The Efficacy of Radiomics Feature Selection for Brain Tumor Survival Prediction

Dang-Linh Trinh, Soo-Hyung Kim, Hyung-Jeong Yang, Guee-Sang Lee
Department of Artificial Intelligence Convergence, Chonnam National University, Republic of Korea.
linhtd812@gmail.com, gslee@jnu.ac.kr, shkim@jnu.ac.kr, hyungjeong@gmail.com

요 약

Glioblastoma (also known as glioblastoma multiforme) is the most aggressive type of brain tumor, leading to rapid death for the patients. The longest survival time of the diagnosed person with glioblastoma is about two years; therefore, overall survival prediction is essential for surgical and treatment planning. In this paper, the radiomics features, including the first-order, shape-based, and texture features, are extracted from the Fluid Attention Inversion Recovery (FLAIR) MRI modality. The clinical information such as the age of patients is added, and the random forest is utilized for ranking the importance of features. The 25 most important features are chosen to fit the machine learning model to get the overall survival time. The efficacy of the proposed method is evaluated on Brain Tumor Segmentation 2018 (BraTS18) challenge validation datasets. The proposed method achieved 97,531.8 on mean squared error (MSE) metrics and 0.292 on SpearmanR.

Keywords: Overall survival prediction, feature selection, machine learning, radiomics features

1. Introduction

Gliomas are the most common type of malignant brain and are graded on a scale of one to four. Gliomas are categorized into two groups: low-grade gliomas (LGG) (including grade I and II) and high-grade gliomas (HGG) (including grade III and IV). The grade four gliomas are the most aggressive type and are known as glioblastoma. The average GBM survival time is around 12 to 28 months, and only 5% of GBM patients survive more than five years. Therefore, overall survival (OS) prediction plays a vital role in precise surgery and treatment planning. There is 3 type of sub-tumors in GBM including edema, necrotic and enhancing tumor.

Magnetic resonance imaging (MRI) is a non-invasive method that helps clinicians plan treatment for brain tumor patients. There are four modalities of MRI scans, including T1-weighted, T2-weighted, contrast-enhanced T1-weighted (T1ce), and Fluid Attention Inversion Recovery (FLAIR). While FLAIR gives the best view of the whole tumor, enhancing tumor boundaries is best seen in the T1ce modality [1]. Radiomics features are extracted from radiographic images relevant to the shape, texture, and first-order statistics of brain tumors. The machine learning algorithm is applied to find the association of these features to the survival time of patients.

In this paper, we propose a method that predicts the OS time based on two stages. Because of good segmentation on enhancing tumors, DKNet [2] is utilized for the first stages. For the second stage, radiomics features are extracted from the FLAIR modality, and the random forest is used for features selection based on Gini importance. These features fit the machine learning model to predict the survival day of each patient on the validation dataset. By using the most essential radiomics features and clinical information such as patient ages, we get promising results on the BraTS18 validation dataset [3][4][5].

2. Related Work

OS prediction is an essential tool which support clinician have a suitable treatment plan for glioblastoma patient. Nowadays, there are two main steps for survival prediction: segmentation of the brain tumor parts and then extracted features from segmented maps are fit to prediction model to get the survival days.

Weninger et al. [6] count the number of voxels of tumor segmentation, which means that volume information, location of the tumor, and clinical information like ages are valuable for OS prediction. Shboul et al. [7] also utilize the radiomics features and Euler characteristics-based features from different intra-tumor parts and fit to machine learning model for survival prediction. Li Sun et al. [8] use different convolutional neural networks to get the segmentation maps. Then all the radiomics features are extracted and selected by a decision tree to get potential features for

