Open-H-RF: A Large-Scale Ultrasound Raw Channel Dataset Initiative

T.S.W. Stevens[*,§], S.W. Penninga[*,§], W.L. van Nierop[*,§], O. Nolan[*,§], F. Duelmer[§], B.N. Frey[§], B. Federici, V. van Hal, L.D. van Harten, H. de Hoop, R. Lopata, B. Luijten, M. van Sambeek, V. van der Schaft, H.-M. Schwab, M.F. Azampour, N. Navab, R. Ali, A. Andrzejek, J. Baek, B. Boitnott, T. Brevett, H. Hashemi, D. Hyun, J.H. Park, A.V. Telichko, C. Thomas, L. Zhuang, Y. Aklilu, H. Arora, J. Barker, A.K. Mackanic, P. McDowall, L. McHugh, I. Pirozzi, S. Simon-Trench, A.M. Bronstein, O. Senouf, S. Vedula, D. Zadok, H. Rivaz, M. Sharifzadeh, S. Chen, G.P. Luke, S. Wang, Yujia Wu, J. Verde, R. Göbl, O. Zettinig, E. Boctor, Yixuan Wu, H.-M. Bae, M.-G. Kim, Y.-M. Kim, H. Lee, S.-H. Oh, P. Pinto, M. Berrada-Sounni, M. Carley, R.J. Fedewa, S. Gavin, N. Goldsmith, R. Kassayan, C. Low, R. Mehranfar, G. Bardoscia, M. Boccardo, A. Botter, M. Carbonaro, G.L. Cerone, E. Cesti, K. Meiburger, S. Seoni, F. Truscello, B.J. Raiteri, L. Caron, M. Gasse, N. Ghigo, O. Gulenko, A. Leconte, S.A. Lee, A. Malescot, E. Martineau, J. Porée, J. Provost, G. Ramos-Palacios, B. Rauby, R.L. Rungta, A.F. Sadikot, A. Wu, P. Xing, A. Chavignon, C. Ferlay, V. Hingot, M. Reberol, A. Waelbroeck, H.-S. Kwon, J. Bucholz, I. Guracar, R. Loftman, C. Williams, T. Grutman, T. Ilovitsh, T.A. Aleef, P.B. Chen, W. Chen, Y. Chen, M. Frew, M. Honarvar, Y.C. Lee, Z. Li, H. Moradi, R. Rohling, S.E. Salcudean, Z. Wu, Q. Zeng, P.C. Cattin, S. El Hadramy, N. Bottenus, A.S. Perez, S. Feng, Y. Feng, P. Jin, Y. Lin, O. Ostras, G. Pinton, F. Santibanez, M. Sode, B. Turkbey, H. Wang, B.J. Wood, L. Zhang, S. Luo, J. Wiskin, A. Austeng, Y. Karabiyik, S.P. Näsholm, O.M.H. Rindal, D. Bustamante, M. Mehrmohammadi, H. Raeis, A. Hameddine, G. Lajoinie, H. de Roo, M. Versluis, R. Zorgdrager, J.S. Au, A.J.Y. Chee, J. Cohen, T. Gu, J.Y.-H. Hsu, S. Ling, M. Mourtzakis, H. Nahas, P. de la Torre, D. Xiao, A.C.H. Yu, E. Zafiris, P. Jarosik, P. Karwat, Z. Klimonda, M. Lewandowski, B. Byram, C. Khan, Z. Marshall, Y.-C. Pan, S. Liu, T. Sultan, C. Hoerig, J. Mamou, D. Auerbach, Y.C. Eldar, A. Wegerhoff, G.S. Fischer, S. Katayose, X. Ma, R. Murakami, Y. Tang, R. Tsumura, Y. Wang, H.K. Zhang, M. Azizian, S. Huver, J.J. Dahl[†,§], R.J.G. van Sloun[†,§], W. Simson[†,§]

***Abstract*— Medical ultrasound is a widely accessible, real-time imaging modality, yet progress in data-driven ultrasound research is constrained by the scarcity of large, open collections of raw channel data. Existing public datasets are predominantly based on reconstructed images, while the smaller number of open raw channel datasets is generally specialized to individual reconstruction or imaging tasks, limiting the development of general-purpose methods and foundation models that operate directly on ultrasound signals. While interpretive tasks can be performed on reconstructed images, many of the key technical challenges in ultrasound imaging, such as aberration correction and reverberation reduction, require the raw channel data itself. We introduce *Open-H-RF*, a community-driven open dataset initiative comprising approximately 40 TB of medical ultrasound raw channel data in 19,332 acquisitions, contributed by 38 institutions across academia and industry, spanning *in-vivo*, phantom, and simulated data. Each dataset is annotated with a number of supported tasks, ranging from generalized reconstruction, speed-of-sound estimation, and aberration correction to blood flow imaging, motion estimation, quantitative ultrasound, and image interpretation. Open-H-RF provides a unified, self-contained data format that captures raw channel data together with the acquisition, probe, metadata, and reconstruction information required for reproducible processing across scanners and acquisition schemes. The corpus is publicly hosted under a permissive CC BY 4.0 license and can be streamed programmatically and selectively at scale. By bringing heterogeneous ultrasound data into a common representation, backed by modern data infrastructure, Open-H-RF lowers the barrier to working with raw channel data and aims to open medical ultrasound research to new communities. See `https://open-h.github.io/Open-H-RF`.**

***Index Terms*— computational imaging, open data, ultrasound imaging**

I. Introduction

OPEN datasets have repeatedly acted as catalysts for progress in data-driven research. They provide common benchmarks and standardized baselines against which new methods can be compared fairly, form the basis of community challenges, and enable scientific discovery. In computer vision, large open collections of natural images have driven progress for over a decade [1]. Similarly, in medical imaging,

The steering committee of the Open-H-RF consortium consists of NVIDIA, Eindhoven University of Technology (TU/e) and Stanford University. [*]Equal contribution as first authors. [†]Equal contribution as senior authors. [§]Core team.

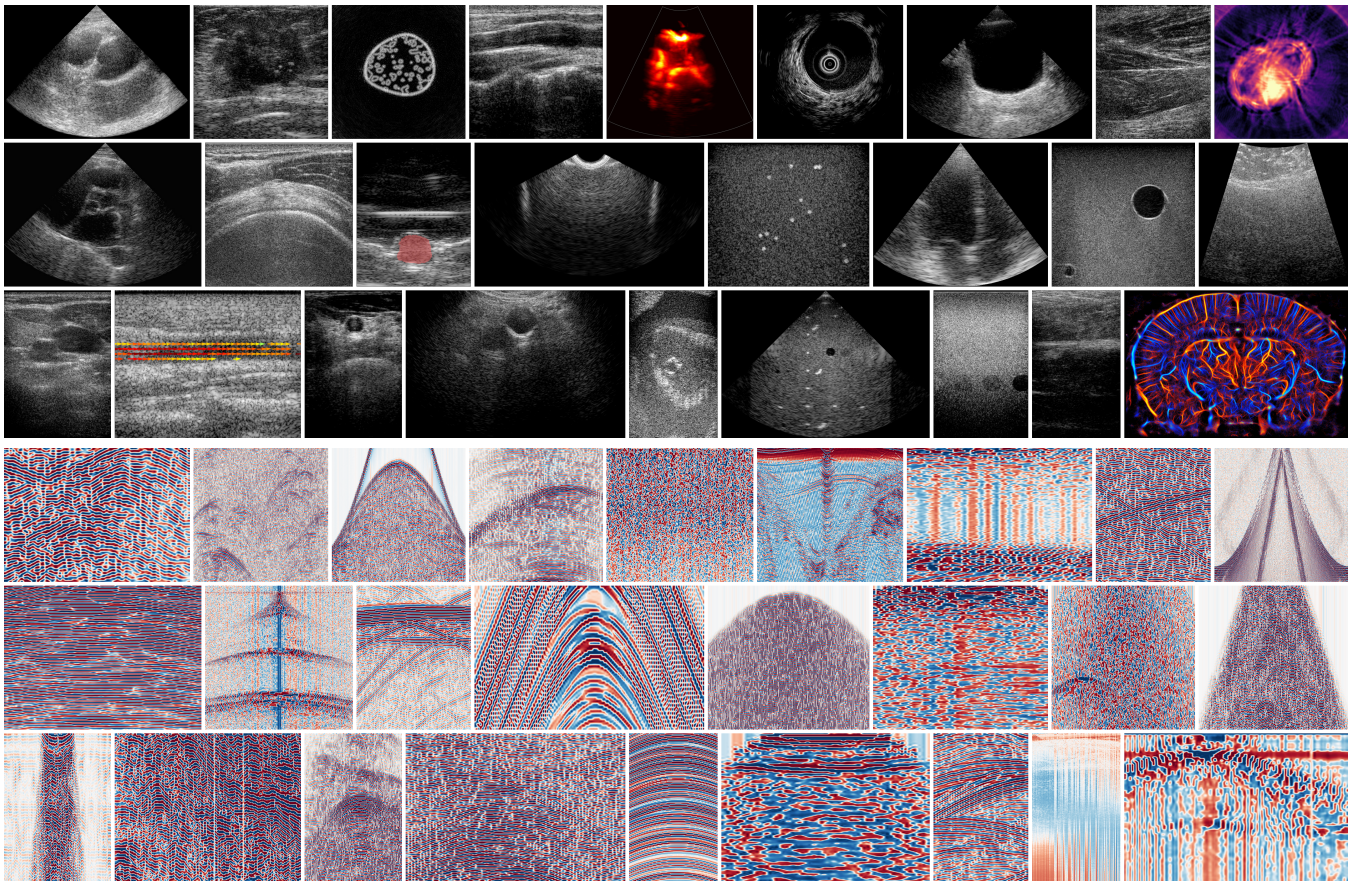

Fig. 1: Impression of the dataset. *Top:* reconstructed images from a selection of samples across the contributions. *Bottom:* the corresponding (demodulated) raw channel data from the same samples, in matching order. Because the full raw channel data required for reconstruction is a far larger tensor (spanning all transmits, receive elements, and axial samples) that cannot be displayed directly, each radio-frequency (RF) panel shows only a single, randomly selected transmit, cropped for visualization, and is meant as an impression of the underlying signal. The dataset itself stores the complete raw channel data, alongside all parameters and metadata needed to derive the reconstructions.

open datasets have enabled progress across modalities. In magnetic resonance imaging (MRI), for instance, the *fastMRI* dataset [2], consisting of thousands of raw *k-space* samples, has spurred the development of dozens of novel approaches to MRI reconstruction [3], [4]. Large, standardized, openly shared datasets have also become the key enabler of foundation models: models pretrained on large and diverse corpora whose weights can be fine-tuned for a wide range of downstream tasks [5]. In medical imaging, such models are now trained on large collections of computed tomography (CT) and MRI data [6], [7]. In ultrasound imaging, USFM was pretrained on more than two million brightness-mode (B-mode) images across twelve organs [8] and EchoCLIP on over one million echocardiogram images with paired reports [9], but in both cases, the full training dataset is not publicly available. USF-MAE, on the other hand, was trained on approximately 370k B-mode images pooled from 46 public collections [10]. For echocardiography, there are large datasets of B-mode sequences with corresponding models for functional parameters such as ejection fraction and left ventricular internal diameter [11], [12], but these are narrowly scoped to those tasks only. Combining data from several institutions gives more variety in scanners, protocols, and patients than any one group could collect alone, and under a permissive license, both academic and industry researchers can benefit from each other.

