Human Respiration Detection Under Interference: Challenges and Solutions

Kehan Wu*, Renqi Chen*, Haiyu Wang, and Guang Wu
Department of Electric and Electrical Engineering, Southern University of Science and Technology, Shenzhen, 518055, China

***Abstract*—Recent research has highlighted the detection of human respiration rate using commodity WiFi devices. Nevertheless, these devices encounter challenges in accurately discerning human respiration amidst the prevailing human motion interference encountered in daily life. To tackle this predicament, this paper introduces a passive sensing and communication system designed specifically for respiration detection in the presence of robust human motion interference. Operating within the 60.48GHz band, the proposed system aims to detect human respiration even when confronted with substantial human motion interference within close proximity. Subsequently, a neural network is trained using the collected data by us to enable human respiration detection. The experimental results demonstrate a consistently high accuracy rate over 90% of the human respiration detection under interference, given an adequate sensing duration. Finally, an empirical model is derived analytically to achieve the respiratory rate counting in 10 seconds.**

***Keywords*—Millimeter Wave, Integrated Sensing and Communication (ISAC), Device-free Respiration Monitoring (DFRM), Machine Learning**

I. Introduction

The ability to accurately and reliably detect respiratory patterns has long been a crucial aspect of healthcare monitoring, sleep analysis, and physical activity tracking. A person's respiration provides valuable insights into his overall well-being. For example, a respiratory rate value greater than 27 breaths per minute (bpm) as been identified as a significant predictor of cardiac arrest and is utilized in the prediction of conditions such as pneumonia or lower respiratory tract infections [1–3]. Traditionally, respiration detection has relied on the use of sensors or wearable devices, which often impose constraints on the individual's freedom of movement and may lead to discomfort or inconvenience [4].

Passive sensing, a branch of Integrated Sensing and Communication (ISAC), has paved the way for device-free respiration monitoring (DFRM) [5, 6]. This innovative approach eliminates the need for physical contact by capturing and processing signals influenced by subtle movements associated with respiration. Importantly, passive sensing has demonstrated its ability to accurately capture the motion of the chest caused by respiration without significantly occupying bandwidth.

There have been several existing methods to estimate respiration rate using non-contact techniques, with most of them relying on channel state information (CSI) or received signal strength (RSS) [7–10]. However, the CSI methods typically utilize only a small part of the signal, resulting in increased delay and compromised accuracy. Moreover, CSI method has to cancel the frequency offset between the transmitter and receiver, realizing the synchronization between them.

Currently, most DFRM systems employ commercial WiFi [7–10], offering convenience and ease of implementation. However, the detection precision of these systems is limited by the wavelength with the following restriction:

$$f_d = \frac{v cos(\theta)}{\lambda} \tag{1}$$

where f_d is Doppler Frequency, θ is the angle between velocity direction and signal arrival direction, λ is the wavelength of the signal.

To overcome the limitations associated with lower sensitivity, the utilization of millimeter waves offers distinct advantages in DFRM systems. Millimeter waves have shorter wavelengths, enabling higher precision in detecting subtle respiratory movements. Furthermore, millimeter-wave technology is extensively employed in next-generation communication systems, creating an opportunity for synergistic integration.

In general, there are two types of interference:

1) Static Environmental Interference: Environmental interference encompasses external factors that potentially compromise the accuracy and reliability of a measurement or detection process, such as the presence of reflected signals from the static environment. Within the domain of respiration detection, interference stemming from the static environment can introduce challenges and significantly impact the precision of measurements.
2) Motion Interference: The interference caused by human motion is induced through random human motion around the detected target in the surrounding area during respiration detection. Movements such as walking, running, or even slight body motions can introduce artifacts and distort the respiration signal. Motion interference poses a challenge because it can lead to inaccurate measurements and make it difficult to extract the true respiratory pattern.

