A Comparison of Machine Learning Models for Survival Prediction of Patients with Glioma Using Radiomic Features from MRI Scans

Madhumitha Manjunath[1,*]   Shayana Saravanakumar[1,*]   Shreya Kiran[1,*]   Jhinuk Chatterjee[1]

[1] Department of Biotechnology, People's Education Society University, Bangalore, Karnataka, India

Indian J Radiol Imaging 2023;33:338–343.

Address for correspondence   Jhinuk Chatterjee, PhD, Department of Biotechnology, People's Education Society University, Bangalore, Karnataka, India (e-mail: jhinukchatterjee@pes.edu).

Abstract

Background   Glioma is a primary, malignant, highly aggressive brain tumor, with patients having an average life expectancy of 14 to 16 months after diagnosis. Magnetic resonance imaging (MRI) scans of these patients can be used to extract and analyze quantifiable features with potential clinical significance. We hypothesize that there is a correlation between radiomic features extracted from MRI scans and survival. Along with clinical data, the radiomic features could be used in survival prediction of patients, providing beneficial information for clinicians to design personalized treatment plans.

Methods   In our study, we have utilized 3D Slicer for tumor segmentation and feature extraction and performed survival prediction of patients with glioma using four different machine learning models.

Results and Conclusion   Among the models compared, we have achieved a maximum prediction accuracy of 64.4% using the k-nearest neighbors model, which was trained and tested on a combination of clinical data and radiomic features extracted from MRI images provided in the BraTS 2020 dataset.

Keywords
► survival prediction
► feature extraction
► MRI
► naïve Bayes
► random forest
► k-nearest neighbors

Introduction

Glioma is the most common primary malignant brain tumor, arising from glial cells. Affected patients have an average life expectancy of 14 to 16 months after diagnosis. Intratumor heterogeneity, a characteristic feature of glioma, causes therapeutic resistance, thereby resulting in poor prognosis.

Magnetic resonance imaging (MRI) is considered to be a standard diagnostic technique for detecting brain tumors due to its superior soft-tissue contrast and sensitivity for pathologies.[1] Feature extraction from MRI scans helps in tumor characterization, and these radiomic features could be used to predict patient survival. Predicting patient survival helps clinicians make personalized treatment plans and also stratify patients for clinical trials.

Survival prediction can be done using machine learning models trained on available clinical and radiomic data. Currently, RStudio is being used to implement machine learning models, and its excellent visualization packages are useful for data exploration and presenting results.[2]

In this study, feature extraction from MRI images using the radiomics extension of the 3D Slicer software was

---

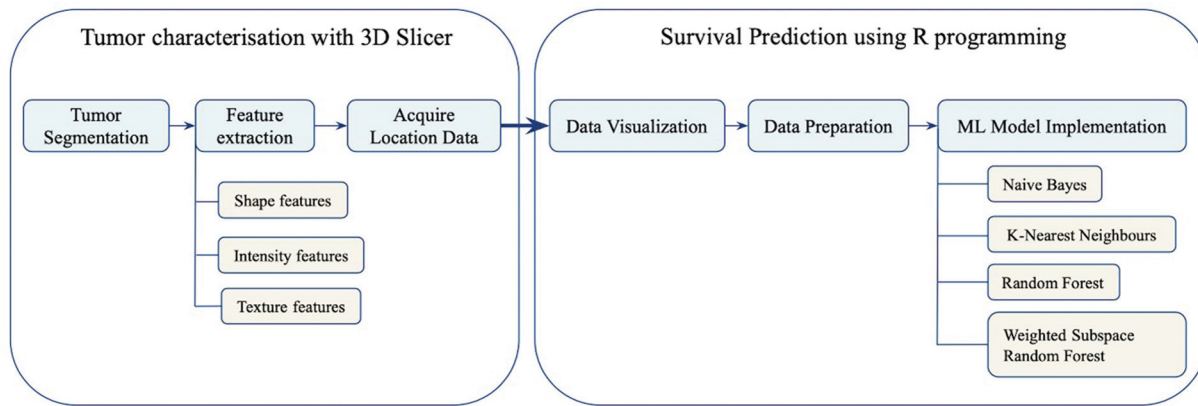* *These authors have contributed equally to the work.*

