



TopoTxR: A topology-guided deep convolutional network for breast parenchyma learning on DCE-MRIs

Fan Wang^{a,1}, Zhilin Zou^{a,1}, Nicole Sakla^b, Luke Partyka^b, Nil Rawal^b, Gagandeep Singh^c, Wei Zhao^e, Haibin Ling^a, Chuan Huang^{d,g}, Prateek Prasanna^{f,*}, Chao Chen^{f,*}

^a Department of Computer Science, State University of New York at Stony Brook, NY, USA

^b Department of Radiology, Newark Beth Israel Medical Center, NJ, USA

^c Department of Radiology, Columbia University Irving Medical Center, NY, USA

^d Department of Radiology and Imaging Sciences, Emory University School of Medicine, GA, USA

^e Department of Radiology, State University of New York at Stony Brook, NY, USA

^f Department of Biomedical Informatics, State University of New York at Stony Brook, NY, USA

^g Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA, USA

ARTICLE INFO

Keywords:

Topology

DCE-MRI

Persistent homology

Spatial attention

3D CNN

pCR prediction

ABSTRACT

Characterization of breast parenchyma in dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a challenging task owing to the complexity of underlying tissue structures. Existing quantitative approaches, like radiomics and deep learning models, lack explicit quantification of intricate and subtle parenchymal structures, including fibroglandular tissue. To address this, we propose a novel topological approach that explicitly extracts multi-scale topological structures to better approximate breast parenchymal structures, and then incorporates these structures into a deep-learning-based prediction model via an attention mechanism. Our topology-informed deep learning model, *TopoTxR*, leverages topology to provide enhanced insights into tissues critical for disease pathophysiology and treatment response. We empirically validate *TopoTxR* using the VICTRE phantom breast dataset, showing that the topological structures extracted by our model effectively approximate the breast parenchymal structures. We further demonstrate *TopoTxR*'s efficacy in predicting response to neoadjuvant chemotherapy. Our qualitative and quantitative analyses suggest differential topological behavior of breast tissue in treatment-naïve imaging, in patients who respond favorably to therapy as achieving pathological complete response (pCR) versus those who do not. In a comparative analysis with several baselines on the publicly available I-SPY 1 dataset (N = 161, including 47 patients with pCR and 114 without) and the Rutgers proprietary dataset (N = 120, with 69 patients achieving pCR and 51 not), *TopoTxR* demonstrates a notable improvement, achieving a 2.6% increase in accuracy and a 4.6% enhancement in AUC compared to the state-of-the-art method.

1. Introduction

Breast cancer imaging faces a critical challenge in accurately modeling complex breast parenchyma structures, which change dynamically due to factors like angiogenesis, radiation therapy, and chemotherapy. Utilizing 3D breast imaging such as MRI to capture and model these changes can significantly impact diagnosis, prognosis, and treatment planning. Traditional cancer imaging studies have primarily focused on tumor texture and shape, overlooking valuable information in the tumor microenvironment. Evidence suggests that diagnostic and prognostic insights lie in the peritumoral stroma and parenchyma, where phenotypic diversity arises from factors like immune infiltration, vascularity, and tissue composition. Parameters like fibroglandular tissue

pattern and parenchymal enhancement also influence breast cancer risk and treatment responses. To tailor diagnosis and treatment strategies to improve patient care, there is a critical need for innovative quantitative methods to comprehensively understand breast cancer biology by exploring the tumor microenvironment and surrounding parenchyma, which can be routinely observed in imaging scans such as breast MRI.

Various approaches for breast image analysis have been proposed. Radiomic approaches learn diagnostic and prognostic signatures from breast tumor and surrounding peritumoral regions using radiomics features (Saha et al., 2018; Van Griethuysen et al., 2017a). These

* Corresponding authors.

E-mail addresses: fanwang1@cs.stonybrook.edu (F. Wang), Prateek.Prasanna@stonybrook.edu (P. Prasanna), chao.chen.1@stonybrook.edu (C. Chen).

¹ Equal contributions.

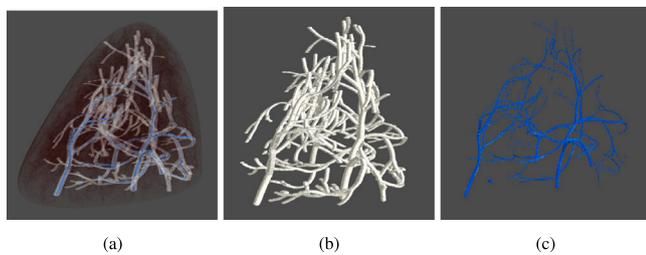


Fig. 1. (a): 3D rendering of a phantom breast with highlighted glandular tissues (white) and topological structures (blue); (b): glandular tissues; (c): topological structures.

handcrafted features, inspired by human knowledge, attempt to capture different measurements such as tumor/peritumor texture (Braman et al., 2017), vessel geometry descriptors (Braman et al., 2022), and other similar characteristics. However, these approaches have two fundamental limitations. First, they usually lack an explicit modeling of the complex structural pattern of peritumoral stroma and parenchyma. Second, these handcrafted features lack sufficient flexibility to model the heterogeneous breast parenchyma and thus cannot provide the desired level of predictive power in practice, despite abundant interpretability.

On the other hand, data-driven approaches, such as deep neural networks (Wu et al., 2023) and convolutional neural networks (CNNs) (Burt et al., 2018; Jarkman et al., 2022; Tack et al., 2018; Subasi et al., 2023; Su et al., 2023; Zhang et al., 2018; Zhu et al., 2017), have shown great promise in various domains, as they learn feature representations in an end-to-end manner. While direct application of CNNs to MRIs seems promising (Liu et al., 2020a; ag Lundervold and Lundervold, 2019; Lyu et al., 2022; Li et al., 2021; Mazurowski et al., 2018; Nawaz et al., 2020; Dalmış et al., 2017; Ranem et al., 2022), CNN methods take the whole MRI as a direct input; a large portion of the input volume may be biologically irrelevant and even noisy enough to bias the prediction model. Additionally, a 3D CNN possesses millions of parameters and requires a substantial amount of training data. Unfortunately, obtaining such amount of data is often impractical for controlled clinical trials like the I-SPY 1 trial (Newitt and Hylton, 2016). Furthermore, CNNs suffer from the limitation of feature interpretability as they lack direct association with the underlying breast tissue structures.

In this paper, we propose *TopoTxR*, a novel method that overcomes the aforementioned disadvantages. Our method extracts the breast parenchyma structures using the mathematical language of topology. It then effectively incorporates these rich topological structures into deep convolutional neural networks, thereby significantly improving predictive power. By explicitly modeling parenchyma, our approach ensures that predictions are firmly based on biological structures, and thus significantly enhances the performance of the data-intensive CNN model, even with a limited training set.

Our method is based on the theory of persistent homology (Edelsbrunner and Harer, 2010), which extracts 1D (loops) and 2D (bubbles) topological structures with guaranteed robustness (Cohen-Steiner et al., 2005). These structures correspond to curvilinear tissue structures (e.g., ducts, vessels, etc.) and voids enclosed by tissues and glands in their proximity. As shown in Fig. 2, compared to previous radiomics features, the topological structures provide a much richer structural context for the modeling of tumor microenvironment. Since these structures are extracted in an unsupervised manner, the quality of their interpretation becomes a key consideration. Using a phantom breast imaging dataset (VICTRE (Badano et al., 2018)), we validate both quantitatively and qualitatively that these topological structures are reasonable approximations of the breast tissue structures. As illustrated in Fig. 1, the extracted topological structures delineate the glandular tissues of a phantom breast image.

To fully exploit the information carried by these topological structures, we propose a topology-guided deep learning model for breast

images. The key idea is to direct the model’s attention to voxels adjacent to these topological structures/tissue structures. As the model is focused on a much smaller set of voxels with high biological relevance, it can be efficiently trained with limited MRI data. Meanwhile, the learning outcomes have the potential to connect the biological causes of various breast pathologies with the manifestations observed in the topology of tissue structures. A method closely related to ours, developed by Du et al. (2022), employs a form of weaker topological information known as Betti curves to enhance prediction accuracy. However, the information carried by Betti curves is much more limited when compared to the explicit topological structures used in our approach, leading to suboptimal performance.

While our approach is task-agnostic, we specifically focus on predicting the response to neoadjuvant chemotherapy (NAC) in breast cancer treatment as a practical application. Correct prediction of pathological complete response (pCR) prior to NAC administration can prevent ineffective treatments, reducing unnecessary patient suffering and healthcare costs. Empirically, we have evaluated our method on the I-SPY 1 dataset (Newitt and Hylton, 2016) and a proprietary dataset. In these evaluations, *TopoTxR* outperforms various baselines, including radiomics approaches, CNNs trained without topological priors, and other state-of-the-art approaches. This highlights the effectiveness of our topology-centric approach in achieving superior predictive performance.

In summary, we present a novel topological method to characterize parenchyma in breast DCE-MRI, aiming to predict pCR. Our method bridges the two extremes of engineered imaging features and completely data-driven CNNs. Our key contributions are:

- Utilization of persistent homology theory to extract topological structures that closely approximate breast fibroglandular tissue.
- Comprehensive evaluation using a phantom breast imaging dataset to validate the accuracy of the extracted topological structures in approximating breast tissues.
- Introduction of a topology-guided spatial attention mechanism designed to direct the focus of 3D CNNs, thereby enhancing their predictive capabilities.

This work builds upon our previous conference paper (Wang et al., 2021) with four key enhancements: (1) A new topology-guided spatial attention (TGSA) module that explicitly directs the attention of the 3D CNN to biologically relevant sets of voxels with a mask loss. This new module effectively eliminates the need for prior persistent homology computing during the inference stage, making it more practical for real-world applications. (2) An exhaustive empirical evaluation on the VICTRE phantom dataset to validate the accuracy of the topological approximation of breast tissues. (3) Model attention visualization that confirms the proposed model’s attention is concentrated on a smaller biologically relevant set of voxels when predicting treatment responses using Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al., 2017). (4) A new proprietary dataset along with comprehensive evaluations to demonstrate the proposed model’s generalization and versatility across different clinical settings. We have made the source code publicly available via this [GitHub repository](#) to facilitate the reproducibility of our research.

1.1. Related work

Quantitative imaging features have been used in conjunction with machine learning classifiers for the prediction of pCR (Cain et al., 2019; Mani et al., 2013). Radiomics approaches, involving analysis of quantitative attributes of tumor texture and shape, have shown promise in the assessment of treatment response. In particular, such features capture the appearance of the tumors and, more recently, peritumor regions (Braman et al., 2019; Grimm, 2015). Such approaches are often limited by their predefined nature, lack of generalizability, reliance on accurate lesion segmentation, and inability to explain phenotypic

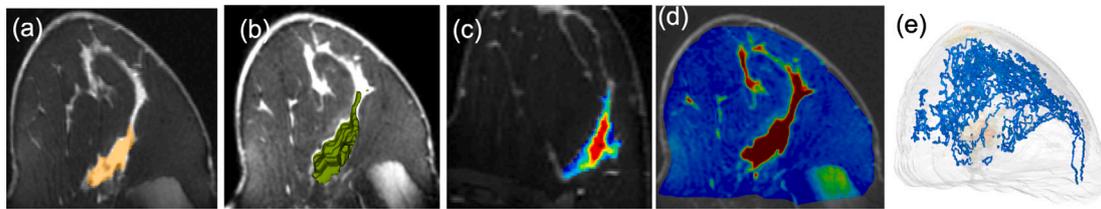


Fig. 2. (a) A example MRI image, and different radiomics features such as (b) 3D shape of a tumor, (c) intratumoral texture (Haralick entropy), and (d) whole breast texture (Haralick energy). In (e), we show topological structures from *TopoTxR*, capturing the geometry of fibroglandular tissues.

differences beyond the peritumoral margin. CNNs have been previously applied to breast DCE-MRI for pCR prediction (Liu et al., 2020b; Ravichandran et al., 2018; Ha et al., 2018; Qu et al., 2020). Owing to the sub-optimal performance of image-only models, image-based CNN approaches have been fused with non-imaging clinical variables to bolster prediction (Duanmu et al., 2020).

