



# Brain Tumor Segmentation and Survival Prediction Using a Cascade of Random Forests

Szidónia Lefkovits<sup>1</sup> , László Szilágyi<sup>2</sup> , and László Lefkovits<sup>2</sup> 

<sup>1</sup> Department of Computer Science,  
University of Medicine, Pharmacy, Sciences and Technology, Tîrgu Mureş, Romania  
[szidonia.lefkovits@science.upm.ro](mailto:szidonia.lefkovits@science.upm.ro)

<sup>2</sup> Department of Electrical Engineering,  
Sapientia Hungarian University of Transylvania, Tîrgu Mureş, Romania  
[lefkolaci@ms.sapientia.ro](mailto:lefkolaci@ms.sapientia.ro)

**Abstract.** Brain tumor segmentation is a difficult task due to the strongly varying intensity and shape of gliomas. In this paper we propose a multi-stage discriminative framework for brain tumor segmentation based on BraTS 2018 dataset. The framework presented in this paper is a more complex segmentation system than our previous work presented at BraTS 2016. Here we propose a multi-stage discriminative segmentation model, where every stage is a binary classifier based on the random forest algorithm. Our multi-stage system attempts to follow the layered structure of tumor tissues provided in the annotation protocol. In each segmentation stage we dealt with four major difficulties: feature selection, determination of training database used, optimization of classifier performances and image post-processing. The framework was tested on the evaluation images from BraTS 2018. One of the most important results is the determination of the tumor ROI with a sensitivity of approximately 0.99 in stage I by considering only 16% of the brain in the subsequent stages. Based on the segmentation obtained we solved the survival prediction task using a random forest regressor. The results obtained are comparable to the best ones presented in previous BraTS Challenges.

**Keywords:** Multi-stage classifier · Random forest · Feature selection · Variable importance · MRI brain tumor segmentation

## 1 Introduction

Image processing is a powerful tool for computer-aided diagnosis especially in the medical field. The most important advantage of medical imaging is the fact that

---

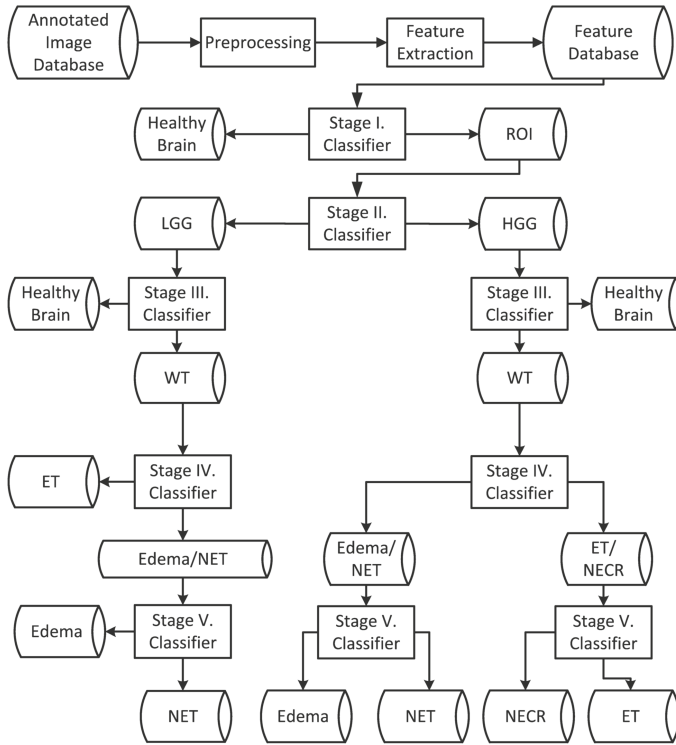
The work of S. Lefkovits in this article was supported by the Communitas Foundation and the work of L. Lefkovits in this article was supported by Sapientia Foundation-Institute for Scientific Research (KPI).

examination performed non-intrusively. MR imaging and diagnosis is increasingly being used for medical investigation. This article is restricted to MRI brain imaging and provides a framework for the automated brain tumor segmentation method proposed, delimiting different types of tumors in multi-modal MRI images. Automatic systems based on machine learning overcomes the laborious, lengthy work of segmentation done manually by experts. It is replicable and much faster than the segmentation performed by experts, which might be fairly different. The most important advantage of such a system is that it can lend assistance in determining the correct diagnosis, surgery or treatment plan and monitor the evolution of the disease. A comprehensive study of all the participating methods at BraTS 2018 Challenges is described in paper [6].

