Journal of Digital Imaging (2019) 32:693–701
https://doi.org/10.1007/s10278-018-0144-1

Prior to Initiation of Chemotherapy, Can We Predict Breast Tumor Response? Deep Learning Convolutional Neural Networks Approach Using a Breast MRI Tumor Dataset

Richard Ha[1] · Christine Chin[2] · Jenika Karcich[1] · Michael Z. Liu[3] · Peter Chang[4] · Simukayi Mutasa[1] · Eduardo Pascual Van Sant[1] · Ralph T. Wynn[1] · Eileen Connolly[2] · Sachin Jambawalikar[3]

Published online: 25 October 2018

Abstract

We hypothesize that convolutional neural networks (CNN) can be used to predict neoadjuvant chemotherapy (NAC) response using a breast MRI tumor dataset prior to initiation of chemotherapy. An institutional review board-approved retrospective review of our database from January 2009 to June 2016 identified 141 locally advanced breast cancer patients who (1) underwent breast MRI prior to the initiation of NAC, (2) successfully completed adriamycin/taxane-based NAC, and (3) underwent surgical resection with available final surgical pathology data. Patients were classified into three groups based on their NAC response confirmed on final surgical pathology: complete (group 1), partial (group 2), and no response/progression (group 3). A total of 3107 volumetric slices of 141 tumors were evaluated. Breast tumor was identified on first T1 postcontrast dynamic images and underwent 3D segmentation. CNN consisted of ten convolutional layers, four max-pooling layers, and dropout of 50% after a fully connected layer. Dropout, augmentation, and L2 regularization were implemented to prevent overfitting of data. Non-linear functions were modeled by a rectified linear unit (ReLU). Batch normalization was used between the convolutional and ReLU layers to limit drift of layer activations during training. A three-class neoadjuvant prediction model was evaluated (group 1, group 2, or group 3). The CNN achieved an overall accuracy of 88% in three-class prediction of neoadjuvant treatment response. Three-class prediction discriminating one group from the other two was analyzed. Group 1 had a specificity of 95.1% ± 3.1%, sensitivity of 73.9% ± 4.5%, and accuracy of 87.7% ± 0.6%. Group 2 (partial response) had a specificity of 91.6% ± 1.3%, sensitivity of 82.4% ± 2.7%, and accuracy of 87.7% ± 0.6%. Group 3 (no response/progression) had a specificity of 93.4% ± 2.9%, sensitivity of 76.8% ± 5.7%, and accuracy of 87.8% ± 0.6%. It is feasible for current deep CNN architectures to be trained to predict NAC treatment response using a breast MRI dataset obtained prior to initiation of chemotherapy. Larger dataset will likely improve our prediction model.

Keywords Breast MRI · Chemotherapy treatment response · Convolutional neural network

Introduction

Neoadjuvant chemotherapy (NAC) has become a widely used treatment approach in the management of breast cancer. In addition to the established benefits of increasing rates of operability and breast-conservation for locally advanced tumors, NAC provides a unique opportunity to assess clinical efficacy of novel systemic combinations and

Work originated from Columbia University Medical Center. This work was recently presented in an oral presentation format at the 2017 RSNA meeting.

✉ Richard Ha
rh2616@columbia.edu

[1] Department of Radiology, Columbia University Irving Medical Center, 622 West 168th Street, PB-1-301, New York, NY 10032, USA

[2] Division of Radiation Oncology, Columbia University Medical Center, Presbyterian Hospital Building, 622 West 168th Street, Level B, New York, NY 10032, USA

[3] Department of Medical Physics, Columbia University Medical Center, 177 Ft. Washington Ave., Milstein Bldg Room 3-124B, New York, NY 10032-3784, USA

[4] Department of Radiology, UC San Francisco Medical Center, 505 Parnassus Ave, San Francisco, CA 94143, USA

targeted therapies in vivo within a treatment-naïve patient population [1].

Several large randomized neoadjuvant trials have demonstrated pathological complete response (pCR) to be a potential surrogate marker for clinical efficacy as there is a significant correlation between patients who achieved a pCR and improved disease-free and overall survival [2, 3]. This association varies among subtypes, with the strongest diagnostic accuracy seen in human epidermal growth factor receptor 2 (HER2)-positive and triple-negative breast cancer [4, 5]. While systemic treatments delivered in the adjuvant setting require many years of follow-up to validate a clinical benefit, pCR serves as an attractive surrogate end point for improved long-term outcome after only several weeks of neoadjuvant therapy [6, 7].

Advances in genomics have demonstrated breast cancer to be a disease with a spectrum of biologically relevant molecular subtypes. This significant disease heterogeneity poses a major challenge in the development of novel treatments. Targeted therapies may only be effective in a small subset of breast cancers, which has contributed to the difficulty establishing a therapeutic benefit in a large, heterogeneous clinical trial [8, 9]. There is potential for significant clinical benefit in streamlining the testing of novel NAC with early response assessment and prediction. This is the goal of the ongoing adaptive neoadjuvant I-SPY 2 (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis 2) trials, which have already "graduated" neratinib in HER2 positive disease and veliparib–carboplatin in triple-negative disease [10, 11]. Timely identification of responders to therapy will reduce the time, cost, and patient numbers needed to identify new beneficial therapies. Furthermore, early identification of non-responders is crucial in minimizing the potential toxicity of ineffective treatments and the delay of further exploration into potential alternative preoperative therapy [12].

