

Auto3DSeg for Brain Tumor Segmentation from 3D MRI in BraTS 2023 Challenge.

Andriy Myronenko, Dong Yang, Yufan He, and Daguang Xu

NVIDIA, Santa Clara, CA

{amyronenko,dongy,yufanh,daguangx}@nvidia.com

Abstract. In this work, we describe our solution to the BraTS 2023 cluster of challenges using Auto3DSeg¹ from MONAI². We participated in all 5 segmentation challenges, and achieved the 1st place results in three of them: Brain Metastasis, Brain Meningioma, BraTS-Africa challenges, and the 2nd place results in the remaining two: Adult and Pediatric Glioma challenges.

Keywords: Auto3DSeg · MONAI · Segmentation · NVAUTO team.

1 Introduction

Multimodal Brain tumor segmentation challenge (BraTS) has established itself as one of the largest MICCAI challenges, where researchers can develop and evaluate their solutions to 3D brain MRI tumor segmentation [3,7,10,13]. BraTS 2023 follows the previous years setup with more data variability and notable new sub-challenges to segment various tumor types in various cohorts. BraTS 2023 segmentation challenges includes the following sub-challenges: Adult Glioma segmentation (same as BraTS 2021), Meningioma segmentation [12], Sub-Sahara-Africa Glioma segmentation (BraTS-Africa) [2], Pediatric tumor segmentation [11], and Metastasis segmentation [14] from brain 3D MRI.

Magnetic Resonance Imaging (MRI) is a key diagnostic tool for brain tumor analysis, monitoring and surgery planning. Usually, several complimentary 3D MRI modalities are acquired - such as T1, T1 with contrast agent (T1c), T2 and Fluid Attenuation Inversion Recover (FLAIR) - to emphasize different tissue properties and areas of tumor spread. For example the contrast agent, usually gadolinium, emphasizes hyperactive tumor subregions in T1c MRI modality.

BraTS aims to evaluate state-of-the-art methods for the segmentation of brain tumors by providing a 3D MRI dataset with ground truth tumor segmentation labels annotated by physicians [7,13,6,4,5]. BraTS training cases include four 3D MRI modalities (T1, T1c, T2 and FLAIR) rigidly aligned, resampled to 1x1x1 mm isotropic resolution and skull-stripped. The input image size is 240x240x155. The data were collected from multiple institutions, using various

¹ <https://monai.io/apps/auto3dseg>

² <https://github.com/Project-MONAI/MONAI>

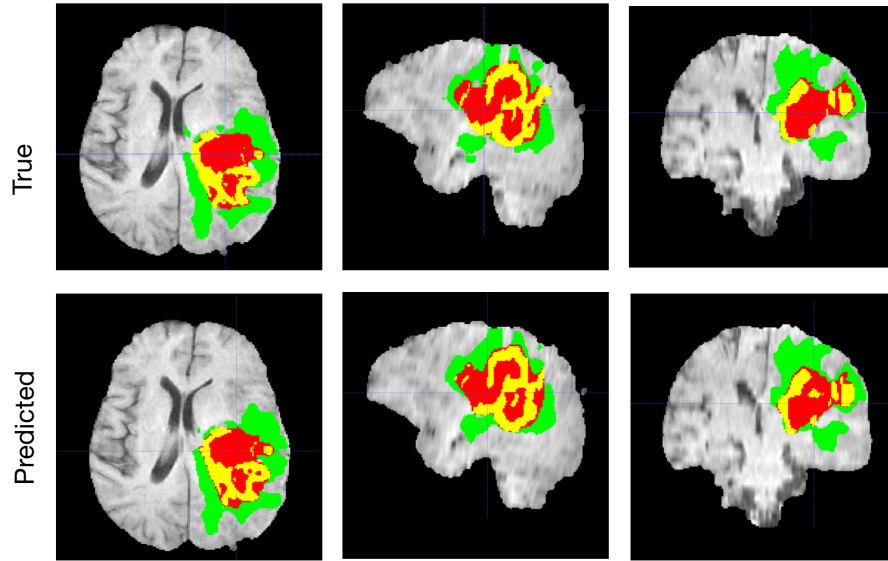


Fig. 1. A typical segmentation example with true and predicted labels overlaid over T1c MRI axial, sagittal and coronal slices. The whole tumor (WT) class includes all visible labels (a union of green, yellow and red labels), the tumor core (TC) class is a union of red and yellow, and the enhancing tumor core (ET) class is shown in yellow (a hyperactive tumor part).

