



HAL
open science

Patch-based Segmentation of Brain Tissues

Nicolas Cordier, Bjoern Menze, Hervé Delingette, Nicholas Ayache

► **To cite this version:**

Nicolas Cordier, Bjoern Menze, Hervé Delingette, Nicholas Ayache. Patch-based Segmentation of Brain Tissues. MICCAI Challenge on Multimodal Brain Tumor Segmentation, Sep 2013, Nagoya, Japan. pp.6 - 17. hal-00917084

HAL Id: hal-00917084

<https://inria.hal.science/hal-00917084>

Submitted on 16 Dec 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Patch-based Segmentation of Brain Tissues

Nicolas Cordier¹, Bjoern Menze^{1,2}, Hervé Delingette¹, and Nicholas Ayache¹

¹ Asclepios Research Project, INRIA Sophia-Antipolis, France

² Computer Vision Laboratory, ETH Zurich, Switzerland

Abstract. We describe our submission to the Brain Tumor Segmentation Challenge (BraTS) at MICCAI 2013. This segmentation approach is based on similarities between multi-channel patches. After patches are extracted from several MR channels for a test case, similar patches are found in training images for which label maps are known. These labels maps are then combined to result in a segmentation map for the test case. The labelling is performed, in a leave-one-out scheme, for each case of a publicly available training set, which consists of 30 real cases (20 high-grade gliomas, 10 low-grade gliomas) and 50 synthetic cases (25 high-grade gliomas, 25 low-grade gliomas). Promising results are shown on the training set, and we believe this algorithm would perform favourably well in comparison to the state of the art on a testing set.

1 Introduction

We describe our submission to the Brain Tumor Segmentation Challenge (BraTS) at MICCAI 2013. This segmentation approach is based on ideas similar to those developed for human brain labelling in [1]. A database of multi-channel patches is first built from a set of training cases, for which label maps are known. Then, given a test case, patches are extracted from several Magnetic Resonance (MR) channels, and similar multi-channel patches are retrieved in the patch database. Since each multi-channel intensity patch from the database is associated with a label patch, a combination of the labels can result in a segmentation map for the test case. First, we detail the pre-processing steps, mainly a global intensity alignment and tumour localization. Second, we recall the currently naive procedure to build the database of patches and retrieve similar patches. Third, we describe the label fusion step. Then, the labelling is performed, in a leave-one-out scheme, for each case of a publicly available training set, which consists of 30 real cases (20 high-grade gliomas, 10 low-grade gliomas) and 50 synthetic cases (25 high-grade gliomas, 25 low-grade gliomas). Finally, results and promising improvements to the current algorithm are discussed.

2 Method: Patch-based Segmentation

In this section, our patch-based segmentation algorithm is described in its most recent state. Current limitations could be bypassed with several promising improvements, which are still work-in-progress at the time of the submission of the article.

2.1 Pre-processing

In order to decrease the computation time, we first sub-sample , with a nearest-neighbour interpolation scheme, all the images to 2-mm isotropic resolution, as in [2]. In the end, to compare the resulting segmentation to the ground truth, we would perform an up-sampling to 1-mm isotropic resolution.

Moreover, we define a bounding box surrounding the tumour, and we crop the images. There are three main reasons to back this pre-processing step:

- computations are faster since there are less patches,
- patches are more relevant since we are only interested in segmenting the tumor and we focus on image parts containing a tumor,
- we avoid any problem related to image parts missing for some channels, since these happen outside these bounding boxes.

While these bounding boxes are known for the training cases, we do not know their localization for a test case. Given more time, we would intend to roughly localize the bounding box by applying a threshold to T_2 FLAIR images. In the meantime, we currently assume a bounding box around the tumor is known.

2.2 Global intensity alignment

Let D be the number of training cases, each of which consists of four MR channels and one label map. Before we build a multi-channel patch database, we would like to normalize each MR image in order to make all the D cases appear more similar to each other. This is a very important step since the similarity criterion, which drives the patch retrieval, is not invariant to affine intensity changes.

Since different Magnetic Resonance Imaging (MRI) scanners could be used for the acquisition of each case, a pre-processing step is mandatory to normalize the image intensities. Similarly to [2], we "align the mean intensities within each channel by a global multiplicative factor" K .

Let c be a channel, among T_1 -weighted, T_2 -weighted, FLAIR T_2 -weighted, and contrast enhanced T_1 -weighted images. Let x be a voxel position in the MR image I_c . Let $\langle I_c \rangle$ be the mean intensity for channel c .

