An Efficient Cascade of U-Net-like Convolutional Neural Networks devoted to Brain Tumor Segmentation

Philippe Bouchet[1], Jean-Baptiste Deloges[1], Hugo Canton-Bacara[1], Gaëtan Pusel[1], Lucas Pinot[2], Othman Elbaz[2], and Nicolas Boutry[3]

[1] EPITA Majeure Santé, Le Kremlin-Bicêtre 94270, France
[2] EPITA Majeure SCIA, Le Kremlin-Bicêtre 94270, France
[3] EPITA Research and Development Laboratory (LRDE), Le Kremlin-Bicêtre 94270, France

Abstract. A *glioma* is a fast-growing and aggressive tumor that starts in the glial cells of the brain. They make up about 30% of all brain tumors, and 80% of all malignant brain tumors. Gliomas are considered to be *rare* tumors, affecting less than 10,000 people each year, with a 5-year survival rate of 6%. If intercepted at an early stage, they pose no danger; however, providing an accurate diagnosis has proven to be difficult. In this paper, we propose a cascade approach using state-of-the-art Convolutional Neural Networks, in order to maximize accuracy in tumor detection. Various U-Net-like networks have been implemented and tested in order to select the network best suited for this problem.

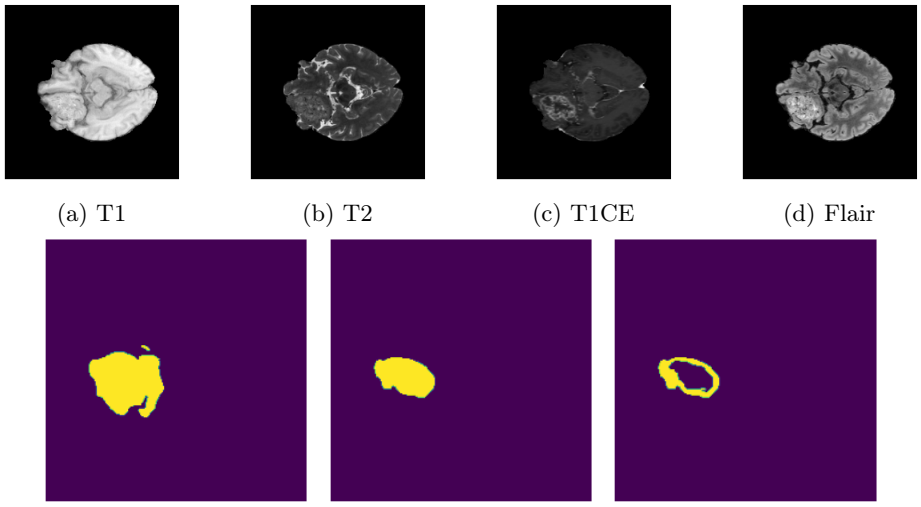
Keywords: Neural Networks Cascade, Deep Learning, Brain Tumor Segmentation, U-Net, Res-U-Net, U-Net++, Attention U-Net, Attention ResU-Net

1 Introduction

The MICCAI BraTS[1] is a challenge that occurs every year since 2012, organized by the Radiological Society of North America, the American Society of Neuroradiology and the Medical Image Computing and Computer Assisted Interventions society (RSNA-ASNR-MICCAI BraTS). The goal of the competition is the evaluation of state-of-the-art methods for the segmentation of intrinsically heterogeneous glioblastoma sub-regions in mpMRI[2] scans using multi-institutional pre-operative baseline multi-parametric magnetic resonance imaging scans[4][3].

As stated in the abstract, intercepting the glioma at an early stage *greatly* increases the patient's survival rate. However, in order to intercept the tumor, it must be accurately diagnosed. Unfortunately, most modern hospitals are not equipped with the technology capable of automatically detecting gliomas[13], thus causing the tumor to grow until it becomes deadly.

We have been provided with clinically acquired training data, as well as the corresponding segmentation labels of the different glioma sub-regions: Enhancing Tumor (ET), Tumor Core (TC) and the Whole Tumor (WT).



(a) T1 (b) T2 (c) T1CE (d) Flair

(e) WT (f) TC (g) ET

Fig. 1: Modalities and corresponding labels.

