AI Augmented Breast Biopsy Diagnostics

Ashbaker Kathleen, Hosny Kareem, MD, MPH

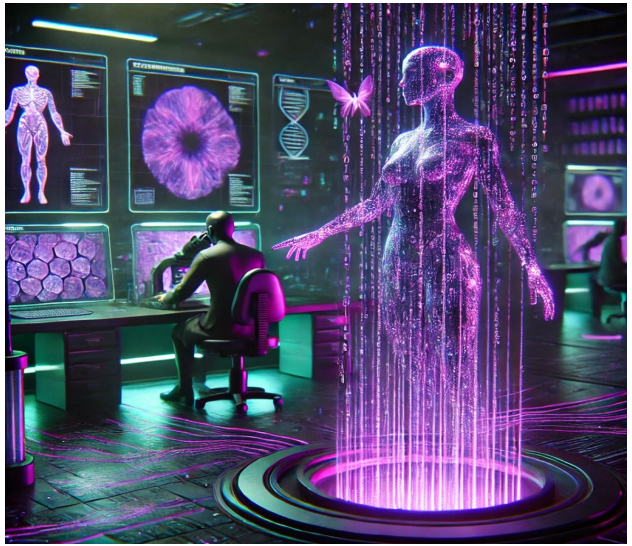
University of Washington School Medicine: Department of Computer Science & Engineering

BE BOUNDLESS

Digital Pathology and AI

Importance of Timely and Accurate Breast Biopsy Interpretation

Digital pathology is revolutionizing diagnostic medicine, especially in breast cancer where time-sensitive decisions are vital.



Artificial Intelligence (AI) has become an indispensable tool for helping manage increasingly complex workflows, enabling triaging and assisting pathologists in interpretation at scale. Our work builds upon these advancements.

UNIVERSITY *of* WASHINGTON

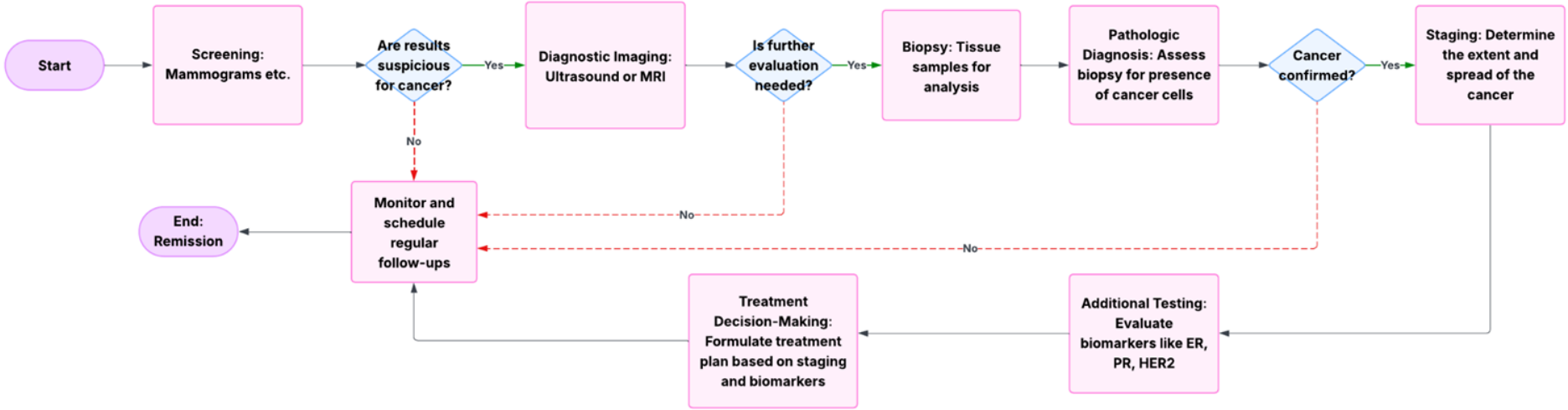
Importance of Timely and Accurate Breast Biopsy Interpretation

Missed and or undiagnosed cases of breast cancer :