Quantitative magnetic resonance imaging (MRI) has emerged as a powerful imaging modality in neoadjuvant treatment response assessment and identification of potential imaging-based biomarkers with successful incorporation into the clinical trial setting [13–16]. Recent approaches have correlated changes in specific morphologic and kinetic parameters between a baseline and interval MRI after the initiation of chemotherapy, as early as after the first cycle, to predict treatment response and pCR. Further integration of clinically relevant mathematical models to account for biologic features of tumor growth and treatment response has enhanced the predictive accuracy of these methods [17, 18]. The vast majority of current models in early-response assessment depend on interval imaging after the initiation of therapy, without the ability to successfully determine a priori treatment response or pCR, prior to the initiation of treatment given the challenges of tumor heterogeneity [13–18].

Deep learning through convolutional neural networks (CNN) has demonstrated strong performance in various image classification tasks in recent years with a growing number of applications [19]. Deep learning methods allow a machine to extract high-level information from raw input images using several non-linear modules to amplify important features for image discrimination and classification. Machine learning may be further supervised using adjustable parameters to intricately correlate specific inputs and outputs.

The purpose of this study is to develop a novel CNN to predict NAC response using a baseline breast MRI tumor dataset and pathological confirmation of treatment response.

Materials and Methods

Patient Selection and Eligibility

A Health Insurance Portability and Accountability Act compliant, institutional review board approved, retrospective review of our database identified 141 patients with the diagnosis of breast cancer between January 1, 2005 and June 1, 2016. All patients met the following criteria: (1) underwent a staging breast MRI prior to the initiation of therapy, (2) received adriamycin-based and/or taxane-based neoadjuvant chemotherapy with additional HER2 directed therapy (trastuzumab/pertuzumab) in patients with HER2 positive tumor, and (3) successfully underwent surgical resection of their primary breast tumor with appropriate lymph node sampling.

Pathologic Analysis

Data on tumor pathologic characteristics were obtained from the original pathology reports of the core biopsy specimen. Breast tumor subtype was determined based on immunohistochemical (IHC) staining of the estrogen receptor (ER) and progesterone receptor (PR) interpreted according to the American Society of Clinical Oncology and College of American Pathologists Guidelines. Tumors were considered receptor positive if either ER or PR demonstrated $\geq 1\%$ positive staining [20]. Tumors were considered HER2 positive if they were 3+ by immunohistochemistry or demonstrated gene amplification with a ratio of HER2/CEP17 ≥ 2 by in situ hybridization [21]. Breast tumor subtypes were defined as follows: Luminal A (ER/PR positive, HER2 negative), luminal B (ER/PR positive, HER2 positive); HER2 positive (ER/PR negative, HER2 positive), and triple negative or basal-like (ER/PR and HER2 negative). Clinical and pathologic staging was determined based on the American Joint Committee on Cancer TNM Staging Manual, 7th edition. Patients were classified into three groups based on their NAC response confirmed on final surgical pathology: pathologic complete response (group 1), partial response (group 2), and no response/progression (group 3). pCR was defined as no

residual invasive disease in the breast or lymph nodes on surgical pathology specimens (ypT0/Tis ypN0).

MRI Methods

MRI was performed on a 1.5-T or 3.0-T commercially available system (Signa Excite, GE Healthcare) using an eight-channel breast array coil. A bilateral sagittal T1-weighted fat-suppressed fast spoiled gradient-echo sequence (17/2.4; flip angle, 35°; bandwidth, 31–25 Hz) was then performed before and after a rapid bolus injection (gadobenate dimeglumine/Multihance; Bracco Imaging; 0.1 mmol/kg) delivered through an IV catheter. Image acquisition started after contrast material injection and was obtained consecutively with each acquisition time of 120 s. Section thickness was 2–3 mm using a matrix of 256 × 192 and a field of view of 18–22 cm. Frequency was in the antero-posterior direction.

Computer-Based Image Analysis: Image Preprocessing

For each breast MRI, tumor was identified on first T1 postcontrast dynamic images (Figs. 1, 2, and 3). The entire breast volume underwent 3D segmentation by a breast fellowship trained radiologist with 8 years of experience using an open source software platform 3D Slicer [22]. A total of 3107 volumetric slices for 141 tumors were collected. The data was normalized by subtracting the mean intensity value of each slice and dividing by the standard deviation of each slice. A 64 × 64 voxel crop of the segmented tumor was then fed into the convolution neural network (Fig. 5). An average of 22 slices of volumetric data per tumor was used, with a threshold of 75 voxels per slice. At the time of training, real-time data augmentation was performed to limit overfitting of data. Using random affine transformation, additional images were created by modifying the images: randomly rotate images (range 10°), horizontal flip, shear (range 0.1), and zoom (range 0.1) the images.

Convolution Neural Network

The convolution neural network was implemented using the keras toolbox (https://github.com/fchollet/keras) with tensorflow backend in python. The models were run on a linux workstation utilizing NVIDIA GTX 1080ti GPUs.

