AirSAS: Controlled Dataset Generation for Physics-Informed Machine Learning

Benjamin Cowen
bmc6220@psu.edu

J. Daniel Park
jdp971@psu.edu

Thomas E. Blanford
teb217@psu.edu

Geoff Goehle
gjg5316@psu.edu

Daniel C. Brown
dcb19@psu.edu

Applied Research Laboratory
The Pennsylvania State University
State College, PA

Abstract

Synthetic aperture sonar (SAS) is an underwater remote sensing technique for applications such as seafloor characterization and object detection. However, underwater SAS datasets are both extremely expensive to collect and difficult to control and repeat. We propose an in-air SAS measurement apparatus (AirSAS) made from commercial off-the-shelf laboratory equipment to generate controlled, repeatable datasets. AirSAS is both flexible and sufficiently delicate to capture the complex acoustic phenomena inherent in SAS measurements. The system allows us to physically control the differences between classes of interest, and observe acoustic phenomenology that is rare or expensive to collect underwater. Accordingly, we can measure and tune which acoustic phenomena deep learning models are sensitive to. AirSAS can generate both circular and linear track collections. The first iteration of the AirSAS dataset is currently under curation.

Synthetic aperature sonar (SAS) generates high resolution imagery by coherently combining time series collected by a moving sensor array [1] and is a central platform for sonar object detection and classification problems. It is believed that particular acoustic scattering phenomena associated with objects, such as resonance and multiple reflections, are important for detection and classification with neural networks [2–5]. However, observations of relevant objects are rare in field data and image quality can be strongly impacted by environmental disturbances[6, 7]. Subject matter expertise is often necessary to label acoustic images, and even still the associated labels frequently suffer due to inexact knowledge of seafloor composition or object locations. The ubiquitous nature of these complications exacerbate the already-difficult problem of understanding the strengths and weaknesses of a given machine learning model in terms of interpretable physical phenomena.

Tightly controlled experimental data sets are necessary for the detailed evaluation of models trained on SAS data. AirSAS was developed to generate relevant object scattering datasets more quickly and inexpensively than in underwater surveys or tank experiments [8–11]. The system is capable of both linear and circular collection geometries (mimicking common SAS survey patterns in fielded applications) and captures key acoustic phenomenology in controlled, repeatable datasets. The system uses commercial off-the-shelf hardware and is also equipped with optical and infrared depth cameras both at the acoustic sensor position and above the imaging stage, allowing for multi-modal analysis.

A key capability AirSAS enables is measurement of the physical differences between classes. In a prototype AirSAS dataset, 16 cylindrical objects with varying physical properties were imaged using

35th Conference on Neural Information Processing Systems (NeurIPS 2021), Sydney, Australia.

