San Antonio Breast Cancer Symposium- December 6-10, 2022.

Artificial Intelligence Detects, Classifies, and Describes Lesions in Clinical Breast Ultrasound Images

Arianna Bunnell[12], Dustin Valdez[12], Thomas Wolfgruber[2], Aleen Altamirano[3], Brenda Hernandez[2], Peter Sadowski[1], John A. Shepherd[2]
University of Hawaii, Honolulu, HI, USA[1], University of Hawaii Cancer Center, Honolulu, HI, USA[2], Instituto Radiodiagnóstico, Managua, Nicaragua[3]

UNIVERSITY OF HAWAI'I CANCER CENTER

ARTIFICIAL INTELLIGENCE PRECISION HEALTH INSTITUTE
UNIVERSITY OF HAWAI'I

Summary

An Artificial Intelligence (AI) model is trained to detect lesions in clinically-acquired breast ultrasound images, using almost 40,000 images from a multiethnic population of women participating in a Hawai'i screening program, and achieves an AUROC of 0.86 on a held out test set. In addition, our system characterizes breast lesions in terms of the ACR BI-RADS mass characteristics.

Background

- Advanced stage breast cancer (stages III and IV) rates in the U.S.-Affiliated Pacific Islands (USAPI) are much higher than in the continental United States. Examples include Guam (60% advanced stage cancer rate), American Samoa (74%), and the Federated States of Micronesia (81%) [1].
- The USAPI suffer from a lack of trained radiologists and poor access to screening mammography. Breast ultrasound (BUS) is a feasible replacement for mammography in low-resource areas.
- The ACR BI-RADS masses lexicon for BUS [2] delineates important characteristics to assist healthcare workers and teleradiologists in differentiating between benign and malignant breast lesions.
- Our overall hypothesis is that introducing portable handheld BUS systems coupled with an AI detection algorithm operated by a general healthcare worker will reduce advanced stage cancer rates.
- In this study, we ask if AI can detect and determine the cancer status and BI-RADS mass characteristics in clinically-acquired breast ultrasound images.

Methods

Study Design: Prospective case-control study of all women who received screening or diagnostic breast ultrasound imaging from 2009 to 2021 (over 60,000) at one clinical site participating in the Hawai'i and Pacific Islands Mammography Registry (HIPIMR).
Case Selection: Participants had to have a BUS visit within 1 year of biopsy-confirmed breast cancer diagnosis. Cancer labels were collected from the Hawai'i Tumor Registry.
Non Case Selection: Three controls for each case matched on BUS machine manufacturer and birth year.
BUS Image Processing: Doppler, elastography, and invalid scans were excluded. Scans were also excluded if the scan area was majority (>75%) black. Scan images were cleaned according to our adaption of the methods presented in [3]. This includes automatically cropping to scan shape and removing artifacts.
Strongly-Labeled Dataset: *Approximately 20% of our overall data were reviewed and delineated by our study radiologist without knowledge of cancer status.* BI-RADS mass characteristics were also described. The images were split into training (70%), validation (20%), and testing (10%) subsets by woman.
Weakly-Labeled Dataset: The remainder of our BUS scans only had cancer/no cancer labels and were used to extend the pretraining of the AI model backbone. See below. Images were split into training (70%) and validation (30%) subsets by woman.
AI Model Design: We chose a modified Detectron2 Mask R-CNN model [4] that consisted of a pre-trained ResNet-101 classifier backbone. We fine-tuned the backbone on the weakly labeled dataset to classify cancer status. Detectron2 performs object localization, segmentation, and classification simultaneously and was further fine-tuned using the strongly-labeled dataset to produce lesion delineations, classification and BI-RADS lesion descriptors to help explain the AI finding to the end user.
Statistics: Model performance was evaluated on the strongly-labeled, held-out test set. Metrics include precision and One vs. Rest Area under the Receiver Operating Characteristic Curve (AUROC).

Results

Table 1: Statistics of the dataset, broken down by data split. This dataset was collected from a single Hawaii health center over a 12-year period. Cancer labels were obtained via biopsy records through the HTR.

