IBRO Neuroscience Reports 13 (2022) 523–532

Contents lists available at ScienceDirect

IBRO Neuroscience Reports

journal homepage: www.sciencedirect.com/journal/IBRO-Neuroscience-Reports

Research Paper

A deep learning framework integrating MRI image preprocessing methods for brain tumor segmentation and classification

Khiet Dang [a,b], Toi Vo [a,b], Lua Ngo [a,b,*], Huong Ha [a,b,*]

[a] School of Biomedical Engineering, International University, Vietnam National University – Ho Chi Minh City, Ho Chi Minh City, Viet Nam
[b] Vietnam National University – Ho Chi Minh City, Ho Chi Minh City, Viet Nam

ARTICLE INFO

Keywords:
Glioma grading
Deep learning
Three-dimensional
Data augmentation
Segmentation

ABSTRACT

Glioma grading is critical in treatment planning and prognosis. This study aims to address this issue through MRI-based classification to develop an accurate model for glioma diagnosis. Here, we employed a deep learning pipeline with three essential steps: (1) MRI images were segmented using preprocessing approaches and UNet architecture, (2) brain tumor regions were extracted using segmentation, then (3) high-grade gliomas and low-grade gliomas were classified using the VGG and GoogleNet implementations. Among the additional pre-processing techniques used in conjunction with the segmentation task, the combination of data augmentation and Window Setting Optimization was found to be the most effective tool, resulting in the Dice coefficient of 0.82, 0.91, and 0.72 for enhancing tumor, whole tumor, and tumor core, respectively. While most of the proposed models achieve comparable accuracies of about 93 % on the testing dataset, the pipeline of VGG combined with UNet segmentation obtains the highest accuracy of 97.44 %. In conclusion, the presented architecture illustrates a realistic model for detecting gliomas; moreover, it emphasizes the significance of data augmentation and segmentation in improving model performance.

1. Introduction

Glioma is the most frequent brain tumor developing from glial cells. The World Health Organization has classified glioma into four grades, including circumscribed astrocytoma (grade I), diffuse astrocytoma (grade II), anaplastic astrocytoma (grade III), and glioblastoma (GBM - grade IV) (Tabibkhooei et al., 2020). Gliomas of grades I, II, and III are classified as low-grade gliomas (LGG), depending on their severity. More specifically, grade I may be treated surgically with adjuvant therapies. Grade II has recurrence risks with a median survival of five to 15 years, with a survival rate of up to 81.6 %, and grade III has a lower survival rate of 57.6 % (Ertosun and Rubin, 2015). Regarding high-grade gliomas (HGG), GBM (Grade IV) is the most common malignant brain tumor among adults, with an average survival time of 12–15 months. In other words, the five-year survival rate of patients with GBM is significantly low, approximately 5.6 % (Tabibkhooei et al., 2020; Decuyper et al., 2021). As a result, the classification between LGG and HGG is critical in treatment management and prognostication (Zlochower et al., 2020).

Despite advances in glioma treatment strategies that have increased the median survival rates of patients, clinical management of these patients is currently confronted with a lack of appropriate tumor-grading technologies (Tabibkhooei et al., 2020). The current standard methods are biopsies and surgical resections, which still concern morbidity and mortality. Furthermore, the biopsy is susceptible to sampling error, and the results are contingent on the experience of neuropathologists (Guzmán-De-Villoria et al., 2014). As a result, Magnetic Resonance Imaging (MRI), a safe, non-radiation, and non-invasive method that provides vital information about brain tumors should be emphasized. The fundamental reason is that MRI can precisely identify the neuroanatomic structures using its improved contrast discrimination and ability to record images in several planes (Matsumoto et al., 2010; Maravilla and Sory, 1986). Furthermore, tumor molecular characteristics such as mutational status can be revealed by the MRI scans in the early process of oncogenesis, allowing clinicians to identify a prognosis marker for patients, develop a treatment plan, and support therapeutic interventions (Gore et al., 2021; Fathi Kazerooni et al., 2020).

However, because of the irregular form and varied composition of gliomas, MRI diagnosis remains a challenge for clinicians (Gore et al., 2021). First, although advanced techniques such as magnetic resonance spectroscopy have significantly improved the prediction performance,

* Corresponding authors at: School of Biomedical Engineering, International University, Vietnam National University – Ho Chi Minh City, Ho Chi Minh City, Viet Nam.
E-mail addresses: ntlua@hcmiu.edu.vn (L. Ngo), htthuong@hcmiu.edu.vn (H. Ha).

https://doi.org/10.1016/j.ibneur.2022.10.014
Received 10 June 2022; Accepted 31 October 2022
Available online 7 November 2022

they are still associated with marked error rates (Pouratian et al., 2007). Secondly, manual MRI diagnosis is inefficient because physicians cannot handle millions of images in a fair amount of time. As a result, MRI-coupled artificial intelligence algorithms, capable of recognizing image patterns and accurately identifying genetic markers of glioma, could enhance efficiency and provide more valuable diagnostic information from routine radiology work (Zlochower et al., 2020; Ahuja, 2019).

Several attempts have been applied before to integrate deep learning into this procedure. Decuyper et al. (2021) trained a pipeline including UNet for segmentation and ResNet for classification, attaining an accuracy rate of 90 % and a sensitivity rate of roughly 93.48 %. Besides, other investigations, including Yang et al. (2018), Ahammed Muneer et al. (2019), Saba et al. (2020), and Cheng et al. (2022), should also be considered due to their outstanding accuracies ranging from 94.5 % to 99.82 %. Furthermore, several studies have also reported their achievements while combining segmentation into classification procedures, such as Decuyper et al. (2021) and Ahammed Muneer et al. (2019), with accuracies of 90 % and 92.86 %, respectively.

However, there are still few articles comparing models with and without segmentation to evaluate the effects of this process, and no studies can claim the influence when combining the segmentation results directly into the classification models with a piece of clear statistical evidence. Therefore, this work proposes combining segmentation findings with several architectures in this sector, such as OM-Net (Zhou et al., 2020) and Segtran (Li et al., 2022). However, UNet and its version, such as nnUNet by Isensee et al. (2021), is still the most commonly utilized model in such a procedure (Decuyper et al., 2021; Baid et al., 2020). As a result, UNet is integrated into this study to confirm the hypothesis of its influence on glioma diagnosis.

Identifying tumor regions is still a problem even in the four MRI modalities, which is required to change the focal visibility and classify different tumor regions efficiently. As a result, several preprocessing methods are utilized to enhance information in these images for the segmentation steps. The first preprocessing technique is gamma correction, which is an intensity invariance approach to increase the generalizability of brain tumor segmentation tasks. By adjusting the overall intensity and contrast of input images, this method can address the issue of enhancing the contrasted structures for tumor-region recognition (Huang et al., 2021a; Tai et al., 2021). Secondly, the window setting optimization (WSO) module, which has been tried in other fields of CT images, such as in identifying acute ischemic stroke and abdominal angiography, is utilized. Such implementation would be hypothesized to enhance the abnormal tumor tissues on MRI images (Arsava et al., 2014; Doerner et al., 2018). The final technique mentioned is data augmentation, a typical approach in deep learning networks that would help generate more data, particularly affine transformation (Nalepa et al., 2019).

An interpretable explanation of the working mechanisms associated with the CNNs pipeline is essential before such an approach can be integrated into clinical settings because it can establish user trust and confidence (Selvaraju et al., 2020). Such understanding is also helpful for researchers to explain the failures of CNN models and focus their efforts on the most successful research routes (Selvaraju et al., 2020). Other similar studies have also tried to apply this method to interpret their models to aid in a better understanding of the imaging characteristics of gliomas (Rajapaksa and Khalvati, 2021; Wei et al., 2022). As a result, this study proposes the application of Gradient-weighted Class Activation Mapping (Grad-CAM), an effective visualization tool for 3D CNN architectures, to increase the interoperability for components contributing to glioma grading (Huang et al., 2019).