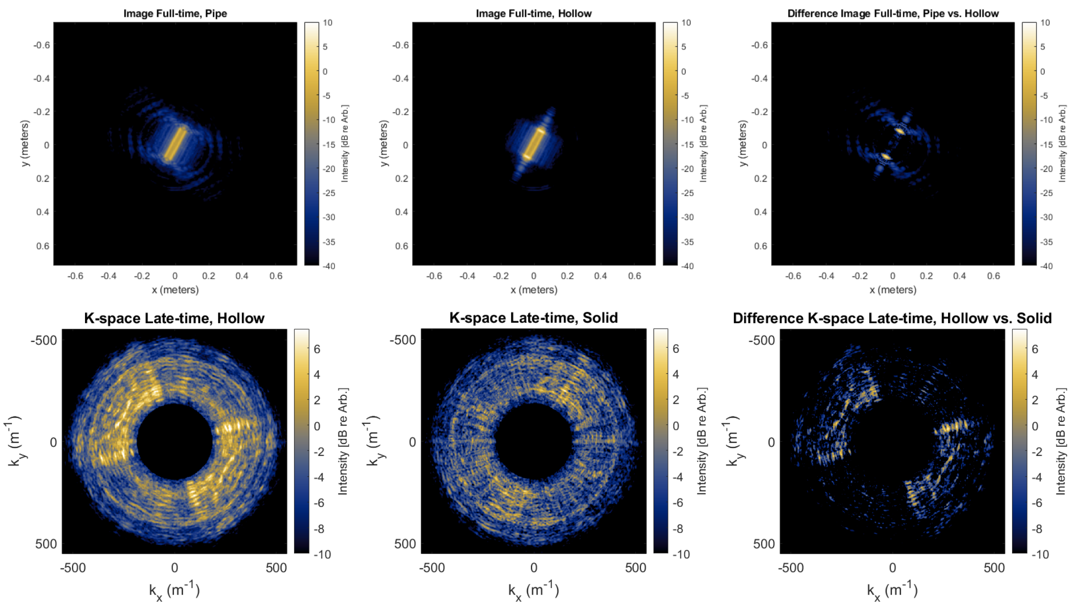

Figure 1: AirSAS image and spatial wavenumber (k-space) representation comparisons of object classes. The top row compares Pipe and Hollow in the image domain, and bottom row compares Solid and Hollow in k-space. Controlled measurements and careful alignment post-processing allow us to compute the difference signals (right column), highlighting key acoustic phenomenology that are expected be discriminatory between respective class pairs. Robust machine learning models trained to distinguish these classes are expected to utilize energy from this signal.

a circular track. The top left panel of Figure 1 shows a SAS image of a cylinder from the Pipe class, and the top-middle panel shows one from the Hollow class. The physical difference between the Pipe and Hollow classes is whether the ends of the cylinder are open or closed, respectively. Because the AirSAS apparatus is so controlled, we can use signal processing to align the images and compute their direct subtraction (visualized in the top-right panel). The resulting difference signal highlights two physical phenomena: (1) the direct reflection off of the closed ends of the Hollow object, which appear as bright, concentrated energy, and (2) the multi-path scattering returns from the insonifying ping traveling through the pipe, which is less focused due to mismatch with the image formation assumptions [1]. The difference signal can thus be used as a pixel-level label of these discriminatory physical phenomena.

The bottom row of Figure 1 shows a similar comparison between a cylinder from the Hollow class (bottom left) and a one from the Solid class (bottom middle) in the spatial wavenumber domain (k-space). In k-space, the radial axis denotes frequency; the center has no signal energy because the data is band-limited from 10kHz-30kHz. The discriminatory acoustic phenomenon that results from the physical differences between Hollow and Solid cylinders is the late-time resonant response, where the insonifying ping has temporarily coupled with the object before returning to the sensor (the Hollow cylinder resonates, the Solid one does not). Resonance manifests as energy concentrated along a radial axis, as is highlighted in the difference signal in the bottom-right panel.

We hypothesize that models relying on verifiable physical phenomena will have more predictable performance as long as that phenomena is available. By generating data that highlights and segments the specific acoustic phenomenology, the extent to which a deep model is focused on those phenomena can be measured [12], and models can even be regularized to utilize this physics-based knowledge [13]. As a preliminary example, we trained tiny CNNs proposed for object detection in SAS imagery [14] to perform binary classification on the Solid vs. Hollow problem. The trained models were evaluated on a test set, and the bottom-level contrastive saliency maps [12] were recorded for each sample (using the model's predicted class as the "top" node). Thanks to the high precision afforded by AirSAS, these maps can be aligned and compared with the corresponding pixel level feature label map (i.e. the feature difference images in the rightmost column of Figure 1). In the left panel of

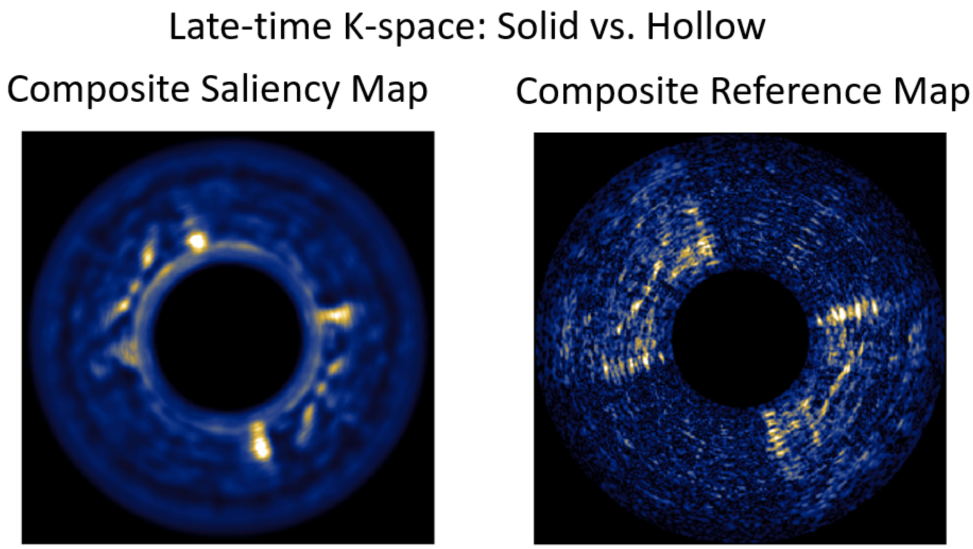

Figure 2: Composite saliency maps (left column) and reference feature maps (right).

