Enhanced Graph Cuts for Brain Tumor Segmentation Using Bayesian Optimization

Mauricio Castaño[1(✉)], Hernán F. García[1,2,3(✉)], Gloria L. Porras-Hurtado[2(✉)], Álvaro A. Orozco[3(✉)], and Jorge I. Marin-Hurtado[4(✉)]

[1] GAMMA Research Group, Universidad del Quindío, Armenia, Colombia
mcastanoa@uqvirtual.edu.co
[2] Grupo de Investigación Salud Comfamiliar, Comfamiliar Risaralda,
Pereira, Colombia
gporras@comfamiliar.com
[3] Grupo de Investigación en Automática, Universidad Tecnológica de Pereira,
Pereira, Colombia
{hernan.garcia,aaog}@utp.edu.co
[4] GDSPROC Research Group, Universidad del Quindío,
Armenia, Colombia
jorgemarin@uniquindio.edu.co

Abstract. Brain tumor segmentation is a difficult task, due to the shape variability that malignancy brain structures exhibit between patients. The main problem in this process is that the tumor contour is usually computed from parametric models that need to be well-tuned to perform an accurate segmentation. In this paper, we propose an enhanced Graph cut on which the model parameters are selected through a probabilistic approach. Here, we use Bayesian optimization to find the optimal hyper-parameters that segment the tumor volume accurately. The experimental results show that by using Bayesian optimization, the graph cut model performs an accurate segmentation over brain volumes in comparison with common segmentation methods in the state-of-the-art.

Keywords: Bayesian optimization · Graph cuts · Gaussian processes
Brain tumor segmentation

1 Introduction

Brain malignancies are the most dangerous pathologies in neurological diseases. These malignancies present different degrees of aggressiveness, different prognosis and heterogeneous histological sub-regions (i.e., peritumoral edema, necrotic core, enhancing and non-enhancing tumor core). This variability (due to intrinsic heterogeneities of gliomas) poses a challenging task in which the imaging phenotype is described by varying appearance and shape profiles across multimodal MRI scans, reflecting varying tumor tissue properties. This variability

poses a challenging task in which the imaging phenotype is described by varying appearance and shape biological descriptors across neurological scans, reflecting varying tumor tissue properties [12].

Localizing tumor areas is a crucial procedure for brain surgery planing. One of the main problems in this case is in fact the time, in practice, radiation oncologists spend a substantial portion of their time performing the segmentation manually using segmentation and visualization tools. Besides, in the literature survey, several techniques are proposed to overcome the challenges of brain tumor segmentation. Specifically, brain tumor segmentation algorithms based on convolutional neural networks (CNNs) have been shown to be at least as effective as other automated tumor segmentation methods [8].

In recent years, the researchers focused on exploring the entire field related to neural networks, a fully automatic pipeline that involves chaining together several unique 3D U-Net, a type of 3D patch-based convolutional neural network [1]. In general, state of the art focuses on models that initiate a process of forced learning given by iterations that cause specific weights to vary until an acceptable result is reached (i.e., to find a plausible tumor contour) [12]. The main problem of these approaches is that although the results are relevant, it is difficult to extract clinical information from these learning processes (i.e., to capture tissue properties from hidden layers as in CNNs) [9].

Patch-based methods make use of energy functions to define a given contour that matches plausible shape structures (i.e., tumor contour) [10]. These energy functions allow us to define the tumor contour from appearance and shape constraints based on the tumor properties [3]. However, the resulting performance of these approaches depends on the correct selection of the model parameters (i.e., graph cuts (GC) for image segmentation) [6]. Global optimization is an essential task in any complex problem where design and choice of model parameters play a key role. In the machine learning field, such problems are found in the tuning of hyperparameters [15] and experimental design [7].

Bayesian optimization (BO) [7,15], proves to outperform state of the art for global optimization algorithms on many challenging optimization benchmark functions [11]. In this context, Bayesian optimization assumes that the objective function is sampled from a Gaussian process, maintaining the posterior distribution for this function as observations (by running learning algorithm experiments with different hyperparameters are observed). In this paper, we propose an enhanced Graph cut on which the model parameters are selected through a probabilistic approach. Here, we use Bayesian optimization to find the optimal hyperparameters that segment the tumor volume accurately. Our contribution is based on the Bayesian optimization process that finds the model parameters for controlling the energy function of the graph cut in a probabilistic way. The rest of the paper proceeds as follows. Section 2 provides a detailed discussion of materials and methods. Section 3 presents the experimental results and some discussions about the proposed method. The paper concludes in Sect. 4, with a summary and some ideas for future research.

