, JARVIS operates with two types of devices: 1) a miniature IoT sensing device that is attachable to gym machines to track machine exercises; and 2) a VR HMD that processes the sensor data as well as visualizes the exercise information and the computer-generated virtual environment in real time. JARVIS is composed of two core components: the *Real-time Exercise Analyzer* and the *VR Synthesizer* that run inside the VR HMD at the backend and the frontend, respectively.

At the backend, the *Real-time Exercise Analyzer* retrieves the sensor data from the IoT sensing device and analyzes the sensor data. Specifically, the *Real-time Exercise Analyzer* consists of two major components: 1) *Exercise Progress Tracker* and 2) *Exercise Type Recognizer*. The role of the *Exercise Progress Tracker* is to segment the streaming sensor data into individual exercise repetitions (i.e., *Repetition Segmentor*), count the number of repetitions (i.e., *Repetition Counter*), and track the progress within each repetition (i.e., *Motion Progress Detector*). Given the segmented repetitions, the *Exercise Type Recognizer* recognizes the exercise type of each repetition.

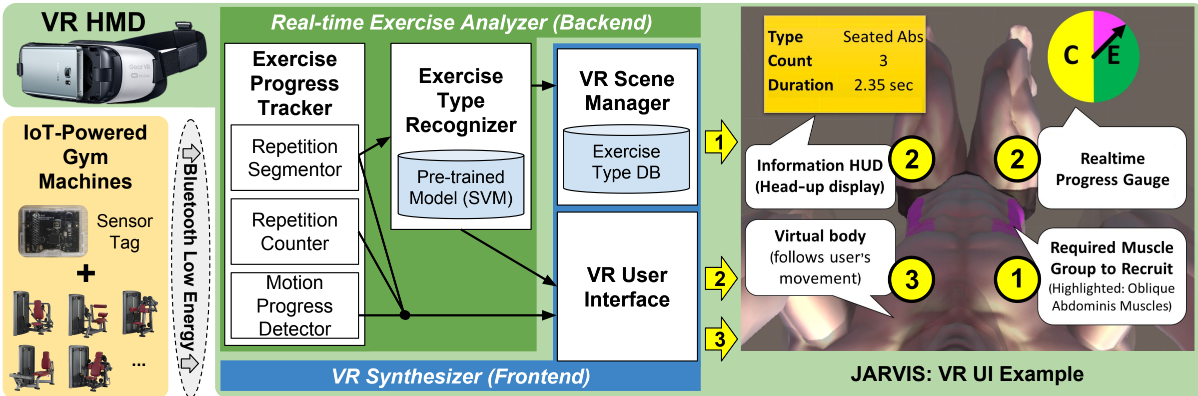


Fig. 2. The system architecture of JARVIS.

Through a suite of techniques including sensor placement location identification, feature selection, and session-wise voting, the *Exercise Type Recognizer* is able to achieve over 99% exercise type recognition accuracy.

At the frontend, the *VR Synthesizer* consists of two major components: 1) *VR Scene Manager* and 2) *VR User Interface*. The *VR Scene Manager* plays two roles: it creates a virtual gym exercise environment that corresponds to the current exercise type; it also presents a virtual body of the exerciser while highlighting the target muscle group based on the current exercise type. The *VR User Interface* renders a head-up display in a virtual space and delivers exercise progress information to a user in real time.

In the following two sections, we describe the *Real-time Exercise Analyzer* and the *VR Synthesizer* in detail.

4 REAL-TIME EXERCISE ANALYZER

4.1 Target Machine Exercises

We target 12 machine exercises recommended by the resistance training guide for healthy adults from the American College of Sports Medicine (ACSM) [5]. These exercises are among the most common machine exercises that target different muscle groups on the body. Each exercise uses a dedicated machine to train a specific muscle group. The 12 machine exercises and their targeting muscle groups are shown in Figure 3.

4.2 Data Acquisition

JARVIS uses the CC2650STK SensorTag developed by Texas Instruments (TI) as the IoT sensing device. SensorTag is TI's state-of-the-art IoT sensing device that integrates high-performance sensors and Bluetooth Low Energy (BLE) communication in a miniature form factor. Specifically, JARVIS utilizes the SensorTag's on-board 3-axis accelerometer and 3-axis gyroscope (i.e., motion sensor) to capture machine exercises. The sensing range of the accelerometer and the gyroscope is set to be ±16g and ±1200 degree per second (dps), respectively, and the sampling rate is set to be 10 Hz. To facilitate a user to attach the SensorTag to exercise machines, we have designed and 3D printed a plastic case ($1.79 \times 2.64 \times 0.55$ inch) with an embedded magnet to host the SensorTag. With the magnet, the case can be easily and firmly attached to exercise machines. The SensorTag and the customized case are illustrated in Figure 4.

4.3 Exercise Progress Tracking

4.3.1 Repetition Segmentation and Counting. The goal of repetition segmentation is to segment the streaming sensor data so that each segment contains one complete repetition of the performed machine exercise. Since a

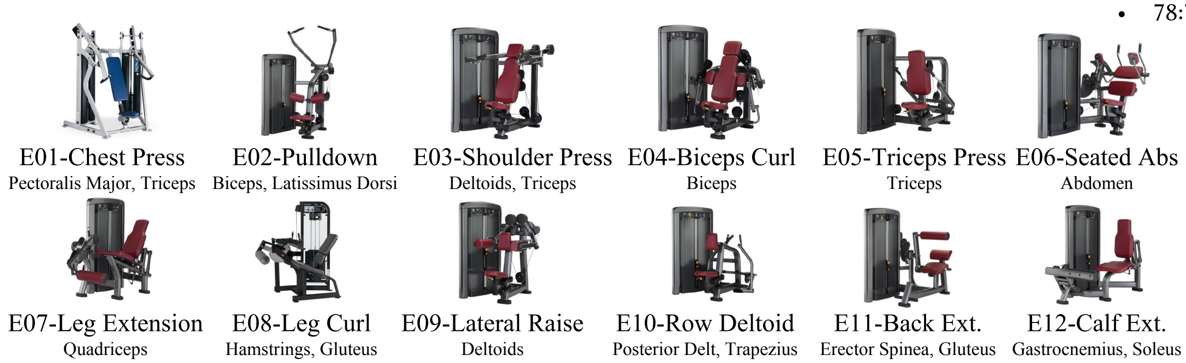


Fig. 3. The twelve machine exercises considered in this work.

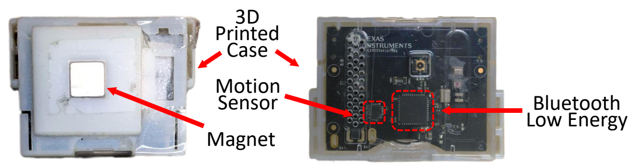


Fig. 4. The miniature SensorTag device and the customized 3D printed case with an embedded magnet.

user can place the SensorTag on exercise machines in different ways which leads to different orientations, one straightforward scheme is to derive the orientation-independent acceleration magnitude signal and then apply peak detection on top of it to segment exercise repetitions. However, such scheme is unsuitable in the context of machine exercises because different machine exercises will have different numbers of peaks and valleys in each repetition. To illustrate this point, Figure 5 (a) and Figure 5 (b) depict the three axes of the accelerometer signal as well as the corresponding acceleration magnitude signal of three repetitions of *Pulldown* and *Seated Abs*, respectively. As shown, within each repetition, the acceleration magnitude signal of *Pulldown* has one peak and two valleys while *Seated Abs* has three peak and two valleys. Without knowing the machine exercise type *a prior*, by blindly applying peak detection, one repetition can be mistakenly segmented into multiple repetitions.