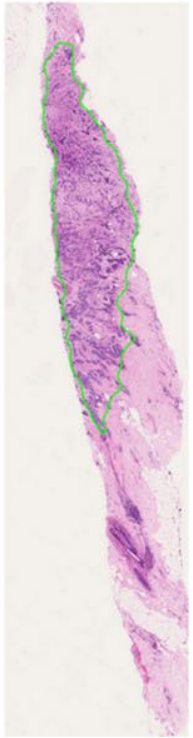
2022: Artificial-intelligence-based, computer-aided diagnosis (AI-CAD); The AI-CAD correctly **localized ~79%missed cancers** on prior mammograms [2]

Importance of Timely and Accurate Breast Biopsy Interpretation

Current State for Diagnostic Workflow for Breast Cancer [1]



Future State



Scanned image is analyzed by trained model

If DCIS

Order ER

If IDC/ILC

Order ER/PR/HER2

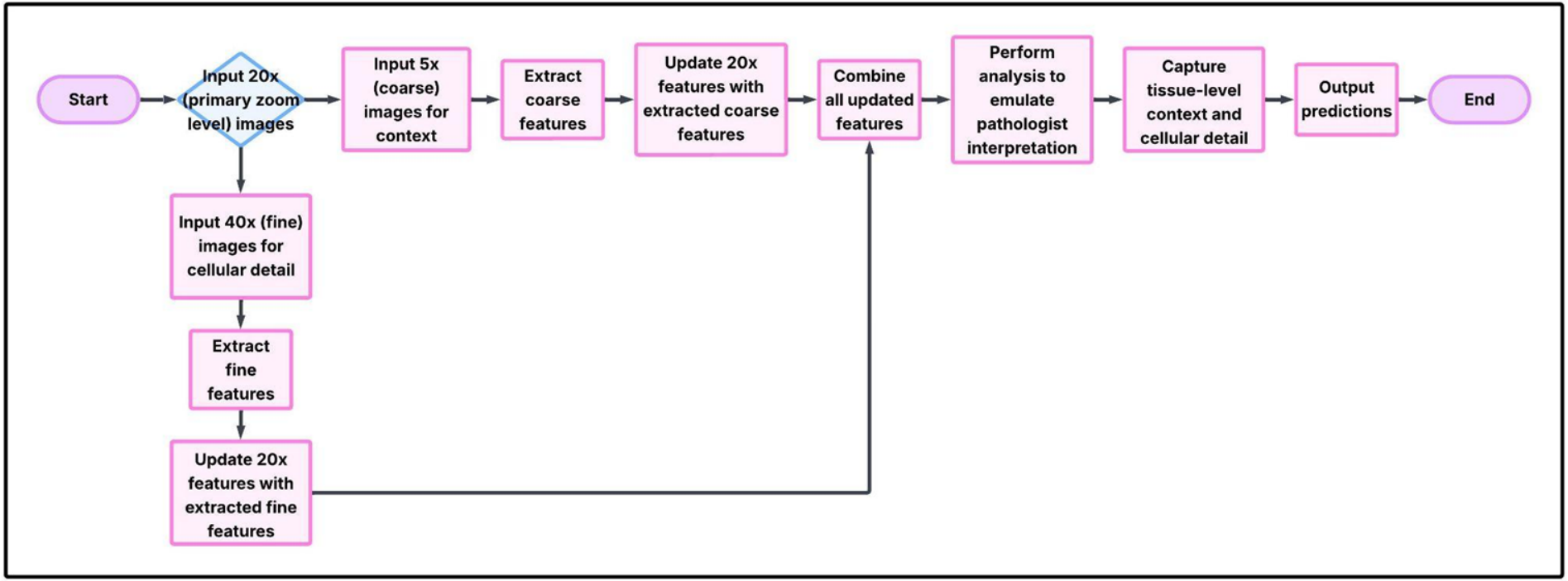
If negative/benign

No further action

MIL-Driven Platform for Streamlined Breast Biopsy Interpretation

K. Ashbaker, R. Soraki, T. Krishnan, M. Nelson, C. Rizkalla, M. Kilgore, J. Henriksen, K. Hosny

- Goal: to develop a platform aimed at improving breast biopsy interpretation TATs by automating IHC orders for neoplastic cases
- Involves using AI-augmented diagnoses on digital slides to stratify breast biopsies into neoplastic and non-neoplastic pathways
- Neoplastic cases are further classified into in situ disease (DCIS & LCIS) and invasive disease (IDC & ILC)
- Breast pathologists will receive IHC and H&E-stained slides simultaneously
- Expected to reduce biopsy TAT by approximately one day


