

## Article

# Deep Learning for Fully Automatic Tumor Segmentation on Serially Acquired Dynamic Contrast-Enhanced MRI Images of Triple-Negative Breast Cancer

Zhan Xu <sup>1</sup>, David E. Rauch <sup>1</sup>, Rania M. Mohamed <sup>2</sup>, Sanaz Pashapoor <sup>2</sup>, Zijian Zhou <sup>1</sup>, Bikash Panthi <sup>1</sup>, Jong Bum Son <sup>1</sup>, Ken-Pin Hwang <sup>1</sup>, Benjamin C. Musall <sup>1</sup>, Beatriz E. Adrada <sup>2</sup>, Rosalind P. Candelaria <sup>2</sup>, Jessica W. T. Leung <sup>2</sup>, Huong T. C. Le-Petross <sup>2</sup>, Deanna L. Lane <sup>2</sup>, Frances Perez <sup>2</sup>, Jason White <sup>3</sup>, Alyson Clayborn <sup>3</sup>, Brandy Reed <sup>4</sup>, Huiqin Chen <sup>5</sup>, Jia Sun <sup>5</sup>, Peng Wei <sup>5</sup>, Alastair Thompson <sup>6</sup>, Anil Korkut <sup>7</sup>, Lei Huo <sup>8</sup>, Kelly K. Hunt <sup>9</sup>, Jennifer K. Litton <sup>3</sup>, Vicente Valero <sup>3</sup>, Debu Tripathy <sup>3</sup>, Wei Yang <sup>2</sup>, Clinton Yam <sup>3</sup> and Jingfei Ma <sup>1,\*</sup>

<sup>1</sup> Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA; zxu5@mdanderson.org (Z.X.)

<sup>2</sup> Department of Breast Imaging, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>3</sup> Department of Breast Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>4</sup> Department of Clinical Research Imaging, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>5</sup> Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>6</sup> Section of Breast Surgery, Baylor College of Medicine, Houston, TX 77030, USA

<sup>7</sup> Department of Bioinformatics & Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>8</sup> Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>9</sup> Department of Breast Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

\* Correspondence: jma@mdanderson.org



**Citation:** Xu, Z.; Rauch, D.E.;

Mohamed, R.M.; Pashapoor, S.; Zhou, Z.; Panthi, B.; Son, J.B.; Hwang, K.-P.; Musall, B.C.; Adrada, B.E.; et al. Deep Learning for Fully Automatic Tumor Segmentation on Serially Acquired Dynamic Contrast-Enhanced MRI Images of Triple-Negative Breast Cancer. *Cancers* **2023**, *15*, 4829.

<https://doi.org/10.3390/cancers15194829>

Academic Editor: Cynthia Ma

Received: 9 August 2023

Revised: 10 September 2023

Accepted: 22 September 2023

Published: 2 October 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Simple Summary:** Quantitative image analysis of cancers requires accurate tumor segmentation that is often performed manually. In this study, we developed a deep learning model with a self-configurable nnU-Net for fully automated tumor segmentation on serially acquired dynamic contrast-enhanced MRI images of triple-negative breast cancer. In an independent testing dataset, our nnU-Net-based deep learning model performed automated tumor segmentation with a Dice similarity coefficient of 93% and a sensitivity of 96%.

**Abstract:** Accurate tumor segmentation is required for quantitative image analyses, which are increasingly used for evaluation of tumors. We developed a fully automated and high-performance segmentation model of triple-negative breast cancer using a self-configurable deep learning framework and a large set of dynamic contrast-enhanced MRI images acquired serially over the patients' treatment course. Among all models, the top-performing one that was trained with the images across different time points of a treatment course yielded a Dice similarity coefficient of 93% and a sensitivity of 96% on baseline images. The top-performing model also produced accurate tumor size measurements, which is valuable for practical clinical applications.

**Keywords:** deep learning; tumor segmentation; triple-negative breast cancer

## 1. Introduction

Triple-negative breast cancer (TNBC) is an aggressive subtype of breast cancer, representing approximately 15% of all breast cancers and contributing to approximately 40% of breast cancer-related deaths [1]. Neoadjuvant systemic therapy followed by surgery is the

standard-of-care treatment for TNBC. However, the responses of patients to neoadjuvant systemic therapy vary, and only approximately 50% of patients achieve a pathological complete response, which is a useful surrogate marker for favorable long-term clinical outcomes. Given the aggressive nature of TNBC and substantial variability in pathologic complete response rates, noninvasive imaging methods for accurate tumor characterization and early prediction of tumor response to therapy will be highly valuable. Quantitative image analyses of tumors are increasingly used for early detection of cancer [2], accurate tumor localization and staging [3], and treatment response assessment [4], or prediction [5].

An important step in quantitative image analyses is tumor segmentation. The most commonly used method of tumor segmentation is manual contouring and annotation by experienced radiologists. However, this process is labor-intensive and tedious, as well as being susceptible to human errors and inter-reader variations [6]. To overcome these challenges, computer-aided diagnosis algorithms have been developed. Such algorithms are usually model-based and involve active contouring [7–9], automated thresholding [10,11], region growing [12,13], or combinations of methods [14]. However, these methods are typically applicable only under specific assumptions since their optimization depends on pre-defined constraint thresholds.

Deep learning techniques, especially convolutional neural networks [3], have been explored for their potential in delineating images or multidimensional features, including automated tumor segmentation [15,16]. A deep learning model [17] trained with over 45,000 mammograms outperformed a previous computer-aided diagnosis system that relied on designed features and selected seed points. U-Net, a dedicated convolutional neural network for medical imaging segmentation, has gained popularity [18]. An implementation using a two-stage U-Net demonstrated accurate breast tumor segmentation across multiple datasets [19]. According to a recent review [20], three out of six referenced studies of breast cancer segmentation were based on U-Net.

Dynamic contrast-enhanced (DCE) MRI is capable of measuring contrast agent kinetics and has shown high sensitivity in detecting breast cancer [21]. The kinetic texture, which represents contrast enhancement characteristics, can be employed to segment regions of tissue with similar vascular properties [22,23]. Early methods of using DCE-MRI for tumor segmentation can be categorized as the atlas-based methods [24,25], cluster methods [26], and classifier-based methods [27]. Recent studies have shown that deep learning-based methods can offer superior tumor segmentation performance, potentially by integrating both kinetic characteristics of the signals and tumor texture information (e.g., voxel/pixel based, including tumor shape and background tissue homogeneity). Using fully convolutional networks, a hierarchical convolutional neural network framework was developed to perform segmentation [28]. Another deep learning approach is to compose a multiple-components U-Net framework to transform the segmentation into a multi-classification task, such as to separate tumors, fibroglandular tissue [29], and other tissues. DeepMedic [30] is one widely-accepted framework that efficiently computes the model parameters via a configurable multi-resolution pathway. It was originally developed for automated brain lesion segmentation, but has since been adapted for breast tumor studies [31,32]. Other deep learning frameworks, such as SegNet [31,33] and ResNet [3,34,35], also achieved good results in automated segmentation.

However, the success rate of the translation of these models trained from one dataset to a different and independent dataset is still very limited. The technical challenge is likely caused by the methods' configuration and parameters optimization that are needed for the dataset diversity. Adapting a model iteratively requires extensive knowledge and effort, which increases the complexity of the clinical workflow. Another challenge is the lack of high-quality datasets for the model training of a specific disease population. MRI datasets are typically much larger in image size but smaller in population size than sonogram or mammogram datasets. The availability of DCE-MRI data specific to TNBC patients is even more limited and poses another barrier to train a functional model.

