FitFone: Tracking Home Workout in Pandemic Times

Denys J.C. Matthies, Thorleif Harder, Franz Bretterbauer, Viktoria Ginter, Horst Hellbrück
Department of Electrical Engineering and Computer Science
Technical University of Applied Sciences Lübeck, Mönkhofer Weg 239, 23562 Lübeck, Germany
{denys.matthies, horst.hellbrueck}@th-luebeck.de {thorleif.harder, franz.bretterbauer, viktoria.ginter}@stud.th-luebeck.de

ABSTRACT

Global pandemics, such as COVID-19, significantly impact mental and physical well-being. Regular home workouts become more critical during these extraordinary times, as physical activity can positively impact our health. Having the ability to track workout progress motivates consistent home workout schedules. In this paper, we demonstrate two proof-of-concept implementations enabling the tracking of home workouts. First, we demonstrate exercise tracking when the user wears a smartphone relying on IMU data. Second, we show tracking when a doffed phone is placed in front of the user while relying on inaudible Doppler sensing.

CCS CONCEPTS

• **Human-centered computing** → **Interactive systems and tools**;
• **Applied computing** → *Consumer health.*

KEYWORDS

Mobile Human Activity Recognition, Activity Tracking, Pandemic, Home Workout, Smartphone, Machine-Learning, Neural Networks

ACM Reference Format:
Denys J.C. Matthies, Thorleif Harder, Franz Bretterbauer, Viktoria Ginter, Horst Hellbrück. 2021. FitFone: Tracking Home Workout in Pandemic Times. In *The 14th PErvasive Technologies Related to Assistive Environments Conference (PETRA 2021), June 29-July 2, 2021, Corfu, Greece.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3453892.3461334

1 INTRODUCTION

Pandemic times have a tremendous impact on our physical and mental well-being [9], given strict social distancing conditions [10] and reduced possibilities for physical activity [13, 20]. Performing regular workouts is key factor in maintaining health, which physicians recommend [24]. Yet it is unclear whether outdoor workouts are significantly healthier than those performed indoors [26]. To date, research has shown that tracking outdoor workouts has increasingly become of interest, as it serves as a motivational factor for consistency in exercise routines [21]. Given this, tracking indoor workouts can be of essential importance to maintain health during pandemic times.

PETRA 2021, June 29-July 2, 2021, Corfu, Greece

ACM ISBN 978-1-4503-8792-7/21/06…$15.00
https://doi.org/10.1145/3453892.3461334

Figure 1: MotionFone: Relying on accelerometer and gyroscope data. The smartphone remains in the pocket. SonicFone: Relying on ultrasonic sensing. The smartphone is placed in front of the user.

In this paper, we contribute with an artifact [30], a smartphone-based indoor workout tracking. We explored two different technology approaches (*see Figure 1*):

- `MotionFone`: Using accelerometer & gyroscope
- `SonicFone`: Using speaker & mic. (Doppler sensing)

2 RELATED WORK

Nowadays, typical Human Activity Recognition (HAR) relies on pattern recognition by machine learning [15]. HAR can be divided into six fields: Gesture, Events, Behaviors, Group actions, Atomix actions, and Human-to-Human/Object Interactions [27]. Traditional sensing methods of workout activity is vision-based [8, 23]. However, in Mobile HAR [16], we affix the sensor to the human body through wearables [5]. The variety of sensors incorporated in smartphones contributed to their prevalence for tracking purposes [2].

2.1 Motion Sensing

An Inertial Measurement Unit (IMU) [22] is a motion sensor device, which is fully microelectromechanical [14]. An IMU usually incorporates accelerometers, gyroscopes, and magnetometers. Using these particular motion sensors, we can identify a variety of human activities following Lara et al. categorical grouping [17]: Ambulation (walking, running, sitting,...), Transportation (riding a bus, cycling,...), Phone Usage (Text messaging, call,...), Daily Activity (Eating, drinking, reading,...), Exercise / Fitness (rowing, weight lifting,...), Military (crawling, kneeling,...), and Upper Body (Chewing, speaking, swallowing,...). Focusing on the motion sensors of smart devices, such as wearables and smartphones, researchers have demonstrated the capability to identify a great number of Ambulation and Transportation activities [11]. Meanwhile, using gesture

Table 1: Comparison of average accuracy (F1 score) for each tree classifier.

