Published: 03 September 2025

RESEARCH-ARTICLE

SpiroSense: Transforming Smartphones into Pulmonary Metrics Monitors with Ultrasonic Technology

LONG FAN, Nanjing University, Nanjing, Jiangsu, China

LEI XIE, Nanjing University, Nanjing, Jiangsu, China

SHIYUAN MA, Nanjing University, Nanjing, Jiangsu, China

YANLING BU, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, China

CHUYU WANG, Nanjing University, Nanjing, Jiangsu, China

WEIBANG PAN, Nanjing University, Nanjing, Jiangsu, China

View all

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Volume 9, Issue 3 (September 2025)
https://doi.org/10.1145/3749489
EISSN: 2474-9567

SpiroSense: Transforming Smartphones into Pulmonary Metrics Monitors with Ultrasonic Technology

LONG FAN, Nanjing University, China
LEI XIE*, Nanjing University, China
SHIYUAN MA, Nanjing University, China
YANLING BU, Nanjing University of Aeronautics and Astronautics, China
CHUYU WANG, Nanjing University, China
WEIBANG PAN, Nanjing University, China
LU KE, Nanjing University, China
SANGLU LU, Nanjing University, China

Chronic respiratory conditions such as Chronic Obstructive Pulmonary Disease (COPD) and asthma often progress insidiously, making early detection vital for effective intervention. Current gold-standard Pulmonary Function Testing (PFT) methods, such as spirometry, evaluate lung function by measuring airflow rates to detect potential obstructions. But, their cost, often several hundred dollars or more, limits their accessibility for regular at-home monitoring. In this paper, we present *SpiroSense*, a novel system that transforms a smartphone into a portable, low-cost, and accurate PFT device for everyday use by integrating a custom 3D-printed attachment costing just a dozen dollars. However, a critical limitation arises from the smartphone's inherent audio sampling rate (typically 48 kHz), which constrains the airflow resolution to 11.9 L/s when using conventional cross-correlation-based time delay estimation. This coarse resolution is insufficient to capture key pulmonary metrics, such as a Peak Expiratory Flow (PEF) of 10 L/s, with high fidelity. To address this, we propose *SonicFlow*, which establishes a foundational airflow rate sensing model based on ultrasonic phase features and improves the airflow rate resolution to 0.148 L/s. Furthermore, airflow-induced high-frequency harmonic noise within the 3D-printed model, combined with ambient environmental noise, further complicates accurate sensing. To mitigate this, we introduce *NoiseClear*, an end-to-end ultrasonic signal enhancement model designed to effectively suppress noise while preserving critical airflow velocity information. We prototype *SpiroSense* and evaluate its performance on a cohort of 59 participants, including 29 healthy individuals and 30 patients. Experimental results show that *SpiroSense* achieves average estimation error of 6.44% for Forced Vital Capacity (FVC), 7.42% for Forced Expiratory Volume in one second (FEV_1), and 3.01% for the FEV_1/FVC ratio.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing systems and tools**.

Additional Key Words and Phrases: Pulmonary Function Tests, Smartphones, Ultrasonic Sensing, Airflow Velocity

*Corresponding author.

Authors' Contact Information: Long Fan, State Key Laboratory for Novel Software Technology, Nanjing University, China, fanl@smail.nju.edu.cn; Lei Xie, State Key Laboratory for Novel Software Technology, Nanjing University, China, lxie@nju.edu.cn; Shiyuan Ma, State Key Laboratory for Novel Software Technology, Nanjing University, China, mashiyuan@smail.nju.edu.cn; Yanling Bu, College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China, byling@nuaa.edu.cn; Chuyu Wang, State Key Laboratory for Novel Software Technology, Nanjing University, China, chuyu@nju.edu.cn; Weibang Pan, Nanjing University, China, pwb15161986019@outlook.com; Lu Ke, Department of Critical Care Medicine, Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, China, kelu@nju.edu.cn; Sanglu Lu, State Key Laboratory for Novel Software Technology, Nanjing University, China, sanglu@nju.edu.cn.

ACM 2474-9567/2025/9-ART79
https://doi.org/10.1145/3749489

ACM Reference Format:
Long Fan, Lei Xie, Shiyuan Ma, Yanling Bu, Chuyu Wang, Weibang Pan, Lu Ke, and Sanglu Lu. 2025. SpiroSense: Transforming Smartphones into Pulmonary Metrics Monitors with Ultrasonic Technology. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 3, Article 79 (September 2025), 26 pages. https://doi.org/10.1145/3749489

1 INTRODUCTION

Pulmonary diseases like Chronic Obstructive Pulmonary Disease (COPD) [34], asthma [28], and other respiratory disorders often develop insidiously, with early-stage symptoms that are subtle and easily overlooked. Pulmonary Function Testing (PFT) [9], the clinical gold standard for assessing ventilatory diseases, quantifies the volume and velocity of exhaled airflow using specialized medical equipment. Key metrics, such as Forced Vital Capacity (FVC) and the ratio of Forced Expiratory Volume in one second to FVC (FEV_1/FVC ratio), offer essential insights into obstructive or restrictive pulmonary impairments[19, 54]. Routine PFT allows individuals and clinicians to detect early abnormalities, monitor disease progression, and assess the impact of lifestyle or environmental factors on respiratory health. For instance, a progressive decline in the FEV_1/FVC ratio may suggest the onset of COPD, prompting early intervention through medical consultation or behavioral adjustments. Despite its clinical importance, regular PFT remains impractical for many individuals due to the reliance on bulky, specialized equipment (e.g., Impulse oscillometry system [33]) operated by trained personnel, typically located in hospitals or large clinics. Moreover, the high cost of devices and the inconvenience of scheduling regular appointments pose significant barriers to routine lung health monitoring, especially for elderly patients, individuals in remote areas, or those managing chronic respiratory conditions. This gap highlights the pressing need for low-cost, accessible, and user-friendly alternatives capable of delivering reliable PFT outside clinical settings, such as at home.

Current PFT approaches can be broadly categorized into *passive* and *active* sensing techniques. Passive methods, including acoustic-based sensing [42, 44], RF-based sensing [4, 18, 21], WiFi-based sensing [17, 20, 26, 48], mmWave-based sensing [12, 13, 16, 27, 46, 50], and vision-based methods [40], typically infer respiratory activity by analyzing chest motion or breathing sounds. While these methods offer convenience, they only provide indirect indicators, such as breathing rate or chest movement, and fail to directly measure airflow velocity, which is critical for accurate diagnosis and disease monitoring [9]. In contrast, active methods such as clinical spirometers [30, 31, 37], directly measure lung airflow and volumes, offering more accurate and clinically meaningful assessments compared with passive techniques. However, clinical PFT equipment remains prohibitively expensive, often costing several hundred dollars, and typically requires specialized training, limiting its accessibility for routine or continuous self-monitoring. Even commercial spirometers (e.g., SP100A Spirometer [7]) designed for home use are not affordable for all households, and often depend on proprietary hardware and software, which creates barriers to the widespread adoption of remote digital health monitoring. Other active methods focus on specific physiological signals, such as airway caliber [47], respiratory events like apnea [32], sleep stages [51], or respiratory wheeze sounds [3, 6, 38], yet they offer only coarse-grained insights and fall short of the clinical standards required for comprehensive lung function assessment [9]. This raises an important question: *Can we design a highly accurate, accessible, and cost-effective PFT solution capable of directly quantifying exhaled airflow velocity and volume in a clinically meaningful manner?*

In this paper, we present *SpiroSense*, a novel system that transforms a standard smartphone into a portable pulmonary assessment device for household use, as shown in Fig. 1. *SpiroSense* requires only a smartphone and a custom 3D-printed attachment, costing just a few dozen dollars, to accurately measure lung airflow by capturing ultrasonic phase shifts. For example, a change in airflow velocity of 1 m/s will result in a phase shift of 0.29 radians in the ultrasonic received signal. *SpiroSense* actively emits ultrasonic signals and uses the received phase-shifted signals to infer airflow velocity, which is then converted into standard pulmonary metrics such as FEV_1 and FVC. The sensing mechanism utilizes Frequency-Modulated Continuous Wave (FMCW) ultrasonic signals, which are transmitted from the smartphone's top speaker into the 3D-printed airflow model when the user exhales into

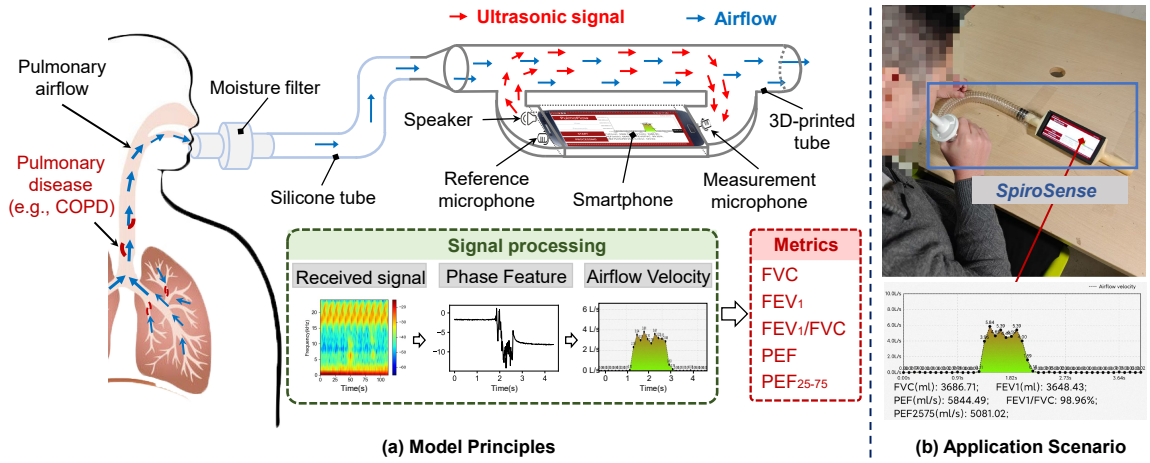


Fig. 1. Smartphone-based pulmonary function monitoring system: (a) Fundamental principles, signal processing workflow, and key PFT metrics; (b) representative application scenarios and measurement results.

the tube. The bottom microphone of the smartphone, designated as the measurement microphone, captures the received signal that carries the pulmonary airflow velocity information. Our system then applies the *SonicFlow* model to extract accurate velocity measurements and convert them into medically relevant indicators based on the calibrated design of the 3D-printed attachment.

There are two major challenges in realizing PFT using smartphones: *The first challenge lies in accurately measuring continuous lung airflow velocity using smartphones, which are constrained by limited sampling rates (typically* 48 kHz*).* When pulmonary air enters the surrounding environment, it diffuses rapidly, leading to swift attenuation of airflow velocity and negatively impacting measurement accuracy. Additionally, traditional Time-of-Flight (ToF) methods, which use cross-correlation to estimate ultrasonic propagation delays caused by airflow velocity changes, are inadequate. For example, at a sampling rate of 48 kHz, the velocity resolution of ToF-based methods is approximately 11.9 L/s. Given that typical pulmonary maximum Peak Expiratory Flow (PEF) is less than 10 L/s, ToF methods fail to provide the required precision. To address the issue of airflow diffusion, we designed a 3D-printed airflow tube that hosts the smartphone's speakers and microphones at its ends. This tube confines both airflow and ultrasonic signal propagation within a narrow path, concentrating the airflow and enhancing measurement accuracy. Based on the 3D-printed tube, to overcome smartphone-based sampling rate limitations, we propose the *SonicFlow* model, which utilizes theoretical derivation to reveal the mechanism between the airflow velocity and the phase change of the ultrasonic received signal. Specifically, following the *SonicFlow* model, we first apply a cross-correlation approach to synchronize and align the received signals, followed by IQ demodulation to extract phase signals from the complex Intermediate Frequency (IF) signals. We then derive fine-granularity airflow velocity features from the phase signal. To ensure precise calculation of PFT metrics, we incorporate Kalman filtering to smooth the airflow velocity features and utilize a sliding standard deviation method for puff activity detection. Consequently, our approach achieves an airflow velocity resolution of 0.148 L/s–approximately 80 times finer than that of the ToF method.

The second challenge arises from interference caused by airflow sounds within the 3D-printed model and ambient noise. When air is blown into the 3D-printed model during the measurement process, the high-frequency harmonic noise produced by the airflow within the pipe generates interference, further complicating the problem. Since the frequency range of these interference noises overlaps with that of the ultrasonic signal, traditional filtering methods prove ineffective in removing them. Additionally, high-frequency harmonics from external sources

such as traffic noise, television, or household appliances also overlap with ultrasonic waves, disrupting signal propagation and reception. The presence of these noises can significantly disrupt the highly sensitive phase characteristics, leading to substantial errors in the phase derived through IQ demodulation or even rendering it impossible to demodulate and extract the phase signal. To address this challenge, we propose a noise elimination method called *NoiseClear*, which leverages a real-time ultrasonic signal enhancement model operating in the waveform domain. *NoiseClear* employs an end-to-end ultrasonic signal enhancement model that processes raw waveforms directly on a smartphone. It features an encoder-decoder architecture with skip connections and is optimized across both time and frequency domains using multiple loss functions. This method effectively suppresses ambient noise and airflow sound interference while preserving the airflow velocity information encoded in the ultrasonic signal.

In summary, our main contributions are as follows:

- We present *SpiroSense*, a smartphone-based system for pulmonary function assessment that uses ultrasonic sensing and a low-cost 3D-printed airflow model to enable accurate measurements with standard smartphones (48kHz sampling rate).
- We introduce *SonicFlow*, a theoretical model that establishes a relationship between airflow velocity and ultrasonic phase shifts. It achieves a resolution of 0.148 L/s, approximately 80× finer than that of conventional ToF methods.
- We propose *NoiseClear*, an end-to-end ultrasonic signal enhancement model capable of real-time operation. It effectively suppresses ambient and airflow noise while preserving critical airflow velocity information, ensuring accurate and reliable signal processing across diverse environmental conditions.
- We implement and evaluate *SpiroSense* in a study involving 59 participants (29 healthy individuals and 30 patients). Results show average error rates of 6.44% (FVC), 7.42% (FEV_1), and 3.01% (FEV_1/FVC ratio), validating its clinical relevance.