In this work, we aimed to develop a model for automated segmentation of TNBC on DCE-MRI images. The deep learning framework we employed to build our model is nnU-Net [36], which is based on the standard U-Net structure but offers a unique feature of automated hyperparameter optimization. In nnU-Net, some model parameters are empirically derived using ten datasets from the Medical Segmentation Decathlon [37] while the remaining parameters are customized with the application-specific training datasets. We hypothesize that the nnU-Net framework, with a self-configuring segmentation model that has demonstrated broad success over a variety of datasets and image modalities, could provide accurate automated segmentation of TNBC using DCE-MRI images. To overcome the challenge of the limited datasets, we combined the data acquired serially throughout a patient's treatment course and hypothesized that tumor progression over different time points will improve the segmentation performance compared to using the data from a single time point. Specifically, we trained and tested models over multiple semiquantitative maps from DCE-MRI images. The models were systematically evaluated in terms of the Dice similarity coefficient (DSC) and sensitivity. Our findings show that the subtraction between pre-contrast and post-contrast images provided the best performance. Using the model trained from data across all time points over the treatment course, we achieved a median DSC of 93% and sensitivity of 96% over our independent testing dataset across all time points.

## 2. Materials and Methods

This study was approved by the Institutional Review Board (IRB) of The University of Texas MD Anderson Cancer Center and was part of an ongoing IRB-approved prospective clinical trial (NCT02276443) of patients with stage I-III TNBC who were being monitored for responses to neoadjuvant systemic therapy. This study followed the ethical guidelines set out in the Declaration of Helsinki, and written informed consent was obtained from each participant.

### 2.1. Dataset

A total of 301 patients with biopsy-confirmed stage I-III TNBC were included in this study. The data inclusion criteria were identical to and described in a previously published work [38]. The imaging protocol for each patient included DCE-MRI, which was acquired at multiple time points during a patient's treatment course: at baseline (BL), after two cycles (C2), and after four cycles (C4) of neoadjuvant systemic therapy. Among all the patients, 299 had BL scans, 221 had C2 scans, and 272 had C4 scans. Data from patients with inflammatory breast cancer, failed manual tumor segmentation due to technical issues, or a complete response to treatment without any visible residual enhancing lesions were excluded from model training. In total, 744 datasets (285 from BL, 207 from C2, and 252 from C4) were used for this study. Patients' demographic and clinical characteristics for these datasets are presented in Table 1.

**Table 1.** Patients' characteristics of all datasets included in this study and datasets from baseline (BL), after 2 cycles (C2), and after 4 cycles (C4) of neoadjuvant systemic therapy.

Characteristic	All Datasets	BL	C2	C4
No. of datasets	744	285	207	252
Age, mean $\pm$ SD, years	50 $\pm$ 11	50 $\pm$ 11	50 $\pm$ 11	50 $\pm$ 11
Longest tumor diameter, mean $\pm$ SD, cm	2.7 $\pm$ 1.6	3.4 $\pm$ 1.5	2.6 $\pm$ 1.4	2.1 $\pm$ 1.5
Clinical stage, <i>n</i> (%)				
I	96 (13)	37 (13)	29 (14)	30 (12)
II	542 (73)	210 (74)	148 (72)	184 (73)
III	106 (14)	38 (13)	30 (14)	38 (15)

**Table 1.** *Cont.*

Characteristic	All Datasets	BL	C2	C4
T category, <i>n</i> (%)				
T1	139 (19)	54 (19)	39 (19)	46 (18)
T2	509 (68)	195 (68)	141 (68)	173 (69)
T3	83 (11)	31 (11)	23 (11)	29 (12)
T4	13 (2)	5 (2)	4 (2)	4 (2)
N category, <i>n</i> (%)				
N0	490 (66)	188 (66)	139 (67)	163 (65)
N1	171 (23)	67 (24)	44 (21)	60 (24)
N2	26 (3)	9 (3)	8 (4)	9 (4)
N3	57 (8)	21 (7)	16 (8)	20 (8)

## 2.2. Image Acquisition

The DCE-MRI images were acquired using a 3T GE 750w MR scanner (GE Healthcare, Waukesha, WI, USA) and an eight-channel bilateral phased array breast coil. The imaging protocol utilized a three-dimensional (3D) T1-weighted DISCO [39] sequence with intravenous bolus injection of contrast agent (Gadovist, Bayer HealthCare, Whippany, NJ, USA) at a rate of 2 mL/s and a dose of 0.1 mL/kg, followed with a saline flush. The imaging parameters included an acquisition matrix size of  $320 \times 320$ , a field-of-view of  $300 \times 300$  mm, a slice thickness of 3.2 mm, a slice gap of  $-1.6$  mm, a TR of 6 ms, a TE of 1.1/2.3 ms, a flip angle of  $12^\circ$ , and a temporal resolution of approximately 12 s. The number of slices ranged from 112 to 192, and the number of temporal phases ranged from 32 to 64.

## 2.3. Data Curation

DCE-MRI images and binary masks were first zero-padded along both sides of the imaging volume to 192 slices, and then the full-field-of-view images were fed for model training without any cropping. Two breast radiologists with 6 years of experience (R.M.M.) and 11 years of experience (S.P.) manually segmented the tumors in consensus using an in-house MATLAB-based software (The MathWorks, Natick, MA, USA). These manually segmented tumors are shown by the reference masks in Figures 2 and 3. The manual segmentation was performed on a subtraction image obtained by subtracting the pre-injection phase (the initial time frame) from the early phase (the frame at approximately 2.5 min after injection). Voxels of high contrast uptake between phases were identified as tumors, and voxels of signal void from biopsy clips or tumor necrosis were excluded. In cases where multiple tumors were present, only the dominant one was labeled at BL, and the same tumor was followed at C2 and C4 if images were acquired at those time points.

In addition, we calculated and used the following three semiquantitative maps: positive enhancement integral (PEI) [40], signal enhancement ratio (SER) [41], and maximum slope of increase (MSI). These maps [42] were calculated on an AW Server using the software provided by the vendor (v3.2, GE Healthcare, Milwaukee, WI, USA).

To evaluate the impact of central necrosis and biopsy clips on segmentation accuracy, we compared models with their inclusion and exclusion. Unless noted otherwise, the default segmentation used for the constructed models described below in Section 2.4 relates to models with all necrosis and clips excluded (Mk\_Excl). The tumor masks including necrosis and clips (Mk\_Incl) were generated by filling the central void voxels within Mk\_Excl automatically. Among all 285 BL cases, 38 cases were excluded as their Mk\_Incl failed to include all voxels of necrosis and clips, resulting in 247 cases for model development.

## 2.4. Automatic Segmentation Framework

We used the default original configuration of nnU-Net (Supplementary Text S1, nnU-Net model training configuration and procedure) without any major architectural changes [36]. During the model training, the input data were first preprocessed to extract the data fingerprint, which included median shapes, signal intensity distribution, spac-

ing distribution, and modality. The preprocessing steps included cropping the images to non-zero regions to improve the computational efficiency. Then, the rule-based parameters, including batch size, resample strategy, intensity normalization, and network topology, were derived through a dictionary lookup approach. The nnU-Net framework also used fixed parameters that were pre-trained from different applications and hardcoded for all new studies, including the learning rate, optimizer, number of epochs, choice of activation function, and loss function. Finally, empirical parameters were calculated on the basis of the ensemble of 2D and 3D results and the integration of the models from five-fold cross-validation, which determined the details in inference and post-processing.