MRI scanners. Annotations include 3 tumor subregions: the enhancing tumor, the peritumoral edema, and the necrotic and non-enhancing tumor core. The annotations were combined into 3 nested subregions: whole tumor (WT), tumor core (TC) and enhancing tumor (ET), as shown in Figure 1.

2 Methods

We describe a solution that is automated and is easy to run even for non-expert users. We use Auto3DSeg utilizes open source components in MONAI [1], offering both beginner and advanced researchers the means to effectively develop and deploy high-performing segmentation algorithms. The minimal user input to run Auto3DSeg for BraTS23 datasets, is

```
1 #!/bin/bash
2 python -m monai.apps.auto3dseg AutoRunner run \
3     --input="./input.yaml" --algos=segresnet
```

where a user provided input config (input.yaml) includes several lines:

```
1 # This is the YAML file "input.yaml"
2 modality: MRI
3 datalist: "./dataset.json"
```

```

4 dataroot: "/data/brats23"
5
6 class_names:
7 - { "name": "wt", "index": [1,2,3] }
8 - { "name": "tc", "index": [1,3] }
9 - { "name": "et", "index": [3] }
10 sigmoid : true

```

When running this command, Auto3DSeg will analyse the dataset, generate hyperparameter configs for several supported algorithms, trains them, and produce inference and ensembling. The system will automatically scale to all available GPUs.

The 3 minimum user options (in input.yaml) are data modality (MRI in this case), location of the downloaded Brats23 dataset (dataroot), and the list of input filenames with an associated fold number (dataset.json). We generate the 5-fold split assignments randomly for each of the segmentation sub-challenges (since each of the BraTS 2023 sub-challenges share the same data structure).

Since BraTS defines its specific label mapping (from integer class labels to 3 tumor subregions) we have to define it in the config, and since these subregions are overlapping, we use "sigmoid: true" to indicate multi-label segmentation, where the final activation is sigmoid (instead of the default softmax).

Currently the supported 3D segmentation algorithms are SegResNet [15], DiNTS [9] and SwinUNETR [8,16] with their unique training recipes. Here we used only the SegResNet algorithms, and trained it using 5-fold cross validation.

The simplicity of Auto3DSeg is a very minimal user input, which allows even non-expert users to achieve a great baseline performance. The system will take care of most of the work to analyse, configure and optimally utilize the available GPU resources. And for expert users, there are many configurations options that can manually provided to override the automatic values, for better performance tuning.

2.1 SegResNet

SegResNet³ is an encode-decoder based semantic segmentation network based on [15] It is a U-net based convolutional neural network (see Figure 2).

We use spatial augmentation including random affine and flip in all axes, random intensity scale, shift, noise and blurring. We use the combined Dice + Focal loss.

The same loss is summed over all deep-supervision sublevels:

$$Loss = \sum_{i=0}^4 \frac{1}{2^i} Loss(pred, target^{\downarrow}) \quad (1)$$

where the weight $\frac{1}{2^i}$ is smaller for each sublevel (smaller image size) i . The target labels are downsized (if necessary) to match the corresponding output size using nearest neighbor interpolation

³ <https://docs.monai.io/en/stable/networks.html>

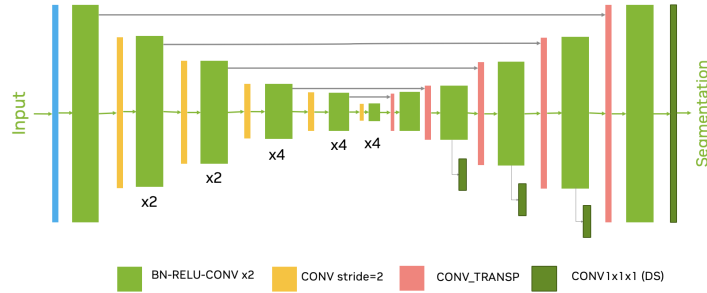


Fig. 2. SegResNet network configuration. The network uses repeated ResNet blocks with batch normalization and deep supervision

We use the AdamW optimizer with an initial learning rate of $2e^{-4}$ and decrease it to zero at the end of the final epoch using the Cosine annealing scheduler. We use batch size of 1 per GPU, We use weight decay regularization of $1e^{-5}$.

