Application of artificial intelligence to mammography-tomosynthesis combined images for breast cancer screening

S. Pacilè, C.Aguilar, A.Iannessi, P.Fillard

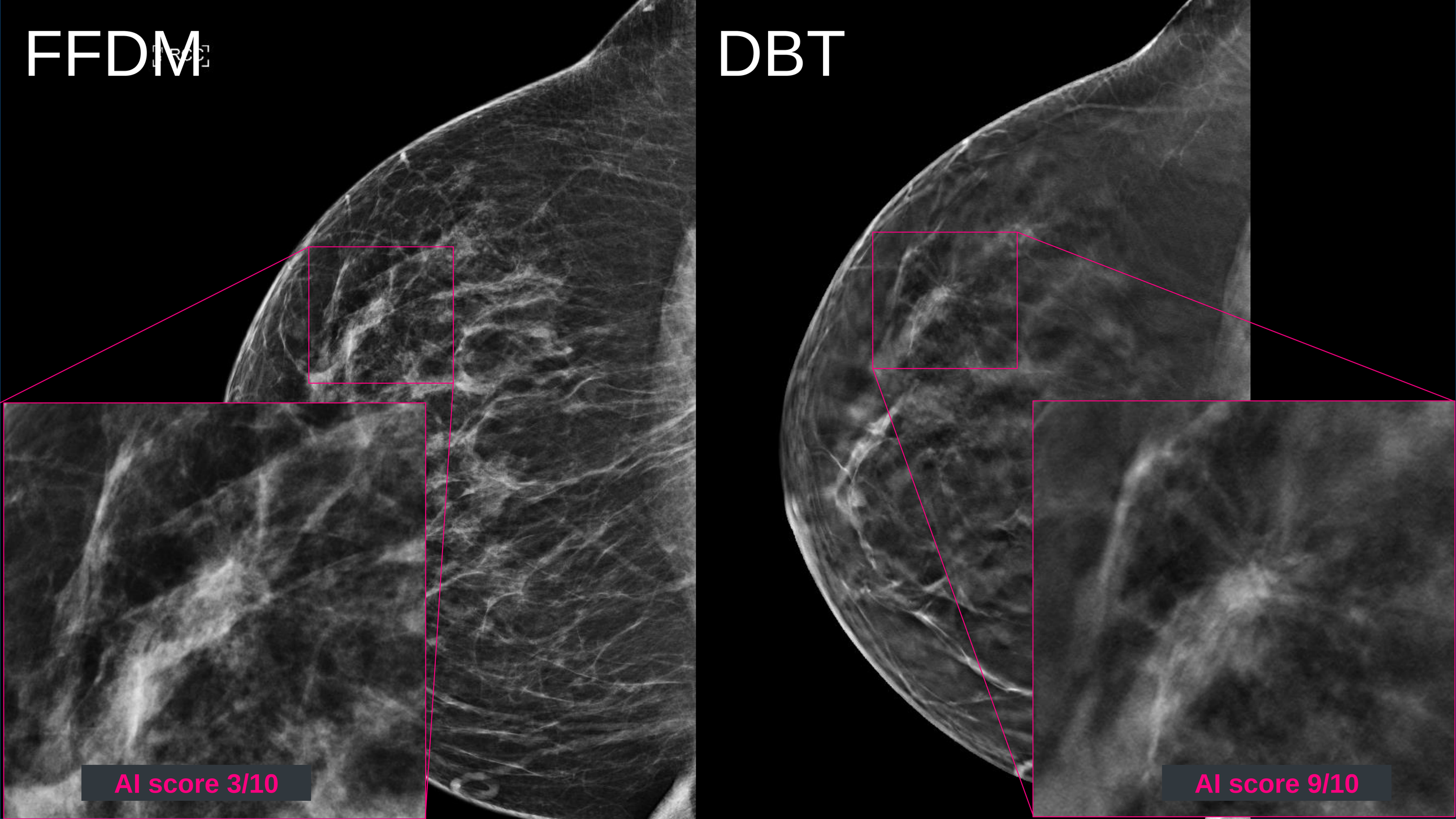
spacile@therapixel.com

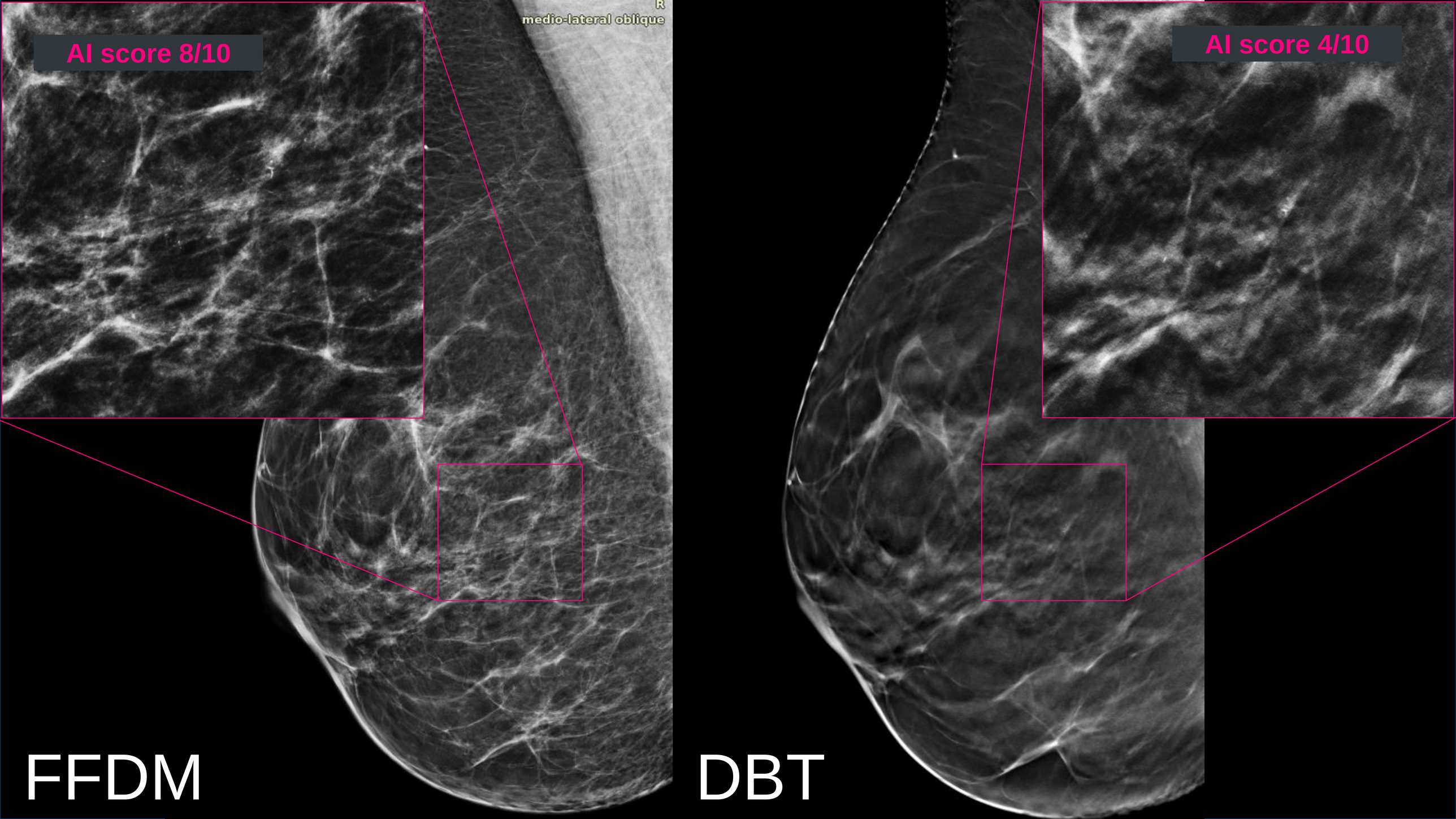
Disclosures

- Employee at Therapixel

Introduction

- Mammography is the standard of care for the early detection of breast cancer and reduces mortality from breast cancer
- The reconstructed quasi–three-dimensional data acquired with digital breast tomosynthesis (DBT) improves detection, characterization, and localization of lesions
- DBT as a standalone technique cannot replace 2D techniques (FFDM or 2DSM) for microcalcification analysis and comparison with priors
- AI performance is linked to lesion visibility:
  - better performance on DBT for soft tissue lesions
  - better performance on FFDM for calcifications