A total of 10 nnU-Net models (Table 2) were constructed using different combinations of input images. The first model utilized only subtraction image data from the BL dataset and was named nnU-Net\_BL. The second, third, and fourth models were created using PEI, SER, and MSI data from the BL dataset and were named nnU-Net\_PEI, nnU-Net\_SER, and nnU-Net\_MSI, respectively. The fifth model, nnU-Net\_Comb, was generated using concatenating subtraction image, PEI, SER, and MSI data from the BL dataset to form an additional data dimension. These five models were designed to identify the most sensitive and accurate imaging metric from DCE-MRI images acquired at the same time point. The sixth and seventh models were trained using subtraction image data from the C2 and C4 datasets and were named nnU-Net\_C2 and nnU-Net\_C4, respectively. The eighth model nnU-Net was created by combining cohorts at three time points (BL, C2, and C4) and was named nnU-Net\_3tpt. The nnU-Net\_BL, nnU-Net\_C2, nnU-Net\_C4, and nnU-Net\_3tpt models were evaluated to determine the accuracy of the models at different time points. The ninth and tenth models were with exclusion and inclusion of central necrosis and biopsy clips and were named nnU-Net\_Excl and nnU-Net\_Incl, respectively.

**Table 2.** Constructed models and their inputs. Sub, subtraction image.

Model Name	Input Dataset	DCE Metrics
nnU-Net_BL	BL	Sub
nnU-Net_PEI	BL	PEI
nnU-Net_SER	BL	SER
nnU-Net_MSI	BL	MSI
nnU-Net_Comb	BL	Sub + PEI + MSI + SER
nnU-Net_C2	C2	Sub
nnU-Net_C4	C4	Sub
nnU-Net_3tpt	BL + C2 + C4	Sub
nnU-Net_Excl	BL	Sub
nnU-Net_Incl	BL	Sub

Each dataset was randomly divided into development and testing sets at a 5:1 ratio. Each development set was further split for five-fold cross-validation at a ratio of 4:1 for training and validation. The development and testing sets of nnU-Net\_3tpt were composed separately by merging corresponding sets from nnU-Net\_BL, nnU-Net\_C2, and nnU-Net\_C4.

Upon completion of training, the models were ensembled by averaging the softmax probabilities from each fold of cross-validation. The resulting ensembled model was used for inference on independent testing data. The training was performed in both 2D and 3D models. During inference, the ensemble of both 2D and 3D prediction was generated by performing voxel-wise majority vote.

All training was performed using an NVIDIA DGX1 system with dual 20-core Intel Xeon E5-2698 2.2-GHz CPUs, 512 GB of DDR4 RAM, and eight NVIDIA Tesla V100 32-GB GPUs with a total of 256 GB of GPU memory (NVIDIA, Santa Clara, CA, USA). The software

environment included an Ubuntu Linux 18.4.6 operating system, Python 3.8.12, CUDA 11.1, cuDNN 7.6.5, and TensorFlow 2.8.0.

### 2.5. Statistical Analysis

The performance of each model was evaluated using four overlap-based segmentation metrics, true positive, false negative, false positive, and true negative, using the manually labeled mask as the reference standard on a per-subject basis. The true positive, false negative, false positive, and true negative are defined as the percentage of the number of voxels within the union between the reference and predicted masks, of the number of voxels within the reference mask but outside the predicted mask, the number of voxels within the predicted mask but outside the reference mask, and the number of background voxels outside the reference mask, respectively. The DSC and sensitivity were calculated and averaged across all subjects as the metrics for overall performance [43].

Given the non-normal distribution of the results, the within-group results of subject-based DSC and sensitivity were summarized using interquartile and median calculations; the paired two-sample comparison was performed with Wilcoxon signed-rank test and the unpaired two-sample comparison was performed with Wilcoxon rank-sum test; the multiple comparison was performed using Kruskal–Wallis test. For all comparisons,  $\alpha = 0.05$  was considered the threshold for statistical significance and adjusted with Bonferroni correction for multiple comparison.

The segmentation performance was examined for primary tumors with the following largest dimensions:  $\leq 2$  cm (T1), 2–5 cm (T2), and  $\geq 5$  cm (T3–4).

## 3. Results

### 3.1. Segmentation Performance of Semiquantitative Parametric Maps

The Kruskal–Wallis test demonstrated significant differences among models with different input metrics (Figure 1). The  $\chi^2$  value was 42.1 for DSC ( $p < 0.05$ ) and 59.9 for sensitivity ( $p < 0.05$ ); post hoc-paired Wilcoxon signed-rank test was performed and adjusted at  $p = 0.05/4 = 0.0125$ . For DSC, nnU-Net\_BL was better than nnU-Net\_PEI ( $p < 0.0125$ ), nnU-Net\_SER ( $p < 0.0125$ ) and nnU-Net\_MSI ( $p < 0.0125$ ) but similar to nnU-Net\_Comb ( $p = 0.90$ ). For sensitivity, nnU-Net\_BL was better than nnU-Net\_SER ( $p < 0.0125$ ) and nnU-Net\_MSI ( $p < 0.0125$ ) but similar to nnU-Net\_PEI ( $p = 0.19$ ) and nnU-Net\_Comb ( $p = 0.58$ ).

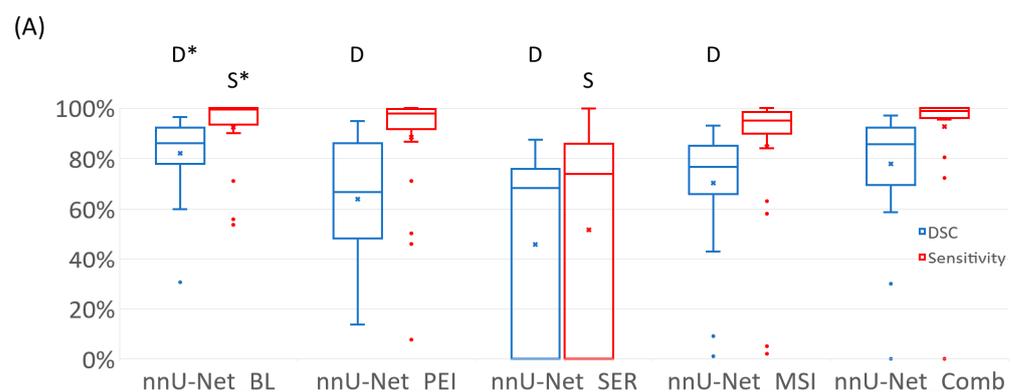


Figure 1. Cont.

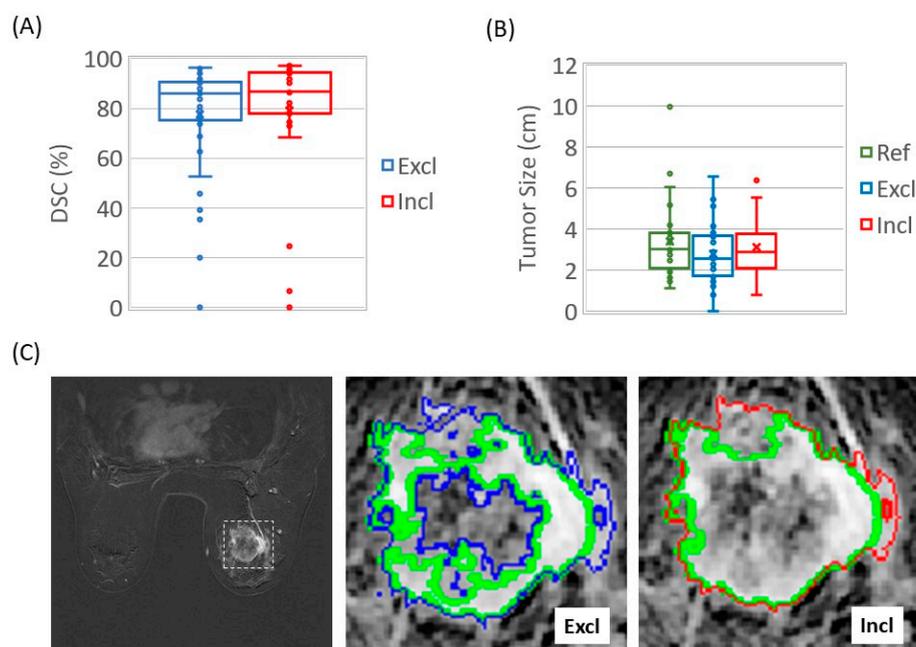
(B)

