



## **A generative model for segmentation of tumor and organs-at-risk for radiation therapy planning of glioblastoma patients**

**Agn, Mikael; Law, Ian; Munck Af Rosenschöld, Per; Van Leemput, Koen**

*Published in:*  
SPIE Medical Imaging 2016: Image Processing

*Link to article, DOI:*  
[10.1117/12.2216814](https://doi.org/10.1117/12.2216814)

*Publication date:*  
2016

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Agn, M., Law, I., Munck Af Rosenschöld, P., & Van Leemput, K. (2016). A generative model for segmentation of tumor and organs-at-risk for radiation therapy planning of glioblastoma patients. In *SPIE Medical Imaging 2016: Image Processing* (Vol. 9784). Article 97841D SPIE - International Society for Optical Engineering.  
<https://doi.org/10.1117/12.2216814>

---

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# A Generative Model for Segmentation of Tumor and Organs-at-Risk for Radiation Therapy Planning of Glioblastoma Patients

Mikael Agn<sup>a</sup>, Ian Law<sup>b</sup>, Per Munck af Rosenschöld<sup>c</sup>, and Koen Van Leemput<sup>a,d</sup>

<sup>a</sup>Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark

<sup>b</sup>Department of Clinical Physiology, Nuclear Medicine and PET, Rigshospitalet, Copenhagen University Hospital, Denmark

<sup>c</sup>Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Denmark

<sup>d</sup>Martinos Center for Biomedical Imaging, MGH, Harvard Medical School, USA

## ABSTRACT

We present a fully automated generative method for simultaneous brain tumor and organs-at-risk segmentation in multi-modal magnetic resonance images. The method combines an existing whole-brain segmentation technique with a spatial tumor prior, which uses convolutional restricted Boltzmann machines to model tumor shape. The method is not tuned to any specific imaging protocol and can simultaneously segment the gross tumor volume, peritumoral edema and healthy tissue structures relevant for radiotherapy planning. We validate the method on a manually delineated clinical data set of glioblastoma patients by comparing segmentations of gross tumor volume, brainstem and hippocampus. The preliminary results demonstrate the feasibility of the method.

## 1. INTRODUCTION

When planning for radiotherapy of brain tumors, several structures need to be segmented from multi-modal magnetic resonance (MR) images. Important structures are healthy sub-cortical structures that should be spared from radiation (so-called organs-at-risk) and the MR-defined gross tumor volume (GTV). Automatic segmentation of tumor-affected brains is challenging since the location, shape and appearance of tumors, as well as the effect of tumors on surrounding healthy tissue, vary greatly across patients. Moreover, brain tumor images often exhibit significant intensity inhomogeneity as well as large intensity variations between patients, particularly when they are acquired with different scanners or at different imaging facilities.

A number of methods have been proposed for automatic brain tumor segmentation, including e.g., [2, 3] and other methods evaluated within the MICCAI brain tumor segmentation (BRATS) challenges [1]. However, these methods have typically been focused solely on the segmentation of tumors and are difficult to extend to also segment healthy organs-at-risk. In contrast, there are many successful methods for healthy brain segmentation which are capable of segmenting even small structures with a fair level of accuracy, such as [4, 5]. However, these methods typically have difficulties handling abnormal brain tissue such as tumors. A method capable of segmenting healthy brain structures in detail while at the same time being able to handle tumors or other types of abnormal tissue still remains an open problem. Some research with atlas-based methods for brain tumor segmentation have started to look in this direction, e.g., in [6] a semi-automatic method was presented to non-linearly register a healthy atlas to brains with tumors; and in [7] an automatic method was presented for simultaneous segmentation of tumors and sub-cortical healthy structures, but the accuracy was not validated on important organs-at-risk.

In this paper we propose a fully automated generative method to segment brain tumors and organs-at-risk simultaneously. The method is a further development of a previously presented method for brain tumor

---

Further author information: (Send correspondence to Mikael Agn)  
E-mail: miag@dtu.dk

segmentation that has been validated within the BRATS challenge on a publicly available data set of gliomas [8]. Similarly to this previous work, we use a spatial tumor prior consisting of convolutional restricted Boltzmann machines (cRBMs), which are a type of Markov random fields (MRFs) capable of modeling more complex interactions than traditionally used first-order MRFs. We learn the cRBM features automatically from expert segmentations without the use of the *intensity information* corresponding to these segmentations. Therefore, the model is able to adapt to varying intensity contrasts during the segmentation phase. In contrast to our previous method [8], here we combine the cRBM tumor shape prior with a whole-brain segmentation method for healthy brains [5]. The resulting probabilistic atlas-based method is capable of spatially adapting to individual brains, enabling us to obtain detailed segmentations of healthy structures, such as organs-at-risk in radiotherapy, in addition to delineations of the tumor itself.

As an initial evaluation, we test the proposed method on a manually delineated clinical data set of 20 glioblastoma patients. We report segmentation accuracy on GTV and two important organs at risk, brainstem and hippocampus, compared to the manual segmentations. These preliminary results demonstrate the feasibility of the method.

## 2. MODELING FRAMEWORK

Let  $\mathbf{D} = (\mathbf{d}_1, \dots, \mathbf{d}_I)$  denote the multi-contrast MR data of a subject after logarithmic transformation, where  $I$  is the number of voxels and  $\mathbf{d}_i \in \mathbb{R}^N$  contains the log-transformed intensities at voxel  $i$ , where  $N$  is the number of MR contrasts. We aim to assign one of the following labels to each voxel  $i$ : tumor (that is to say the MR-defined GTV), peritumoral edema or one of  $K$  healthy tissue labels. For this purpose we build a generative model that describes the image formation process and use this model to obtain a fully automated segmentation algorithm.

