MamaDino: A Hybrid Vision Model for Breast Cancer 3–Year Risk Prediction

Ruggiero Santeramo, Igor Zubarev, Florian Jug
Human Technopole, Milan, Italy

Motivation

Breast cancer screening programmes increasingly aim to move from uniform screening intervals to risk-adapted, personalized strategies, but accurate short- to medium-term risk estimation remains challenging.

Why imaging-based risk?

- Traditional clinical/demographic risk scores provide moderate discrimination and do not fully exploit mammographic morphology and texture.
- Deep learning models operating on FFDM (e.g., Mirai) achieve strong short- to medium-term risk prediction and have been externally validated.

Key modelling gap

- Many high-performing risk models rely on high-resolution pixel inputs and mostly CNN backbones, with limited explicit modelling of contralateral asymmetry (despite contralateral comparison being central in radiology).

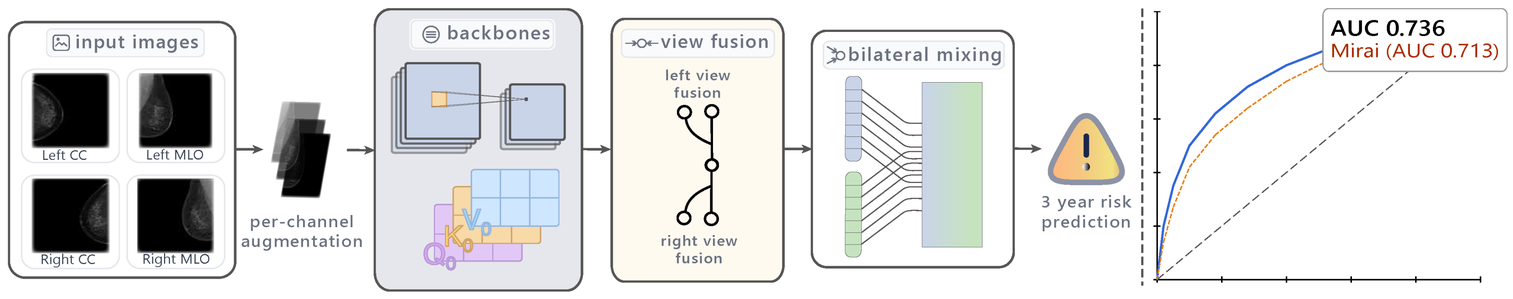
Hypothesis

We hypothesized that, by exploting:

- complementary CNN + transformer inductive biases, and
- explicit bilateral reasoning,

a resolution-efficient hybrid model that combines could match or exceed established high-resolution approaches (e.g., Mirai) while using substantially fewer input pixels and remaining robust across sites/vendors.

