

Fine-grained occupant activity monitoring with Wi-Fi channel state information: Practical implementation of multiple receiver settings



Hoonyong Lee^a, Changbum R. Ahn^{b,*}, Nakjung Choi^c

^a Department of Architecture, College of Architecture, Texas A&M University, TX 77843-3137, United States

^b Department of Construction Science, College of Architecture, Texas A&M University, TX 77843, United States

^c End-to-End Networked System Department, Nokia Bell Labs, NJ 07974-0636, United States

ARTICLE INFO

Keywords:
Occupancy sensing
Non-intrusive monitoring
Wifi signals
Multiple receivers

ABSTRACT

Human activity recognition is essential for various smart-home applications. With the development of sensing technology, various approaches have been proposed for occupancy monitoring indoors. However, such approaches have practical limitations that they require additional occupancy sensors, which may raise privacy issues and obtrude on occupants' daily lives. In this research, a Wi-Fi-based occupancy monitoring system, Wi-Sensing, is proposed to recognize occupant's activities of daily living in a non-intrusive way by exploiting commercial off-the-shelf Wi-Fi devices. Channel State Information (CSI) has been extracted from Wi-Fi signals collected from multiple Wi-Fi devices, which could be replaced by Internet of Things (IoT) devices. While multiple receivers are needed to cover the entirety of an indoor space, previous approaches have been proposed to extract numerous features from a single transmitter-receiver pair. In this context, this study presents a new approach toward extracting spatial-temporal features from multiple receivers deployed throughout an indoor space. In this approach, a Short-Time Fourier Transform (STFT) was used to convert time-series CSI data into image data. The converted image data from each receiver was then integrated as large image data, which preserved the temporal-spatial information of all the receiver data. A Convolutional Neural Network (CNN) was used as a feature extractor for the image data, and Long Short-Term Memory (LSTM) was exploited to classify basic activities in daily life (e.g., personal hygiene, eating, mobility, etc.). Wi-Sensing provides over 96% classification accuracy in two different indoor environments.

1. Introduction

As Wi-Fi networks are now ubiquitous in most buildings and homes, researchers have harnessed Wi-Fi signals as a new sensing source for non-intrusive occupant monitoring systems [1]. Such monitoring systems exploit Wi-Fi variance caused by the occupant's movements to classify activities of daily living. A Wi-Fi transmitter (i.e., an Access Point (AP)) sends Wi-Fi signals, which reach a Wi-Fi receiver via various routes (either by penetrating obstacles or by being reflected by the obstacles), which is referred to as multipath propagation. A human body is also an obstacle reflecting the signals. However, unlike static objects (e.g., walls, furniture, or appliances), occupants are not stationary and can move, which results in signal variations at the receiver over time. Wi-Fi-based occupant monitoring systems exploit these variations in the received Wi-Fi signals to classify occupant's activities of daily living.

In order to represent the signal variations, a Received Signal

Strength Indicator (RSSI) was used. The RSSI superimposes the multi-path signals as a single amplitude. By exploiting the RSSI, coarse-grained activity detection and meter-scaled indoor localization is possible. However, by superimposing the received signals, the RSSI underestimates the effect of physical environmental changes on the signal propagation. Thus, the variation of RSSI becomes insensitive to fine-grained activity detection and centimeter-scaled indoor localization [2]. Recently, Channel State Information (CSI) has been used for occupant monitoring systems, since the CSI is more stable and robust than RSSI for occupant monitoring because of its utilization of Orthogonal Frequency Division Multiplexing (OFDM) [3]. OFDM divides a signal bandwidth into multiple orthogonal subcarriers, and then the CSI is obtained from each subcarrier of the received signals. Thus, the CSI indicates physical environmental changes, such as changes in an occupant's location and motion, at the subcarrier level [2]. By exploiting the CSI, many studies have been performed in the fields of occupant indoor localization/tracking, identification, counting, gesture

* Corresponding author.

E-mail addresses: onarcher@tamu.edu (H. Lee), ryanahn@tamu.edu (C.R. Ahn), nakjung.choi@nokia-bell-labs.com (N. Choi).

recognition, breathing/respiration/heart rate estimation, and activity classification [1]. Such monitoring systems can be utilized for smart-home healthcare systems, which detect either fall accidents or abnormal pattern of daily living for the elderly [4]. Also, they can effectively manage building energy consumption by operating Heating Ventilation and Air Conditioning (HVAC) systems based on the occupants' activities [5,6].

For CSI-based occupant monitoring systems, conventional machine learning algorithms and deep learning algorithms have been used [1]. Since the performance of such learning-based algorithms relies highly on features extracted from the CSI, previous studies [3,4,7–13] have focused on extracting appropriate features in both the time and frequency domains through extensive experiments; many of these experiments were conducted with a transceiver (a transmitter-receiver pair) in a single space (e.g., a laboratory, an office, or a room in an apartment) [3,7,8,11,12,14–19]. However, in any practical application, multiple receivers would need to be deployed in several different locations for several reasons. Firstly, since Wi-Fi signals lose energy as they travel indoors, a transceiver has a limited coverage area for occupant monitoring [20]. In order to detect the presence of an occupant using Wi-Fi signal variation, the reflected signal from the occupant's body needs to reach the receiver with enough magnitude of variation. If the occupant stays in a different room from the one where the receiver is placed, or if the distance between the occupant and the receiver increases, the signal variation caused by the occupant's activity becomes too diminished to accurately distinguish the variation from noise [4]. Secondly, multiple receivers are required to detect the exact location of an occupant [21]. This information about an occupant's location is critical in differentiating activities that involve similar body movements but are performed in different locations. For example, washing hands in the bath and washing dishes in the kitchen may generate similar patterns of signal variations at a receiver. When the information about the occupant location is obtained from the multiple receivers, those two different activities become distinguishable.

Exploiting data from multiple receivers has thus been established as critical in ensuring the practical implementation of CSI-based occupant monitoring system; however, the question of how to best leverage data from multiple receivers has not been adequately investigated. For example, RT-Fall [4], a falls detection system, selected the receiver containing the highest signal variations among all the receivers and used data from only that receiver. E-eyes [22], an activity recognition system, attempted to identify occupant activities using data from multiple receivers, but it constructed the signature distribution of each activity by putting CSI amplitude data from all the receivers into one histogram, which undermined the usefulness of the spatial and temporal features extracted from each receiver data. Depending on the relative location between an occupant and multiple receivers, each receiver exhibits different patterns of CSI amplitude that change over time as the occupant moves to different locations. Because the histogram is constructed by superimposing the distinct CSI patterns of each receiver, the spatiotemporal aspect of the occupant's activity would be undermined.

In this context, this research proposes, Wi-Sensing, an approach to exploit multiple receivers to accurately classify the occupant's activities. Data from multiple receivers was converted into image data using a Short-Time Fourier Transform (STFT) and merged into large image data, which contains both the spatial-temporal aspect and the unique pattern of activity. A hybrid Convolutional Neural Network (CNN)-Long Short-Term Memory (LSTM) was used as the feature extractor and activity classifier.

2. Background

2.1. Conventional occupant monitoring systems

Occupants have been monitored either for building energy

management or for smart-home health care by various occupancy sensor systems [4,6]. These conventional occupant monitoring systems are mainly able to detect occupant presence, count the number of occupants, and track occupant locations via video/image [23,24], motion [25,26], CO₂ [27,28], temperature and humidity [29], and sound [30,31] data collected from occupancy sensors. However, such monitoring systems require additional devices to collect this information, such as cameras, microphones, wearable devices, or CO₂ sensors to only provide coarse-grained activity detection [32]. Also, such approaches are intrusive to occupants' daily lives and raise privacy issues [6]. Therefore, such approaches have faced inherent limitations for practical implementation.

As wireless network systems have been built into commercial buildings and homes, the Wi-Fi signal has been exploited as a new sensing source for collecting occupant information without additional devices or costs, because Wi-Fi-based monitoring systems exploit commercial Wi-Fi devices used in most homes and buildings (e.g., Wi-Fi routers, smartphones, computers, and IoT devices) [33–37]. Moreover, the inability of such monitoring systems to generate an image of, and sound from, an occupant alleviates privacy concerns [38,39]. The received Signal Strength Indicator (RSSI) of the Wi-Fi signal received from Wi-Fi receivers was used for occupant localization and activity detection. Since the intensity of the RSSI is related to the distance between the Access Point (AP) and the Wi-Fi receiver (such as mobile devices), the location of the receiver can be estimated based on the RSSI. These approaches expect that the location of occupants is identical to the location of their mobile devices, and rely on the assumption that an occupant always carries their mobile device. The intensity of the RSSI is also affected by the occupant's activity as the occupant's body reflects the Wi-Fi signals. Therefore, different activities generate the distinct intensities of the RSSI. However, the RSSI presents the received multiple signals as a single amplitude and can fluctuate significantly as a result of even a slight change in the indoor environment, making it difficult to distinguish the fine-grained movements of the occupant from the noise [2]. Consequently, the RSSI-based occupancy monitoring systems could classify coarse-grained occupant activities [35].

