
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Cao, Yangjie; Wang, Fuchao; Lu, Xinxin; Lin, Nan; Zhang, Bo; Liu, Zhi; Sigg, Stephan
Contactless Body Movement Recognition during Sleep via WiFi Signals

Published in:
IEEE Internet of Things Journal

DOI:
[10.1109/JIOT.2019.2960823](https://doi.org/10.1109/JIOT.2019.2960823)

Published: 01/03/2020

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Cao, Y., Wang, F., Lu, X., Lin, N., Zhang, B., Liu, Z., & Sigg, S. (2020). Contactless Body Movement Recognition during Sleep via WiFi Signals. *IEEE Internet of Things Journal*, 7(3), 2028-2037. Article 8936937. <https://doi.org/10.1109/JIOT.2019.2960823>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Contactless Body Movement Recognition during Sleep via WiFi Signals

Nan Lin, Fuchao Wang, Xinxin Lu, Yangjie Cao, *Member, IEEE*, Bo Zhang, *Member, IEEE*, Zhi Liu, *Member, IEEE*, and Stephan Sigg, *Member, IEEE*

Abstract—Body movement is one of the most important indicators of sleep quality for elderly people living alone. Body movement is crucial for sleep staging and can be combined with other indicators such as breathing and heart rate to monitor sleep quality. Nevertheless, traditional sleep monitoring methods are inconvenient and may invade users' privacy. To solve these problems, Contactless Body Movement Recognition (CBMR) method via WiFi signals is proposed. Firstly, CBMR uses the commercial off-the-shelf WiFi devices to collect Channel State Information (CSI) data of body movement and segment the CSI data by sliding window. Then, the context information of the segmented CSI data is learned by a Bi-directional Recurrent Neural Network (Bi-RNN). Bi-RNN can fuse the forward and backward propagation information at some point, and input it into a deeper Independently Recurrent Neural Network (IndRNN) with residual mechanism to extract the deeper features and capture the time dependencies of CSI data. Finally, the type of body movement can be recognized and classified by the softmax function. CBMR can effectively reduce data preprocessing and the delay caused by manually extracting features. The results of the experiment conducted on a complex body movement dataset show that our method gives desirable performance and achieves average accuracy of greater than 93.5%, which implies a prospect application of CBMR.

Index Terms—body movement recognition, WiFi, CSI data, deep learning, RNN

I. INTRODUCTION

ACCORDING to statistics, elderly population over the age of 65 will reach about 1.5 billion by 2050 [1]. Furthermore, the number of elderly people living alone will also increase. The health problems of elderly people living alone have aroused great attention of the society and families. Sleep quality of the elderly is an important indicator to measure whether the body is healthy or not. Investigations show that poor-quality sleep may lead to various health issues such as high blood pressure [2], depression [3], and migraine [4], etc., which seriously threaten the normal life of elderly people living alone.

To monitor people's body movements during sleep, the researchers used a variety of sleep monitoring technologies.

This research is supported by Nature Science Foundation of China (Grant No. 61602230), the Research Foundation Plan in Higher Education Institutions, Education Bureau of Henan Province, China (Grant No. 17A520016), and the Outstanding Young Talent Research Fund of Zhengzhou University (Grant No. 1521337044). (Corresponding author: Yangjie Cao.)

N. Lin, F. Wang, X. Lu, Y. Cao, B. Zhang are with School of Software Engineering, Zhengzhou University, Zhengzhou 450000, China (e-mail: linnan@zzu.edu.cn; wfc117@qq.com; jennifer4529@163.com; caoyj@zzu.edu.cn; zhangbo2050@zzu.edu.cn).

Z. Liu is with at Shizuoka University, Japan (e-mail: Liu@ieee.org). S. Sigg is with Aalto University, Espoo, Finland (e-mail: stephan.sigg@aalto.fi).

One approach is contact-based method, e.g. Polysomnography (PSG) [5], Actigraph [6], [7], Pressure sensors [8]–[10]. These approaches normally result in good performance, but with the trade off of employing numerous sensors. These sensors are normally either highly sensitive and precisely positioned, or to be worn by the monitoring target. A better approach is device-free method. In comparison, device-free method is often more economical and convenient, such as computer vision based sleep monitoring [11], [12]. However, this method may invade people's privacy and is often susceptible to illumination; Some researchers utilize Radio Frequency (RF) technology like the Doppler radar [13], which rely on specialized hardware device, thus limiting its deployment. In contrast, WiFi, another RF technology, overcomes the above disadvantages, by utilizing the existing ubiquitous in-house WiFi signal, freeing the personnel from any on-body equipment, as well as avoiding the effect of illumination and unnecessary personal privacy invasion.

Inspired by the advantages of WiFi signals, Contactless Body Movement Recognition (CBMR) method is proposed in this paper to explore the relationship between body movement and WiFi channel state information (CSI) data collected from commercial off-the-shelf WiFi devices. The deep learning model adopted by CBMR is designed to handle the complex CSI time series and automatically extract features of body movement. CBMR bears several advantages. First of all, it is relieved from laborious and delicate data pre-processing such as denoising and manual feature extraction. Furthermore, CBMR achieves zero burden recognition by requiring no on-body equipment.

The contributions of this paper are summarized as follows:

1. CBMR is proposed to extract deep features of raw CSI data obtained from WiFi signals, which is further used for body movement recognition during sleep.
2. A sleep body movement dataset is built for sleep research. Different actions (turning over, raising a leg, sitting up, etc.) of multiple testee are collected using WiFi devices.
3. To evaluate the performance of CBMR, we conducted extensive experiments concerning parameter tuning, recognition accuracy comparison with state-of-the-arts, model complexity analysis, furniture setup variation adaptability, etc. Experiment results demonstrate that CBMR achieves desirable result and has an average accuracy of up to 93.5%.

The remaining of this paper is organized as follows. Section II provides related work of sleep monitoring technologies and an introduction about WiFi sensing. We will introduce CBMR architecture and analyze CSI characteristics in Section III,

followed by experimental results, evaluation and presents some existing problems in Section IV. We finally conclude the work in Section V.

II. RELATED WORK

In this section, we first summarize sleep monitoring technologies in general, then introduce WiFi sensing in detail.

A. Sleep Monitoring Technologies

Major technologies for sleep monitoring can be categorized into contact-device based approaches and device-free approaches.

