

RESEARCH

Open Access



# Digital image analysis and machine learning-assisted prediction of neoadjuvant chemotherapy response in triple-negative breast cancer

Timothy B. Fisher<sup>1</sup>, Geetanjali Saini<sup>2</sup>, T. S. Rekha<sup>3</sup>, Jayashree Krishnamurthy<sup>3</sup>, Shruti Bhattarai<sup>2</sup>, Grace Callagy<sup>4</sup>, Mark Webber<sup>4</sup>, Emiel A. M. Janssen<sup>5,7</sup>, Jun Kong<sup>6\*</sup> and Ritu Aneja<sup>1,2\*</sup>

## Abstract

**Background** Pathological complete response (pCR) is associated with favorable prognosis in patients with triple-negative breast cancer (TNBC). However, only 30–40% of TNBC patients treated with neoadjuvant chemotherapy (NAC) show pCR, while the remaining 60–70% show residual disease (RD). The role of the tumor microenvironment in NAC response in patients with TNBC remains unclear. In this study, we developed a machine learning-based two-step pipeline to distinguish between various histological components in hematoxylin and eosin (H&E)-stained whole slide images (WSIs) of TNBC tissue biopsies and to identify histological features that can predict NAC response.

**Methods** H&E-stained WSIs of treatment-naïve biopsies from 85 patients (51 with pCR and 34 with RD) of the model development cohort and 79 patients (41 with pCR and 38 with RD) of the validation cohort were separated through a stratified eightfold cross-validation strategy for the first step and leave-one-out cross-validation strategy for the second step. A tile-level histology label prediction pipeline and four machine-learning classifiers were used to analyze 468,043 tiles of WSIs. The best-trained classifier used 55 texture features from each tile to produce a probability profile during testing. The predicted histology classes were used to generate a histology classification map of the spatial distributions of different tissue regions. A patient-level NAC response prediction pipeline was trained with features derived from paired histology classification maps. The top graph-based features capturing the relevant spatial information across the different histological classes were provided to the radial basis function kernel support vector machine (rbfSVM) classifier for NAC treatment response prediction.

**Results** The tile-level prediction pipeline achieved 86.72% accuracy for histology class classification, while the patient-level pipeline achieved 83.53% NAC response (pCR vs. RD) prediction accuracy of the model development cohort. The model was validated with an independent cohort with tile histology validation accuracy of 83.59% and NAC prediction accuracy of 81.01%. The histological class pairs with the strongest NAC response predictive ability were tumor and tumor-infiltrating lymphocytes for pCR and microvessel density and polyploid giant cancer cells for RD.

\*Correspondence:

Jun Kong  
jkong@gsu.edu  
Ritu Aneja  
raneja@uab.edu

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Conclusion** Our machine learning pipeline can robustly identify clinically relevant histological classes that predict NAC response in TNBC patients and may help guide patient selection for NAC treatment.

**Keywords** Triple-negative breast cancer, Neoadjuvant chemotherapy, Machine learning, Digital image analysis, Feature engineering

## Introduction

Triple-negative breast cancer (TNBC) is an aggressive breast cancer subtype that lacks expression of estrogen, progesterone, and human epidermal growth factor 2 receptors [1]. TNBC accounts for 15–20% of all breast cancers, affecting nearly half a million women in the USA each year [2, 3]. The 5-year survival rate of TNBC patients is 15% lower than that of patients with other breast cancer subtypes [1]. At the time of diagnosis, TNBCs tend to have a more advanced histologic grade and larger size compared to hormone-positive breast cancers [4]. TNBCs also have higher recurrence and metastasis rates and typically metastasize to the brain, lung, and liver [5, 6]. No targeted or endocrine therapy is available for TNBC, and neoadjuvant chemotherapy (NAC) is the standard of care. NAC involves the use of chemotherapy prior to surgery to reduce tumor size, downgrade tumors amenable to resection, and improve long-term clinical outcomes. The primary endpoint of NAC is a pathological complete response (pCR), defined as the absence of residual invasive disease (RD) in the breast and axilla.

pCR is an important predictor of disease-free survival and overall survival in patients with TNBC [7]. Only 30–40% of TNBC patients achieve pCR with conventional NAC; the rest (~70%) either do not respond or respond partially to NAC. Non-responders and partial responders can be spared treatment side effects and offered alternative treatment regimens (e.g., a combination of NAC and immunotherapy) to improve outcomes and decrease morbidity [8–10]. More recently, immunotherapy has shown success in TNBC management, and the FDA has approved pembrolizumab for use in combination with NAC in high-risk patients with early-stage TNBC [11].

Although the mechanisms underlying chemoresistance in TNBCs remain elusive, the marked inter- and intratumoral heterogeneity in TNBCs may contribute to variability in NAC response. Currently, there is a lack of multi-modal biomarkers that can stratify TNBC patients into NAC responders, partial or non-responders, hindering personalized approaches for TNBC management. Furthermore, there is limited information on the robustness and accuracy of current biomarkers, e.g., Ki67, pH3, tumor-infiltrating lymphocytes (TILs), and histological features in predicting NAC treatment response

individually or in combination. Traditional staining techniques provide limited information about the immune landscape (e.g., type of TILs). The low reproducibility and objectivity of traditional scoring methods also impair the clinical adoption of these markers. TNBCs are heterogeneous, and their tumor microenvironment (TME) represents a complex ecosystem of cellular components, such as tumor, stromal, and immune cells. Communication between TME components and their spatial relationships affect cancer progression, treatment response, and disease outcomes [12, 13]. Studies have shown that the histomorphological components of the TME, such as a tumor, microvessels (MVD), polyploid giant cancer cells (PGCCs), immune cells, and necrotic areas, can help predict NAC response in TNBC [14]. Advances in computing, imaging, and pathology have created new opportunities to explore the relationships between histology, molecular events, and clinical outcomes to help predict NAC response in patients with TNBC [15].

Manual histomorphological characterizations of hematoxylin and eosin (H&E)-stained tissues is time-consuming and prone to inter- and intra-observer variability and fails to capture the TME spatial architecture, limiting their clinical value. Machine learning (ML) can more accurately and efficiently characterize the TME [16–18]. ML outperforms humans in terms of accuracy and speed and can identify novel predictive features and spatial patterns beyond human recognition [16–18]. The aim of this study was to develop an ML-based model to effectively predict NAC response (pCR or RD) in TNBC patients using spatial histological features from whole slide images (WSIs) of H&E-stained biopsy tissue sections.

## Methods

### Study population

H&E-stained pre-NAC core needle biopsies from treatment-naïve patients with TNBC were acquired from the Decatur Hospital, Georgia, USA, and the University of Galway, Ireland. The Decatur cohort was used as a discovery cohort for model development, and the Galway cohort was used as a validation cohort for the developed model. Patient samples with little to no tissue area, staining issues, or plating artifacts were excluded from the analysis. After this screening process, the final sample sizes of the model development and validation cohorts for prediction analyses were 85 and 79, respectively [19].

Patient clinical information of the Emory and Galway cohorts is presented in Table 1.

**Tumor slide selection and annotation**

H&E-stained slides were scanned by a slide scanner (Hamamatsu NanoZoomer 2.0-HT C9600-13) at 40× magnification (0.23 μm/pixel). All WSIs were reviewed and annotated by board-certified pathologists at maximum resolution with an open-source image processing software (QuPath, ver. 0.1.2) (Fig. 1A). A total of 16 histology labels were annotated, including tumor, stroma, adipocytes, PGCC, normal tissue, stromaTIL (sTIL), blood vessels, benign tumor, MVD, tertiaryTIL (teTIL), tumorTIL (tTIL), in situ carcinoma, hemorrhage, necrosis, apocrine change, and mucinous change. teTILs are TILs close to tertiary lymphoid structures [20]. Each WSI background was labeled separately. Coordinates and

histology class labels of tissue region contours were saved and preprocessed (Figs. 1A, 2B).