This project uses three preprocessing approaches: gamma correction, WSO module, and data augmentation (Tai et al., 2021; Nalepa et al., 2019; Lee et al., 2018). In terms of classification, we propose to grade gliomas using deep learning architectures, with VGG and GoogleNet being chosen because they both demonstrated significant performance in previous studies (Yang et al., 2018; Ahammed Muneer et al., 2019). In addition, we aim to compare the performances between models with and without segmentation to assess the importance of the applied method in the classification task. Our main objective is to explore (1) the effects of WSO and gamma correction on MRI images, (2) the role of segmentation in the classification model, and (3) the working mechanism of CNN through Grad-CAM.

2. Materials and methods

Fig. 1 depicts the overview of this study. First, the original MRI images are preprocessed using three different methods: gamma correction, WSO module, and data augmentation. UNet is used for segmentation to distinguish between the non-tumor regions, necrosis and non-enhancing tumor core, peritumoral edema, and enhancing tumor core. Next, the two classification algorithms, including VGG and GoogleNet, are applied using the two input types after passing through the over-sampling procedure: (a) the cropped-background images and (b) the cropped-non-tumor-region images resulting from the segmented outputs from the UNet. All of the hyperparameters are optimized to improve performance. Finally, the proposed models are assessed using the Dice coefficient, Precision, Recall, and Accuracy performance measures to determine the most efficient segmentation and classification method. Grad-CAM is also employed as salient maps for visualization.

2.1. Dataset description

In this study, the dataset in NifTi file extension (.nii.gz) from Multimodal Brain Tumor Segmentation (BraTS) Challenge 2019 (Bakas et al., 2017a, 2017b, 2017c, 2019; Menze et al., 2015) is used, which consists of 335 clinically routine collected pre-operative MRI images in the two classes of HGG (259 images) and LGG (76 images). To be more specific, there are four different image modalities, including (a) native (T1), (b) contrast-enhancing T1 (T1CE), (c) T2-weighted, and (d) Fluid Attenuated Inversion Recovery (FLAIR) volumes.

All of the tumor masks were labeled manually by experienced neuroradiologists, including non-tumor region (label 0), necrotic and non-enhancing tumor core (NCR/NET – label 1), peritumoral edema (ED – label 2), and GD-enhancing tumor core (ET – label 4) (Bakas et al., 2017a, 2017b, 2017c, 2019; Menze et al., 2015). An experienced local

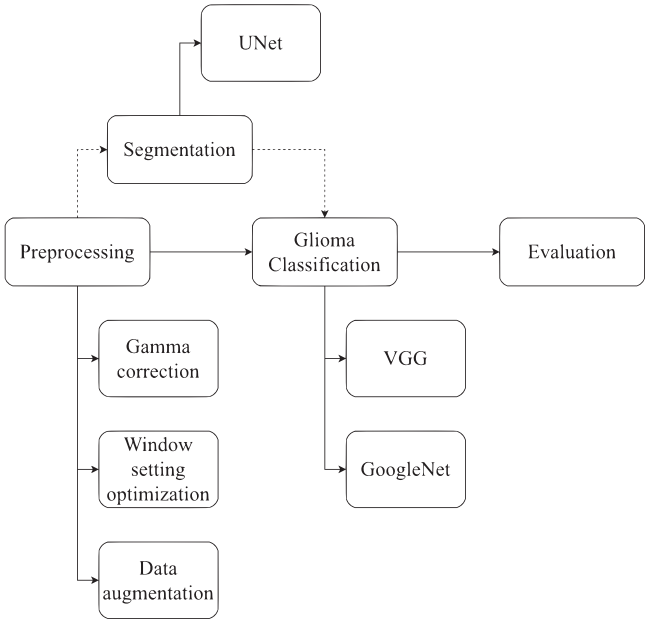

Fig. 1. Study design.

radiologist was also invited to double-check the labels. The data has already been preprocessed, including co-registering to the same anatomical template, interpolating to the identical resolution (1 mm^3), and being skull-stripped.

2.2. Data preprocessing

In this study, background removal is performed before dataset splitting with the ratio of 6:2:2 for training, validation, and testing set, respectively. The distribution of tumor regions is mentioned in Table 1 for each set. Subsequently, these images are preprocessed using gamma correction, WSO, and data augmentation.

Gamma correction, also known as gamma distortion, is a powerful tool of transformation that can be used to enhance images or videos using the exponent "gamma" (γ) (Amiri and Hassanpour, 2012). In this study, each of the MRI modalities would be adjusted by the different values of γ, except the FLAIR images, as they are clear enough to recognize the peritumoral edema with the hyper-intense regions (Fathi Kazerooni et al., 2020). Meanwhile, the other modalities, including T1, T1CE, and T2, would be significantly enhanced when using the gamma values of 2.5, 2.9, and 3.2, respectively, chosen as the most effective ones in improving Dice coefficients in segmentation tasks.

WSO is a module that consists of a convolutional layer with $1 \times 1 \times 1$ filters and a customized activation function. In this study, the intensity range would be set from 0 to U and using the two values of window widths (WW) and window levels (WL), the customized activation functions are determined based on the linear functions thanks to the following equations (Lee et al., 2018):

$$F_{lin}(x) = max((Wx+b, U) \quad , 0)$$

$$\text{In which} W = \frac{U}{WW} \text{and} b = -\frac{U}{WW}\left(WL - \frac{WW}{2}\right).$$

Data augmentation is a common strategy for promoting deep neural network performance, and in most cases, it is regarded as an implicit regularization. It is critical when the high-quality labeled data quantity is limited, and gathering new examples is expensive and time-consuming. As a result, it should be considered a prevalent issue in medical analysis, particularly brain tumor segmentation and classification (Nalepa et al., 2019).

The affine operations are used in this work as augmentation procedures because this method is significantly popular and easy to use in preprocessing images (Nalepa et al., 2019). However, only rotation is applied because the others, such as shifting, frequently lose information after cropping the background. The rotated angles ranging from 0 to 30 degrees are chosen randomly along one of the three dimensions of the 3D pictures using the SciPy library. This technique is implemented in the training, validation, and testing sets of the segmentation phase; however, in the classification one, it would not be applied to ensure the number of samples is equal in all test cases.

2.3. Tumor segmentation using UNet

In this study, UNet is employed for glioma segmentation because of its popularity in other studies (Decuyper et al., 2021; Baid et al., 2020). To be more specific, UNet was developed in 2015 by Olaf Ronneberger et al. to improve an architecture that could function with a few training samples and produce higher-performing segmentations. The UNet architecture is shown in Fig. 2, consisting of a contracting path (left side) and an expansive path (right side). First, the former contains the typical convolutional layers, followed by the Leaky ReLU activation function. In each down-sampling process, instance normalization and max-pooling layers are implemented. The reason for using instance normalization layers is that the batch size in this network is too small, and it may be unstable with the under-fitting problem and less efficient model performance. Consequently, an instance normalization layer is required to address such an issue by dealing with each element in a batch (Yong et al., 2020).

Regarding the expansive path at the right side of the u-shape, the transposed convolutional layers appear as the role of up-sampling features before the concatenation of the up-sampled outputs and the copy of the corresponding feature maps from the contracting path. At the end of the model, the final convolutional layer and the softmax activation function for the multi-class classification is utilized. Besides, the learning rate would be cut in half during the training procedure if the validation loss value does not reduce after 50 epochs. The hybrid loss function is implemented as the following equation:

$$Loss \quad function = \alpha \times (Categorical \quad Cross \quad Entropy) + \beta \times (1 - J_{tumor-region})$$

$\alpha = 14$ and $\beta = 0.7$ are the two parameters chosen through the tuning process of the Keras Tuner. The average Jaccard coefficient of the tumor region ($J_{tumor\text{-}region}$) is defined as a custom function to minimize the influences from the non-tumor area:

$$J_{tumor-region} = \frac{TP_{tumor-region}}{TP_{tumor-region} + FP_{tumor-region} + FN_{tumor-region}}$$

In this study, the ROI of the images will be defined using two different methods. Firstly, as mentioned in the previous part, the whole brain would be cropped out of the pictures. Regarding the second approach, only the tumor regions of the images would be cropped out after the two-class segmentation procedure. Several MRI images with poor performances (Dice coefficients are smaller than 0.69) would be eliminated to ensure the performance of the classification architectures. Besides, the predicted results are also cropped out and put into the training dataset along with other modalities to experiment with whether such information helps boost performance in the classification model.