The framework presented in this paper is an extended segmentation system of greater complexity based on our model presented at MICCAI-BraTS 2016 [15]. This model was built on a feature extraction algorithm [14] and single-staged random forest (RF) [11] classifiers with optimized parameters. The random forest approach was used in few systems presented at BraTS 2017 [10,17–19]. The segmentation results obtained showed that the tumor region is well detected, but the contours of the whole tumor and the interior tumor tissues are not well delimited. The source of the aforementioned errors could be the choice of training samples used, the unbalanced database provided, and its enormous size. These three factors cannot be counteracted by a single-stage RF classifier. Another deficiency in our previous model is that it considered almost any spatial relationship between the tumor tissues, according to the annotation protocol described in [12,16].

In the current work we propose a multi-stage classifier based on the random forest algorithm. In our current experiments we attempt to circumvent the deficiencies of our old framework and improve the segmentation results.

The proposed framework is built around the model given in Fig. 1. The stage I classifier detects the tumorous zone from the entire 3D MRI image. This phase is tuned to have extremely good sensitivity. It considers the tumor zone to be the goal of detection, and therefore this is the positive segmentation zone. Thus, it is able to delimit the image ROI containing the tumor with a sensitivity of approximately 0.99, which is only around 16% of the entire brain. This means that our ROI considered in the subsequent stages is a highly reduced region. In stage II we developed a classifier that is able to separate the images in the two types given, i.e., LGG and HGG. Consequently, at this point our method is split into two structurally similar branches, because the classifiers are trained differently only on LGG or HGG images. In stage III the WT (whole tumor) classifier delimits the whole tumor from healthy brain tissue. In stage IV the role of the classifier is the determination of the ET (enhancing tumor) region. In the case of HGG, this region is a considerable part of the tumor and includes necrotic tissue regions too. The necrotic tissue inside ET is labelled with the same class number as the non-enhancing tumor. These two tissue types had different annotations at BraTS 2016. In the case of LGG, the segmentation of ET is more difficult because of its small size, and it can also be confused with



**Fig. 1.** Discriminative model proposed for segmentation (Each stage-classifier is binary as described in Subject. 2.1.)

other tissues such as vessels. Stage V tries to delimit the edema from the non-enhancing tumor. Because of the similar visual aspects of the two tissues, this segmentation step is error-sensitive.

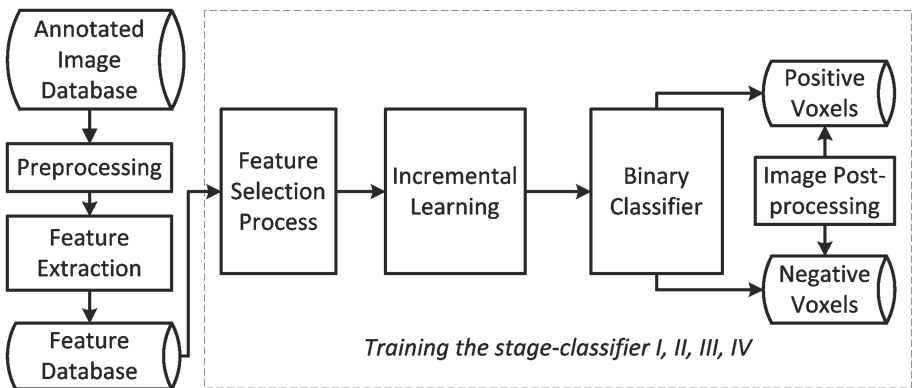
The use of binary classifiers for all these classification decisions follows from the annotation protocol. It states that “the various tissue elements (edema, non-enhancing, enhancing, necrosis) usually follow an outside-inside sequence therefore one should start from the outside and delineate regions within the previous layer. Due to this «Mozart kugel» appearance it is enough to always delimitate what is outside and internal border should not be delimitated [12]”.

The rest of the paper is organized as follows: in Sect. 2 the proposed cascaded binary classification model is described, followed by Sect. 3 presenting the validation and test results both for the segmentation and survival tasks. Finally, conclusions are drawn and discussion and further improvements are proposed.