CNN Architecture

The architecture follows the general structure of the VGG 16 network. The overall network architecture is shown in Fig. 4. In this architecture, a block consists of multiple convolution layers of 3 × 3 convolution kernels that have progressively increasing feature channels in deeper layers. The convolution layers are followed by the non-linear rectified linear unit activation function (RELU [23]). Before each increase of feature channels, a 2 × 2 max pooling layer is applied to reduce the

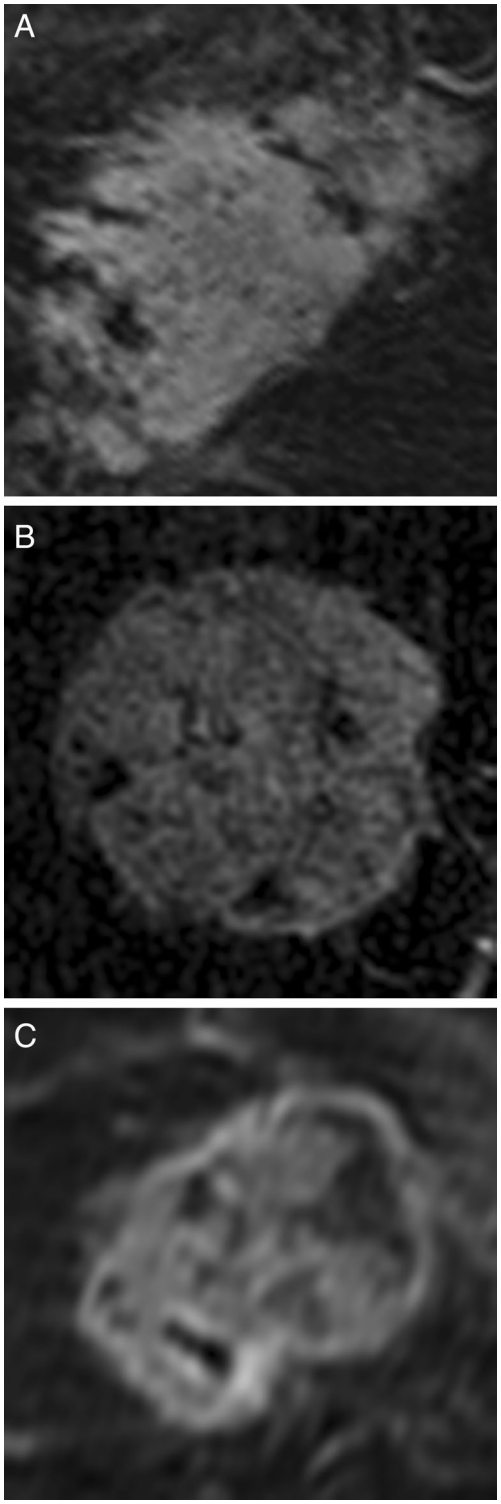

Fig. 1 **a**–**c** Examples of T1 postcontrast breast MRI images of tumors that underwent CNN evaluation in our study from patients with complete pathologic response

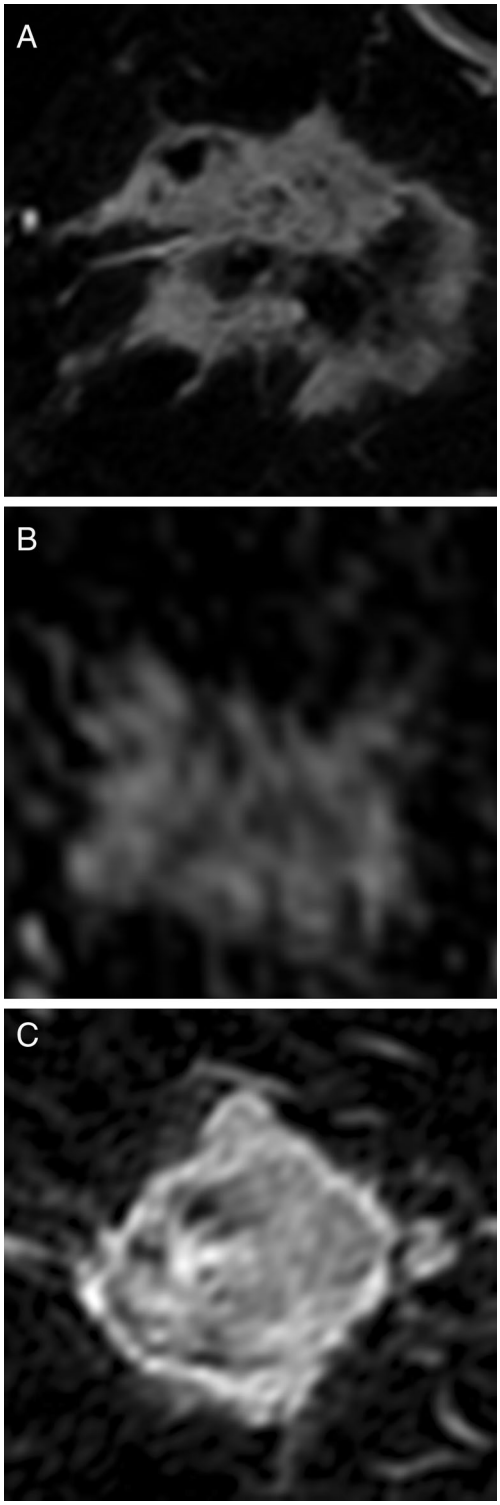

Fig. 2 **a**–**c** Examples of T1 postcontrast breast MRI images of tumors that underwent CNN evaluation in our study from patients with partial response

