Leveraging WiFi Sensing toward Automatic Recognition of Pain Behaviors

1st Xijia Wei
UCL Interaction Centre
University College London
London, UK
xijia.wei.21@ucl.ac.uk

2nd Temitayo Olugbade
School of Engineering & Informatics
University of Sussex
Brighton, UK
t.olugbade@sussex.ac.uk

3rd Fangzhan Shi
Dept of Security & Crime Science
University College London
London, UK
fangzhan.shi.17@ucl.ac.uk

4th Shuang Wu
Dept of CEGE
University College London
London, UK
shuang-wu.22@ucl.ac.uk

5th Amanda Williams
Clinical, Edu & Hlth Psychology
University College London
London, UK
amanda.williams@ucl.ac.uk

6th Nicolas Gold
Dept of Computer Science
University College London
London, UK
n.gold@ucl.ac.uk

7th Youngjun Cho
Dept of Computer Science
University College London
London, UK
youngjun.cho@ucl.ac.uk

8th Kevin Chetty
Dept of Security & Crime Science
University College London
London, UK
k.chetty@ucl.ac.uk

9th Nadia Bianchi-Berthouze
UCL Interaction Centre
University College London
London, UK
n.berthouze@ucl.ac.uk

***Abstract*—WiFi sensing has been well explored for recognizing human activity types. However, research is limited in the possibility of its use in identifying affective expressions such as behaviors associated with pain experience. As critical groundwork, we investigated the use of channel state information from WiFi devices for capturing speed and lateral asymmetry attributes of physical activity. These two attributes are body movement qualities associated with hesitation and guarding, respectively, which are pain behaviors that are valuable to address in physical rehabilitation for people with chronic pain. We obtained mean F1 scores of 0.92 and 0.90 for automatic detection of movement speed levels and lateral asymmetry. These findings suggest that WiFi sensors could be a valuable alternative or supplement to traditional motion capture systems, for unobtrusive, continuous evaluation for hesitation and guarding behaviors in everyday physical activity in the home.**

***Index Terms*—Affect recognition, chronic pain, pain behavior, wireless sensing, machine learning.**

I. Introduction

People with pain employ strategies for performing physical activity in the presence of pain, low movement self-efficacy due to pain, or fear of movement associated with pain [1]–[4]. Some strategies, e.g. initiating sit-to-stand at large knee angles and compensating with the use of the upper limbs for a laterally asymmetrical maneuver during the lift phase [4], [5], are unhelpful and could worsen pain experience. Thus, it can be useful to make people with pain aware of these ‘protective’ pain behaviors during everyday activities, with the aim of guiding them to adopt more valuable coping strategies for dealing with challenging physical activity [5]. Assessment for protective behaviors can also be useful in measuring outcomes of pain conditions and/or management. This paper investigates automatic detection of such behaviors based on the use of WiFi sensing so as to support objective evaluation during everyday physical activity where other movement capture methods have limitations.

Current solutions for capturing body movements rely on cameras or wearable devices. While easy to access, cameras raise privacy concerns and are significantly limited for capture of everyday activity in the home due to the problem of occlusions. Wearable devices, on the other hand, are not as readily available and can be burdensome to keep powered, connected, and worn. These challenges hinder the wide adoption of technologies for physical rehabilitation. WiFi sensing is emerging as an alternative [6]. Their value for human tracking (as well as localization [7]) is based on WiFi signal fluctuations due to movement [8]. These fluctuations can be captured as variations in channel state information (CSI), which describe the scattering, fading, and power decay of wireless signals altered by obstacles. Commodity WiFi routers use CSI to scan signal propagation for the purpose of informing improvement of wireless communication quality. The pervasiveness of WiFi networks makes WiFi sensing a readily available method for capturing human movement data in the home [9]. Previous studies on WiFi sensing have focused on automatic recognition of types of activities (e.g., walking, standing, falling) [10]. Thus, there is still little understanding on how (if at all) WiFi signals can be used for assessing qualities of movements used in executing individual activities.

Speed is one of the movement qualities relevant in the context of pain. Slow movement can be indicative of hesitation or limited joints coordination due to pain or related cognition and affect [5], [11], [12]. For sit-to-stand or stand-to-sit, which are movements critical to everyday activity [13], lateral asymmetry is another pertinent quality [4], [14], [15]. Asymmetry typically involves use of the upper limbs as support in the transfer to/from the seat, with a twist toward one side of body, possibly to minimize loading on the trunk or legs and related to self-efficacy for the movement. As a first step in exploring the possibility of automatic protective behavior detection in people with musculoskeletal pain based on WiFi signals, we investigated the use of data from WiFi routers for differentiating between three speeds of movement in addition to automatic recognition of the type of movement. We focus on sit-to-stand, stand-to-sit, reach-and-grab, and bend-and-pick movements. They are representative of actions in everyday functioning, e.g. reaching forward while vacuuming, and can be challenging for people with pain, especially low back pain. The three classes of speed that we consider (fast, normal, slow) are based on self-definition by individual (healthy) participants, rather than using objective metrics. We further evaluate classification of sit-to-stand and stand-to-sit movement as normal, asymmetric primarily using the left-hand side (asymmetric-left), or asymmetric primarily using the right-hand side (asymmetric-right).