## 2 Method

### 2.1 Segmentation

The delimitation of the brain tumor from the healthy tissues can be achieved by a voxel-wise segmentation. To solve this task we propose a multi-stage discriminative model based mainly on the random forest algorithm and its facilities. Voxel-wise segmentation starts with the construction of the feature database obtained from the annotated image database. The feature database generation process is identical both for the segmentation (classification) and the training phases, as well. It consists of the following steps: preprocessing, local feature definition and extraction (Figs. 1 and 2).



**Fig. 2.** Discriminative model proposed for training each stage

The database used in our segmentation made up of was the training and validation databases created for the BraTS 2018 Challenge [4]. The training set consists of 75 low-grade and 210 high-grade MRI brain images. The image data consists of 4 modalities T1, T1c, T2 and FLAIR, acquired from 19 different MRI scanners using different protocols [4, 7]. All the images had been segmented manually by several experts, and the average annotation is in fact the ground truth given in the database. The modalities are co-registered, interpolated to the same resolution and skull-stripped. The annotated regions [8, 9] are labeled in 4 different classes: 0 for background and healthy tissue, 1 for NCR/NET (necrotic and/or non-enhancing tumor), label 2 for ED (the edema) and label 4 for ET (the enhancing tumor).

During preprocessing we handled three important artifacts: inhomogeneity correction, noise filtering and intensity standardization. For inhomogeneity reduction in MR images, we applied the N4 filter implemented in the ITK package [1]. The anisotropic filtering from the same package was used for noise reduction. Intensity normalization was done by histogram linear transformation in such a way that the first and third quartiles had predefined values.

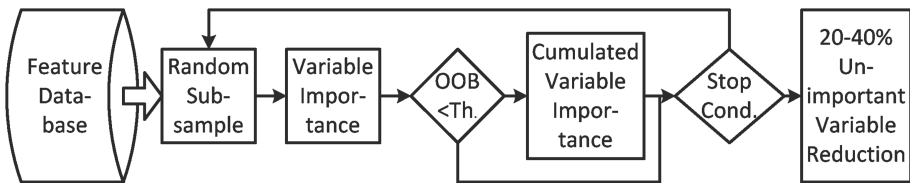
In voxel-wise segmentation it is necessary to define a set of intensity- and local neighboring features. The following features were extracted: first order operators (mean, standard deviation, max, min, median, Sobel, gradient); higher order operators (laplacian, difference of gaussian, entropy, curvatures, kurtosis, skewness); texture features (Gabor filter); spatial context features (symmetry, projection, neighborhoods), – the same as in our previous work.

The segmentation workflow given in Fig. 1 requires nine binary classifiers. Each classifier is trained and evaluated on its own feature database during its training process (Fig. 2). The global training consists of five training stages and each stage is composed of the following four steps:

1. feature selection based on variable importance [13] provided by the random forest;
2. incremental training of the RF stage-classifier;
3. optimization of the classification performance according to the task of the given tumor tissue segmentation;
4. image post-processing, with the role of reducing false detections and implementing the layered structure of tumor tissues.

The first step (1), feature selection based on the variable importance provided by the RF algorithm, and the third step (3), the performance optimization of the random forest classifier, were presented in our previous articles [14, 15]. These approaches were used to create our one-stage segmentation system presented at the previous BraTS Challenge in 2016 [15]. In our current work we use these algorithms in each of the five stages.

In the first step we defined 960 different features for each voxel. The RF classification algorithm is not able to deal with all the input image voxels and all 960 features previously defined, due to hardware and software limits. Therefore, this large amount of data was handled by taking advantage of the random forest variable importance evaluation. Our idea was to implement an iterative feature selection algorithm presented in [14]. The main idea of the algorithm is to evaluate the variable importance several times on a randomly chosen part of the feature database (Fig. 3). If the OOB error of the forest ensemble was below a certain threshold then the variable importance was taken into consideration and cumulated. Averaging the variable importances in the iterations the algorithm was able to eliminate the most unimportant 20–40% of variables in each run.



**Fig. 3.** Feature selection algorithm

In random forest approaches the training set is usually created out of the existing annotated images by random subsampling. In the case of BraTS 2018 the annotated image set contains 285 MR images and each image is made up of about 1,500,000 voxels, which means about 450 million samples. This huge database is, in addition, extremely unbalanced. In consequence we must obtain a well-defined database for training our random forest classifier. The solution to this is the incremental learning procedure that consists of enlarging the current training set by incrementally adding incorrectly classified random subsamples. In the second step (2), this incremental learning is repeated several times until the classification performances are adequate or the upper limit of hardware and/or software is reached. The flowchart of the incremental learning is given in Fig. 4.

