Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

PRISM-Net: Patient-specific reference-guided inter-breast symmetry matching for three-class breast DCE-MRI classification

Boya Zhang
Nankai University
Shuaiwen Zhou
Beijing University of Posts and Telecommunications
Di Kong
Tsinghua University
Mingxu Wang
Chinese PLA General Hospital
Wenbiao Du
Beijing Institute of Technology
Yiman Zhong
Beihang University
Yuexin Duan
Nankai University
Xiawei Yue
Nankai University
Liuquan Cheng
Chinese PLA General Hospital
Xiru Li
2468li@sina.com

Chinese PLA General Hospital

Research Article

Keywords:

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11425380/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

PRISM-Net: Patient-specific reference-guided inter-breast symmetry matching for three-class breast DCE-MRI classification

Boya Zhang[1,2†], Shuaiwen Zhou[4†], Di Kong[6,2], Mingxu Wang[5], Wenbiao Du[7,2], Yiman Zhong[8,2], Yuexin Duan[1,2], Xiawei Yue[1,2], Liuquan Cheng[5*], Xiru Li[1,3*]

[1]Nankai University, No. 94 Weijin Road, Nankai District, Tianjin, 300071, China.
[2]Zhongguancun Academy, Beijing, China.
[3]The First Medical Center of Chinese PLA General Hospital, Beijing, China.
[4]Beijing University of Posts and Telecommunications, Beijing, China.
[5]The Six Medical Center of Chinese PLA General Hospital, Beijing, China.
[6]Tsinghua University, Beijing, China.
[7]Beijing Institute of Technology, Beijing, China.
[8]Beihang University, Beijing, China.

*Corresponding author(s). E-mail(s): 13910209982@139.com; 2468li@sina.com;
Contributing authors: boyazhang@mail.nankai.edu.cn; qdzsw666@bupt.edu.cn; kd24@mails.tsinghua.edu.cn; wmxwork999@163.com; duwenbiao@bit.edu.cn; yeeman@buaa.edu.cn; duanyuexin@mail.nankai.edu.cn; yxw@mail.nankai.edu.cn;
†These authors contributed equally to this work.

Abstract

Breast DCE-MRI classification of breast sides as no lesion, benign, or malignant is complicated by substantial patient-specific variation in breast anatomy, tissue composition, and physiological enhancement. Although radiologists routinely use

the contralateral breast as an internal reference, computational bilateral comparison remains challenging because corresponding contralateral regions are not anatomically identical and bilateral differences are not necessarily pathological. We propose PRISM-Net, a registration-free bilateral framework that treats the contralateral breast as an approximate patient-specific reference. PRISM-Net learns complementary target-appearance and reference-conditioned deviation representations. Adaptive patch-level matching constructs content-dependent contralateral references in a shared feature space without voxel-wise registration, while local-contrast reweighting emphasizes focal deviations over diffuse bilateral mismatch. The resulting deviation features are integrated with intrinsic target-side appearance through residual cross-attention and aggregated across slices for breast-side classification. On ODELIA, PRISM-Net achieved Macro AUC, Micro AUC, and quadratic weighted kappa values of $\mathbf{84.11 \pm 2.33}$, $\mathbf{90.64 \pm 1.61}$, and $\mathbf{60.94 \pm 5.64}$ on the in-distribution test set, and $\mathbf{68.51 \pm 4.54}$, $\mathbf{80.74 \pm 2.68}$, and $\mathbf{43.45 \pm 7.10}$ on the held-out institution, respectively, yielding the highest point estimates among the evaluated methods across the primary metrics. PRISM-Net also demonstrated favorable performance on an independent institutional cohort and background-complexity subsets, while ablation experiments supported the contributions of adaptive reference matching and focal deviation modeling. These findings support patient-specific contralateral reference learning as a clinically grounded strategy for breast DCE-MRI classification under heterogeneous background conditions.

Keywords: Breast DCE-MRI , Patient specific contralateral reference learning , Radiological prior knowledge , Medical image classification , Background parenchymal enhancement

1 Introduction

Breast cancer remains a leading cause of cancer-related mortality among women worldwide (Bray et al. 2024; Siegel et al. 2024). Dynamic contrast-enhanced breast magnetic resonance imaging (DCE-MRI) is highly sensitive for breast cancer detection and is widely used for screening, diagnosis, staging, and treatment assessment (Abdullah et al. 2025; Kuhl 2024; Hirsch et al. 2025). Nevertheless, its interpretation remains challenging because breast anatomy, tissue composition, physiological activity, and imaging appearance vary substantially across patients. These patient-specific background signals may resemble suspicious findings or reduce the conspicuity of true abnormalities, contributing to both over-calling and under-calling (Giess et al. 2014; Ray et al. 2018; Bechyna and Baltzer 2025). Separating lesion-related evidence from background-signal interference is therefore a fundamental problem in breast DCE-MRI interpretation.

Bilateral comparison provides a clinically intuitive strategy for reducing background-signal interference in breast DCE-MRI. Also, previous computer-aided and multiparametric imaging studies indicated that bilateral asymmetry and contralateral healthy-tissue biomarkers can provide complementary diagnostic information (Yang et al. 2014; Leithner et al. 2019). This prior is particularly informative in

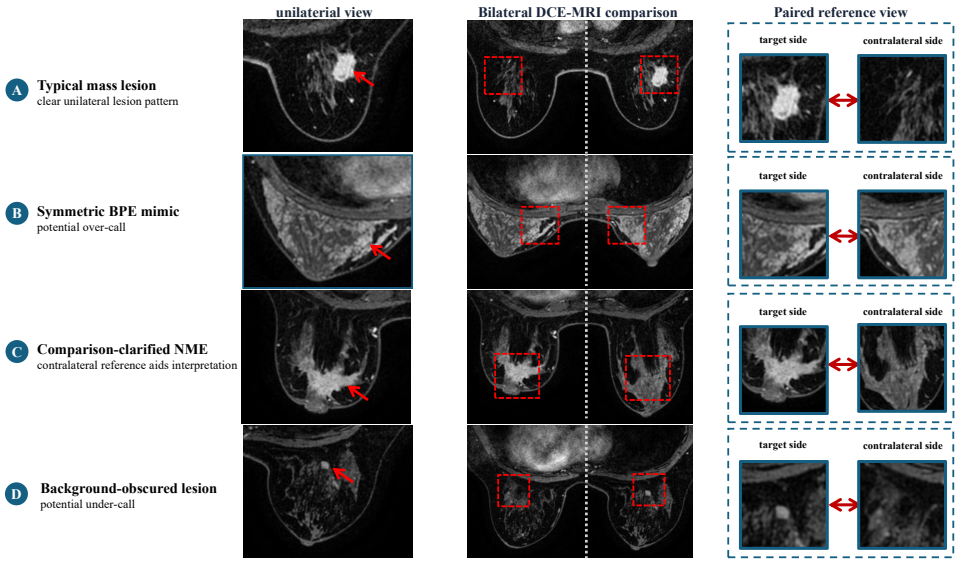


Fig. 1: Representative clinical examples demonstrating bilateral comparison in breast DCE-MRI. Each row presents unilateral and bilateral views, with paired target-side and contralateral reference crops. (A) Typical mass lesion in a 56-year-old woman. An irregular, mildly lobulated right upper-outer breast mass showed asymmetric heterogeneous enhancement. The lesion was assessed as BI-RADS 5, and biopsy confirmed invasive carcinoma. (B) Symmetric BPE mimic in a 23-year-old woman with dense fibroglandular tissue and moderate symmetric BPE. Suspicious enhancement on the unilateral view had a corresponding contralateral pattern, supporting symmetric BPE; both breasts were assessed as BI-RADS 1(no suspicious enhancement). (C) Comparison-clarified NME in a 44-year-old woman. Asymmetric progressive enhancement in the left upper breast comprised NME with multiple small satellite masses. Bilateral comparison showed no corresponding contralateral enhancement, clarifying the finding as NME. The lesion was assessed as BI-RADS 5, and pathology confirmed invasive carcinoma. (D) Background-obscured lesion in a 39-year-old woman with dense fibroglandular tissue and moderate symmetric BPE. A 10-mm right upper-inner breast mass could be overlooked among surrounding enhancement. The absent contralateral counterpart facilitated detection; the right breast was assessed as BI-RADS 3, and pathology confirmed fibroadenoma. BPE, background parenchymal enhancement; DCE-MRI, dynamic contrast-enhanced magnetic resonance imaging; NME, non-mass enhancement.

DCE-MRI, where contrast uptake occurs in both lesions and normal fibroglandular tissue. In the BI-RADS lexicon, fibroglandular tissue (FGT) characterizes the amount and distribution of glandular tissue, whereas background parenchymal enhancement (BPE) describes its physiological enhancement after contrast administration. FGT and BPE therefore provide complementary descriptors of background-signal complexity, which has been associated with reduced diagnostic performance (Bechyna and Baltzer

2025; American College of Radiology 2013; Liao et al. 2020; Bahl 2025). As illustrated in Fig. 1, using the contralateral breast as a patient-specific internal reference can help dismiss bilaterally similar mimics, clarify abnormalities without a contralateral counterpart, and reveal lesions obscured by surrounding background. The diagnostic value of bilateral comparison therefore lies not in asymmetry detection alone, but in leveraging shared physiological characteristics to approximate a patient-specific baseline for interpreting suspicious deviations.

Although deep learning studies in mammography and digital breast tomosynthesis have demonstrated the diagnostic value of bilateral information (Shimokawa et al. 2023; Wang et al. 2023; Zeng et al. 2026; Pelluet et al. 2025), the clinical prior of bilateral comparison remains underused in deep learning models for breast MRI (Hao et al. 2020; Lo Gullo et al. 2024; Gravina et al. 2021, 2024; Oviedo et al. 2025; Zhang et al. 2023). At the modeling level, many methods remain lesion-centric, single-sided, or only implicitly bilateral. Fine-grained bilateral modeling requires correspondence between target regions and contralateral tissue. Conventional correspondence estimation often relies on deformable registration, which remains challenging in breast DCE-MRI because of nonrigid breast motion and enhancement-related temporal intensity changes (Zheng et al. 2015; Truhn et al. 2019; Antropova et al. 2019; Hizukuri et al. 2021; Sun et al. 2023). At the task level, most breast MRI classification models have been developed in lesion-enriched binary settings. However, as breast MRI expands across screening and diagnostic applications, examinations without suspicious findings constitute a clinically relevant part of the case spectrum (Marcon et al. 2024; Verburg et al. 2022). The recently released multicenter ODELIA resource provides breast-side labels for *no-lesion*, *benign*, and *malignant* findings, yet studies in this setting remain scarce (Müller-Franzes et al. 2025; ODELIA-AI 2025). Taken together, the central gap is the lack of an end-to-end breast MRI framework that jointly learns registration-free, spatially adaptive cross-breast correspondence, treats the contralateral breast as an approximate patient specific reference rather than a normal template, and distinguishes diagnostically relevant focal deviations from diffuse or bilaterally shared variation while preserving intrinsic target side appearance. This gap is particularly consequential for breast side three class classification, where the model must distinguish *no lesion*, *benign*, and *malignant* findings under heterogeneous anatomical and physiological backgrounds without assuming that a target lesion has already been identified.

