Multimodality Machine Learning for Breast Cancer Detection: Synergistic Performance with Upstream Data Fusion of Digital Breast Tomosynthesis and Ultrasound

David Porter, PhD, William Walton, PhD, Susan Harvey, MD, Benjamin Tsui, Keith Peyton, Mary Kate Jones, Taek-Soo Lee, PhD, Paul H Yi, MD, Ferdinand K Hui, MD, Haris I Sair, MD, Emily Ambinder, MD, Lisa Mullen, MD, & Seung-Jun Kim, PhD

1Johns Hopkins Applied Physics Laboratory; 2Johns Hopkins Radiology Artificial Intelligence Lab (RAIL), 3University of Maryland, Baltimore County

Background.
Breast ultrasound (US) combined with mammography, increasingly digital breast tomosynthesis (DBT), provides high sensitivity for breast cancer detection. Ultrasound alone is limited by low specificity and mammography alone is limited by dense tissue. Although multimodality imaging for breast cancer detection identifies more cancers compared to mammography alone, there are high false positive rates. We hypothesize that Upstream Data Fusion (UDF) may alleviate these challenges. We aimed to develop and test a pilot multimodality machine learning algorithm for the detection of breast cancer, to discern whether fusion improves diagnostic performance over the individual modalities.

Methods.
DBT and breast US studies were performed in 6 patients with 10 breast lesions, 4 biopsy-proven benign and 6 biopsy-proven malignant. Data augmentation using a sliding window was performed resulting in 39 data points for benign lesions and 67 data points for malignant lesions. Data from the two modalities was combined using UDF, a data fusion technique that has been implemented for military purposes for detection and localization of enemy targets using multiple surveillance modalities. The UDF-combined images were then used to train Support Vector Machines (SVMs), a type of supervised machine learning model, in order to differentiate benign and malignant lesions. Outcome measures sensitivity, specificity and accuracy are presented for the joint UDF model. Accuracy of the joint model was compared to models using DBT alone and US alone using Fisher’s Exact Test.

Results.
In distinguishing between benign and malignant lesions, the UDF-combined imaging model showed joint-sensitivity of 100%, specificity of 90%, and accuracy of 96%. Although not reaching statistical significance, the UDF-combined imaging model demonstrated a trend toward greater diagnostic accuracy than models using each of the modalities individually. For example, the UDF model using DBT alone has accuracy 92% (p=0.55) and the UDF model using US alone has accuracy 94% (p=0.21).

Conclusion.
Despite a small training set in this pilot work, UDF multimodality learning models fusing DBT and US were able to increase accuracy over US or DBT alone.