Model	DSC			Sensitivity		
	Median	Quartile		Median	Quartile	
		Q1	Q3		Q1	Q3
nnU-Net_BL	0.86	0.78	0.92	0.99	0.93	1.00
nnU-Net_PEI	0.67	0.48	0.86	0.98	0.91	1.00
nnU-Net_SER	0.68	0.00	0.76	0.74	0.00	0.86
nnU-Net_MSI	0.77	0.66	0.85	0.95	0.90	0.98
nnU-Net_Comb	0.85	0.69	0.92	0.99	0.96	1.00

**Figure 1.** Segmentation performance of nnU-Net models with different combinations of BL DCE images and semiquantitative parametric maps. The DSC and sensitivity were measured at the subject level using manually labeled masks as the reference standard and were then averaged across the BL test set. (A) The boxplots of each set of results, first and third quartiles (lower and upper ends of box, respectively), the min and max limits (whiskers) at 1.5 interquartile away from the first and third quartiles; median (horizontal line in box), mean (x), and outliers (discrete data points) were presented. (The letters above the boxplots indicated statistical significance between that metric and the reference metric, which was labeled with the same letter and an asterisk on top). (B) The detailed quantitative results used for the boxplots in (A).

### 3.2. Mask Type Comparison

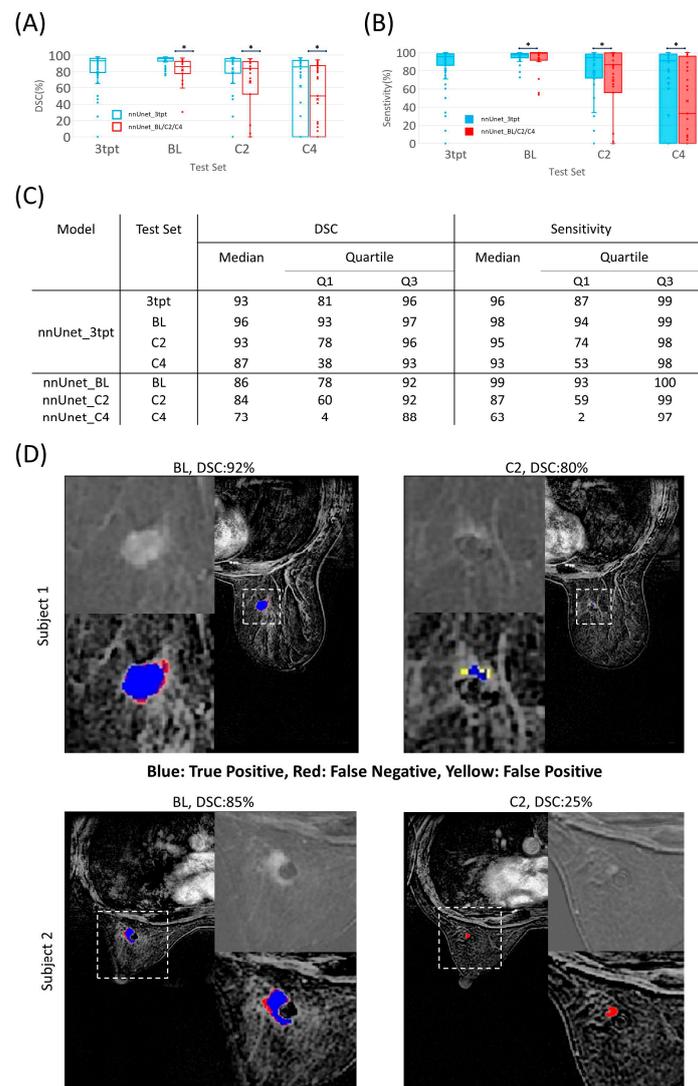
A Wilcoxon signed-rank test showed similar group DSC mean ranks ( $p = 0.17$ ) in nnUnet\_Excl (without central necrosis and biopsy clips) and nnUnet\_Incl (including central necrosis and biopsy clips) (Figure 2). The tumor sizes identified by the DL models were also similar to the tumor sizes of the references (Figure 2B). The segmentation of both models preserved details that corresponded accurately with their respective references (Figure 2C).



**Figure 2.** Automated tumor segmentation with and without inclusion of central necrosis and biopsy clips. (A) nnU-Net\_Excl and nnU-Net\_Incl on the same test cases had similar DSCs ( $p = 0.27$ ). (B) Tumor sizes based on reference masks (ref: green) were similar to those estimated with nnU-Net\_Excl ( $p = 0.14$ ) and nnU-Net\_Incl ( $p = 0.58$ ). (C) Automated masks without inclusion (Excl: blue) and with inclusion (Incl: red) of central necrosis and biopsy clips overlaying corresponding reference (green) mask of a representative subject. The subimage within the dashed box has been zoomed in and displayed as the background in the Excl and Incl images.

### 3.3. Segmentation Performance Using Datasets of Different Time Points

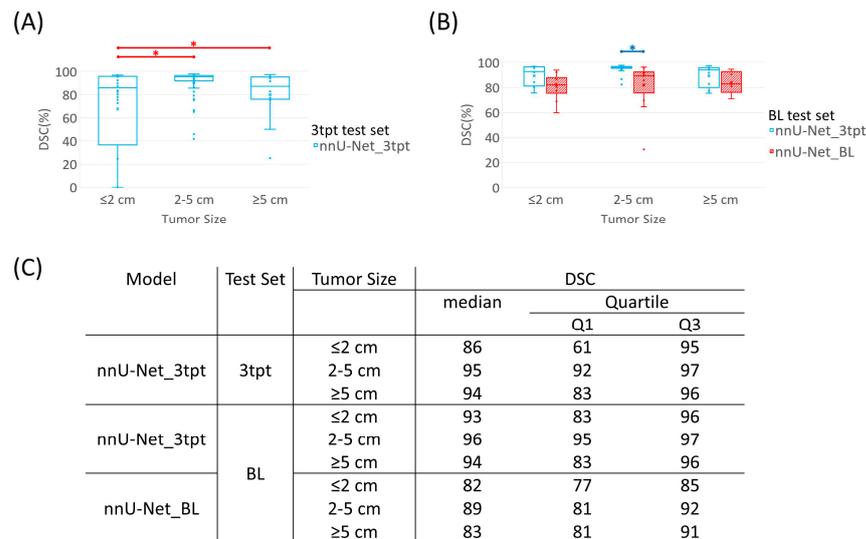
The nnU-Net\_3tpt model, which was trained on a combination of data from all three time points, BL, C2, and C4, demonstrated better segmentation performance than the models trained on data from a single time point (nnU-Net\_BL, nnU-Net\_C2, and nnU-Net\_C4) (Figure 3). The Paired Wilcoxon signed-rank test demonstrated that nnU-Net\_3tpt had a better performance than nnU-Net\_BL on the testing dataset of BL for both DSC ( $p < 0.05$ ) and sensitivity ( $p < 0.05$ ). Similarly, nnU-Net\_3tpt had a better performance than nnU-Net\_C2 on the testing dataset of C2 for both DSC ( $p < 0.05$ ) and sensitivity ( $p < 0.05$ ), and a better performance than nnU-Net\_C4 on the testing dataset of C4 for both DSC ( $p < 0.05$ ) and sensitivity ( $p < 0.05$ ). The decrease in performance over the treatment time points in two representative subjects is shown in Figure 3D.



