



[Ge, Y.](#), [Wang, J.](#), [Li, S.](#), [Qi, L.](#), [Zhu, S.](#), [Cooper, J.](#) , [Imran, M. A.](#) and [Abbasi, Q. H.](#) (2023) WiFi sensing of Human Activity Recognition using Continuous AoA-ToF Maps. In: IEEE Wireless Communications and Networking Conference (WCNC2023), Glasgow, Scotland, UK, 26-29 March 2023, ISBN 9781665491228 (doi: [10.1109/WCNC55385.2023.10118954](https://doi.org/10.1109/WCNC55385.2023.10118954))

There may be differences between this version and the published version.
You are advised to consult the published version if you wish to cite from it.

<https://eprints.gla.ac.uk/287647/>

Deposited on 12 December 2022

Enlighten – Research publications by members of the University of Glasgow

<http://eprints.gla.ac.uk>

WiFi sensing of Human Activity Recognition using Continuous AoA-ToF Maps

Yao Ge¹, Jingyan Wang¹, Shibo Li^{1,2}, Liyuan Qi¹, Shuyuan Zhu², Jonathan Cooper¹
Muhammad Imran¹, Qammer H. Abbasi¹

¹University of Glasgow, James Watt School of Engineering, Glasgow, UK

²University of Electronic Science and Technology of China, Chengdu, China

Abstract—Joint communication and sensing technique has been adopted for smart home design and other applications recently. WiFi sensing, which utilizes mutually orthogonal channel response to monitor the changes in the medium, is regarded as one of key techniques in this field. Human activity recognition using wireless communication system is a key function of future internet of things system. The effective and inexpensive WiFi sensing system can help people for device-free controlling, healthcare monitoring without concern of image information leakage that uses camera system. In this article, we proposed a continuous angle of arrival and time of flight (AoA-ToF) maps based method that adopts multiple signals classification analysis on commercial and off-the-shelf WiFi devices to detect the human activities. The performance of our system achieves 85.6% accuracy in total with 8 activities among 5 users. Meanwhile, we investigated the performance of our system under different conditions, including direction and user identity. The results show the system’s robustness for human activity recognition under such conditions.

Index Terms—WiFi sensing, channel state information, multiple signal classification, deep learning

I. INTRODUCTION

Indoor human activity recognition (HAR) has played a significant role in intelligent internet of things (IoT) controlling and healthcare monitoring. Device-free WiFi sensing in HAR is an emerging research trend for nearly a decade, which spawned various kinds of applications including vitals monitoring, human daily activity recognition, falling detection, and signs recognition [1], [2]. Compared to other equipment of radar-based device-free approaches, the WiFi devices objectively own higher cost-effective which is the better choice for general indoor cases implementation.

Since the WiFi signals are approved to be accessible for wireless sensing, there are various kinds of HAR systems have been developed. WiFall system [3] is designed to detect falling activity for health monitoring using the stability of CSI time and diversity of frequency spectrum of WiFi signals. Lip motion detection system [4] is proposed for under-mask speech recognition, which utilizes WiFi signal’s variation while human speaking to classify different pronunciation. Eyes [5] leverages the distribution information of CSI amplitude for HAR tasks. All the above systems are focused on extracting temporal signals influenced by human activity but don’t utilize the spatial information brought by the WiFi antenna array. Widar3.0 [6] realizes an overall HAR system

that can be transferred to different environments, orientations, and locations. It integrates the Doppler frequency variation from multiple views of WiFi devices in distinct coordinates, with regard to human position. However, the system is only available when the scenario has more than three receivers.

Inspired by the Widar3.0 system, we proposed a continuous angle of arrival (AoA) and time of flight (ToF) maps-based method that adopt single pair of commercial WiFi devices to achieve fine-grained human activity recognition task. Our contributions are summarised in three points.

- 1) We firstly leverage the continuous AoA-ToF maps (CATM) in the HAR task as our best knowledge. Compared to the previous frequency spectrum-based classification method that only leverages temporary information, the CATM reserve both temporal and spatial human body information.
- 2) We design a lightweight temporal neural network for HAR tasks in our CATM system.
- 3) We collect and establish a dataset of human activity with WiFi sensing, and human skeleton information for validation.