To address this gap, we formulate bilateral DCE-MRI analysis as patient-specific reference-guided interpretation, in which the contralateral breast provides approximate internal context for identifying localized deviations from shared anatomical and physiological background patterns. We therefore propose PRISM-Net, a patient-specific reference-guided inter-breast symmetry matching for three-class breast DCE-MRI classification without explicit anatomical registration. PRISM-Net encodes the target and contralateral breasts using a shared-weight backbone to establish a common feature space. Bidirectional patch-level cross-breast matching then retrieves an adaptive contralateral reference for each target region under anatomical variability. The model integrates target features, matched references, directional differences, and cross-breast interactions to construct reference-conditioned representations of bilateral

deviations. A local-contrast reweighting mechanism further suppresses diffuse or bilaterally shared variation while emphasizing diagnostically relevant focal asymmetries. The resulting deviation evidence is fused with intrinsic target-side appearance and aggregated across slices to classify each breast as no-lesion, benign, or malignant. By converting clinical bilateral comparison into a learnable diagnostic prior, PRISM-Net provides a unified framework for patient-specific reference modeling and breast-side classification..

We evaluate PRISM-Net across the public multicenter ODELIA dataset (ODELIA-AI 2025) and independent institutional DCE-MRI cohorts, all with breast-side labels for *no-lesion*, *benign*, and *malignant* findings. Bilateral no-lesion examinations supported by clinical BI-RADS assessments are retained to represent lesion-free settings, while dedicated BPE- and FGT-complexity cohorts capture challenging backgrounds. Together, this multi-cohort design provides a clinically oriented assessment of three-class performance, cross-cohort generalizability, and robustness to background-signal complexity.

The contributions of this work are as follows.

1. We formulate breast-side DCE-MRI classification as a patient-specific bilateral reference learning problem. The contralateral breast is treated as an approximate internal reference for distinguishing shared anatomical and physiological background patterns from localized lesion-related deviations across *no-lesion*, *benign*, and *malignant* findings.
2. We propose PRISM-Net, a registration-free bilateral framework that learns complementary target-appearance and reference-conditioned deviation representations. Its adaptive patch-level matching mechanism constructs content-dependent contralateral references in a shared feature space, avoiding fixed mirrored correspondence and deformable registration.
3. We introduce local-contrast focal deviation modeling and appearance-preserving fusion. The former emphasizes spatially localized target-to-reference deviations over diffuse bilateral mismatch, whereas the latter integrates these cues with intrinsic target-side appearance through a residual pathway for volumetric classification.
4. We evaluate PRISM-Net on the public multicenter ODELIA dataset, a held-out external ODELIA institution, and independent institutional cohorts including BPE- and FGT-complexity subsets. The results show favorable discrimination and ordinal agreement relative to the evaluated baselines, and ablation analyses support the contributions of adaptive contralateral matching and focal deviation modeling.

2 Related work

2.1 Deep learning in breast MRI diagnosis

Existing deep learning research in breast MRI has predominantly focused on lesion-level or binary diagnostic tasks, commonly using localized lesion regions or unilateral breast inputs.

Deep learning in breast MRI has predominantly focused on lesion-level characterization. CNNs, recurrent networks, projection-based models, and multi-task

architectures have been used to distinguish benign from malignant lesions from manually segmented, radiologist-selected, or automatically detected regions. Representative approaches include multiparametric CNN classification, four-dimensional CNN–LSTM aggregation, deep-feature maximum-intensity projections, and joint lesion segmentation and classification (Truhn et al. 2019; Antropova et al. 2019; Hizukuri et al. 2021; Hu et al. 2021; Sun et al. 2023). DAE-CNN and its PBPK-informed extension further incorporate contrast-agent disentanglement and physiologically constrained data augmentation to improve lesion representation under limited training data (Gravina et al. 2021, 2024). Despite differences in architecture and input representation, these methods generally assume that a target lesion has already been identified.

Existing breast MRI diagnostic models also predominantly adopt binary task formulations. Lesion-level studies commonly distinguish benign from malignant findings, whereas breast-level or volume-level models usually predict malignant versus non-malignant outcomes (Müller-Franzes et al. 2025; Hirsch et al. 2025). Such formulations either exclude breasts without detectable lesions or combine no-lesion and benign cases within a single negative category. The ODELIA dataset and benchmark provide a clinically relevant task setting by assigning no-lesion, benign, and malignant labels to each breast in multicentre DCE-MRI examinations (Müller-Franzes et al. 2025; ODELIA-AI 2025). Its Medical Slice Transformer baseline aggregates slice-level features from pre-contrast T1-weighted, first post-contrast subtraction, and T2-weighted volumes, establishing breast-side three-class diagnosis as a direct precedent (Müller-Franzes et al. 2025).Because the two breasts are treated as separate samples, the model does not explicitly learn cross-breast interactions.

Taken together, current breast MRI deep-learning methods remain largely shaped by lesion-level inputs and binary label spaces. Model development would benefit from prediction units and label definitions that better reflect clinical diagnostic tasks. Breast-side three-class classification addresses this requirement by jointly distinguishing no-lesion, benign, and malignant states without presuming that a target lesion has already been identified.

2.2 Bilateral comparison and symmetry-aware learning in breast imaging

Bilateral comparison is a routine interpretive strategy in breast imaging, allowing local morphology and enhancement to be evaluated against a patient-specific internal reference. Because the contralateral breast may itself contain abnormalities and bilateral anatomy is affected by natural asymmetry, nonrigid deformation, positioning variation, and heterogeneous background enhancement, bilateral symmetry is an approximate prior and direct voxel-wise correspondence is unreliable (Alterson and Plewes 2003; Hennessey et al. 2014; Jansen et al. 2011; Liao et al. 2020; Zeng et al. 2026). Early breast MRI studies quantified bilateral symmetry using registration-free handcrafted three-dimensional descriptors or predefined whole-breast kinetic differences, but remained limited to global representations (Alterson and Plewes 2003; Yang et al. 2014).

In subsequent breast MRI deep learning studies, bilateral information has not been entirely absent, but it was often incorporated implicitly. Zhang et al. proposed

a two-stage framework using Mask R-CNN for suspicious lesion detection followed by ResNet50 for malignancy classification in breast MRI (Zhang et al. 2023). In the detection stage, the model used pre-contrast images and subtraction images of the left and right breasts as inputs, allowing breast symmetry to be considered within the whole-breast context. However, the method did not explicitly establish correspondence between the two breasts, nor did it construct asymmetry features, a contralateral reference branch, bilateral attention, or a dedicated bilateral fusion module. The subsequent classification stage mainly focused on local candidate regions detected by Mask R-CNN and used DCE parametric maps for lesion characterization. Therefore, this type of work suggested that breast MRI deep learning had recognized the importance of bilateral information, but such information was mainly incorporated as global contextual input. Within BreastMRI-FCDD, which performed breast-level binary cancer detection on DCE-MRI subtraction maximum intensity projections, Oviedo et al. (Oviedo et al. 2025) further investigated FCDD-Symmetric, in which the contralateral breast representation defined the normal-class center for anomaly scoring. However, its projection-level, center-based formulation did not explicitly establish spatially adaptive correspondence between bilateral three-dimensional feature volumes or use spatially matched contralateral features for background signal baseline normalization.

In mammography, digital breast tomosynthesis, and other breast imaging modalities, bilateral information has been more explicitly embedded into deep learning models. Earlier studies introduced distortion-insensitive, logic-guided bilateral comparison and paired asymmetry modeling for cancer localization and classification, while BilAD subsequently extended bilateral asymmetry learning to DBT (Shimokawa et al. 2023; Liu et al. 2019; Guan et al. 2020). DisAsymNet attempted to disentangle asymmetric abnormalities from symmetric breast structures in bilateral mammograms, reconstructing abnormality-removed symmetric images to improve interpretability (Wang et al. 2023). A recent bilateral information-guided Vision Transformer framework further proposed a registration-free strategy for mammography, directly concatenating bilateral images and introducing soft spatial prompts to model cross-breast structural differences without explicit image registration (Zeng et al. 2026). BiGAM-Net also used asymmetry-aware fusion and inspectable gating mechanisms to quantify model reliance on intrinsic lesion features and contralateral asymmetry evidence in paired breast imaging (Haque et al. 2026). These studies collectively suggest that bilateral information can serve as a structural medical prior to improve diagnostic performance and interpretability.

These studies support the value of explicit bilateral reasoning, while their two-dimensional or quasi-two-dimensional formulations provide limited coverage of the deformable volumetric anatomy and heterogeneous enhancement of breast DCE-MRI. Registration-free, spatially adaptive cross-breast matching coupled with patient-specific background normalization therefore remains the central methodological gap.

3 Methods

PRISM-Net translates radiological bilateral comparison into a registration-free framework that conditions each target breast on an adaptively matched contralateral

reference, as illustrated in Fig. 2. We first formulate bilateral breast-side classification and introduce the clinical rationale for patient-specific reference learning in Sec. 3.1. We then establish a shared bilateral representation space in Sec. 3.2 and construct adaptive patch level contralateral references without voxel-wise registration in Sec. 3.3. Next, Sec. 3.4 introduces focal bilateral-deviation modeling to attenuate diffuse mismatch and summarize localized reference-conditioned evidence. Finally, Sec. 3.5 integrates this evidence with target-side appearance and aggregates it across slices for breast-side three-class classification.

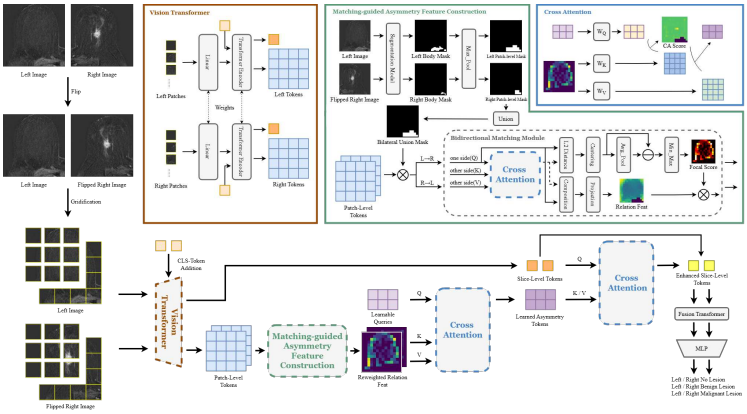


Fig. 2: Overall architecture of the proposed PRISM-Net for breast DCE-MRI classification.