Moreover, there is extensive literature highlighting the use of CNNs for cancer diagnosis in mammography (Kooi et al., 2017; Abdelhafiz et al., 2019; Subasi et al., 2023). However, mammograms are only 2D projections of 3D tissue structures. This results in a loss of interpretability of the extracted structures. 2D mammography is limited by the loss of interpretability of extracted structures and 3D mammography using tomography is limited by the inability of the mammogram to provide information regarding background parenchymal enhancement or tumoral/peritumoral enhancement kinetics. Moreover, MRI has been shown to be superior to mammography in determining the extent of breast cancer. In contrast, mammography and ultrasound fail to accurately evaluate tumor size in 8% of cases post-neoadjuvant chemotherapy, as reported in a retrospective review by Akazawa et al. (2006), Keune et al. (2010) and Londero et al. (2004). The purpose of post-neoadjuvant chemotherapy examination is to evaluate the extent of residual disease. However, mammography, ultrasound, and physical exams accurately detect only 13%–25% of pCR cases, as indicated in a study by Herrada et al. (1997), Vinnicombe et al. (1996) and von Minckwitz et al. (1999). MRI and ultrasound each remains superior to mammography with respect to residual tumor detection. Notably, MRI is significantly more effective than mammography in identifying cases of multifocal or multicentric disease, as reported by Londero et al. (2004).

Topological information, in particular, persistent homology (Edelsbrunner and Harer, 2010), provides a robust way to quantify the topological information in an image. This information, encoded as persistence diagrams or persistence barcodes, has found diverse applications in various image analysis tasks, such as cardiac image analysis (Wu et al., 2017), brain network analysis (Lee et al., 2012; Yao et al., 2024), and neuron image segmentation (Hu et al., 2019). In recent years, it has been combined with deep neural networks to enforce topological constraints in image segmentation tasks (Clough et al., 2020; Hu et al., 2019; Shit et al., 2020; Stucki et al., 2023). Abundant work has been done to learn topology from persistence diagrams, for instance, through vectorization (Adams et al., 2017), kernel machines (Carriere et al., 2017; Kusano et al., 2016; Reininghaus et al., 2015), and deep neural networks (Hofer et al., 2017). Additionally, topology has been formulated as graphs, which are then effectively integrated with Graph Neural Networks (GNN) for applications such as pathology image classification (Wang et al., 2023) and retinal artery/vein classification (Mishra et al., 2021). An even weaker topological information called Betti curve is extracted for learning with breast images (Du et al., 2022). However, the actual geometric realization of the topological structures, e.g., cycles and bubbles, has not yet been fully explored. These topological structures capture the geometric details of breast tissues, such as fibroglandular tissues, and can be mapped back to the original breast volumes to provide biologically relevant information for further CNN analysis. For the first time, we propose a deep learning method that leverages the geometry of topological structures as an explicit attention mechanism in this paper.

2. Methodology

We propose a topological method to extract topological structures of high saliency, approximating tissue structures, and utilize these extracted structures as auxiliary information to train a deep convolutional network with raw MRI inputs. Although we focus on training our model for the pCR prediction task, the methodology is versatile enough to be generalized for other tasks. Our approach is detailed in Fig. 3.

We first compute salient topological structures from the input image utilizing persistent homology theory. Topological structures of dimensions 1 and 2, i.e., loops and bubbles, can both correspond to important tissue structures. 1D topological structures capture curvilinear structures such as ducts, vessels, etc. 2D topological structures represent voids enclosed by the tissue structures and their attached glands. These topological structures directly delineate the critical tissue structures with high biological relevance. Thus we hypothesize that by focusing on these tissue structures and their affinities, we can gain pertinent contextual information for pCR prediction.

Subsequently, we introduce a novel 3D CNN framework tailored for breast MRIs that integrates topological structures via an attention mechanism. Our method constructs a custom loss function, combining a mask-guided loss and a refined classification loss, the latter based on focal loss as detailed in Lin et al. (2017). Notably, we identify two types of pertinent topological structures: loops and bubbles. Our network consists of two separate 3D CNNs, treating the two types of topological structures separately. Empirical evidence demonstrates that both topology types capture complementary structural signatures, proving essential for achieving optimal predictive performance.

Next, we present details of our method, including the background knowledge of persistent homology (Section 2.1), how to compute cycles representing the topological structures (Section 2.2), and our 3D CNN with topology-guided attention (Section 2.3).

2.1. Background: Persistent homology

We review the basics of persistent homology in an intuitive way. Interested readers may refer to Edelsbrunner and Harer (2010) for more details. Persistent homology extracts the topological information of data observed via a scalar function. Given an image domain, X , and a real-valued function $f : X \rightarrow \mathbb{R}$, we can construct a sublevel set $X_t = \{x \in X : f(x) \leq t\}$ where t is a threshold controlling the “progress” of sublevel sets. The family of sublevel sets $\mathcal{X} = \{X_t\}_{t \in \mathbb{R}}$ defines a filtration, i.e., a family of subsets of X nested with respect to the inclusion: $X_\alpha \subseteq X_\beta$ if $\alpha \leq \beta$. As the threshold t increases from $-\infty$ to $+\infty$, topological structures such as connected components, handles, and voids appear and disappear. The *birth time* of a topological structure is the threshold t at which the structure appears in the filtration. Similarly, the *death time* is the threshold t at which the structure disappears. Persistent homology tracks the topological changes of sublevel sets X_t and encodes them in a *persistence diagram*, i.e., a point set in which each point (b, d) represents a topological structure with birth time b and death time d . Its lifespan, i.e., $d - b$, is called the *persistence* of the structure. In practice, we believe persistence is an indicator of the saliency of a topological structure; it has been proven that the persistence of a structure bounds the amount of perturbation one has to inject into the input in order to “shed off” the

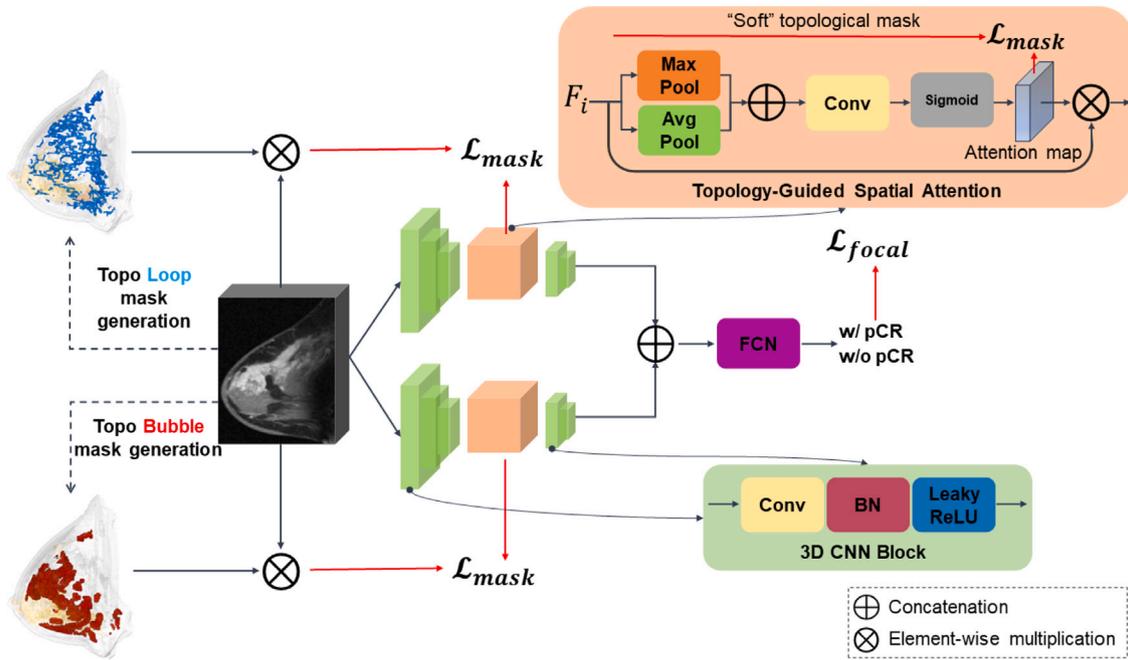


Fig. 3. Our proposed TopoTxr pipeline. We extract 1D and 2D topological structures from breast MRI based on persistent homology. Rather than using binary masks, we extract topological structures with intensity values from raw MRIs (“soft” topological masks) for mask loss \mathcal{L}_{mask} supervision. Each 3D CNN branch includes five 3D CNN blocks and a topology-guided spatial attention module (TGSA). The input to TGSA is the feature map from the third convolution layer, F_t , while its output to the fourth convolution layer is the generated attention map multiplied by F_t . The model features two distinct 3D CNN branches with a fully connected network for pCR prediction.

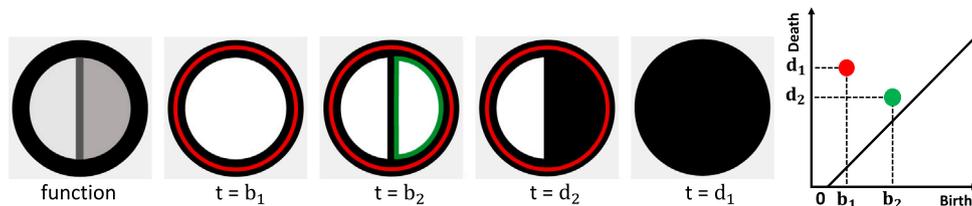


Fig. 4. From left to right: a synthetic image f , sublevel sets at thresholds $b_1 < b_2 < d_2 < d_1$, and the 1D persistence diagram. The red loop represents a 1D structure born at b_1 and killed at d_1 . The green loop represents a 1D structure born at b_2 and killed at d_2 . They correspond to the red and green dots respectively in the diagram.

structure (Cohen-Steiner et al., 2005). One may compute the number of linearly independent topological structures at different thresholds, t . These counts parametrized by t constitute a simpler topological signature called *Betti curves*. We acknowledge that while Betti curves can be directly computed from the persistence diagram, the reverse is not possible.

See Fig. 4 for an example function f and its sublevel sets at different thresholds. At time b_1 , a new handle (delineated by the red cycle c_1) is created. This handle is later destroyed at time d_1 . Another handle delineated by the green cycle c_2 is created and killed at b_2 and d_2 respectively. The topological changes are summarized in a persistence diagram on the right. Each handle corresponds to a 2D dot in \mathbb{R}^2 , whose x and y coordinates are birth and death times. Their persistence values are $d_1 - b_1$ and $d_2 - b_2$ respectively.

2.2. Persistence cycles and their computation

Although the persistence diagram has been used for topological analysis in various dataset (Hu et al., 2019; Wang et al., 2020; Wu et al., 2017), it only records limited information, i.e., the times at which these topological structures appear/disappear. We hypothesize that a detailed geometric realization of these topological structures can be crucial for learning from images. To this end, we propose extracting these topological structures and integrating them into the learning

process. As shown in Fig. 3, we extract loops (blue) to denote 1D topological structures and bubbles (red) for 2D topological structures. These structures are then used to guide the attention mechanism within the neural network.

Next, we formally introduce how different topological structures can be represented with cycles. We also explain how these representative cycles are computed. Intuitively, a topological cycle of dimension p is a p -manifold without boundary. A 1-dimensional (1D) cycle is a loop. A 2-dimensional (2D) cycle is a bubble. A cycle z represents a persistent homology structure if it delineates the structure at its birth. For example, in Fig. 4, the red and the green loops denote the handles born at time b_1 and b_2 , respectively.

We assume a discretization of the image domain into distinct elements, i.e., vertices (corresponding to voxels), edges connecting adjacent vertices, squares, and cubes. These elements are referred to as 0-, 1-, 2-, and 3-dimensional cells, or simply 0-, 1-, 2-, and 3-cells. A set of p -cells constitutes a p -chain. The *boundary* of a p -cell, σ , is the set of its $(p - 1)$ -faces. For example, an edge’s boundary comprises its two adjacent vertices, a square’s boundary includes the four enclosing edges, and a cube’s boundary is made up of the six squares surrounding it. The boundary of any p -chain, c , is the formal sum of the boundaries of all its elements, $\partial(c) = \sum_{\sigma \in c} \partial(\sigma)$, where the sum is the mod-2 sum. In other words, we are taking the exclusive or of the boundaries of all

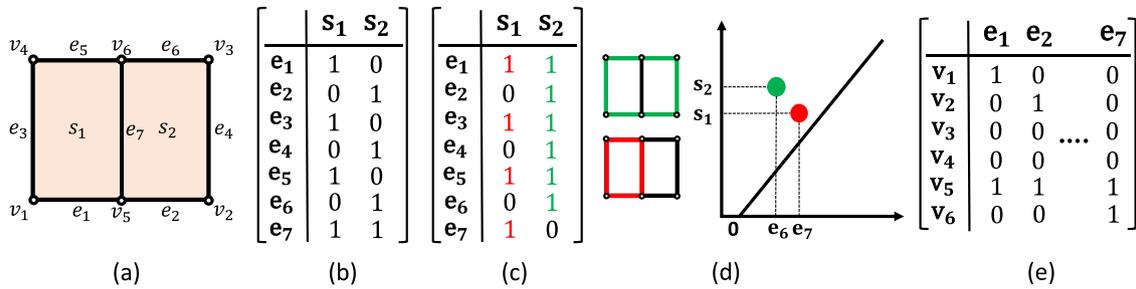


Fig. 5. (a) Example of a cubical complex whose cells are sorted monotonically non-decreasing according to the function values. (b) 2D boundary matrix ∂ . (c) Reduced boundary matrix. (d) Persistence diagram and resulting topological cycles of ∂ . (e) 1D boundary matrix.

the elements.² For a set of edges forming a path, the boundary consists of the two end vertices. Similarly, for a set of squares forming a patch, its boundary is the loop enclosing the patch, and for a set of cubes, the boundary is the bubble enclosing them.