2   Materials and Methods

2.1   Database

In this work we used the Brain Tumor Image Segmentation Challenge (Brats) 2015 [12]. This Database contains high-grade tumors, Low-grade tumors and labels maps made by experts based on landmarks. The tumors of this database are located in different brain regions. The label map showed in Fig. 1, have four different labels 1- for Necrosis (Green), 2- for the Edema (Yellow), 3- for Non-enhancing tumor (Red) and 4- for Enhancing tumor (Blue). We used the MRI T1 images with resolution of 240×240 pixels and $1\,\text{mm} \times 1\,\text{mm} \times 1\,\text{mm}$ voxel size.

(a) Axial View   (b) Sagital view   (c) Coronal view
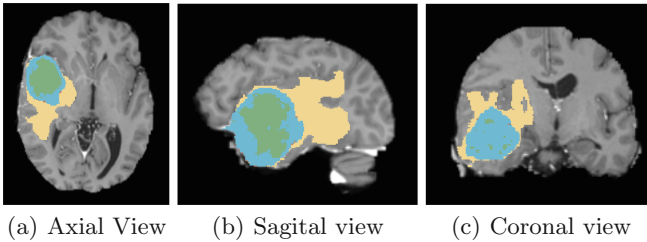
Fig. 1. Sample of a given abnormal tissue for the Brast2015 Database. (Color figure online)

2.2   Graph Cuts

We use a fast approximate energy minimization approach with label costs, that uses the *alpha*-expansion algorithm [5]. This algorithm is commonly used to minimize energies that involve unary, pairwise, and specialized higher-order terms that describes given appearance contour [3]. The segmented image can be modeled as an energy minimization that finds a labeling f (i.e., estimated contour) as,

$$E(f) = E_{\text{smooth}}(f) + E_{\text{data}}(f) + E_{\text{label}}(f), \tag{1}$$

where, $E_{\text{smooth}}(f)$ is a measure of the smoothness by parts of the labeling f, and $E_{\text{data}}(f)$ measures the discrepancy between f and the observed data. As in [5], the term $E_{\text{data}}(f)$ is computed as,

$$E_{data}(f) = \sum_{p \epsilon P} D_p(f_p), \tag{2}$$

where D_p measures how well the label f_p fits the pixel p. Generally, this is evaluated using a quadratic standard, which can be given by $(f_p - i_p)^2$ where

i_p is the original intensity of the pixel. The smoothness cost $E_{\text{smooth}}(f)$, is a standard regularizer which can be modeled as,

$$E_{\text{smooth}}(f) = \sum_{p,q \epsilon \mathcal{N}} V_{p,q}(f_p, f_q), \tag{3}$$

where each $V_{p,q}$ weights all $f_p \neq f_q$. A simple use of this function can be given by $V_{p,q}(f_p, f_q) = K \cdot |f_p - f_q|$ (with K being an arbitrary constant). Hence, if each $V_{p,q}$ define a metric, then the minimization of the Eq. (1) it is known as the problem of metric labeling and can be effectively optimized with the *alpha-expansion* algorithm [5]. The label cost penalize each unique label that appears in f as $E_{\text{label}}(f) = \sum_{L \subseteq \mathcal{L}} h_L \cdot \delta_L(f)$, where h_L is the non-negative label cost of labels L and the indicator function $\delta_L(.)$ is defined on a label subset L as,

$$\delta_L(f) \stackrel{\text{def}}{=} \begin{cases} 1 \ \exists p : f_p \in L \\ 0 \ otherwise \end{cases} \tag{4}$$