**Tile-level preprocessing**

Before model training, representative histology regions in WSIs were annotated with contours for each histology class. For each contour, a bounding box was created within the ground-truth area to extract the annotated tissue region. A sliding window of size 224×224 pixels was used to partition each WSI into image tiles. Only tiles overlapping the annotated areas by at least 90% were retained (Fig. 1C). The spatial containment query was invoked to identify the histology class for each tile. All image tiles were normalized by the stain color prior to model development [21]. Additionally, the image channel associated with the hematoxylin stain was separated from each color image tile by color deconvolution (Fig. 1C) [22].

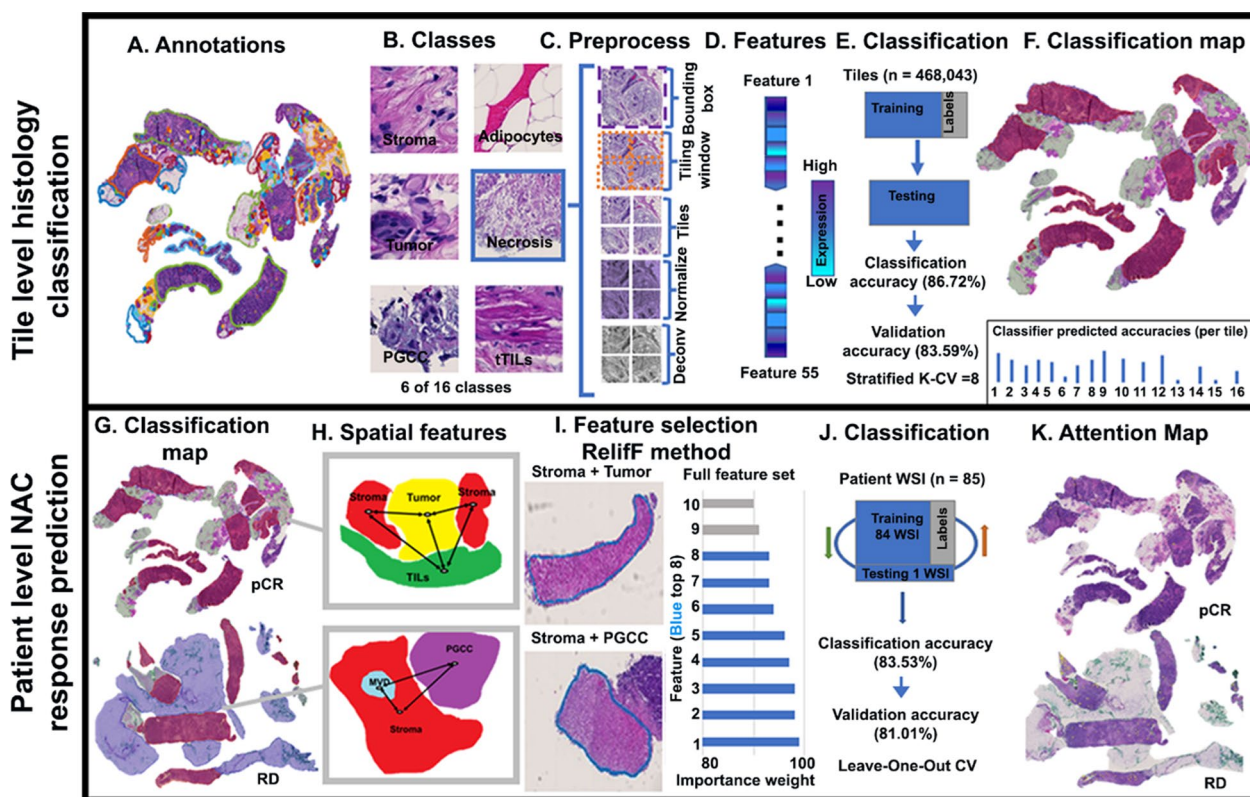
**Tile-level histology feature extraction and classification**

Tile-level histology features were derived from tissue tiles of different histology classes. After the preprocessing step, 468,043 tiles were produced from WSIs of model development cohort. The tile histology classification performance was evaluated by a stratified eightfold cross-validation strategy [23–25]. Additionally, an independent validation set was established with all tiles from the validation cohort. We extracted 80 tile-level features from each tile by six image texture extraction methods: gray-level co-occurrence matrix (GLCM; Method S1A), Gabor filter (Method S1B), local binary patterns (LBP; Method S1C), Tamura (Method S1D), lower-order histogram (Method S1E), and higher-order histogram (Method S1E) [26–34]. These 80 features were further reduced by excluding features that had 0, not applicable (NA), or repetitive values. We reduced the total number of texture features to 55 by exclusion criteria (Additional file 1: Table S1). Four ML classifiers were used to classify image tiles by the resulting feature set, including 1-nearest neighbor (1NN), linear support vector machine (linSVM), radial basis function SVM (rbfSVM), and ensemble tree (ensembleTree) with the RUSBoost method [35–37]. For model development and validation, the eightfold cross-validation mechanism was used. Each time, seven folds of data were used for training while the remaining one fold was used as the testing set (Fig. 1E). The stratified eightfold cross-validation method ensured that each data fold contained representative samples from each class and reliably assessed the tile histology class prediction performance. For each testing image tile, a trained classifier produced 16 probability values, one for each histology class of interest. Using the predicted tile histology class labels and tile spatial coordinates, we

**Table 1** Clinical information of the Emory Hospital and Galway cohorts

	Emory hospital dataset (n = 85)	Galway dataset (n = 79)
<i>Age</i>		
< 55 years old	37	47
= > 55 years old	48	31
NA	0	1
<i>Grade</i>		
1	3	0
2	18	21
3	66	58
<i>Tubule formation (T)</i>		
1	1	0
2	3	5
3	52	74
NA	29	0
<i>Nuclear pleomorphism (P)</i>		
1	2	0
2	7	4
3	47	75
NA	29	0
<i>Mitotic count (M)</i>		
1	6	21
2	16	28
3	34	30
NA	29	0
<i>Vital status</i>		
Alive	53	69
Dead	32	9
NA	0	1

NA Non-available



**Fig. 1** Overall schema of the developed NAC response prediction pipeline. The tile-level histology classification module (first step) consists of **A** training WSI annotation; **B** definition of histology classes of interest; **C** tile preprocessing; **D** feature extraction and selection; **E** classifier training, testing, and validation; and **F** generation of histology classification map. The patient-level NAC response prediction module (second step) consists of **G** graph node identification; **H** TME spatial descriptor computation; **I** graph construction and graph feature selection; **J** machine learning model training, testing, and validation; and **K** generation of an attention map with highlighted tissue regions with full feature set. Abbreviations: FE, feature extraction; sTILs, stromal TILs; tTILs, tumor TILs; Feat, feature; MVD, microvessel; PGCC, polypliod giant cancer cell

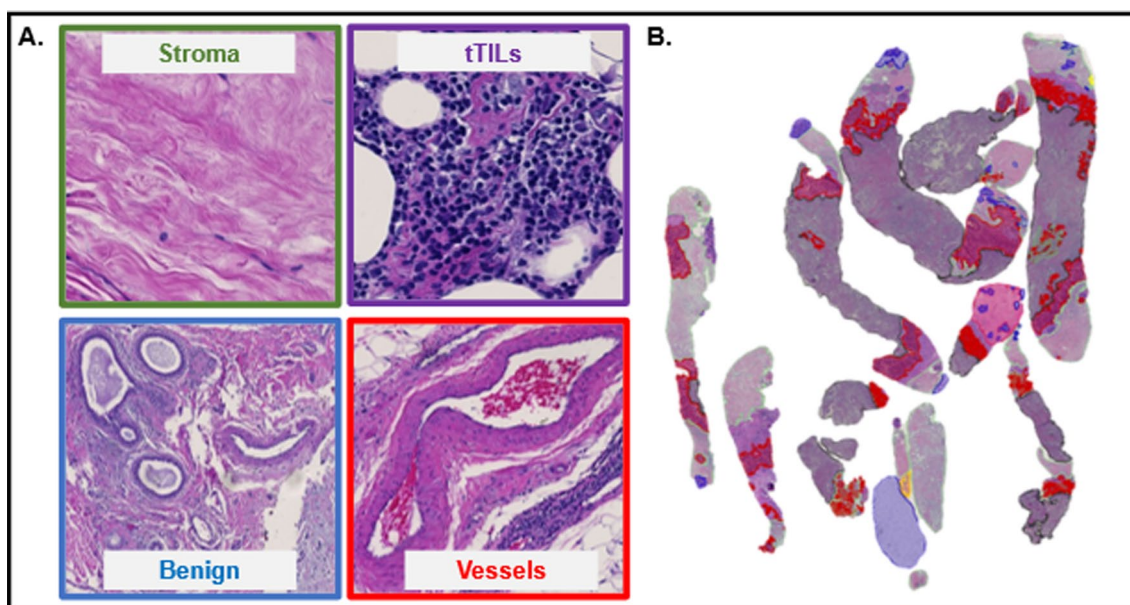
assembled tile-level histology class labels spatially and produced tile-level histology classification maps for each patient (Fig. 2). Each histology class was represented by a unique color in the classification maps, enabling the visualization of TME histology component distributions within a tissue context (Fig. 1F).