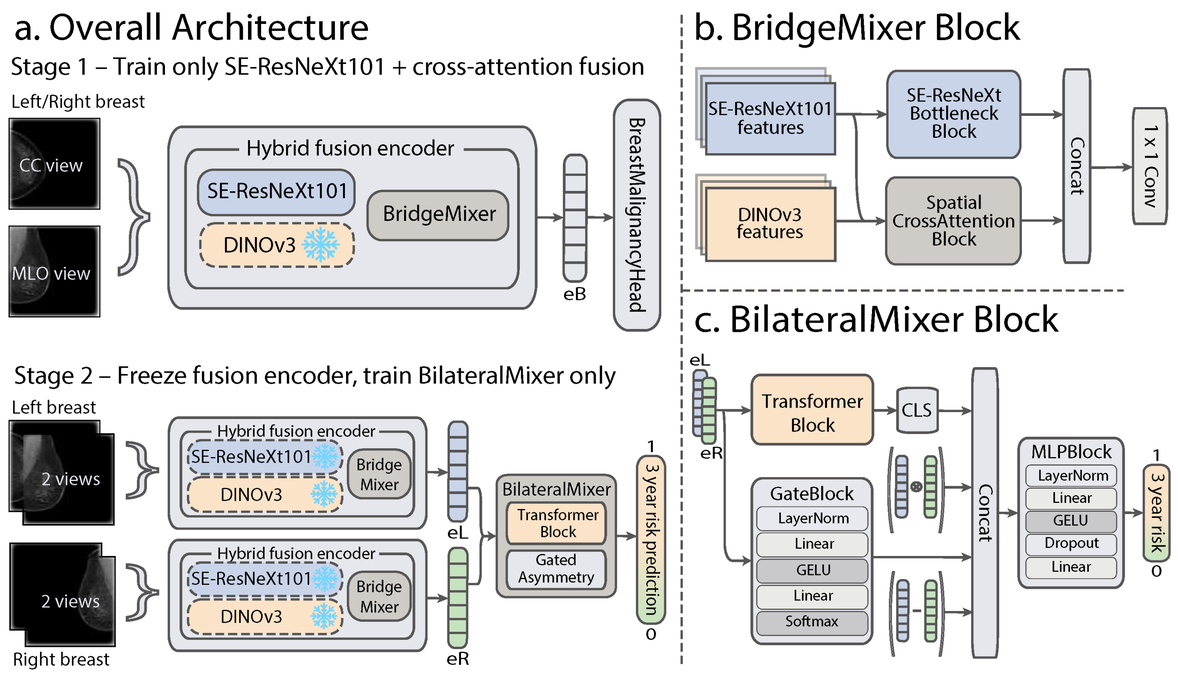
Method



MamaDino predicts 3-year fixed-horizon breast cancer risk from routine screening FFDM using four standard views (L/R × CC/MLO) processed at 512×512.

Two Stage training:

- **Stage 1:** Train a hybrid fusion encoder for breast-level malignancy.
- **Stage 2:** Freeze the encoder; train a patient-level BilateralMixer to fuse left/right breast embeddings.



Results

Resolution-efficient performance:
Despite using 512×512 inputs (≈ 13× fewer pixels than Mirai's 1664×2048), MamaDino achieves comparable or better 3-year risk discrimination on the in-distribution cohort and matches Mirai on the OOD cohort.

Bilateral modelling helps:
Adding the BilateralMixer improves over a breast-level aggregation baseline, supporting the value of explicit contralateral reasoning for patient-level risk prediction.

Stable across subgroups:
Subgroup analyses show consistent performance trends across key strata (e.g., age and scanner vendor/site), suggesting the approach is not driven by a narrow subset of the population or acquisition conditions.

Model	Resolution	In-Distribution Test Set AUC	In-Distribution Test Set (95% CI)	OOD Test Set AUC	OOD Test Set (95% CI)
DINO-only*	512×512	0.621	–	–	–
SEResNeXt-only*	512×512	0.668	–	–	–
Mirai† (Yala et al., 2021)	1664×2048	0.713	(0.686–0.740)	0.676	(0.646-0.707)
MamaDino * (ours)	512×512	0.727	(0.701–0.754)	0.666	(0.635-0.696)
MamaDino † (ours)	512×512	**0.736**	(0.710–0.762)	**0.677**	(0.647-0.707)

* Predictions were at breast-level, and patient risk corresponded to the max of the two breast scores.
† Predictions were at patient-level.

