3D Capsule Networks for Brain Image Segmentation

A. Avesta, Y. Hui, M. Aboian, J. Duncan, H.M. Krumholz and S. Aneja

AJNR Am J Neuroradiol published online 20 April 2023
http://www.ajnr.org/content/early/2023/04/20/ajnr.A7845

Published April 20, 2023 as 10.3174/ajnr.A7845

ORIGINAL RESEARCH
ADULT BRAIN

3D Capsule Networks for Brain Image Segmentation

A. Avesta, Y. Hui, M. Aboian, J. Duncan, H.M. Krumholz, and S. Aneja

ABSTRACT

BACKGROUND AND PURPOSE: Current autosegmentation models such as UNets and nnUNets have limitations, including the inability to segment images that are not represented during training and lack of computational efficiency. 3D capsule networks have the potential to address these limitations.

MATERIALS AND METHODS: We used 3430 brain MRIs, acquired in a multi-institutional study, to train and validate our models. We compared our capsule network with standard alternatives, UNets and nnUNets, on the basis of segmentation efficacy (Dice scores), segmentation performance when the image is not well-represented in the training data, performance when the training data are limited, and computational efficiency including required memory and computational speed.

RESULTS: The capsule network segmented the third ventricle, thalamus, and hippocampus with Dice scores of 95%, 94%, and 92%, respectively, which were within 1% of the Dice scores of UNets and nnUNets. The capsule network significantly outperformed UNets in segmenting images that were not well-represented in the training data, with Dice scores 30% higher. The computational memory required for the capsule network is less than one-tenth of the memory required for UNets or nnUNets. The capsule network is also >25% faster to train compared with UNet and nnUNet.

CONCLUSIONS: We developed and validated a capsule network that is effective in segmenting brain images, can segment images that are not well-represented in the training data, and is computationally efficient compared with alternatives.

ABBREVIATIONS: CapsNet = capsule network; Conv1 = first network layer made of convolutional operators; ConvCaps3 = third network layer made of convolutional capsules; ConvCaps4 = fourth network layer made of convolutional capsules; DeconvCaps8 = eighth network layer made of deconvolutional capsules; FinalCaps13 = final thirteenth network layer made of capsules; FinalCaps13 = final layer; GPU = graphics processing unit; PrimaryCaps2 = second network layer made of primary capsules

Neuroanatomic image segmentation is an important component in the management of various neurologic disorders.[1-3] Accurate segmentation of anatomic structures on brain MRIs is an essential step in a variety of neurosurgical and radiation therapy procedures.[1,3-6] Manual segmentation is time-consuming and is prone to intra- and interobserver variability.[7,8] With the advent of deep learning to automate various image-analysis tasks,[9,10] there has been increasing enthusiasm for using deep learning for brain image autosegmentation.[11-14]

UNets are among the most popular and successful deep learning autosegmentation algorithms.[11,15-17] Despite the broad success of UNets in segmenting anatomic structures across various imaging modalities, they have well-described limitations. UNets perform best on images that closely resemble the images used for training but underperform on images that contain variant

Received September 13, 2022; accepted after revision March 11, 2023.

From the Department of Radiology and Biomedical Imaging (A.A., M.A., J.D.), Department of Therapeutic Radiology (A.A., Y.H., S.A.), Center for Outcomes Research and Evaluation (A.A., Y.H., H.M.K., S.A.), and Division of Cardiovascular Medicine (H.M.K.), Yale School of Medicine, New Haven, Connecticut; and Departments of Statistics and Data Science (J.D.) and Biomedical Engineering (J.D., S.A.), Yale University, New Haven, Connecticut.

Arman Avesta is a PhD student in the Investigative Medicine Program at Yale, which is supported by Clinical and Translational Science Awards grant No. UL1 TR001863 from the National Center for Advancing Translational Science, a component of the National Institutes of Health (NIH). This work was also directly supported by the National Center for Advancing Translational Sciences grant number KL2 TR001862 as well as by the Radiological Society of North America's (RSNA) Fellow Research Grant Number RF2212. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of NIH or RSNA.

The investigators within the Alzheimer's Disease Neuroimaging Initiative contributed to the design and implementation of Alzheimer's Disease Neuroimaging Initiative but did not participate in the analysis or writing of this article.

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of NIH.

Please address correspondence to Sanjay Aneja, MD, Department of Therapeutic Radiology, Yale School of Medicine, Center for Outcomes Research and Evaluation, 195 Church St, 6th Floor, New Haven, CT 06510; e-mail: sanjay.aneja@yale.edu; @SanjayAnejaMD

Indicates open access to non-subscribers at www.ajnr.org

Indicates article with online supplemental data.

http://dx.doi.org/10.3174/ajnr.A7845

Table 1: Study participants tabulated by the training, validation, and test sets

Data Partitions	No. of MR Imaging Volumes	No. of Patients	Age (mean) (yr)	Sex	Diagnosis
Training set	3199	841	76 (SD, 7)	42% F, 58% M	29% CN, 54% MCI, 17% AD
Validation set	117	30	75 (SD, 6)	30% F, 70% M	21% CN, 59% MCI, 20% AD
Test set	114	30	77 (SD, 7)	33% F, 67% M	27% CN, 47% MCI, 26% AD

Note:—F indicates female; M, male; CN, cognitively normal; MCI, mild cognitive impairment; AD, Alzheimer disease.

anatomy or pathologies that change the appearance of normal anatomy.[8] Additionally, UNets have a large number of trainable parameters; hence, training and deploying UNets for image segmentation often requires substantial computational resources that may not be scalable in all clinical settings.[15] There is a need for fast, computationally efficient segmentation algorithms that can segment images not represented in the training data with high fidelity.