### Spatial TME feature extraction and NAC response prediction

The leave-one-out cross-validation (LOO-CV) method was used to evaluate the patient-level NAC response prediction performance [38, 39]. Similar to the tile-level histology class classification step, the model development and validation cohorts were used for model development and validation, respectively. As tumor cells interact closely with immune cells, stroma, PGCCs, and adipocytes in the TME [40–45], histology classification maps were generated for each patient. To better model the spatial relationships of these histology components in biological systems [46–53], we created TME graphs to characterize tissue TME states

and the spatial interactions of tissue regions of paired histology classes. For each pair of histology classes (e.g., tumor and PGCC maps), a TME graph was constructed from the corresponding tile-level histology maps. A simple graph  $G = (V, E)$  is undirected and unweighted, with  $V$  as the graph node set and  $E$  as the graph edge set [46, 49]. Each tile cluster was determined as a spatially connected tile component, with all connected tiles sharing the same histology class label. The centroids of the resulting tile clusters were used as nodes in a graph [47]. A graph edge between a pair of nodes  $u$  and  $v$ , i.e.,  $edge(u, v)$  was established using the Euclidean distance for a given pair of node histology classes. In this way, the spatial histology class distribution was represented by a TME graph structure [46]. Next, a set of TME features was extracted from each TME graph. In total, 20 graph features related to texture feature averages, local node configuration, and global graph connectivity were produced for each patient (Additional file 1: Table S2) [47]. Specifically, the texture feature averages were derived from the





**Fig. 2** Representative image tiles of distinct histological classes and a tile-level histology classification map. **A** Typical examples of image tiles (224×224) capturing stroma, tTILs, benign tissue, and vessels are presented. **B** A histology classification map is presented to visualize the spatial organization of TME components related to stroma (green), benign tumor (blue), tTILs (purple), and blood vessels (red)

tile-level feature set (Method S2A). Local node configuration features were used to characterize the local neighborhood information (Method S2B), while graph connectivity features represented the global graph structures (Method S2C).

For an optimal prediction performance, we retained the top eight TME graph features using the importance weight method. The importance weight of graph features was determined using the ReliefF algorithm, a filter-method approach designed to solve classification problems with discrete or numerical features [54–57]. A feature has a lower importance weight if a feature value difference is observed in a neighboring instance pair with the same histology class (i.e., a ‘hit’). By contrast, a feature presents a higher importance weight if a feature value difference is observed in a neighboring instance pair with different histology classes (i.e., a ‘miss’). The top eight TME graph features ranked by importance weights were retained for patient-level NAC response prediction [58–66].

Four ML classifiers, including 1NN, linSVM, rbf-SVM, and ensemble tree with the RUSBoost method, were used for patient-level NAC response prediction. Each patient was represented by the top eight TME graph-derived features associated with histology class pairs. Each classifier produced two NAC response class probability values, one for pCR and the other for RD. A patient was predicted to belong to the NAC response class associated with a larger class probability.

#### Statistical analysis

Statistical analysis was performed by Python (Python Software Foundation, <https://www.python.org/>), MATLAB 2020a (Natick, MA, USA), and R (R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org/>). The ReliefF importance weights were used to assess the significance of the selected TME graph features [54–56, 67, 68]. The resulting prediction performance was represented by a confusion matrix. The NAC response class pCR and RD were considered positive and negative groups, respectively. A false-positive (FP) was a RD case incorrectly predicted as pCR, while a false-negative (FN) case was a pCR case incorrectly predicted as RD. Multiple evaluation metrics were computed, including accuracy, sensitivity (i.e. Recall), specificity, precision, and F1-Score. The tile-level histology classifier was evaluated by the stratified k-fold cross-validation ( $k=8$ ), while the patient-level NAC response prediction was assessed by the LOO-CV.

#### Results

##### ML classifier provides accurate tile-level histology classification in H&E-stained WSIs

Our results suggest that the 16 histology classes of interest were well differentiated by 55 tile texture features for the model development cohort (Additional file 1: Table S1). With the stratified eightfold cross-validation strategy, we used onefold of image tiles to train the classifier and test it with the remaining seven folds in each

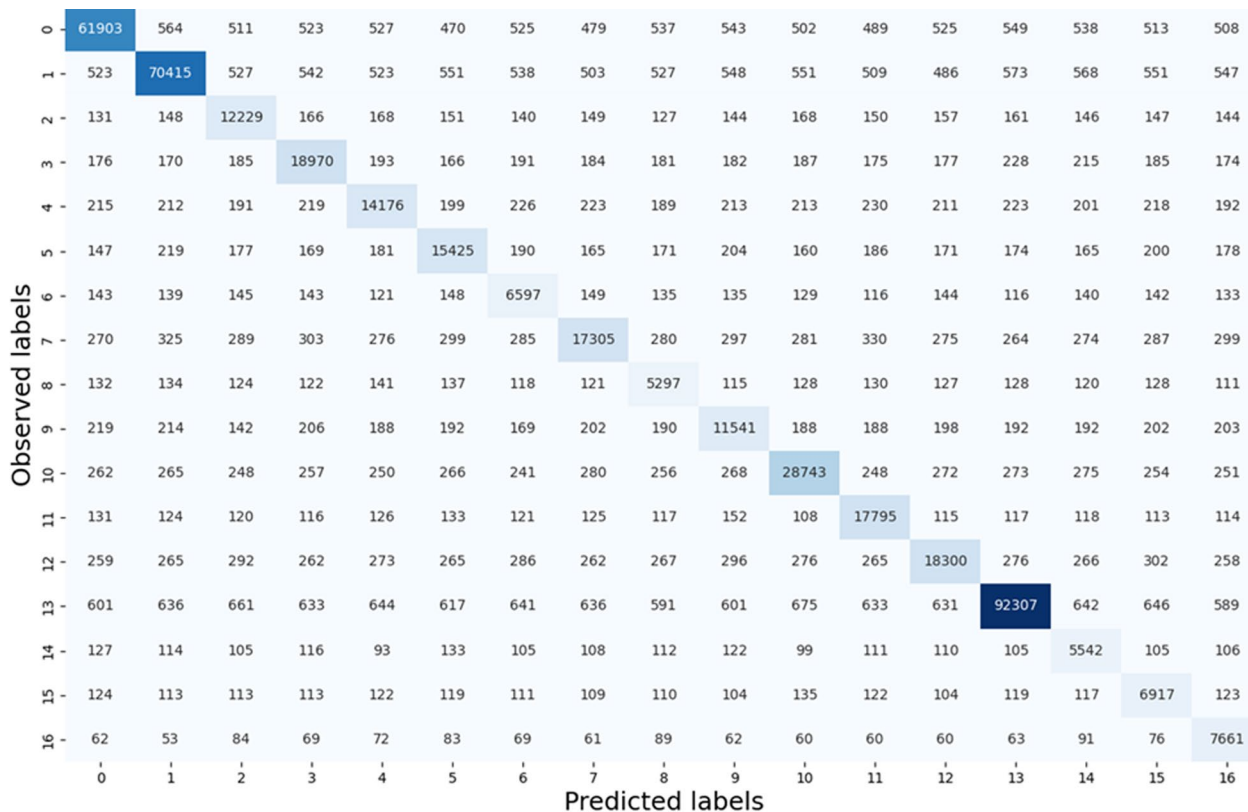
round. The average training accuracies of tile-level histology classification by 1NN, linSVM, and ensemble tree were 63.86%, 61.33%, and 80.08%, respectively. Of the four classifiers that we trained and tested, the best performance was achieved by the rbfSVM with an average training and testing accuracy of 87.16% and 86.72%, respectively (Fig. 3). The individual histology class average testing accuracy ranged from 72.51 to 91.18%. Additionally, the classifier reached an average recall from 75.11 to 92.97%, an average precision from 62.47 to 91.28%, and an average F1-score from 70.21 to 92.81% for all histology classes on the testing dataset (Fig. 4). It was noticed that the rbfSVM classifier was good at recognizing classes such as stroma, tumor and adipocytes, but weak at recognizing apocrine or mucinous change. Detailed classification results with Emory cohort are provided in Additional file 1: Figures S5–S7.

