Cerberus: A Multi-headed Network for Brain Tumor Segmentation

Laura Daza[1](✉), Catalina Gómez[1], and Pablo Arbeláez[1]

Universidad de los Andes, Bogotá, Colombia {la.daza10, c.gomez10, pa.arbelaez}@uniandes.edu.co

Abstract. The automated analysis of medical images requires robust and accurate algorithms that address the inherent challenges of identifying heterogeneous anatomical and pathological structures, such as brain tumors, in large volumetric images. In this paper, we present Cerberus, a single lightweight convolutional neural network model for the segmentation of fine-grained brain tumor regions in multichannel MRIs. Cerberus has an encoder-decoder architecture that takes advantage of a shared encoding phase to learn common representations for these regions and, then, uses specialized decoders to produce detailed segmentations. Cerberus learns to combine the weights learned for each category to produce a final multi-label segmentation. We evaluate our approach on the official test set of the Brain Tumor Segmentation Challenge 2020, and we obtain dice scores of 0.807 for enhancing tumor, 0.867 for whole tumor and 0.826 for tumor core.

Keywords: Semantic segmentation · Brain tumor · MRI

1 Introduction

The use of Magnetic Resonance Imaging (MRI) for detection, treatment planning and monitoring of brain tumors has spurred interest in automatic segmentation of these structures. However, this task comprises many challenges, including the difficulty in annotating tumors with irregular shapes and appearances in large diagnostic images acquired through different protocols and scanners. Consequently, the availability of datasets for this task is highly restricted. These limitations call for automated algorithms that are robust to class imbalance and reduced training sets.

The recent success of Deep Neural Networks for segmentation tasks in natural images has promoted the development of specialized approaches for processing volumetric medical data. One notable example is U-Net [1], an encoder-decoder architecture that is used as foundation in most methods for biomedical segmentation. For instance, a recent variation of the 3D U-Net, dubbed No New-Net [12], achieved the second place on the segmentation task of the Brain Tumor Segmentation (BraTS) challenge [14, 5, 4, 2, 3]. This method uses the standard U-Net with an extensive training procedure that includes cross-validation of five

models aided with additional annotated data from the same task in the Medical Segmentation Decathlon challenge [15].

In the field of brain tumor segmentation, a common approach is the use of a cascade of networks that first segment the coarsest structures, and then use those results as input for the following networks [16, 13]. This approach specializes each network to a specific target, allowing the detection of objects with varying characteristics. However, cascaded networks necessarily require training more than one architecture, which significantly increases the number of parameters and, in most cases, hinders end-to-end training of the system. Deep Cascaded Attention Network (DCAN) [18] attempts to address these issues by sharing a low-level feature extractor followed by three independent encoder-decoder branches specialized for each brain region. The cascaded component is an attention mechanism that inputs the features of the branch for the coarsest categories to the fine-grained ones. Another recent method [9] replaces complex cascades with a multitask learning approach that employs two decoders at different scales for the coarse and fine categories. In contrast, we present a single model with a robust backbone to extract rich features that can be used to obtained specialized segmentations.

A second family of techniques focuses on reducing the computational cost inherent to processing 3D data. Among the best performing methods in the BraTS Challenge 2018 is DMFNet [7], a 3D encoder-decoder architecture that uses dilated multi-fiber units [8] to limit the number of parameters and FLOPS. In [6], Reversible Layers [10] are introduced to the No-New-Net architecture as an alternative to reduce memory consumption. Instead of storing all the activations in the forward pass, they are re-calculated during the backward pass using the next layer's activations. This strategy allows to process complete volumes rather than patches, but the performance gains are small and the training time is increased by 50% with respect to the non-reversible equivalent. Another method that does not require image patching is introduced in [11], in which the method alleviates memory consumption by swapping data from GPU to CPU memory inside the forward propagation.

In this paper, we propose Cerberus, a single lightweight network to address the task of brain tumor segmentation. Figure 1 shows an overview of our model. We hypothesize that learning representations that are common to all the categories, followed by specialized modules to recover brain tumor regions, allows our model to exploit both the shared coarse characteristics of neurological pathologies, and the unique features that are inherent to each structure type. Cerberus leverages a shared encoder to learn a common representation space with sufficient expressive capabilities for identifying heterogeneous brain tumors. Our method also learns how to combine the parameters learned for the specialized tasks to solve the more challenging multi-label segmentation task. With less than 4M of parameters, our model trains faster and can be trained on a single GPU. We demonstrate the competitive performance of Cerberus in the official validation sets of the BraTS Challenge 2018 and 2020. We will make our code publicly available in order to ensure reproducibility and encourage further research.