A p -chain is a p -cycle if its boundary is empty. All p -cycles form the null space of the boundary operator, defined as $\{c : \partial(c) = \emptyset\}$. Every topological structure, formally defined as a *homology class*, can be represented by various cycles, all equivalent under the boundary operator. Formally, given any cycle z_0 representing a homology class h , we can express h as $\{z_0 + \partial(c)\}$ for all possible chains, c , of a higher dimension. We can choose any of the cycles to represent this class. In persistent homology, we can represent each dot in the persistence diagram with one representative cycle at its birth. As depicted in Fig. 4, the red and green cycles represent the two corresponding handles. Note that the choice of representative cycle is not unique. A relevant question is to choose the shortest representative cycle (i.e., one with the least number of edges) for each dot in the diagram (Wu et al., 2017; Zhang et al., 2019). In this paper, we focus on choosing a standard representative cycle for computational efficiency.

Computation of persistent homology and representative cycles. We assume a filtration function on a discretization of the image domain. An example discretization of a 2D image is given in Fig. 5(a). We first sort all cells in increasing order according to their function values. The vertices take the intensity values of their corresponding voxels. For any higher dimensional cell, its function value is the maximum of the values of its vertices. The computation of persistence diagrams is then performed by encoding the p -dimensional boundary operator in a binary matrix named *boundary matrices*, ∂_p . ∂_p maps p -cells to their boundaries. Fig. 5 shows the 1D and 2D boundary matrices of the given complex and its filtration. The 1D boundary matrix is essentially the incidence matrix of the underlying graph (Fig. 5(e)). High dimensional boundary matrices are defined similarly (e.g., a 2D boundary matrix in Fig. 5(b)). It is essential that the rows and columns of the boundary matrices are arranged so that as we increase the row/column indices, the corresponding cell function values are monotonically non-decreasing.

The persistence diagram is computed by reducing the boundary matrix in a manner akin to Gaussian elimination, but without row or column perturbation. This reduction involves column operations on ∂ executed from left to right and from bottom to top. Fig. 5(c) shows the reduced 2D boundary matrix. Upon reducing the boundary matrices, each non-zero column corresponds to a persistent dot in the persistence diagram, as shown in Fig. 5(d). The reduced column itself is the cycle representing the corresponding topological structure. In this paper, we pay attention to both 1D and 2D cycles, corresponding to loops and bubbles. The extracted cycles will be used to explicitly guide 3D CNNs for analysis. The computation of representative cycles is of the same complexity as the computation of persistent homology. In theory, it takes $O(n^\omega)$ time ($\omega \approx 2.37$ is the exponent in the matrix multiplication

² The choice of mod-2 sum is because we focus on homology over \mathbb{Z}_2 field, following the common practice in topological data analysis.

time, i.e., time to multiply two $n \times n$ matrices) (Milosavljević et al., 2011). Here n is the number of voxels in an image. In practice, computing all cycles for an input image of size (256^3) takes approximately 5 min.

2.3. A 3D CNN with topology-guided spatial attention

To incorporate the generated topological structures into a deep learning classifier, we propose a novel topology-guided spatial attention module. Using a new loss that enforces the spatial attention map to be similar to the topological structures, we direct the model’s attention to the vicinity of breast tissues, thus achieving better prediction power. We first explain the generation of the topological mask, i.e., the union of all topological structures. Next, we explain how to use the topological masks to guide the attention of a deep neural network.

Generating topological masks. We initiate by extracting topological structures which are represented as sets of voxels. The union of these voxel sets is then used to form a mask that closely approximates biologically pertinent regions.

To ascertain persistence and topological cycles that resemble breast tissue, we invert the MRI image, denoted as $f = -I$, causing the tissue structures to correspond to regions with lower intensity values. Subsequent to this computation, we pinpoint topological cycles represented by dots on the persistence diagram with high persistence. It is theoretically well understood that dots with lower persistence are less likely to represent genuine signals (Cohen-Steiner et al., 2005). As such, we exclusively focus on high-persistence dots, which are indicative of more prominent structures and have a higher likelihood of representing true tissue structures. The persistence threshold is a hyper-parameter adjusted empirically.

Upon selecting topological structures based on persistence, we use the voxels from their representative cycles to craft topological masks. We generate two distinct binary 3D masks that represent the 1D and 2D topological cycles. Empirical observations have revealed that a “soft” mask conveys more detailed information. Rather than directly employing binary masks, the foreground voxels are populated with their inherent image intensity values. Given that all masked MRIs are padded to a size of 256^3 , we subsequently produce two topological masks, M_{topo}^1 and M_{topo}^2 , which pertain to the 1D and 2D topological cycles, respectively.

Topology-guided spatial attention. In vision-related tasks, spatial attention creates an attention map by harnessing the inter-spatial relationships among features (Woo et al., 2018). In the realm of medical tasks, attention mechanisms have been effectively applied to medical image registration (Song et al., 2022) and drug response prediction (Feng et al., 2021). TopoTxR employs this type of attention to focus on topological structures. Unlike traditional spatial attention, mask-guided attention integrates masks to sharpen model focus which offers a more targeted approach (Pang et al., 2019). With these masks defining explicit regions of interest, the model is adept at selectively concentrating on the most pertinent sections, while minimizing attention to

Table 1

Average percentage of each tissue type over 25 samples from each phantom breast profile.

	Dense	Fatty	Hetero	Scattered
Fat	28.59%	79.82%	49.97%	70.44%
Skin	4.34%	3.09%	3.89%	3.26%
Glandular	42.19%	6.43%	28.45%	12.50%
Nipple	0.11%	0.02%	0.07%	0.03%
Muscle	22.00%	8.71%	15.18%	11.67%
Ligament	1.35%	1.68%	1.51%	1.61%
TDLU ^a	0.31%	0.01%	0.13%	0.03%
Duct	0.22%	0.07%	0.12%	0.06%
Artery	0.38%	0.07%	0.30%	0.17%
Vein	0.49%	0.09%	0.39%	0.21%

^a Terminal Duct Lobular Unit.

regions that might contain noise or are of lesser significance based on mask guidance. This not only streamlines the training process but also bolsters performance. Our work introduces the topology-guided spatial attention module, as illustrated in Fig. 3. This module generates a spatial attention map, which is supervised by the aforementioned “soft” topology masks.

We employ distinct 3D CNN networks for both the 1D and 2D branches, each maintaining the same architecture. Every 3D CNN is structured with five 3D convolution layers, with each layer being succeeded by a batch normalization layer and then a LeakyReLU activation. In the topology-guided attention module, the feature maps from the third convolution layer, referred to as F_i (where i signifies the branch dimension), are initially processed through both average-pooling and max-pooling operations independently. These pooled results are then concatenated, serving as input for a subsequent convolution layer. The convolution layer’s output passes through a sigmoid activation function, resulting in the generation of attention maps, denoted as M_i where $i \in [1, 2]$.

To emphasize the regions highlighted by the attention maps, we conduct an element-wise multiplication of F_i and M_i . The product, $M_i \odot F_i$, is channeled as input into the fourth layer of the 3D CNN. Outputs from both 3D CNN branches are subsequently vectorized and concatenated. This combined output feeds into the fully convolutional network, culminating in a binary classification. In our ablation study, we further investigate the optimal layer from which to source input for the attention module.

The generation of spatial attention masks is guided by “soft” topological masks. This guidance ensures their similarity is captured through a mean squared error (MSE) loss. The loss for our topology-guided spatial attention module, denoted as \mathcal{L}_{mask} , is defined as:

$$\mathcal{L}_{mask} = \left\| M_1 - M_{topo}^1 \right\|_2^2 + \left\| M_2 - M_{topo}^2 \right\|_2^2, \quad (1)$$

where M_{topo}^1 and M_{topo}^2 represent the topological masks for 1D and 2D dimensions, respectively.

The focal loss enforces a well-balanced classification performance (Lin et al., 2017). The focal loss, denoted as \mathcal{L}_{focal} , is defined as:

$$\mathcal{L}_{focal}(p) = -\theta(1-p)^\gamma y_{gt} \log(p) - (1-\theta)p^\gamma (1-y_{gt}) \log(1-p) \quad (2)$$

In this context, y_{gt} is the ground truth of the classification, while p indicates the probability of the model prediction. Focal loss incorporates two hyperparameters: θ and γ . The parameter θ is responsible for balancing the weights of positive and negative sample losses, whereas θ amplifies the loss contribution from hard samples. The mask loss and focal loss together form our overall loss function:

$$\mathcal{L}_{All} = \mathcal{L}_{focal} + \lambda \mathcal{L}_{mask}, \quad (3)$$

where λ is a hyper-parameter.

Table 2

T1/T2 values and standard deviations for each breast tissue type.

	Mean T1 + Std. Dev. (ms)	Mean T2 + Std. Dev. (ms)
Fat	366.78 ± 7.75	52.96 ± 1.54
Skin	887.00 ± 92.00	22.30 ± 7.00
Glandular	1444.83 ± 92.70	54.36 ± 9.35
Nipple	796.00 ± 21.00	63.00 ± 4.00
Muscle	1232.90 ± 255.00	37.20 ± 9.80
Ligament	400.00 ± 10.00	40.00 ± 2.00
TDLU	1444.83 ± 92.70	54.36 ± 9.35
Duct	796.00 ± 21.00	63.00 ± 4.00
Artery	1984.40 ± 146.70	275.00 ± 0.00
Vein	1984.40 ± 146.70	275.00 ± 0.00

3. Topological approximation of breast tissue: Validation on a phantom dataset

The key hypothesis of this paper is that the topological structures, extracted in an unsupervised manner, are good approximations of breast tissue. This section empirically validates this hypothesis.

Ideally, we would like to observe a significant overlap between the true breast tissue and the extracted topological structures. Unfortunately, obtaining ground truth breast tissue from MRIs is very demanding in both time and expertise. Instead, we resort to the VICTRE Breast Phantom data to generate synthetic breast volumes with ground truth breast tissue. This phantom dataset is developed as part of the VICTRE (Virtual Imaging Clinical Trials for Regulatory Evaluation) project by the FDA (Food and Drug Administration) (Badano et al., 2018). The project produces the VICTRE in-silico trials which provide Monte Carlo simulations of digital mammography (DM) and digital breast tomosynthesis (DBT) exams. Although originally designed for DM and DBT, we can synthesize MRIs using the same phantom data. By comparing with the ground truth tissue, we can validate the approximation quality of the extracted 1D and 2D topological structures.

3.1. VICTRE synthetic dataset and the synthesized MRIs

The VICTRE Breast Phantom generates various types of synthesized breast image volumes based on tunable parameters. Here, we adopt the default configurations used in the VICTRE trial to generate four different profiles of breast phantom densities, covering different real world scenarios: dense, heterogeneously dense, scattered density, and fatty. A synthetic 3D breast phantom is a multi-class mapping in which each voxel is assigned with one of 10 tissue types: fat, skin, glandular, nipple, muscle, ligament, terminal duct lobular unit (TDLU), duct, artery, and vein. Table 1 summarizes the composition of tissue types for each profile. Fig. 6 shows sample volumes for each of the four profiles. In experiments, we generate 25 sample volumes for each profile (100 in total) and simulate their corresponding MRIs for validation.

Using the 100 synthetic breast volumes, we perform simulations to generate 100 different synthetic T1-weighted MRIs. The T1/T2 weighted values for each tissue type are listed in Table 2. For each tissue type within a specific breast volume, we sample a single value for both T1 and T2. This ensures that all voxels associated with that particular tissue type have identical values within that breast volume. However, for different breast volumes, we derive a new set of T1 and T2 values for each tissue type based on the known distribution. The simulated MRI signal is then computed using:

$$S = k[H] \frac{\sin \alpha (1 - e^{-TR/T1})}{(1 - (\cos \alpha) e^{-TR/T1})} e^{-TE/T2}$$

where $k[H]$ is the spin density, α stands for the flip angle, and TR (repetition time), and TE (echo time) are MR acquisition parameters. In our experiments, we set $k[H] = 1$, $\alpha = 6^\circ$, $TR = 50$ ms, and $TE = 10$ ms to mimic different T1-weighted images. The simulated MR images are processed by our algorithm to extract topological structures.