Classifier	BFTree	CDT	J48Consol.	RandomTree	SimpleCart	J48	J48graft	SPAARC	JCHAIDStar	DPCTree	REPTree
Acc.	94%	78%	94%	94%	94%	92%	90%	88%	86%	94%	78%
Gyro	86%	80%	82%	64%	86%	86%	84%	84%	88%	80%	80%

recognition with smartwatches is more common in Mobile HAR [29, 31]. Smartwatches became powerful in sensing a great variety of activities, such as detecting the tool being used [19] or tracking cardio-workout activity [25]. Tracking running performance based on the steps collected by the smartwatch of smartphone is most common, albeit accuracy differences are significant [3].

2.2 Sonic Sensing

Identifying performed activities via the emitted sound has been previously demonstrated with devices that emit considerable noise, such as a drill, jackhammer etc. [6, 19] Another method is utilizing an ultrasound wave, which is undetectable to humans. Watanabe at el. [28] instrumented a user with speakers and microphones for movement sensing, and gesture and context-recognition. The same principle can be used with a consumer smartphone to detect motion. Fu et al. [7] demonstrated an exercise monitoring using consumer smartphones. Ultrasonic sensing have also shown the capability to detect the following home workout exercises: Bicycle, Squats, and Toe Touches. Their basic implementation utilizes a frequency of 20kHz, while positioning the smart phone within a two-meter reach of the user.

3 MOTIONFONE

As users usually have a smartphone, the goal was to develop a mobile app that can be downloaded and installed. The app should be able to display activities and have a stopwatch feature to record and save the duration of different activities performed. For this purpose, a machine learning approach is used.

3.1 Scope

Our first implementation demonstrates the utilization of motion sensors from the Inertial Measurement Unit (IMU), namely the accelerometer and gyroscope. In our application, the smartphone should rest in the user's pocket while they perform activities. When choosing our activities, we concentrated on fitness activities that address the entire body. We decided to record four activities and several resting activities. The resting activities were grouped into a single activity named *"default"*. The resulting classes are:

- *climbing*
- *jumping jack*
- *running on spot*
- *stretch jumping*
- *resting (default)*

3.2 Model Development

3.2.1 Data Recording. For the model training, we recorded data on the selected activities. For recording, we implemented a self-developed data collector. The collector is an application that records data from selected IMU-sensors. Many frameworks, such as Flutter [18] only provide a low sample rate. To change that, we used an adapted version of the Flutter's package sensors.

Eventually, we recorded 25 training datasets and 10 test datasets per class/activity. To achieve a robust model, care was taken to ensure that the activities varied accordingly during recording. All activities were recorded at different intensities and speeds. Moreover we varied the surfaces on which we performed the activities, such as jumping jack, running-on-spot, and stretch-jump. The reason is that different surfaces produce distinctive degrees of bounce and damping. Five training data sets and two test data sets were recorded on each of the following surfaces: tiles, long pile carpet, short pile carpet, laminate, and parquet. In the stretch jump activity, we also varied the object that the user jumped on: high couch, low couch, medium-high basket couch, (low) windowsill, and dining room chair. The data was recorded, with a sampling rate of 48 Hz. A window size of a little more than two seconds a total of 98 values were also recorded. Two seconds were chosen because activities, such as the stretch-jump and the climb, can take up to 1.5 seconds, depending on the variation of the execution.