Ultrasound is the most widely deployed diagnostic imaging modality in the world [13]. It is real-time, non-ionizing, comparatively inexpensive, and portable enough to be brought to the patient. This matters for the approximately $47\,\%$ of the global population that has little to no access to diagnostics, a gap that is largest at the primary-care level where cost, infrastructure, and staffing constraints rule out CT and MRI [14]. Ultrasound is also the only clinical modality that images at frame rates high enough to resolve millisecond-scale physiology, which underpins shear-wave elastography, functional ultrasound, and super-resolution imaging [15], [16]. However, ultrasound image quality depends on *operator skill* and the *patient's anatomy*, and ultrasound does not offer standard and reproducible imaging planes, which challenges consistent interpretation.

Every one of the aforementioned ultrasound datasets consists of brightness-mode (B-mode) images. B-mode is the final product of a lossy pipeline: beamforming, typically under an assumed constant speed of sound (SoS), envelope detec-

Fig. 2: Global map of the origin of the contributions to Open-H-RF, also showing the type of contributor (academia, industry or mixed).

tion, log compression, and vendor-specific post-processing. The raw measurements consist of the *raw channel data*: the echo signals recorded by each individual transducer element after every transmit event (Section II). Raw channel data retains the channel-level amplitude and phase enabling the analysis of inter-channel coherence, and frequency information that the displayed image discards. Moreover, the dynamic range is typically two orders of magnitude larger than that of B-mode images. Interpretive tasks, such as segmentation, quantification, and diagnosis, can be performed on B-mode images. However, many of the key technical challenges in ultrasound imaging, such as correcting inaccurate assumptions (e.g., aberration correction and reverberation reduction) and tissue characterization (e.g., tomographic attenuation imaging), require access to the raw channel data (Section III-B).

Image reconstruction reduces the data by roughly two orders of magnitude. Beamforming collapses the signals of all elements into a single value per pixel, discarding the per-element information needed to estimate the true SoS, correct aberrations, or separate reverberation from the direct echoes, and envelope detection further discards the phase on which Doppler and coherence-based methods rely. Accordingly, a substantial body of work has shown that operating directly on raw channel data or pre-envelope beamformed signals enables capabilities beyond what the displayed image can support [17]: adaptive beamforming and image reconstruction [18], speed-of-sound inversion [19], aberration correction [20], [21], dehazing [22], microbubble localization for super-resolution imaging [23], and clutter filtering and flow estimation [24], [25]. To date, the community has not explored how data scale and variety can support solving these fundamental problems in medical ultrasound imaging.

The scarcity of open-access raw channel data can be explained by three factors: *hardware access*, *fragmented formats*, and *data rates*. First, *hardware access*: clinical ultrasound scanners generally do not permit the extraction of raw channel data. Their processing is performed in dedicated hardware created for clinical real-time imaging. Access to raw channel data is commonly considered only useful for device engineering and not readily available. Furthermore, the raw data rates exceed the design specifications of clinical storage and networking. Raw channel data is therefore available mainly from research scanners requiring dedicated hardware. Second, *fragmented formats*: acquisition parameters and storage conventions vary by vendor and by lab, for example, between the Verasonics MATLAB environment, the us4us ARRUS package, and open initiatives such as the Ultrasound File Format (UFF) [26], the UltraSound ToolBox (USTB) [27], and `zea` [28]. This makes it challenging to combine and explore data from different sources. Third, *data rates*: ultrasound can sample thousands of elements at multi-megahertz rates, producing on the order of a gigabyte per second and terabytes per session [29]. Transferring and storing raw channel data is therefore traditionally burdensome and has imposed a practical limit on the size of both private and open datasets.

Despite these barriers, several open resources of raw channel data have been released. PICMUS established a standardized benchmark for plane-wave beamforming, consisting of a total of six simulated, phantom, and *in-vivo* acquisitions with common image-quality metrics against which beamformers could be compared [30]. CUBDL extended this to deep-learning-based beamforming by pooling data from seven institutions and 576 acquisition sequences across *in-vivo*, phantom, and simulated sources [31], [32]. PALA and ULMShare made data available for ultrasound localization microscopy, enabling direct comparison of super-resolution and microvascular imaging methods [23], [33]. These efforts demonstrated the value of shared raw channel data. However, each dataset targets a single task, covers a limited range of anatomies, probes, and transmit schemes, and uses its own data format. None were designed for, or approach, the scale and variety required for task-agnostic end-to-end foundation model training.

Addressing these shortcomings requires a pooled global initiative as has been demonstrated in other fields. Open X-Embodiment demonstrated that aggregating heterogeneous robot data under a common schema yields models that outperform any of the contributing single-source policies [34]. Open-H-Embodiment adapted that model to the sensing, safety, and governance constraints of the medical domain, pooling 780 hours of data from more than 50 institutions under a unified format and showing that domain-specific pretraining at that scale transfers across tasks and hardware platforms [35]. Open-H-RF follows this model for raw channel data. By providing access to a large corpus of raw channel data covering a variety of scanning regimes and imaging media, Open-H-RF aims to support open comparison, algorithmic development, and scientific discovery on raw channel data.

This work introduces Open-H-RF, a community dataset initiative comprising $40\,\mathrm{TB}$ of raw channel data from $19{,}332$ acquisitions, spanning *in-vivo*, phantom, and simulated sources (Fig. 1). In terms of **scale and diversity**, it exceeds the largest existing open collection [31] by more than an order of magnitude and covers a wide range of anatomies, phantoms, probe types, and transmit schemes (Section IV-B). A **unified format**, built on existing open standards and the `zea` [28] tooling, makes acquisitions from different scanners, probes, and transmit sequences directly interoperable and documents the conventions needed to interpret them correctly (Section III-C). **Permissive licensing** under CC BY 4.0 removes the access agreements and non-commercial restrictions that limit the reuse of several existing ultrasound datasets, allowing both academic and commercial use. Finally, the contributing groups supplied intended **tasks** with demonstrations, which we organize into five categories: *probe motion*, *tissue motion*, *inaccurate assumptions*, *tissue characterization*, and *therapy* (Section III-B). These include tracked 3D and swept synthetic

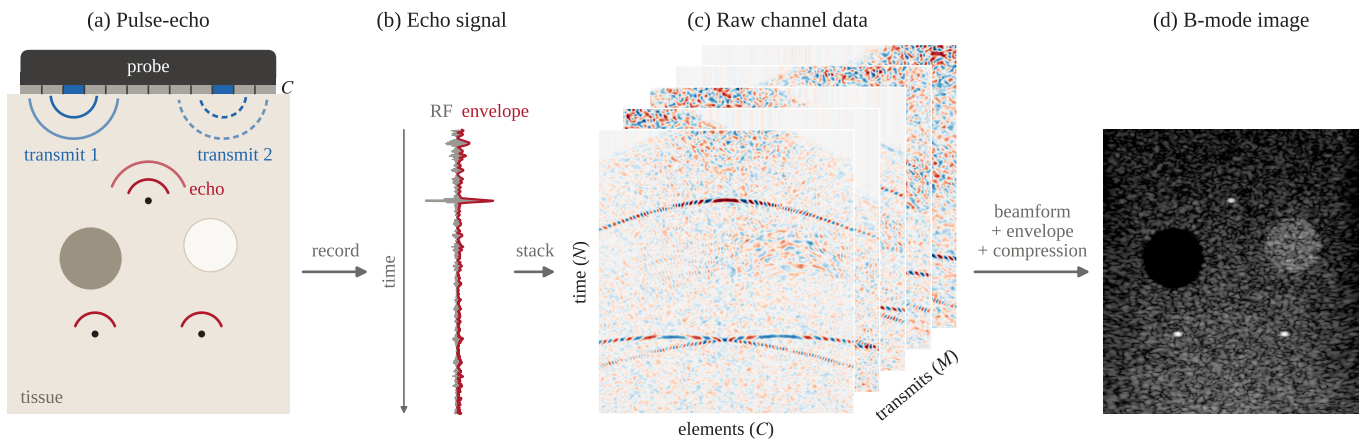

Fig. 3: Simplified overview of ultrasound imaging. (a) In each transmit event, the transducer array emits an acoustic pulse into the tissue; shown here are two synthetic aperture transmits, each fired by a single element. Wherever the acoustic properties in the tissue change, part of the wave is scattered back as echoes. (b) Each element records the returning echoes over time as a radio-frequency (RF) signal, whose envelope shows the echo strength. (c) Recording all C elements for each of the M transmit events gives one frame of raw channel data with N samples per element ($C \times M \times N$, shown demodulated). Each point scatterer appears as a hyperbola, because its echo reaches each element at a slightly different time. (d) Beamforming aligns and sums these signals for every pixel, and envelope detection and log compression then give the B-mode image. Here we show a simulated phantom with three point targets, an anechoic cyst, and a hyperechoic inclusion.

aperture (SSA) imaging, blood flow imaging, SoS estimation, and aberration correction, supported by complementary metadata such as reference probe positions, flow maps, and SoS maps stored alongside the raw channel data.

II. Ultrasound Imaging

Ultrasound imaging is based on the pulse-echo principle [17], [36], [37], see Fig. 3. A transducer array of C elements emits a short acoustic pulse into the body. We refer to this as a *transmit event*. By firing the elements with relative delays, a wavefront can be focused at a point, steered as a plane wave, diverged from a virtual source behind the array, or emitted from a single element. As the wave propagates, it is partially reflected and scattered wherever the acoustic properties of the tissue change, and the reflected pressure waves return to the transducer as *echoes*. Each element measures the incident pressure integrated over its surface, and this signal is digitized by an analog-to-digital converter at tens of MHz, yielding N fast-time samples per element. These digitized pressure measurements are the *raw channel data*, stored either as real-valued RF signals or as demodulated, complex-valued in-phase and quadrature (IQ) signals. A *frame* is the set of M transmit events that together are used to reconstruct a single image, so that one frame of raw channel data can be written as a tensor with dimensions $C \times M \times N$. We use the term *acquisition* to refer to a sequence of frames acquired in a single imaging session. Ultrasound imaging is typically two-dimensional, but volumetric (3D) imaging can be achieved using 2D matrix arrays, which enable electronic beam steering and focusing in both the lateral and elevational dimensions [38]. However, fully sampled matrix arrays require substantially more transducer elements and electronic channels than conventional 1D arrays, resulting in increased hardware complexity and data throughput [39].

