

Ultrasound Lung Aeration Map via Physics-Aware Neural Operators

Jiayun Wang^{1*}, Oleksii Ostrasz^{2,3*}, Masashi Sode^{2,3},

Bahareh Tolooshams¹, Zongyi Li¹, Kamyar Azizzadenesheli⁴,

Gianmarco F. Pinton^{2,3‡}, Anima Anandkumar^{1‡}

¹Department of Computing and Mathematical Sciences, California Institute of Technology, 1200 E California Blvd, Pasadena, 91125, CA, United States.

²Department of Biomedical Engineering, University of North Carolina at Chapel Hill, 103 South Building, Chapel Hill, 27514, NC, United States.

³Department of Biomedical Engineering, North Carolina State University, Campus Box 7625, Raleigh, 27695, NC, United States.

⁴ NVIDIA, 2788 San Tomas Express Way, Santa Clara, 95051, CA, United States.

Contributing authors: peterw@caltech.edu; oleksii@email.unc.edu; msode@email.unc.edu; btoloosh@caltech.edu; zongyili@caltech.edu; kamyara@nvidia.com; gia@email.unc.edu; anima@caltech.edu;

*Equal contribution.

‡Equal advising.

Abstract

Lung ultrasound (LUS) is a growing modality in clinics for diagnosing and monitoring acute and chronic lung diseases due to its low cost and accessibility. LUS works by emitting diagnostic pulses, receiving pressure waves and converting them into radio frequency (RF) data, which are then processed into B-mode images for radiologists to interpret. However, unlike conventional ultrasound for soft tissue anatomical imaging, LUS interpretation is complicated by complex reverberations from the pleural interface caused by the inability of ultrasound to penetrate air. The indirect B-mode images make interpretation highly dependent on reader expertise, requiring years of training, which limits its widespread use despite its potential for high accuracy in skilled hands.

047 To address these challenges and democratize LUS as a reliable diagnostic tool, we
048 propose LUNA, an AI model that directly reconstructs lung aeration maps from
049 RF data, bypassing the need for indirect interpretation of B-mode images. LUNA
050 uses a Fourier neural operator, which processes RF data efficiently in Fourier
051 space, enabling accurate reconstruction of lung aeration maps. From reconstructed
052 aeration maps, we calculate lung percent aeration, a key clinical metric, offering a
053 quantitative, reader-independent alternative to traditional semi-quantitative LUS
054 scoring methods. Trained on abundant simulated data and fine-tuned with a small
055 amount of real-world data, LUNA achieves robust performance, demonstrated by
056 an aeration estimation error of less than 10% in ex-vivo swine lung scans. We
057 demonstrate the potential of directly reconstructing lung aeration maps from RF
058 data, providing a foundation for improving LUS interpretability, reproducibility
059 and diagnostic utility.

060 **Keywords:** ultrasound, lung imaging, lung aeration, deep learning, physics-aware
061 learning, neural operator, operator learning

064 1 Introduction

066 Lung ultrasound (LUS) is an important non-invasive real-time imaging modality widely
067 used for diagnostics and monitoring of lung disease in its acute and chronic phases,
068 such as respiratory and extrapulmonary diseases [4–6]. Compared to X-ray imaging and
069 computed tomography (CT), LUS has inherent advantages: it is non-ionizing, portable,
070 low-cost and suitable for frequent or bedside monitoring. Furthermore, the unique
071 acoustic interaction between soft tissue and air provides distinct contrast mechanisms,
072 potentially offering complementary diagnostic information to X-ray and CT imaging
073 – expert users can achieve a sensitivity and specificity of 90% to 100% for disorders
074 including pleural effusions, lung consolidation, pneumothorax and interstitial syndrome
075 [7].

076 Despite its promise, barriers are limiting the widespread adoption of LUS because
077 interpreting lung ultrasound images usually requires years of experience [8, 9]. Unlike
078 other diagnostic ultrasound (e.g. fetal and abdominal ultrasound [10, 11]), LUS images
079 primarily rely on the interpretation of artifacts created by reverberations at the tissue-
080 air interface [12], rather than direct visualization of lung structures [13]. This indirect
081 imaging mechanism makes clinical interpretation highly dependent on the operator’s
082 expertise and the imaging system’s parameters (e.g., central frequency, focal depth
083 gain settings), leading to significant variability and insufficient interobserver agreement
084 [14–16]. Furthermore, current delay-and-sum beamforming methods, designed for soft
085 tissue imaging, are not optimized for the unique acoustic physics of lung tissue, limiting
086 diagnostic accuracy [17].

087 Recent advances in machine learning have introduced automated methods that
088 assist clinicians for LUS image interpretation and diagnostic purposes [18, 19], such as
089 identifying horizontal (A-lines) and vertical (B-lines) artifacts or directly classifying
090 diseases like COVID-19 and pneumonia from brightness-mode (B-mode) images [20–22].
091 However, these approaches that directly process and learn from B-mode images face
092

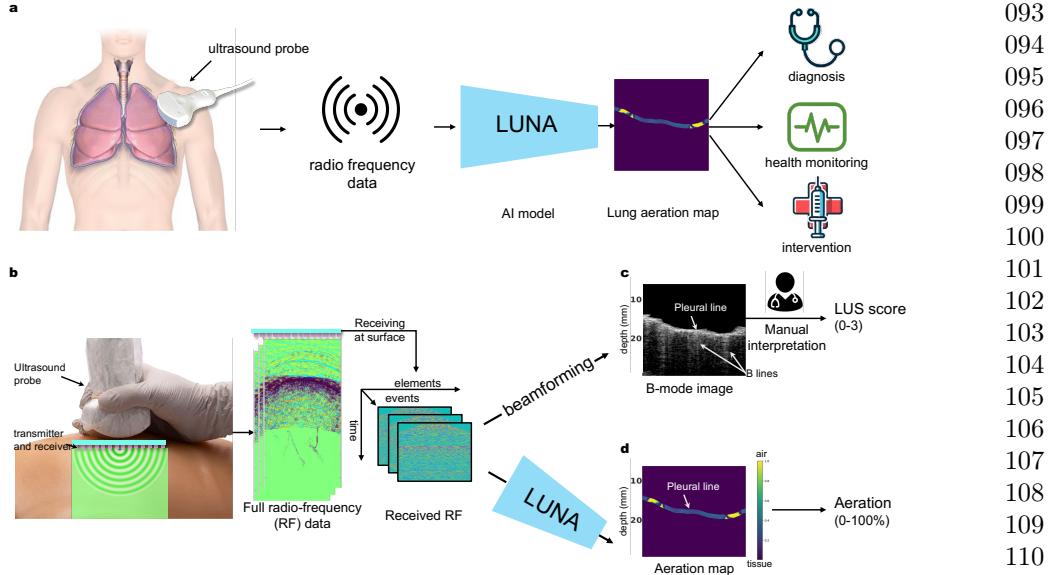


Fig. 1: Overview: LUNA (the Lung Ultrasound Neural operator for Aeration) reconstructs lung aeration maps from ultrasound radio-frequency (RF) data. **a**, Ultrasound devices scan the lung and the RF data is fed to LUNA for reconstructing a lung air-tissue map, or aeration map, which is human-interpretable. The aeration map is used to estimate lung percent aeration, a critical clinical outcome for evaluation, monitoring and diagnostics [1, 2]. **b**, The forward process of ultrasound wave propagation in the lung, whose physics simulation generates the training data for LUNA. **Existing v.s. our approach:** **c**, The current practice, B-mode lung image requires manual interpretation to assign LUS score [3], which has high variations due to the difficulty of identifying artifacts created by the complex wave propagation physics in the lung. LUS score is also coarse, image-level and semi-quantitative. **d**, We reconstruct lung aeration maps that directly depict tissue-air maps, from which clinicians can directly read pixel-level aeration. The method is reproducible and provides quantitative two-dimensional aeration distribution.

several challenges: 1) the information related to lung morphology at mesoscopic/alveolar level is lost in the beamforming process [23]; 2) such models do not generalize well to different ultrasound devices as each device has a setting with different imaging parameters (dynamic range, time gain compensation) and quality of the resulting B-mode images also vary [24].

Our approach: We take a fundamentally different approach to lung ultrasound analysis. Rather than interpreting on B-mode images, either manually or automatically [7, 25, 26], we propose an AI model, LUNA (the Lung Ultrasound Neural operator for Aeration), that directly reconstructs lung aeration maps from delayed ultrasound radiofrequency (RF) data ¹, bypassing the traditional beamforming process. LUNA offers two significant advantages: 1) By directly processing raw RF data, it preserves

¹Delayed RF data is a space-time representation that accounts for the time-of-flight delays of raw wave propagation data in the lung.

139 diagnostic information otherwise lost in beamforming and reduces variability caused by
140 device and parameter differences. 2) By reconstructing lung aeration maps, it eliminates
141 the need for interpreting B-mode artifacts, providing a more direct and quantitative
142 representation of lung disease. Notably, the reconstructed lung aeration map enables
143 the calculation of the lung percent aeration, a critical clinical measurement for diagnosis
144 [1, 2]. 3) The reconstructed lung aeration maps provide a quantitative analysis of
145 lung status, while the existing clinically-adopted LUS score system [3] only provides a
146 coarse and semi-quantitative index for assessing lung aeration. LUNA could establish a
147 quantitative link between ultrasound propagation and the disease state of the lung.

148 LUNA is designed to address two key challenges in reconstructing lung aeration
149 maps from RF data: 1) **The complexity of ultrasound propagation in the**
150 **lung.** The highly reflective tissue-air interface and multiple scattering make the
151 inverse problem ill-posed, necessitating a learning-based approach capable of efficiently
152 capturing subtle changes in RF data. Our LUNA, based on Fourier neural operators
153 [27], learns maps between function spaces and features parameterized directly in
154 Fourier space, capturing subtle RF data variations and enabling efficient extraction of
155 diagnostically relevant features across spatial and temporal scales. 2) **The difficulty**
156 **of collecting real paired data.** Collecting real paired RF-lung aeration map data is
157 time-consuming. We use an experimentally validated simulation approach, Fullwave-
158 2 [28], which solves the full-wave equation to model nonlinear wave propagation,
159 frequency-dependent attenuation and density variations. Its ability to accurately
160 simulate ultrasound propagation in the lung's complex acoustic environment is crucial
161 for understanding the relationship between RF data and lung aeration maps. LUNA is
162 thus trained on the abundant simulated data and fine-tuned on a small number of real
163 data samples.

164 To summarize, the proposed LUNA has following main contributions:

- 165 • **Pioneering ultrasound lung aeration map reconstruction.** This work is the
166 first to reconstruct lung aeration maps from ultrasound RF data, enabling accurate
167 percent aeration estimation and supporting future diagnostic applications.
- 168 • **Effectiveness of training on simulation data and its generalizability**
169 **to real scenarios.** We demonstrate the power of learning wave propagation
170 physics in lung using simulation: Training on abundant simulation data, our model
171 demonstrates satisfactory generalizability on *ex vivo* data.