3.1 Bilateral problem formulation

PRISM-Net is motivated by the clinical use of bilateral comparison in breast DCE-MRI. Rather than interpreting a target-side finding solely from its absolute appearance, radiologists compare it with contralateral tissue to determine whether it represents a meaningful deviation from patient-specific anatomical and physiological background patterns. Because the contralateral breast is neither anatomically identical to the target breast nor guaranteed to be lesion-free, we treat it as an approximate patient-specific reference rather than a normal template.

Operationalizing this reference-based interpretation raises three modeling challenges. First, corresponding bilateral regions are not necessarily located at identical mirrored coordinates because of nonrigid breast deformation, patient positioning, and anatomical asymmetry, making direct left-right subtraction unreliable. Second, bilateral mismatch may arise from moderate/marked BPE, heterogeneous/dense FGT, or other natural asymmetry rather than local pathology. Third, bilateral asymmetry alone is insufficient for three-class diagnosis, because distinguishing no-lesion, benign, and malignant findings also requires intrinsic target-side appearance and cross-slice volumetric context.

Accordingly, PRISM-Net learns two complementary representations for each breast: a target-appearance representation that preserves intrinsic diagnostic features, and a reference-conditioned deviation representation that characterizes how local target features differ from adaptively matched contralateral context.

Formally, let $\mathbf{X}^L$ and $\mathbf{X}^R$ denote the paired left and right breast DCE-MRI volumes from one examination, with $\mathbf{X}^L, \mathbf{X}^R \in \mathbb{R}^{D\times H\times W}$, where D is the number of axial slices and H and W are the in-plane spatial dimensions. Let $\mathcal{S} = \{L, R\}$ denote the set of breast sides. For each side $a \in \mathcal{S}$, the diagnostic label is $y^a \in \{0, 1, 2\}$, corresponding to no-lesion, benign, and malignant findings, respectively. Given the paired bilateral volumes, PRISM-Net learns a function f_θ that jointly predicts the side-specific class probability distributions $(\boldsymbol{\pi}^L, \boldsymbol{\pi}^R) = f_\theta(\mathbf{X}^L, \mathbf{X}^R)$, where $\boldsymbol{\pi}^a \in [0, 1]^3$ and $\sum_{c=0}^{2} \pi_c^a = 1$ for each $a \in \mathcal{S}$.

3.2 Shared bilateral representation

To make the approximate contralateral reference computationally usable, the bilateral volumes must first be mapped into a comparable semantic space. PRISM-Net horizontally flips the right breast volume, denoted by $\bar{\mathbf{X}}^R = \mathrm{Flip}(\mathbf{X}^R)$, to place both breasts in a common orientation. This operation provides only a coarse spatial prior and does not assume voxel-wise correspondence. The left volume $\mathbf{X}^L$ and the flipped right volume $\bar{\mathbf{X}}^R$ are then encoded slice-wise using a shared DINOv3 ViT-S/16 backbone (Siméoni et al. 2025). Weight sharing ensures that bilateral features are represented in the same semantic space, thereby enabling subsequent feature-space matching.

For the d-th slice and breast side $a \in \mathcal{S}$, the shared encoder produces a global slice token and a matrix of local patch tokens:

$$\mathbf{s}_d^a \in \mathbb{R}^C, \qquad \mathbf{P}_d^a \in \mathbb{R}^{N\times C}, \tag{1}$$

where $\mathbf{s}_d^a$ preserves the global target-side appearance, $\mathbf{P}_d^a$ contains the corresponding local patch representations, C is the feature dimension, and N is the number of image patches. The patch-level representations are subsequently used to construct adaptive contralateral references, whereas the global slice tokens are retained for appearance-preserving classification.

3.3 Registration-free adaptive contralateral reference matching

With bilateral features represented in a common semantic space, the next step is to construct a patient-specific reference for each target patch. Because identical mirrored coordinates do not guarantee anatomical correspondence, direct patch-wise subtraction may confound pathological differences with variations in breast shape, positioning, and deformation. PRISM-Net therefore allows each target patch to retrieve an adaptive contralateral reference through soft feature-space matching within the paired axial slice, without requiring voxel-wise registration.

Before matching, unreliable regions are excluded using binary masks. Let $\mathbf{M}^L, \mathbf{M}^R \in \{0, 1\}^{D\times H\times W}$ denote the exclusion masks corresponding to the original bilateral volumes, where 1 indicates a region excluded from matching. The right-side

mask is flipped together with the right breast volume, $\bar{\mathbf{M}}^R = \mathrm{Flip}(\mathbf{M}^R)$, and the bilateral exclusion mask in the common orientation is defined as $\mathbf{M}^U = \mathbf{M}^L \vee \bar{\mathbf{M}}^R$. For slice d, the valid patch set is

$$\mathcal{V}_d = \left\{ i \;\middle|\; \max_{(u,v)\in\mathcal{O}_i} M^U_{d,u,v} = 0 \right\}, \tag{2}$$

where $\mathcal{O}_i$ denotes the spatial support of patch i. This filtering ensures that bilateral matching is performed only between patches that are reliable on both sides.

For one matching direction, let $a \in \mathcal{S}$ denote the target side and let $b \in \mathcal{S}$, $b \neq a$, denote the contralateral reference side. Based on the valid patch-token matrices, the matching weights and adaptive contralateral references are computed as

$$\begin{aligned} \mathbf{A}_d^{a\leftarrow b} &= \mathrm{softmax}\left(\frac{(\mathbf{P}^a_{d,\mathcal{V}_d}\mathbf{W}_Q)(\mathbf{P}^b_{d,\mathcal{V}_d}\mathbf{W}_K)^\top}{\sqrt{d_k}} \right), \\ \mathbf{R}_d^a &= \mathbf{A}_d^{a\leftarrow b}(\mathbf{P}^b_{d,\mathcal{V}_d}\mathbf{W}_V), \end{aligned} \tag{3}$$

where d_k is the projected query–key dimension and $\mathbf{W}_Q$, $\mathbf{W}_K$, and $\mathbf{W}_V$ are learnable projection matrices. The softmax operation is applied over the valid contralateral patches, so each row of $\mathbf{A}_d^{a\leftarrow b}$ describes the correspondence distribution between one target patch and the candidate reference patches. Consequently, each row $\mathbf{r}^a_{d,i}$ of $\mathbf{R}^a_d$ represents a soft, content-adaptive contralateral reference for the corresponding target token $\mathbf{p}^a_{d,i}$, rather than the feature at a fixed mirrored location. The same matching operation is applied in both directions with shared parameters. These adaptively matched reference pairs provide the basis for bilateral relation modeling. However, their discrepancies are not uniformly diagnostic.

3.4 Focal bilateral-deviation modeling

A large target-to-reference difference does not necessarily signify pathology, as it may arise from natural asymmetry in breast morphology, FGT distribution, asymmetric BPE, hormonal or lactational influences, or inflammatory and edematous changes. Therefore, after obtaining the adaptive contralateral reference, PRISM-Net preserves both the target and reference representations and explicitly encodes their bilateral relationship.

For each valid patch $i \in \mathcal{V}_d$, the bilateral relation representation is constructed as

$$\begin{aligned} \mathbf{c}^a_{d,i} &= \mathrm{Concat}\left(\mathbf{p}^a_{d,i}, \mathbf{r}^a_{d,i}, \mathbf{p}^a_{d,i} - \mathbf{r}^a_{d,i}, \mathbf{p}^a_{d,i} \odot \mathbf{r}^a_{d,i}\right), \\ \mathbf{g}^a_{d,i} &= \phi\left(\mathbf{c}^a_{d,i}\right), \end{aligned} \tag{4}$$

where $\mathbf{p}^a_{d,i}$ preserves target-side appearance, $\mathbf{r}^a_{d,i}$ provides patient-specific contralateral context, their difference encodes the directional target-to-reference deviation, and their element-wise product captures feature-wise bilateral interaction. The projection layer $\phi(\cdot)$ maps the concatenated representation to a C-dimensional relation feature $\mathbf{g}^a_{d,i}$.

Although the relation representation captures bilateral differences, diffuse mismatch may reflect background variation rather than focal pathology. PRISM-Net therefore measures how strongly each patch-level discrepancy stands out from its local

spatial neighborhood. Let $\mathcal{N}_d(i) = \mathcal{N}(i) \cap \mathcal{V}_d$ denote the valid patches within the 3×3 neighborhood of patch i. The local-contrast asymmetry score and reweighted relation feature are computed as

$$\begin{aligned} A^a_{d,i} &= \left\| \mathbf{p}^a_{d,i} - \mathbf{r}^a_{d,i} \right\|_2, \\ \bar{A}^a_{d,i} &= \frac{1}{|\mathcal{N}_d(i)|} \sum_{j \in \mathcal{N}_d(i)} A^a_{d,j}, \\ F^a_{d,i} &= \mathrm{ReLU}\left(A^a_{d,i} - \bar{A}^a_{d,i} \right), \\ \widetilde{F}^a_{d,i} &= \frac{F^a_{d,i} - \min_{j \in \mathcal{V}_d} F^a_{d,j}}{\max_{j \in \mathcal{V}_d} F^a_{d,j} - \min_{j \in \mathcal{V}_d} F^a_{d,j} + \epsilon}, \\ \widehat{\mathbf{g}}^a_{d,i} &= \widetilde{F}^a_{d,i} \mathbf{g}^a_{d,i}. \end{aligned} \tag{5}$$

This local-contrast formulation emphasizes discrepancies that are salient relative to their surroundings while attenuating spatially diffuse bilateral mismatch.

Because diagnostically relevant deviations may be sparse or distributed across multiple regions, direct global averaging could dilute their contribution. We therefore stack the reweighted relation features as $\widehat{\mathbf{G}}^a_d \in \mathbb{R}^{|\mathcal{V}_d| \times C}$ and summarize them using a shared set of learnable asymmetry queries $\mathbf{Q}_A \in \mathbb{R}^{K \times C}$:

$$\mathbf{T}^a_d = \mathrm{CA}\left(\mathbf{Q}_A, \widehat{\mathbf{G}}^a_d, \widehat{\mathbf{G}}^a_d \right), \qquad \mathbf{T}^a_d \in \mathbb{R}^{K \times C}. \tag{6}$$

The complete relation-modeling module is applied bidirectionally with shared parameters. The resulting tokens remain side-specific because the target and contralateral reference roles are reversed in the two matching directions. The resulting side-specific asymmetry tokens summarize focal reference-conditioned deviations, but they do not replace the intrinsic appearance of the target breast. We therefore integrate them with the global target-side slice representations for final volumetric classification.