To form an image, the raw channel data is *beamformed*. The most common method, delay-and-sum (DAS), computes for each pixel the time of flight of the wavefront from the transducer to that pixel and back to each element, and sums the correspondingly delayed channel signals. These delays assume a constant SoS, usually $1540\,\mathrm{m/s}$, although soft tissue ranges from about $1450\,\mathrm{m/s}$ in fat to above $1600\,\mathrm{m/s}$ in muscle. The beamformed signal is then envelope-detected and log-compressed to form the familiar *B-mode* (brightness mode) image. Motion can be measured across multiple frames, which is used in Doppler and vector flow imaging. These imaging modes estimate blood velocity from the phase changes between consecutive frames, after clutter filtering separates the weakly scattering blood from the much stronger tissue signal.

Each image formation step discards information or relies on assumptions that are violated in practice: tissue is heterogeneous in SoS, multiple scattering between tissue layers causes reverberation, and attenuation depends on frequency, tissue absorption and scattering. Image formation from raw channel data is therefore an inverse problem with a physically grounded but only approximate forward model. The data is also high-dimensional and high-rate, and the acquisition scheme itself, i.e., which transmit events to fire, is a design choice that can be optimized [40].

III. Methods

A. Call for Contributions

Data collection began with a public announcement and request for proposals (RFP) document. The RFP described the motivation for collecting a dataset of raw channel data and subsequent training of a foundation model, along with a list of task groups for which access to raw channel data is particularly relevant. The RFP specified that the provided data

TABLE I: Open-H-RF contributions: acquisitions by medium and intended tasks.

Contribution	Acq.	Tasks
`colorado-boulder`	★42 ●20	Recon, Motion, IQA, 3D
`concordia` [41]	▲2,000	Recon, Interp, QUS
`dartmouth-uct` [42], [43]	▲2,625	Recon, SoS, QUS
`kaist-snubh-barreleye`	★70	Recon, SoS, Interp, QUS
`mosaic-nuevosono`	★9	Interp, Motion
`nv-raw2insights-us` [44]	▲923	Recon, SoS, Interp
`oslo` [45]	★5 ●23 ▲11	Recon, Motion, IQA
`politorino` [46]	★30	Motion
`resolvestroke`	★40 ●3	Recon, Motion, IQA, 3D, Flow
`siemens-healthineers`	★590	Recon
`stanford-murine` [47], [48], [49], [50]	★102 ●143	Recon, SoS
`strasbourg-basel`	●9	Recon, SoS, 3D
`technion`	★39 ●1	Recon, Interp, Motion, IQA, Clutter
`tel-aviv` [51], [52]	★16 ●6	Interp
`tue-aaa`	★15	Recon, Motion
`tue-cardiac` [53]	★12	Recon
`tue-carotid`	★80	Recon, IQA
`tumunich`	●11	Recon, 3D
`twente-cavitation`	●19	MCD
`twente-microbubblesim` [54]	▲500	Flow
`twente-vortexflow`	●6	Recon, IQA, Flow
`ubc` [55]	▲7,845	Recon, Interp, 3D
`ulmshare` [33], [56]	★99	Motion, Flow, Clutter
`unc-liver` [57], [58], [59]	▲1,906	Recon, SoS, Interp, Clutter
`unc-openpros` [60]	▲248	Recon, SoS, QUS
`us4us`	★16 ●2	Recon
`vanderbilt`	★164 ●1	Recon, IQA, Clutter
`waterloo-carotid` [61]	★93	Recon, Motion, Flow, Clutter
`waterloo-femoralvein` [62]	★15	Recon, Motion, Flow, Clutter
`waterloo-muscle` [63]	★1,248	Recon, SoS
`weillcornell`	●120	Recon, QUS
`weizmann-sampl`	★30	Recon, SoS
`wpi` [64]	●5 ▲190	Recon, 3D
Total (33 **contributions**)	★2,715 ●369 ▲16,248	19,332 acquisitions

Media: ★ *in-vivo*; ● phantom; ▲ simulation.
Tasks: Recon: reconstruction, beamforming; **SoS**: Speed of sound estimation, aberration correction; **Interp**: interpretation, segmentation, detection; **Motion**: Motion tracking; **IQA**: Image quality assessment; **3D**: tracked reconstruction, matrix imaging; **Flow**: Blood flow, ULM; **QUS**: Quantitative ultrasound; **Clutter**: Clutter/reverb suppression; **MCD**: Microbubble cavitation detection.

must be channel data, and provided a data tier list ranging from *in-vivo* human data to digitally simulated data. An important requirement for each contribution was that its data fell under the Creative Commons BY 4.0 license (CC BY 4.0), ensuring that it could be freely shared and reused provided appropriate attribution was given. For *in-vivo* data, the ethical approval under which it was acquired is documented in the data card of each contribution, and all protected health information was removed so that the data are de-identified. Accepted proposals were then provided with a submission guide, which instructed that the data must be supplied in `zea` format (Section III-C), along with a basic reconstruction script to visualize the data, and a license file. After each initial contribution, the Open-H-RF core team reviewed the submitted dataset to spot any spurious or missing metadata entries and to check whether the reconstruction script would indeed produce the expected result. Contributors were also required to provide a reference image against which the data curators could validate reproducibility. Data was collected globally from 38 institutions (Fig. 2). Contributors across academia and industry created a total of 33 contributions (Table I).

B. Task Categories

Contributions can be sorted into five general categories: probe motion compensation, tissue motion estimation, correction of inaccurate image formation assumptions, tissue characterization, and therapy. In the following, we describe each task in the context of these categories, explaining how access to channel data supports research on that task.

1) Probe motion: Conventional reconstruction treats each frame as an independent 2D slice acquired from a fixed probe position. When the probe is moved, knowledge of its pose allows measurements from different positions to be combined. **3D / tracked reconstruction (3D):** Reconstructing volumes from tracked or mechanically moved probes, or from matrix arrays. This allows organs, tumors, and fractures to be measured in 3D. Similarly, when the probe moves in-plane, every frame observes the same scatterers from a new position, and the frames can be summed in phase to synthesize an aperture that is larger than the array by itself. Phase-preserving measurements permit coherent synthetic-aperture focusing across probe positions.

2) Tissue motion: Moving tissue and flowing blood change the received signals from one transmit to the next. This motion can be a source of diagnostic information, but can also produce noise that must be filtered during reconstruction. **Motion estimation / tracking (Motion):** Estimating tissue displacement [65], for example via shear wave elastography, following structures such as muscle fascicles over time, or aneurysm wall mechanics as an indicator of rupture risk. Access to channel data adds control over focusing, enables phase-sensitive estimation of lower-velocity flows, and allows motion estimation and reconstruction to be performed jointly. **Blood flow / ultrasound localization microscopy (Flow):** Estimating blood motion with Doppler methods, or localizing and tracking microbubbles to map the microvasculature. Applications include transcranial imaging in stroke, carotid and venous hemodynamics [61], [62], and micrometer-scale cerebral microvasculature [56]. Channel data access enables optimizing beamforming, angle compounding, and tissue suppression, in addition to estimation or localization.

3) Inaccurate assumptions: Delay-and-sum beamforming relies on simplified propagation assumptions, including a known homogeneous background sound speed and the Born approximation, thereby neglecting effects such as refraction and multiple scattering. These inaccuracies result in image quality degradation, motivating the development of methods with better assumptions. **Generalized reconstruction / beamforming (Recon):** Image formation from the measured signals using delay-and-sum, adaptive, model-based, or learned methods, including tomographic reconstruction from ring and limited-view arrays. Access to channel data before summation allows different reconstruction methods to be compared on the same acquisition. Additionally, it allows computing channel-domain metrics such as lag-one coherence (LOC) for more elaborate evaluation [66]. **Aberration correction (SoS):** Correcting the defocusing caused by an incorrect assumed sound speed. Aberration from fat and muscle layers is a main cause of poor image quality in difficult-to-image patients [67]. Channel data preserve the arrival times and phase across the

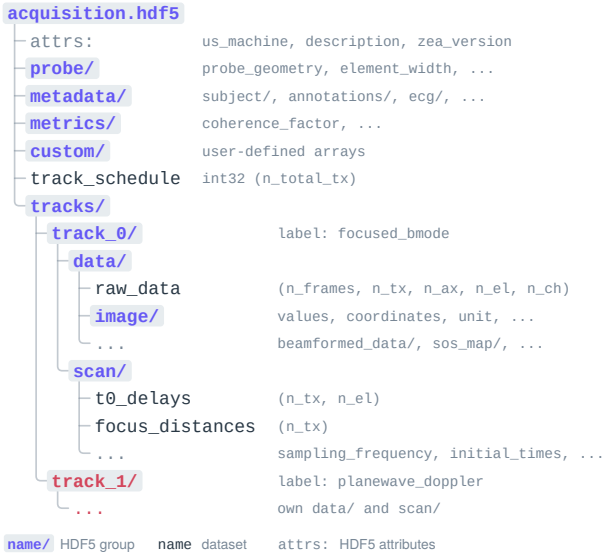

Fig. 4: Layout of a `zea` acquisition file. Channel data and parameters required to reconstruct it are kept together in a single HDF5 file. Interleaved transmit sequences are stored as separate *tracks* that share the probe and metadata of the file; single-track files place `data/` and `scan/` directly at the root. Ellipses denote further fields of the specification, which are documented in full in the `zea` docs.

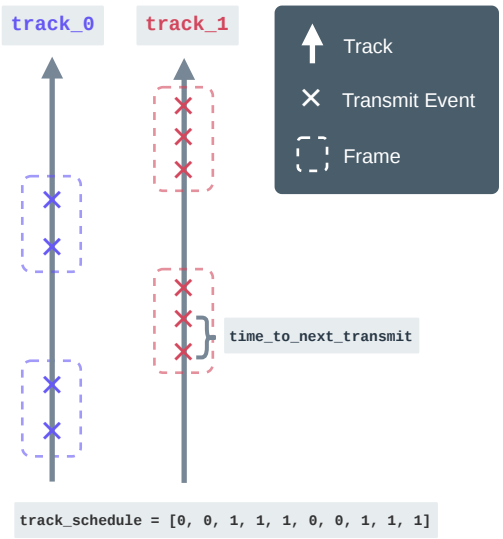

Fig. 5: Illustrative example of a `zea` file with two tracks with interleaving transmits. A global track schedule records which track each transmit belongs to, so the full acquisition timing can be reconstructed across tracks. Tracks allow different types of transmit events to be stored in the same acquisition file. Generally, each track has its own `scan` parameters and is processed by its own pipeline.

aperture needed to estimate and compensate these propagation errors. **Clutter / reverberation suppression (Clutter):** The removal of unwanted contributions such as stationary tissue, off-axis echoes, or multi-path echoes is a major area of research in ultrasound imaging. Clutter suppression is particularly relevant to cardiac and abdominal imaging, where body-wall reverberation degrades images [22], [68], and in Doppler and ULM, where tissue motion must be suppressed [24]. Aperture-domain methods exploit differences in coherence and arrival structure across elements, which are lost after summation. **Image-quality assessment (IQA):** Quantifying resolution, contrast, and artifact levels, and hence the effect of the corrections above, often requires channel data to evaluate alternative beamformers on identical measurements.