The remainder of this paper is organized as follows: Section 2 reviews related work. Section 3 presents the system preliminaries. The theoretical model is presented in Section 4. Section 5 elaborates on our design, while Section 6 describes the implementation and evaluation results. Section 7 presents a case study evaluating *SpiroSense* in a clinical setting. Section 8 discusses the limitations, potential issues, and opportunities for future work. Finally, we conclude our study in Section 9.

2 RELATED WORK

To achieve our goal of developing a portable, accurate, and cost-effective solution for daily PFT, this section reviews relevant research in two key domains: PFT methods and airflow velocity sensing techniques.

2.1 Pulmonary Function Testing

PFT is fundamental for diagnosing and managing respiratory diseases such as asthma and COPD. Traditional medical devices, including spirometers, peak flow meters, and pulmonary gas analyzers, assess lung health by measuring lung capacity, volume, and airflow [35]. Spirometers are the most widely used devices, enabling the calculation of key indices like FEV_1 and the FEV_1/FVC ratio [9], which are critical for detecting airway obstruction and evaluating lung function. There are two main types of spirometers: volume-measurement spirometers [10, 23, 36] that measure volume directly, and flow-measurement spirometers [1] that infer volume by measuring airflow. Although accurate, these devices are expensive and require professional supervision for proper usage, limiting their practicality for routine home monitoring. To overcome these challenges, researchers have explored contactless respiratory assessment methods, including WiFi-based approaches [20, 26, 43, 48, 52], RF-based techniques [4, 21] and mmWave-based techniques [13, 16, 45, 50]. These methods estimate pulmonary metrics by monitoring chest movement; however, their performance is often compromised by real-world environmental

factors such as multipath interference, which undermines their reliability and limits their deployment in diverse home settings. Additionally, while SpiroSmart [25] and SpiroCall [14] calculate the exhaled flow rate by estimating models of the user's vocal tract and the reverberation of sound around the user's head, their sensing principle is limited and highly susceptible to environmental noise in real-world scenarios. Although traditional spirometry remains the gold standard for PFT, the demand for accessible, portable, and cost-effective alternatives for daily monitoring continues to grow, inspiring the development of innovative, user-friendly solutions.

2.2 Airflow Velocity Sensing

The gold standard for PFT typically involves using specialized devices to measure the airflow velocity within the lungs, as accurate airflow velocity measurement is crucial for assessing lung function effectively. Existing techniques span a wide range, from fiber-optic anemometers [53] and graphene-based sensors [22] to piezoelectric transducers [24]. In non-invasive sensing, ultrasonic TOF-based airflow velocity sensing calculates the airflow velocity by employing a cross-correlation method to estimate the time delay of the received signal [44, 47]. For example, Wang et al. [44] developed a non-contact system to monitor respiration rate by sensing airflow with ultrasound signals, though it does not directly measure lung health metrics like FEV_1 or FVC. Other research, such as FlowSense [5], uses audio data to monitor airflow in indoor spaces, and systems like PTEase [47] utilizes active acoustic sensing to measure internal changes in the Cross-Sectional Areas (CSA) at different airway positions. These methods demonstrate the promise of acoustic-based airflow sensing, but they face significant limitations in resolution due to the sampling rate constraints of smartphone microphones (typically 48kHz), resulting in insufficient detail for capturing subtle variations in airflow critical to pulmonary diagnosis.

While these prior efforts have inspired our work, none fully satisfy the requirements for smartphone-based PFT. Existing methods either lack the resolution needed for accurate airflow velocity measurements or rely on external hardware, compromising portability and cost-efficiency. To bridge this gap, we propose *SpiroSense*, a smartphone-based system that leverages ultrasonic sensing for pulmonary airflow velocity measurement. By integrating advanced signal processing techniques with novel sensing capabilities, *SpiroSense* aims to deliver a portable, accurate, and affordable solution for daily pulmonary health monitoring, overcoming the limitations of existing approaches.

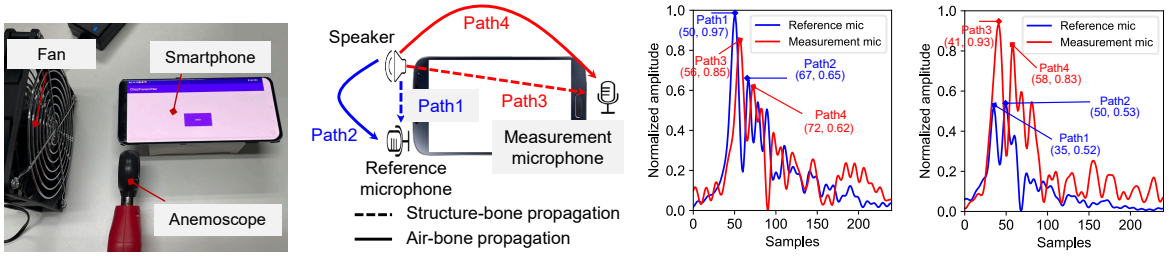
3 EMPIRICAL STUDY

Assessing pulmonary conditions, particularly pulmonary ventilation volume, is crucial for diagnosing various pulmonary diseases. This is typically done by measuring airflow velocity, which poses a challenge for smartphone-based sensing due to its inherent limitations. In this section, we explore the feasibility of using smartphones to measure airflow velocity, which is integral to evaluating pulmonary conditions. The study is divided into two key parts: i) Airflow Velocity Sensing Using ToF Method: We examine the traditional smartphone-based ToF method for airflow velocity sensing, highlighting its limitations through both theoretical modeling and experimental analysis. ii) Ultrasonic Phase-based Airflow Sensing: Given the limitations of smartphones in directly sensing airflow, we design comparative experiments to evaluate the ultrasonic phase method, aiming to demonstrate the potential of smartphones in assessing pulmonary condition despite their limitations.

3.1 Airflow Velocity Sensing Using ToF Method

Observation 1: *The resolution of airflow velocity using the ToF method on smartphones is constrained by the sampling rate (typically 48kHz), limiting the velocity resolution to 16.85 m/s. This is insufficient for capturing subtle airflow changes in pulmonary function assessment, where typical lung airflow velocities are around 1 m/s.*

To evaluate the feasibility of ToF-based airflow sensing on smartphones, we utilize the device's speaker and microphones to transmit and receive ultrasonic signals, respectively. Airflow affects the Doppler shift of the



(a) Experiment setup. (b) Ultrasonic signal propagation. (c) Airflow velocity is 0m/s. (d) Airflow velocity is 3.2m/s.

Fig. 2. Ultrasonic airflow velocity sensing using the ToF method with cross-correlation analysis: (a) Experiment setup; (b) schematic of ultrasonic signal propagation using a smartphone; (c) cross-correlation result at an airflow velocity of 0 m/s, serving as a baseline reference; and (d) cross-correlation result at an airflow velocity of 3.2 m/s, illustrating the phase shift introduced by airflow.

ultrasonic wave, thus influencing the propagation time. In our experiment setup, as shown in Fig. 2(a), a fan generates a controlled airflow field, with the smartphone placed directly within it. We employ the smartphone's default configuration to transmit FMCW ultrasonic signals, sweeping from 17kHz to 23kHz over a 10ms duration. wo microphones capture the signal: a reference microphone (i.e., closer to the speaker) and a measurement microphone (i.e., farther away), as illustrated in Fig. 2(b). We use the cross-correlation ToF method to measure the time difference between the ultrasonic signals received by the two microphones to determine airflow velocity. As shown in Fig. 2(c), the difference in ultrasonic signal propagation time between `Path3` and `Path4` is $\frac{N_d}{f_s}$, where N_d represents the sample difference, and f_s is the sampling rate. Therefore, we can derive the relationship between the ultrasonic velocity in air-bone and structure-bone as $\frac{L}{c} - \frac{L'}{c'} = \frac{N_d}{f_s}$, where L and L' represent the path lengths and c and c' represent the ultrasonic velocity in air-bone and structure-bone, respectively. From this, we can further calculate the ultrasonic velocity in structure-bone as:

$$c' = L' / \left(\frac{L}{c} - \frac{N_d}{f_s} \right). \tag{1}$$

For instance, as shown in Fig. 2(c), assuming that $L' \approx L$ with a value of 0.15m, and given the ultrasonic velocity in air-bone $c = 346\,\text{m/s}$, the sample difference $N_d = 16$, and the sampling rate $f_s = 48\,\text{kHz}$, the ultrasonic velocity in structure-bone can be estimated as 1497.1m/s. Comparison of Fig. 2(c) and Fig. 2(d) reveals that the sample counts of `Path4` and `Path3` differ by 16 and 17 samples, respectively, which are nearly identical. *Results demonstrate that the cross-correlation delay method implemented on smartphones lacks the sensitivity to detect wind velocity changes.*

Analysis: To understand the limitations of this method, we model the scenario theoretically. Assume that the length of the smartphone is L, and the speed of the ultrasonic in the air is c, the airflow velocity at t_0 and t_1 can be represented $u_0 = \frac{L}{t_0} - c$ and $u_1 = \frac{L}{t_1} - c$, respectively. The change in airflow velocity can be expressed as: $\Delta u = u_1 - u_0 = L\frac{t_0 - t_1}{t_1 t_0}$. Assuming a zero airflow velocity at t_0, we establish $t_0 = \frac{L}{c}$, the change of airflow velocity can be expressed as $\Delta u = \frac{c^2}{-c + L/\Delta t}$. To detect changes in delay caused by airflow, the sampling rate should satisfy:

$$f_s \geq \frac{1}{\Delta t} \approx \frac{1}{L}\left(\frac{c^2}{\Delta u} + c \right). \tag{2}$$

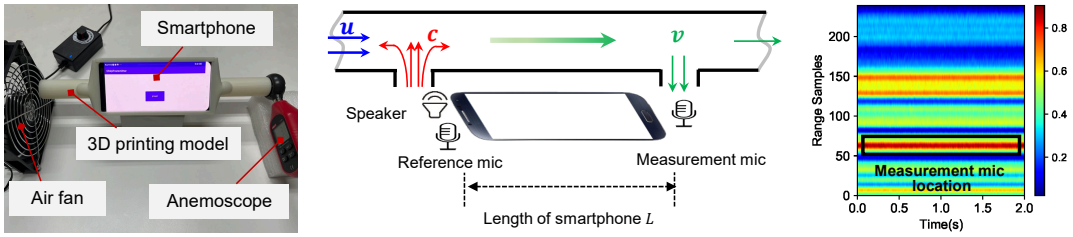
For a smartphone length L of 15cm and speed of sound c of 343m/s, if the desired airflow velocity resolution Δu is 1m/s, the required sampling rate is: $f_s \geq 786.613\,\text{kHz}$. This is approximately 16 times the typical sampling rate of current smartphones, which is 48 kHz.

3.2 Ultrasonic Phase-based Airflow Sensing

Observation 2: *Ultrasonic phase signals are sensitive to airflow velocity. As airflow increases, the IQ signal demonstrates more pronounced phase shifts and linear convergence.*

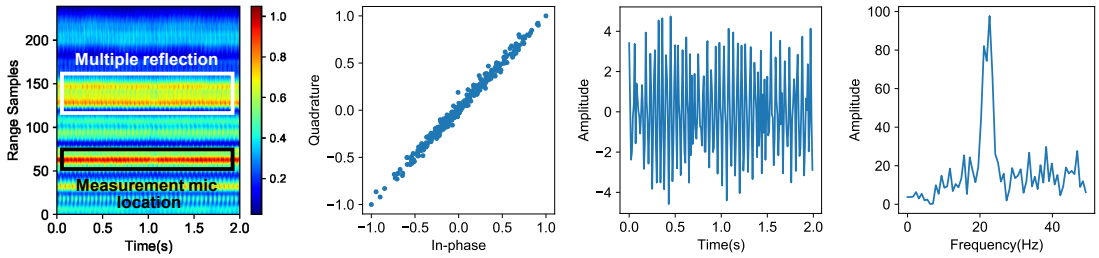
The resolution of the cross-correlation method for smartphone airflow velocity sensing is limited to 16.85 m/s, which is insufficient for fine-grained measurements required in practical applications. As a result, we explore the potential for improving the resolution by utilizing phase information extracted from ultrasonic signals. In our experiment setup, as shown in Fig. 3(a), we design a 3D-printed structure that controls the propagation path of both airflow and ultrasonic signals. As illustrated in Fig. 3(b), the underlying principle of this model is to measure airflow velocity by examining how airflow affects the propagation speed of ultrasonic signals within the 3D-printed model. We employ the default configuration to transmit ultrasonic signals and capture the signals using two microphones: one serving as a reference microphone and the other as a measurement microphone. To investigate the impact of airflow velocity on the phase of the ultrasonic signal, we multiply the measurement microphone's signal by the transmitted signal, followed by bandpass filtering to isolate the baseband signal. The accumulated baseband signal over time for airflow velocities of 0 and 3.2 m/s is shown in Fig. 3(c) and Fig. 4(a), respectively. These results demonstrate how airflow velocity influences the phase shift of the ultrasonic signal, offering insights into its effect on signal propagation.

