MRI Based High-Grade and Low-Grade Brain Gliomas Classification by Using Video Vision Transformers

Muhammed Oguz Tas and Hasan Serhan Yavuz
Electrical-Electronics Engineering
Eskisehir Osmangazi University
Eskisehir, Turkiye

Muhammed Alperen Tas
Doctor of Medicine
Kutahya Health Sciences University
Kutahya, Turkiye

***Abstract*—Gliomas, which constitute the majority of brain tumors, are among the most common and aggressive brain tumors of the central nervous system. Since gliomas are often difficult to treat and significantly affect the quality of life and survival of patients, the successful classification of glioma grade is crucial. In this study, we used the Video Vision Transformer (ViViT) approach, known for its success in object detection problems, to classify gliomas as high-grade gliomas (HGG) or low-grade gliomas (LGG) using MR images obtained from a well-known brain tumor database, BraTS-2019. Instead of utilizing all MR images, the model was trained with fewer images to efficiently represent the dataset. The proposed method performs end-to-end feature extraction and classification by eliminating the need for preprocessing steps such as merging different sequences or segmenting the tumor, which typically yields to extra processing load. In the study, we also utilized a novel loss function to address false positive (FP) and false negative (FN) situations in binary classification problems, and the obtained results were evaluated with the AUC metric, which is widely used for imbalanced data sets. To the best of our knowledge, this is the first study to apply the ViViT architecture for the classification of glioma grades. We obtained promising results (AUC: 0.9410) which suggest significant potential for future research in this area.**

Keywords—glioma grade classification; HGG-LGG; video vision transformer (ViViT); PEN-BCE; artificial intelligence

I. INTRODUCTION

Brain tumors are a heterogeneous group of tumors that originate from various cell groups in the central nervous system or can occur after metastasis of tumors in different parts of the body. According to the report published by the American Cancer Society, it was estimated that more than 24,000 individuals would be diagnosed with brain tumors and more than 18,000 individuals would die from brain tumors in the USA in 2023 [1]. Brain tumors are categorized into diverse types and have different grades depending on the origin of the cell and the aggressiveness of the tumor. Gliomas, which constitute the majority of brain tumors, are classified as low-grade (I-II) and high-grade (III-IV) based on pathological evaluation. High-grade gliomas (HGG) are a rapidly progressing tumors that requires urgent treatment and leads to the patient's death within a short period, whereas low-grade gliomas (LGG) are a slowly growing tumor associated with a longer survival time [2, 3].

Although various medical imaging methods are used in the diagnosis and evaluation of tumors, Magnetic Resonance (MR) imaging is preferred as the most effective technique due to its ability to distinguish different tissues based on contrast levels, utilizing the different sequences it offers (T1, T2, FLAIR). Moreover, MR imaging aids in accurately determining the location, shape, and degree of tumors by displaying brain tissue in high resolution and providing detailed information about the brain's anatomical and functional features.

Accurate and rapid detection of gliomas is critical, as different grades determine the tumor's biological behavior, growth rate, and response to treatment. Clinicians have the expertise to determine tumor grade; however, they must examine large sets of MR images and interpret tissue features or patterns to reach a definitive conclusion. This process requires clinicians to spend a significant amount of time analyzing images as it depends on their subjective evaluation.

Developing technology and artificial intelligence-based methods have enabled numerous studies to be conducted on topics such as tumor diagnosis, classification, and segmentation by automatically interpreting large data sets. There are also many studies in literature on the classification of glioma grades. Some of these consist of machine learning (ML) based methods that extract patterns in the data and then make predictions [4-8]. However, since the manual feature extraction process requires domain expertise in ML-based studies, many neural network-based methods have also been applied to automatically extract features of the data at different levels and model the spatiotemporal relationships in the input images [9-16]. Some studies train the classification model using different MR image sequences [11, 12], others combine various feature groups [13], and some segment the tumor region before classifying the glioma grade [11, 14]. These approaches often result in an extended modeling process and require more processing power.

In this study, we implemented a transformer-based architecture for the classification of high-grade and low-grade gliomas using data from a well-known brain tumor database. By

979-8-3315-2981-9/24/$31.00 ©2024 IEEE

employing a slice selection method that centers on the slice with the largest tumor area, patient data is represented with fewer images. This approach yielded results comparable to those obtained in studies requiring high processing power as available in the literature. The rest of the paper is organized as follows: the proposed method is explained in the second section, experimental studies are given in the third section, and the results are interpreted in the last section.