4) Tissue characterization: Beyond depicting anatomy, many tasks aim to image and interpret the physical properties of tissue. **Speed-of-sound estimation (SoS):** Estimating the speed of sound in the medium. Sound speed decreases with hepatic fat content in fatty liver disease [49] and is linked to muscle health [63]. Channel data preserve the arrival times and phase across the aperture, which allows computing phase-based measures of signal coherence such as common midpoint phase error (CMPE) [21], which sound-speed estimation methods rely on. **Quantitative ultrasound (QUS):** Estimating tissue properties such as attenuation, backscatter coefficients, or echo-envelope statistics, which allow distinguishing benign from malignant breast masses and grading liver steatosis. Channel data provides access to the frequency content of the backscattered signal, which log-compressed B-mode images do not preserve. **Segmentation / interpretation (Interp):** Identifying and/or delineating anatomical structures, or classifying findings, in measurements or images. Applications include breast lesion diagnosis, tumor volumetry, vessel-wall delineation in intravascular imaging, and standard-view recognition in echocardiography and fetal imaging. While conventional image-based methods do not rely on channel data, these contributions enable the development of direct raw-to-insight models and studies of how reconstruction choices affect downstream performance.

5) Therapy: Beyond imaging, ultrasound is used therapeutically, for example to transiently open the blood-brain barrier for targeted drug delivery [69]. **Microbubble cavitation detection (MCD):** Detecting and characterizing acoustic emissions from bubble activity during ultrasound exposure. Monitoring this bubble activity underpins the safety and efficacy of ultrasound drug delivery and blood-brain-barrier opening. Recorded waveforms with sufficient bandwidth are required, and array data further enable passive mapping of emission locations.

C. Data Organization

One of the central tasks of the Open-H-RF initiative is to compile the contributions of all teams into a single, unified format, together with the tooling needed to make the resulting corpus readily accessible, both for a casual inspector and for demanding workflows such as model training. The format builds on the `zea` file specification [28], which was extended in collaboration with the Open-H-RF initiative and is summarized in Fig. 4.

```
import zea
path = "hf://nvidia/Open-H-RF/..." # or local path
num_frames, tx_idx = ..., ...
with zea.File(path) as f:
    raw_data = f.data.raw_data[:num_frames, tx_idx]
```

Fig. 6: Streaming a portion of a `zea` file directly from the Open-H-RF corpus. Accessing only a subset of frames or transmits streams and caches only that part of the data to disk.

1) File format: Raw channel data is a digitized record of the pressure field, integrated over the surface of each element (Section II). Its samples can only be interpreted together with a description of the transducer elements that measured this field and of the transmit events that produced it. Each file therefore holds a single acquisition together with everything needed to process it, organized in a fixed hierarchy of groups (Fig. 4): a `data` group containing the raw channel data; a `scan` group containing the acquisition parameters necessary for reconstruction, such as the transmit delays, the sampling frequency, and the start time of each receive window; and a `probe` group storing the transducer geometry and properties, such as the position and width of every element. Optional groups store provenance and annotations (`metadata`), quality `metrics`, and contributor-specific arrays (`custom`).

The task categories of Section III-B place different demands on the data, yet they share the same underlying measurement: the signal each element recorded after each transmit event. The format therefore separates this measurement from task-specific reference data, which is stored in the same file. The `data` group holds the raw channel data together with derived spatial maps, each stored with its coordinates and unit, such as beamformed data, attenuation maps, images, or the SoS maps used to evaluate aberration correction. Other references, such as probe positions for tracked reconstruction, true microbubble positions for localization microscopy, or segmentations, are stored alongside the raw channel data as well.

Yet, contributions differ not only in their parameters and references, but also in how they organize an acquisition. A single acquisition may interleave several transmit sequences, for example focused transmits for B-mode imaging with plane waves for Doppler. The format represents each sequence in an acquisition as a *track*: a labeled pair of `data` and `scan` groups within the same file. The `probe`, `metadata`, and `metrics` groups are shared across tracks, as they belong to the same acquisition. A global track schedule records which track each transmit belongs to, so that the timing of the full acquisition can be reconstructed (Fig. 5). Each track can thus be processed by its own pipeline, while all the data related to a single acquisition is kept together in one file.

2) Data access: Even with all parameters stored, accessing raw channel data requires *domain knowledge*: reconstructing an image involves a chain of processing steps, such as filtering, demodulation, beamforming, and envelope detection, whose order and settings differ between contributions. We therefore asked each contributor to provide a reproducible reconstruction pipeline together with a reference image, against which we verified that the pipeline produces the intended result (Section III-A). This pipeline is stored with the contribution and specifies the exact reconstruction steps applied to its data, and each file records the version of the reconstruction library (`zea` [28]) with which it was stored. Together, these allow any recipient to use the data exactly as intended by the original authors. Moreover, the pipelines are implemented with graphics processing unit (GPU)-accelerated, differentiable operations in a machine learning framework, so that ultrasound reconstruction can be integrated directly into model training. The corpus itself is versioned as well: Open-H-RF is hosted at `https://huggingface.co/datasets/nvidia/Open-H-RF`, this paper describes `v1.0`, and any release can be directly loaded through `zea` by its tag, so follow-up work can state exactly which version of the corpus was used.

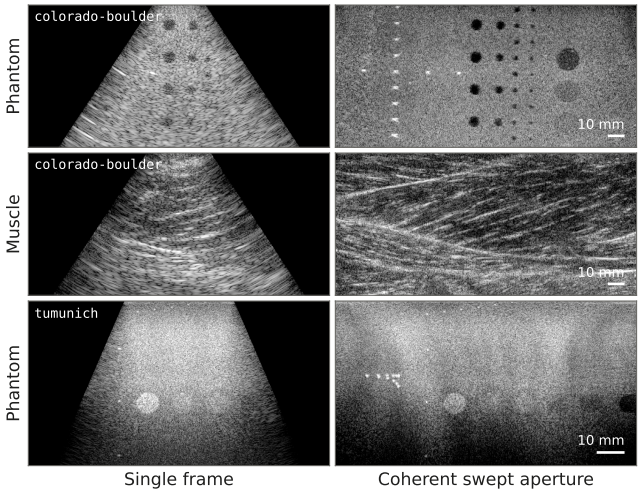

Fig. 7: Tracked SSA. The `colorado-boulder` and `tumunich` contributions demonstrate an SSA (right column) built from multiple individual acquisitions (left column).

The *data rates* associated with raw channel data pose a second barrier to access (Section II). An acquisition of a few frames quickly reaches multiple gigabytes, and the entire corpus spans 40 TB. Downloading the corpus in full is therefore impractical for most users. Instead, users can filter the corpus, for example by anatomy, probe type, or intended task, and access only the parts they need. Within each file, the raw channel data is split into compressed chunks along the transmit axis, so that selected frames or transmits can be read locally or streamed remotely without loading the full acquisition, see Fig. 6 for an example.

IV. Results

In the following, we demonstrate several of the tasks introduced above and how they differ across contributions. Then, we provide an analysis of dataset composition, quantifying its overall size, and breakdowns in terms of anatomy, probe type, and transmit scheme.

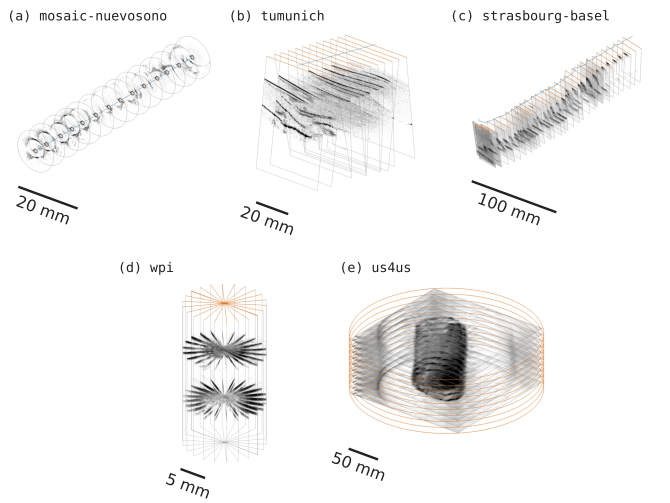

Fig. 8: Variety of 3D acquisition geometries in Open-H-RF. Each plane is one acquired frame at its recorded position, with its outline; orange marks the transducer elements (the rotating element at the catheter center in (a), the array face in (b)–(d), the ring in (e)), and the blue line is the tracked probe path. (a) `mosaic-nuevosono`: IVUS catheter pullback in an *in-vivo* porcine artery (14 of 100 cross-sections over 67 mm). (b) `tumunich`: robot-tracked elevational sweep of a thyroid phantom (12 of 1237 frames). (c) `strasbourg-basel`: robot-tracked sweep along a fractured femur phantom on a curved 276 mm path (32 of 158 frames). (d) `wpi`: 1D array rotated 181° about its axial axis over a phantom (12 of 182 frames). (e) `us4us`: ring-array USCT of an *in-vivo* forearm, 10 slices 9 mm apart; transmission-attenuation reconstructions rather than B-modes.

A. Task Examples

1) 3D / tracked reconstruction (3D): Six of the 33 contributions record where each frame was acquired, alongside the raw channel data. The position comes from a robot (`tumunich`), optical tracking (`colorado-boulder`, NDI Polaris Vega XT), a robot-mounted electromagnetic sensor (`strasbourg-basel`, trakSTAR), a motorized rotation stage (`wpi`), a catheter pullback encoder (`mosaic-nuevosono`), or a positioning system that steps the subject through a ring array (`us4us`). Tracking can turn a small acquisition into a larger 2D or 3D one.

a) Swept Synthetic Aperture: Two contributions contain tracked, in-plane movement, enabling swept synthetic aperture reconstructions (Fig. 7): `colorado-boulder` sweeps 160 mm with a 20 mm array, and `tumunich` 70 mm with a 38 mm array. When measured on the wire targets of their phantoms, coherent summation narrows the −6 dB lateral width by 3.4× and 1.8×, respectively (Table II).

TABLE II: Lateral width (−6 dB) of the wire targets under coherent and incoherent compounding.

