IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 14 May 2023, accepted 26 May 2023, date of publication 31 May 2023, date of current version 7 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3281558

RESEARCH ARTICLE

TS-Net: A Deep Learning Framework for Automated Assessment of Longitudinal Tumor Volume Changes in an Orthotopic Breast Cancer Model Using MRI

YUNKYOUNG JUN[1], SEOKHA JIN[1], NOEHYUN MYUNG[1], JIWOO JEONG[1], JIMIN LEE[2,3], AND HYUNG JOON CHO[1]

[1]Department of Biomedical Engineering, Ulsan National Institute of Science and Technology (UNIST), Eonyang-eup, Ulju-gun, Ulsan 44919, Republic of Korea
[2]Department of Nuclear Engineering, Ulsan National Institute of Science and Technology (UNIST), Eonyang-eup, Ulju-gun, Ulsan 44919, Republic of Korea
[3]Graduate School of Artificial Intelligence, Ulsan National Institute of Science and Technology (UNIST), Eonyang-eup, Ulju-gun, Ulsan 44919, Republic of Korea

Corresponding authors: Jimin Lee (jiminlee@unist.ac.kr) and Hyung Joon Cho (hjcho@unist.ac.kr)

This work was supported in part by the Korea Health Technology Research and Development Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare, Republic of Korea, under Grant HI21C1161; in part by the Institute of Information & Communications Technology Planning & Evaluation (IITP) funded by the Ministry of Science and ICT (MSIT), Republic of Korea, under Grant 2020-0-01336; in part by the Artificial Intelligence Graduate School Program (UNIST); and in part by the National Research Foundation of Korea (NRF) funded by MSIT under Grant NRF-2021R1F1A1057818 and Grant NRF-2022R1A2C201119112.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Animal Care and Use Committee (IACUC) of the Ulsan National Institute of Science and Technology.

ABSTRACT Monitoring tumor volume changes in response to therapeutic agents is a critical step in preclinical drug development. Here, an automated magnetic resonance imaging (MRI)-based approach is proposed using a deep learning framework for tracking longitudinal tumor volume changes in an orthotopic breast cancer model treated with chemotherapy. Longitudinal magnetic resonance images are employed to track changes in tumor volume over time, using an untreated group and a doxorubicin-treated group as the dataset to evaluate treatment effects. Our approach, called Tumor Segmentation-Net (TS-Net), involves replacing the encoder of U-Net with a pre-trained ResNet34 to improve performance. The model was trained using a sample size of n=19 from the untreated group and then subsequently assessed on both the untreated group (n=5) and treated group (n=6). The correlation between the tumor volume determined from the ground truth and that obtained from the trained output was strong (R^2=0.984, slope=0.996). These results can lead to automated three-dimensional visualization of different longitudinal volume changes with and without treatment. Notably, for small tumors with volumes between 2 and 5 mm^3, the proposed TS-Net demonstrated an average Dice similarity coefficient score of 0.85, indicating the ability to reliably detect early tumors that may often be missed. Our approach offers a promising tool for preclinical evaluation of tumor volume changes and treatment efficacy in animal models.

INDEX TERMS Tumor segmentation, deep learning, longitudinal MR imaging, orthotopic breast cancer model, therapeutic effect.

I. INTRODUCTION

Breast cancer, a prevalent and potentially fatal disease [1], [2], highlights the importance of developing and optimizing chemotherapeutic agents. Preclinical studies play a crucial role in the development of new drugs, allowing for the identification of agents with minimal side effects in humans [3]. By helping to ensure drug safety and efficacy before clinical trials, preclinical studies expedite the drug development process and potentially improve cancer treatment outcomes [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Carmelo Militello.

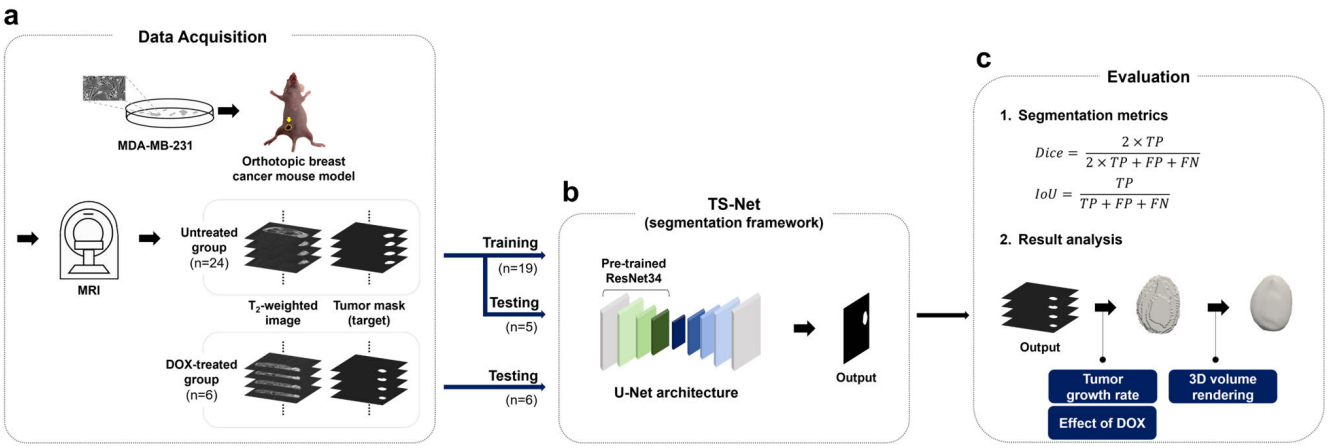

FIGURE 1. Research process depicted in three main stages: (a) data acquisition, (b) TS-Net, and (c) evaluation. The trained TS-Net's segmentation outcomes are assessed using Dice and IoU metrics, and the results are depicted through tumor volume growth rate graphs, effect of doxorubicin (DOX) graphs, and 3D volume renderings.

Therefore, orthotopic rodent models are commonly used to simulate realistic tumor growth and metastasis patterns in humans [5]. These models closely replicate primary human tumors by transplanting tumor cells into the organ of origin in rodents [6].