172 2 Results

173

174 **LUNA, a deep neural operator framework for lung aeration map recon-**
175 **struction.** We implement LUNA (the lung ultrasound neural operator for aeration),
176 a deep learning framework for reconstructing lung aeration map ρ^A from the radio
177 frequency (RF) data/acoustic pressure p received by the lung ultrasound (LUS) trans-
178 ducer (fig. 1a). LUS involves placing the ultrasound probe at the body surface and
179 using the RF data received at the surface for diagnostics and monitoring of lung disease
180 (fig. 1b). The inverse problem of reconstructing lung aeration map ρ^A (obtained from
181 reconstructed lung density image) from RF signal p is challenging due to the complex,
182 high nonlinear wave propagation physics and the ill-posed nature of the problem. LUNA
183

184

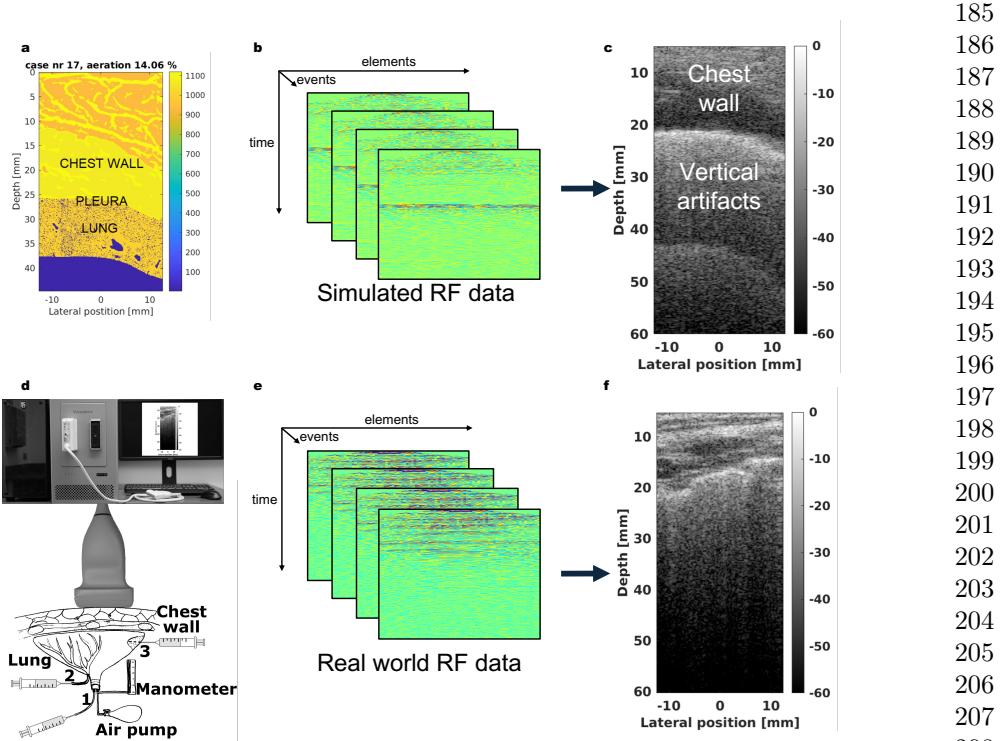


Fig. 2: The process of ultrasound imaging *in silico* (a-c) and *ex vivo* (d-f). LUNA is trained on 10k *in silico* (simulated) data and fine-tuned on 18 *ex vivo* (real) data. **a,** Combined aeration map comprised of the tissue-specifically segmented body wall (top) and underlying lung deformed to conform its internal surface which models realistic pleural interface. **b,** Stack of numerically simulated raw RF data of 128 transmit-receive events visualized as the amplitude of received backscattered signal in form receiver-time. **c,** Corresponding B-mode image formed using simulated RF data demonstrates its proper anatomical part (chest wall thickness and composition) along with multiple coalescent vertical artifacts below the pleural line which are consistent with modeled uniformly distributed fluid retention in lung parenchyma. **d,** Scanning of fresh porcine lungs of known aeration (*displacement method*) through chest wall fragment in the water tank (*ex vivo*) using a programmable ultrasound machine and linear transducer. **e,** Stack of real-world raw RF data of 128 transmit-receive events. **f,** B-mode image formed using real-world RF data demonstrates both anatomical (chest wall tissue, pleural line) and artifactual (multiple coalescent vertical artifacts below the pleural line) parts of the image consistent with modeled significant uniform fluid retention in the lung.

learns to approximate the inverse operator with a non-linear parameterized model f_θ

$$\hat{\rho}^A = f_\theta(p) \quad (1)$$

Comparison to Existing Beamforming Method. In the current practice, the high-dimensional ultrasound RF data is beamformed and compressed into a 2D B-mode

185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230

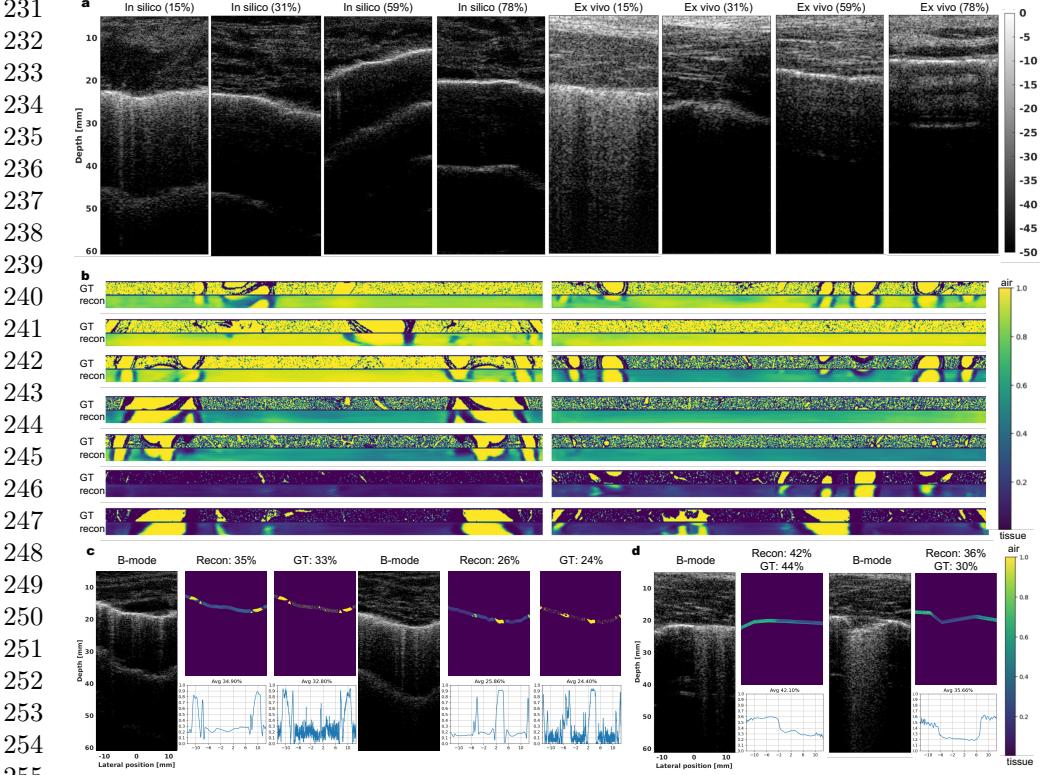


Fig. 3: Visualization of LUS B-mode images and aeration maps, *in silico* and *ex vivo*. **a**, B-mode images corresponding to varying levels of aeration used in *in silico* and *ex vivo* experiments, demonstrating a small domain discrepancy between the two datasets. **b**, Ground-truth and reconstructed aeration maps in *in silico*, showing close visual similarity. **c/d**, Reconstructed aeration maps overlaid on B-mode images for *in silico* (c) and *ex vivo* (d) experiments. The reconstructed aeration maps closely align with the B-mode images, effectively capturing aeration changes. We also provide *in silico* ground-truth aeration maps to which the predictions are similar (c). The lower right plot in each subfigure depicts the reconstructed 1D percent aeration curve.

image (fig. 1c) [7, 29], which is an indirect way to infer the complicated wave propagation physics in the lung and usually requires years of experience for radiologists to effectively interpret such images [8, 9]. In contrast, LUNA’s input is the delayed ultrasound RF data, a space-time representation that accounts for the time-of-flight delays of wave propagation in the lung. LUNA bypasses the beamforming which compresses the RF data and directly estimates the lung aeration map, facilitating easy and interpretation-free clinical uses (fig. 1d).

Neural Operator Framework is Suitable for the Task. LUNA’s architecture is based on the Fourier neural operator (FNO) [27], a deep neural architecture designed to learn maps between function spaces by parameterizing the integral kernel directly in Fourier space. Learning to extract diagnostically relevant features in the image space from

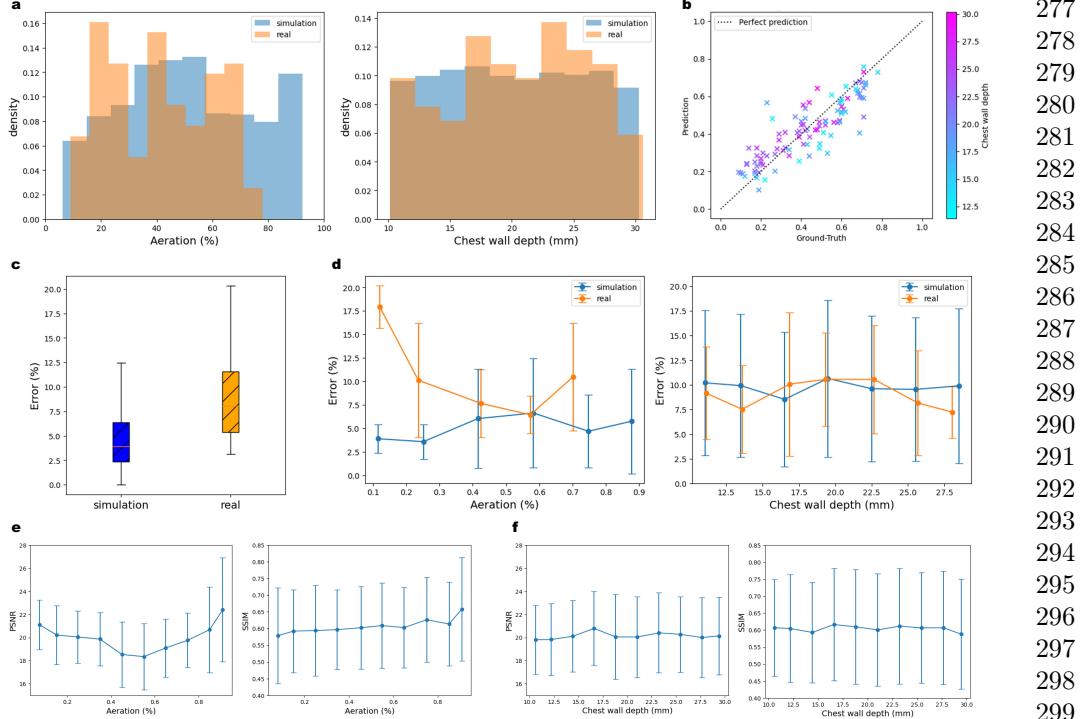


Fig. 4: LUNA performance, *in silico* and *ex vivo*. **a**, The histogram of percent aeration and chest wall depth, two important lung properties, of *in silico* and *ex vivo* data. *In silico* data covers all possible lung properties of *ex vivo* data. **b**, LUNA percent aeration prediction performance of *ex vivo* data. The prediction and ground truth are close to the perfect prediction line, suggesting its capability to estimate lung aeration *ex vivo*. **c**, LUNA percent aeration prediction error of *in silico* and *ex vivo* data. The error gap between them is 4.2%. **d**, LUNA percent aeration prediction error grouped by ground-truth percent aeration and chest wall depth, *in silico* and *ex vivo*. Both sets have consistent errors for different aeration and chest wall depth, indicating the LUNA’s insensitivity to such properties. **e/f**, LUNA aeration map 2D reconstruction performance (PSNR and SSIM) grouped by ground-truth percent aeration (e) and chest wall depth (f) *in silico*. The performance is consistent for samples with different percent aeration and chest wall depth.

the underlying wave propagation physics requires methods that span different time and spatial scales. Ultrasound data measured at the transducer surface is a function of space \times time, but images must represent the correct information in a space \times space by extracting relevant information at different moments in time and transferring them to the correct location in space. The location of this information is not known *a priori*. Resolution agnostic approaches, such as FNO[27], are ideally suited to this task since information is intrinsically learned in Fourier space which does impose specific spatial or temporal constraints on the location of information representation.

