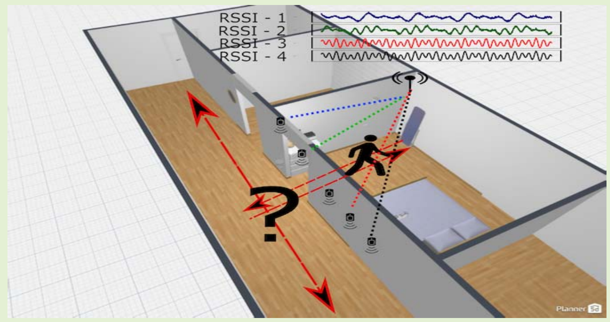
Tracking Human Motion Direction With Commodity Wireless Networks

Habibur Rahaman and Vladimir Dyo

***Abstract*—Detecting when a person leaves a room, or a house is essential to create a safe living environment for people suffering from dementia or other mental disorders. The approaches based on wearable devices, e.g. GPS bracelets may detect such events require periodic maintenance to recharge or replace batteries, and therefore may not be suitable for certain types of users. On the other hand, camera-based systems require illumination and raise potential privacy concerns. In this paper, we propose a device-free walking direction detection approach based on RF-sensing, which does not require a person to wear any equipment. The proposed approach monitors the signal strength fluctuations caused by the human body on ambient wireless links and analyses its spatial patterns using a convolutional neural network to identify the walking direction. The approach has been evaluated experimentally to achieve up to 98% classification accuracy depending on the environment.**



***Index Terms*—Activity recognition, machine learning, Internet of Things, RF-sensing, 802.15.4.**

I. Introduction

WANDERING is a dangerous behavior associated with dementia or Alzheimer's disease as people who wander outside their house can easily get lost or even die from exposure to harsh weather or other risks [1], [2]. Creating a safe environment is known to be essential to cope with this behaviour and may involve equipment such as alarms to detect when the person leaves his place [3]. The traditional measures include deploying alarms on doors or windows, such as pressure-sensitive mats [4] at the door, or having a person carry a GPS location tracking device [5]. The prior work on walking direction detection is based on passive infrared (PIR) sensor grid solutions [6]–[9], RFID based solutions [10], [11] and smartphone-based solutions [12], [13]. However, passive infrared sensor grid solutions require instrumenting the space with specialised sensor equipment, whereas wearable devices or phone-based solutions [12], [13] require periodic maintenance, such as recharging or replacing the batteries, which elderly people cannot always be relied upon.

RF-sensing based approaches [14]–[17] emerged as an alternative to traditional PIR or other solutions. RF-sensing is a technique that can locate human objects by analyzing their impact on the signal strength of ambient wireless links [18]. The technique is based on the fact that the human body attenuates, scatters, and reflects radiofrequency waves due to 70% water content [19]. As the pattern of those fluctuations varies with the object position relative to each wireless link, the object's location can be established using machine learning algorithms. The key advantage of RF-sensing is the ability to localize objects without using wearable devices or deploying specialized sensor equipment within indoor space. As opposed to video-based tracking methods, RF-sensing enables developing sensing solutions that are non-intrusive and with minimal privacy implications. RF-sensing has been recently applied to tracking and monitoring fall detection [20], breath detection [21], [22], gesture [23], [24] even speech recognition [25].

In this work, we propose an RF-sensing-based direction detection approach, which can detect movement direction and distinguish whether a person is entering, leaving the room, or just passing by without entering or leaving the room. The major challenge in applying the RF-sensing approach is the relatively short event duration, which severely constrains the sample size available for statistical analysis. This is because traditional RF-sensing configurations typically involve multiple sensor nodes transmitting beacons to a centralized base station [18], which due to channel collisions and receiver capacity constraint, limits the beacon rate and consequently the sample size. The second challenge is the need to distinguish a pattern of Received Signal Strength (RSS) fluctuations to

Manuscript received July 19, 2021; revised August 29, 2021; accepted September 4, 2021. Date of publication September 7, 2021; date of current version October 18, 2021. The associate editor coordinating the review of this article and approving it for publication was Dr. Amitava Chatterjee. *(Corresponding author: Vladimir Dyo.)*

The authors are with the School of Computer Science and Technology, University of Bedfordshire, Luton LU1 3JU, U.K. (e-mail: habibur.rahaman@study.beds.ac.uk; vladimir.dyo@beds.ac.uk).

Digital Object Identifier 10.1109/JSEN.2021.3111132

identify the direction within the same spatial trajectories. For this reason, a classical fingerprinting analysis used by mapping spatial signatures to a location may not be suitable. What is needed is a method that analyzes a sequence of spatial signatures to detect a direction.

In this paper, we propose and evaluate a *reverse* RF-sensing configuration, which is capable of capturing relatively short-lived events. The reverse-RF sensor involves a single transmitter and multiple receivers, which allows to overcome the channel constraint and significantly increase the sample size necessary for classification. The RSS samples collected from sensor nodes are then analyzed using Convolutional Neural Network for pattern analysis to detect a trajectory direction. The contributions of this paper are as follows:

- Novel reverse RF-sensing method for direction detection for assisted living applications.
- Experimental evaluation and measurements.
- Performance evaluation using a case study in several indoor locations.