To address these challenges, we propose a DFRM system based on passive millimeter radar with a strong anti-interference ability, acceptable processing speed, and improved

*Both authors contributed equally to this work. Email: {12010519, 11910818, 12011331}@mail.sustech.edu.cn, and wug@sustech.edu.cn

accuracy. It will offer a powerful solution for advancing both healthcare monitoring and communication capabilities.

The subsequent sections of this paper are structured as follows. Section II presents an introduction to the mmWave passive sensing and communication system. The signal processing algorithm for passive sensing and the methodology for approximating classification accuracy are expounded in Section III. Section IV outlines the system implementation, accompanied by the demonstration and discussion of experimental results. Lastly, the paper concludes in Section V.

II. Overview of the System

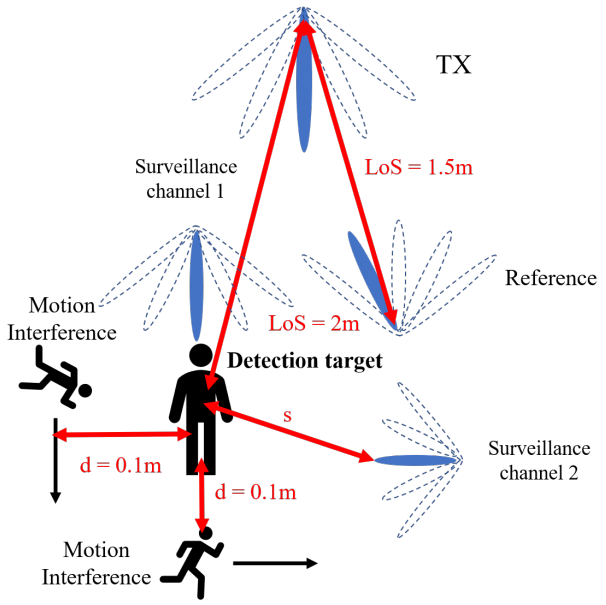

Fig. 1. Model of the Sensing System.

A model of passive sensing and communication system working on the mmWave waveband is elaborated in Fig. 1. The transmitter portion transmit the communication signal with a carrier frequency of 60.48GHz over the downlink.

The transmitter part is composed of a Sivers millimeter wave phased array block with 16 antennas and a USRP baseband processing block. The receiver portion is consists of 3 different channels. To be specific, the first channel demonstrated as the reference channel is settled to receive the communication signal, the second and third channel denoted as surveillance channel. Both are used to receive the reflected signal from the chest. The transmitter delivers the communication signal to the receiver without Doppler frequency in the reference channel, and the receiver towards the reference channel receive signal to obtain the transmission signal without interference. The surveillance channels receive the echo signal from the moving target in the environment, and the reflected signal is processed by the receiver to obtain the Doppler information of the target [11]. The surveillance channels are used to detect the Doppler frequency in the surroundings, and the Doppler frequency can derive the approximate human respiration rate. And s represents the distance between the target and the closest receiver which is a variable in different experiment.

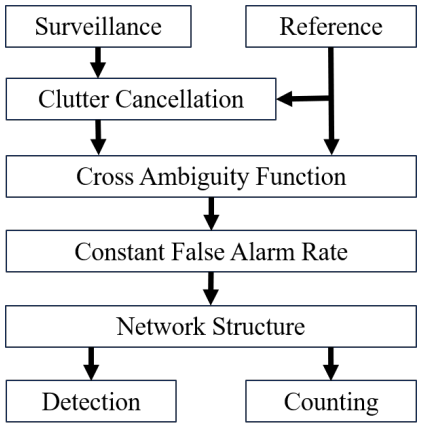

Fig. 2. Signal Processing Diagram.

Subsequently, a clutter cancellation method is applied to mitigate the internal interference within the signals obtained from both the surveillance channels and the reference channel. The processed signals from these channels are then subjected to the Cross Ambiguity Function (CAF) and Constant False Alarm Rate (CFAR) algorithms to extract the Doppler frequency information. These resulting Doppler frequency values are subsequently fed into a neural network that has been trained using the collected data, enabling the detection of human respiration.