In this work, we design a Principle Component Analysis (PCA)-based scheme to segment repetitions. The key observation behind our scheme is illustrated in Figure 5 (c) and (d). Specifically, Figure 5 (c) and Figure 5 (d) illustrate the first principle component (PC) extracted from the three axes of the accelerometer signal of *Pulldown* and *Seated Abs*, respectively. We observe that even though the exercise type is different, each repetition intersects the mean crossing line of the first PC signal exactly twice. The same observation holds true for all the 12 target machine exercises considered in our work. This is because one repetition of any type of the target machine exercises consists of one concentric phase (i.e., muscle shortening) and one eccentric phase (i.e., muscle lengthening). The first PC reliably captures both two phases across different machine exercises.

Based on this key observation, our PCA-based repetition segment scheme first extracts the first PC of the 3-axis accelerometer data and finds the mean crossing point of the first PC. Our scheme then finds out whether the first PC is going downward or upward at the first mean crossing point. If going downward, the lowest minima between every non-overlapping pair of mean crossing points is the peak of the repetition and the two closest maxima on the left and right side of the two mean crossing points are the start and end point of the repetition. If going upward, the highest maxima between every non-overlapping pair of mean crossing points is the peak of the repetition and the two closest minima on its left and right are the start and end of the repetition.

After a new repetition is segmented, the number of segmented repetitions is updated in real time.

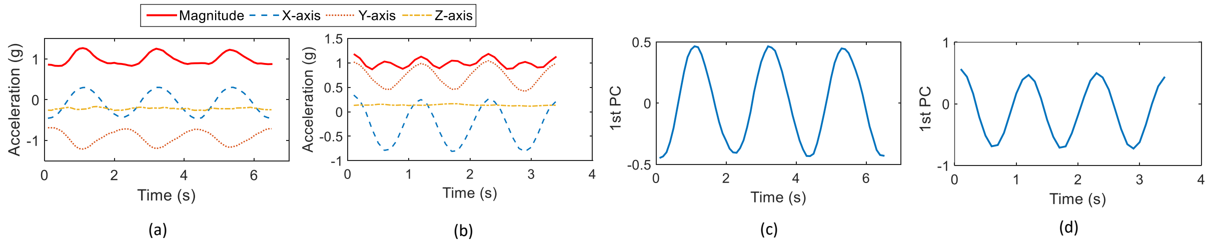


Fig. 5. The illustration of the principle of the repetition segmentation algorithm. (a) and (b): 3-axis accelerometer data and the corresponding acceleration magnitude signal of three repetitions of (a) *Pulldown* and (b) *Seated Abs*. (c) and (d): the first principle component (PC) of the three axes of the accelerometer signal of (c) *Pulldown* and (d) *Seated Abs*.

4.3.2 Motion Progress Estimation. To provide the user with high interactivity, JARVIS tracks the progress of the exerciser within each exercise repetition in real time. However, providing the exact progress status within each repetition in real time is not feasible because the exact progress status can only be available after seeing the complete repetition. As an alternative, we design a motion progress estimation scheme to provide a reasonable estimation about the progress within each repetition in real time. Specifically, our scheme uses the values of the first PC signal, obtained from the segmentation stage, to estimate the progress status. The progress status starts from 0% and ends at 100% with a step of 10%. The specific values of the first PC signal that correspond to those status percentages are obtained by previously seen repetitions as training data. During online motion progress tracking, if the value of the first PC falls between two status percentages, the higher percentage will be reported as the estimated progress status.

4.4 Exercise Type Recognition

After segmenting the machine exercises into repetitions, the second stage of machine exercise analytics is to identify the type of the exercise for each repetition. Our key observation is that due to the different mechanical constraints of exercise machines, each type of machine exercises has a certain form. Based on this observation, we frame the exercise type recognition problem as a classification problem. As explained before, exercisers could place the SensorTag on exercise machines in different ways which may lead to different orientations. To make our classification algorithm orientation-independent, we compute the magnitude of the three dimensional accelerometer data as well as the magnitude of the three dimensional gyroscope data within each repetition. Based on these two magnitudes, we have extracted a list of features which have been proven to be effective for activity recognition [52, 53]. Table 2 lists all the features. Finally, we stack the extracted features into a feature vector and import the feature vector into a linear kernel Support Vector Machine (SVM) for classification.

Table 2. List of features used for exercise type recognition.

Mean	Median	Standard Deviation
Variance	Skewness	Kurtosis
Energy	Interquartile Range	Spectral Entropy
First Order Derivative	Second Order Derivative	Magnitude of Average Rotational Speed
Dominant Frequency	RMS	Signal Magnitude Area

4.4.1 Sensor Placement Location Identification. Although attaching a sensing device directly to exercise machines provides cleaner sensor data, it is necessary to identify the sensor placement location on each exercise machine that achieves the best exercise type recognition performance. To achieve our goal, without loss of generality, we collect sensor data from two locations on each exercise machine, and find the best combination of locations across the exercise machines by searching through all possible location combinations. Our criteria on choosing sensor placement locations are: 1) the sensor placement locations must be easily accessible by users; and 2) the two sensor placement locations on one machine should have different motion ranges so that the sensor data collected from these two locations are different enough. We label the location on each machine with larger motion range as 'L' and the location on each machine with smaller motion range as 'S'. Since 12 exercise machines are considered in this work, the number of all possible sensor location combinations is 4096.

We envision that in the future, every exercise machine produced by machine manufacturers will have a standardized slot to plug in the SensorTag. Our experiment results will help machine manufacturers find the best locations for those standardized slots on exercise machines.

4.4.2 Feature Selection and Session-wise Voting. To minimize the computational overhead of feature extraction, we utilize the Sequential Floating Forward Selection (SFFS) feature selection algorithm [20] to identify a minimal subset of features that achieves the highest accuracy for recognizing the exercise type of each repetition. In addition, we utilize a voting scheme across repetitions in the same exercise session to further enhance the recognition accuracy. Specifically, we take the majority of the recognized exercise types from all the repetitions within the same exercise session as the recognized exercise type for the whole exercise session. As we will show in the evaluation section, our session-wise voting scheme does not need more than two repetitions on average to achieve the highest accuracy, which provides users with highly interactive experiences.

5 VR SYNTHESIZER

5.1 VR Scene Management

Once a user starts exercising, the *VR Scene Manager* automatically initiates a virtual exercise environment based on the recognized exercise type. When a user is taking a rest between exercise sessions, the *VR Scene Manager* provides an outside (i.e., pass-through) vision using the camera placed at the back of the VR HMD. In the following, we describe three important features of the *VR Scene Manager*.

5.1.1 Virtual Self Representation. The first important feature of *VR Scene Manager* is the virtual self representation created for the user. The *VR Synthesizer* generates a virtual body of the user which follows the user's movements during machine exercises. This virtual body enables the user to intuitively understand the pace and progress of the current exercise repetition.

5.1.2 Muscle Highlighting. The second important feature of *VR Scene Manager* is to visually highlight the target muscle group that corresponds to the exercise being performed. Figure 6 illustrates this feature where the oblique abdomens muscle group is rendered using purple color, which has an effect of attracting users' attention during exercise. This design hypothesizes that visual highlight of target muscle group will lead to greater mind-muscle connection (MMC) [42]. MMC is a practical term denoting the strategy which gives attention to consciously direct neural drive to the target muscle usually achieved through imagination [42]. Increased MMC is known to lead to greater muscle activation [10], which potentially increases muscle protein accretion [49, 50]. In the following user study, we evaluate the efficacy of muscle highlighting in terms of muscle activation using surface electromyography (sEMG) signal analysis [4, 44].