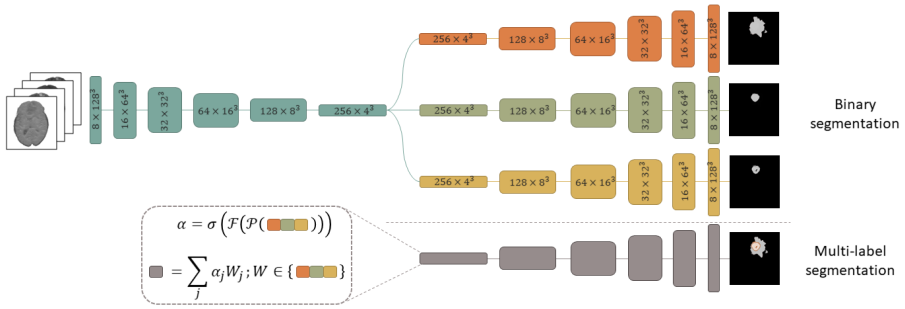

Fig. 1. Overview figure of Cerberus. Our approach uses a shared encoder that learns a common representations which can be then exploited by specialized decoders. Then, a combination module is applied to combine the kernels of each decoder to produce a multi-label segmentation. Skip connections between the encoder and the decoders are omitted for simplicity.

2 Method

We introduce Cerberus, a unified encoder-decoder architecture designed to simultaneously solve multiple binary semantic segmentation tasks and a multi-label task. Our method uses a single encoder to extract rich features at different resolutions of the input image, and then inputs the resulting information to three different decoders that specialize in specific structures. This strategy allows the decomposition of a multi-label segmentation problem into various simpler binary tasks. Finally, Cerberus learns to combine the parameters learned for the sub-tasks to solve a multi-label segmentation problem.

2.1 Cerberus architecture

Our model is composed of two main stages: the first one encodes general features obtained from the multimodal volumes, and the second one specializes in solving three binary tasks and a multi-label task. We conduct exhaustive experiments to optimize every stage of our architecture and the training curriculum, and finally define the best configuration empirically.

Encoder For the first stage we use an encoder with five blocks that contain Depthwise Separable Convolutions (DSC) and a residual connection. Every block duplicates the amount of feature maps and reduces to half the spatial resolution of the input. We define the output feature maps at block i as

$$e_i = \mathcal{E}_i(e_{i-1}) + e_{i-1}, i \in \{1, ..., 5\} \tag{1}$$

where e_0 is the input image and $\mathcal{E}_i$ denotes a sequence of two DSC. Throughout the entire architecture, every convolutional operation is followed by a normalization layer and a ReLU activation. We use Group Normalization [17] instead of

the standard Batch Normalization because the former is less affected by batches of small size. Additionally, similar to U-Net, we use skip connections to connect the features maps from each stage of the encoder with its corresponding block in the decoder.

Decoder In the second stage, each block contains a single $3\times3\times3$ convolutional layer to combine the result of the next deepest decoder block with the input from the corresponding skip connection. We include an upsample module composed of a $1\times1\times1$ convolutional layer followed by a trilinear upsampling operation to double the spatial resolution and halve the number of feature maps of the input from the decoder. The output of block i within the jth branch is calculated as

$$d_{j,i} = \mathcal{D}_{j,i}(U_{j,i}(d_{j,i+1}) \parallel e_i), j = \{1, 2, 3\} \tag{2}$$

where $\mathcal{D}$ denotes the $3 \times 3 \times 3$ convolution, U represents the upsample module and $\parallel$ is a concatenation operation. These decoders are specialized in segmenting the different regions of the tumor.