2.2. CSI-based occupant monitoring systems

Since the Linux CSI 802.11n tool has been published, the CSI has been used for non-intrusive occupant monitoring systems [40]. Mathematically, the CSI is expressed by Eq. (1) in Multiple Input Multiple Output (MIMO)-Orthogonal Frequency Division Multiplexing (OFDM) wireless systems [8]. As the transmitted Wi-Fi signals (X_i) are affected by physical environmental factors, such as fading, scattering, and power loss, the received signals (Y_i) become changed, and the CSI (H_i) can be computed by Eq. (1). As compared to the RSSI—the superimposition of multipath signals—the CSI presents a channel response to the environmental changes at an individual subcarrier level. Consequently, the CSI is more robust and stable than the RSSI, which means that fine-grained occupant activity classification can be obtained by exploiting the CSI [1–3,41]. Although the raw CSI variation contains unique patterns representing the activity, filtering is first applied to the raw CSI before features are extracted. A low-pass Butterworth filter has been frequently used since most occupant activities occur at less than 10 Hz, but noise is generally distributed in the high frequency band [4]. Instead of the low-pass filter, Discrete Wavelet Transform (DWT) or phase offset removal approaches were also used [1].

$$Y_i = H_i X_i + N_i \quad (1)$$

where, i : data packet, $i \in [1, N]$,

N : Number of received packets,

Y_i : Received signal vector, H_i : CSI matrix,

X_i : Transmitted signal vector, N_i : Noise vector

For feature extracting, statistical features in the time-domain have been used for conventional machine learning algorithms, such as K

Nearest Neighbor (KNN), Support Vector Machine (SVM), and Random Forest (RF) [1]. Wi-Chase [8] provided 94% accuracy for classifying running, walking, and hand moving, exploiting the SVM with six statistical features extracted from the CSI in the time-domain: mean, standard deviation, 25th and 75th percentile, median absolute deviation, and maximum of the CSI. Since the features were computed as a single value regardless of the length of the activity, the features did not contain temporal information about the activity. Therefore, the features extracted from two walking activities in opposite directions are identical, and cannot be differentiated. Since the statistical features lose the temporal information, and appropriate feature selection requires extensive study, deep learning algorithms have been used as the activity classifiers [3,42–44]. Instead of the single value of the statistical features, the segmented time-series CSI is inputted to such approaches.

The LSTM is one of the most used activity classifiers because it is a powerful recurrent neural network model for sequential data classification. Previous research showed that LSTM outperformed other classifiers (e.g., Support Vector Machine or SVM and Random Forest or RF) for classifying simple activities: lying down, falling, walking, running, sitting down, and standing up [3]. In order to exploit the CNN, the time-series CSI was also converted to image data. DeepHare [44] (a hybrid CNN-LSTM model-based human activity recognition scheme) was used to convert the time-series CSI into image data which has three axes (time, CSI subcarrier index, and CSI amplitude). The converted image data was first fed into the CNN to extract numerous features, and the LSTM then classified the occupant's activities using those features. DeepHare provided 97.6% classification accuracy for the predefined set of activities, including standing, running, lying down, walking, sitting, and no activity ("empty home") using a single Wi-Fi transmitter-receiver pair. By exploiting deep learning algorithms, the extracted features preserve the temporal information. However, the spatial information is still dismissed, as the CSI was obtained from a single receiver, which cannot monitor the entire space due to its limited coverage area.

In this context, this research proposes a way to utilize multiple receivers for extracting features which contain spatial-temporal information about the occupant's activities. By exploiting multiple receivers, this approach increases the overall performance of the proposed activity classification model. The performance of other benchmark models has also increased with multiple receivers.

3. Methodology

The proposed model exploits one Wi-Fi transmitter (the Access Point or AP) and multiple receivers deployed in different fixed locations; the receivers simultaneously collect Wi-Fi signals transmitted from the AP. Fig. 1 shows the overview of the proposed model. First, the CSI data of different patterns is obtained from the multiple receivers. The time-series data is then converted into image data using a STFT, and further integrated into one large image dataset. The image data is fed into the CNN-LSTM model to extract features and classify the occupant's activities. The following section describes the three phases of the proposed model in detail: (1) data preprocessing; (2) feature extraction; (3) activity classification.

3.1. Data preprocessing

The raw CSI contains high frequency noise not related to the occupant's activities, which themselves generate less than 10 Hz frequency in the Wi-Fi signals [4]. For de-noising, a fifth-order, low-pass Butterworth filter is used with the 10 Hz cut-off frequency. Fig. 2 shows the raw and the filtered CSI collected during walking. In this case, the transmit rate of the Wi-Fi signal was set as 100 Hz. The occupant was walking from the 1,000 to 2,000 packet number, the sections of the graphs where the CSI shows a distinct perturbation. Since the frequency of the walking activity is less than 10 Hz, the filtered CSI retains the distinct pattern of walking as shown in Fig. 2-b. The 30 colored lines represent each subcarrier, which have different amplitudes but similar patterns to each other.

As shown in Fig. 2, the occupant activity generates a unique pattern of variation in the CSI. However, such patterns differ depending on the relative positions of the occupant and the receiver. Fig. 3 shows the multipath propagation of Wi-Fi signal indoors. Wi-Fi signals reach each receiver through different propagation paths, which results in distinct CSI variations at the receiver. Fig. 4 clearly shows the different CSI variations at each receiver during walking. The occupant walked from R1 to R2 as shown in Fig. 3; the grey panels in Fig. 4 indicate when the occupant performed the walking activity. The distinct perturbation appears earliest at R1, and then at R2 as the occupant became closer to R2. The overall amplitude of the CSI is also different before and after the walking activity occurred. Moreover, the CSI for R3 does not contain any perturbation deriving from the walking activity, since R3 was placed in the bedroom and the occupant only walked in the living room. The Wi-Fi signals reflected from the occupant lost their energy before the signals reached R3. Therefore, multiple receivers are shown to provide information about not only the types of activity occurring, but also the location where the activity is performed. If only R3 was in use, walking in the living room would not be detected at all; similarly, if only R1 was in use, the direction of the walking activity in the living room would be unclear.

In order to take full advantage of the multiple-receivers setting, the proposed approach extracts all the necessary spatial-temporal features from multiple receivers. First, data sampling is applied to the sequence of CSI using a moving window whose length is 5 s with a 1 s interval (timeframes which are empirically determined). For extracting features from the sampled data, the CNN is used. Since individuals have different musculoskeletal systems and habitual movements, for the same activity the sampled data may be slightly different for each occupant. The CNN is a powerful tool to identify common features within the different sampled data. The CNN leverages spatial correlation of the input data by using a convolutional layer, which identifies important features regardless of its location within the given input data. To accomplish this, the time-series CSI must be converted into image data. The STFT whose result preserves the variations of the CSI in both the time and the frequency domains is used for converting the CSI into image data. Since the spectrogram expresses such information in three axes (x-axis: time, y-axis: frequency, and z-axis: energy), it can be used as image data. Fig. 5 shows the spectrogram of the sample data for the first subcarrier collected from each receiver. The spectrogram has 9 and 51 data points in the time and the frequency axes, respectively. Each spectrogram contains information about the types of activity and the

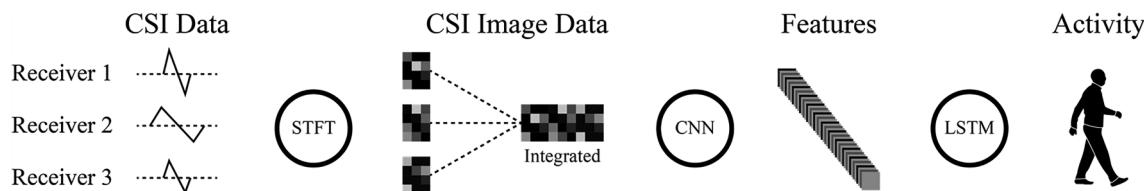


Fig. 1. Overview of the proposed hybrid CNN-LSTM model.

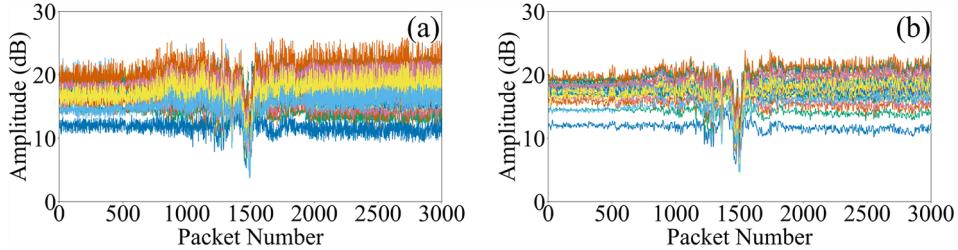


Fig. 2. CSI denoising: (a) the raw CSI and (b) the filtered CSI.