As depicted in 1, the WT describes the complete extent of the tumor. It entails the TC and ET, which is typically depicted by hyper-intense signal in FLAIR. The TC describes the bulk of the tumor. The TC entails the ET, as well as the necrotic parts of the tumor (NCR). The ET is described by areas that show hyper-intensity in T1CE when compared to T1, but also when compared to "healthy" white matter in T1CE. The voxel values for these images are:

- 1: Necrotic Tumor
- 2: Peritumoral edematous/invaded tissue
- 4: Enhancing tumor
- 0: Everything else

Furthermore, by applying the logical bitwise *or* operation on these voxel values, we can find the different tumor regions.

- $1 \cup 2 \cup 4$: Whole Tumor
- $1 \cup 4$: Tumor Core
- 4: Enhanced Tumor

In this paper, we propose a three-stage cascaded network, implementing a state-of-the-art variation of the U-Net at each stage. Each U-Net model will

specialize in segmenting the whole tumor, the tumor core and the enhanced tumor, in that order.

2 State-of-the-Art

2.1 Main architectures in biomedical image segmentation

Since the apparition of the U-Net [19] in 2016 in the biomedical image segmentation community, it has been established as the gold standard of medical image segmentation ever since.[14] We've seen numerous variations of the model, such as the U-Net++ [24], the ResU-Net [5][6], or the Attention U-Net [16][15]. Many different state-of-the-art architectures have since been proposed and benchmarked in previous iterations of the BraTS challenge[14]. Concerning these different architectures, let us recall their main advantages compared to the standard U-Net.

2.2 U-Net++

The U-Net++ [23] is a dense network, rendering it capable to extract rather minute details from an image or a volume. Moreover, the network also presents two output methods: a regular single output image, or deep supervision[9], with multiple outputs. Since each output has its own loss, we are able to calculate a weighted sum of the loss, allowing us to have more accurate results, at the cost of slower computation. In contrast, the single output mode is slightly less accurate than the deep supervision mode, however computation time will be much shorter as there is only a single output. It was selected for its versatility.

2.3 Attention ResU-Net

The Attention ResU-Net [12], is a combination of the Attention U-Net [16][15] and the ResU-Net [5][6], seeking to make use of each of these network's advantages. The ResU-Net is interesting, since it uses residual blocks in order to eliminate gradient related issues, such as the vanishing gradient problem. The Attention U-Net highlights only the relevant activations during training, meaning that the relevant parts of the image get large weights, thus reducing computational resources that are wasted on irrelevant activations of certain weights[17].

2.4 Existing Cascade Networks

Note that we also found several research papers related to cascades of networks used in the context of biomedical image segmentation. Among them, we can cite Wang G. *et al.* [22] who propose a cascade designed to decompose the multi-class segmentation problem into a sequence of three binary segmentation problems according to the sub region hierarchy. Y. Guo et al. [8] proposed a *Bidirectional Symmetric Cascade Network* (BSCN), such that each layer is supervised by vessel contour labels of specific diameter scale instead of the usual supervised approach which consists of using a ground truth to train different network layers.

3 Methodology

As stated in the opening abstract, this task was carried out using the Python machine learning frameworks *Tensorflow 2.9.1* as well as *numpy* and *nibabel* for handling volumes. We used OVHCloud, as well as the ROMEO supercomputer from the University of Reims Champagne-Ardenne in order to train the models. When we worked on the OVH platform, we had 4 Tesla V100S GPU's in parallel with each 32 Gb of VRAM, and when we worked on ROMEO, we got 4 Tesla P100-SXM2 GPU's with each 16Gb of VRAM. We obtained our benchmark results from testing with 200 patients from the dataset.

3.1 Motivation

Due to the nature of the segmentation problem, we propose a three-stage cascaded model, with variations of state-of-the-art U-Net models at each stage. We seek to train each model of our cascaded ensemble to specialize in the segmentation of different sub-regions of the glioblastoma. Since each sub-region of the brain tumor is contained within its predecessor, we can provide the subsequent models of our cascade with additional information pertaining to the location of the tumor. Moreover, using a three-stage cascaded model allows us to approach the problem with modularity in mind. We are not restricted to using the same model for each stage of our cascade. Instead, we can experiment with different types of U-Net-like models at different stages, for optimal performance.

3.2 Memory management

In previous years, the MICCAI BraTS challenge provided competitors with a dataset containing the scans of approximately 350 patients. This year we have 1251 patients. The main issue of these data sets is known to be the heterogeneity of the data. Managing and accessing the entirety of the data posed to us an infrastructure problem, especially for regular computers with very little amounts of RAM and VRAM, as we simply cannot load all of the data into one big array.