The contributions of this paper can be summarized as follows:

- A novel Hybrid-Fusion Stacked LSTM architecture (HFS-LSTM) for assessment of protective behaviors. Our HFS-LSTM is based on stacked long short-term memory (LSTM) neural networks and performs data fusion at two levels: a) a fusion of amplitude and phase CSI data at the input level, and b) a fusion of frozen pretrained and learnable encodings from (a). The pretrained encoder is trained for activity recognition so as to provide mid-level information about activity type for recognizing protective behavior.
- Evaluation of the HFS-LSTM for automatic detection of 3 movement speed levels for 4 activity types (sit-to-stand, stand-to-sit, reach-and-grab, bend-and-pick) and automatic detection of left and right lateral asymmetries in sit-to-stand and stand-to-sit movements. This was based on pilot data captured from healthy participants in acted settings using a setup of four WiFi devices. We compared our HFS-LSTM with traditional machine learning methods (random forest, support vector machine, k-nearest neighbors, multilayer perceptron).

II. Related Works

A. Automatic Detection of Behaviors Associated with Pain Experience

In contrast to gross assessment of movement, e.g. quantification of physical activity levels using pedometers, assessment of movement behaviors associated with pain experience relies on data from multiple regions of the body [4], [5], [15]. Findings in [16] suggest that RGB videos have been the most dominant approach to capture such data. Beyond these and media based on cameras with depth sensing capability, wearable sensor systems that either only provide acceleration and angular velocity data or also output joint positions and angle information have been the main contenders. The widely-used EmoPain dataset [3] that captures movement data for assessing pain-related expressions is based on a high-fidelity full-body inertia sensor system that consists of 18 gyroscope units and an expensive surface electromyography system. The dataset was, for example, used for protective behavior detection in [17] where both the motion capture data and muscle activity data were fused based on graph convolution and LSTM networks. The use of fewer (and low-cost) sensors, which only capture half-body information, explored in [4] and [18] shows promise although automatic detection performance was found to be lower than using the full sensor set of the EmoPain dataset. While WiFi sensing is an alternative as discussed in Section I, there has been no study of its use for detection of pain expressions during physical activity. The focus in WiFi sensing has largely been on activity recognition [6], which is a lower level of movement abstraction.

B. WiFi-based Activity Recognition

A pioneer study in this area involves the CARM [19], which is based on a hidden Markov model that uses sequential transformed (using principal component analysis) features extracted from CSI data. The model was able to identify eight activities including walking, sitting, falling, and boxing, with an average accuracy of 0.93. Another work is Wi-Chase [20]. It employed statistical features such as mean, standard deviation, percentiles, median absolute deviation, and maximum value from CSI amplitude and phase measurements. These features were fed into a k-nearest neighbours (kNN) model to detect three activity types: running, walking, and moving hands. The study showed that statistical features from CSI data can be useful for activity recognition. However, it is representative of existing WiFi-based systems, which predominantly focus on differentiation between very distinct activities. Although their capacity to distinguish between activities with similar actions is expanding, there is still a lack of discussion around the efficacy of statistical features and temporal information for discrimination between fine-grained qualities of movements. Such qualities can especially encode affective modulations in activity execution [21].

The only study relevant to this is based on the EQ-Radio [22] and used a 7.2 GHz FMCW radar (rather than WiFi devices) to capture chest movement associated with respiration and heartbeats, further using the heart rate data for estimating inter-beat interval (IBI) features. Both respiration and IBI features were then applied to classification into anger, sadness, joy, and pleasure based on a support vector machine (SVM) model, with accuracy of 0.72. Some studies have explored wireless sensing for detecting interruption of movement in the context of Parkinson's disease [23], but no study has explored

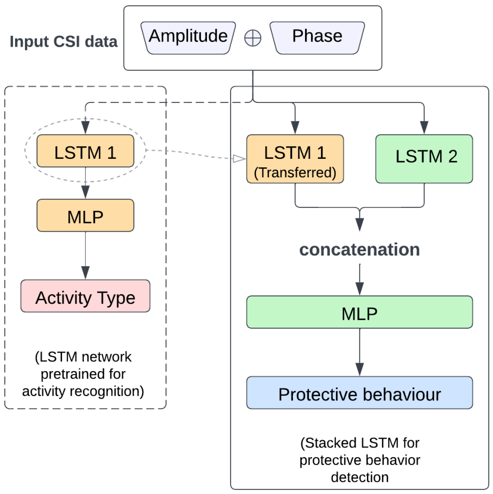


Fig. 1. An overview of the proposed HFS-LSTM architecture for protective behavior detection. In our experiments, we evaluate the HFS-LSTM specifically for automatic recognition of movement speed levels as well as for automatic detection of lateral asymmetry.

the use of WiFi sensing for detection of qualities of body movement associated with pain or related fear.

III. Proposed Hybrid-Fusion Stacked LSTM (HFS-LSTM)

We present a novel Hybrid-Fusion Stacked LSTM architecture (HFS-LSTM), which is illustrated in Figure 1. The use of LSTM [24], [25] to encode temporal relationships is well established in the area of machine learning. So, it was an intuitive building block for our model. The HFS-LSTM fuses two (stacked) LSTM networks such that one of the LSTM networks (LSTM1) captures mid-level information about the type of activity encoded in the signal, while the second LSTM network (LSTM2) captures other encodings relevant for recognition of the given protective behavior.

