Poster: A Millimeter-Wave Wireless Sensing Approach for At-Home Exercise Recognition

Edward M Sitar IV; Moh Sabbir Saadat; Sanjib Sur
Department of Computer Science and Engineering; University of South Carolina, Columbia, USA
esitar@email.sc.edu; msaadat@email.sc.edu; sur@cse.sc.edu

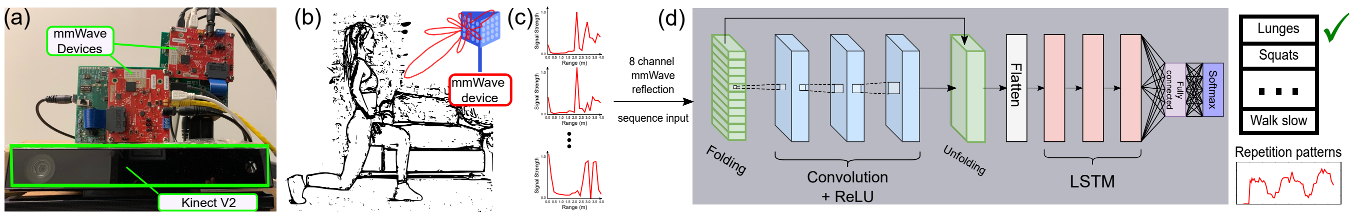

Figure 1: *(a) Data collection setup with two mmWave devices and an RGB-D camera; (b) Example of an individual performing lunges in front of mmWave devices; (c) Reflected signal from one frame of exercise; (d) Classification network architecture.*

ABSTRACT

At-home exercise monitoring is vital to applications like rehabilitative care and physical therapy. In this work, we use millimeter-wave signal reflections to assess the exercise, where we classify the exercise type by designing a supervised deep learning model, and estimate the number of repetitions by leveraging phase information embedded in the reflections.

CCS CONCEPTS

• Human-centered computing → Ubiquitous and mobile computing; **• Computing methodologies → Neural networks**.

KEYWORDS

Millimeter-Wave; Wireless Sensing; Activity Recognition; Deep Learning

ACM Reference Format:
Edward M Sitar IV; Moh Sabbir Saadat; Sanjib Sur, Department of Computer Science and Engineering; University of South Carolina, Columbia, USA, esitar@email.sc.edu; msaadat@email.sc.edu; sur@cse.sc.edu. 2022. Poster: A Millimeter-Wave Wireless Sensing Approach for At-Home Exercise Recognition. In *The 20th Annual International Conference on Mobile Systems, Applications and Services (MobiSys '22), June 25-July 1, 2022, Portland, OR, USA*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3498361.3538781

1 INTRODUCTION

Advances in deep learning frameworks and improvements in sensing devices have enabled smarter at-home personal documentation, such as tracking general activities, recording vital signs without wearables, and detecting falls of elderly individuals. Researchers have used a variety of techniques to build such systems, *e.g.*, optical cameras, wearable sensors, and wireless signals. But the use of RGB cameras for documentation in personal spaces could be privacy-invasive, and the performance largely depends on optimal lighting conditions. Moreover, the use of wearable sensors could be inconvenient since users may be required to wear multiple sensors for detailed data recording. Approaches utilizing Wi-Fi signals operate at low carrier frequency with low bandwidth; so, they are unable to detect small movements.

Thus, we propose to use millimeter-wave (mmWave) wireless signals from 5G-and-beyond smart devices, which have the ability to track human limb movements at finer granularity. Our objective in this work is two-fold: (1) Classify the type of exercise an individual is performing, and (2) Provide a set of temporal assessments of the exercise, such as cycle count, cycle length, rest period, *etc.* Having the ability to classify activity and offer feedback enables applications like remote physical therapy and remote personal training. However, mmWave signals are not readily amenable to generate images, for two reasons: (1) high susceptibility to specular reflections, *i.e.*, the reflected mmWave signals from a body are not entirely directed back at the receiver, and (2) low spatial resolutions due to the use of only a few antennas in commodity devices. Previous works on human activity recognition with mmWave [1, 2] have focused on constructing human silhouettes or skeletons, which, similar to vision-based approaches, raises privacy concerns for users. Others [3, 4] seek to use point cloud data generated from mmWave signals to recognize activity. These approaches require extensive signal processing and generate coarse point clouds which do not contain sufficient information for detailed activity assessment. Inspired by previous works using Wi-Fi [5], our key intuition is that *similar patterns of movement will produce comparable changes in mmWave reflection data*, *i.e.*, we can use the reflected signal directly for our activity monitoring task. We exploit spatial and temporal information in the reflected signals to train a deep learning model for exercise classification, and we leverage the signals' phase information to count the number of repetitions.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
MobiSys '22, June 25-July 1, 2022, Portland, OR, USA
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9185-6/22/06.
https://doi.org/10.1145/3498361.3538781

2 SYSTEM DESIGN AND RESULTS

To achieve spatial resolutions along both the vertical and horizontal directions, we use two millimeter-wave transceivers with the receiving antennas rotated vertically and horizontally. There are four receiving antennas on each of the transceivers and one active transmitting antenna. A co-located RGB-D camera captures ground truth depth images and key joint positions (see Figure 1[a]). The mmWave transceiver uses FMCW chirps with linear frequency sweep along a bandwidth of ≈1.8 GHz. Since the mmWave transceivers capture chirps at a much higher frame rate than the depth sensor, we take the median over all chirps within a frame to keep the frame rate consistent. We process the mmWave data by applying a 256 point FFT, deconstructing the signal into its component frequencies, corresponding to reflections from objects at different ranges. The transceivers capture reflections from a very large range; so, we prune the reflections to discard information about ranges beyond the user's depth, accentuating the exercise signature.