The rest parts of the paper are structured as follows: Section II introduces the basic model of CSI transmission. Section III demonstrates the methods of signals calibration, feature extraction of CATM, and temporal neural network classification. Section IV and V analyse the overall performance and give the conclusion respectively. The structure of the proposed system is shown in Fig. 1.

II. PRELIMINARIES

Channel state information (CSI) is used to characterize the channel attributes of WiFi links in the physical layer as the sample of channel frequency response, which can be disturbed by multipath fading and shading. In the indoor environment, human movement is one of the significant factors that influence the CSI value. Meanwhile, CSI data contains human-related information, like daily movement and rhythmic physiological phenomena, like respiration and heartbeat. WiFi sensing technique aims to extract and analyze this potential human related information. We could use the equation to describe the CSI:

$$H(f, t) = e^{-j2\pi\Delta ft}(H_s(f) + H_a(f, t)) + N(t) \quad (1)$$

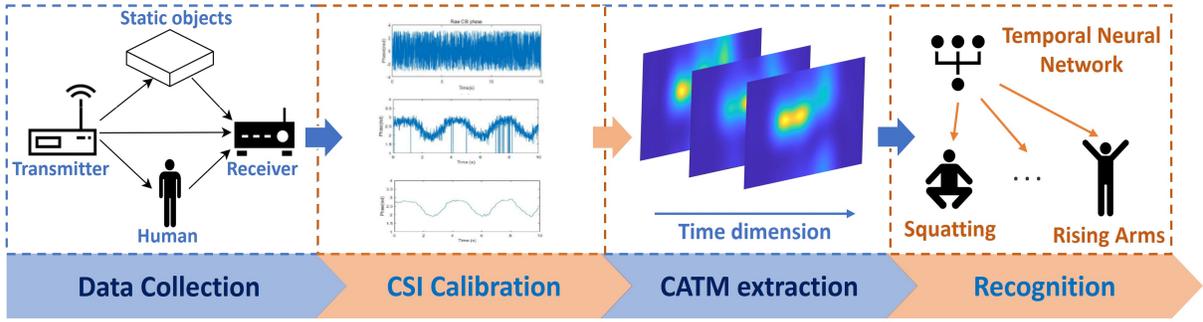


Fig. 1: CATM based HAR system of WiFi sensing

where $e^{-j2\pi\Delta ft}$ is the random phase shift due to the asynchronous sending and receiving process of the WiFi system; H_s and $H_a(f, t)$ represents the CSI signals from all the static paths (including the signals in line of sight (LOS) path and those reflected off the stationary objects) and active paths (including signals reflected from the dynamic objects) respectively. $N(t)$ represents the noise during transmission. CSI signals in active paths can be expressed as:

$$H_a(f, t) = \sum_{i=1}^{N_d} a_i(f, t) e^{-j2\pi \frac{(d_i(t) + d_a(t))}{\lambda}} \quad (2)$$

where N_d is the index of the dynamic path, $a_i(f, t)$ represents the complex attenuation factor of the i^{th} path; $e^{-j2\pi \frac{d_i(t)}{\lambda}}$ represents the phase change of i^{th} path; $d_i(t)$ and $d_a(t)$ are the static length and vibrating length of i^{th} path. λ represents the wavelength of the WiFi signal.

Through coupling the above two equations, it can be found that human related information is reflected from the amplitude and phase shift. However, previous researches found that the raw CSI data is too noisy to extract the feature [7], [8], especially for tiny motion. In this case, we designed a series of experiments of breathing detection to validate the performance of data sanitation techniques. The reason why we choose breathing instead of other activities is the respiration information is periodical and tiny enough, which make it easier to check the correlation between main frequency response and human motion.