Much part of our time time has been spent designing and setting up the framework in which we would tackle the problem. After having come up with a solution to store all the data in numpy arrays on the disk using the NumPy's *memmap* function, we quickly implemented a framework in order to generate the data and build the models.

3.3 Data formatting

The medical volumes of this challenge are of dimensions $(240, 240, 155)$, and we decided to combine each modality (T1, T2, T1CE, FLAIR) into a single volume with 4 channels; thus giving us a dimension of $(240, 240, 155, 4)$ for our input data. The ground truth's dimensions remained the same.

Due to this year's limitations in VRAM (limited to 16GB), running a fully 3-dimensional network has proven to be a difficult and time consuming task. In order to preserve the information pertaining to the shape of the tumor in 3D, we opted for a 2.5D approach, that is, for each modality we fed the network with three horizontal slices of a single patient, instead of the whole 3D volume.

So we compared the 2D and the 2.5D approach for each architecture and each cascade, and found that the 2.5D approach resulted in far more accurate predictions, as more data was being fed into the network at each epoch.

3.4 Normalization

For the normalization step, we proceed in the following manner. Let N be the number of voxels of a given modality x among the T1, T1CE, T2, FLAIR modalities, then the indices i of the voxels of x belonging to $[0, N[$ with $N = 240 \times 240 \times 155$. This leads to:

$$\forall i \in [1, N],\ x_i^{\text{norm}} = \min\left(\max\left(\frac{x_i - \mu}{5\ \sigma}, -1.0\right), 1.0\right) \tag{1}$$

with

$$\mu = \frac{1}{\#(\mathcal{NZ})} \sum_{i \in \mathcal{NZ}} x_i,$$

$$\sigma = \frac{1}{\#(\mathcal{NZ})} \sqrt{\sum_{i \in \mathcal{NZ}} (x_i - \mu)^2}.$$

where the indexes of the non zeros values of the modality x is denoted by $\mathcal{NZ} = \#\{i\ ;\ x_i \neq 0\}$ and $\#$ denotes the cardinality. Note that the operators min and max are used here to ensure that each value of the normalized input x^{norm} lies in $[-1, 1]$.

3.5 Learning step

We used Keras' built-in *Adam* optimizer, with a constant learning rate of $\alpha_0 = 10^{-4}$. The maximal number of epochs was set at 100 and we used an early stopping based on the validation loss with a patience equal to 5. In order to ensure that we preserved the best network weights, we used the checkpoint callback with the best weights restoration to `True`.

3.6 Loss function

For this segmentation task, a voxel can belong to either one of two classes. One label for the tumor, and one for everything else. Knowing this, we can use binary cross-entropy loss at each stage of our cascade to calculate the difference between the network's prediction of the tumor, and the ground truth. We define the binary cross-entropy function BCE:

$$\text{BCE} = -\frac{1}{N}\sum_{i=1}^{N}(y_i \ \log(\hat{y_i}) + (1-y_i) \ \log(1-\hat{y_i})) \tag{2}$$

3.7 Evaluation of the model

Once the model is trained, we can predict an output X, and use the Dice function to check the similarity with the ground truth Y. A value $\epsilon > 0$ is added to avoid a division by zero error, it is set to a very small number as not to modify the accuracy of the Dice function. We recall the Dice formula (up to the added ϵ avoiding divisions by zero):

$$\text{Dice} = \frac{2 * |X| \cap |Y|}{|X| + |Y| + \epsilon}, \text{ with } \epsilon = 10^{-6}. \tag{3}$$

The closer the Dice coefficient is to 1, the more accurate the prediction is.

4 Our contributions

The main architectures that won the first places in the BraTS MICCAI challenge in 2020 [1] were all based on U-Net models [18,20,10].

Having noticed this, we decided that it would be interesting to test variations of U-Net networks in order to find the network that is best suited for the task of brain tumor segmentation. Some of these networks, however, have not yet been implemented in Python, so we took it upon ourselves to implement some of these networks. For example, we designed and implemented an algorithm[4] capable of generating a U-Net++ network of size L, as its size can vary.

The main difference between Wang. G *et al.* [22] and our approach is that here we consider at each step the *optimal* network in a set of networks which will be described in subsection 4.2.