Capsule networks (CapsNets) represent an alternative autosegmentation method that can potentially overcome the limitations of UNets.[18-20] CapsNets can encode and manipulate spatial information such as location, rotation, and size about structures within an image and use this spatial information to produce accurate segmentations. Encoding spatial information allows CapsNets to well generalize on images that are not effectively represented in the data used to train the algorithm.[19,20] Moreover, CapsNets use a smarter paradigm for information encoding, which relies on fewer parameters leading to increased computational efficiency.[18-20]

Capsule networks have shown promise on some biomedical imaging tasks[20] but have yet to be fully explored for segmenting anatomic structures on brain MRIs. In this study, we explore the utility of CapsNets for segmenting anatomic structures on brain MRIs using a multi-institutional data set of >3000 brain MRIs. We compare the segmentation efficacy and computational efficiency of CapsNets with popular UNet-based models.

MATERIALS AND METHODS

Data Set

The data set for this study included 3430 T1-weighted brain MR images belonging to 841 patients from 19 institutions enrolled in the Alzheimer's Disease Neuroimaging Initiative study.[21] The inclusion criteria of the Alzheimer's Disease Neuroimaging Initiative have been previously described.[22] On average, each patient underwent 4 MRI acquisitions. Details of MRI acquisition parameters are provided in the Online Supplemental Data.[21] We randomly split the patients into training (3199 MRI, 93% of the data), validation (117 MR imaging volumes, 3.5% of the data), and test (114 MRI volumes, 3.5% of the data) sets. Data were divided at the patient level to assure that all images belonging to a patient were assigned to either the training, validation, or test set. Patient demographics are provided in Table 1. This study was approved by the institutional review board of Yale School of Medicine (No. 2000027592).

Anatomic Segmentations

We trained our models to segment 3 anatomic structures of the brain: the third ventricle, thalamus, and hippocampus. These structures were chosen to represent structures with varying degrees of segmentation difficulty. Preliminary ground truth segmentations were initially generated using FreeSurfer (http://surfer.nmr.mgh.harvard.edu)[23-25] and then manually corrected by 1 board-eligible radiologist with 9 years of experience in brain image analysis. The Online Supplemental Data detail the process by which ground truth segmentations were established.

Image Preprocessing

MR imaging preprocessing included correction for intensity inhomogeneities, including B_1 field variations.[26,27] We used FSL's Brain Extraction Tool (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET) to remove the skull, face, and neck tissues, resulting in the extracted 3D image of the brain.[28,29] To overcome memory limitations, we performed segmentations on $64 \times 64 \times 64$ voxel patches of the MR imaging volume that contained the segmentation target. The patch was automatically placed over the expected location of the segmentation target using predefined coordinates referenced from the center of the image. The coordinates of each patch were computed during training and were fixed during testing, without any manual input and without using the ground truth segmentations. Details of preprocessing are provided in the Online Supplemental Data.

CapsNets

CapsNets have 3 main components: 1) capsules that each encode a structure together with the pose of that structure: the pose is an n-dimensional vector that learns to encode orientation, size, curvature, location, and other spatial information about the structure; 2) a supervised learning paradigm that learns how to transform the poses of the parts (eg, head and tail of the hippocampus) to the pose of the whole (eg, the entire hippocampus); and 3) a clustering paradigm that detects a whole if the poses of all parts transform into matching poses of the whole. Further details regarding differences between CapsNets and other deep learning models are provided in the Online Supplemental Data.

2D CapsNets were previously introduced by LaLonde et al[20] to segment 1 section of the image at a time. We developed 3D CapsNets for volumetric segmentation of a 3D volume, with the architecture shown in Fig 1*A*.[20] We developed 3D CapsNets for volumetric segmentation of a 3D volume, with the architecture shown in Fig 1*A*. The first layer, Conv1, performs 16 convolutions ($5 \times 5 \times 5$) on the input volume to generate 16 feature volumes, which are reshaped into 16D vectors at each voxel. The 16D vector at each voxel is reshaped into a pose that learns to encode spatial information at that voxel. The next layer, PrimaryCaps2, has 2 capsule channels that learn two 16D-to-16D convolutional transforms ($5 \times 5 \times 5$) from the poses of the previous-layer parts to the poses of the next-layer wholes. Likewise, all capsule layers (green layers in Fig 1*A*) learn m- to n-dimensional transforms from the poses of parts to the poses of wholes.