III. METHODOLOGY

A. Calibration technique

At first we collected CSI data using Intel 5300 network interface cards (NIC) and omnidirectional antennas. A volunteer is asked to sit statically between a WiFi receiver and a WiFi transmitter and keep breathing evenly. The plots of raw CSI amplitude and phase are shown in Fig. 2. In comparison with CSI amplitude, CSI phase information of human body motion is more sensitive, especially for respiration detection [8], [9]. However, the random phase shift from asynchronous links pollutes informative phase variation as Fig. 2 shows. Phase sanitation is significant for the post-classification task of WiFi sensing. According to the previous analysis of WiFi

CSI transmission [10], noise components can be divided into sampling frequency offset (SFO) and packet detection delay (PDD) caused by the receiver:

- SFO: Clocks of transmitter and receiver are out of synchronization, which leads the sampling frequency of the receiver to shift randomly. This will cause the phase of collected signal to be unpredictable shifted.
- PDD: Detection of WiFi signals in receiver side cause the delay.

Next, the conjugation multiplication is adopted in our system to reduce the SFO which has been proved in IndoTrack [11]. It considers the phases of two transmission paths are shifted by the same value from the receiver. In this case, conjugate multiplication can help to reduce SFO and PDD. Secondly, due to the multi-path effect of WiFi signals, receivers can collect multiple impulse signals while the transmitter only sends them once. It shows the time dispersion of wireless transmission, which means the ToF should be various for one transmitted packet. Based on the line-of-sight path, we adopt the ToF-based filter to eliminate the high delay components of CSI signals. To explore the influence of ToF corresponding to the respiration signals, we adopt the power delay profile (PDP). The PDP method utilizes the influence of time delay due to the multi-path effect. PDP can be calculated by inverse fast Fourier transform (IFFT) of raw CSI data. Due to the multi-path effect, physical WiFi signals are not only transferred by the LOS path but also can be reflected by any objects around the circumstances, which is in NLOS paths. The signals with a long time delay can be regarded as uncertain data that carried more environmental information. Meanwhile, considering the sampling frequency, the time delay should be smaller than the time gap between the two samples. Otherwise, the signals will be attenuated by other samples' information in the hardware or transmission medium. In this case, we calculated the PDP of CSI signals and set the threshold of time to the sampling time to filter the long ToF signals out. The experiment result is shown in Fig. 3 tested in distinct subcarriers.

B. AoA-ToF Maps Construction

For our activity classification, the directional movement feature of the human body is the main parameter for researchers

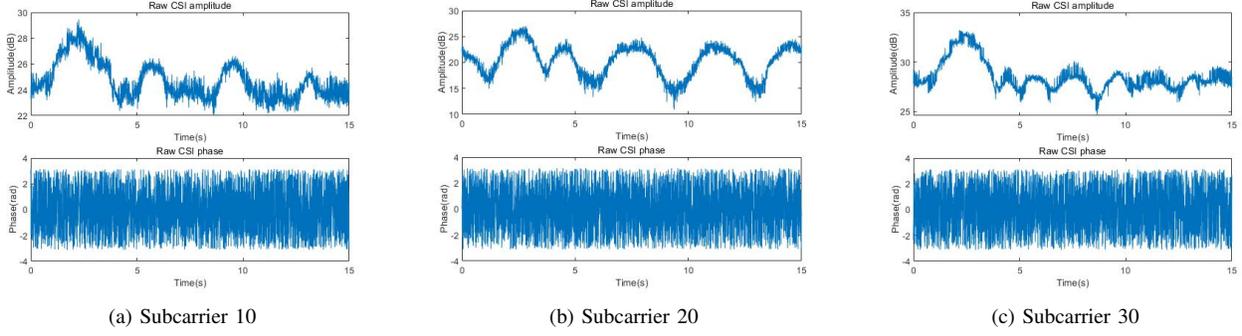


Fig. 2: Raw CSI data within volunteer keep breathing evenly for 15 seconds of different subcarriers

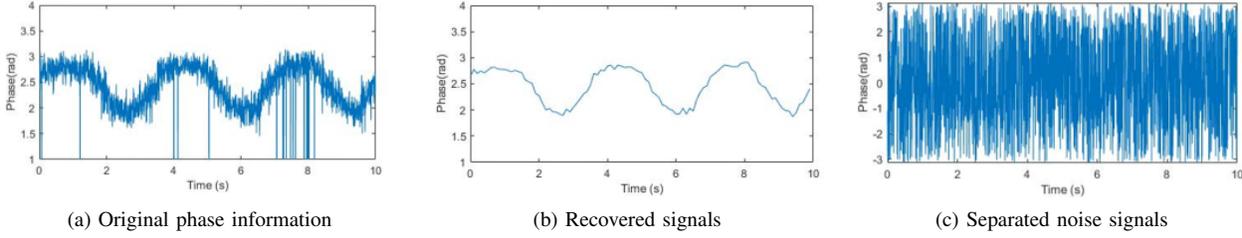


Fig. 3: The PDP filtered result of CSI.