Let  $\mathbf{l} = (l_1, \dots, l_I)^T$  denote the healthy segmentation, where  $l_i \in \{1, \dots, K\}$ . We introduce a binary map  $\mathbf{y} = (y_1, \dots, y_I)^T$  to indicate voxels that are part of a tumor, and another binary map  $\mathbf{z} = (z_1, \dots, z_I)^T$  to indicate all tumor-affected voxels, i.e., voxels that are part of either tumor or peritumoral edema. This means that an edema voxel will be indicated by  $z_i = 1$  and  $y_i = 0$ , and a tumor voxel by  $z_i = 1$  and  $y_i = 1$ . Our generative model then consist of a likelihood function  $p(\mathbf{D}|\mathbf{l}, \mathbf{y}, \mathbf{z})$  that links labels to MR intensities, and a prior distribution on labels  $p(\mathbf{l}, \mathbf{y}, \mathbf{z}) = p(\mathbf{l})p(\mathbf{y}, \mathbf{z})$ , where  $p(\mathbf{l})$  encodes prior knowledge of healthy brain anatomy and  $p(\mathbf{y}, \mathbf{z})$  models the shape of tumor and edema.

For the likelihood  $p(\mathbf{D}|\mathbf{l}, \mathbf{y}, \mathbf{z})$ , we use Gaussian mixture models (GMMs) to model the relationship between tissue labels and MR intensities. Furthermore, we model bias fields that typically corrupt MR scans as linear combinations of spatially smooth basis functions added to the scans. Letting  $\boldsymbol{\theta}$  denote all GMM and bias field parameters with a prior  $p(\boldsymbol{\theta})$ , the resulting likelihood is given by

$$p(\mathbf{D}|\mathbf{l}, \mathbf{y}, \mathbf{z}) = \int_{\boldsymbol{\theta}} p(\mathbf{D}|\mathbf{l}, \mathbf{y}, \mathbf{z}, \boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta} \quad \text{with}$$

$$p(\mathbf{D}|\mathbf{l}, \mathbf{y}, \mathbf{z}, \boldsymbol{\theta}) = \prod_i \begin{cases} p_i(\mathbf{d}_i|\boldsymbol{\theta}_{l_i}, \mathbf{C}) & \text{if } z_i = 0 \text{ and } y_i = 0, \text{ (healthy)} \\ p_i(\mathbf{d}_i|\boldsymbol{\theta}_e, \mathbf{C}) & \text{if } z_i = 1 \text{ and } y_i = 0, \text{ (edema)} \\ p_i(\mathbf{d}_i|\boldsymbol{\theta}_t, \mathbf{C}) & \text{if } z_i = 1 \text{ and } y_i = 1, \text{ (tumor)} \\ (y_i = 1 \text{ and } z_i = 0 \text{ prohibited by prior, see Eq. (1)}) \end{cases}$$

Here  $\boldsymbol{\theta}_x$  denotes the parameters of the GMM connected to tissue  $x$  and  $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_N)$ , where  $\mathbf{c}_n$  denotes the parameters of the bias field model for MR contrast  $n$ . We define a Gaussian mixture model, with  $G_x$  Gaussian components, as  $p_i(\mathbf{d}_i|\boldsymbol{\theta}_x, \mathbf{C}) = \sum_{g=1}^{G_x} \gamma_{xg}\mathcal{N}(\mathbf{d}_i - \mathbf{C}^T \boldsymbol{\phi}^i | \boldsymbol{\mu}_{xg}, \boldsymbol{\Sigma}_{xg})$ , where subscript  $g$  denotes a Gaussian component within the Gaussian mixture model;  $\mathcal{N}(\cdot)$  denotes a normal distribution; and the parameters  $\gamma_{xg}$ ,  $\boldsymbol{\mu}_{xg}$  and  $\boldsymbol{\Sigma}_{xg}$  are the weight, mean and covariance of the corresponding Gaussian. Furthermore,  $\boldsymbol{\phi}^i$  evaluates the basis functions of the bias field model at voxel  $i$ . We restrict healthy structures with similar intensity properties to have the same GMM parameters. We use the following superstructures: non-brain tissues, cerebrospinal fluid, gray matter structures (including hippocampus) and white matter structures (including brainstem). For the prior  $p(\boldsymbol{\theta})$ , we follow the formulations in [8]; with uniform priors on  $\mathbf{C}$  and most mean vectors, conjugate

priors on weights and covariances, and a prior with a linear constraint on the mean vectors related to edema and healthy white matter structures.

For the healthy segmentation prior  $p(\mathbf{l})$ , we use a probabilistic atlas learned from manual annotations in 39 subjects as described in [5]. Among many other structures, the atlas includes spatial probability maps of brainstem and hippocampus. The atlas is parametrized by a sparse tetrahedral mesh with node positions  $\boldsymbol{\eta}$ . Assuming conditional independence of the labels between voxels given  $\boldsymbol{\eta}$ , the prior is given by

$$p(\mathbf{l}) = \int_{\boldsymbol{\eta}} p(\mathbf{l}|\boldsymbol{\eta})p(\boldsymbol{\eta})d\boldsymbol{\eta} \quad \text{with}$$

$$p(\mathbf{l}|\boldsymbol{\eta}) = \prod_{i=1}^I p_i(l_i|\boldsymbol{\eta}),$$

where  $p(\boldsymbol{\eta})$  is a topology-preserving deformation prior [9].

