

# Preoperative Grading of Brain Gliomas Using 3D-ResNet18 Based on Multimodal MRI and Attention Mechanism

He Yuanzhe, Qian Chengyi and Wang Yuanjun\*

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

**Abstract:** In response to the prevalent challenge of imprecise preoperative glioma grading prediction, a multimodal Magnetic Resonance Imaging (MRI) image with attention mechanism prediction model based on Residual Network (ResNet) is proposed for accurate prediction. To achieve this, publicly accessible datasets from BraTS 2019 and BraTS 2020 were employed. MRI images across three modalities underwent preprocessing, cropping, and subsequent stacking to generate comprehensive multimodal representations. Meanwhile, An Efficient Channel Attention (ECA) mechanism based on 3D data (3D-ECA) is proposed and introduced to reduce the problem of slow network convergence and overfitting. Notably, experiments demonstrate that the 3D-ECA attention mechanism added to the model improves the network training speed and the classification accuracy of the model, and the classification accuracy reached 88.1%, with an impressive area under the subject working characteristic curve of 0.918. Therefore, the 3D-ResNet18 deep learning model, incorporating multimodal MRI and 3D-ECA attention mechanisms, demonstrates high accuracy and robustness in distinguishing between high-grade and low-grade gliomas. Thus, showcasing significant clinical potential for glioma classification and diagnosis.

**Keyword:** Glioma classification, 3D-residual networks, Magnetic resonance imaging, Deep learning, Multimodal magnetic resonance imaging.

## INTRODUCTION

Gliomas are the most prevalent primary malignant tumours in the adult brain, with a global annual incidence of 5.6 cases/100,000 people, and the median Overall Survival (OS) for glioblastoma is only 14.4 months [1]. According to the 2016 World Health Organization classification of tumours of the central nervous system, gliomas are classified as low-grade glioma (LGG) and high-grade gliomas (HGG). However, accurate and noninvasive prediction of glioma grading is essential for the development of effective clinical treatment plans, as gliomas of different grades require different clinical treatment strategies, which have a significant impact on patient prognosis and outcome [2]. Although conventional single sequence magnetic resonance imaging still plays an important role in the clinical diagnosis of glioma grading, its sensitivity and specificity are relatively low [3].

In recent years, the application of artificial intelligence in the field of medical image processing has become a hotspot of clinical research, especially machine learning and deep learning have shown great potential in the prediction of tumour grading, differential diagnosis, tumour segmentation, and survival prediction, etc. [4]. Ahammed *et al.* [5] demonstrated a great potential of machine learning in the prediction of tumour grading, differential diagnosis, tumour segmentation, and survival prediction by using classifiers based on the Wndchrm tool and deep conv-

olutional neural networks based on the VGG-19 Convolutional Neural Networks (CNN) classifier to classify and compare LGG, oligodendroglioblastoma, mesenchymal glioma and glioblastoma multiforme with an accuracy of 92.86% and 98.25% respectively. Their experiments also observed that increasing the number of image inputs improves the performance of the classifier. Naser *et al.* [6] collected pre-T1 enhancement, FLAIR and post-T1 enhancement MRI images of 110 patients with lower grade gliomas based on migration learning of pre-trained convolutional bases and fully connected classifiers with Vgg16 for tumour grading. The accuracy, sensitivity and specificity of the grading model to classify lower grade gliomas into grade II and III were 0.95, 0.97 and 0.98, respectively. However, most of the previous studies used unimodal MRI images as the input data for convolutional networks, which failed to make full use of the data information from different medical images. In this paper, we propose to develop a 3D-ResNet18 deep learning model based on multimodal MRI images with attention mechanism to stack three modal MRI images to generate multimodal MRI images. In order to solve the problems of slow network convergence and over fitting, ECA attention mechanism based on 3D data is proposed and introduced. Considering the limitations of computing power and computational memory, the images are cropped and shrunk to ensure that the network can extract all the important features in the images. The problem of imbalance in the number of samples between image classes is solved by oversampling the data samples. The multimodal MRI images processed above were fed into the network to enable the network to learn the image features, and through the steps of convolution, local maximum