In orthotopic models, it is important to monitor tumor progression longitudinally following treatment to support the determination of the therapeutic effects of new agents [7]. Specifically, tumor volume reduction is directly related to the response to therapy, making magnetic resonance imaging (MRI) scans a valuable tool for repeatable and accurate monitoring of tumor volume changes at multiple time points [8], [9]. MRI-based approaches are especially useful for investigating therapeutic effects on small tumors in rodent models, which can be quantified using longitudinal analysis of tumor volume changes. However, manual segmentation of conventional MR images for orthotopic tumor models still requires expert intervention [10]. Several studies have investigated the sensitivity of MRI in detecting small tumors, showing that the detection of small tumors below 5–10 mm^3 by MRI remains challenging [11], [12].

In order to facilitate automated segmentation of tumors, deep learning-based methods are currently being extensively studied [13], [14]. Convolutional neural networks (CNNs) have demonstrated exceptional performance in this regard, with the U-Net architecture [15], a CNN variant, being successfully applied to medical image segmentation [16]. U-Net^{++} [17] is an extension of U-Net that uses a nested structure to improve feature extraction and introduces some additional techniques such as deep supervision to improve performance in semantic segmentation tasks. To address the complexity of hyperparameter tuning, nnU-Net [18] has been reported. This method is capable of automatically optimizing itself for a specific medical segmentation dataset without requiring any manual hyperparameter tuning. Studies have shown nnU-Net to be highly effective, setting a new state-of-the-art on medical datasets [18], [19]. Another promising architecture is ResNet [20], which has been widely adopted and extended for various computer vision tasks, including object detection and segmentation [21], [22]. Although the potential exists to utilize longitudinal medical images to generate deep learning–based models, relatively few studies have been conducted due to the unique characteristics of medical images. As a result, applications of these neural networks have remained limited, particularly in tracking longitudinal therapeutic responses in orthotopic rodent models, especially during the early stages of tumor development.

The aim of this study is to optimize a deep learning-based tumor segmentation framework, called TS-Net, for an orthotopic breast cancer model using longitudinal magnetic resonance (MR) images. The framework is designed to effectively segment small tumors, track tumor growth, and monitor the efficacy of therapeutic interventions. To achieve these goals, a hypothesis was formulated that training a deep learning model exclusively with MR images of an untreated group would lead to a segmentation framework that can demonstrate proficient performance with both untreated and chemotherapy-treated groups. Such a framework could enable comprehensive investigations targeting the chemotherapeutic agent doxorubicin (DOX) in an orthotopic breast cancer model. Accordingly, longitudinal MR images of an untreated group were utilized for the purpose of training the deep learning model, after which the model performance was evaluated for both untreated and DOX-treated groups. Our results showed that the framework can automatically evaluate three-dimensional (3D) tumor volumes at each time point, enabling the identification of differences in tumor growth rates between untreated and treated groups. This process is depicted in Fig. 1. Our study provides a promising tool for more accurate and efficient analysis of orthotopic breast cancer models using longitudinal MR imaging.

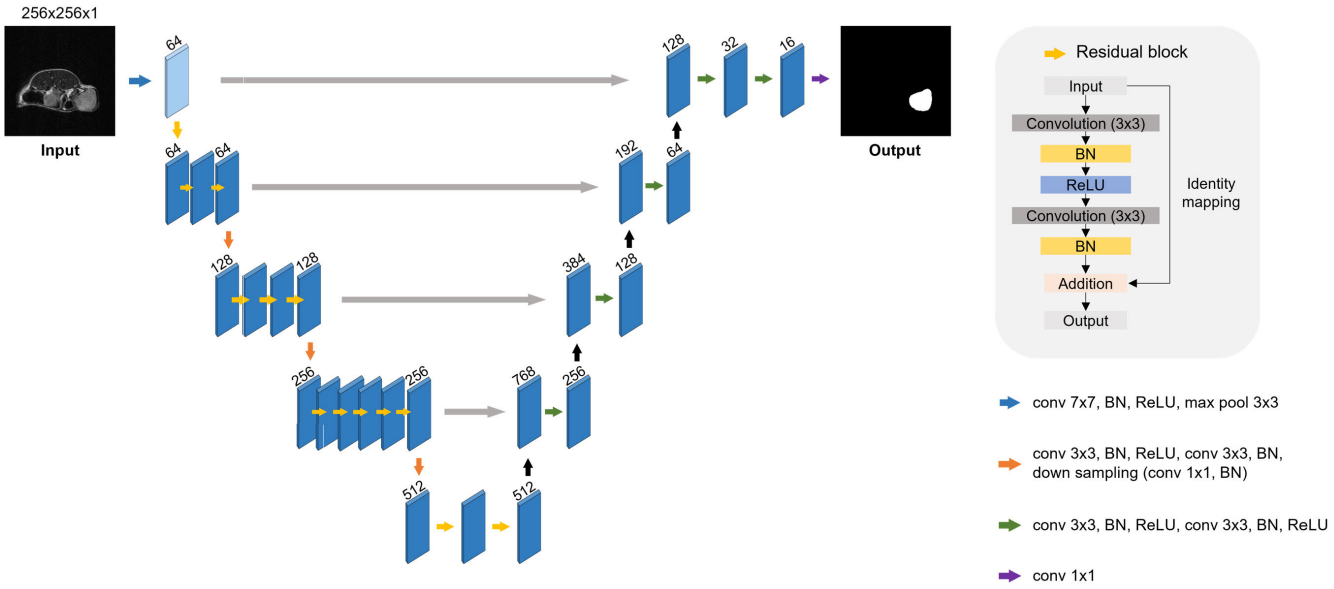

FIGURE 2. Architecture of the proposed TS-Net.

II. MATERIALS AND METHODS

A. PREPARATION OF CELLS

The MDA-MB-231 line (HTB-26; ATCC, VA, USA), a well-established human triple-negative breast carcinoma cell line commonly employed in anticancer research to evaluate chemosensitivity to DOX, was utilized to generate the orthotopic breast cancer model. The cells were maintained in Dulbecco's Modified Eagle Medium (DMEM; High Glucose, Pyruvate; Gibco, Thermo Fisher Scientific, MA, USA) supplemented with 10% fetal bovine serum (FBS-22A; Capricorn Scientific, Ebsdorfergrund, Germany) and 1% penicillin-streptomycin (PS-B; Capricorn Scientific, Ebsdorfergrund, Germany) and incubated at 37°C under 5% CO_2.

