



3D U-Net Improves Automatic Brain Extraction for Isotropic Rat Brain Magnetic Resonance Imaging Data

Li-Ming Hsu^{1,2,3*†}, Shuai Wang^{4†}, Lindsay Walton^{1,2,3}, Tzu-Wen Winnie Wang^{1,2,3}, Sung-Ho Lee^{1,2,3} and Yen-Yu Ian Shih^{1,2,3*}

¹ Center for Animal Magnetic Resonance Imaging, The University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, ² Biomedical Research Imaging Center, The University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, ³ Department of Radiology, The University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, ⁴ School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, China

OPEN ACCESS

Edited by:

Yizhang Jiang,
Jiangnan University, China

Reviewed by:

Jan Egger,
Graz University of Technology, Austria
Yang Lei,
Emory University, United States

*Correspondence:

Li-Ming Hsu
limingh@email.unc.edu
Yen-Yu Ian Shih
shihy@unc.edu

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 24 October 2021

Accepted: 15 November 2021

Published: 16 December 2021

Citation:

Hsu L-M, Wang S, Walton L,
Wang T-WW, Lee S-H and Shih Y-YI
(2021) 3D U-Net Improves Automatic
Brain Extraction for Isotropic Rat Brain
Magnetic Resonance Imaging Data.
Front. Neurosci. 15:801008.
doi: 10.3389/fnins.2021.801008

Brain extraction is a critical pre-processing step in brain magnetic resonance imaging (MRI) analytical pipelines. In rodents, this is often achieved by manually editing brain masks slice-by-slice, a time-consuming task where workloads increase with higher spatial resolution datasets. We recently demonstrated successful automatic brain extraction via a deep-learning-based framework, U-Net, using 2D convolutions. However, such an approach cannot make use of the rich 3D spatial-context information from volumetric MRI data. In this study, we advanced our previously proposed U-Net architecture by replacing all 2D operations with their 3D counterparts and created a 3D U-Net framework. We trained and validated our model using a recently released CAMRI rat brain database acquired at isotropic spatial resolution, including T2-weighted turbo-spin-echo structural MRI and T2*-weighted echo-planar-imaging functional MRI. The performance of our 3D U-Net model was compared with existing rodent brain extraction tools, including Rapid Automatic Tissue Segmentation, Pulse-Coupled Neural Network, SHape descriptor selected External Regions after Morphologically filtering, and our previously proposed 2D U-Net model. 3D U-Net demonstrated superior performance in Dice, Jaccard, center-of-mass distance, Hausdorff distance, and sensitivity. Additionally, we demonstrated the reliability of 3D U-Net under various noise levels, evaluated the optimal training sample sizes, and disseminated all source codes publicly, with a hope that this approach will benefit rodent MRI research community.

Significant Methodological Contribution: We proposed a deep-learning-based framework to automatically identify the rodent brain boundaries in MRI. With a fully 3D convolutional network model, 3D U-Net, our proposed method demonstrated improved performance compared to current automatic brain extraction methods, as shown in several qualitative metrics (Dice, Jaccard, PPV, SEN, and Hausdorff). We trust that this tool will avoid human bias and streamline pre-processing steps during 3D high resolution rodent brain MRI data analysis. The software developed herein has been disseminated freely to the community.

Keywords: 3D U-Net, MRI, brain extraction, segmentation, brain mask, rat brain

INTRODUCTION

Magnetic resonance imaging (MRI) is a commonly utilized technique to noninvasively study the anatomy and function of rodent brains (Mandino et al., 2019). Among the data pre-processing procedures, brain extraction is an important step that ensures the success of subsequent registration processes (Uhlich et al., 2018). Automating brain extraction (a.k.a. skull stripping) is particularly challenging for rodent brains as compared to humans because of differences in brain/scalp tissue geometry, image resolution with respect to brain-scalp distance, tissue contrast around the skull, and sometimes signal artifacts from surgical manipulations (Hsu et al., 2020). Additionally, rodent brain MRI data is typically acquired at higher magnetic fields (mostly > 7T), where stronger susceptibility artifacts and field biases represent challenges to the rodent brain extraction process. Further, most rodent MRI studies do not utilize a volume receiver and therefore exhibit higher radiofrequency (RF) coil inhomogeneity. For these reasons, the extraction tools that work with human brains (Kleesiek et al., 2016; Fatima et al., 2020; Wang et al., 2021) cannot be directly adopted for rodent brain applications. In practice, rodent brain extraction is often achieved by manually drawing brain masks for every slice, making it a time-consuming and operator-dependent process. Datasets with high through-plane resolution make manual brain extraction a particularly daunting task. Therefore, a robust and reliable automatic brain extraction tool would streamline the pre-processing pipeline, avoid personnel bias, and significantly improve research efficiency (Babalola et al., 2009; Lu et al., 2010; Gaser et al., 2012; Feo and Giove, 2019; Hsu et al., 2020).

To date, the most prominent tools to address rodent MRI brain extraction include Pulse-Coupled Neural Network (PCNN)-based brain extraction proposed by Chou et al. (2011), Rapid Automatic Tissue Segmentation (RATS) proposed by Oguz et al. (2014), and SHape descriptor selected External Regions after Morphologically filtering (SHERM) proposed by Liu et al. (2020), as well as a convolutional deep-learning based algorithm, 2D U-Net, proposed by Hsu et al. (2020). While PCNN, RATS, and SHERM have demonstrated remarkable success [detailed introduction and comparisons discussed in Hsu et al. (2020)], their performance is subject to brain size, shape, texture, and contrast; hence, their settings often need to be adjusted per MRI-protocol for optimal results. In contrast, the U-Net algorithm explores and learns the hierarchical features from the training dataset, and provides a user-friendly and more universally applicable platform (Ronneberger et al., 2015; Yogananda et al., 2019).

As a specific type of convolutional neural network (CNN) architecture (Krizhevsky et al., 2012), U-Net has proven valuable in biomedical image segmentation (Ronneberger et al., 2015; Yogananda et al., 2019). U-Net utilizes the encoder/decoder structure that easily integrates multi-scale information and has better gradient propagation during training. Our previous 2D U-Net approach uses 2D convolutional kernels to predict brain boundaries on a single slice basis. However, since the 2D framework only takes a single slice as input and does not utilize information

across slice direction, it inherently fails to leverage context from adjacent slices in a volumetric MRI dataset. To improve upon and possibly outperform 2D U-Net, a 3D U-Net framework using 3D convolutional kernels to predict segmentation predictions on volumetric patches must be explored (Çiçek et al., 2016).

In this work, we demonstrated the use of 3D U-Net for brain extraction in high resolution 3D rat brain MRI data. The whole network is implemented based on Keras (Chollet, 2015) with TensorFlow (Martín et al., 2016) as the backend. We trained and tested the 3D U-Net model for brain extraction performance using a recently released CAMRI rat brain database (Lee et al., 2021), including T2-weighted (T2w) rapid acquisition with relaxation enhancement (RARE) structural MRI (0.2 mm isotropic resolution) and T2*-weighted (T2*w) echo-planar-imaging (EPI) functional MRI (0.4 mm isotropic resolution). The 3D U-Net model was compared with existing rodent brain extraction tools, including RATS, PCNN, SHERM, and 2D U-Net. To benchmark the utility of this approach to rodent MRI data of various quality, we assessed 3D U-Net performance under various noise levels and evaluated the optimal training sample sizes. It is our hope that the information provided herein, together with dissemination of source codes, will benefit the rodent MRI research community.