II. Methodology

In this study, we utilized a transformer-based architecture to classify glioma grades, representing MR images from a renowned brain tumor database with a reduced number of samples. The general structure of the method applied in the study is summarized in Figure 1.

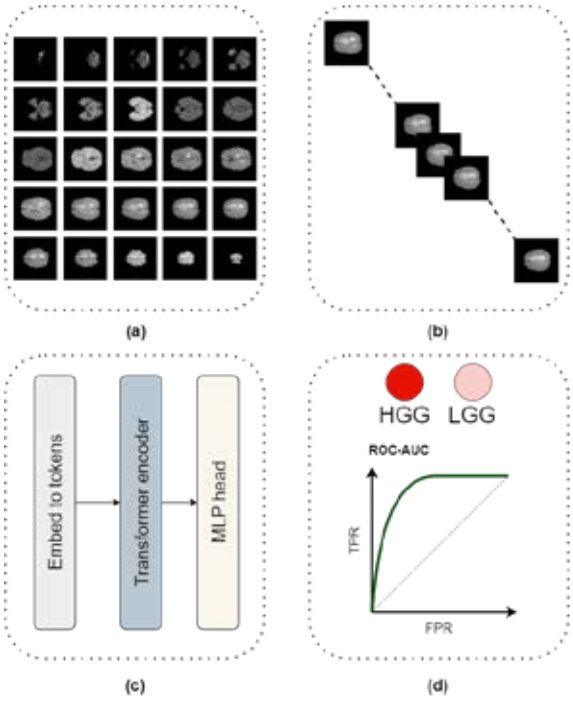


Fig. 1. Block diagram of the proposed method. (a) Original MR images in the database. (b) Selection of MR slices. (c) Training of the ViViT model. (d) Classification of HGG-LGG and evaluation of the model with ROC-AUC

A. Brain Tumor Segmentation 2019 (BraTS-2019) Database

The Multimodal Brain Tumor Segmentation Challenge (BraTS) 2019 database is a database containing information on 259 HGG and 76 LGG patients [17-21]. The csv file in the database contains information on age, survival time and resection status of the patients. The MR images in the database are stored in NIFTI format, offering an integrative view of different components of tumor biology through various MR sequences. For example, T1-weighted images are sensitive to nuclei and necrotic tumor nuclei, while FLAIR and T2 emphasize peritumoral edema [17]. The obtained images were manually segmented by different experienced neuroradiologists following the same annotation protocol. Each sample in the database contains 155 slices with dimensions of 240×240 pixels.

B. Slice Selection and Adapting Input Images to the Model

BraTS-2019 database contains 155 MR slices from all sequences for each patient. While performing the tests, the input images in the FLAIR sequence are fed to the network as 240×240 gray level and 11-slice (240×240×11). Using more slices can increase performance by allowing the model to learn more details. However, since increasing the number of slices will elevate the computational costs and data processing times, an 11-slice structure was experimentally applied in this study for optimum performance. To adjust the 11-slice input images, the slice with the largest tumor area among the slices containing the tumor regions, along with the five neighboring slices before and five after this slice, are selected. All input images are normalized between 0 and 1 to enhance the network's performance at last.

C. Video Vision Transformer (ViViT) Architecture

The success of attention-based models in natural language processing has recently inspired studies on approaches to integrate transformer structures into convolutional neural networks (CNNs) in computer vision. While transformers used with convolutional mechanisms do not significantly affect the classification success, the Vision Transformer (ViT) study [22], which proposes image classification with pure transformers, has demonstrated that it uses much fewer training resources than CNN structures and achieve more success. With this approach making a breakthrough in image classification, the Video Vision Transformer (ViViT) model, which uses pure transformers for video classification and whose architecture is given in Figure 2, has been proposed by Arnab et. al [23]. In this study, the classification of glioma degrees was performed using ViViT.

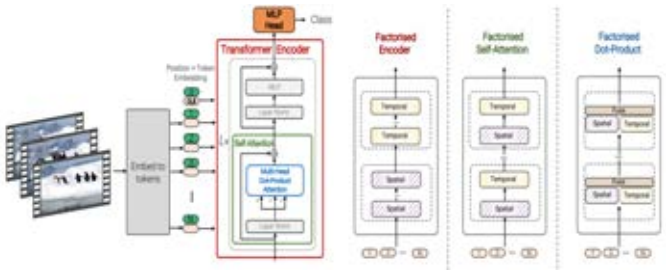


Fig. 2. ViViT architecture [23]

III. Experimental Results

To classify glioma grades, which is crucial for determining treatment plans and providing prognosis estimates, MR images from the BraTS-2019 database were utilized, and training was conducted using the ViViT architecture.

