A Novel Domain Adaptation Framework for Medical Image Segmentation

Amir Gholami[1(✉)], Shashank Subramanian[2], Varun Shenoy[1], Naveen Himthani[2], Xiangyu Yue[1], Sicheng Zhao[1], Peter Jin[1], George Biros[2], and Kurt Keutzer[1]

[1] University of California Berkeley, Berkeley, USA
amirgh@berkeley.edu
[2] The University of Texas at Austin, Austin, USA

Abstract. We propose a segmentation framework that uses deep neural networks and introduce two innovations. First, we describe a biophysics-based domain adaptation method. Second, we propose an automatic method to segment white matter, gray matter, glial matter and cerebrospinal fluid, in addition to tumorous tissue. Regarding our first innovation, we use a domain adaptation framework that combines a novel multispecies biophysical tumor growth model with a generative adversarial model to create realistic looking synthetic multimodal MR images with known segmentation. These images are used for the purpose of training time data augmentation. Regarding our second innovation, we propose an automatic approach to enrich available segmentation data by computing the segmentation for healthy tissues. This segmentation, which is done using diffeomorphic image registration between the BraTS training data and a set of pre-labeled atlases, provides more information for training and reduces the class imbalance problem. Our overall approach is not specific to any particular neural network and can be used in conjunction with existing solutions. We demonstrate the performance improvement using a 2D U-Net for the BraTS'18 segmentation challenge. Our biophysics based domain adaptation achieves better results, as compared to the existing state-of-the-art GAN model used to create synthetic data for training.

Keywords: Segmentation · Neural network · Machine learning · Glioblastoma multiforme · Tumor growth models · Image registration

1 Introduction

Automatic segmentation methods have the potential to provide accurate and reproducible labels leading to improved tumor prognosis and treatment planning, especially for cases where access to expert radiologists is limited.

A. Gholami and S. Subramanian—Equally contributed.

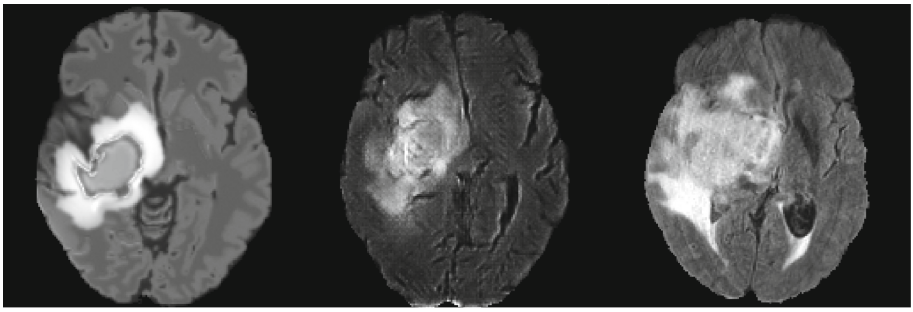
A. Crimi et al. (Eds.): BrainLes 2018, LNCS 11384, pp. 289–298, 2019.
https://doi.org/10.1007/978-3-030-11726-9_26

In the BraTS competition [4], we seek to segment multimodal MR images of glioma patients. Common brain MRI modalities include post-Gadolinium T1 (used to enhance contrast and visualization of the blood-brain barrier), T2 and FLAIR (to highlight different tissue fluid intensities), and T1. We use the data for these four modalities to generate the segmentations using a methodology that we outline below.

In most image classification tasks deep neural networks (DNNs) have been a very powerful technique that tends to outperform other approaches and BraTS is no different. From past BraTS competitions two main DNN architectures have emerged: DeepMedic [15] and U-Net [20]. How can we further improve this approach? Most research efforts have been on further improving these architectures, as well as coupling them with post-processing and ensembling techniques. In our work here, we propose a framework to work around the relatively small training datasets used in the BraTS competition. Indeed, in comparison to other popular classification challenges like ImageNet [5] (which consists of one million images for training), the BraTS training set contains only 285 instances (multimodal 3D MR images), a number that is several orders of magnitude smaller than the typical number of instances required for DNNs to work well. These observations have motivated this work, whose contributions we summarize below.