When the best tile-level histology classifier (rbf-SVM), trained using the model development cohort, was applied to the validation cohort, an average validation accuracy of 83.59% was achieved (Additional file 1: Fig. S1). The

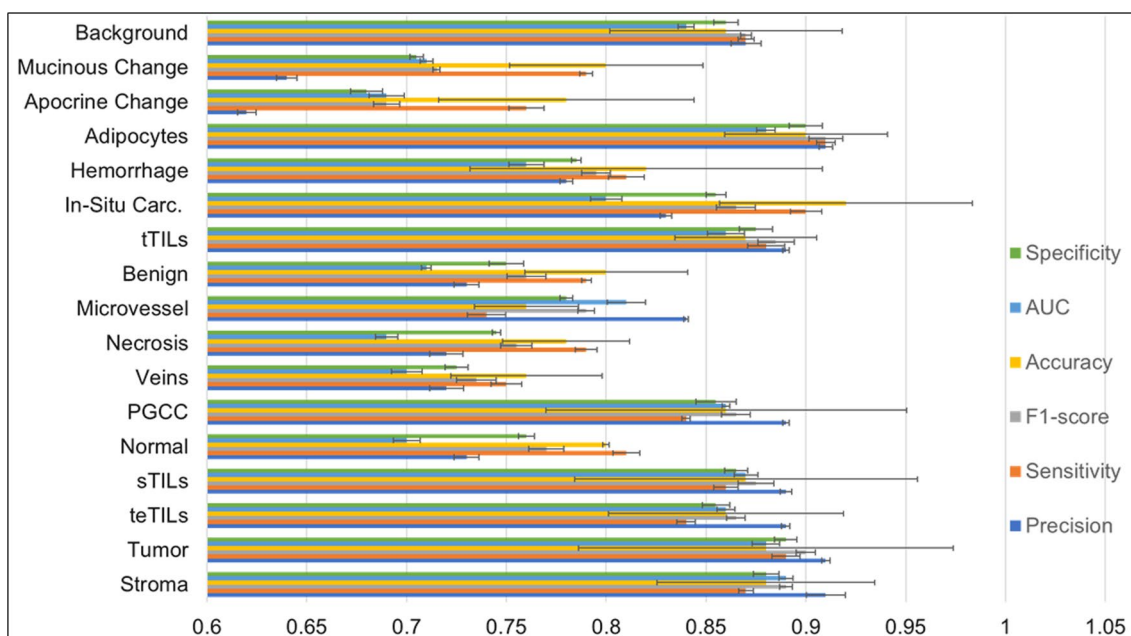
validation accuracy for individual histology classes in the validation cohort ranged from 69.74 to 87.88%. Additionally, the classifier reached an average recall ranging from 72.0 to 90.32%, an average precision from 81.43 to 84.19%, and an average F1-score from 76.11 to 86.92% for all histology classes (Additional file 1: Fig. S2). The true histology class, the predicted histology class, and the image tile spatial coordinates with respect to each WSI were recorded for each tile. Histology classification maps were generated for visualizing spatial distribution of histological classes. Each histology class was assigned a unique color, and the predicted histology class results were spatially assembled by the image tile spatial order (Additional file 1: Fig. S3). Detailed classification results with Galway cohort can be found in Additional file 1: Figures S8–S10.

**Spatial TME features of paired histology classes predict NAC response**

Using the importance weights from the ReliefF algorithm, we ranked all TME features from paired histology classes.



**Fig. 3** Testing tile-level histology classification performance in the model development cohort. Confusion matrix showing the aggregated performance of the rbfSVM model for tile-level histology class prediction (i.e., 0, stroma; 1, tumor; 2, tertiary TILs; 3, stroma TILs; 4, normal tissue; 5, PGCCs; 6, blood vessels; 7, necrosis; 8, microvessel; 9, benign tumor; 10, tumor TILs; 11, in situ carcinoma; 12, hemorrhage; 13, adipocytes; 14, apocrine change; 15 mucinous change; and 16, background). Abbreviations: 1NN, 1-nearest neighbor; linSVM, linear support vector machine SVM; PGCC: polypliod giant cancer cells; rbfSVM, radial basis function SVM



**Fig. 4** Testing tile-level histology classification performance of the rbfSVM classifier with the Emory Hospital development cohort. Each bar represents the weighted average of tiles and their predicted probabilities during testing for a histology class. Specificity (green) measures how often the rbfSVM classifier correctly predicted true negatives. AUC (blue) reflects the model’s ability to distinguish between positive and negative classes. Accuracy (yellow) indicates the proportion of correct predictions out of total predictions. F1-score (gray) presents a balanced view of rbfSVM classifier performance. Sensitivity (orange) suggests how often the rbfSVM classifier correctly identifies positive instances. Precision (blue) indicates how often the rbfSVM classifier correctly predicts true positives. Error bars represent the 95% confidence interval in each case

For each histology class pair, we computed 20 TME features that were related to texture feature averages using a random geometric construct, local node-based features using a spectral construct, and global graph-based features using a minimum spanning tree construct (Additional file 1: Table S2). The top eight histology pairs are presented in Fig. 5. For example, the histological class pair of tumors and tTILs with the largest importance weight is known to be strongly associated with pCR. In contrast, the histological class pair of microvessel density and PGCCs strongly correlate with RD. These results are in line with previously published studies [58–66].

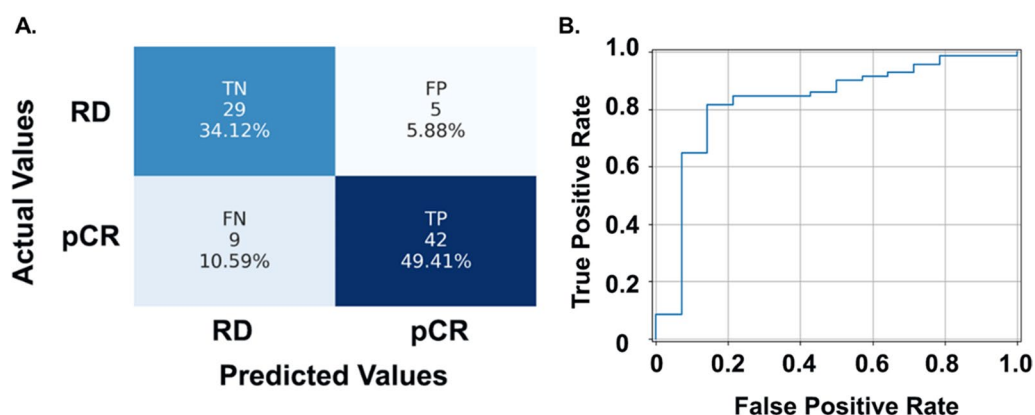
With each TNBC patient represented by selected TME features, we trained and tested the classifier for NAC response prediction using LOO-CV strategy. The prediction accuracies by 1NN, linSVM, and ensemble tree were 72.90%, 62.43%, and 70.61%, respectively. Similar to the tile-level histology classification, rbfSVM achieved the best NAC response prediction at the patient level with a prediction accuracy of 83.53%. Out of 51 cases, 42 were correctly predicted as pCR. Twenty-nine out of 34 patients were correctly predicted as RD (Fig. 6A). The FP and FN groups included five and nine misclassified patients, respectively, resulting in a specificity of 85.29% and sensitivity of 82.35% (Fig. 6A). Additionally, the receiver operating characteristic (ROC) curve of the