The system's primary objective is twofold: firstly, to detect the presence of human respiration in the surrounding environment, and secondly, to accurately count the respiratory rate.

III. System Model

A. Signal Model

The transmitter transmits the communication signal $s(t)$ to the receiver, the received signal via the reference channel is expressed in (2)

$$y_r(t) = \sum_{i=1}^{L_r} \alpha_{i,r} s(t - \tau_{i,r}) e^{-j2\pi f_\Delta t} + n_{i,r}(t) \quad (2)$$

where L_r is the number of paths, $\alpha_{i,r}$ is the attenuation coefficient of the i-th path, $\tau_{i,r}$ is the delay of the i-th path, f_Δ is the frequency offset between the transmitter and receiver, $n_{i,r}(t)$ is the noise of the i-th path.

The received signal via the surveillance channel is expressed in (3)

$$y_s(t) = \sum_{i=1}^{L_s} \alpha_{i,s} s(t - \tau_{i,s}) e^{-j2\pi f_\Delta t} e^{-j2\pi f_i t} + n_{i,s}(t) \quad (3)$$

where L_s is the number of paths, $\alpha_{i,s}$ is the attenuation coefficient of the i-th path, $\tau_{i,s}$ is the delay of the i-th path, f_i is the Doppler frequency of the i-th path, $n_{i,s}(t)$ is the noise of the i-th path.

B. Signal Processing of Passive Sensing

The received signals from surveillance and reference beams are sampled at the baseband with a period T_s, which can be expressed by

$$y_s[n] = y_s(nT_s) \tag{4}$$

and

$$y_r[n] = y_r(nT_s) \tag{5}$$

The received signal undergoes clutter cancellation to mitigate interference originating from the static environment. Subsequently, it is processed by the Cross Ambiguity Function and Constant False Alarm Rate algorithms to calculate the Doppler Frequency within the vicinity. The resulting information is then forwarded to a network architecture to facilitate both the detection and counting of human respiration.

1) Adaptive Clutter Cancellation Algorithm: The Adaptive Clutter Cancellation Algorithm is used to eliminate the interference of reflected signal from the static environment [12]. The algorithm encompass the Least Mean Square (LMS) , recursive minimum algorithm, and normalized least mean square algorithm. In this paper, an adaptive direct path clutter cancellation algorithm based on LMS is used.

The $\mathbf{V}_r$ is the matrix of delayed reference signal of the reference channel:

$$\mathbf{V}_r = \begin{bmatrix} y_r[1] & y_r[0] & \cdots & y_r[-P+2] \\ y_r[2] & y_r[1] & \cdots & y_r[-P+3] \\ \vdots & \vdots & \vdots & \vdots \\ y_r[N] & y_r[N-1] & \cdots & y_r[N-P+1] \end{bmatrix} \tag{6}$$

where P is the number of delays, N is the length of Coherent Integration Time (CIT).

Suppose $y_s[n]$ is the received signal vector of the surveillance channel, then $\mathbf{V}_s$ is the matrix of the surveillance channel:

$$\mathbf{V}_s = [y[1], y[2], \cdots, y[N]]^T \tag{7}$$

According to the LMS method, the weight matrix can be represented by

$$\mathbf{K} = \left(\mathbf{V_r}^H \mathbf{V_r}\right)^{-1} \mathbf{V_r}^H \mathbf{V_s} \tag{8}$$

The surveillance channel signal after eliminating clutter interference $\mathbf{y}_\mathrm{s}^*$ can be represented by

$$\mathbf{y}_\mathrm{s}^* = \mathbf{y}_s - \mathbf{V_r K} \tag{9}$$

2) Doppler Frequency Estimation Algorithm: The CAF is a crucial radar signal processing tool that plays a vital role in representing the matched filter response. It provides valuable information about the time delays present in the echo signal as well as the energy of the Doppler-shifted echo component. By estimating the correlation between signals collected from two channels, the CAF enables the estimation of target parameters such as time delay and Doppler frequency shift. The expression for this correlation estimation of continuous time form is as follows:

$$R(f_d) = \max_{\tau} \sum_{n=0}^{N-1} y_s[n] \{y_r[n-\tau]\}^\dagger \cdot e^{-j2\pi f_d n T_s} \tag{10}$$

where $\dagger$ denotes the complex conjugate, f_d is the Doppler frequency, N is number of the sample points in CIT, and τ is the matched time delay.