Related work: Recently, deep learning approaches using convolutional neural networks (CNNs) have demonstrated excellent performance in semantic segmentation tasks in medical imaging. Seminal works for segmentation stem from fully-convolutional networks (FCNs) [14]. U-Net [20] is another popular architecture for medical segmentation, which merges feature maps from the contracting path of an FCN to its expanding path to preserve local contextual information. Multi-scale information is often incorporated by using parallel convolutional pathways of various resolutions [12] or by using dilated convolutions and cascading network architectures [23]. Post-processing and ensemble methods are also usually used after training with these models. The most commonly used post processing step is Conditional Random Fields (CRF) [12], which has been found to significantly reduce false positives and sharpen the segmentation. Ensembling is also very important to reduce overfitting with deep neural networks. The winning algorithm of the Multimodal Brain Tumor Image Segmentation Benchmark (BraTS) challenge in 2017 was based on Ensembles of Multiple Models and Architectures (EMMA) [11], which bagged a heterogeneous collection of networks (including DeepMedic (winner of ISLES 2015 [15]), U-Nets and FCNs) to build a robust and generic segmentation model.

There are established techniques to address training with small datasets, such as regularization, or ensembling, which was the approach taken by the winning team of BraTS'17. However, in this paper we propose an orthogonal method to address this problem.



(a) (b) (c)

Fig. 1. Domain adaptation results: (a) represents a synthetic FLAIR brain image, (b) represents the domain adapted synthetic FLAIR image, (c) represents the real BraTS FLAIR image. As we can see from the intensity distributions, the values in the adapted images are qualitatively closer to the real images.

Contributions: Our main contributions are as follows:

1. **Data augmentation:** We propose a biophysics-based domain adaptation strategy to add synthetic tumor-bearing MR images to the training examples. There have been many notable works to simulate tumor growth (see [6–8,10,13,19]). We use an in-house PDE based multispecies tumor growth model [22] to simulate synthetic tumors. Since simulated data does not contain the correct intensity distribution of a real MR image, we train an auxiliary neural network to transform the simulated images to match real MRIs. This network gets a multimodal input and transforms this data to match the distribution of BraTS images by imposing certain cycle consistency constraints. As we will show, this is a very promising approach.
2. **Extended segmentation:** We extend the segmentation to the healthy parenchyma. This is done in two steps. First, we segment the training dataset using an atlas-based ensemble registration (using an in-house diffeomorphic registration code). Second, we train our DNN network to segment both tumor and healthy tissue (Four classes: cerebrospinal fluid, gray and white matter, glial matter). Our approach adds important information about healthy tissue delineation, which is actually used by radiologists. It also reduces the inherent class imbalance problem.

Our data augmentation strategy is different from the recent work of [21], which uses GANs to automatically generate data. To the best of our knowledge, our work here is the first to use a biophysics-based domain adaptation framework for automatic data generation, and our approach achieves five percentage points higher dice score as compared to [21], even though we use a 2D neural network architecture (which has a sub-optimal performance as compared to the 3D network used in [21]).

Limitations: Currently, our framework only supports 2D domain transformations. Hence, we are limited to transforming 3D brains slice-by-slice and using

only 2D neural network architectures. This is sub-optimal as 3D CNNs can demonstrably utilize volumetric medical imaging data more efficiently, leading to better and more robust performance (see [9,11,12]). Hence, extending our framework to 3D is the focus of our future work and can potentially lead to greater improvements in performance.

The outline of the paper is as follows. In Sect. 2, we discuss the methodology for domain adaptation (Sect. 2.1), and the whole brain segmentation (Sect. 2.2). In Sect. 2.2 we present preliminary results for the BraTS'18 challenge [1–3,15]. Our method achieves a Dice score of [79.15, 90.81, 81.91] for enhancing tumor, whole tumor and tumor core, respectively for the BraTS'18 validation dataset.

2 Methods

2.1 Domain Adaptation

As mentioned above, one of the main challenges in medical imaging is the scarcity of training data. To address this issue, we use a novel domain adaptation strategy and generate synthetic tumor-bearing MR images to enrich the training dataset. This is performed by first solving an in-house PDE based multispecies tumor model using an atlas brain [22]. This model captures the time evolution of enhancing and necrotic tumor concentrations along with tumor-induced brain edema. The governing equations for the model are reaction-diffusion-advection equations for the tumor species along with a diffusion equation for oxygen and other nutrients. We couple this model with linear elasticity equations with variable elasticity material properties to simulate the deformation of surrounding brain tissue due to tumor growth, also known as "mass effect". However, this data cannot be used directly due to the difference in intensity distributions between a BraTS MRI scan and a synthetic MRI scan. Directly using synthetic data during the training process will adversely guide the neural network to learn features which do not exist in a real MR image, resulting in poor performance.