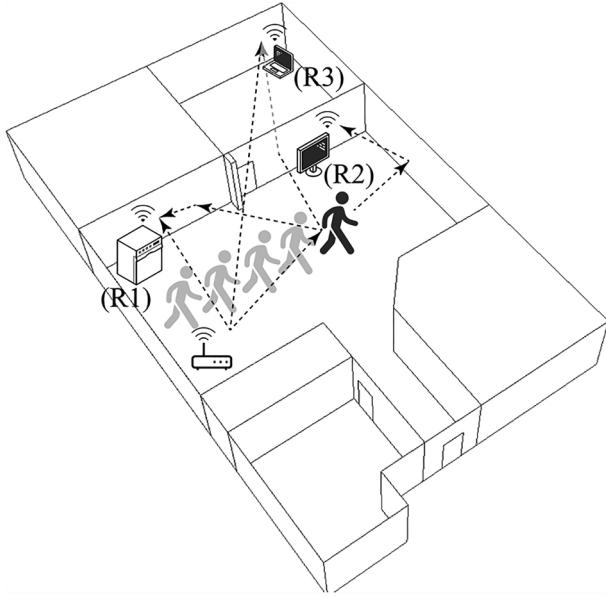


Fig. 3. Multipath propagation of Wi-Fi signals indoors.

relative locations of the occupant and each receiver. All the image data is then attached side-by-side to generate integrated large image data, which preserves spatial information about the occupant's activity (see Fig. 5-d).

3.2. Feature extraction

The integrated image data is then fed into the CNN to extract 1,024 features which contain temporal-spatial aspects of the occupant's activity. First, the input shape is set as $30 \times 51 \times 27$, which represent the number of subcarriers, data points in the frequency-domain, and the time-domain. Fig. 6 represents the shape of the input image data. In order to extract the 1,024 features, the architecture of the CNN is reconstructed from VGGNet [45]. Fig. 7 shows the CNN architecture, which consists of 16 layers. All the convolutional layers have a 3×3 kernel size with zero padding, which is appropriate to the size of the input data. In order to address the vanishing gradient problem, Rectified Linear Units (ReLU) are used as the activation function, which returns 0 for negative values. A batch normalization layer is also used to normalize the input data, and the max-pooling layers are used to avoid over-fitting issues. The stride size of the max-pooling layer is identical to the pool size, meaning that all the elements are pooled once. L2 regularization is also used to reduce over-fitting by adding a 1×10^{-5} penalty for the weight of the loss function. Since the pooled features by the CNN are not appropriate to directly input into the LSTM, a flatten layer is used to convert the shape of the features into a single column.

3.3. Activity classification

The extracted features from the STFT by the CNN contain both the

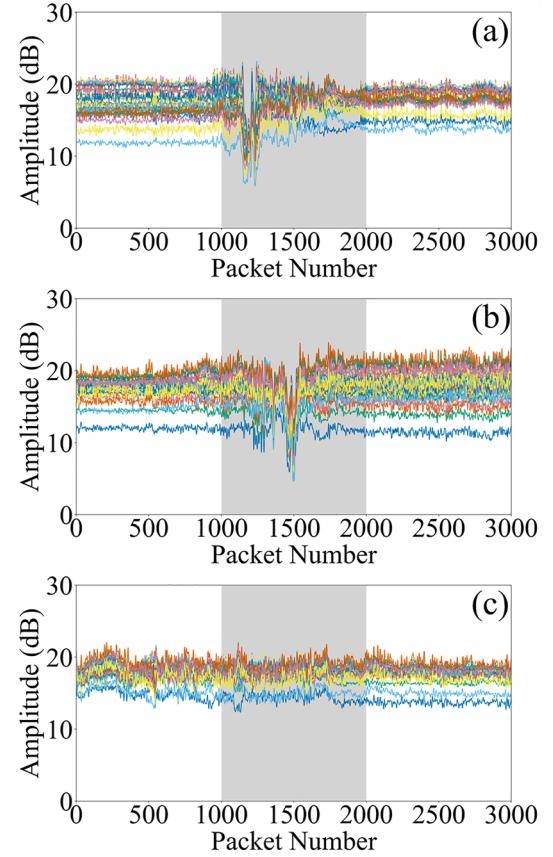


Fig. 4. Comparison of CSI variation collected from different receivers: (a) R1, (b) R2, and (c) R3.

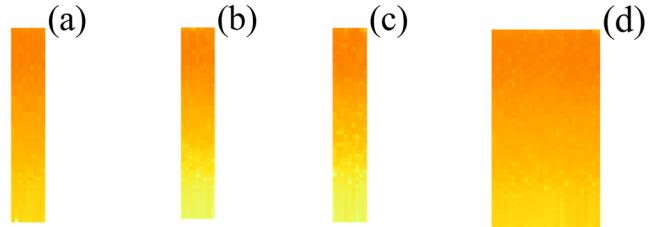


Fig. 5. Spectrogram of sample data: (a), (b), (c), individual receiver data, and (d) integrated receiver data.

spatial and the temporal aspects of the occupant's activity. For the activity classification, one LSTM layer with 512 units and six timesteps (many-to-one LSTM) is used. A batch normalization layer and a fully connected layer are also used to yield the last 8 units, the total number of the occupant's activities for classification. The Softmax activation function is used, which provides each probability that the test activity data matches all the activity; the sum of these probabilities is 1. For compiling the proposed CNN-LSTM model, the optimization algorithm

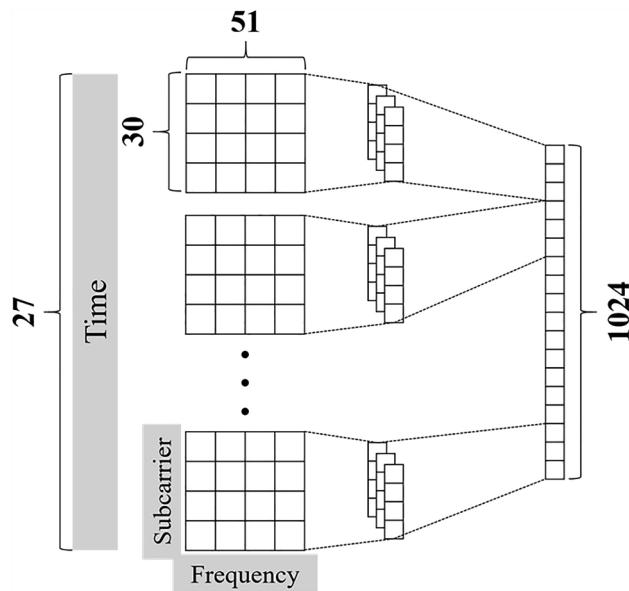


Fig. 6. Shape of the input image data.

is set to Adaptive Moment Estimation (Adam) with 1×10^{-3} learning rate.

4. Evaluation

Although the Wi-Fi signal propagation is mainly affected by the occupant's movements, other indoor environmental factors, such as surrounding Wi-Fi signals, housing layout, or building material also have an impact on the propagation, which results in performance degradation of the CSI-based activity classification models [46]. Therefore, the performance of the proposed model was evaluated in two different housing environments. In the following sections, the experimental design is first described in detail. Next, the model performance is presented and compared to other benchmark models.

4.1. Experiment and data collection

One Lenovo T400 laptops with Intel 5300NIC was used as a Wi-Fi transmitter and three additional Lenovo were used as Wi-Fi receivers. The CSI was simultaneously collected from three different receivers via the Linux CSI 802.11n tool [40]. This occupant monitoring system was implemented identically in two different indoor environments, Unit A and Unit B, with the transceivers deployed at the fixed locations shown

in Fig. 8. Unit A was a wood-frame apartment (632ft²), and Unit B was a reinforced concrete-frame apartment (396ft²). Six subjects (three females and three males) were recruited and asked to perform specified basic activities in the designated locations; the subjects were all in their early 30 s, ranging from 158–182 cm in height and 45–80 kg in weight. On the basis of previous research, basic activities that healthy individuals can perform without assistance were selected, as presented in Table 1 [47]. Since the size and layout of the housing environments differed between Unit A and Unit B, all the activities were not performed identically. For example, Unit B did not include a dining room, so the subject performed the “eating” activity in the living room. Additionally, the subjects in Unit A walked across various locations, but the subjects in Unit B walked only between the bedroom and the living room due to the limited space. The experiment was performed by one subject at a time, meaning that the received signal was affected by a single subject. Also, each subject performed the experiment on different days. Thus, the collected CSI was independent of housing environments, subjects, and time.

4.2. Performance of the proposed model

The total number of data samples was 3,600 for Unit A and 4,320 for Unit B. A randomly-selected 70% of the total data samples were used for the model training and validation at a ratio of 9:1, and the remaining 30% of the sample were used for testing. The number of epochs was set to 500 with 10 batch size for training. The proposed model shows 96.57% and 98.92% accuracy for the basic activity classifications for Unit A and Unit B, respectively. As compared to the existing CSI-based activity classification models [3,4,8,44], the developed approach classifies more complicated activities at higher rate of accuracy (“basic” in our context referring to the fundamental nature of such activities as opposed to their complexity). The previous models mainly classified “simple” activities such as sitting, standing, walking, running, lying down, falling, and hand movements [3,4,8,44]. Those activities are generally performed in one simple movement, but the “basic” activities that were targeted in this study are performed using a combination of various movements. For example, the “cooking” activity consists of chopping, frying, and washing dishes, among other movements. Since the distinct variation in the CSI caused by the simple activities is clearer than that of the basic activities, classifying complicated activities is more challenging than simple activities.