1) *Contact-device based approach*: PSG includes electro-myogram, electro-cardiogram, electro-encephalogram, and electro-oculogram; it is comprehensive and diverse, providing an accurate assessment of sleep architecture and sleep quality [5]. PSG requires a large number of sensors accurately attached to the corresponding parts of the body. Actigraph (e.g., smart watch, smart bracelet) [6], [7] can continuously monitor sleep quality and is convenient to carry and take down. Although actigraph is convenient comparing to PSG, it may cause sensor reading failure duo to unconscious movements of the subject. Sometimes, people may forget to wear or even lose it. Pressure sensors based sleep monitoring technology are in comparison less invasive [8]–[10]. The pressure sensor is placed under a mattress or a customized bed and can detect fine-grained indicators of sleep conditions such as lying postures, breathing and heart rate according to pressure changes. Unfortunately, this technology requires the sensors to be highly sensitive and precise positioned in order to produce satisfactory measurements.

2) *Device-free approach*: Vision-based sleep monitoring technology is a typical contactless sleep monitoring approach [11], [12]. Researchers used an infrared camera to simultaneously analyze respiration, head posture, and body posture during sleep [11]. *Philips vital signs camera* focused on the rise and fall of chest and abdomen to calculate the breathing rate [12]. Nevertheless, the drawbacks of these technologies are obvious, e.g., high expense, susceptible to light and may invade people's privacy. At present, despite the fact that there are some studies using Radio Frequency (RF) technologies like the Doppler radar [13], the frequency modulated carrier waves to detect vital signs [14], these technologies rely on specialized devices, limiting their deployment. In contrast, WiFi, another RF technology, overcome the above disadvantages. WiFi based monitoring technique deduces the state of a person by combining the refraction and reflection of surrounding wireless signals [15], [16]. It is wearable-device free, immune to illumination conditions, and can protect personal privacy. Table I lists the advantages and disadvantages of the above technologies. It is obvious that WiFi based approach surpasses other technologies in almost every aspect, which represents a new trend of sleep quality evaluation method in the future.

B. WiFi Sensing

1) *Applications of WiFi sensing*: WiFi sensing technology has been rapidly developing in recent years, and can be applied

in many fields. For gesture recognition, WiGest [17] used CSI to sense in-air hand gestures around the user's mobile device. HuAc system [18] recognize human activity using CSI. WiFi-ID [19] is among the first body of work that tries to identify people using WiFi signals. WiFall [20] detects a fall activity by combining a one-class support vector machine classifier with a random forest algorithm. EmoSense [21] is a first-of-its-kind WiFi-based emotion sensing system. For sleeping posture changes, Gu et al. proposed *Sleepy* that can recognize both stationary and active states during sleep [22].

2) *Methods of WiFi sensing*: Many researchers have applied CSI to recognize and monitor vital signs [23], [24]. In [23], respiration was monitored using passive WiFi sensing via the Convolutional Neural Network (CNN). A sleep prediction model of body movement via CNN was established by Aarti et al., who also attempted to estimate human sleep patterns using a variant of RNN called Long Short-Term Memory (LSTM), with a moderate 84% accuracy achieved [24].

Hybrid networks are increasingly favored by researchers from a variety of areas. Different hybrid networks are designed for image captioning [25], visual question answering [26], time series classification [27], etc. In [27], LSTM Fully Convolutional Networks (LSTM-FCN) have been proposed for classification of time series in general. Since CSI data belongs to time series, we applied LSTM-FCN to CSI-based sleep movement classification in this paper to evaluate its performance. Following the parameter setup as in [27] and use the CSI data we collected as the training data, LSTM-FCN obtained an accuracy of 82.8%. The possible explanations for this unsatisfactory result lies in that LSTM-FCN is not specifically designed and optimized for our problem. An appropriate hybrid network design for CSI-based sleep movement classification problem is non-trivial, and is left as our future work. We therefore will not discuss hybrid network in the remainder of this paper.

III. CBMR ARCHITECTURE

In this section, an overview of CBMR is first presented, followed by detailed description of major modules.

A. Overview of CBMR

The proposed CBMR characterizes body movement data into levels of movements. CBMR mainly consists of three parts as shown in Fig. 1. In the sensing phase, WiFi signal source and sink are placed at opposite sides of a bed. The sink captures wireless signals along the line-of-sight path and those reflected by static objects, as well as signals distorted by body movements. CSIs of the captured signals are segmented to form short-term CSI data before being fed into the deep learning model. In the data extraction and model building phase, Bi-RNN [28] and IndRNN [29] are utilized with residual mechanism for extracting features of body movement data. Finally, a softmax classifier is used to classify the types of body movement as CBMR's output. Six common kinds of body movements are studied during sleep, including turning over, curving legs, raising an arm, raising a leg, lying down, and sitting up. The components of CBMR are introduced in detail as follows.

TABLE I: Comparison between various sleep technologies.

Technologies	Privacy preservation	illumination independence	Non-wearable	Non-contact	Low cost
PSG [5]	✓	✓			
Actigraph [6]	✓	✓			✓
Pressure Sensor based [8]–[10]	✓	✓	✓		✓
Vision-based [11], [12]			✓	✓	✓
Radar-based [13], [14]	✓	✓	✓	✓	
WiFi-based [15], [16]	✓	✓	✓	✓	✓

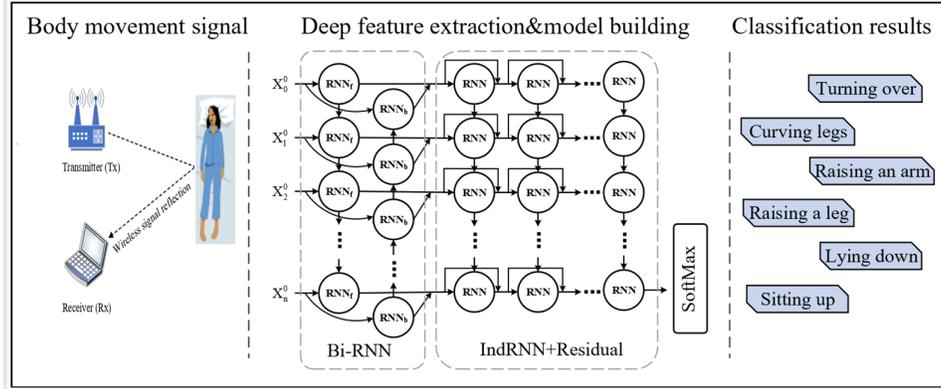


Fig. 1: Overview of CBMR.

