

Translation Resilient Opportunistic WiFi Sensing

Mohammud J. Bocus¹, Wenda Li², Jonas Paulavicius¹, Ryan McConville¹, Raul Santos-Rodriguez¹,
Kevin Chetty² and Robert Piechocki¹

¹ Department of Electrical and Electronic Engineering, University of Bristol, BS8 1UB, UK.

² Department of Electronic and Electrical Engineering, University College London, UK

Abstract—Passive wireless sensing using WiFi signals has become a very active area of research over the past few years. Such techniques provide a cost-effective and non-intrusive solution for human activity sensing especially in healthcare applications. One of the main approaches used in wireless sensing is based on fine-grained WiFi Channel State Information (CSI) which can be extracted from commercial Network Interface Cards (NICs). In this paper, we present a new signal processing pipelines required for effective wireless sensing. An experiment involving five participants performing six different activities was carried out in an office space to evaluate the performance of activity recognition using WiFi CSI in different physical layouts. Experimental results show that the CSI system has the best detection performance when activities are performed half-way in between the transmitter and receiver in a line-of-sight (LoS) setting. In this case, an accuracy as high as 91% is achieved while the accuracy for the case where the transmitter and receiver are co-located is around 62%. As for the case when data from all layouts is combined, which better reflects the real-world scenario, the accuracy is around 67%. The results showed that the activity detection performance is dependent not only on the locations of the transmitter and receiver but also on the positioning of the person performing the activity.

I. INTRODUCTION

WiFi has experienced rapid growth due to an increasing number of wireless devices. Its success is also due to the multiple-input multiple-output (MIMO) technology which has boosted the throughput to meet the growing demand of wireless data traffic. Using the Orthogonal Frequency Division Multiplexing (OFDM) physical layer waveform, WiFi standards such as IEEE 802.11n provide channel state information (CSI) for each received packet. The WiFi signal emitted from the commodity routers interacts with surrounding objects and humans which are present in the propagation environment. Furthermore, CSI captures the changes in wireless characteristics of the environment. Although CSI is included in WiFi since IEEE 802.11n, it is typically not reported by off-the-shelf WiFi Network Interface Cards (NICs). Therefore, dedicated tools need to be used to extract the CSI from the NICs. For example, the Atheros [1] and Linux 802.11n [2] CSI tools and custom modified firmware and open source Linux wireless drivers at the end-user equipment, CSI has been successfully extracted from specific NICs such as Intel 5300 (IWL5300) and some Atheros chipsets.

Using WiFi signal for human sensing has gained much attention over the past decade. Compared to other technologies such as cameras [3], [4] and wearable sensors [5], [6], WiFi based sensing technology is considered as an ideal solution

for future smart homes and healthcare residences since it is not intrusive, provides a wider coverage, is not sensitive to lighting conditions and the users do not need to wear any uncomfortable devices, making this technology completely passive.

The CSI obtained from commodity WiFi devices can be used for multiple indoor monitoring purposes like activity [7]–[9] and gait recognition [10], fall detection [11], [12], gesture [13], [14] and sign recognition [15], human/intrusion detection [16], crowd counting [17], among other applications. Note that there are schemes that use specialised hardware such as Universal Software Radio Peripheral (USRP) [14] and directional antennas [8] to obtain fine-grained signal measurements from WiFi signals. For instance, WiHear [8] makes use of directional antennas to obtain CSI variations caused by the movement of the lips for recognising spoken words. WiHear uses directional antennas to reduce the noise in the CSI data to achieve acceptable accuracy. WiSee [14] uses USRP to capture the OFDM signals and measures the Doppler shift in the signals caused by the reflection from human bodies to recognise nine gestures with an accuracy of 95%. E-eyes [9] uses CSI histograms as fingerprints for recognising nine daily human activities such as brushing teeth, taking shower, washing dishes, walking, etc., with an accuracy of 92%. It detects the activities of a single person using location-oriented time domain features. CARM [7], on the other hand, requires no specialised hardware to achieve high activity recognition accuracy ($\approx 96\%$). It classifies activities performed by a single person such as falling down, walking and sitting down using time-frequency features such as torso and limb velocities by characterising the changing speed of the reflected path length on subcarrier amplitudes. WiFiU [10] extends CARM to recognise human gait from limb and torso velocities. WiFiU can detect a walking human subject at a range as long as 14 meters with an accuracy of 92%.

