From a Point to Hundreds: Embracing LiDAR on Commodity Smartphones for Fine-Grained Pulmonary Function Sensing

Xuefu Dong
The University of Tokyo
dongxuefu@mcl.iis.u-tokyo.ac.jp

Wenwei Li
Peking University
lwenwei@pku.edu.cn

Minhao Cui
Seoul National University
minhaocui@snu.ac.kr

Zilong Wang
Microsoft Research Asia
wangzilong@microsoft.com

Lupeng Zhang
Nanyang Technological University
LUPENG001@e.ntu.edu.sg

Akihito Taya
The University of Tokyo
taya-a@iis.u-tokyo.ac.jp

Yuuki Nishiyama
The University of Tokyo
nishiyama@csis.u-tokyo.ac.jp

Kaoru Sezaki
The University of Tokyo
sezaki@iis.u-tokyo.ac.jp

Lili Qiu
Microsoft Research Asia, UT Austin
lili@cs.utexas.edu

Jie Xiong*
Nanyang Technological University
jie.xiong@ntu.edu.sg

Abstract

Wireless sensing is an emerging technology with a wide range of applications, but most existing systems capture only the motion of a single point, such as in respiration monitoring. This limitation is critical for tasks requiring multi-point data, such as respiratory volume measurement, where different body points provide distinct information, and a single point cannot represent them all. In this paper, we propose LiSen, a smartphone-integrated LiDAR system for multi-point wireless sensing, and demonstrate its contact-free capability for measuring respiratory volume. LiSen uses smartphone LiDAR to track multiple chest and abdominal points, enabling the first ranging-based spirometer system that captures the full volume curve without new-user calibration. We leverage the unique feature of multi-point sensing to address challenges such as body interference, diverse breathing patterns, and pressure differences. Tests with 35 examinees show that LiSen accurately estimates both instantaneous forced expiratory and inspiratory volume, achieving mean absolute errors below 0.24 L and 0.30 L, respectively, and an 8.93% error for four common pulmonary function indices.

CCS Concepts

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

Keywords

Multi-point sensing, pulmonary function test, spirometry, LiDAR

ACM Reference Format:

Xuefu Dong, Wenwei Li, Minhao Cui, Zilong Wang, Lupeng Zhang, Akihito Taya, Yuuki Nishiyama, Kaoru Sezaki, Lili Qiu, and Jie Xiong. 2026. From a Point to Hundreds: Embracing LiDAR on Commodity Smartphones for Fine-Grained Pulmonary Function Sensing. In *ACM/IEEE International Conference on Embedded Artificial Intelligence and Sensing Systems (SenSys '26), May 11–14, 2026, Saint Malo, France.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3774906.3802752

*Corresponding author email: jie.xiong@ntu.edu.sg

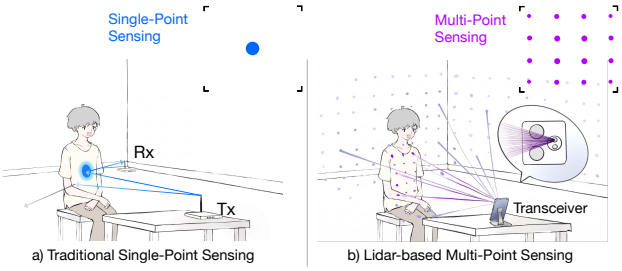


Figure 1: LiSen uses a novel smartphone LiDAR-based multi-point sensing method for Pulmonary Function Test (PFT).

1 Introduction

Wireless signals have been extensively studied for sensing purposes in recent years. Numerous wireless technologies have been utilized for sensing including Wi-Fi [82, 88], RFID [71, 76], mmWave [1, 15], UWB [92, 94], and ultrasound [12, 42, 43, 51]. The basic principle of wireless sensing is that signal propagation is altered when signals are reflected by targets. By analyzing these variations, valuable information can be inferred about the target's presence, location, movement, and even certain physical characteristics. A wide range of wireless sensing applications has been realized, including contact-free vital sign monitoring [88, 93], human gesture recognition [20, 48], and material sensing [74, 76]. Wireless sensing eliminates the need for bulky dedicated sensors by leveraging everyday devices, such as Wi-Fi routers and smartphones, for sensing.

However, existing wireless sensing systems typically treat the reflection path as a single path for analysis [37, 92]. When multiple points on the same object are at similar distances from the transceiver, their reflections arrive nearly simultaneously and merge into a single waveform. In other words, the target motion can only be represented as the motion of one single point. Even with an antenna array of eight antennas to steer the signal in a specific direction, the beamwidth remains too wide, preventing accurate resolution of multiple nearby points [79]. This limitation does not significantly affect certain sensing applications, such as

SenSys '26, Saint Malo, France

ACM ISBN 979-8-4007-2309-4/26/05
https://doi.org/10.1145/3774906.3802752

respiration rate monitoring, where any point on the chest provides the same information, making single-point sensing sufficient. However, it becomes a critical issue for more fine-grained tasks, such as pulmonary function evaluation, where information from multiple reflection points is required, and different points contribute very differently to the measurements. At present, this level of analysis requires professional, bulky equipment [50]. Thus, we pose the question: "Is it possible to leverage portable commercial devices to enable multi-point wireless sensing and support critical applications such as pulmonary function assessment?"

To answer this question, we explore a current trend in smartphones, the most ubiquitous consumer devices. Many of the latest models are equipped with miniature LiDAR modules, which are infrared-based depth sensors. LiDAR emits hundreds of beams and calculates their time-of-flight directly, achieving millimeter-level precision [68], which enables essential applications such as augmented reality, 3D object reconstruction, and photography assistance [36, 83]. We observe that their extremely narrow infrared beams, as narrow as 0.5 degrees, hold great potential for enabling the proposed multi-point wireless sensing. In this work, we focus on enabling pulmonary function testing (PFT) using commercial smartphones, a critical application due to its clinical importance and the unique challenges that can be addressed by moving from single-point to multi-point sensing.

Challenge 1 - Body Motion Interference: One critical issue with existing sensing methods is that the target information can be easily contaminated by irrelevant body movements. This problem becomes even more pronounced during pulmonary function testing (PFT), where vital breathing signals are often buried under large involuntary torso movements, such as leaning forward [78]. Moreover, unlike normal, natural breathing patterns, the respiratory motions during forced efforts in PFT exhibit similarly aperiodic characteristics to torso movements, making them even more difficult to distinguish. Conventional single-point sensing captures a composite motion signal in which respiratory and torso movements are mixed, lacking the spatial context needed to separate them.

Solution 1 - Points Reference: We cancel the body motion interference by leveraging key points as references for other points in the multi-point sensing system. The key insight is that certain areas around the shoulders and pelvis are less affected by respiratory motion. Thus, we can extract cleaner body motion signals from these regions and use the information to build a reference for canceling torso movement in respiration sensing. Specifically, we first track shoulder key points by leveraging dense neighboring points to provide semantic information. We then model the trunk of the seated individual[1] as a plane rotating around the buttocks. Finally, we use the reference points on the shoulders and pelvis to fit the body's rotation and subtract the body motion effect from the respiration signals.

Challenge 2 - Differences in Air Pressure Inside and Outside the Human Body: After obtaining the precise volume changes of the chest and abdomen, we observe that they do not directly correspond to lung capacity. This is because lung capacity is measured by the volume of air under ambient air pressure, which differs from the pressure inside the human body, particularly during forceful exhalation. Therefore, the volume change of the human body is not equal to the lung capacity measured by the air volume under ambient atmospheric pressure. Estimating this hidden parameter is difficult using a single-point sensing method. Even with neural networks, the lack of sufficient spatial information limits the effectiveness of the input.

Solution 2 - Local Coherence Extraction: We estimate *the influence of air pressure inside the human lung* by leveraging shared geometric patterns of local surface deformation. This information can be extracted from the coupling between neighboring points. An intuitive phenomenon inspires us: *A balloon collapses faster when you squeeze it because higher internal pressure speeds up the volume change.* Human breathing follows the same pressure–volume principles, but the relationship is more complex, as both chest geometry and breathing patterns influence lung pressure, even for the same air volume. To capture this, we develop a physics-informed neural network (PINN) that encodes the local geometric coherence of the torso, using relative displacements, surface normals, and patch volumes as inputs to predict indicative pressure values.

Challenge 3 - The Effect of Abdomen: Unlike conventional respiration sensing, where overall torso motion is loosely mapped to breathing activity, PFT requires distinguishing chest and abdominal movements. Because chest and abdomen jointly help determine the true limits of lung volume, yet their body motion relates differently to air volume change [4]. Moreover, respiration mode varies across individuals. For example, some favor chest breathing while others rely more on abdominal breathing [8]. Therefore, the chest and abdomen must be treated as distinct regions, rather than being modeled as a single sensing point. However, segmenting these regions is challenging; unlike hand or facial landmarks, no clear boundary exists between the chest and abdomen as the regions are usually covered by clothing.

Solution 3 - Point Segmentation: To address the above challenge, we leverage another unique advantage of multi-point sensing, i.e., the ability to independently extract information from points at different body locations. The key insight for segmenting the chest and abdomen lies in the physiological structure of the human body: chest movement is constrained by the rigid rib cage, whereas abdominal movement is not. The structural difference between the chest and abdomen creates noticeable changes in respiration-induced displacement around their boundary. By comparing displacement magnitudes across points, we can distinguish chest from abdomen regions. These two sets of points are then fed into parallel input branches of the PINN encoder, which jointly produce the final estimate of lung volume.

Based on the proposed new sensing modality, we developed a smartphone-based PFT sensing system named LiSen. With our proposed system, PFT, which is "often not available in low- and middle-income countries"[2] can become far more convenient and accessible to the general public, as illustrated in Figure 1. We implement LiSen on two COTS smartphones (Apple iPhone 12&13 Pro) and a tablet (Apple iPad Pro M1) respectively. We recruited 35 participants aged 21 to 65 to evaluate the effectiveness of the

[1]Note that the target is usually required to be seated during the PFT.

[2]WHO. Chronic Obstructive Pulmonary Disease (COPD) Fact Sheet. Available at: https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd)

proposed system using the most common pulmonary function test, Forced Vital Capacity (FVC). Unlike prior works that require new-user calibration [64] or bare-chest exposure [53, 61], our system can be used immediately by new users wearing normal clothing.

In summary, our contributions are as follows:

(1) We explore a new wireless-signal-based modality to enable fine-grained multi-point sensing and demonstrate the capability of the novel sensing modality with LiSen, the first smartphone LiDAR-based Pulmonary Function Test system.
(2) We design a series of multi-point–enabled signal processing algorithms to address long-standing practical challenges that cannot be resolved using single-point sensing.
(3) We conducted comprehensive IRB-approved experiments with 35 examinees to evaluate the performance of LiSen. The results demonstrate that the proposed system achieves highly accurate sensing, with mean absolute errors (MAE) below 0.24 L for instantaneous forced expiratory lung volume estimation and below 0.30 L for forced inspiratory estimation.

2 Background

In this section, we first present the basic principles of LiDAR-based multi-point sensing and the volume measurements it enables, followed by a brief overview of Pulmonary Function Testing (PFT).

2.1 LiDAR on Smartphone and Tablet

As a ranging method, LiDAR emits laser beams toward a target and measures the time it takes for them to return, obtaining distances from multiple points. Thanks to their low cost and compact size, LiDAR modules are widely integrated into commercial mobile devices. Taking iPhone as an example, the built-in LiDAR module projects 576 beams to the environment [68], producing non-overlapping speckles arranged in a 24 × 24 array on surrounding objects. The resulting reflections are then captured by a cooperating sensor to measure the time of flight for each individual beam.