Fig. 1 Workflow of tumor characterization and survival prediction in this study. ML, machine learning.

performed. Furthermore, survival prediction was done using image features extracted as well as the available clinical features. The proposed workflow for feature extraction and survival prediction using multimodal MRI scans is illustrated in ►**Fig. 1**. First, the postcontrast T1-weighted (T1 gd) images are used to obtain shape, intensity, and texture features using 3D Slicer. Fourteen features were considered for further analysis with correlation values > 0.015. These radiomic features were then combined with the clinical features of each patient and used as input data for the prediction models, which were designed and implemented in RStudio. The four machine learning models considered are naive Bayes, k-nearest neighbors (KNN), random forest, and weighted random forest with cross-validation. The accuracies achieved with each of these models under different data splitting ratios were compared and a suitable model is suggested for further refinement.

Materials

Dataset
The BraTS 2020 dataset consisting of preoperative multimodal MRI scans in the form of NIfTI files in the .nii.gz format was used for this analysis. There were four modalities available in the dataset: native (T1), postcontrast T1-weighted (T1Gd), T2-weighted (T2) and T2 fluid-attenuated inversion recovery. The images provided were preprocessed by coregistering to the same anatomical template, interpolating to the same resolution (1 mm^3) and were skull-stripped. The clinical data provided included the age, life expectancy (survival in days), and resection status. MRI scans were available for 369 patients, out of which only 237 patients had clinical data and were used for the study.[3]

3D Slicer
3D Slicer is a software used for visualization and analysis of medical images. In our study, tumor segmentation and feature extraction was performed using the 3D Slicer interface and available modules and extensions.[4]

Methodology

Tumor Characterization
The following procedures were performed using the 3D Slicer software for 237 patients with a complete set of clinical data available. The postcontrast T1-weighted (T1Gd) scans were used in this study for tumor characterization, since areas targeted for resection are based on abnormal enhancement visible in these scans.[5]

Tumor Segmentation
The postcontrast enhanced T1-weighted scan (T1Gd) was loaded onto the 3D Slicer interface (►**Fig. 2A**). The Segment Editor module was then used to segment and differentiate the tumor tissue and normal tissue. Segments from three different slices were manually highlighted, using the Paint tool, with tumor tissue in green and normal tissue in yellow, as shown in ►**Fig. 2B**. The manually segmented tumor region was verified by simultaneously viewing the tumor in the previously segmented image provided in the BraTS dataset (see ►**Fig. 2C**). The segmented tumor in the BraTS dataset consists of three parts: the enhancing part of the tumor is shown in blue, the nonenhancing part of the tumor is shown in green, and the peritumoral edema is shown in yellow, as in ►**Fig. 2C**. We have manually segmented the tumor considering only the enhancing and nonenhancing parts of the tumor for our study, since preoperative treatment procedures usually involve only resection of the solid tumor, not the infiltrating parts. Then, using the Fast GrowCut algorithm and default parameters, a 3D model of the tumor was constructed using the chosen segments (►**Fig. 2D**). Due to the similar pixel intensities, additional pixels were also included in the reconstruction. These were removed by erasing the additional regions highlighted in the segment, using the Erase tool.[4] Once a defined tumor shape was obtained, feature extraction was performed.