In the experiments, the database was partitioned into training and test sets with an 85%-15% split. Within the dataset of 335 observations, the proportions of patients labeled as HGG and LGG were 77.3% and 22.7%, respectively.

A. ViViT-Based Classification Results

During model training, a patch size of 24×24×3 was utilized for Tubelet Embedding in the ViViT architecture. Layer_norm_eps coefficient was set as 10^{-6}, projection_dim coefficient was set as 768, num_heads and num_layers coefficients were set as 8. Batch size was set as 16 in the model, and it was trained with Adam optimization method for 300

epochs. The initial learning rate of the model was set as 10^{-5} and if the validation loss remained constant for 25 epochs, learning rate was reduced by 0.9. In addition, Early Stopping was applied to prevent over-fitting of the model. Sigmoid was defined as the activation function in the output layer in model training and binary cross entropy was defined as the loss function. 5-fold cross-validation was performed in model training and model performance was evaluated with AUC metric. In the tests performed, the average test loss was 0.3334 and the average test AUC was 0.9322, and the ROC of the test data in each fold are given in Figure 3.

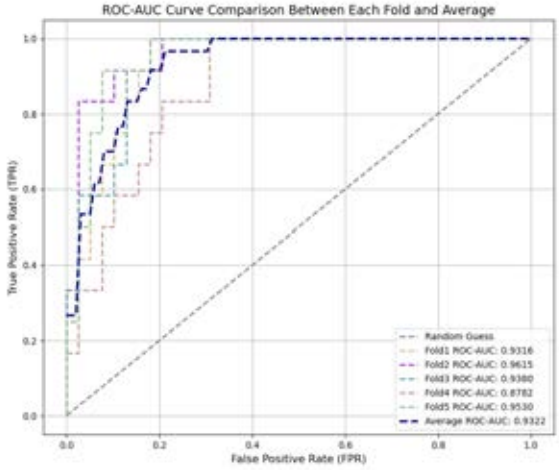


Fig. 3. HGG-LGG classification results with ViViT model

B. *Effect of Slice Selection Technique*

The selection of slices to represent the database directly affects the model's performance. In the study, the selection of 11-slice input images was made in the structure where the largest tumor slice or the median slice was in the center. The obtained results are given in Table 1.

TABLE I. EFFECT OF SLICE SELECTION

Slice Selection	Test Loss	Test AUC
Median slice selection	0.3658	0.9199
Largest slice selection	0.3334	0.9322

C. *Effect of the Loss Function*

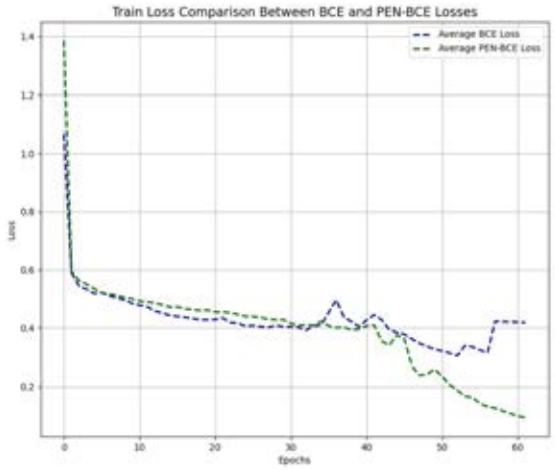
The binary cross entropy (BCE) loss function, which is widely used for binary classification problems, may not be sufficient to directly handle imbalanced datasets because it does not directly account for the false positive (FP) and false negative (FN) cases. However, it has been shown that the penalized BCE (PEN-BCE) loss function, which has four new hyper-parameters (α, β, p_{FN}, p_{FP}) that take into account the FP and FN cases in addition to the BCE loss, gives more successful results [24].

The added penalty parameters α and β represent the FN and FP weights, respectively, and the parameters p_{FN} and p_{FP} represent the FN and FP probability threshold values, respectively. Within the scope of the study, tests were performed with various experimentally defined hyper-parameter combinations as given in Table 2. As seen in Table 2, the highest performance was obtained for the values α=1.0, β=5.0, p_{FN}=0.50, p_{FP}=0.30.