Analysis: Fig. 4 shows the phase changes observed in the ultrasonic signal when the airflow velocity is 3.2 m/s. We first identify the range bin where the measurement microphone is located and extract the corresponding IQ signal, as shown in Fig. 4(b). We extract the phase information by combining the I and Q signals into a complex IF



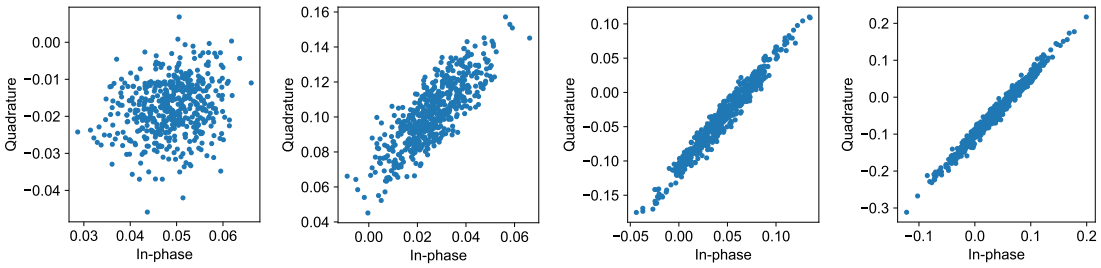
(a) Experiment setup. (b) Airflow velocity sensing modeling. (c) Concatenate chirp signals.

Fig. 3. Ultrasonic airflow velocity sensing based on phase features: (a) Experiment setup; (b) principle of ultrasonic airflow velocity sensing using a 3D-printed model; and (c) cumulative cross-correlation result at an airflow velocity of 0 m/s.



(a) Concatenate chirp signals. (b) IQ signal plan. (c) Phase difference signal. (d) FFT spectrum analysis.

Fig. 4. Ultrasonic airflow velocity sensing based on the phase feature when airflow velocity is 3.2 m/s: (a) Cumulative cross-correlation results; (b) normalized IQ plane signal result at the measurement microphone position; and (c) phase difference signal at the measurement microphone position. (d) FFT result of the phase difference signal.



(a) Airflow velocity is 1 m/s. (b) Airflow velocity is 2 m/s. (c) Airflow velocity is 3 m/s. (d) Airflow velocity is 4 m/s.

Fig. 5. IQ signal analysis of airflow velocity fields in the range of 1 m/s to 4 m/s.

signal, as shown in Fig. 4(c). This periodic behavior can then be analyzed using the Fast Fourier Transform (FFT) to extract frequency information, demonstrating that airflow velocity estimation is feasible through phase change analysis, as shown in Fig. 4(d). To further investigate the correlation between airflow velocity and the ultrasonic phase signal, we design comparative experiments to measure the phase change of the ultrasonic signal as the airflow velocity varies from 1 m/s to 4 m/s. Fig. 5 illustrates that we retained the variation amplitude feature of the original phase change data without normalization. *Experiment results indicate that the ultrasonic phase change gradually converges to linearity as the airflow velocity increases, with an expanding range of change. Therefore, airflow velocity estimation can be achieved by analyzing signal phase information.*

4 SONICFLOW: AIRFLOW SENSING MODELING

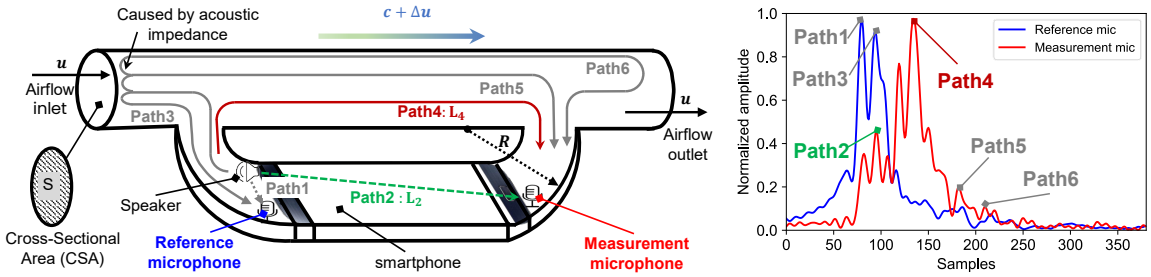
Building on the empirical observations detailed in Sec. 3.2, which demonstrate a clear relationship between airflow velocity and phase changes in the ultrasonic signal, we further investigate the underlying mechanisms by developing models. As shown in Fig. 6, we propose the *SonicFlow* model, designed to measure continuous pulmonary airflow velocity. This model utilizes a custom 3D-printed structure to guide and constrain the propagation of both airflow and ultrasonic signals.

4.1 Transmission and Reception of Ultrasonic Signals

In this model, we utilize a smartphone speaker to transmit FMCW ultrasonic signals, with two microphones employed to receive them. The transmitted chirp can be expressed as follows:

$$S_{\text{tx}}(t) = \cos\left(2\pi f_c t + \pi \frac{B}{T} t^2\right), \tag{3}$$

where f_c is the start frequency of the chirp, B is the bandwidth, and T is the period of the chirp. To facilitate the measurement and analysis of the relationship between airflow velocity and ultrasonic signals, we constrain their propagation paths using the 3D-printed model. Fig. 6(a) illustrates the principle behind the ultrasonic signal propagation path model based on the 3D-printed structural design. In this model, `Path1` and `Path2` represent the structure-borne propagation paths of ultrasonic signals, while the remaining paths correspond to the air-borne propagation paths. As shown in Fig. 6(a), the *left microphone*, located closer to the speaker, is designated as the *reference microphone*, and the ultrasonic signals it receives are defined as *reference signals*; The *left microphone*, located closer to the speaker, is designated as the *reference microphone*, and the ultrasonic signals it receives are referred to as *reference signals*. The microphone on the opposite side is designated as the *measurement microphone*, and the signals it receives are referred to as *measurement signals*. In dynamic scenarios with airflow present in



(a) Ultrasonic signal and airflow propagation model.

(b) Cross-correlation results.

Fig. 6. *SonicFlow* model: (a) Ultrasonic signal and airflow propagation model utilizing a 3D-printed structure; and (b) cross-correlation results between transmitted and received signals.

Path4, the reference signals $S_{\text{ref}}(t)$ and measurement signals $S_{\text{mes}}(t)$ can be expressed as:

$$S_{\text{ref}}(t) = S_{\text{tx}}(t) \cdot \Big[h_1(t) + h_3(t)\Big], \tag{4}$$

$$S_{\text{mes}}(t) = S_{\text{tx}}(t) \cdot \Big[h_2(t) + h_4(t) + h_5(t) + h_6(t)\Big], \tag{5}$$

where h_i denotes the Channel State Information (CSI) corresponding to the i-th propagation path. Taking into account that signal attenuation is influenced by the propagation path, we designate Path1 as the reference signal, with its CSI denoted as $h_{\text{ref}}(t) = h_1(t)$. Correspondingly, the signal attenuation of Path2 can be expressed as the signal of $h_1(t)$ delayed by Δt_2, i.e. $h_2(t) = h_{\text{ref}}(t - \Delta t_2)$. Similarly, the signal attenuations of Path3, Path4, Path5, and Path6 are represented as delayed versions of $h_{\text{ref}}(t)$ by their respective delays Δt_3, Δt_4, Δt_5, and Δt_6. Therefore, the measurement signal $S_{\text{mes}}(t)$ in Eq. 5 can be expressed as:

$$S_{\text{mes}}(t) = S_{\text{tx}}(t) \cdot \underbrace{\Big[h_{\text{ref}}(t - \Delta t_2) + h_{\text{ref}}(t - \Delta t_4) + h_{\text{ref}}(t - \Delta t_5) + h_{\text{ref}}(t - \Delta t_6)\Big]}_{h_{\text{mes}}}, \tag{6}$$

where, $h_{\text{ref}}(t - \Delta t_2)$ represents the CSI of propagation Path2 without airflow information, $h_{\text{ref}}(t - \Delta t_4)$ represents the CSI of propagation Path4 which contains airflow information. The terms $h_{\text{ref}}(t - \Delta t_5)$ and $h_{\text{ref}}(t - \Delta t_6)$ represent multiple reflections caused by acoustic impedance at the tube's entrance and exit. To quantify the differences between h_2 and h_4, we compute the cross-correlation between the transmitted signal and the signals received by the two microphones. As depicted in Fig. 6(b), the blue curve represents the cross-correlation of the transmitted signal $S_{\text{tx}}(t)$ with the reference signal $S_{\text{ref}}(t)$, while the red curve shows its cross-correlation with the measurement signal $S_{\text{mes}}(t)$.

4.2 Extracting CSI-ratio

Our purpose is to obtain the signal attenuation $h_{\text{ref}}(t - \Delta t_2)$ and $h_{\text{ref}}(t - \Delta t_4)$, which are caused by Path2 and Path4, as shown in Fig. 6(a). As shown in Fig. 6(b), the signal energy intensity of Path5 and Path6 is very weak due to multiple reflections. Therefore, based on Eq. (6), the signals arriving from Path2 (i.e., $S_{\text{mes}}(t - \Delta t_2)$) and Path4 (i.e., $S_{\text{mes}}(t - \Delta t_4)$) can be expressed as:

$$\begin{aligned} S_{\text{mes}}(t - \Delta t_2) &= S_{\text{tx}}(t) \cdot h_{\text{ref}}(t - \Delta t_2), \\ S_{\text{mes}}(t - \Delta t_4) &= S_{\text{tx}}(t) \cdot h_{\text{ref}}(t - \Delta t_4). \end{aligned} \tag{7}$$

To extract the CSI values $h_{\text{ref}}(t-\Delta t_2)$ and $h_{\text{ref}}(t-\Delta t_4)$, we multiply the measured signals from Path2 and Path4 by the transmitted signal, then eliminate high-frequency components using low-pass filtering. This process is expressed as:

$$\begin{aligned} h_{\text{ref}}(t-\Delta t_2) &= \mathcal{F}\left[S_{\text{tx}}(t)\cdot S_{\text{mes}}(t-\Delta t_2)\right], \\ h_{\text{ref}}(t-\Delta t_4) &= \mathcal{F}\left[S_{\text{tx}}(t)\cdot S_{\text{mes}}(t-\Delta t_4)\right], \end{aligned} \tag{8}$$

where $\mathcal{F}[\cdot]$ represents a low-pass filter operation with a default cutoff frequency of 6 kHz, and $S_{\text{tx}}(t)$ represents the transmitted signal. The received signal for Path2 is a delayed version of the transmitted signal by Δt_2, and can be expressed as:

$$S_{\text{mes}}(t-\Delta t_2) = \cos\left[2\pi f_c(t-\Delta t_2) + \pi\frac{B}{T}(t-\Delta t_2)^2\right]. \tag{9}$$

We employ IQ demodulation to acquire complex CSI. This approach extracts both the I and Q components of the received signal, capturing detailed information about the phase and amplitude, which is critical for accurately characterizing CSI and extracting features related to airflow velocity. According to Eq. (8), the I and Q components of the Path2 received signal are given by:

$$\begin{aligned} I_{\text{path2}} &= \mathcal{F}\left[S_{\text{mes}}(t-\Delta t_2)\cdot S_{\text{tx}}(t)\right] = \frac{1}{2}\cos\Big(-2\pi f_c\Delta t_2 - 2\pi\frac{B}{T}\Delta t_2 t + \pi\frac{B}{T}\Delta t_2^2\Big), \\ Q_{\text{path2}} &= \mathcal{F}\left[S_{\text{mes}}(t-\Delta t_2)\cdot \mathcal{H}\left[S_{\text{tx}}(t)\right]\right] = \frac{1}{2}\sin\Big(-2\pi f_c\Delta t_2 - 2\pi\frac{B}{T}\Delta t_2 t + \pi\frac{B}{T}\Delta t_2^2\Big), \end{aligned} \tag{10}$$

where $\mathcal{H}\{\cdot\}$ represents the Hilbert transform [2]. Using Euler's formula, the complex CSI for Path2 and Path4 can be expressed as:

$$\begin{aligned} h_2(t) &= \frac{1}{2}e^{-i\left(2\pi\Delta t_2\frac{B}{T}t + 2\pi f_c\Delta t_2 - \pi\frac{B}{T}\Delta t_2^2\right)}, \\ h_4(t) &= \frac{1}{2}e^{-i\left(2\pi\Delta t_4\frac{B}{T}t + 2\pi f_c\Delta t_4 - \pi\frac{B}{T}\Delta t_4^2\right)}. \end{aligned} \tag{11}$$

To eliminate static information and derive airflow velocity, we calculate the CSI ratio r_{h}between Path4 and Path2, based on Eq. (11):

$$r_{\text{h}} = \frac{h_4(t)}{h_2(t)} = e^{-i\left(2\pi(\Delta t_4-\Delta t_2)\frac{B}{T}t + 2\pi f_c(\Delta t_4-\Delta t_2) - \pi\frac{B}{T}(\Delta t_4^2-\Delta t_2^2)\right)}. \tag{12}$$

Here, $\pi\frac{B}{T}(\Delta t_4^2-\Delta t_2^2)$ is a second-order term that can be neglected. Considering the difference in propagation path lengths between Path4 and Path2, we introduce a bias delay time τ. Thus, the CSI ratio is approximated as:

$$r_{\text{h}}(t) \approx \exp\Big[-i\Big(\underbrace{2\pi(\Delta t_4-\Delta t_2-\tau)\frac{B}{T}t}_{\text{IF signal}} + \underbrace{2\pi f_c(\Delta t_4-\Delta t_2-\tau)}_{\text{Phase }\Phi}\Big)\Big]. \tag{13}$$

Here, $2\pi(\Delta t_4-\Delta t_2-\tau)\frac{B}{T}t$ represents the intermediate frequency (IF) signal associated with the microphone measurement range, while $2\pi f_c(\Delta t_4-\Delta t_2-\tau)$ corresponds to the phase Φ, related to airflow velocity.