B. Data Collection and Representation

Commercial off-the-shelf WiFi transmitter (Tx) and receiver (Rx) are deployed for WiFi signal collection. CSI data is extracted from data packets using software tools developed by Halperin et al. [30]. Orthogonal frequency division multiplexing (OFDM) technology is applied to obtain channel response information of multiple subcarriers and to decompose CSI data. The CSI data received is a $N_t \times N_r \times n$ matrix, where N_t and N_r represent the number of transmitter antenna and receiver antenna respectively. The third dimension n stands for subcarrier index in OFDM. A CSI data is sliced into $m = N_t \times N_r$ short term CSI segments, each containing a piece of data from all n subcarriers. CSI segments can be represented as follows:

$$\begin{aligned} \mathbf{CSI}^1 &= \{CSI^{1,1}, CSI^{1,2}, \dots, CSI^{1,n}\} \\ \mathbf{CSI}^2 &= \{CSI^{2,1}, CSI^{2,2}, \dots, CSI^{2,n}\} \\ &\dots\dots\dots \\ \mathbf{CSI}^m &= \{CSI^{m,1}, CSI^{m,2}, \dots, CSI^{m,n}\}, \end{aligned}$$

where $CSI^{m,n}$ represents the n -th subcarrier of the m -th data stream. The data streams shown in Fig. 2 refer to the sequence of amplitude data collected by the receiver over a period of time. The CSI of a single subcarrier is expressed as $CSI^{m,n} = H(f_n) = ||H(f_n)||e^{j\angle H(f_n)}$, $||H(f_k)||$ and $\angle H(f_k)$ indicate its amplitude and phase, respectively. Let $\mathbf{x}^n = \{CSI^{1,n}, CSI^{2,n}, \dots, CSI^{m,n}\}$, all subcarriers at time t can be expressed as $\mathbf{x}_t = \{\mathbf{x}_t^1, \mathbf{x}_t^2, \dots, \mathbf{x}_t^m\}$.

In this paper, the sliding window is used to segment long-term CSI sequence data and to extract jitter of data stream for analyzing short-term body movement. The sliding step length determines overlap percentage between adjacent sliding

windows, which augments CSI data volume and as a result, enhance movement feature extraction. We take all CSI data of a complete body movement as a group of CSI time series denoted by S . We select T as the size of sliding window and set d ($d < T$) as the sliding step length. CSI time series is then divided into $(S - T)/d + 1$ overlapping short time windows. Meanwhile, our method extracts the amplitude of CSI data and uses CSI segments of identical window as a set of input for the neural network.

C. Network Architecture

The continuous and unique motion trajectory of body movement during sleep is composed of the action characteristics of each sampling point in a period of time, and the sampling points are both semantically inseparable and temporally close to each other. Taking the temporality and complexity of body movement data into account, we have proposed a model of deep bi-directional independently recurrent neural network with residual mechanism which is shown in Fig. 3. It consists of the following two parts. The key characteristics and functions of each part are as follows.

The first part is Bi-RNN with ReLU. The regular RNN provides a very elegant approach for solving the problem of time series effectively in virtue of recurrent mechanism that can predict the current state based on the previous data information. Nevertheless, the prediction performance of RNN decreases along with the increase of sequence length, and some important one-directional information cannot be captured. The emergence of bidirectional RNN addresses this issue very well by linking the previous information with the future via current output. Such feature is of great help to learn

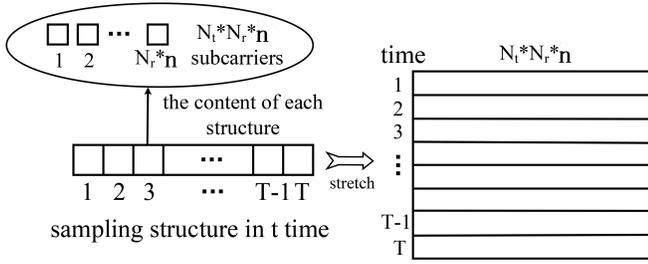


Fig. 2: CSI data stream structure.

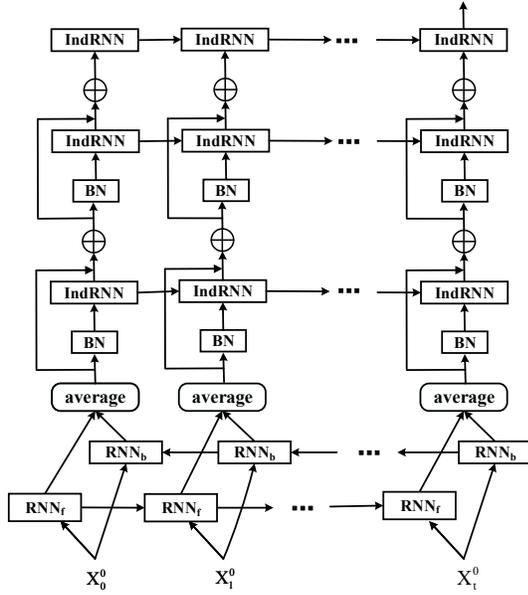


Fig. 3: Proposed model network structure.

the context information of CSI data of body movements. Bi-RNN is made up of two RNN cells, which together determine the output.

In Fig. 3, there is a forward and backward sequence in the hidden layer of Bi-RNN. For the current moment, the forward and backward hidden layer $\vec{\mathbf{h}}_t$, $\overleftarrow{\mathbf{h}}_t$ and the output layer \mathbf{x}_t^1 can be defined as follows:

$$\vec{\mathbf{h}}_t = g(\mathbf{W}_{\vec{h}} \mathbf{x}_t^0 + \mathbf{U}_{\vec{h}} \vec{\mathbf{h}}_{t-1} + \mathbf{b}_{\vec{h}}), \quad (1)$$

$$\overleftarrow{\mathbf{h}}_t = g(\mathbf{W}_{\overleftarrow{h}} \mathbf{x}_t^0 + \mathbf{U}_{\overleftarrow{h}} \overleftarrow{\mathbf{h}}_{t-1} + \mathbf{b}_{\overleftarrow{h}}), \quad (2)$$

$$\mathbf{x}_t^1 = g(\mathbf{V}_{\vec{h}} \vec{\mathbf{h}}_t + \mathbf{V}_{\overleftarrow{h}} \overleftarrow{\mathbf{h}}_t + \mathbf{b}_{\vec{h}\overleftarrow{h}}), \quad (3)$$

