Subtyping invasive carcinomas and high-risk lesions for machine learning based breast pathology

Matthew G Hanna, MD, Patricia Raciti, MD, Alican Bozkurt, PhD, Ran Godrich, MS, Julian Viret, MS, Donghun Lee, MS, Philippe Mathieu, BS, Matthew Lee, PhD, Eugene Vorontsov, PhD, Tomer Sabo, MBA , Felipe C. Geyer, MD, Jorge S. Reis-Filho, MD, PhD, Leo Grady, PhD, Thomas Fuchs, DrSc, Christopher Kanan, PhD

San Antonio Breast Cancer Symposium - December 7-10, 2021

Paige Breast "Beta"

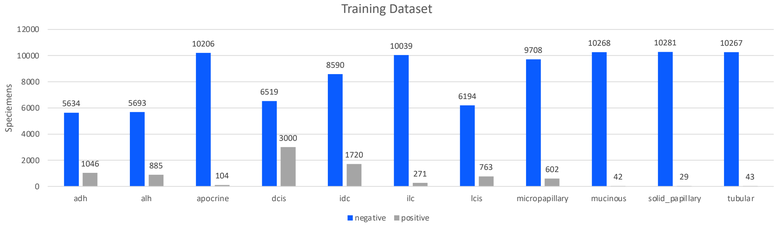
- Pathology multi class operating on H&E WSIs, intended to draw pathologist's attention to concerning features.
- Concerning features: invasive breast cancer (ductal and lobular), in situ breast cancer (DCIS and LCIS), and atypia (atypical ductal hyperplasia, atypical lobular hyperplasia)
- Including distinguishing morphologies: apocrine, micropapillary, mucinous, solid papillary, and tubular
- Biopsy and surgical resection tissues were used to develop and validate the model
- Trained on 40,637 slides from 9,751 breast specimens and validation assessed on 11,447 slides from 3,183 breast specimens (separate from training data)

BACKGROUND

The female mammary gland can develop a myriad of epithelial proliferative lesions including, high risk lesions, in-situ and invasive carcinomas. Identification of these pre-neoplastic and neoplastic conditions in biopsy specimens is crucial for proper patient management and may sometimes pose diagnostic challenges for pathologists. Recent research has shown that machine learning algorithms applied to whole slide images (WSI) can accurately detect and grade various cancers; herein, we devise and test a system that classifies the most common preneoplastic and neoplastic conditions of the female breast from WSIs.

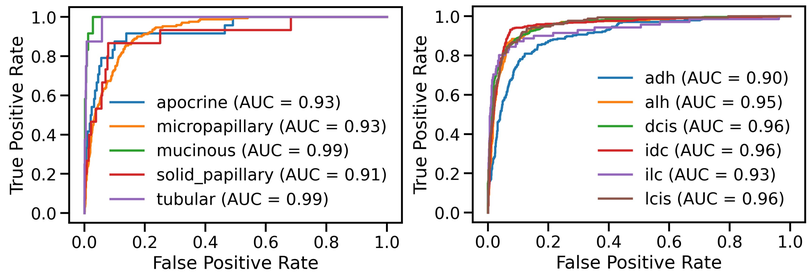
MATERIALS & METHODS

De-identified slides were scanned on Leica AT2 whole slide scanners (20x; 0.5 µm/pixel) from MSK database were retrieved. Clinical diagnostic metadata were extracted from the pathology reports. Using a multi-label multiple-instance learning (ML-MIL) approach, a SE-ResNet50 Convolutional Neural Network (CNN) was trained to classify **atypical lobular hyperplasia (ALH), atypical ductal hyperplasia (ADH), lobular carcinoma in situ (LCIS), ductal carcinoma in situ (DCIS), invasive lobular carcinoma (ILC), invasive ductal carcinoma (IDC)**. In additional morphological subtypes including **apocrine, mucinous, solid papillary, micropapillary,** and **tubular carcinoma** were trained. The system uses the WSI as an input and outputs a slide level class and heatmap for the presence of the trained classes. The CNN was trained on 9,751 surgical specimens (biopsy, 6,289; excision, 3,462) comprising 40,637 slides. A validation dataset of 3,183 breast specimens (biopsy, 1,934; excision, 1,249) comprising 11,447 digital slides, separate from the training set was used to assess performance of the trained model.