3.5 Appearance-preserving volumetric fusion and classification

The preceding module produces side-specific tokens that summarize focal bilateral deviations. For diagnosis, however, this information should complement rather than replace intrinsic target-side appearance, because symmetric findings are not necessarily normal and asymmetric findings are not necessarily pathological. Moreover, the contralateral breast may itself contain abnormalities. PRISM-Net therefore retains the target-side slice token as the primary appearance representation and introduces the bilateral-deviation tokens through a residual cross-attention pathway.

For side $a \in \mathcal{S}$ and slice d, the reference-conditioned slice representation is computed as

$$\widetilde{\mathbf{s}}^a_d = \mathbf{s}^a_d + \lambda\, \mathrm{CA}\left(\mathbf{s}^a_d, \mathbf{T}^a_d, \mathbf{T}^a_d \right), \tag{7}$$

where λ is a fixed residual scaling hyperparameter that controls the contribution of the bilateral-deviation information. The residual scaling factor λ was fixed at 0.8 throughout training and evaluation. This residual formulation preserves the intrinsic target-side appearance while allowing the slice representation to be selectively conditioned on patient-specific contralateral context.

Because diagnostically relevant evidence may extend across multiple slices, the reference-conditioned slice representations are subsequently aggregated to capture volumetric context. To preserve slice order, a positional embedding $\mathbf{e}_d \in \mathbb{R}^C$ is added to each reference-conditioned slice token. A learnable fusion token $\mathbf{u} \in \mathbb{R}^C$ is then appended to the resulting sequence, and a slice-level Transformer produces the side-specific volumetric representation:

$$\mathbf{z}^a = \text{Transformer}\left([\widetilde{\mathbf{s}}_1^a + \mathbf{e}_1, \ldots, \widetilde{\mathbf{s}}_D^a + \mathbf{e}_D, \mathbf{u}]\right)_{\mathbf{u}}, \tag{8}$$

where $(\cdot)_{\mathbf{u}}$ denotes the output representation associated with the fusion token. The slice positional embeddings enable the Transformer to distinguish slice locations and model ordered inter-slice context. The same slice-level Transformer is applied to both breast sides with shared parameters.

Finally, a shared MLP classifier $h(\cdot)$ maps the volumetric representation to side-specific logits $\boldsymbol{\ell}^a = h(\mathbf{z}^a) \in \mathbb{R}^3$, and the corresponding class probability distribution is $\boldsymbol{\pi}^a = \text{softmax}(\boldsymbol{\ell}^a)$. The classifier parameters are shared between the left and right breasts. PRISM-Net is trained using the summed side-level cross-entropy loss:

$$\mathcal{L} = \sum_{a \in \mathcal{S}} \mathcal{L}_{\text{CE}}\left(\boldsymbol{\ell}^a, y^a\right). \tag{9}$$

Thus, the left and right breast labels provide separate side-level supervision within a single bilateral forward pass.

4 Experiments and results

4.1 Dataset and cohort characteristics

Model development used the public ODELIA breast DCE-MRI dataset. Evaluation was performed using held-out ODELIA data and an independent institutional dataset from the General Hospital of the People's Liberation Army. The selection workflow for the private institutional set is shown in Fig. 3.

Public ODELIA dataset. The public ODELIA dataset comprised 741 examinations from 741 women acquired at six European medical centers between December 2006 and May 2024 (ODELIA-AI 2025). Each breast was independently labeled as no-lesion, benign, or malignant. 291 examinations exhibited no lesions, 146 examinations were diagnosed with benign but no malignant lesions, and 304 examinations had malignant lesions. In total, the dataset includes 978 breasts without lesions, 195 with benign lesions, and 309 with a malignant lesion. The original training (n = 408) and validation (n = 103) partitions were retained, while the test partition was split into an internal set (n = 130) and a center-held-out set (n = 100).

Private institutional dataset. An independent institutional cohort was retrospectively assembled to evaluate generalization and performance under complex background conditions. The primary candidate cohort was drawn from the 2025 institutional archive. As examinations with moderate or marked BPE were underrepresented in this cohort, additional candidates meeting this BPE criterion were retrieved from the 2024 and 2026 archives. Examinations were reviewed by two breast

radiologists and one clinician using DCE-MRI findings, BI-RADS assessments, radiology reports, and pathological results. Examinations were excluded for substantial treatment-related or postoperative changes, breast implants, or missing key DCE-MRI phases. After patient-level deduplication, 171 examinations remained (135 from 2025 and 36 BPE-enriched examinations from 2024 and 2026), yielding 342 evaluable breast sides. Overlapping complexity subsets included 55 examinations with moderate/marked BPE and 92 with heterogeneous/dense FGT; 49 met both criteria, as shown in Fig. 3. 79 examinations exhibited no lesions, 32 women were diagnosed with benign but no malignant lesions, and 60 women had malignant lesions.

Breast-side reference standard The left and right breasts were treated as separate prediction targets and labeled as no-lesion, benign, or malignant. Labels were assigned using a pathology-prioritized hierarchy. Malignant or benign labels were assigned according to ipsilateral pathological findings. A no-lesion label was assigned when no ipsilateral enhancing lesion or pathological diagnosis was identified. For breasts with multiple findings, the most severe diagnosis determined the final label, with malignancy taking precedence over benignity and no-lesion. In total, the dataset includes 240 breasts without lesions, 38 with benign lesions, and 64 with a malignant lesion. Cohort characteristics and acquisition protocols are summarized in Tables 1 and Supplementary Table S1.

4.2 Experimental setup

4.2.1 Data preprocessing

A subtraction volume was generated from the pre-contrast and first post-contrast images, and each bilateral volume was split at the midline. The right breast was horizontally flipped to provide coarse bilateral correspondence without voxel-wise registration. Each breast was resized or cropped to a fixed shape, independently intensity-normalized, and masked to suppress non-breast regions. Paired breast volumes and masks were input jointly; no lesion-level annotations or manual registration were required.

4.2.2 Implementation details

PRISM-Net was implemented in PyTorch and optimized end-to-end using the summed cross-entropy losses of the left and right breasts. The shared DINOv3 ViT-S/16 backbone produced 384-dimensional features. The slice fusion Transformer used eight attention heads, and the shared MLP classifier used a hidden dimension of 384 with a dropout rate of 0.25. We used AdamW with a learning rate of 1.5×10^{-5}, weight decay of 0.01, and a batch size of 8. The learning rate followed a cosine-decay schedule with 10 warm-up epochs and a total schedule length of 120 epochs. Training used mixed precision, with random flipping, rotation, noise perturbation, and Mixup ($\alpha = 0.1$, probability = 0.5) applied only to the training set. Hyperparameters, checkpoints, and ordinal calibration thresholds were selected using the validation set, with early stopping after 35 epochs without improvement. Final probabilities were obtained through validation-weighted checkpoint ensembling and test-time augmentation. All models were evaluated using identical preprocessing and evaluation protocols.

Table 1: Demographic, breast background, and lesion characteristics of the final institutional breast DCE-MRI evaluation pool and its partially overlapping evaluation sets.

Characteristic	Final pool	2025 set	BPE set	FGT set
Patients/examinations, n	171	135	55	92
Evaluable breast sides, n	342	270	110	184
Age, years[a]	47.5 ± 13.0	48.8 ± 13.2	42.0 ± 11.0	44.2 ± 11.8
BPE level[a]				
Minimal/mild, n (%)	116 (67.8)	116 (85.9)	0 (0.0)	43 (46.7)
Moderate/marked, n (%)	55 (32.2)	19 (14.1)	55 (100.0)	49 (53.3)
FGT level[a]				
Fatty/scattered, n (%)	79 (46.2)	76 (56.3)	6 (10.9)	0 (0.0)
Heterogeneous/dense, n (%)	92 (53.8)	59 (43.7)	49 (89.1)	92 (100.0)
Lesion laterality[a]				
No-lesion, n (%)	79 (46.2)	66 (48.9)	23 (41.8)	22 (23.9)
Unilateral, n (%)	82 (48.0)	62 (45.9)	27 (49.1)	64 (69.6)
Bilateral, n (%)	10 (5.8)	7 (5.2)	5 (9.1)	6 (6.5)
Patient-level label[a]				
No-lesion, n (%)	79 (46.2)	66 (48.9)	23 (41.8)	22 (23.9)
Benign, n (%)	32 (18.7)	23 (17.0)	13 (23.6)	28 (30.4)
Malignant, n (%)	60 (35.1)	46 (34.1)	19 (34.5)	42 (45.7)
Breast-side label[b]				
No-lesion breast sides, n (%)	240 (70.2)	194 (71.9)	73 (66.4)	108 (58.7)
Benign breast sides, n (%)	38 (11.1)	28 (10.4)	17 (15.5)	33 (17.9)
Malignant breast sides, n (%)	64 (18.7)	48 (17.8)	20 (18.2)	43 (23.4)

[a] Calculated at the patient/examination level. Age is presented as mean ± standard deviation, and categorical variables as n (%).
[b] Calculated at the individual breast-side level.
BPE, background parenchymal enhancement; FGT, fibroglandular tissue; DCE-MRI, dynamic contrast-enhanced magnetic resonance imaging.

4.2.3 Baseline models

We compared PRISM-Net with representative breast MRI diagnostic models and bilateral imaging models, including BreastMRI-FCDD (Oviedo et al. 2025), PBPK DAE-CNN (Gravina et al. 2024), DisAsymNet (Wang et al. 2023), BiGAM-Net (Haque et al. 2026), and the ODELIA Medical Slice Transformer baseline (Müller-Franzes et al. 2025). All models were adapted to the same breast-side three-class label space and evaluated on matched test cases.