Our CapsNet has downsampling and upsampling limbs. The downsampling limb learns what structure is present at each voxel, and the skip connections from downsampling to upsampling limbs

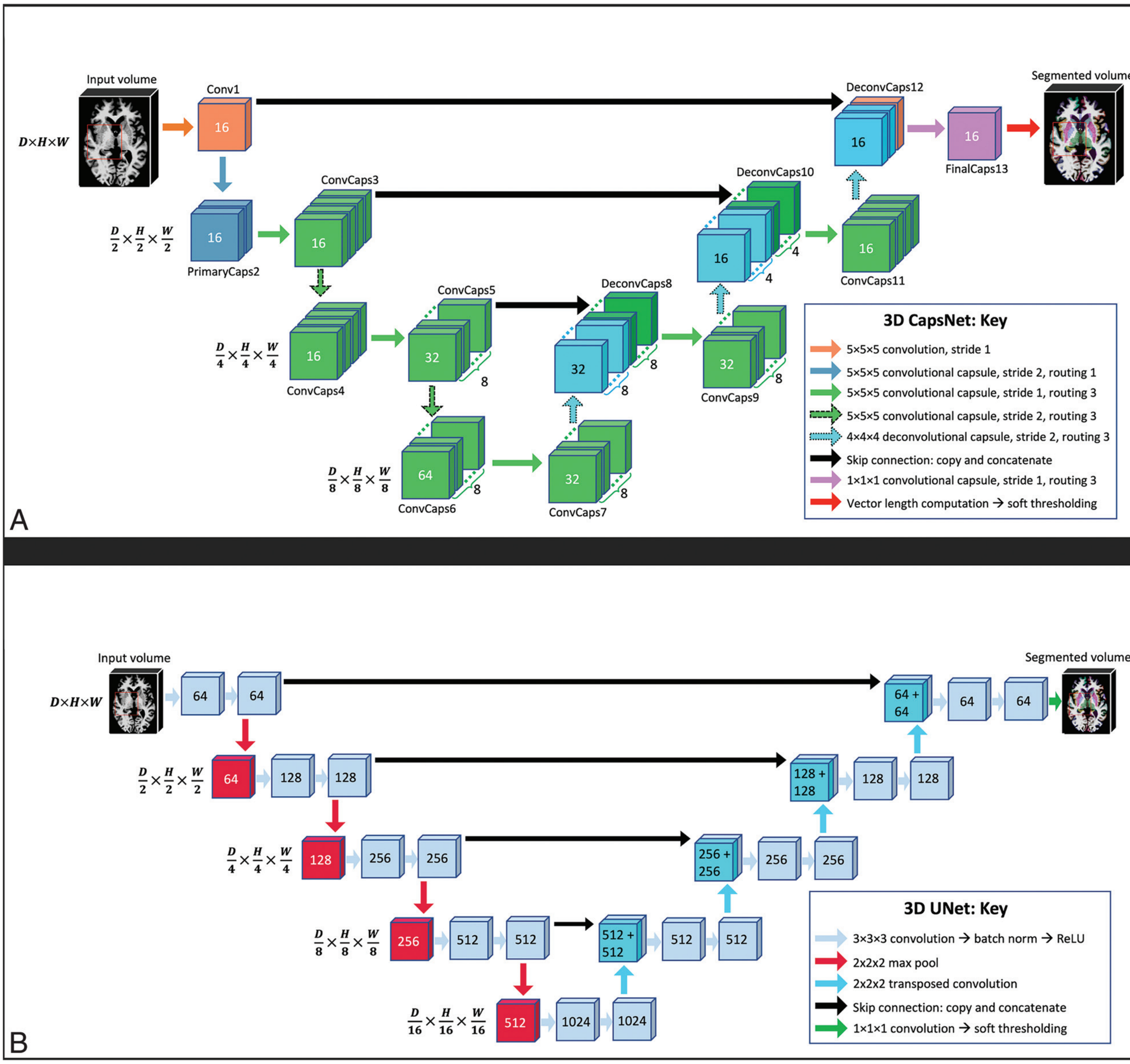

FIG 1. CapsNet (*A*) and UNet (*B*) architectures. The nnUNet architecture was self-configured by the model and is already published.[16] All models process 3D images in all layers, with dimensions shown on the left side. The depth, height, and width of the image in each layer is shown by *D*, *H*, and *W*, respectively. *A*, The number over the Conv1 layer represents the number of channels. The numbers over the capsule layers (ConvCaps, DeconvCaps, and FinalCaps) represent the number of pose components. The stacked layers represent capsule channels. *B*, The numbers over each layer represent the number of channels. In UNet and nnUNet, the convolutions have stride = 1 and the transposed convolutions have stride = 2. Note that the numbers over the capsule layers show the number of pose components, while the numbers over the noncapsule layers show the number of channels.

preserve where each structure is on the image. Downsampling uses $5 \times 5{\times}5$ convolutional transforms with stride = 2. Layers in the deeper parts of CapsNet contain more capsule channels (up to 8) and poses with more components (up to 64) to be able to encode more complex structures, because each capsule in the deeper parts of the model should be able to detect complex concepts in the entire image. Upsampling uses $4 \times 4{\times}4$ transposed convolutional transforms with stride = 2 (turquoise layers in Fig 1*A*). The final layer, FinalCaps13, contains 1 capsule channel that learns to activate capsules within the segmentation target and deactivate them outside the target. The Online Supplemental Data explain the options that we explored for developing our 3D CapsNets and how we chose the best design options. The Online Supplemental Data explain how the final layer activations were converted into segmentations. Details about how the model finds agreeing poses of parts that vote for the pose of the whole are provided in the Online Supplemental Data.