Taking ultrasound RF data p as the input, LUNA first predicts the chest wall v.s. lung tissue segmentation. The segmentation is then combined with the RF data

277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322

323 to reconstruct the lung aeration map $\hat{\rho}^A$. Implementation details including the deep
324 learning architecture, training strategies, hyperparameters and post-processing are
325 provided in Section 4.4. The reconstructed lung aeration map is human-interpretable
326 and can enable many downstream clinical outcomes. Specifically, we consider an
327 important clinical sign, the percent lung aeration γ , which can be calculated from the
328 reconstructed lung aeration map $\hat{\rho}^A$.

329 **Two-Stage Learning of LUNA.** As discussed in the introduction, we adopt a two-
330 stage training and LUNA is firstly trained on the abundant simulated data and fine-tuned
331 on a small number of real data samples, due to the cost of obtaining real data.
332 Specifically, The first stage of the two-stage learning is to train LUNA on 10,150
333 simulated LUS RF data/aeration map pairs. In silico lung aeration map reconstruction
334 is validated with the ground truth to ensure satisfactory performance. In the second
335 stage, LUNA is fine-tuned on 18 ex-vivo data samples. The fine-tuned LUNA is then
336 validated on 103 ex-vivo data samples. Both stages involve data augmentation for
337 improved robustness of LUNA. We present more details on the data below.

338 **LUNA is trained on experimentally-validated simulation data of lung ultra-
339 sound propagation.** Training a machine learning surrogate model for the inverse
340 problem of lung ultrasound wave propagation requires paired data of lung aeration
341 maps and their corresponding RF signals, enabling the model to learn the underlying
342 mapping. Obtaining sufficient real paired data *ex-vivo* or *in-vivo* can be time-consuming,
343 we thus train LUNA on the experimentally validated simulated LUS data and fine-tune
344 the model on real *ex-vivo* data. The simulation tool, Fullwave-2 [28], solves the full-
345 wave equation [30]. In addition to modeling the nonlinear propagation of waves, it
346 describes arbitrary frequency-dependent attenuation and variations in density. The
347 unique ability to accurately model ultrasound propagation in the complex acoustic
348 environment provided by the lung and body wall is a key innovation required for recon-
349 structing lung aeration maps from ultrasound RF data. In the following, we discuss
350 how we obtain the simulated data.

351 **Simulated Data Generation.** Maps of acoustical properties were created by com-
352 bining segmented axial anatomical images of the human chest wall (Visible Human
353 Project, resolution 330 μm) with high-resolution histological images of healthy swine
354 lung tissue (5 μm thickness, 0.55 μm resolution), similar to [31] (fig. 2). First, anatomical
355 chest wall maps and microscopic lung structures were merged to produce acoustical
356 maps. Lung aeration, the ratio of air to non-air in the lung, was quantified from binary-
357 segmented histological images, creating aeration maps where 1 represents air and 0
358 represents non-air ($\rho^A : \mathbb{Z}^2 \rightarrow [0, 1]$). The lung histopathology processing steps are
359 illustrated in fig. A1 of the supplementary materials. Next, simulations of diagnostic
360 ultrasound pulses were conducted using a clinically relevant setup (linear transducer,
361 focused transmit sequence at 5.2 MHz). Corresponding RF signals (p) received by
362 the transducer were collected. These simulations were performed with the Fullwave-2
363 acoustic simulation tool [28] (fig. 2a-c), which has been experimentally validated for
364 reverberation, phase aberration and tissue-specific acoustic effects such as attenuation,
365 scattering and absorption [32, 33]. Details are available in Section 4.1.

367 **LUNA’s performance on lung aeration map reconstruction, in silico.** We
368 report numerical results and visualizations of the lung aeration map reconstruction and

percent aeration estimation *in silico*. We also analyze LUNA’s performance grouped by different factors, including ground-truth percent aeration and chest wall depth. 369
370

Evaluation Metrics. The reconstructed lung aeration map by LUNA is a 2D image, which can be compared with the ground-truth lung aeration map with metrics including PSNR (peak signal-to-noise ratio) and SSIM (structural similarity index measure) *in silico*. As an important clinical outcome, lung aeration γ can be obtained by averaging lung aeration map ρ^A : 371
372
373
374
375
376
377
378
379

$$\gamma = \frac{\sum_i^H \sum_j^W \rho^A(i, j)}{HW} \quad (2)$$

where H, W denotes the spatial size of the aeration map. We use the absolute error between the predicted lung percent aeration $\hat{\gamma}$ and the ground-truth percent aeration γ : $|\hat{\gamma} - \gamma|$, to evaluate the performance of the percent aeration estimation performance. 380
381
382

Lung Aeration Map Reconstruction Performance. The average PSNR and SSIM on the evaluation set of *in silico* data are 20.10 ± 3.51 dB and 0.605 ± 0.128 , respectively, demonstrating acceptable performance of LUNA in reconstructing 2D aeration maps. Visual comparisons of flattened reconstructed and ground-truth aeration maps (fig. 3b) and comparisons of B-mode images with reconstructed aeration maps (fig. 3c) qualitatively validate LUNA’s accuracy. Localized alignment between B-mode image features (A-lines and B-lines) and reconstructed aeration maps further supports the reconstruction’s precision. We also report the average and standard deviation of PSNR and SSIM grouped by ground-truth percent aeration and chest wall depth, as in fig. 4e-f, respectively. The lung aeration map reconstruction performance is consistent over different percent aerations and chest wall depths, suggesting the performance of LUNA was insensitive towards lung properties including percent aeration and chest wall depth. 383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414

Percent Aeration Estimation Performance. As a validation of LUNA’s performance in estimating the clinical outcome, percent aeration, fig. 4c depicts the box plot of the percent aeration estimation error *in silico*. The mean and standard deviation of the percent aeration estimation error is 5.2% and 4.6% and the highest percent aeration error is 12.5%. fig. 4d depicts the percent aeration errors grouped by ground-truth percent aeration and chest wall depth, where the *in silico* percent aeration estimation error is consistent with varying aeration and chest wall depth. 395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414

LUNA has strong generalizability to real ex-vivo data. After training on simulated data, we fine-tune LUNA on a small number of real data (the two-stage training mentioned before) and show its generalizability on the real data. In the following, we discuss the real ex-vivo data acquisition and LUNA’s performance on the real data. 402
403
404
405
406
407
408
409
410
411
412
413
414

Real Ex-Vivo Data. Fresh tissues of 12 swine were hand-held scanned in a water tank using linear transducers and a research ultrasound system, with the same setting as the simulation data. Units obtained out of the right lung and caudal lobe of the left lung were employed in the calculation of its ground-truth percent aeration, while the left cranial lobe was used for scanning (fig. 2d-f). The ground-truth percent aeration can be compared with the predicted one from LUNA, although the ground-truth lung aeration map is unavailable due to the implausibility of obtaining it in reality. The depth of the 407
408
409
410
411
412
413
414

415
416 **Table 1:** Lung ultrasound data used in the study.
417

	Simulation Set				Ex-vivo Set		
	Training	Validation	Evaluation	Sum	Fine-Tuning	Evaluation	Sum
Number of data samples	10,150	1,450	2,900	14,500	18	103	121
Chest wall depth (mm), average \pm SD	20.1 \pm 5.8	20.2 \pm 5.7	20.3 \pm 5.7	/	21.5 \pm 5.6	21.1 \pm 5.6	/
Aeration (%), average \pm SD	59.8 \pm 28.3	53.8 \pm 24.1	60.1 \pm 25.0	/	33.4	43.2	/

424 reconstructed lung aeration map is set to be 2.6 ultrasound wavelength, considering
425 the ultrasound penetration abilities. More details are provided in Section 4.2.

426 The data used in the study is summarized in Table 1. The simulated and real data
427 have a relatively low domain gap in terms of B-mode image similarity (fig. 3a), percent
428 aeration and chest wall depth distribution (fig. 4a).

429 Results. We report the evaluation results of LUNA on ex vivo swine lungs. fig. 4b
430 depicts the predicted and the ground-truth percent aeration. LUNA achieves a lower than
431 10% average percent aeration estimation error on the ex-vivo data. The data samples
432 gather around the perfect prediction line, indicating the robust ex vivo performance of
433 LUNA.

434 Percent Aeration Estimation Performance. fig. 4c depicts the box plot of the per-
435 cent aeration estimation error ex vivo. The mean and standard deviation of the percent
436 aeration estimation error is 9.4% and 5.4%, with the highest error over all cases being
437 20.3%. The ex vivo percent aeration estimation performance is just slightly higher
438 (4.2%) than the in silico performance, indicating the good generalizability of LUNA in
439 real data. fig. 4d depicts the ex vivo percent aeration estimation error is consistent
440 with varying aeration and chest wall depth, where we see LUNA is insensitive to the
441 lung properties like aeration and chest wall depth, ex vivo. Additionally, the error gap
442 between in silico and ex vivo data is also small.

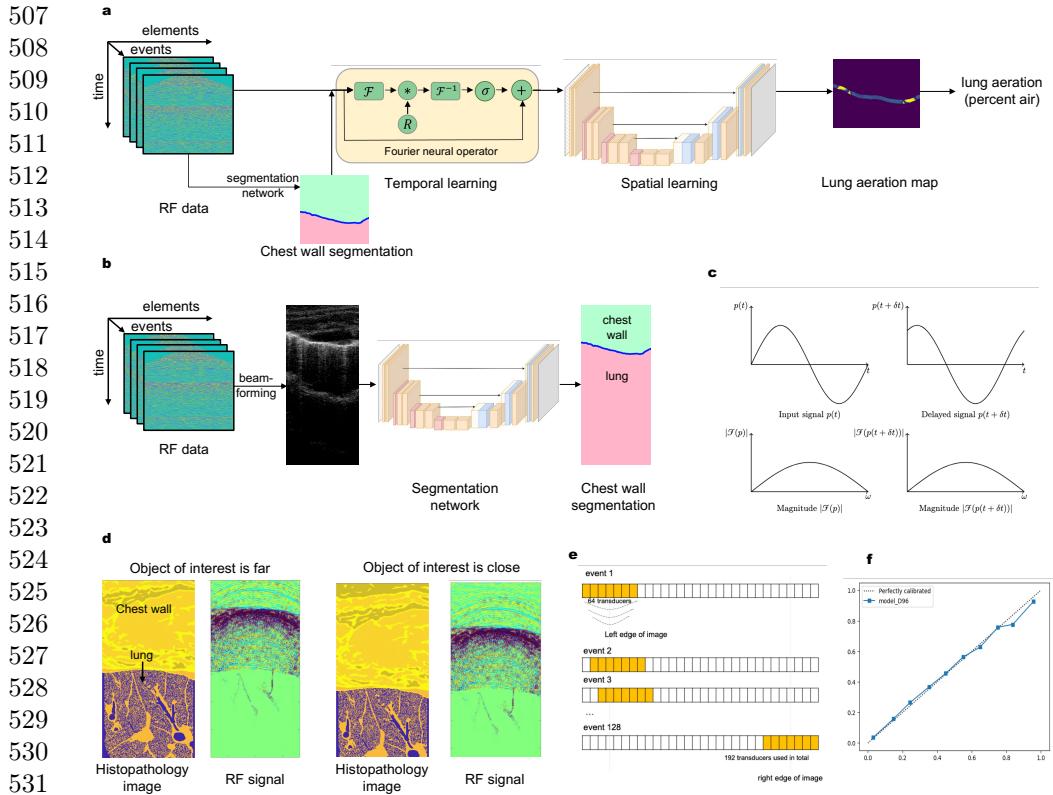
443 Visualization. Although the ground-truth lung aeration maps *ex-vivo* are unavailable,
444 we provide visual comparisons of B-mode images and reconstructed lung aeration
445 maps in fig. 3d—the localized alignment between B-mode image features (A-lines and
446 B-lines) and reconstructed aeration map validates the accuracy of the aeration map
447 reconstruction.