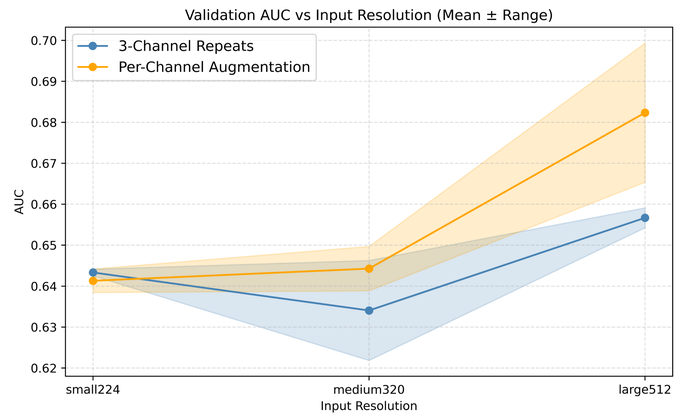
Test Cohorts	In-Distribution Test: MamaDino	In-Distribution Test: Mirai	OOD Test Set: MamaDino	OOD Test Set: Mirai
Age (y)				
< 60	**0.746** (0.710-0.783)	0.727 (0.690-0.764)	0.668 (0.626-0.710)	**0.693** (0.651-0.735)
60–65	**0.705** (0.650-0.760)	0.656 (0.597-0.715)	**0.675** (0.618-0.732)	0.657 (0.596-0.717)
65+	**0.747** (0.696-0.797)	0.743 (0.691-0.794)	**0.691** (0.624-0.758)	0.672 (0.606-0.737)
Ethnicity				
White	0.723 (0.689-0.758)	**0.726** (0.691-0.760)	–	–
Asian	**0.801** (0.723-0.880)	0.660 (0.556-0.764)	–	–
Black/African	0.770 (0.617-0.924)	**0.839** (0.728-0.950)	–	–
Scanner				
GE Med. S.	**0.725** (0.555-0.895)	0.627 (0.443-0.811)	**0.669** (0.618-0.721)	0.622 (0.566-0.677)
Hologic, Inc.	**0.740** (0.713-0.766)	0.722 (0.694-0.749)	0.679 (0.642-0.716)	**0.705** (0.669-0.741)
Siemens	**0.463** (0.217-0.709)	0.457 (0.222-0.691)	–	–
Cancer type**				
DCIS	0.696 (0.636-0.757)	**0.711** (0.650-0.771)	0.659 (0.586-0.731)	**0.689** (0.617-0.761)
Invasive	**0.748** (0.718-0.778)	0.713 (0.682-0.745)	**0.680** (0.647-0.714)	0.673 (0.640-0.707)
Tumour grade**				
G1	**0.755** (0.684-0.826)	0.686 (0.613-0.758)	**0.683** (0.621-0.745)	0.679 (0.614-0.745)
G2	**0.787** (0.751-0.824)	0.756 (0.717-0.796)	**0.695** (0.645-0.745)	0.660 (0.609-0.711)
G3	**0.641** (0.572-0.709)	0.633 (0.561-0.705)	0.664 (0.600-0.728)	**0.686** (0.623-0.749)

* Data are medians, with IQRs in parentheses.
** Cancer data from 3-year follow-up.

Per Channel Augmentation

Mammograms are single-channel, we thus create a pseudo-RGB input by applying different photometric transforms per channel (e.g., independent brightness jitter, contrast jitter, and a CLAHE-enhanced channel) before stacking into 3 channels.

In ablations, this outperforms simple 3-channel replication(commonly used) across three different resolutions (224x224; 320x320; and 512x512), with the largest gain at 512×512.



Data Cohorts

Training cohort:
53,883 women from OPTIMAM (UK), standard bilateral two-view FFDM, FOR-PRESENTATION images.

Test Cohorts:
matched case-control cohorts:

- in-distribution (4 UK sites) - 525 cases + 1,050 controls
- external OOD (unseen site) - 376 cases + 1,504 controls

Conclusion & Future Works

Event at 512×512 (≈ 13× fewer pixels than Mirai), MamaDino matches or slightly exceeds Mirai's discrimination on in-distribution and OOD UK cohorts, with stable performance across demographic, tumour, and scanner strata.
Hence, we believe that increasing input resolution, in conjunction with MamaDino's powerful model structure is likely to yield additional gains.

Limitations:

1. UK-only screening data;
2. Under-representation of Siemens and non-White groups limits subgroup certainty;
3. Single 3-year horizon.

Future Works:

1. Multinational evaluation
2. Extend BilateralMixer to longitudinal and multimodal settings.
3. Systematic ensembles combining MamaDino with Mirai and/or clinical/genomic scores for risk-adapted screening.

References

Yala, Adam, et al. "Toward robust mammography-based models for breast cancer risk." Science Translational Medicine 13.578 (2021): eaba4373.
Santeramo, Ruggiero, et al. "Are better AI algorithms for breast cancer detection also better at predicting risk? A paired case–control study." Breast Cancer Research 26.1 (2024): 25.
Santeramo, Ruggiero, Igor Zubarev, and Florian Jug. "MamaDino: A Hybrid Vision Model for Breast Cancer 3-Year Risk Prediction." Medical Imaging with Deep Learning. 2026.