	Weakly Labeled			Strongly Labeled			
Characteristic, Unit	**Overall**	**Train**	**Validation**	**Overall**	**Train**	**Validation**	**Test**
Women, N	1,755	1,223	532	393	272	76	45
Women with benign findings, N	1,347	937	410	282	195	54	33
Women with malignant findings, N	408	286	122	111	77	22	12
Images, N	33,475	23,437	10,038	4,398	3,136	811	451
Images with benign findings, N	23,276	16,453	6,823	2,785	1,974	515	296
Images with malignant findings, N	10,199	6,984	3,215	1,613	1,162	296	155
Average no. of images per woman, N	19.58	19.69	19.35	11.19	11.53	10.67	10.02

Results (continued)

Table 2: AUROC values for the BI-RADS characteristics and cancer status for detected lesions in the testing set.. *No hyperechoic lesions were present. **< 20 lesions displayed this characteristic.

Target	Category	AUROC
Cancer	**Overall**	**0.86**
Shape	**Overall**	**0.85**
	Oval	0.95
	Round	0.63**
	Irregular	0.97
Orientation	**Overall**	**0.94**
Margin	**Overall**	**0.85**
	Circumscribed	0.95
	Indistinct	0.88
	Angular	0.73
	Microlobulated	0.79**
	Spiculated	0.89
Echo Pattern	**Overall**	*
	Anechoic	0.95
	Hypoechoic	0.83
	Isoechoic	0.98**
	Hyperechoic	*
	Complex cystic & solid	0.95**
	Heterogeneous	0.91
Posterior Features	**Overall**	**0.80**
	No posterior features	0.62
	Enhancement	0.80
	Shadowing	0.73
	Combined pattern	0.79**

The 1,755 women with weakly labeled scans had a total of 33,475 BUS images (average 20 images/woman) and a mean age of 61.11. The 393 women with strongly labeled scans had a total of 4,398 BUS images (average 11 images/woman) with 2,002 benign and 1,437 malignant lesions identified by the study radiologist. The mean age of these women was 64.51. There were 1.47 benign lesions/image and 1.35 malignant lesions/image, on average.

Lesion detections were positive if they overlapped a ground truth lesion by at least 50% (intersection over union = 0.5). A maximum of 4 detections were generated per image. Table 2 values correspond to positive detections only.

Identifying cancerous and benign lesions from the scan area, the AI had an AUROC of 0.70 and 0.75, respectively.

Timing experiments were run on a GPU-enabled (Nvidia Tesla V100) machine on unbatched images. The model made predictions at a rate of 25 images/second time (38.7 milliseconds per image).

Conclusion

Detection, segmentation, and cancer classification of breast lesions are possible in clinically-acquired BUS images using current AI techniques, and accuracy will continue to increase with additional data. Based on our timing experiments, the model is capable of detecting and classifying lesions in real-time during ultrasound capture.

We are currently integrating this model into a portable BUS system and planning to run clinical trials in low-resource areas of the USAPI.

References

[1] Cancer in the U.S. Affiliated Pacific Islands 2007-2018. Pacific Regional Cancer Registry, 2021.
[2] Mendelson EB, Böhm-Vélez M, Berg WA, et al. ACR BI-RADS® Ultrasound. In: ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. Reston, VA, American College of Radiology; 2013.
[3] Shamout FE, Shen Y, Witowski JS, Oliver JR, Kannan K, Wu N, Park J, Beatriu, Reig, Moy L, Heacock L, Geras KJ, editors. The NYU Breast Ultrasound Dataset v1.02021.
[4] Wu Y, Kirillov A, Massa F, Lo W-Y, Girshick R. Detectron2. https://github.com/facebookresearch/detectron2.

BE THE READER

Figure 1 *(right)* shows nine examples of BUS scans from the strongly labeled test set. The leftmost column shows the original scan; the middle column shows the delineation, cancer label, and BI-RADS mass characteristic annotations from the study radiologist; and the rightmost column shows the AI model's predictions for delineation, cancer label, and mass characteristics. The AI model output shows first the benign/cancer prediction, followed by the prediction for the two most probable classes for each of the BI-RADS mass characteristics. Lesion delineation colors are randomly assigned. See how you compare to our AI tool!