5.1.3 Mixed Visual Quality. The third important feature of *VR Scene Manager* is to achieve high perceived VR rendering quality while reducing computation workload of onboard CPU/GPU. The key technique to realize this

Fig. 6. Illustration of three different visual quality options: (a) low quality; (b) mixed quality; and (c) high quality. The oblique abdomen of the avatar is highlighted to induce greater muscle activation.

feature is to use high-quality models that have more polygons and shader effects on the target muscle group while rendering the remaining body parts using low-quality models that have moderate to low number of polygons and/or less expensive shader effects. By doing this, the heat generated by the VR HMD is significantly reduced. To illustrate this feature, Figure 6 shows three different visual quality options: (a) low quality; (b) mixed quality; and (c) high quality, where Figure 6 (b) shows a balanced option between visual quality and computation load by mixing a high-quality shader (e.g., a bumped specular shader with lightmap) on the target muscle group and a low-quality shader for the remaining body parts.

5.2 VR User Interface

The *VR User Interface* collects the continuously tracked exercise information from the backend and displays the information on the VR screen in real time. It employs a head-up display in a fixed location, indicating exercise information including repetition count, elapsed time and pace. It also shows an alert indicating whether the pace of the exercise is too fast or too slow based on the recommendations provided by the American College of Sports Medicine (ACSM) [5].

6 EVALUATION

6.1 Experimental Setup

6.1.1 Participants. We have recruited 15 participants (11 males and 4 females) who helped collect data. The participants are university students, researchers, and staffs with ages ranging from 22 to 48 years old (μ = 27.73; σ = 6.65), weights ranging from 42 kg to 85 kg (μ = 60.51; σ = 8.85), heights ranging from 152 cm to 189 cm (μ = 174; σ = 6.50), and gym exercise experience levels ranging from novice to intermediate levels.

6.1.2 Sensor Deployment. To examine the impact of sensor placement locations on the performance of JARVIS, we attached two SensorTag devices onto two different locations on each exercise machine. Figure 7 shows an example of the two SensorTag locations on two machines respectively.

6.1.3 Data Collection. The data were collected at a fitness center on the university campus. During data collection, the participants were instructed to perform 12 considered machine exercises by following the short instructions on each machine. To capture the intra-subject variability, each participant attended three sessions of data collection on three different days. In each session, each participant performed 10 repetitions of each machine exercise, with their preferred weights put on each machine that they consider appropriate for their strength training. In total, each participant contributed 30 repetitions for each exercise.

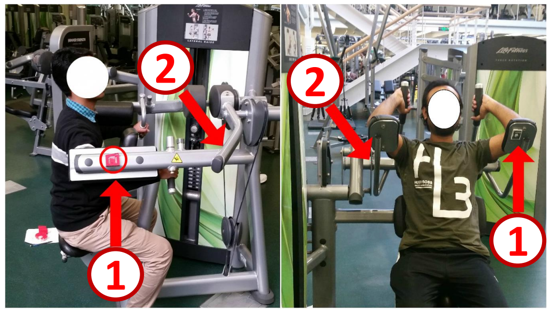


Fig. 7. Illustration of SensorTag placement on the *Lateral Raise* machine (left) and *Seated Abs* machine (right). Circle 1 and 2 indicate the placement locations of two SensorTag devices on each machine.

6.2 Performance of Repetition Segmentation and Counting

6.2.1 Evaluation Metrics. We evaluate the performance of our exercise repetition segmentation scheme using the following three metrics:

- **Miss Rate (MSR).** MSR is defined as the proportion of cases where our scheme misses to detect an exercise repetition.
- **Merge Rate (MGR).** MGR is defined as the proportion of cases where our scheme mistakenly merges two or more exercise repetitions into one repetition.
- **Fragmentation Rate (FR).** FR is defined as the proportion of cases where our scheme mistakenly splits a single exercise repetition into more than one exercise repetitions.

6.2.2 Performance of Repetition Segmentation. Table 3 shows the performance of our exercise repetition segmentation scheme for each type of machine exercises. In terms of MSR, our scheme achieves zero MSR for all machine exercises. In terms of MGR, our scheme achieves zero MGR for all machine exercises except E08 (*Leg Curl*) with a MGR of only 0.11%. Finally, in terms of FR, our scheme achieves zero FR for 6 out of 12 types of machine exercises. Among the other 6 types, the highest FR is only 0.33% for E05 (*Triceps Press*) and E10 (*Row Deltoid*). Taken together, the results indicate that our scheme is able to achieve highly accurate and robust exercise repetition segmentation performance across all types of machine exercises.

Table 3. Performance of repetition segmentation.

Exercise	E 01	E 02	E 03	E 04	E 05	E 06	E 07	E 08	E 09	E 10	E 11	E 12
MSR (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MGR (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00
FR (%)	0.11	0.11	0.11	0.00	0.33	0.00	0.00	0.22	0.00	0.33	0.00	0.00

6.2.3 Performance of Repetition Counting. Based on the repetitions segmented by our scheme, we achieve the repetition counting accuracy of 97.96% out of a total of 5,400 repetitions. Taking a closer look at the counting results, we achieve an accuracy of 99.81% for *within 1* scenario (i.e., ±1 count off compared to the ground truth within a session of 10 repetitions) and 100% for *within 2* scenario.

6.3 Performance of Exercise Type Recognition

6.3.1 Identification of the Best Sensor Placement Locations. We first examine the impact of sensor placement location on the performance of exercise type recognition. Our goal is to find the best sensor placement location combination across all the considered exercise machines. To achieve this goal, we trained three sets of classifiers 1) from 'L' locations only; 2) from 'S' locations only; and 3) from the best location combination, all using leave-one-subject-out cross validation. In particular, classifier 3) is determined by brute-force search among all the possible 4096 sensor placement location combinations.

Table 4 shows the exercise type recognition performance in terms of precision, recall, accuracy, and F1 score. As shown, the best location classifier outperforms the 'L' locations only and 'S' locations only classifiers across all the four metrics by a large margin. Our result demonstrates the impact of sensor placement location on the exercise type recognition performance. More importantly, it can be used to guide the machine manufacturers to find the best sensor placement locations on their exercise machines.

6.3.2 Performance of Session-wise Voting Scheme. Table 4 also shows the exercise type recognition performance with the session-wise voting scheme. As expected, the session-wise voting scheme improves the exercise type recognition performance across all the four metrics. In particular, at the best location, we have achieved a precision of 0.9911 and a recall of 0.9907.

Table 4. Exercise type recognition performance in terms of precision, recall, accuracy, and F1 score at different sensor placement locations with and without session-wise voting.