Compared to previous studies such as [18], [19], this work makes the following contributions:

- We leverage the use of commodity WiFi devices with no specialised hardware for activity recognition in different physical layouts, covering both Line-of-Sight (LoS) and non LoS scenarios.
- The participants performed the activities in a random fashion or different orientations with respect to the transmitter/receiver (not controlled experiment). This is more representative of the real-world scenario.

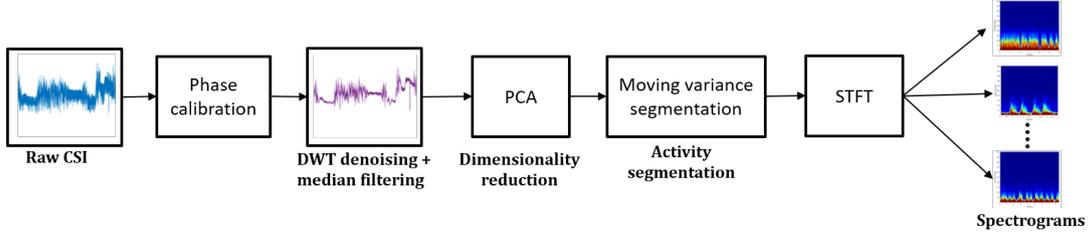


Fig. 1. Block diagram of signal processing techniques applied to CSI data.

- The experimental results evaluate the performance of the WiFi CSI system and identify the layout and coverage sensitivities. The results provide a benchmark for the expected accuracy in different physical transmitter-receiver geometry at different positions.

In this paper, we focus on activity recognition using WiFi CSI. The CSI data was collected alongside video recording during the whole experiment to provide ground truth labels. Both the CSI data and video were timestamped using the Network Time Protocol (NTP) Servers for synchronisation purposes. We demonstrate the detection result of the system for six different activities performed by five participants namely, sitting down, standing up, walking, laying down, standing from the floor and picking up an object from the ground.

The rest of the paper is organised as follows. An overview of the WiFi CSI system is given in Section II. The signal processing for the CSI system is described in Section III. The system implementation and experiment description are provided in Section IV. The activity recognition performance is evaluated in Section V. Finally, conclusions are drawn at the end of this paper.

II. OVERVIEW OF WiFi CSI

OFDM is a widely used technology in many WiFi standards including IEEE 802.11 a/g/n/ac. In an OFDM system, the bandwidth is shared among multiple overlapping but orthogonal subcarriers (subchannels) and therefore they do not interfere with each other. Since the subcarriers have a small bandwidth, they experience only flat fading. Thus, OFDM provides robustness in a frequency-selective fading wireless channel. For a WiFi system with MIMO-OFDM capability, in each packet the CSI is obtained as a 3-dimensional (3D) matrix with $n_t \times n_r \times N$ complex values, where n_t is the number of transmit antennas, n_r is the number of receive antennas and N is the number of subcarriers. Assuming a narrowband flat-fading channel, the frequency domain input-output relation in a MIMO-OFDM system with N subcarriers can be represented as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k, \quad k = 1, 2, \dots, N \quad (1)$$

where

$$\mathbf{x}_k = [x_1, x_2, \dots, x_{n_t}]^T, \quad (2)$$

$$\mathbf{y}_k = [y_1, y_2, \dots, y_{n_r}]^T, \quad (3)$$

$$\mathbf{z}_k = [z_1, z_2, \dots, z_{n_r}]^T, \quad (4)$$

are the transmitted signal, received signal and noise vectors, respectively. For a MIMO system, \mathbf{H}_k has a dimension of $N_r \times N_t$ and is represented as

$$\mathbf{H}_k = \begin{bmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,n_t} \\ h_{2,1} & h_{2,2} & \dots & h_{2,n_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_r,1} & h_{n_r,2} & \dots & h_{n_r,n_t} \end{bmatrix}, \quad (5)$$

where $h_{i,j}$ is the complex-valued channel coefficient between the j th transmit antenna and i th receive antenna. To recover the transmitted signal in a wireless medium, the receiver needs to estimate the channel. In this regard, a training sequence which is known by both the transmitter and receiver is sent in each packet to obtain the channel estimate. This process is often referred to as channel sounding. For a MIMO-OFDM system, the channel estimate (i.e., CSI) is a matrix consisting of complex values for each subcarrier, as in Equation (5). The equaliser uses the CSI to reverse the effects of the channel such as multipath propagation, attenuation, phase shift, scattering, etc., to recover the transmitted signal. In order to have an insight on the equalisation process, consider a simple single-input single-output (SISO) OFDM system, i.e., $n_t = n_r = 1$. Let the received frequency domain signal, following the Fast Fourier transform (FFT) operation at the OFDM receiver, be denoted as