3.2.2 Feature Calculation. We developed our model based on our previously recorded dataset. First, the data from our raw files (containing: *AccX, AccY, AccZ, GyrX, GyrY, GyrZ, className*) stored in a .csv file was cleaned, meaning each number of recorded activity had an exact size of 98 samples. The newly generated file was converted into an .arff file by a self-built feature generator. Initially, the feature generator was developed in Java as a command line program. However, for our real-time classifier, we implemented a feature generator running on the phone using Dart [4]. Our app calculated six statistical features *(mean, median, sum, standard deviation, min element, max element)* for all axis together (3d vector norm), and for each individual axis separately to prevent missing characteristic motion that might only be present on a single axis.

3.2.3 Classifier Selection. The goal was to find a computational inexpensive classifier that can run in real-time on the phone. We generated two separate .arff files containing two independent sets: *training.arff* and *test.arff*. We evaluated the following algorithms: *BFTree, CDT, DPCTree, J48, J48Consolidated, J48graft, JCHAINStar,*

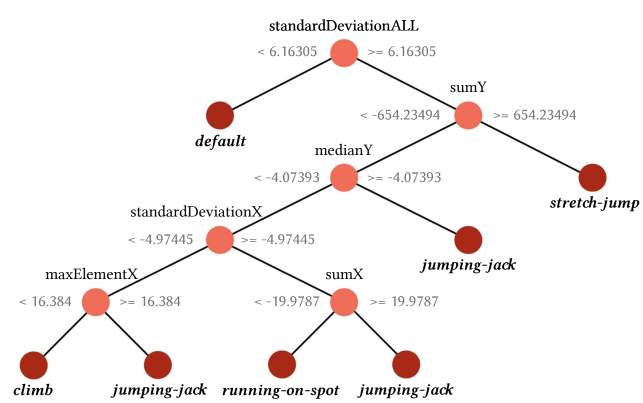

Figure 2: Generated Tree (Classifier: BFTree).

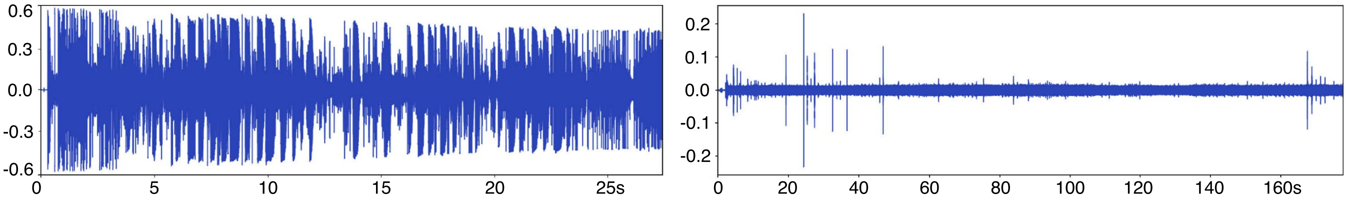

Figure 3: For the data collection, we tested two different phones, a Redmi Note 7 (left) and a Redmi K20 Pro (right). Although the phones had similar equipment, the recorded data varied substantially, which is shown by the plotted data of the jumping jack exercise. The data taken by Redmi Note 7 demonstrates clearer signal patterns.

RandomTree, REPTree, SimpleCart, and *SPAARC* with the Weka Data mining tool [12]. Then, we calculated models and compared their accuracy (*see Table 1*). In considering the overall accuracy and standard deviation, the classifier BFTree showed best results. Figure 2 shows the generated tree.

3.3 Real-Time Classifier

The real-time classifier of our mobile app was developed in Dart and collects motion data with a sampling rate of 48Hz, after the routine *sensor_start()* was manually triggered.As soon as a total of 98 data samples (corresponding to the selected window size of 2s) were collected, the app starts classifying the signal based on the previously calculated BFTree *(see Figure 2)*. We implemented a sliding window approach with an overlap of a single sample (~1%). The detected activity is displayed once detected at least 24 times in a row. This condition is necessary to prevent new activities from being detected all over. We run an evaluation with an unknown user asking him to perform all exercises in a random order for 20 seconds $\approx$ 10 times *(see Figure 4)*.

a	b	c	d	e	<
90	0	0	0	10	**a** - climb
0	100	0	0	0	**b** - default
10	0	90	0	0	**c** - jumping-jack
0	0	0	100	0	**d** - running-on-spot
10	0	0	0	90	**e** - stretch-jump

Figure 4: Confusion Matrix [in %] showing 50 trials (10x each activity) resulting in an accuracy of ~95%.