2.4. Glioma classification using VGG and GoogleNet

Before the glioma grading procedure, including VGG and GoogleNet architecture, the random oversampling technique is implemented to

Table 1
Distribution of tumor regions in the volume (cm^3) for the training, validation, and testing sets. NCR/NET: necrotic and non-enhancing tumor core; ED: peritumoral edema; ET: enhancing tumor core.

	Training set (cm^3)	Validation set (cm^3)	Testing set (cm^3)
NCR/NET	20.10 ± 26.34	22.73 ± 29.62	29.90 ± 41.84
ED	52.91 ± 39.93	62.38 ± 40.78	66.81 ± 41.16
ET	16.57 ± 17.24	16.89 ± 15.71	26.83 ± 23.88

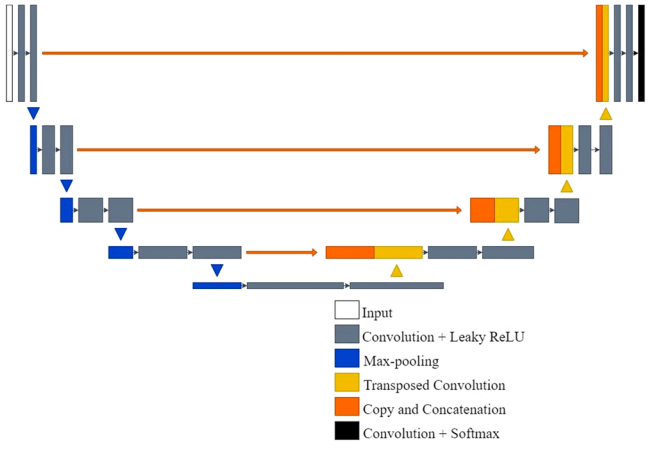

Fig. 2. Three-dimensional UNet architecture.

solve the imbalanced data classification problems (Ling and Li, n.d.). During the oversampling process, several MRI images in LGG class are randomly chosen to ensure the numbers of images in the LGG and HGG are the same with a ratio of 1:1.

Karen Simonyan and Andrew Zisserman pioneered using VGG to increase deep learning performance during the visual recognition procedure in 2014 (Simonyan and Zisserman, 2015). VGG was implemented with the addition of convolutional layers and a relatively tiny kernel size in all filters (Simonyan and Zisserman, 2015). Compared to the original model, the proposed one consists of fewer convolutional and max-pooling layers to optimize the computational costs when implementing the three-dimensional dataset. More specifically, the images are subsequently processed through a stack of convolutional layers where filters with a very narrow receptive field ($3 \times 3 \times 3$) are utilized. Besides, the two max-pooling layers following the convolutional ones with the pool size of $2 \times 2 \times 2$ and the strides (2, 2, 2) are applied to perform the down-sampling procedures.

Furthermore, the above-described stack is followed by a flattened layer and two densely-connected layers. The first contains 512 units, and the second has only two neurons corresponding to the two classes of the given problems. Three dropout layers are also applied to minimize overfitting. Finally, an SGD optimizer is utilized in this designed architecture to optimize the objective function (Fig. 3).

GoogleNet was initially suggested by Christian Szegedy et al. in 2015, along with the release of a state-of-the-art architecture known as Inception, to go deeper by building a network-in-network and using the convolutional construction blocks to discover the best architecture and replicate it spatially. The Inception block has three clusters linked to the preceding layer, followed by a concatenation layer, and finally covered by a $1 \times 1 \times 1$ convolutional layer (Szegedy et al., 2015). Instead of the nine Inception blocks as illustrated in the original model, however, the number of this architecture is reduced to optimize convolutional costs with the three-dimensional inputs. First, the cropped-background input (A7) has six Inception blocks, as described in Fig. 4. Like the other convolutional neuron networks, this network begins with successive layers, including the convolutional and max-pooling ones. Eventually, there is an auxiliary classifier consisting of a global average pooling layer, a flattened layer, and a densely-connected layer with the softmax activation. There are also several dropout layers to prevent overfitting. Secondly, regarding the cropped-non-tumor-region input (A8–A12), the quantity of Inception blocks is significantly reduced to three to optimize the model results.

2.5. Evaluation metrics

Regarding segmentation, the results would be evaluated by four metrics: Precision, Recall, Dice coefficient, and the Hausdorff distance.

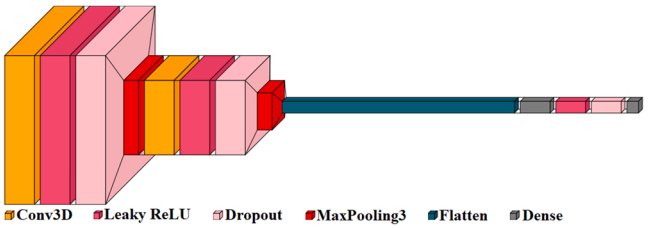

Fig. 3. VGG architecture.

First, precision (also known as positive predictive value) is calculated by the proportion of true positives found among the predicted positives.

$$Precision = \frac{True \quad Positive}{True \quad Positive + False \quad Positive}$$

Secondly, recall (also known as sensitivity) is the proportion of true positives found among the actual positive samples.

$$Recall = \frac{True \quad Positive}{True \quad Positive + False \quad Negative}$$

The Dice similarity coefficient, also known as the Sorensen-Dice index or Dice coefficient, is a statistical tool for determining the similarity of the two data sets. This index is one of the most frequently-used tools in segmentation algorithms of computer vision and can be utilized for various applications (Moore, n.d.).

$$Dice = \frac{2 \times True \quad Positive}{(True \quad Positive + False \quad Positive) + (True \quad Positve + False \quad Negative)}$$

The Hausdorff Distance (HD) is defined as the maximum distance between a point in one set and the most neighboring point in the other set, and it is calculated as:

$$HD(X, Y) = \max\{sup_{x \in X} \quad inf_{y \in Y} \quad d(x, \quad y), \quad sup_{y \in Y} \quad inf_{x \in X} \quad d(x, \quad y)\}$$

where d(x,y) represents the Euclidean distance between $x \in X$ and $y \in Y$. Besides, to avoid concerns with noisy predictions, the 95th percentile is utilized instead of the max operation, which we call Hausdorff95 (Noori et al., 2019).

Regarding classification, accuracy is applied to determine the difference between the result and the "true" value. It is calculated using the ratio of correct predictions across the sample space.

$$Accuracy = \frac{True \quad Positive + True \quad Negative}{Sample \quad size}$$

In the current research, 5-fold cross-validation is also implemented in the classification tasks. This procedure would be evaluated using the one-way analysis of variance (ANOVA), one of the most commonly used statistical tools in medical research, especially when there are more than two groups in the investigation. The mistake of alpha-level inflation, which raises the chance of Type 1 error (false positive) and is induced by numerous comparisons, necessitates ANOVA. To be more specific, ANOVA employs the statistic F to investigate the ratio of variances between and within groups. The primary focus of this analysis is on differences in variances. Several articles have illustrated that ANOVA solves mean difference problems by leveraging variance differences between and within groups (Ross and Willson, 2017; Kim, 2017).

Furthermore, Grad-CAM is utilized as an evaluation method to generate a visual explanation for the decisions made by the convolutional neural network; therefore, it would be helpful to make them more visible and explainable (Selvaraju et al., 2020). In the first phase of the procedure, a model would be created with the original input but two different outputs, including the last convolutional output value and the raw image predictions. In the next phase of the process, the gradients of the raw score to the ultimate convolutional output are computed to gain the heat map. After the heat-map matrix is obtained, it would be rescaled into the original image size and then overlaid onto the picture to visualize the model decision.

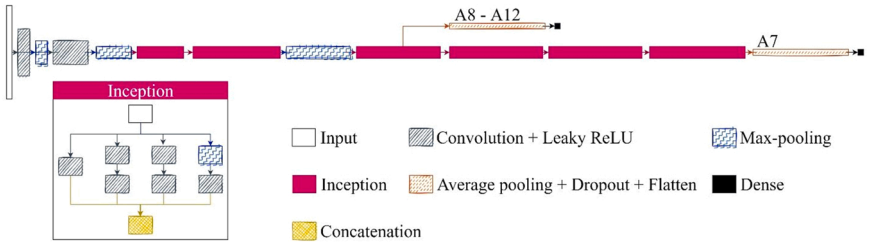

Fig. 4. GoogleNet architecture.