Dataset	Wire targets	−6 dB lateral width [mm]	
		Coherent	Incoherent
`colorado-boulder`	12	0.58 ± 0.28	1.96 ± 0.36
`tumunich`	9	0.29 ± 0.05	0.51 ± 0.18

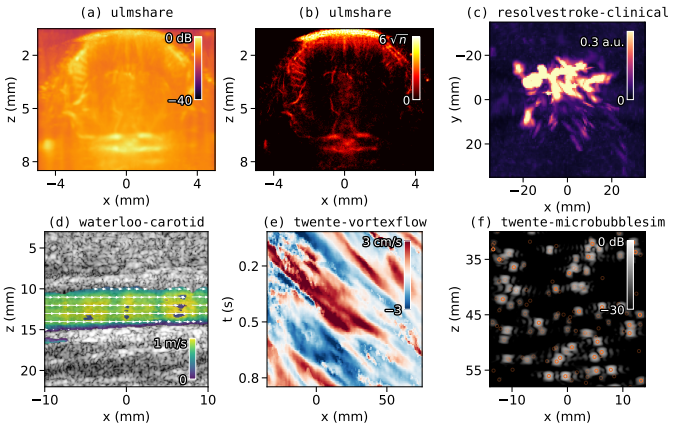

Fig. 9: Blood flow in Open-H-RF. (a), (b) `ulmshare`: mouse brain through the intact skull, from 20,000 frames: power Doppler after singular-value clutter filtering (a) and ULM track density at 10 µm (b). (c) `resolvestroke-clinical`: top-view maximum-intensity projection of the stored 3D microvascular image of a human brain. (d) `waterloo-carotid`: vector velocity in a human common carotid artery at peak systole, drawn where the power Doppler is within 9 dB of its 99.5th percentile and the speed exceeds 0.1 m/s. (e) `twente-vortexflow`, 120 V pump: fluctuating Doppler velocity along the wake of the cylinder against time; the diagonal stripes are vortices carried downstream. (f) `twente-microbubblesim`: B-mode of the short 2.5 MHz pulse, true bubble positions circled, solid within ±2 mm of the imaging plane and faint beyond.

b) 3D Acquisitions: A number of contributions enable positioning their frames in 3D, as shown in Fig. 8. In `mosaic-nuevosono`, an intravascular ultrasound (IVUS) catheter spins a single element inside the vessel and is pulled back along it, so each rotation images a cross-section centered on the transducer, and the cross-sections stack along a 67 mm pullback (a). In `tumunich`, a robot translates a linear array across its elevation, stacking parallel planes over 50 mm of a thyroid phantom (b). In `strasbourg-basel`, a robot instead follows a fractured femur phantom along a curved 276 mm path, so the image planes turn with the bone (c). In `wpi`, the probe stays in place and the array is rotated by 181° about its own axial axis, producing a fan of planes that share one axis, as used for elevational synthetic-aperture focusing (d). In `us4us`, the subject moves rather than the probe, stepping a forearm through a ring array to obtain stacked tomographic slices (e). The frame spacing alone spans more than two orders of magnitude, from 0.04 mm between frames (`tumunich`) through 0.6 mm (`mosaic-nuevosono`) and 1.75 mm (`strasbourg-basel`) to 9 mm between slices (`us4us`).

2) Blood flow / ultrasound localization microscopy: Six of the 33 contributions target flow estimation, across simulation, phantom and *in-vivo*, in 2D and 3D, see Fig. 9. `ulmshare`

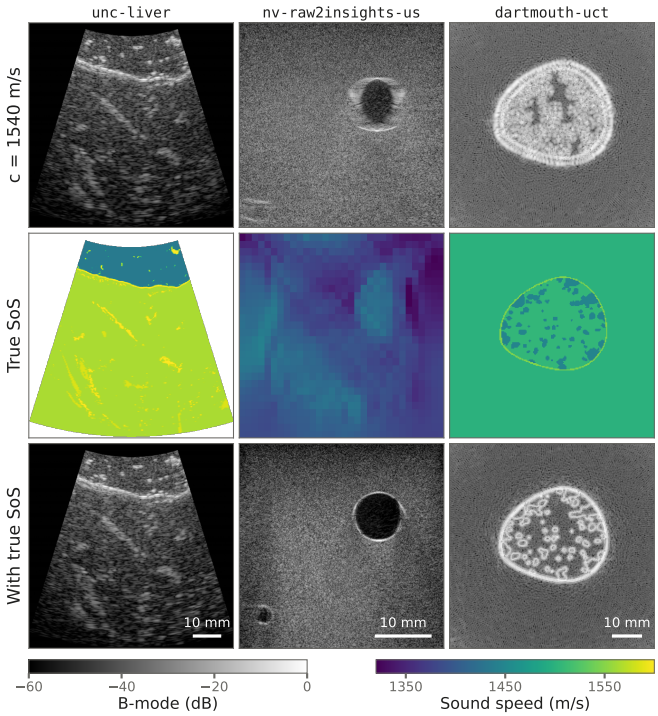

Fig. 10: SoS correction on three of the four simulations that store their SoS map. Columns: `unc-liver` abdominal wall over liver (curvilinear array), `nv-raw2insights-us` phantom with an aberrating layer and cysts (linear array), and `dartmouth-uct` breast in a ring array. Top: DAS at a uniform $1540\,\mathrm{m/s}$. Middle: stored SoS. Bottom: the same beamformer with straight-ray delays through the true map.

contains 99 contrast-enhanced acquisitions of the mouse brain through the intact skull, from 61 animals, enough to develop and train ultrasound localization microscopy (ULM) methods (Fig. 9). From $12.8\,\mathrm{s}$ of one acquisition, power Doppler resolves the vasculature (a) and ULM traces 3016 microbubble tracks (b). `resolvestroke-clinical` and `waterloo-carotid` store their own flow maps alongside the raw channel data, so new methods can be compared with theirs on identical input: 3D transcranial perfusion of the human brain (c), and vector flow in 93 carotid acquisitions from eight participants (d). Two Twente contributions add a flow phantom whose vortex street is visible in the Doppler velocity (`twente-vortexflow`, e) and simulated microbubbles with every true position stored (`twente-microbubblesim`, f).

3) Speed-of-sound estimation / aberration correction: Nine of the 33 contributions name SoS estimation or aberration correction as an intended use. Four simulations store the full SoS map that generated the data: a Fullwave abdominal wall over liver (`unc-liver`), a phantom with an aberrating layer and cysts (`nv-raw2insights-us`), and two ultrasound computed tomography (USCT) sets, a breast in a ring array (`dartmouth-uct`) and a limited-view prostate (`unc-openpros`). Two others provide measured sound speeds: `stanford-murine` images phantoms beneath 10–$15\,\mathrm{mm}$ layers of pork or chicken, whose SoS was measured by through-transmission, and `waterloo-muscle` records a through-transmission SoS of the calf alongside its *in-vivo* muscle acquisitions. The remaining three, *in-vivo* breast lesions with biopsy-based labels (`kaist-snubh-barreleye`) and thyroid (`weizmann-sampl`), and a robot-tracked fractured-femur phantom (`strasbourg-basel`), do not include any numerical reference.

Fig. 10 shows how the stored maps can be used in full-synthetic-aperture simulations to update reconstruction assumptions. Each dataset is beamformed twice with the same straight-ray delay model: once through a uniform $1540\,\mathrm{m/s}$ medium, and once through its true map. In the `nv-raw2insights-us` phantom, the medium is 100–$200\,\mathrm{m/s}$ slower than assumed, which smears the cyst into an ellipse at $1540\,\mathrm{m/s}$. In `dartmouth-uct`, the ring array sees the breast through water, so assuming $1540\,\mathrm{m/s}$ blurs the skin and merges the inclusions. The `unc-liver` abdominal wall changes little: the slow fat layer ($\approx 1440\,\mathrm{m/s}$) and the faster tissue below ($\approx 1565\,\mathrm{m/s}$) average out close to $1540\,\mathrm{m/s}$ along each path.

B. Dataset Composition

Version `v1.0` of Open-H-RF comprises 33 contributions, consisting of 19,332 acquisitions, 25,092,732 frames, and 241,078,230 transmit events. The acquisitions cover a range of imaging media, including both *in-vivo* and phantom targets, as well as a large number of *in-silico* simulations. Fig. 11 categorizes these acquisitions in terms of imaging medium and corresponding subtypes. Fig. 11 shows that each medium contains a wide range of subtypes, encompassing many of the common ultrasound imaging targets. Fig. 12 maps this categorization onto the contributions, identifying, for example, that 4 contributions image the carotid artery.

The contributions also cover a wide variety of probe types and transmit schemes, illustrated by Figs. 13 and 14, respectively. In terms of probe type, the contributions contain linear, curved, ring, phased-array, and IVUS probes, with a range of transducer element counts and center frequencies. Common transmit schemes include plane wave, synthetic aperture, focused, diverging wave, coded, and *mixed*, which refers to cases in which multiple transmit scheme types are used in a single track.

C. Lag-One Coherence (LOC)

The LOC is the normalized cross-correlation between the delayed signals of neighboring array elements [66]. This can give a local, target-free estimate of the achievable contrast and contrast-to-noise ratio (CNR). With Open-H-RF, we supply data with a large LOC spread, which allows evaluating methods under realistic noise and clutter conditions. We plot the LOC for a subset of the collected *in-vivo* anatomies in Fig. 15, ordered by the number of frames. Every dot in the plot denotes an acquisition that was done with full-aperture focused imaging, single-element transmits, or line-scanning, with at least three frames available. The data comes from the `siemens-healthineers`,

`tue-carotid`, `oslo-carotid`, `weizmann-sampl`, `vanderbilt` (both fundamental and harmonic), `technion-cardiac`, `tue-cardiac`, `oslo-cardiac` and `technion-bladder` contributions.

V. Discussion

Raw channel data is difficult and expensive to acquire. It requires research scanners that expose raw channel data, which are costly and not available in every lab. In addition, raw channel data is often considered proprietary, as it reveals details of

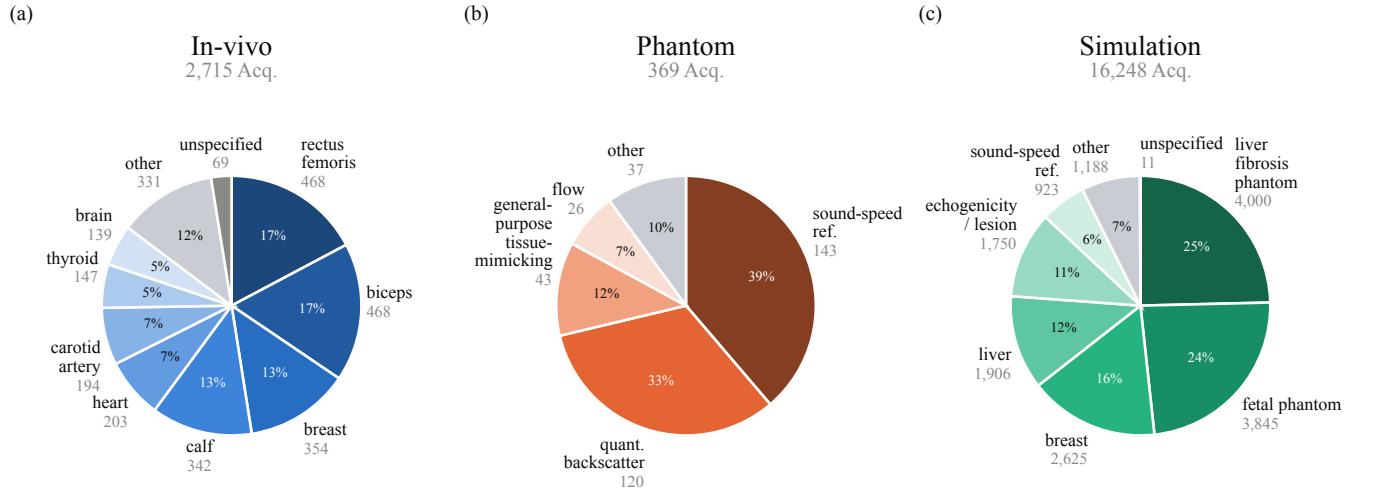

Fig. 11: All 19,332 acquisitions divided by medium into (a) *In-vivo*, (b) *Phantom*, and (c) *Simulation*, and further divided into anatomical / phantom subtypes.