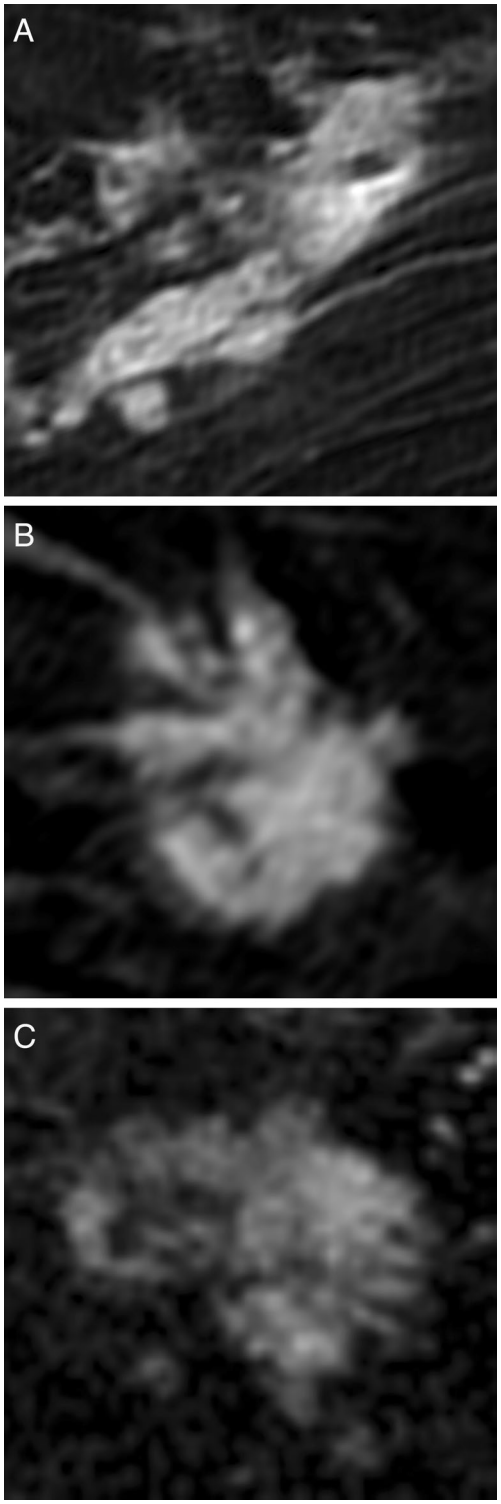

Fig. 3 **a**–**c** Examples of T1 postcontrast breast MRI images of tumors that underwent CNN evaluation in our study from patients with no response/disease progression

amount of parameters and computation in the network, serving the double purpose of controlling overfitting. Four of these blocks are stacked on each other before the architecture flattens out to a full connected dense layer. The fully connected layer acts as a perceptron and is mathematically similar to a least squares regression. Dropout of 25% is applied in the dense layer to prevent overfitting by limiting co-adaptation of parameters [24]. L2 regularization with a beta of 0.01 is

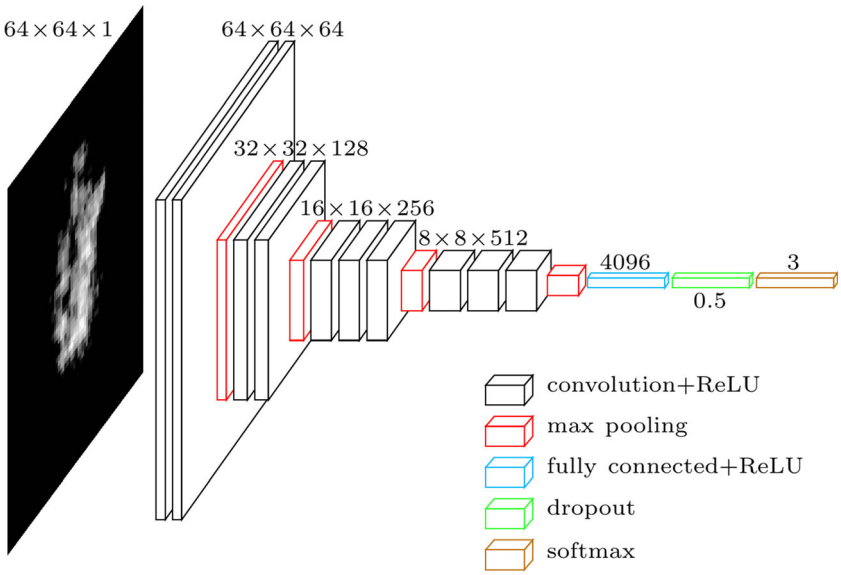

Fig. 4 CNN architecture for neoadjuvant class prediction

used after the dense layer to place a penalty on the squared magnitude of the kernel weights. This penalizes outlier parameters and in encourages generalizable parameters. This reduces overfitting in the model and leads to a more generalizable model. A softmax classifier is used for the loss function.

CNN Training

The data was divided into 80% validation and 20% test. The validation test set was then divided into fivefold, and fivefold cross-validation was performed. Training from scratch without pretrained weights was done over 100 epochs using adam optimizer with nesterov momentum at an initial learning rate of 0.002. Each of the five models was tested against the 20% hold out data to obtain sensitivity, specificity, and accuracy. Receiver operator curves were also calculated for each of the five models.

Results

Patient Tumor Pathology and Response

A total of 141 patients met the criteria for inclusion in this study. Three-class neoadjuvant prediction model was evaluated for the three patient groups. The breakdown of tumor response and molecular subtype is shown in Table 1. Group 1 consisted of 46 patients with pathologic complete response. Group 2 consisted of 57 patients with partial response. Group 3 consisted of 38 patients with no response to progression on chemotherapy. The molecular subtype based on IHC staining included the following: 61 luminal A, 39 luminal B, 16 HER2 positive, and 25 triple negative or basal-like.

The rate of pCR is shown in Table 2, demonstrating 18% (11/61) of the luminal A, 46% (18/39) of the luminal B group, 50% (8/16) of the HER2 positive group, and 36% (9/25) of the triple-negative group achieved pCR. Combined luminal B, HER2 positive, and triple-negative tumors had a significantly higher rate of pCR compared to luminal A, with a rate of 44% (35/80) versus 18% (11/61) respectively ($p = 0.002$).