As a result, this multi-point sensing paradigm provides a unique opportunity to estimate the target's volume. Figure 2 shows an example torso volume of region V enclosed by the front torso wall (FRONT WALL), the xy plane (BACK WALL), and SIDE CROSS SECTIONS. Assuming BACK WALL stays still during respiration, Solemani et al. [65] prove that we can approximate the volume in the form of multi-point depth, using Gauss's Theorem:

$$V_v = \iiint_V (\nabla \cdot \mathbf{F})\, dV = \oiint_S (\mathbf{F} \cdot \mathbf{n})\, dS \approx S_{FW\_z\_proj} \sum z_i\,, \quad (1)$$

where S is the piecewise smooth boundary surface of region V ($\partial V = S$); n is the outward normal vector of S; $\mathbf{F} = [0, 0, z]^T$ is a continuously differentiable vector field; and $S_{FW\_z\_proj}$ is the projection of region V onto the xy-plane, which can be estimated using the convex hull algorithm. LiDAR provides precise $S_{FW\_z\_proj}$ and $\{z\}$ from dense and accurate multi-point information, which further boost the fidelity of the above approximation.

2.2 Pulmonary Function Test (PFT)

Respiratory movement is a coordinated process driven by the diaphragm, intercostal, and abdominal muscles [14, 49]. Inhalation involves the contraction of the diaphragm and the elevation of the ribs, which expand the thoracic cavity and draw air into the lungs. Exhalation occurs passively via elastic recoil, or forcefully with contraction of internal intercostals and abdominal muscles to increase intra-abdominal pressure. Pulmonary Function Testing (PFT) quantitatively measures airflow and lung volume. As summarized in Table 1, the Forced Vital Capacity (FVC) test is the most commonly used method for obtaining pulmonary function indices. A standard FVC maneuver [3] requires the examinee to: **(i)** inhale maximally after three to four normal breaths, **(ii)** exhale as forcefully and completely as possible, and **(iii)** inhale maximally once more.

Table 1: MOST-USED and other important pulmonary function indices

FVC Test Indices	Explanation
FVC | FORCED VITAL CAPACITY	VOLUME OF TOTAL AIR EXHALED
FEV1 | FORCED EXPIRATORY VOLUME IN A SECOND	VOLUME OF AIR EXHALED IN THE FIRST SECOND
FEV1/FVC |	RATIO OF FEV1 AND FVC
FIVC | FORCED INSPIRATORY VITAL CAPACITY	VOLUME OF TOTAL AIR INHALED

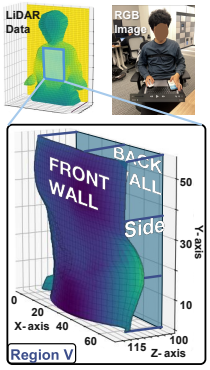


Figure 2: Coordinate Setup.

3 Challenges of Dynamic Volume Sensing

The vanilla ranging-based PFT sensing method [64] provides the basic principle of contact-free PFT. However, it still faces several fundamental challenges: (1) interference from unintentional non-respiratory torso movements; (2) discrepancies between air pressure and lung pressure; and (3) signal diversity due to individual variations in breathing patterns, such as mixed chest and abdominal breathing. In this section, we discuss these challenges in detail.

3.1 C1: Body Motion Interference

Though physicians typically instruct examinees to sit up straight at the beginning of PFT, many of them inevitably lean forward or backward during maximal inhalation and exhalation maneuvers [64]. Such body sway induces distance signal fluctuation on the decimeter scale, which can easily overwhelm respiration-related signals at the centimeter scale [54, 81]. Therefore, we need to estimate the body motion and correct the lung volume measurement as:

$$V_v \approx S_{FW\_z\_proj} \sum (f_{LiDAR}(x, y) - f_{lean}(x, y)), \quad (2)$$

where $f_{LiDAR}(x, y)$ denotes the measured displacement of the subject's front wall, and $f_{lean}(x, y)$ represents the displacement due to body leaning (referred to as MOTION WALL in the following).

Prior ranging-based PFT systems have explored several ways to mitigate the effect of unintentional motion. SpiroSonic [66] compensates contaminated signals with a linear trend, which only addresses minor body movements, while DeepBreath [78] subtracts

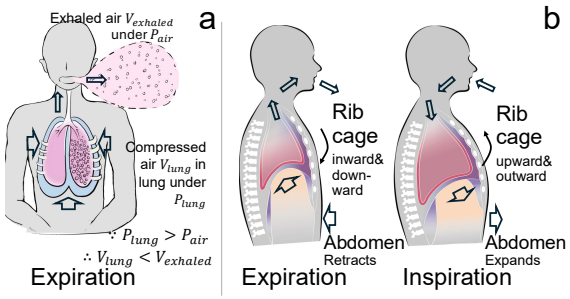


Figure 3: (a) Pressure discrepancy-oriented change in air volume inside and outside body; (b) Different respiratory mechanisms of the chest and abdomen.

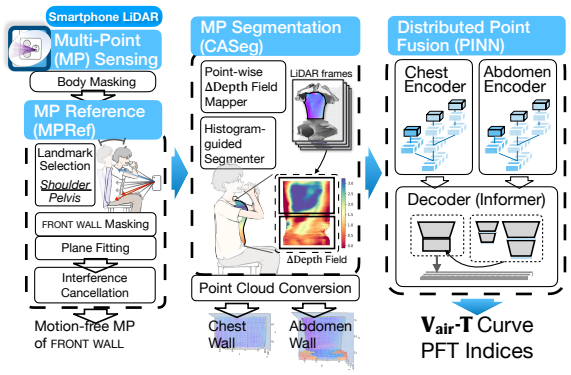


Figure 4: The fourfold pipeline for LiSen.

large-window moving averages to suppress interference from periodic, continuous respiration. Consequently, they fail to account for aperiodic, single-instance forced breaths during an FVC test, when unintentional motions can be more pronounced.

3.2 C2: Different Air Pressure Inside and Outside the Human Body

Another factor we need to consider is the pressure difference. The estimated physical volume of the lung does not directly correspond to the lung capacity, as shown in Figure 3(a). This discrepancy arises because lung capacity in PFT is defined as the volume of air exchanged at ambient atmospheric pressure, whereas the intrapulmonary pressure (the air pressure within the lung) fluctuates during respiration [56]. According to Boyle's law, the volume and pressure of air are inversely proportional. Thus, we have:

$$V_{air,t} = \frac{nRT}{P_{air,t}} = \frac{P_{lung,t}}{P_{air}} \cdot V_{v,t}\ , \tag{3}$$

where P_{air} is the ambient air pressure, n is the number of molecules, T is the temperature, R is a constant, and $P_{lung,t}$ denotes the intrapulmonary pressure at timestamp t. In practice, the pressure difference can reach over 120 cmH_2O or 0.116 atm during forced expiration [2, 35], resulting in more than 10% discrepancies between $V_{air,t}$ and $V_{front,t}$. However, it is difficult to directly measure ground truth $P_{lung,t}$, the intrapulmonary pressure,[3] as it requires a professional physician to insert a catheter into the lung [72].

3.3 C3: The Effect of Abdomen

Finally, it is insufficient to model the measured volume $V_{air,t}$ using only a single representative pressure $P_{lung,t}$, as discussed above. In practice, pressure and volume changes in the chest and abdominal regions can vary significantly and independently during the respiration process [2, 40], resulting in different pressures in the lungs and abdomen. This divergence is partly due to inter-individual differences in breathing patterns. Specifically, some examinees rely more on chest breathing, characterized by rib cage elevation, whereas others predominantly use diaphragmatic (abdominal) breathing, in which the diaphragm descends to expand lung volume, as illustrated in Figure 3(b). Thus, these two breathing modes induce different lung and abdominal pressures. Moreover, even the same subject may exhibit slight variations in the mixture of these two breathing modes across different trials.

Therefore, a chest-abdomen segmentation method is required to enable regional independent processing in lung capacity estimation:

$$V_{air,t} = \frac{P_{abdomen,t}}{P_{air}} \cdot V_{abdomen,t} + \frac{P_{lung,t}}{P_{air}} \cdot V_{chest,t}\ . \tag{4}$$

However, there is no clear boundary between the chest and abdomen, as both regions exhibit similar respiratory movements under clothing. Existing studies have either only qualitatively classified chest- or abdomen-dominated respiration [31, 78], or conducted simple segmentation under ideal conditions (e.g., examinees with a bare upper body) [5, 10].

4 System Design

In the following section, we introduce how LiSen responds to the three challenges with a workflow depicted in Figure 4 using three sequential modules, i.e., eliminating the motion interference using **Point Reference** (for C1), dividing multi-points into chest wall and abdomen using **Point Segmentation** (for C3), and estimating an indicative pressure cue using **Local Coherence Extraction** (for C2). To use the system, the examinee sits on a backless chair, with both hands holding a spirometer to the mouth, and both arms raised to avoid occluding the FRONT WALL. A smartphone LiDAR is placed in the front. To ensure reliable results and capture the maximum effort, the examinee repeats PFT at least three times [62].

4.1 Multi-Point ROI extraction: Eliminate Surrounding Interference

The modern smartphone LiDAR captures reflections of the emitted 576 depth beams and fuses them with optical data to provide multi-point information in a 256×192 depth map[4] D. In addition, synchronized RGB video I, camera intrinsics M are also streamed. However, such a dense multi-point sensing paradigm is a double-edged sword: while it provides meaningful sensing information, it is also inevitably contaminated by reflections from surrounding objects. To address this, we first extract the region of interest (ROI) around the examinee to suppress interferences.

A straightforward approach is to apply neural network–based segmentation to all RGB frames to obtain the subject ROI. However,

[3] Although handheld respiratory pressure meters exist, they measure pressure at the mouth or nose [73], which differs from the intrapulmonary pressure [11].

[4] Even though no direct access to raw LiDAR measurement is available, we can project the 2D depth map into 3D point cloud in the world coordinate system (WCS).

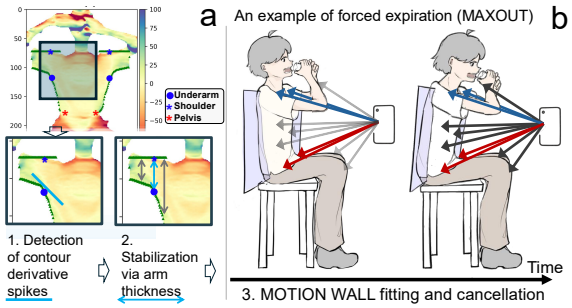


Figure 5: Multi-Point Reference (MPRef) Module: (a) MP tracker; (b) examinees tend to lean forward rigidly.

this method is computationally burdensome and sometimes unstable, as it relies entirely on deep neural networks and dense RGB frames. Instead, we propose a lightweight and stable method that uses simple threshold-based signal processing to extract the subject ROI from LiDAR frames with multi-point information. Only the first RGB frame is required for initialization using a pose estimator.

Specifically, we use AlphaPose [26] to process the first RGB frame and obtain a bounding box around the examinee's torso. LiDAR points within this bounding box are selected as the initial ROI, and the geometric center of the ROI, denoted as $P_{c,1}$, is chosen as the reference point for the subsequent ROI extraction. We then leverage the fact that the human chest-to-back distance is typically less than 30 cm [7], and remove LiDAR points located more than 40 cm deeper than $P_{c,1}$ to eliminate reflections from surrounding objects. To further separate objects at similar depths, we compute frame gradients with a Sobel operator and discard points above the 90th percentile of the frame. The remaining connected components are identified, and the component containing $P_{c,1}$ is defined as the first-frame LiDAR ROI segment, $\Omega_1(x, y)$. We then iterate this process: at each step, the reference point from the previous frame, $P_{c,t}$, guides the segmentation of the next frame to obtain $\Omega_{t+1}(x, y)$, ensuring stable tracking of the examinee over time.