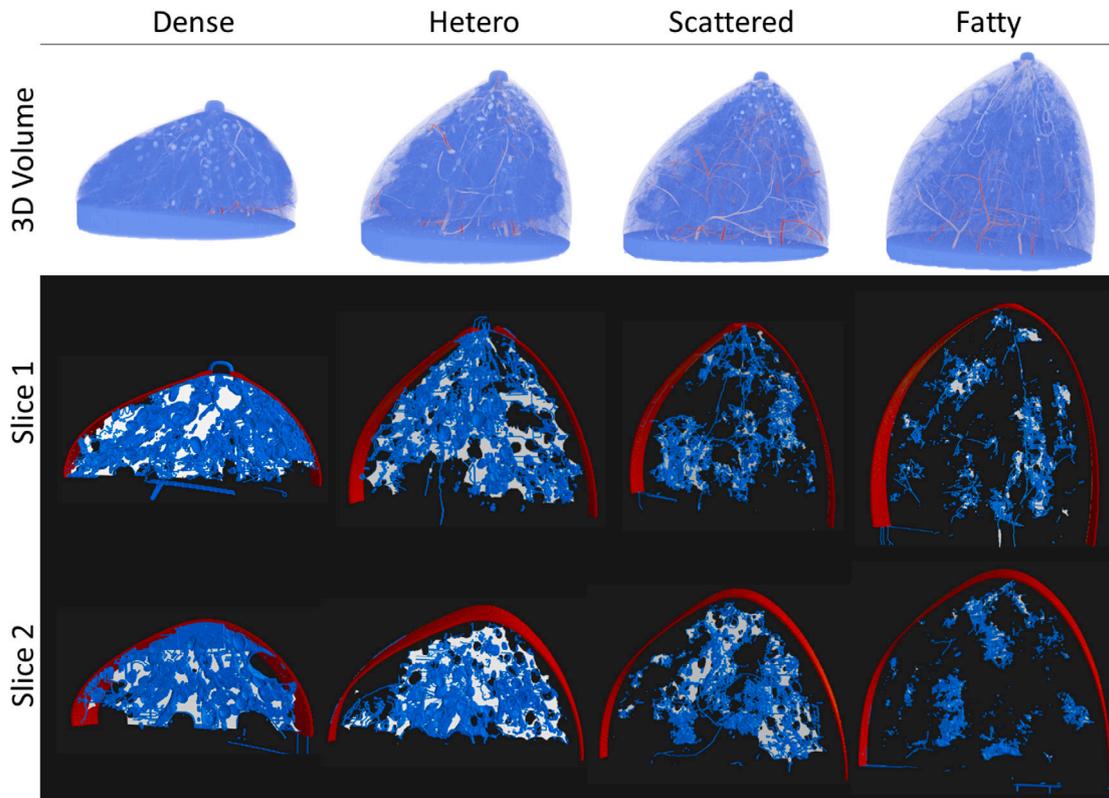


Fig. 6. Row 1: 3D renderings of VICTRE phantom breasts of four distinct profiles. Rows 2 and 3: two slices at different positions of the corresponding breast phantoms. Red: 1-voxel width breast outline; blue: extracted topological structures; white: ground truth breast tissues. Each slice's rendering includes several additional slices around the target cross sections for detailed examination.

Table 3

Evaluation of the approximation quality of topological structures using topological precision and recall.

Topological precision \uparrow				Topological recall \uparrow			
Dense	Fatty	Hetero	Scattered	Dense	Fatty	Hetero	Scattered
0.93 ± 0.01	0.95 ± 0.01	0.97 ± 0.01	0.96 ± 0.02	0.26 ± 0.01	0.20 ± 0.01	0.18 ± 0.03	0.17 ± 0.03

Table 4

Evaluation of the approximation quality of topological structures based on the mean distance from topological structures to breast tissues and from breast tissues to topological structures.

Dist. Tissue to Topo. \downarrow				Dist. Topo. to Tissue \downarrow			
Dense	Fatty	Hetero	Scattered	Dense	Fatty	Hetero	Scattered
1.96 ± 0.08	1.88 ± 0.06	2.15 ± 0.09	2.05 ± 0.07	1.96 ± 0.08	15.67 ± 0.96	5.57 ± 0.33	11.50 ± 1.33

These extracted structures will then be compared with the ground truth tissues, specifically the synthesized breast image volumes. Fig. 6 shows the corresponding MRI images of the four phantom volumes.

3.2. Quantitative and qualitative validation

We show that the extracted topological structures effectively capture the breast tissues. We apply our method introduced in Section 2.2 to the synthetic MRIs to extract topological structures, and then compare these extracted structures with the ground truth tissue in the synthetic volumes. If there is a significant overlap, it indicates that the extracted topological structures closely approximate the actual breast tissue. In our study, breast tissue is defined as the union of all tissue types characterized by rich structures, including glandular tissue, TDLU, ducts, arteries, and veins. Conversely, fat, skin, nipple, muscle, and ligament are categorized separately from breast tissue and are considered as background elements.

The comparison between the topological structures generated by our algorithm and the ground truth breast tissue is detailed in Tables 3

and 4. We denote the predicted topological mask, encompassing the union of 1D and 2D structures, as M_{topo} , and the ground truth mask of breast tissue as M_{tissue} . To assess this comparison, we employ four different quantitative measures:

1. **Topological precision:** the percentage of predicted topological mask being breast tissue,

$$\frac{|M_{topo} \cap M_{tissue}|}{|M_{topo}|};$$

2. **Topological recall:** the percentage of breast tissue that is covered by the predicted topological mask,

$$\frac{|M_{topo} \cap M_{tissue}|}{|M_{tissue}|};$$

3. **Distance from tissue to topological mask:** the mean Euclidean distance from a tissue voxel to its nearest voxel in the topological mask,

$$\frac{1}{|M_{tissue}|} \sum_{x \in M_{topo}} \min_{y \in M_{tissue}} \text{dist}(x, y);$$

4. **Distance from topological mask to tissue:** the mean Euclidean distance from a voxel in topological mask to its nearest tissue voxel,

$$\frac{1}{|M_{topo}|} \sum_{x \in M_{topo}} \min_{y \in M_{tissue}} \text{dist}(x, y).$$

The topological precision and recall results are presented in Table 3. Our experiment was conducted on a total of 100 volumes, comprising 25 volumes for each breast profile. We report the outcomes for these different profiles, providing both the mean and standard deviation. The topological precision numbers indicate that the majority of the extracted topological mask resides within the true breast tissues, showing over 93% alignment for all four profiles. This high percentage demonstrates the accuracy of the extracted topological mask in delineating the actual tissue. Regarding topological recall, the extracted topological mask's one-voxel thickness means it covers only a portion of the true tissue. In essence, the topological mask acts as a 'skeleton' of the breast tissue. This observation is highlighted by the topological recall of 26.28% for dense breast profile. For heterogeneous, scattered, and fatty breast profiles, where the tissue is much sparser, the topological recall shows a lower percentage, ranging between 17% and 20%.

Similar observations are noted in the two distance-based metrics presented in Table 4. The low mean distance from tissue to topological mask, consistent across all profiles, indicates that the topological mask effectively covers the tissue, with a topological mask voxel usually in close proximity to any tissue voxel. Regarding the distance from topological mask to tissue, we observe low values for dense and heterogeneous profiles, suggesting that the topological mask closely approximates the true tissue in these profiles. Conversely, for scattered and fatty breast profiles, this distance is larger, implying that in these profiles, where the tissue is much sparser, the topological structures tend to be over-predicted, leading to false positive structures that are far from the actual tissue.

Qualitative validation. Fig. 6 illustrates how topological structures effectively capture breast tissues. The first row presents 3D renderings of VICTRE-generated synthetic breast phantoms for four distinct profiles. The second and third rows display two slices at different positions of the corresponding breast phantoms, where the topological mask (in blue) and ground truth breast tissues (in white) are shown, alongside the breast outlines (in red). It is important to note that each slice's rendering includes several additional slices around the target cross sections for detailed examination. It is clear that the blue topological structures closely match the white breast tissues, especially noticeable in the scattered and fatty volumes. This observation highlights that our extracted topological masks effectively mirror the actual breast tissues, confirming the biological significance of these structures.

4. Experiments and results

In this section, we evaluate the efficacy of *TopoTxR* in predicting pCR response to neoadjuvant chemotherapy for breast cancers, utilizing data from both the public ISPY-1 dataset and a proprietary dataset from Rutgers.

Given that *TopoTxR* captures topological structures of high relevance approximating breast parenchymal structures, it offers substantial predictive accuracy in gauging response to neoadjuvant chemotherapy. The structure of this section is as follows: Section 4.1 provides details of the ISPY-1 and Rutgers datasets. Section 4.2 outlines implementation details, including experimental settings and preprocessing steps. Section 4.3 provides a detailed discussion of the baseline methods. Subsequently, Section 4.4 presents the quantitative results of our analysis. Finally, Section 4.5 is dedicated to presenting the findings from our ablation studies.

4.1. Dataset description

I-SPY1: The I-SPY1 DCE-MRI dataset, available via 'The Cancer Imaging Archive' (TCIA) from the I-SPY1/ACRIN 6657 study (Newitt and Hylton, 2016), includes longitudinal breast DCE-MRI scans evaluating the impact of neoadjuvant chemotherapy (NAC) on stage II or III breast cancer patients. All MRIs were performed on a 1.5 T scanner with a breast coil. The study involved patients with T3 tumors, measuring at least 3 cm in diameter, who were undergoing NAC. Initial volume acquisition was pre-contrast, followed by subsequent volume acquisition during and post gadolinium-based contrast agent administration. Pathological complete response (pCR) was used as a measure of the effectiveness of NAC. Our method was validated by predicting pCR using the initial MRIs taken four weeks prior to treatment from the I-SPY1 post-contrast DCE-MRI data. Our analysis included 161 patients, with 47 achieving pCR (average age = 48.8) and 114 not achieving pCR (average age = 48.5).

Rutgers proprietary dataset: We also applied our methods to a proprietary dataset comprising 120 patients, among whom 69 achieved pCR and 51 did not. In this study, the first phase T1-weighted fat-suppressed post-contrast DICOM images for each patient, captured 90 s post-contrast, were used. All imaging was performed on 1.5T MRI scanners, using gadolinium contrast agents (Gd-DTPA) at a dose of 0.2 mmol/kg, administered at a rate of 2.0 mL/s, followed by a 10 mL saline flush of 0.9%. This retrospective case-control study analyzed hospital records spanning from January 1, 2008, to June 30, 2023, and was conducted without direct contact with the participants.

We included 120 women from Robert Wood Johnson Barnabas Health hospitals, specifically Newark Beth Israel Medical Center and Cooperman Barnabas Medical Center, all of whom had confirmed diagnoses of invasive breast carcinoma. Tumors had to be clinically and radiologically measurable post-biopsy, at least 10 mm in diameter. Excluded were patients with prior cytotoxic regimens in the breast of interest, but those treated for contralateral breast cancer were included. Participants were above 18 years of age, neither pregnant nor lactating, had undergone a primary mass core biopsy, and had MRI-compatible health conditions. Tumors needed to be stage I-III, T4, any N, M0, inclusive of clinical or pathological inflammatory cancers, as well as regional stage IV, provided the only site of metastasis was the supraclavicular lymph nodes. Exclusion criteria included recent use (within 30 days) of investigational agents and allergies to compounds resembling MRI contrast dye, preventing contrast-enhanced examinations.

We would like to highlight that while our model is contingent on high quality dynamic contrast enhanced MRI imaging, its utility across the majority of breast centers is maintained given that most breast practices undergo American College of Radiology (ACR) accreditation. We agree that our model may not be generalizable for those practices which do not use high quality DCE-MRI or ACR accreditation standards. However, to ensure broad applicability, we have also applied our model to the national dataset used in the I-SPY1 trial.

4.2. Experiment details

We implemented our model using Pytorch (Paszke et al., 2019) and conducted all experiments on a single NVIDIA RTX A6000 GPU. Hyperparameters, including learning rate, momentum, loss weight factor, and dropout rate for the dropout layer, were fine-tuned using a grid search. Specifically, we conducted a 3-fold cross-validation on the I-SPY 1 dataset, testing each combination of parameters in the grid. The model training was executed using the Adam optimizer (Kingma and Ba, 2014), with the learning rate set at 0.01 and momentum at 0.9. The dropout rate in the last layer was fixed at 0.2, and the weight of the topology-guided mask loss, λ , was set at 0.01. For the focal loss, the class weight factors, θ and γ , were adjusted to 0.68 and 2.0, respectively, in line with the ratio of positive to negative samples in

Table 5

Comparative analysis of our proposed method, *TopoTxR*, with baseline methods across four metrics – accuracy, AUC, specificity, and sensitivity – on the I-SPY1 dataset, utilizing 10-fold cross-validation. The performance of *TopoTxR* integrated with non-imaging clinical features (denoted as *TopoTxR+Clinical*) is displayed in the final row of the table.