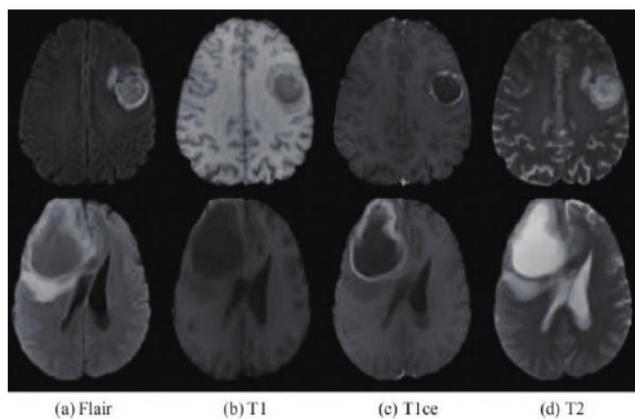
\*Address correspondence to this article at the School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; E-mail: yjusst@126.com

pooling, global average pooling, fully connected layers, batch normalisation, activation function, etc., the output of the model was finally mapped to the probability distributions of the classes through the SoftMax layer in order to predict the pathological grade of the gliomas, and the robustness was tested through the five-fold cross-validation method.

## MATERIAL

### MRI Image Data

The proposed method was assessed in this work utilizing the brain tumor segmentation datasets from BraTS 2019 and BraTS 2020, which are publically available. There are 259 high-grade glioma instances and 76 low-grade glioma cases in the BraTS 2019 training set, and 34 more high-grade glioma cases are included in the BraTS 2020 training set. As shown in Figure 1, each brain tumour data sample has four modalities: T1-weighted Imaging (T1WI), T1 Contrast-Enhanced (T1ce), T2-weighted Imaging (T2WI), and Fluid-Attenuated Inversion Recovery (FLAIR), each with a dimension of  $240 \times 240 \times 155$ .



**Figure 1:** Brain glioma MRI with four modalities.

In order to evaluate the model performance more comprehensively, a five-fold cross-validation set was used. The dataset was divided into five mutually exclusive subsets, and in each training-validation cycle, four subsets served as the training set, resulting in a training set of about 295 cases per fold. It was divided into LGG and HGG groups according to the WHO classification, where the LGG group (WHO class I-II) had about 61 cases, while the HGG group (WHO class III-IV) had about 235 cases. A remaining subset served as the validation set, resulting in a validation set of approximately 74 cases per fold, with approximately 15 cases in the LGG group and 58 cases in the HGG group.

In view of the unbalanced nature of the samples in the dataset, an oversampling operation was performed

on the LGG groups in the dataset, with a 1:4 oversampling method for the LGG groups in the validation set, and a 1:4 oversampling method for the LGG groups in the training set, in order to achieve a balance in the sample data. As a result, the final experimental training set for each fold was about 479 cases, with about 244 cases for the LGG group and 235 cases for the HGG group. The validation set is about 118 cases, with about 60 cases in the LGG group and 58 cases in the HGG group.

By averaging the results of five cross-validations, we were able to more reliably evaluate the performance of the proposed method on different data subsets.

## IMAGE PRE-PROCESSING

### Image Resizing and Image Spacing Adjustment

Load the original data from medical images in nii format. A scaling factor was calculated based on the original image size and the desired new size, and then the image was resized using the nearest neighbour interpolation method. This process ensured that all images are at the same spatial resolution for subsequent analyses while maintaining the accuracy of the pixel values. Accordingly, the spacing of the images was adjusted to maintain the spatial integrity of the data. This step ensured that the spatial relationship between images was maintained when subsequent analyses were performed.

### Image Resampling

Image resampling was performed to ensure that all images were at the same spatial resolution for subsequent analysis. The nearest neighbour interpolation method was chosen, specifically for the binary segmentation mask, to maintain the accuracy of the pixel values.

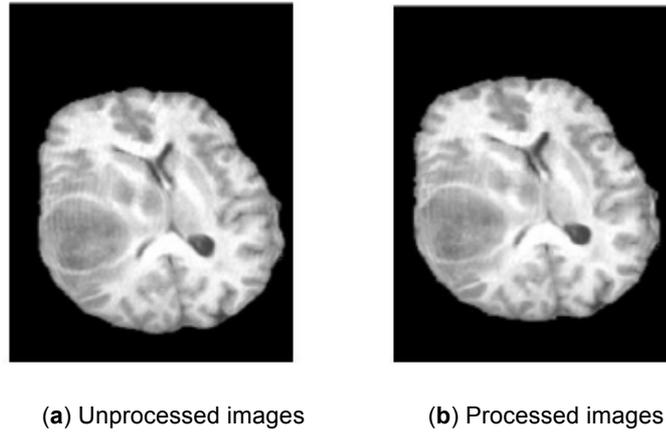
### Image Cropping

Considering the abundance of extraneous background in brain tumor images and constraints like limited computational resources, the original image data underwent cropping to dimensions of  $200 \times 200 \times 120$ . This cropping procedure alleviates computational load without compromising experiment reproducibility or result reliability, as illustrated in Figure 2.