By examining the CAF, it becomes apparent that the CAF traverses various combinations of time delay and Doppler frequency within their respective feasible ranges. When the values of τ and f_d align with the actual time delay and Doppler frequency shift between the surveillance signal and the reference signal, a distinct local peak marked as $R(f_d)$ emerges. The magnitude of this peak indicates the strength of the Doppler frequency component at that particular time delay. Utilizing this information, one can estimate the precise values of the time delay and Doppler frequency difference between the reference signal and the monitor signal.

C. Methodologies to Mitigate Interference

The approach used in this study is the two-dimensional CFAR Algorithm (2D-CA-CFAR) [13]. In the 2D-CA-CFAR algorithm, the unit that needs to be compared with the threshold is referred to as the cell under test (CUT). Typically, an equal number of units in the surrounding of the CUT are used to compute the background noise power, and these units are called training units (w). In addition, there are several units between the CUT and the training units that do not participate in the calculation of the background noise power, and these units are referred to as guard units. The threshold $\beta = \alpha P_n$ is determined by the background noise power P_n and the threshold factor α. The background noise power is the mean of the training units and can be written as:

$$P_n = \frac{1}{N} \sum_{w \in W} w \tag{11}$$

where N is the number of training units, and W is the set of training units.The threshold factor α is determined by the pre-defined false alarm rate and can be expressed as follows:

$$\alpha = N(P_{fa}^{-1/N} - 1) \tag{12}$$

where α represents the false alarm rate.

The Constant False Alarm Rate algorithm (CFAR) dynamically adjusts the threshold for target detection based on the level of noise present in the environment. This adaptive adjustment helps reduce the chances of false detection, missed detection, and false alarms. Furthermore, the CFAR algorithm effectively handles scenarios where no Doppler frequency is present, ensuring accurate detection and improved performance overall.

D. Classification Based On Deep Learning

For the two primary research tasks, namely the relationship between respiratory detection under interference and perception duration, and the determination of the number of respiratory under interference, deep learning methods are chosen for the analysis. For the consideration of generalizability and time cost, ResNet [14] is adopted in this paper.

The basic building block of ResNet is residual block, which consists of a skip connection and batch normalization. The skip connection directly adds the input to the output, enabling the network to learn the residual function, represented by formula (13).

$$\mathbf{y} = F(\mathbf{x}) + \mathbf{x} \tag{13}$$

where $\mathbf{x}$ denotes the input features and $F(\mathbf{x})$ represents the residual function, expressed as:

$$F(\mathbf{x}) = \mathbf{W_2} \cdot ReLU(\mathbf{W_1} \cdot \mathbf{x} + \mathbf{b_1}) + \mathbf{b_2} \tag{14}$$

where $\mathbf{W_1}$ and $\mathbf{W_2}$ denote the weight parameters in the convolutional layers, b_1 and b_2 denote the bias, and $ReLU$ is activation function.

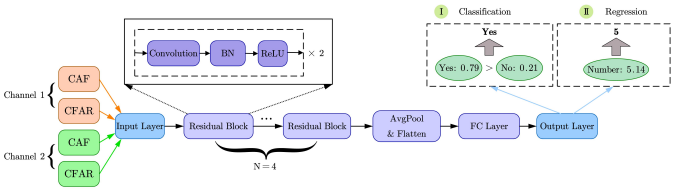

Fig. 3. Network Structure.