$$Y = HX + Z. \quad (6)$$

In 802.11n, out of the 64 subcarriers that occupy the 20 MHz bandwidth, 4 subcarriers are used as pilots and 52 as data-carrying subcarriers. The remaining subcarriers are unused (they serve as DC/null subcarriers). As for the 40 MHz bandwidth, out of a total of 128 subcarriers, only 114 subcarriers are used, out of which 6 serve as pilots. The channel estimate at each pilot position can be obtained as

$$\hat{H}_p(k) = \frac{Y_p(k)}{X_p(k)}, \quad p = 1, 2, \dots, N_p \quad (7)$$

where p is the pilot index, N_p is the total number of pilot subcarriers, $\hat{H}_p(k)$ is estimated the channel frequency response (CFR) at the pilot subcarriers, X_p and Y_p are the transmitted and received pilot signals, respectively. In order to estimate the channel at the data subcarriers, the receiver performs channel interpolation.

III. SIGNAL PROCESSING FOR ACTIVITY SENSING

CSI measurements in the time domain capture the changes in the wireless signal due to the latter's interaction with surrounding objects or human activities and the observed patterns can be used for various purposes. Different WiFi sensing applications have specific requirements in terms of their signal processing techniques and classification/estimation algorithms. This section presents the signal processing techniques applied to the raw CSI data, as illustrated in Fig. 1.

A. Phase Calibration

In WiFi systems, the true phase information in the raw CSI measurements is corrupted due to hardware errors. These include Sampling Frequency Offset (SFO), Packet Detection Delay (PDD) and Carrier Frequency Offset (CFO) [1]. In this work, the phase is calibrated as in [20] where a linear transformation is applied to the raw phase data to eliminate the phase offset. The measured phase $\hat{\phi}_i$ of the i th subcarrier be expressed as:

$$\hat{\phi}_i = \phi_i - 2\pi \frac{k_i}{N} \delta t + \beta + Z, \quad (8)$$

where ϕ is the true phase, β is the phase offset due to CFO, δt is the timing offset between the transmitter and receiver, k_i is the index of the i th subcarrier and Z is noise. In the Intel 5300 CSI tool, $i \in \{1, 30\}$ and N is the FFT size. For example, $N = 64$ for a 20 MHz WiFi channel in IEEE 802.11 a/g/n. The terms δt , β and Z make it difficult to obtain the true phase from WiFi NICs. The phase obtained from the raw CSI measurements is corrected by first unwinding it and then applying a linear transformation. The main idea is to remove the terms δt and β by considering the phase across the whole frequency band [20].

B. Noise Reduction

Since the raw CSI data is too noisy to be directly used in machine learning algorithms for activity sensing, we use the Discrete Wavelet Transform (DWT) technique to filter out in-band noise and preserve the high frequency components, thus avoiding significant distortion to the signal. DWT-based noise filtering consists of transforming the signal into the wavelet domain whereby the signal is divided into several frequency levels called wavelets. More specifically, the signal is passed through a set of high pass and low pass filters at each level. The output from the high pass and low pass filters provides the detailed and approximation coefficients, respectively [11]. The detailed coefficients in the first level contains information about the noise and the abrupt changes caused by human activity. Therefore, the detailed coefficients in the first level are used to compute a threshold. The latter is adapted for lower wavelets and the noise is removed in all levels without introducing significant distortion to the signal. In addition to DWT denoising, 1-D median filtering is also applied to the signal to remove any undesired transients in the signal which are not caused by human motion.

C. Data Compression

The raw CSI measurements were collected on a device equipped with the Intel 5300 NIC. Considering one transmit and three receive antennas, we obtain $1 \times 3 \times 30 = 90$ complex CSI values in each packet. The packet rate was set at 1 kHz. This results in a significant amount of data that needs to be processed and which will serve as input to a learning algorithm for classification. Therefore, dimensionality reduction is required to reduce the computational complexity. In this work, the PCA dimensionality reduction and denoising technique has been used. PCA is used to identify the time-varying correlations between the CSI streams which are then optimally combined to extract components that represent the variation caused by human activities. The number of principal components (PCs) is empirically selected to achieve a good trade-off between classification performance and computational complexity [10]. Following DWT denoising, the first two or three PCs are, on average, sufficient to capture most of the variance (≈ 70 -80%) in the original 30 subcarriers [11]. Similar to [10], in the CSI system we extract the first six PCs. However, the first one is safely discarded since it contains noise due to reflection from stationary objects like furniture, walls etc., and therefore discarding it will not result in any loss of information [7], [10], [17]. Therefore, only the next five PCs are retained for further processing.