to establish the mapping relationships from signals to multiple activities class [12]. However, for complex human motions, like waving hands (left and right side), the behavior patterns can be similar to each other. We assume that the human motion is capable to influence the channel parameters of Angle of arrival (AoA) and Time of flight (ToF). AoA indicates the direction of receiving signals with respect to the receiver and ToF indicates the time cost of transmission from the transmitter to receiver, by LOS and NLOS. In this case, we extract continuous AoA-ToF maps (CATM), which directly represent the behavior coordination in physical space.

Firstly, we regard the double antennas as an antenna array and analysis it with a Multiple Signal Classification (MUSIC) signal processing algorithm [13]. However, the MUSIC algorithm requires the number of transmission paths should be smaller than the number of antennas. In our setup, the number of antennas is limited to the number of ports in the commercial WiFi chip. In the equation, we separate the signals received by each antenna into 30 different groups based on the number of subcarriers of the chip. The separated frequency bands of subcarriers lead to frequency-selective fading. For different frequency components of the signal, the wireless transmission channel will show distinct random responses. According to the frequency-selective effect, we simulate a virtual antenna array composed of 30 independent subcarriers from a single antenna. On the other hand, in the linear antenna array, phase differences between two actual antennas are related to the time cost of receiving the data from the same path. Suppose the

index of antenna and subcarrier are M and K respectively.

$$\mathbf{X}_{CSI} = \begin{bmatrix} x_{csi}(1,1) & x_{csi}(1,2) & \cdots & x_{csi}(1,K) \\ x_{csi}(2,1) & x_{csi}(2,2) & \cdots & x_{csi}(2,K) \\ \vdots & \vdots & \ddots & \vdots \\ x_{csi}(M,1) & x_{csi}(M,2) & \cdots & x_{csi}(M,K) \end{bmatrix} \quad (3)$$

Next, we analyze the phase shift indicator with the AoA and ToF value of the individual path. We suppose the incident angle of receiving signals towards the linear array is θ .

$$\phi_\theta = \frac{2\pi(m-1)d\sin\theta}{\lambda} \quad (4)$$

We mainly adopt the MUSIC-based method which is proposed in [14]. Due to the frequency band f_θ between two adjacent subcarriers, there is a 2.5 radians phase shift. At l_{th} path, the phase shift of k_{th} subcarrier with ToF τ is given by:

$$\phi_\tau = 2\pi(k-1)f_\theta\tau \quad (5)$$

Besides, we leveraged the Smoothed-CSI matrix proposed in SpotFi [14] to increase the virtual unit and enhance the linear array association. The Eq. 6 shows the cases of $M = 2$ and $K = 30$. For simplicity, We use \mathbf{X}_{S-CSI} to call the smoothed CSI matrix.

$$\mathbf{X}_{S-CSI} = \begin{bmatrix} x_{csi}(1,1) & x_{csi}(1,2) & \cdots & x_{csi}(1,15) \\ x_{csi}(1,2) & x_{csi}(2,2) & \cdots & x_{csi}(2,16) \\ \vdots & \vdots & \ddots & \vdots \\ x_{csi}(1,15) & x_{csi}(1,16) & \cdots & x_{csi}(1,30) \\ x_{csi}(2,1) & x_{csi}(2,2) & \cdots & x_{csi}(2,16) \\ x_{csi}(2,2) & x_{csi}(2,3) & \cdots & x_{csi}(2,17) \\ \vdots & \vdots & \ddots & \vdots \\ x_{csi}(2,15) & x_{csi}(2,16) & \cdots & x_{csi}(2,30) \end{bmatrix} \quad (6)$$