Location	Without Session-wise Voting				With Session-wise Voting			
	Precision	Recall	Accuracy	F1 Score	Precision	Recall	Accuracy	F1 Score
Location 'L' Only	0.8102	0.8081	80.82%	0.8091	0.8854	0.8809	88.10%	0.8831
Location 'S' Only	0.8114	0.8121	81.22%	0.8117	0.8829	0.8815	88.17%	0.8822
Best Location	0.9432	0.9430	94.30%	0.9431	**0.9911**	**0.9907**	**99.08%**	**0.9909**

6.3.3 Performance of Subject Independent Model. We examine the exercise type recognition performance of the subject independent model. Figure 8 shows the average recognition accuracy across all the machine exercises using leave-one-subject-out cross validation. As shown, 11 out of 15 subjects achieve 100% accuracy while the other 4 subjects achieve accuracy of 97.24%, 97.22%, 97.22%, and 94.45%, respectively. To provide a detailed look at the results, Table 5 lists the confusion matrix in terms of types of machine exercises. As shown, 10 out of 12 machine exercises achieve 100% precision and recall. Taken together, our results indicate that JARVIS is able to achieve high recognition performance across all the 12 machine exercises in a subject independent manner.

Table 5. Confusion matrix of exercise type recognition.

	E01	E02	E03	E04	E05	E06	E07	E08	E09	E10	E11	E12
E01	451	0	0	0	0	0	0	0	0	0	0	0
E02	0	451	0	0	0	0	0	0	0	0	0	0
E03	0	0	451	0	0	0	0	0	0	0	0	0
E04	0	0	0	450	0	0	0	0	0	0	0	0
E05	0	0	0	0	452	0	0	0	0	0	0	0
E06	0	0	0	0	0	450	0	0	0	0	0	0
E07	0	0	0	0	0	20	420	0	10	0	0	0
E08	0	0	0	0	0	0	0	451	0	0	0	0
E09	0	0	0	0	0	0	0	0	450	0	0	0
E10	0	0	0	0	0	0	0	0	0	430	0	20
E11	0	0	0	0	0	0	0	0	0	0	450	0
E12	0	0	0	0	0	0	0	0	0	0	0	451

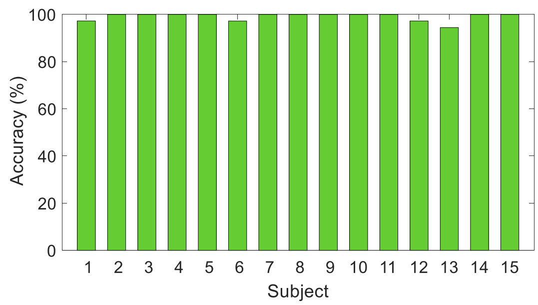


Fig. 8. Performance of subject independent model.

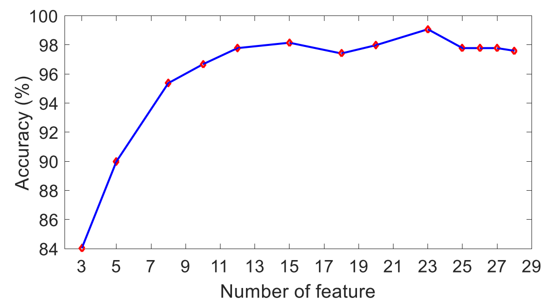


Fig. 9. Accuracy changes over number of features.

6.3.4 Impact of Number of Features. We examine the impact of the number of features on the performance of exercise type recognition. Figure 9 shows the average recognition accuracy as a function of the number of features selected by the sequential floating forward selection (SFFS) algorithm using leave-one-subject-out cross validation. As shown, the recognition accuracy increases in general as the number of features increases. The accuracy reaches the highest when using 23 features out of a total of 28 features. We also observe that the recognition accuracy already reaches 98% when only using 12 features. This result indicates that by using a small number of features, JARVIS can still achieve high recognition performance.

6.3.5 Impact of Amount of Training Data. We analyze the impact of amount of training data in terms of number of subjects as well as number of sessions. First, we use data from different numbers of subjects as training datasets and examine its impact on the performanceof exercise type recognition. Specifically, for one-subject and fourteen-subject cases, we examined all possible fourteen subject selection combinations; for the other cases, we examined 15 random subject selection combinations due to the large number of possible subject selection combinations. Figure 10 (left) shows the average exercise type recognition accuracies in terms of the number of subjects. As shown, the accuracy is 91.39% when trained on data collected from only one subject. The accuracy increases as data from more subjects are added into the training dataset. With training data from six subjects, we have already achieved an accuracy of 99.28%. It should be noted that the there is a drop on accuracy for the fourteen-subject case. We conjecture that this is a result of evaluating a partial subset of all possible combinations.

Next, we use data from different numbers of exercise sessions as training datasets and examine its impact on the performanceof exercise type recognition. Specifically, since each subject performed three exercise sessions on each exercise machine, we used data collected from one session, two sessions, and all the three sessions as the

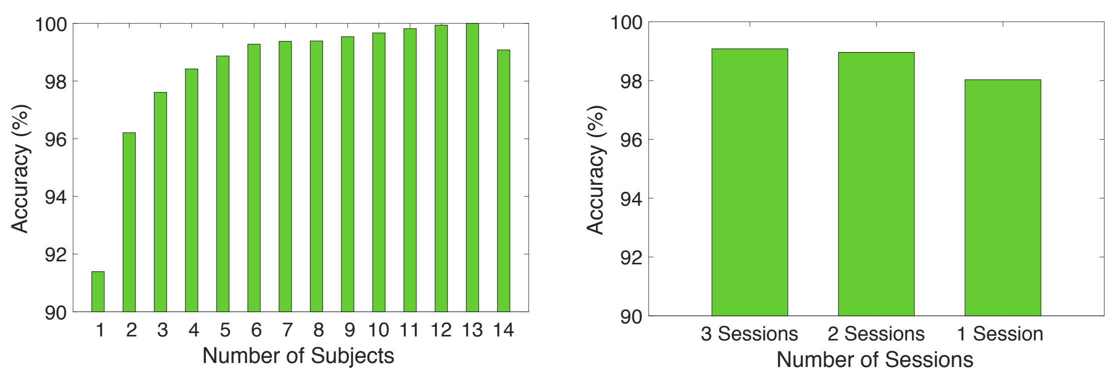


Fig. 10. Impact of amount of training data: the number of subjects (left); and the number of exercise sessions (right).

training datasets, respectively. Figure 10 (right) illustrates the average exercise type recognition accuracies in terms of the number of exercise sessions. As shown, the accuracy is 98.03% when trained on data collected from only one session. The accuracy increases to 98.96% when trained on data collected from two sessions.

6.4 System Performance

We evaluate the system performance of JARVIS in terms of processing time, power consumption, and device temperature. We used a Samsung Galaxy S6 smartphone equipped with a quad-core 2.1 GHz Cortex-A57 processor and 3 GB RAM, mounted to a Samsung Gear VR Headset as our evaluation platform.

6.4.1 Processing Time. To examine the computational demand of JARVIS, we measure the average processing time consumed by each of the key components of the backend *Real-time Exercise Analyzer*. In particular, we used the Android implementation of libSVM[1] for the classification component. We run five sets of trials, with each set including 100 repetitions of different machine exercises from different participants.

Table 6 shows the breakdown of the average processing time for each component. As shown, JARVIS is able to achieve real-time processing performance. Taking a closer look at the result, the classification component takes the most amount of processing time among all the components.

Table 6. Breakdown of the processing time.

Component	Processing Time (ms)
Repetition Segmentation	0.010
Feature Extraction	0.011
Exercise Type Classification	0.073
Total	0.094

6.4.2 Power Consumption. To examine the power consumption of JARVIS, we use the Monsoon Power Monitor[2] to measure the power consumption of both the smartphone mounted to the VR HMD and the SensorTag. The measurement setup is illustrated in Figure 11. During the experiment, we turn off other applications and irrelevant services including GPS, WiFi and cellular services, and continuously transmit sensor data from the SensorTag at 10 Hz to the smartphone through BLE. We profile the power consumption for each component of JARVIS and measure the power consumption for 5 mins and report the averaged values. Table 7 shows the breakdown of power consumption for each component.