4.2.4 Evaluation metrics

Performance was evaluated at the breast-side level using Macro AUC, Micro AUC, and QWK, which summarized class-balanced discrimination, overall discrimination,

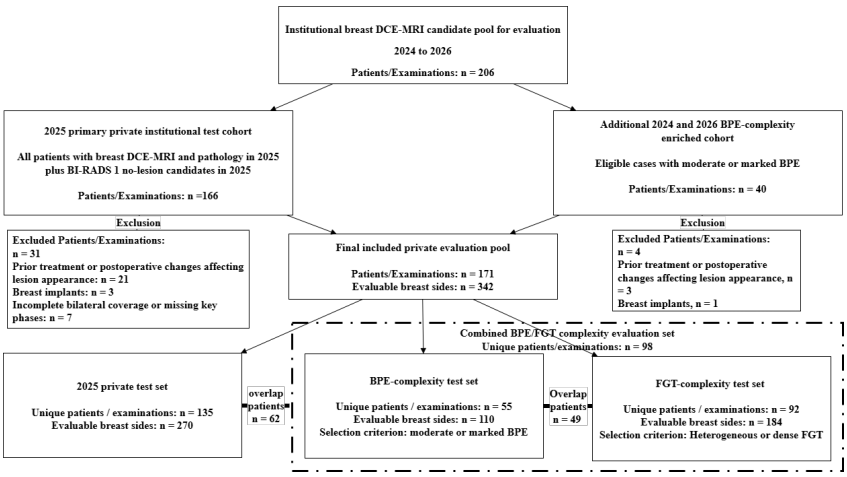


Fig. 3: **Flowchart of institutional patient selection and dataset construction.** The candidate pool comprised 206 patients, each contributing one bilateral DCE-MRI examination. It included a primary 2025 institutional cohort of 166 patients and an additional BPE-enriched cohort of 40 patients from 2024 and 2026, and FGT-complexity membership was subsequently assigned within the final private evaluation pool according to FGT level. After the specified exclusions, the final private evaluation pool comprised 171 patients and 342 evaluable breast sides. Three partially overlapping evaluation sets were defined: the 2025 private test set with 135 patients and 270 breast sides, the BPE-complexity set with 55 patients and 110 breast sides, and the FGT-complexity set with 92 patients and 184 breast sides. The dashed enclosure denotes the union of the two complexity sets, comprising 98 patients. Among them, 62 were also included in the 2025 private test set, and 49 satisfied both complexity criteria. BPE complexity was defined as moderate or marked BPE, whereas FGT complexity was defined as heterogeneous or dense FGT. DCE-MRI, dynamic contrast-enhanced magnetic resonance imaging; BPE, background parenchymal enhancement; FGT, fibroglandular tissue.

and ordinal agreement across no-lesion, benign, and malignant, respectively. All values were expressed on a percentage scale. Results are reported as the point estimate of the final checkpoint ensemble $\pm$ the standard deviation obtained from $2,000$ bootstrap resamples. PRISM-Net was compared with each baseline on matched cases using paired bootstrap testing with $10,000$ iterations and a two-sided significance threshold of p ¡ 0.05. In each bootstrap iteration, resampling was performed at the patient level, with the left and right breast-side predictions from the same patient kept together as one clustered unit. Cohort membership and reference labels were fixed across models, and no test label informed model selection, calibration, or threshold tuning.

Table 2: Performance comparison on the ODELIA public development dataset.

Method	Setting	Macro AUC	Micro AUC	QWK
BreastMRI-FCDD (Oviedo et al. 2025)	ID	$59.90 \pm 6.10^{*}$	$70.20 \pm 9.75^{*}$	$10.77 \pm 9.75^{*}$
	OOD	$54.11 \pm 4.00^{*}$	$73.01 \pm 5.42^{*}$	$2.61 \pm 2.83^{*}$
PBPK DAE-CNN (Gravina et al. 2024)	ID	$56.07 \pm 5.84^{*}$	$68.46 \pm 6.07^{*}$	$3.78 \pm 3.45^{*}$
	OOD	$50.36 \pm 3.69^{*}$	$65.80 \pm 2.39^{*}$	$2.24 \pm 4.01^{*}$
DisAsymNet (Wang et al. 2023)	ID	$69.36 \pm 6.34^{*}$	$81.14 \pm 2.64^{*}$	$24.02 \pm 20.85^{*}$
	OOD	53.71 ± 7.71	75.65 ± 2.74	$12.78 \pm 13.17^{*}$
BiGAM-Net (Haque et al. 2026)	ID	$62.32 \pm 5.85^{*}$	$72.43 \pm 6.87^{*}$	$10.86 \pm 12.13^{*}$
	OOD	$50.00 \pm 3.02^{*}$	$70.87 \pm 4.77^{*}$	$2.96 \pm 3.65^{*}$
ODELIA (Müller-Franzes et al. 2025)	ID	$76.92 \pm 0.48^{*}$	$85.56 \pm 0.91^{*}$	$46.75 \pm 1.62^{*}$
	OOD	60.31 ± 1.07	77.44 ± 2.17	$29.46 \pm 4.58^{*}$
PRISM-Net	**ID**	$\mathbf{84.11 \pm 2.33}$	$\mathbf{90.64 \pm 1.61}$	$\mathbf{60.94 \pm 5.64}$
	OOD	$\mathbf{68.51 \pm 4.54}$	$\mathbf{80.74 \pm 2.68}$	$\mathbf{43.45 \pm 7.10}$

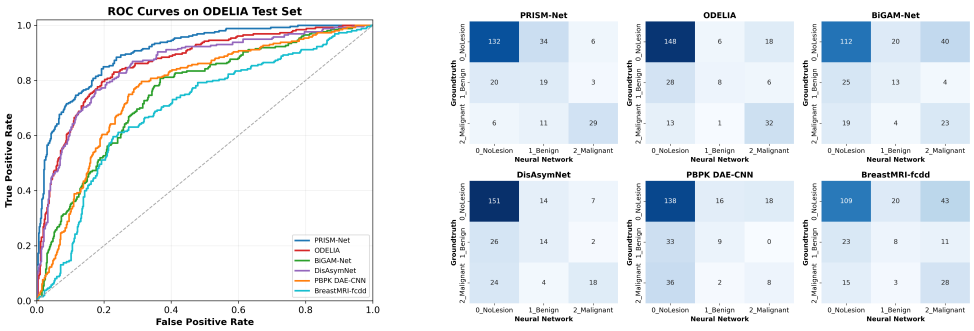
Values are reported as mean ± SD. AUCs are percentages and QWK values are scaled by 100. $^{*}p < 0.05$ versus PRISM-Net. ID, in distribution; OOD, out of distribution.

4.3 Performance on the ODELIA public development dataset

PRISM-Net achieved the best performance on both the ODELIA in-distribution and RSH out-of-distribution sets (Table 2). Its Macro AUC, Micro AUC, and QWK were 84.11±2.33, 90.64±1.61, and 60.94±5.64 in-distribution, and 68.51±4.54, 80.74±2.68, and 43.45 ± 7.10 out-of-distribution. Relative to the strongest baseline, ODELIA, the corresponding gains were 7.19, 5.08, and 14.19 points in-distribution and 8.20, 3.30, and 13.99 points out-of-distribution. ROC curves and confusion matrices are shown in Figs. 4 and 5.

4.4 External validation on the private institutional test set

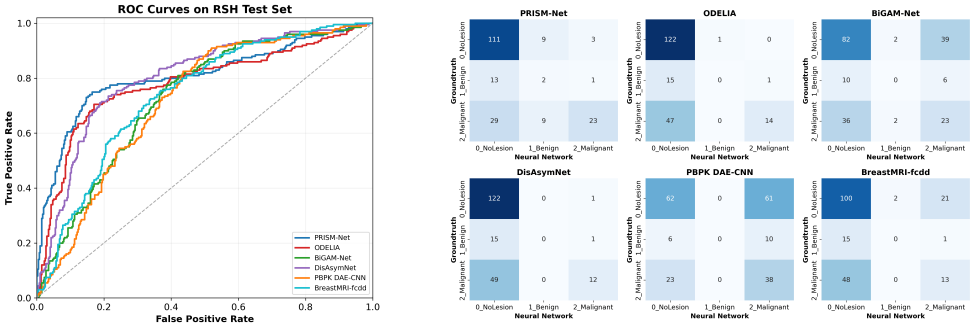
We compared PRISM-Net with DisAsymNet and ODELIA, which were the two strongest baseline methods on the ODELIA public development dataset. The independent 2025 institutional cohort was not used for training, validation, model selection, or tuning. PRISM-Net achieved the highest Macro AUC, Micro AUC, and QWK: 66.85 ± 3.47, 86.38 ± 1.84, and 38.85 ± 6.43, respectively (Table 3). Its QWK exceeded ODELIA and DisAsymNet by 40.86 and 26.41 points, respectively. The corresponding ROC curves and confusion matrices are shown in panels (a) and (d) of Supplementary Fig. S1.



(a) ROC curves.

(b) Confusion matrices.

Fig. 4: **ROC curves and confusion matrices of six models on the ODELIA in-distribution (ID) test set.** The left panel shows the ROC curves, while the right panel presents the corresponding confusion matrices for three-class classification of no-lesion, benign, and malignant.



(a) ROC curves.

(b) Confusion matrices.

Fig. 5: **ROC curves and confusion matrices of six models on the RSH out-of-distribution (OOD) test set.** The left panel shows the ROC curves, while the right panel presents the corresponding confusion matrices for three-class classification of no-lesion, benign, and malignant.

4.5 Performance on the background-complexity external test sets

The comparison was also restricted to DisAsymNet and ODELIA, the two best-performing baselines in the public-dataset evaluation. PRISM-Net also ranked first on both complexity cohorts (Table 3). On the BPE-complexity set, Macro AUC, Micro AUC, and QWK were 66.23 ± 4.90, 82.66 ± 2.91, and 31.48 ± 9.36; on the

FGT-complexity set, they were 67.71 ± 3.38, 80.11 ± 2.37, and 40.88 ± 6.83. Corresponding ROC curves and confusion matrices are shown in panels (b,e) and (c,f) of Supplementary Fig. S1.

Table 3: External validation performance on the private institutional test set and background-complexity test sets.

Cohort	Method	Macro AUC	Micro AUC	QWK
2025 private external test set	ODELIA (Müller-Franzes et al. 2025)	62.66 ± 3.63	$81.64 \pm 1.95^{*}$	$-2.01 \pm 1.12^{*}$
	DisAsymNet (Wang et al. 2023)	63.14 ± 4.21	$82.62 \pm 2.18^{*}$	$12.44 \pm 5.47^{*}$
	PRISM-Net	$\mathbf{66.85 \pm 3.47}$	$\mathbf{86.38 \pm 1.84}$	$\mathbf{38.85 \pm 6.43}$
BPE-complexity test set	ODELIA (Müller-Franzes et al. 2025)	64.87 ± 5.63	78.55 ± 3.28	$1.91 \pm 2.19^{*}$
	DisAsymNet (Wang et al. 2023)	52.69 ± 6.15	$75.90 \pm 3.75^{*}$	$11.75 \pm 7.55^{*}$
	PRISM-Net	$\mathbf{66.23 \pm 4.90}$	$\mathbf{82.66 \pm 2.91}$	$\mathbf{31.48 \pm 9.36}$
FGT-complexity test set	ODELIA (Müller-Franzes et al. 2025)	62.23 ± 3.30	$73.10 \pm 2.26^{*}$	$-0.58 \pm 1.07^{*}$
	DisAsymNet (Wang et al. 2023)	64.64 ± 4.16	$75.53 \pm 2.48^{*}$	$15.21 \pm 5.55^{*}$
	PRISM-Net	$\mathbf{67.71 \pm 3.38}$	$\mathbf{80.11 \pm 2.37}$	$\mathbf{40.88 \pm 6.83}$

Values are mean ± SD. AUCs are percentages; QWK values are scaled by 100. $^{*}p < 0.05$ versus PRISM-Net. BPE, background parenchymal enhancement; FGT, fibroglandular tissue.