Comparisons: UNets and nnUNets

Optimized 3D UNets and nnUNets were also trained on the same training data,[11-13,30] and their segmentation efficacy and computational efficiency were compared with our CapsNet using the same test data. UNets and nnUNets have shown strong autosegmentation performance across a variety of different imaging modalities and anatomic structures and are among the most commonly used segmentation algorithms in biomedical imaging.[11-13,15,31,32] Figure 1*B* shows the architecture of our UNet. The input image undergoes 64 convolutions ($3 \times 3 \times 3$) to generate 64 feature maps. These maps then undergo batch normalization and rectified linear unit activation. Similar operations are performed again, followed by downampling using max pooling ($2 \times 2 \times 2$). The downsampling and upsampling limbs each include 4 units. Upsampling uses $2 \times 2 \times 2$ transposed convolutions with stride = 2. The final layer performs a $1 \times 1 \times 1$ convolution to aggregate all 64 channels, followed by soft thresholding using the sigmoid function. The model learns to output a number close to 1 for each voxel inside the segmentation target and a number close to zero for each voxel outside the target. We also trained self-configuring nnUNets that automatically learn the best architecture as well as the optimal training hyperparameters.[16]

Model Training

The CapsNet and UNet models were trained for 50 epochs using the Dice loss and the Adam optimizer.[33] The initial learning rate was set at 0.002. We used dynamic paradigms for learning rate scheduling, with a minimal learning rate of 0.0001. The hyperparameters for our UNet were chosen on the basis of the best-performing model over the validation set. The hyperparameters for the nnUNet were self-configured by the model.[16] The training hyperparameters for CapsNet and UNet are detailed in the Online Supplemental Data.

Model Performance

The segmentation efficacy of the 3 models was measured using Dice scores. To compare the performance of each segmentation model when training data are limited, we also trained the models using subsets of the training data with 600, 240, 120, and 60 MRIs. We then compared the segmentation efficacy of the models using the test set. The relative computational efficiency of the models was measured by the following: 1) the computational memory required to run the model (in megabytes), 2) the computational time required for training each model, and 3) the time that each model takes to segment 1 MR imaging volume.

Out-of-Distribution Testing

To evaluate the performance of CapsNet and UNet models on the images that were not represented during training, we trained the models using images of the right hemisphere of the brain that only contained the right thalamus and right hippocampus. Then, we evaluated the segmentation efficacy of the trained models on the images of the left hemisphere of the brain that contained the contralateral left thalamus and left hippocampus. Because the left-hemisphere images in the test set are not represented in the right-hemisphere images in the training set, this experiment evaluates the out-of-distribution performance of the models. We intentionally did not use any data augmentation during training to assess out-of-distribution performance of the models. Given that nnUNet paradigm requires data augmentation, the nnUNet was not included in this experiment. We additionally tested whether the fully-trained models can generalize to segment raw images that did not undergo steps of preprocessing. The Online Supplemental Data summarize the results of these experiments.

Implementation

Images were preprocessed using Python (Version 3.9) and FreeSurfer (Version 7). PyTorch (Version 1.11; https://pytorch.org/) was used for model development and testing. Training and testing of the models were run on graphics processing unit (GPU)-equipped servers (4 virtual CPUs, 61 GB RAM, 12 GB NVIDIA GK210 GPU with Tesla K80 accelerators; https://www.nvidia.com/). The code used to train and test our models, our pretrained models, and a sample MR imaging is available on the GitHub page of our lab (www.github.com/Aneja-Lab-Yale/Aneja-Lab-Public-CapsNet).

RESULTS

All 3 segmentation models showed high performance across all 3 neuroanatomic structures with Dice scores of >90% (Fig 2). Performance was highest for the third ventricle (95%–96%) followed by the thalamus (94%–95%) and hippocampus (92%–93%). Dice scores between the CapsNet and UNet-based models were within 1% for all neuroanatomic structures (Table 2).

Although both CapsNet and UNet had difficulty segmenting contralateral structures, the CapsNet significantly outperformed the UNet (thalamus P value $<$ .001, hippocampus P value $<$ .001) (Table 3). CapsNet models frequently identified the contralateral structure of interest but underestimated the size of the segmentation, resulting in Dice scores between 40% and 60%. In contrast, the UNet models frequently failed to identify the contralateral structure of interest, resulting in Dice scores of <20% (Fig 3).

Segmentation performance for each model remained high across training data sets of varying sizes (Fig 4). When trained on 120 brain MRIs, all three models maintained their segmentation accuracy within 1% compared to models trained on 3199 brain MRIs. However, segmentation performance did decrease for all three models when trained on 60 brain MRIs (83% for CapsNet, 84% for UNet, and 88% for nnUNet).

The CapsNet was more computationally efficient compared with UNet-based models (Fig 5). The CapsNet required 228 MB, compared with 1364 MB for UNet and 1410 MB for nnUNet. The CapsNet trained 25% faster than the UNet (1.5 versus 2 seconds per sample) and 100% faster than the nnUNet (1.5 versus 3 seconds per sample). When we compared the deployment times of the fully-trained models, CapsNet and UNet could segment images equally fast (0.9 seconds per sample), slightly faster than the nnUNet (1.1 second per sample).