B. ESTABLISHMENT OF THE ORTHOTOPIC BREAST CANCER MODEL

MDA-MB-231 cells (5×10^6) were suspended in 50 μL of phosphate-buffered saline (PBS, pH 7.4; Gibco, Thermo Fisher Scientific, MA, USA) and mixed with 50 μL of Matrigel (Cat No. 354234; Corning Inc., Corning, NY, USA) in a 1:1 ratio [23]. The cell mixture was injected into the inguinal right mammary fat pad of female CAnN.Cg-Foxn1nu/Crl mice (6 weeks old; Orient Bio, Seongnam, Republic of Korea) [24]. To monitor the growth of tumors, their volumes were estimated twice a week by measuring their length and width using a vernier caliper [25], [26].

C. DATA ACQUISITION

When tumor volumes reached 75 mm^3 [27], the orthotopic breast cancer model was divided into two groups, untreated and treated, and the in vivo study (MRI imaging) was started. At the same time, the treated group started to receive a weekly intravenous injection of 0.1 mg/20 g DOX (HY-15142; MedChemExpress, Monmouth Junction, NJ, USA) through the tail vein. All experimental procedures were performed with the approval of the Institutional Animal Care and Use Committee (IACUC) of the Ulsan National Institute of Science and Technology.

To obtain a dataset for developing the deep learning model, subjects from both untreated and treated groups were scanned using a non-invasive 7T MRI scanner (Pharmascan 7T; Bruker BioSpin, Ettlingen, Germany) every 5 days for a total of 30 days. All subjects from both groups were given 1–2% isoflurane anesthesia in a mixture of O_2/N_2O (3:7) during the scanning process. T_2-weighted axial images were obtained using the rapid acquisition with relaxation enhancement (RARE) pulse sequence, with the MR parameters echo time (TE), repetition time (TR), RARE factor, matrix size, field of view (FOV), slice thickness, and number of slices set to 35 ms, 4000 ms, 4, 256 × 256, 35 × 35 mm^2, 0.5 mm, and 50, respectively [27]. Ground truth masks of the breast tumor were generated through manual binary annotation by two preclinical MRI researchers. From the 50 MR slices obtained from the RARE pulse sequence per subject, only those including tumors were included in the dataset.

D. NETWORK ARCHITECTURE

This study proposes an approach called TS-Net, which has the U-Net architecture with a ResNet34 pre-trained on ImageNet. ResNet34 is a deep neural network with 34 layers of residual blocks and identity mappings that provide a more direct connection between the input and output compared with the plain U-Net. Each residual block consists of two 3 × 3 convolutional blocks, batch normalization (BN),

TABLE 1. Details of the dataset used in this study. The training set comprised 19 subjects from the untreated group. In order to perform five-fold cross-validation, the train and validation sets were separated at the subject level. The test set included 5 subjects from the untreated group and 6 subjects from the DOX-treated group.

	Train	Test	
	Untreated group	Untreated group	Dox-treated group
No. of subjects	19	5	6
No. of MR slices	1979	531	579

a rectified linear unit (ReLU), and an identity mapping to effective address the vanishing gradient problem [28], [29], [30], [31]. To optimize the TS-Net architecture, the U-Net encoder is replaced with a pre-trained ResNet34, which increases the number of feature maps from 64 to 512. At the final layer of the network, a 1×1 convolution with sigmoid activation is applied to generate the segmentation map. Fig. 2 displays the architecture of TS-Net. The code and dataset of TS-Net can be found via the following link: https://github.com/ykj97/TS-Net.

E. TRAINING

Table 1 lists the details of the dataset used in this study. A total of 1979 two-dimensional (2D) MR slices of the untreated group (19 subjects) were utilized as the training set. Among the 19 subjects in the untreated group, the train and validation sets were divided at the subject level at an 80:20 ratio for five-fold cross-validation. As the proportion of tumors in individual slices of MR images is relatively low, the MRI dataset of this work exhibited class imbalance. To address this issue during the training of our deep learning segmentation model, focal loss was implemented as the loss function, which is specifically designed to deal with class imbalance [32], [33]. The parameters of focal loss, namely α and γ, were set to 0.75 and 2, respectively. The adaptive moment estimation (ADAM) optimizer was also utilized to update the model weights [34]. Additionally, the fixed learning rate was set to 0.0003, epoch to 500, and batch to 16. To prevent overfitting, data augmentation was performed on the MR slices of the training set during training to enhance the quality and size of the data [35]. Data augmentation was randomly applied to 30% of the total training set using the methods of center crop (240×240), resize (256×256), elastic transform, contrast-limited adaptive histogram equalization (CLAHE), horizontal flip, and vertical flip [36]. TS-Net takes 2D breast tumor images (256×256) as input and produces a corresponding 2D segmentation map (256×256).

To verify TS-Net's performance, various network architectures were utilized. U-Net was used as the baseline model, and variations of U-Net, namely U-Net^{++}, original and pre-activation residual block, ResNet34, ResNet50, and nnU-Net were applied to confirm competitive performance. All networks were trained and evaluated using the same dataset distribution and environment. Also, all networks except nnU-Net were trained with the same hyperparameters. The architecture was implemented using Pytorch [37] and performed on a single GPU (NVIDIA GeForce RTX 3090).

F. EVALUATION

To evaluate the effectiveness of the trained model, two assessments were conducted. The first involved the untreated group of 531 MR images that was not included in the training process, and the second assessed the performance of the trained model for the DOX-treated group of 579 MR images. To evaluate model performance, soft voting was utilized with the threshold of calculated average probability set to 0.4. The mean and standard deviation (SD) of the Dice similarity coefficient (Dice) and intersection over union (IoU) metrics were calculated for each group. The Dice and IoU [38] metrics were used as quantitative scores of segmentation performance and expressed in terms of true positive (TP), false positive (FP), and false negative (FN).

$$\text{Dice} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{1}$$

$$\text{IoU} = \frac{TP}{TP + FP + FN} \tag{2}$$

Based on the assessment process, a comparison of segmentation performance using various network architectures was performed by calculating Dice and IoU scores as well as by measuring the number of parameters and inference times.