## MODELLING METHODS

### ResNet18 Model

In order to solve the problems of gradient vanishing and limited expressiveness in deep neural network



**Figure 2:** Preprocessed MRI images.

training, HE *et al.* [7] from Microsoft Research proposed a deep residual network (ResNet) in 2015. ResNet18 is a smaller version of the ResNet family of models, which has a moderate number of layers, fast convergence rate. Compared with ResNet50 and ResNet101, ResNet18 has lower network complexity and computational resource consumption. The structure of ResNet18 network is shown in Table 1.

ResNet18 can effectively use residual connectivity to optimise the gradient flow during training, avoiding the problem of overfitting. In order to ensure the accuracy of brain glioma classification, so the ResNet18 network model was chosen for this study.

### 3D CNN

2D CNNs are typically used to perform two-dimensional convolutional operations at the convolutional layer to extract features in the local receptive field and generate an output feature map by applying additive bias and nonlinear activation

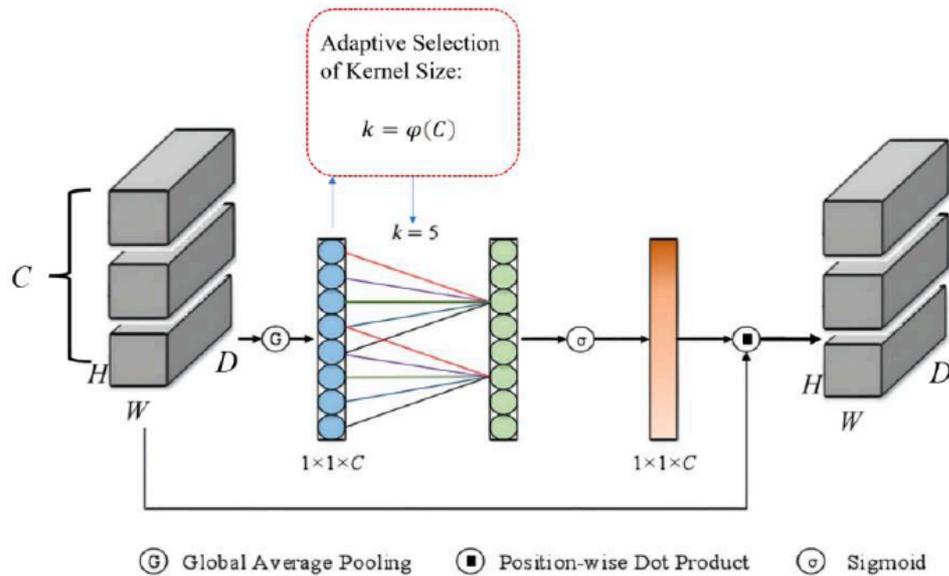
functions. The pooling layer is typically used to reduce the size of the feature map by downsampling to enhance invariance. CNN models are typically constructed by alternately stacking multiple convolutional and pooling layers and using a backpropagation algorithm to update the weighting parameters. However, when processing medical data, 2D networks only take into account the 2D image information in the slices, whereas 3D networks can make full use of the spatial structure information of the images. 3D convolution obtains spatial information by applying a 3D convolution kernel on a cube consisting of multiple consecutive frames. The value of 3D convolution for the  $j$ -th feature map at position  $(x,y,z)$  in layer  $i$  of the network is denoted as:

$$v_{ij}^{xyz} = \tanh \left( b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} \omega_{ijm}^{(pqr)} v_{(i-1)m}^{(x+p)(y+q)(z+r)} \right) \quad (1)$$

$\tanh(\cdot)$  : activation function;  $b_{ij}$  : Additive bias;  $\omega_{ijm}^{(pqr)}$  : the convolution value at the  $m$ -th feature map

**Table 1: Resnet18 Network Structure Diagram**

Layer name	Output size	18-layer
conv1	112×112	7×7, 64, stride 2
		3×3max pool, stride 2
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
	1×1	
FLOPs		$1.8 \times 10^9$



**Figure 3:** 3D-ECA Attention Module Flow chart.

$(p, q, r)$  position of the previous layer of the network ;  $P_i$ ,  $Q_i$ ,  $R_i$ : Height, Width and Depth of the convolution kernel

### 3D-ECA Attention Module

As shown in Figure 3, we propose an ECA attention mechanism based on 3D data called 3D-ECA. Compared with the traditional attention mechanism based on 2D data, the 3D ECA module has better applicability and effectiveness in dealing with 3D data.

