3D Brain Tumor Survival Days Prediction using Knowledge Distillation

Tien-Bach-Thanh Do
Chonnam National University, South Korea
thanhdo.work@gmail.com

Soo-Hyung Kim
Chonnam National University, South Korea
shkim@jnu.ac.kr

Huyng-Jeong Yang
Chonnam National University, South Korea
hjyang@jnu.ac.kr

Guee-Sang Lee
Chonnam National University, South Korea
gslee@jnu.ac.kr

ABSTRACT

The brain tumor is the most dangerous cancer recently. For that reason, overall survival prediction plays an important role in diagnosis and treatment planning for brain tumor patients. The main target of our research is to demonstrate the effectiveness of features extracted from the combination of the whole tumor, tumor core, and enhancing tumor to overall survival prediction. We want to focus on patients that have resection status is gross total resection to see the effect of this feature on the patients. For segmentation, we used BiTr-Unet due to the advantage of the CNN-Transformer combined model. Dice coefficients for enhancing tumor, tumor core, and the whole tumor are 0.734, 0.78, and 0.885 respectively on the validation dataset. Furthermore, numerical features including the ratio of tumor size to brain size and the area of tumor surface as well as the age of subjects are extracted from predicted tumor labels. We used C-index to choose the best-affected feature radiomics to use for survival prediction. The accuracy could be 0.483 on the validation dataset.

KEYWORDS

Brain tumor segmentation, 3D U-Net, Survival days prediction, Distillation learning

1 INTRODUCTION

One of the primary brain tumors that kill the majority of brain tumor patients is glioma. The glioma has a few sub-regional structures, including enhancing tumor, necrotic core, and peritumoral edema (ET). The World Health Organization (WHO) [5] categorizes glioma tumors into four tiers based on their behaviors and microscopic pictures. Low Grade Glioma (LGG) and High Grade Glioma are two of the four levels of glioma tumors (HGG). LGG are benign tumors that come in grades I and II and grow gradually. The right LGG treatment could add many years to the patient's life. Additionally, HGG is an aggressive and cancerous tumor of Grade III and Grade IV. As the HGG tumor grows, quick surgery is necessary. For HGG tumors, the patients are unable to live longer than 15 months. The standard course of treatment for glioma patients, particularly HGG patients, is tumor excision. As a result, tumor segmentation is crucial for the detection and management of gliomas. Planning a course of treatment and diagnosing a condition both need the use of magnetic resonance imaging (MRI), a non-invasive technique.

T1-weighted, T2-weighted, contrast-enhanced T1-weighted (T1ce), and Fluid Attenuation Inversion Recovery (FLAIR) MRI sequences are used to detect brain tumors, as shown in Figure 1 below. Each modality provides a benefit for each subregion of

the tumor, such as FLAIR having a better overall image and T1 having a stronger contrast on the tumor core.

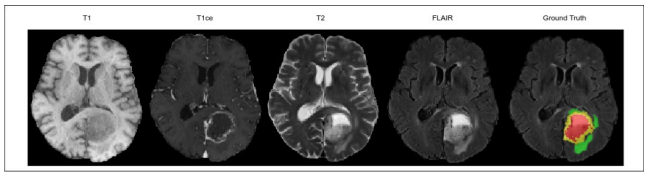

Figure 1: A representation of BraTS 2020 dataset where green part is edema, yellow parted is enhancing tumor, and red part is necrotic.

In this paper, we proposed a method to predict the survival time of patients, including two stages. First, due to the advantage of the CNN-Transformer combined model, we used the BiTr-Unet model for brain tumor segmentation. After the segmented task, we used the Pyradiomics package to generate radiomic features. Then we use knowledge distillation to learn and predict the survival days of patients.