For a statistical comparison of the various networks, the Mann–Whitney U test [39] was utilized. The null hypothesis was that there is no difference between the distributions of the two Dice scores from TS-Net and the others. The significance level of the decision of a hypothesis test, namely alpha, was set at 0.05.

G. VOLUME QUANTIFICATION AND VISUALIZATION

1) TUMOR VOLUME

Segmentation results of the trained TS-Net were used for automated tumor volumetric measurements. The obtained tumor volumes were utilized to monitor the tumor growth and effectiveness of therapy at each time point, and additionally to evaluate the performance of the trained TS-Net against the ground truth. To calculate a tumor volume in a single 2D MR image, the total number of tumor voxels is counted and multiplied by the unit volume of the voxel. In this study, the unit volume of the voxel was 0.009 mm^3. Even though 2D MR images were employed, the measured volume has a unit of mm^3 because of the slice thickness. For example, to measure the volume of tumors distributed over 10 to 15 MR images on day 30, the tumor volumes of each MR image were summed.

To observe the growth rate of the tumor volume in graph form, the subject with the largest tumor volume was selected from among the 5 subjects in the untreated group, and the subject with the smallest tumor volume was selected from among the 6 subjects in the treated group. These were determined

TABLE 2. Comparison of segmentation performance with P-values from Mann–Whitney U tests.

Network	Untreated group		DOX-treated group		Total		P-value	The number of parameters	Inference time (ms/image)
	Dice(±SD)	IoU(±SD)	Dice(±SD)	IoU(±SD)	Dice(±SD)	IoU(±SD)			
U-Net	0.908 ± 0.160	0.856 ± 0.174	0.925 ± 0.086	0.868 ± 0.106	0.918 ± 0.128	0.864 ± 0.143	0.788	23,379,969	5.65
U-Net^{++}	0.906 ± 0.169	0.856 ± 0.182	0.928 ± 0.072	0.871 ± 0.094	0.918 ± 0.129	0.864 ± 0.143	0.617	26,072,337	9.04
U-Net with pre-activation Residual block	0.910 ± 0.159	0.858 ± 0.173	0.923 ± 0.091	0.866 ± 0.109	0.917 ±0.129	0.862 ± 0.143	0.579	32,432,003	6.73
U-Net with original Residual block	0.911 ± 0.156	0.860 ± 0.168	0.924 ± 0.085	0.888 ± 0.107	0.918 ± 0.124	0.864 ± 0.140	0.656	32,432,129	6.67
U-Net with ResNet50	0.910 ± 0.153	0.858 ± 0.169	0.926 ± 0.080	0.869 ± 0.101	0.919 ± 0.120	0.863 ±0.138	0.970	32,514,833	9.43
nnU-Net	0.913 ± 0.028	0.863 ± 0.030	0.930 ± 0.015	0.875 ± 0.025	0.923 ± 0.024	0.870 ± 0.028	0.118	33,471,820	7.31
TS-Net	**0.912 ± 0.148**	**0.860 ± 0.165**	**0.927 ± 0.075**	**0.870 ± 0.097**	**0.920 ± 0.116**	**0.865 ± 0.134**	-	**24,430,097**	**6.13**