MATERIALS AND METHODS

Dataset Descriptions

This study utilizes a recently disseminated CAMRI dataset (Lee et al., 2021), available at <https://doi.org/10.18112/openneuro.ds003646.v1.0.0>. The CAMRI dataset consisted of rats aged P80–210 weighing 300–450 g across cohorts of three different rat strains: Sprague-Dawley ($n = 41$), Long-Evans ($n = 13$), and Wistar ($n = 33$); and both sexes: male ($n = 65$) and female ($n = 22$). Detailed information about each cohort can be found in Lee et al. (2021). Each subject contains a T2w RARE and an T2*w EPI, and resolutions were 0.2 mm isotropic and 0.4 mm isotropic, respectively. To train our U-Net model, we first established a training dataset by randomly selecting 80% of the T2w RARE and T2*w EPI images in the CAMRI rat database (69 subjects), leaving the remaining 20% of the data as final performance testing dataset (18 subjects). In the training process, we further randomly selected 80% of the data from the training dataset (55 subjects). The remaining 20% of the data from the training dataset was used to validate the training of U-Net model (14 subjects). We repeated the training-validation process five times to avoid potential bias in data splitting. For each U-Net algorithm, the U-Net model with the highest averaged validation accuracy was then used as the final model for testing (**Supplementary Figure 2**).

U-Net

The 3D U-Net framework is shown in **Figure 1**. The contracting path includes 32 feature maps in the first convolutional block, 64 in the second, then 96, 128, and 256 in the third, fourth,

and fifth, respectively. Compared to the configuration described by Ronneberger et al. (2015), we replaced the cross-entropy loss function with the Dice coefficient loss (Wang et al., 2020) to free the optimization process from a class-imbalance problem (Millettari et al., 2016). Since we evaluated T2w RARE and T2*w EPI data, we performed spatial normalization for distinct resolutions. For spatial normalization, we resampled all images into the same spatial resolution at $0.2 \text{ mm} \times 0.2 \text{ mm} \times 0.2 \text{ mm}$ using nearest-neighbor (NN) interpolation. The NN was chosen because it is most suitable for binary images (brain mask). In U-Net training, the voxels belonging to the rat brain were labeled as 1 and other voxels (background) were labeled as 0. Our network was implemented using Keras (Chollet, 2015) with TensorFlow (Martín et al., 2016) as the backend. The initial learning rate and batch size were $1e^{-3}$ and 16, respectively. We used Adam (Kingma and Ba, 2015) as the optimizer and clipped all parameter gradients to a maximum norm of 1. During the training process, we randomly cropped $64 \times 64 \times 64$ sized patches from all directions as the input. Patches were chosen to be $64 \times 64 \times 64$ to maintain maximum patch size and cover as much of the rat brain as possible ($\sim 2,000 \text{ mm}^3$) without requiring excessive computational demand. During the inference process, the extracted and overlapped patches were input into the trained model with a $16 \times 16 \times 16$ stride. The overlapped predictions were averaged and then resampled back to the original resolution using nearest-neighbor interpolation to generate final output. We also trained (1) 2D U-Net with patch size of 64×64 (2D U-Net64) and a 16×16 stride (Hsu et al., 2020) and (2) 3D U-Net with patch size of $16 \times 16 \times 16$ and a $4 \times 4 \times 4$ stride (3D U-Net16) to compare the segmentation performance with the proposed 3D U-Net of patch size $64 \times 64 \times 64$ and a $16 \times 16 \times 16$ stride (3D U-Net64).

Reproducibility of 3D U-Net Performance

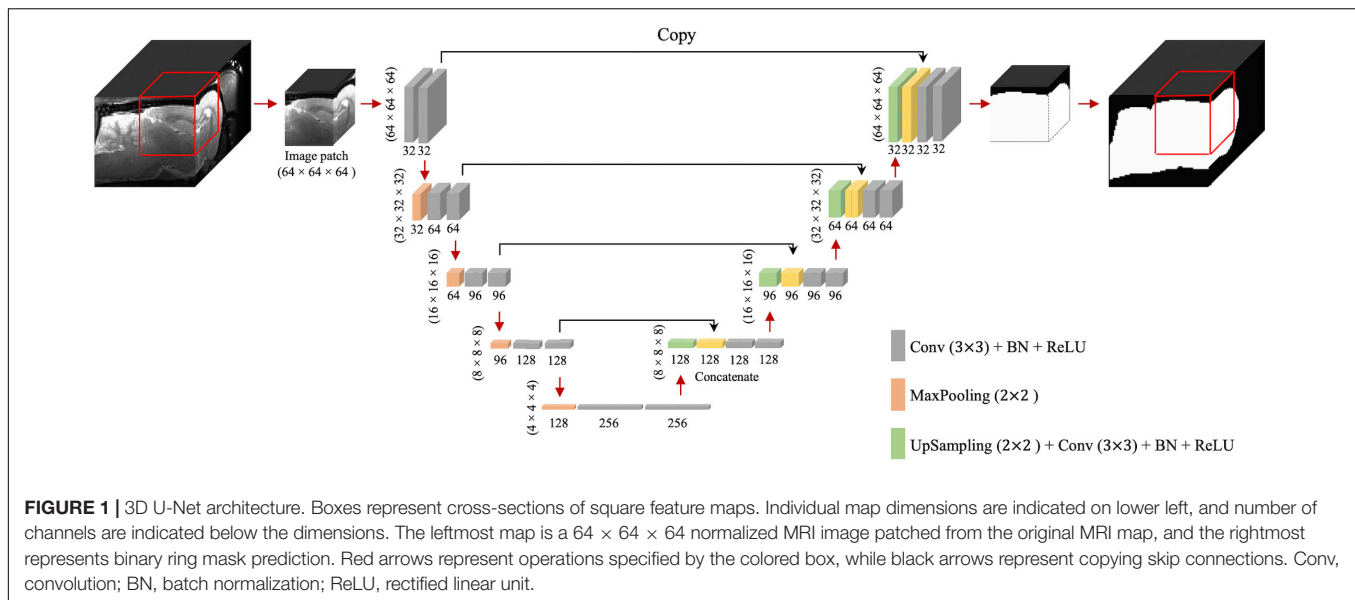
To evaluate the 3D U-Net64 performance reproducibility, we reexamined its accuracy under two conditions: (1) Adding different Gaussian noise in the testing images, and (2) using a different number of rats from the original cohort in the training process. Specifically, as for the first validation analysis, we added Gaussian white noise in the normalized testing images with variance from 5×10^{-4} to 5×10^{-5} with step 5×10^{-5} to investigate the segmentation performance. For each testing image, the signal-to-noise ratio (SNR) was estimated to represent the image noise levels by calculating the ratio of signal intensities in the area of interest and the background (**Supplementary Figure 1**). We estimated the average signal by using two spherical volumes of interest (VOI) with diameter = 1 mm in bilateral striatum. We also put two VOIs at background to extract noise in standard deviation of signal. To evaluate the optimal subject number in 3D U-Net64 training process, we randomly selected 5–55 training subjects in increments of 5 subjects from the total 55 training datasets, and 2, 8, and 14 validation subjects from the total 14 validation datasets. This random selection was repeated 5 times to avoid bias. We then calculated accuracy of the derived brain masks across each testing dataset.

Evaluation Methods

To demonstrate the reliability of our proposed method, we compared our 3D U-Net method with the most prominently used methods for rat brain segmentation: RATS (Oguz et al., 2014), PCNN (Chou et al., 2011), SHERM (Liu et al., 2020), and 2D U-Net (Hsu et al., 2020). All images were bias-corrected for field inhomogeneities using Advanced Normalization Tools (N4ITK) (Avants et al., 2009; Tustison et al., 2010). The parameters in each method were chosen according to the best parameters suggested in the literature. For the RATS algorithm, the intensity threshold (T) was set to the average intensity in the entire image and the brain size value (Vt) was set to $1,650 \text{ mm}^3$ (Oguz et al., 2014). For the PCNN algorithm, the brain size range was set to $1,000\text{--}3,000 \text{ mm}^3$ (Chou et al., 2011). For SHERM, the brain size range was set to $500\text{--}1,900 \text{ mm}^3$ (Liu et al., 2020). The convexity threshold in SHERM, defined as the ratio between the volume of a region and that of its convex hull, was set to 0.5.