**Figure 3.** Segmentation performance of nnU-Net models using data from various time points. DSC (A) and sensitivity (B) of the different models applied to the corresponding testing dataset. Blue bars on top of the dataset indicate significant difference in paired Wilcoxon signed-rank test ( $p < 0.05$ , indicated by black asterisks). (C) The detailed quantitative results used for the boxplots (A,B). (D) Two representative subjects and the predicted segmentation performance using the nnU-Net\_3tpt model. The reference mask is the union of blue and red masks.

The Kruskal–Wallis test demonstrated significant differences in DSC among various tumor sizes over the test set of 3tpt ( $\chi^2 = 15.7, p < 0.05$ ). The DSC had higher mean ranks

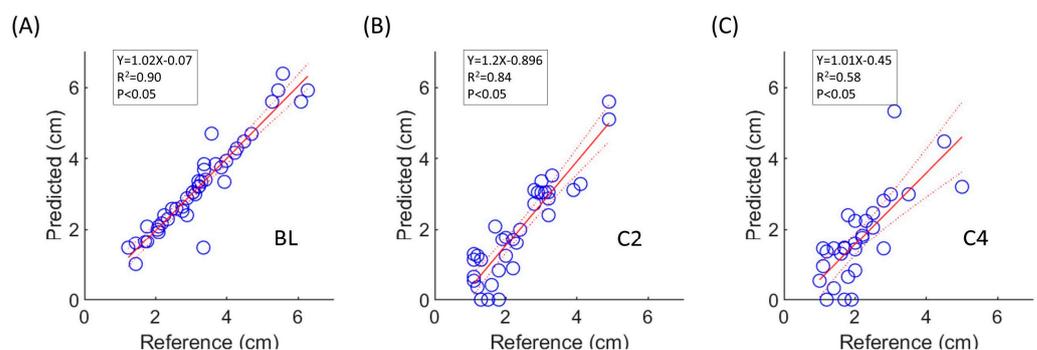
for tumors larger than 2 cm but smaller than 5 cm ( $p < 0.05$ ) in relation to tumors smaller than 2 cm. Similarly, the tumors larger than 5 cm had a higher mean rank of DSC than tumors smaller than 2 cm ( $p < 0.05$ ), as shown in Figure 4A. In the BL test set (Figure 4B), nnU-Net\_3tpt had higher mean ranks of DSC than nnU-Net\_BL ( $p < 0.05$ ) for tumors larger than 2 cm but smaller than 5 cm. For tumors that were 2 cm or smaller ( $p = 0.07$ ) and tumors at 5 cm or larger ( $p = 0.29$ ), the performance of the two models was similar. The higher mean ranks of DSC and sensitivity for nnU-Net\_3tpt than for nnU-Net\_C2 and nnU-Net\_C4 across tumor sizes are presented in Supplementary Figure S1.



**Figure 4.** Segmentation performance of nnU-Net models by tumor size. (A) DSC of nnU-Net\_3tpt across various tumor sizes in the 3tpt test set. Red bars indicate significant difference in unpaired Wilcoxon rank-sum test adjusted at  $p < 0.016$  (indicated by red asterisks). (B) DSC of nnU-Net\_3tpt and nnU-Net\_BL applied on BL test sets across various tumor sizes. Blue bar indicates statistically significant difference in paired Wilcoxon signed-rank test ( $p < 0.05$ , indicated by blue asterisks). (C) The detailed quantitative results used for the boxplots in (A,B).

### 3.4. Tumor Size Comparison

The tumor size of the predicted segmentation from nnU-Net\_3tpt was also evaluated using the reference standard (Figure 5). The intraclass correlation coefficient was 0.95 ( $p < 0.05$ ) between predicted segmentation with nnU-Net\_3tpt and the reference standard (Figure 5A), which demonstrated accurate tumor size estimation on the BL test set. The correlation coefficient was 0.88 at C2 test set ( $p < 0.05$ ) and 0.73 at C4 test set ( $p < 0.05$ ).



**Figure 5.** Comparison of tumor size between predicted segmentation using nnU-Net\_3tpt and reference tumor mask. Shown are linear relationships between tumor size of predicted segmentation and reference at BL (A), C2 (B), and C4 (C). The best-fit linear regression was represented by the solid line, while the 95% confidence interval bands were denoted by dashed lines.

#### 4. Discussion

In this study, we employed an automated deep learning framework for medical imaging segmentation, nnU-Net, in conjunction with 744 datasets to develop an accurate segmentation model specifically for TNBC under treatment. We evaluated a range of semiquantitative maps from DCE-MRI as model inputs and found that the subtraction images of the initial phases from the peak arterial phases yielded the optimal contrast for model training. Additionally, using images from multiple time points during a patient's treatment resulted in significantly better segmentation performance than using images from a single time point during treatment. This improved performance of our top-performing model likely stemmed from a greater diversity in terms of tumor size, shape, location, signal intensity patterns, and other morphological characteristics. Notably, our model accurately estimated tumor size, which is important as it is often necessary to measure changes in tumor size during treatment.

Our results indicate that DCE subtraction images provide sufficient image information to achieve good tumor segmentation with our deep learning model. To our knowledge, most reported breast cancer segmentation models for DCE-MRI use the original multiphasic images [32,44] or the subtraction between pre-contrast and post-contrast images as the input [45]. We investigated the former approach by incorporating the entire series of DCE-MRI data at BL. However, we were able to input only about 26 patient datasets during training due to the large computational memory consumption required to store 4D image series and optimize millions of model parameters. In contrast, using subtraction images effectively reduced the dimensionality of the data from 4D to 3D and required much less memory. Even though some breast cancer studies with DCE-MRI showed the use of PEI [42,46] or MSI [47,48] parametric maps for diagnosis or treatment prediction purposes, our study showed that the use of the simple subtraction images could produce a model with better or equivalent performance for tumor segmentation compared to the use of incorporating other DCE parametric maps. Further, using subtraction images for segmentation is advantageous because they are easier to generate than the other parametric maps, whose generation may require specialized software.

To the best of our knowledge, our study is the first to perform model training using datasets including images of patients from multiple treatment time points. The nnU-Net\_3tpt model trained using this approach demonstrated better performance than the model trained using data from only a single treatment time point. By combining data from three treatment time points, the training datasets were effectively tripled. Since all the model training parameters remained identical, it is possible that the increased dataset size contributed to the improved model performance. With the expanded training dataset, our nnUNet\_3tpt model exhibited performance similar to or even better than that of recent breast tumor segmentation studies. For instance, when Yue et al. evaluated model performance on a dataset of 1000 subjects ( $n_{\text{training}} = 800$ ,  $n_{\text{testing}} = 200$ ), their own model, Res-UNet, achieved a DSC of 0.894, and their implementation of nnU-Net achieved a DSC of 0.887 [45], whereas our nnU-Net\_3tpt model achieved a DSC of 0.93 in the BL test set. Other notable studies include one in which an nnU-Net trained on a training dataset of 102 subjects achieved a DSC of 0.87 (median value, mean was not reported) on a test set of 55 subjects [49]. Additionally, a regional convolutional neural network model trained on a dataset of 241 patients, including over 10,000 slices, achieved a DSC of 0.79 on a test set of 98 patients, including approximately 9000 slices, by splitting the 3D dataset into 2D space to increase dataset size [3]. A 3D U-Net model from a full dataset of 141 subjects ( $n_{\text{test}} = 30$ ) achieved a mean DSC of 0.78 [44]. The aforementioned models were developed for a variety of subtypes of breast cancer, not exclusively TNBC. In contrast, our nnU-Net\_3tpt model, which was trained exclusively on a large sample of TNBC subjects, holds significant potential for application within the TNBC population.