Finally, for  $p(\mathbf{y}, \mathbf{z})$  we use a cRBM model, defined below.

## 2.1 Spatial tumor prior using convolutional RBMs

We model the shape of tumor-affected tissue by restricted Boltzmann machines (RBMs), a type of MRFs that are capable of modeling higher-order interactions between voxels through local connections to hidden units. Connections directly between voxels or between hidden units do not exist, which facilitates inference with the model. To allow for more efficient inference over large images without a predefined size, we use *convolutional* RBMs (cRBMs), where the connection weights are shared among all locations [10]. As in [8], we combine two separate cRBMs to form the tumor prior: one that models interactions in the tumor-affected map  $\mathbf{z}$  and one that models interactions in the tumor map  $\mathbf{y}$ .

In particular, the cRBM that models  $\mathbf{z}$  is defined by  $p(\mathbf{z}) = \sum_{\mathbf{H}} p(\mathbf{z}, \mathbf{H})$ , with  $p(\mathbf{z}, \mathbf{H}) \propto e^{-E(\mathbf{z}, \mathbf{H})}$  and the energy term

$$E(\mathbf{z}, \mathbf{H}) = - \sum_k \mathbf{h}_k \bullet (\mathbf{w}_k * \mathbf{z}) - \sum_k b_k \sum_j h_{kj} - a \sum_i z_i,$$

where  $\bullet$  denotes element-wise product followed by summation and  $*$  denotes convolution. Here the model is defined in 1D to avoid cluttered equations; it is straightforward to extend it to 3D images. Each hidden group  $\mathbf{h}_k \in \mathbf{H}$  is connected to the visible units in  $\mathbf{z}$  with a convolutional filter  $\mathbf{w}_k$ , which models interactions between the hidden and visible units, effectively detecting specific features in  $\mathbf{z}$ . Furthermore, each hidden group has a bias  $b_k$  and visible units have a bias  $a$ , encouraging units to be enabled or disabled.

The two cRBMs, defined by  $p(\mathbf{z})$  with hidden units  $\mathbf{H}$  and  $p(\mathbf{y})$  with hidden units  $\mathbf{G}$ , are trained separately. During the training phase we learn the filters and bias terms from expert segmentations of tumor and edema obtained from training data. This is done by stochastic gradient ascent with the contrastive divergence (CD) approximation of the log-likelihood gradients with one block-Gibbs sampling step [11]. We use the so-called enhanced gradient together with the CD approximation to obtain more distinct filters [12, 13].

After the training phase we combine the two cRBMs to form the joint tumor shape prior:

$$p(\mathbf{y}, \mathbf{z}) = \sum_{\mathbf{G}, \mathbf{H}} p(\mathbf{y}, \mathbf{z}, \mathbf{G}, \mathbf{H})$$

$$\text{with } p(\mathbf{y}, \mathbf{z}, \mathbf{G}, \mathbf{H}) \propto e^{-E_{\text{comb}}(\mathbf{y}, \mathbf{z}, \mathbf{G}, \mathbf{H})}$$

$$\text{where } E_{\text{comb}}(\mathbf{y}, \mathbf{z}, \mathbf{G}, \mathbf{H}) = E(\mathbf{y}, \mathbf{G}) + E(\mathbf{z}, \mathbf{H}) + \sum_i f(y_i, z_i), \quad (1)$$

which models both the tumor and the surrounding edema simultaneously. Here,  $f(z_i, y_i) = \infty$  if  $z_i = 0$  and  $y_i = 1$ , and otherwise 0, restricting tumor to be within the tumor-affected region.

## 2.2 Inference

Exact inference of  $p(\mathbf{l}, \mathbf{y}, \mathbf{z}|\mathbf{D})$  requires an intractable integration over all possible combinations of model parameters. To side-step this difficulty, we use the approximation  $p(\mathbf{l}, \mathbf{y}, \mathbf{z}|\mathbf{D}) \simeq p(\mathbf{l}, \mathbf{y}, \mathbf{z}|\mathbf{D}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}})$ , where  $\{\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}}\}$  are the parameter values that maximize  $p(\boldsymbol{\theta}, \boldsymbol{\eta}|\mathbf{D})$ . We first use the same model parameter optimization algorithm as in [5] to maximize  $p(\boldsymbol{\theta}, \boldsymbol{\eta}|\mathbf{D})$ . After this, we sample from  $p(\mathbf{l}, \mathbf{y}, \mathbf{z}|\mathbf{D}, \hat{\boldsymbol{\eta}})$  by using Markov chain Monte Carlo sampling (MCMC) and subsequently use voxel-wise majority voting to obtain the final segmentation.