To obtain a final multi-label segmentation we include a fourth path that combines the kernels learned in the other decoders as follows:

$$\begin{aligned} W_{ml} &= \sum_j \alpha_j W_j \\ \alpha &= \sigma(\mathcal{F}(\mathcal{P}(W_1 \parallel W_2 \parallel W_3))) \end{aligned} \tag{3}$$

where α denotes the weight given to each kernel W. To calculate α we introduce an attention module that concatenates the kernels, performs an average pooling $\mathcal{P}$ along the spatial dimensions, and assign the weights using a fully connected layer $\mathcal{F}$ and a softmax activation σ. Since we use the kernels learned from the binary branches, the only additional parameters incurred in the multi-label path come from the attention modules.

2.2 Loss function

Cerberus produces three binary segmentation outputs (WT, TC, ET) and a final multi-label (ML) output. We define a combination of the Dice loss and the Cross-entropy loss calculated for each generated mask, and perform a weighted sum of all four to optimize the network, as shown in Equation 4.

$$Loss = \sum_\ell \beta_\ell \left(L_\ell^{Dice} + L_\ell^{CE}\right); \ell \in \{WT, TC, ET, ML\} \tag{4}$$

We empirically found that the weights that maximize our results are $\beta = [0.1, 0.3, 0.2, 0.4]$.

2.3 Pre-processing

Normalization of MRI intensity values is crucial for processing different modalities in Neural Networks. Hence, we follow a standard normalization per patient: for each modality, we subtract the mean and divide by the standard deviation of the brain region, while setting the intensities outside the brain to zero.

2.4 Implementation details

We train our models with Adam optimizer with an initial learning rate of $1e-3$ and include an L_2 regularization coefficient of $1e-5$. We reduce the initial learning rate by a factor of 0.1 whenever the validation loss has not decreased for 30 epochs. We adopt the on-the-fly data augmentation strategy proposed by [12]. Our transformations include rotations, scaling, mirroring and gamma correction. To address data imbalance, we define a patch-based sampling strategy such that the center voxel in a patch has equal probability of belonging to any category. Given the memory limitations of processing 3D images, we use patches of size $128 \times 128 \times 128$ and set the batch size to 6 to maintain the memory consumption under 12GB

2.5 Inference

During inference, the multi-label output corresponds to the segmentation of the tumor and the three binary outputs are used to obtain the uncertainty of the predictions for each region evaluated. For the uncertainty, we apply a sigmoid function to the predictions, calculate the complement of the probabilities and re-scale the values between 0 and 100. The result of this process is a pixel-wise map with values close to zero where the network predicted a category with high confidence. Also, we use patches extracted from the images at uniform intervals, insuring that all pixels are covered. To reconstruct the final image, we assign higher weights to the central voxel of each patch and combine all the predictions. In the ablation studies no further processing is made. For the final results in the official validation server, we train the models in a 5-fold cross-validation fashion and perform an additional test time augmentation (TTA) step that consists of flipping the patches along all axes and averaging the predictions. Finally, we define a simple final post-processing step consisting on the elimination of any component smaller than a threshold by assigning them to the nearest label.

3 Experiments and Results

3.1 Database

We develop our model on the BraTS 2020 dataset, which comprises MRI scans of high grade glioblastomas and low grade gliomas. The annotation labels manually provided by one to four raters are edema, necrosis and non-enhancing tumor (NCR/NET), and enhancing tumor (ET). The challenge evaluates the

following overlapping regions: whole tumor (WT), which includes all the three labels; tumor core (TC) that comprises the ET and NCR/NET; and enhancing tumor (ET). The training and validation sets contain 369 and 125 patients respectively, each with four MRI modalities available: T1 weighted, post-contrast T1 weighted, T2 weighted and FLAIR. The challenge provides an official server for evaluation. Besides, we compare our performance to recent published papers with results on the official validation set of BraTS 2018.

To conduct the ablation experiments, we split the training dataset into training and validation subsets with the 80% and 20% of the patients, respectively. We choose the model's weights that achieved the best Dice score on our validation subset to obtain results on the official validation sets.

3.2 Evaluation

The BraTS challenge evaluates the performance of segmentation algorithms using the Dice score (DS), Hausdorff distance (H95%), sensitivity (recall) and specificity. We report the average Dice score and Hausdorff distance over the patients in the corresponding evaluation set. For the uncertainty evaluation, the Dice score is calculated at different confidence measurements and the area under the Dice vs. uncertainty threshold curve is reported as the Dice AUC. An additional integrated score considers the Dice AUC and the area under the curve of the filtered true negatives and true positives for different thresholds.