To address this issue, we use CycleGAN [24] to perform domain adaptation from the generated synthetic data to the real BraTS images. This is done by learning a mapping $G : X \rightarrow Y$ such that the distribution of images from $G(X)$ is indistinguishable from the distribution Y using an adversarial loss. Here, X is the simulated tumor data, and Y is the corresponding data which matches the BraTS distribution. Because this mapping is highly under-constrained, it is coupled with an inverse mapping $F : Y \rightarrow X$ and a cycle consistency loss is introduced to enforce $F(G(X)) \approx X$ (and vice versa).

For training the domain adaptation network, we first computationally simulate synthetic tumors in a healthy brain atlas, located approximately at the whole tumor center taken from each BraTS image. Hence, every synthetic tumorous brain is paired with the corresponding data from a real BraTS image. Then, we perform a pre-processing step to transform our synthetic results to intensities. We produce a segmentation map for every tissue (healthy and tumorous) class and sample intensities for each class from a real MRI scan. We assign these sampled intensities to every voxel in our synthetic segmentation map to finally obtain

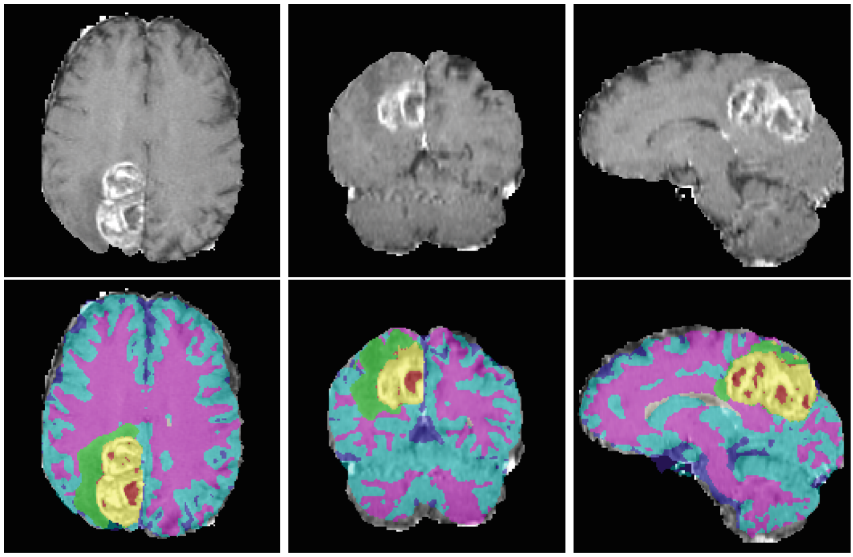


Fig. 2. (*Top row*): The original T1ce image for Brats18_TCIA02_135_1 training data is shown for different views (axial, coronal and sagittal). (*Bottom row*): The corresponding extended segmentation for healthy cells computed by solving a 3D registration problem with a segmented atlas. We overlay the BraTS tumor segmentation with the registered segmentation to get the final results.

our synthetic MRI scans. Then, we train with these pre-processed synthetic MRI scans and their corresponding BraTS images. Samples of our adaptation results are shown Fig. 1, which demonstrate an almost indistinguishable adaptation of the simulated data with the real images.

2.2 Whole Brain Segmentation with Healthy Tissues

An orthogonal approach that we propose for data augmentation, is an extended segmentation BraTS training data. That is, we segment the healthy parenchyma into gray/white matter, cerebrospinal fluid, and glial cells. The delineation of these healthy tissues contain important information which is actually used by radiologists. For example, the delineation of the tissues could be compressed due to tumor growth in the confined space of the brain. Providing this information to the classifier can help in better segmenting tumorous regions. However, such data is not readily available in the BraTS training dataset, since labelling the tumorous regions itself is laborious, let alone annotating full healthy tissues which is orders of magnitude more time consuming. We propose a novel automated approach to compute this information through image registration. In our method, we only need one (or preferably a few) fully segmented brains.