To quantitatively evaluate the segmentation performance of 3D U-Net64, 2D U-Net64, 3D U-Net16, RATS, PCNN, and SHERM, we estimated the similarity of the generated brain segmentation results compared to manual brain masks (ground truth) drawn by an anatomical expert according to the Paxinos and Watson (2007) rat atlas. The manual segmentation was performed at the original MRI resolution before data resampling to $0.2 \text{ mm} \times 0.2 \text{ mm} \times 0.2 \text{ mm}$ for U-Net training. Evaluations included: (1) volumetric overlap assessments via Dice, the similarity of two samples; (2) Jaccard, the similarity of two samples where Dice doesn't satisfy the triangle inequality; (3) positive predictive value (PPV), the rate of true positives in prediction results; (4) sensitivity (SEN), the rate of true positives in manual delineation; (5) center-of-mass distance (CMD), the Euclidean distance to the center of mass of two samples; and (6) a surface distance assessment by Hausdorff distance, the distance of two samples. The following definitions were used for each: $Dice = 2(|A \cap B|) / (|A| + |B|)$, $Jaccard = (|A \cap B|) / (|A \cup B|)$, $PPV = (|A \cap B|) / B$, $SEN = (|A \cap B|) / A$, and $Hausdorff = \max\{h(A, B), h(B, A)\}$ and $h(A, B) = \max_{a \in A, b \in B} \{ \min d(a, b) \}$ where A denotes the voxel set

of the manually delineated volume, B denotes the voxel set of the predicted volume, and $d(a, b)$ is the Euclidian distance between a and b . The Hausdorff distance was only estimated in-plane to avoid confounds from non-uniformly sampled data. The maximal Hausdorff distance (i.e., worst matching) across slices for each subject was then used for comparison. Superior performance was indicated by higher Dice, Jaccard, PPV, and SEN, and lower Hausdorff values. We also reported the computation time on a Linux-based [Red Hat Enterprise Linux Server release 7.4 (Maipo)] computing system (Intel E5-2680 v3 processor, 2.50 GHz, 256-GB RAM) for each method. The computation times reported do not include pre-processing steps (i.e., signal normalization, image resampling, and bias correction). Paired t -tests were used for statistical comparisons between different algorithms, and two-sample t -tests were used to compare T2w RARE and T2*w EPI images in each algorithm. The threshold for significance was set to the alpha level ($p < 0.05$).



RESULTS

Figure 2 illustrates the performance of our trained 3D U-Net64 algorithm compared to 2D U-Net64, 3D U-Net16, RATS, PCNN, and SHERM for rat brain segmentation in the CAMRI dataset. Across all measures except PPV (which had no significant differences), 3D U-Net64 showed superior T2w RARE brain segmentation performance over other existing methods (RATS, PCNN, and SHERM). Notably, although all the U-Net-based approaches produced ideal results with Dice > 0.90, the 3D U-Net64 showed significantly higher accuracy ($p < 0.05$) versus the remaining U-Net approaches (2D U-Net64 and 3D U-Net16). The high PPV (>0.90) and low SEN (<0.90) from other existing methods (RATS, PCNN, and SHERM) indicate that brain segmentation was underestimated. In contrast, the low PPV (0.91 on anisotropic T2w RARE and 0.87 on anisotropic T2*w EPI) and high SEN (>0.95) from 2D U-Net16 suggest an overestimation. The significantly lower Hausdorff distance (4.29 on T2w RARE and 4.71 on T2*w EPI) and CMD (0.24 on T2w RARE and 0.25 on T2*w EPI) in 3D U-Net64 further supports its superior ability to match the ground-truth. Notably, all U-Net methods provide excellent (Dice > 0.95) or high (Dice > 0.90) accuracy on T2w RARE and T2*w EPI images. All methods showed significant lower Dice in T2*w EPI images as compared to T2w RARE images ($p < 0.05$). Specifically, the 3D U-Net64 still reached outstanding T2*w EPI brain segmentation performance metrics (Dice, Jaccard, PPV, and SEN > 0.95), whereas RATS, PCNN, and SHERM had lower T2*w EPI brain segmentation metrics (Dice, Jaccard, and SEN < 0.90). The compromised performance in the T2*w EPI image compared with T2w RARE indicates that these methods are less effective at handling data with lower spatial resolution. Further, 3D U-Net64 showed high accuracy (Dice > 0.95) in a validation study (**Supplementary Figure 2**). Together, these results suggest that 3D U-Net64 is a reliable and reproducible approach for rat brain segmentation.

Figures 3, 4 illustrate the best and worst Dice score cases on T2w RARE images from the CAMRI dataset using all six methods. In the best case, both 2D U-Net64 and 3D U-Net16 provided nearly perfect segmentation (Dice > 0.98), and although the RATS, PCNN, and SHERM algorithms also showed high performance (Dice = 0.95), the inferior brain boundaries were less accurate. In the worst case, all U-Net methods still achieved a satisfactory segmentation with Dice > 0.90, but the RATS, PCNN, and SHERM algorithms failed to identify the brainstem, olfactory bulb, and inferior brain regions where the MRI signal was weaker.

Figure 5 illustrates the best and worst Dice score cases on T2*w EPI from the CAMRI dataset using all six methods. In the best case, all U-Net methods provide nearly perfect segmentation (Dice > 0.98). The RATS (Dice = 0.89), PCNN (Dice = 0.93), and SHERM (Dice = 0.83) algorithms showed acceptable segmentation performance but underestimated inferior brain regions. In the worst case, all U-Net methods still achieve a satisfactory segmentation with Dice > 0.90, but the RATS, PCNN, and SHERM algorithms failed to accurately identify the inferior brain boundaries and incorporated excessive tissue outside the brain (Dice < 0.85).

We performed two validation analyses to evaluate the reliability of 3D U-Net segmentation performance. First, we added Gaussian noise to each testing image to evaluate 3D U-Net64 brain segmentation under different noisy environments. **Figure 6** shows that the SNR and 3D U-Net64 segmentation performance decrease with increasing Gaussian white noise variance. The original mean SNR is about 26 for both T2w RARE and T2*w EPI data. When the SNR > 10 (i.e., reduced to 38% of original SNR) in T2w RARE images and SNR > 14 (i.e., reduced to 54% of original SNR) in T2*w EPI images, the segmentation accuracy could still reach an excellent Dice > 0.95. **Supplementary Figures 3, 4** illustrate the best and worst examples on original T2w RARE and T2*w EPI images.