DISCUSSION

Neuroanatomic segmentation of brain structures is an essential component in the treatment of various neurologic disorders. Deep learning–based autosegmentation methods have shown the ability to segment brain images with high fidelity, which was previously a time-intensive task.[13,14,17,34] In this study, we compared the segmentation efficacy and computational efficiency of CapsNets with UNet-based autosegmentation models. We

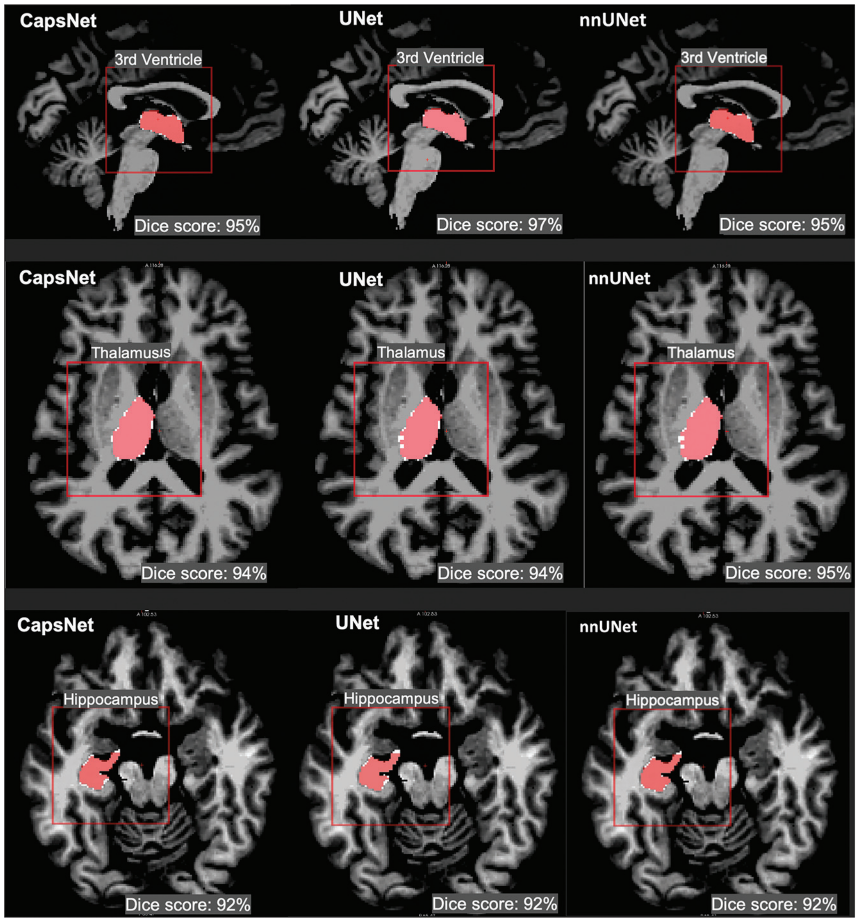

FIG 2. CapsNet, UNet, and nnUNet segmentation of brain structures that were represented in the training data. Segmentations for three structures are shown: third ventricle, thalamus, and hippocampus. Target segmentations and model predictions are, respectively, shown in red and white. Dice scores are provided for the entire volume of the segmented structure in this patient (who was randomly chosen from the test set).

Table 2: Comparing the segmentation efficacy of CapsNets, UNets, and nnUNets in segmenting brain structures that were represented in the training data[a]

Brain Structure	CapsNet Dice (95% CI)	UNet Dice (95% CI)	nnUNet Dice (95% CI)	Repeated Measures ANOVA *P* Value[b]
Third ventricle	95% (94–96)	96% (95–97)	96% (95–97)	.03
Thalamus	94% (93–95)	95% (94–96)	94% (92–96)	.1
Hippocampus	92% (91–93)	93% (92–94)	92% (91–93)	.1

[a] The segmentation accuracy was quantified using Dice scores on the test set (114 brain MRIs). The third ventricle, thalamus, and hippocampus, respectively, represent easy, medium, and difficult structures to segment.
[b] $df = 3 - 1 = 2$.

Table 3: Comparing the efficacy of CapsNets and UNets in segmenting images that were not represented in the training data[a]

Brain Structure	CapsNet Dice (95% CI)	UNet Dice (95% CI)	CapsNet vs UNet (*P* Value[b])
Thalamus	52% (46–58)	16% (11–21)	< .01
Hippocampus	43% (38–48)	10% (6–14)	< .01

[a] Both models were trained to segment the right thalamus and hippocampus. Then, they were tested on segmenting the contralateral left thalamus and hippocampus.
[b] $df = 114 - 1 = 113$.

found CapsNets to be reliable and computationally efficient, achieving segmentation accuracy comparable with commonly used UNet-based models. Moreover, we found CapsNets to have higher segmentation performance on out-of-distribution data, suggesting an ability to generalize beyond their training data.

Our results corroborate previous studies demonstrating the ability of deep learning models to reliably segment anatomic structures on diagnostic images.[11,12,14] UNet-based models have been shown to effectively segment normal anatomy across a variety of different imaging modalities including CT, MR imaging, and x-ray images.[15,31,32,35-37] Moreover, Isensee et al[16] showed the ability of nnUNets to generate reliable segmentations across 23 biomedical image-segmentation tasks with automated hyperparameter optimization. We have extended prior work by demonstrating similar segmentation efficacy between CapsNets and UNet-based models, with CapsNets being notably more computationally efficient. Our CapsNets require <10% of the amount of memory required by UNet-based methods and train 25% faster.