3.3 Ablation experiments

We describe in detail the empirical choices within our model. We analyze the advantages of our approach by testing three types of architectures: multi-category networks (shared encoder and decoder), independent binary networks (separated encoder and decoder), and our proposed method with a shared encoder and separate decoders. We report the results in our validation subset from the BraTS 2020 patients in Table 1.

Table 1. Ablation experiments on our validation set. Parenthesis indicate the labels used for training. The models with * was trained optimizing the annotated labels (edema, NCR/NET and ET).

	Dice			Hausdorff95		
Model	ET	WT	TC	ET	WT	TC
Single network*	0.770	0.889	0.823	9.24	13.82	13.21
Separate networks*	0.773	0.884	0.796	12.85	8.20	15.89
Separate networks	0.771	0.896	0.811	6.64	12.47	11.02
Cascaded networks	0.782	0.896	0.808	13.35	12.47	10.73
Separate decoders	0.783	0.895	0.836	6.42	15.07	14.08
Communication modules	0.773	0.888	0.844	7.44	16.34	10.07
Cerberus	**0.794**	**0.897**	**0.845**	**3.59**	**6.19**	**6.47**

In the first setting we optimize the annotations provided by the challenge (edema, NCR/NET and ET) instead of the regions given that our loss function is designed for non-overlapping labels. Table 1 shows that this model obtains lower performance for ET, the fine grained region. This phenomenon is probably because the tasks are highly unbalanced, and a single decoder is not capable of learning a proper representation for the smaller categories.

In the second setting we train independent networks optimizing the annotations and the regions. The results demonstrate that directly optimizing the regions results in better predictions, specially for the WT and TC regions. In addition, we train cascaded networks by using the output from the coarsest regions as input to the following networks. In this case the results for ET, the smallest region. This happens because the coarsest regions guide the segmentation of smaller structures by limiting the search space within the image, but comes at the cost of training as many models as categories.

In the last setting, we train three models with shared encoders and different specialized decoders. The first approach has completely independent decoders that specialize to their corresponding task using only information from the encoder. In this case there is an improvement with respect to using separate networks, which proves that learning a unique representation space results in richer features. In the second approach we share information during the decoding stage as well by adding communication modules between the separate decoders. These modules take information from the three paths and combine it as presented in [8]. Finally, our Cerberus outperforms all methods in TC and ET and has results comparable to the cascaded networks in WT.

3.4 Comparison with the State-of-the-Art

In Table 2, we compare our best model with similar methods that have results in the official BraTS validation set and a published paper. Since all the methods are evaluated in the 2018 challenge, we retrain our method using this dataset. The metrics of the other methods are retrieved from the original papers.

Table 2. Comparison of Cerberus performance against competitive methods on the 2018 validation set.

				Dice score			Hausdorff95		
Model	Patch	Params (M)	FLOPS (G)	ET	WT	TC	ET	WT	TC
No New-Net [12]	128	10.36	202.25	0.796	0.908	0.843	3.12	4.79	8.16
Rev. U-Net [6]	Full	12.37	–	0.803	0.910	0.862	2.58	4.58	6.84
DMFNet [7]	128	3.88	27.04	0.801	0.906	0.845	3.06	4.66	6.44
DCAN [18]	128	–	–	0.817	0.912	0.862	3.57	4.26	6.13
Cerberus (ours)	128	4.02	49.82	0.797	0.895	0.835	4.22	7.77	10.3

Table 2 shows the competitive performance of our method in comparison to the state-of-the-art. Cerberus obtains similar results to four top performing

methods, even using a single network for the evaluation, against the five fold cross-validation used by most of the other methods. If we compare Cerberus' performance to the No New-Net results, the Dice scores for the three categories are remarkably similar.

Cerberus presents minor differences in performance with respect to Reversible U-Net, a single model trained on the 80% of the training data without model ensembles. The major difference in performance is on the TC category, 3.1%, and minor for ET (0.7%) and WT (1.6%). However, note that Reversible U-Net has three times more parameters than our method.