4.1 Data analysis

In an effort to understand the data that was provided to us by the MICCAI BraTS challenge, we took it upon ourselves to compute the histogram of the cardinality of non-zero voxels in each patient's volume (we used the FLAIR modality as a reference). For the sake of clarity, we ordered these cardinalities in an increasing manner. We found that there was a sizeable amount of corrupted data, with patients either lacking in voxels, or having an over abundance of voxels in their volumes. We applied this analysis on the entirety of the 1251 patients, as well as the 219 patients in the validation data.

As expected, the distribution of the validation data was the same as the training data, as depicted on Figure 2. However we were not able to define a threshold corresponding to the minimal size of the brain volume ensuring that data is not "corrupted".

[4] `https://github.com/sudomane/unetpp/blob/main/unetpp/unetpp.py`

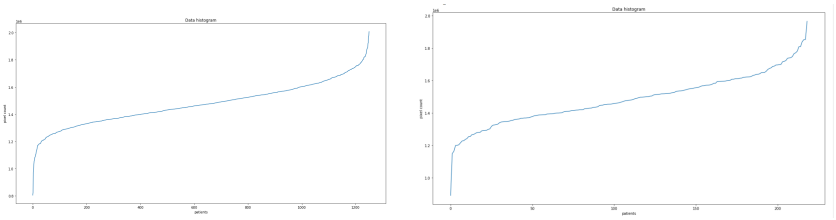


Fig. 2: Brains volumes distributions of training (left) and validation (right) patients

4.2 Proposing a new Cascade of Neural Networks

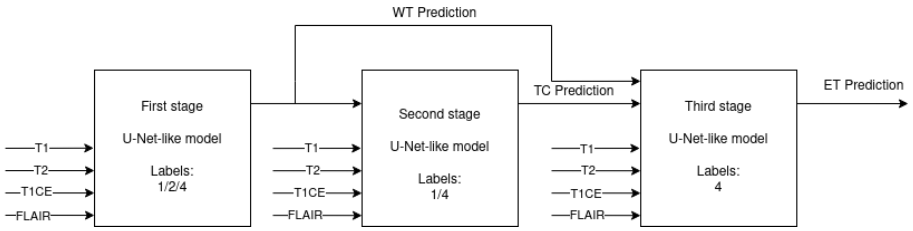


Fig. 3: Flow of data in a cascade model with 3 networks

The idea in this challenge is to build a cascaded model implementing three different U-Net-like networks with sigmoidal outputs (so that each output remains between 0 and 1). Each step of the cascade corresponds to the U-Net-like model with the best accuracy. The accuracy is evaluated with the Dice function in Eq. 3 defined earlier. Ultimately, the cascade will be composed of nothing but the most accurate networks at each step (see Figure 5).

The training procedure is carried out in four steps: one for each network fed with the predictions of the previous ones, and a fourth step to combine the outputs of each networks. Let us define Ξ_k as the function corresponding to the k^{th} network, we obtain the following formulas:

$$\widetilde{WT} = \Xi_1(FLAIR, T1, T1CE, T2),$$
$$\widetilde{TC} = \Xi_2(FLAIR, T1, T1CE, T2, \widetilde{WT}),$$
$$\widetilde{ET} = \Xi_3(FLAIR, T1, T1CE, T2, \widetilde{WT}, \widetilde{TC}).$$

We then apply a voxel-wise binarization operator that sets the value of our voxel to 1 if its original value is greater than 0.5, and 0 otherwise. We define this binarization function b:

$$\forall x \in [0,1], b(x) = \begin{cases} 1 & \text{if } x > 0.5 \\ 0 & \text{otherwise} \end{cases}$$

Now that we have defined our binarization function, we can proceed to reconstruct the 3D volume of our prediction by using the output from each stage of our cascade. Let us establish the formula to reconstruct the prediction volume:

$$4*b(\widetilde{ET})*b(\widetilde{TC})*b(\widetilde{WT})+2*(1-b(\widetilde{TC}))*b(\widetilde{WT})+(1-b(\widetilde{ET}))*b(\widetilde{TC})*b(\widetilde{WT}).$$

5 Results

5.1 Scores on training data set

Table 1: WT Dice results

	Mean	Std. Dev	1st Quartile	3rd Quartile
U-Net	86.42%	13.77%	82.68%	95.51%
U-Net++	**88.54%**	12.04%	86.95%	95.75%
ResU-Net	88.41%	12.59%	87.38%	95.82%
Attention U-Net	83.39%	20.88%	83.00%	95.22%
Attention ResU-Net	86.02%	17.81%	85.46%	95.38%