Feature Extraction
The Segment Statistics module was used to compute tumor volume using a binary labelmap representation of the segment. To compute the shape, intensity, and texture
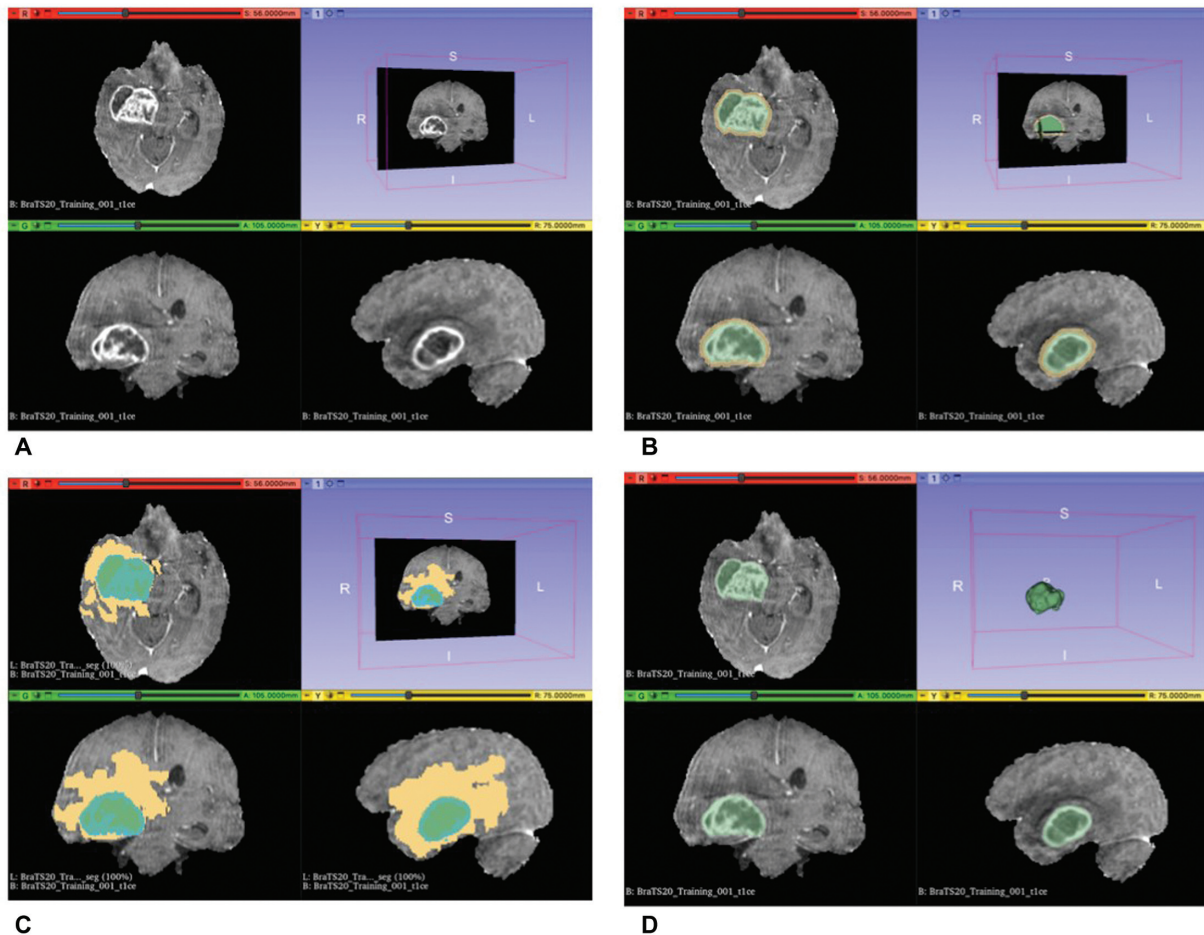
Fig. 2 Manual tumor segmentation performed using 3D Slicer. (**A**) Initial MRI scan loaded in the 3D Slicer interface. (**B**) Manual segmentation of the tumor (tumor tissue in *green*, surrounding tissue in *yellow*). (**C**) Verifying the segmented region by comparison with BraTS segmented image. The *blue* region indicates the enhancing part of the tumor, *green* region indicates the nonenhancing part, and *yellow* region indicates peritumoral edema. (**D**) Reconstructed volume of the tumor using Fast GrowCut.

features,[6] the radiomics extension (based on the PyRadiomics package in Python) was used.[7] Texture features quantify intratumor heterogeneity, which is an important factor that determines the prognosis. Among the available feature classes, "first order," "GLCM" (gray level co-occurrence matrix), and "3D shape" features were selected to be calculated. The bin width was set to 25 and symmetrical GLCM was enforced (default parameters). A total of 79 features were calculated, out of which 16 features were considered for further analysis.