The 3D-ECA module uses  $1 \times 1 \times 1$  convolutional layers directly after the global average pooling layer, removing the fully connected layers. The module avoids dimensionality reduction and captures cross-channel interactions efficiently with good results involving only a few parameters. 3D-ECA accomplishes cross-channel information interactions through a one-dimensional convolution, where the size of the convolution kernel is adaptively varied by a function that allows more cross-channel interactions in layers with a larger number of channels.

Firstly, through the global average pooling layer, the large 3D features  $(D \times H \times W)$  of each channel are compressed to a real number, and the feature map dimension changes:  $(C, D, H, W) \rightarrow (C, 1, 1, 1)$ .

Then, the adaptive 1D convolutional kernel size ( $k$ ) is calculated as follows.

$$k = \varphi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{odd} \quad (2)$$

where  $\gamma = 2$ ,  $b = 1$ ,  $C$  is the number of channels

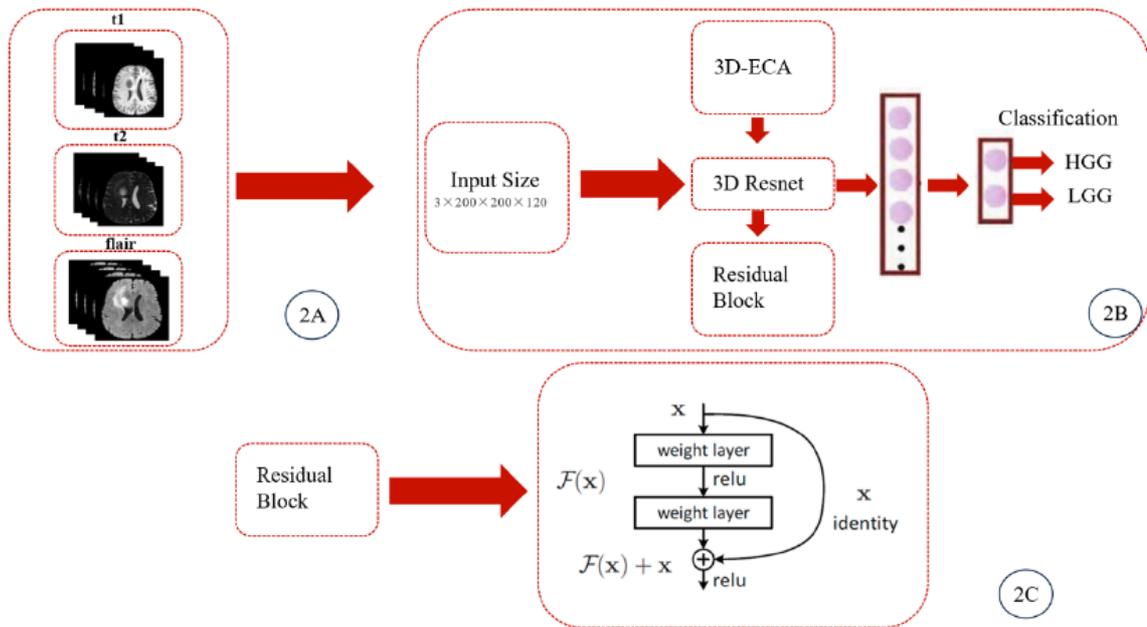
Next,  $k$  is used in a one-dimensional convolution kernel to obtain, for each channel of the feature map, a vector of weights, varying in dimension:  $(C, 1, 1, 1) \rightarrow (C, 1, 1, 1)$ .

Finally, the normalised weights and the original input feature maps are multiplied channel by channel to obtain the weighted output features:  $(C, D, H, W) * (C, 1, 1, 1) \rightarrow (C, D, H, W)$ .