Table 2: TC Dice results

	Mean	Std. Dev	1st Quartile	3rd Quartile
U-Net	**89.26%**	17.26%	89.97%	97.54%
U-Net++	84.31%	21.35%	84.11%	95.42%
ResU-Net	82.97%	24.20%	81.22%	96.71%
Attention U-Net	81.14%	26.83%	82.42%	96.01%
Attention ResU-Net	85.10%	23.86%	88.13%	97.07%

Table 3: ET Dice results

	Mean	Std. Dev	1st Quartile	3rd Quartile
U-Net	88.19%	20.84%	90.94%	97.44%
U-Net++	80.57%	23.67%	79.12%	93.71%
ResU-Net	89.59%	20.14%	92.39%	98.09%
Attention U-Net	89.62%	20.44%	91.70%	98.36%
Attention ResU-Net	**89.94%**	20.61%	93.00%	98.37%

Note: The U-Net++ model was tested in single output mode.

After having run the predictions on the entirety of the train data set, we collected the information pertaining to the accuracy of the models given to us by the Dice evaluation (see Tables 1,2,3).

5.2 Scores on validation data

Table 4: Dice and HD95 of our optimal model.

Dice	WT (Mean)	WT (Std)	TC (Mean)	TC (Std)	ET (Mean)	ET (Std)
	88.68%	12.43%	80.67%	25.07%	75.35%	27.67%
HD95	WT (Mean)	WT (Std)	TC (Mean)	TC (Std)	ET (Mean)	ET (Std)
	11.54	21.23	17.79	56.83	27.79	87.84

6 Discussion

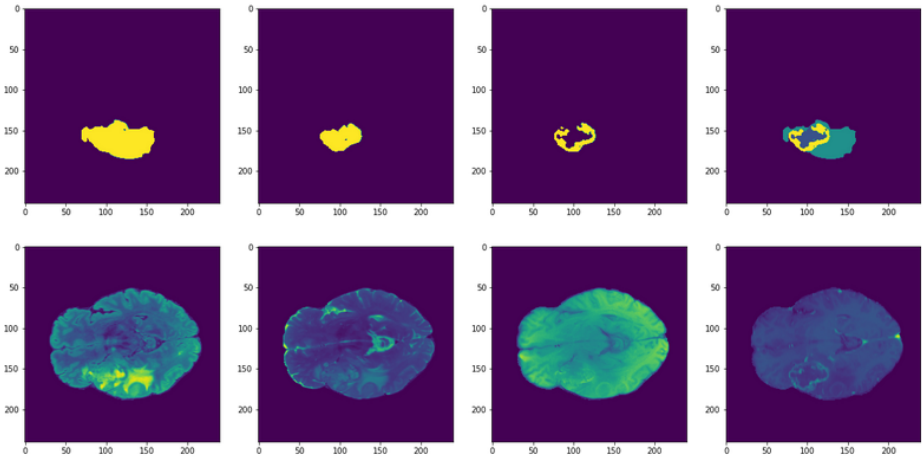


Fig. 4: Prediction on validation data with the complete cascade implementation. In the top row, from left to right: WT prediction, TC prediction, ET prediction, complete prediction.

At the first step (WT), we notice that the network with the best results is the U-Net++, with a mean Dice score of 88.54%. Then at the second step (TC), we notice that the network with the best results is a standard U-Net, with a mean Dice score of 89.26%. Finally, at the third step (ET), we notice that the network with the best results is the Attention ResU-Net, with a mean dice score of 89.94%.

The label for the extended tumor (ET) has proven to be the most difficult to segment, as it is defined with minute detail. However, the Attention ResU-Net performs the best for this label (see Table 3). This can be due to the mechanisms present in the model, such as the use of the attention mechanism, as well as residual blocks which help in eliminating the vanishing gradient problem[12].

Because of its inherent advantages, the Attention ResU-Net seems to be the best suited for the task of segmenting fine detail in MRI scans.

In summary, a cascade model made up of a U-Net++ for the first step, a standard U-Net for the second step, and an Attention ResU-Net for the third step, would provide us with the best results (see Figure 4).

7 Proposed solution

Based on the benchmark results, we have chosen the following networks for each region:

- Whole Tumor: U-Net++[23]
- Tumor Core: U-Net[19]
- Enhanced Tumor: Attention ResU-Net[12]

Now that we are able to determine the final architecture for the optimal cascaded U-Net model, here is a graph of the proposed architecture.

