Device-free Personalized Fitness Assistant Using WiFi

XIAONAN GUO, Indiana University-Purdue University Indianapolis
JIAN LIU, WINLAB, Rutgers University
CONG SHI, Stevens Institute of Technology
HONGBO LIU, Indiana University-Purdue University Indianapolis
YINGYING CHEN, WINLAB, Rutgers University
MOOI CHOO CHUAH, Lehigh University

There is a growing trend for people to perform regular workouts in home/office environments because work-at-home people or office workers can barely squeeze in time to go to dedicated exercise places (e.g., gym). To provide personalized fitness assistance in home/office environments, traditional solutions, e.g., hiring personal coaches incur extra cost and are not always available, while new trends requiring wearing smart devices around the clock are cumbersome. In order to overcome these limitations, we develop a device-free fitness assistant system in home/office environments using existing WiFi infrastructure. Our system aims to provide personalized fitness assistance by differentiating individuals, automatically recording fine-grained workout statistics, and assessing workout dynamics. In particular, our system performs individual identification via deep learning techniques on top of workout interpretation. It further assesses the workout by analyzing both short and long-term workout quality, and provides workout reviews for users to improve their daily exercises. Additionally, our system adopts a spectrogram-based workout detection algorithm along with a Cumulative Short Time Energy (CSTE)-based workout segmentation method to ensure its robustness. Extensive experiments involving 20 participants demonstrate that our system can achieve a 93% accuracy on workout recognition and a 97% accuracy for individual identification.

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

Additional Key Words and Phrases: personalized fitness assistant, device-free, channel state information

1 INTRODUCTION

Nowadays, people pay more attention to their physical health because the accelerated pace of life and increasing work pressure have compelled many people to adapt to a sedentary lifestyle, either sitting all day working on computers or spending long hours with phones. Such sedentary lifestyles often lead to many chronic illnesses (e.g., obesity), which have significantly negative impacts on the quality of people's lives. As a result, there is an increasing trend for people to exercise regularly. However, there are few workout options for the work-at-home

Authors' addresses: Xiaonan Guo, Indiana University-Purdue University Indianapolis, IN, USA, 46202, xg6@iupui.edu; Jian Liu, WINLAB, Rutgers University, NJ, USA, 08901, jianliu@winlab.rutgers.edu; Cong Shi, Stevens Institute of Technology, NJ, USA, 07030, cshi5@stevens.edu; Hongbo Liu, Indiana University-Purdue University Indianapolis, IN, USA, 46202, hl45@iupui.edu; Yingying Chen, WINLAB, Rutgers University, NJ, USA, 08901, yingche@scarletmail.rutgers.edu; Mooi Choo Chuah, Lehigh University, PA, USA, 18015, chuah@cse.lehigh.edu.

people or office workers due to their difficulties on squeezing time to exercise at dedicated places (e.g., gym). Thus, many people gradually resort to home/office, which is more convenient without space and time constraints, to exercise and maintain their health. Furthermore, people also would like to record his/her exercise details to keep track of their fitness plans. For example, a user performs various exercises such as free-weight with dumbbells, stretch, riding a stationary bike and running on a treadmill, and the corresponding exercise statistics could infer meaningful health-related information (e.g., calories burned). It is common that people cannot always stick to their fitness plans due to the lack of guidance from exercise experts in home/office environments. Therefore, in this paper, we take one step further to develop a smart exercise assistant system that facilitates personal workout monitoring and assessment targeting home/office environment.

Traditional solutions on workout assistance rely on hiring personal coaches, which incur high cost, and are mostly not available in home/office environments. There is a new trend of utilizing smartphones and fitness trackers [1] to perform workout monitoring including step counting and multi-sport tracking. These approaches require users to wear smart devices all the time and cannot judge whether the users perform exercise correctly. The details on workout monitoring are critical to help the people correct their exercise postures, and thereby achieve their fitness goals. Incorrect postures may cause unnecessary injury or degrade the efficiency of muscle building. Recently, dedicated hardware and sensors (i.e., inertial sensors, RFID tags [11]) are deployed to track users' exercise records. In particular, FEMO [11] proposes a free-weight exercise monitoring system by attaching passive RFID tags on the dumbbells and leveraging the Doppler shift profile. FitCoach [12] develops a workout recognition and review scheme leveraging the inertial sensors embedded in smartphones or smartwatches. However, it is hard to require people, especially seniors, to wear these devices during exercises, and it is desirable to automatically recognize workout and provide recommendations. Therefore, we seek a solution that is device-free to provide fine-grained personalized fitness assistance. We find that Channel State Information (CSI) [13, 14, 17, 31], embedded in WiFi signals, is a desirable candidate to facilitate workout interpretation and workout recommendation without attaching any devices to users. Because of the prevalence of WiFi infrastructures in home/office environments, it is possible to capture people's activities without their participation. Furthermore, to provide a personalized fitness assistance, it is critical to identify each individual since people may choose a shared environment, either in a family environment or in an office space, to exercise. Therefore, our proposed system should derive salient features that can perform individual identification even at a shared space.

In order to extract unique features from CSI that can identify each individual's movement, we resort to deep learning technique, which has demonstrated its capability in other fields such as image processing [29, 33] and natural language processing [10, 24]. We believe the multi-layer design in deep learning [16, 25] could provide distinctive features for individual identification and workout interpretation. Furthermore, people perform non-workout activities throughout the day such as typing in front of computers and answering phone calls, so it is important that the proposed system has the capability to distinguish workout activities from non-workout ones. We propose to differentiate workout activities by searching for repetitive patterns in CSI readings using an autocorrelation-based method. Additionally, CSI is sensitive to various environmental factors (e.g., environmental noise and locations), thus our system should be robust to external interference other than workout activities. In particular, to adapt to small location changes (e.g., people may not always exercise at the same spot even in the same location), the exercise training profiles are collected around a certain range of a profile location and hence make the trained model more robust. Finally, the system aims to assess workout dynamics to boost the workout efficiency and avoid unnecessary injury. We propose to measure workout intensity and strength by defining two new metrics, and thereby provide desirable workout recommendations to users.

To summarize, our system provides two high-level functionalities to enable personalized fitness assistance: **(1) Personalized workout interpretation.** We develop a Deep Neural Network (DNN) based model to provide personalized workout interpretation with two tiers of information. Specifically, in the first tier, our system will identify workout types and estimate workout frequencies. Then, the second tier leverages fine-grained latent

representatives to further identify individuals. The abstract representation in the deeper layer of a DNN model is beneficial to our system in differentiating individuals because individual identification requires more complex knowledge than workout recognition. **(2) Smart workout assessment.** Our system analyzes and assesses a user's workout in terms of its intensity and strength. Specifically, we define two new metrics that depict the short-term and long-term pictures of a user's workout pattern, respectively. *Work-to-Rest-Ratio* measures the complete cycle of a repetition within one set and Repetition-tempo-rate represents the tempo (or speed) at which a user performs a repetition. Then, our system provides workout trends to each user in terms of these two metrics, which the user can correct the exercise form accordingly. This method could also be an accompaniment to just watching and following fitness videos [3] from professionals.

The contributions of our work are summarized as follows:

- We develop a device-free fitness assistant system for home/office scenarios by utilizing existing WiFi infrastructure without active user participation. The system can differentiate individuals to enable personalized fitness assistance with comprehensive workout analysis and smart workout assessment.
- We propose a deep learning-based approach that identifies individuals on top of users' workout interpretations leveraging fine-grained latent representations. This is important for home/office environments where family members and colleagues may share the same space.
- We provide a quantitative workout analysis framework to infer various fine-grained information including the workout type, the number of sets and repetitions. A spectrogram-based method is adopted to offer accurate exercise segmentation.
- Our system is capable of providing smart fitness recommendation by automatically assessing exercise postures and offering feedback to a user for posture correction. It can also be used as an early alert of possible physical health issues.
- We deploy our system prototype with a pair of laptops and conduct extensive experiments with 20 people during a period of 10 months. We demonstrate that our system can achieve high accuracy for both individual identification and exercise recognition, as well as provide smart workout assessment.

2 RELATED WORK

In general, activity recognition and fitness assistant systems can be broadly categorized as motion sensor-based, vision-based and RF-based. In this part, we review existing studies and compare with our personalized fitness assistant system.

Motion sensor-based. Motion sensors (i.e., accelerometer and gyroscope) embedded in body-attached mobile/wearable devices have been used for both activity recognition [9, 18] and exercise recognition [7, 12]. Specifically, Nuactive [9] and Lasagna [18] recognize human activities through semi-supervised learning and deep neural networks on the captured motion sensor readings, respectively. Additionally, the sensors in customized workout gloves [7] and wearable devices (e.g., smartwatches) [12] are used to recognize exercise types and the number of repetitions. These approaches, however, require users to wear additional devices and are limited to recognize the exercise movements only involving upper body parts such as wrist and arm.

Vision-based. To overcome the aforementioned weaknesses, a couple of studies attempt to recognize human activities with the help of cameras [23, 32]. These approaches either use unsupervised learning or pattern matching on the recorded video frames to achieve fine-grained activity recognition. In addition, video analysis has also been applied to measuring workout dynamics and providing exercise guidance [6, 26]. For instance, Su *et al.* [26] develop a workout recognition system which assists patients in conducting rehabilitation exercises through profile matching on Kinect readings. Celiktutan *et al.* [6] develop an exercise recognition and assessment scheme which tracks skeleton geometries with Kinect and quantifies the action goodness. However, these vision-based schemes only work when a user stays within the device's line-of-sight zone and can be significantly affected by the environmental light conditions. In addition, video-based techniques also bring privacy concerns.