4.3 Extracting Airflow Velocity

According to Eq. (13), it is evident that the IF component in the CSI-ratio depends on both Path2 and Path4, as well as time. However, the phase component Φ is exclusively influenced by Path2 and Path4. Thus, leveraging the phase component Φ offers a more reliable means of calculating airflow velocity. Based on the structure illustrated

in Fig. 6(a), where $\Delta t_2 = \frac{L_2}{c'}$ and $\Delta t_4 = \frac{L_4}{c}$, the relationship between phase rotation $\Delta\Phi$ and changes in airflow velocity Δu can be expressed as follows:

$$\Delta\Phi = 2\pi f_c \left(\frac{L_4}{c + \Delta u} - \frac{L_2}{c'} - \tau \right), \tag{14}$$

where f_c is the starting frequency of the chirp signal, c is the speed of ultrasonic in air-borne media, and c' is the speed of ultrasonic in structure-borne media, as defined in Eq. (1). Here, Δu denotes the change in airflow velocity, and $\tau = \frac{L_4 - L_2}{c}$ represents the time bias caused by the difference in propagation path lengths between `Path4` and `Path2` in a static scene. For instance, given that the $L_4 = 0.322$ m, $L_2 = 0.165$ m, $c = 34$ m/s ($25^\circ C$), and $c' = 1497.1$ m/s, when the airflow velocity Δu increases by 1 m/s, the phase rotation $\Delta\Phi$ becomes 38.88 rad. Furthermore, given the cross-sectional area $\mathcal{A}$ of the tube, the airflow velocity can be determined as follows:

$$\Delta Q_{\text{air}}(t) = \left[\frac{L_4}{\frac{\Delta\Phi(t)}{2\pi f_c} + \frac{L_2}{c'} + \tau} - c \right] \cdot \mathcal{A}. \tag{15}$$

Here, L_2 denotes the distance between the smartphone's top speaker and bottom microphone (i.e., the smartphone's length), and L_4 represents the signal propagation path length through the 3D-printed tube model. The path length L_4 can be calculated as $L_4 = \pi R + L_2$, where R is the corner arc radius, as shown in Fig. 6(a). $\mathcal{A}$ denotes the cross-sectional area of the 3D-printed model tube. These parameters can be determined once the smartphone and the 3D-printed model are finalized. Thus, the change in airflow rate $\Delta Q_{\text{air}}(t)$ (m^3/s) can be obtained by measuring the phase change $\Delta\Phi$. However, achieving high-precision phase change measurements presents challenges. Specifically, it is crucial to accurately capture fine-grained phase changes when the transmitting and receiving signals are not perfectly synchronized. Additionally, interference from environmental noise and tube airflow noise must be addressed.

5 SYSTEM DESIGN

We propose *SpiroSense*, a groundbreaking system that transforms a standard smartphone into a household pulmonary examination device, bridging the gap between clinical requirements and existing sensing methodologies. The system overview, shown in Fig. 7, comprises three key stages: *i) Signal synchronization and alignment*, ensuring coherent demodulation of the received signal; *ii) Noise elimination and phase signal extraction*, leveraging the *NoiseClear* model to suppress high-frequency harmonic noise caused by airflow turbulence within the device and employing IQ synchronous demodulation to extract ultrasonic phase characteristics; and *iii) Lung condition indicators estimation*, utilizing the *SonicFlow* model to convert phase characteristics into precise measurements of lung airflow velocity and associated pulmonary metrics.

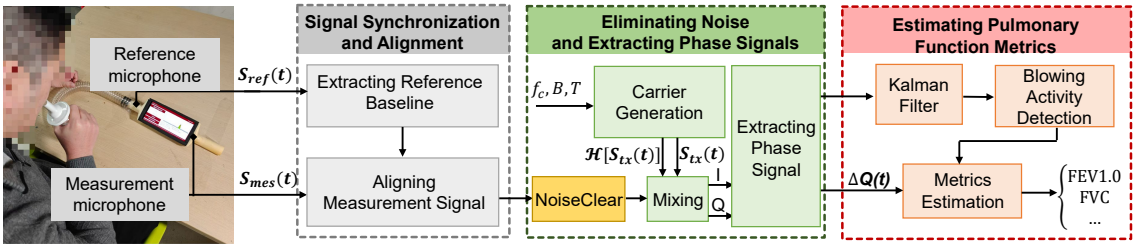


Fig. 7. Overview of the *SpiroSense* system, which transforms a smartphone into a pulmonary examination device.

5.1 Signal Synchronization and Alignment

Previous ultrasonic airflow velocity measurement methods primarily rely on cross-correlation to estimate the ToF between the transmitted and echo signals, enabling airflow velocity estimation. However, as discussed in Sec. 3.1, this approach faces limitations for fine-grained airflow velocity measurements on smartphones due to the relatively low microphone sampling rate (typically 48 kHz). Leveraging the relationship between airflow velocity and ultrasonic signal phase rotation established in the *SonicFlow* model (Sec. 4), we synchronize and demodulate the echo signals to extract the baseband phase information and subsequently derive airflow velocity characteristics. Achieving synchronization between transmitted and received signals is challenging because commercial smartphones lack a signal-synchronous interface between their speakers and microphones.

To overcome this, we propose a reference signal synchronization approach based on the cross-correlation method, as shown in Fig. 8. This method calculates the cross-correlation between the reference and transmitted signals to establish a baseline for aligning the measurement signal. Since the start time of the received signal at the measurement microphone cannot be precisely determined, a reference signal is used as a baseline. Ideally, there would be no cross-correlation delay between the transmitted signal and the reference signal received by the reference microphone. However, as shown in Fig. 9, a gap exists between the cross-correlation results of the reference microphone and the ideal result. To address this, we calculate the cross-correlation offset between the reference signal and the transmitted signal using the following equation: offset = $(p + \frac{l}{2}) \mod l$, where p represents the peak position of the cross-correlation result, and l is the length of the transmitted signal. Once the offset is determined, the measured signal is adjusted by removing the initial incomplete chirp segment, ensuring proper alignment for further processing. As illustrated in Fig. 10, this adjustment guarantees correct alignment between the measured and transmitted signals.

5.2 Eliminating Noise and Extracting Phase Signals

Once the measurement signal is properly aligned, the next step is to perform IQ demodulation on the received signal according to Eq. (10). This process involves extracting the in-phase (I) and quadrature (Q) components of

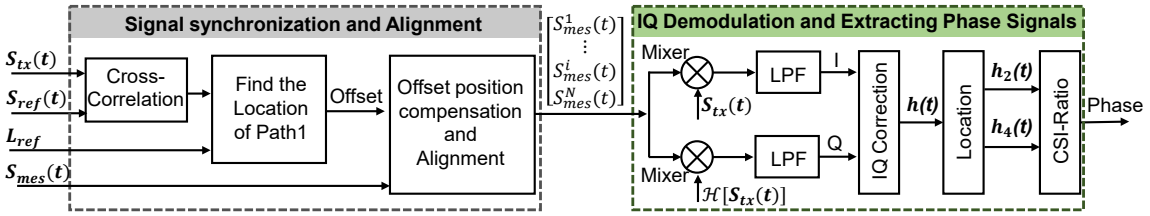
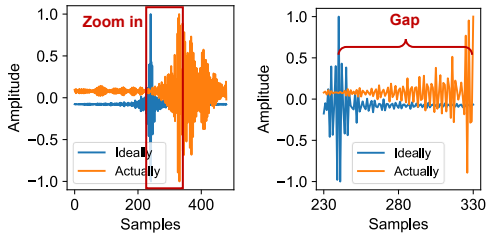
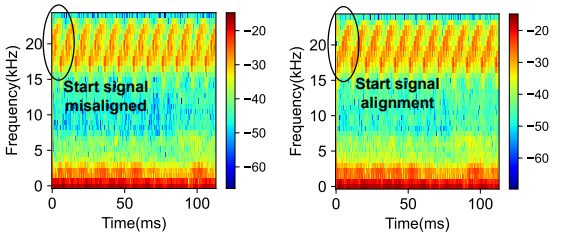


Fig. 8. Pipeline of signal synchronization and IQ demodulation.



(a) Cross-correlated signals. (b) Zoom in signals.

Fig. 9. Ideal and actual measured cross-correlation signals.



(a) Before alignment. (b) After alignment.

Fig. 10. Measurement signal alignment.

the signal. By applying Euler's formula, we can calculate the phase angle of the complex signal, which provides valuable insights into airflow velocity. However, several challenges arise when a user exhales into the tube. High-frequency harmonics generated by airflow sounds can significantly distort the received signal. Furthermore, ultrasonic signals are highly sensitive to interference from ambient sounds commonly encountered in everyday environments, such as car horns, televisions, and household appliances. These external noises often generate harmonic frequencies similar to the ultrasonic signal, disrupting both its propagation and reception. To mitigate these challenges before IQ-demodulation, we propose the *NoiseClear* model, as illustrated in Fig. 11, to eliminate ambient and harmonic noise effectively.

5.2.1 Ambient Noise Eliminating Based on NoiseClear. When air is blown into the tube, turbulence and eddy currents generate airflow noise, as shown in Fig. 12. This figure illustrates the spectrum after applying a high-pass filter with a cutoff frequency of 15 kHz. The noise spans frequencies that overlap with those of the ultrasonic signal, significantly hindering the accurate detection of airflow characteristics. To overcome this challenge, we focus on enhancing the ultrasonic signal in a manner compatible with smartphone systems. Specifically, we model the temporal ultrasonic signal $x \in \mathbb{R}^{\mathrm{T}}$ as a combination of the clean ultrasonic signal $y \in \mathbb{R}^{\mathrm{T}}$, which carries airflow features, and ambient harmonic noise $n \in \mathbb{R}^{\mathrm{T}}$, represented as: $x(t) = y(t) + n(t)$. Our goal is to define an enhancement function f that eliminates the harmonic noise, such that $y \approx f(x)$. To achieve this, we introduce *NoiseClear*, a novel ultrasonic signal enhancement model inspired by DEMUCS [11], as depicted in Fig. 11. Using an encoder-decoder architecture, the model is designed to eliminate high-frequency harmonic noise on smartphones. It consists of a convolutional encoder, a Long Short-Term Memory (LSTM) module, and a convolutional decoder, with the encoder and decoder linked via skip connections. Both the encoder and decoder consist of five layers, arranged in reverse order to ensure that the corresponding layers in the decoder align with those in the encoder at the same scale and index.

The signal processing flow is as follows: *High-pass Filtering*: We apply high-pass filtering with a 15 kHz cutoff frequency to remove low-frequency speech signals, ensuring the input only includes ultrasonic signals in the 17 kHz–23 kHz range. *Feature Extraction and Compression*: The normalized input signal S_{in} is encoded into a 768-dimensional feature vector $F_{\text{in}}^{\text{lstm}}$, which captures and compresses time-series ultrasonic features. *Signal Dynamics Modeling*: We use a 2-layer LSTM module to model long-term time dependencies in the input feature vector $F_{\text{in}}^{\text{lstm}}$, improving the understanding of the signal dynamics. *Signal Reconstruction*: We employ a 5-layer

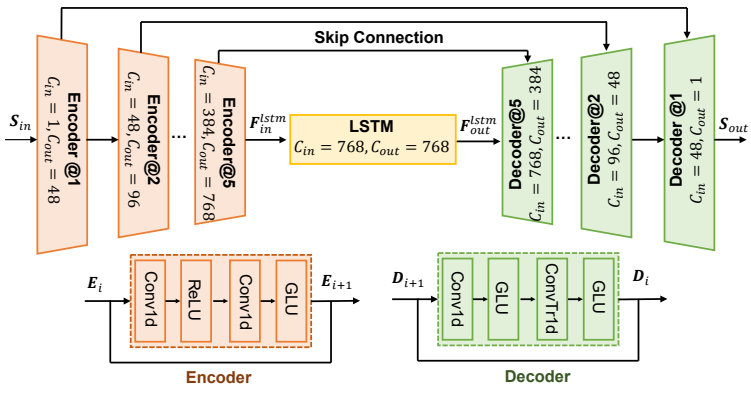


Fig. 11. Architecture of the *NoiseClear* ultrasonic signal enhancement network, featuring an encoder-decoder structure for suppression of ambient noise.

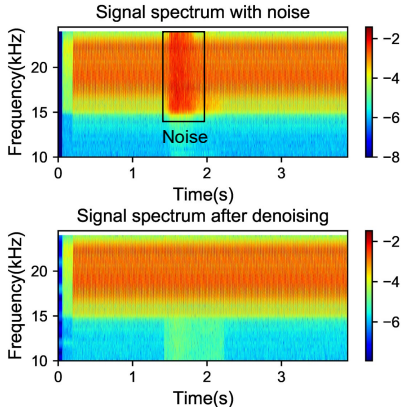
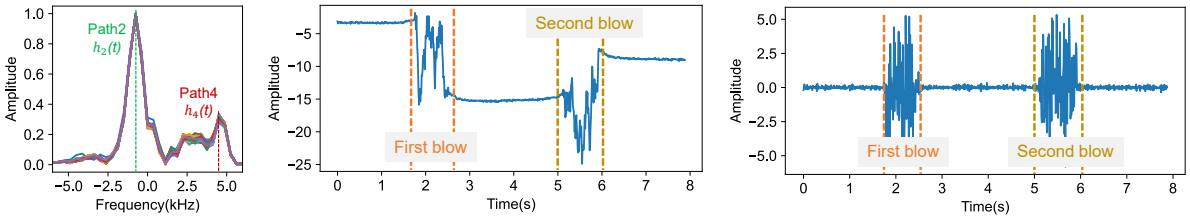


Fig. 12. Frequency spectrum comparison before and after *NoiseClear* processing.



(a) FFT results of IF signal. (b) Extracted phase signal. (c) Phase differential signal.