The global intensity alignment is such that:

$$\forall c, \forall x, I_c(x) \leftarrow I_c(x) \times \frac{K}{\langle I_c \rangle}$$

where K is arbitrarily set to 700, which is a value close to the mean intensities. Currently, this step is performed with *cropped* images.

2.3 Building a database of patches and retrieving similar patches

For each case I of the D training cases, for each voxel x of I , we extract a 3x3x3 patch P_x^{Ic} per channel c , and we concatenate these patches to get a multi-channel patch P_x^I .

Given a new case I^{test} , which consists of the four MR channels previously mentioned, we compute a multi-channel patch $P_x^{I^{\text{test}}}$ for each voxel x .

For each patch in I^{test} , for each of the D reference cases, we find the k -nearest-neighbours with respect to our similarity measure, which is here the sum of squared differences $d(\cdot, \cdot)$ between multi-channel patches. In the following, $k = 5$.

For each voxel x of the test case I^{test} , we have retrieved the k most similar patches for each reference case, which makes $D \times k$ similar multi-channel patches in total, for which we know the corresponding 3D label patches.

2.4 Label fusion

In order to define the label map S for the test case I , a label fusion method is defined.

Vote weight for each retrieved patch For each voxel x in I^{test} , each retrieved similar patch $(P_{y_{n,m}}^{I_{n,m}})_{n \in [1,D], m \in [1,k]}$ contributes to a voting scheme, with a vote weight $w(x, \cdot)$ which depends on the similarity measure $d(\cdot, \cdot)$:

$$w(x, P_y^I) = \exp\left(-\frac{d(P_y^I, P_x^{I^{\text{test}}})}{\sigma^2}\right)$$

where σ^2 is the maximum (over every voxel $x \in I^{\text{test}}$ and every training case $n \in [1, D]$) of the L_2 error made with first-neighbour patches only ($m = 1$).

Aggregation of the votes Given vote weights for each retrieved patch, there are two possibilities to aggregate the votes. First, if we consider a voxel x , the simplest method increments the votes for the label found at the center of each retrieved patch. Second, the method we use in this paper makes use of the labels found in the whole patch, i.e. at the center x and in the $3 \times 3 \times 3$ neighbourhood of x of each label patch. Thus, we increment the votes for 27 voxels for each retrieved patch. This results in regularized vote maps.

Vote maps For each label, the aggregation of votes results in a *vote map*, which consists in a map of the aggregated vote weights. These vote maps can be seen as probability maps: they can be scaled by the total of vote increments so that the total of votes for each foreground voxel is 1, in which each vote is a non-negative value between 0 and 1. This rescaling does not affect the segmentation procedure though, as opposed to the rescaling described in the following.

Label selection In the Brain Tumor Segmentation Challenge, we are interested in the segmentation of three regions (complete tumor, tumor core, enhancing tumor) which show an interesting property: these regions are interlocked. Therefore, during label selection, we proceed with three consecutive steps:

1. distinguish between the background and the complete tumor,
2. inside the complete tumor, distinguish between the tumor core and what remains of the complete tumor (the edema),
3. inside the tumor core, distinguish between the enhancing tumor and what remains of the tumor core (the non-enhancing tumor and the necrotic core).

For each of these steps, we want to give more weights to labels which are less represented in the reference dataset, typically tumor classes. Indeed, less represented labels are less likely to be picked at the patch retrieval step, and in the case of tumor classes, the number of false negatives could be increased. Given a pool of two region labels to select from, we proceed as follows in order to achieve a penalization of the biggest region: for each region separately, we divide voxel-wise region votes by the number of voxels for which the region votes are non-zero. Therefore, we slightly favour labels for which only a few voxels have non-zero votes. Finally, for each voxel x , the label with the highest scaled vote is picked.

3 Evaluation

Our segmentation procedure is evaluated on 30 real cases (20 high-grade gliomas, 10 low-grade gliomas) and 50 synthetic cases (25 high-grade gliomas, 25 low-grade gliomas). The dataset is publicly available through the MICCAI 2013 Brain Tumor Segmentation challenge. We perform the labelling of each image of the training set, in a leave-one-out framework. The Dice score, Jaccard index, specificity and sensitivity are computed by the online evaluation tool provided by the organizers of the Brain Tumor Segmentation Challenge. Average scores are given in tables 1 to 4, per-patient scores are given in figures 1 and 2. The best segmentation maps for each of the four training datasets are shown in figures 3 to 10.