RF-based. Recently, RF-based sensing attracts a considerable attention of many researchers [2, 11, 22] due to the prevalent usage of wireless-enabled devices. WiSee [22] and WiTrack [2] can recognize human gestures and track $3D$ motion of a user using the minute Doppler shift and the round-trip time of flight (ToF) of wireless signals, respectively. Additionally, FEMO [11] recognizes and assesses free-weight workouts by attaching passive RFID tags on the dumbbells and analyzing the Doppler shifts of the received signals. These approaches, however, require dedicated devices (e.g., Universal Software Radio Peripheral (USRP), RFID readers) which limits their applications in practice. Recent studies have shown the potential of using off-the-shelf WiFi to recognize human daily activities [30, 31]. They examine CSI in WiFi signals to capture the environmental changes associated with human activity motions. Although existing WiFi based approaches show intriguing results in identifying general activity types (e.g., walking, standing, sitting), the problem of assessing workout qualities (e.g., strength, speed) and providing exercise feedback remains open. Additionally, comparing with the existing activity-recognition studies (e.g., [30, 31]), which mainly focus on recognizing location-oriented daily activities (e.g., cooking in a kitchen, brushing teeth in a bathroom), this work explores the feasibility of using WiFi to recognize various exercise activities at the same or nearby place and further identify the individual who performs the exercises.

Different from existing work, we propose a device-free personalized exercise recognition and assessment scheme leveraging existing WiFi infrastructures. The proposed system provides comprehensive workout interpretation, and differentiates individuals without active user participation. Specifically, we examine channel state information (CSI) embedded in WiFi readings which capture workout dynamics of both free weight and body stretching exercises. Moreover, a workout assessment algorithm is developed to evaluate exercise quality by examining motion speed and strength through frequency domain analysis.

3 SYSTEM DESIGN

3.1 Challenges

In order to design a system that enables personalized fitness assistance using WiFi, a number of challenges regarding how to design mechanisms for comprehensive workout interpretation and smart workout assessment need to be addressed:

Fitness Individual Differentiation. It is common that family members or colleagues take turns to perform workout at a shared environment (e.g., home/office). To realize personalized fitness assistance, our system should have the capability of differentiating different individuals. This is challenging especially for a device-free system facilitating personalized fitness assistance, since no dedicated device is attached to each individual to capture signals associated with his/her corresponding movements. We seek a general solution that can work for a group of people in such a shared environment. This requires our system to differentiate individuals and offer personalized workout interpretation and assessment.

Fine-grained Workout Interpretation. Providing fine-grained workout interpretation using WiFi is also challenging. First, people perform non-workout activities throughout the day. Thus, the proposed system should have the capability of distinguishing workout related activities from non-workout ones. Second, in order to provide fine-grained workout interpretation, our system needs to accurately decompose an exercise into pieces of repetitions. Thus, a robust CSI segmentation scheme needs to be integrated into our system. Third, our system also needs to handle a common case where a specific user may perform workout at a slightly shifted location (e.g., one foot away) from the location where his training profiles are collected. This is challenging because CSI readings are sensitive to location changes, which could significantly degrade the system performance if careful design consideration is not taken.

Smart Workout Assessment. Incorrect and inconsistent workout activities will degrade the efficiency of muscle building and cause unnecessary injuries. However, it is challenging to design mechanisms for workout assessment using CSI readings in a device-free design. To overcome such challenge, we define two workout-related
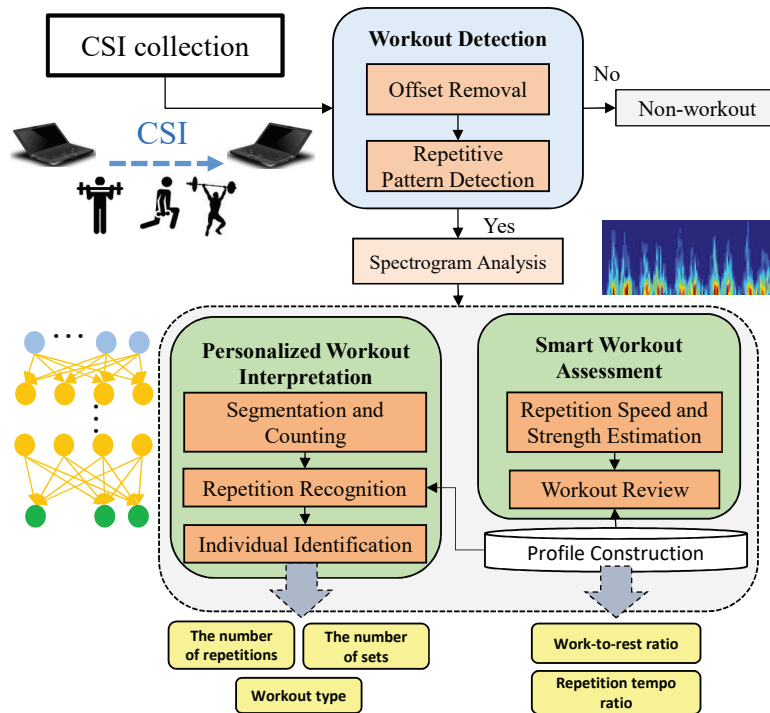
Fig. 1. Proposed system flow.

metrics, one at the repetition-level and the other considering the perspective of a whole exercise to provide workout assessment. Such assessment can be used to provide users with feedback to correct their postures in subsequent exercises.

3.2 System Overview

In this work, we develop a personalized device-free fitness assistant system utilizing the existing WiFi infrastructure. The key idea is to use the fine-grained CSI, which is readily available from commodity WiFi devices, to accurately depict personalized workout statistics and assess workout dynamics. We first define two specific terms characterizing the workout activity: *rep* and *set*. A rep (repetition) is one complete motion cycle of a particular exercise. A set is a group of consecutive repetitions. Typically, a user is expected to perform a workout exercise consisting of several sets to achieve efficient muscle building and maintain a fit body. Given these inherent characteristics of a workout, we anticipate that we can observe repetitive CSI patterns because of the repetitive movements from different body parts during workouts. Furthermore, the minute movement of different workouts have distinct impacts on CSI and thus these unique features of a particular workout can be extracted via CSI analysis.

As illustrated in Figure 1, the proposed system takes as input CSI readings which can be directly extracted from WiFi cards [13]. Our system first performs *Workout Detection* to differentiate workout-related activities from non-workout ones based on the discovery of periodical patterns. Specifically, it first applies a sliding window on the CSI time series and then removes the offset (e.g., mean, linear or polynomial trend) of the corresponding CSI data through *Offset Removal*. After that, *Repetitive Pattern Detection* further detects workout activities based on an autocorrelation calculation of the CSI readings. Once the workout-related activities are identified from

(a) CSI amplitude of one subcarrier with corresponding activity time frame.

(b) Workout raw data

(c) Offset removal
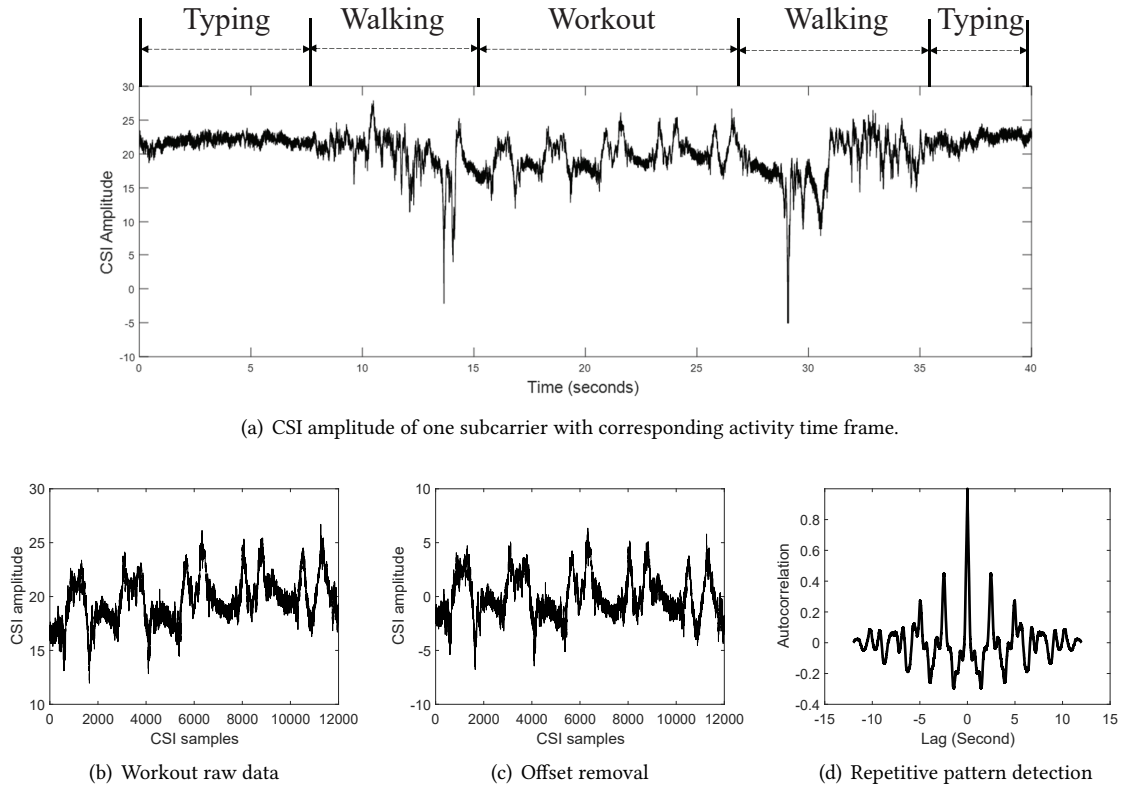
(d) Repetitive pattern detection

Fig. 2. An example illustrates the process of workout detection.

CSI readings, the system will then perform *Spectrogram Analysis* by converting the CSI time series to frequency-domain. The spectrogram of CSI readings reflects how the workout energy of each frequency component evolves with time.