Table 7. Breakdown of the power consumption.

Device	Component		Current (mA)	Power (mW)
Smartphone	BLE Communication		61.46	245.7
	Processing Backend		17.65	70.57
	VR Frontend	Low Quality	800.3	3198
		Mixed Quality	815.6	3259
		High Quality	867.9	3468
SensorTag	10 Hz Data Transmission		3.967	11.90

Finally, we estimate the battery lifetimes of both the smartphone and the SensorTag. With a battery capacity of 2550 mAh, JARVIS can run about 2.85 hours on a Galaxy S6 smartphone. With a 240 mAh lithium coin cell battery, the SensorTag can keep sending in sensor data for about 60.5 hours.

[1]http://www.csie.ntu.edu.tw/~cjlin/libsvm/
[2]https://www.msoon.com/LabEquipment/PowerMonitor/

6.4.3 Device Temperature. We measure temperature changes of the smartphone CPU over time under three rendering quality levels: 1) high, 2) mixed, and 3) low, respectively. During rendering, the high, mixed, and low quality level has 155.8k, 30.1k and 13.8k triangles, respectively. We run JARVIS for 30 mins and use the log feature of the Android Debug Bridge to track the temperature changes of the smartphone CPU.

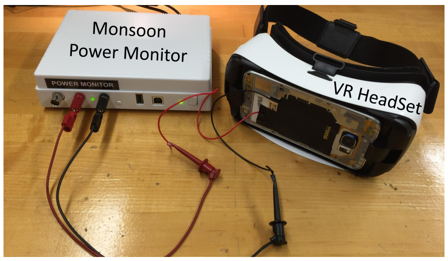


Fig. 11. Power consumption measurement setup with Samsung Galaxy S6, Gear VR and Monsoon Power Monitor.

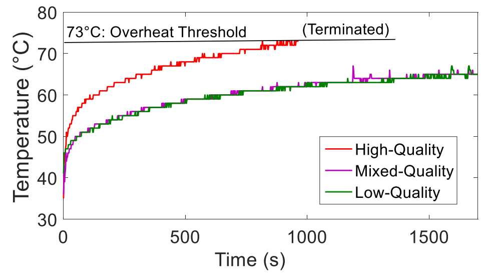


Fig. 12. CPU temperature changes over time. High-Quality mode overheats the smartphone CPU after 15 min.

Figure 12 shows the temperature changes of the smartphone CPU under three rendering quality levels. As shown, under the high quality level, the CPU is overheated up to 73 degree within 15 minutes. As a consequence, JARVIS is terminated for device safety. In contrast, under the low and mixed quality levels, JARVIS is able to run for more than 30 minutes while keeping the CPU temperature below the overheat threshold. This result indicates that our mixed visual quality technique effectively reduces the heat generated during rendering, and thus successfully addresses the overheat problem.

7 USER STUDY

We evaluate the efficacy and overall user experience of JARVIS, and compare it to the traditional machine exercises without JARVIS. Our evaluation methodology consists of a questionnaire survey using Intrinsic Motivation Inventory (IMI), a quantitative analysis of sEMG signals, and an open-ended interview analysis regarding overall user experience.

To evaluate the effectiveness of JARVIS, we use *Seated Abs* as a representative type of machine exercises, due to its health benefits and high chance of being misunderstood by users in terms of its goal. That is, abdominal training is known to be beneficial for building the rectus muscle group around belly button. However, abdominal muscles including *rectus* and *oblique* ones are very sensitively connected to each other, and to lumbar spine as well. Therefore, abdominal muscle training should be performed in a way to build overall muscle balance and lumbar stability [18]. Moreover, although the oblique abdominis muscles are considered to be more important contributors to lumbar stability, abs training without correct guidance predominantly activate the rectus abdominis muscles may lead to imbalance in muscle building [26]. One example of effective intervention for balanced abdominal training is verbal instructions, which showed positive statistically significant differences in terms of balanced abdominis muscle activation [21].

It is worthwhile to note that the degree of muscle recruitment and neuromuscular drive during exercise can be measured and compared by the sEMG activity in sports analytics [4, 44]. Ratio of sEMG activity has been used to evaluate their relative efficacy in eliciting higher levels of muscle recruitment [15]. Specifically, our user study adopts sEMG-based method to evaluate the effectiveness of exercise instructions in terms of oblique abdominis muscle recruitment [21].

7.1 Method

7.1.1 Participants. We recruited 10 university students who are different from the participants in the earlier data collection through department email distribution lists and an on-site recruitment at a fitness center on the university campus. All participants had prior experience with the *Seated Abs* machine exercise.

7.1.2 Data Collection. sEMG data were collected using Trigno wireless sEMG device[3], which consists of one base station and multiple wireless sEMG sensors. Each sEMG sensor has signal bandwidth of 20-450 Hz, transmission range of 40 meters and sampling rate of 4,000 samples/sec with 16-bit resolution. The base station is capable of streaming data to an analysis software over USB wired connection. The transmitted data was saved into computer storage for further analysis.

Following the abdominal exercise instruction study in the literature of orthopedic and sports physical therapy [21], four sEMG sensors were placed at upper/lower rectus abdominis (URA/LRA) as well as external/internal oblique abdominis (EOA/IOA). Figure 13 shows the placements of these four sEMG sensors.

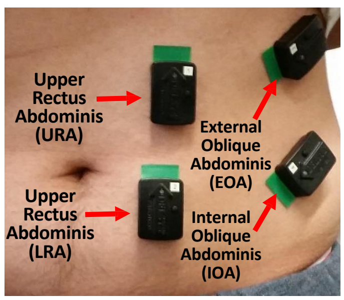


Fig. 13. Illustration of sEMG sensor placement at upper/lower rectus abdominis (URA/LRA), and external/internal oblique abdominis (EOA/IOA).

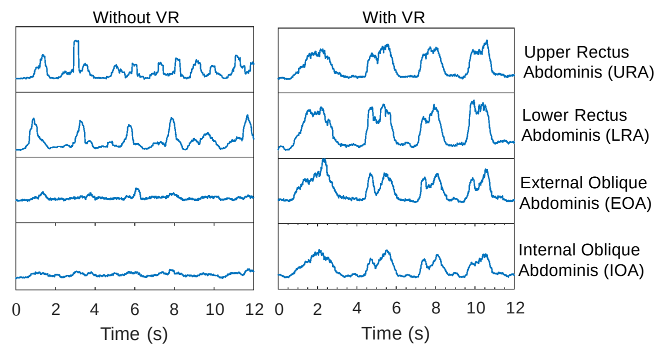


Fig. 14. Without VR vs. With VR in terms of the root-mean-square (RMS) values of sEMG signals.

7.1.3 Study Design. A within-subject design is employed. Participants were instructed in two conditions: one (manipulated) condition with the VR exercise environment emphasizing activation of oblique abdominis, and one (controlled) condition without the VR exercise environment. Prior to actual study participation, participants were told about the basic user interfaces, and were asked to take a look at the VR space with their virtual body representation.