The proposed model also offers methodological improvement. Many of the existing studies [3,7,8,11,12,14–16,18,19] exploited a single transceiver to monitor the occupant. However, a Wi-Fi transmitter–receiver pair has a limited coverage area. Therefore, it may not be possible for one receiver to detect the occupant's activities throughout the entire indoor space, which could degrade the overall performance of

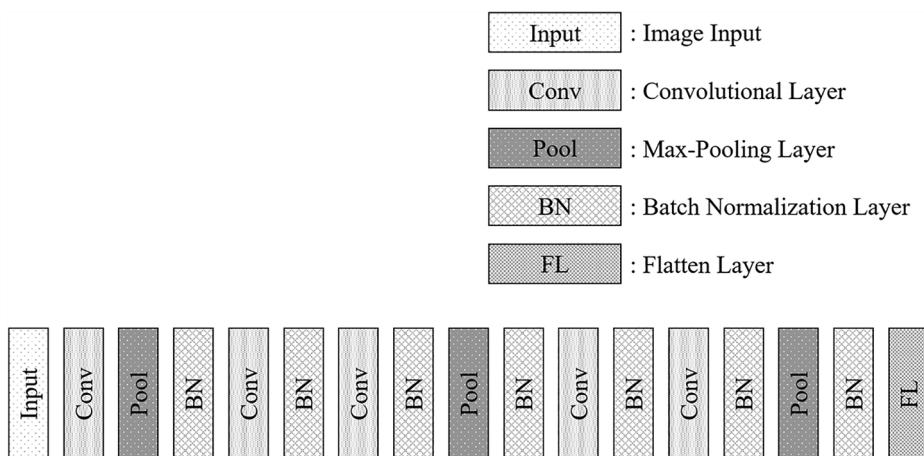


Fig. 7. The CNN architecture.

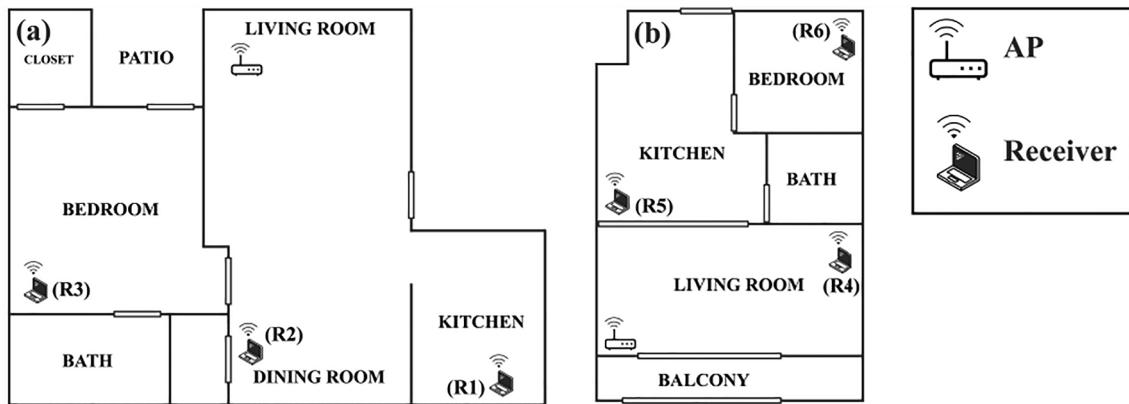


Fig. 8. Experimental test beds: (a) Unit A and (b) Unit B.

Table 1
Basic activities.

No activity (empty home)	Bathing in the bathroom
Walking in various directions	Using the toilet in the bathroom
Cooking in the kitchen	Sleeping in the bedroom
Eating in the dining room for Unit A	Falling in various locations
Eating in the living room for Unit B	

Table 2
Performance comparison between transmitter–multiple receiver pairs and a transmitter–receiver pair.

Testbed	Receiver	Accuracy	Testbed	Receiver	Accuracy
Unit A	Multiple	96.57%	Unit B	Multiple	98.92%
	R1	84.26%		R4	78.16%
	R2	73.15%		R5	81.56%
	R3	71.57%		R6	85.19%

activity classification model. Table 2 compares the model performance of multiple receivers with each transmitter–receiver pair. The accuracy of each individual receiver is lower than the accuracy when the multiple receivers were used. The performance for Unit B was also higher than the performance for Unit A. Since Unit B has a smaller footprint than Unit A, the area not covered by a single transmitter–receiver pair in Unit B was smaller than Unit A. Therefore, the individual receivers accumulated more information about the occupant activity in Unit B than A, which increased the overall performance of the monitoring systems in that unit.

Fig. 9 shows the confusion matrix of the results for Unit A. R1, which was placed in the kitchen, shows higher accuracy for the cooking and eating activities, which were performed near R1. R3, placed in the bedroom, detects the sleeping activity performed in the bedroom better than other activities performed outside of the bedroom. For the walking activity, however, as the subject walked in various directions during the experiment, all the receivers retrieved the Wi-Fi signals reflected by the subject. Therefore, all of the individual receivers show high accuracy for walking classification. This implies that due to its limited coverage area, a receiver is most accurate in classifying activities performed nearby rather than activities performed in other locations.

In order to monitor the occupant throughout the entire space, this research utilizes a setup with multiple receivers and one AP. Previous studies showed that using multiple receivers increases the performance of CSI-based occupant monitoring systems [4,22]. However, the model performance is dependent on precisely how these multiple receivers are exploited. When multiple sensors are used, there are three data fusion methodologies: data-level, feature-level, and decision-level integration [48]. Fig. 10 shows three different ways to use the multiple receivers. Fig. 10-a illustrates the proposed way to combine multiple receivers’

data at the data level (data-level integration), which creates one image by connecting spectrogram images from each receiver. The features will then be extracted from such an integrated data space. On the other hand, Fig. 10-b illustrates the second way to combine multiple receivers’ data at the feature level (feature-level integration). The features are extracted from each spectrogram image of the receivers, and the classifiers leverage those features. The number of features in this case will be triple that of the proposed data-level way, but spatial information for an occupant’s activity may not be well-captured by each single feature. Fig. 10-c illustrates the decision-level integration, which extracts features and classifies activity at each receiver. The final activity is then decided based on each classified activity.

Table 3 compares the performance of the data-level, feature-level, and decision-level integration models. First, the data-level integration model demonstrated the highest level of accuracy, while still creating lower computational costs than the feature-level and decision-level integration models. The feature-level integration model showed very low accuracy, and the decision-level integration model often demonstrated lower accuracy than compared with the outcomes from a single transceiver pair (though still higher than the feature-level model). Second, the data-level integration model generated a more consistent performance in different environments (Unit A and B). The difference in accuracy between Unit A and B was larger with the feature-level and decision-level integration models. This indicates that the data-level integration model would best contribute to reducing environmental dependency (e.g., home size, layout). The data-level integration model shows that the accuracy for each subject for Unit A is a minimum of 96.11%, a maximum of 100%, an average of 98.73%, and a standard deviation of 1.39%. For Unit B, the accuracy is a minimum of 99.53%, a maximum of 100%, an average of 99.47%, and a standard deviation of 0.79%.

4.3. Performance comparison with benchmark models

The model performance was also compared with the two benchmark models. The first benchmark model was selected to be Wi-Chase [8], which exploits the SVM algorithm, the most frequently used occupant activity classifier [3]. Wi-Chase extracts six statistical features from the CSI in the time-domain: mean, standard deviation, the 25th and 75th percentiles, median absolute deviation, and the maximum of the CSI. The second benchmark model was DeepHare [44]. DeepHare exploits the CNN-LSTM model, which is similar approach to the proposed model. In order to convert the time-series CSI into image data, DeepHare represents the CSI as three axes of time, the subcarrier index, and CSI amplitude instead of using the STFT. Also, to remove the high frequency noise and outliers, Autoencoder (AE) is used. The filtered CSI is then fed into the CNN, which has two convolutional layers, two max-pooling layers, and two fully connected layers with ReLU as the