Start
Input 20x (primary zoom level) images
Input 5x (coarse) images for context
Extract coarse features
Update 20x features with extracted coarse features
Combine all updated features
Perform analysis to emulate pathologist interpretation
Capture tissue-level context and cellular detail
Output predictions
End
Input 40x (fine) images for cellular detail
Extract fine features
Update 20x features with extracted fine features

Methods: AI-Augmented Diagnoses, Triaging

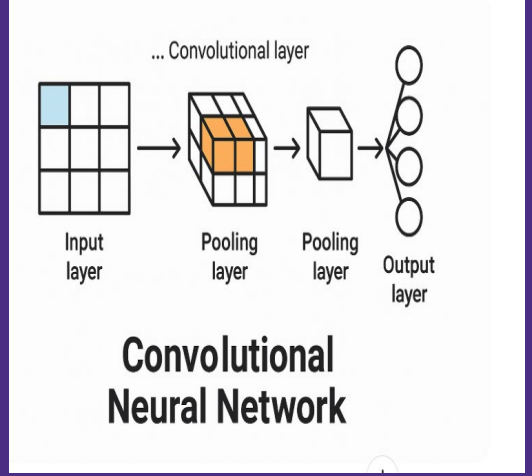
Utilizing AI to automate immunohistochemistry orders for neoplastic cases.

Integration of Multi-Scale Co-attention Convolution for whole slide imaging, considering multiple magnifications simultaneously for nuanced analysis.

Outcomes:

- Streamlining Pathological Analysis
    - Anticipated reduction in turnaround times by about one day.
    - Facilitation of same-day diagnostic reporting and signout, enhancing workflow efficiency.

Methods: (cont)



This study uses Multi-Scale Co-attention Convolution

We selected x20 magnification level as the primary input.

Then used x5 magnification level, and x40 magnification level to update the primary input.

Accounts for primary input features, coarse-grained features, and fine-grained visual cues.

Categories classification

Category	Subtype	Description
In Situ Disease	Ductal Carcinoma In Situ (DCIS)	Non-invasive cancer starting in the milk ducts
	Lobular Carcinoma In Situ (LCIS)	Abnormal cells in the milk-producing lobules
Invasive Disease	Invasive Ductal Carcinoma (IDC)	Cancer that began in the ducts and invaded surrounding tissue
	Invasive Lobular Carcinoma (ILC)	Cancer that started in the lobules and invaded nearby tissue

Report (5x → 10x ← 20x):

Class	Precision	Recall	F1-Score	Support
Normal	0.96	0.98	0.97	190
DCIS	0.93	0.76	0.84	17
LCIS	0.67	0.60	0.63	10
IDC	0.95	0.98	0.97	43
ILC	0.94	0.88	0.91	17
Accuracy			0.95	277
Macro Avg	0.89	0.84	0.86	277
Weighted Avg	0.94	0.95	0.94	277

Report (10x → 20x ← 40x):

Class	Precision	Recall	F1-Score	Support
Normal	0.98	0.94	0.96	190
DCIS	0.79	0.88	0.83	17
LCIS	0.80	0.80	0.80	10
IDC	0.86	0.98	0.91	43
ILC	0.88	0.88	0.88	17
Accuracy			0.93	277
Macro Avg	0.86	0.90	0.88	277
Weighted Avg	0.94	0.93	0.93	277

Combined Confusion Matrix (5x → 10x ← 20x):

Confusion Matrix		Predicted Class				
		Normal	**DCIS**	**LCIS**	**IDC**	**ILC**
Actual Class	**Normal**	189	1	1	2	0
	DCIS	3	13	1	0	0
	LCIS	3	0	6	0	1
	IDC	1	0	0	42	0
	ILC	1	0	1	0	15

Combined Confusion Matrix (10x → 20x ← 40x):

Confusion Matrix		Predicted Class				
		Normal	**DCIS**	**LCIS**	**IDC**	**ILC**
Actual Class	**Normal**	178	3	1	7	1
	DCIS	2	15	0	0	0
	LCIS	1	1	8	0	0
	IDC	0	0	0	42	1
	ILC	1	0	1	0	15

Metrics calculated on the test set:

Method	Accuracy	$\text{Precision}_{(\text{Macro})}$	$\text{Recall}_{(\text{Macro})}$	$\text{F1-Score}_{(\text{Macro})}$	AUC
$\text{AMIL}_{(10x)}$	92.42%	0.90	0.83	0.85	0.98
$\text{AMIL}_{(20x)}$	89.17%	0.79	0.84	0.81	0.96
$\text{AMIL}_{(40x)}$	88.45%	0.82	0.74	0.77	0.93
$\text{DSMIL}_{(10x)}$	92.78%	0.85	0.81	0.83	0.97
$\text{DSMIL}_{(20x)}$	89.17%	0.81	0.78	0.79	0.95
$\text{DSMIL}_{(40x)}$	80.14%	0.60	0.68	0.63	0.90
$\text{TransMIL}_{(10x)}$	93.50%	0.86	0.86	0.86	0.99
$\text{TransMIL}_{(20x)}$	93.14%	0.87	0.83	0.85	0.98
$\text{TransMIL}_{(40x)}$	94.22%	0.87	0.83	0.85	0.98
$\text{Ours}_{(5x \rightarrow 10x \leftarrow 20x)}$	94.58%	0.89	0.84	0.86	0.98
$\text{Ours}_{(10x \rightarrow 20x \leftarrow 40x)}$	93.14%	0.86	0.90	0.88	0.99

Results(cont): Takeaways

Performance Metrics: AUC =0.98; 0.99 for 5,10,20X mag and 10,20,40X mag respectively.

Overall Model Evaluation:

- The weighted average metrics show a high consistency in model performance with **Accuracy, Precision, Recall, and F1-Score all around 0.93**.
- The results suggest that our approach maintains high reliability and accuracy across different breast cancer classes, potentially aiding in better triage and diagnosis in clinical settings.

Results(cont): Takeaways

Precision and Recall Highlights:

- High precision for Normal class
- **LCIS** class achieved perfect recall and precision, showing the model's effectiveness in accurately identifying less common categories.
- **DCIS** class saw significant improvement in recall (0.88) with our method, indicating better identification of potential early-stage cancers.

Discussion and Caveats

Confusion Matrix Insights:

- Strong correct predictions in **Normal** and **IDC** classes with few misclassifications, demonstrating the model's robustness in distinguishing between benign and invasive conditions.

- Misclassifications were minimal and mostly confined to similar categories, suggesting that the models are generally reliable but could be improved for specificity between closely related classes, i.e. invasive and non-invasive.

Conclusion


Same-Day Diagnostic Reporting
Human-AI Collaboration & Validation
AI Platform Potential and Integration
Turnaround Time Reduction
Automatic Triage and IHC Recommendations

Thank you, questions?

References

1. Ahn JS, Shin S, Yang SA, et al. Artificial Intelligence in Breast Cancer Diagnosis and Personalized Medicine. *J Breast Cancer*. 2023;26(5):405-435. doi:10.4048/jbc.2023.26.e45
2. Park GE, Kang BJ, Kim SH, Lee J. Retrospective Review of Missed Cancer Detection and Its Mammography Findings with Artificial-Intelligence-Based, Computer-Aided Diagnosis. *Diagnostics*. 2022; 12(2):387. https://doi.org/10.3390/diagnostics12020387

Subtype Pair	Clinical Relationship	Confusion (5x→20x)	Confusion (10x→40x)	Insight
DCIS vs. IDC	Same ductal origin; non-invasive vs. invasive	No direct confusion; DCIS mostly confused with Normal/LCIS	No direct confusion; DCIS → Normal	Model reliably distinguishes invasive from non-invasive ductal lesions
LCIS vs. ILC	Same lobular origin; non-invasive vs. invasive	LCIS → ILC (1); ILC → LCIS (1)	No confusion observed	Higher magnification reduces confusion between lobular subtypes
DCIS vs. LCIS	Both in situ types; ductal vs. lobular	Some mutual confusion; both misclassified as Normal	Reduced confusion; clearer distinction	In situ cancers are harder to distinguish due to subtler morphological features
IDC vs. ILC	Both invasive; ductal vs. lobular	No mutual confusion	ILC → IDC (1 case)	Invasive cancers show distinct features; minor overlap remains
In Situ vs. Normal	Non-invasive pathology vs. healthy tissue	Frequent misclassification of DCIS/LCIS as Normal	Misclassification persists, especially for DCIS	Subtlety of in situ lesions causes overlap with Normal
Invasive vs. Normal	Aggressive vs. healthy tissue	Very few errors; IDC and ILC well classified	Slight increase in Normal → IDC (7 cases)	Invasive cancers are more distinguishable; model is robust