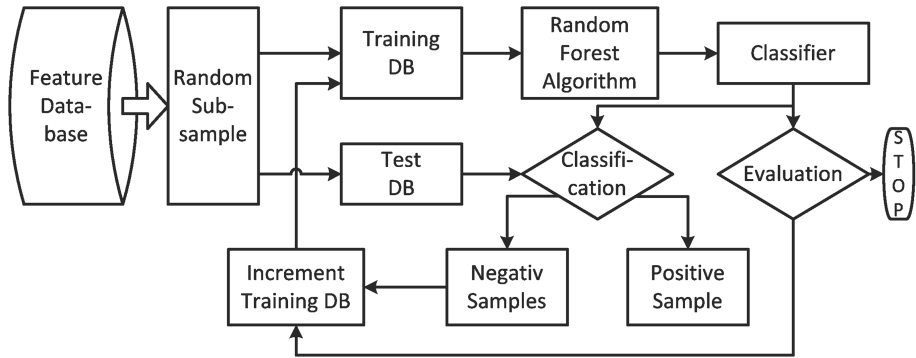


Fig. 4. Incremental learning

The classifier performance optimization (step 3) is in strong correlation with the segmentation task. This assumes the correct choice of training parameters. The random forest classification performance can be tuned via three important parameters:  $m_{trees}$  – the number of randomly chosen features used as a splitting criterion in each node of the trees; the  $n_{trees}$  – the number of trees in the forest;  $n_{nodes}$  – the maximum number of nodes in each tree. These parameters determine the size of the random forest ensemble, the segmentation performances, training time and system complexity, as well. In our experiments [15] these durations can be drastically reduced without any loss in segmentation accuracy.

The last step (4), after the training of each stage-classifier (Fig. 2), is an image post-processing step to do with the segmentation goal of the current stage. Here we managed to incorporate some knowledge about the tumor, such as the number of distinct tumors in a brain, one tumor is a connected zone within the healthy brain tissue, the tumor core is inside the edema, the enhancing tumor is a connected zone inside the whole tumor, etc. By applying this post-processing step we succeeded to eliminate the most of the false detections and improve the quality of segmentation.

## 2.2 Survival Prediction

Survival prediction has a considerable role, especially from a medical perspective, as well as that of the life expectancy of the patient. It helps in monitoring the effects of the medication and treatment applied. This prediction has to be correlated with the disease state and physical well-being of the patient. In this task the only information available were the MRI scans and the age of a small number of patients (59). In the case of this reduced dataset, prediction becomes difficult and leads to a high margin of error. In our survival prediction approach, we evaluated first-order statistics of feature values used in the segmentation phase. The mean and standard deviation of features were computed in the three segmented regions: edema, enhancing tumor and non-enhancing tumor. In order to include the size of each region, the statistical values were weighted by the number of voxels detected over the size of the brain in voxels. During the segmentation task, we determined a total of 120 local features with high importance values. Hence, the means and standard deviations of these 120 features are computed for the 3 tumoral regions, giving a total of 720 features. In our survival prediction method, we trained a random forest regressor with these features, limited the number of trees to 300 and considered the mean squared error as a split criterion. In order to reduce the effect of overfitting, the number of leaves on each tree was also maximized to 128.

## 3 Results

### 3.1 Segmentation

The proposed discriminative model is quite laborious and the proposed classifiers have to be tuned separately (Fig. 1). For training we used well-chosen samples, provided by incremental learning, from the entire BraTS 2018 training dataset. In training, beginning with Stage II, we created different classifiers for HGG and LGG as explained above. The results obtained were evaluated on both sets, namely, the complete training and validation sets, and on the test set within the challenge.

The stage I classifier determines the ROI (region of interest) that contains the tumor region with a high probability. This binary classifier was trained on the whole brain in order to delimit the healthy region from the tumoral region. In this step we used an incrementally trained classifier and applied a post-processing step consisting of a region dilation of 3 voxels. In addition, the two most important connected zones were taken into consideration. The results of the last three incremental steps are given in Table 1 and the improvement brought by the post-processing in Table 2. In this way the ROI obtained is about twice as large as the whole tumor, but the sensitivity reached on the complete training set is 0.989. The correct determination of this ROI has a crucial role in the subsequent stages. Table 1 and Fig. 5a show the average of the sensitivity after the binary segmentation and post-processing in stage I. The ROI obtained reduces the image region