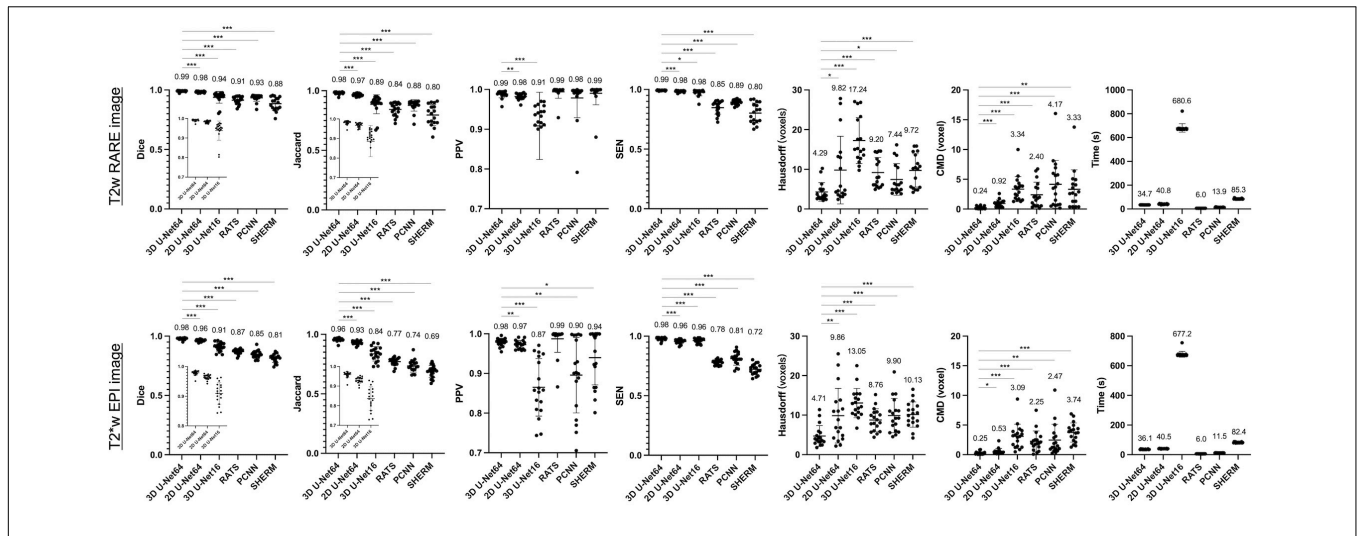


FIGURE 2 | Brain segmentation performance metrics for 3D U-Net64, 2D U-Net64, 3D U-Net16, RATS, PCNN, and SHERM on the CAMRI T2w RARE (upper row) and T2*w EPI (lower row) data. Average value is shown above each bar. Two-tailed paired *t*-tests were used for statistical comparison between 3D U-Net64 with other methods (**p* < 0.05, ***p* < 0.01, and ****p* < 0.001).

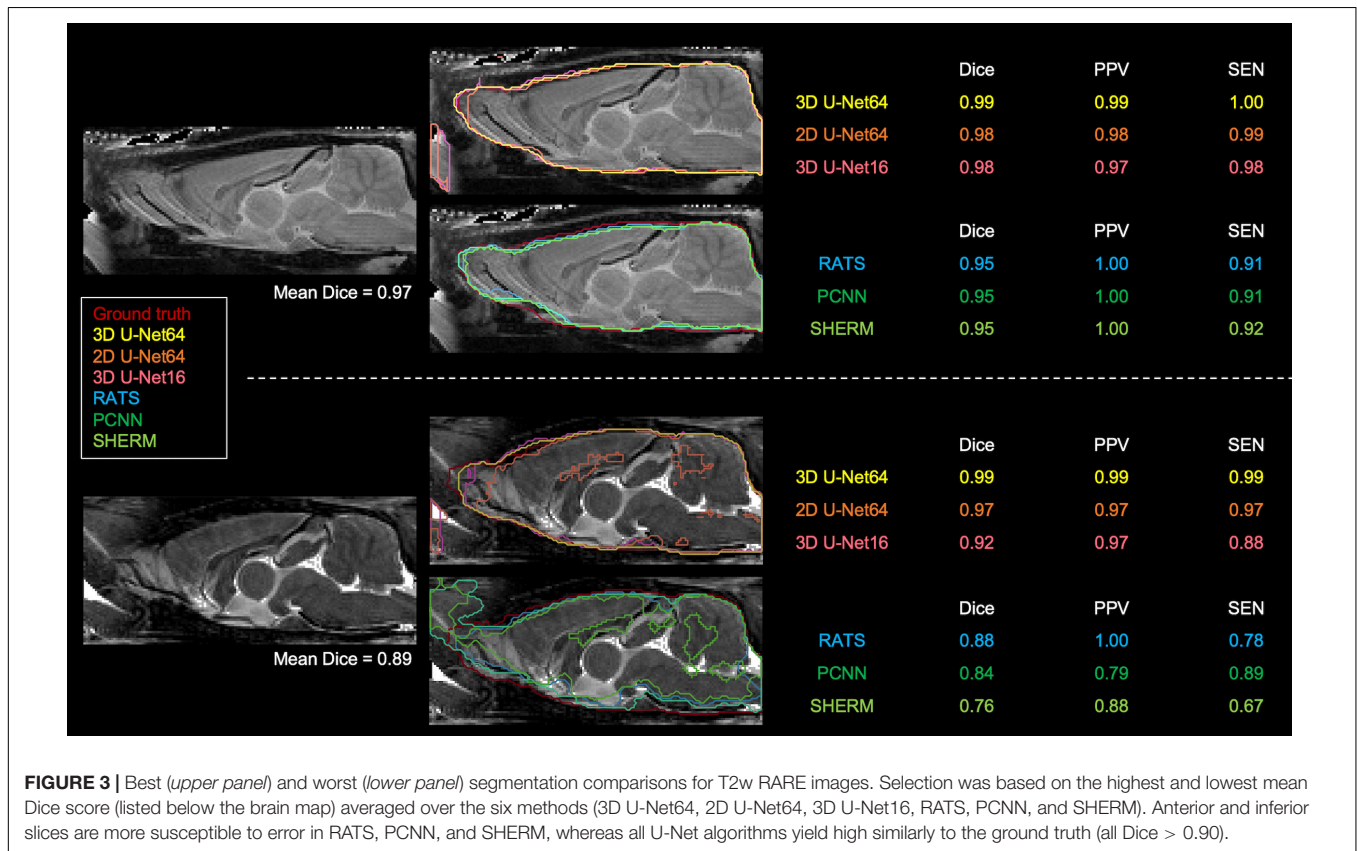
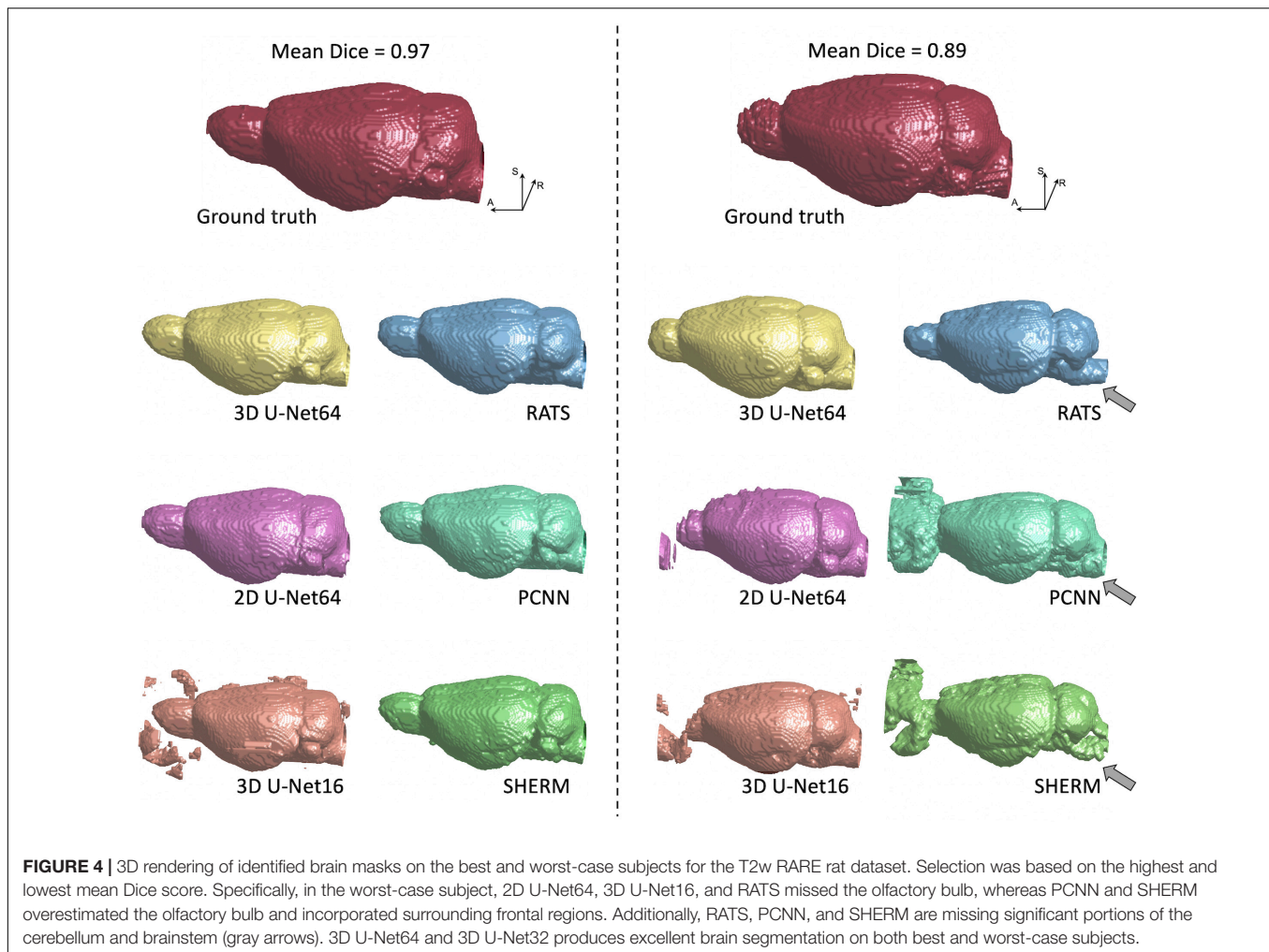