To maximize  $p(\boldsymbol{\theta}, \boldsymbol{\eta}|\mathbf{D})$  we can use coordinate ascent as in [5], where the atlas deformation parameters  $\boldsymbol{\eta}$  are optimized with a conjugate gradient algorithm, and the intensity model parameters  $\boldsymbol{\theta}$  with a generalized expectation-maximization (GEM) algorithm. The optimization is done by iteratively alternating between keeping the deformation parameters fixed while optimizing the intensity model parameters and vice versa until convergence. However, the cRBM model introduces non-local dependencies between the voxels. We therefore temporarily replace the combined cRBM's energy  $E_{\text{comb}}(\mathbf{z}, \mathbf{y}, \mathbf{H}, \mathbf{G})$  with a simple energy of the form

$$E_{\text{tmp}}(\mathbf{z}) = - \sum_{i=1}^I [l_i \neq NB] \left( z_i (y_i \log(wv) + (1 - y_i) \log(w(1 - v))) + (1 - z_i) \log(1 - w) \right),$$

where  $NB$  denotes non-brain tissues, and  $w$  and  $v$  are user-specified parameters that essentially define uniform spatial prior probabilities for tumor-affected tissue and tumor, respectively, to occur within the brain. This reduces the model to the same form as in [5], so we can use the same optimization method.

After the initial parameter estimation, we replace the temporary energy with the original cRBM energy and use MCMC to generate samples of  $\mathbf{l}$ ,  $\mathbf{y}$  and  $\mathbf{z}$  from  $p(\mathbf{l}, \mathbf{y}, \mathbf{z}|\mathbf{D}, \hat{\boldsymbol{\eta}})$ . Note that we keep the parameters  $\boldsymbol{\eta}$  fixed to  $\hat{\boldsymbol{\eta}}$  during sampling, as these atlas deformation parameters are difficult to sample from. In particular, we generate samples of  $\mathbf{l}$ ,  $\mathbf{y}$ ,  $\mathbf{z}$ ,  $\mathbf{G}$ ,  $\mathbf{H}$  and  $\boldsymbol{\theta}$  by block-Gibbs sampling from the distribution  $p(\mathbf{l}, \mathbf{y}, \mathbf{z}, \mathbf{G}, \mathbf{H}, \boldsymbol{\theta}|\mathbf{D}, \hat{\boldsymbol{\eta}})$ . After that, we discard the samples of  $\mathbf{G}$ ,  $\mathbf{H}$  and  $\boldsymbol{\theta}$  as they are of no interest to us, and perform voxel-wise majority voting across the collected samples of  $\mathbf{l}$ ,  $\mathbf{y}$  and  $\mathbf{z}$  to obtain the final segmentation. Block-Gibbs sampling is straightforward to implement as each of the conditional distributions  $p(\mathbf{l}, \mathbf{y}, \mathbf{z}|\mathbf{D}, \mathbf{G}, \mathbf{H}, \boldsymbol{\theta}, \hat{\boldsymbol{\eta}})$ ,  $p(\mathbf{G}|\mathbf{y})$ ,  $p(\mathbf{H}|\mathbf{z})$  and  $p(\boldsymbol{\theta}|\mathbf{D}, \mathbf{l}, \mathbf{y}, \mathbf{z}, \hat{\boldsymbol{\eta}})$  factorizes over its components. We initialize the MCMC sampler with a *maximum a posteriori* segmentation obtained by evaluating the posterior probability  $p(\mathbf{l}, \mathbf{y}, \mathbf{z}|\mathbf{D}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}})$  of the initial optimization method.

## 3. EXPERIMENTS

We demonstrate the performance of the proposed method on a clinical data set of 20 glioblastoma patients that have undergone radiotherapy treatment at Rigshospitalet at Copenhagen University Hospital. These patients have been scanned with a CT scanner and a Siemens Magnetom Espree 1.5T MRI scanner. The data set includes three MR-sequences: FLAIR, T2 and contrast-enhanced T1 (T1c); with a voxel size of  $(1 \times 1 \times 3)$ ,  $(1 \times 1 \times 3)$  and  $(0.5 \times 0.5 \times 1)$  mm respectively. The CT scans have a voxel size of  $(0.5 \times 0.5 \times 3)$  mm. As part of the treatment planning, the MR-defined GTV and several organs-at-risk (such as hippocampus, brainstem, eyeballs and optic nerve) have been manually delineated on the transversal slices of the CT scan of each patient, with the MR-sequences transformed and resampled to the CT scan.

In our method, we use the three MR-sequences to automatically segment the GTV, hippocampus and brainstem. Note that our method do not currently segment eyeballs and the optic nerve, as they are not part of the healthy atlas we are using. These structures will be included in the future. As the only pre-processing step before running the algorithm, we co-register the MR scans and resample them to 1 mm isotropic resolution. As a post-processing step, we transform and resample the resulting segmentations to the CT scan, before comparing with the manual segmentations.

To learn the parameters of the cRBM model, we used the expert segmentations of the training data for the MICCAI Brain Tumor Segmentation challenge of 2013, consisting of edema and tumor segmentations of 20 high-grade gliomas and 10 low-grade gliomas [1]. As this training data set is small, we augmented it by flipping the segmentations in 8 different directions.

### 3.1 Implementation

For the initial model parameter optimization method, we closely follow the implementation details of the whole-brain segmentation method described in [5]. For the sampling and the training of the cRBM model, we closely follow the implementation details in [8]. We used 40 filters of size  $(7 \times 7 \times 7)$  for each cRBM, corresponding to 40 hidden groups. Furthermore, to reduce the number of parameters to be estimated, we let each element in an cRBM filter model two neighboring elements in  $\mathbf{y}$  or  $\mathbf{z}$ , i.e., a filter of size 7 will span over 14 visible units.