The network structure is elaborated in Fig. 3, which uses four residual blocks. The input of the ResNet are Time-Doppler spectrogram obtained from CAF and CFAR results of Time-Doppler spectrogram of two surveillance channels as raw CAF results provide more micro information and CFAR provide more macro information. The output of the ResNet are different for respective tasks. For the detection of breathing presence or absence within perception duration, treat it as a binary classification task and then the output layer has two nodes. On the other hand, for the respiratory rate problem, treat it as a regression task and output layer has only one node. Since we only consider the case where the number of breaths is an integer, the output of the model will be rounded.

IV. Experiment

A. Implementation

The system diagram, as depicted in Fig. 4, presents an overview of the overall system configuration. In the transmitter section, a NI USRP-2954R is utilized to generate an intermediate frequency (IF) signal centered at 500 MHz. This IF signal is subsequently upconverted to 60.48 GHz and transmitted using a Sivers 60.48 GHz phased array. On the receiver side, three phased arrays are connected to two USRP, enabling the reception of signals from both the reference channel and two surveillance channels. The phased arrays are controlled by laptops, facilitating collaborative beam switching functionality.

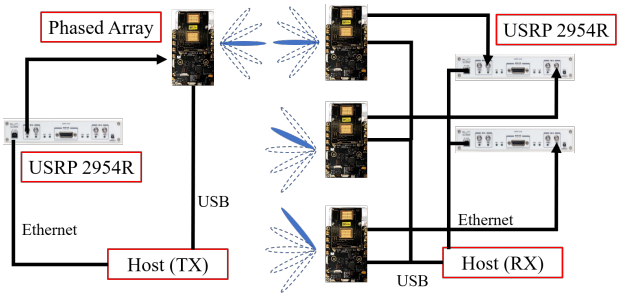

Fig. 4. Block diagram of system implementation..

The transmission signal comprises a training sequence lasting 16 μs and is succeeded by an Orthogonal Frequency Division Multiplexing (OFDM)-modulated data payload with a duration of 200 μs.

Fig. 5 depicts the experimental layout, which corresponds to the model presented in Fig. 1, wherein the Line-of-Sight (LoS) channel is utilized for the purpose of respiration detection. The distance from the transmitter to the detection target (experimenter 1) is 2 m. And experimenter 2 moving around in the vicinity serves as interference. The distance between the target and the closest receiver is changing in different experimental design.

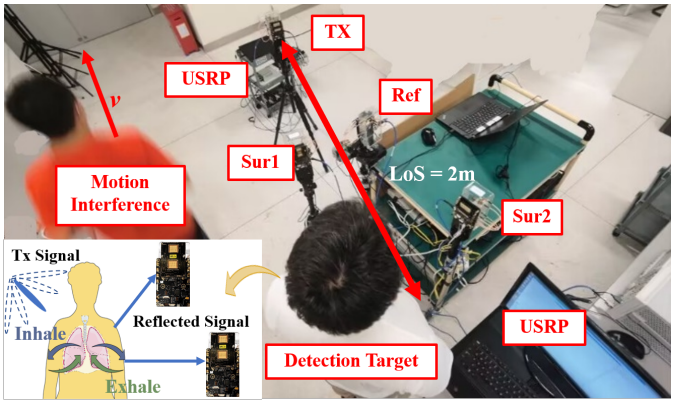

Fig. 5. Experiment Layout.

B. Dataset

The experiment involves sensing two types of human respiration: detection and counting the human respiration under interference. The interference caused by human motion is induced through random walking around the detected target in the surrounding area. Noticed that there are no restrictions on the respiratory intensity of the target, i. e. both deep and calm breathing are allowed. In experimental scenario, the detection target can breath freely. To ensure the robustness of respiratory detection system, different volunteers (variations in height, gender, body shape and so forth) are invited as the detection target. Furthermore, a group of volunteers is instructed to walk around the target with varying directions and speeds.