The segmentation models using masks with or without central necrosis and biopsy clips exhibited similar DSC and sensitivity, indicating that our models based on the nnU-Net framework are flexible and stable. To the best of our knowledge, most published

studies on breast tumor segmentation employed reference masks that included central necrosis and biopsy clips [3,28,44,50]. However, necrosis and biopsy clips may need to be excluded for certain applications, such as functional tumor volume measurements. Our findings indicate that our models can directly output both types of masks without added processing, and segmentation performance is similar with and without exclusion of central necrosis and biopsy clips. In contrast to a recent study [45], in which the tumor was segmented first and then an intensity-based method was applied to delineate a low signal intensity of necrosis within the segmented tumor region, our fully automated model provides an easier approach.

Our nnU-Net\_3tpt model had good performance but can be further improved in several aspects. First, nnU-Net\_3tpt produced increasingly better segmentation results for tumors with larger sizes, which may explain the better performance at BL than at C2 and C4 because tumors at BL are untreated and tend to be larger. A similar trend was noticed in other studies [45,49,51]. The performance of our model on smaller tumors could be improved by including a more diverse range of samples, and the generalizability of the model could be validated by including public datasets for more comprehensive training and independent testing. Second, the nnU-Net\_3tpt model failed to identify tumors smaller than 2 cm in several instances. To avoid such failures, it may be necessary to refine the nnU-Net framework configuration by modifying the loss function to prioritize false negatives. In our training, the model training loss term is guided by DSC, which emphasizes both sensitivity and precision. Other metrics for training loss may be designed to penalize false negatives with heavier weighting. DSC may not be optimal to address the signal heterogeneity of TNBC. An alternative metric for training loss is focal Dice loss, which could alleviate the imbalance between empirically defined subtypes [52]. Models that extract semantic features could integrate spatial information to improve sensitivity to tumors at smaller sizes and tissue boundaries, making it worthwhile to validate their efficacy in the TNBC population [53,54]. Finally, a systematic comparison of our model to the conventional models using the same datasets would better evaluate our model. In addition to the static imaging features used in our study, integrating tumor-specific dynamic information into the nnU-Net framework could also help to reduce false positives [50].

## 5. Conclusions

We developed a fully automated, high-performance segmentation model for TNBC patients using deep learning and a large cohort of DCE-MRI images acquired longitudinally over the patients' treatment course. Of the various types of images used for model training, we found that the simple subtraction images had the best performance. Our model was also capable of reliably segmenting tumors with either exclusion or inclusion of the central necrosis and biopsy clips. The performance of our model, especially for small tumors, may be further improved in a future investigation.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers15194829/s1>, references [36,55]. Figure S1: Segmentation performance of nnU-Net models by tumor size across different datasets; Text S1: nnU-Net model training configuration and procedure.

**Author Contributions:** Conceptualization, Z.X. and J.M.; data curation, B.E.A., R.P.C., A.C. and B.R.; formal analysis, Z.X., D.E.R., R.M.M., S.P., Z.Z., B.P., J.B.S., H.C., J.S. and J.M.; investigation, Z.X. and J.M.; methodology, Z.X., D.E.R., R.M.M., S.P., Z.Z., B.P., J.B.S., B.C.M., B.E.A., R.P.C. and J.M.; supervision, J.M.; writing—original draft preparation, Z.X.; writing—review and editing, Z.X., K.-P.H., B.E.A., R.P.C., J.W.T.L., H.T.C.L.-P., D.L.L., F.P., J.W., P.W., A.T., A.K., L.H., K.K.H., J.K.L., V.V., D.T., W.Y., C.Y. and J.M.; project administration, A.C. and B.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the NIH/NCI under award number P30CA016672.

**Institutional Review Board Statement:** This study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of The University of Texas M. D. Anderson Cancer Center (protocol code: 2014-0185 and date of approval: 16 July 2014).

**Informed Consent Statement:** Written informed consent was received from all patients.

**Data Availability Statement:** Data are available from the corresponding author upon request.

**Acknowledgments:** The authors acknowledge the support from the M.D. Anderson Moon Shots Program and the Robert D. Moreton Distinguished Chair Funds in Diagnostic Radiology. The authors thank Stephanie Deming, senior scientific editor, Research Medical Library, for editing the manuscript.

**Conflicts of Interest:** The authors would like to make the following disclosures:

- K.K.H. serves on the Medical Advisory Board for ArmadaHealth, AstraZeneca, and receives research funding from Cairn Surgical, Eli Lilly&Co., and Lumicell.
- K.H. is currently receiving research funding from Siemens Healthineers and has received research funding from GE.
- J.K.L. received grant or research support from Novartis, Medivation/Pfizer, Genentech, GSK, EMD-Serono, AstraZeneca, Medimmune, Zenith, Merck; participated in Speaker’s Bureau for MedLearning, Physician’s Education Resource, Prime Oncology, Medscape, Clinical Care Options, Medpage; and receives royalty from UpToDate.
- Spouse of A.T works for Eli Lilly.
- D.T. declares research contracts with Pfizer, Novartis, and Ployphor and is a consultant of AstraZeneca, GlaxoSmithKline, OncoPep, Gilead, Novartis, Pfizer, Personalis, and Sermonix.
- W.Y. receives royalties from Elsevier.
- J.M. is a consultant of C4 Imaging, L.L.C., and an inventor of United States patents licensed to Siemens Healthineers and GE Healthcare.
- For the remaining authors, none were declared.

The funders had no role in the design of this study; in the collection, analyses, or interpretation of data; in the writing of manuscript; or in the decision to publish results.