For the GMMs connected to the different structures, we used the following number of Gaussians: three for non-brain tissues, two for cerebrospinal fluid, one for gray matter structures (which includes hippocampus), one for white matter structures (which includes brainstem), one for tumor and one for edema. The healthy GMM parameters were initialized based on the structure probabilities given by the atlas prior model after affine registration to the subject. We initialized the mean vectors of the tumor and edema Gaussians to the percentiles  $\{90,70,95\}$  and  $\{90,70,50\}$  respectively, in FLAIR, T2 and T1c. Before starting the MCMC sampler, we added two more Gaussian for tumor. We initialized the Gaussians by randomly setting  $y_i = 1$  to a fraction of the voxels with  $z_i = 1$  and  $y_i = 0$  in the segmentation used to initialize the sampler, as detailed in [8].

We implemented the algorithm in MATLAB 2015a, except for the atlas mesh deformation part which was implemented in C++. The full segmentation algorithm (including both the initial parameter estimation and the subsequent sampling) was performed on a Core i5-2400 3.1 GHz CPU, taking roughly 2 hours per subject. After a burn-in period of 100 samples, we generated 100 samples and obtained the final segmentation by majority voting on these 100 samples.

### 3.2 Results

Figure 1 shows four exemplary segmentations by the method. Although the method can segment many more structures [5], we focus here on the structures that we can validate against the manual segmentations and that are important for radiotherapy, i.e., brainstem, hippocampus and GTV. As can be seen, the atlas is able to deform well to fit the subjects. The method also captures GTV well although the segmentations are slightly underestimated compared to the manual segmentations. Furthermore, brainstem and hippocampus are captured well, although the hippocampus is slightly larger than in the manual segmentations.

Figure 2 shows a brainstem segmentation by the method from a sagittal view compared to the corresponding manual segmentation. The method captures brainstem well, but as can be seen from this view a part of the midbrain is missing in the method's segmentation. This is due to the different protocols used by the experts at the clinic and for the healthy subjects used to construct the atlas, where this part of the midbrain has been labeled as ventral diencephalon.

Figure 3 shows three hippocampus segmentation by the method from a sagittal view compared to the corresponding manual segmentations. Note that the manual segmentation appear blocky as they are delineated on the transversal slices of the CT scans. In the two first examples, the hippocampus is well captured by the method. In contrast, the manual segmentations are missing parts of the hippocampus. It is clear that the protocol used for the manual segmentations is quite conservative. In the last example however, the method slightly overestimate the hippocampus.

In figure 4, we show quantitative results evaluating the spatial overlap in the structures brainstem, hippocampus and GTV; for the 20 test subjects. In order to compare our method to the manual segmentations, we use Dice score  $D = \frac{TP}{(FN+FP+2 TP)/2}$ , true positive rate  $TPR = \frac{TP}{TP+FN}$  and positive prediction value  $PPV = \frac{TP}{TP+FP}$  as performance metrics. Here, TP, FP and FN count the true positive, false positive and false negative voxels compared to the manual segmentations. For all metrics, a perfect overlap would give a score of 1 and no overlap a score of 0. From left to right, the columns in each plot shows the scores for brainstem, hippocampus and GTV, respectively. The scores for brainstem are the most consistent, with a fairly high average Dice score. The trend of the TPR and PPV shows that brainstem is clearly underestimated compared to the manual segmentations, which is mainly due to the different protocols. Hippocampus, on the other hand, has a low Dice score as in general it is overestimated compared to the manual segmentations. This is mainly due to the conservative protocol of the manual segmenters, but the hippocampus is also sometimes overestimated by the method. It should be noted that hippocampus is a small structure, with an average volume which is 10 times less than the other two



structures in this study. This makes the scores more sensitive to segmentation differences. For most subjects, GTV gets fairly well segmented, but fails for a few subjects which drags down the average scores considerably.

Figure 5 shows two failed segmentation. In the first subject, the method has difficulties to include a resection with a similar intensity distribution as cerebrospinal fluid. This issue affected the segmentations of three subjects in the data set. In the second subject in the figure, the method captures an old infarct in another region of the brain and adapts to its intensity distribution, which has the unfortunate effect that the enhanced ring of the tumor gets labeled as peritumoral edema. In one other subject, the method segmented the GTV fairly well but also segmented a tumor that was not to be included for radiotherapy.