2.2 Pre-processing

We normalize input images to zero mean, unit standard deviation.

2.3 Training

We train the method on 8 GPUs 16GB NVIDIA V100 machine for 600 epochs. We trained the models from scratch, with the exception of BraTS-Africa dataset, which is very small. There, we initialized models from the checkpoints trained on the Glioma segmentation subtask (which was allowed by the competition).

Generally for all 5 segmentation challenges the training process was the same. One exception was the Brain Metastasis challenge datasets, where the data from two sources (UCSF and Stanford) had a missing T2 MRI modality, with only one ground truth sub-region annotated - ET (TC and WT were missing). To be able to use such data, we added a channel Dropout layer (prob=0.5) on T2 input channel. The dropout simulates channel to be all zeros, and help to make network robust to the cases when this channel is missing. For the missing ground truth labels, we computed the loss only over the available ground truth classes (e.g. only for ET). Since we use a multi-label formulation (with the final sigmoid activation), we can easily compute loss only over the available classes (in-contrast to softmax activation). This was something unique only to the Metastasis dataset.

In a few cases, we re-trained (repeated training from the previous checkpoint) again, which showed a very slight validation performance improvement. In retrospect, training longer (for more epochs) should achieve the same effect. During 5-fold training we maintained not only the best average checkpoint, but also the best 3 sub-region checkpoints (ET, TC, WT). We observed that e.g. training

twice (with a random initialization) will result in the same average dice, but individual dice scores per ET, TC, WT may differ between the train runs.

2.4 Ensembling

Our final result is an ensemble of 15 models (3 models from each fold) per each sub-region. Since during training we observed that the best average checkpoint may not be the best in all 3 sub-regions, we decided to ensemble models for each subregion (ET, TC, WT) separately using best sub-region checkpoints. Such approach, in the worst case, may require a total of 45 models, however in practice in most cases the best average checkpoint had also the best performance for all 3 or 2 of the sub-regions, so the total number of different model checkpoints used in the ensembling was about 25.

3 Results

Based on our 5-fold data split, average Dice (average of 3 sub-regions) for 5-folds cross-validation results of SegResNet model is shown in Table 1.

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
Adult Glioma	0.8076	0.925	0.9211	0.9188	0.9205	0.9196
Meningioma	0.9398	0.9403	0.939	0.9321	0.9251	0.93526
BraTS-Africa	0.9144	0.8332	0.9005	0.7199	0.8322	0.84004
Pediatric	0.8051	0.7844	0.794	0.7099	0.7393	0.76654

Table 1. Avg Dice metric using internal 5-fold cross-validation.

We have submitted 5 docker containers (for the 5 segmentation challenges), that vary in the model checkpoint weights. The organizers ran our method on the hidden testing set, and returned back the metrics shown in Tables 2, 3, 4, 5, 6, which we discuss below.

	Dice	HD95	Sensitivity	Specificity
ET	0.8076±0.25	40.88±94.81	0.8874±0.20	0.9997±0.0003
TC	0.8278±0.27	38.85±91.82	0.9020±0.21	0.9996±0.0010
WT	0.8356±0.22	42.01±82.24	0.9357±0.09	0.9995±0.0007
Avg	0.8236	40.58	0.9083	0.9996

Table 2. Task 1 Segmentation - Adult Glioma testing results (provided by the organizers).

Adult Glioma results based on the hidden test set (Table 2), are much lower on average compared to our internal validation, indicating the test set is substantially different from our 5-fold random split validation subsets (except for

	Dice	HD95	Sensitivity	Specificity
ET	0.7903±0.26	48.33±102.4	0.8738±0.21	0.9995±0.0004
TC	0.8844±0.26	33.96±97.62	0.8785±0.22	0.9999±0.0003
WT	0.9144±0.15	21.49±57.63	0.9479±0.03	0.9994±0.0004
Avg	0.8497	35.59	0.9001	0.9995

Table 3. Task 2 Segmentation - BraTS-Africa testing results (provided by the organizers).

the Fold-1, which matched the performance). This result achieves the overall 2nd rank in the Adult Glioma segmentation challenge.

BraTS-Africa results based on the hidden test set are shown in Table 3. The BraTS-Africa dataset was relatively small, but unique due its population. First, most lesions were late stage, as the disease is detected and diagnosed at relatively late stage, unfortunately, in some areas of Africa. Secondly, the MRI scanners are often outdated, which results in lower image quality. Nevertheless, our method achieves the 1st rank in the BraTS-Africa segmentation challenge.