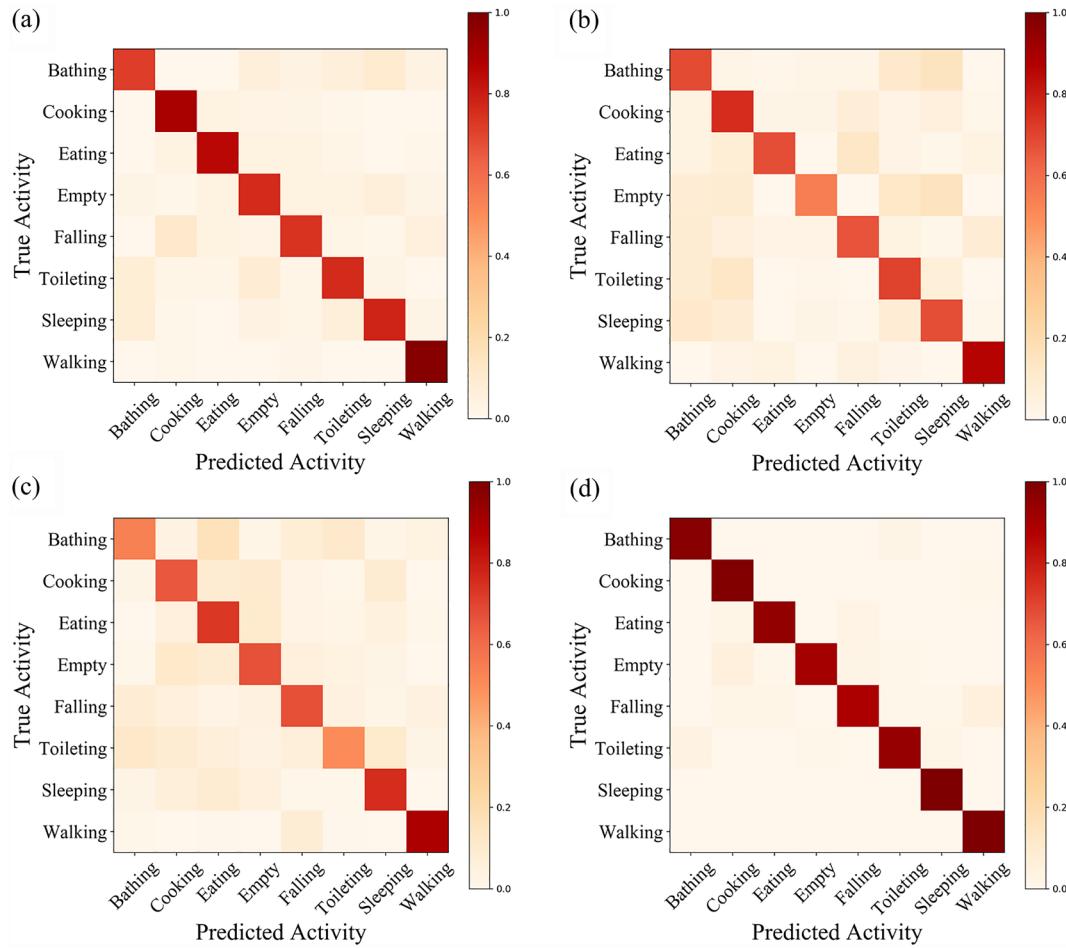


Fig. 9. Confusion matrix of the results for Unit A: (a) R1, (b), R2, (c) R3, and (d) multiple receivers.

activation function. For the activity classifier, the LSTM is used with Softmax activation function. For the performance comparison, the structure of DeepHare was fine-tuned by modifying the model architecture to be identical to the architecture of the proposed model, since the original DeepHare model shows very low accuracy for classifying CSI obtained from our experiment. There are two possible reasons for this low accuracy. First, DeepHare obtained the CSI with 114 subcarriers, but the CSI obtained from our experiment had only 30 subcarriers, which may not be appropriate for the original DeepHare framework. Second, the CNN structure of DeepHare may be too intrinsically shallow for extracting appropriate features from our experimental data. Although DeepHare shows a high degree of accuracy for classifying simple activities (e.g., sitting, standing, walking, running, and lying down), the original DeepHare structure may not be appropriate for classifying our more complex target activities. Therefore, we modified the original structure of DeepHare to match our proposed structure, which was fine-tuned by our experimental data. The difference between the proposed model and the modified DeepHare is in the types of input data, the spectrogram in both the time and the frequency-domain, and the represented CSI in the time-domain.

Table 4 shows the performance of the benchmark models for a single transmitter-receiver pair. The overall performance of the proposed model outperforms the performance of the benchmark models. The two benchmark models also show higher accuracy for Unit B than Unit A, as with the proposed model. Although the difference between the proposed model and the modified DeepHare is in the type of image data, which exploits deep learning algorithms, the performance of the modified DeepHare is not as strong as that of the Wi-Chase model in most cases. This result shows that image data converted by STFT could

preserve necessary information for accurate activity classification using the hybrid CNN-LSTM model.

Fig. 11 compares the performance between the proposed model and the benchmark models when the multiple receivers are used. The benchmark models did not exploit multiple receivers in their previous studies, and our proposed methods of integrations (i.e., data-level, feature-level, and decision-level) were attempted on the benchmark models.

First, the data-level integration model shows an improved performance for both the proposed model and the modified DeepHare compared to other integration models, and the delta of the accuracy between the larger housing environment (Unit A) and the smaller one (Unit B) becomes reduced. However, the Wi-Chase model shows an degraded performance compared to other integration models because the six statistical features extracted from the data-level integration model lose activity information rather than preserving the spatial information. Second, the feature-level integration model provides lower accuracies for the proposed model and the modified DeepHare compared to other integration models. The spatial-temporal information as well as activity information could be lost when integrating features extracted from the image data, which may degrade the model performance. Third, the decision-level integration model does not combine data or features from each receiver, which means that the features do not lose any information, but also do not exploit the spatial-temporal information. Therefore, the accuracy is similar to single-receiver data use, as summarized in Tables 2 and 4. The comparison of the model performances shows that the proposed model outperforms the benchmark models. Also, the comparison of different integration models indicates that the data-level integration we proposed performs better than

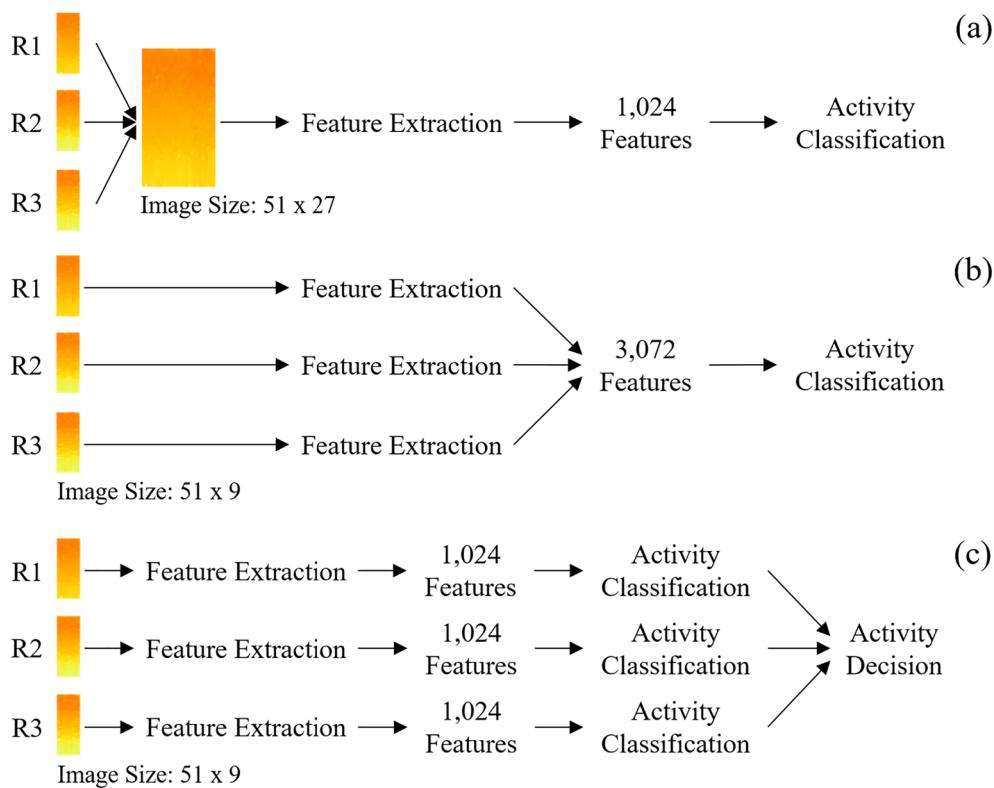


Fig. 10. Three ways for exploiting the multiple receivers: (a) data-level integration, (b) feature-level integration, and (c) decision-level integration.

Table 3

Model performance of the data-level, feature-level, and decision-level integration.

Testbed	Data-level integration	Feature-level integration	Decision-level integration
Unit A	96.57%	62.31%	82.37%
Unit B	98.92%	82.79%	88.43%

the decision- or feature-level integrations, and that decision- or feature-level integrations may not provide performances that are much better than even those obtained from a single transmitter–receiver setting.

5. Discussion

5.1. Practical application

The proposed model is used to classify fine-grained occupant activities for building occupants, and occupant activity is closely linked to the energy consumption of buildings. Fig. 12 shows the occupant's activity profiles for three different days as predicted by the proposed model; in it, "DAQ" means data acquisition process when the occupancy monitoring was stopped twice a day for 1 h at a time, and "Outlier" means that the predefined activity was not found. The occupant slept six or eight hours from 00:00, left home at 11:00, and came back home at 21:00. While staying home, the occupant performed various indoor

activities. With this occupancy information, the energy management systems can control the lighting and HVAC systems to provide a comfortable indoor environment with efficient energy use. Moreover, the proposed approach is a non-intrusive occupancy monitoring system, meaning that the occupant can be monitored for a long period of time without causing discomfort. The accumulated occupant activity information provides key clues to finding the routines of the occupant's daily life, which results in the activity prediction. The routine and prediction of the occupant's activity also enable us to analyze the patterns of the building energy use and predict future energy consumption.