In the controlled condition, the participants are asked to perform the exercise naturally, as they usually do at gyms. In each condition, the participants performed two sets of *Seated Abs* exercises with ten repetitions, respectively. sEMG signal was recorded from the beginning to the end of each set. The participants were allowed to rest for up to two minutes between sets. The sequence of the conditions were balanced across the participants.

After each condition, the participants were asked to respond to 19 selected and adjusted questions from intrinsic motivation inventory (IMI) [37, 41], which is a widely used and extensively validated set of questionnaire to evaluate user experiences. Among five major categories of IMI, we chose three categories related to the application and exercise context: interest/enjoyment, perceived competence, and value/usefulness, consisting of seven, five, and seven questions respectively. The responses were collected in 7-point Likert scale and then averaged in each category for statistical analysis.

After all conditions and surveys, a semi-structured interview followed. We asked the participants to talk about their overall exercise experiences, including the VR application and its interface, the effectiveness of the

[3]http://www.delsys.com/products/wireless-emg/

application, as well as their suggestions in design. All interview data were transcribed and analyzed by open coding [46] and axial coding [19] to discover common themes and patterns.

7.2 Results

7.2.1 sEMG Signal Analysis. Figure 14 shows representative sEMG signals from individual trials performed under conditions without and with JARVIS, respectively. Specifically, it shows that sEMG amplitudes of the abdominis muscles change in response to repetitions. As shown, JARVIS shows its potential in activating the oblique muscles.

Figure 15 summarizes the group means and standard deviations of the normalized sEMG activity of each of the 4 muscle groups. As shown, there is a significant effect of VR application for the two oblique muscle groups ($p < 0.01$ and $p < 0.05$ for external and internal oblique abdominis respectively). This result indicates the potential of effectiveness of JARVIS in activating the target muscle groups.

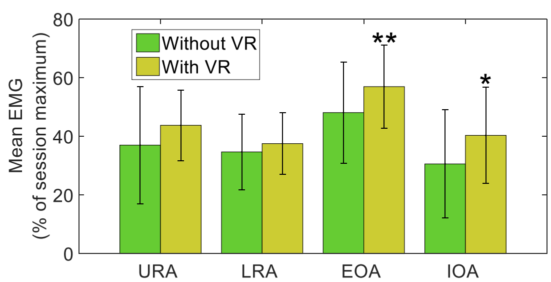


Fig. 15. Mean normalized electormyography values for four muscle groups. With-VR significantly different from Without-VR at oblique abdominis (**$p < .01$; *$p < .05$).

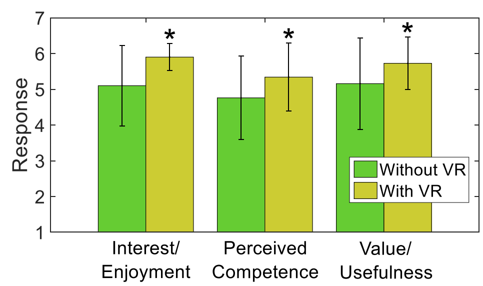


Fig. 16. IMI survey results for three categories. With-VR significantly different from Without-VR (*$p < .05$ for all three categories).

7.2.2 Survey Analysis. Figure 16 shows the summary of IMI questionnaire responses. From all three subcategories of IMI, there was a significant effect of VR application found ($p < 0.05$). In-depth analysis revealed that the students recruited through the mailing list, who had less experiences with strength training machines, showed greater differences in all three categories, than ones hired on-site at the the fitness center. That is, the VR application has a potential in motivating novice or less-experienced exercisers to engage in strength training, implying an opportunity of broader impact to public health.

7.2.3 Interview Analysis. The participants shared their experience while trying JARVIS including its benefits and usefulness, as well as their suggestions to provide more enjoyable machine exercises experiences through JARVIS. They enjoyed and valued the application and its features, which is consistent with the survey results.

Virtual Self and its Movement. The participants liked JARVIS which enabled them to *"see myself moving"* (P8). Seeing their virtually represented body in VR helped them in doing *"the exercise a lot more correctly"* (P6) and in better *"focusing on (my) body"* (P6). Some participants talked about the discrepancy between real self and virtual self, specifically in terms of their appearance: *"[The man in VR] was not exactly the same (to me), but the motion I was doing was like the same with his."* (P4) Interestingly, the discrepancy showed a positive potential toward increased user motivation: *"looking at the muscle man really attracted me to follow him."* (P1) In overall, 'virtual self' contributed to *"more enjoyable"* (P8) exercise experience and user motivation, consistent to the IMI questionnaire survey responses as well as the previous findings in the literature of virtual avatar use and user motivation [9, 51].

Benefits of Muscle Highlighting. As for the visual guide for the muscle groups, the participants found that the muscle highlighting feature was helpful, otherwise they *"would never know how to do it properly"* (P8); *"What the machine is targeting is the muscle, so highlighting (it) really helped me to perform well (and) instructed me."* (P1) It naturally instructed the participants in *"how to (give) focus (on their) abdominal parts"* (P8), and assisted them to better *"conceptualize the way of exercise"* (P10), contributing to the efficacy of the application validated in the muscle activation analysis. We conjecture that the visual highlight of target muscle facilitated greater Mind-muscle connection (MMC) [42], and further studies are encouraged to confirm the relationship.

Benefits of User Interface for Exercise Progress. The participants liked the real-time informative head-up display user interface which provides *"consistent feedback about time and progress"* (P4); *"You could definitely better simply because it tells you the progress of the workout you are doing."* (P7) They recalled how hard it was to *"count the repetitions while focusing on the exercise"* (P2), and talked that *"seeing how many repetitions I had"* (P3) was very helpful.

Points of Improvements and Design Suggestions. The participants provided points of improvements and design suggestions for JARVIS interface design. One participant mentioned a soft competition with another pacemaker avatar or friend online in the virtual environment. This is related with Köhler effect, a well-known theory regarding the benefits of social competition in the literature of sports science [23, 32]; *"I would make a goal to reach to keep it up with a guy in the VR."* (P8) They also wanted to try JARVIS for different machine exercises: *[it would be helpful for] other popular exercises I know like chest press. That will be very helpful because a lot of people do need better form in that. So those compound exercises would definitely help people."* (P10).

8 DISCUSSION

Overall, the contribution of this work is three-folded. First, our work proposes an immersive and interactive virtual exercise assistant based on the combination of IoT and VR technologies, and exemplifies its effectiveness through a pilot user study. We envision that a series of usability and effectiveness studies will follow up toward the realization of effective and practical ubiquitous virtual reality exercise systems and their applications. Second, our work proposes a machine-attachable approach for the placement of motion sensing device, as well as a suite of sensor signal processing algorithms that fully harness the advantages of clearer sensor signals provided by the machine-attachable approach. Third, our work reports the potential of enjoyable, engaging, and effective exercise experiences supported by the significant increases of IMI responses and qualitative interview analysis. The user study also shows the potential effectiveness of JARVIS in facilitating better mind-muscle connection.

In the development of exercise recognition components of JARVIS, leveraging a machine-attachable IoT sensor device brought an advantage of cleaner sensor data. It is expected that the machine-attached sensor would generate cleaner signals while suppressing body movement noises which is one of the major challenges in human activity and movement sensing using IMU sensors [34]. As a result, the repetition segmentation could precisely detect repetitions by detecting local minimum and maximum peaks of the first principle component, eliminating needs of extensive analysis methodologies such as Autocorrelation [27], which requires relatively larger number of repetitions for accurate repetition detection.