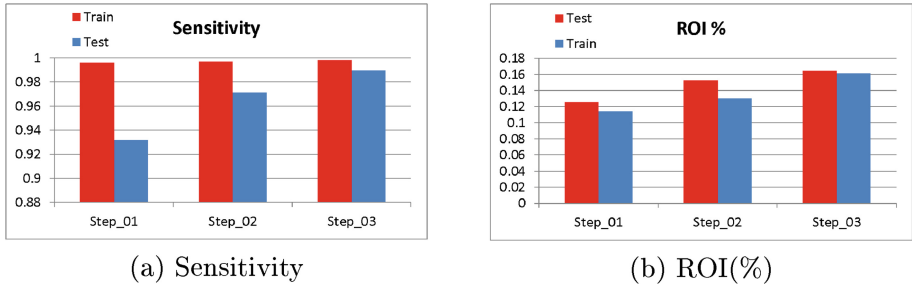


Fig. 5. Stage I segmentation

Table 1. Incremental training of Stage I segmentation

Stage I	Training			Test		
	Sens.	PPV	ROI%	Sens.	PPV	ROI%
Step_01	0.9959	0.5184	0.1143	0.9319	0.508	0.1257
Step_02	0.9971	0.4526	0.1301	0.9712	0.444	0.1523
Step_03	0.9984	0.4299	0.1615	0.9898	0.4257	0.1642

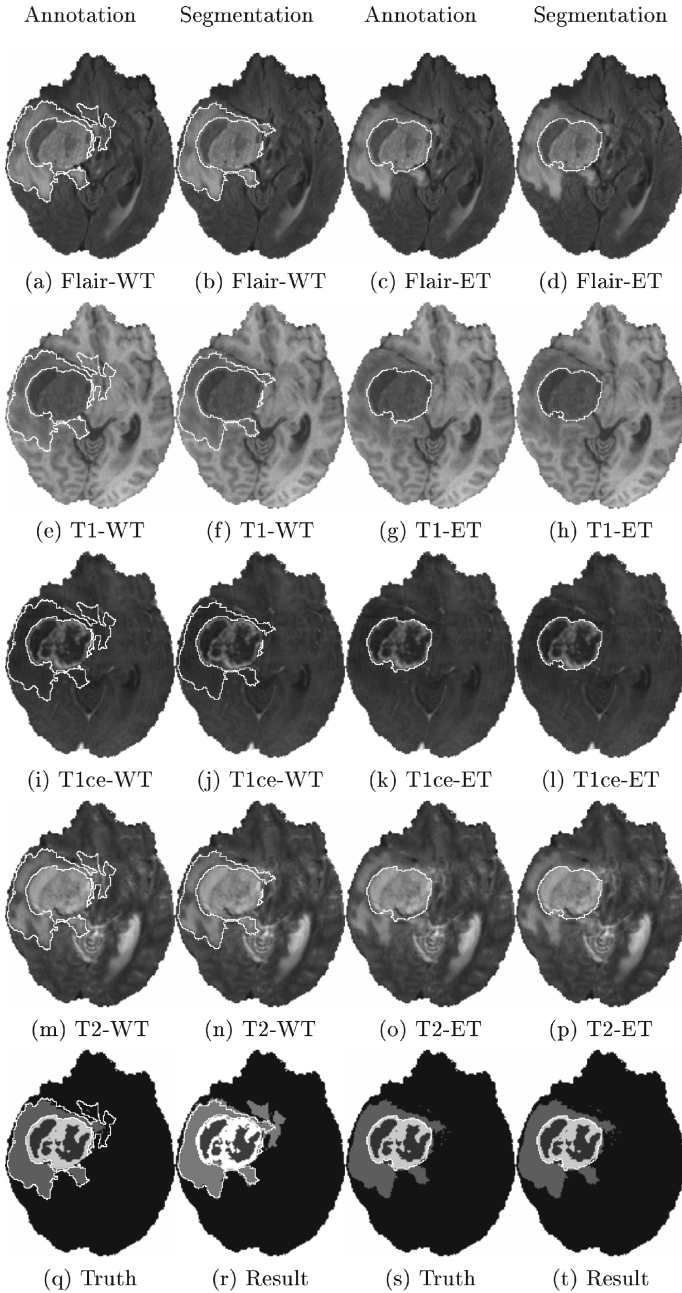
and, implicitly, the feature database, by about 8 times (Fig. 5b). This allows us to create a more precise classifier in the next stages.