LSTM1 is set up to encoder activity type information by pretraining it separately for activity recognition. The LSTM layers are then re-used in the HFS-LSTM. In training the HFS-LSTM, the weights of LSTM1 are frozen, while only the weights of LSTM2 are updated. The outputs from both LSTM networks are concatenated and fed into a four-layer multilayer perceptron (MLP), for classification of the given protective behavior label. In our experiments below, the labels are either movement speed level or lateral asymmetry.

An additional level of fusion occurs earlier in the HFS-LSTM with input-level fusion of amplitude and phase signals extracted from the CSI data for each of the two LSTM networks (LSTM1 and LSTM2). The use of both CSI amplitude and phase data as the input follows the approach of previous work, e.g. [20] using CSI data for activity recognition.

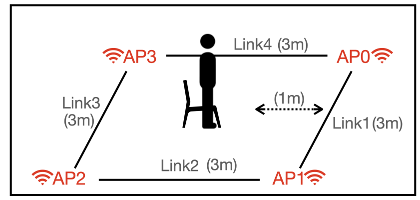


Fig. 2. An illustration of the layout of our WiFi sensor system in a room. The layout consists of four WiFi devices place in the four corners of the room respectively. The participant performs the activities of interest within the space enclosed by the sensor layout.

TABLE I
Summary of The Mean Speeds in Seconds (and Number of Instances) over All Participants Associated with the 3 Different Movement Speed Levels Captured in our Data

Speed level	Sit-to-stand	Stand-to-sit	Bend	Reach
Fast	0.9s (39)	1.3s (38)	2.7s (12)	1.9s (14)
Normal	3.2s (69)	1.4s (71)	2.9s (23)	2.3s (24)
Slow	4.31s (129)	1.5s (126)	3.0s (21)	2.8s (21)

s represents time in seconds.

IV. Experimental Setup

A. Data Collection

1) Sensor system setup and feature extraction approach: We implemented a custom two-way communication protocol proposed in [26] and systematically elaborated in [27] with the Nexmon system [28], [29]. This protocol addresses the lack of synchronization for commercial radio devices which limits the extraction of valuable data from CSI. In the rest of this subsection, we first provide a brief overview of feature extraction from CSI to highlight the need for synchronization before further elaborating on the sensor system setup that we used.

CSI is a set of complex numbers used to describe the radio propagation channel. Since complex numbers have two dimensions, they can depict the phase and amplitude response of electromagnetic wave propagation channels simultaneously. CSI is estimated from the fixed preambles in the protocol, and in the idealized case that the hardware is perfect and synchronized, the CSI would be:

$$H(i,t) = \sum_{k=1}^{N} a_k(i,t) e^{-j2\pi(f_c+f_i)\tau_k(t)} \tag{1}$$

where t represents time, i is the subcarrier index, N is the number of paths, f_c is the centre frequency, f_i is the baseband frequency of subcarrier i, a_k is the complex number representing the attenuation and initial phase offset of path k, and τ_k is the propagation delay of path k. In general, the magnitude of a_k depends on the size of the reflection surface and the incident angle, while τ_k is very sensitive to distance. Since the final CSI (Equation 1) is the sum of the channel response of propagation paths, it is influenced by signal reflections against the human body. Thus, the amplitude of CSI encodes constructive and destructive signal interference patterns. The phase of the CSI could further be valuable, however, as it is dependent on time, it requires synchronization of the radio devices.

The protocol that we used synchronizes two devices using wireless communication that eliminates phase errors caused by clock asynchronization. In order to capture body movement simulataneously from multiple viewpoints, we built a sensor system with 4 WiFi devices (see the system layout in Figure 2). One WiFi device was placed in each corner of the room to allow for capture of movement qualities dependent on spatial configuration, thereby enhancing sensing capabilities for enriched information.

2) Data Capture Procedure: Our study was approved by the local research ethics committee at University College London, and participants gave consent for their data to be collected and processed for the purpose of this study. We invited 10 healthy participants (7 male, 3 female) to our lab. They were university students between 18 and 30 years old (height=155-190cm).

As illustrated in Figure 3, the participants were asked to perform four activity types selected for their relevance to everyday physical activity: sit-to-stand, stand-to-sit, reach-and-grab, and bend-and-pick. The sit-to-stand and stand-to-sit movements were executed as normal or with acted left and right asymmetries that are illustrative of lateral asymmetries due to protective behavior in people with chronic pain. These variants and the other two activity types were done at three different speed levels, fast, normal, and slow, to capture slowness in movement (or hesitation) that has been associated with chronic pain.

B. Data Preprocessing

The raw CSI data recorded by the WiFi sensor (exceeding 4GB) were converted into amplitude and phase signals (resulting in data less than 300MB) based on the method described in Section IV-A1. Figure 4-left shows example amplitude signals for sit-to-stand, bend-and-pick, and reach-and-grab movements, with the 3 different variants of the sit-to-stand (i.e. with and without lateral asymmetries), for the same participant. Distinction in the signal pattern for each of these movements can be seen in the figure. The total number of data instances for training the activity recognition model, lateral asymmetry model and speed levels model are 587, 472, and 587 respectively.