2.3   Bayesian Optimization with Gaussian Process Priors

Since we want to compute the graph cuts hyperparameters in a probabilistic way, our goal is to find the minimum of a cost function $f(\mathbf{x})$ (i.e., the performance index between the ground truth labels and the segmented tumor) on some bounded set $\mathcal{X}$ that controls the model parameters. To this end, Bayesian optimization builds a probabilistic framework for $f(\mathbf{x})$ with the aim to exploit this model to make predictions of the model parameters $\mathcal{X}$ evaluated in the cost function [15]. The main components of the Bayesian optimization framework are the prior of the function to optimize, as well as the acquisition function that will allow us to determine the next point to evaluate the cost function [13]. In this work, we use a Gaussian process prior, due to its flexibility and tractability. A Gaussian Process (GP) is an infinite collection of scalar random variables indexed by an input space such that for any finite set of inputs $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$, the random variables $\mathbf{f} \stackrel{\triangle}{=} [f(\mathbf{x}_1), f(\mathbf{x}_2), \cdots, f(\mathbf{x}_n)]$ are distributed according to a multivariate Gaussian distribution $\mathbf{f}(\mathbf{X}) = \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$. A GP is completely specified by a mean function $m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})]$ (usually defined as the zero function) and a positive definite covariance function given by $k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\left[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}')^T)\right]$ (see [15] for further details).

Let us assume that $f(\mathbf{x})$ is drawn from a Gaussian process prior and that our observations are set as $\{\mathbf{x}_n, y_n\}_{n=1}^{N}$, where $y_n \sim \mathcal{N}(f(\mathbf{x}_n), \nu)$ and ν is the noise variance. The acquisition function is denoted by $a : \mathcal{X} \rightarrow \mathbb{R}^+$ and establishes the point in $\mathcal{X}$ that is evaluated in the optimization process as $x_{\text{next}} = \arg\max_{\mathbf{x}} a(\mathbf{x})$. Since the acquisition function depends on the GP hyperparameters, θ, and the predictive mean function $\mu(\mathbf{x}; \{\mathbf{x}_n, \mathbf{y}_n\}, \theta)$ (as well as the predictive variance function), the best current value is then $x_{\text{best}} = \arg\min_{\mathbf{x}_n} f(\mathbf{x}_n)$.

2.4   Enhanced Graph Cuts with Bayesian Optimization

Our approach is based on the Bayesian optimization process for estimating the model parameters of the graph cut model that segments a given brain tumor accurately in a probabilistic way. In this work, we choose to optimize the foreground seed, $\Omega_1 = \{x_f, y_f, z_f\}$, the background seed $\Omega_2 = \{x_b, y_b, z_b\}$, and the α-parameter of the swap algorithm $\Omega_3 = \alpha$ [5]. For the graph cuts implementation we use the *imcut*[1] toolbox. Besides, as for the Bayesian optimization process, we use as a cost function, the Euclidean distance between the labels of the segmented tumor and the ground truth labels. We use the *GPyOpt*[2] toolbox for python, developed by the Machine Learning group of the University of Sheffield. In this work, we report results for the expected improvement (EI), and the probability of improvement (PI) and some other relevant acquisition functions [15]. Figure 2 shows the block diagram of the proposed model used in this work.
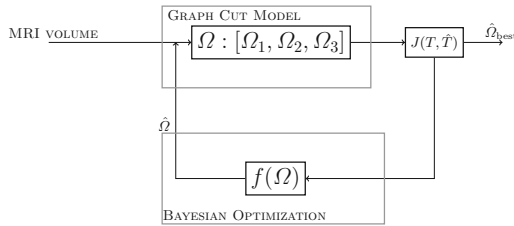
Fig. 2. Block diagram of the proposed approach for the enhanced graph cuts with Bayesian optimization

3   Results and Discussions

In this section, we show the results of our framework for optimizing the graph cuts hyperparameters. We show a comparison between a given manual tuning and an automatic tuning using Bayesian optimization. Besides, we report a comparison of the different segmentation performances of the acquisition functions, as well as some qualitative and quantitative results compared with relevant works in the state-of-art.

---

[1]   *imcut* is a Segmentation tool based on the graph cut algorithm available at https://github.com/mjirik/imcut.