Fig. 13. IQ demodulation and phase feature extraction process: (a) FFT results of the IF signal used to identify propagation paths; (b) extracted phase signals from the IF components; and (c) phase differential signals derived for airflow velocity characterization.

decoder network to decode the feature vector of LSTM output $F_{\text{out}}^{\text{lstm}}$ into an audio signal S_{out}. The effectiveness of *NoiseClear* is demonstrated in Fig. 12, where the signal spectrum before and after processing showcases the model's strong denoising capabilities.

Training details. We train *NoiseClear* on two NVIDIA GeForce RTX 3090 GPUs for 400 epochs, observing significant convergence. The Adam optimizer is used to compute model gradients, with a learning rate set to 3×10^{-4} and a batch size of 32. For dataset partitioning, we randomly select 70% of the data for training, 15% for validation (to monitor model performance during training), and 15% for testing (to evaluate the final performance of *NoiseClear*).

5.2.2 IQ Demodulation and Phase Feature Extraction. To extract the phase features from the received signal, we perform IQ demodulation on the denoised chirp signal S_{mes}, as described in Eq. (10), to derive the Intermediate Frequency (IF) signal. This process involves separating the In-phase (I) and Quadrature (Q) components of the received signal and performing additional signal processing to extract the relevant features. As illustrated in Fig. 8, each aligned measurement signal S_{mes}^i is multiplied by the transmitted signal S_{tx} and its *Hilbert* transform $\mathcal{H}[S_{\text{tx}}(t)]$, respectively. Following this, we apply low-pass filtering to obtain the baseband IF signal for each chirp, denoted as S_{IF}^i. Next, we perform an FFT operation on the complex IF signal of each chirp to estimate the locations of `Path2` and `Path4`, as illustrated in Fig. 13(a). The corresponding phase signals $h_2(t)$ and $h_4(t)$ are extracted by concatenating the IF signal samples at the `Path2` and `Path4` positions. For instance, if the period of the transmitted signal is 0.01 seconds, the phase signal extraction is performed at a sampling rate of 100 Hz. Fig. 13(b) shows the phase signal extracted from two consecutive air blows, which highlight significant phase features during air blowing. Subsequently, as described in Eq. (14), we establish the relationship between the phase differential signal and airflow velocity. By performing differential operations on the phase signal, we can describe the airflow velocity characteristics $\Delta\Phi$, as shown in Fig. 13(c).

5.3 Pulmonary Function Metrics Estimation

As shown in Fig. 14, the process for obtaining pulmonary function metrics from the phase signal is illustrated. Based on the findings in Sec. 3.2, where airflow velocity is related to phase frequency, we compute the airflow velocity by analyzing the frequency energy distribution of the phase signal. Eq. (15) defines the relationship between phase frequency and airflow velocity. According to Parseval's theorem, the energy in the time-domain phase signal is equivalent to that in the frequency domain, meaning that the sum (or integral) of the square of a function is equal to the sum (or integral) of the square of its transform. To derive the signal's short-term energy distribution, we segment the phase signal and perform FFT operations on each segment to obtain its frequency-domain representation. We construct the airflow velocity-time curve by summing the energy across different frequency components simultaneously, creating an energy distribution curve. To smooth the airflow

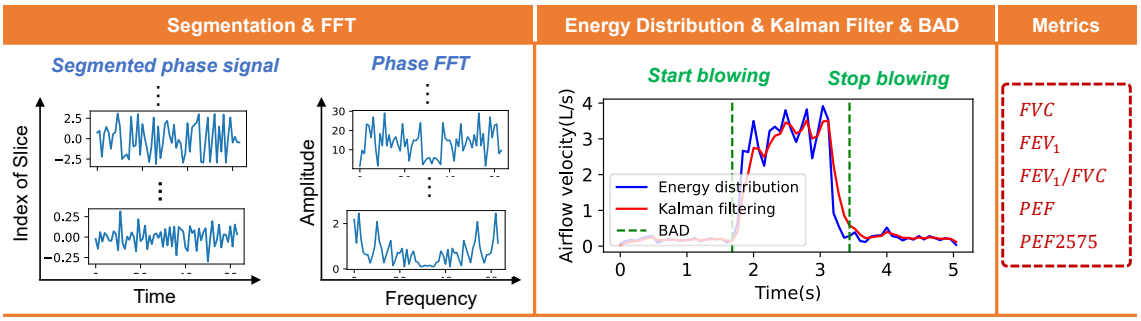


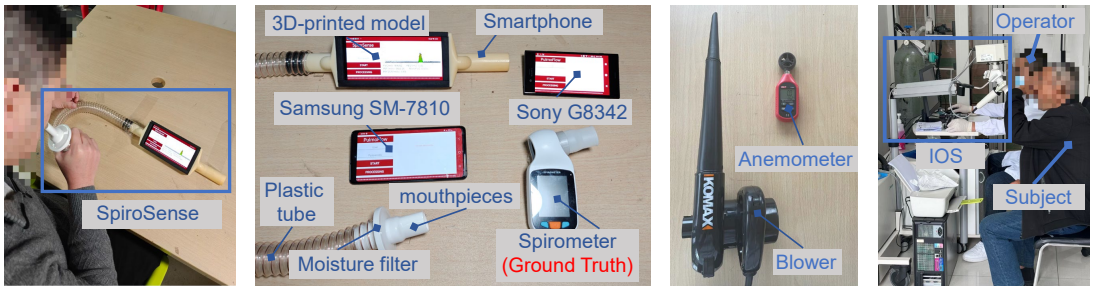
Fig. 14. Pipeline for estimating pulmonary function metrics from phase signals, including phase segmentation, frequency-domain analysis, airflow velocity curve construction, and pulmonary parameter calculation.

velocity and enhance the detection of blowing activity, we apply Kalman filtering to the airflow velocity-time curve. Next, the sliding standard deviation method is used to detect the filtered signal and estimate the airflow activity range (i.e., the start and end times of the blowing), which is crucial for calculating the corresponding lung function metrics. Finally, we estimate pulmonary function metrics based on the airflow velocity-time curve. For example, FVC can be obtained by integrating the airflow velocity-time curve. Similarly, FEV_1 is derived by integrating the airflow velocity-time curve over the first second from the start of blowing. The FEV_1/FVC ratio is a commonly used indicator to diagnose asthma and COPD. For example, COPD is primarily characterized by difficulty in exhaling, leading to the FEV_1/FVC ratio of less than 70%.

6 EVALUATION

6.1 Experiment Setup

We develop the *SpiroSense* application for smartphones to analyze received signals and measure airflow velocity, enabling estimation of PFT metrics. The experiment setup, illustrated in Fig. 15, includes smartphones, 3D-printed models, plastic tubes, moisture filters, and mouthpieces, effectively creating a home-based lung function testing device. Following established clinical standards [9, 54], we standardize the measurement process to ensure reliable and repeatable results. The recommended testing procedure and default conditions are as follows: a) *SpiroSense* is intended for indoor use, where ambient acoustic and airflow conditions are more stable and controlled; b) Testing should be conducted in environments with still or minimal airflow (e.g., air conditioning). Strong air currents (e.g., from fans or open windows directly facing the subject) should be avoided, as they may disrupt both the airflow of exhaled and ultrasonic signal propagation; c) Subjects should remain stationary, ideally seated upright, and hold the smartphone at a fixed position using the provided 3D-printed model to maintain consistent measurement geometry, as shown in Fig. 15(a); d) Subjects should firmly seal the mouthpiece with their lips, wear a nose clip to prevent nasal exhalation, and then exhale forcefully and completely through the tube, emulating the expiratory effort required in clinical spirometry. We adopt this setup not only for subject testing but also for evaluating *SpiroSense*'s performance under varying indoor conditions, including background noise (e.g., conversations, television sounds, traffic) and different smartphones, to validate the robustness of the system in realistic home-use scenarios. To evaluate *SpiroSense*, we conduct experiments on six smartphone models, including the Samsung SM-7810, Sony XPERIA XZ1, Huawei P30 Pro, and Xiaomi 10 and 11, with the default volume configured at 20%. We recruited 29 healthy participants (11 females and 18 males), aged between 18 and 33 years, with Body Mass Indices (BMIs) ranging from 15.8% to 32.4%. For the case study, we recruited an additional 30 participants diagnosed with lung diseases such as COPD or asthma (9 females and 21 males), aged between 29 and 83 years, with BMIs ranging from 15.6% to 29.7%. To reduce experiment errors caused by the



(a) Measurement scenarios. (b) Components of experiment assessment. (c) Dataset device. (d) Gold-standard IOS.

Fig. 15. Experiment setup for lung function assessment using *SpiroSense*: (a) Experiment scenario description; (b) the essential hardware components, including a smartphone, a 3D-printed model, a plastic tube, a moisture filter, and a mouthpiece; (c) additional devices include a variable-speed blower for generating pure airflow signals and an anemometer for measuring airflow; and (d) the gold-standard Impulse Oscillometry System (IOS) for PFT.

manipulation, we instruct each subject to perform three air-blowing activities for each test. Ground truth data are collected using conventional household devices, such as the CONTEC SP70B spirometer [8] (Fig. 15(b)). For the clinical case study, ground truth data are collected using the gold-standard Impulse Oscillometry System (IOS) [33], as illustrated in Fig. 15(d). This study is conducted in accordance with the ethical standards and guidelines for human-subject research. All participants are fully informed about the purpose and procedures of the study and provide their written informed consent prior to participation. Participation is voluntary, and participants are free to withdraw at any time without penalty. The study protocol is reviewed and approved by our ethical Institutional Review Board (IRB) and the hospital's IRB, ensuring compliance with institutional and international ethical guidelines. All collected data are anonymized to protect the privacy and confidentiality of the participants. The experiment measurement protocol used with *SpiroSense* complies with medical standards.

Software. We develop SpiroSense's user interface and manage microphone and speaker operations for ultrasonic signal transmission and reception using Java. The ultrasonic signal processing algorithms are implemented in C/C++ and compiled into a native Android library using the Android Native Development Kit (NDK). For the ultrasonic *NoiseClear* model, we convert the trained TensorFlow model into the TensorFlow Lite (TFLite) format [29] for efficient on-device inference. The native library is integrated into the application and invoked at runtime via Android's Java Native Interface (JNI).

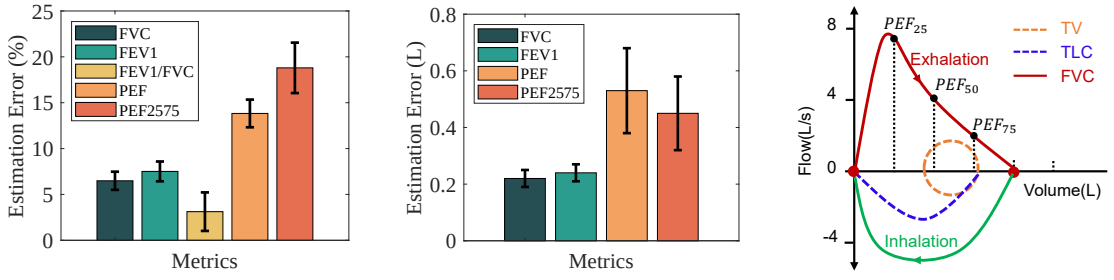
Dataset. To develop the end-to-end ultrasonic denoising model, *NoiseClear*, we construct a comprehensive and reproducible dataset by systematically capturing ultrasonic signals under controlled experimental conditions. As illustrated in Fig. 15(c), we utilize a variable-speed blower to generate stable and interference-free ultrasonic airflow signals, which serve as the clean target outputs for model training. To simulate real-world acoustic disturbances, we recruit ten volunteers (i.e., approximately one-sixth of the total participants) to collect a diverse set of environmental noise recordings, including air-blowing sounds and various ambient interferences encountered in daily life. These recorded noise samples are subsequently superimposed onto the clean airflow signals to synthesize noisy ultrasonic inputs. All signal generation and recording are performed using a Xiaomi 11 smartphone, with the sampling rate configured at 48 kHz, ensuring high-fidelity signal capture and consistent data quality. The final dataset comprises 10,000 input-output sample pairs, each with a duration of 8 seconds, consisting of a noisy input and its corresponding clean reference (output). This systematic data generation process enables the construction of high-quality, temporally aligned training pairs for supervised learning. By incorporating a wide range of noise profiles and adhering to a controlled and repeatable data collection procedure, the dataset ensures both diversity and consistency. This design enhances the robustness and generalizability of

the trained model in real-world applications. All procedures and parameters are carefully documented to support reproducibility and facilitate future benchmarking and evaluation efforts.

Metrics. To effectively assess the performance of *SpiroSense*, we employ both relative and absolute error metrics across key pulmonary function indicators, including FVC, FEV_1, and FEV_1/FVC ratio. The relative estimation error for FVC is computed as: $FVC_{err} = \frac{|FVC_{mes} - FVC_{gt}|}{FVC_{gt}} \times 100\%$, where FVC_{mes} denotes the measured value and FVC_{gt} represents the corresponding ground truth. The absolute estimation error for FVC is defined as: $FVC_{abs\_err} = |FVC_{mes} - FVC_{gt}|$. In addition, we calculate the Standard Error of the Mean (SEM) to assess the variability and stability of the estimation errors across multiple measurements. To provide a comprehensive assessment, we perform a statistical analysis of five key pulmonary function indicators collected from all participating volunteers. This multifaceted evaluation enables us to quantify both the accuracy and consistency of *SpiroSense*'s lung function estimation under realistic usage scenarios.