There were data quality issues for the first participant's recording. One WiFi router lost connection due to environmental noise and interference from other WiFi sensors during that data capture. Consequently, the dataset is missing values (from the offline WiFi sensor) for that participant. Nevertheless, we included this participant's data in our experiments. This is to enable robust evaluation of WiFi sensing for protective behavior detection given that noise and interference are real problems that must be tackled in its use. In future work, missing values could be addressed with fusion of data from multiple sensor types as explored in [30]. There were two forms of additional preprocessing done on the amplitude and phase signals:

TABLE II
HFS-LSTM ARCHITECTURE SUMMARY

Layers	Output Shape
Transferred LSTM1 network	(256,)
LSTM2 Layer	(256,)
Concatenation	(256,)
MLP	(256, 128, 64, 32)
Output Layer	(One-hot encoded output)
Batch Size	32
Learning Rate	0.005
Optimizer	Adam

1) Removing noise due to stationary objects from the phase signal: As both static (e.g., furniture, wall) and moving obstacles (human targets) contribute to the variation in the signal multipath reflection that affects the CSI, variations contributed by stationary obstacles needs to be removed from the phase signal. For this, we used a classic Butterworth filter of order of 5, cut-off frequency of 25Hz, normalized cut-off frequency of 0.5 (cut-off frequency/Nyquist frequency), and sampling frequency of 100Hz.

2) Removing noise due to phase drift from the phase signal: Phase drift is inevitable due to changes in the hardware's working temperature, ageing of components, and the piezoelectric effect in the crystal oscillator [31]. To check for phase drift, we attached two directional antennas close to each other with no gap in between (to avoid any obstacle's impact on signal communication) and recorded two minutes of CSI data. We implemented a 1Hz lowpass filter (order=5) on this recorded signal to remove moving target reflections. We observed a linear phase drift over time (see Figure 5 shows 3D and 2D visualizations of the phase drift).

To remove the phase drift from collected data, we implemented a simple moving average filter with a time window of 100 milliseconds. The moving average filter is commonly used for the time-series signal denoising [32]. By setting the time window of 100 milliseconds, we assumed that the phase drift is negligible. Figure 6 compares the same phase signal with and without implementing a simple moving average filter applied. We found that in the raw signal, phase drift obscures the phase signatures of body movement considerably, while the filtered signal shows a reasonable phase change fluctuating over zero radians and reflects motion orientation toward or away from the antennae.

Figure 4-right shows example phase signals, highlighting distinction in the signal pattern for sit-to-stand (and its lateral asymmetry variants), bend-and-pick, and reach-and-grab movements for the same participant.

V. RESULTS AND DISCUSSION

A. Automatic Detection of Lateral Asymmetry and Speed Levels with Transfer Learning

We used a 4-layer LSTM network (with an additional 1-layer MLP as the output layer) for activity recognition. On satisfactory finetuning of the model, we extracted the network's 4 LSTM layers and transferred them to the HFS-LSTM models for asymmetry detection and speed level detection

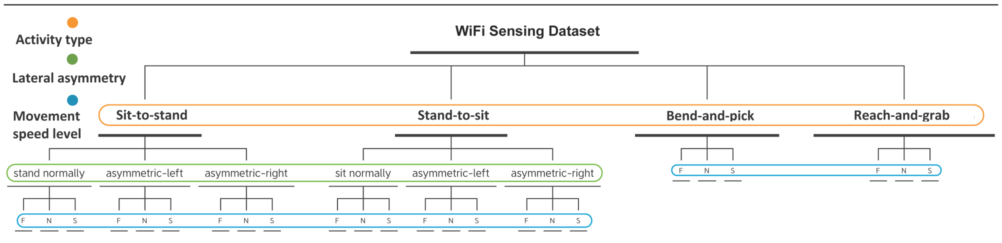


Fig. 3. An overview of the data capture procedure. There are four activity types (Level 1 from the top) performed by the participants: sit-to-stand, stand-to-sit, reach-and-grab, and bend-and-pick. The sit-to-stand and stand-to-sit movements were further executed as normal or with left and right asymmetries (Level 2). Each activity (and the different symmetry variants) was done in three different speeds, fast (F), normal (N) and slow (S) (Level 3).

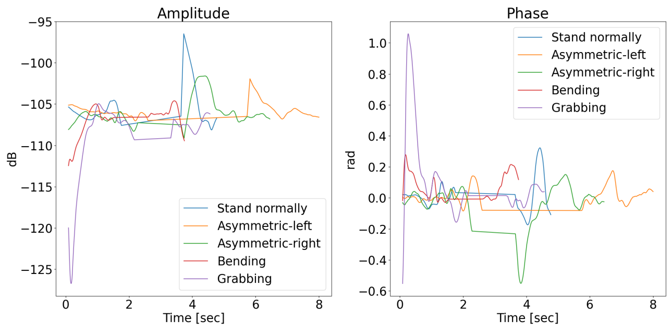
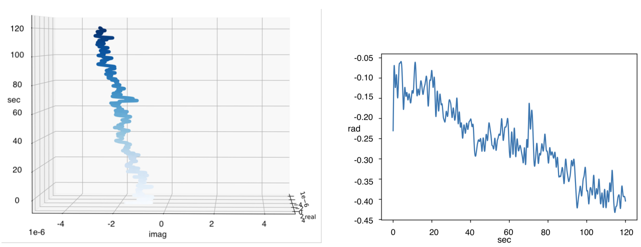
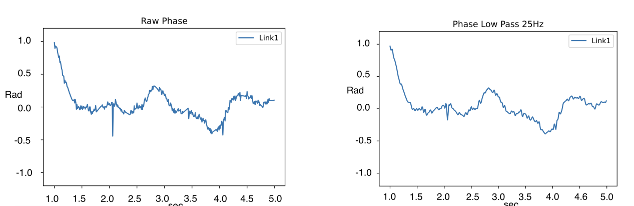


Fig. 4. Amplitude (left) signals, in dB for decibels, and phase (right) signals, in rad for radians, from CSI data for sit-to-stand (normal and in left and right asymmetries), bend-and-pick, and reach-and-grab movements for a single participant. sec denotes time in seconds.