[2]   *Gpyopt* is a Bayesian optimization framework in python available at http://github.com/SheffieldML/GPyOpt.
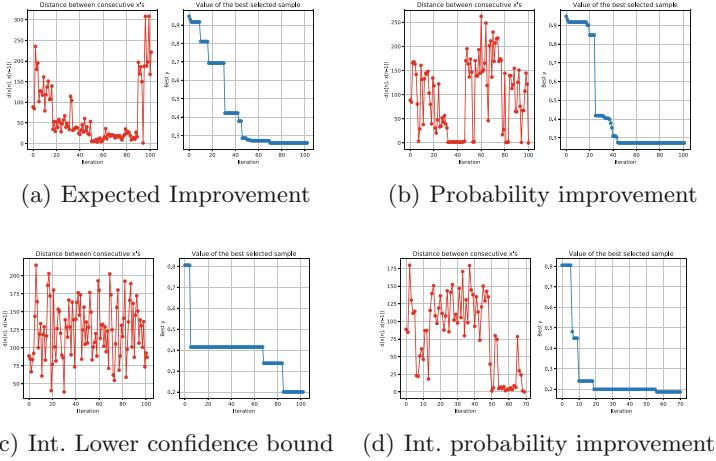
As we can see in Fig. 3, Bayesian optimization can eliminate certain inconveniences that arise for manual tunning of the graph cuts. The figure shows that the optimal parameters allow us to segment the tumor contour more appropriately (i.e., avoiding false negatives derived from the segmentation process).


(a) GC   (b) GC-BO   (c) GT
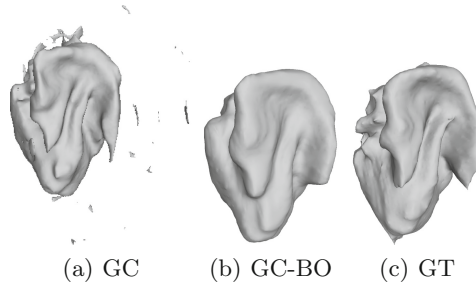
Fig. 3. Comparison between Bayesian optimization (middle) and manual adjustment, of graph cuts hyperparameters (left). The figure shows that the manual tuning of the graph cut model (left) derived in more false negatives in comparison with ground truth tumor (GT right).

Figure 4 shows the convergence of each acquisition function of the BO process. The red plots show the distance between the hyperparameters on each iteration. As a result, we can differentiate the stages of exploration and exploitation of the hyperparameters. Here, the more variation found between each consecutive hyper-parameter indicates the stage of exploration and small distances between consecutive hyperparameters indicates the stage of exploitation. The figures outlined in blue indicate the error convergence of each method. Furthermore, the results also show that the acquisition functions: integrated lower confidence bound and integrated probability improvement perform the tumor segmentation more accurately (see bottom row of Fig. 4).

Figure 5 shows the curvature computed for three different tumors: ground truth volumes (left) and segmented tumors with BO (right). The results show that regions with high saliency (red areas in the tumor volumes), matches the segmented tumors with the optimal hyperparameters (the segmentation process preserves the relative curvature of the original tumor). Finally, Table 1 shows a comparison of different approaches reported in the state-of-the-art (BraTS 2017 challenge [12]). The results show that our approach outperforms some relevant methods in the state-of-art, which are based on deep learning approaches. Hence, since our approach performs the segmentation in an unsupervised manner, the probabilistic tuning of the model parameters sets an important result for these image segmentation approaches.

(a) Expected Improvement

(b) Probability improvement

(c) Int. Lower confidence bound

(d) Int. probability improvement

Fig. 4. Convergence of the Bayesian optimization process for different acquisition functions. The figure shows the distance between values of x selected consecutively (red plots), and the minimum value of the performance index obtained in each iteration (blue plots). (Color figure online)

(a) Ground truth

(b) GC-BO
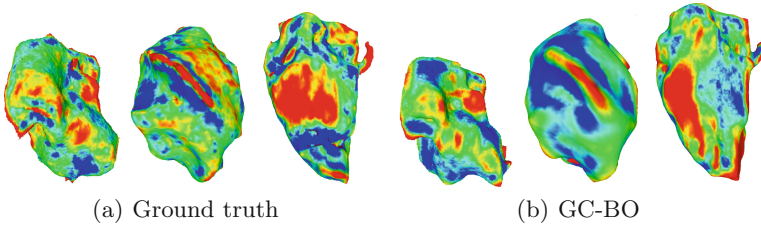
Fig. 5. Curvature comparison of three different subjects (left to right). The bottom row shows the curvature obtained from the segmented tumors with GC-BO. The figure shows the curvature obtained from the ground truth volumes. The figure shows that regions with high saliency (i.e., red colors) match the ground truth data. (Color figure online)

Table 1. Comparison with different methods proposed in [12] for brain tumor segmentation. We report the results as in the BRATS challenge: Dice coefficient, Hausdorff distance, sensitivity, and specificity (Some articles do not report some metrics, so they are assigned to N.R).