FIGURE 3 | Best (upper panel) and worst (lower panel) segmentation comparisons for T2w RARE images. Selection was based on the highest and lowest mean Dice score (listed below the brain map) averaged over the six methods (3D U-Net64, 2D U-Net64, 3D U-Net16, RATS, PCNN, and SHERM). Anterior and inferior slices are more susceptible to error in RATS, PCNN, and SHERM, whereas all U-Net algorithms yield high similarly to the ground truth (all Dice > 0.90).

Next, we evaluated the performance of 3D U-Net64 with regards to the model-training and model-validating sample sizes used during the model training process. We re-trained the 3D U-Net64 model with randomly selected training subgroups. Compared to the “standard model” trained on 55 subjects and validated on 14 subjects, the model with training rats < 25

showed significantly decreased accuracy for T2w RARE images, and the model with training rats < 50 showed significantly decreased accuracy for T2*w EPI images (Figure 7). However, it is worth noting that the model still reached excellent and stable segmentation performance (Dice > 0.95) with training subjects ≥ 10 for both T2w RARE and T2*w EPI MRI data.



There were no significant differences in each training sample selection across varied validation sample sizes in T2w RARE and T2*w EPI image segmentation, suggesting that a large number of model-validation subjects is perhaps unnecessary.

DISCUSSION

U-Net is a neural network that is mainly used for classification and localization (Ronneberger et al., 2015; Çiçek et al., 2016). In this study, we proposed a 3D U-Net architecture for automatic rat brain segmentation. Our results indicate that our proposed brain extraction framework based on 3D U-Net64 represents a robust method for the accurate and automatic extraction of rat brain tissue from MR images.

In this study where isotropic T2w RARE and T2* EPI data were segmented, RATS, PCNN, and SHERM still performed quickly and accurately (Dice > 0.8). 3D U-Net64 showed superior performance (average Dice = 0.99) over these methods and also outperformed our previously proposed 2D U-Net64 [Hsu et al., 2020; please also refer to Hsu et al. (2020) for detailed discussion about RATS, PCNN, and SHERM]. The 3D U-Net

algorithm remains “parameter free” in the segmentation process, as all parameters are automatically learned from the data itself. It should be noted that the U-Net architecture generally requires longer processing times and needs a higher level of computational power for model-training. For single patch inference, 2D U-Net takes less time than 3D U-Net. However, the patch size of 3D U-Net64 is one more dimension than that of 2D U-Net64, so to obtain the whole volume prediction, 3D U-Net needs less inference time and less time overall.

3D U-Net is a variant of 2D U-Net where the inputs are 3D volume (Çiçek et al., 2016). It has the same encoder and decoder structure as in 2D U-Net. However, the encoder uses 3D convolution followed by 3D max-pooling to extract the features and the decoder uses 3D upsampling to reconstruct the annotated images (Çiçek et al., 2016). The key advantage of 3D U-Net is its ability to utilize interslice contextual information. The 3D U-Net16 model tends to have higher SEN but lower PPV than the 2D U-Net64 model, as it misclassifies more background areas as brain (i.e., has more false positives). In other words, 2D U-Net64 may achieve higher precision because it makes fewer false positive errors. Although 3D U-Net16 can extract 3D context from the additional dimension that 2D U-Net64 lacks, its