The rate of no response/progression of disease is shown in Table 3, demonstrating 43% (26/61) of the luminal A group, 10% (4/39) of the luminal B group, 13% (2/16) of the HER2 positive group, and 24% (6/25) of the triple-negative group showed no treatment response or progression of disease. Luminal A tumors had a significantly higher rate of no response/progression compared to the other three groups, with a rate of 43% (26/61) versus 15% (12/80) respectively ($p = 0.0005$).

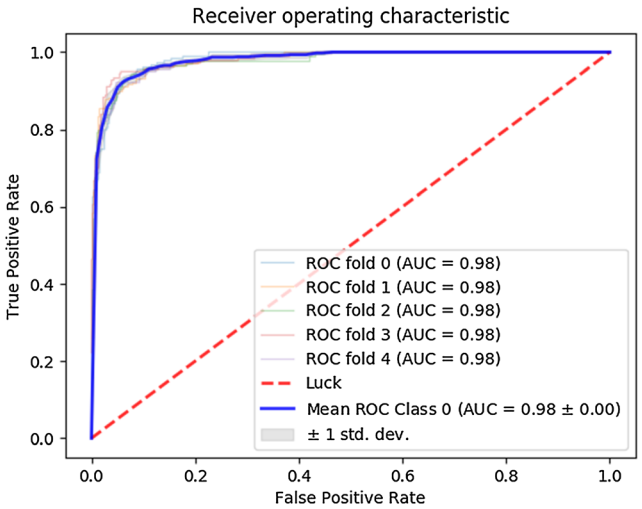

Fig. 5 ROC plot for three-class prediction of NAC treatment response

Table 1 Pathologic tumor response and molecular subtype

Molecular subtype	Pathologic response			
	Complete	Partial	No response/progression	Total
Luminal A	11	24	26	61
Luminal B	18	17	4	39
HER2+	8	6	2	16
Triple−	9	10	6	25
Total	46	57	38	141

Table 3 Rate of no response/progression per molecular subtype

Molecular subtype	Pathologic response	
	No response/progression	
Luminal A	26/61	43%
Luminal B	4/39	10%
HER2+	2/16	13%
Triple−	6/25	24%

CNN Statistical Analysis

The confusion matrix (Table 4) shows the convolution neural network predicted class of the hold out test data versus the true class of the hold out test data. The values represent the average number of slices over the fivefold of cross-validation plus or minus the standard deviation. A final softmax score threshold of 0.5 was used for classification. The CNN achieved an overall mean accuracy of 88% (95% CI, ± 0.6%) in three-class prediction of NAC treatment response on a fivefold validation accuracy test. The ROC plot is shown in Fig. 5. Three-class prediction discriminating one class from the other two was analyzed. Group 1 (complete response) had a specificity of 95.1% ± 3.1%, sensitivity of 73.9% ± 4.5%, and accuracy of 87.7% ± 0.6%. Group 2 (partial response) had a specificity of 91.6% ± 1.3%, sensitivity of 82.4% ± 2.7%, and accuracy of 87.7% ± 0.6%. Group 3 (no response/progression) had a specificity of 93.4% ± 2.9%, sensitivity of 76.8% ± 5.7%, and accuracy of 87.8% ± 0.6%.

Conclusion and Discussion

Prior to initiation of therapy, the CNN algorithm used in our study achieved an overall accuracy of 88% in predicting NAC response in patients with locally advanced breast cancer. Our results demonstrate that it is feasible to utilize CNN to predict NAC response prior to initiation of therapy (Fig. 6). This represents an improved approach to early treatment response assessment based on a baseline breast MRI obtained prior to the initiation of treatment and significantly improves on current prediction methods that rely on interval imaging after the initiation of therapy.

Although there has been significant progress in MRI to assess therapy response, the vast majority of studies thus far depend on interval imaging after initiation of therapy. Quantitative imaging techniques have become an active area of research given the limitations of qualitative tumor response assessment using the Response Evaluation Criteria in Solid Tumors (RECIST) [25]. Quantitative methods of response assessment have examined changes in kinetic parameters (i.e., volume transfer constant K_{trans}, exchange rate constant k_{ep}) in dynamic contrast-enhanced MRI (DCE-MRI) [26–28], as well as morphologic changes (three-dimensional volume, signal enhancement ratio, tissue cellularity) using DCE-MRI, and diffusion-weighted MRI (DW-MRI) with predictive value after one or more cycles of therapy [14, 15, 29]. The limitations of these methods include the often delayed morphologic-based changes that occur despite treatment-induced biologic response that are not reflected by imaging performed during or shortly after completion of therapy. By incorporating a mechanically coupled reaction–diffusion model using patient-specific imaging data to drive a biomechanical model of tumor growth, Weis et al. showed improved prediction of therapy response compared to prior techniques, achieving a sensitivity and specificity of 92% and 84%, respectively [30]. While significant advances in response assessment have been shown, the previously described studies all rely on interval imaging after initiation of therapy.