448 Runtime. The processing time for each case is 0.11 seconds on an NVIDIA 4090
449 GPU, which means the framework can achieve an ideal aeration map display rate at 9
450 Hz.

452 3 Discussion

454 In this work, we introduce LUNA, a machine learning framework that, for the first time,
455 reconstructs lung aeration maps directly from lung ultrasound (LUS) data. Unlike
456 traditional approaches that rely on human interpretation of B-mode images, LUNA
457 bypasses this step and addresses challenges such as the expertise required to effectively
458 interpret the complex physics of wave propagation in the lung and inconsistencies
459 introduced by varying beamforming settings across ultrasound machines. Additionally,
460

LUNA improves the efficiency of lung disease screening, which can potentially improve clinical interpretability, reproducibility and diagnostic capabilities.	461
	462
LUNA also provides a robust quantitative evaluation of 2D lung aeration distribution, addressing limitations of current clinical practices. Currently, clinicians rely on semi-quantitative LUS scoring systems ranging from 0 to 3 [3], which are prone to high individual variability, lack repeatability and offer only one coarse estimation score of lung conditions per scan. In contrast, LUNA's reconstructed aeration maps provide pixel-level 2D lung aeration information derived directly from raw ultrasound signals per scan—data that is currently inaccessible to clinical interpretation. The advancement lays the groundwork for a more precise and reproducible method of linking ultrasound imaging to lung pathology, enabling improved diagnostic accuracy and clinical decision-making.	463
	464
	465
	466
	467
	468
	469
	470
	471
	472
Our approach differs significantly from existing works that analyze LUS B-mode images to extract features like A-lines and B-lines [34, 35] or classify diseases such as COVID-19 and pneumonia [20–22]. Instead, LUNA operates on delayed RF data, preserving frequency domain information and minimizing variations introduced by imaging settings like dynamic range and time gain compensation. This distinction allows for more consistent and precise analyses across diverse clinical setups.	473
	474
	475
	476
	477
	478
A key technical strength of LUNA is its ability to generalize from simulated data to real LUS data, a notable achievement given the challenges in lung imaging and the lack of paired real-world RF and aeration map data. Using a full-wave acoustic pressure field simulator, Fullwave-2 [28], we trained LUNA on simulated lung aeration maps and verified its robust performance in clinically relevant setups, including ex vivo swine lung experiments.	479
	480
	481
	482
	483
	484
While Fullwave-2 [28] effectively models ultrasound propagation in the human body, the feasibility of a learning-free iterative solver for the inverse problem remains questionable. Such solvers have not been explored, likely due to two main challenges: 1) Ill-posedness and instability. RF data is captured only at the transducer surface, providing partial wave information. This makes it difficult to impose the full-wave equation as a constraint, resulting in unstable optimization. 2) Unrealistic computational demands. Unlike geophysical applications [36, 37], the lung's spatial complexity, air-filled structures and phenomena such as reverberation, multiple scattering and nonlinearity render optimization-based solvers computationally infeasible. LUNA addresses these challenges by using machine learning to directly map ultrasound measurements to lung aeration maps, providing an efficient and practical surrogate for this complex inverse problem.	485
	486
	487
	488
	489
	490
	491
	492
	493
	494
	495
LUNA has certain limitations. 1) The lack of aligned ex-vivo lung aeration maps prevents full verification of our 2D reconstructions in such setups. Furthermore, the current model has yet to be validated in vivo with human subjects. 2) LUS is inherently real-time and interactive, relying on operator adjustments during imaging. LUNA currently has a runtime of 0.1 seconds, which should be further accelerated for real-time use. 3) LUNA has not been linked to final diagnosis, such as ARDS and cardiogenic pulmonary edema (CPE). Addressing these limitations will allow LUNA to further advance the interpretability, reproducibility and diagnostic utility of LUS for acute and chronic lung diseases, paving the way for broader clinical adoption.	496
	497
	498
	499
	500
	501
	502
	503
	504
	505
	506



4 Methods	553
4.1 Ultrasound Data Simulation	554
	555

In this section, we discuss the forward problem of lung ultrasound (LUS): for a specific lung (with acoustic properties such as speed of sound c_0 and density ρ), we model the ultrasound propagation and simulate the corresponding acoustic pressure/RF data p received at the body surface. The paired lung acoustic properties and received RF data can be used for training LUNA as a surrogate to solve the inverse problem.

Numerical Simulation of Radio Frequency Data. For generation of channel data closely matching RF signals received in human body ultrasound scanning, acoustic pressure field simulator Fullwave-2 [28] was used. This numerical tool is based on the first principles of sound wave propagation in heterogeneous attenuating medium and accounts for its effects such as distributed aberration (wavefront distortion), reverberation (multiple reflections), multiple scattering and refraction (change of wavefront direction at media interface). Fullwave-2 was successfully validated in various diagnostic and therapeutic scenarios, such as *ex vivo* abdominal measurements [38, 39], water tank experiments [32], *in vivo* human abdominal measurements [40], transcranial brain therapy [33, 41, 42] and traumatic brain injury modeling [43].

The Fullwave2 model employs a nonlinear full-wave equation defined by the following equation:

$$\nabla_1 p + \rho \frac{\partial \mathbf{v}}{\partial t} = 0 \quad (3)$$

$$\nabla_2 \cdot \mathbf{v} + \kappa \frac{\partial p}{\partial t} = 0 \quad (4)$$

where $p(\mathbf{x}, t)$ and $\mathbf{v}(\mathbf{x}, t)$ represent the pressure and velocity wavefield at a given position \mathbf{x} at a given time t respectively, and $\rho(\mathbf{x})$ and $\kappa(\mathbf{x})$ denote the density and the compressibility of the medium at position \mathbf{x} , respectively. When the nonlinear effects are included, the density ρ and the compressibility κ are modified by the pressure p and the nonlinearity coefficient $\beta = 1 + B/A$ as follows:

$$\rho = \rho_0 [1 + \kappa_0 p] \quad (5)$$

$$\kappa = \kappa_0 [1 + \kappa_0 (1 - 2\beta)p] \quad (6)$$

where ρ_0 is the equilibrium density and κ_0 is the equilibrium compressibility. Now, ∇_1 and ∇_2 in equations (3) and (4) are used to denote the complex spatial derivatives that model attenuation and dispersion while maintaining the pressure-velocity formulation of the wave equation. The complex spatial derivatives ∇_1, ∇_2 are written as a scaling of the partial derivative and a sum of convolutions with N relaxation functions, which for ∇_1 can be written as

$$\frac{\partial}{\partial \tilde{x}_1} = \frac{1}{\kappa_{x_1}} \frac{\partial}{\partial x} + \sum_{\nu=1}^N \zeta_{x_1}^\nu(t) * \frac{\partial}{\partial x} \quad (7)$$

599 $\frac{\partial}{\partial \tilde{y}_1} = \frac{1}{\kappa_{y_1}} \frac{\partial}{\partial y} + \sum_{\nu=1}^N \zeta_{y_1}^\nu(t) * \frac{\partial}{\partial y}$ (8)
 600
 601

602 $\frac{\partial}{\partial \tilde{z}_1} = \frac{1}{\kappa_{z_1}} \frac{\partial}{\partial z} + \sum_{\nu=1}^N \zeta_{z_1}^\nu(t) * \frac{\partial}{\partial z}$ (9)
 603
 604

605 $\zeta_{x_1}^\nu$ is the convolution kernel for the N relaxations, indexed by ν . The convolution
 606 kernels are defined as:

607
 608 $\zeta_{x_1}^\nu(t) = -\frac{d_{x_1}^\nu}{\kappa_{x_1}^2} e^{-\left(\frac{d_{x_1}^\nu}{\kappa_{x_1}} + \alpha_{x_1}^\nu\right)t} H(t)$ (10)
 609
 610

611 $\zeta_{y_1}^\nu(t) = -\frac{d_{y_1}^\nu}{\kappa_{y_1}^2} e^{-\left(\frac{d_{y_1}^\nu}{\kappa_{y_1}} + \alpha_{y_1}^\nu\right)t} H(t)$ (11)
 612
 613

614 $\zeta_{z_1}^\nu(t) = -\frac{d_{z_1}^\nu}{\kappa_{z_1}^2} e^{-\left(\frac{d_{z_1}^\nu}{\kappa_{z_1}} + \alpha_{z_1}^\nu\right)t} H(t)$ (12)
 615

616 Note that $\kappa_{x_1}(\mathbf{x})$, $\kappa_{y_1}(\mathbf{x})$, $\kappa_{z_1}(\mathbf{x})$ represent a linear scaling of the derivative at position
 617 \mathbf{x} . This scaling parameter modifies the wave velocity in the x , y and z directions. The
 618 variables $d_{x_1}^\nu$, $d_{y_1}^\nu$, $d_{z_1}^\nu$, represent a scaling-dependent damping profile; and $\alpha_{x_1}^\nu$, $\alpha_{y_1}^\nu$,
 619 $\alpha_{z_1}^\nu$ denote a scaling-independent damping profile. $H(t)$ is the Heaviside or unit step
 620 function. The transformation set for the ∇_2 operator is identical to that of ∇_1 and
 621 the variables associated with this second transformation are denoted by the subscript
 622 2. These relaxation mechanisms incorporated in ∇_1 , ∇_2 are introduced to empirically
 623 model attenuation based on observations of the attenuation laws and parameters
 624 observed in soft tissue. These mechanisms can be generalized to arbitrary attenuation
 625 laws through a process of fitting the relaxation constants.

626 Numerically, Fullwave2 uses the finite-difference time-domain (FDTD) method
 627 to solve the nonlinear full-wave equations (3) and (4). In order to ensure numerical
 628 stability and accuracy in heterogeneous media with high contrast, Fullwave2 utilizes the
 629 staggered-grid finite difference (FD) discretization, whose FD operator has $2M$ -th order
 630 accuracy in space and fourth-order accuracy in time. [28] provides a detailed numerical
 631 solution for the computation of the spatial derivatives ∇_1 , ∇_2 and relaxation functions.
 632

633 In the inverse problem setting, the RF data p_S is only received at surface $p_S = Sp$,
 634 where S is a linear operator. For simplicity, we refer to the received RF data as p in the
 635 paper. Density map ρ contains two parts: chest wall and lung. Lung aeration map ρ^A
 636 is only the lung segmentation from ρ , which is flattened and thresholded to a binary
 637 map, with 1 denoting air pixels and 0 denoting non-air (tissue) pixels (details in the
 638 later part of this subsection).

639 Key acoustical properties of the air, such as almost total reflectivity and impermeability
 640 due to high impedance mismatch with soft tissue, were modeled via constant
 641 zero pressure in air inclusions similar to [31, 44]. This approach is beneficial in terms of
 642 computational cost and simulation stability compared to pulse propagation in medium
 643 of low sound speed and mass density.