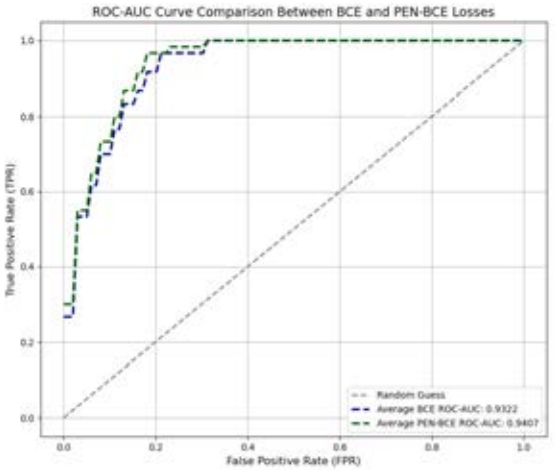
TABLE II. EFFECT OF PEN-BCE LOSS FUNCTION PARAMETERS

Loss Parameters	Test Loss	Test AUC
α=5.0, β=1.0, p_{FN}=0.40, p_{FP}=0.50	0.3299	0.9335
α=5.0, β=1.0, p_{FN}=0.60, p_{FP}=0.50	0.3491	0.9263
α=5.0, β=1.0, p_{FN}=0.70, p_{FP}=0.50	0.3748	0.9169
α=1.0, β=5.0, p_{FN}=0.50, p_{FP}=0.20	0.3781	0.9246
α=1.0, β=5.0, p_{FN}=0.50, p_{FP}=0.30	0.3723	0.9410
α=1.0, β=5.0, p_{FN}=0.50, p_{FP}=0.40	0.3583	0.9246
α=5.0, β=5.0, p_{FN}=0.50, p_{FP}=0.50	0.3510	0.9250
α =5.0, β=5.0, p_{FN}=0.40, p_{FP}=0.50	0.3344	0.9308
α =5.0, β=5.0, p_{FN}=0.50, p_{FP}=0.30	0.3676	0.9308

The comparison of PEN-BCE and BCE results in terms of loss and ROC curves are given in Figure 4.



(a)



(b)

Fig. 4. Performance comparison of BCE and PEN-BCE loss functions. (a) Test loss. (b) Test ROC-AUC

IV. Conclusions

Gliomas, which are among the most common and aggressive brain tumors of the central nervous system, are often difficult to treat and significantly affect the quality of life and survival of patients. Therefore, accurate determination of the glioma grade is crucial for the development of effective treatment strategies and the prognosis of patients. Due to the subjective assessments of clinicians and the slow examination of large image sets in determining the glioma degree, artificial intelligence-based techniques that automatically interpret tissue features or patterns in large image sets have recently been used.

In this study, we utilized a ViViT model trained with fewer images to perform end-to-end, automatic, and low-computation classification of glioma grades (HGG and LGG) using single-sequence MR images from the well-known BraTS-2019 brain tumor database. The proposed method not only eliminated the need for manual feature extraction as required in many ML-based studies but also bypassed preprocessing steps such as merging different sequences, fusing various data types, and segmenting the tumor, which typically add computational load.

To the best of the authors' knowledge, this study represents the first application of the ViViT architecture for glioma grade classification. Instead of the commonly used accuracy metric, this study employed the AUC metric for classifying glioma grades, as it is better suited for imbalanced datasets. The accuracy metric can misleadingly indicate high performance for models that predict the majority class well but fail to accurately predict the minority class.

Although the median slice selection method provides a consistent reference point by taking the slice located at the exact center of the MR images, more successful results were achieved using the largest slice selection method. This approach was more effective in cases where the tumor was not apparent or was small, providing sufficient information for classification. Additionally, the PEN-BCE loss function was observed to enhance classification performance at specific parameters and effectively handled false positive and false negative cases.

References

[1] Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. Cancer statistics, 2023. Ca Cancer J Clin, 2023, 73(1), 17-48.

[2] Tonarelli, L. Magnetic resonance imaging of brain tumor. CEwebsource.com, 2013, 1-12.

[3] Das, S., Bose, S., Nayak, G. K., Satapathy, S. C., & Saxena, S. Brain tumor segmentation and overall survival period prediction in glioblastoma multiforme using radiomic features. Concurrency and Computation: Practice and Experience, 2022, 34(20), e6501.

[4] Cho, H. H., & Park, H. (2017). Classification of low-grade and high-grade glioma using multi-modal image radiomics features. In 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) Jesu Island, Korea, 11-15 June 2017, 3081-3084.