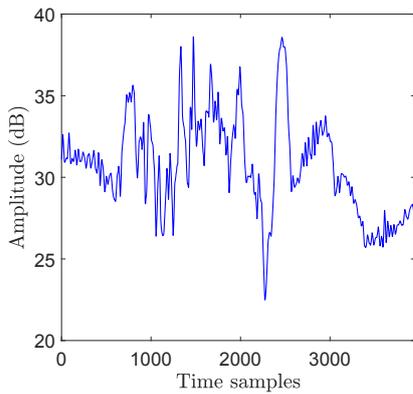
D. Moving Variance Segmentation

In this step, the meaningful CSI variation caused by a human activity is segmented using the Moving Variance Segmentation (MVS) [21] method. The main idea lies in computing the moving variance along the entire CSI stream in a stepwise manner. More specifically, in each step, a moving variance is calculated over a sliding window of length L across neighbouring CSI samples (packets), and the window is centred about the CSI sample in the current position. A high moving variance value represents significant variation in the CSI stream due to human motion, whereas a low value represents only slight fluctuations as in a stable environment. For a CSI stream consisting of M packets, the moving variance is defined as

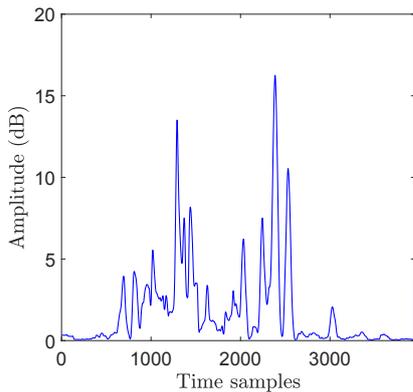
$$\text{CSI}_{mv} = \sum_{m=1}^M \left[\frac{1}{L-1} \sum_{l=1}^L |\text{CSI}_{l \in L} - \mu|^2 \right], \quad (9)$$

$$\mu = \frac{1}{L} \sum_{l=1}^L \text{CSI}_l,$$

where μ and l denote the mean and packet number in the sliding window of length L , respectively, and m is the current sample position in the CSI stream. An example of the CSI stream for the standing activity and its corresponding moving variance stream is shown in Fig. 2, where a sliding window of length $L = 100$ is considered. The sliding window length L needs to be chosen empirically to obtain the best results [21]. As can be observed in Fig. 2(b), the meaningful variations due to human motion become greater in magnitude while the insignificant fluctuations (not caused by human activity)



(a)



(b)

Fig. 2. MVS for standing activity (a) CSI stream (b) Corresponding moving variance sequence.

become smaller. Therefore, the start and end points of an activity can be easily identified and thus segmentation can be performed to improve the system's performance.

E. Spectrogram Generation

The CSI is highly sensitive to the surrounding environment and signal reflections from the human body result in different frequencies when performing different activities. These frequencies can be distinguished in the time-frequency domain (spectrogram) by applying the Short-time Fourier Transform (STFT) to the segmented PCA-denoised signal. Basically, the STFT applies a sliding window to obtain equally-sized segments of the signal and then performs FFT on the samples in each segment. The spectrogram has three dimensions, namely, time, frequency, and FFT amplitude. The Doppler spectrogram from STFT identifies the change of frequencies over time. The window size for FFT determines the trade-off between frequency and time resolutions. For instance, a larger window size results in a higher frequency resolution but lower time resolution. The spectrograms are generated from the five PCs which are then averaged to obtain the final spectrogram.

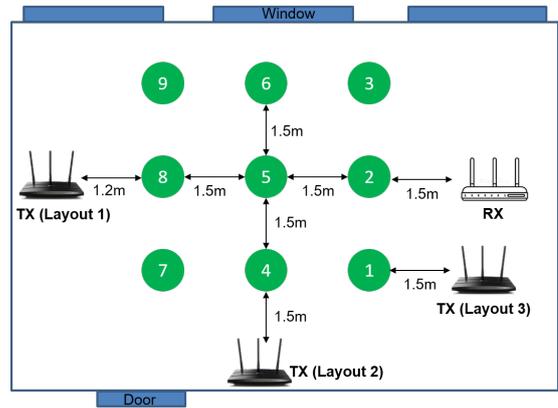


Fig. 3. Experiment layouts.