In our system, we assume the phase difference of different subcarriers is ϕ_τ . Therefore, we replace $x_{csi}(m, k)$ with $e^{-j\phi_\tau}$ and $e^{-j\phi_\theta}$ to explore the phase shift. Meanwhile, we label the transmission path index as L . The steering vector \mathbf{a} and steering matrix \mathbf{A} can be written as:

$$\mathbf{a}(\theta, \tau) = [1, e^{-j\phi_\tau}, \dots, e^{-j\phi_\tau^{14}}, e^{-j\phi_\theta}, \dots, e^{-j(\phi_\theta + \phi_\tau^{14})}] \quad (7)$$

$$\mathbf{A} = [\mathbf{a}(\theta_1, \tau_1), \mathbf{a}(\theta_2, \tau_2), \dots, \mathbf{a}(\theta_L, \tau_L)]^\top \quad (8)$$

In the MUSIC algorithm, the received data is used for eigen-decomposition to separate the signal subspace and the noise subspace. The orthogonality of the signal direction vector and the noise subspace is used to form a spatially scanned spectrum for a full-domain search of the spectral peaks, which enables parameter estimation of the signal. To apply the algorithm, we firstly establish a linear model of a received signal vector of \mathbf{X} , in the presence of a steering matrix of \mathbf{A} , Gaussian white noise of \mathbf{N} , and Smoothed-CSI matrix of \mathbf{S} :

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \quad (9)$$

The basic idea of the MUSIC algorithm is to analyze the covariance matrix $\mathbf{R}_\mathbf{x}$ of received signals:

$$\begin{aligned} \mathbf{R}_\mathbf{x} &= \mathbb{E} \{ \mathbf{X}\mathbf{X}^H \} \\ &= \mathbf{A}\mathbb{E} \{ \mathbf{S}\mathbf{S}^H \} \mathbf{A}^H + \mathbb{E} \{ \mathbf{N}\mathbf{N}^H \} \\ &= \mathbf{A}\mathbf{R}_\mathbf{s}\mathbf{A}^H + \mathbf{R}_\mathbf{N} \end{aligned} \quad (10)$$

where $\mathbf{R}_\mathbf{s}$ and $\mathbf{R}_\mathbf{N}$ represents the correlation matrix of multi-path incident signals and noise respectively. In the K -size eigenvalues of correlation matrix, the smallest $K - L$ eigenvalues are corresponding to the noise and others eigenvalues correspond to L incident signals. The noise subspace can be constructed: $\mathbf{E}_\mathbf{n} = [\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{K-L}]$. Therefore, the power spectrum with respect to the AoA θ and doppler frequency dop can be expressed as:

$$\mathbf{P}(\theta, \tau)_{MUSIC} = \frac{1}{\mathbf{a}^H(\theta, \tau) \hat{\mathbf{E}}_N \hat{\mathbf{E}}_N^H \mathbf{a}(\theta, \tau)} \quad (11)$$

Besides, attributed to the short time delay of subcarriers on a single antenna and limitation of resolution in MUSIC, there are negative ToF value signal components shown in the matrix. Generally, the MUSIC algorithm only focuses on the peak of the power spectrum which reflects the sources'

direction. However, in our system, we set static location for each device, and extract the variation of peak influenced by human activity. Although it's not possible for ToF value to be negative in real world, we adopt these parts and suppose the shape of heatmaps in the negative region is helpful for neural network classification. In the discussion of Section. IV-B4 we compare the influence on performance between non-negative ToF CATM (NT-CATM) and full CATM.

There are two activities, i.e. walking and jogging, that are posted in the time domain as Fig. 4. Because the patterns of these human movements are similar, it's can be proved that the effectiveness of feature extraction with the similarity of two CATM. The red boxes represent the similarity between the two activities, which shows the robustness via different time slots.

C. Temporal classification neural network structure

The structure of our proposed neural network is illustrated in Fig. 5. It contains 3 components: frame embedding, sequential fusion, and classification. The frame embedding consists of one convolutional layer, one residual block, and one pooling layer. We implement the residual block following the function in [15], where an identity mapping is introduced for faster back propagation of gradient. The following is the sequential fusion, which intends to collect the information from the temporal sequence. In the module, we adopt an LSTM structure, which has a robust ability to retain the long-term information in time sequence. In the final step, a simple linear projection with SoftMax is utilized as the classification module that generates the probability distribution of predicted motions. The model is trained by CrossEntropy loss and Adam optimizer.