where \mathbf{h}_{t-1} denotes the output of the hidden layer in the previous moment, \mathbf{W}_h , \mathbf{U}_h , \mathbf{V}_h and \mathbf{b}_h present the input weight, recurrent weight, output weight and bias, respectively. We use non-saturated ReLU as the activation function of Bi-RNN. ReLU function can reduce the computation of neural network and solve the problem of gradient vanishing, and it will make the output of some neurons 0, which will cause the sparseness of network, and reduce the interdependence of parameters, alleviating the occurrence of overfitting problems. Moreover,

we average the two results for maintaining consistent input and output dimensions, which are calculated as

$$\vec{\mathbf{h}}_t = \text{ReLU}(\mathbf{W}_{\vec{h}} \mathbf{x}_t^0 + \mathbf{U}_{\vec{h}} \vec{\mathbf{h}}_{t-1} + \mathbf{b}_{\vec{h}}), \quad (4)$$

$$\overleftarrow{\mathbf{h}}_t = \text{ReLU}(\mathbf{W}_{\overleftarrow{h}} \mathbf{x}_t^0 + \mathbf{U}_{\overleftarrow{h}} \overleftarrow{\mathbf{h}}_{t-1} + \mathbf{b}_{\overleftarrow{h}}), \quad (5)$$

$$\mathbf{x}_t^1 = \text{Add}(\vec{\mathbf{h}}_t, \overleftarrow{\mathbf{h}}_t)/2. \quad (6)$$

The second part is IndRNN with Residual Mechanism. The IndRNN with ReLU activation function whose network architecture, shown in Fig. 4 (a), was proposed by Li et al.. It differs from standard RNN, shown as Fig. 4 (b), in that its recurrent weight \mathbf{u} is represented by a vector instead of a matrix. Each neuron in a layer is independent of another neuron, and each neuron accepts only the current input and its own hidden state from the previous time. The current moment can be described as

$$\mathbf{h}_t^l = \text{ReLU}(\mathbf{W} \mathbf{x}_t^{l-1} + \mathbf{u} \odot \mathbf{h}_{t-1}^{l-1} + \mathbf{b}), \quad (7)$$

where \mathbf{x}_t^{l-1} denotes the output of previous layer; l ($l > 2$) is the layer index of IndRNN; the recurrent weight \mathbf{u} is a vector; \odot represents Hadamard product; \mathbf{W} , \mathbf{u} are shared parameters at different times. The standard RNN can be regarded as a multi-layer perceptron that shares parameters over time, while the IndRNN can be thought of as an independently aggregating space (i.e. through \mathbf{W}) over time (i.e. through \mathbf{u}), utilizing two or more layers to exploit the correlation between different neurons.

In addition, we add residual mechanism based on IndRNN to build a deeper network architecture. The residual mechanism transmits the lower information to the upper layer directly through a highway, merging the underlying features into subsequent operations. The highway architecture with skip connections can skip many layers in height. In this way, the complex problem can be simplified as two addition problems, which also helps counter this problem of gradient vanishing. One layer is skipped here as shown in Fig. 3, and the output of hidden layer can be defined as

$$\mathbf{h}_t^l = \text{ReLU}(\mathbf{W} \mathbf{x}_t^{l-1} + \mathbf{u} \odot \mathbf{h}_{t-1}^{l-1} + \mathbf{b}) + \mathbf{h}_t^{l-1}. \quad (8)$$

Batch Normalization (BN) is adopted in the proposed model. It can normalize the data, which keeps the input of each layer of neural networks in the same distribution and accelerates the convergence and training speed [31].

Combining the above two parts, a deep neural network is constructed. The proposed model reduces the number of parameters and solves the problem of gradient vanishing to some extent. It is good for learning the deep semantics and extracting effective features of CSI data generated by WiFi signals affected by body movement.

D. Classification

Assume that body movements are divided into $y \in (1, 2, \dots, k)$ categories, CBMR aims to create a deep learning model that can be used to accurately predict label y based on the input CSI time series \mathbf{X} of which the size of sliding

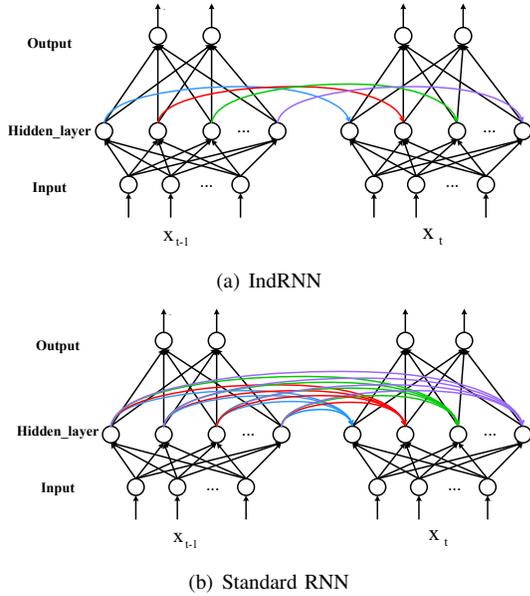


Fig. 4: The architectures of IndRNN and standard RNN.

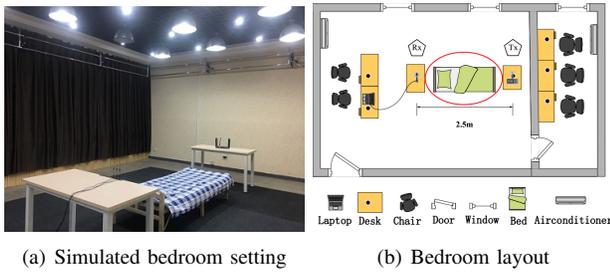


Fig. 5: Bedroom layout in the experiment.

windows is T . The activation function of softmax is used for classification; it calculates the probability of each output \tilde{y} as

$$\tilde{y}_i = P(y|\mathbf{X}) = \frac{\exp(z_y)}{\sum_{y=1}^k \exp(z_y)}, y \in [1, k], \quad (9)$$

where z_y denotes the final output of IndRNN with residual mechanism.

The overall model can be trained by minimizing the cross-entropy loss function using Adam optimizer [32] between the probabilistic outputs and the real labels, which can be defined as

$$J_{y_i}(\tilde{y}_i) = - \sum y_i \log(\tilde{y}_i). \quad (10)$$

IV. EXPERIMENT AND EVALUATION

In this section, we experimentally confirm the effectiveness of the proposed neural network model using real human body movement data. Firstly, the dataset we collected is described in Section IV-A. Then, the experimental setup is introduced in Section IV-B, followed by the evaluation metrics in Section IV-C. Finally, the experimental results are presented and discussed in detail in Section IV-D.

TABLE II: Details of our deep neural network.