IV. SYSTEM IMPLEMENTATION & EXPERIMENT DESCRIPTION

A. System Implementation

The Intel 5300 [2] NIC has been used in our CSI system. Recall that the Intel 5300 Linux CSI tool extracts the CSI measurements from 30 out of 56 subcarriers for each transmit-receive antenna pair. The CSI data and video ground truth were synchronised to an external NTP server to provide timestamped measurements. The transmitter (TX) was a TP-Link access point (AP) and the receiver (RX) was an Intel Next Unit of Computing (NUC) device equipped with the Intel 5300 NIC, from which CSI was extracted and stored for off-line processing. The CSI was collected over 3 antennas (30 subcarriers each) on the receiver side (NUC) in the 2.4 GHz band (20 MHz bandwidth) by pinging the AP at a rate of 1 kHz. This rate was selected to capture noticeable changes or patterns in the time domain signal which are caused by human motion.

B. Experiment Layouts

The CSI data collection was carried out in an office space. The experiment layouts are shown in Fig 3. The monitoring area was approximately 8 m × 6 m with computers, chairs, cabinets and desks in the surroundings. The receiver (NUC) location was kept unchanged while the WiFi AP (TX) was moved in each layout as per Fig. 3, where 3 different layouts were considered. In Layout 1, the TX and RX were facing each other (LoS) while in Layout 2, the TX and RX were at 90° to each other. Finally, in Layout 3 the TX and RX were co-located (placed next to each other). 9 testing positions were used during the experiments and they were separated by 1.5 m from each other. These points are used to evaluate the effect of the system geometry on the activity classification accuracy. In this study, we conducted 6 activities, namely, walking, standing from a chair, sitting on a chair, laying down on the floor, standing from the floor and picking a small object from the floor. The activity descriptions are given in Table I. 5 participants of different age groups performed the 6 activities,

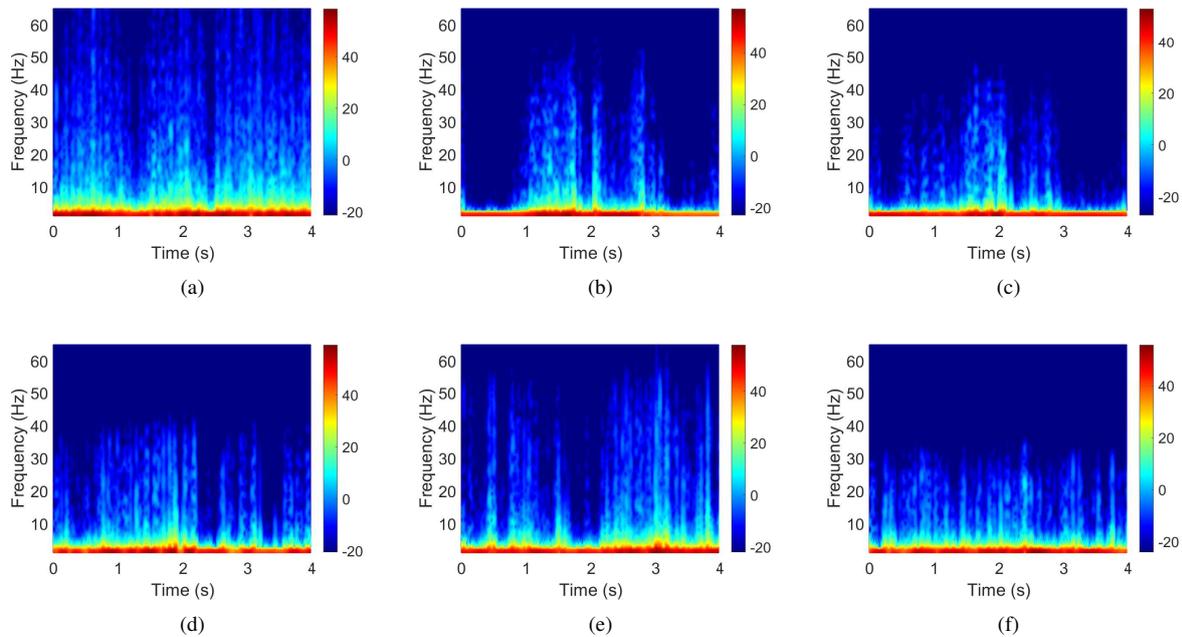


Fig. 4. Layout 2 spectrograms: (a) walking, (b) sitting, (c) standing, (d) laying down, (e) standing from floor, (f) picking up.

TABLE I
EXPERIMENT DESCRIPTION.