Hinton et al. [11] first suggested knowledge distillation as a method of translating an ensemble's collective knowledge to a single neural network. The authors developed a novel sort of ensemble that could discriminate between fine-grained classes that the "large" models were misclassifying by combining numerous big models with many specialized models. During the final model's training, the soft labels might be used as a regularization. The MNIST dataset and the speech recognition challenge were used to show the usefulness of the suggested approach. Radosavovic et al. [12] examined omni-supervised learning, a specific example of semi-supervised learning using readily available labeled data as well as unlabeled internet-scale data sources, and presented the variation of knowledge distillation known as data distillation.

The authors suggested that by averaging the predictions generated by variously altered input data, unlabeled data may be annotated using a single model. A student model was then trained on the merged dataset using the automatically labeled data. Student models outperform models trained just on labeled data when it comes to human key point detection and general object detection, as was shown in the data distillation demonstration.

Inspired by previous papers, we used the knowledge distillation method to train the student model with labeled data from the BraTS 2020 dataset.

2 Related works

To forecast the survival days of patients, a lot of studies have been done recently. The two stages of the majority of trending methods are segmentation and using radiomics information from the segmented map to forecast survival time. The segmented map is obtained by Shboul et al. using an ensemble of random forest and convolutional neural networks, and the RF predictive model is then used to predict the survival days using the 240 most crucial features out of a total of 1366 features based on the Kaplan-Meier curve [13]. To produce the segmented maps in [14], Baid et al. employ a 3D Unet model with a three-stage encoder-decoder architecture. Following that, radiomics characteristics are retrieved and fitted to an MLP network to predict OS. To make the segmentation task's findings more reliable, Feng et al. combine 3D UNets with different input sizes, the number of encoder/decoder blocks, and layer counts. Volume, shape, and surface area are manually determined based on the segmentation result, along with clinical data, to determine OS days using linear regression [15]. The final segmentation result was obtained in [16] by applying a boosting ensemble of three distinct networks, including Unet [17], DFKZnet [18], and CA-CNN [19]. For survival prediction, 14 radiomics features carefully chosen are fitted to the random forest. Location-based features effectively improve the prediction performance, according to Madjid [20] et al.

They determine the location-based features by the distance between the tumor and the center of the brain, the tumor's biggest diameter, and the vertical slide number with the largest tumor diameter. The overall survival prediction result marginally improves when combined with the chosen radiomics features. [21] claims that the size, shape, and texture of the tumor are described by deep features that are retrieved from MRI modalities using a CNN network. To strengthen the segmentation findings, the NLSE model is further proposed with the integration of the non-local module and the squeeze-and-excitation module. The factor analysis is then used to decrease the radiomics and deep features, and all of the remaining features fit the RF regression to determine overall survival days.

Additionally, several researchers suggested employing the derived features from MRI modalities directly due to the time-consuming nature of tumor segmentation for ground-truth labeling. A saliency map that pinpoints the site of the tumor is used by Renato et al. in [22] to offer an OS prediction approach without segmentation. The OS prediction results are promising compared to previous approaches when combined with clinical data. [23] claims that to produce OS prediction results, Linmin Pei et al. directly extract high-dimensional characteristics from the CNN network; ages are the supplementary features used to enhance the findings. The LASSO approach is used to choose relevant features, and all features are then fitted to linear regression to produce results.

3 PROPOSED METHOD

3.1 Segmentation

BiTr-UNet [24] is used for the brain tumor segmentation task to segment into tumor sub-regions, including edema, necrotic, and enhancing tumors. BiTr-Unet [24] differs from TransBTS [9] in that the model adds two ViT layers in the deep skip-connection part to model global features. Due to the advantage of the CNN-Transformer model, BiTr-Unet [24] can learn local information while maintaining long-range information together.

3.2 Radiomics features

We have four 3D brain volumes in the Nifty (nii.gz) format for each patient, with a 240 x 240 x 155 matrix size for each data set. We created radiomic features from the original four MRI images (T1, T1ce, T2, and FLAIR) and the mask of the projected tumor core using the pyradiomics module (https://www.radiomics.io) [25] offered by the Computational Imaging and Bioinformatics Lab at Harvard Medical School. A npy file (a python file format) was used to store each subject's 2065 radiomics features for later analysis. Hereafter, the file is referred to as MRI-npy.