644

A clinically relevant scenario of use of linear transducer L12-5 50 mm (ATL, Bothell, Washington, USA) and 128 sequential focused transmit (2-cycle 2.5 MPa pulses, walking 64 element aperture) at center frequency 5.2 MHz was simulated [12, 45]. This transducer is compatible with research ultrasound systems, such as Vantage 256 (Verasonics Inc, Kirkland, Washington, USA) and a list of diagnostic ultrasound imaging systems Philips. The focal depth varied from 1 to 3 cm and was set to the lung surface position (pleura depth) to maintain the maximum energy deposition at the acoustical channels entrance and for optimal visualization of diagnostic features - vertical artifacts [46, 47]. Accordingly, transmit f-number varied from 0.8 to 2.4. The RF signals were received at locations of the same fired 64 transducer elements (subaperture) and sampled at a rate of 20.8 MHz with 70% fractional bandwidth which is comparable to modern clinical ultrasound imaging systems. All the simulations were performed in 2D space, each independent transmit-receive event in a field of 2.5 cm width and 4.5 cm depth with 0.195 mm lateral translation between neighbor subapertures/events. The duration of the simulations was limited to 87.6 μ s for compliance with existing clinical recommendations [48], to be able to visualize at least the second reflection of the ballistic pulse between surfaces of transducers and the lung-horizontal artifact in healthy cases [13]. A spatial grid of 12 points per wavelength (24.6 μ m) and temporal discretization was set to 8.0 ns accounting for a reference speed sound of 1540 m/s, which corresponded to a Courant-Friedrichs-Lowy condition of 0.5 [49].

Acoustical Properties of the Human Body Wall and the Lung. Maps of acoustical properties were combined out of segmented axial anatomical images of human body wall (Visible Human Project [50], resolution 330 μ m) and histological images (hematoxylin and eosin staining, section thickness 5 μ m) of healthy swine lung, resolution 0.55 μ m (c2a-c), similar to [31]. Segmented slides from these sources were interpolated (nearest neighbor) to fit the simulation elements size (24.6 μ m). Fresh tissue for histological processing and *ex vivo* ultrasound scanning was provided by Tissue Sharing Program of North Carolina State University and processed by Pathology Services Core, University of North Carolina at Chapel Hill. Slide scanner SlideViewTM VS200 (Olympus Corporation, Tokyo, Japan) was used for bright-field microscopy at 10x magnification. Photographic images of human body cryosections were segmented into three types of tissue (connective, adipose and muscle) using custom tissue-specific probability distribution functions in 3 color channels (RGB). Parietal and visceral pleura borders were segmented manually in the original cryosection and histological images. It allowed axial column translation for flattening of the lung and its deformation to conform with various parietal pleura curvature. Histological lung images were binary segmented (air / non-air) based on the arbitrarily chosen brightness threshold of 0.8 in the green channel.

A wide variety of alveolar size (median 94 [IQR 72-132] μ m) and alveolar wall thickness (median 16.5 [IQR 5.5-38.5] μ m) was assured by use of lung tissue from 10 animals of different age (median 3.5 [IQR 3-6] months) and weight (median 78 [IQR 60-185] kg). While the lateral position of vertical artifacts in B-mode images correlates with the location of acoustical channels [31, 51], transfer of the superficial layer (200 μ m) of the histological part of the map allowed significant increase variability of the observed artifacts and underlying RF data (fig. 2). Leveled and cropped rectangular

691 lung tissue images of size 1.8x1.2 cm were randomly selected from the 521 unit dataset
 692 and tailed to obtain a continuous 5 cm wide lung layer for each simulated image. Both
 693 flattened (X cases) and naturally curved (Y cases) lung layers were employed (fig.). To
 694 cover a range of diagnostic ultrasound image features (horizontal and vertical artifacts),
 695 values of aeration and depth of the lung parenchyma were randomly selected out of
 696 uniform distributions in the range [10 90] % and [1 3] cm, respectively. Lung aeration
 697 was calculated as a percent of air elements out of all lung parenchymal elements in
 698 2D. Alveolar size and alveolar wall thickness were calculated as linear intercepts based
 699 on standard method [52] using custom Matlab (Mathworks, Natick, MA, USA) code.
 700 Uniform alveolar derecruitment characteristic for Acute Respiratory Distress Syndrome
 701 (ARDS) at the air-alveoli interface was modeled numerically. First, the subpopulation
 702 of air pixels A neighboring with non-air ones was found. Second, the number of air
 703 pixels n necessary to convert into non-air to achieve target aeration was calculated.
 704 Finally, n random pixels were drawn out of subpopulation A and converted into non-
 705 air. In case if n was more than the number of pixels in A , all the subpopulation was
 706 converted and the algorithm was repeated while the target aeration is achieved. If target
 707 aeration was higher than the initial one, the same sequence was performed finding
 708 non-air pixels in step 1 and converting to air in step 3. These alterations facilitated
 709 practically continuous variable aeration.

710 **Beamforming.** We introduce the beamforming used in the paper as a baseline
 711 to transform RF data into conventional human-readable B-mode images. Using a
 712 conventional delay-and-sum (DAS) beamforming algorithm, a dataset of delayed RF
 713 signals and corresponding reconstructed 2.5 cm wide B-mode images were composed
 714 out of both numerically simulated and scanned *ex vivo* data. Each tensor in the form
 715 [time, receiver, transmit-receive event] represented a single scan/image consisting of
 716 signals from 64 receivers in 128 independent events. Individual receivers' vector signals
 717 s_m were delayed by precalculated τ_m samples to ensure proper focusing, compensate
 718 differences in their (receivers) spatial/lateral position and synchronize these readings
 719 [53] assuming a homogeneous speed of sound 1540 m/s. The beamformed signal of a
 720 single transmit-receive event can be defined in the discrete-time domain as:
 721

$$722 \quad 723 \quad bfs(t) = \sum_{m=1}^M s_m(t - \tau_m) \quad (13)$$

724

725 Where $bfs(t)$ is output signal, m - number of the receiver (1..64), t - discrete time.
 726 DAS has multiple advantages, such as i) grounding on the basic principles of
 727 wave propagation; ii) simple implementation; iii) low computational cost and possible
 728 parallelization, which makes it applicable in real-time; and iv) statistics of real envelopes
 729 [54] and temporal coherence [55] are preserved. Numerical robustness and data-
 730 independency make this beamforming technique highly generalizable and the most
 731 widely used not only in ultrasound imaging, but also in telecommunication.

732 After DAS beamforming, RF signals were envelope-detected using Hilbert transform
 733 and log compressed to a dynamic range appropriate for human visualization [-60 0]
 734 dB. Such processing is conventional and allows to emphasize on weak scattering along
 735 with high-intensity reflections using the same scale [56].
 736

For smoother visual representation, isotropic 2D interpolation (bicubic, factor 4) was applied to B-mode images. To preserve the generalizability of acquired RF data, TGC was not applied to them [57]. The image processing was performed in Matlab (Mathworks, Natick, MA, USA). The numerical simulations were done on a Linux-based computer cluster running GPU NVIDIA® V-100. Individual simulations (128 per B-mode image) were running in parallel and took up to 6 hours.

4.2 Ex Vivo Data

Ex-Vivo Data Acquisition. Fresh tissues of 12 animals (median age 3.5 [IQR 1.75-6] months), median weight 85 [IQR 50-180] kg) were hand-held scanned in a water tank using linear transducer L12-5 50 mm (ATL, Bothell, Washington, USA) and research ultrasound system Vantage 256 (Verasonics Inc, Kirkland, Washington, USA). The sequence, time discretization and bandwidth parameters were identical to those described in simulations. The focal depth varied from X to Y cm and was manually set by the operator corresponding to the visualized pleural depth. Time gain compensation (TGC) was applied and individually adjusted for proper visualization of relevant diagnostic features in the scanning process, however RF data were saved and analyzed without compensation for generalizability reasons. 124 scans were performed (118 for assessment of network performance and 6 for calibration of the simulations). To reproduce intercostal views of transthoracic lung ultrasound imaging, two tissue layers were placed in the water tank - body wall (chest or abdominal) at the top and lung on the bottom. The former was completely submerged in degassed and deionized water and fixed/immobilized to the water tank walls with a custom plastic fixture and nylon strings sutured through the tissue (outside the field of view). The latter was placed costal surface toward the body wall and transducer in anatomical position and free-floating. The bronchus was intubated with custom plastic fitting hermetically connected to i) manual rubber bulb air pump with bleed valve, ii) airway manometer (leveled water column pressure gauge) and iii) line with 3-way stopcock and syringe port outside the water tank for fluid instillation in airways. Pulmonary arteries were ligated and sutured twice and veins were sutured twice at harvesting to prevent aeration of lung vasculature. Lobes of left swine lungs were scanned due to their closest gross anatomical similarity to human ones [58]. Both lungs were dissected into lobes (two in the left, three in the right lung) and segments when technically possible using intersegmental veins and inflation-deflation lines as guides [59]. The dissected lobes/segments were tested for aerostasis (hermeticity) and units with air leaks were excluded from further study (success rate 48%).

Aeration Calculation. Units obtained out of the right lung and caudal lobe of the left lung were employed in the calculation of its bulk aeration, while the left cranial lobe was used for scanning. Such split was necessary because of the destructive nature of degassing after which lung tissue is altered and does not represent normal anatomy at micro- and mesoscopic levels [60]. The lung lobe was short-term inflated to an airway pressure of 20 cm H₂O for alveolar recruitment and static continuous positive airway pressure of 10 cm H₂O [61, 62] was maintained during volume measurements and scanning. Weighing was performed with the bronchus cross-clamped and the airway

fitting disconnected. Bulk lung aeration was calculated as $A_l = (V_l - V_t)/V_l$, where V_l is the total volume of the lobe measured using the fluid displacement method and V_t - volume of the lobe after de-aeration. Lung tissue was degassed three times [63] for 10 minutes in a vacuum chamber at pressure -25 mm Hg [64]. In addition to aeration, lung mass density was calculated as $\rho_l = m_l/V_l$, where m_l - total weight of the lobe. To model ARDS-like distributed deaeration airway/bronchial instillation of isotonic NaCl solution was employed. The severity of lesion was leveraged by adding different volumes of the fluid.

Summary: Data used in the Study. Our machine learning (ML) framework LUNA is mainly trained on simulated data and evaluated on real ex-vivo data. We summarize the data used in the paper in Table 1.

4.3 LUNA Architecture

Overview. LUNA reconstructs lung aeration map ρ^A from the measured RF data p . The overall LUNA pipeline is in fig. 5a. The RF signal is three-dimensional, denoted as $p \in \mathbb{R}^{T \times N_t \times N_e}$, where T, N_t, N_e refer to the number of temporal steps, the number of transducer elements and the number of events, the distinct ultrasound pulse emissions in a scan. Before the reconstruction of ρ^A , we use a segmentation model \mathcal{S} to estimate the chest wall segmentation map $\mathcal{S}(p)$ as an auxiliary task to inform the reconstruction model of the chest wall structure. We then combine the predicted chest wall segmentation map with RF data and predict the lung aeration map $\hat{\rho}^A$ with a neural operator f_θ :

$$\hat{\rho}^A = f_\theta(p, \mathcal{S}(p)) \quad (14)$$

where θ are the parameters of the neural operator. To train the neural operator, we use a loss \mathcal{L} to penalize the difference between predicted lung aeration map $\hat{\rho}^A$ and the ground truth ρ^A , i.e. $\min_\theta \mathcal{L}(\hat{\rho}^A, \rho^A)$. Next, we provide more details of each module of LUNA, followed by the simulation to real domain adaption and the model calibration. Ablation study that demonstrates the empirical contribution of different modules and design choices are in the Section C of the Supplementary.