[5] Ain, Q. U., Duaa, I., Haroon, K., Amin, F., & ur Rehman, M. Z. (2021). Mri based glioma detection and classification into low-grade and high-grade gliomas. In 15th International Conference on Open Source Systems and Technologies (ICOSST), Lahore, Pakistan, 15-16 December 2021, 1-5.

[6] Cui, G., Jeong, J., Press, B., Lei, Y., Shu, H. K., Liu, T., ... & Yang, X. Machine-learning-based classification of lower-grade gliomas and high-grade gliomas using radiomic features in multi-parametric MRI. arXiv preprint, 2019, arXiv:1911.10145.

[7] Cho, H. H., Lee, S. H., Kim, J., & Park, H. Classification of the glioma grading using radiomics analysis. PeerJ, 2018, 6, e5982.

[8] Kumar, R., Gupta, A., Arora, H. S., Pandian, G. N., & Raman, B. CGHF: A computational decision support system for glioma classification using hybrid radiomics-and stationary wavelet-based features. IEEE Access, 2020, 8, 79440-79458.

[9] Mzoughi, H., Njeh, I., Wali, A., Slima, M. B., BenHamida, A., Mhiri, C., & Mahfoudhe, K. B. Deep multi-scale 3D convolutional neural network (CNN) for MRI gliomas brain tumor classification. Journal of Digital Imaging, 2020, 33, 903-915.

[10] Khazaee, Z., Langarizadeh, M., & Ahmadabadi, M. E. S. Developing an artificial intelligence model for tumor grading and classification, based on MRI sequences of human brain gliomas. International Journal of Cancer Management, 2022, 15(1).

[11] Özkaya, Ç., & Sağiroğlu, Ş. Glioma grade classification using CNNs and segmentation with an adaptive approach using histogram features in brain MRIs. IEEE Access, 2023, 11, 52275-52287.

[12] Ali, M. B., Gu, I. Y. H., & Jakola, A. S. Multi-stream convolutional autoencoder and 2D generative adversarial network for glioma classification. In Computer Analysis of Images and Patterns: 18th International Conference, CAIP 2019, Salerno, Italy, September 3–5, 2019, 234-245.

[13] Sivakumar, J., Rajesh Kannan, S., & Manic, K. S. Automated classification of brain tumors into LGG/HGG using concatenated deep and handcrafted features. In Frontiers of Artificial Intelligence in Medical Imaging, Bristol, UK, 2022 (pp. 7-1). Bristol, UK: IOP Publishing.

[14] Ge, C., Gu, I. Y. H., Jakola, A. S., & Yang, J. Deep learning and multi-sensor fusion for glioma classification using multistream 2D convolutional networks. In 40th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), Honolulu, Hawaii, USA, 17-21 July 2018, 5894-5897.

[15] Mehmood, Y., Bajwa, U. I., & Anwar, M. W. Brain tumor grade classification using multi-step pre-training. International Journal of Imaging Systems and Technology, 2024, e23008.

[16] Zhuge, Y., Ning, H., Mathen, P., Cheng, J. Y., Krauze, A. V., Camphausen, K., & Miller, R. W. Automated glioma grading on conventional MRI images using deep convolutional neural networks. Medical physics, 2020, 47(7), 3044-3053.

[17] H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 2015, 34(10), 1993-2024, DOI: 10.1109/TMI.2014.2377694

[18] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", Nature Scientific Data, 2017, 4:170117, DOI: 10.1038/sdata.2017.117

[19] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge", arXiv preprint, 2018, arXiv:1811.02629 ()

[20] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, et al., "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-GBM collection", The Cancer Imaging Archive, 2017. DOI: 10.7937/K9/TCIA.2017.KLXWJJ1Q

[21] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, et al., "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-LGG collection", The Cancer Imaging Archive, 2017. DOI: 10.7937/K9/TCIA.2017.GJQ7R0EF

[22] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint, 2020, arXiv:2010.11929.

[23] Arnab, A., Dehghani, M., Heigold, G., Sun, C., Lučić, M., & Schmid, C. Vivit: A video vision transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, Canada, 11-17 October 2021, 6836-6846.

[24] Tas, M.O., & Yavuz, H.S. Enhancing Lung Cancer Survival Prediction: 3D CNN Analysis of CT Images Using Novel GTV1-SliceNum Feature and PEN-BCE Loss Function. Diagnostics, 2024, 14(12), 1309.