	Activity	Description
1	walking	walking along positions 1-2-3, 4-5-6, 7-8-9, 1-4-7, 2-5-8 and 3-6-9
2	sitting	sitting on a chair at positions 2,4,5,6,8
3	standing	standing from a chair at positions 2,4,5,6,8
4	laying down	laying down on the floor at positions 2,4,5,6,8
5	standing from floor	standing from the floor at positions 2,4,5,6,8
6	picking up	picking up objects from the floor at positions 2,4,5,6,8

one at a time, at the various positions for each layout. Note that the activities were performed in a random fashion or different orientations in a natural way, as would be the case in the real-world scenario.

V. EXPERIMENTAL RESULTS

In this section, the performance of activity recognition in different layouts is presented. Fig. 4 presents the spectrograms for the 6 activities in the Layout 2 configuration. As observed in Fig. 4, the activities that involve rapid body motion such as walking have high energy in the higher frequencies in the spectrogram. It is to be noted that the spectrograms contain information about instantaneous frequency (positive) changes but not motion direction, unlike radar systems. However, the frequency-domain features obtained from a CSI spectrogram are usually enough to identify a given activity. In this work, instead of manually identifying a set of features, we use a 2D Convolutional Neural Network (CNN) for classification purposes. The architecture of the 2D CNN is shown in Fig.

TABLE II
CLASSIFICATION PERFORMANCE FOR EACH LAYOUT.

Layout	Precision	Recall	F1-score	Accuracy
1	90.0%	89.5%	89.1%	90.8%
2	73.9%	74.4%	73.8%	75.7%
3	62.6%	61.9%	61.0%	61.5%
1, 2, 3	67.5%	66.3%	66.0%	67.3%

5. The network consists of a convolutional layer with 64 filters and 2×2 kernel size and the rectifier as the activation function. The next layer is the max-pooling layer with a stride of 2. The output from the max-pooling layer is flattened to create a single long feature vector (1D). The latter is fed to the first fully-connected layer consisting of 64 filters and using rectifier as the activation function. The second fully-connected layer consists of 32 filters with the same activation function as the previous layer. Finally, a softmax layer is used for classification of the 6 classes of activities. 80% of the dataset was randomly chosen and used for training while the remaining 20% was used for testing. As shown in Fig. 6(a), the overall accuracy for the case when data from all layouts are combined is 67%. The reason for the low accuracy is because the LoS and non-LoS data from the different layouts are mixed together and these have different Doppler signatures. The best classification result is obtained for activity 1 (walking) which is more than 90%. This is because the walking activity has higher Doppler shifts than other activities irrespective of the direction or layout. The second best result is observed for activity 6 (picking up) which is more than 70%. The incorrect predictions most happen between the pair of activities such as sitting and standing from a chair, laying down and standing from the floor. The confusion matrix for Layout 1 (LoS) is

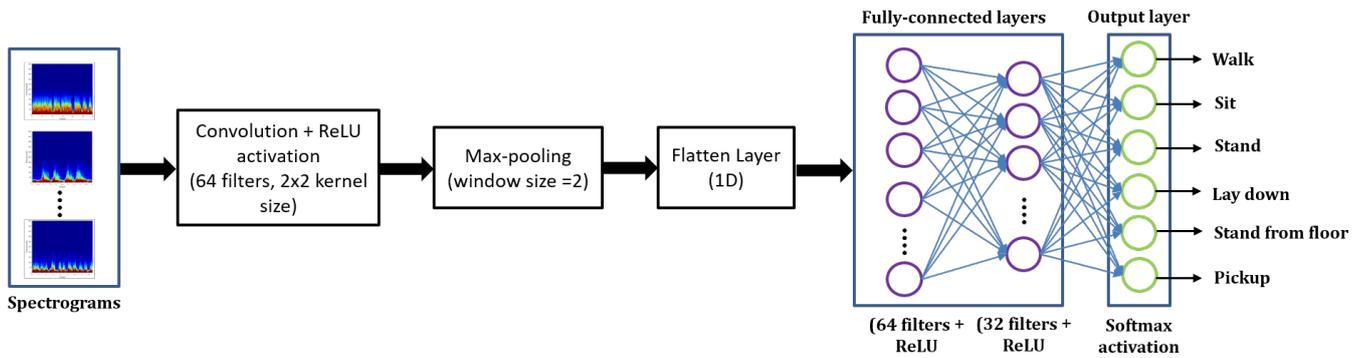


Fig. 5. 2D CNN for activity recognition classification.

also shown in Fig. 6(b). An accuracy around 91% is achieved in this case and this is in line with previous studies such as [7], [10], where an overall accuracy above 90% is obtained for various activities. We also provide the precision, recall and F1 scores for the different layouts in Table II.