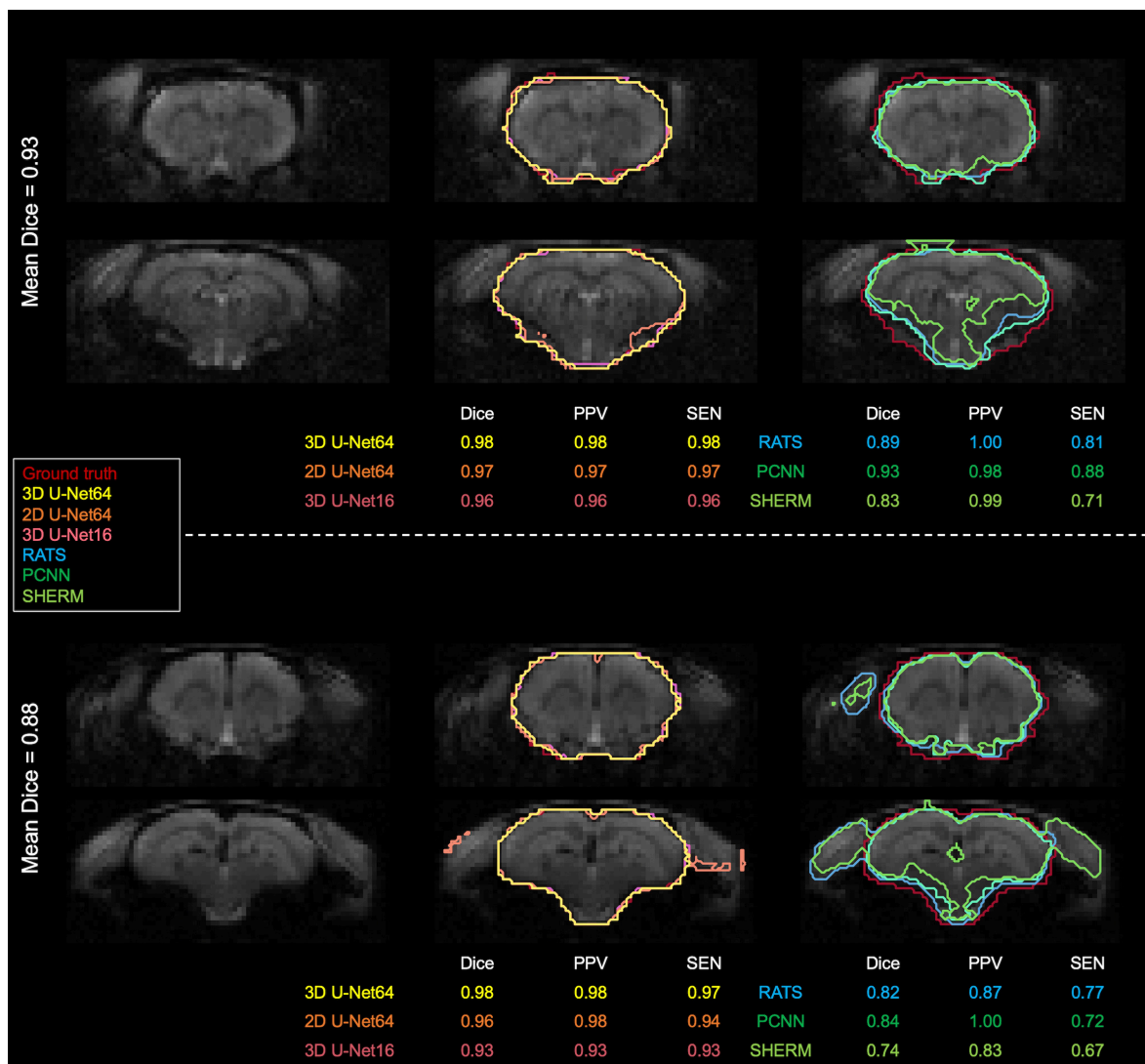


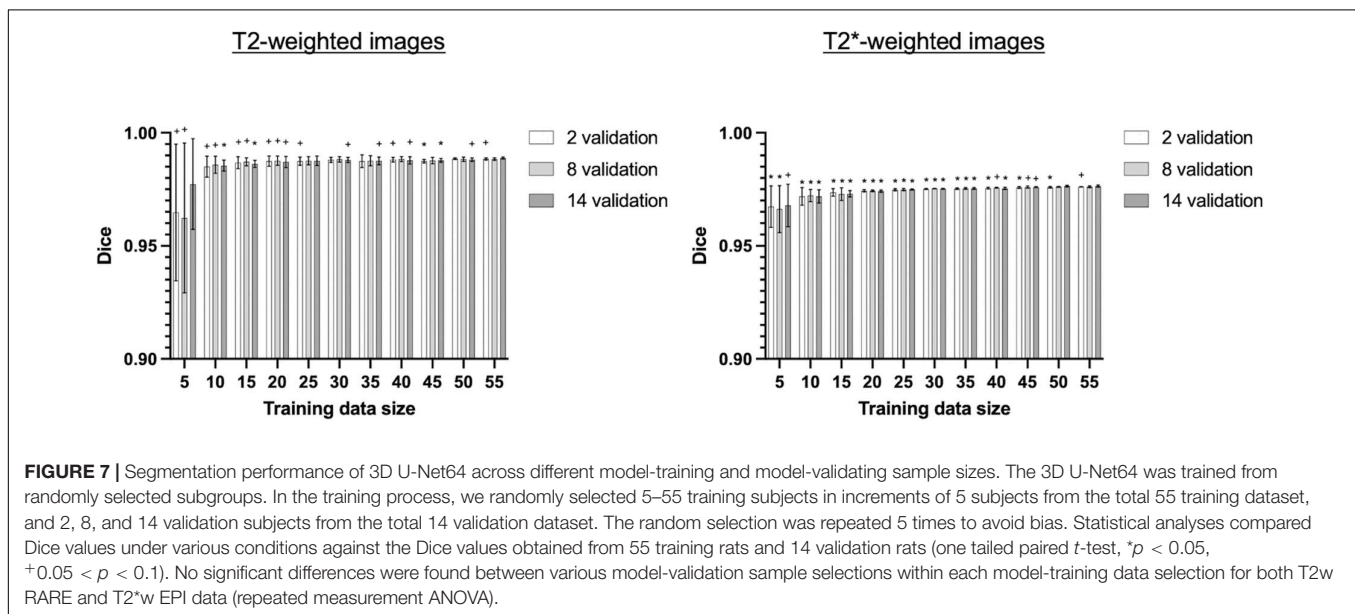
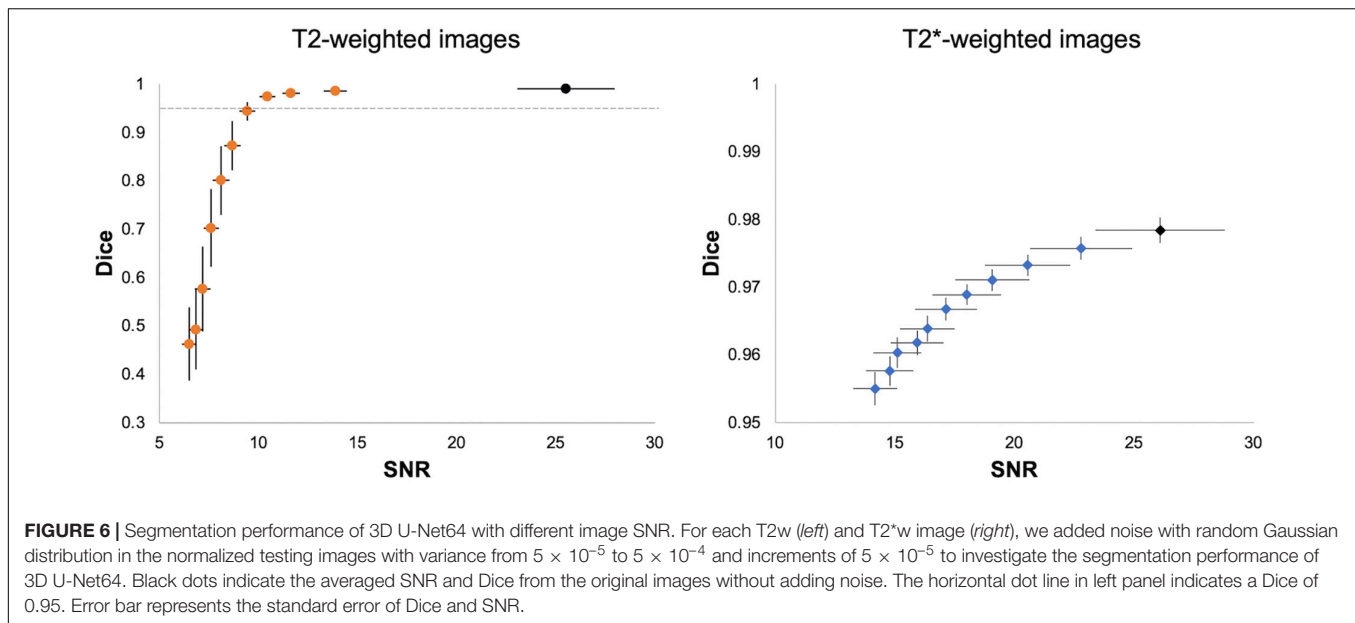
FIGURE 5 | Best and worst segmentation comparisons for T2*w EPI images. Selection was based on the highest and lowest mean Dice score (listed above the brain map) averaged over the six methods (3D U-Net64, 2D U-Net64, 3D U-Net16, RATS, PCNN, and SHERM). Posterior and inferior slices are more susceptible to error in RATS, PCNN, and SHERM, whereas all U-Net algorithms are more similar to the ground truth (all Dice > 0.90).

original 2D dimension resolution is also lost which means that the receptive field is decreased. Therefore, 2D U-Net64 achieves better performance than 3D U-Net16.

To illustrate the reliability of our proposed 3D U-Net64 architecture, we examined its accuracy under two conditions: conducting different Gaussian noise in the testing images and using a different number of rats from the original cohort in the training process. The results highlight the stability and robustness of 3D U-Net64 in segmenting the rat brain in both isotropic T2w RARE and T2*w EPI data. It is encouraging to observe excellent 3D U-Net64 performance (Dice > 0.95), even when the original image SNR is halved. Such performance is comparably higher than the conventional non-UNet-based methods using original SNR (RATS, PCNN, and SHERM). This demonstrates the noise-resistant capacity of 3D U-Net64 and suggests this approach may

handle MRI data over a wide range of SNR quality. In addition, we also retrained the 3D U-Net64 model using a different number of subjects from the original cohort in the training process. This information is useful since the number of samples in a dataset are usually limited in rodent MRI studies. The Dice coefficient increases and stabilizes rapidly with >10 training subjects for both T2w RARE and T2*w EPI. This finding demonstrates the utility of 3D U-Net64 with a limited number of training datasets.