Chest Wall Segmentation. Considering the spatial correspondence between the B-mode image and the chest wall segmentation map (fig. 5b), we first beamform the LUS RF data p to B-mode images $\mathcal{B}(p)$ and then feed the B-mode images to the segmentation model (see beamforming details in Section 4.1). During training, the B-mode images $\mathcal{B}(p)$ and ground-truth lung chest wall segmentation map S are resized to 400×400 for the efficient training of the network. The segmentation network adopts a UNet [65] architecture, with 4 downsampling layers with stride 2. The downsampling path reduces the resized B-mode images $\mathcal{B}(p)$ down to 25×25 at the bottleneck. The downsampling path also increases the feature channel from 1 to 64, 128, 256, 512 sequentially. The upsampling path then reconstructs the spatial dimensions, combining features from the corresponding downsampling layers to preserve spatial information. Finally, an output convolutional layer produces the segmentation map $\mathcal{S}(p)$ of size $400 \times 400 \times 2$, where the first channel corresponds to the chest wall region and the

second one to the lung region. To train the network, we penalize the difference between the prediction and the ground truth with a cross-entropy loss $\mathcal{L}_{\text{CE-seg}}(\mathcal{S}(p), S)$.	829
	830
On the in silico test set, our model achieves 95.8% under Dice similarity coefficient, a well-accepted metric for evaluating image segmentation performance [66]. Additional results of the auxiliary task, chest wall segmentation, can be found in the Supplementary (Section B).	831
	832
	833
	834
	835
	836
	837
	838
	839
	840
	841
	842
	843
	844
	845
	846
	847
	848
	849
	850
	851
	852
	853
	854
	855
	856
	857
	858
	859
	860
	861
	862
	863
	864
	865
	866
	867
	868
	869
	870
	871
	872
	873
	874
Aeration Map Reconstruction. The measured LUS RF data p is combined with chest wall segmentation $\mathcal{S}(p)$ to reconstruct the lung aeration map ρ^A . The reconstruction network is designed as a temporal neural operator which enables invariance learning to the temporal delay (fig. 5c), followed by a spatial network to capture the local features for the aeration map reconstruction. To improve the robustness of the model, we adopt a data augmentation strategy involving both depth and lateral masking (fig. 5d,e).	835
<i>Background: Fourier Neural Operator.</i> FNO (Fourier neural operator) [27] is a powerful neural operator framework that efficiently learns mappings in function spaces, with many applications as surrogate models for solving partial differential equations (PDEs) with many applications [67–69]. In our approach, we decompose 3D ultrasound RF data into temporal and transducer dimensions (element \times event) to achieve computational efficiency by handling the high-dimensional data in a structured way. Applying FNO specifically to the temporal dimension leverages FNO’s strengths in capturing time-dependent patterns, enabling efficient and precise modeling of the data while preserving computational speed and accuracy.	836
<i>Fourier Neural Operator on Temporal Dimension for Invariance to Temporal Delay.</i> Our input data p ’s first dimension is a temporal signal and the output is invariant to delay, i.e., $p(t)$ and $p(t + \delta t)$ are mapped to the same output. In our network, we propose to use the Fourier transform to get the Fourier feature $\mathcal{F}(p)$. Specifically, p has a size of $T \times N_t \times N_e$, where T, N_e, N_t represent the temporal dimension, the number of transducer elements and the the number of events. We perform a Fourier transform on the temporal dimension T only, treating the other dimensions N_t and N_e as batches. The Fourier transform of p along the T dimension can be denoted as $p(f, n_e, n_t) = \mathcal{F}\{p(t, n_t, n_e)\}$, where \mathcal{F} denotes the Fourier transform on the temporal domain, f is the frequency domain corresponding to the temporal dimension. T, n_e is the index in the N_e dimension and n_t is the index in the N_t dimension. The Fourier transformed data P can be represented in terms of its magnitude and phase as follows:	837
$ p(f, n_e, n_t) = \text{Magnitude of } p(f, n_e, n_t)$	838
$\angle p(f, n_e, n_t) = \text{Phase of } p(f, n_e, n_t)$	839
Features are then learned from both the magnitude and phase of the Fourier-transformed data. Because the magnitude $ \mathcal{F}(p) $ is linear time-invariant(invariant to delay), it can learn the delay-invariance mapping effectively.	840
The FNO is implemented with a one-dimensional process on the temporal dimension (by treating other dimensions as different samples of a batch). The FNO module consists of 2 layers, each of which has 32 output channels and 87 modes in the Fourier domain.	841

875 *Spatial Network on Transducer Element Dimensions for Lateral Reconstruction.*
 876 We use a spatial network consisting of a series of convolutions after the FNO module
 877 for processing the RF data's transducer element dimensions (element $n_t \times$ event n_e).
 878 The spatial network also allows learning local features, which is important for image
 879 reconstruction and augments FNO's global feature learning, as FNO's Fourier space
 880 feature learning can be considered as applying global convolution to the RF signal.
 881 The network consists of a series of downsampling and upsampling layers, with a similar
 882 design to [65]. The downsampling path includes layers with channel transitions of
 883 64 to 64, 64 to 128, 128 to 256 and 256 to 256 (due to bilinear interpolation). The
 884 upsampling path includes layers with channel transitions of 512 to 256, 256 to 128,
 885 128 to 64, 96 to 32 and a final transposed convolutional layer from 32 to 32, followed
 886 by a batch normalization layer for 32 channels.

887 *Masking-Based Data Augmentation.* To increase the robustness of LUNA and to
 888 make the aeration map ρ^A reconstruction invariant to partial masking m of the RF
 889 signal p , we adopt local-masking-based data augmentation strategies during training.
 890 That is, the masked input $m(p)$ and unmasked input p should map to the same
 891 reconstruction. During training, we consider two different masking strategies:

- 892 1. **Temporal Masking in RF for depth dimension of the reconstruction.**
 893 The goal is to strengthen the invariance to temporal delay and chest wall depth
 894 from FNO's design via additional data augmentation. The temporal masking is
 895 adopted to enable other layers to also be invariant. Specifically, we uniformly
 896 masked the initial M_t steps ($M_t \leq 200$, or 13% of the total temporal length). The
 897 reason is that the initial temporal steps correspond to the reflections in the chest
 898 wall, to which we want to make the reconstruction invariant.
- 899 2. **Transducer Element Masking in RF for the lateral dimension of the**
 900 **reconstruction.** To encourage lateral correspondence, we randomly mask M_s
 901 elements of the flattened transducer element dimension of size $N_e \cdot N_t$. $M_s \leq 2000$,
 902 or 24% of the total number of equivalent transducer size.

903 **Loss Functions.** We optimize LUNA by minimizing the training loss, which consists
 904 of the 2D aeration map reconstruction loss and aeration prediction loss. The 2D aeration
 905 map reconstruction loss \mathcal{L}_{CE} is a cross-entropy loss that measures the difference between
 906 the predicted and ground-truth aeration maps:
 907

$$908 \quad 909 \quad \mathcal{L}_{CE} = - \sum_{i,j} [\rho_{ij} \log \hat{\rho}_{ij} + (1 - \rho_{ij}) \log(1 - \hat{\rho}_{ij})] \quad (17) \\ 910$$

911 where ρ_{ij} is the $i^{\text{th}}, j^{\text{th}}$ pixel of ground-truth aeration map ρ^A and $\hat{\rho}_{ij}$ is the $i^{\text{th}}, j^{\text{th}}$
 912 pixel of reconstructed aeration map $\hat{\rho}^A$
 913

914 The aeration prediction loss is a ℓ_1 loss that penalizes the difference between the
 915 predicted aeration $\hat{\gamma}$ and ground-truth γ .

$$916 \quad 917 \quad \mathcal{L}_\gamma = \|\gamma - \hat{\gamma}\| \quad (18) \\ 918 \\ 919 \\ 920$$

The total loss is a combination of 2D aeration map reconstruction loss and aeration prediction loss: 921
922

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \eta \mathcal{L}_{\gamma} \quad (19) \quad 923$$

where η is the weight of the aeration loss, which we set to be 0.5 in the experiments. 924
925

Simulation to Real Data Adaption. White Gaussian noise is added to simulated RF data during training on the simulated data to reduce the gap between simulated and real data. After finishing training the model with simulated data, we fine-tune with 18 real data samples, with details available in Table 1. During fine-tuning, we only use aeration loss \mathcal{L}_{γ} as the 2D aeration map of the real samples is unavailable. 927
928
929
930
931
932

Model Calibration. Model calibration measures how well the model’s confidence in its predictions matches its actual accuracy [70]. Ideally, a well-calibrated model has confidence levels proportional to its prediction accuracy. We conduct a calibration test of LUNA for reconstructing aeration maps, showing that it performs well in simulation (fig. 5f). 933
934
935
936
937

To refine confidence levels, we applied a simple scaling method called Platt scaling [71], which adjusts predictions using a sigmoid function: $\hat{\rho}_{ij}^A = \sigma(\hat{p}_{ij}^A)$, where $\sigma(x) = \frac{1}{1+e^{-x}}$. This ensures predictions better reflect true aeration percentages. 938
939
940

The calibration results are reported in fig. 5f. We observe that the model is well-calibrated in silicon. Note that the calibration is not available on the real data as the ground-truth 2D aeration maps are unavailable. The calibration techniques help in adjusting the confidence scores of LUNA’s predictions to better reflect their true accuracy, ensuring that the confidence level is a reliable indicator of the model’s performance. 941
942
943
944
945
946
947

4.4 Implementation Details and Evaluation Protocols 948

Fullwave-2 Solver for Simulated LUS Data Generation. We simulate the wave propagation in the lung with the finite difference method with nonlinear wave propagation [28]. With a batch size of 300, the solver takes approximately 36 hours on 16 NVIDIA V100 GPUs. 949
950
951
952
953

Chest Wall Segmentation Model. We use RMSprop optimizer [72] with a learning rate of 1×10^{-5} , weight decay of 1×10^{-8} and momentum of 0.999. The batch size is 5. We train the model for 10 epochs on one NVIDIA 4090 GPU. 954
955
956

Aeration Reconstruction Model. We use Adam optimizer [73] with a learning rate of $\eta \cdot 0.002$ and $\beta = (0, 0.99^{\eta})$. The model is implemented with the PyTorch framework. The batch size is 26. We first train the model with simulated data for 90 epochs. We then fine-tune the model with real data for another 10 epochs. In total, our training took 1.5 days on one NVIDIA A100 GPU. 957
958
959
960
961
962

Evaluation Protocols. We adopt several metrics to evaluate the performance of LUNA. The first one is the error of the predicted lung aeration. 963
964
965
966

$$\text{aeration error} = \|\gamma - \hat{\gamma}\| \quad (20)$$

967 where $\gamma, \hat{\gamma}$ refers to the aeration of ground-truth and predicted lung aeration map
968 $\rho^A, \hat{\rho}^A$. For simulation data, ρ^A can be calculated according to E.q.2. For real data, the
969 aeration is measured with weighting, with details in Section 4.2.

970 As the ground-truth lung aeration map is also available for simulation, we consider
971 two other metrics to evaluate the 2D reconstruction quality.

972 1. The Normalized Mean Squared Error (NMSE) measures the average of the
973 squares of the errors normalized by the ground truth image's energy: $NMSE = \frac{\sum_{i,j}^{H,W} (\rho_{i,j}^A - \hat{\rho}_{i,j}^A)^2}{\sum_{i,j}^{H,W} (\rho_{i,j}^A)^2}$.

974 2. The Peak Signal-to-Noise Ratio (PSNR) measures the ratio between the maximum
975 possible power of a signal and the power of corrupting noise that affects the fidelity
976 of its representation: $PSNR = 10 \cdot \log_{10} \left(\frac{\max(\rho_{i,j}^A)^2}{\frac{1}{HW} \sum_{i,j} (\rho_{i,j}^A - \hat{\rho}_{i,j}^A)^2} \right)$. Note that H and
977 W are the height and width of the image, respectively.