Meningioma results are shown in Table 4. It was the first time, brain meningioma data was used in BraTS, and the organizers assembled a large dataset. Our method performance is better than for Adult Glioma segmentation (which had a similar dataset size), indicating that perhaps there is less variability in Meningioma tumors, or that our 5-fold cross validation matched the hidden test set better. Our method achieved the 1st rank in the brain Meningioma segmentation challenge.

	Dice	HD95	Sensitivity	Specificity
ET	0.8985±0.18	23.86±68.45	0.9364±0.12	0.9999±0.0001
TC	0.9035±0.17	21.82±64.66	0.9412±0.11	0.9999±0.0004
WT	0.8709±0.19	31.39±71.80	0.9317±0.12	0.9998±0.0002
Avg	0.8910	25.69	0.9364	0.9999

Table 4. Task 3 Segmentation - Meningioma testing results (provided by the organizers).

Metastasis results are shown in Table 5. Again it was the first time, brain metastasis (secondary tumor) data were used in BraTS. This data had the most variability, with some institutional sources had incomplete data/labels. Metastatic tumors are especially challenging since there could be many of them present in the brain, and some of them could be very small and barely visible in the 3D MRI. For these reasons, the average dice metric results are much lower than the corresponding metric for Glioma and Meningioma data. Nevertheless, our method achieved the 1 rank in the brain Metastasis segmentation challenge.

Finally, the Pediatric Glioma results are shown in Table 6. This data type also had its first ever appearance at BraTS. The pediatric glioma data is very rare,

	Dice	HD95	Sensitivity	Specificity
ET	0.6034±0.23	90.64±96.88	0.7541±0.23	0.9999±0.0001
TC	0.6513±0.23	83.11±93.26	0.7952±0.24	0.9999±0.0001
WT	0.6210±0.23	88.14±92.74	0.7700±0.24	0.9999±0.0004
Avg	0.6253	87.30	0.7731	0.9998

Table 5. Task 4 Segmentation - Metastasis testing results (provided by the organizers).

compared to the adult gliomas, hence the dataset size was small. Our method achieved the 2nd rank in the Pediatric glioma segmentation challenge.

	Dice	HD95	Sensitivity	Specificity
ET	0.5458±0.36	115.31±144.03	0.7740±0.20	0.9994±0.0010
TC	0.7815±0.19	27.09±72.11	0.7376±0.13	0.9997±0.0002
WT	0.8361±0.16	18.104±62.77	0.8041±0.09	0.9997±0.0003
Avg	0.7211	53.48	0.7719	0.9996

Table 6. Task 5 Segmentation - Pediatric testing results (provided by the organizers).

4 Discussion and Conclusion

We participated in 5 segmentation challenges of BraTS23 using Auto3DSeg from MONAI. We customized the ensembling workflow (to ensemble each sub-region), and the training for brain metastasis (to account for missing modalities). We achieved a strong performance, with the 1st place in the 3 (out of 5) challenges, and the 2nd place in the remaining two challenges.

References

1. Project-monai/monai, <https://doi.org/10.5281/zenodo.5083813>
2. Adewole, M., Rudie, J.D., Gbadamosi, A., Toyobo, O., Raymond, C., Zhang, D., Omidiji, O., Akinola, R., Suwaid, M.A., Emegoakor, A., Ojo, N., Aguh, K., Kalaiwo, C., Babatunde, G., Ogunleye, A., Gbadamosi, Y., Iorpagher, K., Calabrese, E., Aboian, M., Linguraru, M., Albrecht, J., Wiestler, B., Kofler, F., Janas, A., LaBella, D., Kzerooni, A.F., Li, H.B., Iglesias, J.E., Farahani, K., Eddy, J., Bergquist, T., Chung, V., Shinohara, R.T., Wiggins, W., Reitman, Z., Wang, C., Liu, X., Jiang, Z., Familiar, A., Leemput, K.V., Bukas, C., Piraud, M., Conte, G.M., Johansson, E., Meier, Z., Menze, B.H., Baid, U., Bakas, S., Dako, F., Fataade, A., Anazodo, U.C.: The brain tumor segmentation (brats) challenge 2023: Glioma segmentation in sub-saharan africa patient population (brats-africa) (2023)
3. Baid, U., Ghodasara, S., Bilello, M., Mohan, S., Calabrese, E., Colak, E., Farahani, K., Kalpathy-Cramer, J., Kitamura, F.C., Pati, S., Prevedello, L.M., Rudie, J.D., Sako, C., Shinohara, R.T., Bergquist, T., Chai, R., Eddy, J., Elliott, J., Reade,