Setting items	Detail
Learning rate	0.0002
Training iters	3000
Batch size	64
subcarrier_num	90
sliding_window_size	800
n_hidden	200
n_classes	6
Dropout rate	0.5

A. Dataset Description

We simulate the bedroom scene in a laboratory. The authentic environment and room layout is shown in Fig. 5. The distance between the Wireless Router and the laptop is about 2.5m and they are placed 1m above the ground. Between these two devices, a bed is placed, with 1m×2m in size and 0.5m high.

Eight volunteers (five male and three female; age: 19-26; height: 1.58-1.85m; weight: 51-80kg) are monitored in the above setting to collect CSI data of their body movements. They lie on the bed and try their best to simulate the movements during sleep to ensure the authenticity of the experimental data. Each volunteer repeats six body movements for 80 times in 3 hours, without interruption. Finally, 3840 records are collected, and each of them lasts 3 seconds. All raw CSI data is remixed through the sliding window processing, and the amount of data is increased by $(S - T)/d + 1$ times. Moreover, the body movement data is evaluated with the following ten-fold cross-validation method. The remixed CSI data of body movements is divided into ten parts, of which nine parts are taken as training data and the rest one is used for testing in turn, without data cross. The experimental results return average value of ten results.

B. Experimental Setup

CBMR consists of two WiFi devices: one is the TP_LINK AC1750 wireless router as the Tx, and the other is the Think-Pad X201 laptop equipped with an Intel 5300 802.11n WiFi NIC as the Rx. The Tx has one detectable antenna, and the Rx has three external antennas that helps to solve the problem of data loss caused by the internal current of the laptop and hence to improve the quality of CSI data captured. With these antennas, a pair of Tx and Rx hence forms a 1×3 CSI data streams. Each CSI data stream consists of 30 subcarriers, distributed evenly in the 56 subcarriers of a 20MHz channel [33], so that 90 (1×3×30) CSI data streams can be collected for each time instance. All the experiments conducted in this research are performed in 5GHz frequency band for its ability to obtain better body movement resolution. The sample rate of CSI data is set to 1000 packets/s.

Unless otherwise mentioned, we set the size T of sliding window to 800 and choose $d = 200$ as the sliding step length to construct sample data, with 75% overlap percentage, which increases the amount of CSI data and improves the generalization capacity of the model. The deep learning model described here is implemented by Tensorflow in python, and

its training and classification are run on a GPU with 2880 cores, 875MHz clock speed and 24GB RAM. The detailed parameter setup in our neural network are shown in Table II.

C. Evaluation Metrics

We use the following four metrics to evaluate the performance of CBMR. The first is the most commonly used accuracy, and others are Precision, Recall and F1. Take turn over as a True or Positive (TP) example and not turn over (or one of other body movements) as a False or Negative (FN) example. TP refers to the number of samples that turn over is correctly judged as turn over; FN refers to the number of samples that turn over is judged as not turn over; FP refers to the number of sample that not turn over is incorrectly judged as turn over; and TN refers to the number of samples that not turn over is correctly judged as not turn over.

Precision is defined as the ratio of the number of correctly predicted as turn over to the number of predicted as both turn over and not turn over, which is computed as

$$Precision = \frac{TP}{TP + FP}. \quad (11)$$

Recall refers to the proportion that turn over is truly predicted when the real label is turn over, which is computed as

$$Recall = \frac{TP}{TP + FN}. \quad (12)$$

To avoid extreme situations in which the precision or recall is 1 and the other one is 0, the harmonic average of precision and recall, F1, is used to evaluate the performance of CBMR, which is computed as

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}. \quad (13)$$

The above formulas can solve a two-class problem, but a multi-class classification is required in this paper. Since the number of each body movement in our dataset is relatively balanced, the average F1 value can be defined as

$$F1 = \frac{2}{k} \frac{N_k}{N_{total}} \sum_k \frac{Precision_k \times Recall_k}{Precision_k + Recall_k}, \quad (14)$$

where k is class index of body movement, N_k is the number of samples of k -th class, and N_{total} is the total number of dataset. $Precision_k$ and $Recall_k$ are the Precision and Recall of body movement of k -th class, respectively.

D. Experimental Results

To evaluate whether the angle (the bed against LOS between Transmitter and Receiver) of bed placement has an effect on recognition for CBMR, we collect six different types of body movement data under different angles (i.e. 0, 30, 60, 90) shown in Fig. 6. The red arrow refers to the angle that the bed meets the two devices. Fig. 7 presents the average performance of classification under different angles. As can be seen, classification at 0 degree achieves poor results with the recall of 93.5% and the precision of 93.7%, but the classification performance under different angles differs by less

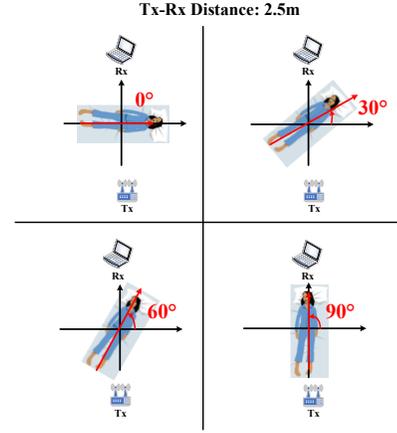


Fig. 6: Data collection of different angles.

than 1% between 93.5% and 94.3%. The results demonstrate that different angles of the bed have minimal effect on the experimental results. Therefore, all the experiments are done at 90 degree.

In addition, to verify the ability of CBMR to adapt to the environment. We did the following experiments. After training CBMR with the data of which the bed is set at 60 degree, we directly test the trained model with the 90 degree data, with the resulting accuracy of only 67.5%. Subsequently, we using merely 50 data sequences of 90 degree to fine-tune the model trained for 60 degree setup, then we again use the 90 degree of data to test the fine-tuned model and the accuracy rises to 92.8%. This shows that CBMR has excellent adaptability in scenarios with minor environmental changes.

In CBMR, the most important parameter is the number of hidden layers. In order to determine the number of the most appropriate hidden layers, the effect of different numbers of hidden layers are examined and compared with baseline LSTM under the dataset we collected. The baseline LSTM is well-known because it improves the overall performance of vanilla RNN and adds the idea of gate architecture for solving the gradient vanishing problem. The results of our observation are shown in line chart Fig. 8, the deeper the network layer is, the higher the classification performance of CBMR and the baseline LSTM can be achieved. When the layers exceed a certain number, the accuracy of the model begins to drop. When the layer number is small, the model is not sensitive enough to capture signal changes. Conversely, overfit of the training data leads to a decrease in test set accuracy, but even when the average accuracy drops, CBMR decreases in a slower speed than the baseline LSTM. The advantages of Bi-RNN, residual mechanism and BN mentioned in Section III-C improve the performance of CBMR in feature extraction of CSI data, so CBMR achieves better classification performance than baseline LSTM. The research result shows that CBMR can effectively solve the problem of gradient vanishing and achieve good performance.