IV. EVALUATION

A. Experiment Setup

We adopted four PCs equipped with Intel 5300 COTS WiFi devices to sense 8 different activities of 5 volunteers. As Fig. 6(d),(h) shows, we set up 1 transmitter and 3 receivers to collect CSI data. Meanwhile, we equipped a Kinect V2 depth camera by the side of the transmitter to capture the human skeleton coordinates, as Fig. 4 shows. For CSI data collection, each profile contains 3-second human activity with a sampling frequency of 1000 Hz.

After CSI calibration and CATM extraction, the dimension of each profile is set to $S \times H \times W$, where S represents the length of the sequence in one packet, H and W both refer to the dimension of a single ATM. We set the size of CATM profile to $100 \times 90 \times 90$. In total, we collected 7998 profiles. To standardize the training step, we set 100 epochs with 32 batch sizes for our system and other comparison tests. The whole dataset is randomly picked 20% of profiles as test set.

B. Overall evaluation and Discussion

The overall classification results are composed of 3 pairs of Tx and Rx, which are calculated to be 85.8%, 85.4%, and 86.8% respectively. For simplicity, we classify the data

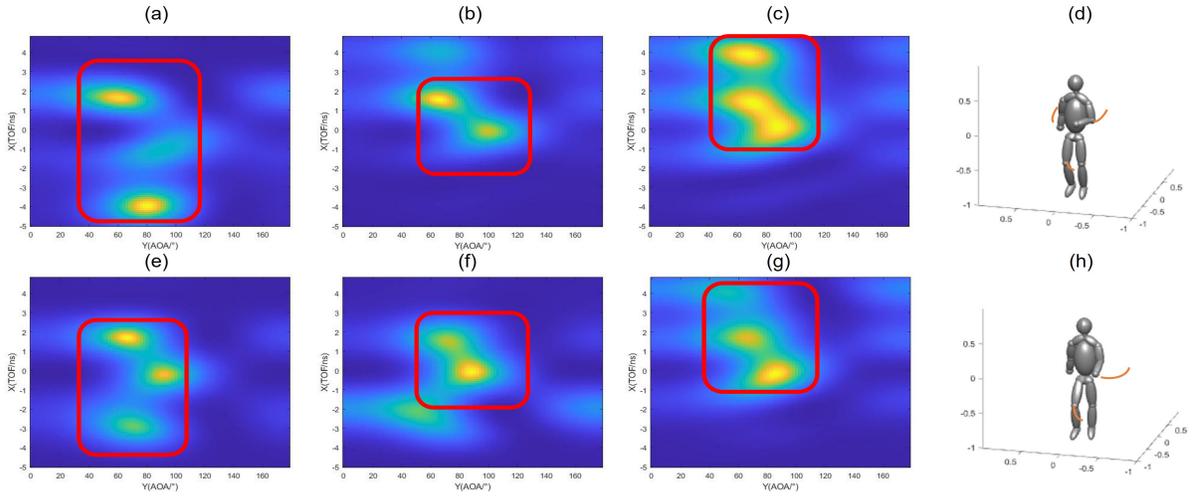


Fig. 4: ATM result of CSI while human jogging (a,b,c) and walking (e,f,g).