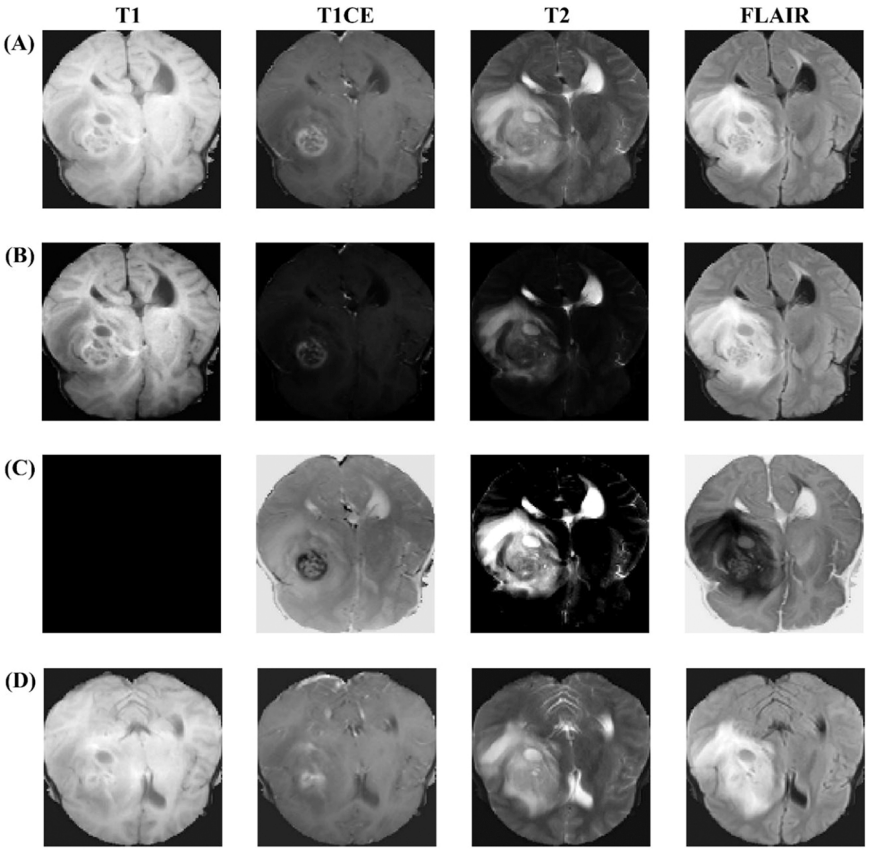

Fig. 5. MRI images (A) before preprocessing and after being implemented (B) Gamma correction, (C) WSO, and (C) Data augmentation.

3. Results

3.1. Preprocessing of MRI images using Gamma correction, WSO, and Data augmentation

The comparison of visual outputs illustrated in Fig. 5 is essential to comprehensively understand the roles of three different preprocessing methods, including gamma correction, WSO, and data augmentation. The first two methods successfully enhance the intensity difference between peritumoral edema and other brain regions in the MRI images; however, the second one (WSO) removes the information of one of the modalities thoroughly (Fig. 5A–C). Fig. 5D also shows the transformation of newly-generated MRI images through the three-dimensional rotation operation in data augmentation while their intensities do not experience significant changes.

Table 2
Result summarization of four-class segmentation task for the testing dataset using three-dimensional UNet. ET = Enhancing Tumor core, WT = Whole Tumor, TC = Tumor Core, DC = Dice coefficient, Pr = Precision, Rc = Recall, HD: Hausdorff distance. Network: N1: 3D; N2: 3D + GC; N3: 3D + WSO; N4: 3D + DA.

Network	ET				WT				TC			
	DC	Pr	Rc	HD	DC	Pr	Rc	HD	DC	Pr	Rc	HD
N1	0.78	0.86	0.72	12.8	**0.91**	0.94	0.88	2.2	0.69	0.59	0.84	7.8
N2	0.80	0.86	0.75	13.7	**0.91**	0.94	0.88	2.1	0.71	0.63	0.82	8.3
N3	0.81	0.84	0.79	12.5	0.79	0.82	0.76	4.4	**0.73**	0.71	0.76	6.8
N4	**0.82**	0.84	0.81	9.0	**0.91**	0.95	0.89	1.2	0.72	0.64	0.82	6.3

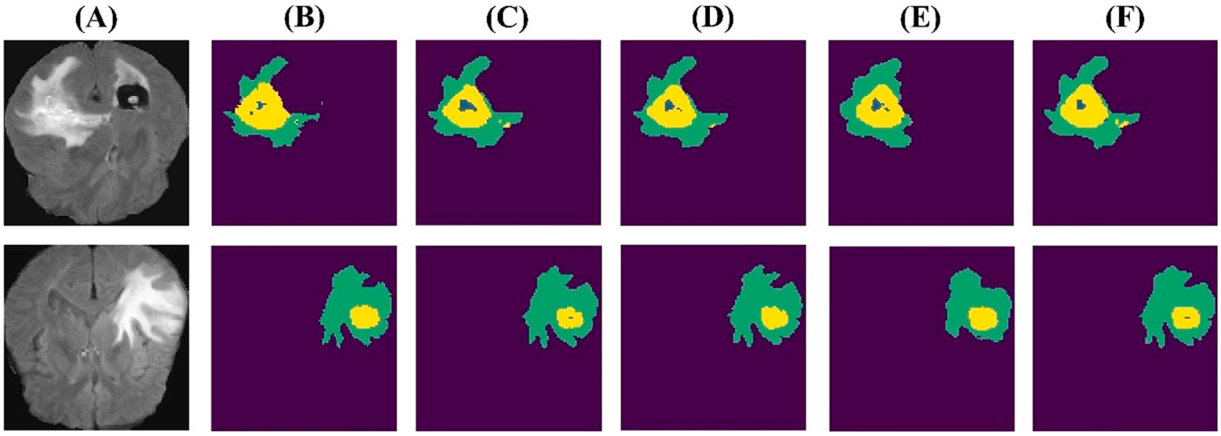

Fig. 6. Visual results of segmentation task, including (A) Original MRI images, (B) Ground truths, (C) N1 – 3D network (No preprocessing), (D) N2 – 3D + GC, (E) N3 – 3D + WSO, (F) N4 – 3D + DA. Blue: necrosis and non-enhancing tumor core, Teal: peritumoral edema, Yellow: enhancing tumor core.

3.2. Glioma segmentation using UNet

This section aims to understand the performance difference of UNet architectures with (N2, N3, and N4) and without (N1) the combination of preprocessing strategies. The segmentation results are shown in Table 2 and Fig. 6, including the performance metrics and visualization of segmentation results. Three models, including N1, N2, and N4, proclaim the highest Dice coefficient of 0.91 for the whole tumor with the similarly detected tumor shape as in the ground truth (Fig. 6B–D, F). However, because N4 achieves higher performance in tumor core and enhancing tumor with Dice coefficients of 0.72 and 0.82 as well as the significant low Hausdorff distance in the three tumor regions (9.0, 1.2, and 6.3 concerning enhancing tumor core, whole tumor, and tumor core), this model is the optimal one. It is also recognized that, when employed in the segmentation networks, data augmentation, instead of any other preprocessing strategy, remains an effective tool for differentiating various tumor regions. Another architecture to explore is N3, which has a Dice coefficient of tumor core of 0.73 and ranks first in the prediction of this region but only occupies the second position when using the Hausdorff distance at 6.3. However, it is deficient in recognizing the whole tumor, with a Dice coefficient of 0.79. Fig. 6E also indicates the poor performance of N4 in peritumoral-edema prediction with the inaccurate shape compared to the ground truth (Fig. 6B). From the results above, we could conclude that data augmentation is the most efficient tool in boosting the Dice coefficient and Hausdorff distance of the outcome.