FFDM
DBT
AI score 3/10
AI score 9/10


AI score 8/10
R
medio-lateral oblique
AI score 4/10
FFDM
DBT

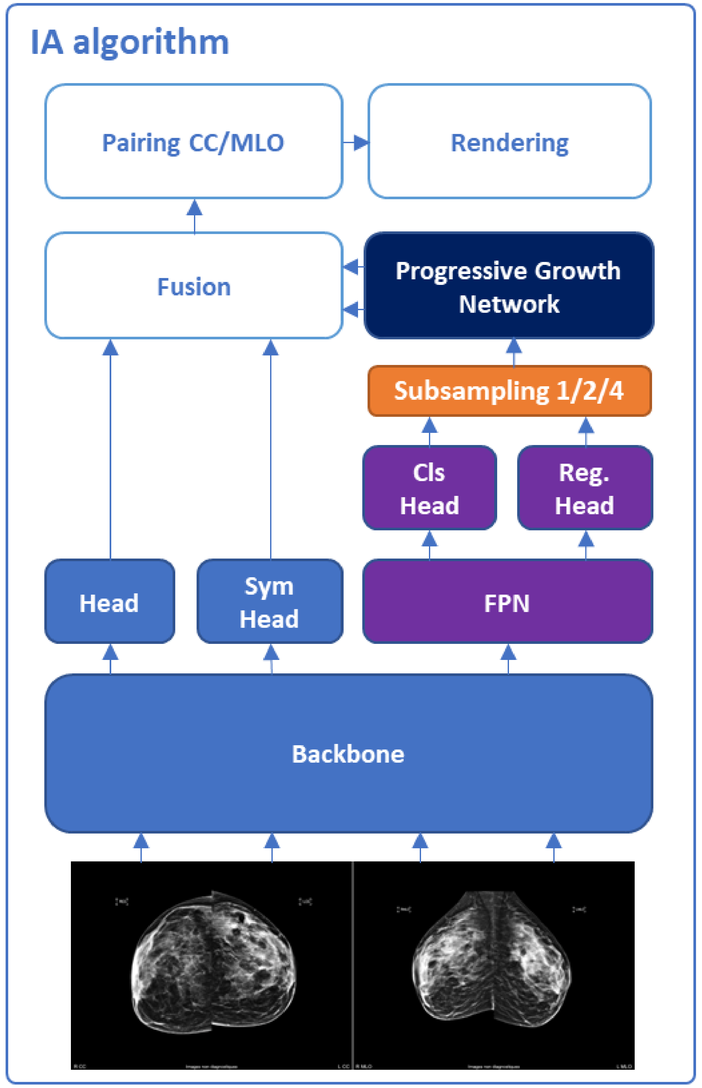
Research Questions

1. Can AI combine the detections of the two modalities?

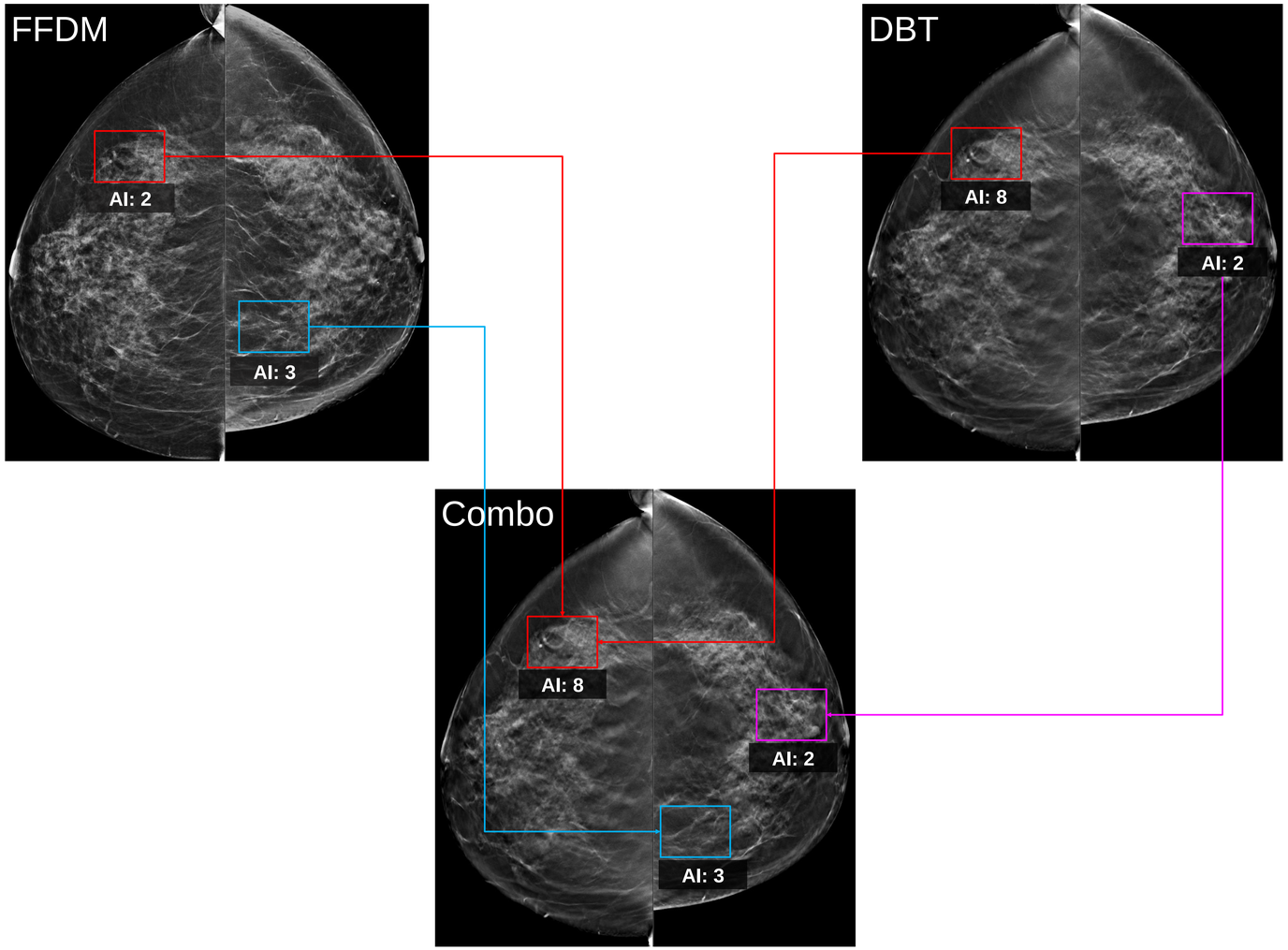
2. Which combination paradigm reaches higher performance?