To the best of our knowledge, this is the first work that proposes a reverse RF-sensing for direction detection applications. The rest of the paper is structured as follows. Section II describes a system model including a scenario and a wireless setup. Section III explains the data pre-processing and a deep learning based classification approach. Section IV presents results, discussion, and a small-scale case study. Section V contains related work, followed by Conclusion in Section VI.

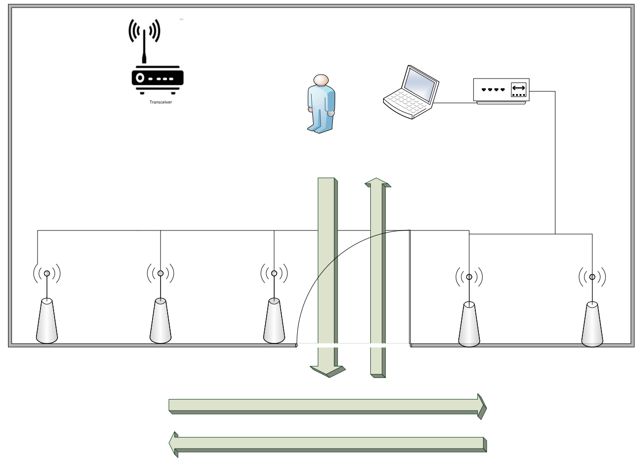


Fig. 1. The layout of the room and sensor placement. 5 receiver nodes are located on both sides of the doorway with a transmitter located inside the room.

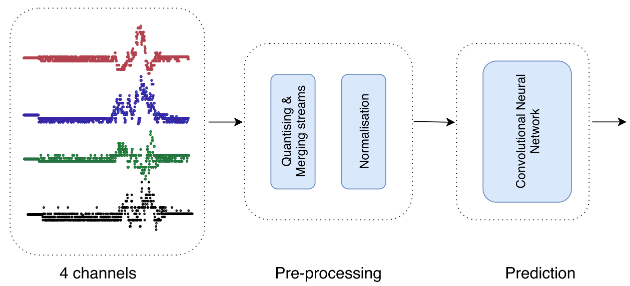


Fig. 2. A high-level overview of an approach.

II. System Model

Our scenario contains a doorway through which a person can enter or leave a room, or can pass by along the corridor, as shown in Figure 1. To track user movements, multiple wireless nodes are located on both sides of the doorway with the transmitter located inside the room. The sensor nodes are equipped with IEEE 802.15.4 CC2420 transceivers [26] operating in the 2.4 GHz band. The IEEE 802.15.4 standard defines 16 channels, and uses Direct Sequence Spread Spectrum (DSSS) to map 4-bits data symbols to a 32-chip pseudo-random direct sequence, which is then modulated using offset quadrature phase shift keying (OQPSK).

A single node is selected as a transmitter that broadcasts beacons at a very high packet rate. The beacons are received by MTM-CM500-MSP nodes connected via a USB port to a laptop running Ubuntu. As the person moves within the room or the corridor, he or she attenuates the wireless links, which is detected as a change in Received Signal Strength (RSS) levels. The pattern of RSS changes across all links can then be analyzed to detect the person's movement direction. The transmit power was reduced to -10 dBm to make sure that only activities in the close vicinity of the room impact the wireless signal strength. Each sensor node logs the data in the following format:

```
<timestamp> <senderID><SEQNO><RSS>
...
```

where RSS and SEQNO are beacon signal strength and sequence number respectively. The RSS value is measured by the transceiver over approximately 128 μs, which corresponds to 8 symbol periods. The received signal strength values for CC2420 transceiver generally range from -100 dBm to 0 dBm with $\pm$3 dB linearity. The timestamp with a millisecond precision is generated by a laptop upon receiving a packet from each USB port.

The proposed RF-sensing configuration uses a single transmitter that broadcasts at very high packet rate to maximize the RF-sensing sampling rate and enable capture of relatively short-lived events. This is in contrast to existing RF-sensing techniques described in the literature, in which multiple nodes broadcast and receive beacons from other neighbours at the same time, with one of the nodes collecting data and streaming it to a laptop [18]. While normal configuration allows sampling of an RF signal from multiple transmitters at multiple locations, it also results in a much lower sampling rate as the receiver throughput is effectively divided among multiple transceivers, reducing the sampling rate on each individual link accordingly. If the nodes use contention-based MAC protocols that further limits the maximum possible packet rate due to packet collisions and channel access time. As an illustration, a system of 4 nodes configured with 1Hz beacon would capture only 15 samples per 15-second event, whereas reverse configuration used in the experiments captured data at a sample rate of 75 packets per second, with approximately 1150 samples per wireless link per event.

III. Direction Detection

A high-level overview of the approach is shown in Figure 2. The RSS values from k wireless links are pre-processed to obtain a $k \times n$ matrix, where each row corresponds to an event