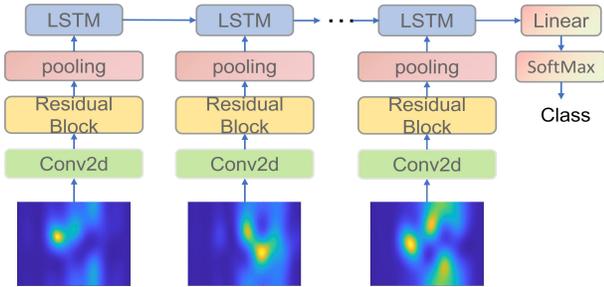


Fig. 5: Res-LSTM Temporal neural network structure

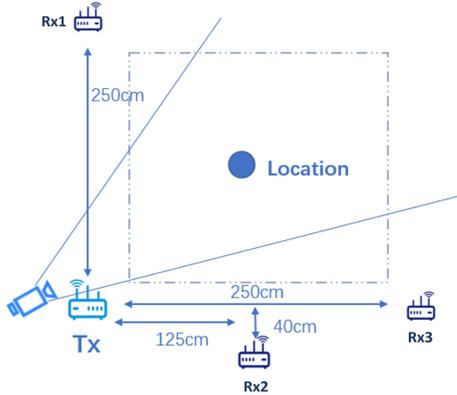


Fig. 6: Experimental setup

collected from the different paths into 3 directions. For example, D1 represents the first pair of Tx and Rx, with the same for others. To evaluate the factors that influence system performance, we analyze the system with a control variable method of different direction paths and user identity.

1) *Influence from direction*: From the accuracy comparison among the 3 directions, we observe that the performance of the classification system of D1 and D3 works slightly better than

TABLE I: Recognition accuracy vs Activity type

	Accuracy (%)			Total
	D1	D2	D3	
Walking	90.3	88.2	87.1	88.5
Jogging	78.6	72.4	79.3	76.7
Squatting	91.4	82.9	84.4	86.3
Chest expansion	83.9	82.4	96.2	86.8
Rising arms	84.4	89.5	83.3	85.7
Lunge press (Right)	85	82.8	91.2	86.4
High-step stretch	86.1	88.2	86.8	87
Lunge press (Left)	85.3	94.1	80	86.4
Total	85.8	85.4	86.8	85.6

D2. We suppose that the LOS path components of D2 cause a stronger effect on our CATM, such as static LOS response.

2) *Influence from user identity*: For distinct users, the behavior habits are various. To validate the performance of our system among different persons, we invite 5 volunteers to participant in our experiments and compare their results to the dataset. The performance is shown in the Table. II. Identities indeed influence the performance with a maximum gap of 12.8%.

3) *Comparison of other HAR systems*: We replicate and compare other three different WiFi HAR systems to validate the performance of our work: Widar3.0 [6], THAT [16], Dop_HAR [17]. The comparison results are illustrated in Fig. 7. The Widar3.0 system gets the highest HAR accuracy of 91.4% compared to ours CATM of 86.8% in D3. However, to adopt Widar3.0 system, there should be at least 3 WiFi receivers for establishment. In our system, only one receiver is needed. From the comparison results with other two systems, our CATM system gains the highest classification accuracy.

4) *Comparison of CATM and NT-CATM*: In our comparison test, we reproduce the NT-CATM described in

TABLE II: Recognition accuracy vs User identity

	Accuracy (%)				
	U1	U2	U3	U4	U5
Walking	85.7	95.2	77.8	95.5	87.5
Jogging	85	83.3	77.8	66.7	70.6
Squatting	83.3	95	73.7	95	84
Chest expansion	91.3	94.1	78.9	85	83.3
Rising arms	90.1	92	90	91.3	63.6
Lunge press (Right)	90.5	95	79.2	81.3	86.4
High-step stretch	92.6	100	83.3	93.3	70.4
Lunge press (Left)	80	77.8	93.8	92.3	87
Total	87.8	91.9	81.6	88.5	79.1

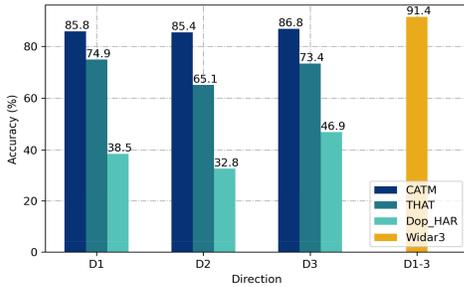


Fig. 7: Accuracy comparison of different WiFi HAR systems in 3 directions

Section.III-B. The results in Fig. 8 validate our assumption that negative ToF values in CATM contain the effective information related the human behaviors.