4 SONICFONE

In our second prototype, we design and implement a home workout-tracking approach without having the smartphone in physical contact with the exercising user.

4.1 Scope

We developed a smartphone application that utilizes ultrasound. Ultrasound waves are emitted through the smartphone's speaker and picked up by the microphone. Making use of the Doppler effect concept and machine learning, we can identify the following movement activities around the device:

- *jumping jack*
- *squat*
- *push up*
- *resting (default)*

4.2 Model Development

4.2.1 Data Recording. We developed a mobile app based on an Android, similar to our previous app. However, we used Flutter to create the user interface of the application, audioplayers library to play the 20kHz sound, and flutter_sound for recording with a sample rate of 44100. The 20 kHz sound was created using AudaCity [1], both the played and recorded file had the .wav format. Although Flutter supports both IOS and Android, the libraries are only compatible with Android.

All activities were recorded multiple times in different rooms on hard ground surfaces. At the beginning, the user places the smartphone in front of him. After choosing the activity, the user pushes 'start' and performs the exercises. 'Stop' is pushed when the participant is done. Depending on the device, the signal can show significant differences in signal quality *(see Figure 3)*.

At the beginning, we check the microphone data for corrupted samples and exclude sounds below 17kHz using a high pass filter. First, we recorded a large amount of data containing 520 exercise executions. Then, we labeled and divided data into segments with a sampling rate of 88,200Hz to catch any frequency shifts.

4.2.2 Neural Network Training. Next, we divided our recorded dataset into training and test data, with a 20% test size fed to the sequential model. A sequential model API is a recurrent neural network, which is powerful for modeling sequence data like sinusoidal waves. The sequential model used contains a stack of linear layers (12), which are responsible for an automated machine learning in neural networks. Our first Dense layer connects our input data (128 units) with an Activation layer, a ReLu. The Activation layer finds whether a neuron is activated or not and transforms the input making it non-linear, thus improving the learning. While there are several kinds of activation functions, the two used here are ReLu and Softmax. A Relu checks whether the values are zero or smaller and sets it to zero once it is negative. The bigger the value the larger the activation. An additional ReLu is connected with another Dense Layer of 256 units. Then, we introduce a Batch Normalization to enable a better comparison when all units are put on the same scale; this can improve training speed and prevent imbalanced gradients. Via a 256 Dense layer, we connect to a Softmax Activation layer. A Softmax function outputs the probability to which class the data may belong. Before introducing a Dropout of 0.5, we connect another ReLu with a Dense layer of 256 units. A Dropout is strongly suggested, as it aims to reduce the complexity of the model to prevent overfitting. We now reduce Dense layer by the number of labels. The output is re-scaled at a final Softmax Activation layer.

4.3 Accuracy Results

The results of this study might not be representative, as we only tested our app with two devices. Not every smartphone is capable of achieving such results based on the difference in sensor quality. Furthermore, execution styles might also differ. In our case, we trained our model based only on two different users. We validated our model with a set of 108 trials *(see Figure 5).*

a	b	c	d	<
97	0	0	3	a - jumping-jack
0	96,2	0	3,8	b - squat
0	0	100	0	c - default
2,13	0	0	98	d - push up

Figure 5: Confusion Matrix [in %] showing 108 trials resulting in an accuracy of ~98%.

With the implementation of a Batch-Normalization-Layer, the accuracy improved from 96% to around 98%. We performed several other tests at which we trained our model with fewer trials and also with both mobile devices (Redmi Note 7 and a Redmi K20 Pro). The confusion somewhat remained at 2%, as most confusions occurred with push-ups.