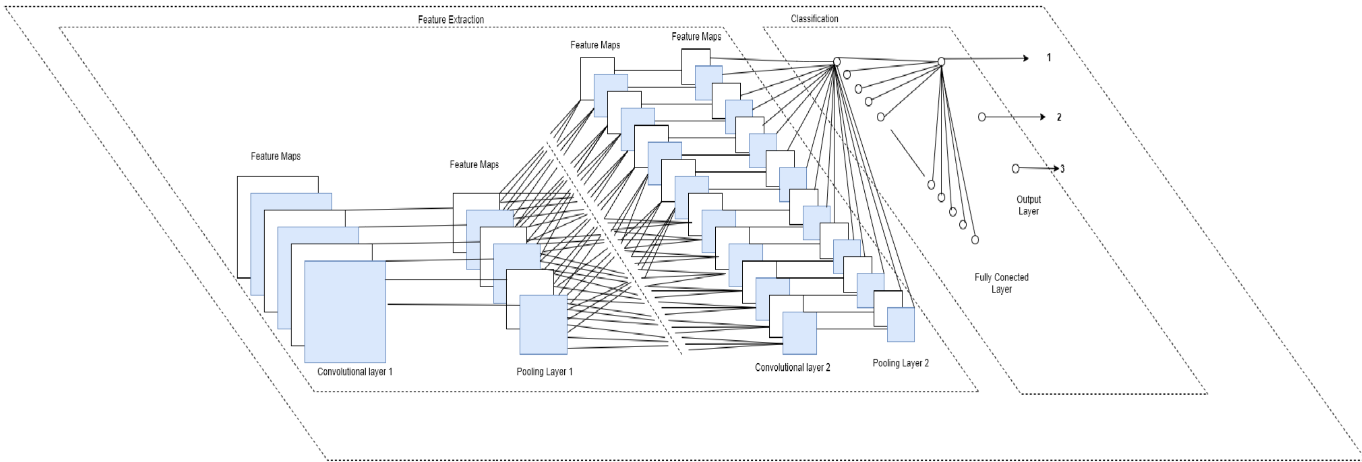
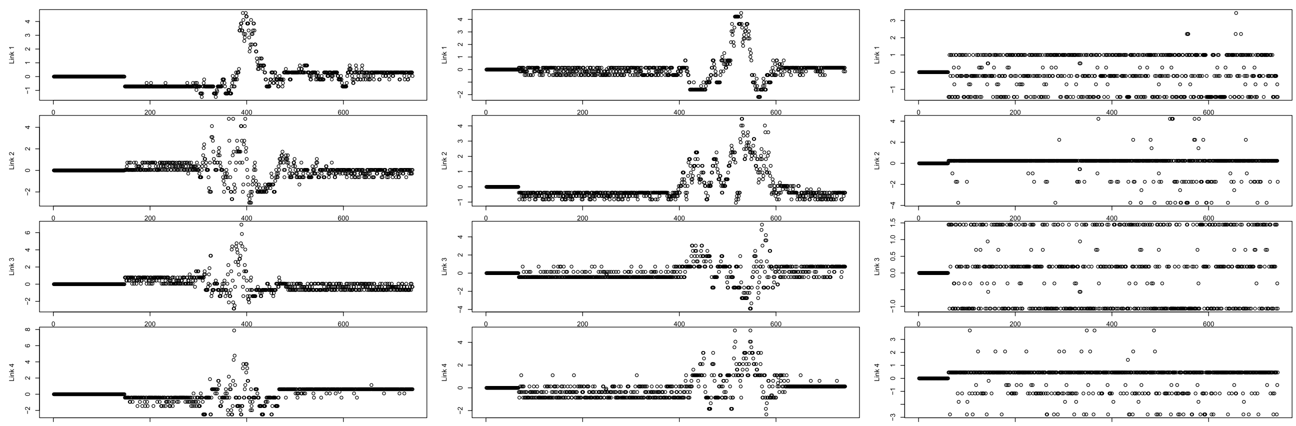


Fig. 3. A convolutional neural network architecture for wireless direction detection.



(a) Walking inside (b) Walking outside (c) Passing by

Fig. 4. RSS data from 4 sensor nodes for each activity.

signature in time. After normalization, the matrix is then fed into a convolutional neural network classifier as described in Section III-B.

A. Data Pre-Processing

Each receiver produces a series of RSS values, so each user activity measurement, such as walking inside or walking outside, represents k RSS streams, one for each wireless link. Although the streams are synchronous, i.e. each beacon is received simultaneously by all k sensor nodes, occasionally lost packets, jitter associated with USB logging in combination with timestamping performed locally on a laptop means that the received streams are not strictly aligned. To synchronize, the timestamps have been quantized to N-milliseconds epochs, and the RSS value within each epoch was substituted by its average value. Consequently, each activity measurement produced a stream of $k+1-$element tuples:

```
(EpochID, RSS_1, RSS_2, RSS_i... RSS_k)
...
```

where k is the number of directed wireless links in the network. The RSS values within each epoch can also be aggregated by selecting a maximum or minimum value for each link within an epoch, however, as shown in our prior work [18], selecting the mean value showed the best classification accuracy. The records containing missing values have been removed from the analysis. We have then computed a running standard deviation for each stream, which was then zero-meaned and normalized. The ordered sequence of spatial signatures forms a signature for each event.

Figure 4 shows the raw RSS data for each activity with each subplot showing the data from each wireless link. A simple visual inspection shows that activity 3 results in quite different RSS fluctuations compared to the other two activities. The signal signatures from moving inside and outside (Figs 4a and 4b) are more difficult to distinguish visually but can be separated using a convolutional neural network approach as shown in our analysis later.

B. Convolutional Neural Network

Convolutional Neural Networks (CNN) is a biologically inspired feedforward neural network that represents the mammalian visual cortex and is widely used in image, video, and signal classification. As a deep learning algorithm, CNN automatically extracts features from input data using multiple levels of abstraction [27].