## References

1. Wu, Q.; Siddharth, S.; Sharma, D. Triple Negative Breast Cancer: A Mountain Yet to Be Scaled Despite the Triumphs. *Cancers* **2021**, *13*, 3697. [[CrossRef](#)]
2. Dent, R.; Trudeau, M.; Pritchard, K.I.; Hanna, W.M.; Kahn, H.K.; Sawka, C.A.; Lickley, L.A.; Rawlinson, E.; Sun, P.; Narod, S.A. Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. *Clin. Cancer Res.* **2007**, *13*, 4429–4434. [[CrossRef](#)]
3. Zhang, Y.; Chan, S.; Park, V.Y.; Chang, K.-T.; Mehta, S.; Kim, M.J.; Combs, F.J.; Chang, P.; Chow, D.; Parajuli, R.; et al. Automatic Detection and Segmentation of Breast Cancer on MRI Using Mask R-CNN Trained on Non-Fat-Sat Images and Tested on Fat-Sat Images. *Acad. Radiol.* **2022**, *29*, S135–S144. [[CrossRef](#)]
4. El Adoui, M.; Drisis, S.; Benjelloun, M. Multi-input deep learning architecture for predicting breast tumor response to chemotherapy using quantitative MR images. *Int. J. Comput. Assist. Radiol. Surg.* **2020**, *15*, 1491–1500. [[CrossRef](#)]
5. Ha, R.; Chin, C.; Karcich, J.; Liu, M.Z.; Chang, P.; Mutasa, S.; Van Sant, E.P.; Wynn, R.T.; Connolly, E.; Jambawalikar, S. Prior to Initiation of Chemotherapy, Can We Predict Breast Tumor Response? Deep Learning Convolutional Neural Networks Approach Using a Breast MRI Tumor Dataset. *J. Digit. Imaging* **2018**, *32*, 693–701. [[CrossRef](#)]
6. Kohli, M.D.; Summers, R.M.; Geis, J.R. Medical Image Data and Datasets in the Era of Machine Learning—Whitepaper from the 2016 C-MIMI Meeting Dataset Session. *J. Digit. Imaging* **2017**, *30*, 392–399. [[CrossRef](#)]
7. Kupinski, M.A.; Giger, M.L. Automated seeded lesion segmentation on digital mammograms. *IEEE Trans. Med. Imaging* **1998**, *17*, 510–517. [[CrossRef](#)]
8. Yuan, Y.; Giger, M.L.; Li, H.; Suzuki, K.; Sennett, C. A dual-stage method for lesion segmentation on digital mammograms. *Med. Phys.* **2007**, *34*, 4180–4193. [[CrossRef](#)]
9. Horsch, K.; Giger, M.L.; Venta, L.A.; Vyborny, C.J. Automatic segmentation of breast lesions on ultrasound. *Med. Phys.* **2001**, *28*, 1652–1659. [[CrossRef](#)]
10. Rojas Domínguez, A.; Nandi, A.K. Detection of masses in mammograms via statistically based enhancement, multilevel-thresholding segmentation, and region selection. *Comput. Med. Imaging Graph.* **2008**, *32*, 304–315. [[CrossRef](#)]
11. Pereira, D.C.; Ramos, R.P.; do Nascimento, M.Z. Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm. *Comput. Methods Programs Biomed.* **2014**, *114*, 88–101. [[CrossRef](#)] [[PubMed](#)]
12. Timp, S.; Karssemeijer, N. A new 2D segmentation method based on dynamic programming applied to computer aided detection in mammography. *Med. Phys.* **2004**, *31*, 958–971. [[CrossRef](#)] [[PubMed](#)]
13. Petrick, N.; Chan, H.P.; Sahiner, B.; Helvie, M.A. Combined adaptive enhancement and region-growing segmentation of breast masses on digitized mammograms. *Med. Phys.* **1999**, *26*, 1642–1654. [[CrossRef](#)] [[PubMed](#)]

14. Huang, Q.; Luo, Y.; Zhang, Q. Breast ultrasound image segmentation: A survey. *Int. J. Comput. Assist. Radiol. Surg.* **2017**, *12*, 493–507. [[CrossRef](#)]
15. Hu, Y.; Guo, Y.; Wang, Y.; Yu, J.; Li, J.; Zhou, S.; Chang, C. Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model. *Med. Phys.* **2019**, *46*, 215–228. [[CrossRef](#)]
16. Al-Antari, M.A.; Al-Masni, M.A.; Choi, M.T.; Han, S.M.; Kim, T.S. A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. *Int. J. Med. Inform.* **2018**, *117*, 44–54. [[CrossRef](#)]
17. Kooi, T.; Litjens, G.; van Ginneken, B.; Gubern-Merida, A.; Sanchez, C.I.; Mann, R.; den Heeten, A.; Karssemeijer, N. Large scale deep learning for computer aided detection of mammographic lesions. *Med. Image Anal.* **2017**, *35*, 303–312. [[CrossRef](#)]
18. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Munich, Germany, 5–9 October 2015; pp. 234–241.
19. Baccouche, A.; Garcia-Zapirain, B.; Castillo Olea, C.; Elmaghraby, A.S. Connected-UNets: A deep learning architecture for breast mass segmentation. *NPJ Breast Cancer* **2021**, *7*, 151. [[CrossRef](#)]
20. Balkenende, L.; Teuwen, J.; Mann, R.M. Application of Deep Learning in Breast Cancer Imaging. *Semin. Nucl. Med.* **2022**, *52*, 584–596. [[CrossRef](#)]
21. Kuhl, C.K.; Mielcareck, P.; Klaschik, S.; Leutner, C.; Wardelmann, E.; Gieseke, J.; Schild, H.H. Dynamic breast MR imaging: Are signal intensity time course data useful for differential diagnosis of enhancing lesions? *Radiology* **1999**, *211*, 101–110. [[CrossRef](#)]
22. Agner, S.C.; Xu, J.; Fatakdawala, H.; Ganesan, S.; Madabhushi, A.; Englander, S.; Rosen, M.; Thomas, K.; Schnall, M.; Feldman, M.; et al. Segmentation and classification of triple negative breast cancers using DCE-MRI. In Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Boston, MA, USA, 28 June–1 July 2009; pp. 1227–1230.
23. Woods, B.J.; Clymer, B.D.; Kurc, T.; Heverhagen, J.T.; Stevens, R.; Orsdemir, A.; Bulan, O.; Knopp, M.V. Malignant-lesion segmentation using 4D co-occurrence texture analysis applied to dynamic contrast-enhanced magnetic resonance breast image data. *J. Magn. Reson. Imaging* **2007**, *25*, 495–501. [[CrossRef](#)]
24. Aljabar, P.; Heckemann, R.A.; Hammers, A.; Hajnal, J.V.; Rueckert, D. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage* **2009**, *46*, 726–738. [[CrossRef](#)]
25. Wang, H.; Yushkevich, P.A. Multi-atlas segmentation without registration: A supervoxel-based approach. In Proceedings of the Medical image computing and computer-assisted intervention: MICCAI International Conference on Medical Image Computing and Computer-Assisted Intervention 2013, Nagoya, Japan, 22–26 September 2013; Volume 16, pp. 535–542.
26. Chen, W.; Giger, M.L.; Bick, U. A fuzzy c-means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. *Acad. Radiol.* **2006**, *13*, 63–72. [[CrossRef](#)]
27. Keller, B.M.; Nathan, D.L.; Wang, Y.; Zheng, Y.; Gee, J.C.; Conant, E.F.; Kontos, D. Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation. *Med. Phys.* **2012**, *39*, 4903–4917. [[CrossRef](#)]
28. Zhang, J.; Saha, A.; Zhu, Z.; Mazurowski, M.A. Hierarchical Convolutional Neural Networks for Segmentation of Breast Tumors in MRI With Application to Radiogenomics. *IEEE Trans. Med. Imaging* **2019**, *38*, 435–447. [[CrossRef](#)]
29. Dalmis, M.U.; Litjens, G.; Holland, K.; Setio, A.; Mann, R.; Karssemeijer, N.; Gubern-Merida, A. Using deep learning to segment breast and fibroglandular tissue in MRI volumes. *Med. Phys.* **2017**, *44*, 533–546. [[CrossRef](#)]
30. Kamnitsas, K.; Ledig, C.; Newcombe, V.F.J.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; Rueckert, D.; Glocker, B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **2017**, *36*, 61–78. [[CrossRef](#)]
31. El Adoui, M.; Mahmoudi, S.A.; Larhmam, M.A.; Benjelloun, M. MRI Breast Tumor Segmentation Using Different Encoder and Decoder CNN Architectures. *Computers* **2019**, *8*, 52. [[CrossRef](#)]
32. Hirsch, L.; Huang, Y.; Luo, S.; Saccarelli, C.R.; Gullo, R.L.; Naranjo, I.D.; Bitencourt, A.G.V.; Onishi, N.; Ko, E.S.; Leithner, D.; et al. Radiologist-Level Performance by Using Deep Learning for Segmentation of Breast Cancers on MRI Scans. *Radiol. Artif. Intell.* **2021**, *15*, e200231. [[CrossRef](#)]
33. Zhang, L.; Mohamed, A.A.; Chai, R.; Guo, Y.; Zheng, B.; Wu, S. Automated deep learning method for whole-breast segmentation in diffusion-weighted breast MRI. *J. Magn. Reson. Imaging* **2019**, *51*, 635–643. [[CrossRef](#)]
34. Chen, X.; Men, K.; Chen, B.; Tang, Y.; Zhang, T.; Wang, S.; Li, Y.; Dai, J. CNN-Based Quality Assurance for Automatic Segmentation of Breast Cancer in Radiotherapy. *Front. Oncol.* **2020**, *10*, 524. [[CrossRef](#)]
35. Gao, J.; Zhong, X.; Li, W.; Li, Q.; Shao, H.; Wang, Z.; Dai, Y.; Ma, H.; Shi, Y.; Zhang, H.; et al. Attention-based Deep Learning for the Preoperative Differentiation of Axillary Lymph Node Metastasis in Breast Cancer on DCE-MRI. *J. Magn. Reson. Imaging* **2023**, *57*, 1842–1853. [[CrossRef](#)]
36. Isensee, F.; Jaeger, P.F.; Kohl, S.A.A.; Petersen, J.; Maier-Hein, K.H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **2021**, *18*, 203–211. [[CrossRef](#)]
37. Antonelli, M.; Reinke, A.; Bakas, S.; Farahani, K.; Kopp-Schneider, A.; Landman, B.A.; Litjens, G.; Menze, B.; Ronneberger, O.; Summers, R.M.; et al. The Medical Segmentation Decathlon. *Nat. Commun.* **2022**, *13*, 4128. [[CrossRef](#)]
38. Panthi, B.; Adrada, B.E.; Candelaria, R.P.; Guirguis, M.S.; Yam, C.; Boge, M.; Chen, H.; Hunt, K.K.; Huo, L.; Hwang, K.-P.; et al. Assessment of Response to Neoadjuvant Systemic Treatment in Triple-Negative Breast Cancer Using Functional Tumor Volumes from Longitudinal Dynamic Contrast-Enhanced MRI. *Cancers* **2023**, *15*, 1025. [[CrossRef](#)]