In stage II the images are classified into two types with regard to segmentation. In the HGG images, the enhancing tumor is a considerable part of the whole tumor and may include some necrotic tissue. In LGG images, the greater part of the tumor consists of edema and non-enhancing tumor. By applying this kind of LGG-HGG separation, we could reduce the effects of unbalanced data (especially the ET in LGG). The image classes obtained correspond in a proportion of more than 95% to the medical LGG-HGG classification given for the training set. The subsequent stages (III, IV and V) are trained differently for LGG and HGG images.

The stage III classifier is applied only on the ROI. Its segmentation task is to delimit the remainder of healthy tissue from the WT. In this stage the segmentation with post-processing creates two disjunct regions, considering the tumor zone a connected region inside the healthy tissue. The Dice scores obtained for the whole tumor (WT) are 0.911 on the training and 0.885 on the validation sets (Table 3).

The stage IV classifier is applied only inside the WT region, and its task is to delimit the enhancing tumor. In the case of HGG images, the ET forms a significant connected region including some necrotic tissue. In the case of LGG images, the region is only a small piece in the WT and may easily be confused with vessels. If the segmentation obtains an ET region of less than 100 voxels, it will be neglected and considered to be a vessel. If the ET is near vessels, false detections are often obtained.





**Fig. 6.** Segmentation example (Rows represent the 4 MRI modalities and the segmentation obtained. Column 1 is the annotated contour of edema; column 2 is the segmentation result of it. Column 3 is the annotation contour of the  $ET \supset NECR$ ; column 4 is the segmentation results of it.)

**Table 2.** Post-processing: Improvements in sensitivity

	Sensitivity	PPV	ROI%	Dice
Classif.	0.963	0.581	0.112	0.707
Step_03	0.990	0.426	0.164	0.574

**Table 3.** Segmentation results of Stage IV & V on the validation database

Label		Dice			Hausdorff95		
		ET	WT	TC	ET	WT	TC
Training	Mean	0.880	0.913	0.911	3.433	4.721	5.434
HGG	StdDev	0.067	0.064	0.061	6.973	7.580	7.244
Training	Mean	0.808	0.916	0.908	3.057	2.725	2.907
LGG	StdDev	0.320	0.045	0.120	5.974	2.050	3.399
Validation	Mean	0.801	0.883	0.786	5.811	7.410	11.511
HGG	StdDev	0.161	0.082	0.180	9.418	12.644	14.141
Validation	Mean	0.489	0.891	0.415	10.265	4.741	15.858
LGG	StdDev	0.420	0.081	0.285	8.864	2.951	9.875
Final	Mean	0.730	0.885	0.702	6.349	6.803	12.499
Validation	StdDev	0.275	0.081	0.259	9.392	11.230	13.349
Final	Mean	0.684	0.830	0.657	6.186	9.180	11.649
Test	StdDev	0.302	0.193	0.308	9.394	13.062	12.670

The stage V classifier has the most difficult task working on the WT, excluding the ET obtained in stage IV. It has to delimit the edema from the non-enhancing tumor tissues. In the case of HGG images, this stage contains another classifier that finds the necrotic tissues inside the ET. The results obtained on the training validation and test images are given in Table 3. A visual segmentation sample is depicted in Fig. 6.

### 3.2 Survival Prediction

The results obtained in survival prediction are in strong correlation with the segmentation performances. Concerning the MSE parameter (Mean Squared Error), meaning the squared difference in number of days, we managed to take the first place on the validation database, as shown on the leaderboard [5] and in Table 4. The individual results of the test have no basis for comparison to the other teams owing to the lack of their results. The comparative study will be published by the organizers of BraTS 2018 [6].

**Table 4.** Survival prediction of the validation database

Team	Cases Eval.	Acc.	MSE	median SE	std SE	Spearman R
lefko	28	0.429	76081.29	24352	111586.8	0.342
Ranking		13	1	4	2	5