Our CNN architecture consists of an input layer, convolutional layers, pooling layers, a fully connected layer, and the output layer as shown in Figure 3. The classifier maps patterns of signal strength fluctuations to a specific trajectory such as going in or outside of the room. Below we describe the configuration of each layer in more detail:

1) Input Layer: The input to CNN is a 1D array containing a stream of RSS values from a certain wireless link with each wireless link representing a separate channel. Each array is zero-meaned and normalised. As every activity measurement resulted in slightly different number of samples, the arrays have been zero-padded from the beginning of the array.

2) Convolutional Layer: The convolutional layer automatically extracts higher-level features from RSS data by applying mathematical 1D convolution. $c_i^k = \sum_{n=0}^{2} w_n x_{i+n} + b^k$, where w is a 1×24 feature detector, which slides over an input data vector performing a convolution operation over the corresponding patch from 1D input data array x. k is the number of feature detectors also known as filters or kernels learned during the training process. The number of filters and their dimension in both convolutional layers were selected experimentally as described later in Section IV. Finally, i, n are location indexes for the original 1D array and the convolution kernel matrix respectively. The result of the convolution c is called a feature map, which is fed into the next layer after an activation function. The proposed model contains one convolutional layer. The performance has been evaluated with a higher number of layers without significant gains in classification accuracy.

The model uses a rectifier linear unit (*ReLU*) activation function, due to the absence of gradient vanishing problem and fast training speed. The ReLU function is defined as: $f(x) = max(0, x)$. The convolutional neural network uses weight sharing to make feature detection translation invariant, reduce the network parameters and reduce training time.

3) Pooling Layer: A pooling layer (also called downsampling) after the convolutional layer reduces the data dimensionality, reduces overfitting, and shortens the training time. The model uses a *Max Pooling*, which takes the largest element from each spatial window within a rectified feature map. The pooling operation is applied separately to each feature map. The output of convolutional and pooling layers represents a high-level feature map, which is then flattened into a single vector of values.

4) Fully Connected Layer: The fully connected neural network takes the output from the convolutional and pooling layers and makes the final classification decision about user activity. The layer contains 64 units (neurons) with every unit connected to every unit in the previous layer. To prevent overfitting and memorization, the model uses a dropout layer with a probability 0.1.

5) Output Layer: The output layer uses *softmax* function to produce output class probabilities from class scoring function. The softmax function is defined as: $f(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{K} x_j}$ where x_i is the scoring function for class i and K is the number of classes. The number of units in output layer is equal to the number of activities to be identified, which is 3 in our experiments.

TABLE I
COMPARISON OF CNN, LSTM AND DTW

	CNN	LSTM	DTW
Accuracy	95.2	50.8	79.6
Recall	95.2	50.8	79.7
Precision	95.2	49.7	81.9
F1-score	95.2	48.9	79.6

LSTM [27] and Dynamic Time Warping [28] based classifiers have been used for comparison. The LSTM model contains 3 layers with 32, 16, and 3 elements in the hidden, fully connected and output layers. The dropout and recurrent dropout values in the hidden layer were set to 0.1. The output layer similarly uses softmax activation function with one output neuron per class value. The Dynamic Time Warping based classifier computes the DTW distance to every trajectory in the training set and then selects a class of the trajectory with the minimum distance. The classification accuracy was computed separately for each wireless link and then averaged for all 1, 2, and 3 link combinations. Similarly to CNN, 4-fold cross-validation has been adopted.

IV. EXPERIMENTAL EVALUATION AND DISCUSSION

The section describes the experimental evaluation of the proposed approach in terms of classification accuracy. The impact of hyperparameters, the performance, and limitations of the approach have also been discussed in this section.

A. Experiment Setup

The experiment has been carried out using three different movement trajectories: leaving the room, entering the room, and passing by without entering the room. Each experiment round lasted for about 15 seconds and was repeated 36 times resulting in a total of 144 wireless link measurements. Although the measurements have been collected with 5 sensor nodes, the reading from node ID 4 had to be discarded due to unstable operation. The model has been implemented using Keras [29] and DTW library [30] in *R* statistical environment [31]. The prediction model was evaluated using repeated k-fold cross-validation with $k = 4$ with final results averaged across 20 repetitions as described below. The data are split into k partitions, then the training is done on the first $k - 1$ partitions and validation on the last one. The process is repeated for all k partitions and the final result is averaged. The data have been processed on 2.3GHz Intel core i9-9880H.

Figure 5 shows a loss and accuracy during the training and validation phases. The curve indicates that the model does not overfit. Table I compares the classification accuracy for convolutional neural network, LSTM, and DTW-based classifiers. As can be seen, CNN performs much better compared to LSTM. The reason for higher accuracy is that because CNN can capture the features of RSS signals corresponding to movements in each direction. The features are not known apriori, but the convolutional neural network is able to extract such sequences from RSS data. DTW-based classifier shows higher accuracy than LSTM but requires a considerably higher amount of computation to compute DTW distances.