Dice score	Complete tumor	Tumor core	Enhancing tumor
mean	0.79	0.60	0.59
std-dev	0.17	0.26	0.25
median	0.87	0.72	0.69

Table 1. Evaluation summary for 20 real high-grade cases.

Dice score	Complete tumor	Tumor core	Enhancing tumor
mean	0.76	0.64	0.44
std-dev	0.18	0.21	0.40
median	0.80	0.71	0.45

Table 2. Evaluation summary for 10 real low-grade cases. Regarding the right-most column, 3 cases show enhancing tumor, and the Dice score is only computed for these.

Dice score	Complete tumor	Tumor core
mean	0.84	0.78
std-dev	0.08	0.12
median	0.85	0.83

Table 3. Evaluation summary for 25 synthetic high-grade cases.

Dice score	Complete tumor	Tumor core
mean	0.83	0.67
std-dev	0.04	0.19
median	0.83	0.74

Table 4. Evaluation summary for 25 synthetic low-grade cases.

4 Discussion

We have described a generic framework for brain tumor segmentation, which relies on similarities between multi-channel patches. The algorithm performs well on the training cases provided for the Brain Tumor Segmentation Challenge. We believe this algorithm would perform favourably well in comparison to the state of the art on a testing set. Promising improvements to the current algorithm are still in progress.

Regarding the pre-processing step, we do not expect the results to differ significantly with slightly different bounding boxes. Indeed, bounding boxes restrict the images in coronal and sagittal views, but the axial view is not cropped at all in order to keep more examples of healthy tissues, as is shown in figures 3 to 10. Moreover, let us assume we make an error during the localization of the bounding box:

- if we find a bounding box containing the true bounding box, then no false negative would be added, and we would expect the number of additional false positives to be low since the specificity of the classifier is close to 1 in practice,
- if we find a bounding box included in the true bounding box, then this may result in less false positives and this would necessarily result in more false negatives since parts of the tumour are necessarily wrongly omitted by the classifier. This is the case for synthetic high-grade cases, due to a human error.

Regarding the global intensity alignment step, we believe results would not be fundamentally different if the full brain mask were used instead of a bounding box. In fact, a global intensity alignment applied to the full images would be the first pre-processing step in order to localize the tumor.

Regarding the retrieval of similar patches in the patch database, a cluster is used since the search is naive and exhaustive. This paper in its current state is mostly a proof-of-concept, and we expect in the foreseeable future to tremendously decrease the computation time by relying on approximate hash-based searches.

Acknowledgements

Part of this work was funded by the European Research Council through the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

References

1. Rousseau, F., Habas, P.A., Studholme, C.: A supervised patch-based approach for human brain labeling. *Medical Imaging, IEEE Transactions on* **30**(10) (2011) 1852–1862
2. Zikic, D., Glocker, B., Konukoglu, E., Criminisi, A., Demiralp, C., Shotton, J., Thomas, O., Das, T., Jena, R., Price, S.: Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel mr. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*. Springer (2012) 369–376