	Accuracy \uparrow	AUC \uparrow	Specificity \uparrow	Sensitivity \uparrow
Radiomics	0.563 \pm 0.085	0.593 \pm 0.098	0.552 \pm 0.180	0.575 \pm 0.081
PD	0.549 \pm 0.081	0.567 \pm 0.097	0.551 \pm 0.167	0.547 \pm 0.071
Radiomics+PD	0.563 \pm 0.093	0.587 \pm 0.099	0.592 \pm 0.178	0.534 \pm 0.087
ViT ^a	0.770 \pm 0.035	0.676 \pm 0.060	0.508 \pm 0.020	0.874 \pm 0.087
ConvNext ^a	0.745 \pm 0.047	0.688 \pm 0.069	0.572 \pm 0.155	0.805 \pm 0.070
ResNext ^a	0.740 \pm 0.036	0.643 \pm 0.088	0.373 \pm 0.229	0.842 \pm 0.108
DenseNet, (Huang et al., 2017)	0.814 \pm 0.057	0.821 \pm 0.031	0.827 \pm 0.023	0.816 \pm 0.079
DenseNet ^a w/ focal	0.789 \pm 0.043	0.729 \pm 0.053	0.627 \pm 0.136	0.830 \pm 0.071
DenseNet-KD, (Du et al., 2022)	0.905 \pm 0.021	0.874 \pm 0.016	0.825 \pm 0.043	0.923 \pm 0.030
TopoTxR w/o focal or TGSA	0.856 \pm 0.046	0.807 \pm 0.046	0.664 \pm 0.123	0.949 \pm 0.056
TopoTxR w/o focal	0.913 \pm 0.031	0.918 \pm 0.031	0.944 \pm 0.057	0.892 \pm 0.051
Self-Attention w/ focal	0.920 \pm 0.030	0.922 \pm 0.030	0.916 \pm 0.065	0.927 \pm 0.051
TopoTxR(Ours)	0.931 \pm 0.027	0.917 \pm 0.036	0.885 \pm 0.068	0.950 \pm 0.033
TopoTxR+Clinical	0.925 \pm 0.023	0.926 \pm 0.028	0.911 \pm 0.056	0.940 \pm 0.035

^a Models represent the 3D versions we developed specifically for this analysis.

our training set. All input MRIs were of size 256^3 , and the model was trained for 200 epochs with a batch size of 2, amounting to a total training time of approximately four hours. Inference for a single MRI takes less than one minute.

Preprocessing. Preprocessing comprises two phases to prepare the data from both the I-SPY1 and Rutgers proprietary datasets for effective network training. Initially, we focus on noise elimination and standardization of the samples into a uniform shape. Specifically, images with excessive background noise were excluded. For the remaining data, we employed dilation and erosion techniques to further reduce any residual background noise. Subsequently, each MRI volume was cropped to the foreground, ensuring a 2-voxel width margin in each dimension. If any dimension of the cropped MRI volume exceeded 256, we resized it while preserving the original aspect ratio, ensuring that the largest dimension did not exceed 256. To center the foreground in each volume, we applied appropriate padding, resulting in a standardized volume size of $256 \times 256 \times 256$. This uniformity facilitates seamless integration with deep neural network training. We normalized the volumes' intensity and inverted the intensity values by multiplying each voxel by -1 , thus preparing them for subsequent persistent homology computation. The second step involves applying a persistence threshold, selecting dimensions of topological structures, and applying dilation to these structures before using them as inputs for the network training. An ablation study of this second step is detailed in Section 4.5.

4.3. Baselines

We conduct a comparative analysis using various baseline methods on both the I-SPY1 and Rutgers proprietary datasets. Due to significant differences between the Rutgers dataset and I-SPY1, we evaluate them separately.³ We perform a 10-fold cross-validation on the I-SPY1 dataset and the Rutgers dataset. The baselines are listed below:

- **Radiomics:** We compute a 92-dimensional radiomic signature (Van Griethuysen et al., 2017b) and train a classifier on this signature. Features are extracted solely from the tumor region.

³ The I-SPY1 and Rutgers datasets underwent preprocessing protocols with different hyperparameters, including the selection of minimum tumor size and application of distinct contrast enhancements. Furthermore, the inclusion of patients with stage I or T4 in the Rutgers dataset contributes to variations in the final image formation compared to I-SPY1.

- **Persistent Diagram (PD):** A classifier is trained using topological features derived from the persistence diagrams (PDs) of the MRI images. While various classifier options are available and behave similarly, we use the sliced Wasserstein kernel distance for PDs as the feature vector (Carriere et al., 2017).
- **Radiomics+PD:** This method involves a combination of both radiomic and PD features for classifier training.

We implement feature selection for all the aforementioned methods using Mutual Information Difference (MID) and Mutual Information Quotient (MIQ). For all baseline features, an exhaustive search is conducted across all combinations of feature selection schemes and a range of classifiers, including Random Forests, Linear Discriminant Analysis, Quadratic Discriminant Analysis, and Support Vector Machine. We report the results based on the best-performing combinations.

- **DenseNet, DenseNet with focal loss, DenseNet-KD:** We perform a comparative analysis of our method against DenseNet, DenseNet with focal loss, and the state-of-the-art method, DenseNet-KD, using the I-SPY1 dataset. DenseNet (Huang et al., 2017) processes raw 3D MRI images directly as input. In contrast, DenseNet-KD (Du et al., 2022) not only uses 3D MRI images but also integrates topological pseudo labels, particularly Betti number curves, for joint training with the DenseNet backbone. This approach leads to noticeable improvements. The results of this comparative analysis are detailed in Table 5.
- **ConvNext, ResNext, ViT:** We compare our method against state-of-the-art architectures—ConvNext (Liu et al., 2022), ResNext (Xie et al., 2017), and transformer-based ViT (Dosovitskiy et al., 2020)—utilizing both the I-SPY 1 dataset and our proprietary Rutgers dataset. The results are presented in Table 5 for the I-SPY 1 dataset and Table 6 for the Rutgers dataset. It is important to note that, like DenseNet, ConvNext, ResNext, and ViT were originally designed for 2D inputs. We have developed 3D versions of these models to conduct a comprehensive comparative analysis.
- **TopoTxR+Clinical:** Inspired by the work of Duanmu et al. (2020), we gather non-imaging clinical data for the I-SPY1 dataset, which includes demographic information (age, race), Estrogen Receptor Status (ER), Progesterone Receptor Status (PR), Hormone Receptor Status (HR), Human Epidermal Growth Factor Receptor 2 (HER2) Status, and a 3-level HR/HER2 categorization. We utilize this information to construct a non-imaging clinical feature vector. This vector is then concatenated with the feature maps from both 3D CNN branches, forming the input for the fully connected (classification) layer to facilitate pCR prediction.

Table 6

Comparative analysis of our proposed method, *TopoTxR*, with baseline methods across four metrics – accuracy, AUC, specificity, and sensitivity – on the proprietary Rutgers dataset, utilizing 10-fold cross-validation.

	Accuracy↑	AUC↑	Specificity↑	Sensitivity↑
DenseNet ^a	0.783 ± 0.047	0.740 ± 0.047	0.781 ± 0.096	0.698 ± 0.151
ViT ^a	0.800 ± 0.041	0.757 ± 0.058	0.819 ± 0.094	0.696 ± 0.132
ConvNext ^a	0.725 ± 0.066	0.719 ± 0.075	0.784 ± 0.097	0.655 ± 0.154
ResNext ^a	0.758 ± 0.049	0.745 ± 0.046	0.795 ± 0.132	0.695 ± 0.138
<i>TopoTxR</i> w/o focal or TGSA	0.883 ± 0.041	0.862 ± 0.051	1.000 ± 0.000	0.724 ± 0.102
<i>TopoTxR</i> w/o TGSA	0.900 ± 0.031	0.878 ± 0.042	0.988 ± 0.023	0.769 ± 0.093
<i>TopoTxR</i> (Ours)	0.925 ± 0.028	0.900 ± 0.055	0.934 ± 0.051	0.867 ± 0.127

^a Models represent the 3D versions we developed specifically for this analysis.

4.4. Quantitative results

The quantitative results on the I-SPY1 dataset are presented in [Table 5](#). Our method, which incorporates focal loss and topology-guided spatial attention, is compared against various baselines across 4 metrics, namely, accuracy, area under the curve (AUC), specificity, and sensitivity. Notably, our proposed method outperforms the state-of-the-art DenseNet-KD method in all four metrics. This achievement highlights the superiority of the topological mask over the Betti curve, which is the main feature used by DenseNet-KD. The richer information content of the topological mask compared to the Betti curve likely contributes to this enhanced performance. Additionally, our method demonstrates a more balanced specificity and sensitivity, attributed to the effective use of focal loss.

We substitute the topology-guided spatial attention module in our framework with a self-attention module and compare against it. As shown in row 12 of [Table 5](#), the topology-guided spatial attention proves to be significantly more effective for pCR prediction. This effectiveness can be attributed to the topological mask’s ability to accurately capture biologically relevant breast tissue. Additionally, the results of incorporating non-imaging clinical features into our framework are presented in the last row of [Table 5](#). These results indicate that the inclusion of clinical data does not significantly impact the efficacy of our model.

Given the distinct imaging characteristics between the I-SPY1 and Rutgers datasets (as detailed in [Section 4.1](#)), it is not feasible to train *TopoTxR* on one dataset and test it on the other. Consequently, we conduct an independent evaluation using the Rutgers dataset, with the results presented in [Table 6](#). The results presented in both tables demonstrate that our proposed method consistently outperforms the baselines, including state-of-the-art models such as DenseNet, DenseNet-KD, ConvNext, ResNext, and ViT across all datasets. This underscores the effectiveness of our approach and its adept integration of topological knowledge for predicting pCR.

Performance analysis across breast tissue density groups. We evaluated the performance of the proposed model across various breast tissue densities to assess its clinical applicability. Due to the lack of density labels in the public I-SPY 1 dataset, we classified the samples from our proprietary Rutgers dataset into four groups based on the density of fibroglandular tissue (FGT), as assessed by our radiologists (N.S. and L.P.). The categorization is as follows:

- **Dense:** >75% of the breast comprises FGT.
- **Heterogeneous:** 51%–75% of the breast comprises FGT.
- **Scattered:** 25%–50% of the breast comprises FGT.
- **Fatty:** <25% of the breast comprises FGT.

We now present the performance of our model across the various breast tissue density groups. According to [Table 7](#), our method consistently delivers superior performance across all density groups, achieving perfect results on the fatty volumes. This highlights the robustness of our proposed model across different breast tissue types.

Table 7

TopoTxR pCR prediction results for four different densities of fibroglandular tissue in the Rutgers dataset.

	Accuracy↑	AUC↑	Specificity↑	Sensitivity↑
Dense	0.952	0.929	0.857	1.000
Hetero	0.913	0.912	0.905	0.920
Scattered	0.902	0.898	0.875	0.920
Fatty	1.000	1.000	1.000	1.000

4.5. Ablation studies

***TopoTxR* w/o focal loss or TGSA.** To highlight the improvements introduced by the focal loss and the topology-guided spatial attention module (TGSA), we have performed a comparative analysis with the initial version of *TopoTxR*. This earlier version, which lacks both focal loss and TGSA, was published in a previous conference paper ([Wang et al., 2021](#)). The initial *TopoTxR* also employs persistent homology for extracting 1D and 2D topological cycles, as elaborated in [Section 2.2](#). The primary distinction lies in its approach to utilizing these masks; the initial version of *TopoTxR* simply inputs the “soft” topological mask (the product of MRI images and binary topological masks) into the CNNs. In contrast, the method proposed in this paper utilizes the topological masks as an explicit guiding mechanism to train the topology-guided spatial attention module. As demonstrated in rows 10 and 13 of [Table 5](#), our proposed method, which incorporates both the focal loss during training and the TGSA module, surpasses the performance of the initial *TopoTxR* and its variants.

Observations from [Table 5](#) reveal that even in the absence of focal loss or TGSA, the initial *TopoTxR* markedly outperforms the baseline methods, specifically those utilizing Radiomics, PD features, or both. Furthermore, when compared to DenseNet, a deep learning method devoid of topology information, the initial *TopoTxR* still shows superior performance. These results underscore the significant value of the proposed usage of topological masks in pCR classification, which is not captured by radiomics, simple topological features like PD, or conventional CNN application.