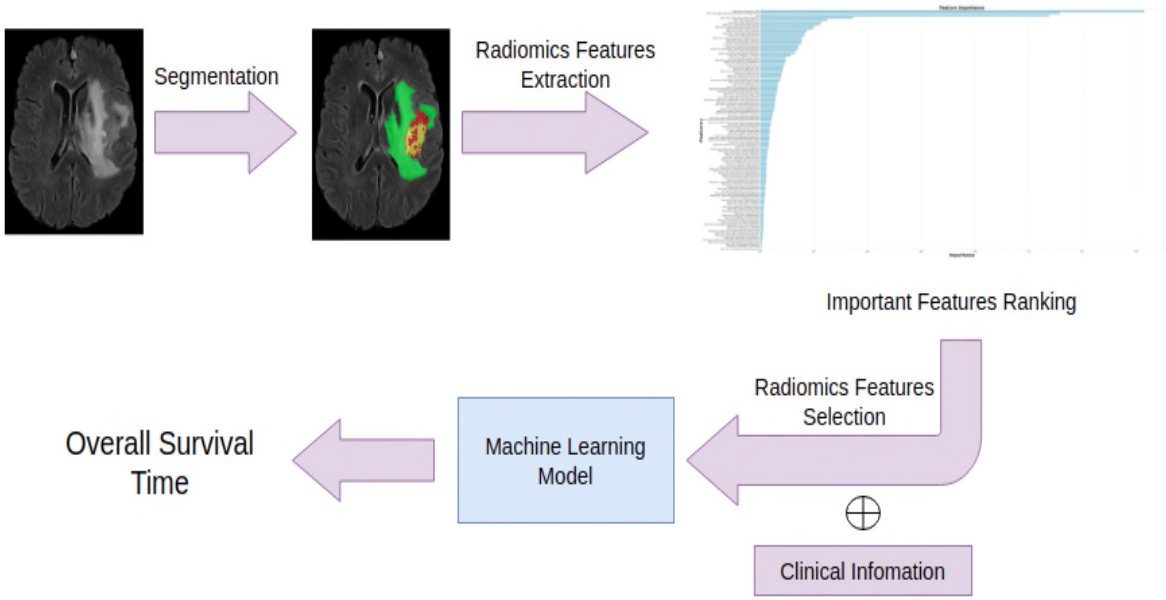

Figure 1. Architecture of proposed method.

OS prediction. According to [9], after tumor segmentation by a multi-planer spatial convolutional neural network, all radiomics features are fit to multilayer perceptron to predict OS time (Banerjee et al., 2019).

3. Proposed Method

Our proposed method contains the following steps, which are described in Figure 1. Enhancing tumors is hard for segmentation in brain tumors; therefore, DKNet [2] is implemented to tackle segmentation tasks. Next, all the radiomics features are extracted from the FLAIR modality and combined with the clinical information to predict the survival time of patients.

DKNet is based on a variant of 3D UNet called the dilated multi-fiber network (DMFNet) [10]. The biggest advantage of DKNet is that it could detect small-sized tumors, so the whole tumor is seen in more detail.

All radiomics features are extracted from the whole tumor for the survival prediction task because the whole tumor is best seen in the FLAIR modality. In the end, we added the clinical information and had 108 features in total. In order to reduce the number of features, a random forest is utilized for ranking the importance of these features. The most 25 critical features are selected based on the Gini index. Finally, we fit selected features on Light Gradient Boosting (LightGBM) regressor to predict the OS time.

4. Experimental Results

We used two metrics to evaluate the results, including Mean Squared Error (MSE) and spearman. While MSE calculates the difference between predicted survival days and ground truth, SpearmanR shows the association between our results and the ground truth. The range of SpearmanR is between -1 and 1, which 0 is no correlation. Thus, for survival prediction, the closer to +1 of SpearmanR, the better for the variable association. The fomulas of MSE and SpearmanR metrics is shown as below.

$$MSE = \frac{\sum_{k=1}^{n} \left(d_k - d_k^*\right)^2}{n}$$

where d_k and d_k^* are predicted days and ground truth, respectively, n is number of samples.

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)},$$

where d_i^2 is the difference ranking of survival days and ground truth, n is number of samples.

Table 1 shows our results, which achieve 97,531.8 and 0.292 on MSE and SpearmanR metrics, respectively. Baid [12] shows a good result on the SpearmanR metric; however, the MSE results are not good compared to other methods. The reason is that their model is better in the middle range and has lower performance at both extremes [12].