Our findings are consistent with prior studies demonstrating the efficacy of CapsNets for image segmentation.[20,38] LaLonde et al[20] previously demonstrated that 2D CapsNets can effectively segment lung tissues on CT images and muscle and fat tissues on thigh MRIs. Their group similarly found that CapsNets can segment images with performance rivaling UNet-based models while requiring <10% of the memory required by UNet-based models. Our study builds on prior studies by showing the efficacy of CapsNets for segmenting neuroanatomic substructures on brain MRIs. Additionally, when we compared this work with prior work, we have implemented 3D CapsNet architecture, which has not been previously described in the literature.

Previous studies have suggested that CapsNets are able to generalize beyond their training data.[19,20] Hinton et al[19] demonstrated that CapsNets can learn spatial information about the objects in the image and can then generalize this information beyond what is present in the training data, which gives CapsNet out-of-distribution generalization capability. The ability to segment out-of-distribution images was also shown by LaLonde et al[20] for their 2D CapsNet model, which segments images. We built on previous studies by demonstrating out-of-distribution generalizability of 3D CapsNets for segmenting medical images.

Although we found CapsNets to be effective in biomedical image segmentation, previous studies on biomedical imaging have shown mixed results.[38] Survarachakan et al[38] previously found 2D

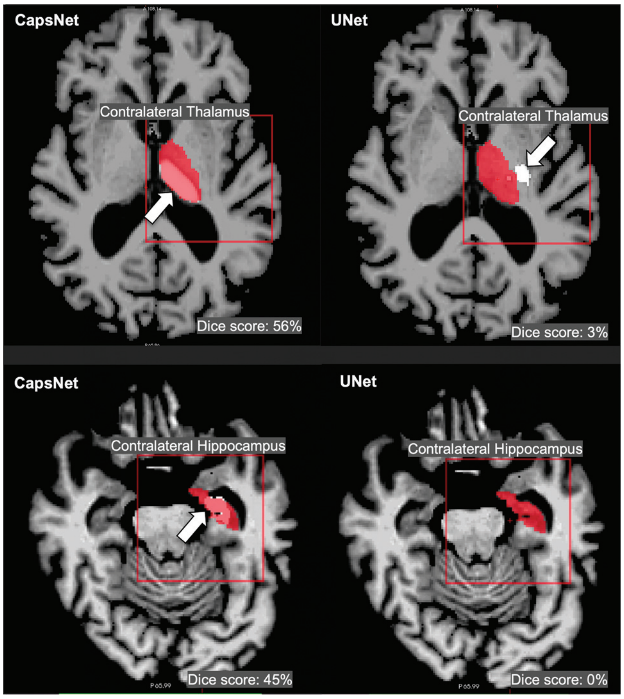

FIG 3. CapsNets outperforms UNets in segmenting images that were not represented in the training data. Both models were trained to segment right-brain structures and were tested to segment contralateral left-brain structures. Target segmentations and model predictions are, respectively, shown in red and white. Dice scores are provided for the entire volume of the segmented structure in this patient. The CapsNet partially segmented the contralateral thalamus and hippocampus (*white arrows*), but the UNet poorly segmented the thalamus (*white arrow*) and entirely missed the hippocampus.

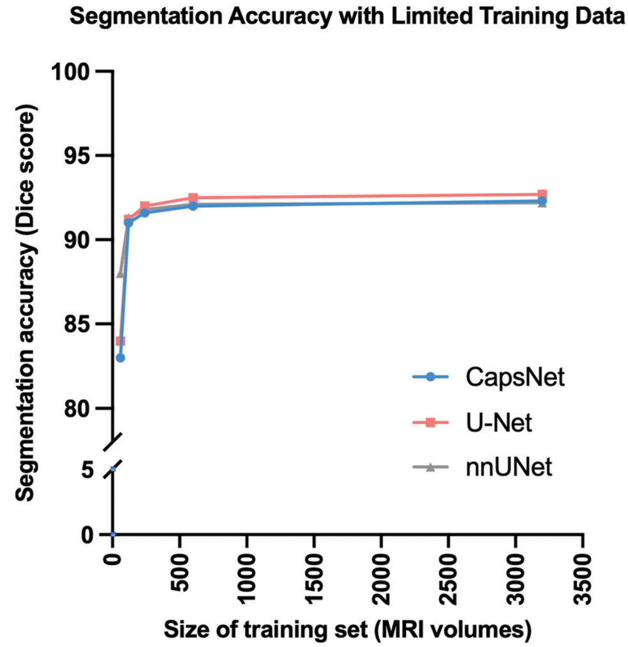

FIG 4. Comparing CapsNets, UNets, and nnUNets when training data are limited. When the size of the training set was decreased from 3199 to 120 brain MRIs, hippocampus segmentation accuracy (measured by Dice score) of all 3 models did not decrease >1%. Further decrease in the size of the training set down to 60 MRIs led to worsened segmentation accuracy.

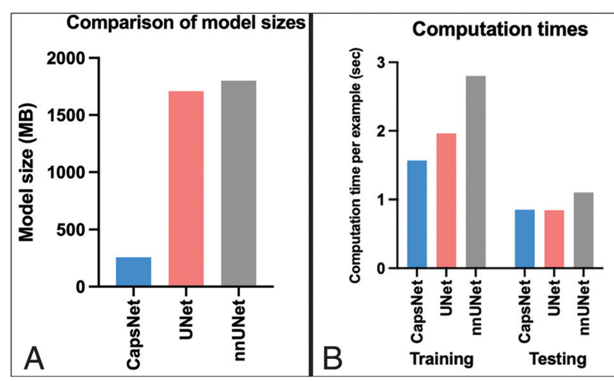

FIG 5. Comparing the computational efficiency among CapsNets, UNets, and nnUNets, in terms of memory requirements (*A*) and computational speed (*B*). *A*, The *bars* represent the computational memory required to accommodate the total size of each model, including the parameters plus the cumulative size of the forward- and backward-pass feature volumes. *B*, CapsNet trains faster, given that its trainable parameters are 1 order of magnitude fewer than UNets or nnUNets. The training times represent the time that each model took to converge for segmenting the hippocampus, divided by the number of training examples and the training epochs (to make training times comparable with test times). The test times represent how fast a fully-trained model can segment a brain image.