The machine-attached sensor also brought the challenge of finding the best sensor placement to achieve the highest accuracy of exercise type detection. That is, because the sensor is placed on a exercise machine, the sensor signals may not fully deliver the unique human body movement characteristics from each exercise, compared to ones from body-worn sensors. As a result, the exercise type recognition algorithm of JARVIS yields only 88.88% of accuracy without the consideration of sensor placement. After finding the best sensor placement combination across the exercise machines, the system achieved 99.08% of accuracy. This approach is comparable to the previous works for finding the best wearable sensor locations for human activity tracking [8, 31], and usability and convenience of sensor placement should be further considered in practical application usage scenarios.

Our evaluation results also revealed the impacts of the amount of training data in terms of number of subjects and number of sessions. As for the number of subjects, the proposed approach yielded 91.39% and 99.28% of accuracy with one and six subjects respectively, showing that reduced number of subjects. The number of sessions employed for training showed 98.96% and 98.03% of accuracy with two and one sessions respectively, implying the room to reduce the burden of training data collection and the marginal differences of data characteristics between sessions. These results have been achieved from the dataset collected while letting the subjects to set their own preferred weights, implying that the proposed approach may have a basic level of robustness against weight variances. Further study is needed to evaluate impacts of weights, perceived load, and user fatigue.

Our user study revealed the potentials of more enjoyable and effective exercise experience enhanced by IoT and VR technologies. For its more practical uses, follow-up studies are needed with close collaboration with researchers in the field of kinesiology, sports science, rehabilitation, and physical therapy. Besides the specific abdomen exercise to reduce lower back pain [21] employed in this study, we believe that a variety of resistance training can benefit from the similar approaches, for example, bench press [45] and elbow flexion (i.e., biceps curl) [25]. To show the potential of the efficacy of specific exercise intervention, we employed sEMG. It helps identify more effective ways of intervention leading to higher level of muscle recruitment and neuromuscular drive [4, 15, 44]. Therefore, sEMG may be employed as a convenient method to evaluate instant muscle recruitment changes, and different methodologies are required to track longitudinal changes of muscle and strength development, such as 1 Repetition Maximum (1RM) [16].

We admit that safety should be the foremost consideration especially in the cases of machine exercises. As mentioned earlier, we aimed to provide immersive VR exercise experience for machine exercises, considering their lower attention demand of body balance and enhanced safety [12] compared to free-weight exercises. However, JARVIS may need more extensive consideration for safety, for example, allowing a user to observe surrounding environment and to proactively handle potential threat. We are currently working on solutions to tackle this challenge. Specifically, we are implementing a mixed-reality (MR) feedback by combining the virtual reality with surrounding environment through the pass-through camera of HMD. In the longer term, JARVIS will be able to run on MR devices such as Microsoft Hololens, and the current technical components of JARVIS will be able to support it with minor frontend modifications such as augmented muscle highlighting on a real body.

For future work, we plan to evaluate the efficacy of JARVIS for other exercise types with a broader population. We also plan to extend JARVIS as a full-pledged gym context monitoring and fitness management platform and design entertainment-oriented games on top of it. Finally, JARVIS represents our initial effort toward ubiquitous mixed reality [35]. We plan to extend JARVIS to enable truly ubiquitous mixed reality experiences in the future.

9 CONCLUSION

In this paper, we present the design, implementation, and evaluation of JARVIS, a virtual exercise assistant based on a miniature IoT sensing device and a mobile VR headset to enable immersive and interactive gym exercise experience. The realization of such virtual experience requires the VR headset to retrieve accurate information of the machine exercise performed by the user in real time. JARVIS achieves this by attaching a miniature IoT sensing device on gym machines and developing a suite of lightweight sensor signal processing algorithms to recognize exercise type and track exercise progress. Based on the extracted exercise information, JARVIS creates an immersive and interactive gym exercise experience. With JARVIS, we envision that a series of usability and effectiveness studies will follow up toward the realization of effective and practical ubiquitous virtual reality exercise system and applications.

10 ACKNOWLEDGEMENT

We would like to thank the anonymous reviewers for their valuable reviews and insightful comments. This research was partially funded by NSF awards #1565604 and #1617627.

REFERENCES

[1] Htc vive. In *https://www.vive.com/us/*.
[2] Oculus rift. In *https://www.oculus.com/en-us/*.
[3] Samsung gear vr. In *http://samsung.com/galaxy/wearables/gear-vr*.
[4] P. Aagaard, E. B. Simonsen, J. L. Andersen, P. Magnusson, and P. Dyhre-Poulsen. Increased rate of force development and neural drive of human skeletal muscle following resistance training. *Journal of applied physiology*, 93(4):1318–1326, 2002.
[5] K. Adams, E. Cafarelli, A. Gary, C. Dooly, S. Matthew, S. J. Fleck, A. C. Fry, J. R. Hoffman, R. U. Newton, J. Potteiger, M. H. Stone, N. A. Ratamess, and T. Triplett-mcbride. Progression models in resistance training for healthy adults, 2009.
[6] G. Allison, P. Godfrey, and G. Robinson. Emg signal amplitude assessment during abdominal bracing and hollowing. *Journal of Electromyography and Kinesiology*, 8(1):51–57, 1998.
[7] F. Anderson, T. Grossman, J. Matejka, and G. Fitzmaurice. Youmove: enhancing movement training with an augmented reality mirror. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pages 311–320. ACM, 2013.
[8] L. Atallah, B. Lo, R. King, and G.-Z. Yang. Sensor positioning for activity recognition using wearable accelerometers. *IEEE transactions on biomedical circuits and systems*, 5(4):320–329, 2011.
[9] A. L. Baylor. Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1535):3559–3565, 2009.
[10] J. Calatayud, J. Vinstrup, M. D. Jakobsen, E. Sundstrup, M. Brandt, K. Jay, J. C. Colado, and L. L. Andersen. Importance of mind-muscle connection during progressive resistance training. *European journal of applied physiology*, 116(3):527–533, 2016.
[11] K.-H. Chang, M. Y. Chen, and J. Canny. *Tracking free-weight exercises*. Springer, 2007.
[12] M. L. Cotterman, L. A. Darby, and W. A. Skelly. Comparison of muscle force production using the smith machine and free weights for bench press and squat exercises. *Journal of Strength and Conditioning Research*, 19(1):169, 2005.
[13] J. Deber, R. Jota, C. Forlines, and D. Wigdor. How much faster is fast enough?: User perception of latency & latency improvements in direct and indirect touch. In *Proceedings of the Annual ACM Conference on Human Factors in Computing Systems*, pages 1827–1836, 2015.
[14] H. Ding, L. Shangguan, Z. Yang, J. Han, Z. Zhou, P. Yang, W. Xi, and J. Zhao. Femo: A platform for free-weight exercise monitoring with rfids. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, pages 141–154. ACM, 2015.
[15] C. T. A. ER and S. M. MCILL. Low back loads over a variety of abdominal exercises: searching for the safest abdominal challenge. 1997.
[16] S. J. Fleck and W. Kraemer. *Designing Resistance Training Programs, 4E*. Human Kinetics, 2014.
[17] C. Frankish, P. Morgan, and R. Hull. Recognition accuracy and usability of pen-based interfaces. In *IEE Colloquium on Interfaces-The Leading Edge (Digest No. 1996/126)*, pages 7–1. IET, 1996.
[18] M. G. Gardner-Morse and I. A. Stokes. The effects of abdominal muscle coactivation on lumbar spine stability. *Spine*, 23(1):86–91, 1998.
[19] K. Holtzblatt and S. Jones. Contextual inquiry: A participatory technique for system design. *Participatory design: Principles and practices*, pages 177–210, 1993.
[20] A. Jain and D. Zongker. Feature selection: Evaluation, application, and small sample performance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(2):153–158, 1997.
[21] G. M. Karst and G. M. Willett. Effects of specific exercise instructions on abdominal muscle activity during trunk curl exercises. *The Journal of orthopaedic and sports physical therapy*, 34(February 2001):4–12, 2004.
[22] L. Karvitz. Exercise Motivation: What Starts and Keeps People Exercising. *University of New Mexico*, 2011.
[23] O. Köhler. Über den Gruppencrirkungsgrad der menschlichen Körperarbeit und die Bedingung optimaler kollektivkraftreaktion [Human physical performance in groups and conditions for optimal collective performance]. *Industrielle Psychotechnik*, 4:209–226., 1927.
[24] C. L. Lox, K. A. M. Ginis, and S. J. Petruzzello. *The psychology of exercise: Integrating theory and practice*. 2006.
[25] D. C. Marchant, M. Greig, and C. Scott. Attentional focusing instructions influence force production and muscular activity during isokinetic elbow flexions. *The Journal of Strength & Conditioning Research*, 23(8):2358–2366, 2009.
[26] M. I. Miller and J. M. Medeiros. Recruitment of internal oblique and transversus abdominis muscles during the eccentric phase of the curl-up exercise. *Physical therapy*, 67(8):1213–1217, 1987.
[27] D. Morris, T. S. Saponas, A. Guillory, and I. Kelner. Recofit: using a wearable sensor to find, recognize, and count repetitive exercises. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3225–3234. ACM, 2014.
[28] B. J. Mortazavi, M. Pourhomayoun, G. Alsheikh, N. Alshurafa, S. I. Lee, and M. Sarrafzadeh. Determining the single best axis for exercise repetition recognition and counting on smartwatches. In *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*, pages 33–38. IEEE, 2014.
[29] M. Muehlbauer, G. Bahle, and P. Lukowicz. What can an arm holster worn smart phone do for activity recognition? In *Wearable Computers (ISWC), 2011 15th Annual International Symposium on*, pages 79–82. IEEE, 2011.
[30] C. Munteanu, R. Baecker, G. Penn, E. Toms, and D. James. The effect of speech recognition accuracy rates on the usefulness and usability of webcast archives. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 493–502. ACM, 2006.