Materials and Methods: AI system

- **Detects** region of interest (ROI) on the 4 views (2D or DBT)
- **Characterizes** ROI with a score from 0 (no suspicion) to 1 (high suspicion for malignancy)
- **Discrete score** from 1 to 10
- Trained on **1M examinations** (4M images):



Materials and Methods: Combination Paradigms



Example of AI score assignment using combination paradigm 1

PARADIGM 1: COMBO conservative

- Same ROI detected on both modalities → maximum score of the two modalities is kept
- ROI detected on one modality only → the score of this modality is kept

PARADIGM 2: COMBO stratified

- Same ROI detected on both modalities:
  - Maximum score 2D for calcifications
  - Maximum score DBT for soft tissue lesions
- ROI detected on one modality only → the score of this modality is kept

Materials and Methods: Dataset & Analysis

DATASET:

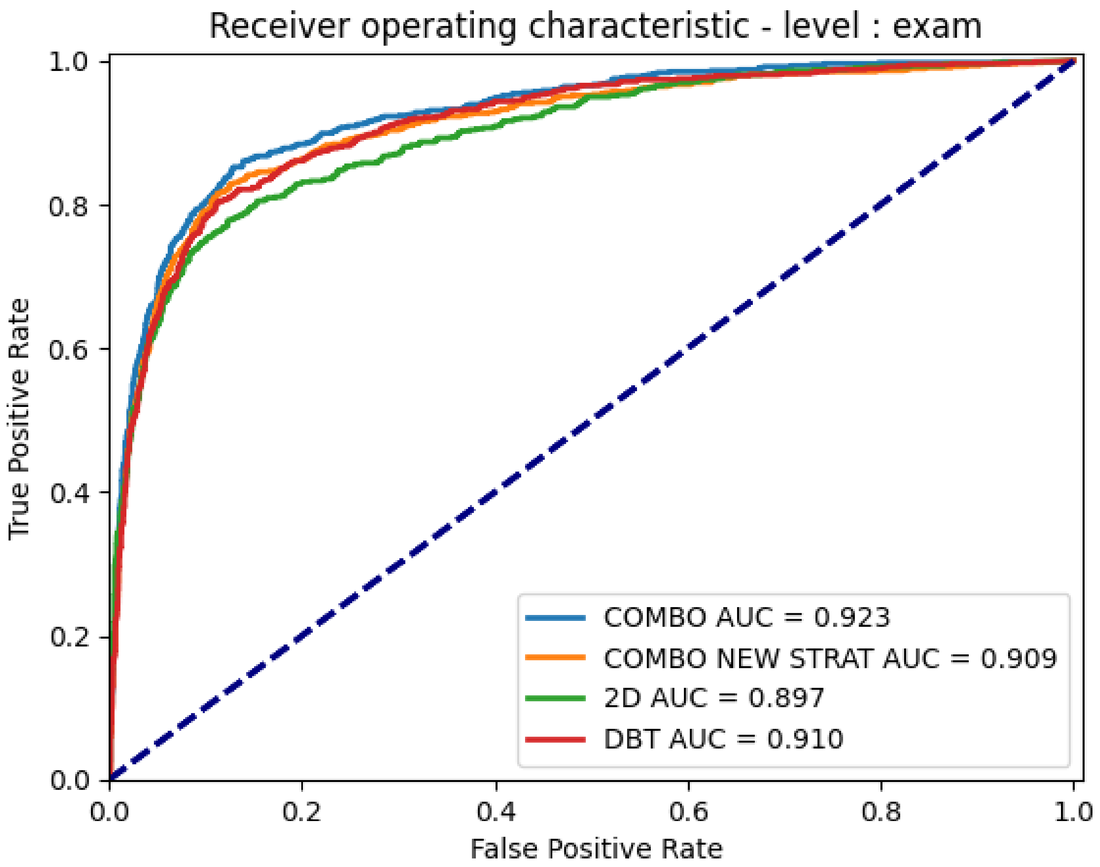
- 3083 included patients
- 753 biopsy proven cancer cases
- 2330 confirmed negative screens