Method	Dice	Hausdorff	Sensitivity	Specificity
3D U-Net [12]	0.9111	19.8741	N.R	N.R
Tensor based feature extraction [2]	0.833	N.R	0.815	N.R
Volumetric Multi-modality Neural Network [4]	0.87	21.32	0.87	0.99
Label Distribution Learning [14]	0.86771	N.R	0.88816	0.99208
3D-SegNet architecture [12]	0.79767	23.64621	0.91027	0.97916
Our approach	0.88664	33.73749	0.88765	0.99849

4   Conclusions and Future Work

In this paper, we propose a Bayesian optimization framework for tuning the model parameters of a graph cuts method. Our method seeks to find the best parameters that segment a given tumor contour in a probabilistic way. The experimental results show that our approach derives in more accurate contours than a given classical procedure in image segmentation with graph cuts. Besides, since the model parameters are optimized, the resulting curvature of the segmented tumor preserves main saliency regions that match the ground truth data. Finally, our approach outperforms some important methods in the state-of-art that use deep learning frameworks.

As for the future works, we plan to extend the classical Energy minimization problem of the region growing approaches to propose a new function that can be optimized with probability black-box functions.

References

1. Beers, A., et al.: Sequential 3D U-Nets for biologically-informed brain tumor segmentation. CoRR abs/1709.02967 (2017)
2. Bharath, H.N., Colleman, S., Sima, D.M., Van Huffel, S.: Tumor segmentation from multi-parametric MRI using random forest with superpixel and tensor based feature extraction. In: Brats Challenge 2017 (2017)
3. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient N-D image segmentation. Int. J. Comput. Vis. **70**, 109–131 (2006)
4. Castillo, L.S., Daza, L.A., Rivera, L.C., Arbeláez, P.: Volumetric multimodality neural network for brain tumor segmentation (2017)
5. Delong, A., Osokin, A., Isack, H.N., Boykov, Y.: Fast approximate energy minimization with label costs. Int. J. Comput. Vis. **96**(1), 1–27 (2012)
6. Dogra, J., Jain, S., Sood, M.: Segmentation of MR images using hybrid kmean-graph cut technique. Proc. Comput. Sci. **132**, 775–784 (2018). International Conference on Computational Intelligence and Data Science
7. Gonzalez, J., Longworth, J., James, D., Lawrence, N.: Bayesian optimisation for synthetic gene design. In: NIPS Workshop on Bayesian Optimization in Academia and Industry (2014)
8. Havaei, M., et al.: Brain tumor segmentation with deep neural networks. Med. Image Anal. **35**, 18–31 (2017)
9. Isin, A., Direkoglu, C., Sah, M.: Review of MRI-based brain tumor image segmentation using deep learning methods. Proc. Comput. Sci. **102**, 317–324 (2016). 12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2016, 29–30 August 2016, Vienna, Austria
10. Jiřík, M., Lukes, V., Svobodova, M., Železný, M.: Image segmentation in medical imaging via graph-cuts (2013)

11. Jones, D.R.: A taxonomy of global optimization methods based on response surfaces. J. Glob. Optim. **21**(4), 345–383 (2001)
12. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., et al.: The multimodal brain tumor image segmentation benchmark (brats). IEEE Trans. Med. Imaging **34**(10), 1993–2024 (2015)
13. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning). The MIT Press (2005)
14. Shengcong Chen, C.D., Zhou, C.: Brain tumor segmentation with label distribution learning and multi-level feature representation. In: Brats Challenge 2017 (2017)
15. Snoek, J., Larochelle, H., Adams, R.P.: Practical Bayesian optimization of machine learning algorithms. In: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems, vol. 25, pp. 2951–2959. Curran Associates, Inc. (2012)