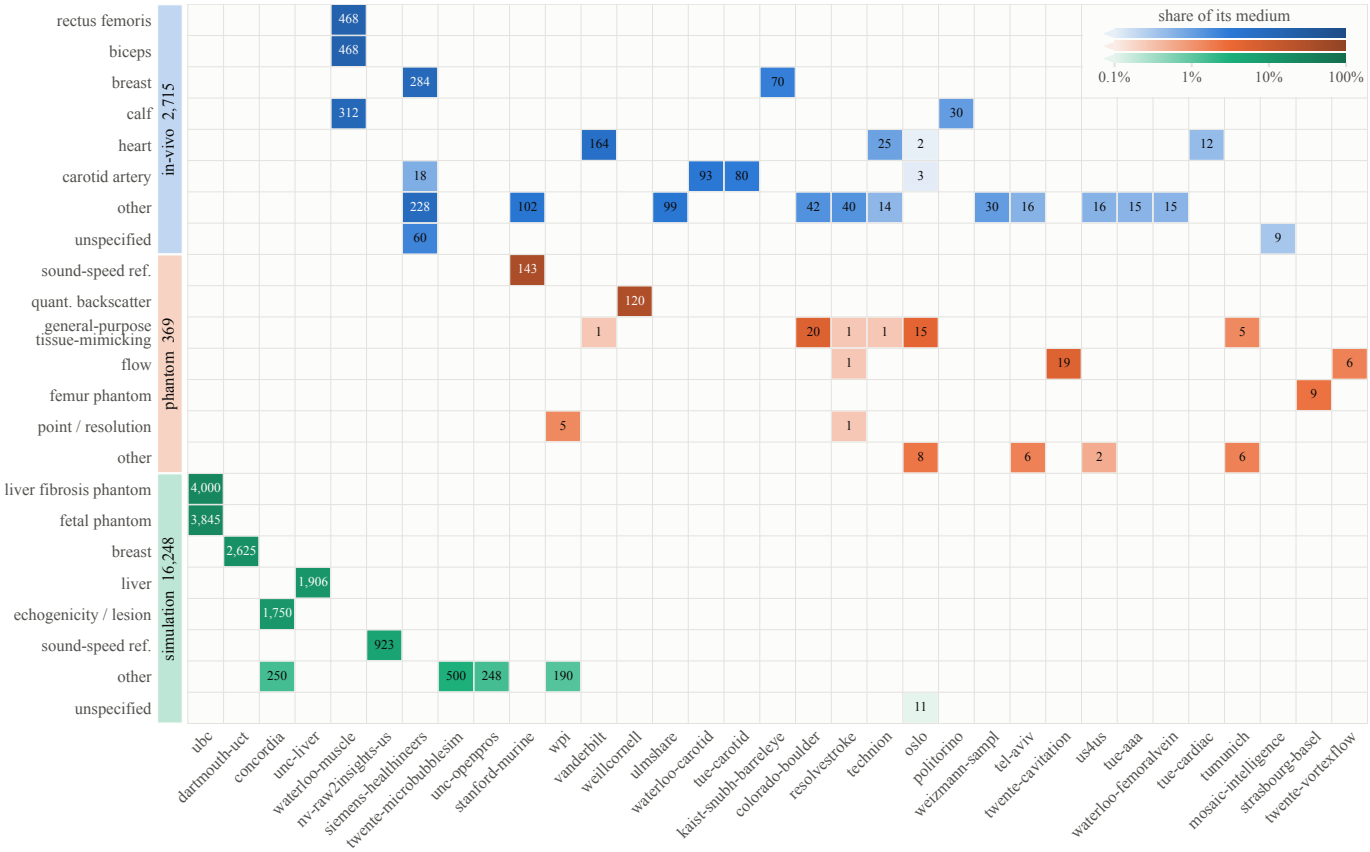

Fig. 12: A matrix indicating which anatomy / phantom subtypes are contained by each contribution to Open-H-RF. The brightness of each color indicates how much of a medium's share is contained by that contribution. Additionally, the number of acquisitions is displayed in each square.

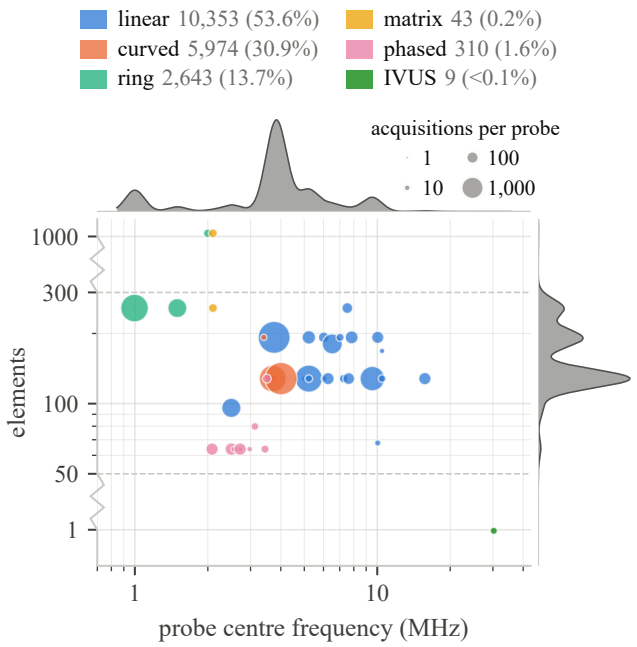

Fig. 13: Acquisitions divided by probe type, with probe center frequencies plotted against element counts.

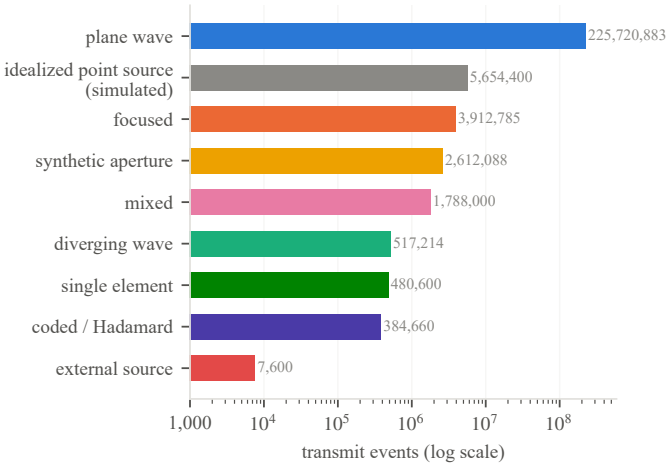

Fig. 14: Transmit events sorted by transmit scheme. *Mixed* refers to acquisitions that interleave multiple transmit schemes in one track; *single element* refers to mechanically scanned single-element transmits; *external source* marks tracks in which a separate probe is used for insonification.

the transmit sequences and system design. As a result, research on raw channel data has largely been confined to groups with their own acquisition setup, with methods developed and evaluated on in-house data rather than across institutions. Open-H-RF offers a shared, permissively licensed alternative, contributed by both academic and industrial partners.

The contents of Open-H-RF may be broken down in a number of ways, providing different views on its composition. For example, in terms of acquisitions, the *simulation* medium dominates, covering 84 % of all acquisitions. This could be explained by the relative ease of generating novel simulated subjects. If we consider the number of frames, however, *in-vivo* dominates, with 98 % of all frames being *in-vivo*, and 85 % of those *in-vivo* frames coming from `ulmshare`, which contains long acquisitions imaged at a high frame rate. Consecutive frames from such acquisitions are highly correlated, so frame counts overstate how much independent information a contribution adds; for anatomical and physiological diversity, the number of distinct subjects, animals, or phantoms is a better proxy for dataset size. The unevenness in representation of different imaging media or anatomical types therefore depends heavily on the level of abstraction we use to count the data: *acquisition*, *frame*, or *transmit event*, and should be taken into account when sampling across contributions for benchmarking or model training.

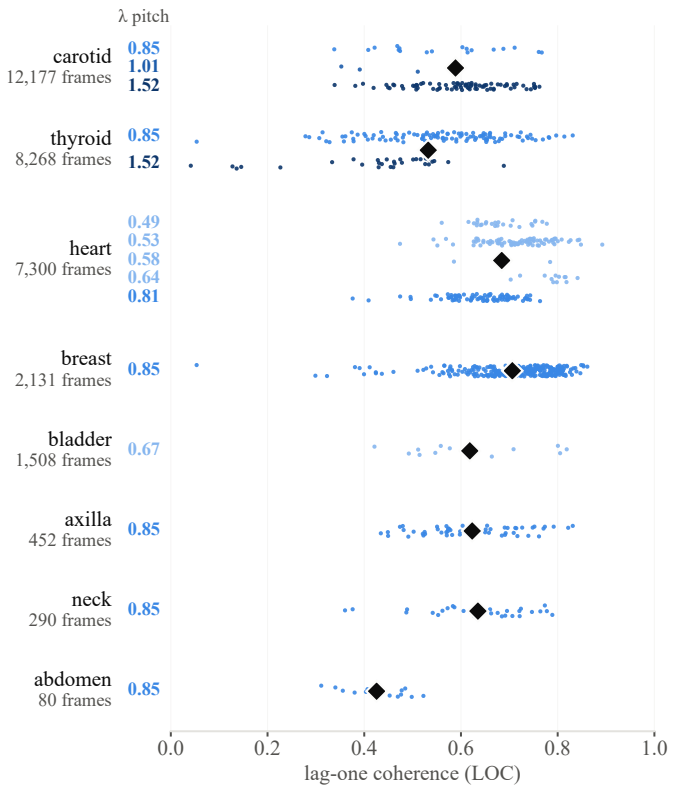

Fig. 15: Lag-one coherence per anatomy across the focused *in-vivo* datasets. Every dot denotes an acquisition available in Open-H-RF, the black diamond shows the mean for a given anatomy. Transducer pitch is expressed as a fraction of transmit wavelength λ and separates the acquisitions vertically per anatomy.

Open-H-RF is designed to support a number of tasks, as presented in Table I. Table III shows the *task coverage*, measured by the number of contributions supporting each task, and each task category. There is broad coverage of most of the task categories, with *Generalized reconstruction / beamforming* being supported by most of the contributions (27), and all other tasks being covered by at least 5 contributions. The exception is the *Therapy* category, containing only one task, *Microbubble cavitation detection*, which is covered only by the `twente-cavitation` dataset. This insight can serve to guide future data collection, to improve the representation and coverage of therapeutic tasks.

TABLE III: Number of Open-H-RF contributions declaring each task, grouped by task category.

Category	Task	Contributions
Probe motion	3D / tracked reconstruction	6
Tissue motion	Motion estimation / tracking	10
	Blood flow / ultrasound localization microscopy	6
Inaccurate assumptions	Generalized reconstruction / beamforming	27
	Aberration correction	9†
	Clutter / reverberation suppression	6
	Image-quality assessment	7
Tissue characterization	Speed-of-sound estimation	9†
	Quantitative ultrasound	5
	Segmentation / interpretation	8
Therapy	Microbubble cavitation detection	1

†Aberration correction and speed-of-sound estimation are a single annotation category; the same 9 contributions are counted under both.