We achieve comparable Dice scores in the three categories with respect to DMFNet, presenting differences of 0.4%, 1.2% and 1.2% for ET, WT and TC, respectively. Finally, compared to DCAN, the major performance gap is for TC category (3.1%), and our model does not require training with cross-validations and ensembles across different data partitions to achieve competitive scores.

Results on BraTS 2020 Validation and Test sets We further address the competitiveness of Cerberus by evaluating its performance on the BraTS 2020 validation and test sets. Tables 3 and 4 show the evaluation metrics in both sets for the Segmentation and Uncertainty Tasks, respectively.

Table 3. Cerberus performance in the Segmentation task on the official Validation and Test sets of the BraTS 2020 Challenge.

	Dice			Hausdorff95		
BraTS 2020 set	ET	WT	TC	ET	WT	TC
Validation	0.748	0.898	0.828	31.82	5.50	9.68
Test	0.807	0.867	0.826	12.34	6.14	21.20

Table 4. Cerberus performance in the Uncertainty task on the official Validation and Test sets of the BraTS 2020 Challenge.

	Dice AUC			Score		
BraTS 2020 set	ET	WT	TC	ET	WT	TC
Validation	0.912	0.826	0.762	-	-	-
Test	0.883	0.841	0.819	0.947	0.935	0.929

3.5 Qualitative results

In Figure 2, we show qualitative examples of the predictions using Cerberus for patients in our validation subset to allow a visual comparison with the annotations. These examples show the accurate localization of brain tumors, specially

the largest region (edema in green) and its internal structures with our single model.

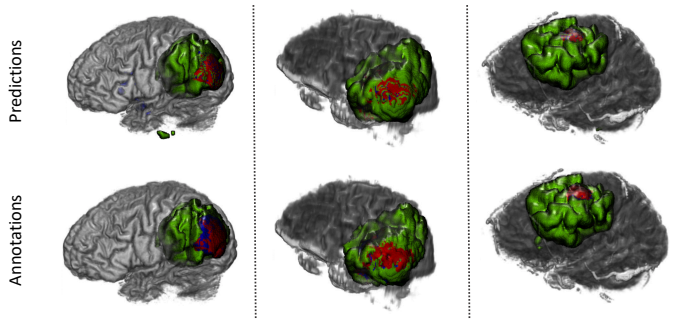

Fig. 2. Qualitative results of Cerberus' predictions over T1 in our validation subset. Top: predictions made by Cerberus; Bottom: annotations made by experts. Each column corresponds to a different patient. Edema is shown in green, ET in red and NCR/NET in blue.

4 Conclusions

We present Cerberus, a method for brain tumor segmentation that uses a single encoder to learn a shared representation space, independent decoders to solve multiple binary tasks, and learns to combine the decoders parameters to solve a multi-label segmentation task. Cerberus achieves competitive performance against the state-of-the-art in the BraTS Challenge 2018, proving the advantages of sharing a model for the feature extraction stage, and training separate reconstruction networks to obtain specialized segmentations. We also demonstrate the superiority and advantages of our approach by comparing it with similar architectures that do not share the encoder or use a single decoder.

References

1. U-net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., III, W.M.W., Frangi, A.F. (eds.) Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 - 18th International Conference Munich, Germany, October 5 - 9, 2015, Proceedings, Part III. Lecture Notes in Computer Science, vol. 9351, pp. 234–241. Springer (2015). https://doi.org/10.1007/978-3-319-24574-4_28, https://doi.org/10.1007/978-3-319-24574-4_28

2. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., Davatzikos, C.: Segmentation labels and radiomic features for the pre-operative scans of the tcga-lgg collection. The cancer imaging archive **286** (2017)
3. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., Davatzikos, C.: Segmentation labels and radiomic features for the pre-operative scans of the tcga-gbm collection. the cancer imaging archive. Nat Sci Data **4**, 170117 (2017)
4. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J.S., Freymann, J.B., Farahani, K., Davatzikos, C.: Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. Scientific data **4**, 170117 (2017)
5. Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R.T., Berger, C., Ha, S.M., Rozycki, M., Prastawa, M., Alberts, E., Lipková, J., Freymann, J.B., Kirby, J.S., Bilello, M., Fathallah-Shaykh, H.M., Wiest, R., Kirschke, J., Wiestler, B., Colen, R.R., Kotrotsou, A., LaMontagne, P., Marcus, D.S., Milchenko, M., Nazeri, A., Weber, M., Mahajan, A., Baid, U., Kwon, D., Agarwal, M., Alam, M., Albiol, A., Albiol, A., Varghese, A., Tuan, T.A., Arbel, T., Avery, A., B., P., Banerjee, S., Batchelder, T., Batmanghelich, K.N., Battistella, E., Bendszus, M., Benson, E., Bernal, J., Biros, G., Cabezas, M., Chandra, S., Chang, Y., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. CoRR **abs/1811.02629** (2018), http://arxiv.org/abs/1811.02629
6. Brügger, R., Baumgartner, C.F., Konukoglu, E.: A partially reversible u-net for memory-efficient volumetric image segmentation. CoRR **abs/1906.06148** (2019), http://arxiv.org/abs/1906.06148
7. Chen, C., Liu, X., Ding, M., Zheng, J., Li, J.: 3d dilated multi-fiber network for real-time brain tumor segmentation in MRI. In: Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P., Khan, A. (eds.) Medical Image Computing and Computer Assisted Intervention - MICCAI 2019 - 22nd International Conference, Shenzhen, China, October 13-17, 2019, Proceedings, Part III. Lecture Notes in Computer Science, vol. 11766, pp. 184–192. Springer (2019). https://doi.org/10.1007/978-3-030-32248-9_21, https://doi.org/10.1007/978-3-030-32248-9_21
8. Chen, Y., Kalantidis, Y., Li, J., Yan, S., Feng, J.: Multi-fiber networks for video recognition. CoRR **abs/1807.11195** (2018), http://arxiv.org/abs/1807.11195
9. Cheng, G., Cheng, J., Luo, M., He, L., Tian, Y., Wang, R.: Effective and efficient multitask learning for brain tumor segmentation. Journal of Real-Time Image Processing pp. 1–10 (2020)
10. Gomez, A.N., Ren, M., Urtasun, R., Grosse, R.B.: The reversible residual network: Backpropagation without storing activations. CoRR **abs/1707.04585** (2017), http://arxiv.org/abs/1707.04585
11. Imai, H., Matzek, S., Le, T.D., Negishi, Y., Kawachiya, K.: High resolution medical image segmentation using data-swapping method. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 238–246. Springer (2019)
12. Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., Maier-Hein, K.H.: No new-net. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Pa-

pers, Part II. pp. 234–244 (2018). https://doi.org/10.1007/978-3-030-11726-9_21, https://doi.org/10.1007/978-3-030-11726-9_21
13. Li, X., Luo, G., Wang, K.: Multi-step cascaded networks for brain tumor segmentation. CoRR **abs/1908.05887** (2019), http://arxiv.org/abs/1908.05887
14. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al.: The multimodal brain tumor image segmentation benchmark (brats). IEEE transactions on medical imaging **34**(10), 1993–2024 (2014)
15. Simpson, A.L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., van Ginneken, B., Kopp-Schneider, A., Landman, B.A., Litjens, G.J.S., Menze, B.H., Ronneberger, O., Summers, R.M., Bilic, P., Christ, P.F., Do, R.K.G., Gollub, M., Golia-Pernicka, J., Heckers, S., Jarnagin, W.R., McHugo, M., Napel, S., Vorontsov, E., Maier-Hein, L., Cardoso, M.J.: A large annotated medical image dataset for the development and evaluation of segmentation algorithms. CoRR **abs/1902.09063** (2019), http://arxiv.org/abs/1902.09063
16. Wang, G., Li, W., Ourselin, S., Vercauteren, T.: Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In: Crimi, A., Bakas, S., Kuijf, H.J., Menze, B.H., Reyes, M. (eds.) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers. Lecture Notes in Computer Science, vol. 10670, pp. 178–190. Springer (2017). https://doi.org/10.1007/978-3-319-75238-9_16, https://doi.org/10.1007/978-3-319-75238-9_16
17. Wu, Y., He, K.: Group normalization. CoRR **abs/1803.08494** (2018), http://arxiv.org/abs/1803.08494
18. Xu, H., Xie, H., Liu, Y., Cheng, C., Niu, C., Zhang, Y.: Deep cascaded attention network for multi-task brain tumor segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 420–428. Springer (2019)