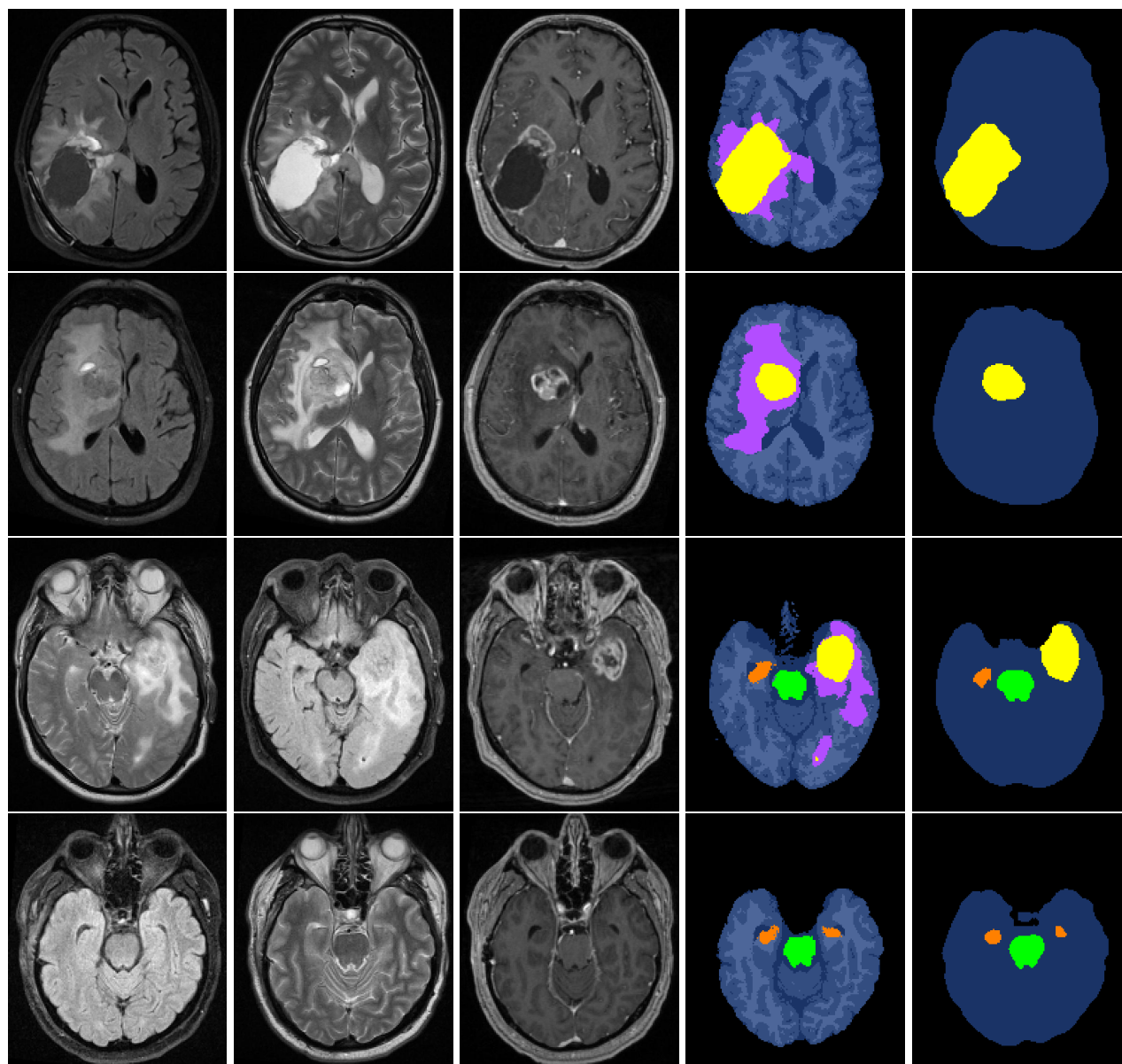


Figure 1. Slices of four exemplary subjects. From left to right: MR-contrasts: FLAIR, T2 and T1c; segmentation by method; and manual segmentation. GTV is in yellow, brainstem is in green, hippocampus is in orange, edema is in lilac, non-brain tissue is in black and healthy superstructures are in shades of blue.

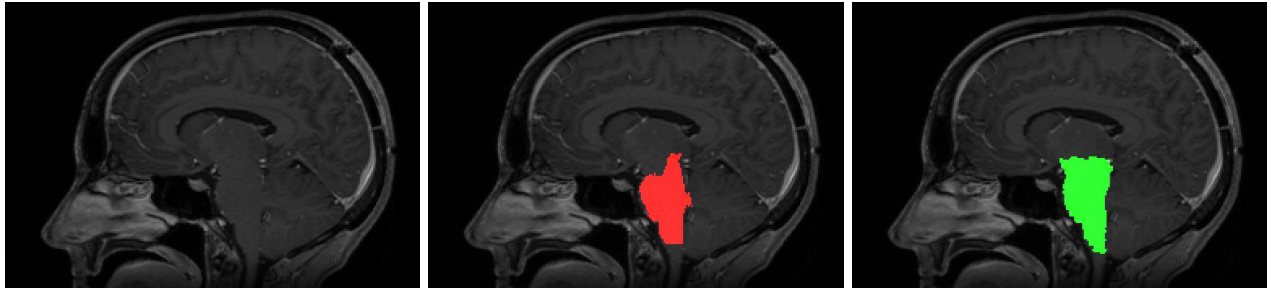


Figure 2. Sagittal view of brainstem. From left to right: T1c, segmentation by the proposed method, and manual segmentation.