After that, we used C-index to see what radiomic features have the most effective on a brain tumor. We saw that 't1_nec_original_glcm_MaximumProbability' feature has the highest value in C-index.

Finally, we concatenate the above radiomic feature with clinical information (Age, Resection Status).

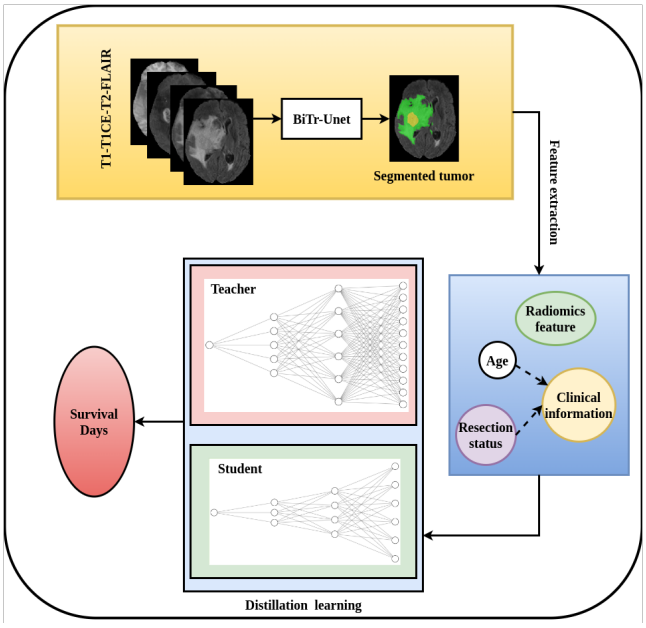

Figure 2: Overall proposed method

3.3 Survival Prediction

For the survival prediction task, we used distillation on learning methods. We used deeper MLP for the teacher model to learn. After that, we used the learned weight from the teacher model to improve the student model. To prevent the over-fitting issue, the training dataset is divided into training and validation datasets in an 80:20 ratio.

4 EVALUATION METRICS FOR SURVIVAL PREDICTION

The accuracy, mean squared error (MSE), and SpearmanR metrics are used to evaluate the results of the survival prediction. All results are uploaded to a website link for review (https://ipp.cbica.upenn.edu/). The key objective is classification, which can be broken down into three categories: short survivors (less than 10 months), medium survivors (between 10 and 15 months), and long survivors (more than 15 months). The following equation is used to obtain the accuracy metrics:

$$Accuracy = \frac{Totalof Correctedsurvivalpredictions}{Totalpatients} \quad (1)$$

Additionally, because this is a supervised learning method, the MSE metric displays the discrepancy between the predicted patient survival days and the actual patient survival days (ground truth) presented in the dataset. Prediction values are better when the MSE values are smaller. The intensity and direction of the monotonic connection between the anticipated results and the ground truth are described by the SpearmanR metric (GT). The value of SpearmanR falls between -1 and +1. The ranking of the patient's survival times will be closer to GT's order if the SpearmanR value is close to 1. The following equations are used to calculate the MSE and SpearmanR metrics:

$$MSE = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n} \quad (2)$$

where Y_i and $\hat{Y}_i$ represent the anticipated and actual patient survival days (GT), respectively.

$$SpearmanR = 1 - \frac{6\sum T^2}{n(n^2-1)} \quad (3)$$

where n is total number of patients, T is difference in paired ranks between predicted days and group truth days.

5 EXPERIMENTAL RESULTS

5.1 Dataset

Based on the BraTS 2020 dataset, we tested our model. Four different types of MR images were used to create the 3D volumes of the medical imaging: T1-weighted (T1), post-contrast T1-weighted (T1ce), T2-weighted (T2), and fluid-attenuated inversion recovery (FLAIR). 369 patients make up the training dataset and 125 patients make up the validation dataset. 19 imaging centers provided multi-modal scans.