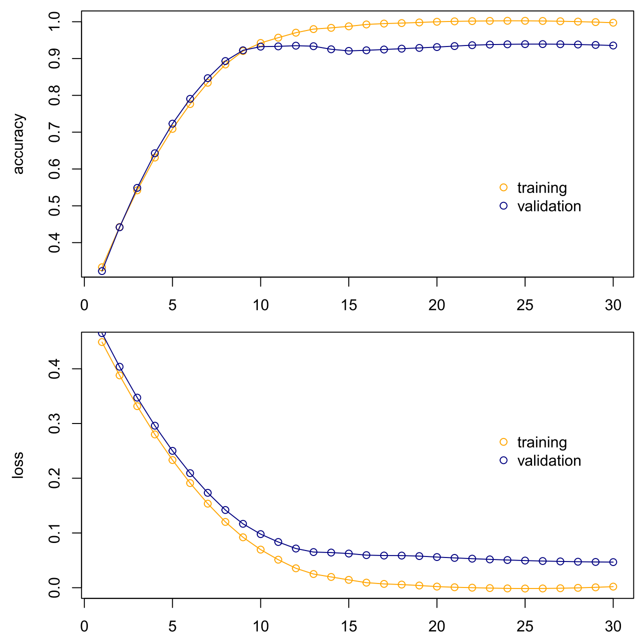


Fig. 5. Learning curve.

TABLE II
THE IMPACT OF LAYER CONFIGURATION

CNN layers	Time to train, s	Accuracy, %
1 conv - 1 maxpool	163	81.9
2 conv - 1 maxpool	226	85.8
1 conv-1 maxpool-1 conv-1 maxpool	175	95.2

B. Hyper-Parameter Tuning

1) The Number of Convolutional and Pooling Layers: Table II shows the impact of the number of convolutional layers on classification accuracy. The experiments have been performed with 12 filters, 1×24 downsampling. The basic configuration with just one convolutional and 1 pooling layer resulted in the lowest performance. The two 2 convolutional layers used in the second configuration allow for extracting potentially more complex features but did not significantly affect the performance. The configuration with two layers of convolutional and max-pooling resulted in the highest accuracy.

The downsampling rate was found to have a significant impact on classification accuracy. The best performance is achieved with the pooling configuration of 1×24, which is rather unusual given prior work in this area typically selected ratios of 1×2 with varying stride sizes. Increasing beyond 1×24 did not bring any performance improvement; moreover, a global max pool network configuration, where an entire input array is summarised by a single value, resulted in significantly lower accuracy.

2) The Number of Filters and Filter Size: Table III shows the impact of filter size on classification accuracy. 1D convolutional neural networks can have relatively large filter sizes compared to their 2D and 3D counterparts. As can be seen, the accuracy increases gradually until the filter size of 24 and then reaches a plateau. The results show that selecting a filter size of 12 is a reasonable compromise between accuracy and model complexity, i.e. further increase in filter size provides minimal marginal accuracy improvement.

TABLE III
THE IMPACT OF FILTER SIZE

CNN configuration	Filter size	Time to train, s	Accuracy,%
2 layer CNN, 12 filters, maxpool x24	1x2	174	81.1
	1x4	171	86.2
	1x6	171	89.1
	1x12	175	95.2
	1x24	207	96.26

TABLE IV
THE IMPACT OF NUMBER OF FILTERS

CNN configuration	Filters	Time to train, s	Accuracy,%
2 layer CNN, filter dimension 1x24, max pool 1x24	3	181	86.3
	6	180	93.1
	12	186	94.3
	24	207	96.25

TABLE V
THE IMPACT OF NUMBER OF WIRELESS LINKS

CNN configuration	Wireless links	Accuracy,%
2 layer CNN, 12 filters, filter dimension 1x24, max pool 1x24	1	83.4
	2	89.8
	3	93.6
	4	95.2

Table IV shows the impact of the number of filters on classification accuracy. The performance gradually increases with the number of filters but the marginal improvement reduces after the number of filters reaches 12.

3) The Number of Wireless Links: Intuitively, the classification accuracy should increase as the number of wireless links increases. In this experiment, we quantify our intuition by measuring classification accuracy against 1, 2, 3, and 4 wireless links. The performance was evaluated for all combinations of links with an average computed for each number of links.

Interestingly, with just one wireless link, the classification accuracy is 83.4%, significantly higher than anticipated. This is in contrast to localization systems, which require data from at least several wireless links to achieve reasonable accuracy. The reason for relatively high accuracy with just one wireless link is possibly the fact that in a direction detection scenario, the person's trajectory is fixed. Therefore, the sequential patterns from signal strength fluctuations from one link can be sufficient to detect the direction. As expected the classification accuracy increases gradually with the number of links. Although the increase is not significant, the higher number of links may help the system to be more robust against noise and interference.

C. Case Study

In this section, we present the performance of our classification method using a separate dataset collected over 3 different locations within the same house, which included a front door to the house (Figure 1) and 2 internal doorways. In each location, the participant walked along one of the following trajectories: through the doorway in each direction, and by the doorway outside of the house for the front door (or along the corridor for internal doors) in both directions. The wireless sensors were placed on both sides of the door with the transmitter located inside the room as shown in Fig. 6. The duration of each measurement was 15 seconds with the experiment repeated 12 times for each direction resulting in a total of 240 measurements. The direction was determined

Fig. 6. The receiving nodes are located from both sides of the doorway and connected via USB hub to a laptop. The transmitting node is located indoor at the same height approximately 2 meters from the door.

using a convolutional neural network algorithm, and 4-fold cross-validation.

1) Metrics: The classifier accuracy was measured using three metrics: accuracy, sensitivity, and specificity as defined below:

Accuracy: The metric shows the total proportion of data that has been correctly classified and is computed as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP, FP, TN and FN refer to true positive, false positive, true negative, and false negative respectively.