3.3. Classification of HGG and LGG

This section aims to identify the most optimal pipeline for glioma grading and test the hypothesis on the importance of segmentation in the classification models. As analyzed in Fig. 7, the difference in performance between VGG and GoogleNet should be considered. More specifically, testing accuracies obtained from VGG (A1) and GoogleNet (A7) without segmentation are identical, with values of about 93 %, indicating that the performances of these two techniques are equally effective. Regarding those integrated with segmentation, whereas the accuracies of GoogleNet (A8–A12) receive a minor drop in testing accuracy (excluding A12), the ones of VGG (A2–A6) witness a tremendous gain of 97.96 % in its figure (Fig. 7). As a result, the role of segmentation would change in each classification model to some extent. Besides, although A6 appears to be more helpful in improving classification performance, A2 should be considered more. The fundamental reason is that while the accuracy of A6 slightly improves over A2, its prediction time is significantly longer due to the increased number of input channels (Fig. 7). Taken all together, the model of VGG combined with segmentation (A2) takes advantage of the comparison to the other approaches mentioned above in terms of performance and computational cost. Besides, we can confirm the critical role of segmentation in the glioma grading procedure.

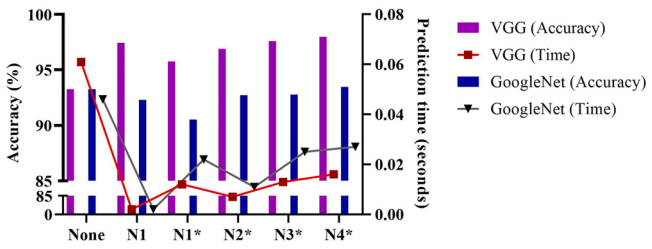

Fig. 7. Result summarization of glioma classification task. N1: 3D UNet, N2: 3D UNet + GC, N3: 3D UNet + WSO, N4: 3D UNet + DA. *: Combined with segmentation results as the fifth channel in the input of classification models.

We also execute the 5-fold cross-validation on VGG models with one-way ANOVA to compare the performance of the original VGG (None) with the one of the others. Table 3 reveals that segmentation is not always helpful in boosting the prediction between LGG and HGG, while only the combinations with N2 and N3 significantly increase accuracy ($p = 0.0087$).

Table 4 presents the results of tunable hyperparameters using Hyperband optimization with a maximum trial count of 20. All sets of hyper-parameters are chosen using the highest performance for each model (Dice coefficient for segmentation tasks and accuracy for classification tasks). The only exception is the batch size of segmentation models, chosen as one due to the shortage of computational memory.

3.4. Grad-CAM explanations

To comprehensively explain the difference in the performance of VGG and GoogleNet (Fig. 7), it is hypothesized that Grad-CAM is applicable in explaining the working mechanism of convolutional networks (Tai et al., 2021). Fig. 8 shows how the Grad-CAM discrepancies resulting from VGG and GoogleNet are utilized to address the problem. More explicitly, the former directly focuses on the lesions and the other notable parts, such as the lateral ventricle. In contrast, the latter focuses on broad regions that roughly account for a hemisphere where the lesions are noticed (Fig. 8A, B). The decision areas of poor-performance pictures in GoogleNet are approximately identical to those of excellent images; in contrast, Grad-CAM of the image containing inaccurately predicted outcomes in VGG obtains dispersed light spots that may have no information for the prediction problem (Fig. 8C). This Grad-CAM visual explanation is consistent with the classification result affected

Table 3
. 5-fold cross-validation results on the VGG model with one-way ANOVA comparing the original VGG (None) with the others. N1: 3D UNet, N2: 3D UNet + GC, N3: 3D UNet + WSO, N4: 3D UNet + DA. *: Combined with segmentation results as the fifth channel in the input of classification models.

Network	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean	Std	P-value
None	0.9327	0.8846	0.9135	0.9320	0.9709	0.9267	0.0315	–
N1	0.9487	0.9103	0.9615	0.8974	0.9487	0.9333	0.0278	0.7351
N1*	0.9570	0.9785	0.9255	0.8667	0.9681	0.9392	0.0451	0.6273
N2*	0.9789	0.9895	0.9792	0.9896	0.9583	0.9791	0.0128	**0.0087**
N3*	0.9512	0.9518	0.9639	0.9518	0.9398	0.9517	0.0085	**0.0087**
N4*	0.9592	0.9184	0.9490	0.9388	0.9898	0.9510	0.0264	0.2226

Table 4
Summarization of tunable hyperparameters in each model.

Hyper-parameters	N1 – 4	A1	A2 – 6	A7	A8 – 12
Leaky ReLU (α)	0.1	0.1	0.2	0.1	0
Dropout (p)	0	0.3	0.1	0.4	0.2
L2 regularization	0.000001	0.00872	0.00976	0.00144	0.00909
Learning rate	0.0001	0.0017	0.00152	0.0003	0.00058
Batch size	1	2	1	2	18

by segmentation when comparing VGG and GoogleNet in Fig. 7.

4. Discussion

The primary purpose of this research is to establish an automatic pipeline for grading gliomas based on MRI images. Besides, the current study aims to evaluate the influence of segmentation and different preprocessing methods and to explain the decisions of the convolutional neural network thanks to Grad-CAM. Specifically, preprocessing techniques, including gamma correction, WSO, and data augmentation, reveal their crucial role during the segmentation procedure with the highest Dice coefficients of 0.91 and 0.82 for the whole tumor and enhancing tumor core (Table 2). Regarding classification, the study compares performance between different classification models and ends up with the best result of 97.44 % in the VGG model combined with segmentation (Fig. 7).

4.1. *Effects of preprocessing techniques on segmentation results*

Regarding preprocessing techniques used in this work, data augmentation should be considered the most notable due to its significant improvement in the segmentation performance (Table 2). Data augmentation is also a state-of-the-art approach often employed in deep learning algorithms, particularly in the medical field, due to the scarcity of data with ground-true labels. This approach changes images in the aspects of directions that would allow more information to be generated.