5.2. Multiple receiver deployment

As shown in the Evaluation section, multiple receivers are essential for monitoring an occupant's activities performed in various locations throughout an indoor space. In any practical implementation, it is important to decide where the multiple receivers should be deployed to optimize the coverage area. The coverage area of the single transmitter–receiver pair has been represented by Fresnel zones [41]. Fresnel zones are the series of concentric ellipsoids with foci of the Wi-Fi receiver and the receiver, as shown in Fig. 13. The n^{th} Fresnel zone can be described as Equation (2).

$$|F_1 D_n| + |D_n F_2| - |F_1 F_2| = n\lambda/2 \quad (2)$$

where F_1 : Point of the AP

F_2 : Point of the receiver

D_n : A point in n^{th} ellipse

Table 4

Model performance for a single transmitter–receiver pair.

Testbed	Receiver	Wi-Chase	Modified DeepHare	Testbed	Receiver	Wi-Chase	Modified DeepHare
Unit A	R1	75.00%	63.15%	Unit B	R4	71.14%	59.72%
	R2	68.98%	51.85%		R5	73.14%	78.86%
	R3	73.70%	51.30%		R6	77.31%	80.94%

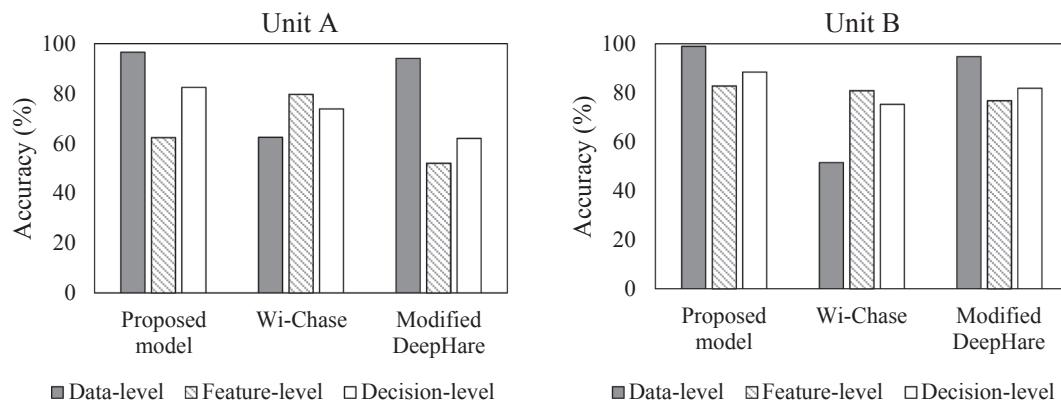


Fig. 11. Comparison of model performance for multiple receivers.

λ : Wavelength of Wi-Fi signal

Although the shape of Fresnel zone may vary depending on the indoor environments [49], we provide guidelines for multiple receiver deployment based on Fresnel zones, since an accurate guideline for multiple receiver deployment has not so far been published. If a limited number of receivers is available, it is recommended that the transmitter-receiver pair should be placed so that the Fresnel zones cover the whole indoor space. In the Evaluation section, when the single transceiver pair was used, Receiver 6 (R6) provided the highest accuracy for Unit B (see Table 2) even though R6 was placed far from the AP. The Fresnel zones generated by the AP and R6 covered almost whole footprint of Unit B, but other two receivers were closer to the AP, which resulted in smaller Fresnel zones compared to that of R6. However, if the size of the indoor space is large, the occupant might not be detected even within the Fresnel zones because the signal might lose most of its energy in traveling long distances. Since Wi-Fi signals travel via the $F_1 - D_1 - F_2$ path outlined in Fig. 13, if the area of the Fresnel zones increases, the travel length increases, which results in further energy loss. In this case, the transmitter-receiver pair should be placed to generate a Fresnel zone which covers the target area. For Unit A, Receiver 1 (R1) provided the highest accuracy compared to other two receivers. The Fresnel zones generated by the AP and R1 covered the kitchen, dining room, and living room, where almost all of the activities were performed. In contrast, the other two receivers generated Fresnel zones which covered areas where only certain activities were performed.

5.3. Locational dependency of receiver

The proposed Wi-Sensing could be operated by various IoT devices, which receive Wi-Fi signals and transfer the signal information to the main hub, which are connected to all the IoT devices. The main hub could extract CSI from the transferred Wi-Fi information collected from multiple IoT devices. The extracted CSI would be converted into image

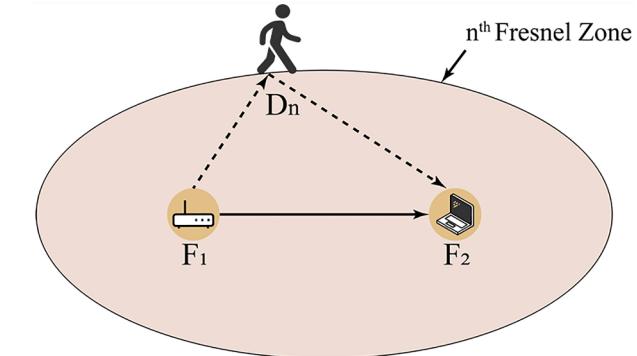


Fig. 13. Coverage area of a single transmitter-receiver pair.

data and the multiple image data would be attached side by side to generate a large image data. Since Wi-Sensing extracts features and classifies occupant's activities from the large image data, the order of each image data is important to generate the large image data. In real life, the position of IoT devices are not fixed. As the location of the device changes, the order in which the trained large image data was made may differ from the test image data. If the location of the device changes, the received signal will also change, but this research shows the effect of the order of image generation on the performance of Wi-Sensing. For Unit A, the pre-trained Wi-Sensing with the image data made in the order of R1-R2-R3 shows 31.94% and 31.85% accuracies for the test image data made in the order of R2-R3-R1 and R3-R1-R2, respectively.

By using a transfer learning algorithm, the performance can increase. From the test data, 5 samples for each activity were collected and the pre-trained Wi-Sensing was re-trained with the sample activity data. Since the Wi-Sensing was already trained to extract important

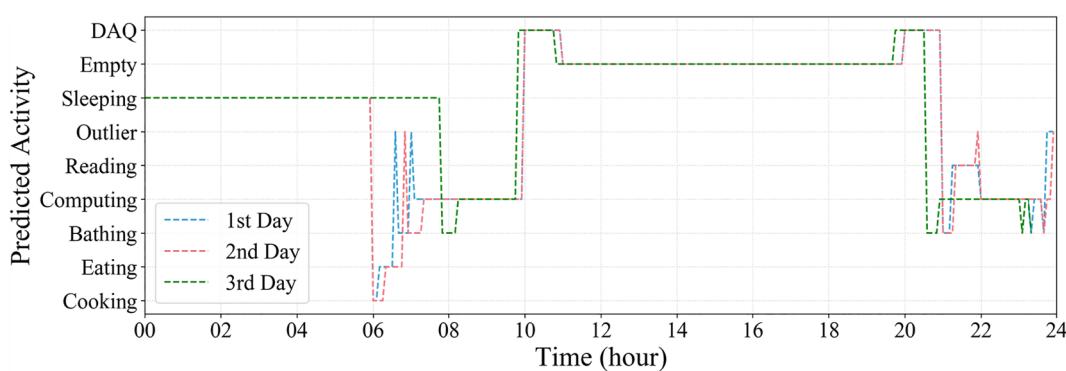


Fig. 12. Occupant's predicted activity profiles for three different days by the proposed model.

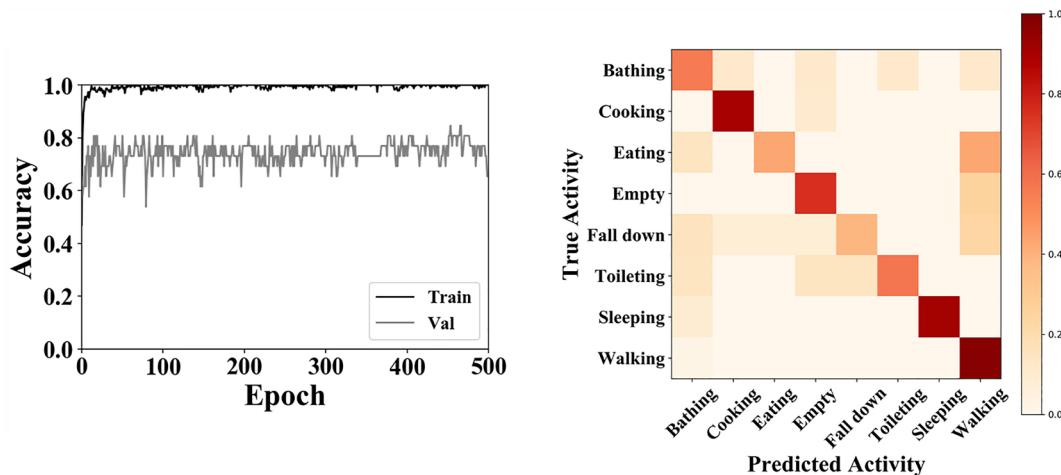


Fig. 14. Model accuracy profile and confusion matrix for the order of R2-R3-R1.