Figure 2, we show the composite saliency map, obtained by averaging aligned sample saliency maps by class. The right panel shows the average difference image, computed by averaging the differences between each test set image and its complement (where all features are kept the same except for that which distinguishes the classes). This highlights the features expected to be discriminatory. Using image comparison techniques, we can measure the consistency between the utilization map (left) and the pixel-level feature labels (right). These data-centric consistency metrics values represent how sensitive a model is to specific physical phenomena. We are currently investigating how to use these metrics for regularization, augmentation, and interpretable analysis of object detection models trained on large underwater SAS datasets.

The AirSAS tool will enable researchers to study, and eventually control, how models react to signatures typically found in the "long tail" of large underwater datasets. The preliminary dataset is currently available to interested readers by contacting the authors. The first published version of the AirSAS dataset is currently under curation.

References

[1] David William Hawkins. *Synthetic Aperture Imaging Algorithms: with application to wide bandwidth sonar*. PhD thesis, University of Canterbury. Electrical and Computer Engineering, 10 1996.

[2] JA Bucaro, BH Houston, M Saniga, LR Dragonette, T Yoder, S Dey, L Kraus, and L Carin. Broadband acoustic scattering measurements of underwater unexploded ordnance (uxo). *The Journal of the Acoustical Society of America*, 123(2):738–746, 2008.

[3] BH Houston, JA Bucaro, T Yoder, L Kraus, J Tressler, J Fernandez, T Montgomery, and T Howarth. Broadband low frequency sonar for non-imaging based identification. In *OCEANS'02 MTS/IEEE*, volume 1, pages 383–387. IEEE, 2002.

[4] A Tesei, JA Fawcett, and R Lim. Physics-based detection of man-made elastic objects buried in high-density-clutter areas of saturated sediments. volume 69, pages 422–437. Elsevier, 2008.

[5] David P Williams. Acoustic-color-based convolutional neural networks for uxo classification with low-frequency sonar. In *John S. Papadakis (Hg.): UACE2019-Conference Proceedings.*

5th Underwater Acoustics Conference and Exhibition. Hersonissos, volume 30, pages 421–428, 2019.

[6] Andrea Bellettini and Marc A Pinto. Theoretical accuracy of synthetic aperture sonar micronavigation using a displaced phase-center antenna. *IEEE journal of oceanic engineering*, 27(4): 780–789, 2002.

[7] Daniel C Brown, Isaac D Gerg, and Thomas E Blanford. Interpolation kernels for synthetic aperture sonar along-track motion estimation. *IEEE Journal of Oceanic Engineering*, 45(4): 1497–1505, 2019.

[8] J Daniel Park, Thomas E Blanford, and Daniel C Brown. Late return focusing algorithm for circular synthetic aperture sonar data. *JASA Express Letters*, 1(1):014801, 2021.

[9] Thomas E Blanford, John D McKay, Daniel C Brown, Joonho D Park, and Shawn F Johnson. Development of an in-air circular synthetic aperture sonar system as an educational tool. In *Proceedings of Meetings on Acoustics 177ASA*, volume 36, page 070002. Acoustical Society of America, 2019.

[10] J Daniel Park, Thomas E Blanford, Daniel C Brown, and Daniel Plotnick. Alternative representations and object classification of circular synthetic aperture in-air acoustic data. *The Journal of the Acoustical Society of America*, 148(4):2661–2661, 2020.

[11] Thomas E Blanford, Joonho D Park, Shawn F Johnson, and Daniel C Brown. Experimental analysis of the effects of bistatic target scattering on synthetic aperture sonar imagery. In *Proceedings of Meetings on Acoustics 178ASA*, volume 39, page 070001. Acoustical Society of America, 2019.

[12] Jianming Zhang, Sarah Adel Bargal, Zhe Lin, Jonathan Brandt, Xiaohui Shen, and Stan Sclaroff. Top-down neural attention by excitation backprop. *International Journal of Computer Vision*, 126(10):1084–1102, 2018.

[13] Ruth Fong and Andrea Vedaldi. Occlusions for effective data augmentation in image classification. In *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pages 4158–4166. IEEE, 2019.

[14] David P Williams. On the use of tiny convolutional neural networks for human-expert-level classification performance in sonar imagery. *IEEE journal of oceanic engineering*, 46(1): 236–260, 2020.