	Average	Importance Rank	Feature Measure	
Tumor+ tTILs	97	1	Vertices	pCR
Stroma+ sTILs	96	2	Vertices	
tTILs+ PGCCs	92	3	Node Degrees	
Normal+ Tumor	92	4	Shortest Path to Highest TF	
Adipocytes+ Microvessel	87	5	Histogram of Gradient Mean	RD
Tumor+ PGCCs	86	6	Eigenvector Centrality	
Adipocytes+ sTILs	85	7	Nearest Neighbor	
Microvessel + PGCCs	83	8	GLCM Mean	

**Fig. 5** TME graph feature selection by ReliefF. TME graph features histology class pairs sorted by their importance weights from the ReliefF algorithm in the model development cohort. Abbreviations: TF, Texture Feature; sTILs, stromal TILs; tTILs, tumor TILs; PGCCs, polyploid giant cancer cells; GLCM, gray-level co-occurrence matrix

best-performing NAC prediction pipeline is presented in Fig. 6B with area under the curve (AUC) reaching 0.83. Detailed prediction performance with the Emory cohort is presented in Table 2.

The top eight graph-derived features capture histological information from the TME that is critical for NAC



**Fig. 6** Patient-level NAC response prediction performance of the best classifier by LOO-CV in the model development cohort. **A** Confusion matrix showing performance of the rbfSVM model. **B** ROC curve of the best NAC response prediction pipeline. Abbreviations: LOO-CV, leave-one-out cross-validation; rbfSVM, radial basis function SVM; ROC, receiver operating characteristic; TN, true negative; FP, false positive; FN, false negative

**Table 2** NAC prediction performance with Emory Hospital cohort

	Accuracy	AUC	Sensitivity	Specificity	Precision	F1
1NN	0.729 [0.703, 0.756]	0.735 [0.690, 0.779]	0.745 [0.693, 0.797]	0.706 [0.676, 0.736]	0.792 [0.755, 0.828]	0.768 [0.747, 0.788]
Linear SVM	0.624 [0.586, 0.660]	0.709 [0.654, 0.764]	0.490 [0.425, 0.555]	0.824 [0.801, 0.846]	0.806 [0.774, 0.839]	0.610 [0.581, 0.638]
RBF SVM	0.835 [0.808, 0.862]	0.827 [0.786, 0.868]	0.824 [0.768, 0.879]	0.853 [0.818, 0.887]	0.894 [0.849, 0.938]	0.857 [0.823, 0.892]
EnsembleTree RUSBoost	0.706 [0.686, 0.725]	0.681 [0.648, 0.714]	0.784 [0.735, 0.833]	0.588 [0.566, 0.611]	0.741 [0.705, 0.777]	0.762 [0.740, 0.784]

The average and 95% confidence interval of accuracy, AUC, sensitivity, specificity, precision, and F1-Score are presented for each model by a leave-one-out cross-validation strategy

response prediction. Additional file 1: Figure S4 highlights representative tissue regions from where image tiles of the associated paired histology classes were derived for TME graph nodes. The resulting attention maps are tissue areas with high discriminating values for NAC response prediction, i.e., pCR vs RD (Additional file 1: Fig. S4).

**Validating the NAC response prediction performance in an independent cohort**

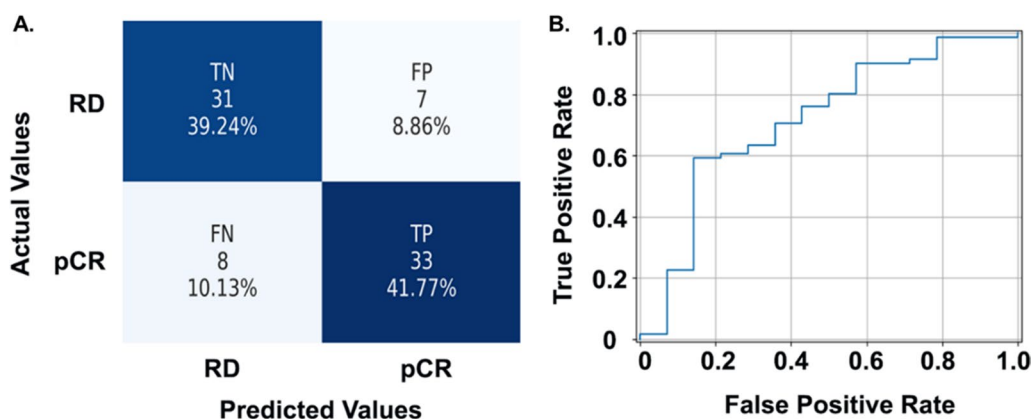
We validated the prediction performance of the pipeline in an independent cohort consisting of 79 WSIs. Using the same image quality checks and preprocessing steps, we analyzed each WSI using the previously trained classifiers. The positive and negative group included 41 and 38 patients, respectively. Our prediction method correctly predicted 33 and 31 patients from the positive and negative groups, respectively, with a prediction accuracy of 81.01%. The FP and FN groups included seven and eight misclassified patients, respectively, resulting in a specificity of 81.58% and sensitivity of 80.49% (Fig. 7A). Additionally, the ROC curve of the best NAC prediction

pipeline is presented in Fig. 7B, with AUC reaching 0.83. Detailed prediction performance with the Galway cohort is presented in Table 3.

**Discussion**

Women with TNBC exhibit significantly worse 5-year survival rates than those with non-TNBC, regardless of the tumor stage at diagnosis [69]. No targeted or endocrine therapy is available for TNBC, and NAC is the cornerstone of treatment. However, only 30–40% of TNBC patients achieve pCR with NAC, and there is a dire need for early identification of the nearly 70% of patients who should be offered alternative regimens to improve treatment outcomes. In this study, we used ML approaches to predict NAC response and stratify patients into NAC responders and non-responders based on H&E-stained WSIs of tissue biopsies. We developed a two-step prediction model: in the first step, the histology class of each H&E image tile was determined using a tile-level classification pipeline; in the second step, the spatial graph-derived features associated with histology class pairs were used to predict patient-level NAC response (pCR





**Fig. 7** The NAC prediction model's performance in the validation cohort. **A** Confusion matrix showing performance of the best NAC response prediction model. **B** ROC curve of the best NAC response prediction pipeline. Abbreviations: LOO-CV, leave-one-out cross-validation; rbfSVM, radial basis function SVM; ROC, receiver operating characteristic; TN, true negative; FP, false positive; FN, false negative; TP, true positive

**Table 3** NAC prediction performance with Galway cohort

	Accuracy	AUC	Sensitivity	Specificity	Precision	F1
1NN	0.684 [0.645, 0.717]	0.671 [0.524, 0.824]	0.585 [0.522, 0.648]	0.789 [0.765, 0.814]	0.75 [0.712, 0.788]	0.653 [0.619, 0.696]
Linear SVM	0.608 [0.572, 0.643]	0.617 [0.574, 0.660]	0.488 [0.428, 0.548]	0.737 [0.707, 0.766]	0.667 [0.643, 0.690]	0.563 [0.542, 0.585]
RBF SVM	0.810 [0.783, 0.837]	0.832 [0.792, 0.873]	0.805 [0.749, 0.860]	0.816 [0.779, 0.853]	0.825 [0.787, 0.863]	0.815 [0.786, 0.844]
EnsembleTree RUSBoost	0.671 [0.643, 0.699]	0.691 [0.654, 0.727]	0.683 [0.625, 0.741]	0.658 [0.622, 0.694]	0.683 [0.645, 0.721]	0.683 [0.652, 0.714]

The average and 95% confidence interval of accuracy, AUC, sensitivity, specificity, precision, and F1-Score are presented for each model by a leave-one-out cross-validation strategy

vs RD). Our model unveils and leverages novel NAC response predictive features and spatial patterns of TME histology components from WSIs of TNBC tissue biopsies. This study also highlights the role of various TME components in accurately predicting NAC response.