Next, our system performs *Personalized Workout Interpretation* and *Smart Workout Assessment*. In particular, the *Personalized Workout Interpretation* performs quantitative analysis of CSI readings to produce workout statistics, which can be used to track if a user follows his personal fitness plan correctly. It contains three modules: *Segmentation and Counting*, *Repetition Recognition* and *Individual Identification*. The *Segmentation and Counting* counts the exercise repetitions based on the cumulative power spectral density analysis of CSI spectrogram. *Repetition Recognition* provides fine-grained exercise recognition at the repetition level via a DNN-based method. Then, our system differentiates different individuals through the *Individual Identification* module (i.e., the deeper layer of the proposed DNN model). The *Smart Workout Assessment* aims to provide workout feedback to the users. In particular, besides desiring to check fitness statistics, a user also cares about his workout quality to ensure the muscle building efficiency from his workout and avoid possible risks of injury. The workout assessment consists of two modules, *Repetition Speed and Strength Estimation* and *Workout Review*. *Repetition Speed and Strength Estimation* focus on the estimation of the workout speed and strength at the repetition level, which are especially useful in both quality assessment and abnormal workout detection (e.g., a senior may have a sudden change in his workout speed due to his physical health situation). *Workout Review* module provides workout recommendations to users through workout assessment at repetition level. Before our system is deployed, we collect the CSI traces

as the training profiles from the experts who can perform different exercise types with standard intensities and strengths. Then, the system compares each repetition of users' exercise profiles against their standard training profiles or experts who visit gym frequently so that it can provide desirable feedback to the users.

4 WORKOUT RECOGNITION

In this part, we will discuss the details of the proposed system on differentiation between workout and non-workout activities, and exercise repetition recognition.

4.1 Differentiation between Workout and Non-workout

In home/office environments where people conduct different daily activities, CSI readings related to non-workout activities (e.g., typing in front of a computer, walking around) will also be captured beside the workout activities. In order to achieve personalized fitness assistance, it is critical for our system to first differentiate workout from other activities. To identify unique features of CSI from workout-related activities, we conduct a preliminary experiment by recording a series of activities of a specific individual in an office environment as follow: A person is typing in front of a table and then walks to a location where he performs five repetitions of standing biceps curl. After that, he walks back to the same table and continues typing. Figure 2(a) shows the corresponding CSI amplitude of one subcarrier (i.e., 5th subcarrier) with the marked time frame of each activity. We observe obvious repetitive patterns in CSI readings during workouts while none for non-workout activities.

Inspired by the above observations, we propose to detect workout activities by searching for the repetitive patterns in CSI readings with autocorrelation method. Autocorrelation is used to provide a measure of similarity between a signal and itself at a given lag, and thus a local maximum is expected if a repetitive pattern exists. Specifically, the detection process of repetitive patterns is as follows: our system first applies a sliding window on the time series of CSI amplitude. Figure 2(b) shows an example of an 8-second sliding window which contains the raw CSI readings associated with a workout. Next, we remove the offset (e.g., mean, linear trend or polynomial trend) of the raw readings within this window. Specifically, we first fit a low-order polynomial to the raw CSI data and subtract it together with the mean value from the raw CSI readings. Figure 2(c) denotes the data after offset removal. Figure 2(d) shows the results of repetitive pattern detection. Then, a peak (i.e., local maximum) detection algorithm is adopted with an empirical threshold (i.e., 0.2 in this example) to derive the number of repetitions by $N_r = \dfrac{N_p - 1}{2}$, where N_r is the number of repetitions and N_p is the number of peaks from autocorrelation. Finally, the workout is detected within the specific sliding-window when N_r is larger than a threshold T. Here we assume the workout activities have at least 3 repetitions, thus T is fixed at 3 in our prototype system if not mentioned otherwise.

4.2 Exercise Repetition Recognition

After workout activity is detected, our system will recognize each exercise repetition to further provide workout statistics. Workout statistics such as the number of set and repetition imply meaningful exercise effects (e.g., calories burned), and are crucial to guide the users to design, follow and adjust their fitness plans. In the proposed design, we adopt spectrogram analysis instead of time domain analysis which is usually negatively impacted by the presence of interferences and noises and consistent CSI fluctuations. To validate the effectiveness of spectrogram analysis, we conduct several preliminary experiments by asking volunteers to perform four sets of *lateral raise* with five repetitions per set. The corresponding spectrogram in time-frequency plane presents a clear repetitive pattern of workout activities in Figure 3(a). We then adopt a *C*umulative *S*hort *T*ime *E*nergy (CSTE) method to provide workout statistics and further accurately segment each exercise at repetition level. Specifically, our system first accumulates all power spectral density (PSD) along the frequency dimension in the spectrogram as shown in Figure 3(b). Next, the energy of the cumulative PSD is accumulated again in short sliding window

(a) Spectrogram of lateral raise

(b) Corresponding Cumulative Power Spectral Density (CPSD)

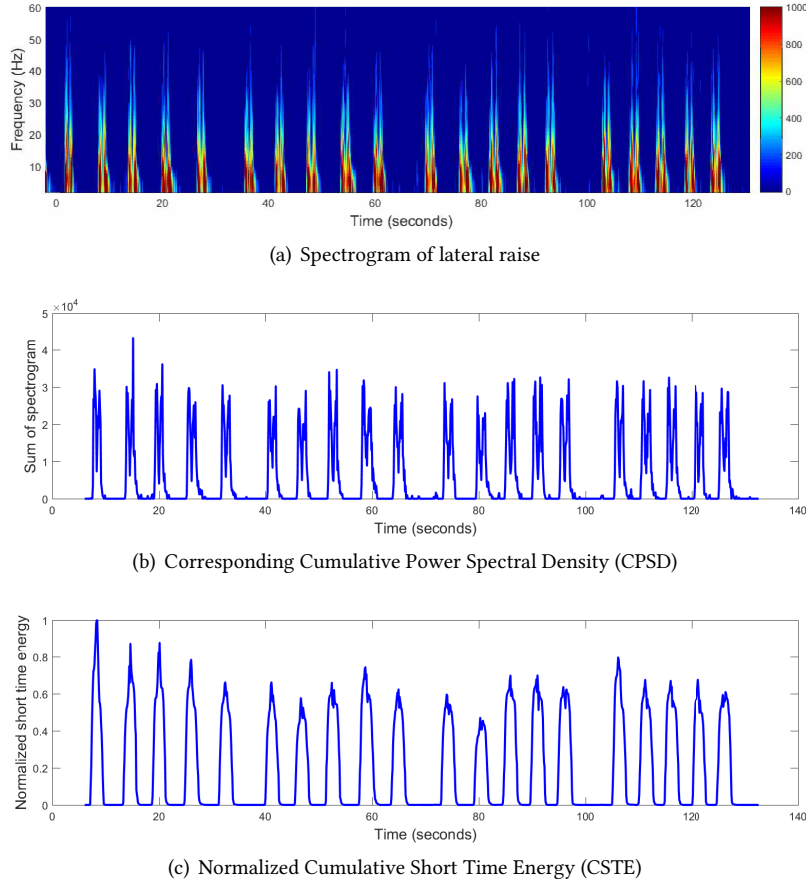(c) Normalized Cumulative Short Time Energy (CSTE)

Fig. 3. The process of exercise repetition recognition.

to further make each repetition more distinguishable. Given the accumulated PSD, the normalized short time energy (STE) [4] is derived as:

$$E_{sqr} = \sum_{i=-\infty}^{\infty} [V(i)W(n-i)]^2, \tag{1}$$

where $V(i)$ is cumulative PSD. $W(n)$ represents the windowing function and n represents the frame shift of samples. Our system identifies all the peaks (i.e., local maximum points) on the normalized STE as shown in Figure 3(c), and the number of peaks represents the number of repetitions. Further, we segment each repetition by searching both left and right-hand side of each peak to find zero energy points, and the CSI data between two zero points corresponds to a repetition. We also cluster multiple repetitions into one exercise set if these peaks are far away from others as a user normally takes a longer rest between two consecutive sets. Note that, our system leverages CSI, which provides fine-grained information of wireless channel, and thus it is sensitive to body movement. Therefore, both upper and lower body movement show significant changes in our spectrogram analysis.
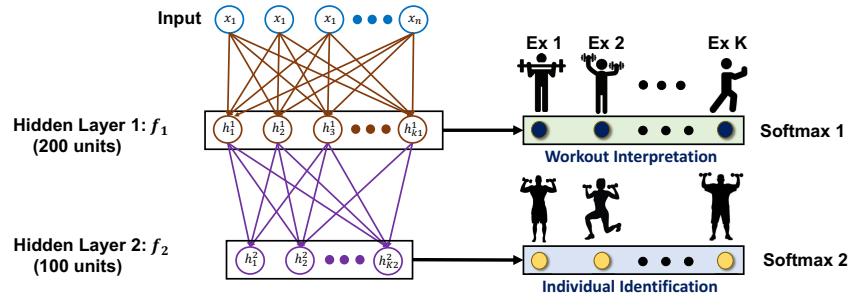
Fig. 4. Illustration of the proposed DNN architecture.

5 PERSONALIZED WORKOUT INTERPRETATION VIA DEEP LEARNING

In this section, we provide a Deep Neural Network (DNN) based method to recognize exercise type and associate the detected exercises with the corresponding user in a shared space. The high level hidden representations produced by DNN are more stable and robust to small-scale variations of input signals as indicated in previous work [28], which meets the hierarchical requirements of our system. Therefore, the proposed DNN structure involves two hidden layers, each including an autoencoder and a softmax classifier as show in Figure 4, for workout interpretation and individual identification, respectively. Since each set of workout usually involves multiple repetitions, our system performs repetition-level workout recognition more frequently than set-level individual identification. Furthermore, to reduce computational cost, instead of using the second layer hidden representatives, we leverage the feature abstractions from the first hidden layer to recognize the workout type of each repetition. Specifically, the input vector of DNN consists of features extracted from 30 subcarriers of raw CSI data. At each hidden layer, the autoencoder learns hidden representations from the input vector, and the softmax classifier, which is connected at the end of autuencoder, builds the activity/individual classification model together with the autoencoder with labeled instance. The autoencoders are trained in an unsupervised fashion while the softmax layer is trained with the labeled representations extracted from the autoencoders.