### Modelling

As shown in Figure 4, when constructing the 3D-ResNet18 model, we improved it based on 2D-ResNet by replacing all 2D convolutional layers with 3D convolution to better adapt to the 3D structure of medical images. The input data consists of T1WI, T2WI and FLAIR images stacked to form an input image with three channels. Considering the limitations of the graphics card and computational memory, we fixed the input image size at  $3 \times 200 \times 200 \times 120$ . The model utilized the Adam optimizer, which employs two critical hyperparameters,  $\beta_1$  and  $\beta_2$ , to regulate the decay rates of the first-order moment estimates (mean) and second-order moment estimates (variance), respectively. The initial learning rate was set to 0.0001, with  $\beta_1$  at 0.9 and  $\beta_2$  at 0.999. The batch size of the data during training was 4 with cross entropy as the loss function, and the maximum training epoch was 35.

To improve the generalisation performance of the model and reduce overfitting, we introduced 3D-ECA attention module after each convolutional layer. This mechanism help the model to better learn the key features in the input data and alleviated the problem of computational redundancy during training.



**Figure 4:** 3D-ResNet18 model architecture.2A: Stacking input model data into three-channel input data;2B: Resampling the image to a size of  $3 \times 200 \times 200 \times 120$  voxels, using residual structures to construct the 3D-ResNet model for feature extraction, transforming features into vectors, and performing glioma classification through fully connected layers;2C: ResNet network residual block structure.3D-ResNet: 3D Residual neural network; 3D-ECA: 3D-ECA Attention Module; LGG: Low-grade Glioma; HGG: High-grade Glioma; X: Data Input; Relu: Activation function.

## RESULTS

The 3D-ResNet18 model we used was developed on the PyTorch 2.1.1+cu118 platform based on Python 3.9 and trained on an NVIDIA-RTX-3060TI graphics processing unit. After the training was completed, we validated the model on the test set and on an independent test set. When evaluating the model performance, we focused on macro F1 scores, Accuracy (ACC), and Receiver Operating Characteristic (ROC) curves to comprehensively assess the efficacy of the model.

When maximum accuracy was obtained on the test set data, we preserved all weight parameters of the model and performed additional validation on the 50 per cent discounted cross-validation set data to ensure that the model performs well on unseen data as well. This comprehensive evaluation approach helps to validate the robustness and generalisation of the model.

## EVALUATION CRITERIA

In the performance evaluation of glioma classification results on BraTS 2019 and BraTS 2020 datasets, Accuracy, Precision, and Recall were taken to evaluate the quality of the results.

This is expressed as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

where TP denotes true positive, TN denotes true negative, FP denotes false positive and FN denotes false negative.

## DIAGNOSTIC EFFICACY OF THE MODEL

Figure 5 showed the confusion matrices obtained using the 3D-ResNet18 model in the five-fold cross-validation set, where the HGG group had the highest classification accuracy in the test sets, and the worst performance is in the test set LGG. In the one-fold validation set with the highest accuracy, the model correctly classified 50 out of 60 cases of LGG data with an ACC of 83.3% and 54 out of 58 cases of HGG data with an ACC of 93.1%.

The performance of this model in different datasets was shown in Table 2. The average area under the curve (AUC) and ACC of 3D-Res Net18 model in the training set were 1.00 and 97.8%, respectively; the average AUC and ACC in the validation set are 0.88 and 83.2%, respectively, as shown in Figure 6.

## 3D-ECA Ablation Experiments

In this study, to verify the effectiveness of the proposed improved method, we conducted ablation experiments and compared the classification results of