Method	MSE	SpearmanR
Weninger et al. [6]	101,012.0	0.258
Banerjee et al. [9]	180,959.4	0.273
Xue Feng et al [11]	103,839.3	0.247
Ujjwal Baid et al [12]	59,550,213.1	0.427
Ours (all features)	102,587.0	0.279
Ours (features selection)	**97,531.8**	**0.292**

Table 1 The OS prediction on our method compared other methods on BraTS 2018 validation dataset

5. Conclusion

In this paper, we propose a method for overall survival prediction on glioblastoma patients. Radiomics features are extracted from the FLAIR modality, and features selection is applied to get the most 25 crucial features. Our method minimizes the day error compared to other methods (shown in Table 1). In the future, instead of machine learning techniques, a deep learning model is applied to predict survival time, and more essential features such as location or clinical information are added to improve the results.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A3B05049058) and This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF)& funded by the Korean government (MSIT) (NRF-2019M3E5D1A02067961). The corresponding author is Guee-Sang Lee.

References

[1] Ujjwal Baid, Swapnil U Rane, Sanjay Talbar, Sudeep Gupta, Meenakshi H Thakur,Aliasgar Moiyadi, and Abhishek Mahajan.2020.Overall survival predictionin glioblastoma with radiomic features using machine learning.Frontiers incomputational neuroscience14 (2020), 61.

[2] Duc-Ky Ngo, Minh-Trieu Tran, Soo-Hyung Kim, Hyung-Jeong Yang, and Guee-Sang Lee. 2020. Multi-task learning for small brain tumor segmentation from MRI.Applied Sciences10, 21 (2020), 7790.

[3] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki,Justin S Kirby, John B Freymann, Keyvan Farahani, and Christos Davatzikos. 2017. Advancing the cancer genome atlas glioma MRI collections with expertsegmentation labels and radiomic features.Scientific data4, 1 (2017), 1 - 13.

[4] Spyridon Bakas, Mauricio Reyes, Andras Jakab, Stefan Bauer, Markus Rempfler,Alessandro Crimi, Russell Takeshi Shinohara, Christoph Berger, Sung Min Ha,Martin Rozycki, et al.2018. Identifying the best machine learning algorithmsfor brain tumor segmentation, progression assessment, and overall survivalprediction in the BRATS challenge.arXiv preprint arXiv:1811.02629(2018).

[5] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, KeyvanFarahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, RolandWiest, et al.2014. The multimodal brain tumor image segmentation benchmark(BRATS). IEEE transactions on medical imaging34, 10 (2014), 1993 - 2024.

[6] Minh-Trieu, T., Soo-Hyung, K., Hyung-Jeong, Y., Guee-Sang, L., In-Jae, O., Sae-Ryung, K. (2021). A deep learning-based method for glioblastoma survival prediction using local context. The 10th International Conference on Smart Media and Applications SMA2021.

[7] Shboul, Z. A., Alam, M., Vidyaratne, L., Pei, L., and Iftekharuddin, K. M. (2019). "Glioblastoma survival prediction," in Brainlesion Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (Quebec City, QC), 508 - 515.

[8] Li Sun, Songtao Zhang, and Lin Luo. 2018. Tumor segmentation and survival prediction in glioma with deep learning. In International MICCAI Brainlesion Workshop. Springer, 83 - 93.

[9] Subhashis Banerjee, Sushmita Mitra, and B Uma Shankar.2018.Multi-planarspatial-ConvNet for segmentation and survival prediction in brain cancer. InInternational MICCAI Brainlesion Workshop. Springer, 94 - 104.

[10] Chen Chen, Xiaopeng Liu, Meng Ding, Junfeng Zheng, and Jiangyun Li. 2019. 3D dilated multi-fiber network for real-time brain tumor segmentation in MRI. In International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 184 - 192.

[11] Xue Feng, Nicholas J Tustison, Sohil H Patel, and Craig H Meyer. 2020. Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic features. Frontiers in computational neuroscience 14 (2020), 25.

[12] Ujjwal, B., Sanjay, T., Swapnil, R., Sudeep, G., Meenakshi, H. T., Aliasgar, M., Siddhesh, T., & Abhishek, M. (2018). Deep learning radiomics algorithm for gliomas (drag) model: a novel approach using 3d unet based deep convolutional neural network for predicting survival in gliomas. In International MICCAI Brainlesion Workshop. Springer, 369 - 379.