In contrast to traditional classifiers (e.g., support vector machine (SVM)), our two-layer DNN based approach could robustly recognize workout types and human identities despite the impacts of wireless measurement variations. Typically, CSI readings suffer from complex signal shifting introduced by minute propagation path variations, which can hardly be suppressed with a single feature space transformation (i.e., SVM). Therefore, traditional classifiers, such as SVM and logistic regression requiring the testing data to be located near the hyperplanes that are learnt using the profile samples, fail to robustly perform individual identification and workout recognition. We thus develop a DNN model with multiple hidden layers to capture compositional/robust feature abstractions [27] and to mitigate the multi-path effect. We further compare our DNN-based method with four other traditional classifiers, including k-nearest neighbors (k-NN), decision tree, random forest and SVM, and discuss the results in Section 7.4.4.

5.1 Feature Extraction

To capture the unique characteristics of each type of exercises, it is necessary to extract reliable and efficient features from CSI measurements as the input vector of DNN model. Since the raw CSI measurements are sensitive to environmental changes, the impact of such ambient interference will be eliminated through obtaining reliable features. In addition, using the features extracted from the raw CSI measurements as input of DNN model also reduces the computational overhead of our system. There are 8 time domain features extracted from each subcarrier, including *maximum, minimum, mean, kurtosis, skewness, variance, median* and *standard deviation*.

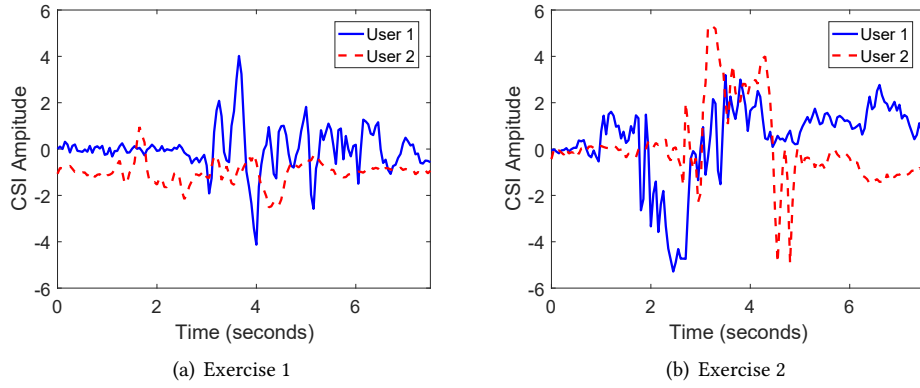(a) Exercise 1   (b) Exercise 2

Fig. 5. CSI amplitudes of two users performing two types of exercises.

Such CSI features are extracted from 30 subcarriers within a single exercise segment, and thus the length of the input vector X is 8×30. In addition, the proposed system adopts a high sampling rate of 1000Hz to ensure the sufficient granularity on capturing human exercise dynamics.

5.2 Autoencoder-based Latent Representation Extraction

Given the extracted CSI features, our DNN model first recognizes the exercise type and then identifies individual through analyzing the hierarchical latent representation derived from the input vector. Figure 4 illustrates our DNN structure consisting of a two-layer stacked autoencoder [28] with two corresponding softmax classifiers. To determine the number of neurons in each hidden layer, our system iterates through various number of neural units and selects the number that provides the highest cross validation accuracy. The proposed DNN structure takes CSI features set X as the input vector and encodes features into hierarchical latent representations (i.e., y_i ($i = 1, 2$)) through activation functions f_i ($i=1, 2$), where the function outputs can be utilized for workout/identity classification.

The neural units in each DNN layer learn a set of latent representations from the CSI feature inputs in an unsupervised fashion. Specifically, the two hidden layers have 200 and 100 neural units, respectively, and force the network to learn a compressed and sparse representation of the inputs through setting the number of hidden units fewer than the number of inputs n.

The hidden units in each autoencoder map the input vector X into a set of latent representations Y through the following equation:

$$Y = \sigma(wX + b), \tag{2}$$

where $\sigma()$ is a logistic sigmoid function formulated as $\sigma(z) = \frac{1}{1+e^{-z}}$, and w and b characterize the weight and bias of the autoencoder, respectively. The autoencoders are trained by minimizing the reconstruction errors between X and Y. Particularly, to train an autoencoder, we initialize weights and bias with random numbers. Then, the optimization function Γ in the autoencoder training is defined as follows:

$$\Gamma(X, X') = \frac{1}{K} \sum_{k=1}^{K} (X_k - X_k')^2 + \lambda \times \Omega_{weights}$$
$$+ \beta \times \Omega_{sparsity}, \tag{3}$$

where K is the number of training samples and X' represents the reconstructed sample from Y using a decoder function. We use $\Omega_{weights}$ and $\Omega_{sparsity}$ to denote the parameters of L_2 regulariser and sparse regulariser [20], which prevent low output values of the neural units. And λ and β are the coefficients of L_2 regulariser and sparse regulariser. Specifically, we denote $\lambda = 0.002$ and $\beta = 4$ for controlling mean squared error for the optimization function Γ. In order to limit the training time while preserving low reconstruction errors for the units in hidden layers, we train the proposed DNN model leveraging a learning rate of 10% and 300 epochs in an unsupervised manner.

5.3 Individual Identification and Workout Interpretation

Individual identification on top of workout interpretation is challenging because gestures of different people from one exercise type are similar. However, we find that different people perform non-identical micro-gestures that can be captured and discriminated from CSI. Figure 5 shows CSI measurement of two users for two different exercises. We observe minute differences in CSI measurement from two users performing the same exercise, because the distinct characteristics (e.g., body shape and distinct behaviors) of each user have unique impacts on CSI measurements. Thus, we can rely on CSI feature representations to identify each individual. Specifically, we construct the two-layer DNN model by stacking one autoencoder layer on the top of another. Each autoencoder layer outputs latent representations which are abstractions of the input CSI features. Given the feature representations from the two layers, we use two SoftMax [5] functions to perform individual identification and workout Interpretation. The first SoftMax function takes the layer 1 hidden representations as input and recognizes exercise type, while the second SoftMax layer is attached to the second autoencoder and identifies the individual identity. The SoftMax functions are defined as follows:

$$P(C_n|Y) = \frac{P(Y|C_n)P(C_n)}{\sum_{j=1}^{n} P(Y|C_j)P(C_j)}, \tag{4}$$

where $P(C_n|Y)$ represents the posterior probability of user identity label C_n given a hidden latent representations Y, and we denote the prior of the same label as $P(C_n)$. $P(Y|C_n)$ indicates the likelihood of the representations Y given class C_n. The values of $P(C_n|Y)$ are constrained by $\sum_{n=1}^{N} P(C_n|Y) = 1$. Each SoftMax function outputs the probability distribution over N profiled users/exercises. Each SoftMax layer is trained through minimizing a mean squared error function, which is defined as:

$$E = \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{k} (t_{ij} - y_{ij})^2, \tag{5}$$

where n is the number of observations, and k is the number of users. t_{ji} is the element of a ground truth matrix for training samples and y_{ij} is the hidden representations for the jth observation. Then, we use an optimized function $n = argmax_{n \in N} P(C_n|Y)$ to find a predicted class for the latent representations Y. The training process of Softmax layer runs 400 epochs with the learning rate fixed at 0.15 in a supervised manner to control the training time while preserving classification accuracy.

5.4 Profile Construction

The design of our profile construction should be in line with the objective of this project that we aim to deploy our system at home or in the office where multiple users may perform exercises. We need to take the following two aspects into consideration. First, our system should have the capability of differentiating different individuals on top of workout recognition in a shared environment. Second, the system needs to accommodate the existence of possible concurrent activities, such as workout or non-workout activities (e.g., walking, typing) performed by other people. To accommodate environment sharing and achieve personalized workout interpretation, our system
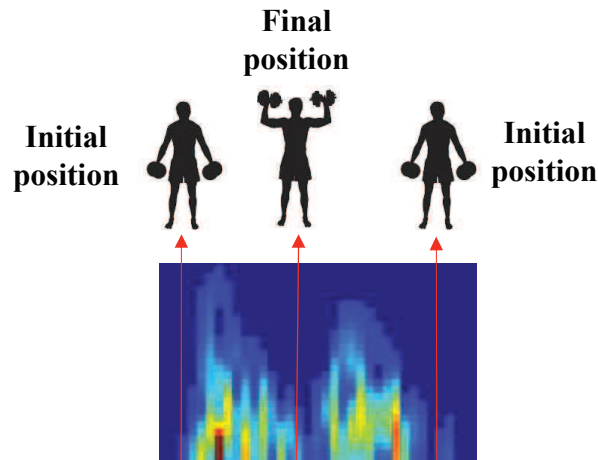
Fig. 6. Spectrogram of one repetition.

collects CSI readings of each workout type from each individual. Therefore, our system adopts a personalized profile construction model. It is possible for a user to perform workout at a slightly shifted location. To adapt to such possible location change, the personalized training profiles are collected at several locations within a certain range (e.g., one foot away from the original location), which makes the trained model in our system robust to location changes. In the case of concurrent activities, non-workout activities may also contain repetitive body movements of other people nearby. However, due to the limited space for home and office environment, it usually only allows one person to perform exercise within a certain range of space during a time period. We model such space using the concept of Fresnel zone [15], which is defined as a series of concentric prolate ellipsoidal regions between a transmitting antenna and a receiving antenna. Given a pair of transceivers close to an exercise area in a room, the Fresnel zone is constructed around that deployment area. We define the area within the $12th$ Fresnel zone as the Sensitive Region because the RF signal transmission is mainly concentrated within the first $8 - 12$ zones [15]. For example, in a typical room size (e.g., $8.3m \times 6.6m$), the distance between a transmitter and a receiver can be $6m$. Then, the derived radius of the Fresnal zone is less than $1.7m$. Therefore, other concurrent activities outside the Fresnel zone do not affect our system performance. Detailed analysis of both the Fresnel zone and experiment results of concurrent activities can be found in Section 8.