Sensitivity: The metric also known as true positive rate or recall shows the proportion of relevant items that have been selected and is defined as follows:

$$Sensitivity = \frac{TP}{TP + FN}$$

Specificity: The metric shows the proportion of negatives that have been correctly classified and is defined as follows:

$$Specificity = \frac{TN}{TN + FP}$$

2) Results: The resulting classification accuracy is 98.3%, 93.2%, and 82.3% for each door respectively as shown in Table VI. The accuracy was the highest for the front door, which represents a relatively simple environment, as the passing trajectory is perpendicular to the doorway with a door in the middle of the trajectory. In the case of internal doors, the environment is more complex, especially if the door is located close to the end of the corridor, in which case 'pass' and 'in/out' trajectories may partially overlap. For doorway 3, both the recall and specificity drop below 90. In particular the recall for 'out' trajectory reduces to 64.5

Tables VII-IX show confusion matrices for each scenario. Although the approach shows relatively good accuracy in distinguishing the in-out movements, the ability to separate between left and right along the corridor is quite limited, which could be due to the fact that all the sensors including a transmitter are located inside a room. This in combination with low transmit power, limits the ability to track movements through the walls, hence left-right and right-left trajectories have been merged under a single category 'pass' in the presented results.

TABLE VI
OVERALL ACCURACY BY LOCATION

Location	Accuracy, %	Specificity			Sensitivity		
		In	Out	Pass	In	Out	Pass
Door 1	98.8	100	100	98	95.5	100	100
Door 2	93.2	100	98	91	91	90	96
Door 3	82.3	93	94	84	88.6	64.5	88.2

TABLE VII
CONFUSION MATRIX. DOOR 1

	In	Out	Pass
In	191	0	9
Out	0	200	0
Pass	0	0	400

TABLE VIII
CONFUSION MATRIX. DOOR 2

	In	Out	Pass
In	182	1	17
Out	0	180	20
Pass	3	13	384

TABLE IX
CONFUSION MATRIX. DOOR 3

	In	Out	Pass
In	195	6	19
Out	27	142	51
Pass	17	35	388

D. Discussion

The experiments show that reverse RF-sensing configuration is able to capture motion direction due to a combination of high sampling rate, multiple RF receivers, and deep learning. The results demonstrate that the approach is feasible and the accuracy depends on the doorway location within the house. However, a practical application of the proposed approach would require its adaptation and evaluation for more complex trajectories associated with non-entry and non-leaving events. This is when a person approaches a door for example, but then turns around and moves in the opposite direction, a scenario which to the best of our knowledge, has not been explored yet in related work. For example, [14] evaluates motion across 8 directions, [15] and [16] evaluate motion in 2 directions, with most works focusing on the straight or simple trajectory direction functionality. Developing such solutions, however, is outside of the scope of the paper and is a potential future work.

V. RELATED WORK

Indoor localization and activity recognition are well-researched topics. Most prior work on wireless device detection rely on smartphones or wearables such as passive RFID tags or other equipment. WalkCompass developed by Roy *et al.* [12] analyzes a smartphone's accelerometer and magnetic sensors to determine the direction in which a user is walking. Experimental results from 15 different buildings show that the system requires around 5 steps to detect the direction, with a median error of 8%. DirectMe [13] uses a real-time algorithm to determine a user's facing direction irrespective of the phone's orientation. The system has been implemented on Galaxy Nexus phone and tested in three different phone orientations and was shown an average and a maximum angular error of 10% and 18% respectively. Both systems require a user carrying a smartphone, which may not be possible in the assisted living scenarios we consider.

Reference [32] developed an infrastructure-less localization system, and uses ambient cues from light flicker to detect indoor location, however, it requires a wearable device.

A. PIR-Array Based

Yun and Song [6] use an array of 4 PIR sensors whose elements are orthogonally aligned to determine the relative direction of human movement. The raw data from PIR sensors are analyzed using machine learning algorithms and the results show the direction accuracy between 89 to 95%. Earlier, Lue *et al.* [7] propose an indoor human tracking system based on 4 sensor clusters each consisting of 5 PIR sensors mounted on the ceiling. The system's feasibility was shown in the simulation. Similarly, [8] and [9] propose a distributed PIR array mounted on the walls for indoor human and robotic tracking. The proposed RF-based approach is based on exploiting commodity RF hardware as motion and direction sensors and demonstrates the feasibility of an alternative to the traditional PIR based solutions.

B. RF-Based

WiDIR [14] was the first work to utilize Wi-Fi signals to detect walking direction. The work is based on the idea is that human motion changes multipath signal distribution, which can be detected by monitoring Channel State Information (CSI) at the receiver. The evaluation using one transmitter and 2 receivers located in 3 corners of the room, showed the accuracy over 8 basic directions with the overall mean accuracy ranging from 10.5 to 15.5 degrees depending on the room. However, the approach relies on Channel State Information, which is supported by the relatively small number of adapters.

Liu *et al.* [15] propose 2 methods to determine a user's walking direction passing through a matrix of 18 XM2110 sensor nodes mounted on both sides of the corridor at 3 different heights. The unsupervised method uses a short-term window RSS variance as a feature, and estimates walking direction through computing a time-delay between signal fluctuations on two selected links. The supervised method uses Dynamic Time Warping (DTW) method to compare the measured signal on a selected link with one of the reference signals on the same link. The accuracy ranges from 40% to nearly 99% depending on link pair combinations. The limitation of the approach is that it relies on a small subset of links for walking direction determination, but cannot predict which subset will provide the optimal performance. In contrast, our approach uses deep learning to analyse a stream of all available wireless links and does not require selecting a specific subset.