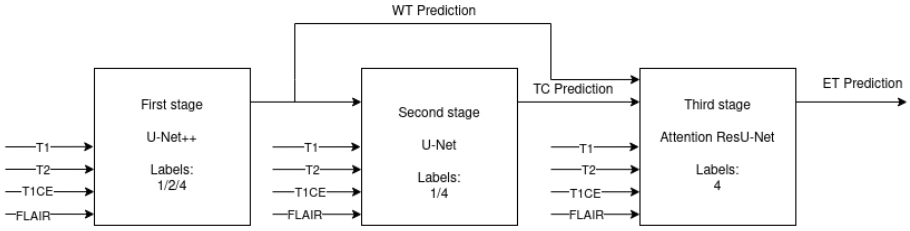


Fig. 5: Optimal cascade with a U-Net++, U-Net and an Attention ResU-Net

8 Final results

Here are the results on the test data set that MICCAI uses to evaluate the model.

Table 5: Results on African Data

	Dice ET	Dice TC	Dice WT	Hausdorff ET	Hausdorff TC	Hausdorff WT
Mean	0.8440	0.8943	0.9370	3.324	6.034	4.280
Std.	0.1450	0.1148	0.0615	2.911	8.454	6.695
Median	0.9042	0.9286	0.9540	2.911	4	2.236
1st quartile	0.8553	0.8874	0.9368	1.414	1.414	1.414
3rd quartile	0.9283	0.9623	0.9656	3.741	6.034	3.162

Table 6: Testing cohort

	Dice ET	Dice TC	Dice WT	Hausdorff ET	Hausdorff TC	Hausdorff WT
Mean	0.7938	0.8227	0.8803	22.19	21.70	11.69
Std.	0.2449	0.2628	0.1357	80.33	71.84	19.56
Median	0.8781	0.9268	0.9250	1.414	3	3.741
1st quartile	0.7912	0.8438	0.8686	1	1.414	2
3rd quartile	0.9316	0.9619	0.9514	3	8.124	8.747

Table 7: Results on Pediatric Data

	Dice TC	Dice WT	Hausdorff TC	Hausdorff WT
Mean	0.2288	0.6657	154.97	30.47
Std.	0.3275	0.2725	167.89	51.04
Median	0	0.7592	53.074	14.90
1st quartile	0	0.5662	10.998	5.692
3rd quartile	0.4429	0.8669	373.12	34.79

9 Conclusion

9.1 Future works

Despite the scores we obtained, we strongly believe in the potential of the three-stage cascade method for the BraTS challenge. Each network of the cascade being able to develop specific features devoted to segmenting a specific sub-region of the glioblastoma. Moreover, each network providing additional information to the subsequent stages in the cascade, causing a guiding effect for each network.

In the future, we would like to implement and test additional features in our three-stage cascaded network. Here are some of the features we would like to test:

- Implement and benchmark the nn-U-Net [11].
- Benchmark with 851 patient instead of 200.
- Improve the overall quality of the training data set by preprocessing data.
- Implement cross-validation.

9.2 Summary

In this paper, we proposed a three-stage cascaded network devoted to brain tumor segmentation. The U-Net-like network employed at each step of the cascade was determined through benchmarking several U-Net variants in order to determine which network performs the best at a certain stage. Moreover, our 2.5D method allows us to optimize the usage of the GPU's VRAM, with minimal performance loss.

9.3 Acknowledgements

We would like to thank Guy Fournier, Bastien Verdebout, Kevin Amil and Christina De Azevedo of OVHcloud, for providing us with the infrastructure and resources that have proven to be detrimental to our success. Without their support, we likely would not have succeeded in our task.

We would also like to thank Arnaud Renard and his team for granting us access to the ROMEO Supercomputer of the University of Reims.

Moreover, we would like to thank Philippe Dewost, Christian Chabrerie, and Patrick DeMichel for supporting us throughout the entirety of this project.