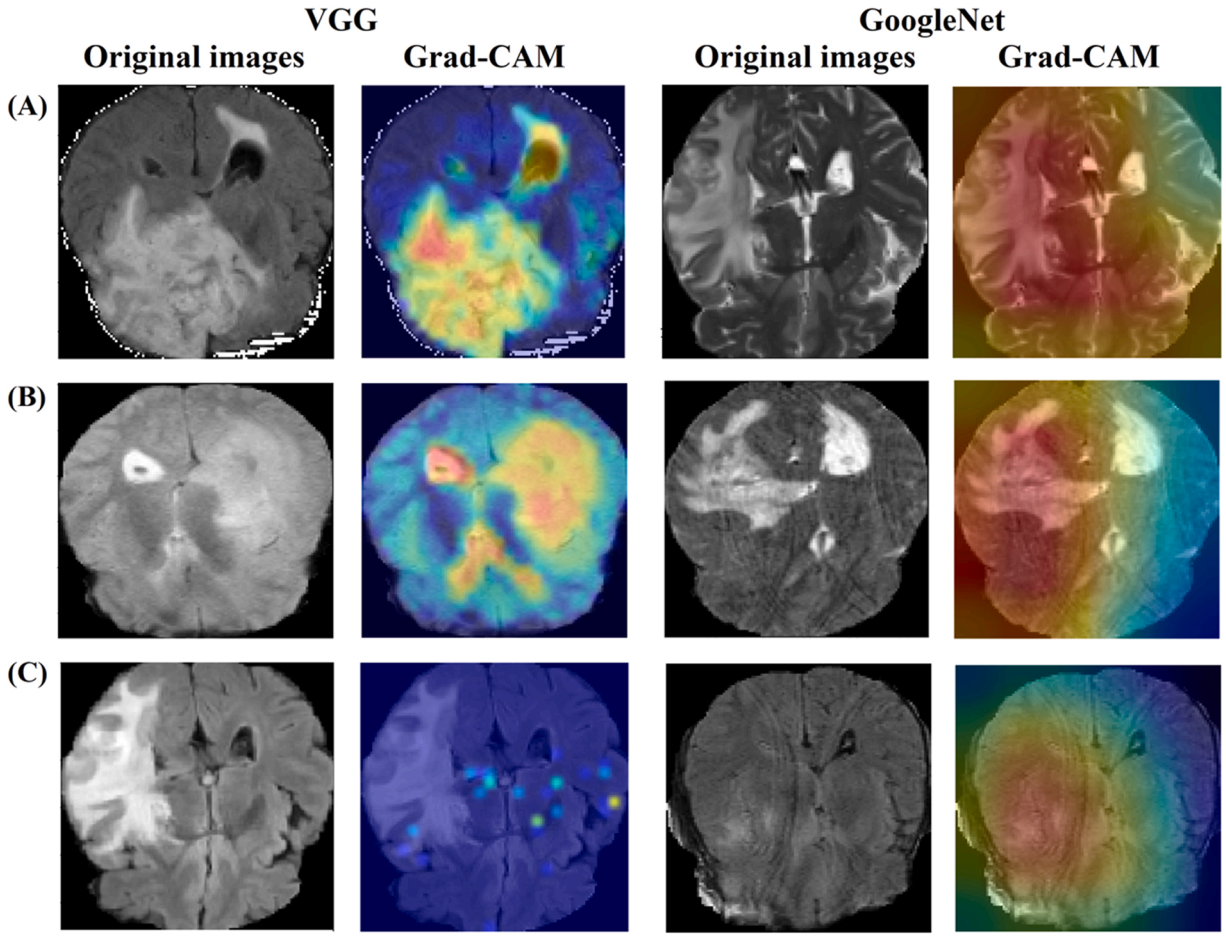

Fig. 8. Grad-CAM comparison between VGG and GoogleNet, including examples of (A, B) good performances and (C) lousy performance.

Table 5
Comparison of other similar studies related to glioma segmentation.

Methods	Data sources	ET		WT		TC	
		DC	HD	DC	HD	DC	HD
kNN-CRF (Havaei et al., 2014)	BraTS 2013	0.63	–	0.85	–	0.75	–
kNN-MRF (Havaei et al., 2014)	BraTS 2013	0.66	–	0.82	–	0.71	–
DeepMedic (Kamnitsas et al., 2016)	BraTS 2015	0.72	–	0.90	–	0.75	–
FCN and CRF-RNN (Zhao et al., 2018)	BraTS 2015	0.62	–	0.84	–	0.73	–
Res-UNet (Noori et al., 2019)	BraTS 2018	0.81	**2.9**	0.90	4.1	0.82	6.3
NVDLMED (Myronenko, 2019)	BraTS 2018	0.78	3.8	0.88	5.9	0.82	**4.8**
OM-Net + CGAp (Zhou et al., 2020)	BraTS 2018	0.80	**2.9**	**0.91**	4.9	0.82	6.9
Segtran (Li et al., 2022)	BraTS 2019	0.74	–	0.90	–	0.82	–
nnUNet (Isensee et al., 2021)	BraTS 2020	**0.82**	17.8	0.89	8.5	**0.85**	17.3
Multi-Scale Feature Fusing Network (Huang et al., 2021b)	BraTS 2015	0.61	–	0.86	–	0.73	–
UNet (Decuyper et al., 2021)	BraTS 2019	0.76	3.9	0.90	5.7	0.80	7.0
Proposed	BraTS 2019	**0.82**	9.0	**0.91**	**1.2**	0.72	6.3

Thanks to the enormous amount of ground-true data, overfitting problems could be avoided in large-capacity algorithms (Nalepa et al., 2019). Several prior studies, such as Decuyper et al. (2021), have also indicated an increase in brain tumor segmentation as a result of the use of this approach (Table 5).

The other two methods have also proved their power. Gamma correction, for example, promotes the Dice coefficients in all tumor regions thanks to the adjustment of the intensity variance but is not sensitive in detecting the tumor boundary with the higher Hausdorff distance (Table 2). Only a few studies have used this approach previously, including those by Huang et al. (2021a) and Tai et al. (2021). However, these two publications used gamma correction on the two-class segmentation and obtained Dice coefficients of 0.86 and 0.93, respectively (Huang et al., 2021a; Tai et al., 2021). Besides, WSO is another noticeable approach that significantly increases the performance of four-class segmentation problems but is also not effective in the boundary recognition of the enhancing tumor core and the whole tumor (Table 2). Specifically, the two hyper-parameters, WW and WL, are necessary for WSO application to limit the intensity of tumor regions in a specific range and enhance them in the images. In this study, these two hyper-parameters would be chosen by Keras Tuner automatically; as a result, one of the modalities is evaluated as unnecessary during the segmentation procedure and removed completely, which might help boost the performance of models (Fig. 5C). Unfortunately, the performance metrics of the whole tumor, such as the Dice coefficient and Hausdorff distance, are drastically worsened, which may be due to information loss in edema zones, particularly when selecting the proper hyper-parameters for the custom activations in WSO. However, the effectiveness in the tumor core and enhancing tumor core demonstrates that WSO is applicable in various medical images, not just for CT images.

4.2. Comparison to other studies related to glioma segmentation

Table 5 compares the performance of different studies related to glioma segmentation, which mainly focuses on the BraTS dataset. Our proposed pipeline achieves outstanding results in enhancing tumor and whole tumor regions. However, it significantly lowers the Dice coefficient in the prediction of tumor core when compared with other studies, especially the study of Isensee et al. using nnUNet in 2021 (Isensee et al., 2021). In terms of Hausdorff distance, the only brain tumor region predicted outstandingly in this study is the whole tumor. In contrast, other designs, such as OM-Net by Zhou et al. (2020) and Res-UNet by Noori et al. (2019), demonstrate their power in boosting tumor-core prediction. The Hausdorff distance in the diagnosis of tumor core, otherwise, is noticeable in the performance of NVDLMED byMyronenko (2019).

4.3. Selection of classification models

When comparing VGG and GoogleNet without segmentation, it is reasonable to conclude that their performances are about equivalent, with values of approximately 93 %. Besides, the accuracies of these two models are relatively close to other similar research, such as Ahammed Muneer et al. (2019) and Yang et al. (2018) (94.64 % and 94.50 %, respectively). However, VGG should be paid more attention because its training period is significantly shorter than GoogleNet (Fig. 7).

4.4. Grad-CAM explanations

There is a substantial gap between VGG and GoogleNet regarding classification models combined with segmentation. While the accuracy of the former rises substantially to 97.44 %, to be more exact, the latter shows little improvement or even a slight fall (Fig. 7). This occurrence might be explained by the discrepancy in their Grad-CAM. In other words, VGG focuses on the brain tumor areas and other notable parts, such as the lateral ventricle. Consequently, when these irrelevant parts are deleted thanks to the segmentation procedure, it may reduce the confusion about the decision of VGG architecture. However, the region of decision for GoogleNet accounts for half of the brain regions, which means that when the surrounding areas of tumors are eliminated, the information supporting the prediction of GoogleNet is lost. It might cause a decline in the performance of this network (Fig. 8). Several studies, such as Ahammed Muneer et al. (2019) and Decuyper et al. (2021), have already examined the combination of classification networks with segmentation which achieves significant performance (Table 6).

4.5. Comparison to other articles related to glioma grading

Our proposed classification models achieve a noticeable result but still need significant improvement. Table 6 describes state-of-the-art techniques, which should also be considered thanks to their

Table 6
Comparison of other similar studies related to glioma grading.

Methods	Data sources	Accuracy (%)
Linear SVM (Zhang et al., 2017)	Tangdu Hospital, China	94.5
SMO (Zhang et al., 2017)	Tangdu Hospital, China	94.5
Logistics (Cho et al., 2018)	BraTS 2017	88.77
RF (Cho et al., 2018)	BraTS 2017	88.77
GBDT (Wang et al., 2019)	Shandong Provincial Hospital, China	87
Pre-trained AlexNet (Yang et al., 2018)	Tangdu Hospital, China	92.7
Pre-trained GoogleNet (Yang et al., 2018)	Tangdu Hospital, China	94.5
kNN (Saba et al., 2020)	BraTS 2015	**99.82**
Gabor-modulated CNN (Singh et al., 2021)	BraTS 2017	**98.68**
Multimodal disentangled variational autoencoder (Cheng et al., 2022)	BraTS 2019	**98.46**
Proposed	BraTS 2019	**97.44**

outstanding accuracies, such as Saba et al. (2020), Singh et al. (2021), and Cheng et al. (2022). Their problem investigation approaches should be considered valuable references for our further studies. For example, Saba et al. (2020) have succeeded in extracting characteristics of MRI images, including deep features (from VGG 19), local binary pattern, and histogram orientation gradient, and applied them in various machine learning algorithms to find the best result of 99.82 % with kNN classifier. Singh et al. (2021) emphasized the role of Gabor orientation filters for shallow layers of deep learning architectures. Cheng et al. (2022) established an autoencoder to reconstruct the features extracted from MRI images, including intensity, wavelet, Laplacian of Gaussian, and local binary pattern, before combining these features through two hidden layers to get the final prediction results (Saba et al., 2020; Cheng et al., 2022; Singh et al., 2021). From the analysis of other research, it is recognized that our model still depends on traditional artificial intelligent architectures and needs original development to find the most suitable algorithms for glioma-grading architecture.