Huang and Lin [16] design Dynamic Time Warping based algorithm to determine walking direction along the corridor using RSS signals from 3 Wi-Fi Access points. The algorithm is also based on the idea that the delay in signal fluctuations on certain link pairs depends on the movement direction. The authors report a classification accuracy of around 70%-82% depending on the link pair combination. Similar to [15] the approach cannot automatically select an optimal subset of links for classification. Rather than trying to monitor the delay in signal fluctuations between certain link pairs, our approach is based on learning the signal patterns that are specific to each walking direction.

Chowdhury proposed a method that analyses channel state information (CSI) for human activity classification [33] and claimed that the proposed system could affect the wellbeing of elderly people. The focus of the research was human movement detection and activity classification rather than walking direction detection. Finally, Hillyard *et al.* [17] propose a method to detect border crossing events using RF transceivers deployed along the border. The system was evaluated using nine CC2531 nodes deployed along 29.26 m border, logging RSS values each time user crossed the links and shown to have 86% correct classification rate. The system, however, does not detect the movement direction. It is also worth noting that most works in the area investigate movement direction using a single sensor modality (i.e. either PIR, RF, accelerometer, but not a combination).

VI. Conclusion and Future Work

The ability to detect movement direction through the doorway is essential for assisted living applications. In this paper, we propose a novel method for movement direction detection based on RF-sensing. The proposed reverse RF-sensing configuration is able to capture and analyze short-duration events, such as people walking inside or outside of the room. The movement direction is determined through the analysis of spatial patterns from RSS signal strength fluctuations using a convolutional neural networks algorithm. The experimental evaluation on real measurements has demonstrated the accuracy of up to 82-98% depending on the number of wireless links and the indoor environment.

Although the approach was motivated by assisted living applications, it would be interesting to explore its potential for monitoring and controlling the pedestrian traffic flows within stores, offices, restaurants and other indoor spaces in general, as part of COVID-19 measures implemented by some organizations [34] to limit the spread of the virus. We leave this as potential future work.

Acknowledgment

The authors would like to thank the reviewers for their helpful comments and suggestions.

References

[1] L. Barnard-Brak and P. Parmelee, "Measuring risk of wandering and symptoms of dementia via caregiver report," *J. Appl. Gerontol.*, vol. 2020, Aug. 2020, Art. no. 733464820947287.

[2] N. A. Neubauer and L. Liu, "Influence of perspectives on user adoption of wander-management strategies," *Dementia*, vol. 20, Mar. 2020, Art. no. 1471301220911304, doi: 10.1177/1471301220911304.

[3] A. Hammoud, M. Deriaz, and D. Konstantas, "Wandering behaviors detection for dementia patients: A survey," in *Proc. 3rd Int. Conf. Smart Sustain. Technol. (SpliTech)*, 2018, pp. 1–5.

[4] L. Ault, R. Goubran, B. Wallace, H. Lowden, and F. Knoefel, "Smart home technology solution for night-time wandering in persons with dementia," *J. Rehabil. Assistive Technol. Eng.*, vol. 7, Sep. 2020, Art. no. 2055668320938591, doi: 10.1177/2055668320938591.

[5] V. Varadharajan, U. Tupakula, and K. Karmakar, "Secure monitoring of patients with wandering behavior in hospital environments," *IEEE Access*, vol. 6, pp. 11523–11533, 2018.