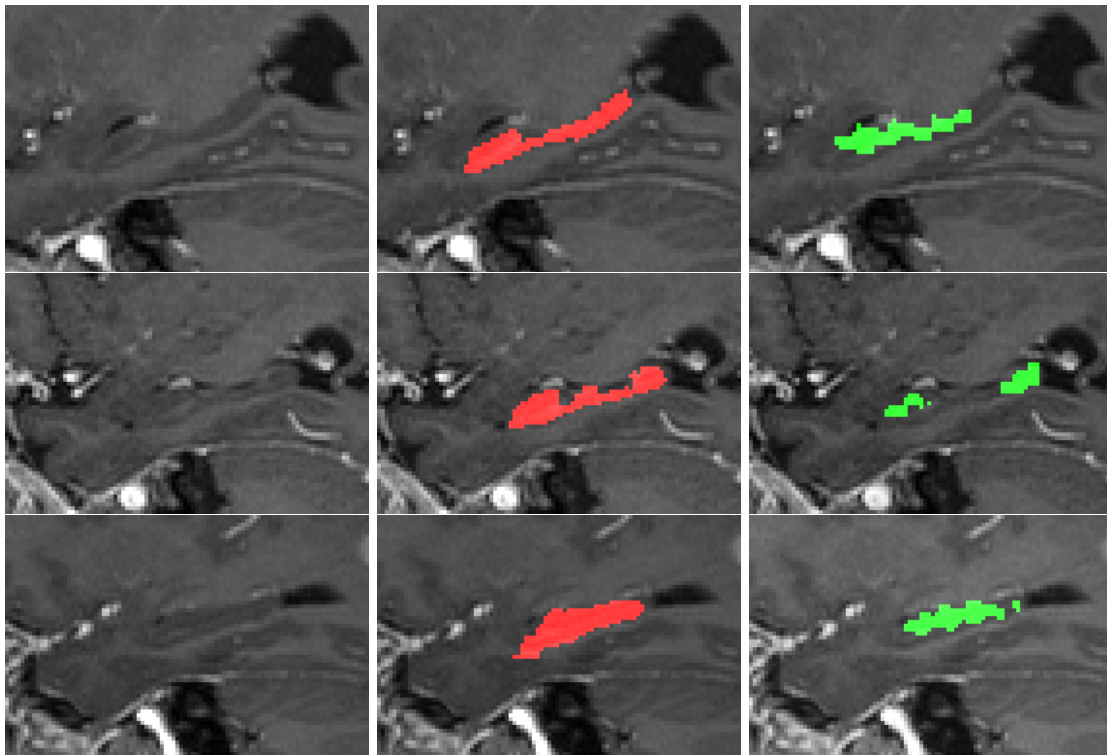


Figure 3. Sagittal view of hippocampus in three subjects. From left to right: T1c, segmentation by the proposed method, and manual segmentation.



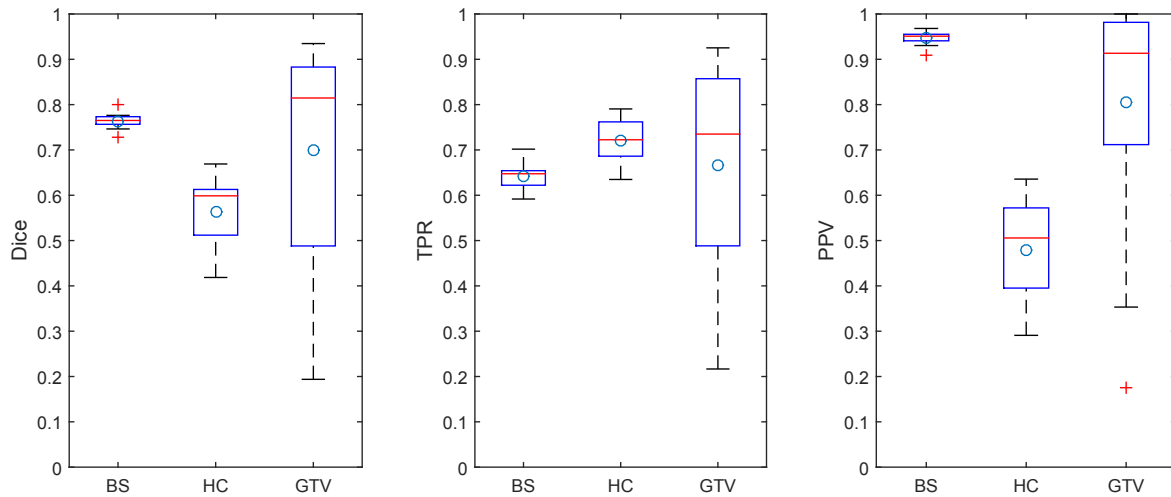


Figure 4. Box plots showing Dice scores, TPR and PPV of brainstem (BS), hippocampus (HC) and GTV. Circles show mean values, central lines show medians, edges of boxes show the 25th and 75th percentiles, and outliers are marked with '+'. '.