Initially, we detect the presence of human respiration amidst the interference, considering periods of 2.5, 5, 7, and 10

TABLE I
DATASETS SETUP

Detection Dataset						
Positive Samples		Negative Samples		Train:Test		K-Fold Method
100		100		7:3		5-Fold
Counting Dataset						
2 Respirations	3Respirations	4 Respirations	5 Respirations	6 Respirations	Train:Test	K-Fold Method
100	100	100	100	100	7:3	5-Fold

seconds. Subsequently, we quantify the occurrence of human respiration under the influence of interference, conducting multiple counts ranging from 2 to 6 instances. Each situation of the experiment is sampled 100 times with the passive sensing system in LoS scenarios, resulting in two datasets, illustrated in Table I.

C. CAF and CFAR Results

Fig. 6(a) and Fig. 6(e) illustrate the time-Doppler spectrograms and CFAR results for the detection of human respiration in the presence of motion interference. Conversely, Fig. 6(b) and Fig. 6(f) display the time-Doppler spectrograms and CFAR results for human respiration in the presence of static environment interference. In contrast, Fig. 6(c)(d)(g)(h) solely depict the spectrograms and CFAR results of interference signals. These spectrograms provide clear visual representations of distinct patterns that distinguish human respiration across various time periods and interference, thereby emphasizing noticeable disparities between interference signals and respiration signals.

D. Classification and Regression Results

1) Detection: This paper investigates the binary classification detection of breathing presence in the environment across various detection time intervals. As elaborated in Fig. 7, the analysis reveals that when the detection time exceeds 7 seconds, accurate identification of potential breathing signals within the detection environment can be achieved with a 100% accuracy rate. Also, the experiment verify the robustness of the proposed approach by changing the the distance between the target and the closest receive. We use different curves to represent different distance cases. Note that analytical approximation is achieved by logistic fitting curve, which can better reveal the tendency of real accuracy. The definition of the accuracy is shown in (15). Moreover, the classification accuracy increases with respect to the length of detection time, in line with the expectations.

$$Accuracy = \frac{TN + TP}{All\ Predicted\ Labels} \tag{15}$$

where TN and TP denote the number of correctly classified positive samples, and negative samples, respectively.

2) Counting: In counting the human respiration task, five experimental scenarios are performed: 2 respirations, 3 respirations, 4 respirations, 5 respirations and 6 respirations within 10 seconds. In order to achieve this counting task, we train the network with 400 samples (80 samples per type) and treat remaining samples as the test set. The results with the closest distance is 0.1m are illustrated in Fig. 8(a).