***TopoTxR* w/o focal loss.** We carry out an additional experiment on I-SPY1 dataset to assess the significance of the TGSA module. As indicated in rows 11 and 13 of [Table 5](#), TGSA effectively directs the attention of the CNNs towards biologically relevant regions for pCR prediction, resulting in a substantial enhancement of performance. However, this approach still falls short of the performance achieved by the model that incorporates focal loss.

***TopoTxR* w/o TGSA.** The performance of *TopoTxR* when using only focal loss has been assessed on the Rutgers dataset, with the results detailed in row 6 of [Table 6](#). It is evident from these results that the application of focal loss contributes to an improvement in performance.

Persistence Threshold. Recall that the persistence of a topological structure is the difference between its birth and death times. To ensure accuracy, we eliminate topological structures with low persistence, as these are typically noise-induced and could adversely affect our

Table 8
Ablation study results. All numbers are reported based on a 10-fold cross-validation on the I-SPY1 dataset.

	Accuracy \uparrow	AUC \uparrow	Specificity \uparrow	Sensitivity \uparrow
Persistence threshold				
90% Remain	0.913 \pm 0.026	0.873 \pm 0.058	0.817 \pm 0.113	0.928 \pm 0.052
60% Remain	0.906 \pm 0.036	0.864 \pm 0.058	0.769 \pm 0.128	0.960 \pm 0.039
Dimension				
Dimension 1	0.900 \pm 0.026	0.896 \pm 0.035	0.880 \pm 0.070	0.912 \pm 0.036
Dimension 2	0.863 \pm 0.038	0.845 \pm 0.060	0.792 \pm 0.141	0.899 \pm 0.055
MRI+Dim1+Dim2	0.863 \pm 0.042	0.825 \pm 0.058	0.675 \pm 0.124	0.954 \pm 0.043
MRI	0.633 \pm 0.200	0.621 \pm 0.102	0.570 \pm 0.322	0.673 \pm 0.354
Dilation radius				
Radius 2	0.894 \pm 0.043	0.873 \pm 0.058	0.817 \pm 0.113	0.928 \pm 0.052
Radius 4	0.863 \pm 0.034	0.860 \pm 0.063	0.827 \pm 0.152	0.893 \pm 0.045
Radius 8	0.913 \pm 0.031	0.913 \pm 0.031	0.911 \pm 0.075	0.907 \pm 0.061
Insertion point for the Topology-guided spatial attention module				
After 2nd Conv	0.894 \pm 0.018	0.908 \pm 0.025	0.919 \pm 0.064	0.898 \pm 0.042
After 4th Conv	0.889 \pm 0.042	0.879 \pm 0.050	0.882 \pm 0.104	0.876 \pm 0.085

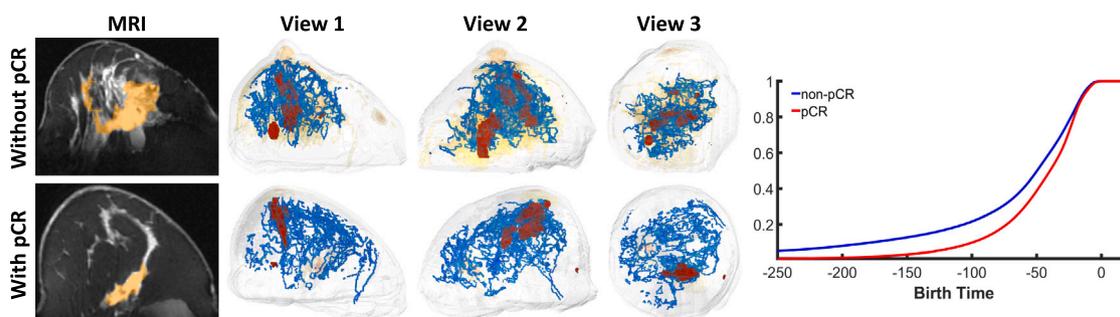


Fig. 7. Qualitative comparison of patients with and without pCR. First column: Slices of breast DCE-MRIs with tumor masked in orange (tumor masks are not used in *TopoTxR*). Columns 2–4: 3D renderings of topological structures from three different views. 1-D structures (loops) are rendered in blue and 2-D structures (bubbles) in red. Right: cumulative density function of topological structures' birth times.

results. We assess the effects of persistence on the I-SPY1 dataset by applying three distinct thresholds, retaining 90%, 60%, and 30% of the structures respectively. These results are detailed in Table 8. As indicated in the table, retaining 30% of the topological structures, which is our default setting as shown in row 13 of Table 5, offers an optimal equilibrium between the quantity and quality of these structures.

Dimension of Topological Structures. We also evaluate our method using only 1D structures (loops) and only 2D structures (bubbles). Each of these approaches outperform the baseline methods that do not incorporate topology information, as evidenced in row 3 and 4 of Table 8. However, they are still not as effective as *TopoTxR*, which utilizes both 1D and 2D topological structures. This outcome indicates that 1D and 2D structures provide complementary information, which is crucial for accurate pCR prediction. For comparison, we extract 1D and 2D soft topological masks as approximations of fibroglandular tissues and stack them alongside the MRI volumes to form three-channel inputs for training a 3D CNN. The results are documented in row 5 of Table 8. While this method shows notable improvement over using MRIs alone (as shown in row 6), due to the additional information provided by the fibroglandular tissue approximation, it does not direct the model's attention as intentionally as the proposed model, resulting in inferior outcomes.

Dilation Radius. In the process of generating topological masks, rather than using binary masks directly, we create 'soft' masks by multiplying these binary masks with the corresponding MRIs. We conduct an ablation study using I-SPY1 dataset to examine the impact of varying the radius of dilation operation on binary masks prior to this multiplication. The data presented in row 7, 8, and 9 of Table 8 indicates that the best

performance is achieved when no dilation operation is applied to the binary masks.

Insertion Point for Topology-Guided Spatial Attention Module.

Finally, we examine how the insertion point of the TGSA module after different layers affects performance. In addition to our default setting, where TGSA receives feature maps from the third convolution layer, we also examine its placement after the second and fourth layers, adjusting the size of the topological masks accordingly. The findings of this study are detailed in the last two rows of Table 8. Our observations indicate that the default setting, as shown in row 13 of Table 5, yields significantly better results.

In summary, the most effective results are obtained when 30% of the topological structures are retained, by employing a combination of both 1D and 2D structures without any dilation, and by applying the topology-guided spatial attention module after the third convolution layer.

5. Qualitative analysis

5.1. Topological feature interpretation

The topological structures identified through persistent homology effectively capture the breast tissue structures. Learning directly from these structures and their vicinity provides the opportunity for interpreting the learning outcomes and drawing novel biological insights. Here we provide some visual analysis as a proof of concept.

Fig. 7 presents topological structures from three distinct perspectives, showcasing one representative sample each for cases with and without pCR. We observe that the 1D and 2D structures are sparser in

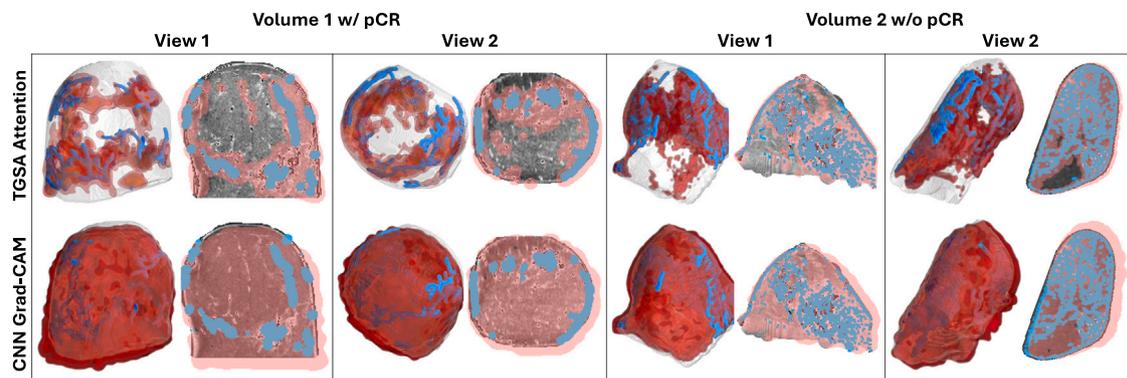


Fig. 8. Attention maps of CNNs when making decisions about pCR predictions. Red: voxels contributing to decisions of CNNs; blue: voxels corresponding to extracted topological structures. Columns 1, 3, 5, and 7 display 3D renderings of the attention maps, while columns 2, 4, 6, and 8 present cross-sections corresponding to the 3D renderings on their left.

the case with a pCR response, while they appear denser in the non-pCR case. Examining the corresponding MRI images, we notice that the breast with pCR exhibits scattered fibroglandular tissue and minimal background parenchymal enhancement. In contrast, the non-pCR breast displays more heterogeneous fibroglandular tissue accompanied by moderate background parenchymal enhancement. This observation suggests that the topological structures may be capturing the intricate fibroglandular structure, potentially serving as an indicator of treatment response.

We also analyze the topological characteristics of two patient populations by comparing their birth times, which indicate the threshold at which a cycle first appears. Given our use of the inverse image in the experiments, denoted as $f = -I$, meaning the birth time essentially corresponds to -1 times the brightness of a structure. For each population (pCR and non-pCR), we compute the Cumulative Density Function (CDF) curve using the average birth time from all samples in that group. As depicted in Fig. 7 (right), the CDF for pCR patients (in red) compared to non-pCR patients (in blue) reveals that topological structures in pCR patients' tissues tend to be less bright or distinguishable than those in non-pCR patients, an observation that aligns with our qualitative assessments. To statistically validate these differences in CDFs, we conduct a Kolmogorov–Smirnov test (Massey, 1951). The resulting p -value of 0.0002 indicates a significant disparity in the birth time distributions between the pCR and non-pCR patient groups.

5.2. Validation on the model attention

The key idea in TopoTxR lies in its ability to direct CNN's attention to a much smaller set of voxels with high biological relevance manifested by the extracted topological structures. As shown in Section 3, these structures have a strong connection with the breast tissue. In this section, we prove that voxels corresponding to topological structures play an essential role when CNN is making a pCR prediction. We visualize the spatial attention using a post-hoc heatmap visualization technique – Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al., 2017). Grad-CAM generalizes Class Activation Mapping (CAM) (Zhou et al., 2016) by lifting its constraints on model complexity and provides good visual explanations for model decisions. It generates a coarse localization map that highlights key regions in the imaging data pivotal for predicting a specific concept. Our aim is to utilize Grad-CAM to gain interpretative insights into the decision-making process of TopoTxR. It is important to note that Grad-CAM does not alter the learning mechanism of the model. It is applied post-training to a model with fixed weights, serving purely as a tool to shed light on how the model arrives at its decisions.

We employ the Grad-CAM implementation from the M3d-CAM library (Gotkowski et al., 2021) to visualize attention maps following

the Topology-Guided Spatial Attention (TGSA) module in the best-performing *TopoTxR* model. For comparison, we modify the *TopoTxR* model to exclude the TGSA module, utilizing a single 3D CNN branch that processes only raw MRI data. In this setup, Grad-CAM is applied after the third convolution layer to visualize attention maps, as depicted in the second row—'CNN Grad-CAM' in Fig. 8. Conversely, the attention maps from our proposed model are shown in the first row—'TGSA Attention'. In Fig. 8, we provide a comparative analysis of the attention maps from two views, showcasing one representative sample each for cases with and without pCR. This comparison highlights that the TGSA module effectively directs the CNN's focus to specific, biologically relevant regions, as delineated by topological structures. Conversely, the model lacking topology information shows a more scattered and dispersed attention distribution across the MRI volume. It is interesting to note that the topological structures appear sparser in cases with a pCR, whereas they are denser in cases without pCR.

6. Discussion

Breast composition in MRI is characterized by the amount of fibroglandular tissue (FGT) and the extent of Background Parenchymal Enhancement (BPE) following contrast administration, in accordance with the American College of Radiology (ACR) Breast Imaging Reporting and Data System (BI-RADS) standards. FGT is classified into four categories: fatty fibroglandular tissue ($\leq 25\%$), scattered fibroglandular tissue (25%–50%), heterogeneous fibroglandular tissue (50%–75%), and extreme fibroglandular tissue ($\geq 75\%$). BPE is rated as minimal, mild, moderate, or marked, based on the FGT enhancement relative to the total FGT volume during the initial phase of T1-weighted fat-suppressed post-contrast imaging, which is captured 90 s after contrast agent administration.