- W., Schaffter, T., Yu, T., Zheng, J., Annotators, B., Davatzikos, C., Mongan, J., Hess, C., Cha, S., Villanueva-Meyer, J.E., Freymann, J.B., Kirby, J.S., Wiestler, B., Crivellaro, P., Colen, R.R., Kotrotsou, A., Marcus, D.S., Milchenko, M., Nazeri, A., Fathallah-Shaykh, H.M., Wiest, R., Jakab, A., Weber, M., Mahajan, A., Menze, B.H., Flanders, A.E., Bakas, S.: The RSNA-ASNR-MICCAI brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. *CoRR abs/2107.02314* (2021), <https://arxiv.org/abs/2107.02314>
4. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., John Freymann, K.F., Davatzikos, C.: Segmentation labels and radiomic features for the pre-operative scans of the tcga-gbm collection. *The Cancer Imaging Archive* (2017), <https://doi.org/10.7937/K9/TCIA.2017.KLXWJJ1Q>
 5. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., John Freymann, K.F., Davatzikos, C.: Segmentation labels and radiomic features for the pre-operative scans of the tcga-lgg collection. *The Cancer Imaging Archive* (2017), <https://doi.org/10.7937/K9/TCIA.2017.GJQ7R0EF>
 6. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., Davatzikos, C.: Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data* **4** (9 2017)
 7. Bakas, S., Reyes, M., et Int, Menze, B.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. In: *arXiv:1811.02629* (2018)
 8. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In: *International MICCAI Brainlesion Workshop*. pp. 272–284. Springer (2021)
 9. He, Y., Yang, D., Roth, H., Zhao, C., Xu, D.: Dints: Differentiable neural network topology search for 3d medical image segmentation. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 5841–5850 (2021)
 10. Karagyris, A., Umeton, R., Sheller, M., Aristizabal, A., George, J., Wuest, A., Pati, S., et al.: Federated benchmarking of medical artificial intelligence with med-perf. *Nature Machine Intelligence* **5**, 799–810 (2023)
 11. Kazerooni, A.F., Khalili, N., Liu, X., Haldar, D., Jiang, Z., Anwar, S.M., Albrecht, J., Adewole, M., Anazodo, U., Anderson, H., Bagheri, S., Baid, U., Bergquist, T., Borja, A.J., Calabrese, E., Chung, V., Conte, G.M., Dako, F., Eddy, J., Ezhov, I., Familiar, A., Farahani, K., Haldar, S., Iglesias, J.E., Janas, A., Johansen, E., Jones, B.V., Kofler, F., LaBella, D., Lai, H.A., Leemput, K.V., Li, H.B., Maleki, N., McAllister, A.S., Meier, Z., Menze, B., Moawad, A.W., Nandolia, K.K., Pavaine, J., Piraud, M., Poussaint, T., Prabhu, S.P., Reitman, Z., Rodriguez, A., Rudie, J.D., Shaikh, I.S., Shah, L.M., Sheth, N., Shinohara, R.T., Tu, W., Viswanathan, K., Wang, C., Ware, J.B., Wiestler, B., Wiggins, W., Zapaishchykova, A., Aboian, M., Bornhorst, M., de Blank, P., Deutsch, M., Fouladi, M., Hoffman, L., Kann, B., Lazow, M., Mikael, L., Nabavizadeh, A., Packer, R., Resnick, A., Rood, B., Vossough, A., Bakas, S., Linguraru, M.G.: The brain tumor segmentation (brats) challenge 2023: Focus on pediatrics (cbtbn-connect-dipgr-asnr-miccai brats-peds) (2023)
 12. LaBella, D., Adewole, M., Alonso-Basanta, M., Altes, T., Anwar, S.M., Baid, U., Bergquist, T., Bhalerao, R., Chen, S., Chung, V., Conte, G.M., Dako, F., Eddy, J., Ezhov, I., Godfrey, D., Hilal, F., Familiar, A., Farahani, K., Iglesias, J.E., Jiang, Z., Johanson, E., Kazerooni, A.F., Kent, C., Kirkpatrick, J., Kofler, F., Leemput,