4.2 Multi-Point Reference: Cancel Body Motion

To suppress unintentional whole-body motion during PFT, as detailed in Section 3.1, we exploit four landmarks: the left and right pelvis, which remain constant as the examinees remain seated, and the left and right shoulders, which remain fixed during normal respiratory motion but shift with unintended body movements. These four points further aid in extracting the respiration-related ROI, $\omega_t(x, y) \subset \Omega_t(x, y)$, from the selected torso LiDAR points by retaining only the points bounded by the four landmarks. More importantly, they allow us to suppress interference by estimating depth changes caused solely by unintentional body movements from the landmarks and subtracting them from the measurements.

4.2.1 Landmark Selections. Pelvis landmarks are relatively easy to obtain, as examinees remain seated. They can be calculated using an off-the-shelf pose estimator (e.g., AlphaPose) in the first frame and kept fixed over time. However, these estimators often fail to reliably select the shoulder landmarks because there could be dozens of nearby points in an image meeting the vague definition of "shoulder". In addition, smartphones may capture only a partial

Algorithm 1: Multi-Point Tracker

Input: Body masks Ω ($shape\ T \times (H \times W)$)
Output: Pelvis $\{L\_pelvis,\ R\_pelvis\}$;
shoulders $\{L\_shlder,\ R\_shlder\}$;
1 **Criteria/thresholds:** $\lfloor 0.33H \rfloor$, p=70%.
/* 1) Pelvis (fixed) */
2 $L\_pelvis, R\_pelvis \leftarrow$ PoseEstimator($\Omega[0]$).
/* 2) Shoulders. */
3 **for** $t = 0$ **to** $T-1$ **do**
4 $x_L[y] \leftarrow$ get_left_contour($\Omega[t]$);
5 $y_{arm} \leftarrow \arg\max(\text{get\_width}(\Omega[t](y < 0.8H)))$;
/* 2)a, Find underarm as steepest shrink on contour under y_{arm}. */
6 $Y = [\,y_{arm} : y_{arm}+\lambda\,]$; // ROI under arm
7 $y^L_{ax} \leftarrow \arg\max(-\nabla x_L[\,y \in Y\,])$;
/* 2)b, Locate shoulder in the column where arm thickness hits p-th percentile of mask vertical span. */
8 $\{vs(x)\} \leftarrow$ calculate vertical span for all near-underarm columns of $\Omega[t]$;
9 $L\_shlder\_x \leftarrow \arg\_\text{percentile}[\text{p}]\ (\{\text{vs(x)}\})$;
10 $L\_shlder\_y \leftarrow$ find nearest row on contour ($x_L[L\_shlder\_x]$) above underarm;
/* R_shlder can be acquired in the same way */
11 $shoulder[t] \leftarrow (L\_shlder, R\_shlder)$
12 **end**
13 **return** *pelvis, shoulder*

view of the body, as they are typically positioned close to the examinee. This can confuse existing pose estimators, causing errors in distinguishing the front and back of subjects and potentially swapping the left and right shoulders. These shoulder swaps, as well as even small-scale jitter, can significantly interfere with the system, since the shoulder points serve as key landmarks.

To address this issue, we design a lightweight geometry-based algorithm that uses only depth data, consisting of two phases, as illustrated in Figure 5. The key insight is that the underarm point can be used to infer the shoulder point and is easier to locate, as it consistently appears at the body corner when the examinee raises their arms. We first coarsely identify the underarm point using rows 4–7 of algorithm 1 by comparing slopes along the body outline. Next, we refine the shoulder point by searching around this coarse underarm location, leveraging the fact that the vertical line connecting the shoulder and underarm corresponds to the widest segment of the arm (as detailed in rows 6–11 of algorithm 1).

4.2.2 Interference Cancellation. After obtaining the four landmarks, we use them as boundary points to extract the ROI from the torso LiDAR points, $\omega_t(x, y) \subset \Omega_t(x, y)$, covering only the chest and abdominal regions directly involved in respiratory motion. The next step is to cancel interference from unintentional body movements during PFT. A straightforward approach would be to isolate the depth changes caused by unintentional movements and subtract them from the measurements. However, since unintentional and respiratory motions occur simultaneously, it is challenging to separate depth variations caused solely by unintentional movements.

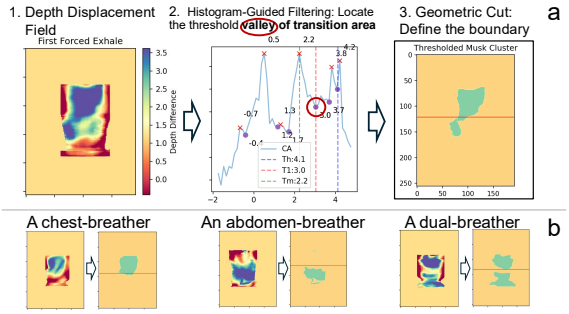


Figure 6: (a) Multi-Point Semantic Segmentation (CASeg) Module; (b) results under varying breathing patterns.

To address this, we leverage the observation that during forceful breathing, the body leans forward or backward almost as a rigid block [52, 69]. In other words, depth changes of all points caused by unintentional movements can be approximated as a plane, with the four landmarks (left/right shoulders and pelvis) serving as coplanar samples of that plane. Thus, we fit a plane using ordinary least squares on the landmark set $\{(x, y, z)|L\_shlder, R\_shlder, L\_pelvis, R\_pelvis\}$ to represent such a plane caused by the unintentional movement:

$$(A, B, C) = \arg\min_{A,B,C} \sum_{k=1}^{4} \left(A\,x_k \;+\; B\,y_k \;+\; C \;-\; z_k\right)^2. \qquad (5)$$

Each point on the estimated plane has a depth coordinate z which is subtracted from the corresponding LiDAR measurements to eliminate the influence of unintentional motions.

4.3 Multi-Point Segmentation: Subdivide Torso into Chest and Abdomen

After canceling unintentional body motion during breathing, we further separate the chest and abdomen regions, as these areas contribute differently to PFT. The key observation is that during forceful breathing, chest movement is constrained by rib-cage expansion, whereas abdominal movement is less restricted. Consequently, a displacement discontinuity tends to appear near the boundary, which can be used to differentiate the two regions. To exploit this observation, we first determine the start and end timestamps of the breathing movement for PFT. We then calculate the depth displacement field over this period, which indicates the magnitude of depth displacement at each point. Finally, we construct a histogram of the depth displacement field and use it to guide boundary detection between the chest and abdomen.

4.3.1 Flow–based Time Parsing. The standard PFT requires the examinee to perform a sequence of breathing maneuvers: a preparatory forced inspiration (MAXIN$_1$), a forced expiration (MAXOUT), and a second forced inspiration (MAXIN$_2$). To obtain the FVC test results, we focus on the last two phases, which require accurately identifying their timestamps. Rather than using volume changes directly, we estimate airflow speed, which provides a more accurate and straightforward measurement.

We use the Hampel filter [55] and moving average smoother to get the flow based on the measured volume trace $V_{v,t}$:

$$F_{front,t} \;=\; \text{movavg}\left[\text{Hampel}\left[\left(\,V_{v,t} - V_{v,t-1}\,\right)/f_s\right]\right]. \qquad (6)$$

Algorithm 2: Multi-Point Segmentation (CASeg)

Input: Depthmap frames D (shape $T \times (H \times W)$);
FRONT WALL masks M (length-T list);
point tracker PT (shape $T \times (H \times W) \times 2$);
MAXOUT interval $[t_{start}, t_{end}]$.
Output: Masks *CHE* (length-T list), *ABD* (length-T list).

/* 1) **Build Depth Displacement Field** */
$\omega'_{end} \leftarrow PT[t_{end}](\omega[t_{stard}])$;
$\Delta d \leftarrow D[t_{start}](\omega[t_{start}]) - D[t_{end}](\omega'_{end})$;
/* 2) **Leverage histogram of Δd to split 0th-frame mask into chest/abdomen** */
$modes\_\Delta d_{all} \leftarrow$ arg_peaks (hist(Δd));
$modes\_\Delta d_{\uparrow} \leftarrow$ arg_peaks (hist($\Delta d_{y<H/2}$));
$modes\_\Delta d_{\downarrow} \leftarrow$ arg_peaks (hist($\Delta d_{y\geq H/2}$));
$thr_{seg} \leftarrow$ Select threshold to segment dominant motion area (i.e. coarse mask of chest or abdoman), using majority displacements $\{modes\_\Delta\}$ with criteria in Sec. 4.3.3;
$M_{coarse} \leftarrow \arg(\Delta d \geq thr_{seg})$
/* 3) **Find steepest shrink row to split the mask into chest/abdomen** */
$w[y] \leftarrow$ get_width(M_{coarse});
$y_{seg} \leftarrow \arg\max_y(-\nabla w[y])$;
$CHE_0 \leftarrow \{\omega[0] : y < y_{seg}\}$; $ABD_0 \leftarrow \{\omega[0] : y \geq y_{seg}\}$;
for $t \leftarrow 0$ **to** $T-1$ **do** // Project masks to all frames
  $CHE[t] \leftarrow PT[t](CHE_0)$; $ABD[t] \leftarrow PT[t](ABD_0)$;
end
return *CHE*, *ABD*

// *CHEST = chest, ABD = abdomen.

Then we use the peak finder algorithm [70] to select positive peaks, which represent the maximal expiratory flow in MAXOUT; negative peaks, which represent the maximal inspiratory flow in MAXIN$_2$. Phase boundaries are defined as the samples nearest to zero around these peaks, yielding the corresponding timestamps ($t_{\text{out,start}}$, $t_{\text{in2,start}}$, $t_{\text{in2,end}}$) for a full FVC cycle (as represented in the 1st rows of algorithm 2).

4.3.2 Point-wise Depth Displacement (ΔD) Field Mapper. After obtaining the start and end timestamps of the breathing, we calculate the depth displacement field over this period, which indicates the magnitude of depth change at each point. However, LiDAR points at different timestamps do not necessarily correspond to the same physical locations on the torso due to unintentional body movements. Thus, we leverage the four landmarks (two shoulders and two pelvis points), stably tracked as described in Section 4.2, as anchors. Each LiDAR point in different frames is then matched based on its relative distance to these four anchors, which remain constant under body motion. For every tracked point, the depth displacement during forced expiration (MAXOUT) is computed as:

$$\Delta d \;=\; z(t_{out,start}) - z'(t_{out,end}), \qquad (7)$$

yielding a dense depth displacement field $\Delta D(x, y)$, that represents respiration-induced movement (2nd row of algorithm 2).

4.3.3 Histogram-guided segmentation. Finally, we use the obtained depth displacement field $\Delta D(x, y)$ to segment the chest region and the abdomen region. The principle is to identify the boundary based

on the distinct movement patterns: chest motion is constrained by rib-cage expansion, whereas abdominal motion is less restricted. In other words, during forceful breathing, points within the chest region exhibit similar depth displacements, as do points within the abdomen region. Thus, it is necessary to examine the distribution of depth displacements. Specifically, we analyze this distribution using the histogram $H(\Delta D)$ of all points, as shown in Figure 6:

- Two large displacement histogram peaks represent the point clusters in chest and/or abdomen, reflecting a plateau-like displacement profile [19, 40].
- Small-displacement peaks ($\lesssim 2$ cm, often smoothed by clothing) correspond to background clusters—body parts that remain stationary during breathing.
- The sparse histogram valleys between the chest and abdomen peaks indicate the boundary we aim to identify.