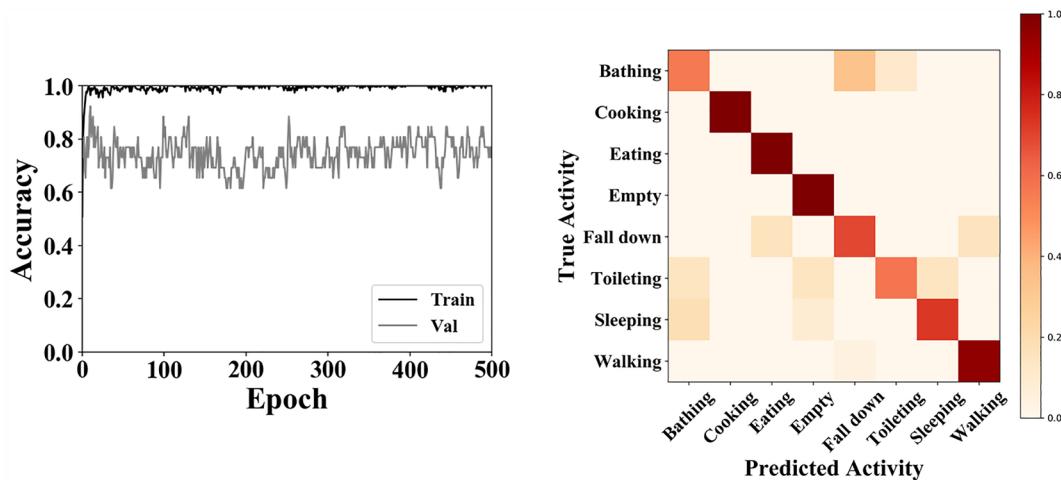


Fig. 15. Model accuracy profile and confusion matrix for the order of R3-R1-R2.

features for activity classification, the feature extraction part, the CNN, was frozen, but the activity classification part, the LSTM, was re-trained with the activity sample data. Figs. 14 and 15 show the model accuracy profile and confusion matrix by re-training the pre-trained model for the order of R2-R3-R1 and R3-R1-R2, respectively. The re-trained model with sample activity data shows 77.77% accuracy for the order of R2-R3-R1 and 85.18% accuracy for the order of R3-R1-R2. The re-trained model by using the transfer learning algorithm shows the potential that the performance of Wi-Sensing could be retained at certain level of accuracy although the locations of receivers (IoT devices) is changed over time.

5.4. Limitations and future research

Although the model proposed here shows a higher degree of accuracy in occupant activity classification compared to the benchmark approaches, the proposed model still has limitations for practical implementation since it classifies the occupant's activities based on the training data of predefined activities. First, the proposed model is only available for a single occupant. Even if multiple occupants perform only the predefined activities, the variation in the Wi-Fi signal caused by multiple occupants will be different than in the predefined activities. Therefore, the proposed model is appropriate for homes where the occupant lives alone, or for individual offices in commercial buildings. A second area of limitation is the environmental and occupant dependencies. The proposed approach classifies an occupant's activity

based on Wi-Fi signal variations as a signal travel indoors. Wi-Fi signal propagation can be affected by various housing environmental factors, such as building materials and unit layouts. Also, occupants may have different habits for performing activities as well as different skeletal-muscular systems, which may have different impacts on the Wi-Fi propagation, and which would require rich training data from the end-users in their indoor environments to overcome. However, it is not always possible to obtain a large amount of training data from the end-user, especially for the elderly. Therefore, an approach needs to be developed to reduce the environmental and occupant dependencies of the proposed model.

6. Conclusion

This research proposed an approach to exploit multiple receivers for CSI-based occupant activity monitoring systems. The temporal-spatial features of this approach increase the overall performance of such monitoring systems. A hybrid CNN-LSTM model, Wi-Sensing, was also proposed to classify the occupant activities. The performance of Wi-Sensing offers over 96% accuracy in two different indoor environments. The occupancy information—in particular fine-grained activity information—monitored by Wi-Sensing would significantly improve building automation by associating the control of HVAC, lightings, and other appliances with the occupant's activity patterns. Furthermore, monitoring the long-term patterns of activities of daily life (ADL) could be critical in providing remote healthcare services for living-alone

elderly, as the gradual decline of ADL routine patterns is a major symptom of geriatric cognitive diseases (e.g., dementia) [50].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research described in this paper was financially supported by a grant (18CTAP-C128499-02) from the Technology Advancement Research Program, funded by Ministry of Land, Infrastructure, and Transport of the Korean government.