A key capability enabled by Open-H-RF is the evaluation of task performance across scanners, probes, and institutions. This makes it possible to study how they generalize and gives a common basis for benchmarking. In addition to task-specific methods, Open-H-RF opens new possibilities in developing methodology for imaging efficiency and quality enhancement that generalizes across tasks. For example, *Compressed Sensing* [70], [71] may be used to recover useful information from fewer measurements, or select transmissions according to the information needed for reconstruction or a downstream task. An interesting avenue for research in this direction is to compare the information gain or quality improvement offered by different transmit schemes, such as focused, plane-wave, diverging-wave, and coded transmissions. This is supported by a number of contributions (e.g., `oslo`, `stanford-murine`, `tue-cardiac`, `tue-carotid`, `twente-microbubblesim`), which image the same target using different transmit schemes, stored as parallel `zea` tracks.

The large volume and variety of raw channel data provided by Open-H-RF opens up new possibilities in developing data-driven solutions that operate on raw channel data. Interesting opportunities lie in both supervised and unsupervised learning tasks. On the supervised side, training regression and classification models directly on raw channel data could circumvent errors and information loss introduced by the beamforming step [72]. Research efforts in this direction could explore methods to standardize data across anatomies, probes, etc., as well as neural network architectures that are parameter-efficient and extract useful features. Unsupervised methods, e.g., self-supervised learning, could be used to train generative models of raw channel data, capturing its structure and variation. Potential applications of raw channel data generation include image quality enhancement (e.g., dehazing [22]) and raw channel data compression [73]. Another application of such a model is *cognitive ultrasound* [74], [75], which harnesses data-driven generative models to achieve information-seeking, patient-adaptive ultrasound acquisitions, and would benefit from a generative prior over the raw channel data itself rather than over the reconstructed image [76]. A natural next step is a foundation model for raw channel data. Such a model could be pretrained with self-supervised objectives on the full corpus and conditioned on the probe geometry and transmit parameters stored in every file, allowing one model to handle different probes and transmit schemes.

The dataset also lowers the barrier for researchers outside the ultrasound community. Each acquisition is stored as a self-contained file with its reconstruction pipeline and can be streamed from the Hugging Face Hub, which makes working with raw channel data considerably easier without extensive ultrasound expertise. This makes it an interesting and challenging benchmark for the computational imaging and machine learning communities. We see Open-H-RF as a starting point rather than a finished resource. The format and submission process are designed to accept new contributions, and we invite the community to extend the corpus, define benchmarks on it, and share the models they train on it.

VI. Conclusion

We present Open-H-RF: the largest openly licensed collection of raw medical ultrasound data to date. The dataset is contributed by 38 institutions from both academia and industry and comprises approximately 40 TB of compressed raw channel data from 19,332 ultrasound scans. It covers a diverse set of tasks, ranging from general reconstruction to motion and blood flow estimation, as well as tracked ultrasound. Throughout, we place a strong focus on accessibility for researchers who want to work with this kind of data, as reflected in the permissive license and the open-source ecosystem in which the dataset is hosted. We hope that Open-H-RF will enable researchers to develop methods that require raw channel data at scale, and unlock rigorous evaluation and discovery in this traditionally restricted research space. We hope to make ultrasound more accessible to researchers and build bridges to other communities, e.g., machine learning and computational imaging communities. One direct and promising direction is the development of an ultrasound foundation model trained on raw channel data and conditioned on the provided metadata. Ultimately, we hope Open-H-RF serves as a shared starting point for the community to build upon, moving ultrasound research beyond the beamformed image and towards methods that exploit the full information content of the raw signal.

Acknowledgment

Tristan S.W. Stevens Data Curation, File Format Design, Software, Writing – Original Draft. **Simon W. Penninga** Data Curation, Writing – Original Draft, Visualization, Marketing. **Oisín Nolan** File Format Design, Software, Writing – Original Draft, Visualization. **Wessel L. van Nierop** File Format Design, Software, Writing – Original Draft. **Felix Duelmer** Data Curation, Software, Writing – Original Draft. **Ben N. Frey** Data Curation. **Louis van Harten** Website Creation, Visualization. **Ruud J.G. van Sloun** Conceptualization, Supervision. **Jeremy Dahl** Conceptualization, Supervision. **Walter Simson** Conceptualization, Supervision, Data Curation, Writing – Review & Editing.

All other authors are dataset contributors from the following institutions: *Eindhoven University of Technology, Technical University of Munich, Stanford University, Mosaic Intelligence, Technion – Israel Institute of Technology, Concordia University, Dartmouth College, IHU Strasbourg, ImFusion, Johns Hopkins University, KAIST, Barreleye Inc., National Institutes of Health, NuevoSono, Politecnico di Torino, Ruhr University Bochum, Polytechnique Montréal, Mila, Resolve Stroke, Seoul National University Bundang Hospital, Siemens Healthineers, School of Biomedical Engineering Tel-Aviv University, The University of British Columbia, University of Basel, University of Colorado Boulder, University of North Carolina at Chapel Hill, North Carolina State University, Iowa State University, QT Imaging Inc., University of Oslo, University of Rochester, University of Twente, University of Waterloo, us4us Ltd., Vanderbilt University, Weill Cornell Medicine, SAMPL LAB Weizmann Institute of Science, Worcester Polytechnic Institute, NVIDIA*.

References

[1] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2009, pp. 248–255.

[2] F. Knoll et al., "fastMRI: A Publicly Available Raw k-Space and DICOM Dataset of Knee Images for Accelerated MR Image Reconstruction Using Machine Learning," *Radiology: Artificial Intelligence*, vol. 2, no. 1, e190007, Jan. 1, 2020.

[3] A. Sriram et al., "End-to-End Variational Networks for Accelerated MRI Reconstruction," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, Cham: Springer International Publishing, 2020, pp. 64–73.

[4] L. Pineda, S. Basu, A. Romero, R. Calandra, and M. Drozdzal, "Active MR k-space Sampling with Reinforcement Learning," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, Cham: Springer International Publishing, 2020, pp. 23–33.

[5] R. Bommasani et al., *On the Opportunities and Risks of Foundation Models*, Aug. 2021. arXiv: 2108.07258 [cs.LG].

[6] S. Pai et al., *Vision Foundation Models for Computed Tomography*, Jan. 2025. arXiv: 2501.09001 [cs.CV].

[7] D. Tak et al., "A generalizable foundation model for analysis of human brain MRI," *Nature Neuroscience*, vol. 29, no. 4, pp. 945–956, Apr. 2026.

[8] J. Jiao et al., "USFM: A universal ultrasound foundation model generalized to tasks and organs towards label efficient image analysis," *Medical Image Analysis*, vol. 96, p. 103 202, Aug. 2024.

[9] M. Christensen, M. Vukadinovic, N. Yuan, and D. Ouyang, "Vision–language foundation model for echocardiogram interpretation," *Nature Medicine*, vol. 30, no. 5, pp. 1481–1488, May 2024.

[10] Y. Megahed, R. Ducharme, A. Erman, M. C. Walker, S. Hawken, and A. D. Chan, "USF-MAE: Ultrasound Self-Supervised Foundation Model with Masked Autoencoding," *Biomedical Signal Processing and Control*, vol. 122, p. 110 313, Aug. 2026.

[11] D. Ouyang et al., "Video-based AI for beat-to-beat assessment of cardiac function," *Nature*, vol. 580, no. 7802, pp. 252–256, Apr. 9, 2020.

[12] G. Duffy et al., "High-Throughput Precision Phenotyping of Left Ventricular Hypertrophy With Cardiovascular Deep Learning," *JAMA Cardiology*, vol. 7, no. 4, pp. 386–395, Apr. 1, 2022.

[13] S. Wang, J. A. Hossack, and A. L. Klibanov, "From Anatomy to Functional and Molecular Biomarker Imaging and Therapy: Ultrasound Is Safe, Ultrafast, Portable, and Inexpensive," *Investigative Radiology*, vol. 55, no. 9, pp. 559–572, Sep. 2020.

[14] K. A. Fleming et al., "The Lancet Commission on diagnostics: transforming access to diagnostics," *The Lancet*, vol. 398, no. 10315, pp. 1997–2050, Nov. 2021.

[15] M. Tanter and M. Fink, "Ultrafast imaging in biomedical ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, no. 1, pp. 102–119, Jan. 2014.

[16] C. Errico et al., "Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging," *Nature*, vol. 527, no. 7579, pp. 499–502, Nov. 2015.

[17] R. J. G. van Sloun, R. Cohen, and Y. C. Eldar, "Deep Learning in Ultrasound Imaging," *Proceedings of the IEEE*, vol. 108, no. 1, pp. 11–29, Jan. 2020.

[18] B. Luijten et al., "Adaptive Ultrasound Beamforming Using Deep Learning," *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 3967–3978, Dec. 2020.

[19] M. Feigin, D. Freedman, and B. W. Anthony, "A Deep Learning Framework for Single-Sided Sound Speed Inversion in Medical Ultrasound," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 4, pp. 1142–1151, Apr. 2020.

[20] W. Simson, L. Zhuang, S. J. Sanabria, N. Antil, J. J. Dahl, and D. Hyun, "Differentiable Beamforming for Ultrasound Autofocusing," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*, Cham: Springer Nature Switzerland, 2023, pp. 428–437.

[21] W. Simson, L. Zhuang, B. N. Frey, S. J. Sanabria, J. J. Dahl, and D. Hyun, "Ultrasound Autofocusing: Common Midpoint Phase Error Optimization via Differentiable Beamforming," *IEEE Transactions on Medical Imaging*, vol. 45, no. 2, pp. 681–692, Feb. 2026.

[22] T. S. W. Stevens, F. C. Meral, J. Yu, I. Z. Apostolakis, J.-L. Robert, and R. J. G. van Sloun, "Dehazing Ultrasound Using Diffusion Models," *IEEE Transactions on Medical Imaging*, vol. 43, no. 10, pp. 3546–3558, Oct. 2024.

[23] B. Heiles, A. Chavignon, V. Hingot, P. Lopez, E. Teston, and O. Couture, "Performance benchmarking of microbubble-localization algorithms for ultrasound

localization microscopy," *Nature Biomedical Engineering*, vol. 6, no. 5, pp. 605–616, Feb. 17, 2022.

[24] C. Demené et al., "Spatiotemporal Clutter Filtering of Ultrafast Ultrasound Data Highly Increases Doppler and fUltrasound Sensitivity," *IEEE Transactions on Medical Imaging*, vol. 34, no. 11, pp. 2271–2285, Nov. 2015.

[25] O. Solomon et al., "Deep Unfolded Robust PCA With Application to Clutter Suppression in Ultrasound," *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 1051–1063, Apr. 2020.

[26] O. Bernard et al., "The Ultrasound File Format (UFF) - First Draft," in *2018 IEEE International Ultrasonics Symposium (IUS)*, IEEE, Oct. 2018, pp. 1–4.

[27] A. Rodriguez-Molares et al., "The UltraSound ToolBox," in *2017 IEEE International Ultrasonics Symposium (IUS)*, IEEE, Sep. 2017, pp. 1–4.