(a) Phase Drifting in 3D (b) Phase Drifting in 2D

Fig. 5. Example of the effect of phase drifting issue, visualized in 3D (Left) and 2D (Right).



(a) Raw Phase without MAF (b) Filtered Phase with MAF

Fig. 6. Example of the effect of the moving average filter (MAF) on the phase signal, showing the raw signal (left) and the filtered signal (right).

(see Section III for details about the HFS-LSTM architecture). The HFS-LSTM models for the two tasks are similar in architecture. The number of units in each layer of the HFS-LSTM models and the other hyperparameters are specified in Table II. However, while the model for automatic detection of movement speed levels is trained using the full dataset, the model of asymmetry detection is only trained on sit-to-stand and stand-to-sit movement data. We evaluated all three models using in a leave-one-subject-out cross-validation approach which involves the systematic exclusion of one participant from the dataset in each fold of the training and evaluation. This strategy assesses the models' generalizability to people not included in the training data. All models were trained on a Windows 11 PC with a 12th Generation Intel CPU and an NVIDIA GTX 1660 Super GPU with CUDA acceleration during the training process, and the total training time for all models (including the comparison models discussed in Section V-B) was approximately 50 minutes.

As can be seen in Figure 7 and in the last rows of Tables III and IV, the HFS-LSTM performs very well for both asymmetry and speed level detection. We obtained mean F1 scores of 0.90 and 0.92 respectively. True positive rates are similar across classes for both tasks. F1 scores are also similar across participants although performance for Participant A is considerably low. This is due to the missing values in the signals for that participant as discussed in Section IV-B. This suggests that missing values need to be carefully addressed so that they do not undermine the efficacy of WiFi sensing for protective behavior detection. Nonetheless, performance for the other participants are very high, despite the inclusion of the signals with missing values in the training data used for prediction. Although activity recognition performance is also high, it is slightly lower than the performance for the other two tasks. The HFS structure of the LSTM model used in the asymmetry and speed level detection tasks, compared to the vanilla LSTM used for activity recognition, may have contributed to the higher performance. This is especially likely given the higher level of abstraction of the two tasks.

B. Comparison with Standard Machine Learning Approaches

We compared the performance of our HFS-LSTM with traditional machine learning algorithms: random forests (RF) [33], kNN, SVM [34], and MLP. For these algorithms,

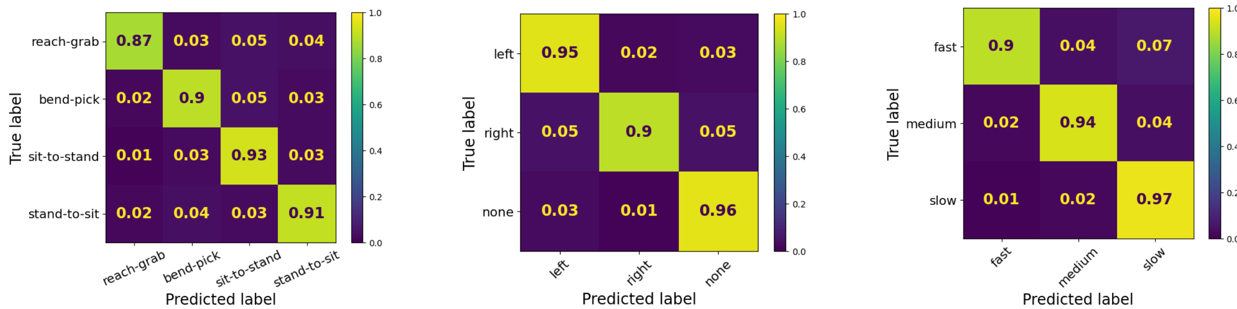


Fig. 7. Confusion matrices (showing proportions across respective rows) for automatic recognition of activity type (left), lateral asymmetry (middle), and movement speed levels (right).

we extracted six aggregate features from the amplitude and phase timeseries based on previous studies on WiFi sensing for activity recognition (see Section II-B). We computed minimum, maximum, standard deviation, variance, skewness, and kurtosis for each of the two signals.

The performance of the four models for activity recognition are shown in the top rows of Tables III and IV. Performance is better than chance level classification (0.25), i.e. random guess, for all four algorithms although true positive is consistently poor for stand-to-sit. The poor performance for stand-to-sit is due to confusion with the other three activity types especially sit-to-stand and reach-and-grab (confusion matrix not shown here). This is not surprising given that all four activity types involve trunk flexion. True positive is highest for bend-and-pick (with the RF and MLP) and for sit-to-stand (with the SVM and kNN). Performance also varies widely across subjects particularly for the SVM where mean F1 score ranges between 0.19 and 0.56 across people. Overall, all four algorithms have similar performances although the kNN has the lowest mean true positive rate and F1 score. Neither of them matches the very good performance of the LSTM in terms of mean F1 score, highlighting the importance of temporal information for activity recognition. The better performance of the four models for Participant A compared to the LSTM suggests that aggregate features minimized the impact of the missing values in the data signals captured for that subject, while the absence of those values was especially amplified in the LSTM that focused on temporal relations.