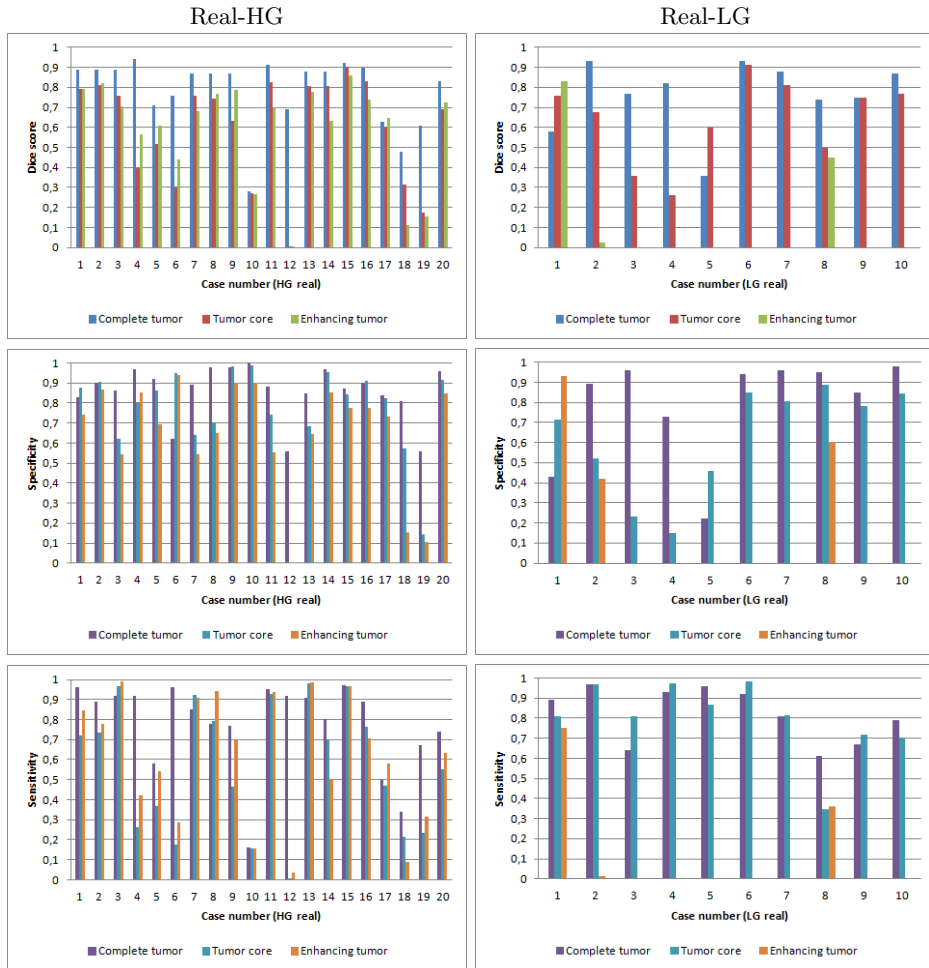


Fig. 1. Per-patient evaluation for the two real BraTS data sets (Real-HG, Real-LG). We show the results for three regions: complete tumor, tumor core, and, for all high-grade cases and a few low-grade cases only, enhancing tumor. We report the following measures: Dice score, Specificity, and Sensitivity.