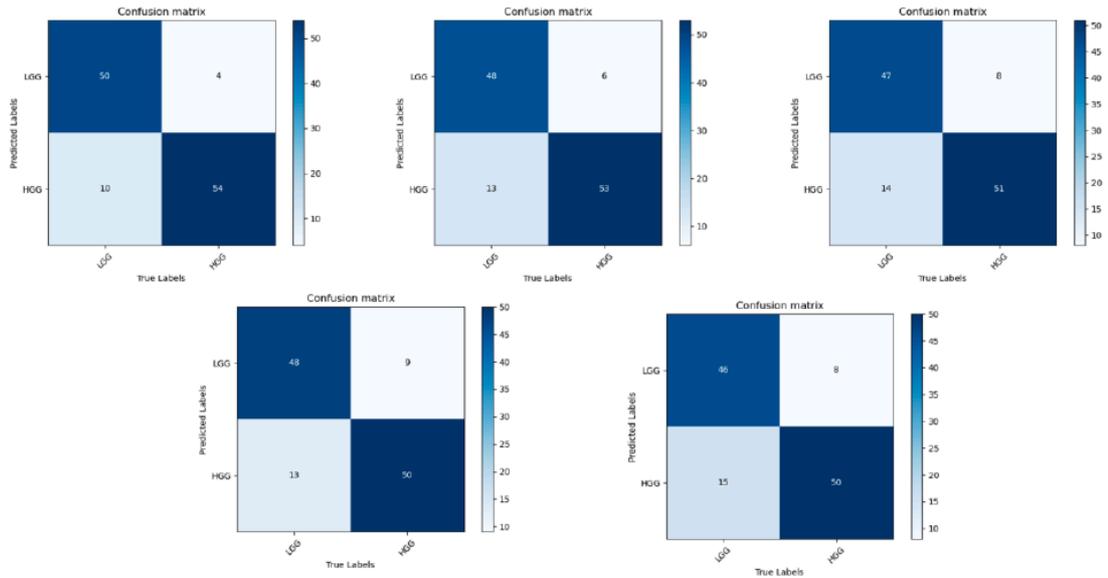


Figure 5: Confusion Matrix in the Five-fold Validation Set.

Table 2: Performance of 3D-ResNet18 Model

Data Set	Mean AUC	Mean ACC/%
Training Set	0.98(0.98-1.00)	99(98-100)
Independent cross validation set	0.88(0.86-0.92)	83.2(80.6-88.1)

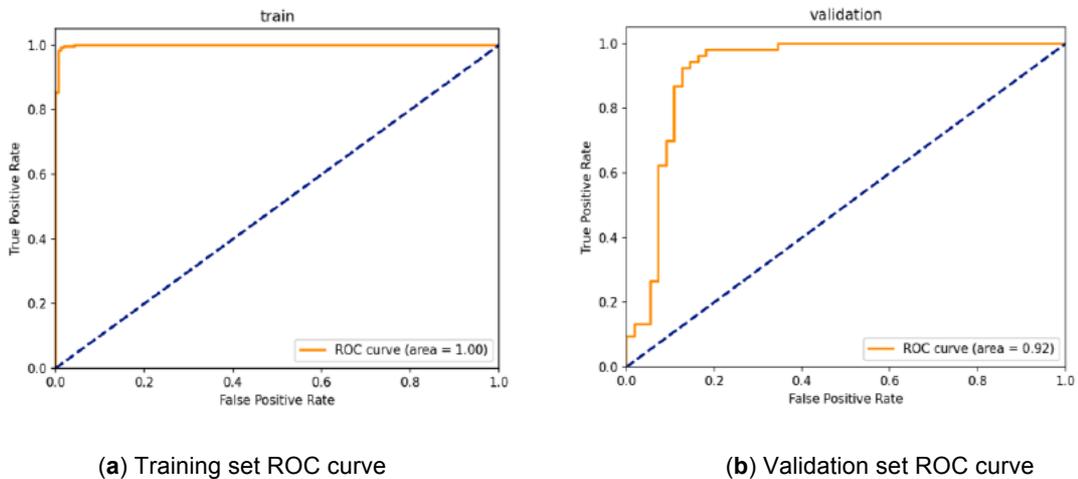


Figure 6: ROC curve of 3D-ResNet18 model.

different combination methods. In the experiments, we adopted 3D-Resnet as the base training method and introduced 3D-ECA attention module to further improve the model performance.

First, we compared the classification results without introducing 3D-ECA attention module with the results after introducing 3D-ECA attention module. For the classification task of tumour grade, we evaluated the performance for HGG and LGG respectively.

As shown in Figure 7, further analyses showed that the model exhibits faster convergence during training