981 Acknowledgments

982 This work is supported in part by NIH-R21EB033150 and ONR (MURI grant
983 N000142312654 and N000142012786). A.A. is supported in part by Bren endowed chair
984 and the AI2050 senior fellow program at Schmidt Sciences. B. T. is supported in part
985 by the Swartz Foundation Fellowship. Z. Li is supported in part by NVIDIA Fellowship.
986 The authors thank Zezhou Cheng and Julius Berner for their helpful discussions.

987

988 Competing Interests

989 The authors declare no competing interests.

990

991

992

993

994

995

996

997

998

999

1000

1001

1002

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

References

- [1] Lee, J. Lung ultrasound as a monitoring tool. *Tuberculosis and Respiratory Diseases* **83**, S12 (2020). 1013
1014
1015
1016
1017
1018
1019
- [2] Kalkanis, A. *et al.* Lung aeration in covid-19 pneumonia by ultrasonography and computed tomography. *Journal of Clinical Medicine* **11**, 2718 (2022). 1020
1021
1022
- [3] Volpicelli, G. *et al.* International evidence-based recommendations for point-of-care lung ultrasound. *Intensive care medicine* **38**, 577–591 (2012). 1023
1024
1025
- [4] Bouhemad, B., Mongodi, S., Via, G. & Rouquette, I. Ultrasound for “lung monitoring” of ventilated patients. *Anesthesiology* **122**, 437–447 (2015). 1026
1027
1028
1029
- [5] Lacedonia, D. *et al.* The role of transthoracic ultrasound in the study of interstitial lung diseases: high-resolution computed tomography versus ultrasound patterns: our preliminary experience. *Diagnostics* **11**, 439 (2021). 1030
1031
1032
1033
1034
1035
1036
1037
- [6] Manolescu, D. *et al.* Ultrasound mapping of lung changes in idiopathic pulmonary fibrosis. *The Clinical Respiratory Journal* **14**, 54–63 (2020). 1038
1039
1040
1041
1042
1043
1044
- [7] Lichtenstein, D. A. Lung ultrasound in the critically ill. *Annals of intensive care* **4**, 1–12 (2014). 1045
1046
1047
- [8] See, K. *et al.* Lung ultrasound training: curriculum implementation and learning trajectory among respiratory therapists. *Intensive care medicine* **42**, 63–71 (2016). 1048
1049
1050
- [9] Vitale, J. *et al.* Comparison of the accuracy of nurse-performed and physician-performed lung ultrasound in the diagnosis of cardiogenic dyspnea. *Chest* **150**, 470–471 (2016). 1051
1052
1053
1054
- [10] Whitworth, M., Bricker, L. & Mullan, C. Ultrasound for fetal assessment in early pregnancy. *Cochrane database of systematic reviews* (2015). 1055
1056
1057
1058
- [11] Mattoon, J. S., Berry, C. R. & Nyland, T. Abdominal ultrasound scanning techniques. *Small Animal Diagnostic Ultrasound-E-Book* **94**, 93–112 (2014). 1048
1049
1050
- [12] Demi, L. *et al.* New international guidelines and consensus on the use of lung ultrasound. *J Ultrasound Med* **42**, 309–344 (2023). 1048
1049
1050
- [13] Soldati, G., Smargiassi, A., Demi, L. & Inchegolo, R. Artifactual lung ultrasonography: It is a matter of traps, order, and disorder. *Applied Sciences* **10** (2020). 1051
1052
1053
1054
- [14] Pivetta, E. *et al.* Sources of variability in the detection of b-lines, using lung ultrasound. *Ultrasound in medicine & biology* **44**, 1212–1216 (2018). 1055
1056
1057
1058

- 1059 [15] Gomond-Le Goff, C. *et al.* Effect of different probes and expertise on the
1060 interpretation reliability of point-of-care lung ultrasound. *Chest* **157**, 924–931
1061 (2020).
- 1062
- 1063 [16] Gravel, C. A. *et al.* Interrater reliability of pediatric point-of-care lung ultrasound
1064 findings. *The American Journal of Emergency Medicine* **38**, 1–6 (2020).
- 1065
- 1066 [17] Ostras, O. & Pinton, G. Development and validation of spatial coherence beamform-
1067 ers for lung ultrasound imaging. *2023 IEEE International Ultrasonics Symposium
(IUS)* 1–4 (2023).
- 1068
- 1069 [18] Xing, W. *et al.* Automatic detection of a-line in lung ultrasound images using
1070 deep learning and image processing. *Medical physics* **50**, 330–343 (2023).
- 1071
- 1072 [19] Sagreya, H., Jacobs, M. & Akhbardeh, A. Automated lung ultrasound pulmonary
1073 disease quantification using an unsupervised machine learning technique for covid-
1074 19. *Diagnostics* **13**, 2692 (2023).
- 1075
- 1076 [20] Horry, M. J. *et al.* Covid-19 detection through transfer learning using multimodal
1077 imaging data. *Ieee Access* **8**, 149808–149824 (2020).
- 1078
- 1079 [21] Born, J. *et al.* Pocovid-net: automatic detection of covid-19 from a new lung
1080 ultrasound imaging dataset (pocus). *arXiv preprint arXiv:2004.12084* (2020).
- 1081
- 1082 [22] Roy, S. *et al.* Deep learning for classification and localization of covid-19 markers
1083 in point-of-care lung ultrasound. *IEEE transactions on medical imaging* **39**,
1084 2676–2687 (2020).
- 1085
- 1086 [23] Mento, F., Perpenti, M., Barcellona, G., Perrone, T. & Demi, L. Lung ultrasound
1087 spectroscopy applied to the differential diagnosis of pulmonary diseases: An in
1088 vivo multicenter clinical study. *IEEE transactions on ultrasonics, ferroelectrics,
and frequency control* **71**, 1217–1232 (2024).
- 1089
- 1090 [24] Xu, W. *et al.* Generalizability and diagnostic performance of ai models for thyroid
1091 us. *Radiology* **307**, e221157 (2023).
- 1092
- 1093 [25] Lee, P. M., Tofts, R. P. & Kory, P. Lung ultrasound interpretation. *Point of Care
Ultrasound E-book* 59 (2014).
- 1094
- 1095
- 1096 [26] Balcescu, C. *et al.* Automated lung ultrasound b-line assessment using a deep learn-
1097 ing algorithm. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency
1098 Control* **67**, 2312–2320 (2020).
- 1099
- 1100 [27] Li, Z. *et al.* Fourier neural operator for parametric partial differential equations
1101 (2020). 2010.08895.
- 1102
- 1103
- 1104

- [28] Pinton, G. A fullwave model of the nonlinear wave equation with multiple relaxations and relaxing perfectly matched layers for high-order numerical finite-difference solutions. *arXiv preprint arXiv:2106.11476* (2021). 1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
- [29] Gargani, L. & Volpicelli, G. How i do it: lung ultrasound. *Cardiovascular ultrasound* **12**, 1–10 (2014).
- [30] Pinton, G. F., Dahl, J., Rosenzweig, S. & Trahey, G. E. A heterogeneous nonlinear attenuating full-wave model of ultrasound. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **56**, 474–488 (2009).
- [31] Ostras, O., Shponka, I. & Pinton, G. Ultrasound imaging of lung disease and its relationship to histopathology: An experimentally validated simulation approach. *J Acoust Soc Am* **154**, 2410–2425 (2023).
- [32] Soulioti, D. E., Santibanez, F. & Pinton, G. F. Deconstruction and reconstruction of image-degrading effects in the human abdomen using fullwave: phase aberration, multiple reverberation, and trailing reverberation. *ArXiv* (2021).
- [33] Pinton, G. *et al.* Attenuation, scattering and absorption of ultrasound in the skull bone. *Medical Physics* 455–467 (2011).
- [34] Zhao, L., Fong, T. C. & Bell, M. A. L. Detection of covid-19 features in lung ultrasound images using deep neural networks. *Communications Medicine* **4**, 41 (2024).
- [35] Hou, D., Hou, R. & Hou, J. Interpretable saab subspace network for covid-19 lung ultrasound screening (2020).
- [36] Virieux, J. & Operto, S. An overview of full-waveform inversion in exploration geophysics. *Geophysics* **74**, WCC1–WCC26 (2009).
- [37] Li, L. *et al.* Recent advances and challenges of waveform-based seismic location methods at multiple scales. *Reviews of Geophysics* **58**, e2019RG000667 (2020).
- [38] Trahey, G., Bottensus, N. & Pinton, G. Beamforming methods for large aperture imaging. *The Journal of the Acoustical Society of America* **141**, 3610–3611 (2017).
- [39] Bottensus, N., Pinton, G. & Trahey, G. The impact of acoustic clutter on large array abdominal imaging. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* (2019).
- [40] Dahl, J. J., Jakovljevic, M., Pinton, G. F. & Trahey, G. E. Harmonic spatial coherence imaging: An ultrasonic imaging method based on backscatter coherence. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* **59**, 648–659 (2012).

- 1151 [41] Pinton, G., Aubry, J.-F., Fink, M. & Tanter, M. Numerical prediction of frequency
1152 dependent 3d maps of mechanical index thresholds in ultrasonic brain therapy.
1153 *Medical physics* **39**, 455–467 (2012).
- 1154
- 1155 [42] Pinton, G. F., Aubry, J.-F. & Tanter, M. Direct phase projection and transcranial
1156 focusing of ultrasound for brain therapy. *Ultrasonics, Ferroelectrics and Frequency
1157 Control, IEEE Transactions on* **59**, 1149–1159 (2012).
- 1158
- 1159 [43] Chandrasekaran, S. *et al.* Shear shock wave injury in vivo: High frame-rate
1160 ultrasound observation and histological assessment. *Journal of biomechanics* **166**,
1161 112021 (2024).
- 1162
- 1163 [44] Ostras, O., Soulioti, D. & Pinton, G. Diagnostic ultrasound imaging of the lung: A
1164 simulation approach based on propagation and reverberation in the human body.
1165 *J Acoust Soc Am* **150**, 3904 (2021).
- 1166
- 1167 [45] Rocca, E. *et al.* Lung ultrasound in critical care and emergency medicine: Clinical
1168 review. *Advances in Respiratory Medicine* **91**, 203–223 (2023).
- 1169
- 1170 [46] Mento, F. & Demi, L. On the influence of imaging parameters on lung ultrasound
1171 b-line artifacts, in vitro study. *The Journal of the Acoustical Society of America*
1172 **148**, 975–983 (2020).
- 1173
- 1174 [47] Mento, F. & Demi, L. Dependence of lung ultrasound vertical artifacts on frequency,
1175 bandwidth, focus and angle of incidence: An in vitro study. *J Acoust Soc Am.*
1176 **150**, 4075 (2021).
- 1177
- 1178 [48] Dietrich, C. *et al.* Lung b-line artefacts and their use. *Journal of thoracic disease*
1179 **8**, 1356–1365 (2016).
- 1180
- 1181 [49] Courant, R., Friedrichs, K. & Lewy, H. On the partial difference equations of
1182 mathematical physics. *IBM Journal of Research and Development* **11**, 215–234
(1967).
- 1183
- 1184 [50] Spitzer, V., Ackerman, M., Scherzinger, A. & Whitlock, D. The visible human
1185 male: a technical report. *Journal of the American Medical Informatics Association*
1186 **3**, 118–130 (1996).
- 1187
- 1188 [51] Kameda, T., Kamiyama, N. & Taniguchi, N. The mechanisms underlying vertical
1189 artifacts in lung ultrasound and their proper utilization for the evaluation of
1190 cardiogenic pulmonary edema. *Diagnostics (Basel)* **12**, 252 (2022).
- 1191
- 1192 [52] Knudsen, L., Weibel, E., Gundersen, H., Weinstein, F. & Ochs, M. Assessment of
1193 air space size characteristics by intercept (chord) measurement: an accurate and
1194 efficient stereological approach. *J Appl Physiol* **108**, 412–421 (2010).
- 1195
- 1196