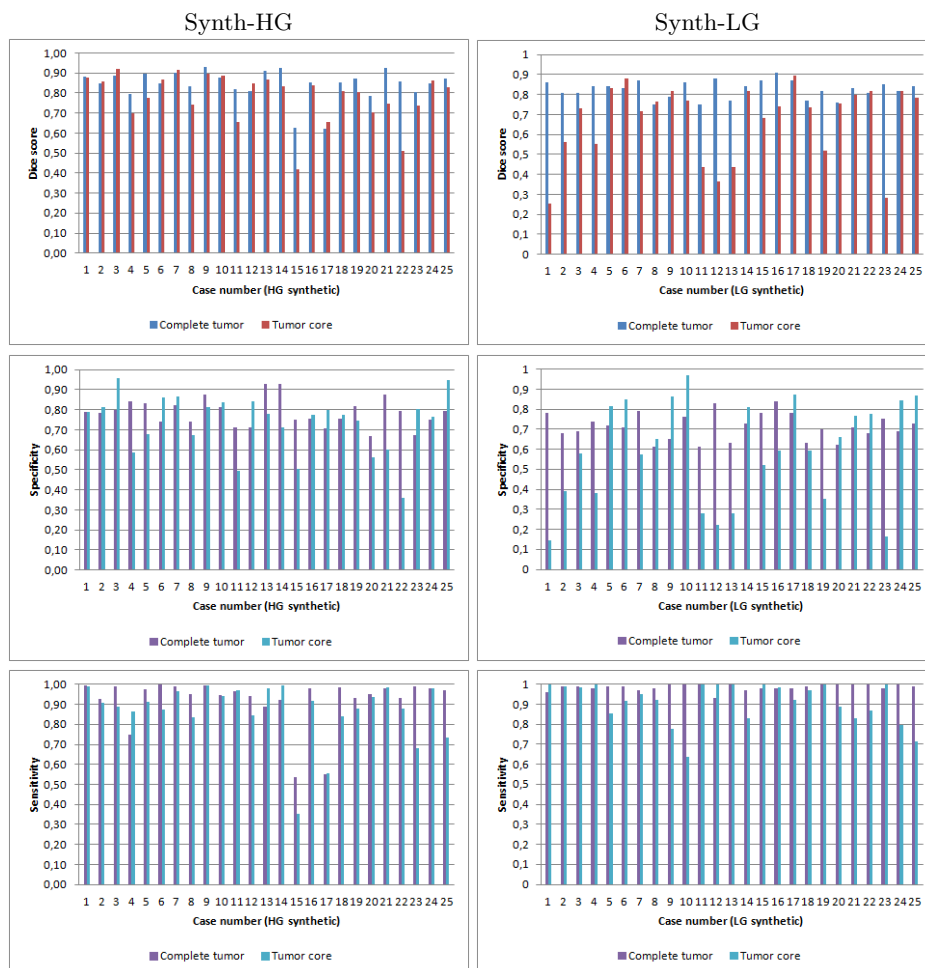


Fig. 2. Per-patient evaluation for the two synthetic BraTS data sets (Synth-HG, Synth-LG). We show the results for two regions: complete tumor, and tumor core. We report the following measures: Dice score, Specificity, and Sensitivity.

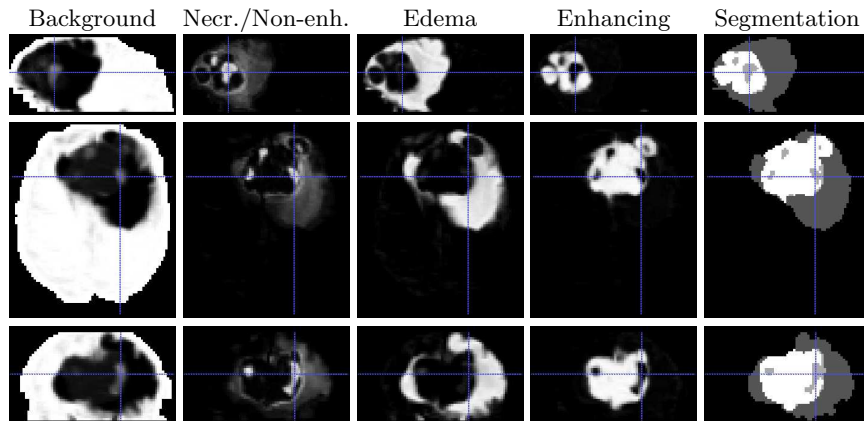


Fig. 3. Real high-grade case 15. From left to right: Vote maps for background, necrosis and non-enhancing tumor (merged), edema, enhancing tumor; Segmentation map. From top to bottom: saggital, axial, and coronal views. Images are cropped.

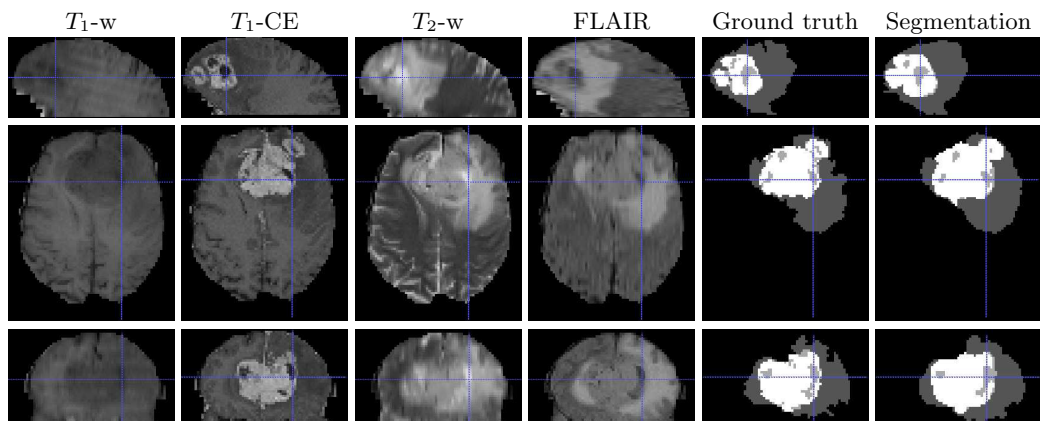


Fig. 4. Real high-grade case 15. From left to right: MR images, ground truth, and segmentation. From top to bottom: saggital, axial, and coronal views. Images are cropped.