4.6. Limitations and future directions

There are still several limitations in this study that we must fill in for the use of glioma segmentation and grading in real life. One of the most significant ones is removing some images in classification combined with segmentation. It is because the segmented regions of these images are outside the ground truths; as a result, it would be meaningless when using the non-tumor regions as the input in classification models. However, this problem also creates a bias in classification results while removing poor performance records. Other different datasets should also be examined to get more information during the learning procedure and evaluate classification performance objectively. Besides, fitting hyper-parameters should be continued to improve tumor core prediction performance. A user interface is also required for the classification process to monitor the pathology and assist clinicians in making timely treatment decisions.

5. Conclusions

This study arrives at several key findings. First, a streamlined pathway is built for developing deep-learning models to differentiate gliomas. The comparison of different networks also illustrates the critical significance of data with ground-true labels, which can be manufactured artificially using data augmentation techniques. Other preprocessing approaches, such as gamma correction and WSO, should be considered in the segmentation procedure. Concerning classification, this study uses two primary methodologies, each of which produces a unique outcome. While most of the suggested architectures obtain equivalent accuracies of about 93 % on the testing dataset, the model of VGG combined with segmentation gets an outstanding result of 97.44 %.

CRediT authorship contribution statements

Khiet Dang: Conceptualization, Methodology, Investigation, Programming, Writing – original draft, Visualization. **Toi Vo:** Funding acquisition, Revision, Supervision. **Lua Ngo:** Writing – review & editing, Supervision. **Huong Ha:** Conceptualization, Writing – review & editing, Supervision.

Compliance with ethical standards

This study was carried out retrospectively using anonymized data from clinical routines. As a result, no ethical approval was necessary. This work is also dependent on the BraTS 2019. There are no ethics declarations required for the usage of this data collection. In our case, informed consent was unnecessary because the dataset was obtained from the public domain.

Conflicts of Interest

The authors have no conflict of interest to declare.

Acknowledgment

This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under Grant no. NCM2020-28-01.

References

Ahammed Muneer, K.V., Rajendran, V.R., Glioma, K.P.J., 2019. Tumor grade identification using artificial intelligent techniques. J. Med. Syst. 43, 113. https://doi.org/10.1007/s10916-019-1228-2.

Ahuja, A.S., 2019. The impact of artificial intelligence in medicine on the future role of the physician. PeerJ 7, e7702. https://doi.org/10.7717/peerj.7702.

Amiri, S.A., Hassanpour, H. Hassanpur, 2012. A preprocessing approach for image analysis using gamma correction. Int. J. Comput. Appl. (0975–8887) 38 (12). https://doi.org/10.5120/4764-6914.

Arsava, E.M., Saarinen, J.T., Unal, A., Akpinar, E., Oguz, K.K., Topcuoglu, M.A., 2014. Impact of window setting optimization on accuracy of computed tomography and computed tomography angiography source image-based Alberta Stroke Program early computed tomography score. J. Stroke Cerebrovasc. Dis. 23, 12–16. https://doi.org/10.1016/j.jstrokecerebrovasdis.2012.05.012.

Baid, U., Talbar, S., Rane, S., Gupta, S., Thakur, M.H., Moiyadi, A., et al., 2020. A novel approach for fully automatic intra-tumor segmentation with 3D U-net architecture for gliomas. Front. Comput. Neurosci. 14, 10. https://doi.org/10.3389/fncom.2020.00010.

Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J.S., et al., 2017a. Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. Sci. Data 4, 170117. https://doi.org/10.1038/sdata.2017.117.

Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., et al., 2017b. Segmentation Labels for the Pre-operative Scans of the TCGA-GBM collection. ⟨https://doi.org/10.7937/K9/TCIA.2017.KLXWJJ1Q⟩.

Bakas, S., Akbari, H., Sotiras A., Bilello M., Rozycki M., Kirby J., et al., 2017c. Segmentation Labels for the Pre-operative Scans of the TCGA-LGG collection. ⟨https://doi.org/10.7937/K9/TCIA.2017.GJQ7R0EF⟩.

Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., et al., 2019. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. ArXiv: 181102629 [Cs, Stat].

Cheng, J., Gao, M., Liu, J., Yue, H., Kuang, H., Liu, J., et al., 2022. Multimodal disentangled variational autoencoder with game theoretic interpretability for glioma grading. IEEE J. Biomed. Health Inform. 26, 673–684. https://doi.org/10.1109/JBHI.2021.3095476.

Cho, H., Lee, S., Kim, J., Park, H., 2018. Classification of the glioma grading using radiomics analysis. PeerJ 6, e5982. https://doi.org/10.7717/peerj.5982.

Decuyper, M., Bonte, S., Deblaere, K., Van Holen, R., 2021. Automated MRI based pipeline for segmentation and prediction of grade, IDH mutation and 1p19q co-deletion in glioma. Comput. Med. Imaging Graph. 88, 101831 https://doi.org/10.1016/j.compmedimag.2020.101831.

Doerner, J., Luetkens, J.A., Iuga, A.-I., Byrtus, J., Haneder, S., Maintz, D., et al., 2018. Poly-energetic and virtual mono-energetic images from a novel dual-layer spectral detector CT: optimization of window settings is crucial to improve subjective image quality in abdominal CT angiographies. Abdom. Radiol. 43, 742–750. https://doi.org/10.1007/s00261-017-1241-1.

Ertosun, M.G., Rubin, D.L., 2015. Automated grading of gliomas using deep learning in digital pathology images: a modular approach with ensemble of convolutional neural networks. AMIA Annu. Symp. Proc. 2015, 1899–1908.

Fathi Kazerooni, A., Bakas, S., Saligheh Rad, H., Davatzikos, C., 2020. Imaging signatures of glioblastoma molecular characteristics: a radiogenomics review. J. Magn. Reson. Imaging 52, 54–69. https://doi.org/10.1002/jmri.26907.

Gore, S., Chougule, T., Jagtap, J., Saini, J., Ingalhalikar, M., 2021. A review of radiomics and deep predictive modeling in glioma characterization. Acad. Radiol. 28, 1599–1621. https://doi.org/10.1016/j.acra.2020.06.016.

Guzmán-De-Villoria, J.A., Mateos-Pérez, J.M., Fernández-García, P., Castro, E., Desco, M., 2014. Added value of advanced over conventional magnetic resonance imaging in grading gliomas and other primary brain tumors. Cancer Imaging 14, 35. https://doi.org/10.1186/s40644-014-0035-8.

Havaei, M., Jodoin, P., Larochelle, H. , 2014. Efficient interactive brain tumor segmentation as within-brain kNN classification. In: Proceedings of the 2014 22nd International Conference on Pattern Recognition, pp. 556–61. ⟨https://doi.org/10.1109/ICPR.2014.106⟩.

Huang, Z., Liu, Y., Song, G., Zhao, Y., 2021a. GammaNet: an intensity-invariance deep neural network for computer-aided brain tumor segmentation. Optik 243, 167441. https://doi.org/10.1016/j.ijleo.2021.167441.

Huang, P., Li, D., Jiao, Z., Wei, D., Li, G., Wang, Q., et al., 2019. CoCa-GAN: common-feature-learning-based context-aware generative adversarial network for glioma grading. In: Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., et al. (Eds.), Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Springer International Publishing, Cham, pp. 155–163. https://doi.org/10.1007/978-3-030-32248-9_18.