Acquiring Location Data
The axial view of the MRI scan was divided into 9 regions based on slice 72 of the axial, slice 112 of the coronal, and slice 120 of the sagittal plane. With respect to these planes, the tumor locations were categorized as being in the right anterior, right center, right posterior, left anterior, left center, left posterior, center anterior, exact center, or center posterior region of the brain. The nine regions considered are shown in ►**Fig. 3**.

Analysis Using R Programming
A total of 229 patients were considered for analysis and survival prediction, out of the 237 patients whose clinical data were available. The data of patient 84 were not considered as it was the only patient that survived. The feature extraction for patients 87 and 177 could not be performed due to poor quality of the scans. The clinical data and feature data were combined to obtain a complete dataset to use for analysis with R.

Data Visualization
The packages used for data visualization were ggplot2[8] and RColorBrewer.[9] Initially, the relationship between age, location, and number of days of survival of the patient was plotted, as shown in ►**Fig. 4**. The nine different locations of tumors were differentiated using different colors. Following which, the distribution of patients in the dataset according to the location of the tumor was plotted (►**Fig. 5**). Correlation values were obtained for different factors in the complete data and a correlation plot was constructed using these values (►**Fig. 6**). The range of color and size of the
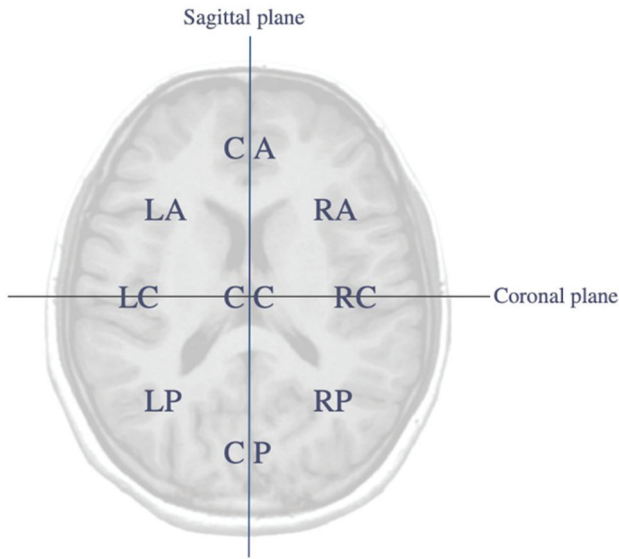
Fig. 3 Regions of the axial scan considered for location mapping. CA, center anterior; CC, exact center; CP, center posterior; LA, left anterior; LC, left center; LP, left posterior; RA, right anterior; RC, right center; RP, right posterior.
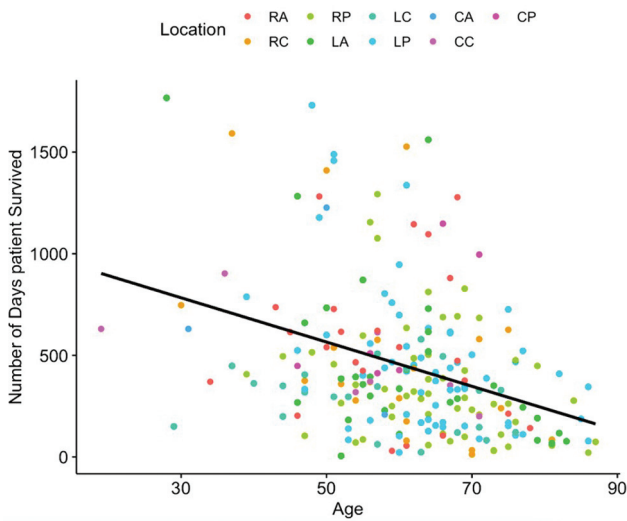

Fig. 4 Relation between patient age, location of tumor, and overall survival.. The *bold line* depicts the negative correlation between patient age and number of days of survival. Each point on the graph represents a patient and the color of the point represents the location of their tumor.

squares in the grid indicate the correlation between the factors against each other.