A single prior study by Ravichandran et al. applied a CNN to 166 breast cancer patients to predict pCR from a pretreatment MRI tumor dataset [31]. They achieved an area under

Table 2 Rate of pCR per molecular subtype

Molecular subtype	Pathologic response	
	Complete	
Luminal A	11/61	18%
Luminal B	18/39	46%
HER2+	8/16	50%
Triple−	9/25	36%

Table 4 Convolution neural network performance confusion matrix

		Predicted response		
		Complete	Partial	No response
True response	Complete	160.2 ± 4.4	10.6 ± 2.9	8.6 ± 2.9
	Partial	9.2 ± 4.2	219.6 ± 5.7	18.2 ± 5.4
	No response	10.8 ± 4.3	19.2 ± 3.1	165 ± 6.7

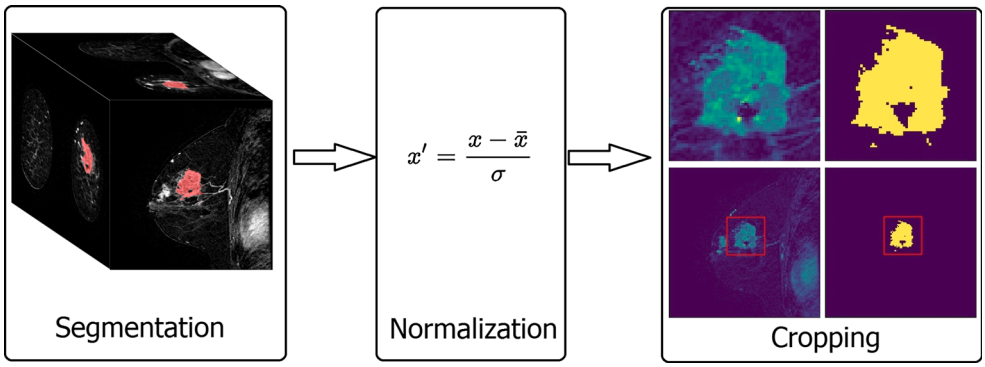

Fig. 6 Block diagram of image preprocessing

the curve (AUC) of 0.77 and an overall accuracy of 82%. Inclusion of clinical variables improved response prediction with an AUC of 0.85 and an accuracy of 85%. While their study put forth a novel approach to pretreatment response prediction, our study further advances prediction strength with an improved overall accuracy of 88%.

Currently available clinical and pathologic data show luminal B, HER2-positive, and triple-negative breast cancer respond best to NAC. A large meta-analysis of 30 studies including 11,695 patients investigating pCR after NAC showed average rates of pCR were 8.3% in luminal A, 18.7% in luminal B, 38.9% in HER2 positive, and 31.1% in triple-negative breast cancer subtypes [32]. Similarly, in our study, HER2-positive and triple-negative tumors achieved significantly higher pCR, compared to the luminal A subtype. While this information is helpful, it cannot be used solely to predict who will respond to NAC given over half of these patients do not have pCR.

In our study, only 18% of luminal A breast tumors achieved pCR, which is consistent with previously reported poor response to NAC in this subtype [32]. In addition, our results show 43% of patients in this group demonstrated no treatment response or progression of disease. In these non-responders, there is a potential for upfront use of novel agents which may be beneficial to the patient, while minimizing toxicity from ineffective treatments and delaying treatment time. However, relying on clinical data alone is not enough to predict who will have a poor response to the current NAC regimen. Given limitation to utilizing available clinical and pathologic data for NAC prediction, our CNN model may have an important clinical utility as a better predictive tool prior to initiation of treatment not relying solely on pathologic data.

The major limitation of our study includes the small, retrospective single-institution nature of our experience. We successfully trained our machine learning algorithm to predict treatment outcome in a subset of patients all treated at our institution—thus inherently selecting for various demographic and clinical features, as well as management biases that may not apply to a larger, more diverse patient population. In addition, patients in this study underwent MRI imaging at different magnetic field strengths (1.5 or 3.0 T). However, selection bias is likely negligible given that the choice of patients undergoing MRI on a 1.5 or 3.0-T magnet was randomly determined purely based on availability of the scanner. In addition, classifications based on ER, PR, and HER2 status are only approximations of the underlying genotype-based breast cancer subtype. However, given the high cost of full gene expression profiling, IHC staining has become a widely used and generally accepted method of approximating true molecular subtypes. [33–35] Other limitations include potential overfitting due to relatively small dataset in our study. However, accepted methods were used to limit overfitting in our study including 50% dropout, augmentation, and L2 regularization.

Prior to initiation of NAC, it is feasible to utilize CNN to predict response using a baseline MRI tumor dataset. A larger dataset will likely improve our model with potential for implementation into the clinical setting. Our early prediction model of treatment response has the potential to impact clinical management in patients with locally advanced breast cancer, including the opportunity to direct appropriate therapy in non-responders, minimize toxicity from ineffective therapies, and facilitate the upfront use of novel targeted treatment in the neoadjuvant setting.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Rastogi P, Anderson SJ, Bear HD, Geyer CE, Kahlenberg MS, Robidoux A, Margolese RG, Hoehn JL, Vogel VG, Dakhil SR, Tamkus D, King KM, Pajon ER, Wright MJ, Robert J, Paik S, Mamounas EP, Wolmark N: Preoperative chemotherapy: updates of National Surgical Adjuvant Breast and Bowel Project Protocols B-18 and B-27. J Clin Oncol 26:778–785, 2008
2. Gianni L, Pienkowski T, Im YH, Tseng LM, Liu MC, Lluch A, Starosławska E, de la Haba-Rodriguez J, Im SA, Pedrini JL, Poirier B, Morandi P, Semiglazov V, Srimuninnimit V, Bianchi GV, Magazzù D, McNally V, Douthwaite H, Ross G, Valagussa P: 5-year analysis of neoadjuvant pertuzumab and trastuzumab in

patients with locally advanced, inflammatory, or early-stage HER2-positive breast cancer (NeoSphere): a multicentre, open-label, phase 2 randomised trial. Lancet Oncol 17:791–800, 2016

3. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, Bonnefoi H, Cameron D, Gianni L, Valagussa P, Swain SM, Prowell T, Loibl S, Wickerham DL, Bogaerts J, Baselga J, Perou C, Blumenthal G, Blohmer J, Mamounas EP, Bergh J, Semiglazov V, Justice R, Eidtmann H, Paik S, Piccart M, Sridhara R, Fasching PA, Slaets L, Tang S, Gerber B, Geyer, Jr CE, Pazdur R, Ditsch N, Rastogi P, Eiermann W, von Minckwitz G: Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet 384:164–172, 2014
4. Esserman LJ, Berry DA, DeMichele A, Carey L, Davis SE, Buxton M, Hudis C, Gray JW, Perou C, Yau C, Livasy C, Krontiras H, Montgomery L, Tripathy D, Lehman C, Liu MC, Olopade OI, Rugo HS, Carpenter JT, Dressler L, Chhieng D, Singh B, Mies C, Rabban J, Chen YY, Giri D, van 't Veer L, Hylton N: Pathologic complete response predicts recurrence-free survival more effectively by cancer subset: results from the I-SPY 1 TRIAL–CALGB 150007/150012, ACRIN 6657. J Clin Oncol 30:3242–3249, 2012
5. von Minckwitz G, Untch M, Blohmer JU, Costa SD, Eidtmann H, Fasching PA, Gerber B, Eiermann W, Hilfrich J, Huober J, Jackisch C, Kaufmann M, Konecny GE, Denkert C, Nekljudova V, Mehta K, Loibl S: Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. J Clin Oncol 30:1796–1804, 2012
6. Wang-Lopez Q, Chalabi N, Abrial C, Radosevic-Robin N, Durando X, Mouret-Reynier MA, Benmammar KE, Kullab S, Bahadoor M, Chollet P, Penault-Llorca F, Nabholtz JM: Can pathologic complete response (pCR) be used as a surrogate marker of survival after neoadjuvant therapy for breast cancer? Crit Rev Oncol Hematol 95:88–104, 2015
7. Broglio KR, Quintana M, Foster M, Olinger M, McGlothlin A, Berry SM, Boileau JF, Brezden-Masley C, Chia S, Dent S, Gelmon K, Paterson A, Rayson D, Berry DA: Association of pathologic complete response to neoadjuvant therapy in HER2-positive breast cancer with long-term outcomes: a meta-analysis. JAMA Oncol 2:751–760, 2016
8. Carey LA: Neoadjuvant clinical trial designs: challenges of the genomic era. Breast 24(Suppl 2):S88–S90, 2015
9. Carey LA, Winer EP: I-SPY 2—toward more rapid progress in breast cancer treatment. N Engl J Med 375:83–84, 2016
10. Park JW, Liu MC, Yee D, Yau C, van 't Veer L, Symmans WF, Paoloni M, Perlmutter J, Hylton NM, Hogarth M, DeMichele A, Buxton MB, Chien AJ, Wallace AM, Boughey JC, Haddad TC, Chui SY, Kemmer KA, Kaplan HG, Isaacs C, Nanda R, Tripathy D, Albain KS, Edmiston KK, Elias AD, Northfelt DW, Pusztai L, Moulder SL, Lang JE, Viscusi RK, Euhus DM, Haley BB, Khan QJ, Wood WC, Melisko M, Schwab R, Helsten T, Lyandres J, Davis SE, Hirst GL, Sanil A, Esserman LJ, Berry DA, I-SPY 2 Investigators: Adaptive randomization of neratinib in early breast cancer. N Engl J Med 375:11–22, 2016
11. Rugo HS, Olopade OI, DeMichele A, Yau C, van 't Veer L, Buxton MB, Hogarth M, Hylton NM, Paoloni M, Perlmutter J, Symmans WF, Yee D, Chien AJ, Wallace AM, Kaplan HG, Boughey JC, Haddad TC, Albain KS, Liu MC, Isaacs C, Khan QJ, Lang JE, Viscusi RK, Pusztai L, Moulder SL, Chui SY, Kemmer KA, Elias AD, Edmiston KK, Euhus DM, Haley BB, Nanda R, Northfelt DW, Tripathy D, Wood WC, Ewing C, Schwab R, Lyandres J, Davis SE, Hirst GL, Sanil A, Berry DA, Esserman LJ, I-SPY 2 Investigators: Adaptive randomization of veliparib-carboplatin treatment in breast cancer. N Engl J Med 375:23–34, 2016
12. von Minckwitz G, Blohmer JU, Costa SD, Denkert C, Eidtmann H, Eiermann W, Gerber B, Hanusch C, Hilfrich J, Huober J, Jackisch C, Kaufmann M, Kümmel S, Paepke S, Schneeweiss A, Untch M, Zahm DM, Mehta K, Loibl S: Response-guided neoadjuvant chemotherapy for breast cancer. J Clin Oncol 31:3623–3630, 2013
13. Gu YL, Pan SM, Ren J, Yang ZX, Jiang GQ: Role of magnetic resonance imaging in detection of pathologic complete remission in breast cancer patients treated with neoadjuvant chemotherapy: a meta-analysis. Clin Breast Cancer 17:245–255, 2017
14. Li W, Arasu V, Newitt DC, Jones EF, Wilmes L, Gibbs J, Kornak J, Joe BN, Esserman LJ, Hylton NM, ACRIN 6657 Trial Team and I-SPY Investigators Network: Effect of MR imaging contrast thresholds on prediction of neoadjuvant chemotherapy response in breast cancer subtypes: a subgroup analysis of the ACRIN 6657/I-SPY 1 TRIAL. Tomography 2:378–387, 2016
15. Hylton NM, Blume JD, Bernreuter WK, Pisano ED, Rosen MA, Morris EA, Weatherall PT, Lehman CD, Newstead GM, Polin S, Marques HS, Esserman LJ, Schnall MD, ACRIN 6657 Trial Team and I-SPY 1 TRIAL Investigators: Locally advanced breast cancer: MR imaging for prediction of response to neoadjuvant chemotherapy—results from ACRIN 6657/I-SPY TRIAL. Radiology 263: 663–672, 2012
16. Hylton NM, Gatsonis CA, Rosen MA, Lehman CD, Newitt DC, Partridge SC, Bernreuter WK, Pisano ED, Morris EA, Weatherall PT, Polin SM, Newstead GM, Marques HS, Esserman LJ, Schnall MD, For the ACRIN 6657 Trial Team and I-SPY 1 TRIAL Investigators: Neoadjuvant chemotherapy for breast cancer: functional tumor volume by MR imaging predicts recurrence-free survival-results from the ACRIN 6657/CALGB 150007 I-SPY 1 TRIAL. Radiology 279:44–55, 2016
17. Weis JA, Miga MI, Yankeelov TE: Three-dimensional image-based mechanical modeling for predicting the response of breast cancer to neoadjuvant therapy. Comput Methods Appl Mech Eng 314:494–512, 2017
18. Yankeelov TE: Integrating imaging data into predictive biomathematical and biophysical models of cancer. ISRN Biomath 2012:1–12, 2012
19. LeCun Y, Bengio Y, Hinton G: Deep learning. Nature 521:436–444, 2015
20. Hammond ME, Hayes DF, Dowsett M et al.: American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol 28:2784–2795, 2010
21. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane L, Allison KH, Allred DC, Bartlett JM, Bilous M, Fitzgibbons P, Hanna W, Jenkins RB, Mangu PB, Paik S, Perez EA, Press MF, Spears PA, Vance GH, Viale G, Hayes DF, American Society of Clinical Oncology, College of American Pathologists: Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. Arch Pathol Lab Med 138:241–256, 2014
22. Pieper S, Halle M, Kikinis R. 3D Slicer. In: 2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano (IEEE Cat No 04EX821). IEEE; p. 632–635.
23. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv:1412.6980
24. Srivastava N, Hinton GE, Krizhevsky A, Sutskever I, Salakhutdinov R: Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res. 15:1929–1958, 2014
25. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, van Glabbeke M, van Oosterom AT, Christian MC, Gwyther SG: New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst 92:205–216, 2000
26. Li X, Arlinghaus LR, Ayers GD, Chakravarthy AB, Abramson RG, Abramson VG, Atuegwu N, Farley J, Mayer IA, Kelley MC, Meszoely IM, Means-Powell J, Grau AM, Sanders M, Bhave SR,