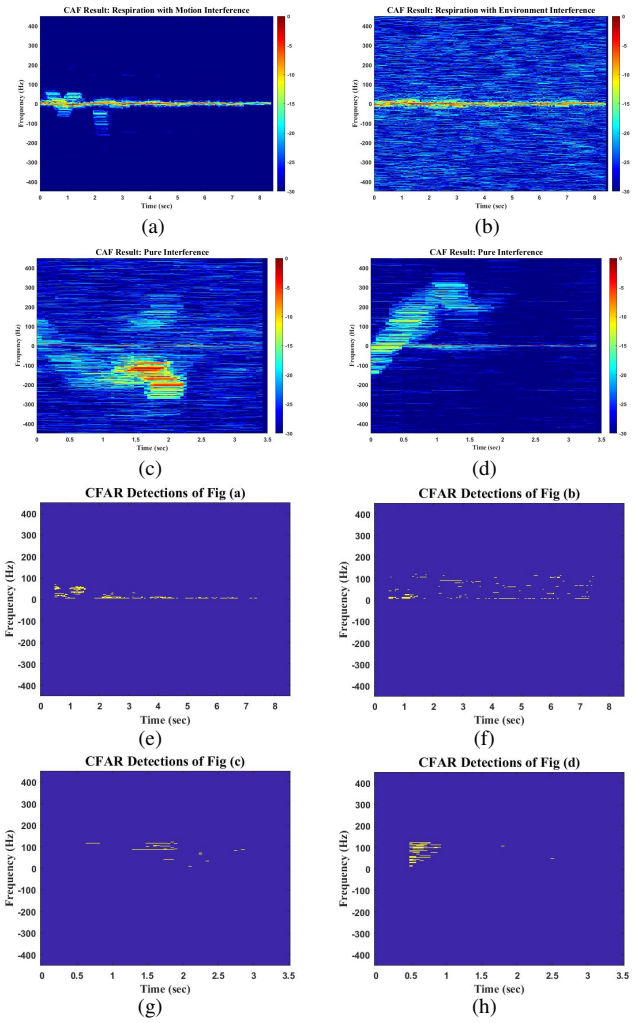

Fig. 6. Fig (a) and (e) represent the results of Cross Ambiguity Function (CAF) and Constant False Alarm Rate (CFAR) algorithms applied to respiration signals in the presence of human motion interference. Fig (b) and (f) display the CAF and CFAR results of respiration signals affected by static environment interference. On the other hand, Fig (c), (d), (g), and (h) exhibit the CAF and CFAR results obtained from environments where no respiration activity is present.

The comparison of other studies is shown in Table II

V. CONCLUSION

This paper presents a comprehensive study on an ISAC system operating in the 60.48 GHz frequency band. The performance of the system in counting human respiration under interference is demonstrated by executing five distinct respiratory rates in the surveillance channel and collecting a

TABLE II
COMPARISON OF STUDIES

Name	Working frequency	System	Interference	Accuracy	Application
Zhang et al.[9]	WiFi	2 commercial WiFi	-	-	Human Respiration
MIT CSAIL[15]	Milimeter Wave	2 rows of FMCW radar	-	96%	Human skeleton
Sun et al.[12]	WiFi	Commercial WiFi and Horn Antenna	-	-	Human Motion
Ali et al.[16]	Milimeter Wave	Simulation	-	-	Vehicle Motion
Ours	**Milimeter**	**4 Sivers phased array**	**Yes**	**90%**	**Human Respiration**

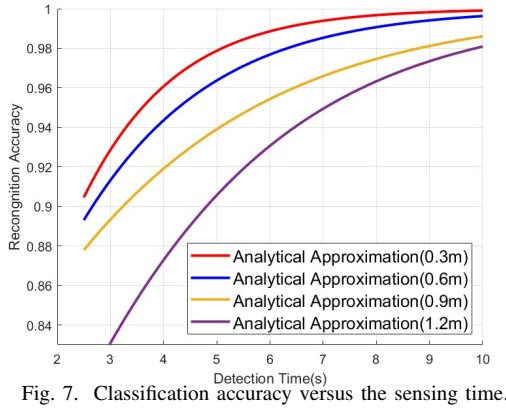

Fig. 7. Classification accuracy versus the sensing time.

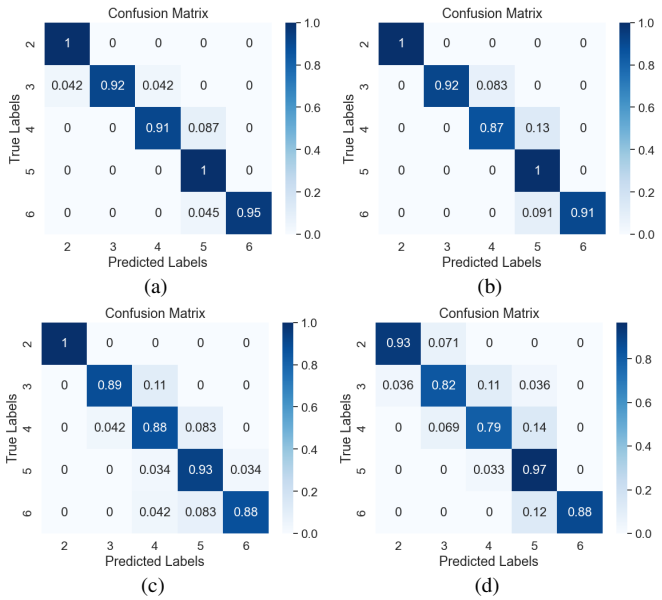

Fig. 8. Fig (a) represents the results of the distance between the target and the closest receiver is 0.3m. The distance in Fig (b), (c) and (d) are 0.6m, 0.9m and 1.2m respectively

dataset of received signals using the aforementioned system. Subsequently, a ResNet model is trained using the collected dataset for both respiration detection and respiratory rate counting.