Data Preparation

First, the relevant features correlating with patient survival (correlation value > 0.015, with the exception of age) were filtered out and a dataframe was assembled. The features used for survival prediction are age, extent of resection, sphericity, surface area to volume ratio, 10th percentile, energy, entropy, kurtosis, mean, median, skewness, autocorrelation, cluster shade, and contrast.
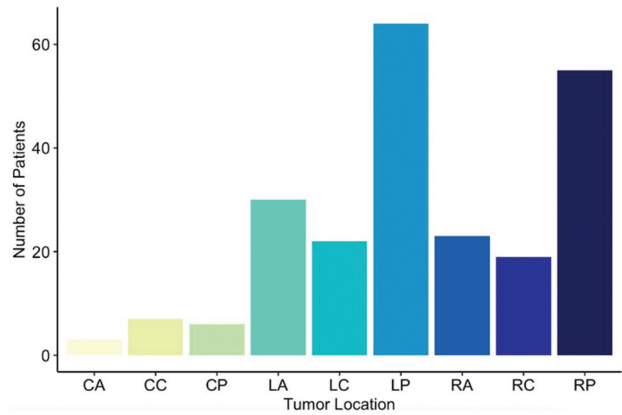

Fig. 5 Distribution of patients by location of their tumor. Color gradient is used only to differentiate the bars.
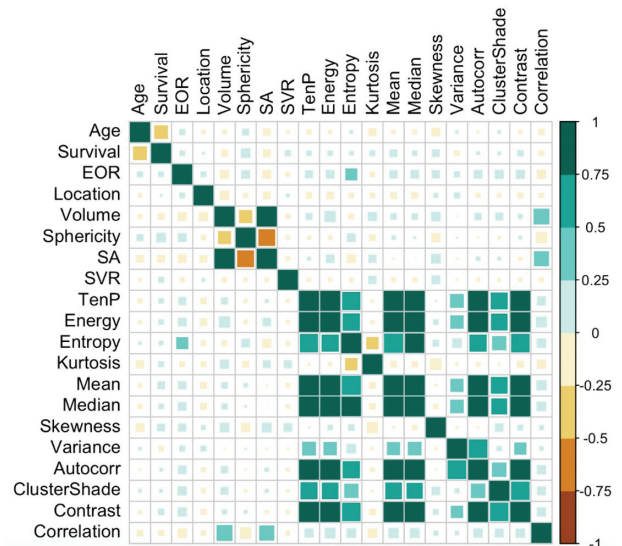

Fig. 6 Correlation plot of different factors. Positive correlation is indicated using hues of *green* and negative correlation is indicated using hues of *brown*.

Next, the number of days of survival of patients was converted to months, so that the patients can be classified into three categories: short survivors (<10 months), mid-survivors (10–15 months), and long survivors (>15 months).[10] In total, 91 patients were categorized as short survivors, 57 patients were midsurvivors, and 81 patients were long survivors.

Extent of resection (gross total resection [GTR], subtotal resection [STR], and NA) and the survival categories were converted to numerical values to be used in prediction. GTR and STR were replaced with 1 and NA values were replaced with 0. Short survivors were indicated with 1, midsurvivors with 2, and long survivors with 3. These categorical variables were then converted to the factor datatype, with extent of resection having two levels and survival having three levels. Patient age was rounded off to the nearest year.