Yankeelov TE: DCE-MRI analysis methods for predicting the response of breast cancer to neoadjuvant chemotherapy: pilot study findings. Magn Reson Med 71:1592–1602, 2014

27. Li X, Abramson RG, Arlinghaus LR, Kang H, Chakravarthy AB, Abramson VG, Farley J, Mayer IA, Kelley MC, Meszoely IM, Means-Powell J, Grau AM, Sanders M, Yankeelov TE: Multiparametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer. Invest Radiol 50:195–204, 2015
28. Ah-See ML, Makris A, Taylor NJ et al.: Early changes in functional dynamic magnetic resonance imaging predict for pathologic response to neoadjuvant chemotherapy in primary breast cancer. Clin Cancer Res 14:6580–6589, 2008
29. Atuegwu NC, Arlinghaus LR, Li X, Chakravarthy AB, Abramson VG, Sanders ME, Yankeelov TE: Parameterizing the logistic model of tumor growth by DW-MRI and DCE-MRI data to predict treatment response and changes in breast cancer cellularity during neoadjuvant chemotherapy. Transl Oncol 6:256–264, 2013
30. Weis JA, Miga MI, Arlinghaus LR, Li X, Abramson V, Chakravarthy AB, Pendyala P, Yankeelov TE: Predicting the response of breast cancer to neoadjuvant therapy using a mechanically coupled reaction-diffusion model. Cancer Res 75:4697–4707, 2015
31. Ravichandran K, Braman N, Janowczyk A, Madabushi A: A deep learning classifier for prediction of pathological complete response to neoadjuvant chemotherapy from baseline breast DCE-MRI. SPIE medical imaging 2018: computer-aided diagnosis 105750C, 2018
32. Houssami N, Macaskill P, von Minckwitz G, Marinovich ML, Mamounas E: Meta-analysis of the association of breast cancer subtype and pathologic complete response to neoadjuvant chemotherapy. Eur J Cancer 48:3342–3352, 2012
33. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE, Borresen-Dale AL: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98:10869–10874, 2001
34. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, Karaca G, Troester MA, Tse CK, Edmiston S, Deming SL, Geradts J, Cheang MCU, Nielsen TO, Moorman PG, Earp HS, Millikan RC: Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 295:2492–2502, 2006
35. Nguyen PL, Taghian AG, Katz MS, Niemierko A, Abi Raad RF, Boon WL, Bellon JR, Wong JS, Smith BL, Harris JR: Breast cancer subtype approximated by estrogen receptor, progesterone receptor, and HER-2 is associated with local and distant recurrence after breast-conserving therapy. J Clin Oncol. 26(14):2373–2378, 2008