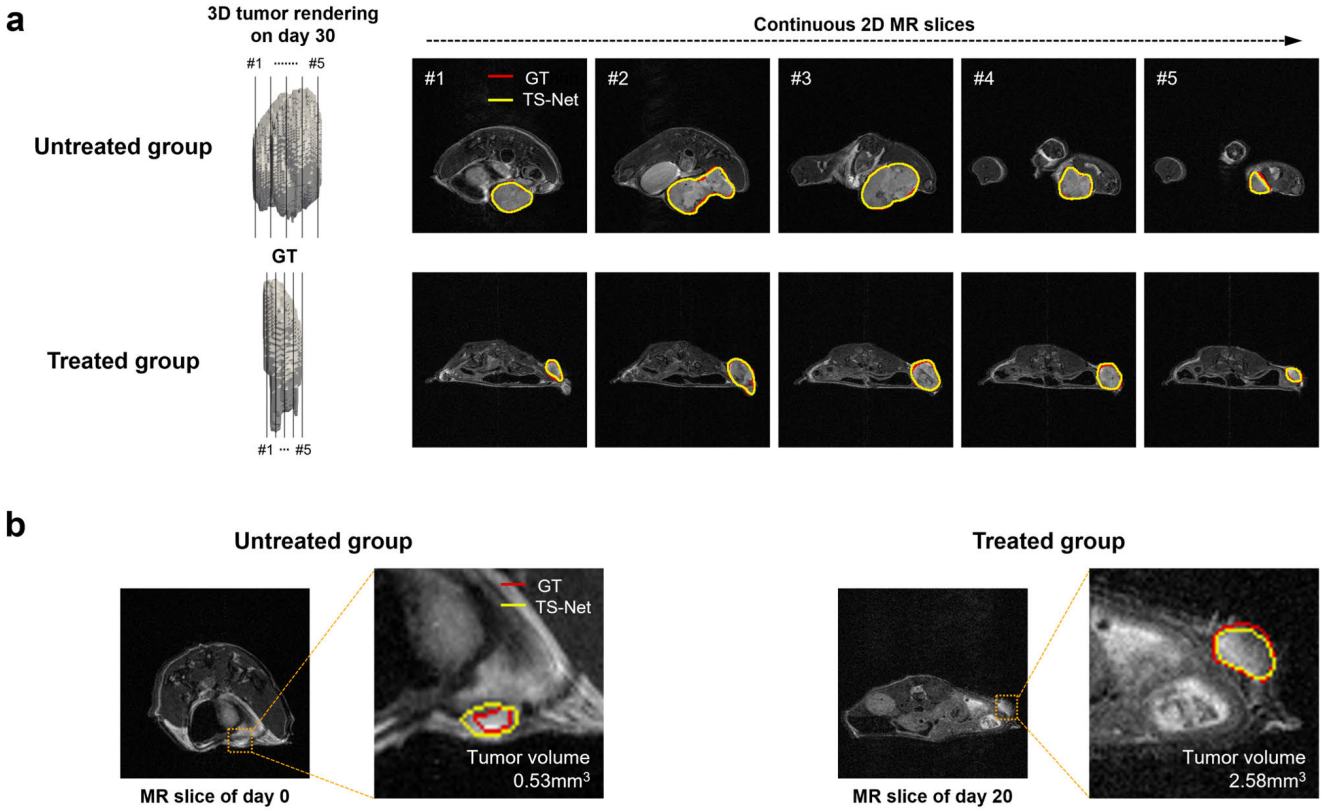

FIGURE 3. Visualization of the segmentation results. (a) A 3D tumor rendering of the ground truth (GT) in untreated and treated groups at day 30 is shown on the left. Segmentation maps corresponding to the continuous slices of the 3D tumor rendering at regular intervals are shown on the right. In the segmentation maps, the ground truth is outlined in red, and the TS-Net output is outlined in yellow. (b) Small tumor regions with ground truth volumes (mm^3) zoomed in to show the ground truth and corresponding segmentation map for untreated (left) and treated (right) groups.

from the total 3D tumor volume obtained on day 30. Then the tumor volume was calculated for each time point (day 0, day 5, day 10, day 15, day 20, day 25, and day 30) using both the ground truth and the TS-Net results for monitoring the growth rate.

Moreover, the effect of DOX at each time point can be confirmed through this graph of tumor volume growth rate. At this time, the mean tumor volume (V) was calculated from all subjects in each group and then normalized using the mean tumor volume at day 0 (V_0) to start at the same value. We also measured the standard error of the two groups.

To verify the performance on small tumors from TS-Net and nnU-Net per slice, a scatter plot and regression analysis were utilized. Also, in order to investigate the sensitivity of TS-Net in detecting small tumors, a box plot was used. Small tumors ($0 <$tumor volume (mm^3) ≤ 5) were divided into each range, and the median, quartile, and whisker were also calculated.

2) 3D VOLUME RENDERING

Visualization of the tumor shape and surface was achieved using 3D volume rendering. This process involved stacking continuous MR slices from each time point into a 3D image, from which isosurface data was extracted and smoothed 10 times for an optimal visualization. The software tools used for 3D volume rendering were Freesurfer (Massachusetts

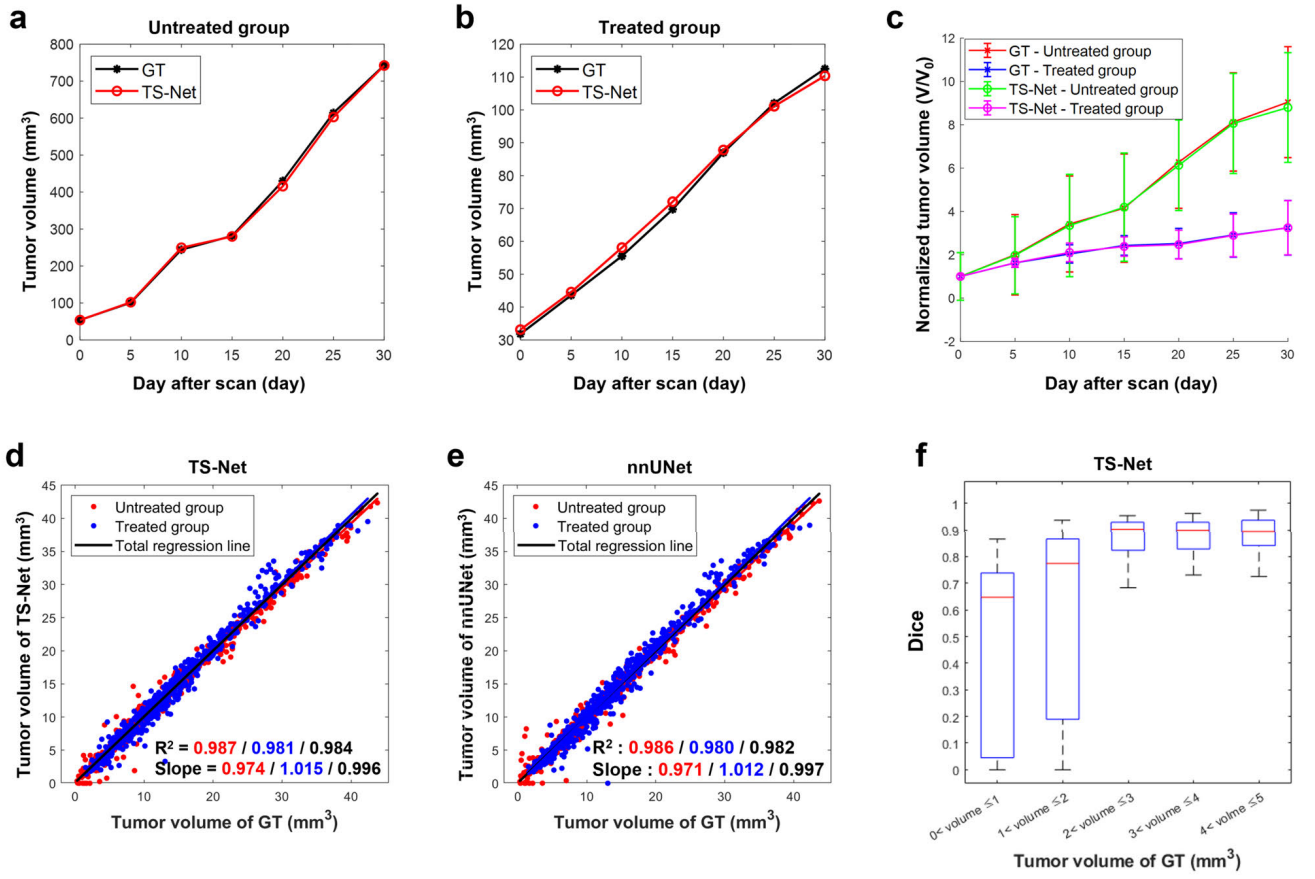

FIGURE 4. Results analysis. (a) Graph depicting the tumor volume growth rate of the largest tumor observed in the untreated group. (b) Graph depicting the tumor volume growth rate of the smallest tumor observed in the treated group. (c) Plot of the effect of DOX. The lines represent the mean normalized tumor volume with standard error. The ground truth (GT) results for the untreated and treated groups are shown in red and blue, and the TS-Net output results are shown in green and pink. (d) Linear regression analysis of TS-Net illustrating the correlation between the tumor volume of the ground truth and the output from TS-Net for the untreated (red) and treated (blue) groups (R^2=0.984). The red and blue line represent each group's regression line, and the total regression line of both groups (black) has a slope of 0.996. (e) Linear regression analysis of nnU-Net illustrating the correlation between the tumor volume of the ground truth and the output from nnU-Net for the untreated (red) and treated (blue) groups (R^2=0.982). The red and blue line represent each group's regression line, and the total regression line of both groups (black) has a slope of 0.997. (f) Box plot of small tumors with a volume less than or equal to 5 mm^3. The median value is represented by the red line.