6.2 Overall Performance

Experimental results demonstrate that SpiroSense achieves high accuracy in estimating key pulmonary function indicators. Specifically, the absolute estimation errors are 0.22L for FVC and FEV_1*, corresponding to relative errors of 6.49% and 7.51%, respectively.* These results highlight SpiroSense's high accuracy and reliability, making it a promising solution for portable and accessible lung function monitoring. The FEV_1/FVC ratio is a clinically significant metric for diagnosing obstructive and restrictive lung conditions. As shown in Fig. 16, the relative estimation error for FEV_1/FVC is 2.72%, further validating the effectiveness of our approach. We also evaluate other key pulmonary function metrics. The absolute estimation errors are: FVC is 0.22L, FEV_1 is 0.24L, PEF is 0.53L, and the PEF during the 25–75% portion of the expiratory curve (PEF_{25-75}) is 0.45L. Correspondingly, the relative estimation errors are: FVC is 6.49%, FEV_1 is 7.51%, PEF is 12.1%, and PEF_{25-75} is 15.9%. As shown in Fig. 16(c), PEF measures the maximum airflow during forced expiration and is sensitive to variations in subject effort, posture, and lung capacity, with normal ranges typically accounting for ±20% of the expected value. Likewise, PEF_{25-75}–which captures the average flow rate during the mid-portion of exhalation–is susceptible to individual physiological differences, often resulting in greater estimation variability. Despite these challenges, *SpiroSense* consistently achieves low estimation errors and small SEM for critical indicators such as FVC and FEV_1, confirming its robustness and reliability in practical scenarios. According to the 2019 spirometry standards from the American Thoracic Society (ATS) and European Respiratory Society (ERS) standards [15], an acceptable difference between repeated spirometry measurements should fall within 0.15L for FEV_1 and FVC. *While SpiroSense's performance may not yet match that of clinical-grade spirometers, its estimation errors fall within the permissible range for use in pre-screening, home-based monitoring, and early detection of pulmonary conditions. This aligns with recent advances advocating the use of digital health tools in telemedicine and chronic disease self-management* [41, 49].



(a) Relative estimate error. (b) Absolute estimate error. (c) Flow-volume curve of FPT.

Fig. 16. The overall performance of measuring lung function indicators based on *SpiroSense*.

Table 1. Comparison with SOTA methods.

Method	FVC ↓	FEV_1 ↓	FEV_1/FVC ↓	PEF ↓	PEF_{25-75} ↓
SpiroSmart [25]	9.67 ± 2.3	9.67 ± 3.3	11.6 ± 5.3	10.3 ± 3.6	-
Spirocall [14]	7.18 ± 4.44	7.25 ± 4.18	9.15 ± 5.27	9.05 ± 6.59	-
PTEase [47]	-	11.31 ± 0.7	15.18 ± 0.52	-	-
SpiroSense	6.49 ± 0.99	7.51 ± 1.08	2.72 ± 0.24	12.1 ± 1.52	15.9 ± 2.2

6.3 Comparison with SOTA Methods

Experimental results reveal that SpiroSense *achieves estimation errors below 8% for key PFT indicators, including* FVC, FEV_1, *and* FEV_1/FVC *ratio, significantly outperforming state-of-the-art (SOTA) smartphone-based approaches in both accuracy and comprehensiveness.*

To rigorously evaluate the performance of *SpiroSense*, we benchmark it against leading SOTA smartphone-based systems under identical experimental conditions (detailed in Section 6.1). Table 1 presents the comparative results, reporting relative errors for FVC, FEV_1, PEF, and PEF_{25-75}, and absolute errors for the FEV_1/FVC ratio. SpiroSmart [25] and SpiroCall [14] estimate airflow by capturing exhalation sounds via the phone's microphone and inferring lung function using a vocal tract model. While convenient, these sound-based approaches are inherently sensitive to ambient noise such as speech, television, and traffic, limiting their accuracy in real-world environments. In contrast, *SpiroSense*'s use of active ultrasonic signals ensures robustness to environmental interference. Compared with SpiroSmart, *SpiroSense* not only achieves lower average errors (e.g., reducing FVC error by 3.18%) across all metrics but also exhibits significantly reduced variance, only one-fifth of SpiroSmart's, highlighting its stability in diverse conditions. Although *SpiroSense* exhibits slightly higher average errors in specific indicators, such as FEV_1 and PEF, its significantly smaller standard deviation compared with SpiroCall suggests that *SpiroSense* provides more stable measurements in real-world scenarios. This robustness stems from *SpiroSense*'s sensing mechanism, which emits and receives ultrasonic signals that are immune to most environmental noise. For a detailed analysis of ambient noise effects, please refer to Section 6.6.

Compared with PTEase [47], which also leverages smartphone hardware but relies on air-pressure cues and simpler signal processing, *SpiroSense* delivers markedly improved estimation accuracy for FEV_1 and the FEV_1/FVC ratio. Notably, *SpiroSense* achieves a FEV_1/FVC ratio estimation error that is 12.46% lower than PTEase and 6.43% lower than SpiroCall. Moreover, unlike PTEase, which focuses on limited indicators, *SpiroSense* supports a broader range of lung function measurements, reinforcing its practical value for comprehensive pulmonary health monitoring.

6.4 Impact of Individual Differences on Estimation Error

Experimental results indicate that individual differences can impact detection errors, but these deviations can be minimized through standardized measurement procedures, thus enhancing the robustness and generalizability of SpiroSense. To assess the impact of individual differences on *SpiroSense*'s performance, we used a Xiaomi 11 smartphone and invited 4 female (F) and 6 male (M) health volunteers. We select two indicators for lung disease detection, FEV_1, and FEV_1/FVC ratio, to evaluate *SpiroSense*'s performance. As shown in Fig. 17, these results demonstrate that more than 70% of participants had errors below the average. The primary cause of detection errors between individuals was found to be leakage due to non-standardized operation methods. For example, operational errors by human subjects M1 and M6 resulted in airflow leakage during the measurement process, leading to an increase in measurement error. This deviation can be minimized by standardizing the measurement process and ensuring consistent operational procedures across individuals.

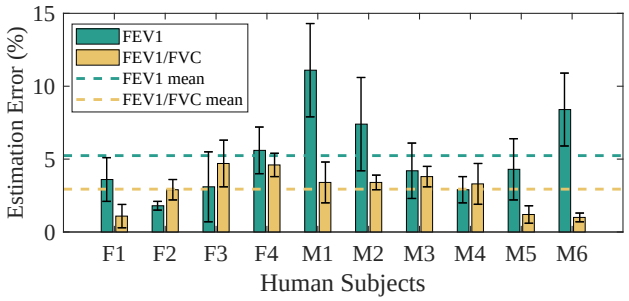


Fig. 17. Estimation errors for different participants.

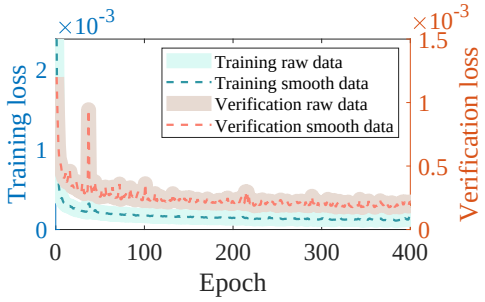


Fig. 18. Loss curves over epochs.

6.5 Performance Evaluation of *NoiseClear* Model

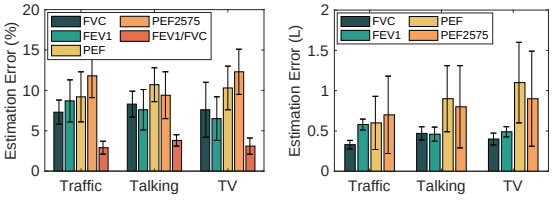
Experimental results demonstrate that the NoiseClear model achieves rapid convergence and maintains stable performance, effectively addressing airflow and ambient noise interference. To mitigate the airflow and ambient noise interference generated by the tube during air blowing, we developed *NoiseClear*, an end-to-end ultrasonic noise denoising model. Fig. 18 shows the loss curves for both the training and validation datasets during the training process. Experimental results show that as the number of training iterations increases, the loss of the *NoiseClear* model decreases rapidly for both datasets and eventually stabilizes after approximately 150 epochs.

6.6 Impact of Ambient Noise

Experimental results confirm that SpiroSense *demonstrates robust performance across various noisy environments, with minimal impact of ambient noise on detection accuracy.* Since the core principle of *SpiroSense* involves using ultrasonic signals to sense lung airflow velocity for lung function detection, evaluating its robustness against environmental noise interference is essential. To examine the impact of different noise sources, such as traffic noise, talking noise, and TV noise, we conducted comparative experiments to assess lung function under noisy conditions, with noise intensity maintained at approximately 80 dB. As illustrated in Fig. 19(a), the relative errors of FEV_1, FVC, and other indicators remain below 10% across all noise scenarios. Similarly, Fig. 19(b) shows that the absolute errors of FEV_1 and FVC are less than 0.6 L in these conditions. These results demonstrate that while noise scenarios slightly influence absolute error, their impact on relative error is negligible, underscoring the system's robustness in noisy environments.

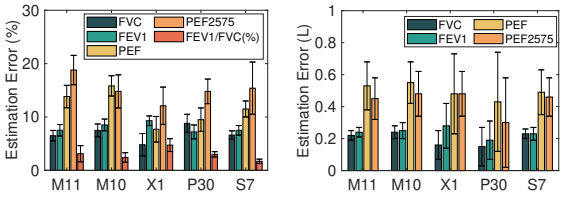
6.7 Impact of Different Smartphones

Experimental results demonstrate that SpiroSense maintains high accuracy and consistent performance across a range of commercial smartphones, with minimal performance degradation due to device-specific variations. To further assess the robustness and generalizability of SpiroSense, we evaluate its performance across five smartphones,



(a) Relative estimate error. (b) Absolute estimate error.

Fig. 19. The impact of different noise scenarios on errors.



(a) Relative estimate error. (b) Absolute estimate error.

Fig. 20. The impact of different smartphones on errors.

including Samsung Galaxy SM-7810 (S7), Sony XPERIA XZ1 (X1), Xiaomi 10 and 11 (M10 and M11), and Huawei P30 Pro (P30). These smartphones differ in both hardware quality and microphone configurations, including microphone type, placement, and the internal signal processing pipeline (e.g., noise suppression mechanisms). We invite nine adult volunteers (i.e., five males and four females), aged between 22 and 30 years and with BMIs ranging from 19% to 32%, to participate in the evaluation, and all participants follow the standardized breathing protocols described in Section 6.1. As shown in Fig. 20(a), the relative errors for key pulmonary function indicators such as FEV_1 and FVC remained below 8% across all tested smartphones. Moreover, Fig. 20(b) illustrates that the absolute errors for FEV_1 and FVC consistently remained under 0.24L, indicating substantial measurement accuracy regardless of smartphones. Despite this overall robustness, we observe subtle variations in performance across models. These are primarily attributed to differences in microphone sensitivity, frequency response characteristics, and built-in noise suppression algorithms. For instance, front-facing microphones (as used in the M10 and M11) demonstrated slightly improved consistency due to more favorable alignment with the exhaled airflow. Additionally, optimal recording gain levels varied across smartphones and operating systems, underscoring the need for adaptive volume calibration during deployment. These findings collectively confirm that SpiroSense is highly adaptable and practical across a variety of smartphone hardware configurations, making it suitable for large-scale, real-world deployment.

6.8 Impact of Smartphone Volume

Experimental results demonstrate that volume settings significantly affect the performance of SpiroSense*, with detection errors increasing as the volume rises. Extremely low volumes also result in higher errors due to insufficient signal strength for accurate detection.* To evaluate the impact of volume on *SpiroSense*, experiments are designed to assess its performance across different volume configurations. Specifically, the volume of the Samsung SM-7810 is adjusted from 10% to 100% in increments, and five volunteers (two males and three females) are invited to participate in the experiment evaluation. As shown in Fig. 21, significant differences are observed in the relative errors of the PFT metrics under various volume settings, with errors increasing as the volume rises. Additionally, lower volumes do not lead to lower relative estimation errors. Surprisingly, the relative estimation errors at 10% volume are higher than at 20%. This is because, at very low volumes, the transmitted signal power decreases, making it more difficult for the microphone to detect the signal.

6.9 Impact of the Run Time of Each Module

Experimental results demonstrate that the synchronization and IQ demodulation modules are sensitive to increased data complexity, resulting in higher processing times, while the NoiseClear module maintains stable performance, demonstrating its efficiency and robustness in signal denoising regardless of data volume. To study the effect of different data volumes on the processing time of each module, as shown in Fig. 22, experiments are designed

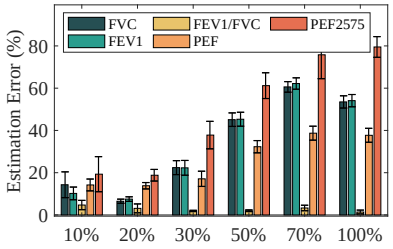
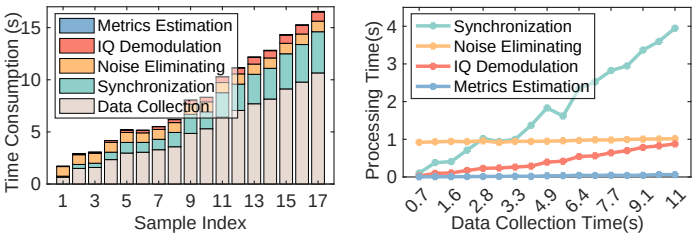


Fig. 21. Impact of smartphone volume on estimation error.



(a) Total time consumption.

(b) Signal processing time.

Fig. 22. Running time of each module in *SpiroSense* on a smartphone.