- K.V., Li, H.B., Liu, X., Mahtabfar, A., McBurney-Lin, S., McLean, R., Meier, Z., Moawad, A.W., Mongan, J., Nedelec, P., Pajot, M., Piraud, M., Rashid, A., Reitman, Z., Shinohara, R.T., Velichko, Y., Wang, C., Warman, P., Wiggins, W., Aboian, M., Albrecht, J., Anazodo, U., Bakas, S., Flanders, A., Janas, A., Khanna, G., Linguraru, M.G., Menze, B., Nada, A., Rauschecker, A.M., Rudie, J., Tahon, N.H., Villanueva-Meyer, J., Wiestler, B., Calabrese, E.: The asnr-miccai brain tumor segmentation (brats) challenge 2023: Intracranial meningioma (2023)
13. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E.R., Weber, M.A., Arbel, T., Avants, B.B., Ayache, N., Buendia, P., Collins, D.L., Cordier, N., Corso, J.J., Criminisi, A., Das, T., Delingette, H., Demiralp, C., Durst, C.R., Dojat, M., Doyle, S., Festa, J., Forbes, F., Geremia, E., Glocker, B., Golland, P., Guo, X., Hamamci, A., Iftekharuddin, K.M., Jena, R., John, N.M., Konukoglu, E., Lashkari, D., Mariz, J.A., Meier, R., Pereira, S., Precup, D., Price, S.J., Raviv, T.R., Reza, S.M.S., Ryan, M.T., Sarikaya, D., Schwartz, L.H., Shin, H.C., Shotton, J., Silva, C.A., Sousa, N., Subbanna, N.K., Szekey, G., Taylor, T.J., Thomas, O.M., Tustison, N.J., Unal, G.B., Vasseur, F., Wintermark, M., Ye, D.H., Zhao, L., Zhao, B., Zikic, D., Prastawa, M., Reyes, M., Leemput, K.V.: The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans. Med. Imaging* **34**(10), 1993–2024 (2015)
 14. Moawad, A.W., Janas, A., Baid, U., Ramakrishnan, D., Jekel, L., Krantchev, K., Moy, H., Saluja, R., Osenberg, K., Wilms, K., Kaur, M., Avesta, A., Pedersen, G.C., Maleki, N., Salimi, M., Merkaj, S., von Reppert, M., Tillmans, N., Lost, J., Bousabarah, K., Holler, W., Lin, M., Westerhoff, M., Maresca, R., Link, K.E., hoda Tahon, N., Marcus, D., Sotiras, A., LaMontagne, P., Chakrabarty, S., Teytelboym, O., Youssef, A., Nada, A., Velichko, Y.S., Gennaro, N., Students, C., of Annotators, G., Cramer, J., Johnson, D.R., Kwan, B.Y.M., Petrovic, B., Patro, S.N., Wu, L., So, T., Thompson, G., Kam, A., Perez-Carrillo, G.G., Lall, N., of Approvers, G., Albrecht, J., Anazodo, U., Lingaru, M.G., Menze, B.H., Wiestler, B., Adewole, M., Anwar, S.M., Labella, D., Li, H.B., Iglesias, J.E., Farahani, K., Eddy, J., Bergquist, T., Chung, V., Shinohara, R.T., Dako, F., Wiggins, W., Reitman, Z., Wang, C., Liu, X., Jiang, Z., Leemput, K.V., Piraud, M., Ezhov, I., Johanson, E., Meier, Z., Familiar, A., Kazerooni, A.F., Kofler, F., Calabrese, E., Aneja, S., Chiang, V., Ikuta, I., Shafique, U., Memon, F., Conte, G.M., Bakas, S., Rudie, J., Aboian, M.: The brain tumor segmentation (brats-mets) challenge 2023: Brain metastasis segmentation on pre-treatment mri (2023)
 15. Myronenko, A.: 3D MRI brain tumor segmentation using autoencoder regularization. In: *International MICCAI Brainlesion Workshop*. pp. 311–320. Springer (2018)
 16. Tang, Y., Yang, D., Li, W., Roth, H.R., Landman, B., Xu, D., Nath, V., Hatamizadeh, A.: Self-supervised pre-training of swin transformers for 3d medical image analysis. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 20730–20740 (2022)