The 229 patients in the combined data were partitioned into training and test sets based on three different training set to testing set ratios: 70:30, 80:20, and 90:10.

Model Design and Survival Prediction

These models were coded for and evaluated in RStudio (RStudio Team, 2022). The packages used for machine learning were tidyverse,[11] Classification And REgression Training (caret),[12] random forest,[13] and wsrf.[14] Four classification models were implemented to predict patient survival by classifying the patients as short, mid, and long survivors. The four models used were: naive Bayes classifier: based on normal distribution; KNN: based on the category with most neighbors; random forest: based on generation of decision trees with different features input; and weighted subspace random forest with 10-fold cross validation: refined random forest model with weightage assigned to different factors.

For all the four models and data partition ratios, a common seed value of 12 was used with the rounding sample kind. For the naive Bayes classifier, the default parameters of the model were used. For the KNN model, we computed the optimal number of neighbors to be considered that gave the highest accuracy, using a sequence of values for the tune grid parameter. The optimum value was found to be 27 neighbors in our case. Default values were considered for all other parameters. After iterating the random forest model with a different number of trees, the number of trees giving highest accuracy was 128, which was used in the final training. The weighted subspace random forest model was implemented using its default parameters.

Results

Feature extraction was performed using 3D Slicer and the computed features were combined with clinical data to generate complete data for 229 patients. The relationship between different factors and patient survival was investigated.

It can be inferred from ►**Fig. 4** that irrespective of the age of the patient, the location of the tumor greatly affects survival. Gliomas have two parts: the solid part and the infiltrating part that coexists with normal, functioning brain tissue. Depending on the location, tumor resection is planned such that there is minimal impact on neurological function. However, due to the infiltrating nature of the tumor tissue, complete resection of the tumor is very difficult without impairing function. Hence, patient survival is very poor in any case.

Per the data we have considered for our study, there are a larger number of patients with tumors in the left and right posterior regions of the brain compared with other locations (►**Fig. 5**). There would have been some bias if the location was to be used in prediction, due to the uneven distribution of patient data. Hence, this was not considered as a reliable factor to predict patient survival.

►**Fig. 6** indicates the positive correlation between different factors and patient survival in green and negative correlation in brown. With respect to the second column in the correlation plot, the factors appearing green with positive correlation in the grid are used for training the survival prediction models along with the age of the patient. The correlation cutoff used in the study was 0.015. Age was considered an additional factor due to strong negative correlation with patient survival.

The values of correlation between the relevant factors and patient survival are summarized in ►**Table 1**. The accuracy obtained after training the machine learning models and evaluating them using the test set, with all three data partition ratios, is captured in ►**Table 2**.

Discussion

The different classifiers used yielded varying accuracies with the three data partitioning ratios, as indicated in ►**Table 2**.

Table 1 Correlation of relevant factors with patient survival

Features	Correlation with survival
Age	−0.363339491
Extent of resection	0.048592138
Sphericity	0.168645561
Surface area to volume ratio	0.059987922
10th percentile	0.046581783
Energy	0.047304826
Entropy	0.030986259
Kurtosis	0.056544171
Mean	0.105506960
Median	0.046930830
Skewness	0.107658714
Autocorrelation	0.016935041
Cluster shade	0.019752254
Contrast	0.073018352

Table 2 Accuracies obtained after evaluation with the test set

Data partition ratio Train set:test set	Accuracy of the models			
	Naive Bayes classifier	K-nearest neighbors	Random forest	Weighted subspace random forest
70:30	39.7%	50%	45.6%	51.5%
80:20	44.4%	64.4%	46.7%	55.6%
90:10	50%	50%	63.6%	50%

Note. The highest accuracy obtained on testing the models with three data partition ratios was 64.4%. This was achieved using the K-nearest neighbors model with an 80:20 training to test set ratio.