As expected, the best performance is obtained in Layout 1 while Layout 3 has the worst performance at around 62%. Layout 2 results in an overall accuracy around 75% which is more than the accuracy in Fig 6(a).

Next, we evaluate the position accuracy as shown in Fig. 7. For the position test, the data is tested for a specific position and trained for all other positions (data from all layouts considered). The walking activity is excluded from this test since three positions were covered while the participant was walking. As can be observed in Fig. 7, positions 4 and 8, have the worst performance where their accuracies are below 60%. The best performance is achieved at position 5 which is the centre position. At positions 4 and 8, the person performing the activity is very close to the transmitter, especially in layouts 1 and 2. Part of the body of the person blocks most of the signals. Therefore the captured direct and reflected WiFi signals at the receiver are not embedded with information related to the movements of the whole human body when performing an activity at these positions.

VI. CONCLUSION

The scope of this study was to evaluate the activity recognition performance in different physical geometry as would be the case in a real-world environment. As expected, the best detection performance, with an accuracy of 91%, is achieved when activities are performed in between the TX and RX in a LoS setup (Layout 1). When the TX and RX are 90° to each other (Layout 2), an accuracy of 75% is achieved. As for Layout 3 where the TX and RX are co-located (non LoS), the accuracy is 62%. The data from all the layouts resulted in an overall combined accuracy of 67%. Finally, it was found that the centre position achieved the highest accuracy in the position test. From the results, it can be inferred that the activity detection performance is dependent not only on the locations of the TX and RX but also on the positioning of the person performing the activity.

ACKNOWLEDGEMENTS

This work was funded under the OPERA Project, the UK Engineering and Physical Sciences Research Council (EPSRC), Grant EP/R018677/1.

REFERENCES

- [1] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity Wi-Fi," *IEEE Transactions on Mobile Computing*, vol. 18, no. 6, pp. 1342–1355, 2018. [Online]. Available: <https://wands.sg/research/wifi/AtherosCSI/>
- [2] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, pp. 53–53, 2011. [Online]. Available: <https://dhalperi.github.io/linux-80211n-csinfo/>
- [3] F. Chin-Shyung, S.-E. Lee, and M.-L. Wu, "Real-time musical conducting gesture recognition based on a dynamic time warping classifier using a single-depth camera," *Applied Sciences*, vol. 9, no. 3, p. 528, 2019.
- [4] R. Saini, P. Kumar, P. P. Roy, and D. P. Dogra, "A novel framework of continuous human-activity recognition using kinect," *Neurocomputing*, vol. 311, pp. 99–111, 2018.
- [5] S. Khawandi, B. Daya, and P. Chauvet, "Integrated monitoring system for fall detection in elderly," in *2012 International Conference on Audio, Language and Image Processing*, 2012, pp. 62–67.
- [6] C. Rodriguez, D. M. Castro, W. Coral, J. L. Cabra, N. Velasquez, J. Colorado, D. Mendez, and L. C. Trujillo, "IoT system for human activity recognition using bioharness 3 and smartphone," in *Proceedings of the International Conference on Future Networks and Distributed Systems*, ser. ICFNDS '17. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: <https://doi.org/10.1145/3102304.3105828>
- [7] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Proceedings of the 21st annual international conference on mobile computing and networking*, 2015, pp. 65–76.
- [8] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with Wi-Fi!" *IEEE Transactions on Mobile Computing*, vol. 15, no. 11, pp. 2907–2920, 2016.
- [9] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-Eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures," in *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 617–628. [Online]. Available: <https://doi.org/10.1145/2639108.2639143>
- [10] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 363–373. [Online]. Available: <https://doi.org/10.1145/2971648.2971670>