[28] T. S. W. Stevens et al., "zea: A Toolbox for Cognitive Ultrasound Imaging," *Journal of Open Source Software*, vol. 11, no. 121, p. 9881, May 2, 2026.

[29] C. Guan, A. P. Rockhill, M. Sode, and G. Pinton, "mach: ultrafast ultrasound beamforming," *Journal of Medical Imaging*, vol. 13, no. 6, p. 062 203, Apr. 6, 2026.

[30] H. Liebgott, A. Rodriguez-Molares, F. Cervenansky, J. A. Jensen, and O. Bernard, "Plane-Wave Imaging Challenge in Medical Ultrasound," in *2016 IEEE International Ultrasonics Symposium (IUS)*, IEEE, Sep. 2016, pp. 1–4.

[31] D. Hyun et al., "Deep Learning for Ultrasound Image Formation: CUBDL Evaluation Framework and Open Datasets," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 68, no. 12, pp. 3466–3483, Dec. 2021.

[32] M. A. L. Bell, J. Huang, D. Hyun, Y. C. Eldar, R. van Sloun, and M. Mischi, "Challenge on Ultrasound Beamforming with Deep Learning (CUBDL)," in *2020 IEEE International Ultrasonics Symposium (IUS)*, IEEE, Sep. 7, 2020, pp. 1–5.

[33] B. Rauby et al., *ULMShare: A Large-Scale In Vivo Ultrasound Localization Microscopy Dataset for Microvascular Imaging*, Jun. 2026. arXiv: 2606.07851 [eess.IV].

[34] Open X-Embodiment Collaboration, "Open X-Embodiment: Robotic Learning Datasets and RT-X Models," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, May 2024, pp. 6892–6903.

[35] Open-H-Embodiment Consortium, *Open-H-Embodiment: A Large-Scale Dataset for Enabling Foundation Models in Medical Robotics*, Apr. 2026. arXiv: 2604.21017 [cs.RO].

[36] J. A. Jensen, "Medical Ultrasound Imaging," *Progress in Biophysics and Molecular Biology*, vol. 93, no. 1-3, pp. 153–165, Jan. 2007.

[37] V. Perrot, M. Polichetti, F. Varray, and D. Garcia, "So you think you can DAS? A viewpoint on delay-and-sum beamforming," *Ultrasonics*, vol. 111, p. 106 309, Mar. 2021.

[38] J. Yu, H. Yoon, Y. M. Khalifa, and S. Y. Emelianov, "Design of a volumetric imaging sequence using a vantage-256 ultrasound research platform multiplexed with a 1024-element fully sampled matrix array," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 67, no. 2, pp. 248–257, 2019.

[39] C. Giangrossi et al., "Requirements and Hardware Limitations of High-Frame-Rate 3-D Ultrasound Imaging Systems," *Applied Sciences*, vol. 12, no. 13, p. 6562, Jun. 28, 2022.

[40] O. Nolan, W. L. van Nierop, L. D. van Harten, T. S. W. Stevens, and R. J. G. van Sloun, "Task-Based Adaptive Transmit Beamforming for Efficient Ultrasound Quantification," in *ICASSP 2026 - 2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2026, pp. 7046–7050.

[41] M. Sharifzadeh, S. Goudarzi, A. Tang, H. Benali, and H. Rivaz, "Mitigating Aberration-Induced Noise: A deep learning-based Aberration-to-Aberration approach," *IEEE Transactions on Medical Imaging*, vol. 43, no. 12, pp. 4380–4392, 2024.

[42] A. Badano et al., "Evaluation of Digital Breast Tomosynthesis as Replacement of Full-Field Digital Mammography Using an In Silico Imaging Trial," *JAMA Network Open*, vol. 1, no. 7, e185474, 2018.

[43] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of Biomedical Optics*, vol. 15, no. 2, p. 021 314, 2010.

[44] NVIDIA Corporation, *NV-Raw2Insights-US*, https://huggingface.co/datasets/nvidia/NV-Raw2Insights-US, 2026.

[45] O. M. H. Rindal, *USTB datasets*, Version v5, 2019. DOI: 10.5281/zenodo.20261898

[46] E. Cesti et al., "Ultrasound-based methods to track skeletal muscle architecture in dynamic tasks: a comparative study," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2026.

[47] A. V. Telichko et al., *Pulse-Echo Ultrasound Murine In Vivo Acquisitions*, 2025. DOI: 10.25452/figshare.plus.28291985.v1

[48] R. Ali et al., *Pulse-Echo Ultrasound Sound Speed Phantom and Meat Aberrating Layer Acquisitions*, 2025. DOI: 10.25452/figshare.plus.28291988.v1

[49] A. V. Telichko et al., "Noninvasive estimation of local speed of sound by pulse-echo ultrasound in a rat model of nonalcoholic fatty liver," *Physics in Medicine & Biology*, vol. 67, no. 1, p. 015 007, Jan. 2022.

[50] R. Ali et al., "Local Sound Speed Estimation for Pulse-Echo Ultrasound in Layered Media," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 2, pp. 500–511, Feb. 2022.

[51] T. Grutman, M. Bismuth, B. Glickstein, and T. Ilovitsh, "Implicit neural representation for scalable 3D reconstruction from sparse ultrasound images," *npj Acoustics*, vol. 1, no. 1, p. 14, Aug. 2025.

[52] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment anything in medical images," *Nature Communications*, vol. 15, no. 1, p. 654, Jan. 2024.

[53] S. W. Penninga and R. J. G. van Sloun, *Deep Bayesian REFoCUS*, Oct. 2026. arXiv: 2610.03419 [cs.LG].

[54] R. Zorgdrager, N. Blanken, J. M. Wolterink, M. Versluis, and G. Lajoinie, "Waveform-Specific Performance of Deep Learning-Based Super-Resolution for Ultrasound Contrast Imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 72, no. 4, pp. 427–439, 2025.

[55] University of British Columbia Open-H-RF contributors, *UBC Module C: Synthetic Channel Proxy and Multimodal Fetal Phantom Ultrasound*, Open-H-RF, 2026.

[56] B. Rauby et al., *ULMShare: A Large-Scale In Vivo Ultrasound Localization Microscopy Dataset for Microvascular Imaging*, 2026. DOI: 10.20383/103.01550

[57] L. Zhuang et al., "Labeled Numerical Phantom of Abdominal Wall for Wave-Physics-Based Ultrasound Imaging: Applications to Image Reconstruction," *IEEE Transactions on Ultrasonics*, vol. 73, no. 1, pp. 24–34, 2026.

[58] M. Sode and G. Pinton, "A fullwave model of the nonlinear wave equation with multiple relaxations and relaxing perfectly matched layers for high-order numerical finite difference solutions," *Physics in Medicine & Biology*, vol. 71, no. 17, p. 175 043, 2026.

[59] M. Sode and G. Pinton, *Spatially heterogeneous power-law attenuation with multiple relaxation mechanisms for ultrasound modeling*, Jun. 2026. arXiv: 2606.11103 [physics.med-ph].

[60] H. Wang et al., "OpenPros: A Large-Scale Dataset for Limited View Prostate Ultrasound Computed Tomography," in *International Conference on Learning Representations*, C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust, Eds., vol. 2026, 2026, pp. 113 741–113 759.

[61] Y. H. Hsu, "High-frame-rate ultrasound characterization of carotid pulse waves to assess cerebrovascular resistance," Ph.D. dissertation, University of Waterloo, 2025.

[62] J. N. Cohen et al., "The influence of aging on graded peripheral venous return and complex blood flow features through human veins," *GeroScience*, 2026.

[63] D. Xiao, P. De la Torre, M. Saif El Nasr, A. J. Chee, M. Mourtzakis, and A. C. Yu, "Live Pulse-Echo Speed-of-Sound Estimation for Quality Assessment of Large Muscles in Humans," *Ultrasound in Medicine & Biology*, vol. 51, no. 11, pp. 1925–1935, 2025.

[64] R. Murakami, Y. Wang, Y. Tang, R. Tsumura, G. S. Fischer, and H. K. Zhang, "Elevational Synthetic Aperture Focusing for Rotated Array-Based Three-Dimensional Ultrasound Imaging," *IEEE Access*, vol. 13, pp. 45 458–45 467, 2025.

[65] M. Ashikuzzaman, A. Héroux, A. Tang, G. Cloutier, and H. Rivaz, "Displacement tracking techniques in ultrasound elastography: From cross correlation to deep learning," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 71, no. 7, pp. 842–871, 2024.

[66] W. Long, N. Bottenus, and G. E. Trahey, "Lag-One Coherence as a Metric for Ultrasonic Image Quality," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, no. 10, pp. 1768–1780, Oct. 2018.

[67] R. Ali et al., "Aberration correction in diagnostic ultrasound: A review of the prior field and current directions," *Zeitschrift für Medizinische Physik*, vol. 33, no. 3, pp. 267–291, 2023.

[68] A. Fatemi, E. A. R. Berg, and A. Rodriguez-Molares, "Studying the origin of reverberation clutter in echocardiography: In vitro experiments and In vivo demonstrations," *Ultrasound in Medicine & Biology*, vol. 45, no. 7, pp. 1799–1813, 2019.

[69] K. Hynynen, N. McDannold, N. Vykhodtseva, and F. A. Jolesz, "Noninvasive MR imaging–guided focal opening of the blood-brain barrier in rabbits," *Radiology*, vol. 220, no. 3, pp. 640–646, 2001.

[70] M. Yousufi et al., "Application of compressive sensing to ultrasound images: A review," *BioMed research international*, vol. 2019, no. 1, p. 7 861 651, 2019.

[71] I. A. Huijben, B. S. Veeling, K. Janse, M. Mischi, and R. J. Van Sloun, "Learning sub-sampling and signal recovery with applications in ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 3955–3966, 2020.

[72] C. K. Abbey, N. Q. Nguyen, and M. F. Insana, "Optimal beamforming in ultrasound using the ideal observer," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 57, no. 8, pp. 1782–1796, 2010.

[73] B. Federici, M. Mischi, and R. J. G. van Sloun, "Variational Ultrasound Channel Data Compression," *IEEE Transactions on Ultrasonics*, vol. 73, no. 8, pp. 927–938, Aug. 2026.

[74] R. J. G. van Sloun, "Active Inference and Deep Generative Modeling for Cognitive Ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 71, no. 11, pp. 1478–1490, Nov. 2024.

[75] W. L. van Nierop, O. Nolan, T. S. W. Stevens, and R. J. G. van Sloun, "Patient-Adaptive Echocardiography Using Cognitive Ultrasound," *IEEE Transactions on Medical Imaging*, vol. 45, no. 7, pp. 4034–4046, Jul. 2026.

[76] T. S. W. Stevens, J. Overdevest, O. Nolan, W. L. van Nierop, R. J. G. van Sloun, and Y. C. Eldar, "Deep Generative Models for Bayesian Inference on High-Rate Sensor Data: Applications in Automotive Radar and Medical Imaging," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 383, no. 2299, p. 20 240 327, Jun. 19, 2025.