Furthermore, the accuracy of the training data and validation data for CBMR and baseline LSTM is presented in Fig. 9. As the number of iterations increases, the train and validation

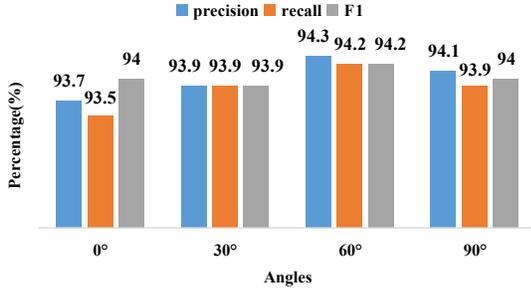


Fig. 7: The classification performance of CBMR under different angles.

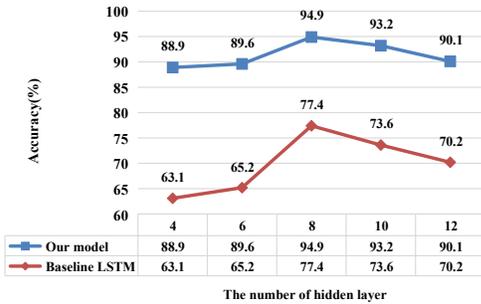


Fig. 8: Comparison of different hidden layers.

results converges. Under the same model parameters, CBMR converges faster than LSTM and is more stable.

The sliding window plays an important role in our experiment and helps remove redundant data labeled as Null. Within a sliding window, if more than 40% of the sampled data is labeled Null (i.e. data without movement), all the data in this sliding window will eventually be tagged Null and discarded. As the sliding window slides, the data labeled as one of six body movements is constantly selected for training and testing. However, the size of sliding window has a great impact on the experimental results. The smaller the sliding window is, the more data after data segmentation, resulting in the increase of training time and the waste of memory.

When the size of sliding window is large, the division of data into without movement and with movement is not accurate enough. The results of classification are compared under three different sizes of sliding window (See Table III).

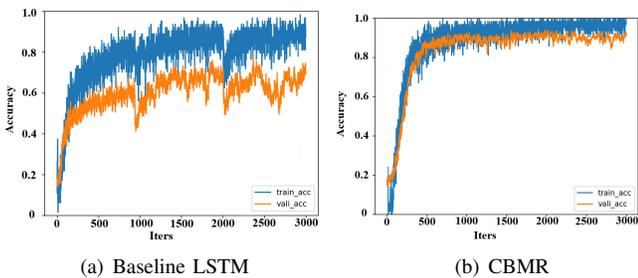


Fig. 9: Accuracy trends of models. The blue line shows train data, and the orange line indicates test data.

TABLE III: Classification results under different sliding windows.

the size of sliding window	Precision	Recall	F1
1000	78.7%	76.26%	87.5%
800	89.1%	90.5%	89.8%
600	86%	86.9%	86.1%

TABLE IV: Model accuracy of different sliding step length and overlap percentage.

sliding step length	100	200	300	400	600	800
overlap percentage	87.5%	75%	62.5%	50%	25%	0%
Accuracy	92.98%	93.34%	92.29%	89.32%	90.7%	90.5%

The values of average precision, average recall and average F1 of all body movements are the highest when $T = 800$ from our experiment. which, therefore, is chosen as the size of sliding window in this paper.

In addition, we did another set of experiments to select the appropriate sliding step length. We selected six sets of data with the sliding step length of 100, 200, 300, 400, 600, and 800 as the sliding step length increases, the number of sequences into which each CSI sequence is divided and the overlap percentage will decrease. As shown in the experimental results in Table IV, we chosen $d = 200$ as the most suitable sliding step length in this paper. Using limited raw data, by verifying different overlap percentage parameters and combining with the optimal sliding window, we find the best combination of sliding window and sliding step length that makes the model have the best generalization capacity.

The confusion matrix of all body movements (TO, CL, RA, RL, LD, and SU represent Turning over, Curving legs, Raising an arm, Raising a leg, Lying down, Sitting up, respectively) using CBMR with testing data is shown in Table V. The row represents the real body movements given in the label, and each column represents body movements tested by CBMR. The values of precision in six types of body movement ranges

TABLE VI: Recognition results of each body movement before and after data filtering.

body movement	data	Precision	Recall	F1
TO	original	96%	91%	93.4%
	filtered	85%	91.7%	88%
CL	original	95.3%	91.7%	93.5%
	filtered	96.5%	90.5%	93.4%
RA	original	95.8%	94.1%	94.9%
	filtered	91.2%	96.8%	93.9%
RL	original	89%	95.8%	87.8%
	filtered	92.9%	97.9%	95.3%
LD	original	94.8%	95.5%	95.1%
	filtered	93%	89.4%	91.2%
SU	original	95.6%	97.9%	96.7%
	filtered	97.9%	96.7%	97.3%

TABLE V: Matrix confusion using CBMR.

	TO	CL	RA	RL	LD	SU	Recall
TO	262	4	0	6	8	8	91%
CL	2	264	1	20	1	0	91.7%
RA	3	1	271	8	5	0	94.1%
RL	2	6	3	276	0	1	95.8%
LD	1	2	6	0	275	4	95.5%
SU	3	0	2	0	1	282	97.9%
Precision	96%	95.3%	95.8%	89%	94.8%	95.6%	94.4%

TABLE VII: The performance comparison between different models.

Models	Turning over	Curving legs	Raising an arm	Raising a leg	Lying down	Sitting up
Baseline LSTM [34]	91.7%	62.5%	66.7%	91.7%	70.8%	81.3%
Bi-RNN [28]	65.6%	68.8%	52.1%	89.6%	52.1%	91.7%
IndRNN [29]	95.8%	60.4%	97.9%	41.7%	97.9%	96.5%
CBMR	88.5%	95.8%	93.8%	96.9%	80.2%	97.9%

TABLE VIII: Number of parameters and FLOPs between different models (in millions).