## 4 Conclusion and Discussion

In this paper we developed a five-stage discriminative model for brain tumor segmentation based on multi-modal MRI data. Our five-stage model implements the layered tissue structure by adequate training of binary classifiers and image post-processing in each segmentation stage. In each stage we attempted to solve the four important issues concerning discriminative models. Our results show that binary classifiers are very efficient for the layered segmentation task. One of the most important results is the determination of a ROI that has to enclose the whole tumor with a very high probability. In stage I, the sensitivity attained is 0.989, with a PPV of 0.426. This step reduces the size of the feature database by about 8 times and provides a reliable ROI for the next segmentation stages. Furthermore, the LGG-HGG separation increased the Dice score by 2%. The WT segmentation reached a Dice score of about 0.885 both on the training and validation sets [5]. This result is comparable to the most well-performing methods. In the test set, the reported Dice decreased by 5%, to 0.83. Analysing the test set, we came to the conclusion that the test set contained many HGG images with different visual aspects compared to the training or validation images. The final results of the survival task in correlation with segmentation performances will be published soon by the BraTS organizers. In our opinion, the MSE score is much more relevant than the accuracy that considers three disjunct time periods as a crisp set (less than 10 months, between 10–15 months, more than 15 months). The system developed is a complex implementation using a large variety of software packages and modules such as ITK in C++ [1], Java, ImageJ and Fiji with Trainable Weka Segmentation [2], the random forest package from R [3], Matlab for performance evaluation and image conversion. Our system is quite complex and still we are working on its dockerized version.

## References

1. ITK. <http://www.itk.org/>
2. Trainable segmentation. [http://imagej.net/Scripting\\_the\\_Trainable\\_Segmentation](http://imagej.net/Scripting_the_Trainable_Segmentation)
3. The R project for statistical computing. <https://www.r-project.org/>
4. Multimodal Brain Tumor Segmentation Challenge 2018. <https://www.med.upenn.edu/sbia/brats2018/>
5. Multimodal Brain Tumor Segmentation Challenge Evaluation Website. <https://www.cbica.upenn.edu/BraTS18/>
6. Bakas, S., Reyes, M., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. arXiv preprint [arXiv:1811.02629](https://arxiv.org/abs/1811.02629) (2018)

7. Bakas, S., et al.: Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci. Data* **4**, 170117 (2017)
8. Bakas, S., et al.: Segmentation labels and radiomic features for the pre-operative scans of the TCGA-GBM collection. *Cancer Imaging Arch.* (2017). <https://doi.org/10.7937/K9/TCIA.2017.KLXWJJ1Q>
9. Bakas, S., et al.: Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection. *Cancer Imaging Arch.* (2017). <https://doi.org/10.7937/K9/TCIA.2017.GJQ7R0EF>
10. 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 *Proceedings of MICCAI-BRATS Challenge* (2017)
11. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
12. Jakab, A.: Segmenting brain tumors with the slicer 3D software manual for providing expert segmentations for the BRATS tumor segmentation challenge. University of Debrecen/ETH Zürich (2012)
13. Goldstein, B.A., Polley, E.C., Briggs, F.: Random forests for genetic association studies. *Stat. Appl. Genet. Mol. Biol.* **10**(1), 1544–6115 (2011)
14. Lefkovits, L., Lefkovits, S., Emerich, S., Vaida M.F.: Random forest feature selection approach for image segmentation. In: *Proceedings of the 9th International Conference on Machine Vision (ICMV)*, vol. 10341, p. 1034117 (2016)
15. Lefkovits, L., Lefkovits, S., Szilágyi, L.: Brain tumor segmentation with optimized random forest. In: Crimi, A., Menze, B., Maier, O., Reyes, M., Winzeck, S., Handels, H. (eds.) *BrainLes 2016*. LNCS, vol. 10154, pp. 80–99. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-55524-9\\_9](https://doi.org/10.1007/978-3-319-55524-9_9)
16. Menze, B.H., Jakab, A., Bauer, S., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* **34**(10), 1993–2024 (2015)
17. Phophalia, A., Maji, P.: Multimodal brain tumor segmentation using ensemble of forest method. In: Crimi, A., Bakas, S., Kuijf, H., Menze, B., Reyes, M. (eds.) *BrainLes 2017*. LNCS, vol. 10670, pp. 159–168. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-75238-9\\_14](https://doi.org/10.1007/978-3-319-75238-9_14)
18. Revanuru, K., Shah N.: Fully automatic brain tumour segmentation using random forests and patient survival prediction using XGBoost. In: *Proceedings of MICCAI-BRATS* (2017)
19. Soltaninejad, M., Zhang, L., Lambrou, T., Yang, G., Allinson, N., Ye, X.: MRI brain tumor segmentation using random forests and fully convolutional networks. In: *Proceedings of MICCAI-BRATS* (2017)