General Hospital, MA, USA), Paraview (Sandia National Labs, Kitware Inc., and Los Alamos National Labs, NY, USA), and MATLAB (The MathWorks, Inc., MA, USA).

III. RESULTS

A. SEGMENTATION PERFORMANCE

As mentioned in the previous section, different CNN architectures were compared with TS-Net for breast tumor segmentation. The mean and SD of Dice and IoU were calculated for each group. The results are presented in Table 2. According to the comparison results in Table 2, the performance of the different networks was comparable in terms of Dice and IoU. It was also observed that all networks showed a slightly better performance with the treated group compared to the untreated group. All networks had P-values from Mann–Whitney U tests larger than 0.05 alpha and thus all networks had similar Dice score distributions. To select trained models, the criteria of a Dice score above 0.920 was set. Among all networks, TS-Net and nnU-Net achieved a Dice score above 0.920 without statistically significant performance differences between them (P-value = 0.118). It should be noted nnU-Net had a relatively slower inference time to predict the test images, likely due to having the largest number of parameters as shown in Table 2.

The segmentation outcomes from the trained TS-Net are illustrated in Fig. 3. The results of the untreated and DOX-treated groups are shown in the first and second rows of Fig. 3a, respectively. The 3D tumor volume was constructed from the continuous ground truth MR slices of day 30, and five sections of the MR slices were chosen at a regular interval. These five MR slices represent the top, middle, and bottom sections of the tumors, and each is shown with the corresponding ground truth (red contour) and segmentation map (yellow contour) in order to illustrate the performance consistency. To focus on small tumors, Fig. 3b shows MR slices with small tumor volumes from the untreated group (ground truth, 0.53 mm^3) and the treated group (ground truth, 2.58 mm^3). The tumor regions are zoomed in showing

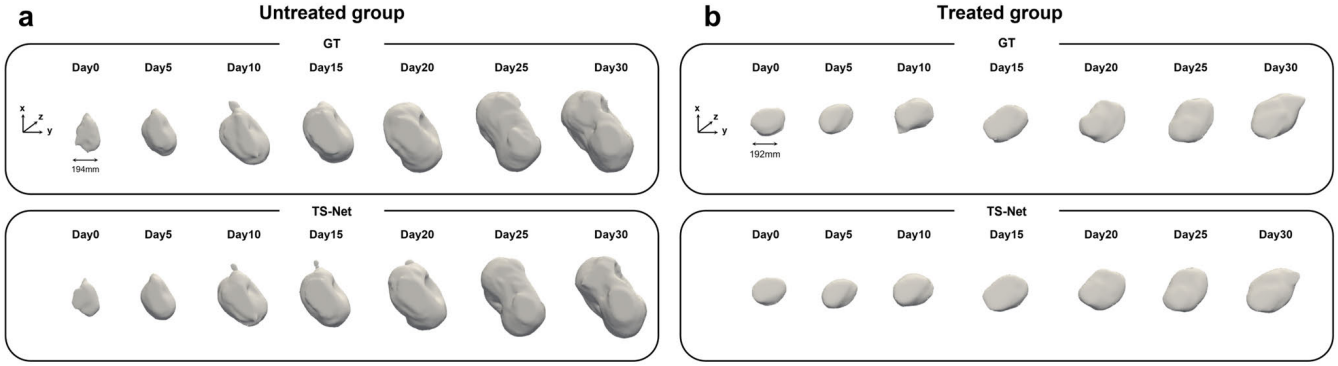

FIGURE 5. 3D volume rendering results. (a) 3D-rendered tumor volumes for the untreated group, and (b) those for the treated group. Ground truth visualizations are shown in the top panels, and TS-Net outputs are shown in the bottom panels.

the corresponding ground truth (red) and segmentation map (yellow) for a clear visualization.

B. VOLUME QUANTIFICATION AND VISUALIZATION

1) TUMOR VOLUME

Fig. 4a and 4b plot the tumor volume growth rate of the untreated and treated groups from TS-Net, respectively. The x-axis denotes the seven time points, while the y-axis represents the corresponding tumor volume. These plots allow us to observe the growth rate of the tumor volume for both the ground truth and the TS-Net results. The tumor growth rate curves gradually increased similarly in the ground truth and TS-Net results of untreated and treated groups. Fig. 4c displays the effect of DOX treatment. As shown in the figure, the normalized mean tumor volume of the untreated group increased rapidly compared to the DOX-treated group.

The tumor volumes were also calculated per 2D image for linear regression analysis and displayed as a scatter plot to compare the performance of TS-Net and nnU-Net in Fig. 4d and 4e. In Fig. 4d, the untreated group (red) has a correlation of 0.987 and a regression line slope of 0.974. The DOX-treated group (blue) has a correlation of 0.981 and a regression line slope of 1.015. As a result, analysis of the two groups from TS-Net yielded a strong correlation ($R^2 = 0.984$) with a total regression line slope of 0.996. In Fig. 4e, the untreated group (red) has a correlation of 0.986 and a regression line slope of 0.971, while the DOX-treated group (blue) has a correlation of 0.980 and a regression line slope of 1.012. As a result, analysis of the two groups from nnU-Net also yielded a strong correlation ($R^2 = 0.982$) with a total regression line slope of 0.987. Comparing the performance results in Fig. 4d and 4e, no significant differences in the values of R^2 or the slope of the regression line were found, indicating a comparable performance of the models. In Fig. 4f, the box plots represent the evaluation results of small tumors ($0 <$ tumor volume (mm^3) ≤ 5). The tumor volumes were divided into ranges to observe the model performance in detail. TS-Net achieved a Dice score of 0.85 for tumor volumes greater than 2 mm^3.