As can be seen in the second rows of Tables III and IV, all four models perform better than chance level detection (0.33) for recognition of lateral asymmetry classes for the sit-to-stand and stand-to-sit movements. The RF has the best performance while the kNN again has the lowest performance. Performance for each class varies widely across the models. For example, right asymmetry has the best true positive rate for RF, while it is best for left asymmetry with the SVM and best for no asymmetry with the MLP. Similar to the findings for activity recognition, the HFS-LSTM outperforms the four models considerably. This is not surprising given that unlike aggregate information based on joint angle or position data, similar metrics (e.g. mean, standard deviation) for CSI amplitude and phase signals do not encode spatial and temporal information. For standard machine learning algorithms, including MLP, kNN, and SVM, which take aggregate information inputs, the algorithms can only make recognition based on overall CSI variation signatures, lacking temporal features. Nevertheless, asymmetry detection is high for Participant D with RF and MLP which have F1 score of 0.82 and 0.83 respectively (performance of the HFS-LSTM for that participant is 0.90).

Tables III and IV show the performance of the four models on speed level detection. As expected given that speed is strongly related to time, none of the four models performs well, with mean F1 score less than chance level classification (0.33) for the kNN and mean F1 scores between 0.38 and 0.41 for the other three models.

TABLE III
TRUE POSITIVE RATES FOR AUTOMATIC RECOGNITION OF ACTIVITY TYPE, LATERAL ASSYMETRY, AND MOVEMENT SPEED LEVELS

	Activity Recognition Performance				
	Sit-to-stand	Stand-to-sit	Bend	Reach	Mean
RF	0.42	0.08	0.63	0.33	0.36
SVM	0.58	0.11	0.39	0.34	0.35
KNN	0.5	0.1	0.37	0.32	0.32
MLP	0.33	0.15	0.61	0.41	0.37
LSTM	0.88	0.89	0.91	0.91	0.89
	Asymmetry Detection Performance				
	None	Left	Right		Mean
RF	0.48	0.38	0.73		0.53
SVM	0.3	0.65	0.52		0.49
KNN	0.44	0.38	0.44		0.42
MLP	0.72	0.24	0.43		0.46
HFS-LSTM	0.93	0.90	0.96		0.93
	Speed Level Detection Performance				
	Fast	Normal	Slow		Mean
RF	0.02	0.11	0.88		0.33
SVM	0.2	0.17	0.69		0.35
KNN	0.28	0.34	0.31		0.31
MLP	0.11	0.19	0.85		0.38
HFS-LSTM	0.91	0.94	0.97		0.93

CONCLUSION AND ETHICAL IMPACT STATEMENT

Our research explores the potential of WiFi CSI for detecting the qualities of body movement during everyday physical

TABLE IV
F1 SCORES BY SUBJECT FOR PARTICIPANTS A TO J

Activity Recognition Performance

	A	B	C	D	E	F	G	H	I	J	Mean
RF	0.45	0.35	0.31	0.29	0.35	0.56	0.38	0.38	0.34	0.54	0.39
SVM	0.56	0.49	0.27	0.45	0.48	0.48	0.19	0.4	0.4	0.4	0.41
KNN	0.37	0.41	0.29	0.27	0.38	0.28	0.27	0.37	0.42	0.35	0.34
MLP	0.34	0.37	0.44	0.44	0.35	0.4	0.43	0.42	0.2	0.34	0.37
LSTM	0.24	0.71	0.71	0.81	0.98	0.98	0.99	0.99	1.00	1.00	0.84

Asymmetry Detection Performance

	A	B	C	D	E	F	G	H	I	J	Mean
RF	0.58	0.72	0.55	0.82	0.52	0.6	0.48	0.47	0.2	0.53	0.54
SVM	0.48	0.67	0.39	0.59	0.5	0.72	0.38	0.43	0.45	0.46	0.50
KNN	0.55	0.52	0.27	0.6	0.38	0.52	0.36	0.41	0.35	0.4	0.43
MLP	0.27	0.26	0.48	0.83	0.49	0.71	0.25	0.43	0.36	0.45	0.45
HFS-LSTM	0.43	0.97	0.77	0.90	1.00	0.99	0.99	0.97	1.00	1.00	0.90

Speed Level Detection Performance

	A	B	C	D	E	F	G	H	I	J	Mean
RF	0.45	0.1	0.49	0.31	0.62	0.31	0.41	0.4	0.42	0.45	0.39
SVM	0.44	0.38	0.2	0.38	0.55	0.28	0.46	0.41	0.23	0.5	0.38
KNN	0.34	0.34	0.33	0.26	0.39	0.31	0.25	0.36	0.24	0.39	0.32
MLP	0.42	0.22	0.44	0.37	0.54	0.34	0.46	0.6	0.26	0.45	0.41
HFS-LSTM	0.40	0.94	0.88	0.98	1.00	0.99	0.99	1.00	1.00	0.995	0.92

activity in the home. Our motivation for pursuing WiFi sensing is for the valuable role that it could play in automatic detection of protective behaviors that enable tailored technology-based interventions for people with chronic pain. For example, a person who consistently uses unhelpful strategies to execute sit-to-stand movements in their home could be encouraged to use a higher seat while confidence in the movement is low and gradually return to the lower seat as they build confidence in the movement. Our findings of true positive rates of 0.90 and 0.96 for detecting left and right asymmetries in sit-to-stand and stand-to-sit movements suggest that WiFi sensing could indeed proffer a solution to the sensing challenge that has undermined progress in that direction [16], [35]. Further, our findings of true positive rates of 0.97 for detecting slow movements highlight the possibility of its use for detecting a useful range of protective behaviors.