6 SMART WORKOUT ASSESSMENT

Correct exercise postures will not only reduce the possibility of injury, but also increase the efficiency of muscle building. In addition, any incorrect exercise form may incur some physical health-related problems, especially for the elderly people. In this section, we will discuss smart workout assessment based on the analysis at repetition level to help users retain correct exercise postures and thereby achieve their fitness goals.

6.1 The Anatomy of a Rep

To perform workout assessment, we first need to have an in-depth understanding of a workout repetition. Basically, a repetition normally consists of a series of body movements from an *initial position* to a *final position* and then back to the *initial position* as shown in Figure 6. The figure shows the spectrogram of a repetition with corresponding movements. In our example, the movement contains *concentric contractions* (i.e., cause muscles to shorten, thereby generating force) and *eccentric contractions* (i.e., cause muscles to elongate in response to a

greater opposing force). Good exercise repetitions need to keep a constant rhythm (i.e., the time ratio between concentric contractions and eccentric contractions in our example).

6.2 Workout Review and Recommendation

In order to achieve effective exercise and avoid injury, it is crucial to customize personal fitness plans by fitness professionals. Such fitness plans often try to regulate the exercise by following the Frequency, Intensity, Time, and Type (FITT) [19] principle of training, which is a set of guidelines that instruct users to set up exercise routines fitting their goals and fitness levels, while maximizing the effects of exercises. Specifically, *Frequency* refers to the frequency of exercise undertaken; *Intensity* refers to the intensity of exercise undertaken; *Time* refers to the time a user spends exercising; *Type* refers to the type of exercise undertaken.

Traditionally, a fitness coach or trainer at gym can keep watching each exercise repetition of a user, and provide advice based on the FITT principle. However, many people who work full time do not have much time to exercise at dedicated place (i.e., the gym). Instead, they often choose to exercise at home or at the office. The proposed *smart workout assessment* module fills the gap between users' needs and fitness plans to meet the FITT principle by providing fine-grained, personalized fitness information and intuitive feedback to users. The module will automatically provide assessments and recommendations, which relies on the measurement of WiFi CSI readings while users are performing exercises. Our system mainly focuses on perceiving *Intensity* and *Time*, because these two guidances are closely related to exercises in repetition-level. Based on the FITT principle and the anatomy of a rep, we define two metrics as follows:

Work-to-rest Ratio. Workout time in repetition-level reflects how much time is spent on a repetition of a set. It measures the complete cycle of a repetition within one set. To gain efficient muscle building, a user should maintain a consistent ratio between the time of repetition and the following time of rest. Thus, we define the work-to-rest ration as:

$$R_{w2r} = \frac{T_w^{(i)}}{T_r^{(i)}},$$

where $T_w^{(i)}$ is the time duration for the ith workout and $T_r^{(i)}$ is the time duration of the rest followed by the ith workout.

Repetition Tempo Ratio (RTR). Workout intensity usually indicates how much energy is consumed, which can be reflected by a relative percentage of a user's maximum endurance. Specifically, we define the intensity as a ratio between the energy from an initial position to a final position. According to the analysis of the anatomy of a repetition, we found that this ratio can be measured through measuring the time duration from an initial position to a final position, and the time duration of from the final position back to the initial position in a repetition. Thus, we define the repetition tempo ratio as follows:

$$R_{RTR} = \frac{T_{i2f}^i}{T_{f2i}^i},$$

where T_{i2f}^i is the time duration from an initial position to a final position of the ith repetition and T_{f2i}^i is the time duration from the final position back to the initial position of the ith repetition.

Workout Recommendation. Based on the assessment results, our system next will provide reasonable workout recommendations to help the user to correct exercise gestures. The basic idea is to compare the exercise assessment results of normal users with the ones obtained from expert users in term of the new metrics.

Specifically, our system not only provides the trend of these two metrics across all repetitions, but also shows the upper and lower bounds of the metrics obtained from expert users. Then a user can have a visualized view of his workout assessment and check whether repetition is correctly performed or not. Figure 7 plots the workout
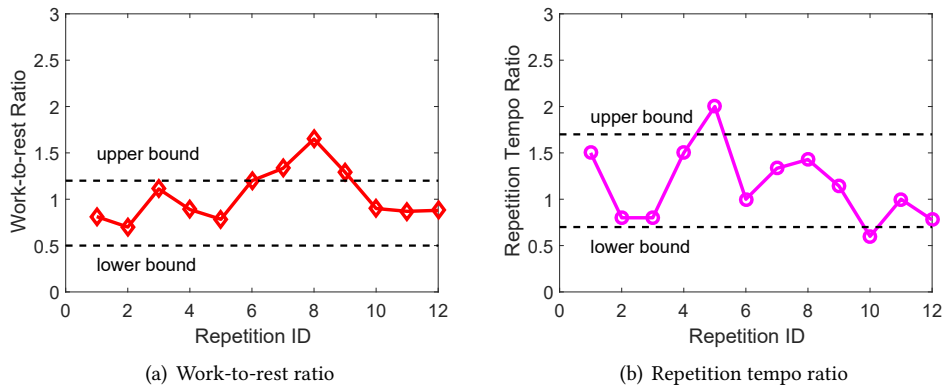
(a) Work-to-rest ratio

(b) Repetition tempo ratio

Fig. 7. Workout assessment for each repetition based on two metrics.

Table 1. Workout activities performed.

Code	Exercise	Code	Exercise
Ex1	Standing biceps curl	Ex6	Tai Chi
Ex2	Lateral raise	Ex7	Dumbbell curl
Ex3	Leg stretch	Ex8	Pile squat
Ex4	Raise and squat	Ex9	Dumbbell triceps extension
Ex5	Leg press	Ex10	Body extension

assessment for a user after he finishes 12 repetitions of lateral raise. The user can observe how each of his repetitions is performed based on work-to-rest ratio and can find the $7^{th}, 8^{th}, 9^{th}$ repetitions are over the upper bound as shown in Figure 7(a). Figure 7(b) provides a different view of each repetition by examining the repetition tempo ratio. The user will find the 5th repetition is over the upper bound and the 10th repetition is below the lower bound. Based on the visualized feedback, the user can adjust his subsequent repetitions by paying attention to those repetitions that are poorly done. In particular, for those repetitions beyond the upper bound of a standard workout, the user needs to extend the rest time after the repetitions, while for those over the upper bound of repetition tempo ratio, the user can increase the intensity from the initial position to the final position.

7 PERFORMANCE EVALUATION

7.1 Experimental Methodology

Experiment Setup. To validate the performance of the proposed system, we implement a prototype by deploying one pair of Dell E6430 laptops equipped with Intel 5300 802.11n WiFi NICs acting as WiFi transceivers. Specifically, the transmitter and receiver are configured to work under *Monitor* mode with fixed packet rate 1000 pkts/sec at 5GHz band. Note that, each CSI measurement of the captured packet consists of 30 subcarriers in the form of complex value.

Data Collection. We conduct the experiments by recruiting 20 volunteers who are college students (18 males and 2 females; age from 22 to 35). The exercise training profiles are collected at several locations within a certain range (i.e., within one foot) of a profile location. Their workout activities are tracked and analyzed at three different indoor venues over a 10-month time period. Note that the venue sizes are $6m \times 4m$, $8.3m \times 6.6m$ and
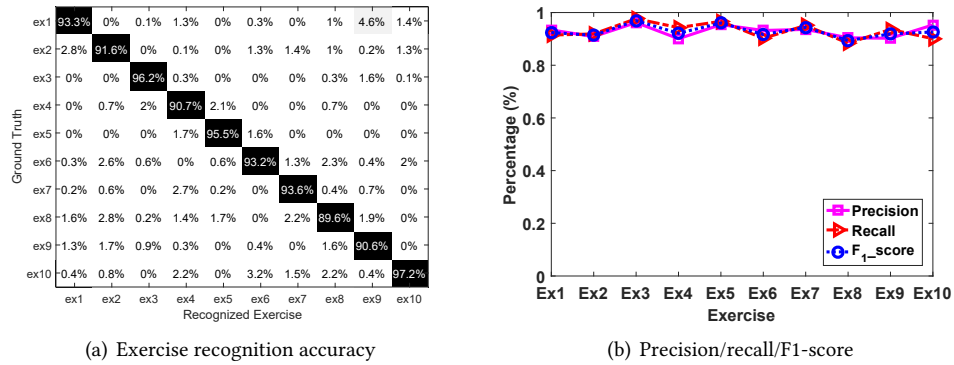
(a) Exercise recognition accuracy

(b) Precision/recall/F1-score

Fig. 8. Accuracy of exercise recognition with the proposed DNN model.

(a) Exercise recognition accuracy
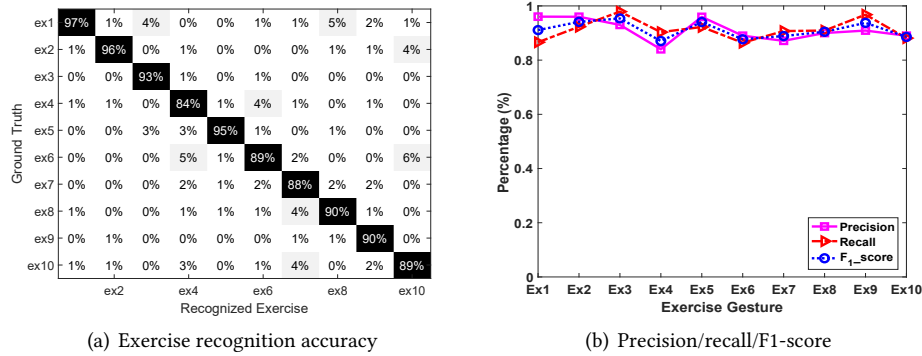

(b) Precision/recall/F1-score

Fig. 9. Accuracy of exercise recognition with a separate two-layer DNN model.