[31] D. O. Olguın and A. S. Pentland. Human activity recognition: Accuracy across common locations for wearable sensors. In *Proceedings of 2006 10th IEEE International Symposium on Wearable Computers, Montreux, Switzerland*, pages 11–14. Citeseer, 2006.
[32] K. a. Osborn, B. C. Irwin, N. J. Skogsberg, and D. L. Feltz. The Köhler effect: Motivation gains and losses in real sports groups. *Sport, Exercise, and Performance Psychology*, 1(4):242–253, 2012.
[33] T. Park, I. Hwang, U. Lee, S. I. Lee, C. Yoo, Y. Lee, H. Jang, S. P. Choe, S. Park, and J. Song. Exerlink: enabling pervasive social exergames with heterogeneous exercise devices. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, pages 15–28. ACM, 2012.
[34] T. Park, J. Lee, I. Hwang, C. Yoo, L. Nachman, and J. Song. E-gesture: a collaborative architecture for energy-efficient gesture recognition with hand-worn sensor and mobile devices. In *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*, pages 260–273. ACM, 2011.
[35] T. Park, M. Zhang, and Y. Lee. When mixed reality meets internet of things: Toward the realization of ubiquitous mixed reality. *ACM GetMobile: Mobile Computing and Communications*, 22(1), 2018.
[36] I. Pernek, K. A. Hummel, and P. Kokol. Exercise repetition detection for resistance training based on smartphones. *Personal and ubiquitous computing*, 17(4):771–782, 2013.
[37] R. W. Plant and R. M. Ryan. Intrinsic motivation and the effects of self consciousness, self-awareness, and ego involvement: an investigation of internally controlling styles. *Journal of Personality*, 53(3):435–449, 1985.
[38] C. Richardson and G. Jull. Muscle control–pain control. what exercises would you prescribe? *Manual therapy*, 1(1):2–10, 1995.
[39] A. Rizzo, J. Difede, B. O. Rothbaum, J. M. Daughtry, and G. Reger. Virtual reality as a tool for delivering ptsd exposure therapy. *Post-Traumatic Stress Disorder: Future Directions in Prevention, Diagnosis, and Treatment*, 2013.
[40] C. Rojek. *Leisure and culture*. Macmillan Press, 2000.
[41] R. M. Ryan, J. P. Connell, and R. W. Plant. Emotions in nondirected text learning. *Learning and Individual Differences*, 2(1):1–17, 1990.
[42] B. J. Schoenfeld and B. Contreras. Attentional Focus for Maximizing Muscle Development. *Strength and Conditioning Journal*, 38(1):27–29, 2016.
[43] C. Seeger, A. Buchmann, and K. Van Laerhoven. myhealthassistant: a phone-based body sensor network that captures the wearer's exercises throughout the day. In *Proceedings of the 6th International Conference on Body Area Networks*, pages 1–7. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011.
[44] O. Shirado, T. Ito, K. Kaneda, and T. E. Strax. Electromyographic analysis of four techniques for isometric trunk muscle exercises. *Archives of physical medicine and rehabilitation*, 76(3):225–229, 1995.
[45] B. J. Snyder and W. R. Fry. Effect of verbal instruction on muscle activity during the bench press exercise. *The Journal of Strength & Conditioning Research*, 26(9):2394–2400, 2012.
[46] A. Strauss and J. Corbin. *Basics of qualitative research: grounded theory procedure and techniques*, volume 13. Sage, London, 1990.
[47] E. Velloso, A. Bulling, and H. Gellersen. Motionma: motion modelling and analysis by demonstration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1309–1318. ACM, 2013.
[48] E. Velloso, A. Bulling, H. Gellersen, W. Ugulino, and H. Fuks. Qualitative activity recognition of weight lifting exercises. In *Proceedings of the 4th Augmented Human International Conference*, pages 116–123. ACM, 2013.
[49] T. Wakahara, A. Fukutani, Y. Kawakami, and T. Yanai. Nonuniform muscle hypertrophy: Its relation to muscle activation in training session. *Medicine and Science in Sports and Exercise*, 45(11):2158–2165, 2013.
[50] T. Wakahara, N. Miyamoto, N. Sugisaki, K. Murata, H. Kanehisa, Y. Kawakami, T. Fukunaga, and T. Yanai. Association between regional differences in muscle activation in one session of resistance exercise and in muscle hypertrophy after resistance training. *European Journal of Applied Physiology*, 112(4):1569–1576, 2012.
[51] N. Yee. The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and virtual environments*, 15(3):309–329, 2006.
[52] M. Zhang and A. A. Sawchuk. A feature selection-based framework for human activity recognition using wearable multimodal sensors. In *Proceedings of the 6th International Conference on Body Area Networks*, pages 92–98. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011.
[53] M. Zhang and A. A. Sawchuk. Motion primitive-based human activity recognition using a bag-of-features approach. In *Proceedings of the 2Nd ACM SIGHIT International Health Informatics Symposium*, IHI '12, pages 631–640. ACM, 2012.

Received November 2017; revised February 2018; accepted June 2018