There are several limitations of the 3D U-Net architecture. First, deep learning is a data driven classification, so segmentation accuracy highly relies on the training dataset. Because we trained our 3D U-Net algorithm by using only T2w RARE and T2*w EPI images in rats, additional training and optimization will be needed to segment brain images with different contrast (e.g., T1-weighted images). Second, patch-based training could lose



information/segmentation consistency or overfit the data if the patch size and number of training samples are imbalanced. The training sample imbalance may be alleviated by using randomly cropped patches to enhance sample diversity and prevent overfitting. Loss of information could be addressed by using a larger patch size. In this study, the image size of our whole rat brain T2w RARE data is $144 \times 144 \times 64$ with resolution of 0.2 mm. Accordingly, we chose the patch size to be $64 \times 64 \times 64$ ($\sim 2,097 \text{ mm}^3$), which could cover a significant portion of the rat brain ($\sim 2,000 \text{ mm}^3$). This patch size was limited by the slice dimension of the imaging matrix. Considering computation efficiency, we chose this patch size to also balance GPU demands for the users of this tool. By extracting

overlapped patches with a $16 \times 16 \times 16$ stride and using the averaged predictions for the final segmentation, we reduced segmentation inconsistency between neighboring patches and mitigated information loss. Our results support the use of a large patch size as we demonstrated superior segmentation performance at $64 \times 64 \times 64$ patch in 3D U-Net than the $16 \times 16 \times 16$ patch. Third, deep learning methods require substantial amounts of manually labeled data (Verbraeken et al., 2020), and their performance can be affected by similarities between the training dataset and the data entering the U-Net brain segmentation process. While more training datasets of varying quality may further improve our current 3D U-Net model, our validation analysis suggested reliable segmentation

performance can be reached with ≥ 10 training subjects. This indicates that the model can be established on a laboratory- or scanner-level basis. Next, our current 3D U-Net architecture uses a patch size of $64 \times 64 \times 64$. As discussed above, this limits the testing image to an image matrix size of at least $64 \times 64 \times 64$. Image resampling to a finer resolution is required if the image matrix size is smaller than $64 \times 64 \times 64$. For a dataset at lower spatial resolution, a patch size of $16 \times 16 \times 16$ as shown in this manuscript could be considered, although inferior performance may be expected. Alternatively, the source codes provided herein can be modified to adapt other patch sizes (such as $32 \times 32 \times 32$) per the user's discretion. For datasets with highly anisotropic resolution (e.g., high in-plane resolution with very few slices), because our results show the 2D U-Net (64×64) significantly outperformed than the 3D U-Net ($16 \times 16 \times 16$), we suggest that the users consider 2D U-Net as described in Hsu et al. (2020). Finally, it should be noted that several advances in U-Net have been proposed, and thus the approach described herein may be improved further (e.g., nnU-Net, cascaded U-Net, U-Net++, double U-Net, and recurrent residual U-Net [R2U-Net]) (Alom et al., 2018; Zhou et al., 2018; Liu et al., 2019; Jha et al., 2020; Isensee et al., 2021). Additionally, several other deep learning-based methods, such as deep attention fully convolutional neural network (FCN) and Mask region-based convolutional neural network (R-CNN), have shown great promises for human medical image segmentation (Liu et al., 2018; Mishra et al., 2018; Chiao et al., 2019; Lei et al., 2020). Although one could fine-tune parameters and apply domain adaption strategies (Ghafoorian et al., 2017; Gholami et al., 2018) to attempt the use of human brain segmentation tools in rodents, no literature has documented success with this approach to the best of our knowledge, likely due to the differences in tissue geometry and MRI contrast features between human and rodent brains (Van Essen et al., 2013; Ma et al., 2018; Grandjean et al., 2020; Lee et al., 2021). Our study describes a step-by-step pipeline for extracting rat brain from isotropic MRI data. We focused primarily on 3D vs. 2D U-Net (Hsu et al., 2020) as well as comparisons against other established techniques developed specifically for rodent brain MRI (Chou et al., 2011; Oguz et al., 2014; Liu et al., 2020). Although additional comparisons with recently developed deep learning-based brain segmentation tools are beyond the scope of the present study, adapting those tools for rodent applications and establishing comprehensive evaluations should bring significant impact to the field. It is our hope that the work presented herein will facilitate such studies in the future.

CONCLUSION

We proposed a 3D U-Net model, an end-to-end deep learning-based segmentation algorithm for brain extraction of 3D high resolution volumetric brain MRI data. The method is fully automated and demonstrates accurate brain mask delineations for isotropic structural (T2w RARE) and functional (T2*w EPI) MRI data of the rat brain. The method was compared against other techniques commonly used for rodent brain extraction

as well as 2D U-Net. 3D U-Net shows superior performance in qualitative metrics including Dice, Jaccard, PPV, SEN, CMD, and Hausdorff distance. We believe this tool will be useful to avoid parameter-selection bias and streamline pre-processing steps when analyzing high resolution 3D rat brain MRI data. The 3D U-Net brain extraction tool can be found at <https://camri.org/dissemination/software/>.

DATA AVAILABILITY STATEMENT

The CAMRI rats dataset is available at doi: 10.18112/openneuro.ds003646.v1.0.0. The 3D U-Net brain extraction tool can be found at <https://camri.org/dissemination/software/>. The 3D U-Net results of testing data are available at doi: 10.17632/gmcf67fw5h.1. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The animal study was reviewed and approved by Institutional Animal Care and Use Committee at the University of North Carolina (UNC) at Chapel Hill.

AUTHOR CONTRIBUTIONS

L-MH and Y-YS designed the study. L-MH and SW implemented U-Net algorithm for rodents and validated the developed methods on various datasets. T-WW and S-HL provided ground-truth brain masks. LW provided data and helped to edit the manuscript. S-HL managed data and software dissemination and helped to design the study. L-MH, SW, and Y-YS wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by National Institute of Neurological Disorders and Stroke (R01NS091236 and R44NS105500), National Institute of Mental Health (R01MH126518, RF1MH117053, R01MH111429, and F32MH115439), National Institute on Alcohol Abuse and Alcoholism (P60AA011605 and U01AA020023), and National Institute of Child Health and Human Development (P50HD103573).

ACKNOWLEDGMENTS

We thank CAMRI members for insightful discussion on this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2021.801008/full#supplementary-material>