4.6 Ablation studies

The full PRISM-Net was best in both settings (Table 4). Removing the bilateral asymmetry branch reduced QWK by 14.56 points in-distribution and 13.40 points out-of-distribution; removing focal score reweighting reduced QWK by 5.55 and 14.38 points, respectively. Macro and Micro AUC also decreased for both variants, supporting the contribution of each component.

Second, we compared different feature extraction backbones while keeping the remaining bilateral learning framework unchanged. With the bilateral framework and evaluation protocol fixed, DINOv3 outperformed ResNet-18, ResNet-50, ViT, and DenseNet-121 on all six metrics (Table 5). DenseNet-121 was the second-best backbone. Relative to DenseNet-121, DINOv3 improved Macro AUC, Micro AUC, and QWK by 2.91, 1.48, and 6.74 points in-distribution and by 2.17, 0.77, and 3.29 points out-of-distribution.

4.7 Interpretability and failure cases

Representative maps (Fig. 6) show that focal reweighting suppresses diffuse or background-related bilateral differences and concentrates responses on localized asymmetry. The no-lesion case showed diffuse activation, the benign case a mild localized response, and the malignant case a compact high-response region corresponding to

Table 4: Ablation study of the main components in PRISM-Net on the ODELIA public development dataset.

Variant	Setting	Macro AUC	Micro AUC	QWK
PRISM-Net w/o bilateral asymmetry branch	ID	$79.75 \pm 1.37^*$	$87.39 \pm 1.34^*$	$46.38 \pm 5.17^*$
	OOD	$64.15 \pm 3.92^*$	$77.88 \pm 3.53^*$	$30.05 \pm 3.88^*$
PRISM-Net w/o focal score reweighting	ID	$79.68 \pm 3.22^*$	$88.12 \pm 2.45^*$	55.39 ± 6.39
	OOD	$64.53 \pm 6.83^*$	79.07 ± 4.58	$29.07 \pm 11.53^*$
PRISM-Net	**ID**	$\mathbf{84.11 \pm 2.33}$	$\mathbf{90.64 \pm 1.61}$	$\mathbf{60.94 \pm 5.64}$
	OOD	$\mathbf{68.51 \pm 4.54}$	$\mathbf{80.74 \pm 2.68}$	$\mathbf{43.45 \pm 7.10}$

Values are mean ± SD. AUCs are percentages; QWK values are scaled by 100. $^*p < 0.05$ versus PRISM-Net. ID, in distribution; OOD, out of distribution.

Table 5: Ablation study of different feature extraction backbones on the ODELIA public development dataset.

Backbone	Setting	Macro AUC	Micro AUC	QWK
ResNet-18 (He et al. 2016)	ID	80.40 ± 1.23	$88.65 \pm 1.35^*$	$51.29 \pm 3.49^*$
	OOD	65.49 ± 4.32	80.34 ± 2.88	$35.84 \pm 5.88^*$
ResNet-50 (He et al. 2016)	ID	$77.95 \pm 2.90^*$	$86.89 \pm 0.81^*$	$48.52 \pm 6.86^*$
	OOD	64.80 ± 3.17	78.95 ± 3.54	$32.25 \pm 6.35^*$
ViT (Dosovitskiy et al. 2021)	ID	$78.15 \pm 2.06^*$	$87.27 \pm 2.23^*$	50.84 ± 6.51
	OOD	$61.49 \pm 4.03^*$	78.61 ± 3.01	$24.27 \pm 8.58^*$
DenseNet-121 (Huang et al. 2017)	ID	81.20 ± 1.67	89.16 ± 1.46	54.20 ± 1.48
	OOD	66.34 ± 4.51	79.97 ± 2.47	40.16 ± 6.79
DINOv3 (Siméoni et al. 2025)	**ID**	$\mathbf{84.11 \pm 2.33}$	$\mathbf{90.64 \pm 1.61}$	$\mathbf{60.94 \pm 5.64}$
	OOD	$\mathbf{68.51 \pm 4.54}$	$\mathbf{80.74 \pm 2.68}$	$\mathbf{43.45 \pm 7.10}$

Values are mean ± SD. AUCs are percentages; QWK values are scaled by 100. $^*p < 0.05$ versus DINOv3. ID, in distribution; OOD, out of distribution..

the enhancing lesion. These maps provide feature-level explanations rather than lesion segmentations.

We further analyzed representative failure cases to characterize the limitations and decision boundaries of PRISM-Net. As shown in Fig. 7, the errors mainly occurred in three challenging scenarios: benign lesions with malignant-like enhancement patterns, bilateral benign lesions that weakened the discriminative value of contralateral symmetry, and small low-suspicion lesions embedded within dense fibroglandular tissue. These cases suggest that the model may still be affected by overlapping enhancement appearances between benign and malignant lesions, confusing bilateral enhancement patterns, and limited lesion conspicuity in complex backgrounds.

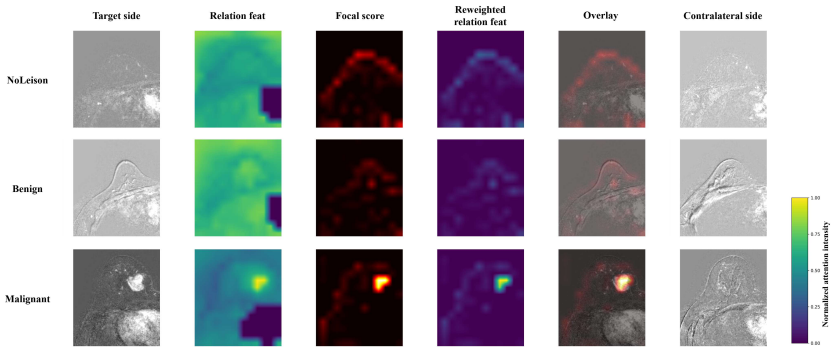


Fig. 6: **Representative interpretability maps generated by PRISM-Net for no-lesion, benign, and malignant cases.** Each row shows one representative case, and each column corresponds to a different stage of the bilateral asymmetry modeling process. From left to right, the columns show the target-side DCE-MRI slice, the relation feature before focal reweighting, the focal score map, the reweighted relation feature, the overlay of the reweighted relation feature on the target-side image, and the contralateral reference side. In the no-lesion case, the focal response is relatively diffuse and does not form a clear lesion-centered activation. In the benign case, the model produces mild and localized asymmetric responses. In the malignant case, the focal score and reweighted relation feature highlight a compact high-response region that spatially corresponds to the suspicious enhancing lesion. These visualizations indicate that PRISM-Net can suppress non-specific bilateral differences and emphasize diagnostically meaningful focal asymmetry.

5 Discussion

This study developed PRISM-Net as a registration-free bilateral learning framework for three-class breast DCE-MRI classification. The main finding is that explicitly modeling the contralateral breast as a patient-specific bilateral reference representation improved both discriminative performance and ordinal diagnostic agreement across public, external, and background-complexity test settings. On the ODELIA in-distribution test set, PRISM-Net achieved a Macro AUC of 84.11 ± 2.33, Micro AUC of 90.64 ± 1.61, and QWK of 60.94 ± 5.64. This advantage was preserved on the RSH out-of-distribution institution. These results suggest that PRISM-Net did not merely improve class separability, but also better preserved the ordered relationship among no-lesion, benign, and malignant.

The observed gains may reflect the way PRISM-Net uses bilateral information for patient-specific background normalization rather than simply treating the contralateral breast as an additional input. Radiological bilateral comparison benefits from shared patient-level physiological and acquisition conditions, but the contralateral breast cannot be assumed to be anatomically identical or disease-free, avoiding the assumption of anatomical symmetry or disease-free status. PRISM-Net therefore treats it as an approximate comparative context rather than a normal template. Shared-weight encoding maps both breasts into a comparable feature space, body-mask-guided

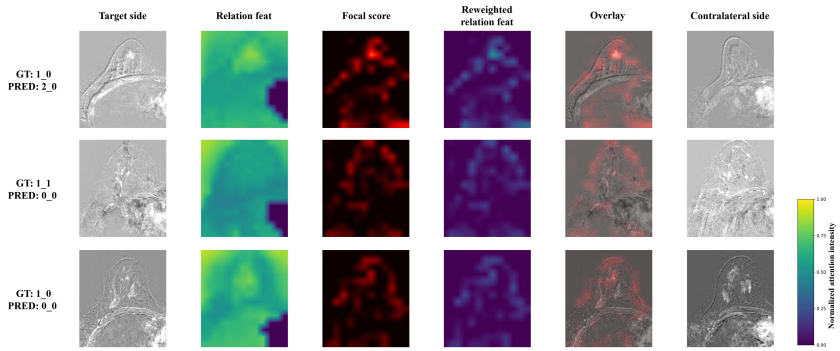


Fig. 7: **Failure-case interpretability and model decision boundaries.** Each row shows one misclassified case. Columns show the target breast, bilateral relation feature, focal score, reweighted relation feature, heatmap, and contralateral breast. GT and PRED denote ground truth and prediction; 0, 1, and 2 represent no-lesion, benign, and malignant, respectively, with suffixes indicating breast side.Row 1 shows a benign BI-RADS 5 lesion predicted as malignant. The irregular, spiculated, heterogeneously enhancing lesion produced a strong focal response, while pathology revealed adenosis, sclerosing adenosis, ductal epithelial hyperplasia, and intraductal papilloma, illustrating overlap between benign proliferation and malignancy.Row 2 shows a benign case in a 39-year-old woman that was predicted as no-lesion. The patient had multiple bilateral fibroadenomas, dense FGT, and minimal BPE. Enhancing masses were present in both breasts, and corresponding enhancement was also observed in the contralateral side, which weakened the discriminative value of bilateral asymmetry for lesion detection. The Row 3 shows an approximately 8-mm fibroadenoma predicted as no-lesion; its small size, circumscribed margin, and dense FGT corresponded to weak focal and reweighted relation responses. These cases highlight three decision boundaries: malignant-like benign proliferation, bilateral lesions with limited asymmetry, and small lesions obscured by complex background signal.

patch filtering restricts matching to anatomically plausible breast tissue, and adaptive patch-level matching reduces dependence on voxel-wise registration despite differences in breast shape, positioning, and deformation. The interpretability analysis in Fig. 6 provides qualitative support for this mechanism. Before focal reweighting, the bilateral relation features contained relatively diffuse responses that could reflect both lesion-related asymmetry and nonspecific background variation. After reweighting, the responses became more spatially concentrated: the malignant example showed a compact high-response region spatially overlapping with the suspicious enhancing region, the no-lesion example lacked a distinct lesion-centered activation, and the benign example exhibited an intermediate pattern. These observations suggest that the bilateral focal response module may help reduce sensitivity to diffuse or bilaterally shared variations while emphasizing localized asymmetric features that are more relevant to classification, which is consistent with the model design. Nevertheless, these visualizations should be interpreted as qualitative feature-level explanations.