References

- [1] Y. Ma, G. Zhou, S. Wang, WiFi sensing with channel state information: a survey, ACM Computing Survey 52 (3) (2019) 46, <https://doi.org/10.1145/3310194>.
- [2] Z. Yang, Z. Zhou, Y. Liu, From RSSI to CSI: Indoor localization via channel response, ACM Computing Survey 46 (2) (2013) 25, <https://doi.org/10.1145/2543592>.
- [3] S. Yousefi, H. Narui, S. Dayal, S. Ermon, S. Valaee, A survey on behavior recognition using WiFi channel state information, IEEE Commun. Magazine 55 (10) (2017) 98–104, <https://doi.org/10.1109/mcom.2017.1700082>.
- [4] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, S. Li, RT-fall: a real-time and contactless fall detection system with commodity WiFi devices, IEEE Trans. Mob. Comput. 16 (2) (2017) 511–526, <https://doi.org/10.1109/tmc.2016.2557795>.
- [5] P. Hoes, J.L.M. Hensen, M.G.L.C. Loomans, B. de Vries, D. Bourgeois, User behavior in whole building simulation, Energy Build. 41 (3) (2009) 295–302, <https://doi.org/10.1016/j.enbuild.2008.09.008>.
- [6] O. Ardakani, A. Bhattacharya, D. Culler, Non-intrusive occupancy monitoring for energy conservation in commercial buildings, Energy Build. 179 (15) (2018) 311–323, <https://doi.org/10.1016/j.enbuild.2018.09.033>.
- [7] H. Zhu, F. Xiao, L. Sun, R. Wang, P. Yang, R-TTWD: robust device-free through-the-wall detection of moving human with WiFi, IEEE J. Sel. Areas Commun. 35 (5) (2017) 1090–1103, <https://doi.org/10.1109/jsac.2017.2679578>.
- [8] S. Arshad, C. Feng, Y. Liu, Y. Hu, R. Yu, S. Zhou, H. Li, Wi-chase: A WiFi based human activity recognition system for sensorless environments, in: Proc. IEEE 18th International Symposium on A World Wireless, Mobile and Multimedia Networks (WoWMoM), Macao, China, 2017, pp. 1–6. doi:10.1109/wowmom.2017.7974315.
- [9] W. Wang, A.X. Liu, M. Shahzad, K. Ling, S. Lu, Device-free human activity recognition using commercial WiFi devices, IEEE J. Sel. Areas Commun. 35 (5) (2017) 1118–1131, <https://doi.org/10.1109/jsac.2017.2679658>.
- [10] Y. Zeng, P.H. Pathak, C. Xu, P. Mohapatra, Your AP knows how you move: fine-grained device motion recognition through WiFi, in: Proc. ACM Workshop on Hot Topics in Wireless (HotWireless), Maui, Hawaii, 2014, pp. 49–54. doi:10.1145/2643614.2643620.
- [11] Y. Zeng, P.H. Pathak, P. Mohapatra, WiWho: WiFi-based person identification in smart spaces, in: Proc. 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), Vienna, Austria, 2016, pp. 1–12. doi:10.1109/ipsn.2016.7460727.
- [12] W. Wang, A.X. Liu, M. Shahzad, Gait recognition using wifi signals, in: Proc. ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, 2016, pp. 363–373. doi:10.1145/2971648.2971670.
- [13] J. Zhang, B. Wei, W. Hu, S.S. Kanhere, WiFi-ID: Human identification using WiFi signal, in: Proc. International Conference on Distributed Computing Sensor Systems (DCOSS), Washington, DC, 2016, pp. 75–82. doi:10.1109/dcos.2016.30.
- [14] Q. Gao, J. Wang, X. Ma, X. Feng, H. Wang, CSI-based device-free wireless localization and activity recognition using radio image features, IEEE Trans. Veh. Technol. 66 (11) (2017) 10346–10356, <https://doi.org/10.1109/tvt.2017.2737553>.
- [15] L. Guo, L. Wang, J. Liu, W. Zhou, B. Lu, HuAc: Human activity recognition using crowdsourced WiFi signals and skeleton data, Wireless Commun. Mobile Comput. 2018 (2018) 1–15, <https://doi.org/10.1155/2018/6163475>.
- [16] W. Jiang, C. Miao, F. Ma, S. Yao, Y. Wang, Y. Yuan, H. Xue, C. Song, X. Ma, D. Koutsounikolas, W. Xu, L. Su, Towards environment independent device free human activity recognition, in: Proc. 24th Annual International Conference on Mobile Computing and Networking (MobiCom), New Delhi, India, 2018, pp. 289–304. doi:10.1145/3241539.3241548.
- [17] J. Wang, L. Zhang, Q. Gao, M. Pan, H. Wang, Device-free wireless sensing in complex scenarios using spatial structural information, IEEE Trans. Wireless Commun. 17 (4) (2018) 2432–2442, <https://doi.org/10.1109/twc.2018.2796086>.
- [18] W. Wang, A.X. Liu, M. Shahzad, K. Ling, S. Lu, Understanding and modeling of WiFi signal based human activity recognition, in: Proc. 21st Annual International Conference on Mobile Computing and Networking (MobiCom), Paris, France, 2015, pp. 65–76. doi:10.1145/2789168.2790093.
- [19] B. Wei, W. Hu, M. Yang, C.T. Chou, Radio-based device-free activity recognition with radio frequency interference, in: Proc. 14th International Conference on Information Processing in Sensor Networks (IPSN), Seattle, Washington, 2015, pp. 154–165. doi:10.1145/2737095.2737117.
- [20] J. Xiao, K. Wu, Y. Yi, L. Wang, L.M. Ni, Pilot: passive device-free indoor localization using channel state information, in: Proc. IEEE 33rd International Conference on Distributed Computing Systems, Philadelphia, 2013, pp. 236–245. doi:10.1109/icdc.2013.49.
- [21] C. Yang, H. Shao, WiFi-based indoor positioning, IEEE Commun. Magazine 53 (3) (2015) 150–157, <https://doi.org/10.1109/mcom.2015.7060497>.
- [22] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, H. Liu, E-eyes: device-free location-oriented activity identification using fine-grained WiFi signatures, in: Proc. 20th Annual International Conference on Mobile Computing and Networking (MobiCom), Maui, Hawaii, 2014, pp. 617–628. doi:10.1145/2639108.2639143.
- [23] V.L. Erickson, M.A. Carreira-Perpiñán, A.E. Cerpa, Occupancy modeling and prediction for building energy management, ACM Trans. Sens. Netw. 10 (3) (2014) 42, <https://doi.org/10.1145/2594771>.
- [24] A. Kamthe, V. Erickson, M.Á. Carreira-Perpiñán, A. Cerpa, Enabling building energy auditing using adapted occupancy models, in: Proc. 3rd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (BuildSys), Seattle, Washington, 2011, pp. 31–36. doi:10.1145/2434020.2434028.
- [25] J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, K. Whitehouse, The smart thermostat: using occupancy sensors to save energy in homes, in: Proc. 8th ACM Conference on Embedded Networked Sensor Systems (SenSys), Zürich Switzerland, 2010, pp. 211–224. doi:10.1145/1869983.1870005.
- [26] Y. Peng, A. Rysanek, Z. Nagy, A. Schlüter, Using machine learning techniques for occupancy-prediction-based cooling control in office buildings, Appl. Energy 211 (1) (2018) 1343–1358, <https://doi.org/10.1016/j.apenergy.2017.12.002>.
- [27] M. Jin, N. Bekiaris-Liberis, K. Weekly, C.J. Spanos, A.M. Bayen, Occupancy detection via environmental sensing, IEEE Trans. Autom. Sci. Eng. 15 (2) (2018) 443–455, <https://doi.org/10.1109/tae.2016.2619720>.
- [28] J. Tanjea, A. Krioukov, S. Dawson-Haggerty, D. Culler, Enabling advanced environmental conditioning with a building application stack, in: Proc. International Green Computing Conference, Arlington, Virginia, 2013, pp. 1–10. doi:10.1109/igcc.2013.6604519.
- [29] A. Beltran, V.L. Erickson, A.E. Cerpa, ThermoSense: Occupancy thermal based sensing for HVAC control, in: Proc. 11th ACM Conference on Embedded Network Sensor Systems (BuildSys), Roma, Italy, 2013, pp. 1–8. doi:10.1145/2528282.2528301.
- [30] R. Valle, ABROA: Audio-based room-occupancy analysis using gaussian mixtures and hidden markov models, in: Proc. Future Technologies Conference, San Francisco, 2016, pp. 1270–1273. doi:10.1109/ftc.2016.7821763.
- [31] O. Shih, A. Rowe, Occupancy estimation using ultrasonic chirps, in: Proc. ACM/IEEE 6th International Conference on Cyber-Physical Systems, Seattle, Washington, 2015, pp. 149–158. doi:10.1145/2735960.2735969.
- [32] W. Wang, J. Chen, G. Huang, Y. Lu, Energy efficient HVAC control for an IPS-enabled large space in commercial buildings through dynamic spatial occupancy distribution, Appl. Energy 207 (1) (2017) 305–323, <https://doi.org/10.1016/j.apenergy.2017.06.060>.
- [33] H. Zou, Y. Zhou, H. Jiang, S.-C. Chien, L. Xie, C.J. Spanos, WinLight: A WiFi-based occupancy-driven lighting control system for smart building, Energy Build. 158 (1) (2018) 924–938, <https://doi.org/10.1016/j.enbuild.2017.09.001>.
- [34] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, Y. Agarwal, Sentinel: occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings, in: Proc. 11th ACM Conference on Embedded Networked Sensor Systems, 2013, Roma, Italy, pp. 1–14. doi:10.1145/2517351.2517370.
- [35] A. Moreira, M.J. Nicolau, F. Meneses, A. Costa, Wi-Fi Fingerprinting in the Real World—RTLS@UM at the EvAAL Competition, in: Proc. International Conference Indoor Positioning and Indoor Navigation (IPIN), Calgary, Canada, 2015, pp. 1–10. doi:10.1109/ipin.2015.7346967.
- [36] S. Bozkurt, G. Elbilo, S. Gunal, U. Yayan, A comparative study on machine learning algorithms for indoor positioning, in: Proc. International Symposium on Innovations in Intelligent Systems and Applications (INISTA), Madrid, Spain, 2015, pp. 1–8. doi:10.1109/inista.2015.7276725.
- [37] J. Torres-Sospedra, R. Montoliu, A. Martínez-Usó, J.P. Avariento, T.J. Arnau, M. Benedito-Bordonau, J. Huerta, UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems, in: Proc. International Conference on Indoor Positioning and Indoor Navigation (IPIN), Busan, Korea, 2014, pp. 261–270. doi:10.1109/ipin.2014.7275492.
- [38] B. Tan, Q. Chen, K. Chetty, K. Woodbridge, W. Li, R. Piechocki, Exploiting WiFi channel state information for residential healthcare informatics, IEEE Commun. Magazine 56 (5) (2018) 130–137, <https://doi.org/10.1109/mcom.2018.1700064>.
- [39] H. Zou, Y. Zhou, R. Arghandeh, C.J. Spanos, Multiple kernel semi-representation learning with its application to device-free human activity recognition, IEEE Internet Things J. 6 (5) (2019) 7670–7680, <https://doi.org/10.1109/jiot.2019.2901927>.
- [40] D. Halperin, W. Hu, A. Sheth, D. Wetherall, Tool release: gathering 802.11n traces with channel state information, ACM SIGCOMM Computer Communication Review 41 (1) (2011) pp. 53–53. doi:10.1145/1925861.1925870.
- [41] D. Zhang, H. Wang, D. Wu, Toward centimeter-scale human activity sensing with WiFi signals, Computer 50 (1) (2017) 48–57, <https://doi.org/10.1109/mc.2017.7>.
- [42] H. Li, K. Ota, M. Dong, M. Guo, Learning human activities through WiFi channel state information with multiple access points, IEEE Commun. Mag. 56 (5) (2018) 124–129, <https://doi.org/10.1109/mcom.2018.1700083>.
- [43] Q. Zhou, J. Xing, J. Li, Q. Yang, A device-free number gesture recognition approach based on deep learning, in: Proc. 12th International Conference Computational Intelligence and Security (CIS), Wuxi, China, 2016, pp. 57–63. doi:10.1109/cis.2016.0022.

- [44] H. Zou, Y. Zhou, J. Yang, C.J. Spanos, Towards occupant activity driven smart buildings via WiFi-enabled IoT devices and deep learning, *Energy Build.* 177 (15) (2018) 12–22, <https://doi.org/10.1016/j.enbuild.2018.08.010>.
- [45] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: Proc. 3rd International Conference on Learning Representations (ICLR), San Diego, California, 2014. doi:arXiv:1409.1556v6.
- [46] H. Lee, C.R. Ahn, N. Choi, T. Kim, H. Lee, The effects of housing environments on the performance of activity-recognition systems using Wi-Fi channel state information: an exploratory study, *Sensors* 19 (5) (2019) 983, <https://doi.org/10.3390/s19050983>.
- [47] R.S. Bucks, D.L. Ashworth, G.K. Wilcock, K. Siegfried, Assessment of activities of daily living in dementia: development of the Bristol activities of daily living scale, *Age Ageing* 25 (2) (1996) 113–120, <https://doi.org/10.1093/ageing/25.2.113>.
- [48] K. Liu, S. Huang, Integration of data fusion methodology and degradation modeling process to improve prognostics, *IEEE Trans. Automation Sci. Eng.* 13 (1) (2016) 344–354, <https://doi.org/10.1109/taes.2014.2349733>.
- [49] Tong Xin, Bin Guo, Zhu Wang, Pei Wang, Jacqueline Chi Kei Lam, Victor Li, Zhiwen Yu, FreeSense: A robust approach for indoor human detection using Wi-Fi signals, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2 (3) (2018) 1–23, <https://doi.org/10.1145/3264953>.
- [50] P. Urwyler, R. Stucki, L. Rampa, R. Müri, U.P. Mosimann, Cognitive impairment categorized in community-dwelling older adults with and without dementia using in-home sensors that recognise activities of daily living, *Sci. Rep.* 7 (2017) 42084, <https://doi.org/10.1038/srep42084>.