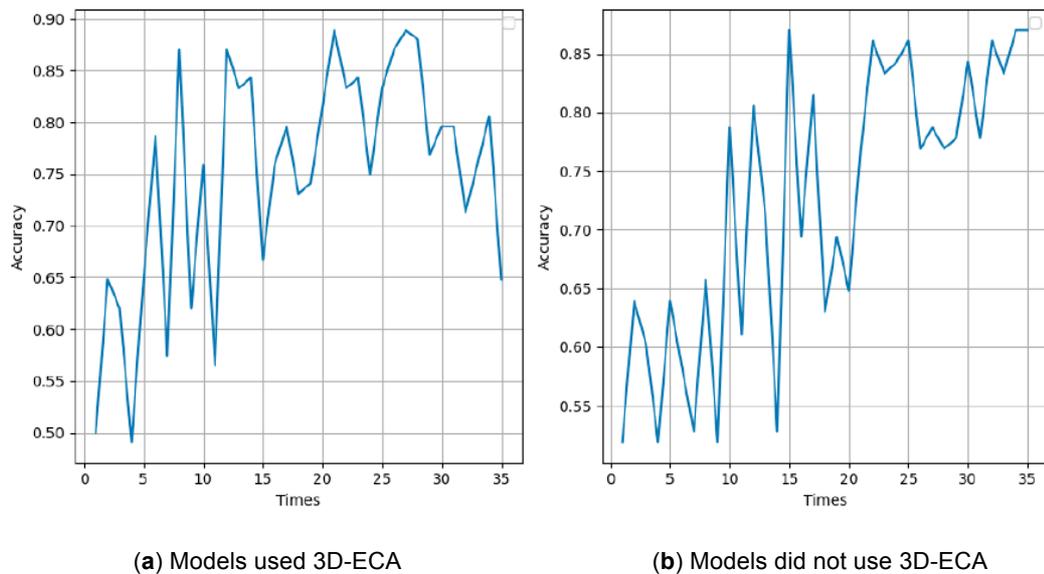
with the introduction of 3D-ECA attention module. The accuracy curve showed that the model using 3D-ECA attention module obtains higher accuracy faster. This indicated that 3D-ECA attention module had a significant positive impact on the performance of the model.

**DISCUSSION**

Gliomas are the most common malignant primary brain tumours in adults originating from glial cells [9]. High-grade gliomas (HGG) are characterised by high aggressiveness and poor prognosis, whereas

**Table 3: Comparison of Glioma Classification Performance of Different Training Methods**

Training Methods	Tumour Grade	Precision/%	Recall/%	Accuracy/%
3DResnet	HGG	80.4	84.9	82.4
	LGG	84.6	80.0	
3DResnet +3D-ECA	HGG	84.3	93.1	88.1
	LGG	92.6	83.3	

**Figure 7:** Accuracy curve of the model when incorporating 3D-ECA or not.

low-grade gliomas (LGG) have a relatively better prognosis and slightly longer survival, but are prone to recurrence, and their grades may progressively increase over time [10]. Currently, histopathology remains the gold standard for tumour grading, however, obtaining tumour tissue samples still relies on invasive surgical procedures or puncture biopsies, which can lead to inherent sampling errors and inter-observer variability in pathohistological analysis. Studies have shown that there is still uncertainty in the pathological diagnosis in 7%-15% of patients, and a combination of imaging is required to describe tumour heterogeneity [11, 12]. Therefore, non-invasive MRI examination techniques play a key role in the diagnosis of glioma grading.

The limitations of various current imaging techniques provide opportunities for deeper image feature extraction and analysis, and imaging histology and deep learning can non-invasively predict glioma pathological grading by high-throughput extraction of deep features such as morphology, organisation and function in voxel-level MRI images. Deep CNN eliminates the steps of feature computation and capturing key features from image features compared to imaging histology, and several studies have shown

that deep learning outperforms imaging histology in predicting glioma pathological grading [13, 14]. The 3D CNN prediction model constructed based on T1WI, T2WI, and FLAIR images in this study has high diagnostic efficacy, which proves the robustness of this deep learning model, and can be used as an effective method for non-invasive prediction of glioma grading in the clinical preoperative period.

#### APPLICATION VALUE

ResNet is one of the most popular deep learning architectures and currently Naser *et al.* [6] collected pre-T1 enhancement, FLAIR and post-T1 enhancement MRI images of 110 patients with lower grade gliomas for tumour grading based on migration learning of pre-trained convolutional bases and fully connected classifiers based on Vgg16. The accuracy, sensitivity and specificity of the grading model to classify lower grade gliomas into grades II and III were 0.95, 0.97 and 0.98, respectively.

In comparison, the difference of our study is that the images input to the model are multimodal MRI images, *i.e.*, T1WI, T2WI, and FLAIR images, instead of inputting individual images, which helps to obtain richer

from different image modalities information from different image modalities, which helps to comprehensively capture and express the multifaceted features of the tumour. This is different from single image input, which can provide more comprehensive and multidimensional data, helping to identify and differentiate different types of brain tumours more accurately.