Our findings are based on acted data in lab settings as a first step. Future work will explore generalizability in settings closer to the real world, e.g., where there may be multiple (moving) people within the home or building beyond the target person, or the activities (such as vacuuming) may involve moving objects. It may be necessary to employ multimodal detection to take advantage of other sensor modalities in addressing the challenge of missing data due to such environmental noise. Another limitation of our work is that we do not explore the influences of different sensor setup attributes (e.g., transmitter-receiver distances) and other movement characteristics beyond speed and lateral asymmetry (e.g., range of motion) on the CSI data. Still, our work and findings are important contributions to advance in the area of WiFi sensing for automatic detection of affective expressions, which is currently largely ignored despite its significance.

Beyond these, it is further important for the community to address the potential risks that could come with the use of WiFi devices for sensing, so as to maximize the value for digital health technology. While real world use of WiFi sensing is still limited and third parties are not currently allowed to directly access CSI from routers installed by broadband service providers, it is critical for the community (research, industry, regulatory) to consider ethical implications before it becomes adopted. This is due to the high surveillance risk that comes with WiFi sensing. Examples of misuse include illicit use of own WiFi sensor systems to spy on neighbors. This is especially relevant for those who live in apartment buildings with thin walls between flats. Unauthorized access into WiFi systems in a person's home could also be used to track their presence at home as well as their activities when at home. The risk is especially high given the increasing cases of cyber crimes globally. There is the possibility that criminals will integrate localization or activity information from WiFi systems gained illegitimately with other personal data, e.g. accessed on the dark web or similarly unlawful sources of personal data, for more targeted attacks on individuals. We suggest a two-pronged solution to mitigate this risk. First, we recommend the use of state-of-the-art technical approaches for security. For example, sensing information from either transmitter or receiver should be encrypted. Random channels could also be used for communication as part of the guard against unauthorized surveillance. Systems can further be set up such that their owners need to give explicit permission (e.g., based on hardware encryption) for ambient sensing purposes. Second, we advocate for increased transparency and public awareness regarding WiFi sensing and its associated risks. Through widespread education about this technology and its risk, individuals, regulators, and other stakeholders could be empowered to make informed decisions that protect personal data privacy.

Acknowledgment

This work was supported by the EU Future and Emerging Technologies Proactive H2020 programme (Grant 824160: EnTimeMent). For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

References

[1] A. Asghari and M. K. Nicholas, "Pain self-efficacy beliefs and pain behaviour: A prospective study," *Pain*, vol. 94, no. 1, pp. 85–100, 2001.

[2] G. Crombez, J. W. S. Vlaeyen, P. H. T. G. Heuts, and R. Lysens, "Pain-related fear is more disabling than pain itself: Evidence on the role of pain-related fear in chronic back pain disability," *Pain*, vol. 80, no. 1-2, pp. 329–339, 1999.

[3] M. S. H. Aung, S. Kaltwang, B. Romera-Paredes, B. Martinez, A. Singh, M. Cella, M. Valstar, and et al., "The automatic detection of chronic pain-related expression: Requirements, challenges and the multimodal emopain dataset," *IEEE Transactions on Affective Computing*, vol. 7, no. 4, pp. 435–451, 2016.

[4] T. Olugbade, N. Bianchi-Berthouze, N. Marquardt, and A. C. d. C. Williams, "Human observer and automatic assessment of movement related self-efficacy in chronic pain: From exercise to functional activity," *IEEE Transactions on Affective Computing*, vol. 11, no. 2, pp. 214–229, 2018.

[5] T. Olugbade, A. Singh, N. Bianchi-Berthouze, N. Marquardt, M. S. H. Aung, and A. C. D. C. Williams, "How can affect be detected and represented in technological support for physical rehabilitation?" *ACM Transactions on Computer-Human Interaction*, vol. 26, no. 1, pp. 1–29, 2019.

[6] X. Wei, Z. Wei, and V. Radu, "Sensor-fusion for smartphone location tracking using hybrid multimodal deep neural networks," *Sensors*, vol. 21, no. 22, p. 7488, 2021.

[7] X. Wei and V. Radu, "Calibrating recurrent neural networks on smartphone inertial sensors for location tracking," in *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2019, pp. 1–8.

[8] X. Wei, Z. Wei, and V. Radu, "Mm-loc: Cross-sensor indoor smartphone location tracking using multimodal deep neural networks," in *2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2021, pp. 1–8.

[9] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A survey on behavior recognition using wifi channel state information," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 98–104, 2017.

[10] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, "Wireless sensing for human activity: A survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1629–1645, 2019.