$7.3m \times 3.6m$, respectively, which correspond to three typical room sizes (i.e., living room, large family room and office space). In order to validate the reliability of the proposed system, a total of 10 typical indoor exercises are performed by the volunteers, including free-weight, body stretching, and aerobic exercises. The details of the exercise types are listed in Table 1. During the 10-month survey, we ask each of the volunteers to conduct 20 sets for each type of exercise per week with 15 repetitions per set. In total, we collect 3013 segments of CSI measurements that are associated with workout activities from the volunteers. For each round of data collection, we ask volunteers to stay in the room for 20 minutes. During this period of time, a volunteer mimics the scenarios of work-at-home or in the office. Specifically, each volunteer is required to perform workout-related activities for 10 minutes and any other activities (e.g., with or without repetitive pattern) for the rest of the time. Therefore, the collected CSI readings could include both regular daily activities (e.g., sitting on a chair at work, typing and walking around) and exercise activities from the volunteer. The time duration ratio of the regular daily activities and exercise activities is around 1 : 1. The design of such a data collection process is to demonstrate that our fitness assistant system has the capability of distinguishing workout-related activities from daily activities.

Evaluation Metrics. To quantify the performance of our system, we define the following four metrics as:

Precision. Given N_w sets of a workout type w in our profile, the precision of recognizing the performed workout N_w is defined as:

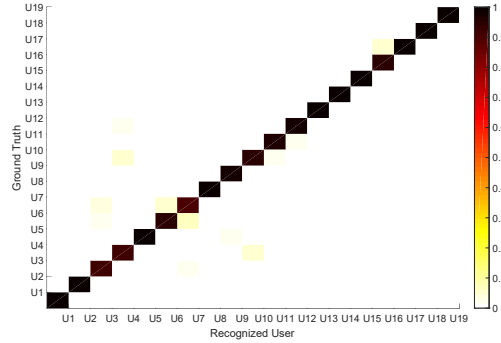$$precision_w = \frac{N_w^T}{N_w^T + M_w^F},$$

Fig. 10. Individual identification accuracy.

where N_w^T is the number of instances correctly recognized as workout w, while M_w^F is the number of sets corresponding to other workouts that mistakenly recognized as workout w.

Recall. Recall of a workout w is defined as the ratio of the repetitions that are correctly recognized as the workout w over all repetitions recognized as workout type w:

$$Recall_w = \frac{N_w^T}{N_w}.$$

F1-score. F1-score is the harmonic mean of precision and recall, and acts as a measure of a classification's accuracy. F1-score ranges from 0 to 1, indicating ascending classification accuracy. In our multi-class scenario, the F1-score for a specific workout w was defined as:

$$F_1^{(w)} = 2 \times \frac{precision_w \times recall_w}{precision_w + recall_w}.$$

Confusion Matrix. The columns of the confusion matrix denote the ground truth of a workout/identity, while the rows represent the classified workout/identity in our system. Each entry in the matrix represents the percentage of correctly classified exercise/identity.

7.2 Personalized Workout Interpretation and Individual Identification

Workout Interpretation. We first examine the performance of workout interpretation based on our DNN model. Figure 8(a) shows the confusion matrix of 10 different workout exercises performed by all volunteers at three different environments. We observe that the average recognition accuracy is around 93% and the standard deviation is 2.6%. Particularly, we find the 8th workout (i.e., pile squat) has a slightly lower accuracy than other types of workout. This is because pile squat involves whole body movement which makes the CSI stream vary a lot from time to time. But our system can still maintain over 89% accuracy. In addition, as shown in Figure 8(b), we find that the corresponding precision, recall and F_1 score are all around 93%. The results show that our DNN-based model can achieve high accuracy in workout recognition and thereby confirm the robustness of our system under different environments. Moreover, we investigate workout interpretation performance of a two-layer neural network, which has the same architecture as the two-layer model described in section 5 but leverages the hidden representatives extracted from the second-layer for workout recognition. As shown in Figure 9, we find that the average recognition accuracy is around 91.1% and the standard deviation is 4.6%. Such recognition accuracy is lower than that of our DNN model, because a two-layer model is over-complicated for our workout recognition task. It demonstrates that our system can recognize exercise types with sufficient high

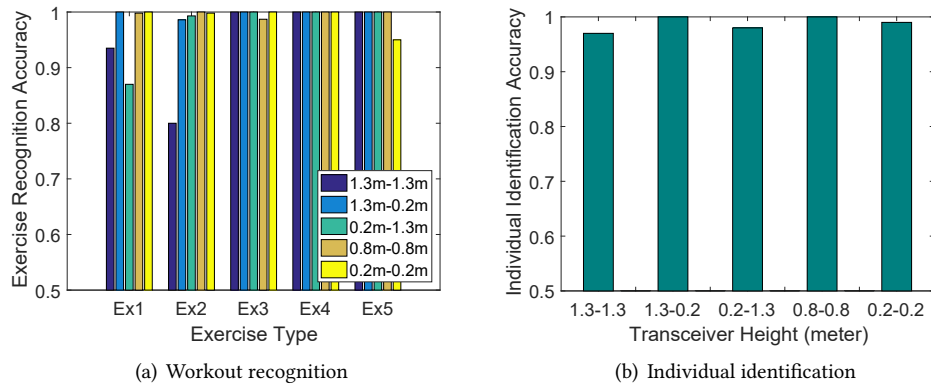(a) Workout recognition   (b) Individual identification

Fig. 11. Impact of transceiver height on exercise recognition and individual identification.

accuracy and low computational complexity leveraging hidden representatives derived from the first layer of the proposed DNN model.

Individual Identification. Next, we perform individual identification by using our DNN model. In Figure 10, the confusion matrix presents the overall accuracy of individual identification. We find that our system always achieves high identification accuracy across all the volunteers. Particularly, the accuracies for all the individuals are over 92% and the standard deviation is 3.34%. This result validates that our system can identify individuals with high accuracy and therefore, supports personalized fitness assistance. Although the proposed DNN model is sufficiently distinctive in a relative large user population (i.e., 20 people in our current experiments), it is not yet clear what the maximum number of users that can be identified by our system. However, our results of 20 people are promising, and this user population size is sufficient for most of the use cases in home or office environments.

7.3 Workout Assessment

Our system provides workout assessment to each user by comparing a user's workout pattern to the experts' in terms of the two metrics defined in section 6.2. Then, recommendations are given to users based on the upper bound and lower bound of the two metrics calculated from the expert users. During this project, all the volunteers try our workout assessment and adjust their postures based on the workout review provided by our system. In particular, we ask 6 expert users, who visit gym regularly and have training experiences before, to perform three sets of workout with 10 repetitions per set for each workout type. Then, we set the upper/lower bound of the specific metric as the maximum/minimum value from the 6 expert users. Overall, all volunteers give positive feedback on our system and confirm that the proposed system can help the users to correct their exercise form.

7.4 Impact of Different Factors on System Performance

There are factors (i.e., heights of device placement, distance between devices, and location shift between profile and testing data) that could affect our system performance. In this part, we will examine and analyze the impact of these factors to discuss the robustness of our system.

7.4.1 Impact of Different Heights of Device Placements. We study the workout interpretation and people identification scheme by placing the wireless transmitter and receiver at different heights. Specifically, we examine three cases where the transceivers are placed at equal heights, $0.2m$, $0.8m$ and $1.3m$, corresponding to the floor, the table, and the top of a cabinet. Furthermore, we also study some real-world scenarios, where the heights of transmitter and receiver are unequal (i.e., $1.3m - 0.2m$, $0.2m - 1.3m$).
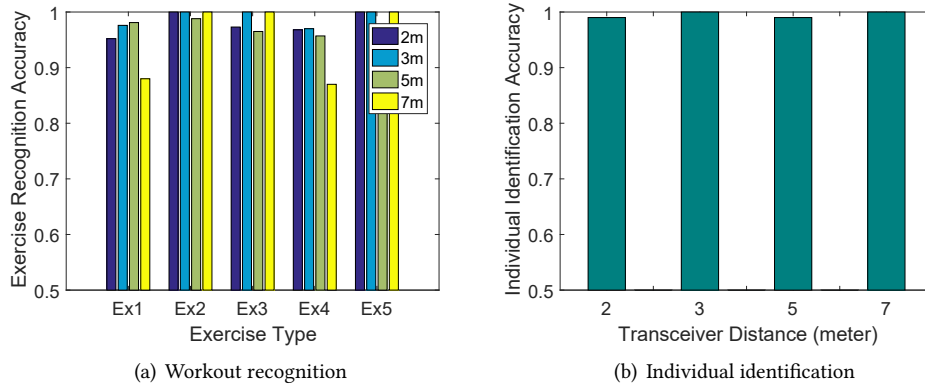
(a) Workout recognition   (b) Individual identification

Fig. 12. Impact of transceiver distance on exercise recognition and individual identification.