Then given an input 3D brain, we perform the following automatic steps to obtain the extended segmentation:

1. Affine registration of each atlas image to the brats image.
2. Diffeomorphic registration of each atlas image to the BraTS image: This step aims to find a deformation map that would "translate" a healthy atlas to match the structure of a given BraTS training example. We compute this deformation by solving a PDE-constrained optimization problem. We refer to [16–18] for details on solving this optimization problem.
3. Majority voting to fuse labels of all deformed atlases to get the final healthy tissue segmentation: The votes are weighted with the quality of diffeomorphic registration measured by the L_2 norm of the residual between each deformed atlas and brats image. This ensures the highest weight for the deformed atlas closest to the BraTS image.

We show an exemplary segmentation for an MRI scan from the BraTS training data in Fig. 2.

3 Setup

Here, we describe our setup and then report our segmentation results in the subsequent section on BraTS'18 dataset.

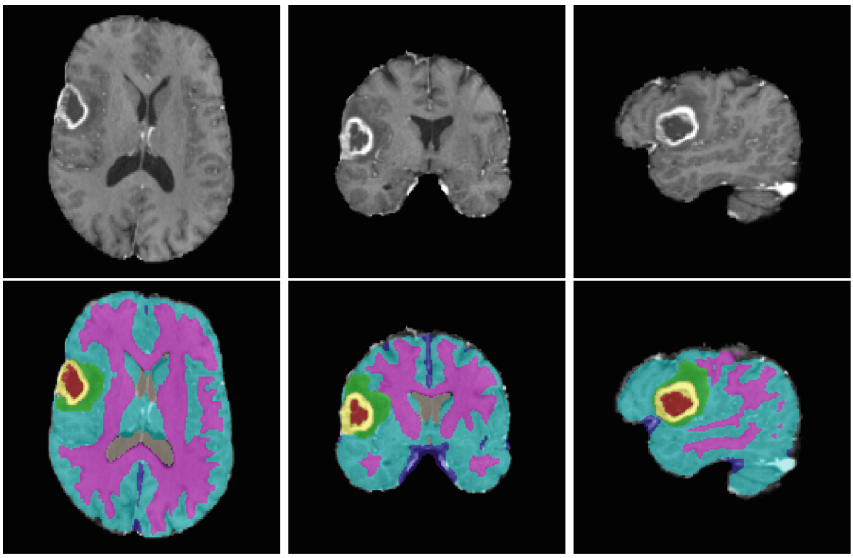


Fig. 3. (*Top row*): The original T1ce validation MR-image for Brats18_CBICA_ABT_1 data is shown for different views (axial, coronal and sagittal). (*Bottom row*): The corresponding segmentation result for healthy cells computed by the neural network. (Color figure online)

Baseline Network for Healthy and Tumor Segmentation. We first obtain the healthy tissue segmentation for all the BraTS training data using the image registration method (with 22 healthy atlases) discussed above, and use the fine grained data to train a neural network. Given that our current domain adaptation framework only supports 2D transformations, we follow a two stage segmentation routine by first localizing the tumor location(s), and then creating 2D slices/crops around the tumor and passing it to a 2D U-Net[1]. We use fixed sizes for our crops (specifically $48 \times 48, 96 \times 96$ and 144×144). This is to ensure no loss of information due to strided operations when we go deeper in the neural network. We train our network using a five-fold validation split of the training data with ADAM optimizer and ensemble the splits to obtain the baseline results. We show the healthy segmentation for a validation MRI scan in Fig. 3.

Data Augmentation Through Domain Adaptation. To augment our data with domain adaptation results, we simulate a synthetic tumor in our atlas corresponding to the whole tumor center of mass of every BraTS training image. We transfer the synthetic brain to the BraTS domain for every axial slice. Hence, our augmented dataset consists of approximately twice the amount of training brains.

Our final neural network is a 2D multi-view U-Net (with the tumor localization strategy described above) with data augmentation using domain adaptation. We train three U-Nets corresponding to the axial, sagittal or coronal view of the MRI scan and ensemble them (similar to the multi-view fusion method outlined in [23]). This is done in order to avoid noisy segmentations and reduce the class imbalance inherent in the BraTS dataset. As before, we train five-fold cross-validation splits and ensemble them to avoid overfitting to the training data.

Table 1. We report the BraTS'18 results for our method for both the baseline model and the final 2D network. Our final submission to the validation portal is highlighted. The last row shows the dice scores for BRaTS'18 testing dataset. Even though we use a sub-optimal 2D network, but we can still achieve significant improvement with the proposed framework.