TME components and their interactions can influence NAC response in patients with TNBC [70–72]. Traditional methods using human annotations are unable to capture these spatial relationships. In contrast, our approach incorporates the spatial relationships of various TME components to predict NAC response in patients with TNBC. Using a graph structure for spatial TME characterization, we identified eight histology component pairs that accurately predicted NAC response. We expect that an investigation with higher-order combinations (e.g., tertiary and quaternary) can further increase NAC response prediction accuracy. The top three TME features captured the spatial interactions between (1) tumor cells and tTILs, (2) stroma and sTILs, and (3) tTILs and PGCCs. Studies have shown the predictive importance of tumor area, immune activation markers, and TILs in TNBC biopsies [73–76].

Our results provide further evidence that the interrelationships between TILs, stroma, adipocytes, and tumor cells can predict NAC response in patients with TNBC. Other recently published studies that have relied on WSI models [77, 78] include one that used a federated learning model to predict NAC response in TNBC, and found hemorrhage, TILs, and necrosis as predictive of pCR and apocrine change, fibrosis, and noncohesive tumor cells being predictive of RD [77]. Another study quantified the stromal and tumor features in a WSI-based multi-omic (WSI, clinical, pathological) ML model and found that high collagenous stroma was best associated with lower pCR rates [78]. Our study used expert annotations that effectively guided the ML models to identify specific histological patterns in spatial TME contexts. While our supervised ML model identified the common histological component of TILs, it did not rank hemorrhage, necrosis, fibrosis, and apocrine change as important predictors due to the lack of annotated training data.

Our NAC response prediction pipeline provides classification accuracy and attention maps that can be highly useful in clinical practice. Attention maps help

pathologists and researchers by identifying tissue regions in a WSI that are highly predictive of NAC response, thereby improving slide review, reducing visual fatigue, and facilitating image data interpretation. Information from attention maps can be combined with other WSI-derived data such as, Ki67 and pH3 immunohistochemistry-stained serial tissue sections, to train deep-learning models for enhanced prediction [79–81]. Ki67 and pH3 are clinical biomarkers with demonstrated NAC response predictive value in TNBC tumors [82, 83]. Furthermore, our predictive model is promising for integrating data from various sources, such as electronic health records, laboratory test results, and demographic information, to provide predictions based on the overall view of the health status of patients.

Limitations of the study include small sample sizes, slide quality issues, and expensive computational processes. Quality checks are necessary to ensure inclusion of adequate samples to develop effective training classifiers. The different slide staining protocols, artifacts, and plating variances from different institutions (e.g., cutting glass slide edges) may have resulted in inconsistencies in slide quality. Thus, although we had a larger number of WSIs to begin with, the final validation cohort was whittled down. Because the sample size was small, there was an imbalance of histology classes presented among different patient slides. More histology classes (e.g., microcalcification, muscle) should be included to improve the training of the tile-level histology classifier in all histology classes. We had two pathologists independently annotate the WSIs; however, more experts can be included in the future to validate the annotations and reduce interobserver variability. Additionally, our pipeline is computationally expensive because multiple processes occur throughout the pipeline such as partitioning gigapixel WSIs, calculating various feature measures for each tile, constructing graphs based on spatial relationships. Computational constraints can stem from institutional high-performance computing (HPC) server data loss, standard maintenance, and outages. Refining the code for faster processing times (parallel processing) based on an advanced computer architecture could help support ML processes and data management. We cannot identify important spatially related histological features using image viewing software alone because the software is not scalable for large datasets. Each digital pathology software is limited in the amount of data processed through its graphical user interfaces before exceeding the computational capabilities.

Future work will include model validation in a larger cohort. Future work will also include the development of prediction models with higher-order feature combinations and graph convolution networks. It is

important to develop an efficient pipeline to increase the amount of image data and decrease the computational time. Additionally, combinations of features with the highest predictive value will be used to increase the predictive power of the full feature set. For example, attention map regions can be leveraged to focus on regions of interest, which can be used for more complex analyses, such as imaging mass cytometry, to distinguish between the various TIL subtypes and to further refine NAC response prediction. We also plan to extend our pipeline to incorporate other tissue stains, including immunohistochemistry. A more efficient pipeline can reduce the frequency of false negatives and thus minimize the risk of undertreating patients, which can result in early relapse and poor outcomes.

## Conclusions

Using feature engineering and supervised ML, we demonstrated the strong discriminating power of TME histological components and their spatial relationships in predicting NAC response in patients with TNBC. Among 120 histology feature pairs, we identified eight with the highest predictive value. The most predictive histology feature pair for pCR was tumor and tTILs, whereas microvessel density and PGCCs was the feature pair most strongly correlated with RD. The proposed ML pipeline can help identify tissue areas in H&E-stained WSIs with a high predictive value for NAC response prediction and can help in clinical decision-making.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13058-023-01752-y>.

**Additional file 1:** The supplementary files include Figures S1–S10 of results, Tables S1 and S2 of the feature list, Methods S1 and S2 of feature extraction, and the supplementary file's bibliography.

## Acknowledgements

Editorial assistance was provided by Christos Evangelou, Ph.D.

## Author contributions

TF, JKo, and RA did conceptualization. TF, RTS, JK, SB, GC, MW, EJ done data procurement and experimentation. TF was involved in data analysis. TF, GS, JKo, RA wrote and edited the manuscript. TF, GS, EJ, JKo, RA were involved in discussion, editing and proof-reading the manuscript.

## Funding

This study was supported by Grants from the National Institutes of Health to RA (R01CA239120) and JKo (1U01CA242936). This research was also supported by Georgia State University Molecular Basis of Disease Doctoral Fellowship and Janssen Scholars of Oncology Diversity Engagement Program awarded to TF.

**Availability of data and materials**

The data underlying this article will be shared on request to the corresponding author.

**Declarations****Ethics approval and consent to participate**

All aspects of this study were approved by the Institutional Review Boards of the institutions involved. Patient consent was not required because all samples were archival.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare no conflicts of interest or financial interest.

**Author details**

<sup>1</sup>Department of Biology, Georgia State University, Atlanta, GA 30302, USA. <sup>2</sup>School of Health Professions, University of Alabama at Birmingham, Birmingham, AL 35294, USA. <sup>3</sup>JSSAHER (JSS Academy of Higher Education and Research) Medical College, Mysuru, Karnataka, India. <sup>4</sup>Discipline of Pathology, University of Galway, Galway, Ireland. <sup>5</sup>Department of Pathology, Stavanger University Hospital, Stavanger, Norway. <sup>6</sup>Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303, USA. <sup>7</sup>Department of Chemistry, Bioscience and Environmental Engineering, University of Stavanger, Stavanger, Norway.