(a) Without opportunistic profiles   (b) With opportunistic profiles
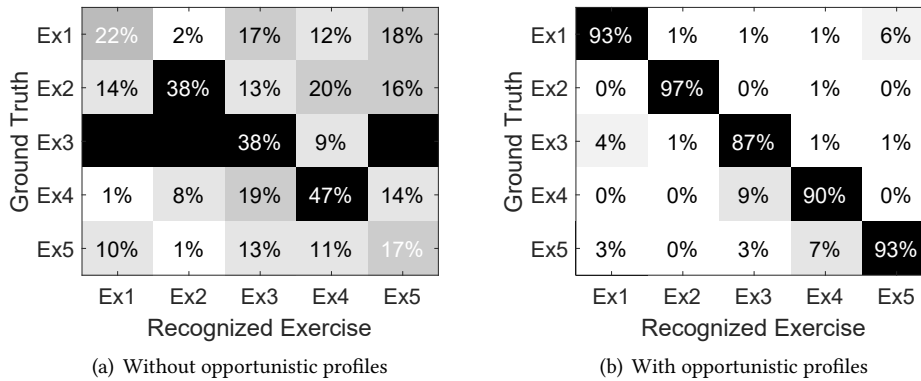
Fig. 13. Improvement of workout recognition accuracy with opportunistic profile construction.

Figure 11(a) shows the workout recognition accuracy for five different exercises. We observe that our DNN workout recognition algorithm achieves comparable high performances with various transceiver placements. In general, the average recognition accuracies for the five pairs of transceiver heights are 94.7%, 99.7%, 97.3%, 99.6%, 98.9%. Specifically, we find that the recognition performances for three transceiver heights (i.e., $1.3m - 0.2m$, $0.8m - 0.8m$, and $0.2m - 0.2m$) achieve over 94% for all five exercises. In addition, we observe that when the transceivers are placed at the height $1.3m - 1.3m$, exercise 1 and 2 have relatively lower recognition accuracy compared with most of the other transceiver heights. This is because both exercise 1 and 2 involve a gesture of raising the dumbbells to the height of the transmitter (i.e., $1.3m$), resulting in similar multipath and line-of-sight effects. Thus, the CSI readings of these two exercises are very similar and thus the recognition performances are slightly degraded. Beside workout recognition, we also present the individual identification accuracy in Figure 11 (b). It is encouraging that the identification accuracies for the five cases with equal transceiver heights are all over 97%. The results confirm the proposed DNN model is robust to different transceiver heights on workout recognition and people identification.
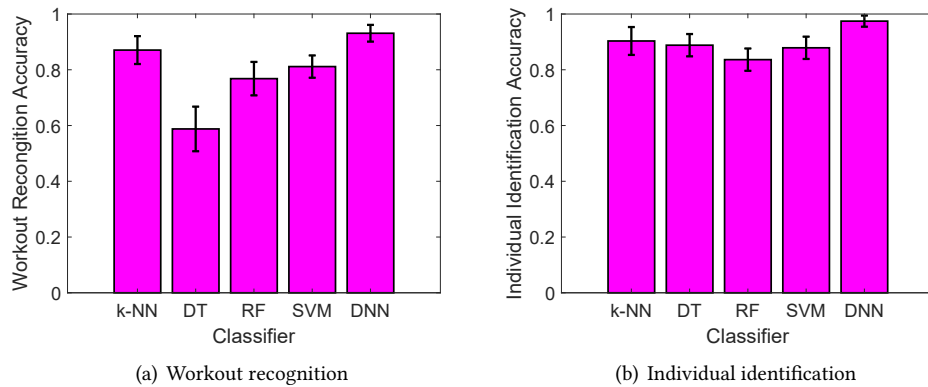
(a) Workout recognition

(b) Individual identification

Fig. 14. Comparison of different classifier

7.4.2 Impact of Different Device Distances. Furthermore, we also study the workout recognition performance under various transceiver distances, which correspond to different room sizes. We show both the exercise recognition and people identification accuracy of five different exercises in Figure 12. As shown in Figure 12(a), our algorithms can achieve good performance for most of the exercises with different transceiver distances. Specifically, the average recognition accuracy for the four distances are: 97.8%, 98.9%, 96.6%, 95.0%. In particular, for the case of the distance 2*m*, 3*m*, 5*m*, the recognition accuracies for the five exercises are all over 94%. For the transceiver distance of 7*m*, some exercises (i.e., Ex1 and Ex4) have relative lower recognition accuracy due to the weak signal strength at such a distance. Figure 12(b) shows the people identification accuracy for the four cases of different transceiver distances. We can find that the identification accuracies are all over 99% which validate the reliability of our people identification algorithm at different distances between transmitter and receiver.

7.4.3 Impact of Location Shift between Profile and Testing Data. Finally, we study the impact of location shift if workout profile and testing data are collected at slightly different locations (e.g., one foot away). As indicated in section 5.4, an opportunistic profile construction technique is adopted to build the user's workout profile at 5 different locations within a certain area, where the user deliberately changes his location when performing the same exercise. As a comparison, we also build another workout profile without opportunistic profile construction (i.e., a user's profile collected from the exact same location only). We present the experimental results for both with and without the opportunistic profile construction as shown in Figure 13. Figure 13(a) reports the workout recognition results in a confusion matrix, and we observe low average recognition accuracy around 32% without opportunistic profile construction. By contrast, with the opportunistic profiles, the workout recognition results significantly improved from 32% to 92% as shown in the confusion matrix in Figure 13(b). This comparison further confirms the robustness of our system.

7.4.4 Comparison of Various Classifiers. We further compare the DNN model with other four traditional classifiers, including k nearest neighbors (k-NN), decision tree (DT), random forest (RF) and support vector machine (SVM). Each classifier is evaluated using 10-fold cross-validation, and the parameters of each classifier are tuned to achieve the best performance. We can observe in Figure 14 that the DNN-based classifier achieves the best workout recognition and individual identification accuracies among the five classifiers. As shown in Figure 14(a), for workout recognition, the other four classifiers have average accuracies of 87%, 58%, 79%, 81%, respectively. And for individual identification accuracies as depicted in Figure 14(b), the four classification models have average accuracies of 90%, 88%, 83%, 87%, which are less effective than the proposed DNN model. Moreover, we observe that
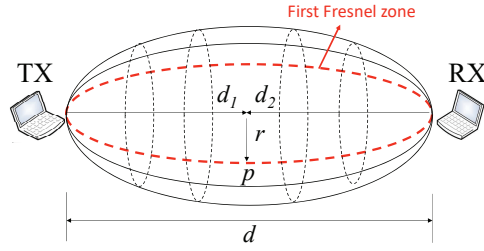
Fig. 15. Sensitive region regulated by Fresnel zone. (We show the first fresnel zone in red)

(a) No people around

(b) 3 meters from the LoS
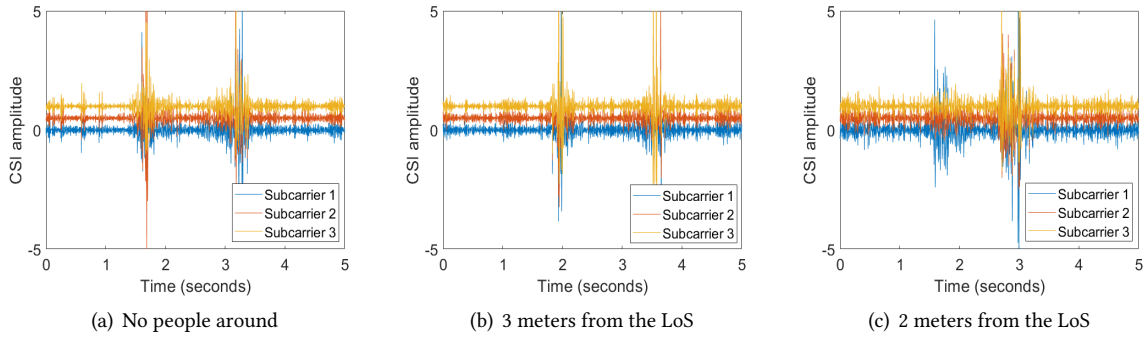
(c) 2 meters from the LoS

Fig. 16. CSI readings of lateral raise under impacts of surrounding people walking.

the proposed DNN model obtains the lowest variances of 3.1% and 1.9% for workout recognition and individual identification respectively. Thus, our DNN model maintain consistent performance in the cross-validation process. This is because the linear-classifiers fail to mitigate the impacts of non-linear value variations introduced by propagation path changes. Therefore, we adopt the DNN model for both workout recognition and individual identification tasks in this work.

8 DISCUSSION

Sensitive Region Between a Pair of WiFi Transceiver. The strength of waves propagating between a pair of transceiver is regulated by the Fresnel zone, which is a series of concentric prolate ellipsoidal regions between a transmitting antenna and a receiving antenna as shown in Figure 15. We define the area within the $12th$ Fresnel zone as the *Sensitive Region* because the RF signal transmission between a pair of transceiver is mainly concentrated within the first $8 - 12$ zones [15]. Specifically, the radius of Fresnel zone at point p can be calculated by:

$$r_n = \sqrt{\frac{n\lambda d_1 d_2}{d_1 + d_2}}, \tag{6}$$

where r_n is the radius of the nth Fresnel zone at point p; λ is the wavelength; d_1 and d_2 are the distance from the transmitter and receiver to p, respectively. For a typical room size as discussed in section 7.1, we assume a user stands in the middle of LoS path (i.e., $d_1 = d_2$). Then, the distance between the user and either deivce ranges 2m to 4m for our system deployment in a typical living or office room. We can derive r_{12} (i.e., the radius of the 12th fresnel zone) to be between 1.2m to 1.7m according to equation 6. These results suggest that the radius of the sensitive region for a typical room size is smaller than 1.7m.