The background-complexity analyses further indicate that bilateral comparison may be particularly useful when unilateral enhancement patterns are confounded by complex parenchymal backgrounds. Moderate or marked BPE has previously been associated with reduced diagnostic performance on breast MRI (Bechyna and Baltzer 2025). In the BPE- and FGT-complexity test sets, PRISM-Net consistently achieved the highest performance among evaluated methods, with the most pronounced relative gains observed for QWK, as shown in Table 3. The confusion matrices in Supplementary Fig. S1 further illustrate fewer severe ordinal misclassifications. As shown in Fig. 6 benign case, dense FGT confounded radiological interpretation in a representative clinical case, leading to a BI-RADS 5 assessment. PRISM-Net, however, correctly classified the affected breast as benign, consistent with postoperative histopathology showing sclerosing adenosis with intraductal papilloma. These findings suggest that patient-specific bilateral reference modeling may provide valuable comparative information under complex background conditions. Physiological and hormonal factors, including menstrual cycle variation, pregnancy, and lactation, can contribute to increased BPE and heterogeneous background enhancement, potentially reducing lesion conspicuity and increasing diagnostic uncertainty. By incorporating contralateral breast information as an individualized comparative reference, PRISM-Net better capture clinically relevant asymmetric patterns beyond absolute enhancement characteristics in such challenging examinations. This capability may be relevant for lesions presenting as NME, where the absence of a discrete mass and substantial overlap between malignant, benign, and physiological enhancement patterns often complicate interpretation. Nevertheless, performance variation across these challenging sets may be influenced by limited sample sizes and the inherent uncertainty associated with complex background enhancement patterns. Further study in larger cohorts is required.

The ablation studies clarify not only the contribution of individual modules but also the design logic underlying PRISM-Net. As shown in Table 4, removing the bilateral asymmetry branch consistently degrades performance, indicating that the contralateral breast contributes more than additional image context. Instead, it serves as a patient-specific reference that enables the model to distinguish shared physiological enhancement patterns from localized abnormal deviations. This result supports adaptive reference matching in the feature space as a more appropriate formulation of bilateral comparison than unilateral analysis or simple bilateral feature aggregation. The contribution of focal score reweighting further suggests that bilateral differences are not uniformly informative. Diffuse enhancement variation and background related mismatch can generate nonspecific responses, whereas deviations that are locally salient are more likely to contain discriminative evidence. Focal reweighting therefore acts as an evidence selection mechanism that suppresses diffuse interference and concentrates the relational representation on diagnostically relevant asymmetry. The backbone comparison provides a complementary observation that reliable registration free correspondence depends on the semantic quality of the shared feature space, since more discriminative representations facilitate stable cross breast matching under anatomical, positional, and background variability. Taken together, these findings reveal a coherent computational pathway in PRISM-Net: robust feature encoding establishes a comparable bilateral representation space, adaptive matching

constructs patient specific references, and focal reweighting identifies the most informative reference conditioned deviations. The performance gains therefore arise not merely from introducing bilateral input, but from structuring bilateral comparison into representation learning, correspondence estimation, and selective evidence extraction.

Several limitations should be acknowledged. First, although the study used a public development dataset, an out-of-distribution validation institution, and an independent private external test set, the private external data were collected from a single institution. Prospective multicenter validation is therefore needed to assess generalizability across centers, imaging devices, and acquisition protocols. Second, the registration-free bilateral design should be interpreted as a feature-space approximation of bilateral comparison rather than as explicit anatomical correspondence. Its effectiveness may be reduced when the contralateral breast provides an unreliable reference, which may be compromised by bilateral disease, marked asymmetric background parenchymal enhancement, implants, postoperative changes, motion artifacts, failed fat suppression, or asymmetric image quality. The observed failure cases were consistent with this boundary (Fig. 7). Third,the model was trained at the breast-side level without requiring lesion-level annotations. Although this weakly supervised setting is practical for large-scale clinical data, it limits direct evaluation of lesion localization. It also does not fully exploit complementary MRI information, including T2-weighted imaging, DWI, ADC maps, ultrafast kinetic features, and clinical risk factors.

Future work should extend the present retrospective evaluation through prospective multicenter studies, broader protocol heterogeneity, and reader studies in realistic clinical workflows. Further studies should also evaluate whether incorporating multiparametric MRI sequences, kinetic information, and clinical risk factors can improve diagnostic performance beyond early DCE-MRI alone. In addition, further comparison between asymmetry maps, model attention, and radiologist-annotated lesion locations may help clarify the spatial interpretability. Overall, the present findings support registration-free bilateral comparison as a potential strategy for improving AI-assisted breast DCE-MRI triage. Larger prospective multicenter studies are required to determine its clinical utility.

6 Conclusion

PRISM-Net presents a registration-free bilateral learning framework for breast-side three-class DCE-MRI classification that treats the contralateral breast as an approximate patient-specific reference rather than a normal template. Through adaptive patch-level matching and focal asymmetry reweighting, the framework leverages bilateral contextual information to characterize diagnostically relevant asymmetry under complex background variation. Across public, external, and background-complexity evaluations, PRISM-Net achieved superior performance compared with existing baselines, while ablation studies demonstrated the complementary contributions of bilateral reference modeling and focal asymmetry reweighting. These findings suggest that patient-specific bilateral comparison may provide a promising strategy for

improving AI-assisted interpretation of challenging breast MRI examinations. Further prospective multicenter validation and reader studies are warranted to assess its potential clinical utility.

CRediT authorship contribution statement

Boya Zhang: Conceptualization, Methodology, Investigation, Data curation, Validation, Formal analysis, Funding acquisition, Writing – original draft. Shuaiwen Zhou: Conceptualization, Methodology, Software, Investigation, Formal analysis, Validation, Visualization, Data curation, Writing – original draft. Di Kong: Review & editing, Methodology, Supervision, Conceptualization. Mingxu Wang: Data curation, Review & editing. Wenbiao Du: Data curation, Review & editing. Yiman Zhong: Data curation, Review & editing. Yuexin Duan: Data curation, Review & editing. Xiawei Yue: Data curation, Review & editing. Liuquan Cheng: Supervision, Resources, Project administration, Conceptualization, Review & editing. Xiru Li: Supervision, Funding acquisition, Resources, Project administration, Review & editing.

Ethics statement

This study involved human subjects and was conducted in accordance with ethical standards. All procedures and protocols were approved by the Ethics Committee of the General Hospital of the People's Liberation Army (Approval No. S2026-078-01).

Funding

This work is supported by the Zhongguancun Academy (Grant No. 02012411).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data and code availability

The public dataset used for model training and evaluation is accessible through the official ODELIA data.The source code will be made publicly available upon acceptance of this article.

References

Antropova, N., Huynh, B., Li, H., Giger, M.L.: Breast lesion classification based on dynamic contrast-enhanced magnetic resonance image sequences with long short-term memory networks. Journal of Medical Imaging **6**(1), 011002 (2019) https://doi.org/10.1117/1.JMI.6.1.011002

American College of Radiology: ACR BI-RADS Atlas: Breast Imaging Reporting and Data System, 5th edn. American College of Radiology, Reston, VA (2013)

Abdullah, K.A., Marziali, S., Nanaa, M., Escudero Sánchez, L., Payne, N.R., Gilbert, F.J.: Deep learning-based breast cancer diagnosis in breast MRI: Systematic review and meta-analysis. European Radiology **35**(8), 4474–4489 (2025) https://doi.org/10.1007/s00330-025-11406-6

Alterson, R., Plewes, D.B.: Bilateral symmetry analysis of breast MRI. Physics in Medicine and Biology **48**(20), 3431–3443 (2003) https://doi.org/10.1088/0031-9155/48/20/011

Bahl, M.: Does background parenchymal enhancement impact the diagnostic accuracy of breast MRI? revisiting the controversy. Radiology **315**(2), 251328 (2025) https://doi.org/10.1148/radiol.251328

Bechyna, S., Baltzer, P.A.T.: Impact of background parenchymal enhancement on diagnostic performance of breast MRI: A systematic review and meta-analysis. Radiology **315**(2), 241919 (2025) https://doi.org/10.1148/radiol.241919

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., Jemal, A.: Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians **74**(3), 229–263 (2024) https://doi.org/10.3322/caac.21834

Dosovitskiy, A., *et al.*: An image is worth 16x16 words: Transformers for image recognition at scale. In: International Conference on Learning Representations (2021)

Gravina, M., Maddaluno, M., Marrone, S., Sansone, M., Fusco, R., Granata, V., Petrillo, A., Sansone, C.: A physiological-informed generative model for improving breast lesion classification in small DCE-MRI datasets. IEEE Journal of Biomedical and Health Informatics **28**(11), 6764–6777 (2024) https://doi.org/10.1109/JBHI.2024.3443705

Gravina, M., Marrone, S., Sansone, M., Sansone, C.: DAE-CNN: Exploiting and disentangling contrast agent effects for breast lesions classification in DCE-MRI. Pattern Recognition Letters **145**, 67–73 (2021) https://doi.org/10.1016/j.patrec.2021.01.023

Guan, Y., Wang, X., Li, H., Zhang, Z., Chen, X., Siddiqui, O., Nehring, S., Huang, X.: Detecting asymmetric patterns and localizing cancers on mammograms. Patterns **1**(7), 100106 (2020) https://doi.org/10.1016/j.patter.2020.100106

Giess, C.S., Yeh, E.D., Raza, S., Birdwell, R.L.: Background parenchymal enhancement at breast MR imaging: Normal patterns, diagnostic challenges, and potential for false-positive and false-negative interpretation. RadioGraphics **34**(1), 234–247