	Dice Score		
	EN	WT	TC
Baseline (Validation)	73.86	89.49	79.94
Proposed (Validation)	79.15	90.81	81.91
Proposed (Testing)	70.96	87.11	76.87

[1] For the localization task, we use a simple 3D U-Net [9], with ten layers and multiclass dice loss.

4 Results

We trained the framework using the BraTS'18 data. The fine-grained segmentation result from the first stage 3D U-Net is shown in Fig. 3. As one can see, this involves both the tumor segmentation, shown in red/yellow/green, as well as

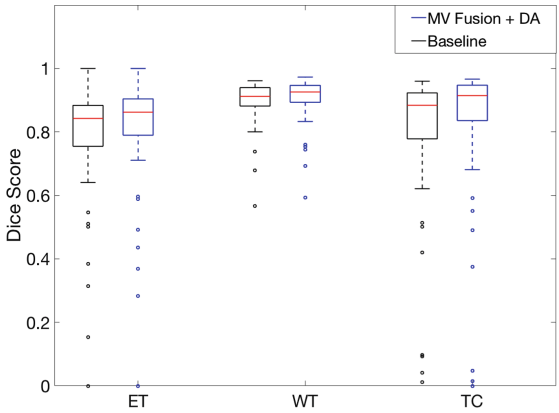


Fig. 4. Box plot for the final model's dice score on the BraTS'18 validation data is shown. This model achieves a mean dice score of (79.15, 90.81, 81.91) percent for (WT, TC, EN), respectively.

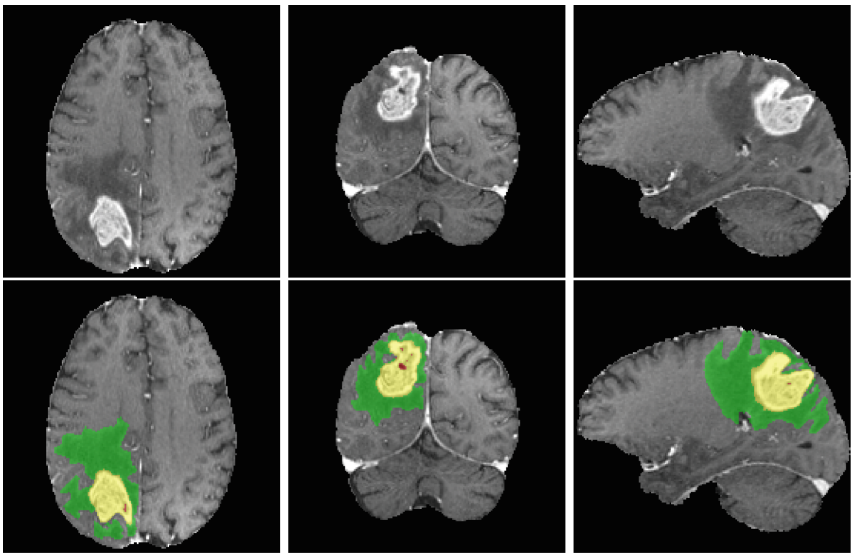


Fig. 5. (*Top row*): The original T1ce validation MR-image for Brats18_CBICA_AAM_1 data is shown for different views (axial, coronal and sagittal). (*Bottom row*): The corresponding tumor segmentation result from the final 2D network.

healthy structure of the brain shown in purple/cyan/gray/dark blue. This data is used for localizing the tumor boundaries. We then use this data and create multi-view slices around the tumor bearing region. Then, this data is passed through the second stage 2D U-Net which was trained along with the domain adaptation data, and fused together to obtain the final segmentation as shown in Fig. 5.

We show quantitative values for the Dice score in Table 1, with the corresponding box plots shown in Fig. 4. The baseline network has a dice score of [73.86, 89.49, 79.94] for Enhancing Tumor (ET), Whole Tumor (WT), and Tumor Core (TC). Using our proposed data augmentation framework leads to a dice score of [79.15, 90.81, 81.91]. These could be further improved by using a 3D network instead of a 2D one, by developing a 3D domain adaptation framework.

5 Conclusion

We presented a new framework for biophysics-based medical image segmentation. Our contributions include an automatic healthy tissue segmentation of the BraTS dataset, and a novel Generative Adversarial Network to enrich the training dataset using a model to generate synthetic phenomenological structures of a glioma. We demonstrated that our approach yields promising results on the BraTS'18 validation dataset. Our framework is not specific to a particular model, and could be used with other proposed neural networks for the BraTS challenge. Extending our domain adaptation framework to 3D can potentially lead to better performance and is the focus of our future work.