2) 3D VOLUME RENDERING

The 3D volume rendering results for the untreated and treated groups at each time point are shown in Fig. 5a and 5b, respectively. The ground truth 3D rendering is displayed in the top panels, and the generated TS-Net output is displayed in the bottom panels. As Fig. 5 shows, the tumor volume gradually expands over time. In addition, the change in tumor shape and surface can be observed by 3D volume rendering. In this way, the performance of TS-Net can be determined through such comparison. These results support that the developed approach is proficient to segment and visualize tumors, and thus by employing TS-Net, researchers can monitor orthotopic breast cancer models over time with a noninvasive method.

IV. DISCUSSION

In this study, a CNN-based approach, TS-Net, was proposed for segmenting breast tumors in an orthotopic breast cancer model using MR imaging. At the beginning of the research, it was hypothesized that if TS-Net trained on longitudinal MR images of an untreated group performed well on a DOX-treated group, then the approach may have the potential for wide utilization in orthotopic breast cancer model research. To test this hypothesis, several networks were evaluated. While U-Net with ResNet50, U-Net with original residual block, and U-Net^{++} had Dice and IoU scores similar to TS-Net, the models had larger numbers of parameters and slower inference times. Therefore, the networks with a Dice score above 0.920 were selected for comparison, a criteria met by TS-Net and nnU-Net. Between them, nnU-Net had higher but not statistically significant Dice and IoU scores (P-value = 0.118). Moreover, linear regression analysis in terms of R^2 and slope from both models showed no meaningful differences. Accordingly, the factors of inference time and number of parameters can be considered as main aspects in determining the proper network, which led us to TS-Net.

The use of a longitudinal MR dataset from an orthotopic breast cancer model involving treatment for the purpose of generating tumor segmentation is meaningful to the medical

imaging field. Through MRI, a series of continuous images were obtained that depicted different sections of tumors. This indicates that our framework is generally effective for segmenting entire tumors. These images also revealed the presence of small tumors that have traditionally been challenging for experts to detect in single MR slices. Despite this difficulty, our proposed approach was able to accurately segment even small tumors, with volumes more than 2 mm^3, in a single MR image. Taken together, the high evaluation scores, longitudinal MR imaging, and ability to detect small tumors for both groups suggest that TS-Net can be widely utilized for orthotopic breast cancer research using doxorubicin.

V. CONCLUSION

In order to automatically track longitudinal tumor volume changes in an orthotopic breast cancer model, a deep learning framework, called TS-Net, was proposed. The framework was developed using longitudinal MR images of an untreated group and a DOX-treated group as the dataset, an approach that has not previously been reported. Results showed that our framework achieved high Dice scores, indicating its potential for monitoring tumor growth rates over time and evaluating the therapeutic effect of DOX. A strong correlation (R^2 = 0.984, slope = 0.996) was also found between the tumor volumes obtained from the ground truth and model output. Additionally, the proposed framework could effectively segment small tumors (2 mm^3) in a single MR image. These results indicate that TS-Net could be a valuable tool for researchers and be applied to the preclinical validation of drug development using orthotopic breast tumor models.

REFERENCES

[1] S. Lei, R. Zheng, S. Zhang, S. Wang, R. Chen, K. Sun, H. Zeng, J. Zhou, and W. Wei, "Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020," *Cancer Commun.*, vol. 41, no. 11, pp. 1183–1194, Nov. 2021.

[2] R. Roslidar, A. Rahman, R. Muharar, M. R. Syahputra, F. Arnia, M. Syukri, B. Pradhan, and K. Munadi, "A review on recent progress in thermal imaging and deep learning approaches for breast cancer detection," *IEEE Access*, vol. 8, pp. 116176–116194, 2020.

[3] T. Vandamme, "Use of rodents as models of human diseases," *J. Pharmacy Bioallied Sci.*, vol. 6, no. 1, p. 2, 2014.

[4] A. G. Polson and R. N. Fuji, "The successes and limitations of preclinical studies in predicting the pharmacodynamics and safety of cell-surface-targeted biological agents in patients," *Brit. J. Pharmacol.*, vol. 166, no. 5, pp. 1600–1602, Jul. 2012.

[5] J. J. Killion, R. Radinsky, and I. J. Fidler, "Orthotopic models are necessary to predict therapy of transplantable tumors in mice," *Cancer Metastasis Rev.*, vol. 17, pp. 279–284, Sep. 1998.

[6] G. D. Ayers, E. T. McKinley, P. Zhao, J. M. Fritz, R. E. Metry, B. C. Deal, K. M. Adlerz, R. J. Coffey, and H. C. Manning, "Volume of preclinical xenograft tumors is more accurately assessed by ultrasound imaging than manual caliper measurements," *J. Ultrasound Med.*, vol. 29, no. 6, pp. 891–901, Jun. 2010.

[7] J. L. Whitwell, "Longitudinal imaging: Change and causality," *Current Opinion Neurol.*, vol. 21, no. 4, pp. 410–416, 2008.

[8] S. K. Lyons, "Advances in imaging mouse tumour models in vivo," *J. Pathol., A J. Pathological Soc. Great Britain Ireland*, vol. 205, no. 2, pp. 194–205, 2005.

[9] H.-H. Dubben, H. D. Thames, and H.-P. Beck-Bornholdt, "Tumor volume: A basic and specific response predictor in radiotherapy," *Radiotherapy Oncol.*, vol. 47, no. 2, pp. 167–174, May 1998.

[10] S. R. Sternberg, "Biomedical image processing," *Computer*, vol. 16, no. 1, pp. 22–34, Jan. 1983.

[11] M. Graeser, S. Schrading, O. Gluz, K. Strobel, C. Herzog, L. Umutlu, A. Frydrychowicz, D. Rjosk-Dendorfer, R. Würstlein, R. Culemann, and C. Eulenburg, "Magnetic resonance imaging and ultrasound for prediction of residual tumor size in early breast cancer within the ADAPT subtrials," *Breast Cancer Res.*, vol. 23, no. 1, pp. 1–13, Dec. 2021.