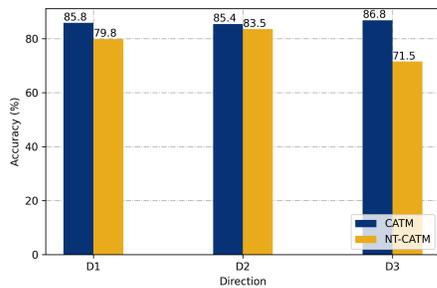


Fig. 8: Accuracy comparison of CATM and NT-CATM in 3 directions

V. CONCLUSION

This paper proposes a CTAMs based system for human activity monitoring in the indoor environment, which adopts spatial phase variation information of CSI. In this system, we firstly analyzed the noise components in raw CSI data and leverage various techniques to filter the phase offset of CSI information for data sanitation. Then, we extracted the features with continuous AoA-ToF maps. Finally, we established a temporal neural network for the HAR task and achieved 85.6%

accuracy among 5 volunteers and 8 activities. The comparison results demonstrate the ubiquitous performance under the conditions of different directions and user identities. Meanwhile, from the comparison tests with other WiFi HAR systems, we get the highest accuracy in the systems considering single WiFi receiver.

ACKNOWLEDGEMENTS

This work was supported in parts by Engineering and Physical Sciences Research Council (EPSRC) grants EP/T021020/1 and EP/T021063/1.

REFERENCES

- [1] Y. Ge, A. Taha, S. A. Shah, K. Dashtipour, S. Zhu, J. M. Cooper, Q. Abbasi, and M. Imran, "Contactless wifi sensing and monitoring for future healthcare-emerging trends, challenges and opportunities," *IEEE Reviews in Biomedical Engineering*, 2022.
- [2] M. Usman, J. Rains, T. J. Cui, M. Z. Khan, M. A. Imran, Q. H. Abbasi *et al.*, "Intelligent wireless walls for contactless in-home monitoring," *Light: Science & Applications*, vol. 11, pp. 212–224, 2022.
- [3] Y. Wang, K. Wu, and L. M. Ni, "WiFall: Device-Free Fall Detection by Wireless Networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2017.
- [4] H. Hameed, M. Usman, A. Tahir, A. Hussain, H. Abbas, T. J. Cui, M. A. Imran, and Q. H. Abbasi, "Pushing the limits of remote rf sensing by reading lips under the face mask," *Nature Communications*, vol. 13, pp. 5168–5176.
- [5] 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, 2014, p. 617–628.
- [6] Y. Zhang, Y. Zheng, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, "Widar3. 0: Zero-effort cross-domain gesture recognition with wi-fi," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.
- [7] F. Wang, S. Panev, Z. Dai, J. Han, and D. Huang, "Can WiFi estimate person pose?" *arXiv:1904.00277*, no. March, 2019.
- [8] D. Zhang, Y. Hu, Y. Chen, and B. Zeng, "Breathtrack: Tracking indoor human breath status via commodity wifi," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3899–3911, 2019.
- [9] X. Wang, C. Yang, and S. Mao, "Tensorbeat: Tensor decomposition for monitoring multiperson breathing beats with commodity wifi," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 9, no. 1, pp. 1–27, 2017.
- [10] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-invasive detection of moving and stationary human with wifi," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2329–2342, 2015.
- [11] X. Li, D. Zhang, Q. Lv, J. Xiong, S. Li, Y. Zhang, and H. Mei, "Indotrack: Device-free indoor human tracking with commodity wi-fi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–22, 2017.
- [12] M. Z. Khan, A. Taha, W. Taylor, M. A. Imran, and Q. H. Abbasi, "Non-invasive localization using software-defined radios," *IEEE Sensors Journal*, vol. 22, no. 9, pp. 9018–9026, 2022.
- [13] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.
- [14] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using wifi," *SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 4, p. 269–282, aug 2015. [Online]. Available: <https://doi.org/10.1145/2829988.2787487>
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European conference on computer vision*. Springer, 2016, pp. 630–645.
- [16] W. W. L. Z. C. Bing Li, Wei Cui and M. Wu, "Two-stream convolution augmented transformer for human activity recognition," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.
- [17] Y. Ge, S. Li, M. Shentu, A. Taha, S. Zhu, J. Cooper, M. Imran, and Q. Abbasi, "A doppler-based human activity recognition system using wifi signals," in *2021 IEEE Sensors*. IEEE, 2021, pp. 1–4.