The boundary is estimated based on two observations: (1) its displacement should lie between the chest and abdominal displacements, as it is physically located between the two regions; and (2) the number of points in the boundary area should be much smaller than in the chest or abdomen regions. The corresponding algorithm is shown in the 3rd–7th rows of algorithm 2. Usually, the boundary obtained from the above method is sufficient to distinguish the chest and abdomen regions as shown in Figure 6b. However, in some cases, clothing can blur this boundary, as illustrated in the upper-right subfigure of Figure 6a. To address this, we compute the width of each row of boundary points; the row exhibiting the steepest top-to-bottom width drop indicates the anatomical transition between the chest and abdomen.

4.4 Local Coherence Extraction: Compensate Pressure Difference

To fully exploit the multi-point geometry of LiDAR data, we convert depth maps into 4D point cloud sequences (3D space + time). Point clouds maintain permutation invariance and scale consistency, allowing per-point attributes, such as surface normals, to capture local geometric context. A temporal–spatial kernel then clusters the 4D point cloud sequence into patches, providing structured input for our physics-informed neural network (PINN) as depicted in Figure 7. This pipeline aligns global and local features, compensates for pressure variations, and decodes accurate respiration volumes.

4.4.1 Conversion of Depth Map to Point Cloud. 3D point clouds preserve invariance and metric consistency, enabling local descriptors for fine respiratory motion. We unwrap depth $D(u, v)$ to (x, y, z) using camera intrinsics [39], and downsample by a same-index policy (odd rows) for stability and efficiency. For a 190 cm tall examinee seated 80 cm from the lidar, this yields an average of ~1318 points for THE CHEST WALL and ~1620 points for THE ABDOMEN WALL. Representation of each point is enriched with a surface normal ($\mathbf{n} = (n_x, n_y, n_z)$) estimated via PCA over its 35 nearest neighbors. Finally, frames are centered by translating their centroid to the origin, and the normalized CHEST WALL and ABDOMEN WALL clouds serve as dual input to our PINN pipeline.

4.4.2 Physics-Informed Neural Network (PINN). While pressure difference is usually ignored in prior ranging-based lung function

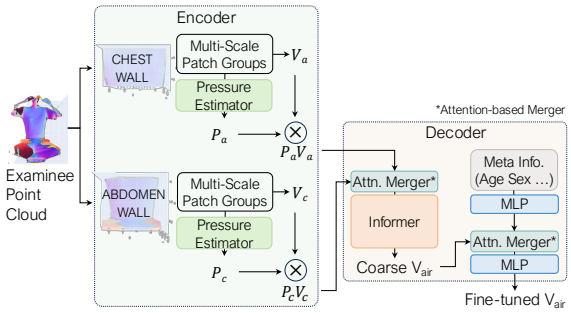


Figure 7: Physics-informed neural network of LiSen, consisting of a dual-input point cloud encoder, an indicative pressure-cue, and an Informer [97]-based decoder.

measurements, our idea is to exploit the multi-point information of the FRONT WALL, and the region volume estimate $V_{v,t}$, to recover an instantaneous suggestive intra-pulmonary pressure $P_{\text{lung},t}$ that corrects the lung volume estimate V_{air} according to Boyle's law.

An intuitive observation suggests the feasibility of a physical correlation between $\{P_{lung,t}\}$ and $\{V_{v,t}\}$: *for a deflating balloon, the volume shrinking would accelerate when the inner pressure is increased, e.g., by squeezing it by hand.* We theorize the phenomenon with two complementary descriptions of the same airflow F:

$$F = \underbrace{\frac{\left(P_{lung,t} - P_{air}\right)^{\frac{2}{3}}}{R}}_{\text{Fluid Mechanics [57]}} = \underbrace{\dot{V}_{air,t} = \frac{\left(\dot{P}_{lung,t}\, V_{v,t} + P_{lung,t}\, \dot{V}_{v,t}\right)}{P_{air}}}_{\text{General Definition}}. \quad (8)$$

Although respiration is an open system, we apply Boyle's law at each sampling instant by approximating the lungs as a quasi-static system with a fixed gas amount. This normalizes lung volume to ambient atmospheric pressure, aligning with standard pulmonary function tests (PFT) [33] and grounding the general definition of airflow above. This yields a differential equation for $P_{lung,t}$:

$$\dot{P}_{lung,t} + P_{lung,t} \cdot \frac{\dot{V}_{v,t}}{V_{v,t}} - \frac{P_{air}}{R \cdot V_{v,t}} \cdot \left(P_{lung,t} - P_{air}\right)^{\frac{2}{3}} = 0. \quad (9)$$

Because R varies with airway geometry and flow regime, we solve this data–physics coupling with a physics-informed neural network (PINN) [9] shown in Figure 7. For dataflow, two 4D sequences from chest/abdomen are encoded via point-cloud backbones [24, 25] into pressure-normalized volume embeddings. These embeddings are fused in a decoder and further refined by an Informer, together with metadata-based personalization, to regress the PFT volume curve $\{V_{air}\}$. For supervision, the model is optimized against ground-truth spirometer measurements using a mean squared error (MSE) loss. The ODE penalty (Equation 9) is omitted, as enforcing this regularization leads to instability and hinders convergence.

Multi-Scale Patch Inputs: Following point cloud preprocessing with the point 4D convolutional (P4D-Conv) layer [24], we partition each frame's chest and abdominal input sequences into $K_0 = 40$ patches (32 points per patch). This patch-based grouping captures local geometry at a defined scale. To capture multi-scale features,

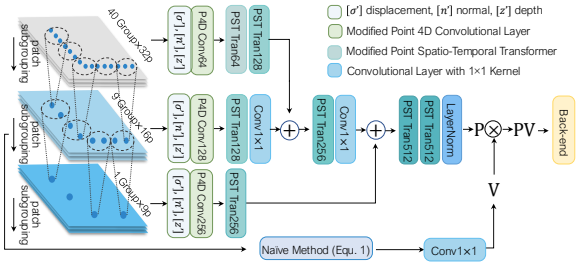


Figure 8: Multi-branch encoder to accommodate multi-scale patches.

the centroids of the primary patches are hierarchically pooled ($K_1 = 9$, $K_2 = 1$), yielding complementary spatial granularity.

Figure 8 shows our multi-branch encoder inspired by mSilent [87], in which features from each scale are fused through element-wise addition rather than channel concatenation. This design enforces projection into a shared feature space while reducing overfitting and parameter count. Notably, our architecture is fundamentally adapted to handle distinct data modalities compared to [87] (4D point clouds Vs. 2D time-frequency spectrograms), including modified preprocessing and encoder layers.

Each patch sequence is processed with a modified P4D-Conv layer [24], using as input: (i) relative spatio-temporal displacement $\boldsymbol{\sigma} = [\sigma_{x'}, \sigma_{y'}, \sigma_{z'}, \sigma_{t'}]^\top$, (ii) surface normals $\mathbf{n}$, and (iii) absolute displacement along the z axis, which directly reflects lung volume changes. The output is:

$$F = \sum_{t\_win} \sum_{group} \sum_{patch} (\mathbf{W}_d \cdot \boldsymbol{\sigma} + \mathbf{W}_n \cdot \mathbf{n} + \mathbf{W}_z[z']), \qquad (10)$$

where $\mathbf{W}_d, \mathbf{W}_n, \mathbf{W}_z$ are kernel weights. In our intended postures, clothing air gaps often appear as thin, stripe-like *wrinkles* in depth, which are more locally abrupt than the torso's smooth depth variation. The patch features incorporate local coherence and facilitate suppression of wrinkle outliers by leveraging nearby regions with tighter garment–body contact. Finally, a spatio-temporal transformer (PST-tran) [25] fuses patch features across space and time. We supplement its encoding with an embedding of instantaneous patch volume, boosting pressure regression efficacy.

Pressure cue and decoding: Parallel heads estimate compartmental pressures $P_{lung,t}$, $P_{abdomen,t}$ and volumes $V_{chest,t}$, $V_{abdomen,t}$ (the latter from the naïve geometric estimator in Equation 1). Motivated by Boyle's law, we inject a physically meaningful cue by multiplying pressure and volume for both chest and abdomen, and let an attention merger learn time-varying mixture weights to accommodate inter-examinee and inter-trial variability:

$$O_{enc} = \text{Attn\_Merger}\{P_{lung,t}V_{chest,t},\ P_{abdomen,t}V_{abdomen}\}. \qquad (11)$$

Sequences are sampled at 30 Hz and often exceed 200 time steps during a single MAXOUT or MAXIN$_2$, requiring long-range temporal modeling. We therefore decode O_{enc} with a 5-layer Informer [97] using ProbSparse self-attention, which retains only the most informative queries, achieving $O(n \log n)$ memory complexity and near-$O(1)$ per-step latency.

Metadata personalization and training: Finally, the Informer output and meta information (sex, height, weight, age, smoking) are projected by MLPs into a shared space and fused via multi-head attention to fine-tune the prediction; a final MLP predicts

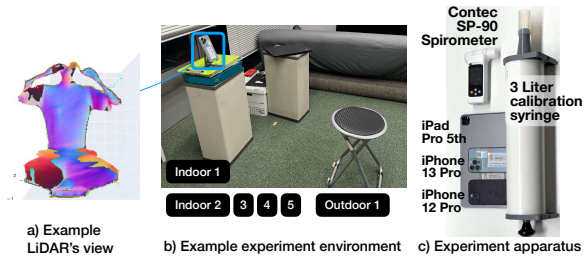


Figure 9: Setup of experiment.

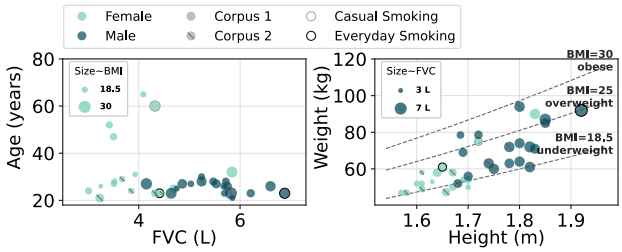


Figure 10: Demographics of cohorts 1 and 2.

$V_{\text{lung},t}$. This metadata-based personalization is calibration-free: it does not require paired ground-truth lung-volume measurements, which are often unavailable in low-resource settings. Moreover, such metadata is routinely collected in spirometry to interpret PFT results. Two identical PINNs (expiration/inspiration) are trained in the same way, using MSE loss and the Adam optimizer [38], with an initial learning rate of 5×10^{-4} and OneCycleLR [63], for 30 epochs.

5 Study Design

In this section, we first present the implementation of LiSen followed by the data collection protocol and evaluation metrics.

5.1 Implementation

5.1.1 Hardware. LiSen is implemented on two commodity smartphones (iPhone 13 Pro, iPhone 12 Pro) and one tablet (iPad Pro M1) as shown in Figure 9c. We access the LiDAR readings using a modified data collection application [83] built on the ARKit framework. The ground truth is collected by a medical-level spirometer (Contec SP-90), which records real-time volume curves and computes pulmonary function indices. Consistent with hospital practice and the ATS/ERS 2019 spirometry standard [33], we calibrate the spirometer daily using a medical-grade 3-liter syringe.

Power and storage budget: Practical daily health monitoring requires minimal impact on device resources. We therefore evaluate the proposed LiDAR sensing method in terms of battery consumption and storage overhead. Because a single PFT session uses LiDAR only briefly, direct battery measurements over short windows are unreliable. Instead, we conduct full-drain tests with continuous LiDAR sensing on an iPhone 12 Pro (2,420 mAh), iPhone 13 Pro (3,150 mAh), and iPad Pro M1 (8,825 mAh), and analyze the logs using Battery Life Profile. The devices were depleted in 100, 139, and 141 minutes, respectively. We want to note that this test is only an approximation, as the results can vary with system version, individual model differences, and environmental factors such as temperature. These results suggest that LiDAR recording for a

standard three-trial spirometry protocol of up to one minute per trial would consume at most 2.1%–3.0% of the battery. Thus, the battery impact is likely negligible.