All of the modalities in this dataset were constructed using the same anatomical template and sampled using a 1 mm x 1 mm x 1 mm dimension. Four specialists divided the MRIs of all the patients, and seasoned neuroradiologists reviewed and accepted their findings. Based on the intratumoral structures in gliomas, these labels were annotated. According to the tumor subregions, including the entire tumor, the tumor core, and the enhancing tumor (see Fig. 2), the segmentation results were assessed. Three zones make up the segmentation of a brain tumor: the augmenting tumor (yellow part), the tumor core (yellow part and red part), and the entire tumor (yellow part, red part, and green part).

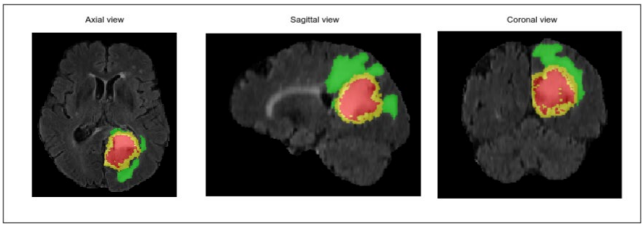

Figure 3: A visualization of Brain Tumor based on regions.

Additionally, the training data includes the ground truth for segmentation, which includes labels 1 for nET/necrotic, 2 for edema, 4 for ET, and 0 for background. Additionally, each patient's surviving days are provided in the training data. Due to the labeling accuracy, we use training data to determine how well the shape radiomics characteristic affects patient survival time. This dataset include contains clinical data, including age, tumor grade, and status of resection. For patients with HGG, the resection status includes gross total resection (GTR), subtotal resection (STR), and NA (not available).

5.2 Implementation Details

In our proposed network, we developed the proposed model on the Pytorch and Keras platform with python programming language. We trained our segmentation model with a batch size is 4 and an epoch is 400, and our survival prediction model with a batch size is 4 and an epoch is 100. We used NVIDIA GeForce RTX 3090 GPU processor.

5.3 Results

5.3.1 Segmentation results

We used the BraTS 2020 dataset for the training stage and then we evaluated the segmentation results on the validation set. The segmentation performances of BiTr-UNet model are 0.737, 0.807, and 0.894 in the Dice metric for the enhancing tumor, tumor core, and whole tumor respectively. We also compared the performance BiTr-UNet method with other methods such as 3D U-Net, Basic V-Net, Deeper V-Net Residual 3D U-Net, TransBTS, and nn-Unet, detailed in table 1.

Table 1: Segmentation results on the BraTS 2020 validation dataset

Method	ET_Dice	TC_Dice	WT_Dice
3D U-Net [6]	0.688	0.791	0.841
Basic V-Net [7]	0.618	0.753	0.846
Deeper V-Net [7]	0.689	0.779	0.861
Residual 3D U-Net [8]	0.716	0.765	0.825
TransBTS [9]	0.785	0.812	0.89
Nn-Unet [10]	0.820	0.851	0.889
BiTr-UNet[24] (we used)	0.737	0.807	0.894

5.3.2 Survival prediction results

We applied knowledge distillation learning for the survival prediction task. The student model got only 0.31, 212563.379, and 0.043 in accuracy MSE and SpearmanR, respectively. After applying knowledge from the teacher model, the student model got 0.483, 111201.448, and 0.104 in accuracy MSE and SpearmanR, respectively.

Table 2: Survival prediction results on the BraTS 2020 validation dataset

Method	Accuracy	MSE	SpearmanR
Russo et al. [2]	0.586	99776.58	0.338
Pang et al. [3]	0.517	151013.386	0.155
SrUNet [4]	0.552	101697	0.329
Student (ours)	0.31	212563.379	0.043
Learned_student (ours)	0.483	111201.448	0.104

6 CONCLUSION

This study suggests a two-stage methodology for estimating patients with glioblastoma's overall survival time. The segmentation of brain tumors based on BiTr-UNet is the first stage. The second stage involves using the Pyradiomics software to extract radiomics features from the combination of the enhancing tumor, the entire tumor, and the tumor core. To forecast the overall survival time, we applied the weight to the student model after they were combined with the clinical data and fitted into a deep multi-layer perceptron teacher model. Results from experiments using the BraTS 2020 dataset have demonstrated this. Although we applied knowledge distillation learning to survival prediction, our result is not as well as other top methods. We will try to make the teacher model learn better.