(2014) https://doi.org/10.1148/rg.341135034

Hao, W., Gong, J., Wang, S., Zhu, H., Zhao, B., Peng, W.: Application of mri radiomics-based machine learning model to improve contralateral bi-rads 4 lesion assessment. Frontiers in oncology **10**, 531476 (2020) https://doi.org/10.3389/fonc.2020.531476

Hennessey, S., Huszti, E., Gunasekura, A., Salleh, A., Martin, L., Minkin, S., Chavez, S., Boyd, N.F.: Bilateral symmetry of breast tissue composition by magnetic resonance in young women and adults. Cancer Causes & Control **25**(4), 491–497 (2014) https://doi.org/10.1007/s10552-014-0351-0

Haque, A.K.A., Li, X., Li, Y., Yuan, X.: BiGAM-Net: A bilateral gated fusion framework for imbalanced breast imaging. Expert Systems with Applications **328**, 132746 (2026) https://doi.org/10.1016/j.eswa.2026.132746

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700–4708 (2017)

Hizukuri, A., Nakayama, R., Nara, M., Suzuki, M., Namba, K.: Computer-aided diagnosis scheme for distinguishing between benign and malignant masses on breast DCE-MRI images using deep convolutional neural network with bayesian optimization. Journal of Digital Imaging **34**(1), 116–123 (2021) https://doi.org/10.1007/s10278-020-00394-2

Hirsch, L., Sutton, E.J., Huang, Y., Kayis, B., Hughes, M., Martinez, D., Makse, H.A., Parra, L.C.: High-performance open-source AI for breast cancer detection and localization in MRI. Radiology: Artificial Intelligence **7**(5), 240550 (2025) https://doi.org/10.1148/ryai.240550

Hu, Q., Whitney, H.M., Li, H., Ji, Y., Liu, P., Giger, M.L.: Improved classification of benign and malignant breast lesions using deep feature maximum intensity projection MRI in breast cancer diagnosis using dynamic contrast-enhanced MRI. Radiology: Artificial Intelligence **3**(3), 200159 (2021) https://doi.org/10.1148/ryai.2021200159

He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

Jansen, S.A., Lin, V.C., Giger, M.L., Li, H., Karczmar, G.S., Newstead, G.M.: Normal parenchymal enhancement patterns in women undergoing mr screening of the breast. European Radiology **21**(7), 1374–1382 (2011) https://doi.org/10.1007/s00330-011-2080-z

Kuhl, C.K.: Abbreviated breast MRI: State of the art. Radiology **310**(3), 221822

(2024) https://doi.org/10.1148/radiol.221822

Lo Gullo, R., Ochoa-Albiztegui, R.E., Chakraborty, J., Thakur, S.B., Robson, M., Jochelson, M.S., Varela, K., Resch, D., Eskreis-Winkler, S., Pinker, K.: Development of an MRI radiomic machine-learning model to predict triple-negative breast cancer based on fibroglandular tissue of the contralateral unaffected breast in breast cancer patients. Cancers **16**(20), 3480 (2024) https://doi.org/10.3390/cancers16203480

Leithner, D., Horvat, J.V., Bernard-Davila, B., Helbich, T.H., Ochoa-Albiztegui, R.E., Martinez, D.F., Zhang, M., Thakur, S.B., Wengert, G.J., Staudenherz, A., Jochelson, M.S., Morris, E.A., Baltzer, P.A.T., Clauser, P., Kapetas, P., Pinker, K.: A multiparametric [18F]FDG PET/MRI diagnostic model including imaging biomarkers of the tumor and contralateral healthy breast tissue aids breast cancer diagnosis. European Journal of Nuclear Medicine and Molecular Imaging **46**(9), 1878–1888 (2019) https://doi.org/10.1007/s00259-019-04331-6

Liao, G.J., Henze Bancroft, L.C., Strigel, R.M., Chitalia, R.D., Kontos, D., Moy, L., Partridge, S.C., Rahbar, H.: Background parenchymal enhancement on breast MRI: A comprehensive review. Journal of Magnetic Resonance Imaging **51**(1), 43–61 (2020) https://doi.org/10.1002/jmri.26762

Liu, Y., Zhou, Z., Zhang, S., Luo, L., Zhang, Q., Zhang, F., Li, X., Wang, Y., Yu, Y.: From unilateral to bilateral learning: Detecting mammogram masses with contrasted bilateral network. In: Medical Image Computing and Computer Assisted Intervention - MICCAI 2019. Lecture Notes in Computer Science, vol. 11769, pp. 477–485. Springer, ??? (2019). https://doi.org/10.1007/978-3-030-32226-7_53

Marcon, M., Fuchsjäger, M.H., Clauser, P., Mann, R.M.: ESR essentials: Screening for breast cancer - general recommendations by EUSOBI. European Radiology **34**(10), 6348–6357 (2024) https://doi.org/10.1007/s00330-024-10740-5

Müller-Franzes, G., Escudero Sánchez, L., Payne, N., Athanasiou, A., Kalogeropoulos, M., Lopez, A., Soro Busto, A.M., Camps Herrero, J., Rasoolzadeh, N., Zhang, T., Mann, R., Jutz, D., Bode, M., Kuhl, C., Veldhuis, W., Saldanha, O.L., Zhu, J., Kather, J.N., Truhn, D., Gilbert, F.J.: A European Multi-Center Breast Cancer MRI Dataset. arXiv preprint (2025). https://doi.org/10.48550/arXiv.2506.00474 . https://arxiv.org/abs/2506.00474

Müller-Franzes, G., Khader, F., Siepmann, R., Han, T., Kather, J.N., Nebelung, S., Truhn, D.: Medical slice transformer for improved diagnosis and explainability on 3d medical images with DINOv2. Scientific Reports **15**(1), 23979 (2025) https://doi.org/10.1038/s41598-025-09041-8

ODELIA-AI: ODELIA Challenge 2025. Hugging Face Datasets. Accessed April 17, 2026 (2025). https://huggingface.co/datasets/ODELIA-AI/ODELIA-Challenge-2025

Oviedo, F., Kazerouni, A.S., Liznerski, P., Xu, Y., Hirano, M., Vandermeulen, R.A., Kloft, M., Blum, E., Alessio, A.M., Li, C.I., Weeks, W.B., Dodhia, R., Lavista Ferres, J.M., Rahbar, H., Partridge, S.C.: Cancer detection in breast MRI screening via explainable AI anomaly detection. Radiology **316**(1), 241629 (2025) https://doi.org/10.1148/radiol.241629

Pelluet, G., Rizkallah, M., Tardy, M., Mateus, D.: Enhancing breast cancer screening: Unveiling explainable cross-view contributions in dual-view mammography with sparse bipartite graphs attention networks. Computerized Medical Imaging and Graphics **125**, 102620 (2025) https://doi.org/10.1016/j.compmedimag.2025.102620

Ray, K.M., Kerlikowske, K., Lobach, I.V., Hofmann, M.B., Greenwood, H.I., Arasu, V.A., Hylton, N.M., Joe, B.N.: Effect of background parenchymal enhancement on breast MR imaging interpretive performance in community-based practices. Radiology **286**(3), 822–829 (2018) https://doi.org/10.1148/radiol.2017170811

Siméoni, O., et al.: DINOv3 (2025). https://arxiv.org/abs/2508.10104

Siegel, R.L., Giaquinto, A.N., Jemal, A.: Cancer statistics, 2024. CA: A Cancer Journal for Clinicians **74**(1), 12–49 (2024) https://doi.org/10.3322/caac.21820

Shimokawa, D., Takahashi, K., Kurosawa, D., Takaya, E., *et al.*: Deep learning model for breast cancer diagnosis based on bilateral asymmetrical detection (BilAD) in digital breast tomosynthesis images. Radiological Physics and Technology **16**(1), 20–27 (2023) https://doi.org/10.1007/s12194-022-00686-y

Sun, L., Zhang, Y., Liu, T., Ge, H., Tian, J., Qi, X., Sun, J., Zhao, Y.: A collaborative multi-task learning method for BI-RADS category 4 breast lesion segmentation and classification of MRI images. Computer Methods and Programs in Biomedicine **240**, 107705 (2023) https://doi.org/10.1016/j.cmpb.2023.107705

Truhn, D., Schrading, S., Haarburger, C., Schneider, H., Merhof, D., Kuhl, C.: Radiomic versus convolutional neural networks analysis for classification of contrast-enhancing lesions at multiparametric breast MRI. Radiology **290**(2), 290–297 (2019) https://doi.org/10.1148/radiol.2018181352

Verburg, E., Gils, C.H., Velden, B.H.M., Bakker, M.F., Pijnappel, R.M., Veldhuis, W.B., Gilhuijs, K.G.A.: Deep learning for automated triaging of 4581 breast MRI examinations from the DENSE trial. Radiology **302**(1), 29–36 (2022) https://doi.org/10.1148/radiol.2021203960

Wang, X., Tan, T., Gao, Y., Han, L., Zhang, T., Lu, C., Beets-Tan, R., Su, R., Mann, R.: DisAsymNet: Disentanglement of asymmetrical abnormality on bilateral mammograms using self-adversarial learning. In: Medical Image Computing and Computer Assisted Intervention - MICCAI 2023, pp. 57–67. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-43990-2_6

Yang, Q., Li, L., Zhang, J., Shao, G., Zhang, C., Zheng, B.: Computer-aided diagnosis of breast DCE-MRI images using bilateral asymmetry of contrast enhancement between two breasts. Journal of Digital Imaging **27**(1), 152–160 (2014) https://doi.org/10.1007/s10278-013-9617-4

Zhang, Y., Liu, Y.-L., Nie, K., Zhou, J., Chen, Z., Chen, J.-H., Wang, X., Kim, B., Parajuli, R., Mehta, R.S., Wang, M., Su, M.-Y.: Deep learning-based automatic diagnosis of breast cancer on MRI using mask R-CNN for detection followed by ResNet50 for classification. Academic Radiology **30**(Suppl 2), 161–171 (2023) https://doi.org/10.1016/j.acra.2022.12.038

Zheng, Y., Wei, B., Liu, H., Xiao, R., Gee, J.C.: Measuring sparse temporal-variation for accurate registration of dynamic contrast-enhanced breast MR images. Computerized Medical Imaging and Graphics **46**, 73–80 (2015) https://doi.org/10.1016/j.compmedimag.2015.05.004

Zeng, T., Zeng, Y., Zhang, Z., Zhang, X., Ichiji, K., Chou, S.-Y., Bukovsky, I., Vrba, J., Homma, N.: Bilateral information-guided diagnosis of breast masses in mammography using vision transformer. IEEE Journal of Biomedical and Health Informatics **30**(5), 4226–4237 (2026) https://doi.org/10.1109/JBHI.2025.3626200

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary.pdf
- highlights.pdf