For storage, the proposed system records up to one minute of continuous LiDAR data along with a single RGB frame, generating 179.7–216.9 MB of temporary cache files. The file size is modest, and importantly, they are deleted immediately after analysis, imposing no long-term storage burden on the user's device.

5.1.2 Software. The preprocessing modules (MPRef, CASeg) were implemented in Python, and our deep learning PINN was trained using PyTorch on a desktop computer with an Intel i9-10980XE 3.00 GHz CPU, 128 GB memory, and an Nvidia A6000 GPU. Note that *all reported results* are from models trained using other examinees' data. To accelerate the process, we adopt a Leave-Three-Subject-Out strategy (LTSO). We present the following implementation details to facilitate reproduction of LiSen:

- **AlphaPose parameter:** We use AlphaPose with the released model weights (no additional training) solely to initialize the body mask and coarse ROI. Because shoulder localization can be unstable under partial visibility, we subsequently rely on our geometry-based shoulder tracker (Algorithm 1).
- **ARKit parameter:** The ARKit API provides time-aligned RGB and depth streams, which we resample to 30 Hz.
- **Smoothing and normalization:** We reduce shoulder landmark jitter using a Hampel filter ($win\_size = 60$, σ) and moving-average smoothing ($win\_size = 5$). After depthmap-to-point-cloud conversion, each frame of chest/abdomen is normalized by translating coordinates so the chest/abdomen mask centroid is at the origin.

5.2 Study Procedures and Datasets

A dataset was constructed following approval from our Institutional Review Board (IRB). We recruited two cohorts comprising 35 healthy examinees in total (18 males and 17 females) from three countries. Figure 10 shows the wide age range of 21 to 65 years (mean 30.8) and a diverse FVC from 3.01 to 6.88 L. Cohort 1 (25 examinees, 13 male and 12 female) was used to validate each component of the proposed system and to evaluate its robustness across variations in LiDAR–subject distance, device orientation, and examinee clothing. Another 10 additional examinees (5 male and 5 female) in cohort 2 were recruited to examine the system's generalizability across varying usage scenarios, including new environments, user postures, and lighting conditions.

LiSen targets user-initiated, short PFT-style maneuvers (e.g., FVC) at close range in lower-limb-stable postures (e.g., seated, standing still, or high-Fowler's). Our data collection followed the safety principles and operational recommendations of the ATS/ERS spirometry standards [33]. For both cohorts, each data-collection session consisted of up to six FVC maneuvers. To ensure safety, a mandatory rest of at least one minute was required between maneuvers within each session, and a longer break of at least five minutes was provided after every six maneuvers. Each new attempt began only after the examinee confirmed readiness, and stopped immediately if any discomfort was reported or observed. Gold-standard spirometric data were collected under the oversight of a licensed physician.[5] During all remote data-collection sessions, the supervising physician provided protocol-level guidance, ensured compliance with the ATS/ERS 2019 spirometry standard [33], and validated the acceptability and repeatability of each maneuver prior to inclusion in the dataset.

An example collection scenario is illustrated in Figure 9b, with a smartphone placed in front of the examinee. The examinee removes heavy outerwear and stretches and wraps the garment hem to reduce slack, mitigating cloth–body air-gap artifacts. The examinee then performs the FVC test using a spirometer with a new mouthpiece and a sanitized nose clip to prevent nasal air leakage.

5.3 Evaluation Preparation

To assess the quality of the estimated V–T curve and evaluate the performance, we adopt the following metrics and benchmarks.

5.3.1 Performance Metrics.

MAXOUT MAE. The mean–absolute error (MAE) evaluated on the *expiration* segment.

MAXIN2 MAE. The MAE on the subsequent *forced inspiration* segment beginning at the end of V_{maxout}.

Pulmonary-function indices. From each regressed volume–time curve, we derive the standard spirometric indices {FVC, FEV_1, FEV_1/FVC, FIVC} (defined in Table 1). We report absolute error for the scalar quantities (FVC, FEV_1, FIVC) and percentage error for the ratio.

5.3.2 Baselines.

DeepBreath[†] [78]. Depthmap-based model for volume regression (ResNet+CBAM feature extractor + 1D U-Net regressor). Since the original motion canceller is designed for steady, periodic breathing and unsuitable for single forced maneuvers, we either remove the module or replace it with our motion-cancellation solution (**DeepBreath w/ MPRef**).

LiSen w/o Metadata. The input and processing module for examinees' metadata is removed to assess the impact of metadata personalization.

LiSen w/o MPRef. Raw FRONT WALL points are fed directly into CASeg, exposing downstream stages to global torso translation and posture drift.

LiSen w/o CASeg. Motion-cancelled points are fed straight into a single-input version of PINN, revealing the value of explicit multi-region segmentation.

LiSen w/o PINN. The PINN is replaced with a vanilla point-cloud backbone [25] (P4DConv+PST) using single-scale patches (memory limits cap at 20 patches per region vs. 40), emphasizing the gains achieved through physics-guided design.

LiSen-Sparse ($H_S \times W_S$). To simulate sparsity without access to raw LiDAR data, we decimate the iPhone depth map into a virtual grid of size $H_S \times W_S$, aggregating valid depths within each cell using the median, while preserving the original resolution so the MPRef-CASeg-PINN pipeline remains unchanged. A cell is marked valid if it overlaps the chest/abdomen mask. We evaluate 12×16 (1/256

[5]The physician's role was limited to research supervision and quality assurance and did not constitute clinical practice.

[†]The implementation was provided by the authors and retrained on our dataset for a fair comparison, excluding the multi-task learning and motion cancellation modules.

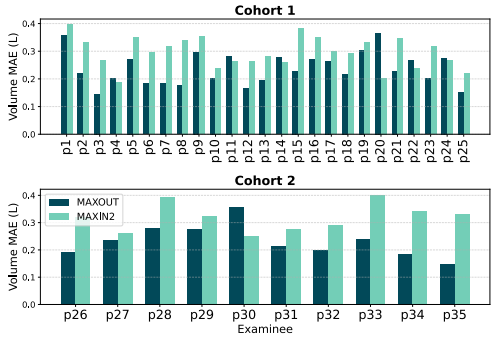


Figure 11: Per-user performance of LiSen.

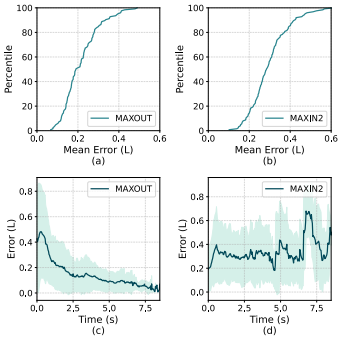


Figure 12: CDFs and temporal error distributions of volume estimation.

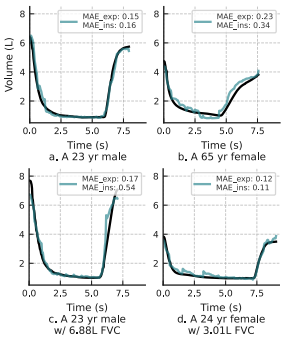


Figure 13: Example estimation across different examinees.

Method	MAXOUT (MAE)	MAXIN_2 (MAE)	FVC (%)
DeepBreath	0.57±0.13	0.40±0.16	26.22±4.68
DeepBreath w/MPRef	0.46±0.13	0.30±0.10	25.71±7.00
LiSen w/o Metadata	0.36±0.11	0.42±0.14	11.02±5.76
LiSen w/o MPRef	0.36±0.16	0.37±0.15	13.79±6.02
LiSen w/o CASeg	0.34±0.12	0.34±0.10	13.60±5.57
LiSen w/o PINN	0.27±0.09	0.44±0.14	13.10±4.99
LiSen-Sparse (12×16)	0.35±0.13	0.42±0.16	13.59±5.34
LiSen	**0.24±0.07**	**0.30±0.08**	**9.07±4.98**

Table 2: Baseline comparison.

points), indicating merit of dense multi-point and feasibility of sparser point sensing (e.g., using multi-zone laser focusing modules in commodity smartphones.)

6 Evaluation

In this section, we first present a baseline study to validate the contribution of the key components in our system. We then evaluate LiSen's generalizability across users, environments, postures, and illumination conditions. Finally, we assess its robustness under additional variations encountered in real-world settings.

6.1 Baseline Study (Cohort 1)

Table 2 presents a performance comparison between LiSen, prior-work baselines, and ablation models. With all components enabled, LiSen attains **0.24 L** MAXOUT MAE and **9.07%** FVC error, surpassing the best competing result. We also observe that LiSen achieves similarly promising performance on Cohort 2.

Compared to depth map-based DeepBreath [78], our point cloud approach improves both the expiratory (MAXOUT) and inspiratory (MAXIN_2) estimation for more than 0.10 L. However, DeepBreath performs relatively better on MAXIN_2 than on MAXOUT, showing an opposite trend to LiSen. We attribute this difference to the phase characteristics. During MAXOUT, rapid and forceful exhalation compresses volume changes into a short period, producing more consistent spatial features across users and making the richer distance-invariant features of point clouds easier to exploit. In contrast, MAXIN_2 has no speed requirement and involves averagely longer inhalation at variational rates as shown in Figure 12d. We suspect that the projected 3D point clouds are inherently more sensitive to these motion variances than the 2D depth maps.

Fortunately, our proposed modules constrain the MAXIN_2 MAE to a highly competitive level. Adding basic metadata (also required for medical assessment) personalizes the estimation, reducing both MAXOUT and MAXIN_2 errors by more than 30%. Ablating MPRef increases the MAXOUT error (where motion is heavily coupled) by roughly 50%. Removing CASeg degrades performance across both phases, highlighting the need to model the chest and abdomen separately to account for diverse breathing patterns. Finally, removing the PINN mainly increases inspiratory error, suggesting that physics guidance is particularly helpful for more diverse inputs. Furthermore, while point-density downsampling degrades performance on both phases, the resulting MAXOUT MAE and FVC still outperform DeepBreath, underscoring the robustness of our multi-point sensing workflow.

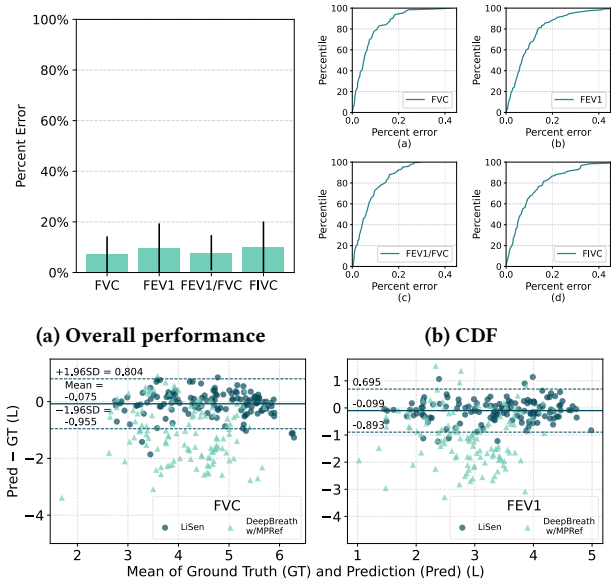
6.2 User Generalization (Cohort 1&2)

We further investigate the uncalibrated generalizability of LiSen, a critical requirement especially for low-resource settings, by evaluating per-user performance of our LTSO-trained model on both pulmonary curve and indices.