Exercise Classification: We leverage the reflections received by multiple antennas and use them in a customized Convolutional Neural Network (CNN) with Long Short-Term Memory (LSTM) classifier to predict the exercise. The core purpose of CNN is to extract relevant features from the input data using sliding convolutional filters that capture the spatial variations in signals, and the core purpose of LSTM to learn from unique temporal changes for different exercises. At a high level, the network takes mmWave reflections received by 8 antennas as input and outputs the exercise type. The input data is folded so we can apply convolutions to each frame independently, and after applying several convolutions with ReLU activations on each frame, we unfold the data to restore its original sequence structure and flatten it to acquire a feature vector for each frame. We use LSTM layers as gated recurrent components in our network design to capture the temporal changes inherent to each activity. Following the LSTM, we use one fully connected layer followed by a softmax to assign a probability for each of the 18 exercises. The network outputs the exercise with the highest probability as its predicted class. Figure 1(d) shows our network architecture.

Identifying Repetitions within Exercise: We use the difference in phases of the received signal between two receiving antennas to identify the spatio-temporal repetitions in the exercise. Intuitively, the limb and torso movements during activity cause fast and frequent variations of phase while resting periods will cause much slower variation in phase (see Figure 2[b] for lunges exercise).

The reflected mmWave signals have a large bandwidth, but not all these frequencies are strongly reflected by the human body. The weakly reflected components show low variance in phase across the full exercise routine. Thus, we compute variance across all the frames for each of the existing frequencies, and empirically set a threshold of 80% of the maximum variance to filter out the weakly reflected components. To locate the start and end of an exercise repetition, we use a sliding window of 1.2 seconds and 90% overlap in the filtered signal to compute a series of variances across time. To reflect the relative importance between the range bins, we use a weighted summation of the variances for the multiple range bins, where the weights are proportional to the reflection strength in the corresponding range bin.

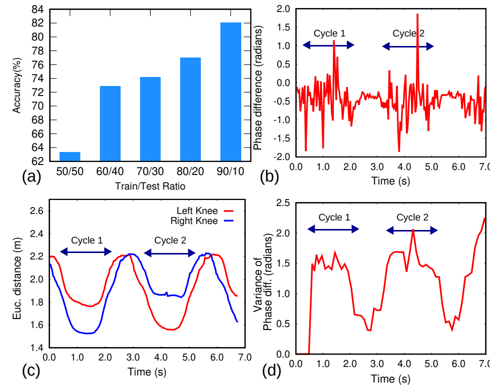

Figure 2: *(a) Classification accuracy for different train/test split; (b) Phase difference; (c) Euclidean distance of knee joints; (d) Variance of phase difference.*

Preliminary Results: We collect 460 data samples of one individual performing 18 exercises, such as forward lunges, push-ups, lateral arm raises, squats, *etc.* We train the CNN-LSTM and find the average prediction accuracy on the test set when using different train/test ratios. We use 20% of the training data for validation, and keep all hyperparameters consistent across all experiments. Figure 2(a) shows that our classification network is able to achieve 82% prediction accuracy on the test set, using 331 total data samples for training. This validates our intuition that different human activities produce sufficiently distinct patterns in mmWave signals for deep learning models to classify activity. Figure 2(d) shows the phase variance for a lunge exercise, which conforms with the ground truth repetitions estimated by the RGB-D camera (Figure 2[c]).

3 CONCLUSION AND FUTURE WORKS

In this work, we propose and preliminarily evaluate a mmWave signal based exercise classification and cycle estimation system. The system does not invade privacy like a camera-based approach and provides finer-grained exercise information. In the future, we plan to extend this work to classify between subjects, evaluate the performance with different input signal representations, such as range-angle and range-doppler responses, estimate duration, tempo, *etc.*, in exercises and deploy and evaluate in practical at-home settings.

ACKNOWLEDGMENTS

We sincerely thank the reviewers for their comments. This work is partially supported by the NSF under grants CAREER-2144505, MRI-2018966, and CNS-1910853.

REFERENCES

[1] A. Adhikari, et al., "MilliPose: Facilitating Full Body Silhouette Imaging from Millimeter-Wave Device," in *ACM UbiComp*, 2021.
[2] A. Sengupta, et al., "mm-Pose: Real-Time Human Skeletal Posture Estimation Using mmWave Radars and CNNs," *IEEE Sensors Journal*, vol. 20, no. 17, 2020.
[3] P. Zhao, et al., "mID: Tracking and Identifying People with Millimeter Wave Radar," in *IEEE DCOSS*, 2019.
[4] A. Singh, et al., "RadHAR: Human Activity Recognition from Point Clouds Generated through a Millimeter-Wave Radar," in *ACM mmNets*, 2019.
[5] Y. Zhu, et al., "FitAssist: Virtual Fitness Assistant Based on WiFi," in *MobiQuitous*, 2019.