[6] J. Yun and M.-H. Song, "Detecting direction of movement using pyroelectric infrared sensors," *IEEE Sensors J.*, vol. 14, no. 5, pp. 1482–1489, May 2014.
[7] X. Luo, B. Shen, X. Guo, G. Luo, and G. Wang, "Human tracking using ceiling pyroelectric infrared sensors," in *Proc. IEEE Int. Conf. Control Autom.*, Dec. 2009, pp. 1716–1721.
[8] Q. Hao, D. J. Brady, B. D. Guenther, J. B. Burchett, M. Shankar, and S. Feller, "Human tracking with wireless distributed pyroelectric sensors," *IEEE Sensors J.*, vol. 6, no. 6, pp. 1683–1696, Dec. 2006.
[9] R. C. Luo, J. Yi-Huei, and O. Chen, "Robotics human tracking system through wireless pyroelectric sensor system," in *Proc. IEEE Workshop Adv. Robot. Its Social Impacts*, Aug. 2008, pp. 1–6.
[10] Y. Oikawa, "Tag movement direction estimation methods in an RFID gate system," in *Proc. 6th Int. Symp. Wireless Commun. Syst.*, Sep. 2009, pp. 41–45.
[11] R. Sankarkumar and D. Ransinghe, "Watchdog: A novel, accurate and reliable method for addressing wandering-off using passive RFID tags," in *Proc. 11th Int. Conf. Mobile Ubiquitous Syst., Comput., Netw. Services*, 2014, pp. 208–217, doi: 10.4108/icst.mobiquitous.2014.258040.
[12] N. Roy, H. Wang, and R. R. Choudhury, "I am a smartphone and i can tell my user's walking direction," in *Proc. 12th Annu. Int. Conf. Mobile Syst., Appl., Services*, New York, NY, USA, Jun. 2014, pp. 329–342, doi: 10.1145/2594368.2594392.
[13] A. Mariakakis, "DirectMe: A mobile phone algorithm for direction detection," Duke Univ., Durham, NC, USA, Tech. Rep., 2013. [Online]. Available: https://ece.duke.edu/sites/ece.duke.edu/files/GWDD2013_Mariakakis_0.pdf
[14] D. Wu, D. Zhang, C. Xu, Y. Wang, and H. Wang, "WiDir: Walking direction estimation using wireless signals," in *Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput.*, New York, NY, USA, Sep. 2016, pp. 351–362, doi: 10.1145/2971648.2971658.
[15] T. Liu, Z.-M. Chen, and Z.-Q. Liang, "Walking direction detection using received signal strengths in correlated RF links," in *Proc. 11th Asian Control Conf. (ASCC)*, Dec. 2017, pp. 2872–2877.
[16] H. Huang and S. Lin, "WiDet: Wi-Fi based device-free passive person detection with deep convolutional neural networks," in *Proc. 21st ACM Int. Conf. Modeling, Anal. Simul. Wireless Mobile Syst.*, New York, NY, USA, Oct. 2018, pp. 53–60, doi: 10.1145/3242102.3242119.
[17] P. Hillyard, N. Patwari, S. Daruki, and S. Venkatasubramanian, "You're crossing the line: Localizing border crossings using wireless RF links," in *Proc. IEEE Signal Process. Signal Process. Educ. Workshop (SP/SPE)*, Aug. 2015, pp. 249–254.
[18] H. Rahaman and V. Dyo, "Counting calories without wearables: Device-free human energy expenditure estimation," in *Proc. 16th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2020, pp. 1–6.
[19] N. Patwari, L. Brewer, Q. Tate, O. Kaltiokallio, and M. Bocca, "Breathfinding: A wireless network that monitors and locates breathing in a home," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 1, pp. 30–42, Feb. 2014.
[20] Y. Tian, G.-H. Lee, H. He, C.-Y. Hsu, and D. Katabi, "RF-based fall monitoring using convolutional neural networks," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–24, Sep. 2018, doi: 10.1145/3264947.
[21] A. M. Ashleibta, Q. H. Abbasi, S. A. Shah, M. A. Khalid, N. A. Abuali, and M. A. Imran, "Non-invasive RF sensing for detecting breathing abnormalities using software defined radios," *IEEE Sensors J.*, vol. 21, no. 4, pp. 5111–5118, Feb. 2021.
[22] H. Abdelnasser, K. A. Harras, and M. Youssef, "Ubibreathe: A ubiquitous non-invasive wifi-based breathing estimator," in *Proc. 16th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, New York, NY, USA, 2015, pp. 277–286, doi: 10.1145/2746285.2755969.
[23] M. A. A. Haseeb and R. Parasuraman, "Wisture: RNN-based learning of wireless signals for gesture recognition in unmodified smartphones," 2017, *arXiv:1707.08569*. [Online]. Available: https://arxiv.org/abs/1707.08569
[24] Z. Chi *et al.*, "EAR: Exploiting uncontrollable ambient RF signals in heterogeneous networks for gesture recognition," in *Proc. 16th ACM Conf. Embedded Netw. Sensor Syst.*, New York, NY, USA, Nov. 2018, pp. 237–249, doi: 10.1145/3274783.3274847.
[25] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with Wi-Fi!" *IEEE Trans. Mobile Comput.*, vol. 15, no. 11, pp. 2907–2920, Nov. 2016, doi: 10.1109/TMC.2016.2517630.
[26] *TI CC2420 ZigBee-Ready RF Transceiver*. Accessed: Jul. 18, 2021. [Online]. Available: https://www.ti.com/lit/ds/symlink/cc2420.pdf
[27] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.
[28] M. Müller, "Dynamic time warping," in *Information Retrieval for Music and Motion*. Berlin, Germany: Springer, 2007, pp. 69–84.
[29] F. Chollet *et al.* (2015). *Keras*. [Online]. Available: https://keras.io
[30] T. Giorgino, "Computing and visualizing dynamic time warping alignments in R: The dtw package," *J. Stat. Softw.*, vol. 31, no. 7, pp. 1–24, 2009. [Online]. Available: https://www.jstatsoft.org/v031/i07
[31] R Core Team. (2019). *R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing*. Vienna, Austria. [Online]. Available: https://www.R-project.org/
[32] B. Munir and V. Dyo, "Passive localization through light flicker fingerprinting," *IEEE Sensors J.*, vol. 19, no. 24, pp. 12137–12144, Dec. 2019.
[33] T. Z. Chowdhury, "Using Wi-Fi channel state information (CSI) for human activity recognition and fall detection," M.S. thesis, Univ. British Columbia, Vancouver, BC, Canada, 2018. [Online]. Available: http://hdl.handle.net/2429/65593, doi: 10.14288/1.0365967.
[34] HM Government. (2020). *Keeping Workers and Customers Safe During COVID-19 in Restaurants, Pubs, Bars and Takeaway Services. COVID-19 Secure Guidance for Employers, Employees and the Selfemployed*. [Online]. Available: https://assets.publishing.service.gov.uk/media/5eb96e8e86650c278b077616/working-safely-during-covid-19-restaurants-pubs-takeaway-services-091120.pdf