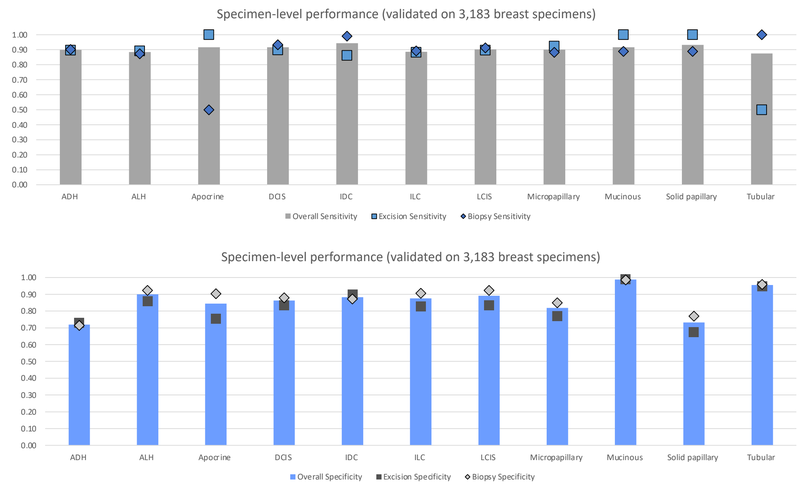


RESULTS

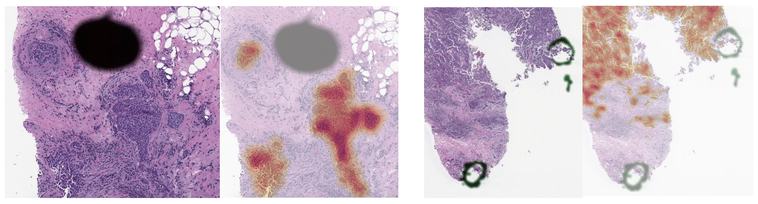
The system was validated on 3,183 breast specimens (biopsy, 1,934; excision, 1,249) comprising 11,447 digital slides that were not included in the training of the CNN model.



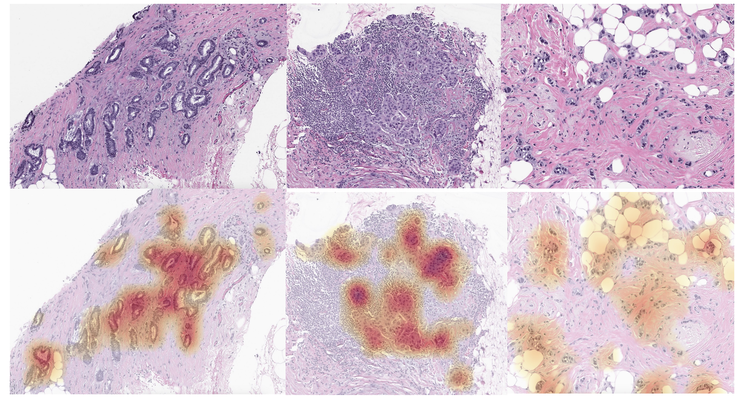
Left, AUC curves for morphological subtype detection
Right, AUC curves for diagnostic category detection



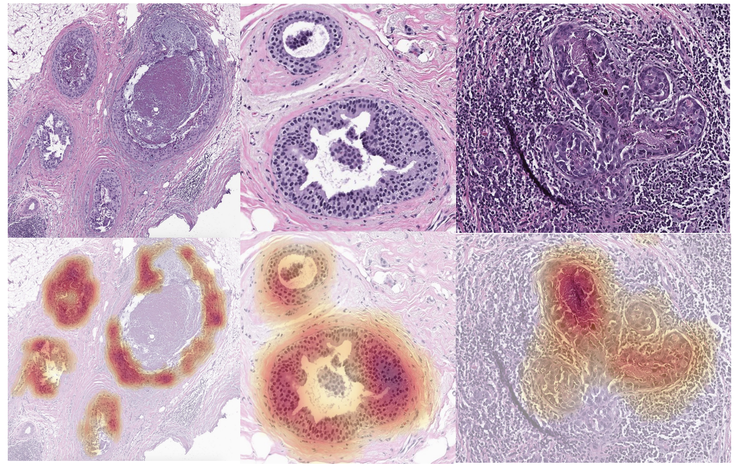
Top, Sensitivity for diagnostic categories & morphologies
Bottom, Sensitivity for diagnostic categories & morphologies



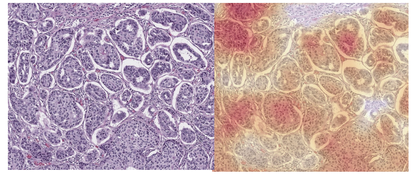
Paige Breast shows robust detection with ink markings



Matched representative slides with invasive carcinoma
Top, H&E; Bottom, Paige Breast heatmap



Matched representative slides with ductal carcinoma in situ
Top, H&E; Bottom, Paige Breast heatmap



Detection of micropapillary morphology

CONCLUSIONS

The trained CNN had a high performance in identifying the presence of ADH, ALH, DCIS, IDC, ILC, LCIS, and, apocrine, micropapillary, mucinous, solid papillary, and tubular carcinomas. Further studies expanding classes to include all clinically relevant lesions and morphologies are underway. In addition, the same approach can be used to detect microinvasions and calcifications in breast tissue.