Models	Baseline LSTM	Bi-RNN	IndRNN	CBMR
Number of parameters	2.48	1.24	0.3	0.44
FLOPs	7.42	3.71	0.91	1.38

from 89% to 96%, and the values of recall are between 91% and 97.9%. The results show that turning over is the most difficult to be recognized, because its range of movement is the least in all body movements. It is worth noting that the motion trajectory of curving legs is almost the same as that of raising a leg, so it is difficult to distinguish these two types of body movement. Fortunately, CBMR, as shown in Table V, has obtained considerable recognition results on these two types of body movement.

To verify CBMR can accurately recognize body movement without processing CSI data, we use cauer filter to process the CSI data and remove some signal noise to make amplitude fluctuations smoother. Fig. 10 shows the filtering effect of six body movements which have different effects on WiFi signals. The amplitude changes of all body movements are clearer after being filtered. The experimental results before and after filtering do not show much difference, as shown in Table VI. The three types of body movements of curving legs, raising an arm and sitting up have similar degree of recognition before and after filtering. The recognition results of raising a leg in the case of filtering is better than not filtering. Unexpectedly, the recognition of turning over and lying down after data filtering is not good. The reason may be that CSI data information of these two body movements is removed by the filter, resulting in the loss of important features. No matter whether the experiments are conducted with raw data or filtered data, our deep learning model has achieved desirable results. To sum up, CBMR helps eliminate the trouble of data pre-processing, reduce the experimental operation process and operation time.

We compare the experimental results across different models. Table VII presents the performance of CBMR and the existing neural network models such as baseline LSTM, Bi-

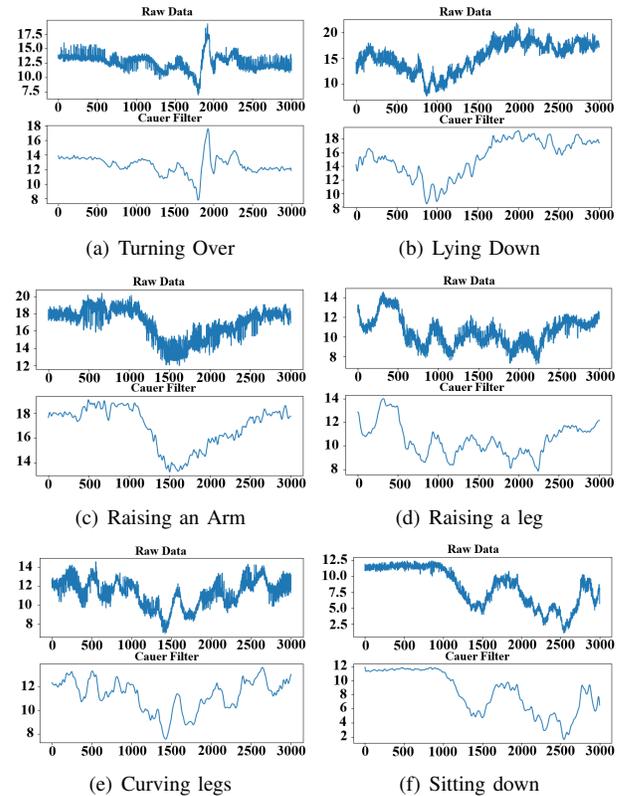


Fig. 10: Amplitude information of all body movements before and after data filtering.

RNN and IndRNN. Within all the models, sitting up is the most accurately recognized, especially in CBMR, with 97.9%. Surprisingly, raising an arm only obtains an accuracy of 41.7% in IndRNN, yet higher in other models. It is obvious that CBMR outperforms the others, and achieves better recognition accuracy for each body movement. The main reason is that CBMR combines the advantages of the other models, and revises the details such as activation function and BN to make it more suitable for handling time series data of tiny body movement. In a different experiment in which the testee is with a cover on his/her body, CBMR still achieves an average accuracy of 87.5%, which verifies the robustness of CBMR as well.

Otherwise, we take the number of parameters and FLOPs as indicators of the computational complexity of the model. In Table VIII we report the number of parameters and FLOPs for each model. From the Table VIII we can see that CBMR also has obvious advantages in complexity. The internal mechanism of IndRNN determines that the number of parameters and FLOPs are reduced and the accuracy does not decrease [29]. CBMR integrates the IndRNN network, thus reducing the number of parameters and FLOPs.

In the field of WiFi sensing research in a single-person environment [35], [36], there are always some dynamic factors like presence of multiple human or mobility from surrounding objects. These factors will indeed affect the research results, removing these irrelevant dynamic factors is a hard problem widely acknowledged in WiFi sensing research. To investigate the effect of these factors on CBMR, we did the following two sets of experiments. In the first group of experiments, when collecting the data of the subjects, an irrelevant person walks back and forth outside the red circle (far from the line of sight) as shown in Fig. 5 (b), making some slight movements. Retraining CBMR using this set of data, the results of the experiment is 92.4%, which is 1.1% different from the accuracy of single-person environment. In the second group of experiments, when collecting the data of the subjects, an irrelevant person does irregular actions slightly in the red circle (near the line of sight). In this case, the accuracy of the model is very poor, only 62%. The above two experiments show that dynamic factors appearing near the line of sight have a greater impact on the model results. In our proposed application scenario, however, it's rarely the case that presence of multiple human or motion from surrounding objects is involved. Instead, the condition of the PoI, who almost always lives alone, is our major concern, e.g. Solitary elderly. In our future work, we plan to solve the hard problem.

V. CONCLUSION

In this paper, CBMR, a Contactless Body Movement Recognition model, has been proposed. It determines the types of tiny body movement during sleep using CSI data collected from a commercial off-the-shelf WiFi router and a laptop, without complex equipment deployment. CBMR greatly reduces time of manually extracting features. Moreover, we conducted extensive experiments to evaluate CBMR, the experiment results indicate that the model gets desirable performance and achieves average accuracy of greater than 93.5%.