Received: 7 August 2023 Accepted: 11 December 2023

Published online: 18 January 2024

**References**

- Craig DW, O'Shaughnessy JA, Kiefer JA, Aldrich J, Sinari S, Moses TM, et al. Genome and transcriptome sequencing in prospective metastatic triple-negative breast cancer uncovers therapeutic vulnerabilities. *Mol Cancer Ther.* 2013;12(1):104–16.
- Howlader NNA, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds. SEER Cancer Statistics Review. 1975–2017).
- Criscitiello C, Azim HA Jr, Schouten PC, Linn SC, Sotiriou C. Understanding the biology of triple-negative breast cancer. *Ann Oncol.* 2012;23 Suppl(6):vi13–8.
- Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res.* 2007;13(15 Pt 1):4429–34.
- Liedtke C, Mazouni C, Hess KR, André F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol.* 2008;26(8):1275–81.
- Niikura N, Hayashi N, Masuda N, Takashima S, Nakamura R, Watanabe K, et al. Treatment outcomes and prognostic factors for patients with brain metastases from breast cancer of each subtype: a multicenter retrospective analysis. *Breast Cancer Res Treat.* 2014;147(1):103–12.
- Gass P, Lux MP, Rauh C, Hein A, Bani MR, Fiessler C, et al. Prediction of pathological complete response and prognosis in patients with neoadjuvant treatment for triple-negative breast cancer. *BMC Cancer.* 2018;18(1):1051.
- Biswas T, Efrid JT, Prasad S, Jindal C, Walker PR. The survival benefit of neoadjuvant chemotherapy and pCR among patients with advanced stage triple negative breast cancer. *Oncotarget.* 2017;8(68):112712.
- Chen VE, Gillespie EF, Zakeri K, Murphy JD, Yashar CM, Lu S, et al. Pathologic response after neoadjuvant chemotherapy predicts locoregional control in patients with triple negative breast cancer. *Adv Radiat Oncol.* 2017;2(2):105–9.
- Gamucci T, Pizzuti L, Sperduti I, Mentuccia L, Vaccaro A, Moscetti L, et al. Neoadjuvant chemotherapy in triple-negative breast cancer: a multicentric retrospective observational study in real-life setting. *J Cell Physiol.* 2018;233(3):2313–23.
- Valencia GA, Rioja P, Morante Z, Ruiz R, Fuentes H, Castaneda CA, et al. Immunotherapy in triple-negative breast cancer: a literature review and new advances. *World J Clin Oncol.* 2022;13(3):219–36.
- Baghban R, Roshangar L, Jahanban-Esfahlan R, Seidi K, Ebrahimi-Kalan A, Jaymand M, et al. Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Commun Signal.* 2020;18(1):59.
- Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med.* 2013;19(11):1423–37.
- Bilous M. Breast core needle biopsy: issues and controversies. *Mod Pathol.* 2010;23(2):S36–45.
- Cooper LAD, Kong J, Gutman DA, Dunn WD, Nalisnik M, Brat DJ. Novel genotype-phenotype associations in human cancers enabled by advanced molecular platforms and computational analysis of whole slide images. *Lab Invest.* 2015;95(4):366–76.
- van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nat Med.* 2021;27(5):775–84.
- Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial intelligence and machine learning in pathology: the present landscape of supervised methods. *Academic Pathology.* 2019;6:2374289519873088.
- Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: challenges and opportunities. *Med Image Anal.* 2016;33:170–5.
- Cindy Sampias GR. H&E Staining Overview: A Guide to Best Practices [Webpage]. Leica Biosystems2023. <https://www.leicabiosystems.com/us/knowledge-pathway/he-staining-overview-a-guide-to-best-practices/>.
- Kang BW, Seo AN, Yoon S, Bae HI, Jeon SW, Kwon OK, et al. Prognostic value of tumor-infiltrating lymphocytes in Epstein-Barr virus-associated gastric cancer. *Ann Oncol.* 2016;27(3):494–501.
- Reinhard E, Adhikhmin M, Gooch B, Shirley P. Color transfer between images. *IEEE Comput Graphics Appl.* 2001;21(5):34–41.
- Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol.* 2001;23(4):291–9.
- Zeng X, Martinez TR. Distribution-balanced stratified cross-validation for accuracy estimation. *J Exp Theor Artif Intell.* 2000;12(1):1–12.
- Prusty S, Patnaik S, Dash SK. SKCV: stratified K-fold cross-validation on ML classifiers for predicting cervical cancer. *Front Nanotechnol.* 2022;4:972421.
- Szeghalmy S, Fazekas A. A comparative study of the use of stratified cross-validation and distribution-balanced stratified cross-validation in imbalanced learning. *Sensors.* 2023;23(4):2333.
- Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern.* 1973;SMC-3(6):610–21.
- Jain AK, Farrokhnia F. Unsupervised texture segmentation using Gabor filters. *Pattern Recognit.* 1991;24(12):1167–86.
- Wang X, Ding X, Liu C. Gabor filters-based feature extraction for character recognition. *Pattern Recognit.* 2005;38(3):369–79.
- Haghighat M. Gabor Feature Extraction GitHub2022 [February 8, 2022]. <https://github.com/mhaghighat/gabor>.
- Nosaka R, Ohkawa Y, Fukui K. Feature extraction based on co-occurrence of adjacent local binary patterns. In: Proceedings of the 5th Pacific Rim conference on Advances in Image and Video Technology—Volume Part II; Gwangju, South Korea: Springer; 2011. p. 82–91.
- Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell.* 2002;24(7):971–87.
- Tamura H, Mori S, Yamawaki T. Textural features corresponding to visual perception. *IEEE Trans Syst Man Cybern.* 1978;8(6):460–73.
- Julesz B. Textons, the elements of texture perception, and their interactions. *Nature.* 1981;290(5802):91–7.
- Kather JN, Weis C-A, Bianconi F, Melchers SM, Schad LR, Gaiser T, et al. Multi-class texture analysis in colorectal cancer histology. *Sci Rep.* 2016;6(1):27988.
- Mounce SR, Ellis K, Edwards JM, Speight VL, Jakomis N, Boxall JB. Ensemble decision tree models using RUSBoost for estimating risk of iron failure in drinking water distribution systems. *Water Resour Manag.* 2017;31(5):1575–89.
- Seiffert C, Khoshgoftaar TM, Hulse JV, Napolitano A. RUSBoost: a hybrid approach to alleviating class imbalance. *IEEE Trans Syst Man Cybern Part A Syst Hum.* 2010;40(1):185–97.