4.3.1 Limitations. Although accuracy seems high, one needs to consider that our conditions were optimal. An ultrasound tracking is limited by its distance and functionality of the phone's hardware capabilities as Figure 3 demonstrates. Additionally, many electronic devices generate ultrasound frequencies that could interfere and reduce accuracy.

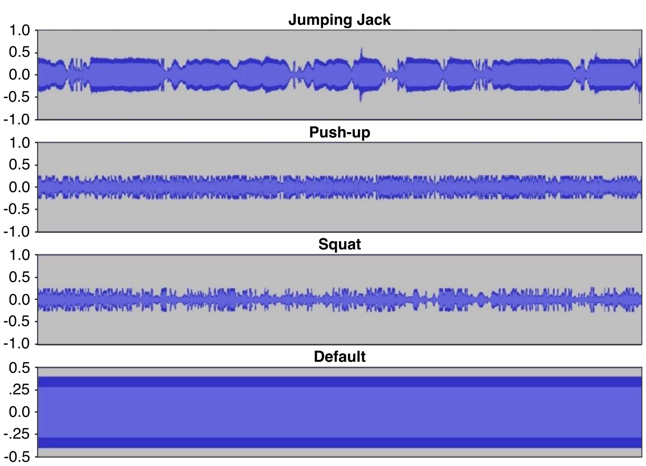

Figure 6: Display the received signal for all activities and the default sinus wave.

5 CONCLUSION & FUTURE WORK

In this paper, we presented two approaches to track home workout using a consumer smartphone. This use case becomes increasingly important during pandemic times, such as the current COVID-19 pandemic. With our first example application, we demonstrate the tracking via motion sensors (accelerometer and gyroscope) when the device is worn by utilizing a conventional machine learning approach. Our second example application demonstrates tracking when the device is doffed in front of the user.

We envision future mobile apps to utilise similar tracking approaches with a greater range of supported exercises. A major challenge seems to be the technical support of different types of devices. An opportunity with body-worn devices is to additionally consider gathered vital signs to adjust the workout plan.