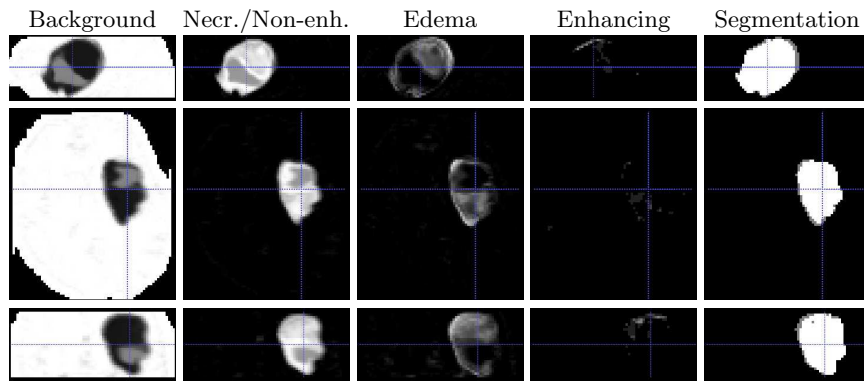


Fig. 5. Real low-grade case 06. From left to right: Vote maps for background, necrosis and non-enhancing tumor (merged), edema, enhancing tumor; Segmentation map. From top to bottom: sagittal, axial, and coronal views. Images are cropped.

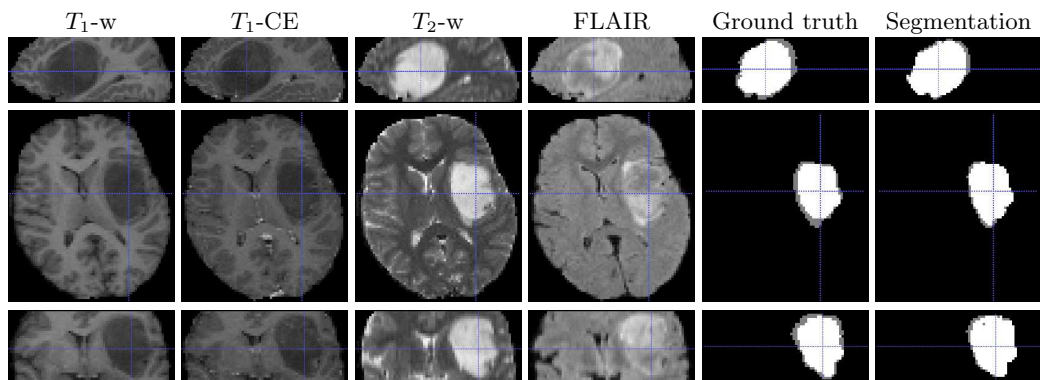


Fig. 6. Real low-grade case 06. From left to right: MR images, ground truth, and segmentation. From top to bottom: sagittal, axial, and coronal views. Images are cropped.

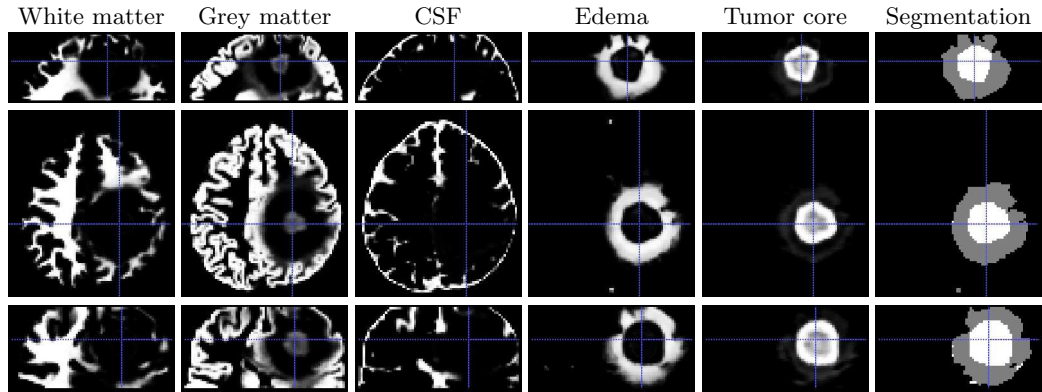


Fig. 7. Synthetic high-grade case 07. From left to right: Vote maps for white matter, grey matter, cerebro-spinal fluid (CSF), edema, and tumor core; Segmentation map. From top to bottom: sagittal, axial, and coronal views. Images are cropped.