The experimental results showcase remarkable classification accuracy, approaching 100%, for detection times exceeding 6 seconds. The accuracy of counting the frequency of human respiration achieve 90% overall.

REFERENCES

[1] A. Nicolò, C. Massaroni, E. Schena, and M. Sacchetti, "The Importance of Respiratory Rate Monitoring: From Healthcare to Sport and Exercise," *Sensors*, vol. 20, no. 21, p. 6396, Nov. 2020.

[2] J. Rumiński, "Analysis of the parameters of respiration patterns extracted from thermal image sequences," *Biocybernetics and Biomedical Engineering*, vol. 36, no. 4, pp. 731–741, Jan. 2016.

[3] C. Chourpiliadis and A. Bhardwaj, "Physiology, Respiratory Rate," in *StatPearls*. StatPearls Publishing, 2023.

[4] C. Massaroni, A. Nicolò, D. Lo Presti, M. Sacchetti, S. Silvestri, and E. Schena, "Contact-Based Methods for Measuring Respiratory Rate," *Sensors*, vol. 19, no. 4, p. 908, Jan. 2019.

[5] E. Vanegas, R. Igual, and I. Plaza, "Sensing Systems for Respiration Monitoring: A Technical Systematic Review," *Sensors (Basel, Switzerland)*, vol. 20, no. 18, p. 5446, Sep. 2020.

[6] J. Wang, N. Varshney, C. Gentile, S. Blandino, J. Chuang, and N. Golmie, "Integrated Sensing and Communication: Enabling Techniques, Applications, Tools and Data Sets, Standardization, and Future Directions," *IEEE Internet of Things Journal*, vol. 9, no. 23, pp. 23 416–23 440, Dec. 2022.

[7] T. Nakamura, M. Bouazizi, K. Yamamoto, and T. Ohtsuki, "Wi-Fi-Based Fall Detection Using Spectrogram Image of Channel State Information," *IEEE Internet of Things Journal*, vol. 9, no. 18, pp. 17 220–17 234, Sep. 2022.

[8] X. Wang, C. Yang, and S. Mao, "Resilient Respiration Rate Monitoring With Realtime Bimodal CSI Data," *IEEE Sensors Journal*, vol. 20, no. 17, pp. 10 187–10 198, Sep. 2020.

[9] Y. Zeng, D. Wu, R. Gao, T. Gu, and D. Zhang, "FullBreathe: Full Human Respiration Detection Exploiting Complementarity of CSI Phase and Amplitude of WiFi Signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 148:1–148:19, Sep. 2018.

[10] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, "MultiSense: Enabling Multi-person Respiration Sensing with Commodity WiFi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 3, pp. 102:1–102:29, Sep. 2020.

[11] C. Wang, J. Liu, Y. Chen, H. Liu, L. Xie, W. Wang, B. He, and S. Lu, "Multi - Touch in the Air: Device-Free Finger Tracking and Gesture Recognition via COTS RFID," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*. Honolulu, HI: IEEE, Apr. 2018, pp. 1691–1699.

[12] H. Sun, L. G. Chia, and S. G. Razul, "Through-Wall Human Sensing With WiFi Passive Radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, pp. 2135–2148, Aug. 2021.

[13] M. Kronauge and H. Rohling, "Fast Two-Dimensional CFAR Procedure," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1817–1823, Jul. 2013.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[15] M. Zhao, T. Li, M. A. Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, "Through-Wall Human Pose Estimation Using Radio Signals," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, Jun. 2018, pp. 7356–7365.

[16] Z. Ali, A. Duel-Hallen, and H. Hallen, "Early Warning of mmWave Signal Blockage and AoA Transition Using sub-6 GHz Observations," *IEEE Communications Letters*, vol. 24, no. 1, pp. 207–211, Jan. 2020.