REFERENCES

[1] Installing Audacity. 2013. What is Audacity. (2013).
[2] Rafael Ballagas, Jan Borchers, Michael Rohs, and Jennifer G Sheridan. 2006. The smart phone: a ubiquitous input device. *IEEE PerCom* 5, 1 (2006), 70–77.
[3] Meredith A Case, Holland A Burwick, Kevin G Volpp, and Mitesh S Patel. 2015. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *Jama* 313, 6 (2015), 625–626.
[4] Dart. 2020. Programming language. https://dart.dev
[5] Don Samitha Elvitigala, Denys J. C. Matthies, Chamod Weerasinghe, and Suranga Nanayakkara. 2021. GymSoles++: Combining Google Glass with Smart Insoles to Improve Body Posture when Performing Squats. In *PETRA'21*.
[6] Shaikh Faizan. 2017. Getting Started with Audio Data Analysis (Voice) using Deep Learning, Analytics Vidhya, 23 August 2017. 25 (2017).
[7] Biying Fu, Dinesh Vaithyalingam Gangatharan, Arjan Kuijper, Florian Kirchbuchner, and Andreas Braun. 2017. Exercise monitoring on consumer smart phones using ultrasonic sensing. In *iWOAR'17*.
[8] Preetham Ganesh, Reza Etemadi Idgahi, Chinmaya Basavanahally Venkatesh, Ashwin Ramesh Babu, and Maria Kyrarini. 2020. Personalized system for human gym activity recognition using an RGB camera. In *PETRA'20*. 1–7.
[9] Marina Garriga, Isabel Agasi, Ester Fedida, Justo Pinzón-Espinosa, Mireia Vazquez, Isabella Pacchiarotti, and Eduard Vieta. 2020. The role of mental health home hospitalization care during the COVID-19 Pandemic. *Acta Psych. Scand.* (2020).
[10] Michael Greenstone and Vishan Nigam. 2020. Does social distancing matter? *University of Chicago, Becker Friedman Institute for Economics* 2020-26 (2020).
[11] Marian Haescher, John Trimpop, Denys J. C. Matthies, Gerald Bieber, Bodo Urban, and Thomas Kirste. 2015. aHead: considering the head position in a multi-sensory setup of wearables to recognize everyday activities... In *HCII*. Springer, 741–752.
[12] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I.H. Witten. 2009. The WEKA data mining software: an update. *ACM SIGKDD* 11, 1 (2009), 10–18.
[13] Marc Ashley Harris. 2018. The relationship between physical inactivity and mental wellbeing. *Health psychology open* 5, 1 (2018).
[14] Jack W Judy. 2001. Microelectromechanical systems (MEMS): fabrication, design and applications. *Smart materials and Structures* 10, 6 (2001), 1115.
[15] Eunju Kim, Sumi Helal, and Diane Cook. 2009. Human activity recognition and pattern discovery. *IEEE pervasive computing* 9, 1 (2009), 48–53.
[16] Oscar D Lara and Miguel A Labrador. 2012. A mobile platform for real-time human activity recognition. In *IEEE CCNC*. IEEE, 667–671.
[17] Oscar D Lara and Miguel A Labrador. 2012. A survey on human activity recognition using wearable sensors. *IEEE Communications* 15, 3 (2012), 1192–1209.
[18] Prajyot Mainkar and Salvatore Giordano. 2019. *Google Flutter Mobile Development Quick Start Guide*. Packt Publishing Ltd.
[19] D.J.C. Matthies, G. Bieber, and U. Kaulbars. 2016. AGIS: automated tool detection & hand-arm vibration estimation using an unmodified smartwatch. In *iWOAR'16*.
[20] Frances McKee-Ryan, Zhaoli Song, Connie R Wanberg, and Angelo J Kinicki. 2005. Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of applied psychology* 90, 1 (2005), 53.
[21] Maria D Molina and S Shyam Sundar. 2020. Can mobile apps motivate fitness tracking? A study of technological affordances and workout behaviors. *Health communication* 35, 1 (2020), 65–74.
[22] Melvin M Morrison. 1987. Inertial measurement unit. US Patent 4,711,125.
[23] A. Nagarkoti, R. Teotia, A.K. Mahale, and P.K. Das. 2019. Realtime indoor workout analysis using machine learning & computer vision. In *EMBC*. IEEE, 1440–1443.
[24] T. Rath, J.K. Harter, and J. Harter. 2010. *Wellbeing: The five essential elements.*
[25] Chenguang Shen, Bo-Jhang Ho, and Mani Srivastava. 2017. Milift: Efficient smartwatch-based workout tracking using automatic segmentation. *IEEE TMC* 17, 7 (2017), 1609–1622.
[26] Jo Thompson Coon, Kate Boddy, Ken Stein, Rebecca Whear, Joanne Barton, and Michael H Depledge. 2011. Does participating in physical activity in outdoor natural environments have a greater effect on physical and mental wellbeing than physical activity indoors? A systematic review. *ES&T* 45, 5 (2011), 1761–1772.
[27] Michalis Vrigkas, Christophoros Nikou, and Ioannis A Kakadiaris. 2015. A review of human activity recognition methods. *Frontiers in Robotics and AI* 2 (2015), 28.
[28] Hiroki Watanabe, Tsutomu Terada, and Masahiko Tsukamoto. 2013. Ultrasound-based movement sensing, gesture-, and context-recognition. In *ISWC'13*. 57–64.
[29] Hongyi Wen, Julian Ramos Rojas, and Anind K Dey. 2016. Serendipity: Finger gesture recognition using an off-the-shelf smartwatch. In *CHI'16*. 3847–3851.
[30] Jacob O Wobbrock. 2012. Seven research contributions in HCI. 1 (2012), 52–80.
[31] P. Zhu, H. Zhou, S. Cao, P. Yang, and S. Xue. 2018. Control with Gestures: A Hand Gesture Recognition System Using Off-the-Shelf Smartwatch. In *BIGCOM*.