REFERENCES

- Alom, M. Z., Hasan, M., Yakopcic, C., Taha, T. M., and Asari, V. K. (2018). Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. *arXiv [preprint]*. doi: 10.1109/NAECON.2018.8556686
- Avants, B. B., Tustison, N., and Song, G. (2009). Advanced normalization tools (ANTS). *Insight J.* 2, 1–35.
- Babalola, K. O., Patenaude, B., Aljabar, P., Schnabel, J., Kennedy, D., Crum, W., et al. (2009). An evaluation of four automatic methods of segmenting the subcortical structures in the brain. *Neuroimage* 47, 1435–1447. doi: 10.1016/j.neuroimage.2009.05.029
- Chiao, J.-Y., Chen, K.-Y., Liao, K. Y.-K., Hsieh, P.-H., Zhang, G., and Huang, T.-C. (2019). Detection and classification the breast tumors using mask R-CNN on sonograms. *Medicine* 98:e15200. doi: 10.1097/MD.00000000000015200
- Chollet, F. (2015). *Keras Documentation [WWW Document]*. San Francisco: GitHub
- Chou, N., Wu, J., Bai Bingren, J., Qiu, A., and Chuang, K.-H. (2011). Robust automatic rodent brain extraction using 3-D pulse-coupled neural networks (PCNN). *IEEE Trans. Image Process.* 20, 2554–2564. doi: 10.1109/TIP.2011.2126587
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, Lecture Notes in Computer Science*, eds S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells (Cham: Springer International Publishing), 424–432. doi: 10.1007/978-3-319-46723-8_49
- Fatima, A., Shahid, A. R., Raza, B., Madni, T. M., and Janjua, U. I. (2020). State-of-the-Art Traditional to the Machine- and Deep-Learning-Based Skull Stripping Techniques, Models, and Algorithms. *J. Digit. Imaging* 33, 1443–1464. doi: 10.1007/s10278-020-00367-5
- Feo, R., and Giove, F. (2019). Towards an efficient segmentation of small rodents brain: a short critical review. *J. Neurosci. Methods* 323, 82–89. doi: 10.1016/j.jneumeth.2019.05.003
- Gaser, C., Schmidt, S., Metzler, M., Herrmann, K.-H., Krumbein, I., Reichenbach, J. R., et al. (2012). Deformation-based brain morphometry in rats. *Neuroimage* 63, 47–53. doi: 10.1016/j.neuroimage.2012.06.066
- Ghafoorian, M., Mehrtash, A., Kapur, T., Karsemeyer, N., Marchiori, E., and Pesteie, M. (2017). “Transfer learning for domain adaptation in MRI: Application in brain lesion segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, (Quebec City: Springer), 516–524.
- Gholami, A., Subramanian, S., and Shenoy, V. (2018). “A novel domain adaptation framework for medical image segmentation,” in *International MICCAI Brainlesion Workshop 2018*. (Cham: Springer) 289–298
- Grandjean, J., Canella, C., Anckaerts, C., Ayranc, G., Bougacha, S., Bienert, T., et al. (2020). Common functional networks in the mouse brain revealed by multi-centre resting-state fMRI analysis. *Neuroimage* 205:116278. doi: 10.1016/j.neuroimage.2019.116278
- Hsu, L.-M., Wang, S., Ranadive, P., Ban, W., Chao, T.-H. H., Song, S., et al. (2020). Automatic Skull Stripping of Rat and Mouse Brain MRI Data Using U-Net. *Front. Neurosci.* 14:568614. doi: 10.3389/fnins.2020.568614
- Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., and Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* 18, 203–211. doi: 10.1038/s41592-020-01008-z
- Jha, D., Riegler, M. A., and Johansen, D. (2020). “Doubleu-net: A deep convolutional neural network for medical image segmentation,” in *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, (Piscataway: IEEE), 558–564.
- Kingma, D. P., and Ba, J. L. (2015). “Adam: A method for stochastic optimization,” in *3rd International Conference on Learning Representations (ICLR)*, (Ithaca: arXiv.org), 13.
- Kleesiek, J., Urban, G., Hubert, A., Schwarz, D., Maier-Hein, K., Bendszus, M., et al. (2016). Deep MRI brain extraction: a 3D convolutional neural network for skull stripping. *Neuroimage* 129, 460–469. doi: 10.1016/j.neuroimage.2016.01.024
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing System*, eds F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Red Hook: Curran Associates, Inc).
- Lee, S.-H., Broadwater, M. A., Ban, W., Wang, T.-W. W., Kim, H.-J., Dumas, J. S., et al. (2021). An isotropic EPI database and analytical pipelines for rat brain resting-state fMRI. *Neuroimage* 243:118541. doi: 10.1016/j.neuroimage.2021.118541
- Lei, Y., Dong, X., Tian, Z., Liu, Y., Tian, S., Wang, T., et al. (2020). CT prostate segmentation based on synthetic MRI-aided deep attention fully convolution network. *Med. Phys.* 47, 530–540. doi: 10.1002/mp.13933
- Liu, H., Shen, X., Shang, F., Ge, F., and Wang, F. (2019). “CU-Net: Cascaded U-Net with Loss Weighted Sampling for Brain Tumor Segmentation,” in *Multimodal Brain Image Analysis and Mathematical Foundations of Computational Anatomy: 4th International Workshop, MBIA 2019, and 7th International Workshop, MFCA 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings, Lecture Notes in Computer Science*, eds D. Zhu, J. Yan, H. Huang, L. Shen, P. M. Thompson, and C.-F. Westin (Cham: Springer International Publishing), 102–111. doi: 10.1007/978-3-030-33226-6_12
- Liu, M., Dong, J., Dong, X., and Yu, H. (2018). “Segmentation of lung nodule in CT images based on mask R-CNN,” in *2018 9th International Conference on Awareness Science and Technology, iCAST*, (Piscataway: IEEE), 95–100.
- Liu, Y., Unsal, H. S., Tao, Y., and Zhang, N. (2020). Automatic brain extraction for rodent MRI images. *Neuroinformatics* 18, 395–406. doi: 10.1007/s12021-020-09453-z
- Lu, H., Scholl, C. A., Zuo, Y., Demny, S., Rea, W., Stein, E. A., et al. (2010). Registering and analyzing rat fMRI data in the stereotaxic framework by exploiting intrinsic anatomical features. *Magn. Reson. Imaging* 28, 146–152. doi: 10.1016/j.mri.2009.05.019
- Ma, Z., Perez, P., Ma, Z., Liu, Y., Hamilton, C., Liang, Z., et al. (2018). Functional atlas of the awake rat brain: a neuroimaging study of rat brain specialization and integration. *Neuroimage* 170, 95–112. doi: 10.1016/j.neuroimage.2016.07.007
- Mandino, F., Cerri, D. H., Garin, C. M., Straathof, M., van Tilborg, G. A. F., Chakravarty, M. M., et al. (2019). Animal functional magnetic resonance imaging: trends and path toward standardization. *Front. Neuroinform.* 13:78. doi: 10.3389/fninf.2019.00078
- Martin, A., Paul, B., Jianmin, C., Zhifeng, C., and Andy, D. (2016). “Tensorflow: A system for large-scale machine learning,” in *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*, (California: USENIX Association), 265–283.
- Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation,” in *2016 Fourth International Conference on 3D Vision (3DV)*, (Piscataway: IEEE), 565–571. doi: 10.1109/3DV.2016.79
- Mishra, D., Chaudhury, S., Sarkar, M., and Soin, A. S. (2018). Ultrasound Image Segmentation: a Deeply Supervised Network with Attention to Boundaries. *IEEE Trans. Biomed. Eng.* 66, 1637–1648. doi: 10.1109/TBME.2018.2877577
- Oguz, I., Zhang, H., Rumpel, A., and Sonka, M. (2014). RATS: rapid Automatic Tissue Segmentation in rodent brain MRI. *J. Neurosci. Methods* 221, 175–182. doi: 10.1016/j.jneumeth.2013.09.021
- Paxinos, G., and Watson, C. (2007). *The Rat Brain in Stereotaxic Coordinates: Hard Cover Edition*, 6th Edn. Amsterdam: Academic Press.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI), Lecture Notes in Computer Science*, eds N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi (Cham: Springer International Publishing), 234–241. doi: 10.1007/978-3-319-24574-4_28
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., et al. (2010). N4ITK: improved N3 bias correction. *IEEE Trans. Med. Imaging* 29, 1310–1320. doi: 10.1109/TMI.2010.2046908
- Uhlisch, M., Greiner, R., Hoehn, B., Woghiren, M., Diaz, I., Ivanova, T., et al. (2018). Improved Brain Tumor Segmentation via Registration-Based Brain Extraction. *Forecasting* 1, 59–69. doi: 10.3390/forecast1010005
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., et al. (2013). The WU-Minn Human Connectome Project: an overview. *Neuroimage* 80, 62–79. doi: 10.1016/j.neuroimage.2013.05.041
- Verbraken, J., Wolting, M., Katzy, J., Kloppenburg, J., Verbelen, T., and Rellermeyer, J. S. (2020). A survey on distributed machine learning. *ACM Comput. Surv.* 53, 1–33. doi: 10.1145/3377454

- Wang, S., Nie, D., Qu, L., Shao, Y., Lian, J., Wang, Q., et al. (2020). CT male pelvic organ segmentation via hybrid loss network with incomplete annotation. *IEEE Trans. Med. Imaging* 39, 2151–2162. doi: 10.1109/TMI.2020.2966389
- Wang, X., Li, X.-H., Cho, J. W., Russ, B. E., Rajamani, N., Omelchenko, A., et al. (2021). U-net model for brain extraction: trained on humans for transfer to non-human primates. *Neuroimage* 235:118001. doi: 10.1016/j.neuroimage.2021.118001
- Yogananda, C. G. B., Wagner, B. C., Murugesan, G. K., Madhuranthakam, A., and Maldjian, J. A. (2019). “A deep learning pipeline for automatic skull stripping and brain segmentation,” in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, (Piscataway: IEEE), 727–731. doi: 10.1109/ISBI.2019.8759465
- Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., and Liang, J. (2018). “UNet++: A Nested U-Net Architecture for Medical Image Segmentation,” in *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support*, ed. D. Stoyanov (Cham: Springer), 3–11. doi: 10.1007/978-3-030-00889-5_1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Hsu, Wang, Walton, Wang, Lee and Shih. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.