Impact of Ambient Interference. In real world scenarios (e.g., office and home), which are usually shared by multiple people, the CSI dynamics with respect to the workout pattern of one specific user may also be distorted by daily activities of sounding people. However, according to our theoretical analysis on Fresnel zone, the interferences from surrounding people should be negligible if they are moving outside the sensitive region. To validate our theoretical analysis, we conduct a set of experiments to inspect the impact of the interferences from surrounding people. Particularly, we ask one participant to conduct exercise on the WiFi LoS path and ask another participant to keep walking 3 meters (i.e., out of the sensitive region) away from the LoS path. This scenario is common since people usually will not walk too close to the exerciser due to the safety reason in a shared home/office environment. In addition, we choose walking activity because human walking involves movements of different body parts (e.g., leg, arm, trunk) and thus has strong impacts on CSI readings than stationary activities (e.g., typing, sitting down). As shown in Figure 16 (a) (b), the CSI patterns of lateral raise has little difference between the cases where no people walking and several people walking 3 meters away. Although we observe noisy measurements in Figure 16 (c), where the other participant walking closely (i.e., 2m) to the LoS path, the CSI measurements still be able to capture the exercise dynamics of the target user. For both cases, our system can successfully classify the lateral raise with 92% accuracy which is only 3% lower than the case without interference from other people. As demonstrated above, the CSI readings are dominated by WiFi signals within the sensitive region. If the LoS path is blocked or the WiFi enable equipments are moved, our system needs to start a new round of training data collection, which requires the cooperation from users, and then updates the profiles accordingly.

Limitation and Future Work. Our system takes the first step to provide personalized quantitative and qualitative evaluation of users' fitness using WiFi. However, there are several limitations that require further future work to fine tune our system. In our system design, we assume our system normally runs in a relatively stable indoor environment where the furniture, WiFi transmitter and receiver are fixed. However, the layout can be different from when users first use our system to create profiles due to different furniture or device location. To address this limitation in our future work, we can accommodate possible environment changes (e.g., location changes of surrounding desks, chairs) via automatic calibration of personalized exercise profile based on autoencoder [8]. In addition, it is also possible that the face orientation of a user is different between the profile and later exercise. For example, a user may perform workout with a different face orientation that is 90 degree away from the orientation in profile. Such difference in face orientation could also reduce the recognition accuracy. To adapt to such possible changes, one possible solution is to collect training profiles with representative user orientations (e.g., 0, 90, 180, and 270 degree). Thus, the profile constructed based on training instances with different orientations makes the system robust. However, this may increase the size of training sets. Another promising technique that can be leveraged to address this issue is transfer learning[21]. We plan to devise an algorithm based on transfer learning to eliminate the extra efforts in the reconstruction of training profile in our future work.

9 CONCLUSION

In this paper, we study the feasibility of using ubiquitous WiFi signals to help users perform effective exercise in home/office environments by providing personalized fine-grained workout statistics (i.e., workout type, the number of sets, and the number of repetitions) and dynamic assessments (i.e., stability of workout intensity and strength). Toward this end, we design and implement a device-free personalized fitness assistant system based on CSI measurements, which is available from existing WiFi infrastructure. Our system is capable of differentiating individuals on top of fine-grained workout interpretation and providing smart workout assessment. More specifically, our system uses a DNN-based model to perform workout interpretation with comprehensive workout analysis, and then extracts fine-grained latent representations to further identify individuals. In addition,

our system assesses workout by analyzing both short and long-term workout quality, and provides workout reviews for users to ameliorate subsequent exercises and avoid unnecessary injury. Extensive experiments involving 20 volunteers are conducted under three typical indoor venues over a 10-month time period. The experimental results demonstrate that the proposed system achieves over 93% and 97% accuracy to identify the workout type and corresponding individuals with only a single pair of WiFi transmitter and receiver, respectively. We envision that the proposed system will not only provide more convenience for personalized fitness assistance, but also can be extended to various application scenarios and benefit people's daily lives.

ACKNOWLEDGMENTS

REFERENCES

[1] 2014. Fitbit. http://www.fitbit.com/.
[2] Fadel Adib, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2014. 3D Tracking via Body Radio Reflections. In *11th USENIX Symposium on Networked Systems Design and Implementation (USENIX NSDI)*, Vol. 14. 317–329.
[3] Juliane Arney. 2005. You should be in pictures! Experts share tips on how to produce a quality fitness video for any purpose. *IDEA Fitness Journal* 2, 1 (2005), 86–90.
[4] RG Bachu, S Kopparthi, B Adapa, and Buket D Barkana. 2010. Voiced/unvoiced decision for speech signals based on zero-crossing rate and energy. In *Advanced Techniques in Computing Sciences and Software Engineering*. Springer, 279–282.
[5] Christopher M Bishop. 2006. Pattern recognition. *Machine Learning* 128 (2006).
[6] Oya Çeliktutan, Ceyhun Burak Akgul, Christian Wolf, and Bülent Sankur. 2013. Graph-based analysis of physical exercise actions. In *ACM international workshop on Multimedia indexing and information retrieval for healthcare (ACM MIIRH)*. 23–32.
[7] Keng-hao Chang, Mike Y Chen, and John Canny. 2007. Tracking Free-Weight Exercises. In *Proceedings of the 9th international conference on Ubiquitous computing (ACM UbiComp)*.
[8] Xi Chen, Chen Ma, Michel Allegue, and Xue Liu. 2017. Taming the inconsistency of Wi-Fi fingerprints for device-free passive indoor localization. In *INFOCOM 2017-IEEE Conference on Computer Communications, IEEE*. IEEE, 1–9.
[9] Heng-Tze Cheng, Feng-Tso Sun, Martin Griss, Paul Davis, Jianguo Li, and Di You. 2013. Nuactiv: Recognizing unseen new activities using semantic attribute-based learning. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services (ACM Mobisys)*.
[10] Ronan Collobert and Jason Weston. 2008. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*. ACM, 160–167.
[11] Han Ding, Longfei Shangguan, Zheng Yang, Jinsong Han, Zimu Zhou, Panlong Yang, Wei Xi, and Jizhong Zhao. 2015. Femo: A platform for free-weight exercise monitoring with rfids. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (ACM Sensys)*. 141–154.
[12] Xiaonan Guo, Liu Jian, and Yingying. Chen. 2017. FitCoach: Virtual Fitness Coach Empowered by Wearable Mobile Devices. In *Proceedings of the International Conference on Computer Communications (IEEE INFOCOM)*.
[13] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2010. Predictable 802.11 packet delivery from wireless channel measurements. In *ACM SIGCOMM Computer Communication Review*, Vol. 40. ACM, 159–170.
[14] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool release: Gathering 802.11 n traces with channel state information. *ACM SIGCOMM Computer Communication Review* 41, 1 (2011), 53–53.
[15] Hristo D Hristov. 2000. *Fresnal Zones in Wireless Links, Zone Plate Lenses and Antennas*. Artech House, Inc.
[16] Honglak Lee, Roger Grosse, Rajesh Ranganath, and Andrew Y Ng. 2009. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In *Proceedings of the 26th annual international conference on machine learning*. ACM, 609–616.
[17] Haochao Li, Eddie CL Chan, Xiaonan Guo, Jiang Xiao, Kaishun Wu, and Lionel M Ni. 2015. Wi-counter: smartphone-based people counter using crowdsourced wi-fi signal data. *IEEE Transactions on Human-Machine Systems* 45, 4 (2015), 442–452.
[18] Cihang Liu, Lan Zhang, Zongqian Liu, Kebin Liu, Xiangyang Li, and Yunhao Liu. 2016. Lasagna: towards deep hierarchical understanding and searching over mobile sensing data. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking (ACM MobiCom)*. 334–347.
[19] American College of Sports Medicine et al. 2013. *ACSM's guidelines for exercise testing and prescription*. Lippincott Williams & Wilkins.

[20] Bruno A Olshausen and David J Field. 1997. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision research* 37, 23 (1997), 3311–3325.

[21] Sinno Jialin Pan, Qiang Yang, et al. 2010. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22, 10 (2010), 1345–1359.

[22] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home gesture recognition using wireless signals. In *Proceedings of the 19th annual international conference on Mobile computing and networking (ACM MobiCom)*. 27–38.

[23] Sreemanananth Sadanand and Jason J Corso. 2012. Action bank: A high-level representation of activity in video. In *IEEE Conference on Computer Vision and Pattern Recognition (IEEE CVPR)*. 1234–1241.

[24] Ruhi Sarikaya, Geoffrey E Hinton, and Anoop Deoras. 2014. Application of deep belief networks for natural language understanding. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)* 22, 4 (2014), 778–784.

[25] Richard Socher, Yoshua Bengio, and Chris Manning. 2013. Deep learning for NLP. *Tutorial at Association of Computational Logistics (ACL), 2012, and North American Chapter of the Association of Computational Linguistics (NAACL)* (2013).

[26] Chuan-Jun Su, Chang-Yu Chiang, and Jing-Yan Huang. 2014. Kinect-enabled home-based rehabilitation system using Dynamic Time Warping and fuzzy logic. *Applied Soft Computing* 22 (2014), 652–666.

[27] Yi Sun, Xiaogang Wang, and Xiaoou Tang. 2014. Deep learning face representation from predicting 10,000 classes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1891–1898.

[28] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre-Antoine Manzagol. 2010. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research* 11 (2010), 3371–3408.

[29] Naiyan Wang and Dit-Yan Yeung. 2013. Learning a deep compact image representation for visual tracking. In *Advances in neural information processing systems*. 809–817.

[30] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st annual international conference on mobile computing and networking (ACM MobiCom)*. 65–76.

[31] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. 2014. E-eyes: device-free location-oriented activity identification using fine-grained WiFi signatures. In *Proceedings of the 20th annual international conference on Mobile computing and networking (ACM MobiCom)*. 617–628.

[32] Jianxin Wu, Adebola Osuntogun, Tanzeem Choudhury, Matthai Philipose, and James M Rehg. 2007. A scalable approach to activity recognition based on object use. In *11th International Conference on Computer Vision (IEEE ICCV)*. 1–8.

[33] Junyuan Xie, Linli Xu, and Enhong Chen. 2012. Image denoising and inpainting with deep neural networks. In *Advances in Neural Information Processing Systems*. 341–349.