Huang, D., Wang, M., Zhang, L., Li, H., Ye, M., Li, A., 2021b. Learning rich features with hybrid loss for brain tumor segmentation. BMC Med. Inf. Decis. Mak. 21, 63. https://doi.org/10.1186/s12911-021-01431-y.

Isensee, F., Jäger, P.F., Full, P.M., Vollmuth, P., Maier-Hein, K.H., 2021. nnU-Net for brain tumor segmentation. In: Crimi, A., Bakas, S. (Eds.), Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer International Publishing, Cham, pp. 118–132. https://doi.org/10.1007/978-3-030-72087-2_11.

Kamnitsas, K., Ferrante, E., Parisot, S., Ledig, C., Nori, A.V., Criminisi, A., et al., 2016. DeepMedic for brain tumor segmentation. In: Crimi, A., Menze, B., Maier, O., Reyes, M., Winzeck, S., Handels, H. (Eds.), Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer International Publishing, Cham, pp. 138–149. https://doi.org/10.1007/978-3-319-55524-9_14.

Kim, T.K., 2017. Understanding one-way ANOVA using conceptual figures. Korean J. Anesth. 70, 22–26. https://doi.org/10.4097/kjae.2017.70.1.22.

Lee, H., Kim, M., Do, S. , 2018. Practical window setting optimization for medical image deep learning. ArXiv:181200572 [Cs].

Ling, C.X., Li, C. , n.d. Data Mining for Direct Marketing: Problems and Solutions, 7. ⟨https://doi.org/10.1007/3-540-39205-X_83⟩.

Li, S., Sui, X., Luo, X., Xu, X., Liu, Y., Goh, R.S.M. , 2022. Medical Image Segmentation using Squeeze-and-Expansion Transformers.

Maravilla, K.R., Sory, W.C., 1986. Magnetic resonance imaging of brain tumors. Semin. Neurol. 6, 33–42. https://doi.org/10.1055/s-2008-1041445.

Matsumoto, N., Nagao, K., Hirayama, A., Sato, Y., 2010. Non-invasive assessment and clinical strategy of stable coronary artery disease by magnetic resonance imaging, multislice computed tomography and myocardial perfusion SPECT. Circ. J. 74, 34–40. https://doi.org/10.1253/circj.CJ-09-0791.

Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., et al., 2015. The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imaging 34, 1993–2024. https://doi.org/10.1109/TMI.2014.2377694.

Moore, C.M. , n.d. Dice similarity coefficient | Radiology Reference Article | Radiopaedia.org. Radiopaedia. ⟨https://doi.org/10.53347/rID-75056⟩.

Myronenko, A., 2019. 3D MRI brain tumor segmentation using autoencoder regularization. In: Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (Eds.), Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer International Publishing, Cham, pp. 311–320. https://doi.org/10.1007/978-3-030-11726-9_28.

Nalepa, J., Marcinkiewicz, M., Kawulok, M., 2019. Data augmentation for brain-tumor segmentation: a review. Front. Comput. Neurosci. 13, 83. https://doi.org/10.3389/fncom.2019.00083.

Noori, M., Bahri, A., Mohammadi, K. , 2019. Attention-guided version of 2D UNet for automatic brain tumor segmentation. In: Proceedings of the 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), pp. 269–75. ⟨https://doi.org/10.1109/ICCKE48569.2019.8964956⟩.

Pouratian, N., Asthagiri, A., Jagannathan, J., Shaffrey, M.E., Schiff, D., 2007. Surgery insight: the role of surgery in the management of low-grade gliomas. Nat. Rev. Neurol. 3, 628–639. https://doi.org/10.1038/ncpneuro0634.

Rajapaksa, S., Khalvati, F., 2021. Localized perturbations for weakly-supervised segmentation of glioma brain tumours. ArXiv. https://doi.org/10.48550/arXiv.2111.14953 (211114953 [Cs, Eess]).

Ross, A., Willson, V.L., 2017. One-way anova. In: Ross, A., Willson, V.L. (Eds.), Basic and Advanced Statistical Tests: Writing Results Sections and Creating Tables and Figures. SensePublishers, Rotterdam, pp. 21–24. https://doi.org/10.1007/978-94-6351-086-8_5.

Saba, T., Sameh Mohamed, A., El-Affendi, M., Amin, J., Sharif, M., 2020. Brain tumor detection using fusion of hand crafted and deep learning features. Cogn. Syst. Res. 59, 221–230. https://doi.org/10.1016/j.cogsys.2019.09.007.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2020. Grad-CAM: visual explanations from deep networks via gradient-based localization. Int. J. Comput. Vis. 128, 336–359. https://doi.org/10.1007/s11263-019-01228-7.

Simonyan, K., Zisserman, A. , 2015. Very deep convolutional networks for large-scale image recognition. ArXiv:14091556 [Cs]. ⟨https://doi.org/10.48550/arXiv.1409.1556⟩.

Singh, R., Goel, A., Raghuvanshi, D.K., 2021. Computer-aided diagnostic network for brain tumor classification employing modulated Gabor filter banks. Vis. Comput. 37, 2157–2171. https://doi.org/10.1007/s00371-020-01977-4.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al., 2015. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9. ⟨https://doi.org/10.1109/CVPR.2015.7298594⟩.

Tabibkhooei, A., Izadpanahi, M., Arab, A., Zare-Mirzaei, A., Minaeian, S., Rostami, A., et al., 2020. Profiling of novel circulating microRNAs as a non-invasive biomarker in diagnosis and follow-up of high and low-grade gliomas. Clin. Neurol. Neurosurg. 190, 105652 https://doi.org/10.1016/j.clineuro.2019.105652.

Tai, Y.-L., Huang, S.-J., Chen, C.-C., Lu, H.H.-S., 2021. Computational complexity reduction of neural networks of brain tumor image segmentation by introducing fermi–Dirac correction functions. Entropy 23, 223. https://doi.org/10.3390/e23020223.

Wang, X., Wang, D., Yao, Z., Xin, B., Wang, B., Lan, C., et al., 2019. Machine learning models for multiparametric glioma grading with quantitative result interpretations. Front. Neurosci. 12, 1046. https://doi.org/10.3389/fnins.2018.01046.

Wei, Y., Li, C., Chen, X., Schöinlieb, C.-B., Price, S.J. , 2022. Collaborative learning of images and geometrics for predicting isocitrate dehydrogenase status of glioma. In: Proceedings of the 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), pp. 1–4. ⟨https://doi.org/10.1109/ISBI52829.2022.9761407⟩.

Yang, Y., Yan, L.-F., Zhang, X., Han, Y., Nan, H.-Y., Hu, Y.-C., et al., 2018. Glioma grading on conventional MR images: a deep learning study with transfer learning. Front. Neurosci. 12, 804. https://doi.org/10.3389/fnins.2018.00804.

Yong, H., Huang, J., Meng, D., Hua, X., Zhang, L., 2020. Momentum batch normalization for deep learning with small batch size. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M. (Eds.), Computer Vision – ECCV 2020. Springer International Publishing, Cham, pp. 224–240. https://doi.org/10.1007/978-3-030-58610-2_14.

Zhang, X., Yan, L.-F., Hu, Y.-C., Li, G., Yang, Y., Han, Y., et al., 2017. Optimizing a machine learning based glioma grading system using multi-parametric MRI histogram and texture features. Oncotarget 8, 47816–47830. https://doi.org/10.18632/oncotarget.18001.

Zhao, X., Wu, Y., Song, G., Li, Z., Zhang, Y., Fan, Y., 2018. A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. Med. Image Anal. 43, 98–111. https://doi.org/10.1016/j.media.2017.10.002.

Zhou, C., Ding, C., Wang, X., Lu, Z., Tao, D., 2020. One-pass multi-task networks with cross-task guided attention for brain tumor segmentation. IEEE Trans. Image Process. https://doi.org/10.1109/tip.2020.2973510.

Zlochower, A., Chow, D.S., Chang, P., Khatri, D., Boockvar, J.A., Filippi, C.G., 2020. Deep learning AI applications in the imaging of glioma. Top. Magn. Reson. Imaging 29. https://doi.org/10.1097/RMR.0000000000000237 (115–00).