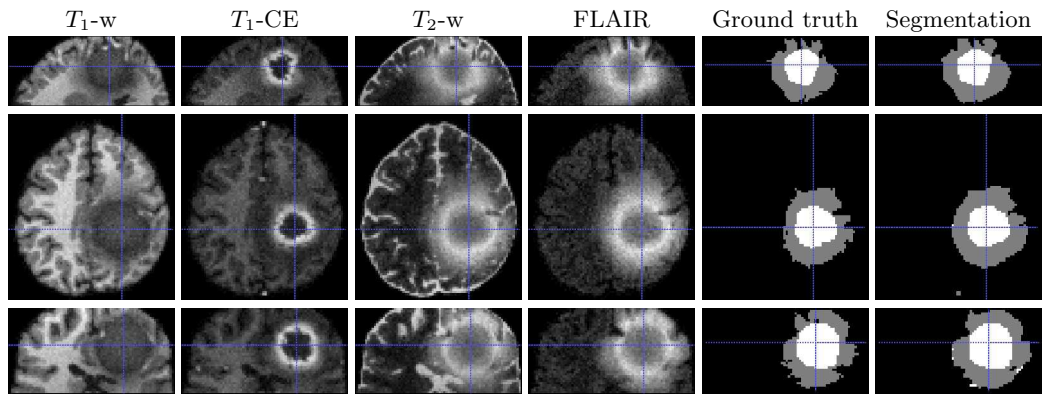


Fig. 8. Synthetic high-grade case 07. From left to right: MR images, ground truth, and segmentation. From top to bottom: sagittal, axial, and coronal views. Images are cropped.

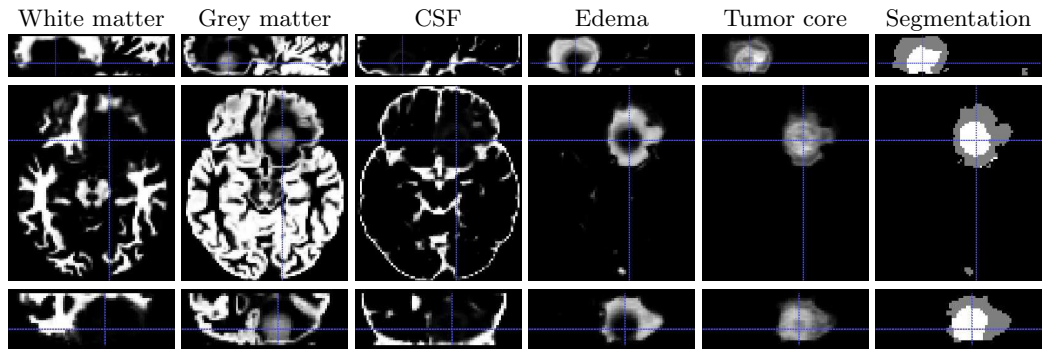


Fig. 9. Synthetic low-grade case 17. From left to right: Vote maps for white matter, grey matter, cerebro-spinal fluid (CSF), edema, and tumor core; Segmentation map. From top to bottom: sagittal, axial, and coronal views. Images are cropped.

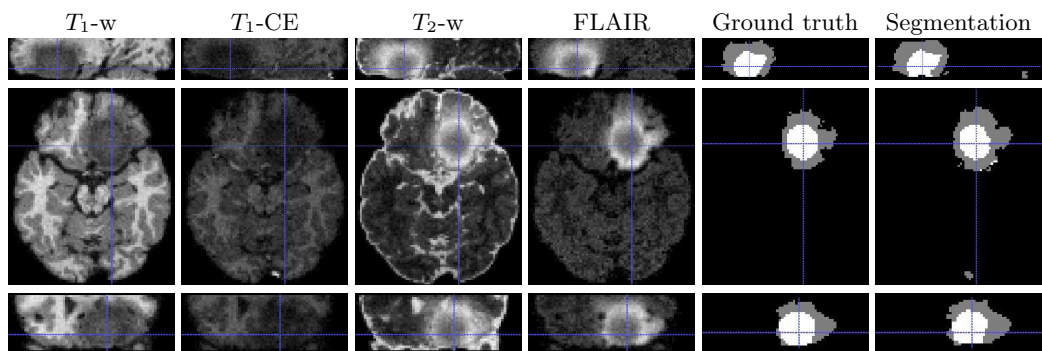


Fig. 10. Synthetic low-grade case 17. From left to right: MR images, ground truth, and segmentation. From top to bottom: sagittal, axial, and coronal views. Images are cropped.