[11] G. L. K. Shum, J. Crosbie, and R. Y. W. Lee, "Effect of low back pain on the kinematics and joint coordination of the lumbar spine and hip during sit-to-stand and stand-to-sit," *Spine*, vol. 30, no. 17, pp. 1998–2004, 2005.

[12] T. J. Somers, F. J. Keefe, J. J. Pells, K. E. Dixon, S. J. Waters, P. A. Riordan, J. A. Blumenthal, and et al., "Pain catastrophizing and pain-related fear in osteoarthritis patients: Relationships to pain and disability," *Journal of Pain and Symptom Management*, vol. 37, no. 5, pp. 863–872, 2009.

[13] P. M. Dall and A. Kerr, "Frequency of the sit to stand task: An observational study of free-living adults," *Applied Ergonomics*, vol. 41, no. 1, pp. 58–61, 2010.

[14] C. L. Christiansen and J. E. Stevens-Lapsley, "Weight-bearing asymmetry in relation to measures of impairment and functional mobility for people with knee osteoarthritis," *Archives of Physical Medicine and Rehabilitation*, vol. 91, no. 10, pp. 1524–1528, 2010.

[15] C. Wang, M. Peng, T. Olugbade, N. D. Lane, A. C. d. C. Williams, and N. Bianchi-Berthouze, "Learning temporal and bodily attention in protective movement behavior detection," in *Proceedings of the International Conference on Affective Computing and Intelligent Interaction Workshops and Demos*, 2019, pp. 324–330.

[16] T. Olugbade, M. Bieńkiewicz, G. Barbareschi, V. D'amato, L. Oneto, A. Camurri, C. Holloway, and et al., "Human movement datasets: An interdisciplinary scoping review," *ACM Computing Surveys*, vol. 55, no. 6, pp. 1–29, 2022.

[17] G. Cen, C. Wang, T. A. Olugbade, A. C. d. C. Williams, and N. Bianchi-Berthouze, "Exploring multimodal fusion for continuous protective behavior detection," in *2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2022, pp. 1–8.

[18] T. Olugbade, R. A. Buono, A. C. d. C. Williams, S. de Ossorno Garcia, N. Gold, C. Holloway, and N. Bianchi-Berthouze, "Emopain (at) home: Dataset and automatic assessment within functional activity for chronic pain rehabilitation," in *Proceedings of the International Conference on Affective Computing and Intelligent Interaction*, 2022, pp. 1–8.

[19] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, vol. 2015-September, pp. 65–76, 9 2015.

[20] S. Arshad, C. Feng, Y. Liu, Y. Hu, R. Yu, S. Zhou, and H. Li, "Wi-chase: A wifi based human activity recognition system for sensorless environments," in *2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. IEEE, 2017, pp. 1–6.

[21] M. Karg, A.-A. Samadani, R. Gorbet, K. Kühnlenz, J. Hoey, and D. Kulić, "Body movements for affective expression: A survey of automatic recognition and generation," *IEEE Transactions on Affective Computing*, vol. 4, no. 4, pp. 341–359, 2013.

[22] M. Zhao, F. Adib, and D. Katabi, "Emotion recognition using wireless signals," in *Proceedings of the 22nd annual international conference on mobile computing and networking*, 2016, pp. 95–108.

[23] A. Tahir, J. Ahmad, S. A. Shah, G. Morison, D. A. Skelton, H. Larijani, Q. H. Abbasi, M. A. Imran, and R. M. Gibson, "Wifreeze: Multiresolution scalograms for freezing of gait detection in parkinson's leveraging 5g spectrum with deep learning," *Electronics*, vol. 8, no. 12, p. 1433, 2019.

[24] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[25] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," *Neural computation*, vol. 12, no. 10, pp. 2451–2471, 2000.

[26] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single WiFi access point," in *Proceedings of the 13th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2016*, 2016.

[27] F. Shi, W. Li, C. Tang, P. Brennan, and K. Chetty, "Doppler sensing using wifi round-trip channel state information," in *2023 IEEE Wireless Communications and Networking Conference (WCNC23)*, 2023.

[28] M. Schulz, D. Wegemer, and M. Hollick, "Nexmon: The C-based Firmware Patching Framework," 2017. [Online]. Available: https://nexmon.org

[29] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free Your CSI: A Channel State Information Extraction Platform For Modern Wi-Fi Chipsets," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, ser. WiNTECH '19, 2019, pp. 21–28. [Online]. Available: https://doi.org/10.1145/3349623.3355477

[30] X. Wei and V. Radu, "Leveraging transfer learning for robust multimodal positioning systems using smartphone multi-sensor data," in *2022 IEEE 12th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2022, pp. 1–8.

[31] D. R. Brown, R. Mudumbai, and S. Dasgupta, "Fundamental limits on phase and frequency tracking and estimation in drifting oscillators," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2012, pp. 5225–5228.

[32] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Proceedings of the 21st annual international conference on mobile computing and networking*, 2015, pp. 65–76.

[33] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[34] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273–297, 1995.

[35] T. Olugbade, "Where are movement sensing, emotion inferring technologies in the real world?" https://uclic.ucl.ac.uk/research/affective-computing/entimement/where-are-movement-sensing-emotion-inferring-technologies-in-the-real-world/, 2021, [Online; accessed 17-May-2023].