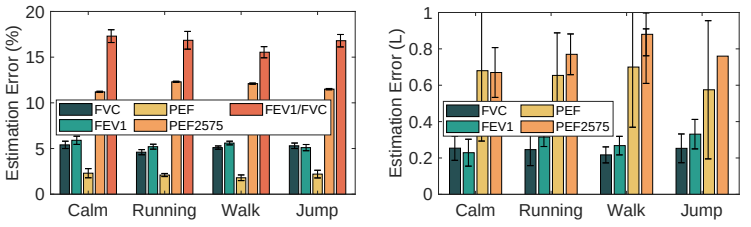
where the data acquisition time is increased from 650 ms to 10.66 s. As the data acquisition time increases, the processing times for both the synchronization and IQ demodulation modules rise significantly, although the total signal processing time remains less than half of the data collection time. This suggests that synchronizing and demodulating larger or more complex data sets require more computational resources. As shown in Fig. 22(b), it is noteworthy that *NoiseClear* processing remains relatively stable, with minimal variance between 650 ms and 1,016 ms across all samples. This consistency highlights the module's efficiency in signal denoising, remaining unaffected by the increasing data volume.

6.10 Impact of Different Conditions of User

Experimental results indicate that SpiroSense *maintains reliable measurement accuracy across different user conditions, with minimal impact from physical activity-induced fluctuations in respiratory patterns, demonstrating its robustness in real-world scenarios.* To assess the performance of the *SpiroSense* system under varying user conditions, experiments are designed to simulate real-life scenarios, including both calm and physically exerted states. Specifically, five volunteers aged between 20 and 30 are invited to participate in the study. Lung function indicators are measured under different physical activities, such as running, walking, and jumping, to replicate post-exercise conditions. As shown in Fig. 23, the experiment results reveal slight deviations in the PFT metrics measured immediately after different exercises, likely due to reduced breath stability following physical exertion. However, it is important to note that despite these variations, the overall perception accuracy of *SpiroSense* remains minimally affected by physical activity. These findings suggest that while physical activity may induce temporary fluctuations in respiratory patterns, the system retains its ability to reliably measure lung function metrics, ensuring its robustness across various real-world conditions.

6.11 Ablation Study of the *NoiseClear* Module

The ablation study confirms the critical role of the NoiseClear *denoising module in* SpiroSense*, as its inclusion significantly reduces estimation errors and stabilizes measurement accuracy across various pulmonary function indicators.* To assess the significance of the *NoiseClear* denoising module, an ablation study is conducted to evaluate its impact on the performance of *SpiroSense*. Ten volunteers are invited to participate in the experiment, comparing system performance under two configurations: with and without *NoiseClear*. In the configuration without *NoiseClear*, IQ demodulation is applied directly to the synchronized and aligned signals without the denoising step. As illustrated in Fig. 24, these results demonstrate that the estimation errors across various lung function indicators are consistently lower when the *NoiseClear* module is utilized compared with the configuration without it. Moreover, the absence of the *NoiseClear* module leads to a significant increase in the variance of estimation errors for key metrics, such as PEF_{25-75}, thereby reducing the confidence in the measurements. This variance highlights the critical role of *NoiseClear* in stabilizing the accuracy of the evaluations.



(a) Relative estimate error. (b) Absolute estimate error.

Fig. 23. Impact of different user conditions on estimation errors.

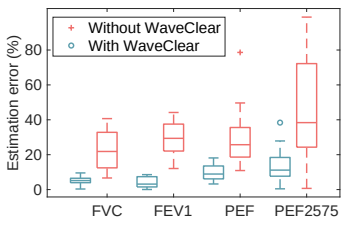


Fig. 24. Impact of the *NoiseClear* module on PFT accuracy.

Table 2. Estimation error of PFT metrics.

Category	FEV_1 (%)	FVC (%)	FEV_1/FVC (%)
Asthma	6.61%	6.90%	3.1%
COPD	6.16%	7.74%	3.74%
Average	6.39%	7.32%	3.41%

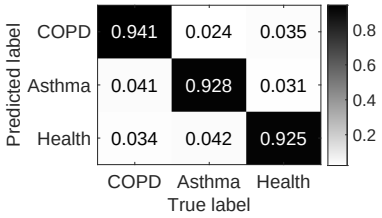


Fig. 25. Confusion matrix.

7 CASE STUDY

The SpiroSense *system achieves accurate pulmonary function estimation with errors of* 4.39% *for* FEV_1*,* 5.32% *for FVC, and* 6.41% *for* FEV_1/FVC*, and demonstrates strong disease classification performance, with* 92.8% *accuracy for asthma and* 94.1% *for COPD detection.*

The clinical study involves human subjects with different conditions (asthma and COPD), focusing on two disease prediction classification tasks: i) distinguishing asthma patients from healthy individuals, and ii) distinguishing COPD patients from healthy individuals. As shown in Table 2, the error in lung function estimation for both asthma and COPD patients remains within an acceptable range, with average errors for FEV_1, FVC, and FEV_1/FVC ratio being 6.39%, 7.32%, and 3.41%, respectively. These values demonstrate the reliability of the *SpiroSense* system in accurately estimating pulmonary health metrics. The slight variation in errors between asthma and COPD suggests that, while the system performs well across both conditions, the precision of estimation may be influenced by the severity and specifics of each disease. In the disease prediction classification tasks, pulmonary function metrics (e.g., FEV_1, FVC, and FEV_1/FVC ratio) are used to train a Support Vector Machine (SVM) classifier aimed at distinguishing asthma and COPD patients from healthy individuals. The models are evaluated using cross-validation, showing strong classification performance with accuracy rates of 92.8% for asthma detection and 94.1% for COPD detection.

The SpiroSense *system shows high accuracy and reliability in detecting asthma and COPD, making it an effective noninvasive tool for the early detection and management of pulmonary diseases.* As shown in Fig. 25, the confusion matrix illustrates the classification performance for asthma and COPD detection. Most patients are correctly classified, with minimal misclassification between conditions, indicating high accuracy and reliability of the *SpiroSense* system. The model distinguishes asthma and COPD patients well from healthy individuals, demonstrating high sensitivity and specificity. These results underscore that *SpiroSense* is a robust, noninvasive tool for monitoring pulmonary health and predicting diseases. It offers significant clinical benefits for the early detection and management of pulmonary conditions. Further analysis reveals that the system's ability to predict asthma or COPD is consistent across a diverse group of participants, reinforcing the robustness of *SpiroSense* in real-world settings. These findings suggest that *SpiroSense* is an effective tool for noninvasive, smartphone-based pulmonary health monitoring and disease prediction, providing both clinical and practical advantages for early detection and management of pulmonary diseases.

8 LIMITATION AND DISCUSSION

While *SpiroSense* demonstrates strong overall performance in pulmonary function estimation, practical deployment still faces several challenges. In this section, we discuss the system's potential limitations and propose feasible solutions, along with key considerations for real-world applications.

Limitation. Although the proposed smartphone-based pulmonary function detection system, *SpiroSense*, achieves significant advantages in overall system performance compared to SOTA methods, it still has some limitations in real environments. Specifically, i) *Detection Accuracy*: The primary goal of *SpiroSense* is to detect

early lung function diseases or perform long-term monitoring of chronic lung diseases. Ensuring a low missed detection rate or detection error is crucial for practical application. However, there is a slight gap in detection accuracy between *SpiroSense* and medical equipment. ii) *System Deployment*: The essence of the system is to use smartphones to send and receive ultrasonic signals to sense airflow velocity characteristics, thereby obtaining lung function-related indicators. However, differences in the frequency response of audio devices in different smartphones and volume control bring restrictions to system deployment. These limitations need to be addressed for *SpiroSense* to realize its full potential in clinical applications.

Discussion. This section explores other potential practical concerns of our proposed system, *SpiroSense*, which transforms a standard smartphone into a household pulmonary examination device. One of the key advantages of *SpiroSense* lies in its use of 3D-printed models that enable smartphones to accurately detect the volume and speed of exhaled airflow, advancing the concept of home-based smart medical care for lung disease detection. *(a) Low-cost 3D-printed model:* Typically costing only a few dollars when mass-produced—and around ten dollars individually—this component is significantly more affordable than clinical spirometry devices. Furthermore, since the model only channels airflow and ultrasonic signals without any mechanical or electronic components, it experiences negligible wear and tear, ensuring long-term stability without accumulating measurement drift over time. *(b) Calibration over time:* We acknowledge that environmental factors such as temperature, humidity, or mechanical stress over long-term use may introduce minor measurement deviations. To mitigate this, *SpiroSense* can be periodically calibrated using a reference airflow with a known constant flow rate, ensuring sustained measurement accuracy over time. *(c) Hygiene concerns: SpiroSense* also addresses hygiene concerns, especially for users with respiratory conditions or compromised immune systems. To ensure safe usage, it supports medical disposable or easily washable mouthpieces (as shown in Fig. 15(b)), similar to those used in traditional spirometers. Users can replace or sanitize the attachment after each use to maintain hygiene. In addition, since the 3D-printed model is designed without electronic components and is easily disassembled, it can be conveniently cleaned and disinfected. This not only enhances the overall durability and usability of the system but also contributes to safer long-term use, especially for users with respiratory vulnerabilities. *(d) Clinical risks:* In clinical practice, the severity of measurement errors in PFT, such as spirometry, is evaluated based on their potential to mislead diagnoses and treatment decisions [9, 54]. SpiroSense is designed not to replace professional medical equipment or directly guide treatment decisions, but rather to serve as a home-based daily pulmonary function monitoring system, and specific disease diagnosis still requires clinicians. Furthermore, most pulmonary diseases are chronic in nature, and clinical diagnoses are typically based on longitudinal trends rather than a single measurement. As shown in Fig. 25, the misidentification rate of *SpiroSense* for common pulmonary conditions is less than 7.5%. By following standard medical protocols [39], such as performing three repeated measurements, the misidentification rate drops to 1.6%. Over time, with continuous use, the system's ability to track trends rather than isolated events further minimizes the clinical impact of occasional errors. This longitudinal robustness makes *SpiroSense* well-suited for routine self-monitoring without posing significant clinical risks. *(e) Integration with telemedicine and home monitoring:* SpiroSense is well-suited for long-term monitoring of chronic respiratory conditions such as asthma or COPD. Its compatibility with smartphones enables real-time data collection and remote transmission, supporting telehealth consultations and continuous disease management. Additionally, the platform can be expanded to integrate with other health metrics (e.g., symptom scores, cough detection, oximetry), providing a comprehensive toolkit for personalized and proactive care at home. *(f) Future work:* To further enhance the accuracy and generalizability of *SpiroSense*, we plan to involve more subjects, including patients with varying degrees of pulmonary conditions, in experimental validation. In parallel, we aim to optimize software compatibility across diverse smartphones and improve the application's user interface for elderly users, such as incorporating voice-based interactions.

9 CONCLUSION

Changes in pulmonary ventilation are crucial for the early detection and diagnosis of lung diseases such as asthma and COPD. However, existing solutions either require expensive medical equipment or indirectly reflect lung function, rather than directly assessing lung disease from pulmonary ventilation. We propose *SpiroSense*, a novel system design that transforms a standard smartphone into a pulmonary household examination device, aiming to bridge the gap between clinical needs and current sensing techniques. The diffusion effect of airflow occurs when pulmonary air mixes with the surrounding air, causing rapid spread and attenuation of airflow velocity, which impacts measurement accuracy. To tackle this challenge, *SpiroSense* utilizes a 3D-printed model to control the propagation path of both airflow and ultrasonic signals. To overcome the limitations of smartphone sampling rates, *SpiroSense* leverages the fine-grained phase features of ultrasonic signals to measure airflow velocity within the tube. Although this design ensures accurate airflow measurement, blowing air into the 3D-printed tube introduces additional turbulence noise. The high-frequency harmonics this turbulence generates overlap with the ultrasonic frequency, affecting phase feature extraction. To this end, we propose an end-to-end ultrasonic signal noise elimination model, *NoiseClear*, based on a real-time ultrasonic signal enhancement approach in the waveform domain. We implement a prototype of *SpiroSense* and verify its effectiveness and robustness through comprehensive experimental evaluation. Experimental results demonstrate that the average error rates of *SpiroSense* for PFT metrics, including FVC, FEV_1, and FEV_1/FVC ratio, are 6.44%, 7.42%, and 3.01%, respectively.

Acknowledgments

This work is supported in part by National Natural Science Foundation of China under Grant No. 92467202; Natural Science Foundation of Jiangsu Province (Key Program) under Grant No. BK20243040; National Natural Science Foundation of China under Grant Nos. 62272216, 62372224, 62402217; Natural Science Foundation of Jiangsu Province under Grant No. BK20241377; and Collaborative Innovation Center of Novel Software Technology and Industrialization.

References

[1] Shikha Ambastha, Sharath Umesh, Uma Maheshwari K, and Sundarrajan Asokan. 2016. Pulmonary Function Test Using Fiber Bragg Grating Spirometer. *Journal of Lightwave Technology* 34, 24 (2016), 5682–5688.
[2] E. Bedrosian. 1963. A Product Theorem for Hilbert Transforms. *Proc. of the IEEE* 51, 5 (1963), 868–869.
[3] Soujanya Chatterjee, Md Mahbubur Rahman, Tousif Ahmed, Nazir Saleheen, Ebrahim Nemati, Viswam Nathan, Korosh Vatanparvar, and Jilong Kuang. 2020. Assessing Severity of Pulmonary Obstruction from Respiration Phase-Based Wheeze-Sensing Using Mobile Sensors. In *Proc. of the 38th ACM CHI*. 1–13.
[4] Zhe Chen, Tianyue Zheng, Chao Cai, and Jun Luo. 2021. MoVi-Fi: Motion-Robust Vital Signs Waveform Recovery via Deep Interpreted RF Sensing. In *Proc. of the 27th ACM MobiCom*. 392–405.
[5] Bhawana Chhaglani, Camellia Zakaria, Adam Lechowicz, Jeremy Gummeson, and Prashant Shenoy. 2022. FlowSense: Monitoring Airflow in Building Ventilation Systems Using Audio Sensing. *Proc. of the ACM IMWUT*. 6, 1 (2022).
[6] Keum San Chun, Viswam Nathan, Korosh Vatanparvar, Ebrahim Nemati, Md Mahbubur Rahman, Erin Blackstock, and Jilong Kuang. 2020. Towards Passive Assessment of Pulmonary Function from Natural Speech Recorded Using a Mobile Phone. In *Proc. of the 18th IEEE PerCom*. 1–10.
[7] Contec. 2024. CONTEC SP100A Spirometer. https://www.contecmed.com/productinfo/859505.html. [Accessed 15-01-2024].
[8] Contec. 2024. CONTEC SP70B Spirometer. https://www.contecmed.com/productinfo/603431.html. [Accessed 15-01-2024].
[9] Robert O. Crapo. 1994. Pulmonary-Function Testing. *New England Journal of Medicine* 331, 1 (1994), 25–30.
[10] F. de Jongh. 2008. Spirometers. *Breathe* 4, 3 (2008), 251–254.
[11] Alexandre Défossez, Nicolas Usunier, Léon Bottou, and Francis Bach. 2019. Music Source Separation in The Waveform Domain. *arXiv preprint arXiv:1911.13254* (2019).
[12] Long Fan, Lei Xie, Xinran Lu, Yi Li, Chuyu Wang, and Sanglu Lu. 2023. mmMIC: Multi-Modal Speech Recognition based on mmWave Radar. In *Proc. of the 42nd IEEE INFOCOM*. 1–10.