Recent studies indicate that elevations in BPE correlate with an increased risk of breast cancer development and changes in BPE can predict NAC response (Chen et al., 2015; Liao et al., 2020; Oh et al., 2018; Preibsch et al., 2016; Van Der Velden et al., 2015; You et al., 2017, 2018). Our work hinges on the idea that breast composition elucidated on MRI can predict treatment response in pre-neoadjuvant chemotherapy MRI scans. Two notable studies (You et al., 2017, 2018) highlight that post-chemotherapy reductions in BPE significantly correlate with pCR, both in HER2-positive (You et al., 2018) and HER2-negative breast cancers (You et al., 2017). The prevailing hypothesis is that tumor-adjacent vascular permeability might echo BPE fluctuations before and after NAC administration, with a general decline in BPE after NAC (Rella et al., 2020).

FGT resembles the concept of mammographic density, a recognized independent breast cancer risk factor (Duffy et al., 2018). When assessing both FGT and BPE using MRI, research suggests that elevations in BPE, regardless of FGT presence (e.g., moderate to marked BPE in

breasts that are primarily fatty or have scattered FGT), can indicate an increased cancer risk (Arasu et al., 2019). Thus, BPE is considered more instrumental than FGT for evaluating NAC response and shaping prediction models (Arasu et al., 2019).

7. Conclusion

This paper introduces *TopoTxR*, a novel topological biomarker that capitalizes on the rich geometric information inherent in structural MRI to enhance downstream CNN processing. To harness the intrinsic topological information effectively, we integrate a topology-guided spatial attention mechanism. Our model combines information from raw MRIs and topological masks, addressing the sample imbalance problem in datasets using focal loss. Specifically, we compute 1D cycles and 2D bubbles from breast DCE-MRIs employing the theory of persistent homology. These topological structures are then utilized to guide neural network attention by supervising the generation of attention maps. Furthermore, we demonstrate the predictive power of *TopoTxR* in forecasting pCR using treatment-naïve imaging.

CRedit authorship contribution statement

Fan Wang: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Zhilin Zou:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Nicole Sakla:** Writing – review & editing, Data curation. **Luke Partyka:** Data curation. **Nil Rawal:** Data curation. **Gagandeep Singh:** Data curation. **Wei Zhao:** Writing – review & editing, Data curation. **Haibin Ling:** Writing – review & editing, Methodology. **Chuan Huang:** Writing – review & editing, Data curation. **Prateek Prasanna:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization. **Chao Chen:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

All authors disclosed no relevant relationships.

Acknowledgments

This work received partial support from the following grants: National Science Foundation (NSF), USA 2144901 and National Cancer Institute (NCI), USA R21CA258493, R03CA223052, and R01CA297843. Ling was supported in part by National Science Foundation (NSF) Division of Information and Intelligent Systems (IIS), USA grants 2331769 and 2006665.

Data availability

Data will be made available on request.

References

Abdelhafiz, D., Yang, C., Ammar, R., Nabavi, S., 2019. Deep CNN for mammography: advances, challenges and applications. *BMC Bioinform.*

Adams, H., Emerson, T., Kirby, M., Neville, R., Peterson, C., Shipman, P., Chepushanova, S., Hanson, E., Motta, F., Ziegelmeier, L., 2017. Persistence images: A stable vector representation of persistent homology. *J. Mach. Learn. Res.* 18, 3990–4031.

Akazawa, K., Tamaki, Y., Taguchi, T., Tanji, Y., Miyoshi, Y., Kim, S.J., Ueda, S., Yanagisawa, T., Sato, Y., Tamura, S., et al., 2006. Preoperative evaluation of residual tumor extent by three-dimensional magnetic resonance imaging in breast cancer patients treated with neoadjuvant chemotherapy. *Breast J.* 12 (2), 130–137.

Arasu, V.A., Miglioretto, D.L., Sprague, B.L., Alsheik, N.H., Buist, D.S., Henderson, L.M., Herschorn, S.D., Lee, J.M., Onega, T., Rauscher, G.H., et al., 2019. Population-based assessment of the association between magnetic resonance imaging background parenchymal enhancement and future primary breast cancer risk. *J. Clin. Oncol.* 37 (12), 954.

Badano, A., Graff, C.G., Badal, A., Sharma, D., Zeng, R., Samuelson, F.W., Glick, S.J., Myers, K.J., 2018. Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an in silico imaging trial. *JAMA Netw. Open* 1 (7), e185474. <http://dx.doi.org/10.1001/jamanetworkopen.2018.5474>.

Braman, N.M., Etesami, M., Prasanna, P., Dubchuk, C., Gilmore, H., Tiwari, P., Plecha, D., Madabhushi, A., 2017. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. *Breast Cancer Res.*

Braman, N., Prasanna, P., Bera, K., Ailou, M., Khorrami, M., Leo, P., Etesami, M., Vulchi, M., Turk, P., Gupta, A., et al., 2022. Novel radiomic measurements of tumor-associated vasculature morphology on clinical imaging as a biomarker of treatment response in multiple cancers. *Clin. Cancer Res.* 28 (20), 4410–4424.

Braman, N., Prasanna, P., Whitney, J., Singh, S., Beig, N., Etesami, M., Bates, D.D., Gallagher, K., Bloch, B.N., Vulchi, M., et al., 2019. Association of peritumoral radiomics with tumor biology and pathologic response to preoperative targeted therapy for HER2 (ERBB2)-positive breast cancer. *JAMA Netw. Open* 2 (4).

Burt, J.R., Torosdagli, N., Khosravan, N., RaviPrakash, H., Mortazi, A., Tissavirasingham, F., Hussein, S., Bagci, U., 2018. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. *Br. J. Radiol.* 91 (1089), 20170545.

Cain, E., Saha, A., Harowicz, M., Marks, J., Marcom, P., Mazurowski, M., 2019. Multivariate ML models for prediction of pCR to NAC in BCa using MRI features: a study using an independent validation set. *BCa Res. Treat.*

Carriere, M., Cuturi, M., Oudot, S., 2017. Sliced Wasserstein kernel for persistence diagrams. In: *International Conference on Machine Learning*. PMLR, pp. 664–673.

Chen, J.H., Hon, J.Y., Hsu, C., Mehta, R.S., Carpenter, P.M., Su, M.Y., 2015. Background parenchymal enhancement of the contralateral normal breast: association with tumor response in breast cancer patients receiving neoadjuvant chemotherapy. *Transl. Oncol.* 8 (3), 204–209.

Clough, J., Byrne, N., Oksuz, I., Zimmer, V.A., Schnabel, J.A., King, A., 2020. A topological loss function for deep-learning based image segmentation using persistent homology. TPAMI.

Cohen-Steiner, D., Edelsbrunner, H., Harer, J., 2005. Stability of persistence diagrams. In: *Proceedings of the Twenty-First Annual Symposium on Computational Geometry*. pp. 263–271.

Dalmış, M.U., Litjens, G., Holland, K., Setio, A., Mann, R., Karssemeijer, N., Gubern-Mérida, A., 2017. Using deep learning to segment breast and fibroglandular tissue in mri volumes. *Med. Phys.* 44 (2), 533–546.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2020. An image is worth 16x16 words: Transformers for image recognition at scale.

Du, S., Lao, Q., Kang, Q., Li, Y., Jiang, Z., Zhao, Y., Li, K., 2022. Distilling knowledge from topological representations for pathological complete response prediction. In: *Medical Image Computing and Computer Assisted Intervention—MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part II*. Springer, pp. 56–65.

Duanmu, H., Huang, P.B., Brahmavar, S., Lin, S., Ren, T., Kong, J., Wang, F., Duong, T.Q., 2020. Prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer using deep learning with integrative imaging, molecular and demographic data. In: *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23*. Springer, pp. 242–252.

Duffy, S.W., Morrish, O.W., Allgood, P.C., Black, R., Gillan, M.G., Willsher, P., Cooke, J., Duncan, K.A., Mitchell, M.J., Dobson, H.M., et al., 2018. Mammographic density and breast cancer risk in breast screening assessment cases and women with a family history of breast cancer. *Eur. J. Cancer* 88, 48–56.

Edelsbrunner, H., Harer, J., 2010. *Computational Topology: an Introduction*. Am. Math. Soc.

Feng, R., Xie, Y., Lai, M., Chen, D.Z., Cao, J., Wu, J., 2021. AGMI: Attention-guided multi-omics integration for drug response prediction with graph neural networks. In: *2021 IEEE International Conference on Bioinformatics and Biomedicine. BIBM*, pp. 1295–1298. <http://dx.doi.org/10.1109/BIBM52615.2021.9669314>.

Gotkowsky, K., Gonzalez, C., Bucher, A., Mukhopadhyay, A., 2021. M3d-CAM: A PyTorch library to generate 3D attention maps for medical deep learning. In: *Bildverarbeitung FÜR Die Medizin 2021: Proceedings, German Workshop on Medical Image Computing, Regensburg, March 7–9, 2021*. Springer, pp. 217–222.

Grimm, L., 2015. Breast MRI radiogenomics: Current status and research implications. *J. Magn. Reson. Imaging*.

Ha, R., Chin, C., Karcich, J., Liu, M., Chang, P., Mutasa, S., Sant, E., Wynn, R., Connolly, E., Jambawalikar, S., 2018. Prior to initiation of chemotherapy, can we predict breast tumor response? Deep learning convolutional neural networks approach using a breast MRI tumor dataset. *J. Digit. Imaging* <http://dx.doi.org/10.1007/s10278-018-0144-1>.