[12] A. Jethava, S. Ali, D. Wakefield, R. Crowell, J. Sporn, and J. Vrendenburgh, "Diagnostic accuracy of MRI in predicting breast tumor size: Comparative analysis of MRI vs histopathological assessed breast tumor size," *Conn Med*, vol. 79, no. 5, pp. 261–267, 2015.

[13] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and its variants for medical image segmentation: A review of theory and applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021.

[14] A. Işın, C. Direkoğlu, and M. Şah, "Review of MRI-based brain tumor image segmentation using deep learning methods," *Proc. Comput. Sci.*, vol. 102, pp. 317–324, 2016.

[15] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.

[16] G. Du, X. Cao, J. Liang, X. Chen, and Y. Zhan, "Medical image segmentation based on U-Net: A review," *J. Imag. Sci. Technol.*, vol. 64, no. 2, p. 20508, Mar. 2020.

[17] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction With MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer, 2018, pp. 3–11.

[18] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "NnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021.

[19] L. Huo, X. Hu, Q. Xiao, Y. Gu, X. Chu, and L. Jiang, "Segmentation of whole breast and fibroglandular tissue using nnU-Net in dynamic contrast enhanced MR images," *Magn. Reson. Imag.*, vol. 82, pp. 31–41, Oct. 2021.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[21] X. Ou, P. Yan, Y. Zhang, B. Tu, G. Zhang, J. Wu, and W. Li, "Moving object detection method via ResNet-18 with encoder–decoder structure in complex scenes," *IEEE Access*, vol. 7, pp. 108152–108160, 2019.

[22] Y. Jiang, L. Chen, H. Zhang, and X. Xiao, "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module," *PLoS ONE*, vol. 14, no. 3, Mar. 2019, Art. no. e0214587.

[23] R. Mukhopadhyay, R. Sen, B. Paul, J. Kazi, S. Ganguly, and M. C. Debnath, "Gemcitabine co-encapsulated with curcumin in folate decorated PLGA nanoparticles; a novel approach to treat breast adenocarcinoma," *Pharmaceutical Res.*, vol. 37, no. 3, pp. 1–19, Mar. 2020.

[24] G.-L. Zhang, Y. Zhang, K.-X. Cao, and X.-M. Wang, "Orthotopic injection of breast cancer cells into the mice mammary fat pad," *J. Visualized Exp.*, no. 143, Jan. 2019, Art. no. e58604.

[25] V. Kersemans, B. Cornelissen, P. D. Allen, J. S. Beech, and S. C. Smart, "Subcutaneous tumor volume measurement in the awake, manually restrained mouse using MRI," *J. Magn. Reson. Imag.*, vol. 37, no. 6, pp. 1499–1504, Jun. 2013.

[26] M. M. Tomayko and C. P. Reynolds, "Determination of subcutaneous tumor size in athymic (nude) mice," *Cancer Chemotherapy Pharmacol.*, vol. 24, no. 3, pp. 148–154, Sep. 1989.

[27] N. Myung, S. Jin, H. J. Cho, and H.-W. Kang, "User-designed device with programmable release profile for localized treatment," *J. Controlled Release*, vol. 352, pp. 685–699, Dec. 2022.

[28] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[29] A. F. Agarap, "Deep learning using rectified linear units (ReLU)," 2018, *arXiv:1803.08375*.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 630–645.

[31] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 8609–8613.

[32] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.
[33] C. Ye, W. Wang, S. Zhang, and K. Wang, "Multi-depth fusion network for whole-heart CT image segmentation," *IEEE Access*, vol. 7, pp. 23421–23429, 2019.
[34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[35] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019.
[36] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," *Information*, vol. 11, no. 2, p. 125, Feb. 2020.
[37] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, and A. Desmaison, "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–11.
[38] K. Xia, H. Yin, P. Qian, Y. Jiang, and S. Wang, "Liver semantic segmentation algorithm based on improved deep adversarial networks in combination of weighted loss function on abdominal CT images," *IEEE Access*, vol. 7, pp. 96349–96358, 2019.
[39] N. Nachar, "The Mann–Whitney U: A test for assessing whether two independent samples come from the same distribution," *Tuts. Quant. Methods Psychol.*, vol. 4, no. 1, pp. 13–20, Mar. 2008.

YUNKYOUNG JUN received the B.S. degree in electronic engineering from Kongju National University, Cheonan, Republic of Korea, in 2021. She is currently pursuing the M.S. degree in biomedical engineering with the Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea. Her research interests include deep learning and medical imaging.

SEOKHA JIN received the B.S. degree in biomedical science and the Ph.D. degree in biomedical engineering from the Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea, in 2014 and 2021, respectively. He is currently a Sejong Research Fellow with UNIST. His research interest includes perfusion and diffusion MRI in preclinical disease models.

NOEHYUN MYUNG received the B.S. degree in biomedical engineering in biology from the School of Life Science, Ulsan National Institute of Science and Technology (UNIST), Republic of Korea, in 2017, where he is currently pursuing the combined M.S. and Ph.D. degree in biomedical engineering. His research interests include biofabrication, drug delivery systems, medical devices, xenotransplantation, and disease modeling.

JIWOO JEONG received the B.S. degree in biomedical engineering from Ulsan University, Ulsan, Republic of Korea, in 2022. She is currently pursuing the M.S. degree in biomedical engineering with the Ulsan National Institute of Science and Technology (UNIST), Ulsan. Her research interest includes convert MRI into CT using deep learning.

JIMIN LEE received the B.S. degree in nuclear and quantum engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea, in 2015, and the Ph.D. degree in biomedical radiation sciences from the Department of Transdisciplinary Studies, Seoul National University, Seoul, Republic of Korea, in 2021.

Since 2021, she has been an Assistant Professor with the Department of Nuclear Engineering and the Graduate School of Artificial Intelligence, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea. Her research interest includes deep learning approaches for solving biomedical imaging problems.

HYUNG JOON CHO received the B.S. degree in physics from Seoul National University, Seoul, Republic of Korea, in 2000, and the Ph.D. degree in radiological science and technology from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2005.

Since 2022, he has been a Professor with the Department of Biomedical Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea. His research interest includes quantitative MR imaging for disease biomarkers.