37. Seiffert C, Khoshgoftaar TM, Hulse JV, Napolitano A, editors. RUSBoost: improving classification performance when training data is skewed. In: 2008 19th international conference on pattern recognition; 2008 8–11. 2008.
38. Varoquaux G, Colliot O. Evaluating machine learning models and their diagnostic value. In: Olivier C, editor. *Machine Learning for Brain Disorders*; Springer; 2023.
39. Cheng H, Garrick DJ, Fernando RL. Efficient strategies for leave-one-out cross validation for genomic best linear unbiased prediction. *J Anim Sci Biotechnol*. 2017;8(1):38.
40. Saini G, Joshi S, Garlapati C, Li H, Kong J, Krishnamurthy J, et al. Polyploid giant cancer cell characterization: new frontiers in predicting response to chemotherapy in breast cancer. *Seminars in Cancer Biology*. 2021.
41. Oshi M, Tokumar Y, Angarita FA, Lee L, Yan L, Matsuyama R, et al. Adipogenesis in triple-negative breast cancer is associated with unfavorable tumor immune microenvironment and with worse survival. *Sci Rep*. 2021;11(1):12541.
42. Yamaguchi J, Moriuchi H, Ueda T, Kawashita Y, Hazeyama T, Tateishi M, et al. Active behavior of triple-negative breast cancer with adipose tissue invasion: a single center and retrospective review. *BMC Cancer*. 2021;21(1):434.
43. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, et al. A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. *Nat Med*. 2009;15(1):68–74.
44. Mosieniak G, Sliwinka MA, Alster O, Strzeszewska A, Sunderland P, Piechota M, et al. Polyploidy formation in doxorubicin-treated cancer cells can favor escape from senescence. *Neoplasia*. 2015;17(12):882–93.
45. Mittal K, Donthamsetty S, Kaur R, Yang C, Gupta MV, Reid MD, et al. Multinucleated polyploidy drives resistance to Docetaxel chemotherapy in prostate cancer. *Br J Cancer*. 2017;116(9):1186–94.
46. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. *IEEE Rev Biomed Eng*. 2009;2:147–71.
47. Yener B. Cell-graphs: image-driven modeling of structure-function relationship. *Commun ACM*. 2016;60(1):74–84.
48. Chen JM, Li Y, Xu J, Gong L, Wang LW, Liu WL, et al. Computer-aided prognosis on breast cancer with hematoxylin and eosin histopathology images: a review. *Tumour Biol*. 2017;39(3):1010428317694550.
49. Acar E, Plopper GE, Yener B. Coupled analysis of in vitro and histology tissue samples to quantify structure-function relationship. *PLoS ONE*. 2012;7(3):e32227.
50. Bilgin CC, Ray S, Baydil B, Daley WP, Larsen M, Yener B. Multiscale feature analysis of salivary gland branching morphogenesis. *PLoS ONE*. 2012;7(3):e32906.
51. Li G, Semerci M, Yener B, Zaki MJ. Effective graph classification based on topological and label attributes. *Stat Anal Data Min ASA Data Sci J*. 2012;5(4):265–83.
52. Bilgin C, Demir C, Nagi C, Yener B, editors. Cell-graph mining for breast tissue modeling and classification. In: 2007 29th annual international conference of the IEEE engineering in medicine and biology society; 2007. pp. 22–26.
53. Bilgin CC, Bullough P, Plopper GE, Yener B. ECM-aware cell-graph mining for bone tissue modeling and classification. *Data Min Knowl Discov*. 2009;20(3):416–38.
54. Wang Z, Zhang Y, Chen Z, Yang H, Sun Y, Kang J, et al., editors. Application of ReliefF algorithm to selecting feature sets for classification of high resolution remote sensing image. In: 2016 IEEE international geoscience and remote sensing symposium (IGARSS); 2016. pp. 10–15.
55. Mathworks. Rank importance of predictors using ReliefF or RRelieff algorithm 2022. <https://www.mathworks.com/help/stats/relieff.html>
56. Robnik-Šikonja M, Kononenko I. Theoretical and empirical analysis of ReliefF and RRelieff. *Mach Learn*. 2003;53(1):23–69.
57. Urbanowicz RJ, Meeker M, La Cava W, Olson RS, Moore JH. Relief-based feature selection: Introduction and review. *J Biomed Inform*. 2018;85:189–203.
58. Rapoport BL, Nayler S, Mlecnik B, Smit T, Heyman L, Bouquet I, et al. Tumor-infiltrating lymphocytes (TILs) in early breast cancer patients: high CD3(+), CD8(+), and immunoscore are associated with a pathological complete response. *Cancers (Basel)*. 2022;14(10):2525.
59. Sasanpour P, Sandoughdaran S, Mosavi-Jarrahi A, Malekzadeh M. Predictors of pathological complete response to neoadjuvant chemotherapy in Iranian breast cancer patients. *Asian Pac J Cancer Prev*. 2018;19(9):2423–7.
60. Zgura A, Galesa L, Bratila E, Anghel R. Not available. *Maedica (Bucur)*. 2018;13(4):317–20.
61. Song IH, Heo SH, Bang WS, Park HS, Park IA, Kim YA, et al. Predictive value of tertiary lymphoid structures assessed by high endothelial venule counts in the neoadjuvant setting of triple-negative breast cancer. *Cancer Res Treat*. 2017;49(2):399–407.
62. Qian X-L, Xia X-Q, Li Y-Q, Jia Y-M, Sun Y-Y, Song Y-M, et al. Effects of tumor-infiltrating lymphocytes on nonresponse rate of neoadjuvant chemotherapy in patients with invasive breast cancer. *Sci Rep*. 2023;13(1):9256.
63. Mohammed RAA, Ellis IO, Mahmmod AM, Hawkes EC, Green AR, Rakha EA, et al. Lymphatic and blood vessels in basal and triple-negative breast cancers: characteristics and prognostic significance. *Mod Pathol*. 2011;24(6):774–85.
64. Krüger K, Silwal-Pandit L, Wik E, Straume O, Stefansson IM, Borgen E, et al. Baseline microvessel density predicts response to neoadjuvant bevacizumab treatment of locally advanced breast cancer. *Sci Rep*. 2021;11(1):3388.
65. White-Gilbertson S, Voelkel-Johnson C. Giants and monsters: unexpected characters in the story of cancer recurrence. *Adv Cancer Res*. 2020;148:201–32.
66. Sirois I, Aguilar-Mahecha A, Lafleur J, Fowler E, Vu V, Scriver M, et al. A unique morphological phenotype in chemoresistant triple-negative breast cancer reveals metabolic reprogramming and PLIN4 expression as a molecular vulnerability. *Mol Cancer Res*. 2019;17(12):2492–507.
67. Robnik-Šikonja M, Kononenko I. An adaptation of Relief for attribute estimation in regression. In: Proceedings of the fourteenth international conference on machine learning: Morgan Kaufmann Publishers Inc.; 1997. pp. 296–304.
68. Kononenko I, Šimec E, Robnik-Šikonja M. Overcoming the myopia of inductive learning algorithms with RELIEFF. *Appl Intell*. 1997;7(1):39–55.
69. Al-Ejeh F, Simpson PT, Sanus JM, Klein K, Kalimutho M, Shi W, et al. Meta-analysis of the global gene expression profile of triple-negative breast cancer identifies genes for the prognostication and treatment of aggressive breast cancer. *Oncogenesis*. 2014;3(4):e100-e.
70. Thagaard J, Stovgaard ES, Vognsen LG, Hauberg S, Dahl A, Ebstrup T, et al. Automated quantification of sTIL density with H&E-based digital image analysis has prognostic potential in triple-negative breast cancers. *Cancers (Basel)*. 2021;13(12):3050.
71. Hong J, Rui W, Fei X, Chen X, Shen K. Association of tumor-infiltrating lymphocytes before and after neoadjuvant chemotherapy with pathological complete response and prognosis in patients with breast cancer. *Cancer Med*. 2021;10(22):7921–33.
72. Lee HJ, Cho SY, Cho EY, Lim Y, Cho SI, Jung W, et al. Artificial intelligence (AI)-powered spatial analysis of tumor-infiltrating lymphocytes (TIL) for prediction of response to neoadjuvant chemotherapy (NAC) in triple-negative breast cancer (TNBC). *J Clin Oncol*. 2022;40(16\_suppl):595.
73. Holanek M, Selingerova I, Bilek O, Kazda T, Fabian P, Foretova L, et al. Neoadjuvant chemotherapy of triple-negative breast cancer: evaluation of early clinical response, pathological complete response rates, and addition of platinum salts benefit based on real-world evidence. *Cancers (Basel)*. 2021;13(7):1586.
74. Eguchi Y, Nakai T, Kojima M, Wakabayashi M, Sakamoto N, Sakashita S, et al. Pathologic method for extracting good prognosis group in triple-negative breast cancer after neoadjuvant chemotherapy. *Cancer Sci*. 2022;113(4):1507–18.
75. Kolberg-Liedtke C, Feuerhake F, Garke M, Christgen M, Kates R, Grischke EM, et al. Impact of stromal tumor-infiltrating lymphocytes (sTILs) on response to neoadjuvant chemotherapy in triple-negative early breast cancer in the WSG-ADAPT TN trial. *Breast Cancer Res*. 2022;24(1):58.
76. Chang S-W, Abdul-Kareem S, Merican AF, Zain RB. Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods. *BMC Bioinform*. 2013;14(1):170.
77. Ogier du Terrail J, Leopold A, Joly C, Bégulier C, Andreux M, Maussion C, et al. Federated learning for predicting histological response to neoadjuvant chemotherapy in triple-negative breast cancer. *Nat Med*. 2023;29(1):135–46.



78. Hacking SM, Karam J, Singh K, Gamsiz Uzun ED, Brickman A, Yakirevich E, et al. Whole slide image features predict pathologic complete response and poor clinical outcomes in triple-negative breast cancer. *Pathol Res Pract.* 2023;246:154476.
79. Cooper L, Sertel O, Kong J, Lozanski G, Huang K, Gurcan M. Feature-based registration of histopathology images with different stains: an application for computerized follicular lymphoma prognosis. *Comput Methods Programs Biomed.* 2009;96(3):182–92.
80. Roy M, Wang F, Teodoro G, Bhattarai S, Bhargava M, Rekha TS, et al. Deep learning based registration of serial whole-slide histopathology images in different stains. *J Pathol Inform.* 2023;14:100311.
81. Zamboglou C, Kramer M, Kiefer S, Bronsert P, Ceci L, Sigle A, et al. The impact of the co-registration technique and analysis methodology in comparison studies between advanced imaging modalities and whole-mount-histology reference in primary prostate cancer. *Sci Rep.* 2021;11(1):5836.
82. Keam B, Im SA, Lee KH, Han SW, Oh DY, Kim JH, et al. Ki-67 can be used for further classification of triple negative breast cancer into two subtypes with different response and prognosis. *Breast Cancer Res.* 2011;13(2):R22.
83. Kim JY, Jeong HS, Chung T, Kim M, Lee JH, Jung WH, et al. The value of phosphohistone H3 as a proliferation marker for evaluating invasive breast cancers: a comparative study with Ki67. *Oncotarget.* 2017;8(39):65064–76.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