CapsNets to be effective for segmenting heart structures but ineffective for segmenting the hippocampus on brain images. Our more favorable results in segmenting the hippocampus are likely because of the 3D structure of our CapsNet, which can use the contextual information in the volume of the image rather than just a section of the image to better segment the complex shape of the hippocampus.[39]

Our study has several limitations. Our models were tested on only 3 brain structures that are commonly segmented on brain MRIs, meaning that our findings may not generalize across other imaging modalities and anatomic structures. Nevertheless, our findings show the efficacy of CapsNets on brain structures with different levels of segmentation difficulty, suggesting the potential utility for a variety of scenarios. Computational efficiency across models was measured using the same computing resources and GPU memory, and our findings may not translate to different computational settings. Future studies can further explore the relative computational efficiency of CapsNets compared with other autosegmentation models across different computing environments. We only compared the efficacy of CapsNets with UNet-based models. While there are multiple other autosegmentation models, UNet-based models are currently viewed as the most successful deep learning models for segmenting biomedical images. Further studies comparing the CapsNet with other deep learning models are an area of future research. Last, we found CapsNets to outperform UNet models when segmenting contralateral structures not represented in the training data. Techniques like data augmentation have shown the ability to improve the generalizability of UNet models in this scenario. Nevertheless, our findings demonstrate the ability of CapsNets to encode spatial information without the need for such techniques, which often require additional computational resources. This result further highlights the potential computational advantages of CapsNets for medical image segmentation.