6.2.1 Volume-Time Curve Estimation. For each pair of estimated V-T curve and its ground truth, we compute the absolute error per point to compute the distribution of estimated errors. The detailed result in Figure 12 shows that 80% of samples having a MAXOUT MAE less than 0.27 L and a MAXIN_2 MAE less than 0.36 L. This distribution of volume MAE aligns with our argument in Section 6.1. While MAXOUT MAE steadily declines after the initial peak as most air is expelled, MAXIN_2 MAE remains around 0.3 L, reflecting the greater challenge of fine regression compared to MAXOUT.

We also show the per-user performance in Figure 11 and detailed sample curves in Figure 13, drawn from examinees across different age groups and FVC index categories. The largest inhalation errors occur for P1 and P33, both with a 0.40 L MAXIN_2 MAE, which remains close to the DeepBreath baseline average and demonstrates the strong user generalizability of LiSen. P1 and P33 are the tallest and shortest examinees in the cohort, suggesting that performance degradation could be mitigated with a more diverse training set.

6.2.2 Pulmonary Function Indices Estimation. We observe that LiSen not only achieves promising results in volume curve regression but also accurately estimates key pulmonary function indices. For three key pulmonary function indices—FVC, FEV1,



(a) Overall performance **(b) CDF**

(c) Bland-Altman analysis

Figure 14: Pulmonary function index estimation.

FEV1/FVC—plus FIVC, LiSen achieves mean errors of 7.10%, 9.82%, 7.84%, and 10.05%, respectively, as shown in Figure 14a. The detailed cumulative distributions are further shown in Figure 14b. The small FVC error is even lower than audio-based measurement reported in [77] (9.9%), where the tube-blowing sound is more directly connected to airflow and cumulated volume compared to the complex relation of FRONT WALL.

Furthermore, we evaluate the clinical reliability using Bland-Altman analysis as shown in Figure 14c. For both FVC and FEV1, LiSen achieves a mean bias within 0.10 L and 95% limits of agreement (LoA) of [−0.96, 0.80] L for FVC and [−0.89, 0.70] L for FEV1, which is substantially more stable than the DeepBreath w/MPRef baseline. From a clinical perspective, the near-zero mean bias indicates minimal systematic error. Although the LoA is wider than the stringent diagnostic tolerances of clinical spirometers, it remains practical for large-scale contactless screening and long-term pulmonary monitoring in home-care settings.

6.3 Scenario Generalization (Cohort 2)

To assess whether LiSen generalizes beyond a fixed lab setting (*Indoor 1 Stool*), we recruited Cohort 2 and evaluated the system across six scenarios varying in environment (indoor vs. outdoor), posture (stool, standing, wheelchair, high Fowler), and illumination (daylight vs. low-light) as shown in Figure 16. We report the performance of our model trained exclusively on Cohort 1 (*Indoor 1 Stool*) and tested on Cohort 2 under the 6 new scenarios. This cross-cohort approach ensures that both the examinees and the specific scenario are strictly unseen during training. For comparison, we also provide a baseline result of an LTSO-trained PINN tested on the *Indoor 2 Stool* scenario. Figure 15 summarizes the percentage errors for three key spirometric indices across these settings.

Across all indoor conditions, the mean FVC error remains between 8.3–10.5%. FEV1 errors are similarly bounded within 7.4–11.5%, and FEV1/FVC errors within 5.9–13.8%, with overlapping standard deviations across scenarios, indicating our system can adapt to environmental variation and common postures. The greatest FVC error of 10.5 ± 4.7% is attained in *Indoor 4 Wheelchair*, suggesting that the less stable seated posture in a wheelchair may increase lower-body instability during the test. However, we do not observe systematic degradation when moving from the indoor baseline to outdoor use, nor when changing postures from neutral stool sitting to standing, wheelchair sitting, or high-Fowler posture.

The outdoor experiments further highlight the robustness of LiDAR-based approaches under varying illumination in open-air environments. The FVC and FEV1 errors under *Outdoor 1 Stool & Dark* (8.0 ± 8.4% and 11.6 ± 5.1%) are even better than those under normal daylight (*Outdoor 1 Stool*: 9.6 ± 3.0% and 13.0 ± 4.2%). We suspect the increased noise arises from infrared-band interference from the sunlight. Overall, these results indicate that our multi-point LiDAR sensing and motion-robust PFT estimation generalizes across diverse real-world scenarios without recalibration, supporting the feasibility of at-home and out-of-clinic FVC testing.

6.4 Effect of Real-World Setup Diversities

We further evaluate the impact of real-world setup diversities, using the average performance of LiSen under the setup in Figure 9 as the neutral baseline. The model is trained solely on data from the baseline setup and tested on data from additional setups (partly shown in Figure 18) using a leave-three-subject-out (LTSO) protocol, with all hyperparameters identical to those described in Section 4.4.2. The performance is reported in Figure 17.

LiDAR–subject distance. Moving the phone 20 cm–40 cm further away (1.0 m, 1.2 m) has negligible impact: MAXOUT MAE even drops slightly (from 0.30 L to 0.26 L) thanks to a wider field of view, while MAXIN_2 increases by only 0.02 L. These results suggest that users do not need to position the device with centimeter-level precision.

LiDAR angle. At a 15° yaw offset, the errors remain statistically unchanged, and at 30° the increase is limited to ≤0.02 L, indicating that the multi-point reference effectively compensates for moderate perspective distortion under orientation variations.

Clothing. As Figure 17 shows, typical indoor garments (e.g., a *mild* cardigan or midweight sweater, shown in Figure 18) introduce marginal error (MAXOUT MAE ≤ 0.28 L, MAXIN_2 MAE ≤ 0.39 L). A *loose oversized sweater* further degrades performance (MAXOUT MAE 0.34 L, MAXIN_2 MAE 0.42 L), as the loose fabric easily decouples from the body and obscures subtle respiratory displacements. However, the degradation remains bounded, confirming that LiDAR's spatial features effectively compensate for the clothing-induced air gap.

Devices. Curves from the iPad Pro and iPhone 12 Pro differ by less than 0.03 L from the baseline (iPhone 13 Pro), demonstrating that LISEN generalizes across different models without re-calibration.

Summary. Across all 15 perturbed conditions, MAXOUT MAE never exceeds 0.42 L and MAXIN_2 never exceeds 0.48 L—only 0.12–0.18 L above baseline. These bounded drifts confirm that motion cancellation, region segmentation, and physics-based constraints remain effective under common real-world variations.

7 Related Work

We first survey wireless-sensing systems, and then mobile device-based PFT methods, before positioning our laser ranging approach.

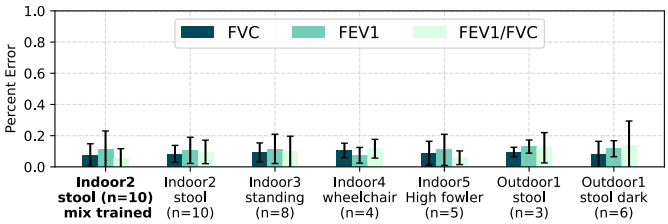


Figure 15: Generalizability to scenarios (environment, posture, illumination).

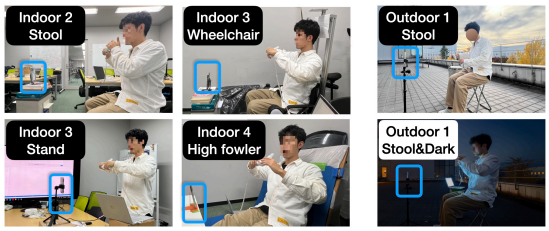


Figure 16: Experiment scenarios (environment, posture, illumination).

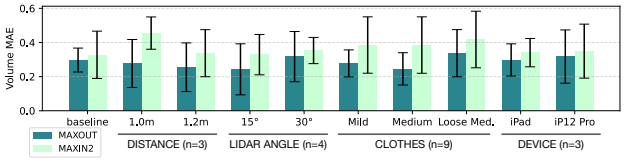


Figure 17: Performance under setup diversity.

Figure 18: Part of real-world setup diversity; (a) diverse clothing; (b) different range and LiDAR orientation).

7.1 Conventional Wireless Sensing

A wide range of wireless signals, such as WiFi, LoRa, RFID, UWB/mmWave, and acoustics, have been employed for sensing [22, 57, 96]. In particular, ubiquitous WiFi infrastructure has been used for respiration monitoring [89, 90], indoor localization and tracking [45–47], and activity recognition [28]. LoRa's long-range capability enables large-area sensing [16, 75, 91]. With wider bandwidth, UWB/mmWave supports fine-grained perception, including body point clouds [58, 67] and posture estimation [60]. Commodity speakers/microphones also enable ultrasonic sensing at inaudible bands, powering diverse applications [17, 20, 21, 27, 29, 43, 86]. However, typical RF and acoustic transmitters produce wide beams, resulting in strong multipath effects and coarse angular resolution—limitations that we address with narrow-beam LiDAR.

7.2 Mobile PFT

Prior systems using LiDAR [34, 59, 95] only detect normal respiration patterns without quantifying lung volumes for full PFT. Among PFT-oriented works, audio-based PFT [30, 41, 66, 77] is vulnerable to audio interference, and those built with earphones [13, 80] can be intrusive for some examinees. Range-based PFT systems [1, 64, 66, 84], such as those using Kinect camera [18], face challenges such as motion interference. For example, DeepBreath [78] relies on periodic signals for motion removal, and does not explicitly separate chest vs. abdomen. In contrast, LiSen enables uncalibrated FVC prediction by canceling motion without relying on periodicity, compensating for intra-ambient pressure differences, and explicitly modeling chest and abdomen contributions. To our knowledge, LiSen is the first mobile FVC system that addresses these challenges.

8 Discussion

Privacy concerns. High-precision depth sensing inevitably captures information about a user's body and environment, raising potential privacy concerns. Future work could incorporate privacy-preserving approaches such as differential privacy [23], federated learning [44], or encryption [32].

Other sensing modalities. Camera-based methods depend on sufficient lighting and raise privacy concerns [6]. The coarse spatial resolution of mmWave radar and Wi-Fi limits their ability to capture fine-grained body motion. For example, MSense [15] tracks only a few points rather than dense multi-points, restricting detailed spatial analysis. Acoustic methods can capture airflow sounds but cannot distinguish between chest and abdominal contributions, and are vulnerable to ambient noise [30, 66, 77, 85].

Limitations of LiDAR. Despite its advantages, smartphone LiDAR has limitations—short sensing range, narrow field of view, and bundled LiDAR–RGB API access—resulting in higher energy consumption than RF or audio-based continuous monitoring. This limits its suitability for long-term, always-on respiratory tracking. However, at-home PFTs like spirometry require only brief, close-range measurements (e.g., 80 cm), making smartphone LiDAR well suited for such tasks.

9 Conclusion

We introduce the concept of multi-point sensing, and propose the first smartphone-LiDAR system that transforms a ubiquitous LiDAR-equipped phone into a calibration-free at-home pulmonary function sensor, recovering full volume curves from multi-point torso measurements. We envision that the multi-point sensing modality could inspire other applications requiring fine-grained, multi-point spatial information.

Acknowledgments

This work was supported by JST PRESTO, Japan, under Grant Number JPMJPR2527, by JST SPRING, Grant Number JPMJSP2108, and by a research fund endowed by Kubota Co., Ltd. This research was also supported by the NTU SUG-NAP. This research was also supported by the MSIT (Ministry of Science and ICT), Korea, under the Global Scholars Invitation Program (RS-2024-00456287) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation). We thank Dr. Chuang Yang for his valuable advice regarding the neural network design.