ACKNOWLEDGMENTS

In the future, 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) and by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2018R1D1A3B05049058 & NRF-2020R1A4A1019191).

REFERENCES

[1] Patricia S. Abril and Robert Plant. 2007. The patent holder's dilemma: Buy, sell, or troll? *Commun. ACM* 50, 1 (Jan. 2007), 36–44. DOI: http://dx.doi.org/10.1145/1188913.1188915

[2] Russo et al. 2020. Impact of spherical coordinates transformation pre-processing in deep convolution neural networks for brain tumor segmentation and survival prediction. International MICCAI Brainlesion Workshop.

[3] Pang et al. 2020. Glioma segmentation using encoder-decoder network and survival prediction based on cox analysis. International MICCAI Brainlesion Workshop.

[4] Pei et al. 2020. Multimodal brain tumor segmentation and survival prediction using a 3D self-ensemble ResUNet. International MICCAI Brainlesion Workshop

[5] Louis et al. 2021. The 2021 WHO classification of tumors of the central nervous system: a summary. Neuro-oncology.

[6] Cicek et al. 2016. 3d u-net: ¨ learning dense volumetric segmentation from sparse annotation. In: International conference on medical image computing and computer-assisted intervention. pp. 424–432.

[7] Milletari et al. 2016. S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV). pp. 565–571. IEEE

[8] Zhang et al. 2018. Road extraction by deep residual u-net. IEEE Geoscience and Remote Sensing Letters 15(5), 749–753

[9] Wang et al. 2021. Transbts: Multimodal brain tumor segmentation using transformer. International Conference on Medical Image Computing and Computer-Assisted Intervention

[10] Isensee et al. 2020. nnU-Net for brain tumor segmentation. International MICCAI Brainlesion Workshop

[11] Hinton et al. 2015. Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531

[12] Radosavovic et al. 2015. Data distillation: towards omni-supervised learning. arXiv e-prints arXiv:1712.04440

[13] Shboul et al. 2017. Glioblastoma and survival prediction. Int. MICCAI Brainlesion Workshop, 10670, 358–368

[14] Baid et al. 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: Cham, Switzerland; pp. 369–379

[15] Feng et al. 2020. Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic features. Front. Comput. Neurosci. 14, 25.

[16] Sun et al. 2018. Tumor segmentation and survival prediction in glioma with deep learning. In International MICCAI Brainlesion Workshop; Springer: Cham, Switzerland, pp. 83–93.

[17] Ronneberger et al. 2015. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, pp. 234–241.

[18] Isensee et al. 2017. Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge. In International MICCAI Brainlesion Workshop; Springer: Cham, Switzerland, pp. 287–297.

[19] Wang et al. 2017. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In International MICCAI

Brainlesion Workshop; Springer: Cham, Switzerland, pp. 178–190.

[20] Soltani et al. 2021. Efficacy of Location-Based Features for Survival Prediction of Patients with Glioblastoma Depending on Resection Status. Front. Oncol. 11,2509.

[21] Huang et al. 2021. Overall Survival Prediction for Gliomas Using a Novel Compound Approach. Front. Oncol. 3150, 724191.

[22] Hermoza et al. 2021. Post-hoc Overall Survival Time Prediction from Brain MRI. In Proceedings of the 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 13-16 April 2021; pp. 1476-1480

[23] Pei et al. 2020. Context aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images. Sci.Rep. 1-11.

[24] Jia et al. 2022. Bitr-unet: a cnn-transformer combined network for mri brain tumor segmentation. International MCCAI Brainlesion Workshop. Springer.

[25] Van Griethuysen et al. 2017. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research, 77(21), e104-e107.