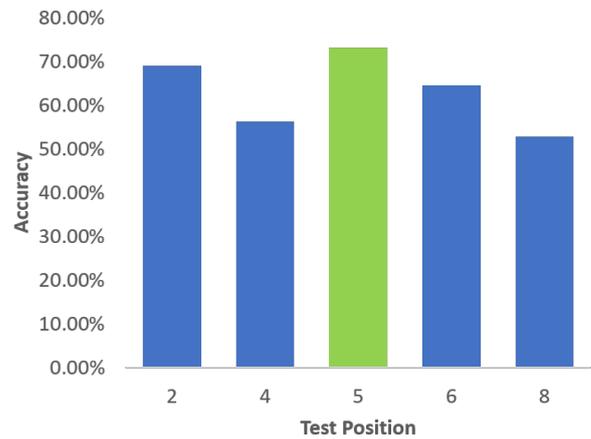
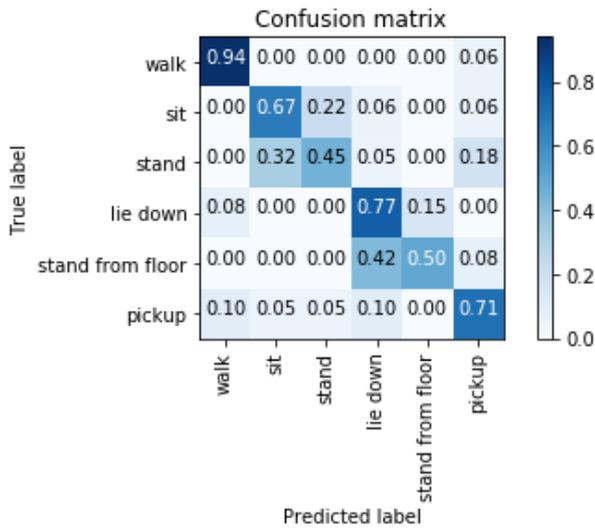


Fig. 7. Position test classification accuracy.

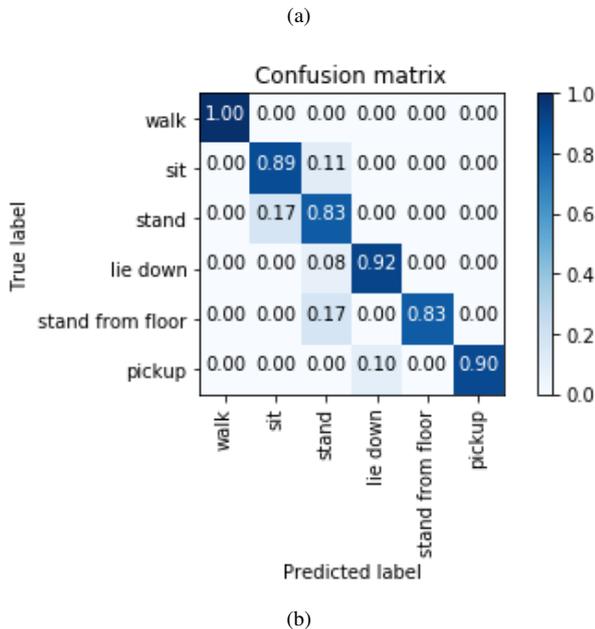


Fig. 6. Confusion matrix (a) Combined layouts 1, 2, 3 (b) Layout 1.

[11] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "FallDeFi: Ubiquitous fall detection using commodity wi-fi devices," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 4, Jan. 2018. [Online]. Available: <https://doi.org/10.1145/3161183>

[12] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "RT-Fall: A real-time and contactless fall detection system with commodity wifi devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 511–526, 2017.

[13] S. Tan and J. Yang, "WiFinger: leveraging commodity wifi for fine-grained finger gesture recognition," in *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 201–210. [Online]. Available: <https://doi.org/10.1145/2942358.2942393>

[14] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*, ser. MobiCom '13. New York, NY, USA: Association for Computing Machinery, 2013, pp. 27–38. [Online]. Available: <https://doi.org/10.1145/2500423.2500436>

[15] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "SignFi: Sign language recognition using wifi," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 1, Mar. 2018. [Online]. Available: <https://doi.org/10.1145/3191755>

[16] E. Soltanaghaei, A. Kalyanaraman, and K. Whitehouse, "Peripheral WiFi vision: Exploiting multipath reflections for more sensitive human sensing," in *Proceedings of the 4th International on Workshop on Physical Analytics*, ser. WPA '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 13–18. [Online]. Available: <https://doi.org/10.1145/3092305.3092308>

[17] S. Liu, Y. Zhao, F. Xue, B. Chen, and X. Chen, "DeepCount: Crowd counting with WiFi via deep learning," 2019.

[18] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Device-free human activity recognition using commercial wifi devices," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 5, pp. 1118–1131, 2017.

[19] W. Li, R. J. Piechocki, K. Woodbridge, and K. Chetty, "Physical activity sensing via stand-alone WiFi device," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[20] X. Dang, X. Tang, Z. Hao, and Y. Liu, "A device-free indoor localization method using csi with wi-fi signals," *Sensors*, vol. 19, no. 14, p. 3233, 2019.

[21] Z. Dong, F. Li, J. Ying, and K. Pahlavan, "Indoor motion detection using Wi-Fi channel state information in flat floor environments versus in staircase environments," in *Sensors*, 2018.