ANALYSIS:

For each modality (2D, DBT, Combo conservative, Combo stratified):

1. Area under the Receiver Operating Characteristic curve (**AUC ROC**)
2. Area under the Localization ROC curve (**AUC LROC**)
3. Average Precision = weighted mean of precision achieved at each threshold (**AP PR**)

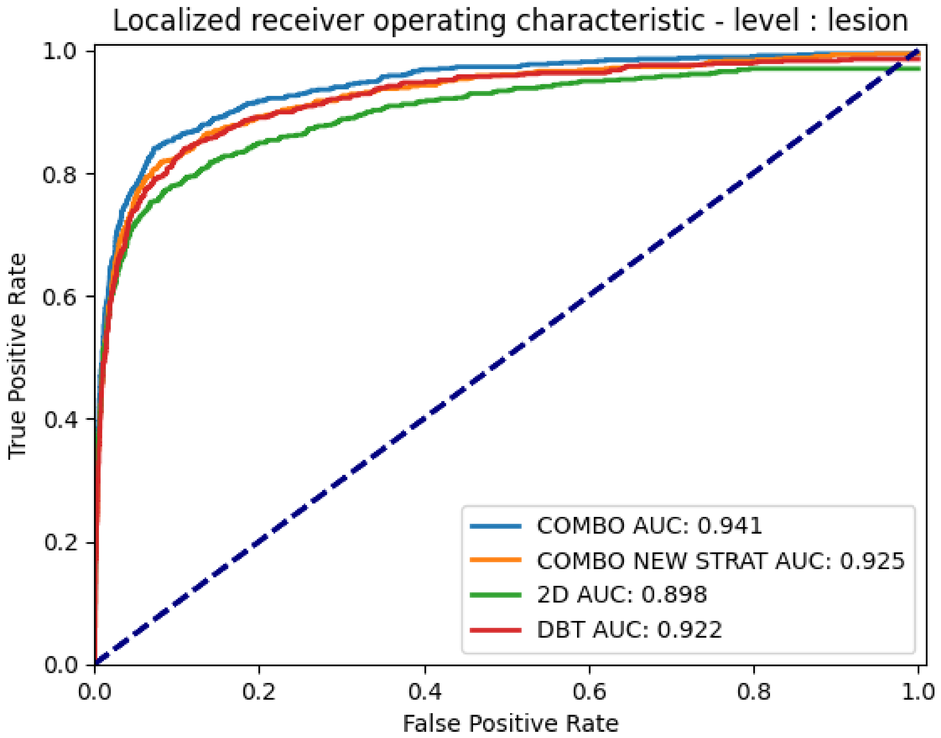
Results – ROC



- COMBO conservative AUC higher than AUC of individual modalities
- COMBO conservative better than COMBO stratified

Combination paradigm	Δ AI-2D	Δ AI-DBT
COMBO conservative	+0.027 (0.018 – 0.036)	0.013 (0.007 – 0.019)
COMBO stratified	+ 0.013 (0.001 – 0.025)	- 0.001 (-0.005 – 0.003)