- [53] McKeighen, R. & Buchin, M. New techniques for dynamically variable electronic delays for real time ultrasonic imaging. *1977 Ultrasonics Symposium* 250–254 (1977). 1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
- [54] Destrempes, F. & Cloutier, G. A critical review and uniformized representation of statistical distributions modeling the ultrasound echo envelope. *Ultrasound in Medicine and Biology* **36**, 1037–1051 (2010).
- [55] Salles, S. *et al.* Experimental evaluation of spectral-based quantitative ultrasound imaging using plane wave compounding. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **61**, 1824–1834 (2014).
- [56] Park, S. & Yoo, Y. A new fast logarithm algorithm using advanced exponent bit extraction for software-based ultrasound imaging systems. *Electronics* **12** (2023). URL <https://www.mdpi.com/2079-9292/12/1/170>.
- [57] Vara, G. *et al.* Texture analysis on ultrasound: The effect of time gain compensation on histogram metrics and gray-level matrices. *J Med Phys* **45**, 249–255 (1920).
- [58] Judge, E. *et al.* Anatomy and bronchoscopy of the porcine lung. a model for translational respiratory medicine. *American journal of respiratory cell and molecular biology* **51**, 334–343 (2014).
- [59] Oizumi, H. *et al.* Techniques to define segmental anatomy during segmentectomy. *Annals of Cardiothoracic Surgery* **3** (2014). URL <https://www.annalscts.com/article/view/3507>.
- [60] Scarpelli, E. The alveolar surface network: a new anatomy and its physiological significance. *The Anatomical record* **251**, 491–527 (1998).
- [61] Braithwaite, S. A., van Hooijdonk, E. & van der Kaaij, N. P. Ventilation during ex vivo lung perfusion, a review. *Transplantation Reviews* **37**, 100762 (2023). URL <https://www.sciencedirect.com/science/article/pii/S0955470X23000162>.
- [62] Lee, J., Fang, X., Gupta, N., Serikov, V. & Matthay, M. Allogeneic human mesenchymal stem cells for treatment of e. coli endotoxin-induced acute lung injury in the ex vivo perfused human lung. *Proc Natl Acad Sci USA* **106**, 16357–16362 (2009).
- [63] Stengel, P., Frazer, D. & Weber, K. Lung degassing: An evaluation of two methods. *Journal of applied physiology* **48**, 370–375 (1980).
- [64] Edmunds, L. & Huber, G. Pulmonary artery occlusion. i. volume-pressure relationships and alveolar bubble stability. *Journal of applied physiology* **22**, 990–1001 (1967).

- 1243 [65] Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for
1244 biomedical image segmentation (2015).
- 1245
- 1246 [66] Dice, L. R. Measures of the amount of ecologic association between species.
1247 *Ecology* **26**, 297–302 (1945).
- 1248
- 1249 [67] Rashid, M. M., Pittie, T., Chakraborty, S. & Krishnan, N. A. Learning the
1250 stress-strain fields in digital composites using fourier neural operator. *Iscience* **25**
1251 (2022).
- 1252 [68] Pathak, J. *et al.* Fourcastnet: A global data-driven high-resolution weather model
1253 using adaptive fourier neural operators. *arXiv preprint arXiv:2202.11214* (2022).
- 1254
- 1255 [69] Guan, S., Hsu, K.-T. & Chitnis, P. V. Fourier neural operator network for fast
1256 photoacoustic wave simulations. *Algorithms* **16**, 124 (2023).
- 1257
- 1258 [70] Guo, C., Pleiss, G., Sun, Y. & Weinberger, K. Q. On calibration of modern neural
1259 networks (2017).
- 1260
- 1261 [71] Platt, J. *et al.* Probabilistic outputs for support vector machines and comparisons
1262 to regularized likelihood methods. *Advances in large margin classifiers* **10**, 61–74
1263 (1999).
- 1264
- 1265 [72] Hinton, G., Srivastava, N. & Swersky, K. Neural networks for machine learning
1266 lecture 6a overview of mini-batch gradient descent. *Cited on* **14**, 2 (2012).
- 1267
- 1268 [73] Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv
1269 preprint arXiv:1412.6980* (2014).
- 1270
- 1271
- 1272
- 1273
- 1274
- 1275
- 1276
- 1277
- 1278
- 1279
- 1280
- 1281
- 1282
- 1283
- 1284
- 1285
- 1286
- 1287
- 1288

Appendix A Lung Histology Generation Details

This section describes the creation of lung histology maps for simulating lung ultrasound data. LUNA is trained on abundant simulation data.

Acoustical property maps were generated by combining human chest wall anatomical images (Visible Human Project, 330 μm resolution) with high-resolution histological images of healthy swine lung tissue (5 μm thick, 0.55 μm resolution), following [31] (fig. 2). Binary-segmented histological images quantified lung aeration (air vs. non-air), producing maps where 1 represents air and 0 represents non-air ($\rho^A : \mathbb{Z}^2 \rightarrow 0, 1$). Details of this process are shown in fig. A1 in the supplementary materials.

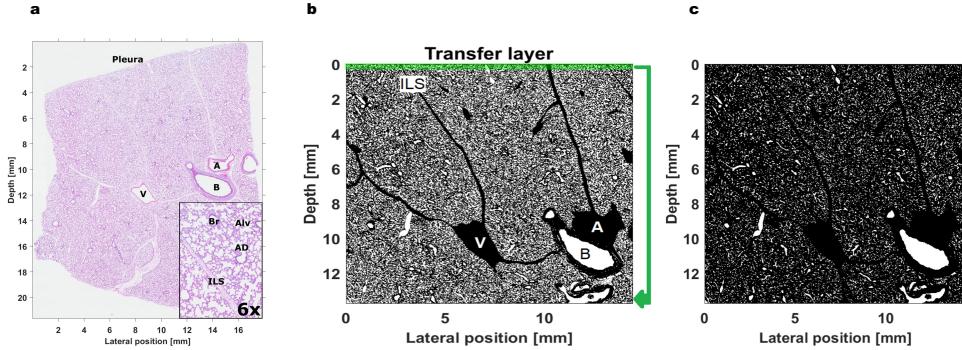
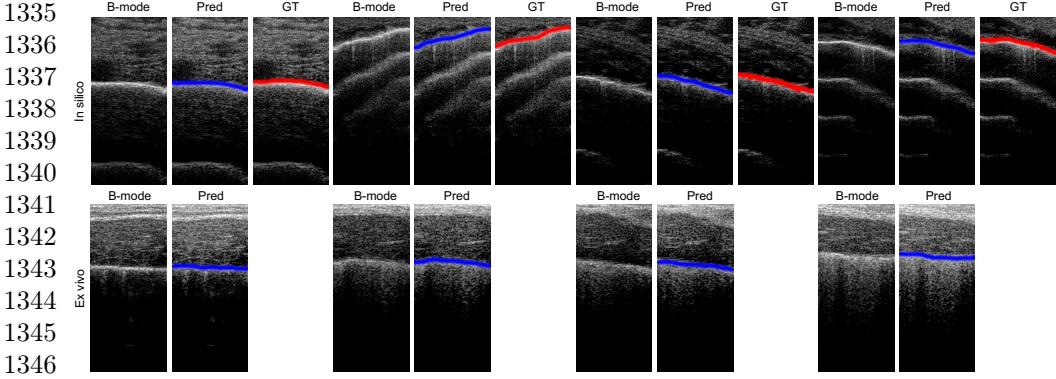


Fig. A1: Formation and processing of lung aeration map for in silicon simulation. a, Formation of the acoustical maps as a stationary input for Fullwave-2 simulation tool: histology (H&E) of healthy swine lung. Insertion shows a magnified part of it and the ability to identify lung tissue architecture. B - bronchus, Br - bronchiole, A - artery, V - vein, AD - alveolar duct, Alv - alveolus, ILS - interlobular septum. b, Binary-segmented and leveled histology image deformed to conform to linear pleural line and cropped to a rectangular shape. Its superficial 0.1 mm thick layer (green) is used for repetitive transfer procedures to increase the spatial variability of lung structures. c, Segmented histology after applied algorithmic modification modeling ARDS (added fluid/non-air pixels are evenly distributed among tissue-air interfaces) with target aeration of 14%.

Appendix B Chest Wall Segmentation Results

This section presents the results of chest wall segmentation, highlighting the separation line (pleural line) between the chest wall and the lung, which is consistently located below the chest wall in the input B-mode ultrasound images (fig. C3). For visualization, the separation line is overlaid onto the B-mode images to demonstrate the segmentation performance. On the in silico test set, our model achieves an impressive Dice similarity coefficient [66] of 95.8%, indicating near-perfect alignment with the ground truth segmentation obtained from the simulation. In the ex vivo test set, the model performs reliably despite the inherent domain differences between simulation and real-world

1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334



1347 **Fig. B2: Chest wall segmentation results.** The separation line (pleural line) between the
 1348 chest wall and the lung is overlaid onto B-mode ultrasound images for visualization. Upper:
 1349 For in silico evaluation samples, the predicted segmentation (Pred) aligns closely with the
 1350 ground truth (GT) obtained from the simulation. Lower: For ex vivo evaluation samples, the
 1351 segmentation performance remains robust, with a small performance gap observed between
 1352 the simulated and real-world datasets.

1352

1353

1354 data. The results suggest a minimal performance gap between the in silico and ex vivo
 1355 conditions, demonstrating the robustness of our method.

1356

1357 Appendix C Ablation Study on Network Design

1359 In this ablation study, we investigate the impact of different components of our proposed
 1360 model on the ex vivo percent aeration prediction performance.

- 1361 1. **Temporal augmentation.** As discussed in Section 4.3 of the main paper, tem-
 1362 poral masking strengthens the invariance of LUNA to temporal delay. Empirically,
 1363 removing temporal augmentation leads to 6.2% performance loss.
- 1364 2. **Spatial augmentation.** As discussed in Section 4.3 of the main paper, spatial
 1365 masking improves LUNA’s ability to learn lateral correspondence. Empirically,
 1366 removing spatial augmentation leads to 0.9% performance loss.
- 1367 3. **FNO module.** The inclusion of the temporal Fourier Neural Operator (FNO)
 1368 significantly reduces prediction error compared to replacing it with a ResNet
 1369 module with equivalent parameters, highlighting the effectiveness of FNO in
 1370 capturing temporal dynamics. Empirically, replacing FNO with ResNet leads to
 1371 1.2% performance loss.
- 1372 4. **The percent aeration loss.** We set the aeration loss’s weight $\eta = 0$ in Eqn. 19.
 1373 We show that \mathcal{L}_γ further refines the model’s predictions by directly optimizing
 1374 the percent aeration accuracy. Empirically, removing percent aeration loss leads
 1375 to 2.1% performance loss.

1376 These components collectively contribute to the performance improvement as shown in
 1377 fig. C3.

1378

1379

1380

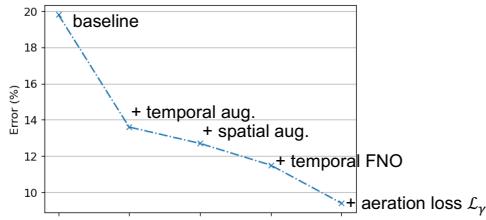


Fig. C3: Ablation study results of LUNA. The ex vivo percent aeration prediction performance is shown as different components of the proposed model are progressively added. Temporal augmentation and spatial augmentation improve generalization, while the temporal Fourier Neural Operator (FNO) outperforms a ResNet module with equivalent parameters. The addition of aeration loss \mathcal{L}_γ further refines predictions, achieving the lowest error.

1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403
 1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426