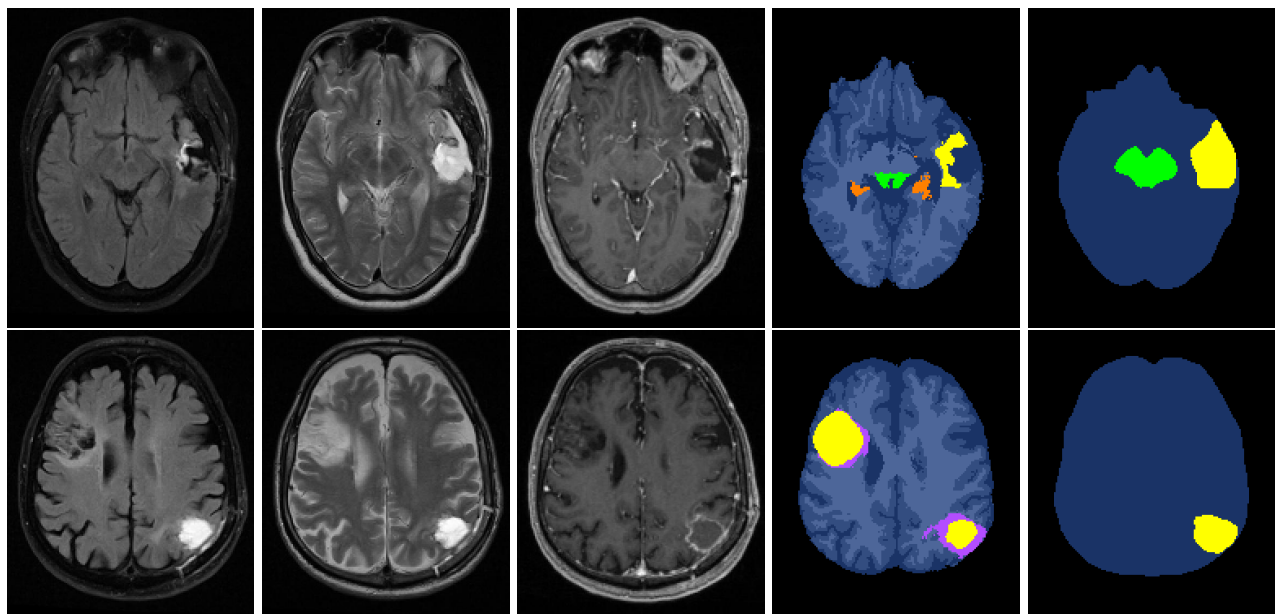


Figure 5. Slices of two failed subject. From left to right: MR-contrasts: FLAIR, T2 and T1c; segmentation by method; and manual segmentation. GTV is in yellow, brainstem is in green, hippocampus is in orange, edema is in lilac, non-brain tissue is in black and healthy superstructures are in shades of blue.

#### 4. DISCUSSION

In this paper, we have proposed a fully automated generative method for simultaneous brain tumor and organs-at-risk segmentation for radiotherapy planning. The method combines an existing whole-brain segmentation method with a tumor prior that uses convolutional restricted Boltzmann machines to model tumor shape. It is not tied to any specific imaging protocol as the parameters of the tumor model are estimated only from expert *segmentations* of annotated training images, without using intensity information. To initially test the feasibility of the method, we have evaluated its performance on a manually delineated clinical data set of 20

glioblastoma patients. We have reported the segmentation accuracy on gross tumor volume and two organs at risk in radiotherapy, i.e., brainstem and hippocampus; with varying but promising results. The evaluation revealed some misalignment between the protocols used by the segmentation experts of the test data set and the protocols used when constructing the healthy atlas. Future work will involve further experiments to improve the tumor prior and use more informed priors on the GMM parameters; inclusion of other important organs at risk in the atlas, such as eyeballs and the optic nerve; and validation on a larger data set.

## ACKNOWLEDGMENTS

This research was supported by NIH NCRR (P41-RR14075), NIBIB (R01EB013565) and the Lundbeck Foundation (R141-2013-13117).

## REFERENCES

- [1] Menze, B. H., et al.: The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging* 34(10), 1993-2024 (2015)
- [2] Tustison, N., Wintermark, M., Durst, C., and Avants, B.: ANTs and Árboles. In: *Proc MICCAI-BRATS 2013* (2013)
- [3] Havaei, M., et al.: Brain Tumor Segmentation with Deep Neural Networks. *arXiv preprint arXiv:1505.03540* (2015)
- [4] Fischl, B.: FreeSurfer. *Neuroimage*, 62(2), 774-781 (2012)
- [5] Puonti, O., Iglesias, J. E., and Van Leemput, K.: Fast, Sequence Adaptive Parcellation of Brain MR Using Parametric Models. In: *MICCAI 2013. Lecture Notes in Computer Science*, vol. 8149, 727-734 (2013)
- [6] Kwon, D., Shinohara, R. T., Akbari, H., and Davatzikos, C.: Combining generative models for multifocal glioma segmentation and registration. In: *MICCAI 2014. Lecture Notes in Computer Science*, vol. 8673, 763-770 (2014)
- [7] Bauer, S., Lu, H., May, C. P., Nolte, L. P., Büchler, P., and Reyes, M.: Integrated segmentation of brain tumor images for radiotherapy and neurosurgery. *International Journal of Imaging Systems and Technology*, 23(1), 59-63 (2013)
- [8] Agn, M., Puonti, O., Law, I., Munck af Rosenschöld, P., and Van Leemput, K.: Brain Tumor Segmentation Using a Generative Model with an RBM Prior on Tumor Shape. To appear in: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - Brainles workshop at MICCAI 2015, Revised Selected Papers. Lecture Notes in Computer Science*, vol. 9556 (2016) (pre-print: <http://people.compute.dtu.dk/miag/brainles-agn.pdf>)
- [9] Van Leemput, K.: Encoding Probabilistic Brain Atlases Using Bayesian Inference. *IEEE Transactions on Medical Imaging* 28(6), 822-837 (2009)
- [10] Lee, H., Grosse, R., Ranganath, R., and Ng, A. Y.: Unsupervised learning of hierarchical representations with convolutional deep belief networks. *Communications of the ACM* 54(10), 95-103 (2011)
- [11] Fischer, A., and Igel, C.: Training restricted Boltzmann machines: An introduction. *Pattern Recognition* 47(1), 25-39 (2014)
- [12] Melchior, J., Fischer, A., and Wiskott, L.: How to Center Binary Restricted Boltzmann Machines. *arXiv preprint arXiv:1311.1354* (2013)
- [13] Cho, K., Raiko, T., and Ilin, A.: Enhanced gradient for training restricted Boltzmann machines. *Neural computation* 25(3), 805-831 (2013)