References

1. Baid, U., Ghodasara, S., Mohan, S., Bilello, M., Calabrese, E., Colak, E., Farahani, K., Kalpathy-Cramer, J., Kitamura, F.C., Pati, S., et al.: The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. arXiv.org (Sep 2021), `https://arxiv.org/abs/2107.02314`
2. 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. Nature Scientific Data **4** (09 2017). https://doi.org/10.1038/sdata.2017.117
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 (07 2017). https://doi.org/10.7937/K9/TCIA.2017.KLXWJJ1Q
4. 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 (07 2017). https://doi.org/10.7937/K9/TCIA.2017.GJQ7R0EF
5. Diakogiannis, F., Waldner, F., Caccetta, P., Wu, C.: Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS Journal of Photogrammetry and Remote Sensing **16**, 94–114 (02 2020). https://doi.org/10.1016/j.isprsjprs.2020.01.013

6. Diakogiannis, F.I., Waldner, F., Caccetta, P., Wu, C.: Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data (Jan 2020), https://arxiv.org/abs/1904.00592
7. Greenwade, G.D.: The Comprehensive Tex Archive Network (CTAN). TUGBoat **14**(3), 342–351 (1993)
8. Guo, Y., Peng, Y.: Bscn: bidirectional symmetric cascade network for retinal vessel segmentation. BMC medical imaging **20**(1), 1–22 (2020), https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880-020-0412-7
9. Lee, C.Y., Xie, S., Gallagher, P., Zhang, Z., Tu, Z.: Deeply-supervised nets (2014). https://doi.org/10.48550/ARXIV.1409.5185, https://arxiv.org/abs/1409.5185
10. Luu, H.M., Park, S.H.: Extending nn-unet for brain tumor segmentation (Dec 2021), https://arxiv.org/abs/2112.04653
11. Luu, H.M., Park, S.H.: Extending nn-unet for brain tumor segmentation. In: International MICCAI Brainlesion Workshop. pp. 173–186. Springer (2022)
12. Maji, D., Sigedar, P., Singh, M.: Attention res-unet with guided decoder for semantic segmentation of brain tumors. Biomedical Signal Processing and Control **71**, 103077 (2022). https://doi.org/https://doi.org/10.1016/j.bspc.2021.103077, https://www.sciencedirect.com/science/article/pii/S1746809421006741
13. MD Prados, W.C.: Surgery for low-grade glioma: Rationale for early intervention, https://pubmed.ncbi.nlm.nih.gov/8846605/
14. 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., 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, Ã., 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., Sarikaya, D., Schwartz, L., Shin, H.C., Shotton, J., Silva, C.A., Sousa, N., Subbanna, N.K., Szekely, G., Taylor, T.J., Thomas, O.M., Tustison, N.J., Unal, G., Vasseur, F., Wintermark, M., Ye, D.H., Zhao, L., Zhao, B., Zikic, D., Prastawa, M., Reyes, M., Van Leemput, K.: The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Transactions on Medical Imaging **34**(10), 1993–2024 (2015). https://doi.org/10.1109/TMI.2014.2377694
15. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., et al.: Attention u-net: Learning where to look for the pancreas (May 2018), https://arxiv.org/abs/1804.03999
16. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D.: Attention u-net: Learning where to look for the pancreas (2018). https://doi.org/10.48550/ARXIV.1804.03999, https://arxiv.org/abs/1804.03999
17. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D.: Attention u-net: Learning where to look for the pancreas (2018). https://doi.org/10.48550/ARXIV.1804.03999, https://arxiv.org/abs/1804.03999
18. Peiris, H., Chen, Z., Egan, G., Harandi, M.: Reciprocal adversarial learning for brain tumor segmentation: A solution to brats challenge 2021 segmentation task (Jan 2022), https://arxiv.org/abs/2201.03777
19. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation (May 2015), https://arxiv.org/abs/1505.04597

20. Siddiquee, M.M.R., Myronenko, A.: Redundancy reduction in semantic segmentation of 3d brain tumor mris (Nov 2021), https://arxiv.org/abs/2111.00742
21. Vu, M.H., Nyholm, T., Löfstedt, T.: Tunet: End-to-end hierarchical brain tumor segmentation using cascaded networks. In: International MICCAI Brainlesion Workshop. pp. 174–186. Springer (2019)
22. Wang, G., Li, W., Ourselin, S., Vercauteren, T.: Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In: International MICCAI brainlesion workshop. pp. 178–190. Springer (2017), https://arxiv.org/pdf/1709.00382.pdf
23. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: Unet++: A nested u-net architecture for medical image segmentation (2018). https://doi.org/10.48550/ARXIV.1807.10165, https://arxiv.org/abs/1807.10165
24. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: Unet++: Redesigning skip connections to exploit multiscale features in image segmentation (Jan 2020), https://arxiv.org/abs/1912.05074