CONCLUSIONS

In this study, we showed that 3D CapsNets can accurately segment neuroanatomic structures on brain MR images with segmentation accuracy similar to that of UNet-based models. We also showed that CapsNets outperformed UNet-based models in segmenting out-of-distribution data. CapsNets are also more computationally efficient compared with UNet-based models because they train faster and require less computation memory.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Feng CH, Cornell M, Moore KL, et al. **Automated contouring and planning pipeline for hippocampal-avoidant whole-brain radiotherapy.** *Radiat Oncol* 2020;15:251 CrossRef Medline
2. Dasenbrock HH, See AP, Smalley RJ, et al. **Frameless stereotactic navigation during insular glioma resection using fusion of three-dimensional rotational angiography and magnetic resonance imaging.** *World Neurosurg* 2019;126:322–30 CrossRef Medline
3. Dolati P, Gokoglu A, Eichberg D, et al. **Multimodal navigated skull base tumor resection using image-based vascular and cranial nerve segmentation: a prospective pilot study.** *Surg Neurol Int* 2015;6:172 CrossRef Medline
4. Thompson RF, Valdes G, Fuller CD, et al. **Artificial intelligence in radiation oncology: a specialty-wide disruptive transformation?** *Radiother Oncol* 2018;129:421–26 CrossRef Medline
5. Kotecha R, Aneja S. **Opportunities for integration of artificial intelligence into stereotactic radiosurgery practice.** *Neuro Oncol* 2021;23:1629–30 CrossRef Medline
6. Aneja S, Chang E, Omuro A. **Applications of artificial intelligence in neuro-oncology.** *Curr Opin Neurol* 2019;32:850–56 CrossRef Medline
7. Nalepa J, Marcinkiewicz M, Kawulok M. **Data augmentation for brain-tumor segmentation: a review.** *Front Comput Neurosci* 2019;13:83 CrossRef Medline
8. Despotović I, Goossens B, Philips W. **MRI segmentation of the human brain: challenges, methods, and applications**. *Comput Math Methods Med* 2015;2015:e450341 CrossRef Medline
9. Joel MZ, Umrao S, Chang E, et al. **Using adversarial images to assess the robustness of deep learning models trained on diagnostic images in oncology.** *JCO Clin Cancer Inform* 2022;e2100170 CrossRef Medline
10. Chang E, Joel MZ, Chang HY, et al. **Comparison of radiomic feature aggregation methods for patients with multiple tumors.** *Sci Rep* 2021;11:9758 CrossRef Medline
11. Rudie JD, Weiss DA, Colby JB, et al. **Three-dimensional U-Net convolutional neural network for detection and segmentation of intracranial metastases.** *Radiol Artif Intell* 2021;3:e200204 CrossRef Medline
12. Rauschecker AM, Gleason TJ, Nedelec P, et al. **Interinstitutional portability of a deep learning brain MRI lesion segmentation algorithm.** *Radiol Artif Intell* 2022;4:e200152 CrossRef Medline
13. Weiss DA, Saluja R, Xie L, et al. **Automated multiclass tissue segmentation of clinical brain MRIs with lesions.** *Neuroimage Clin* 2021;31:102769 CrossRef Medline
14. Rudie JD, Weiss DA, Saluja R, et al. **Multi-disease segmentation of gliomas and white matter hyperintensities in the BraTS data using a 3D Convolutional Neural Network.** *Front Comput Neurosci* 2019;13:84 CrossRef Medline
15. Punn NS, Agarwal S. **Modality specific U-Net variants for biomedical image segmentation: a survey.** *Artif Intell Rev* 2022; 55:5845 Medline
16. Isensee F, Jaeger PF, Kohl SA, et al. **nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation.** *Nat Methods* 2021;18:203–11 CrossRef Medline
17. Duong MT, Rudie JD, Wang J, et al. **Convolutional neural network for automated FLAIR lesion segmentation on clinical brain MR imaging.** *AJNR Am J Neuroradiol* 2019;40:1282–90 CrossRef Medline
18. Sabour S, Frosst N, Hinton GE. **Dynamic routing between capsules.** In: *Proceedings of the 31st International Conference on Neural Information Processing Systems. NIPS'17*, Red Hook, New York. December 4–9, 2017:3859–69
19. Hinton GE, Sabour S, Frosst N. **Matrix capsules with EM routing.** In: *International Conference on Learning Representations*, Vancouver, British Columbia, Canada. April 30 to May 3, 2018
20. LaLonde R, Xu Z, Irmakci I, et al. **Capsules for biomedical image segmentation.** *Med Image Anal* 2021;68:101889 CrossRef Medline
21. Crawford KL, Neu SC, Toga AW. **The Image and Data Archive at the Laboratory of Neuro Imaging.** *Neuroimage* 2016;124:1080–83 CrossRef Medline
22. Weiner M, Petersen R, Aisen P. **Alzheimer's Disease Neuroimaging Initiative**. September 16, 2014. https://clinicaltrials.gov/ct2/show/NCT00106899. Accessed March 21, 2022
23. Clerx L, Gronenschild EH, Echavarri C, et al. **Can FreeSurfer compete with manual volumetric measurements in Alzheimer's disease?** *Curr Alzheimer Res* 2015;12:358–67 CrossRef Medline
24. Ochs AL, Ross DE, Zannoni MD, et al. **Comparison of automated brain volume measures obtained with NeuroQuant and FreeSurfer.** *J Neuroimaging* 2015;25:721–27 CrossRef Medline.
25. Fischl B. **FreeSurfer.** *Neuroimage* 2012;62:774–81 CrossRef Medline
26. Fischl B, Salat DH, Busa E, et al. **Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain.** *Neuron* 2002;33:341–55 CrossRef Medline
27. Ganzetti M, Wenderoth N, Mantini D. **Quantitative evaluation of intensity inhomogeneity correction methods for structural MR brain images.** *Neuroinformatics* 2016;14:5–21 CrossRef Medline
28. Somasundaram K, Kalaiselvi T. **Automatic brain extraction methods for T1 magnetic resonance images using region labeling and morphological operations.** *Comput Biol Med* 2011;41:716–25 CrossRef Medline
29. Popescu V, Battaglini M, Hoogstrate WS, et al. **Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis.** *Neuroimage* 2012;61:1484–94 CrossRef Medline
30. Cardenas CE, Yang J, Anderson BM, et al. **Advances in auto-segmentation.** *Semin Radiat Oncol* 2019;29:185–97 CrossRef Medline
31. Elguindi S, Zelefsky MJ, Jiang J, et al. **Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy.** *Phys Imaging Radiat Oncol* 2019;12:80–86 CrossRef Medline
32. Francis S, Jayaraj PB, Pournami PN, et al. **ThoraxNet: a 3D U-Net based two-stage framework for OAR segmentation on thoracic CT images.** *Phys Eng Sci Med* 2022;45:189–203 CrossRef Medline
33. Yaqub M, Jinchao F, Zia MS, et al. **State-of-the-art CNN optimizer for brain tumor segmentation in magnetic resonance images.** *Brain Sci* 2020;10:427 CrossRef Medline
34. Guha Roy A, Conjeti S, Navab N, et al. **QuickNAT: a fully convolutional network for quick and accurate segmentation of neuroanatomy.** *Neuroimage* 2019;186:713–27 CrossRef Medline
35. Yahyatabar M, Jouvet P, Cheriet F. **Dense-Unet: a light model for lung fields segmentation in chest x-ray images.** *Annu Int Conf IEEE Eng Med Biol Soc* 2020;2020:1242–45 CrossRef Medline
36. Chi J, Zhang S, Han X, et al. **MID-UNet: multi-input directional UNet for COVID-19 lung infection segmentation from CT images.** *Signal Process Image Commun* 2022;108:116835 CrossRef Medline
37. Agnes SA, Anitha J. **Efficient multiscale fully convolutional UNet model for segmentation of 3D lung nodule from CT image.** *J Med Imaging Bellingham Wash* 2022;9:052402 CrossRef Medline
38. Survarachakan S, Johansen JS, Aarseth M, et al. **Capsule Nets for Complex Medical Image Segmentation Tasks.** In: *Proceedings of the 10th Colour and Visual Computing Symposium 2020 (CVCS 2020)*, Gjøvik, Norway; Virtual. September 16–17, 2020
39. Avesta A, Hossain S, Lin M, et al. **Comparing 3D, 2.5D, and 2D approaches to brain image auto-segmentation.** *Bioengineering (Basel)* 2023;10:181 CrossRef Medline