References

[1] Aakriti Adhikari, Austin Hetherington, and Sanjib Sur. 2021. mmFlow: Facilitating At-Home Spirometry with 5G Smart Devices. In *2021 18th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 1–9. doi:10.1109/SECON52354.2021.9491616

[2] Emilio Agostoni and Hermann Rahn. 1960. Abdominal and thoracic pressures at different lung volumes. *Journal of applied physiology* 15, 6 (1960), 1087–1092.

[3] Richard K Albert, Stephen G Spiro, and James R Jett. 2008. *Clinical Respiratory Medicine: Expert Consult-Online and Print*. Elsevier Health Sciences.

[4] Andrea Aliverti, Raffaele Dellacà, Paolo Pelosi, Davide Chiumello, Luciano Gattinoni, and Antonio Pedotti. 2001. Compartmental analysis of breathing in the supine and prone positions by optoelectronic plethysmography. *Annals of biomedical engineering* 29 (2001), 60–70.

[5] Andrea Aliverti, Davide Lacca, and Antonella LoMauro. 2022. Quantitative analysis by 3D graphics of thoraco-abdominal surface shape and breathing motion. *Frontiers in Bioengineering and Biotechnology* 10 (2022), 910499.

[6] Apple Inc. 2021. *Shot on iPhone 12: Portraits, cityscapes, the night sky, and more*. Apple Newsroom article stating that the LiDAR scanner enables up to 6x faster autofocus in low-light scenes on iPhone 12 Pro models.

[7] Robert C Augusteyn, Derek Nankivil, Ashik Mohamed, Bianca Maceo, Faradia Pierre, and Jean-Marie Parel. 2012. Human ocular biometry. *Experimental eye research* 102 (2012), 70–75.

[8] Gila Benchetrit. 2000. Breathing pattern in humans: diversity and individuality. *Respiration physiology* 122, 2-3 (2000), 123–129.

[9] Shengze Cai, Zhiping Mao, Zhicheng Wang, Minglang Yin, and George Em Karniadakis. 2021. Physics-informed neural networks (PINNs) for fluid mechanics: A review. *Acta Mechanica Sinica* 37, 12 (2021), 1727–1738.

[10] SJ Cala, CM Kenyon, Giancarlo Ferrigno, P Carnevali, Andrea Aliverti, Antonio Pedotti, PT Macklem, and DF Rochester. 1996. Chest wall and lung volume estimation by optical reflectance motion analysis. *Journal of Applied Physiology* 81, 6 (1996), 2680–2689.

[11] Miles Campbell and Amit Sapra. 2023. Physiology, airflow resistance. In *StatPearls [Internet]*. StatPearls Publishing.

[12] Shirui Cao, Dong Li, Sunghoon Ivan Lee, and Jie Xiong. 2023. PowerPhone: Unleashing the Acoustic Sensing Capability of Smartphones. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*. ACM, New York, NY, USA, 1–16.

[13] Yetong Cao, Dong Ma, Wentao Xie, Qian Zhang, and Jun Luo. 2025. ESPIRO: Natural Pulmonary Function Monitoring via Earphone-Acquired Speech. In *Proceedings of the 31st Annual International Conference on Mobile Computing and Networking*. ACM, New York, NY, USA, 329–344.

[14] Leon Chaitow, Dinah Bradley, and Chris Gilbert. 2014. The structure and function of breathing. *Recognising and treating breathing pattern disorders: a multidisciplinary approach. 2nd ed.* (2014), 23–43.

[15] Zhaoxin Chang, Fusang Zhang, Jie Xiong, Weiyan Chen, and Daqing Zhang. 2024. MSense: Boosting Wireless Sensing Capability Under Motion Interference. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 108–123.

[16] Zhaoxin Chang, Fusang Zhang, Jie Xiong, Junqi Ma, Beihong Jin, and Daqing Zhang. 2022. Sensor-free soil moisture sensing using lora signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–27.

[17] Haiming Cheng, Wei Lou, Yanni Yang, Yi-pu Chen, and Xinyu Zhang. 2023. TwinkleTwinkle: Interacting with Your Smart Devices by Eye Blink. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 2 (2023), 1–30.

[18] Willem de Boer, Joan Lasenby, Jonathan Cameron, Rich Wareham, Shiraz Ahmad, Charlotte Roach, Ward Hills, and Richard Iles. 2010. SLP: A Zero-Contact Non-Invasive Method for Pulmonary Function Testing.. In *BMVC*. 1–12.

[19] Anne De Groote, Muriel Wantier, Guy Chéron, Marc Estenne, and Manuel Paiva. 1997. Chest wall motion during tidal breathing. *Journal of Applied Physiology* 83, 5 (1997), 1531–1537.

[20] Xuefu Dong, Yifei Chen, Yuuki Nishiyama, Kaoru Sezaki, Yuntao Wang, Ken Christofferson, and Alex Mariakakis. 2024. ReHEarSSE: Recognizing Hidden-in-the-Ear Silently Spelled Expressions. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–16.

[21] Xuefu Dong, Liqiang Xu, Lixing He, Zengyi Han, Kenneth Christofferson, Yifei Chen, Akihito Taya, Yuuki Nishiyama, and Kaoru Sezaki. 2025. Poster: Recognizing Hidden-in-the-Ear Private Key for Reliable Silent Speech Interface Using Multi-Task Learning. In *Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 211–215.

[22] Di Duan, Shengzhe Lyu, Mu Yuan, Hongfei Xue, Tianxing Li, Weitao Xu, Kaishun Wu, and Guoliang Xing. 2025. Argus: Multi-View Egocentric Human Mesh Reconstruction Based on Stripped-Down Wearable mmWave Add-on. In *Proceedings of the 23rd ACM Conference on Embedded Networked Sensor Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3715014.3722045

[23] Cynthia Dwork. 2006. Differential privacy. In *International colloquium on automata, languages, and programming*. Springer, 1–12.

[24] Hehe Fan, Yi Yang, and Mohan Kankanhalli. 2021. Point 4d transformer networks for spatio-temporal modeling in point cloud videos. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 14204–14213.

[25] Hehe Fan, Yi Yang, and Mohan Kankanhalli. 2022. Point spatio-temporal transformer networks for point cloud video modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 2 (2022), 2181–2192.

[26] Hao-Shu Fang, Jiefeng Li, Hongyang Tang, Chao Xu, Haoyi Zhu, Yuliang Xiu, Yong-Lu Li, and Cewu Lu. 2022. Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time. *IEEE transactions on pattern analysis and machine intelligence* 45, 6 (2022), 7157–7173.

[27] Yongjian Fu, Shuning Wang, Linghui Zhong, Lili Chen, Ju Ren, and Yaoxue Zhang. 2022. SVoice: Enabling Voice Communication in Silence via Acoustic Sensing on Commodity Devices. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*. Association for Computing Machinery, New York, NY, USA, 622–636.

[28] Ruiyang Gao, Mi Zhang, Jie Zhang, Yang Li, Enze Yi, Dan Wu, Leye Wang, and Daqing Zhang. 2021. Towards Position-Independent Sensing for Gesture Recognition with Wi-Fi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 2 (2021), 1–28.

[29] Tian Gao, Xuefu Dong, Akihito Taya, Yuuki Nishiyama, and Kaoru Sezaki. 2025. Expression Recognition Based on Ear Canal Shape Detection Using Earbud and Ultrasound. In *IEICE Conferences Archives*. The Institute of Electronics, Information and Communication Engineers.

[30] 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 *Proceedings of the 2016 CHI conference on human factors in computing systems*. 5675–5685.

[31] Yanbin Gong, Qian Zhang, Bobby H.P. NG, and Wei Li. 2022. BreathMentor: Acoustic-based Diaphragmatic Breathing Monitor System. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 2, Article 53 (jul 2022), 28 pages. doi:10.1145/3534595

[32] Thore Graepel, Kristin Lauter, and Michael Naehrig. 2012. ML confidential: Machine learning on encrypted data. In *International conference on information security and cryptology*. Springer, 1–21.

[33] 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.

[34] Junying Hu, Yongjian Fu, Lili Chen, Xinyi Li, Xue Sun, Ju Ren, and Yaoxue Zhang. 2024. Unveiling Contactless Sensing with LiDAR Mobility. In *2024 IEEE Smart World Congress (SWC)*. IEEE, 487–494.

[35] ROBERT E Hyatt and RAYMOND E Flath. 1966. Relationship of air flow to pressure during maximal respiratory effort in man. *Journal of Applied Physiology* 21, 2 (1966), 477–482.

[36] Junseok Jang and Kwangjae Sung. 2024. Design of Indoor Space Guidance System Using LiDAR and Camera on iPhone. *Journal of the Semiconductor & Display Technology* 23, 1 (2024), 71–78.

[37] Hongbo Jiang, Jingyang Hu, Daibo Liu, Jie Xiong, and Mingjie Cai. 2021. Driversonar: Fine-grained dangerous driving detection using active sonar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–22.

[38] Diederik P. Kingma and Jimmy Ba. 2017. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs.LG]

[39] Kris Kitani. 2023. 16-385 Computer Vision. Course slides, Carnegie Mellon University. https://www.cs.cmu.edu/~16385/.

[40] Kimio Konno and Jere Mead. 1967. Measurement of the separate volume changes of rib cage and abdomen during breathing. *Journal of applied physiology* 22, 3 (1967), 407–422.

[41] 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 *Proceedings of the 2012 ACM Conference on ubiquitous computing*. ACM, New York, NY, USA, 280–289.

[42] Dong Li, Shirui Cao, Sunghoon Ivan Lee, and Jie Xiong. 2022. Experience: practical problems for acoustic sensing. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 381–390.

[43] Dong Li, Jialin Liu, Sunghoon Ivan Lee, and Jie Xiong. 2022. LASense: Pushing the Limits of Fine-grained Activity Sensing Using Acoustic Signals. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 1, Article 21 (mar 2022), 27 pages. doi:10.1145/3517253

[44] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. 2020. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine* 37, 3 (2020), 50–60.

[45] Wenwei Li, Ruiyang Gao, Jie Xiong, Jiarun Zhou, Leye Wang, Xingjian Mao, Enze Yi, and Daqing Zhang. 2024. WiFi-CSI difference paradigm: Achieving

efficient doppler speed estimation for passive tracking. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 2 (2024), 1–29.

[46] Wenwei Li, Jiarun Zhou, Jie Xiong, Yuhui Xie, Leye Wang, Duo Zhang, and Daqing Zhang. 2025. Rethinking WiFi-based Angle Estimation for Robust Passive Indoor Localization. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 4 (2025), 1–28.

[47] Xiang Li, Daqing Zhang, Qin Lv, Jie Xiong, Shengjie Li, Yue Zhang, and Hong Mei. 2017. IndoTrack: Device-free indoor human tracking with commodity Wi-Fi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 72.

[48] Jinyi Liu, Wenwei Li, Tao Gu, Ruiyang Gao, Bin Chen, Fusang Zhang, Dan Wu, and Daqing Zhang. 2023. Towards a Dynamic Fresnel Zone Model to WiFi-based Human Activity Recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 2, Article 65 (jun 2023), 24 pages. doi:10.1145/3596270

[49] John M Luce and Bruce H Culver. 1982. Respiratory muscle function in health and disease. *Chest* 81, 1 (1982), 82–90.

[50] Andrew B Lumb and Caroline R Thomas. 2020. *Nunn's Applied Respiratory Physiology eBook: Nunn's Applied Respiratory Physiology eBook*. Elsevier Health Sciences.

[51] Wenguang Mao, Mei Wang, Wei Sun, Lili Qiu, Swadhin Pradhan, and Yi-Chao Chen. 2019. Rnn-based room scale hand motion tracking. In *The 25th Annual International Conference on Mobile Computing and Networking*. ACM, New York, NY, USA, 1–16.