[13] Long Fan, Lei Xie, Wenhui Zhou, Chuyu Wang, Yanling Bu, and Sanglu Lu. 2024. Beamforming for Sensing: Hybrid Beamforming based on Transmitter-Receiver Collaboration for Millimeter-Wave Sensing. *Proc. of the ACM IMWUT*. 8, 2 (2024).
[14] Mayank Goel, Elliot Saba, Maia Stiber, Eric Whitmire, Josh Fromm, Eric C. Larson, Gaetano Borriello, and Shwetak N. Patel. 2016. SpiroCall: Measuring Lung Function over a Phone Call. In *Proc. of the 2016 CHI*. 5675–5685.
[15] Brian L Graham, Irene Steenbruggen, Martin R Miller, Igor Z Barjaktarevic, Brendan G Cooper, Graham L Hall, Teal S Hallstrand, David A Kaminsky, Kevin McCarthy, Meredith C McCormack, et al. 2019. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. *American Journal of Respiratory and Critical Care Medicine* 200, 8 (2019), e70–e88.
[16] Shijie Han, Dongheng Zhang, Jinbo Chen, Haoyu Wang, Jinli Zhang, Qibin Sun, and Yan Chen. 2023. Fine-grained Lung Function Sensing based on Millimeter-Wave Radar. In *Proc. of the 15th IEEE IC-WCSP*. 471–476.
[17] Yinghui He, Jianwei Liu, Mo Li, Guanding Yu, and Jinsong Han. 2024. Forward-Compatible Integrated Sensing and Communication for WiFi. *IEEE Journal on Selected Areas in Communications* 42, 9 (2024), 2440–2456.
[18] Yinghui He, Guanding Yu, Yunlong Cai, and Haiyan Luo. 2024. Integrated Sensing, Computation, and Communication: System Framework and Performance Optimization. *IEEE Trans. on Wireless Communications* 23, 2 (2024), 1114–1128.
[19] James C Hogg. 2004. Pathophysiology of Airflow Limitation in Chronic Obstructive Pulmonary Disease. *The Lancet* 364, 9435 (2004), 709–721.
[20] Jingyang Hu, Hongbo Jiang, Tianyue Zheng, Jingzhi Hu, Hongbo Wang, Hangcheng Cao, Zhe Chen, and Jun Luo. 2024. M2-Fi: Multi-person Respiration Monitoring via Handheld WiFi Devices. In *IEEE INFOCOM 2024*. 1221–1230.
[21] Jingzhi Hu, Tianyue Zheng, Zhe Chen, Hongbo Wang, and Jun Luo. 2023. MUSE-Fi: Contactless MUti-person SEnsing Exploiting Near-field Wi-Fi Channel Variation. In *Proc. of the 29th ACM MobiCom*.
[22] Libei Huang, Yong Liu, Geng Li, Yun Song, Jianjun Su, Le Cheng, Weihua Guo, Ganggang Zhao, Hanchen Shen, et al. 2023. Ultrasensitive, Fast-Responsive, Directional Airflow Sensing by Bioinspired Suspended Graphene Fibers. *Nano Letters* 23, 2 (2023), 597–605.
[23] Junko Izawa, Kouki Matsuoka, Prarinya Siritanawanan, and Shinji Fukusawa. 2024. A Method of Vital Capacity Measurement Using Three-Dimensional Depth Sensor. In *Proc. of the 39th IEEE ITC-CSCC*. 1–5.
[24] Lu Jin, Zekun Liu, Mucahit Altintas, Yan Zheng, Zhangchi Liu, Sirui Yao, Yangyang Fan, and Yi Li. 2022. Wearable Piezoelectric Airflow Transducers for Human Respiratory and Metabolic Monitoring. *Acs Sensors* 7, 8 (2022), 2281–2292.
[25] Eric C. Larson, Mayank Goel, Gaetano Boriello, Sonya Heltshe, Margaret Rosenfeld, and Shwetak N. Patel. 2012. SpiroSmart: using a microphone to measure lung function on a mobile phone. In *Proc. of the ACM IMWUT*. 280–289.
[26] Xin Li, Hongbo Wang, Zhe Chen, Zhiping Jiang, and Jun Luo. 2024. UWB-Fi: Pushing Wi-Fi towards Ultra-wideband for Fine-Granularity Sensing. In *Proc. of the 22nd ACM MobiSys*. 42–55.
[27] Yi Li, Chuyu Wang, Lei Xie, Qiancheng Jin, Long Fan, Jingyi Ning, and Sanglu Lu. 2024. Facial Landmark Detection Based on High Precision Spatial Sampling via Millimeter-wave Radar. *Proc. of the ACM IMWUT* 8, 4 (2024), 1–26.
[28] O. Löwhagen. 2012. Diagnosis of asthma – a new approach. *Allergy* 67, 6 (2012), 713–717.
[29] Abadi Martín, Agarwal Ashish, Barham Paul, Brevdo Eugene, Chen Zhifeng, Citro Craig, S Corrado Greg, Davis Andy, Dean Jeffrey, Devin Matthieu, et al. 2015. TensorFlow: Large-scale Machine Learning on Heterogeneous Systems.
[30] M R Miller, S A Dickinson, and D J Hitchings. 1992. The Accuracy of Portable Peak Flow Meters. *Thorax* 47, 11 (1992), 904–909.
[31] Kathleen M. Mortimer, Andre Fallot, John R. Balmes, and Ira B. Tager. 2003. Evaluating the Use of a Portable Spirometer in a Study of Pediatric Asthmaa. *Chest* 123, 6 (2003), 1899–1907.
[32] Rajalakshmi Nandakumar, Shyamnath Gollakota, and Nathaniel Watson. 2015. Contactless Sleep Apnea Detection on Smartphones. In *Proc. of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. 45–57.
[33] n.d. 2023. Impulse Oscillometry System. https://intl.vyaire.com/products/ios-impulse-oscillometry. [Accessed 19 Nov. 2023].
[34] Romain A. Pauwels and Klaus F. Rabe. 2004. Burden and Clinical Features of Chronic Obstructive Pulmonary Disease (COPD). *The Lancet* 364, 9434 (2004), 613–620.
[35] Nayia Petousi, Nick P Talbot, Ian Pavord, and Peter A Robbins. 2019. Measuring Lung Function in Airways Diseases: Current and Emerging Techniques. *Thorax* 74, 8 (2019), 797–805.
[36] David A. Rebuck, Nicola A. Hanania, Anthony D. D'Urzo, and Kenneth R. Chapman. 1996. The Accuracy of a Handheld Portable Spirometer. *Chest* 109, 1 (1996), 152–157.
[37] James S Robertson, William E Siri, Hardin B Jones, et al. 1950. Lung Ventilation Patterns Determined by Analysis of Nitrogen Elimination Rates; Use of the Mass Spectrometer as a Continuous Gas Analyzer. *The journal of clinical investigation* 29, 5 (1950), 577–590.
[38] Nazir Saleheen, Tousif Ahmed, Md Mahbubur Rahman, Ebrahim Nemati, Viswam Nathan, Korosh Vatanparvar, Erin Blackstock, and Jilong Kuang. 2020. Lung Function Estimation from a Monosyllabic Voice Segment Captured Using Smartphones. In *Proc. of the 22nd ACM MobileHCI*. 1–11.
[39] Sanja Stanojevic, David A Kaminsky, Martin R Miller, Bruce Thompson, Andrea Aliverti, Igor Barjaktarevic, Brendan G Cooper, Bruce Culver, Eric Derom, Graham L Hall, et al. 2022. ERS/ATS Technical Standard on Interpretive Strategies for Routine Lung Function Tests. *European Respiratory Journal* 60, 1 (2022).

[40] Andrea Valenzuela, Nicolás Sibuet, Gemma Hornero, and Oscar Casas. 2021. Non-Contact Video-Based Assessment of the Respiratory Function Using a RGB-D Camera. *Sensors* 21 (2021).
[41] Maroula Vasilopoulou, Andriana I Papaioannou, Georgios Kaltsakas, Zafeiris Louvaris, Nikolaos Chynkiamis, Stavroula Spetsioti, Eleni Kortianou, Sofia Antiopi Genimata, Anastasios Palamidas, Konstantinos Kostikas, et al. 2017. Home-based Maintenance Telerehabilitation Reduces the Risk for Acute Exacerbations of COPD, Hospitalisations and Emergency Department Visits. *European Respiratory Journal* 49, 5 (2017).
[42] Haoran Wan, Shuyu Shi, Wenyu Cao, Wei Wang, and Guihai Chen. 2021. RespTracker: Multi-user Room-scale Respiration Tracking with Commercial Acoustic Devices. In *IEEE INFOCOM 2021*. 1–10.
[43] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. 2016. Human Respiration Detection with Commodity WiFi Devices: Do User Location and Body Orientation Matter?. In *Proc. of the ACM IMWUT*. 25–36.
[44] Tianben Wang, Daqing Zhang, Leye Wang, Yuanqing Zheng, Tao Gu, Bernadette Dorizzi, and Xingshe Zhou. 2019. Contactless Respiration Monitoring Using Ultrasound Signal With Off-the-Shelf Audio Devices. *IEEE Internet of Things Journal* 6, 2 (2019), 2959–2973.
[45] Yao Wang, Tao Gu, Tom H. Luan, and Yong Yu. 2022. Your Breath Doesn't Lie: Multi-user Authentication by Sensing Respiration Using mmWave Radar. In *2022 19th IEEE SECON*. 64–72.
[46] Xiangyu Xu, Jiadi Yu, Chengguang Ma, Yanzhi Ren, Hongbo Liu, Yanmin Zhu, Yi-Chao Chen, and Feilong Tang. 2022. mmECG: Monitoring Human Cardiac Cycle in Driving Environments Leveraging Millimeter Wave. In *IEEE INFOCOM 2022*. 90–99.
[47] Xiangyu Yin, Kai Huang, Erick Forno, Wei Chen, Heng Huang, and Wei Gao. 2023. PTEase: Objective Airway Examination for Pulmonary Telemedicine Using Commodity Smartphones. In *Proc. of the 21st ACM MobiSys*. 110–123.
[48] Gu Yu, Meng Wang, Peng Zhao, Yantong Wang, Hao Zhou, Yusheng Ji, and Celimuge Wu. 2022. SpiroFi: Contactless Pulmonary Function Monitoring using WiFi Signal. In *Proc. of the 30th IEEE/ACM IWQoS*. 1–10.
[49] Paolo Zanaboni, Birthe Dinesen, Hanne Hoaas, Richard Wootton, Angela T Burge, Rochelle Philp, Cristino Carneiro Oliveira, Janet Bondarenko, Torben Tranborg Jensen, Belinda R Miller, et al. 2023. Long-term Telerehabilitation or Unsupervised Training at Home for Patients with Chronic Obstructive Pulmonary Disease: A Randomized Controlled Trial. *American Journal of Respiratory and Critical Care Medicine* 207, 7 (2023), 865–875.
[50] Duo Zhang, Xusheng Zhang, Yaxiong Xie, Fusang Zhang, Hongliu Yang, and Daqing Zhang. 2024. From Single-Point to Multi-Point Reflection Modeling: Robust Vital Signs Monitoring via mmWave Sensing. *IEEE Trans. on Mobile Computing* 23, 12 (2024), 14959–14974.
[51] Feng Zhang, Chenshu Wu, Beibei Wang, Min Wu, Daniel Bugos, Hangfang Zhang, and K. J. Ray Liu. 2021. SMARS: Sleep Monitoring via Ambient Radio Signals. *IEEE Trans. on Mobile Computing* 20, 1 (2021), 217–231.
[52] Fusang Zhang, Daqing Zhang, Jie Xiong, Hao Wang, Kai Niu, Beihong Jin, and Yuxiang Wang. 2018. From Fresnel Diffraction Model to Fine-grained Human Respiration Sensing with Commodity Wi-Fi Devices. *Proc. of the ACM IMWUT*. 2, 1 (2018).
[53] Zhang Zhang, Yuran Kang, Ni Yao, Jing Pan, Wen Yu, Yao Tang, Yue Xu, Liqiang Wang, Lei Zhang, and Limin Tong. 2021. A Multifunctional Airflow Sensor Enabled by Optical Micro/Nanofiber. *Advanced Fiber Materials* 3 (2021), 359–367.
[54] Joseph D Zibrak, Carl R O'Donnell, and Keith Marton. 1990. Indications for Pulmonary Function Testing. *Annals of Internal Medicine* 112, 10 (1990), 763–771.