39. Saranathan, M.; Rettmann, D.W.; Hargreaves, B.A.; Clarke, S.E.; Vasanawala, S.S. Differential subsampling with cartesian ordering (DISCO): A high spatio-temporal resolution dixon imaging sequence for multiphase contrast enhanced abdominal imaging. *J. Magn. Reson. Imaging* **2012**, *35*, 1484–1492. [[CrossRef](#)]
40. Khat, A.; Gianfelice, D.; Amara, M.; Boulanger, Y. Influence of post-treatment delay on the evaluation of the response to focused ultrasound surgery of breast cancer by dynamic contrast enhanced MRI. *Br. J. Radiol.* **2006**, *79*, 308–314. [[CrossRef](#)]
41. Yang, W.; Qiang, J.W.; Tian, H.P.; Chen, B.; Wang, A.J.; Zhao, J.G. Multi-parametric MRI in cervical cancer: Early prediction of response to concurrent chemoradiotherapy in combination with clinical prognostic factors. *Eur. Radiol.* **2017**, *28*, 437–445. [[CrossRef](#)]
42. Zhou, Z.; Adrada, B.E.; Candelaria, R.P.; Elshafeey, N.A.; Boge, M.; Mohamed, R.M.; Pashapoor, S.; Sun, J.; Xu, Z.; Panthi, B.; et al. Prediction of pathologic complete response to neoadjuvant systemic therapy in triple negative breast cancer using deep learning on multiparametric MRI. *Sci. Rep.* **2023**, *13*, 1171. [[CrossRef](#)]
43. Taha, A.A.; Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med. Imaging* **2015**, *15*, 29. [[CrossRef](#)]
44. Rahimpour, M.; Saint Martin, M.J.; Frouin, F.; Akl, P.; Orlhac, F.; Koole, M.; Malhaire, C. Visual ensemble selection of deep convolutional neural networks for 3D segmentation of breast tumors on dynamic contrast enhanced MRI. *Eur. Radiol.* **2023**, *33*, 959–969. [[CrossRef](#)]
45. Yue, W.; Zhang, H.; Zhou, J.; Li, G.; Tang, Z.; Sun, Z.; Cai, J.; Tian, N.; Gao, S.; Dong, J.; et al. Deep learning-based automatic segmentation for size and volumetric measurement of breast cancer on magnetic resonance imaging. *Front. Oncol.* **2022**, *12*, 984626. [[CrossRef](#)]
46. Dogan, B.E.; Turnbull, L.W. Imaging of triple-negative breast cancer. *Ann. Oncol.* **2012**, *23*, vi23–vi29. [[CrossRef](#)]
47. Milon, A.; Vande Perre, S.; Poujol, J.; Trop, I.; Kermarrec, E.; Bekhouche, A.; Thomassin-Naggara, I. Abbreviated breast MRI combining FAST protocol and high temporal resolution (HTR) dynamic contrast enhanced (DCE) sequence. *Eur. J. Radiol.* **2019**, *117*, 199–208. [[CrossRef](#)]
48. Onishi, N.; Sadinski, M.; Hughes, M.C.; Ko, E.S.; Gibbs, P.; Gallagher, K.M.; Fung, M.M.; Hunt, T.J.; Martinez, D.F.; Shukla-Dave, A.; et al. Ultrafast dynamic contrast-enhanced breast MRI may generate prognostic imaging markers of breast cancer. *Breast Cancer Res.* **2020**, *22*, 58. [[CrossRef](#)]
49. Janse, M.H.A.; Janssen, L.M.; van der Velden, B.H.M.; Moman, M.R.; Wolters-van der Ben, E.J.M.; Kock, M.; Viergever, M.A.; van Diest, P.J.; Gilhuijs, K.G.A. Deep Learning-Based Segmentation of Locally Advanced Breast Cancer on MRI in Relation to Residual Cancer Burden: A Multi-Institutional Cohort Study. *J. Magn. Reson. Imaging* **2023**. *online ahead of print.* [[CrossRef](#)]
50. Wang, S.; Sun, K.; Wang, L.; Qu, L.; Yan, F.; Wang, Q.; Shen, D. Breast Tumor Segmentation in DCE-MRI With Tumor Sensitive Synthesis. *IEEE Trans. Neural Networks Learn. Syst.* **2021**, *34*, 4990–5001. [[CrossRef](#)]
51. Zhou, Z.; Sanders, J.W.; Johnson, J.M.; Gule-Monroe, M.; Chen, M.; Briere, T.M.; Wang, Y.; Son, J.B.; Pagel, M.D.; Ma, J.; et al. MetNet: Computer-aided segmentation of brain metastases in post-contrast T1-weighted magnetic resonance imaging. *Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol.* **2020**, *153*, 189–196. [[CrossRef](#)]
52. Zhao, R.; Qian, B.; Zhang, X.; Li, Y.; Wei, R.; Liu, Y.; Pan, Y. Rethinking Dice Loss for Medical Image Segmentation. In Proceedings of the 2020 IEEE International Conference on Data Mining (ICDM), Sorrento, Italy, 17–20 November 2020; pp. 851–860.
53. Li, Y.; Han, G.; Liu, X. DCNet: Densely Connected Deep Convolutional Encoder-Decoder Network for Nasopharyngeal Carcinoma Segmentation. *Sensors* **2021**, *21*, 7877. [[CrossRef](#)]
54. Zhang, H.; Gao, Z.; Zhang, D.; Hau, W.K.; Zhang, H. Progressive Perception Learning for Main Coronary Segmentation in X-Ray Angiography. *IEEE Trans. Med. Imaging* **2023**, *42*, 864–879. [[CrossRef](#)]
55. Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.