References

1. Bakas, S., et al.: Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. Nat. Sci. Data **4**, 170117 (2017)
2. Bakas, S., et al.: Segmentation Labels for the Pre-operative Scans of the TCGA-GBM Collection (2017). http://doi.org/10.7937/k9/tcia.2017.klxwjj1q
3. Bakas, S., et al.: Segmentation Labels for the Pre-operative Scans of the TCGA-LGG Collection (2017). http://doi.org/10.7937/k9/tcia.2017.gjq7r0ef
4. Bakas, S., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. arXiv preprint arXiv:1811.02629 (2018)
5. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: a large-scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, pp. 248–255. IEEE (2009)
6. Gholami, A.: Fast algorithms for biophysically-constrained inverse problems in medical imaging. Ph.D. thesis, The University of Texas at Austin (2017)
7. Hawkins-Daarud, A., Rockne, R.C., Anderson, A.R.A., Swanson, K.R.: Modeling tumor-associated edema in gliomas during anti-angiogenic therapy and its imapct on imageable tumor. Front. Oncol. **3**, 66 (2013)
8. Hawkins-Daarud, A., van der Zee, K.G., Tinsley Oden, J.: Numerical simulation of a thermodynamically consistent four-species tumor growth model. Int. J. Numer. Methods Biomed. Eng. **28**(1), 3–24 (2012)

9. Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., Maier-Hein, K.H.: Brain tumor segmentation and radiomics survival prediction: contribution to the BRATS 2017 challenge. CoRR abs/1802.10508 (2018). http://arxiv.org/abs/1802.10508
10. Ivkovic, S., et al.: Direct inhibition of myosin II effectively blocks glioma invasion in the presence of multiple motogens. Mol. Biol. Cell **23**(4), 533–542 (2012)
11. Kamnitsas, K., et al.: Ensembles of multiple models and architectures for robust brain tumour segmentation. CoRR abs/1711.01468 (2017). http://arxiv.org/abs/1711.01468
12. Kamnitsas, K., et al.: Efficient multi-scale 3d CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. **36**, 61–78 (2017). http://www.sciencedirect.com/science/article/pii/S1361841516301839
13. Lima, E., Oden, J., Hormuth, D., Yankeelov, T., Almeida, R.: Selection, calibration, and validation of models of tumor growth. Math. Models Methods Appl. Sci. **26**(12), 2341–2368 (2016)
14. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. CoRR abs/1411.4038 (2014). http://arxiv.org/abs/1411.4038
15. Maier, O., et al.: ISLES 2015 - a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. Med. Image Anal. **35**, 250–269 (2017)
16. Mang, A., Biros, G.: A semi-Lagrangian two-level preconditioned Newton-Krylov solver for constrained diffeomorphic image registration. SIAM J. Sci. Comput. **39**(6), B1064–B1101 (2017)
17. Mang, A., Gholami, A., Biros, G.: Distributed-memory large deformation diffeomorphic 3d image registration. In: SC16: International Conference for High Performance Computing, Networking, Storage and Analysis (2016)
18. Mang, A., Gholami, A., Davatzikos, C., Biros, G.: CLAIRE: a distributed-memory solver for constrained large deformation diffeomorphic image registration. arXiv preprint arXiv:1808.04487 (2018)
19. Oden, J.T., et al.: Toward predictive multiscale modeling of vascular tumor growth. Arch. Comput. Methods Eng. **23**(4), 735–779 (2016)
20. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. CoRR abs/1505.04597 (2015). http://arxiv.org/abs/1505.04597
21. Shin, H.-C., et al.: Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In: Gooya, A., Goksel, O., Oguz, I., Burgos, N. (eds.) SASHIMI 2018. LNCS, vol. 11037, pp. 1–11. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00536-8_1
22. Subramanian, S., Gholami, A., Biros, G.: Simulation of glioblastoma growth using a 3d multispecies tumor model with mass effect. arXiv preprint arXiv:1810.05370
23. Wang, G., Li, W., Ourselin, S., Vercauteren, T.: Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. CoRR abs/1709.00382 (2017). http://arxiv.org/abs/1709.00382
24. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV) (2017)