[52] Tomoyuki Ogino, Kyoshi Mase, Masafumi Nozoe, Tomohiro Wada, Yuki Uchiyama, Yoshihiro Fukuda, and Kazuhisa Domen. 2015. Effects of arm bracing on expiratory flow limitation and lung volume in elderly COPD subjects. *Respiratory care* 60, 9 (2015), 1282–1287.

[53] KyeongTaek Oh, Cheung Soo Shin, Jeongmin Kim, and Sun K Yoo. 2018. Level-set segmentation-based respiratory volume estimation using a depth camera. *IEEE Journal of Biomedical and Health Informatics* 23, 4 (2018), 1674–1682.

[54] Tadsawiya Padkao and Orachorn Boonla. 2020. Relationships between respiratory muscle strength, chest wall expansion, and functional capacity in healthy nonsmokers. *Journal of exercise rehabilitation* 16, 2 (2020), 189.

[55] Ronald K Pearson, Yrjö Neuvo, Jaakko Astola, and Moncef Gabbouj. 2016. Generalized hampel filters. *EURASIP Journal on Advances in Signal Processing* 2016 (2016), 1–18.

[56] TJ Pedley, RC Schroter, and MF Sudlow. 1970. The prediction of pressure drop and variation of resistance within the human bronchial airways. *Respiration physiology* 9, 3 (1970), 387–405.

[57] Helinyi Peng, Akihito Taya, Yuuki Nishiyama, and Kaoru Sezaki. 2026. AED-Hunter: Investigating AED Retrieval in the Real World via Gamified Mobile Interaction and Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 10, 1, Article 19 (March 2026), 30 pages. doi:10.1145/3789689

[58] Kun Qian, Zhaoyuan He, and Xinyu Zhang. 2020. 3D point cloud generation with millimeter-wave radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–23.

[59] Omar Rinchi, Ahmad Alsharoa, and Denise A Baker. 2024. Remote breathing monitoring using lidar technology. In *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 1–6.

[60] Arindam Sengupta, Feng Jin, Renyuan Zhang, and Siyang Cao. 2020. mm-Pose: Real-time human skeletal posture estimation using mmWave radars and CNNs. *IEEE Sensors Journal* 20, 17 (2020), 10032–10044.

[61] Tiina M Seppänen, Janne Kananen, Kai Noponen, Olli-Pekka Alho, and Tapio Seppänen. 2015. Accurate measurement of respiratory airflow waveforms using depth data. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 7857–7860.

[62] Yun Su Sim, Ji-Hyun Lee, Won-Yeon Lee, Dong In Suh, Yeon-Mok Oh, Jong-seo Yoon, Jin Hwa Lee, Jae Hwa Cho, Cheol Seok Kwon, and Jung Hyun Chang. 2017. Spirometry and bronchodilator test. *Tuberculosis and respiratory diseases* 80, 2 (2017), 105–112.

[63] Leslie N Smith and Nicholay Topin. 2019. Super-convergence: Very fast training of neural networks using large learning rates. In *Artificial intelligence and machine learning for multi-domain operations applications*, Vol. 11006. SPIE, 369–386.

[64] Vahid Soleimani, Majid Mirmehdi, Dima Damen, James Dodd, Sion Hannuna, Charles Sharp, Massimo Camplani, and Jason Viner. 2016. Remote, depth-based lung function assessment. *IEEE Transactions on Biomedical Engineering* 64, 8 (2016), 1943–1958.

[65] Vahid Soleimani, Majid Mirmehdi, Dima Damen, Sion Hannuna, Massimo Camplani, Jason Viner, and James Dodd. 2015. Remote pulmonary function testing using a depth sensor. In *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 1–4.

[66] Xingzhe Song, Boyuan Yang, Ge Yang, Ruirong Chen, Erick Forno, Wei Chen, and Wei Gao. 2020. SpiroSonic: monitoring human lung function via acoustic sensing on commodity smartphones. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*. ACM, New York, NY, USA, 1–14.

[67] Yongkun Song, Ke Zhang, Tianxing Yan, Tian Jin, Yongpeng Dai, and Tao Wang. 2025. Dual-Modal Human Pose Reconstruction using 3D UWB Radar Images and Point Clouds. *IEEE Transactions on Instrumentation and Measurement* (2025).

[68] Gledson Rodrigo Tondo, Charles Riley, and Guido Morgenthal. 2023. Characterization of the iPhone LiDAR-based sensing system for vibration measurement and modal analysis. *Sensors* 23, 18 (2023), 7832.

[69] Nutsupa Ubolnuar, Anong Tantisuwat, Premtip Thaveeratitham, Somrat Lertmaharit, Chathipat Kruapanich, Jaturong Chimpalee, and Witaya Mathiyakom. 2020. Effects of pursed-lip breathing and forward trunk lean postures on total and compartmental lung volumes and ventilation in patients with mild to moderate chronic obstructive pulmonary disease: an observational study. *Medicine* 99, 51 (2020), e23646.

[70] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* 17, 3 (2020), 261–272.

[71] Ju Wang, Jie Xiong, Xiaojiang Chen, Hongbo Jiang, Rajesh Krishna Balan, and Dingyi Fang. 2017. TagScan: Simultaneous Target Imaging and Material Identification with Commodity RFID Devices. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking* (Snowbird, Utah, USA) *(MobiCom '17)*. Association for Computing Machinery, New York, NY, USA, 288–300. doi:10.1145/3117811.3117830

[72] Watts R Webb, Josef R Smith, and Guy D Campbell. 1961. Peri-alveolar pressures: measurement and significance in pulmonary dynamics. *Annals of Surgery* 153, 5 (1961), 650.

[73] Wikipedia. 2025. Respiratory pressure meter — Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/w/index.php?title=Respiratory%20pressure%20meter&oldid=1262912358. [Online; accessed 17-April-2025].

[74] Chenshu Wu, Feng Zhang, Beibei Wang, and KJ Ray Liu. 2020. mSense: Towards mobile material sensing with a single millimeter-wave radio. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–20.

[75] Binbin Xie, Deepak Ganesan, and Jie Xiong. 2022. Embracing lora sensing with device mobility. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*. ACM, New York, NY, USA, 349–361.

[76] Binbin Xie, Jie Xiong, Xiaojiang Chen, Eugene Chai, Liyao Li, Zhanyong Tang, and Dingyi Fang. 2019. Tagtag: material sensing with commodity RFID. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems* (New York, New York) *(SenSys '19)*. Association for Computing Machinery, New York, NY, USA, 338–350. doi:10.1145/3356250.3360027

[77] Wentao Xie, Qingyong Hu, Jin Zhang, and Qian Zhang. 2023. EarSpiro: Earphone-based Spirometry for Lung Function Assessment. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–27.

[78] Wentao Xie, Chi Xu, Yanbin Gong, Yu Wang, Yuxin Liu, Jin Zhang, Qian Zhang, Zeguang Zheng, and Shifang Yang. 2024. DeepBreath: Breathing Exercise Assessment with a Depth Camera. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (2024), 1–26.

[79] Jie Xiong and Kyle Jamieson. 2013. ArrayTrack: A Fine-Grained Indoor Location System. In *10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*. 71–84.

[80] Chi Xu, Wentao Xie, Baichen Yang, Yizhen Zhang, Yanbin Gong, Jin Zhang, Wei Li, Shifang Yang, and Qian Zhang. 2025. EasySpiro: Assessing Lung Function via Arbitrary Exhalations on Commodity Earphones. In *Proceedings of the 31st Annual International Conference On Mobile Computing and Networking*. ACM, New York, NY, USA, 513–528.

[81] Xiaoyan Yang, Haishuang Sun, Mei Deng, Yicong Chen, Chen Li, Pengxin Yu, Rongguo Zhang, Min Liu, Huaping Dai, and Chen Wang. 2022. Characteristics of diaphragmatic and chest wall motion in people with normal pulmonary function: a study with free-breathing dynamic MRI. *Journal of Clinical Medicine* 11, 24 (2022), 7276.

[82] Zhiyun Yao, Xuanzhi Wang, Kai Niu, Rong Zheng, Junzhe Wang, and Daqing Zhang. 2024. WiProfile: Unlocking Diffraction Effects for Sub-Centimeter Target Profiling Using Commodity WiFi Devices. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 185–199.

[83] Chandan Yeshwanth, Yueh-Cheng Liu, Matthias Nießner, and Angela Dai. 2023. Scannet++: A high-fidelity dataset of 3d indoor scenes. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 12–22.

[84] 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 *2022 IEEE/ACM 30th International Symposium on Quality of Service (IWQoS)*. 1–10. doi:10.1109/IWQoS54832.2022.9812913

[85] Kuang Yuan, Dong Li, Hao Zhou, Zhehao Li, Lili Qiu, Swarun Kumar, and Jie Xiong. 2025. WindDancer: Understanding Acoustic Sensing under Ambient Airflow. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 2 (2025), 1–25.

[86] Sangki Yun, Yi-Chao Chen, Huihuang Zheng, Lili Qiu, and Wenguang Mao. 2017. Strata: Fine-grained acoustic-based device-free tracking. In *Proceedings of the 15th annual international conference on mobile systems, applications, and services*. Association for Computing Machinery, New York, NY, USA, 15–28.

[87] Shang Zeng, Haoran Wan, Shuyu Shi, and Wei Wang. 2023. mSilent: Towards general corpus silent speech recognition using COTS mmWave radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–28.

[88] Youwei Zeng, Dan Wu, Jie Xiong, Jinyi Liu, Zhaopeng Liu, and Daqing Zhang. 2020. MultiSense: Enabling Multi-person Respiration Sensing with Commodity WiFi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 102 (sep 2020), 29 pages. doi:10.1145/3411816

[89] Youwei Zeng, Dan Wu, Jie Xiong, Jinyi Liu, Zhaopeng Liu, and Daqing Zhang. 2020. MultiSense: Enabling multi-person respiration sensing with commodity WiFi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–29.

[90] Youwei Zeng, Dan Wu, Jie Xiong, Enze Yi, Ruiyang Gao, and Daqing Zhang. 2019. Farsense: Pushing the range limit of wifi-based respiration sensing with csi ratio of two antennas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–26.

[91] Fusang Zhang, Zhaoxin Chang, Kai Niu, Jie Xiong, Beihong Jin, Qin Lv, and Daqing Zhang. 2020. Exploring lora for long-range through-wall sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–27.

[92] Fusang Zhang, Zhaoxin Chang, Jie Xiong, Junqi Ma, Jiazhi Ni, Wenbo Zhang, Beihong Jin, and Daqing Zhang. 2023. Embracing consumer-level UWB-equipped devices for fine-grained wireless sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–27.

[93] Fusang Zhang, Zhi Wang, Beihong Jin, Jie Xiong, and Daqing Zhang. 2020. Your Smart Speaker Can" Hear" Your Heartbeat! *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–24.

[94] Fusang Zhang, Jie Xiong, Zhaoxin Chang, Junqi Ma, and Daqing Zhang. 2022. Mobi2Sense: empowering wireless sensing with mobility. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 268–281.

[95] Mingshi Zhang, Kangning Liu, Yubing Wang, Yuqing Chen, Yuxin Lei, Lei Liang, Jianwei Zhang, Jun Zhang, Li Qin, and Lijun Wang. 2025. Phase-modulated coherent LiDAR for contactless, privacy-preserving respiration monitoring. *IEEE Sensors Journal* (2025).

[96] Hao Zhou, Kuang Yuan, Mahanth Gowda, Lili Qiu, and Jie Xiong. 2024. Rethinking orientation estimation with smartphone-equipped ultra-wideband chips. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. ACM, New York, NY, USA, 1045–1059.

[97] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. 2021. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 35. 11106–11115.