Using the 70:30 training to test set ratio, the highest accuracy was obtained using the weighted subspace random forest model (51.5%). Following which, the models achieving highest accuracies for the 80:20 and 90:10 ratio were the KNN and random forest models, with an accuracy of 64.4 and 63.6%, respectively. The prediction model designed by Sun et al[14] achieved an accuracy of 61% for three-class classification. We suggest the KNN model achieving 64.4% accuracy with an 80:20 data partition for future studies.

Although the accuracies obtained by these models are on par with prior work done in survival prediction, there are some limitations to them. First, our dataset for machine learning was limited in size with only 229 data points that could be considered for analysis. If applied to a larger dataset, the models could have been trained more precisely. Second, there were very little clinical data about the patients available in the BraTS dataset. Only the patient age and extent of resection were provided in the data. If other factors such as gender, family history, genetics, occupation, and prior radiation exposure were included, better risk factor identification would have been possible and correlation with feature data could have been investigated. Furthermore, we have only considered 15 factors contributing to patient survival, per the correlation values obtained, from the 79 features extracted due to limited computing power. A larger number of features could be extracted, which includes the 2D shape features, gray level run length matrix, gray level size zone matrix, neighboring gray tone difference matrix, and gray level dependence matrix. This would help in identifying more factors that could be correlated with patient survival. Lastly, we have only used basic classification models to predict patient survival in our study. The use of deep learning techniques in feature selection and prediction could help in refining the results obtained.

The results obtained in our study for the different models show promise of developing more accurate and robust models to predict patient survival. By increasing the sample size and refining the machine learning models, including other factors such as gender, family history, genetics, occupation, and prior radiation exposure, a more accurate prediction could be obtained, which would in turn have more immediate clinical application.

Author Contributions
The idea was conceived by all the authors. S.K. and M.M. performed feature extraction using the 3D Slicer software. S.S. performed the data visualization, coding for the four machine learning models, and implementation of survival prediction using R programming. The manuscript was prepared by all the authors. J.C. supervised the study and also reviewed and finalized the manuscript.

References

1 Ahola M, Uusitalo S, Palva L, Sepponen R. Scaling the magnetic resonance imaging through design research. In: Ahram T, Taiar R, eds. Human Interaction, Emerging Technologies and Future Systems V. IHIET 2021. Lecture Notes in Networks and Systems. Vol. 319. Cham: Springer; 2022

2 R Core Team. R: A Language and Environment for Statistical Computing. ViennaR Foundation for Statistical Computing 2021. Accessed on December 9, 2022 at: https://www.R-project.org/

3 Menze BH, Jakab A, Bauer S, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Trans Med Imaging 2015;34(10):1993–2024

4 Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 2012;30(09):1323–1341

5 Donahue MJ, Blakeley JO, Zhou J, Pomper MG, Laterra J, van Zijl PC. Evaluation of human brain tumor heterogeneity using multiple T1-based MRI signal weighting approaches. Magn Reson Med 2008;59(02):336–344

6 Rathi VP, Palani S. Brain tumor MRI image classification with feature selection and extraction using linear discriminant analysis. 2012. arXiv:1208.2128

7 van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017;77(21):e104–e107

8 Wickham H. Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag; 2016

9 Neuwirth E. RColorBrewer: ColorBrewer Palettes. R package version 1.1–3. 2022. Accessed on December 9, 2022 at: https://CRAN.R-project.org/package=RColorBrewer

10 Sun L, Zhang S, Chen H, Luo L. Brain tumor segmentation and survival prediction using multimodal MRI scans with deep learning. Front Neurosci 2019;13:810

11 Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. J Open Source Softw 2019;4(43):1686

12 Kuhn M. caret: Classification and Regression Training. R package version 6.0–93. 2022. Accessed on December 9, 2022 at: https://CRAN.R-project.org/package=caret

13 Liaw A, Wiener M. Classification and regression by randomForest. R News 2002;2(03):18–22

14 Zhao H, Williams GJ, Huang JZ. Wsrf: an R package for classification with scalable weighted subspace random forests. J Stat Softw 2017;77(03):1–30