- Herrada, J., Iyer, R.B., Atkinson, E.N., Sneige, N., Buzdar, A.U., Hortobagyi, G.N., 1997. Relative value of physical examination, mammography, and breast sonography in evaluating the size of the primary tumor and regional lymph node metastases in women receiving neoadjuvant chemotherapy for locally advanced breast carcinoma. *Clin. Cancer Res.: Off. J. Am. Assoc. Cancer Res.* 3 (9), 1565–1569.
- Hofer, C., Kwitt, R., Niethammer, M., Uhl, A., 2017. Deep learning with topological signatures. *Adv. Neural Inf. Process. Syst.* 30.
- Hu, X., Li, F., Samaras, D., Chen, C., 2019. Topology-preserving deep image segmentation. *Adv. Neural Inf. Process. Syst.* 32.
- Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 4700–4708.
- Jarkman, S., Karlberg, M., Pocevičiūtė, M., Bodén, A., Bándi, P., Litjens, G., Lundström, C., Treanor, D., van der Laak, J., 2022. Generalization of deep learning in digital pathology: Experience in breast cancer metastasis detection. *Cancers* 14 (21), 5424.
- Keune, J.D., Jeffe, D.B., Schootman, M., Hoffman, A., Gillanders, W.E., Aft, R.L., 2010. Accuracy of ultrasonography and mammography in predicting pathologic response after neoadjuvant chemotherapy for breast cancer. *Am. J. Surg.* 199 (4), 477–484.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kooi, T., Litjens, G., Van Ginneken, B., Gubern-Mérida, A., Sánchez, C.I., Mann, R., den Heeten, A., Karssemeijer, N., 2017. Large scale deep learning for computer aided detection of mammographic lesions. *Med. Image Anal.* 35, 303–312.
- Kusano, G., Hiraoka, Y., Fukumizu, K., 2016. Persistence weighted Gaussian kernel for topological data analysis. In: *International Conference on Machine Learning*. PMLR, pp. 2004–2013.
- Lee, H., Kang, H., Chung, M.K., Kim, B.-N., Lee, D.S., 2012. Persistent brain network homology from the perspective of dendrogram. *IEEE Trans. Med. Imaging*.
- Li, H., Menegaux, A., Schmitz-Koep, B., Neubauer, A., Bäuerlein, F., Shit, S., Sorg, C., Menze, B., Hedderich, D., 2021. Automated claustrum segmentation in human brain MRI using deep learning. *Hum. Brain Mapp.* 42 (18), 5862–5872. <http://dx.doi.org/10.1002/hbm.25655>, URL: <https://eurekamag.com/research/079/334/079334358.php>.
- Liao, G.J., Henze Bancroft, L.C., Strigel, R.M., Chitalia, R.D., Kontos, D., Moy, L., Partridge, S.C., Rahbar, H., 2020. Background parenchymal enhancement on breast MRI: a comprehensive review. *J. Magn. Reson. Imaging* 51 (1), 43–61.
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal loss for dense object detection. In: *Proceedings of the IEEE International Conference on Computer Vision*. pp. 2980–2988.
- Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., Xie, S., 2022. A convnet for the 2020s. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 11976–11986.
- Liu, M.Z., Mutasa, S., Chang, P., Siddique, M., Jambawalikar, S., Ha, R., 2020a. A novel CNN algorithm for pathological complete response prediction using an I-SPY TRIAL breast MRI database. *Magn. Reson. Imaging* 73, 148–151.
- Liu, M.Z., Mutasa, S., Chang, P., Siddique, M., Jambawalikar, S., Ha, R., 2020b. A novel CNN algorithm for pathological complete response prediction using an I-SPY TRIAL breast MRI database. *Magn. Reson. Imaging*.
- Londero, V., Bazzocchi, M., Del Frate, C., Puglisi, F., Di Loreto, C., Francescutti, G., Zuiani, C., 2004. Locally advanced breast cancer: comparison of mammography, sonography and MR imaging in evaluation of residual disease in women receiving neoadjuvant chemotherapy. *Eur. Radiol.* 14, 1371–1379.
- ag Lundervold, A.S., Lundervold, A., 2019. An overview of deep learning in medical imaging focusing on MRI. *Z. Med. Phys.* 29 (2), 102–127.
- Lyu, Q., Namjoshi, S., McTyre, E., Topaloglu, U., Barcus, R., Chan, M., Cramer, C., Debinski, W., Gurcan, M., Lesser, G., Lin, H.-K., Munden, R., Pasche, B., Solingapuram Sai, K.K., Strowd, R., Tatter, S., Watabe, K., Zhang, W., Wang, G., Whitlow, C., 2022. A transformer-based deep-learning approach for classifying brain metastases into primary organ sites using clinical whole-brain MRI images. *Patterns* 3, 100613. <http://dx.doi.org/10.1016/j.patter.2022.100613>.
- Mani, S., Chen, Y., Li, X., Arlinghaus, L., Chakravarthy, A.B., Abramson, V., Bhavne, S.R., Levy, M.A., Xu, H., Yankeelov, T.E., 2013. Machine learning for predicting the response of breast cancer to neoadjuvant chemotherapy. *JAMIA*.
- Massey, Jr., F.J., 1951. The Kolmogorov-Smirnov test for goodness of fit. *J. Amer. Statist. Assoc.*
- Mazurowski, M., Buda, M., Saha, A., Bashir, M., 2018. Deep learning in radiology: an overview of the concepts and a survey of the state of the art. *J. Magn. Reson. Imaging* 49, <http://dx.doi.org/10.1002/jmri.26534>.
- Milosavljević, N., Morozov, D., Skraba, P., 2011. Zigzag persistent homology in matrix multiplication time. In: *SoCG*. pp. 216–225.
- von Minckwitz, G., Costa, S.D., Eiermann, W., Blohmer, J.-U., Tulusan, A.H., Jackisch, C., Kaufmann, M., 1999. Maximized reduction of primary breast tumor size using preoperative chemotherapy with doxorubicin and docetaxel. *J. Clin. Oncol.* 17 (7), 1999.
- Mishra, S., Wang, Y.X., Wei, C., Chen, D., Hu, X., 2021. VTG-Net: A CNN based vessel topology graph network for retinal artery/vein classification. *Front. Med.* 8, 750396. <http://dx.doi.org/10.3389/fmed.2021.750396>.
- Nawaz, A., Anwar, S.M., Liaqat, R., Iqbal, J., Bagci, U., Majid, M., 2020. Deep convolutional neural network based classification of alzheimer's disease using MRI data. In: *2020 IEEE 23rd International Multitopic Conference. INMIC, IEEE*. pp. 1–6.
- Newitt, D., Hylton, N., 2016. Multi-center breast DCE-mri data and segmentations from patients in the I-SPY 1/ACRIN 6657 trials. *Cancer Imaging Arch. doi*.
- Oh, S.J., Chae, E.Y., Cha, J.H., Shin, H.J., Choi, W.J., Kim, H.H., 2018. Relationship between background parenchymal enhancement on breast mri and pathological tumor response in breast cancer patients receiving neoadjuvant chemotherapy. *Br. J. Radiol.* 91 (1088), 20170550.
- Pang, Y., Xie, J., Khan, M.H., Anwer, R.M., Khan, F.S., Shao, L., 2019. Mask-guided attention network for occluded pedestrian detection. In: *2019 IEEE/CVF International Conference on Computer Vision. ICCV*, pp. 4966–4974. <http://dx.doi.org/10.1109/ICCV.2019.00507>.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimselshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. Pytorch: An imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems* 32. Curran Associates, Inc., pp. 8024–8035, URL: <http://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- Preibsch, H., Wanner, L., Bahrs, S., Wietek, B., Siegmund-Luz, K., Oberlecher, E., Hahn, M., Staebler, A., Nikolaou, K., Wiesinger, B., 2016. Background parenchymal enhancement in breast MRI before and after neoadjuvant chemotherapy: correlation with tumour response. *Eur. Radiol.* 26, 1590–1596.
- Qu, Y.-H., Zhu, H.-T., Cao, K., Li, X.-T., Ye, M., Sun, Y.-S., 2020. Prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer using a deep learning (DL) method. *Thorac. Cancer* <http://dx.doi.org/10.1111/1759-7714.13309>.
- Ranem, A., Kalkhof, J., Özer, C., Mukhopadhyay, A., Oksuz, I., 2022. Detecting respiratory motion artefacts for cardiovascular MRIs to ensure high-quality segmentation. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, pp. 447–456.
- Ravichandran, K., Braman, N., Janowczyk, A., Madabhushi, A., 2018. A deep learning classifier for prediction of pathological complete response to neoadjuvant chemotherapy from baseline breast DCE-mri. In: *Medical Imaging 2018: Computer-Aided Diagnosis*. Vol. 10575, SPIE, pp. 79–88.
- Reininghaus, J., Huber, S., Bauer, U., Kwitt, R., 2015. A stable multi-scale kernel for topological machine learning. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 4741–4748.
- Rella, R., Bufi, E., Belli, P., Petta, F., Serra, T., Masiello, V., Scrofani, A.R., Barone, R., Orlandi, A., Valentini, V., et al., 2020. Association between background parenchymal enhancement and tumor response in patients with breast cancer receiving neoadjuvant chemotherapy. *Diagn. Interv. Imaging* 101 (10), 649–655.
- Saha, A., Harowicz, M., Mazurowski, M., 2018. Breast cancer MRI radiomics: An overview of algorithmic features and impact of inter-reader variability in annotating tumors. *Med. Phys.* 45, <http://dx.doi.org/10.1002/mp.12925>.
- Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2017. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In: *ICCV. IEEE Computer Society*, pp. 618–626, URL: <http://dblp.uni-trier.de/db/conf/iccv/iccv2017.html#SelvarajuCDVPB17>.
- Shit, S., Paetzold, J.C., Sekuboyina, A.K., Ezhov, I., Unger, A., Zhykka, A., Plum, J.P.W., Bauer, U., Menze, B.H., 2020. Cldice - a novel topology-preserving loss function for tubular structure segmentation. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR*, pp. 16555–16564, URL: <https://api.semanticscholar.org/CorpusID:232428560>.
- Song, X., Chao, H., Xu, X., Guo, H., Xu, S., Turkbey, B., Wood, B.J., Sanford, T., Wang, G., Yan, P., 2022. Cross-modal attention for multi-modal image registration. *Med. Image Anal.* 82, 102612. <http://dx.doi.org/10.1016/j.media.2022.102612>, URL: <https://www.sciencedirect.com/science/article/pii/S1361841522002407>.
- Stucki, N., Paetzold, J.C., Shit, S., Menze, B., Bauer, U., 2023. Topologically faithful image segmentation via induced matching of persistence barcodes. In: *International Conference on Machine Learning*. PMLR, pp. 32698–32727.
- Su, Z., Niazi, M.K.K., Tavolara, T.E., Niu, S., Tozbikian, G.H., Wesolowski, R., Gurcan, M.N., 2023. BCR-Net: A deep learning framework to predict breast cancer recurrence from histopathology images. *Plos One* 18 (4), e0283562.
- Subasi, A., Kandpal, A.D., Raj, K.A., Bagci, U., 2023. Breast cancer detection from mammograms using artificial intelligence. In: *Applications of Artificial Intelligence in Medical Imaging*. Elsevier, pp. 109–136.
- Tack, A., Mukhopadhyay, A., Zachow, S., 2018. Knee menisci segmentation using convolutional neural networks: data from the osteoarthritis initiative. *Osteoarthr. Cartil.* 26 (5), 680–688.
- Van Der Velden, B.H., Dmitriev, I., Loo, C.E., Pijnappel, R.M., Gilhuijs, K.G., 2015. Association between parenchymal enhancement of the contralateral breast in dynamic contrast-enhanced MR imaging and outcome of patients with unilateral invasive breast cancer. *Radiology* 276 (3), 675–685.
- Van Griethuysen, J.J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G., Fillion-Robin, J.-C., Pieper, S., Aerts, H.J., 2017a. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 77 (21), e104–e107.

- Van Griethuysen, J.J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G., Fillion-Robin, J.-C., Pieper, S., Aerts, H.J., 2017b. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.*
- Vinnicombe, S.J., MacVicar, A.D., Guy, R.L., Sloane, J.P., Powles, T.J., Knee, G., Husband, J.E., 1996. Primary breast cancer: mammographic changes after neoadjuvant chemotherapy, with pathologic correlation. *Radiology* 198 (2), 333–340.
- Wang, H., Huang, G., Zhao, Z., Cheng, L., Juncker-Jensen, A., Nagy, M.L., Lu, X., Zhang, X., Chen, D.Z., 2023. CCF-GNN: A unified model aggregating appearance, microenvironment, and topology for pathology image classification. *IEEE Trans. Med. Imaging*.
- Wang, F., Kapse, S., Liu, S., Prasanna, P., Chen, C., 2021. TopoTxR: A topological biomarker for predicting treatment response in breast cancer. In: *Information Processing in Medical Imaging: 27th International Conference, IPMI 2021, Virtual Event, June 28–June 30, 2021, Proceedings*. Springer, pp. 386–397.
- Wang, F., Liu, H., Samaras, D., Chen, C., 2020. Topogan: A topology-aware generative adversarial network. In: *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16*. Springer, pp. 118–136.
- Woo, S., Park, J., Lee, J.-Y., Kweon, I.S., 2018. Cbam: Convolutional block attention module. In: *Proceedings of the European Conference on Computer Vision. ECCV*, pp. 3–19.
- Wu, P., Chen, C., Wang, Y., Zhang, S., Yuan, C., Qian, Z., Metaxas, D., Axel, L., 2017. Optimal topological cycles and their application in cardiac trabeculae restoration. In: *Information Processing in Medical Imaging: 25th International Conference, IPMI 2017, Boone, NC, USA, June 25–30, 2017, Proceedings 25*. Springer, pp. 80–92.
- Wu, J., Ye, X., Mou, C., Dai, W., 2023. Fineehr: Refine clinical note representations to improve mortality prediction. In: *2023 11th International Symposium on Digital Forensics and Security. ISDFS*, pp. 1–6.
- Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K., 2017. Aggregated residual transformations for deep neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1492–1500.
- Yao, J., Hagemann, N., Xiong, Q., Chen, J., Hermann, D.M., Chen, C., 2024. Topological analysis of mouse brain vasculature via 3d light-sheet microscopy images. *arXiv: 2402.16894*.
- You, C., Gu, Y., Peng, W., Li, J., Shen, X., Liu, G., Peng, W., 2018. Decreased background parenchymal enhancement of the contralateral breast after two cycles of neoadjuvant chemotherapy is associated with tumor response in HER2-positive breast cancer. *Acta Radiol.* 59 (7), 806–812.
- You, C., Peng, W., Zhi, W., He, M., Liu, G., Xie, L., Jiang, L., Hu, X., Shen, X., Gu, Y., 2017. Association between background parenchymal enhancement and pathologic complete remission throughout the neoadjuvant chemotherapy in breast cancer patients. *Transl. Oncol.* 10 (5), 786–792.
- Zhang, J., Saha, A., Zhu, Z., Mazurowski, M., 2018. Hierarchical convolutional neural networks for segmentation of breast tumors in MRI with application to radiogenomics. *IEEE Trans. Med. Imaging* PP, 1. <http://dx.doi.org/10.1109/TMI.2018.2865671>.
- Zhang, X., Wu, P., Yuan, C., Wang, Y., Metaxas, D.N., Chen, C., 2019. Heuristic search for homology localization problem and its application in cardiac trabeculae reconstruction. In: *IJCAI*. pp. 1312–1318.
- Zhou, B., Khosla, A., A., L., Oliva, A., Torralba, A., 2016. Learning deep features for discriminative localization. *CVPR*.
- Zhu, Q., Du, B., Turkbey, B., Choyke, P.L., Yan, P., 2017. Deeply-supervised CNN for prostate segmentation. In: *2017 International Joint Conference on Neural Networks. IJCNN, IEEE*, pp. 178–184.