REFERENCES

- [1] Z. Zimmer, C. Jagger, C. T. Chiu, M. B. Ofstedal, F. Rojo, and Y. Saito, "Spirituality, religiosity, aging and health in global perspective: A review," *SSM - Population Health*, vol. 2, pp. 373–381, 2016.
- [2] C. J. Bathgate and J. Fernandez-Mendoza, "Insomnia, short sleep duration, and high blood pressure: Recent evidence and future directions for the prevention and management of hypertension," *Current Hypertension Reports*, vol. 20, no. 6, p. 52, 2018.
- [3] M. J. Murphy and M. J. Peterson, "Sleep disturbances in depression," *Sleep Med Clin*, vol. 10, no. 1, pp. 17–23, 2015.
- [4] T. J. Song, C. H. Yun, S. J. Cho, W. J. Kim, K. I. Yang, and M. K. Chu, "Short sleep duration and poor sleep quality among migraineurs: A population-based study," *Cephalalgia*, vol. 38, no. 5, p. 333102417716936, 2018.
- [5] C. Armon, A. Roy, and W. Nowack, "Polysomnography: overview and clinical application," <http://emedicine.medscape.com/article/1188764-overview>.
- [6] L. Granovsky, G. Shalev, N. Yacovzada, Y. Frank, and S. Fine, "Actigraphy-based sleep/wake pattern detection using convolutional neural networks," 2018.
- [7] J. A. Vitale, G. Banfi, A. Galbiati, L. Ferini-Strambi, and A. Torre, "Effect of night-game on actigraphy-based sleep quality and perceived recovery in top-level volleyball athletes," *International Journal of Sports Physiology and Performance*, vol. 14, no. 2, pp. 1–14, 2018.
- [8] H.-M. C. T.-H. Kim, S.-J. Kwon and Y.-S. Hong, "Determination of lying posture through recognition of multitier body parts," *Wireless Communications and Mobile Computing*, vol. 2019, 2019.
- [9] C. Ma, W. Li, R. Gravina, J. Cao, Q. Li, and G. Fortino, "Activity level assessment using a smart cushion for people with a sedentary lifestyle," *Sensors*, vol. 17, no. 10, p. 2269, 2017.
- [10] L. Rosales, B. Su, M. Skubic, and K. C. Ho, "Heart rate monitoring using hydraulic bed sensor ballistocardiogram1," *Journal of Ambient Intelligence and Smart Environments*.
- [11] D. Fei, J. Dong, X. Wang, F. Ying, and C. Feng, "Design and implementation of a noncontact sleep monitoring system using infrared cameras and motion sensor," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 7, pp. 1–9, 2018.
- [12] "Philips vital signs camera [online]," <http://www.vitalsignscamera.com/>.
- [13] C. Li, J. Ling, J. Li, and J. Lin, "Accurate doppler radar noncontact vital sign detection using the relax algorithm," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 3, pp. 687–695, 2010.
- [14] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, 2015, pp. 837–846.
- [15] R. Ravichandran, E. Saba, K.-Y. Chen, M. Goel, S. Gupta, and S. N. Patel, "Wibreathe: Estimating respiration rate using wireless signals in natural settings in the home," in *Pervasive Computing and Communications (PerCom)*, 2015 IEEE International Conference on. IEEE, 2015, pp. 131–139.
- [16] X. Liu, J. Cao, S. Tang, J. Wen, and P. Guo, "Contactless respiration monitoring via off-the-shelf wifi devices," *IEEE Transactions on Mobile Computing*, vol. 15, no. 10, pp. 2466–2479, 2016.
- [17] H. Abdelnasser, M. Youssef, and K. A. Harras, "Wigest: A ubiquitous wifi-based gesture recognition system," in *IEEE Conference on Computer Communications*, 2015, pp. 1472–1480.
- [18] L. Guo, L. Wang, J. Liu, W. Zhou, and B. Lu, "HuaC: Human activity recognition using crowdsourced wifi signals and skeleton data," *Wireless Communications & Mobile Computing*, vol. 2018, no. 5, pp. 1–15, 2018.
- [19] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-id: Human identification using wifi signal," in *International Conference on Distributed Computing in Sensor Systems*, 2016, pp. 75–82.
- [20] C. Han, K. Wu, Y. Wang, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2017.
- [21] Y. Gu, T. Liu, J. Li, F. Ren, Z. Liu, X. Wang, and P. Li, "Emosense: Data-driven emotion sensing via off-the-shelf wifi devices," in *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, pp. 1–6.
- [22] Y. Gu, J. Zhan, Z. Liu, J. Li, Y. Ji, and X. Wang, "Sleepy: Adaptive sleep monitoring from afar with commodity wifi infrastructures," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2018, pp. 1–5.
- [23] U. M. Khan, Z. Kabir, S. A. Hassan, and S. H. Ahmed, "A deep learning framework using passive wifi sensing for respiration monitoring," in

- GLOBECOM 2017-2017 IEEE Global Communications Conference. IEEE, 2017, pp. 1–6.
- [24] A. Sathyanarayana, S. Joty, L. Fernandez-Luque, F. Offi, J. Srivastava, A. Elmagarmid, S. Taheri, and T. Arora, “Impact of physical activity on sleep: a deep learning based exploration,” arXiv preprint arXiv:1607.07034, 2016.
 - [25] H. Chen, H. Zhang, P. Y. Chen, J. Yi, and C. J. Hsieh, “Show-and-fool: Crafting adversarial examples for neural image captioning,” 2017.
 - [26] J. Andreas, M. Rohrbach, T. Darrell, and K. Dan, “Deep compositional question answering with neural module networks,” Computer Science, vol. 27, pp. 55–56, 2015.
 - [27] F. Karim, S. Majumdar, H. Darabi, and S. Chen, “Lstm fully convolutional networks for time series classification,” IEEE Access, vol. 6, no. 99, pp. 1662–1669, 2018.
 - [28] M. Schuster and K. K. Paliwal, “Bidirectional recurrent neural networks,” IEEE Transactions on Signal Processing, vol. 45, no. 11, pp. 2673–2681, 1997.
 - [29] L. Shuai, W. Li, C. Cook, C. Zhu, and Y. Gao, “Independently recurrent neural network (indrn): Building a longer and deeper rnn,” 2018.
 - [30] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, “Tool release: Gathering 802.11n traces with channel state information,” Acm Sigcomm Computer Communication Review, vol. 41, no. 1, pp. 53–53, 2011.
 - [31] T. Cooijmans, N. Ballas, C. Laurent, Ç. Gülçehre, and A. Courville, “Recurrent batch normalization,” arXiv preprint arXiv:1603.09025, 2016.
 - [32] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” Computer Science, 2014.
 - [33] Y. Xiao, “Ieee 802.11n: enhancements for higher throughput in wireless lans,” IEEE Wireless Communications, vol. 12, no. 6, pp. 82–91, 2005.
 - [34] 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.
 - [35] F. Wang, S. Panev, Z. Dai, J. Han, and D. Huang, “Can wifi estimate person pose?” 2019.
 - [36] D. Huang, R. Nandakumar, and S. Gollakota, “Feasibility and limits of wi-fi imaging,” in Acm Conference on Embedded Network Sensor Systems, 2014.