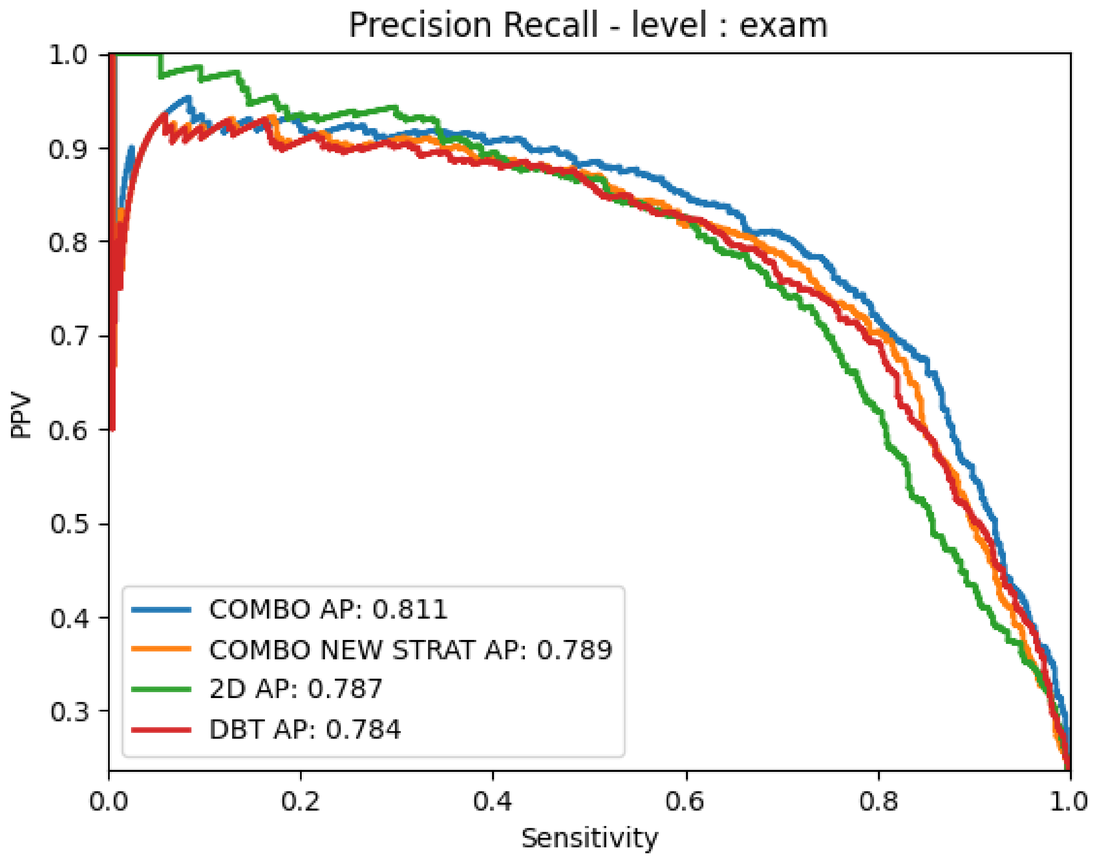
RESULTS – LROC



- Both combination paradigms have higher detection performance
- More pronounced gain with respect to 2D

Combination paradigm		Δ AI-2D	Δ AI-DBT
COMBO conservative	Calcifications	0.013 (0.004 – 0.028)	0.011 (0.0 – 0.027)
	Soft tissue lesions	0.04 (0.017 – 0.064)	0.017 (-0.003 – 0.041)
COMBO stratified	Calcifications	0.013 (0.004 – 0.028)	0.011 (0.0 – 0.027)
	Soft tissue lesions	0.022 (-0.001 – 0.047)	-0.001 (-0.024 – 0.023)

Results – AP PR



- Better performance on COMBO conservative
- 2D still has the highest value of sensitivity without making False Positive results

Combination paradigm	Δ AI-2D	Δ AI-DBT
COMBO conservative	0.022 (-0.013 – 0.056)	0.028 (-0.015 – 0.066)
COMBO stratified	-0.001 (-0.036 – 0.035)	0.005 (-0.038 – 0.043)

KEY POINTS

- The combination of modalities outperforms individual 2D and DBT
- The conservative combination paradigm seems to be the most accurate
- Detection performance is improved for both calcifications and soft tissue lesions
- More pronounced improvement for detection with respect to 2D

Future Developments

- Test alternative combination paradigms
- Investigate separately the contribution of FFDM and 2DM
- Test on a screening based population

THANK YOU

spacile@therapixel.com