Furthermore, we further enhance the advantages and value of constructing predictive models from multimodal MRI images by introducing 3D-ECA attention module, which helps the model to better understand and utilise the information associations between different MRI sequences, and improves the model's attention and weight allocation to important features. In this way, the model can more accurately capture the key features in the image, which in turn improves the diagnostic accuracy and robustness of the tumour.

In addition, previous studies have transformed MRI images into JPG format scene images, and although this method, although obtaining better efficacy, loses the correlation information between layers of the MRI image, preventing further optimisation of the model. Therefore, we replace all 2D CNNs with 3D CNNs to make full use of the correlation information between up and down, left and right, and front and back in the 3D space of MRI image data. In addition, 3D network can improve the accuracy of model classification and make the model easier to train by excluding irrelevant voxels [16-20]. Zhuge *et al.* [12] proposed 2D mask R-CNN and 3D conv Net methods to grade gliomas, respectively, and their experimental results similarly showed that the 3D CNN method outperforms the former, which is also in line with the results of the model we constructed which is also consistent with the results of the model we constructed.

### LIMITATIONS OF THIS STUDY

Our study has some limitations that need to be improved in future studies: the AUC of the model is lower than that of the training set in the 50-fold cross-validation set because the neural network is overfitting. In the future, we plan to introduce more samples to increase the robustness and generalisation ability of the model.

In this study, there was an imbalance between HGG and LGG sample data, mainly due to the high prevalence of HGG in epidemiology. To further improve the robustness of the model, we plan to expand more sample data for LGG to better balance the dataset.

Although we have achieved some results, awareness of these limitations will help guide future

research directions to improve the reliability and applicability of the model.

### CONCLUSION

In summary, our study demonstrates that the 3D-Resnet18 deep learning model based on multimodal MRI and attentional mechanisms used to differentiate between high- and low-grade brain gliomas has high accuracy, which provides a new method for non-invasive prediction of preoperative glioma pathological grading, with a view to assisting clinicians in preoperative decision-making and improving patients' prognostic survival, and also potentially assuring the further entry of deep learning into clinical practice.

### CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

### REFERENCE

- [1] Jiang T, Nam D H, Ram Z, *et al.* Clinical practice guidelines for the management of adult diffuse gliomas[J]. *Cancer Lett*, 2021, 499:60-72. <https://doi.org/10.1016/j.canlet.2020.10.050>
- [2] Louis D N, Perry A, Reifenberger G, *et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System:a summary[J]. *Acta Neuropathol*, 2016, 131(6): 803-820. <https://doi.org/10.1007/s00401-016-1545-1>
- [3] Su C Q, Lu S S, Han Q Y, *et al.* Intergrating conventional MRI,texture analysis of dynamic contrast-enhanced MRI, and susceptibility weighted imaging for glioma grading[J]. *Acta Radiol*, 2019, 60(6): 777-787. <https://doi.org/10.1177/0284185118801127>
- [4] Zeineldin, R. A., Karar, M. E., Coburger, J., Wirtz, C. R. & Burgert, O. Deep Seg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images. *Int. J. Comput. Assist Radiol. Surg.* 15, 909-920. <https://doi.org/10.1007/s11548-020-02186-z>
- [5] Ahammed Muneer K V, Rajendran V R, K P J. Glioma Tumor Grade Identification Using Artificial Intelligent Techniques[J]. *J Med Syst*, 2019, 43(5): 113. <https://doi.org/10.1007/s10916-019-1228-2>
- [6] Naser MA, Deen MJ. Brain tumor segmentation and grading of lowergrade glioma using deep learning in MRI images[J]. *Comput Biol Med*, 2020, 121:103758. <https://doi.org/10.1016/j.combiomed.2020.103758>
- [7] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [8] Wang Q, Wu B, Zhu P, *et al.* ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks.[J]. *CoRR*,2019,abs/1910.03151. <https://doi.org/10.1109/CVPR42600.2020.01155>
- [9] Zha C, Meng X, Li L, *et al.* Neutrophil extracellular traps mediate the crosstalk between glioma progression and the tumor microenvironment via the HMGB1/RAGE/IL-8 axis[J]. *Cancer Biol Med*, 2020, 17(1): 154-168. <https://doi.org/10.20892/j.issn.2095-3941.2019.0353>
- [10] Mzoughi H, Njeh I, Wali A, *et al.* Deep Multi-Scale 3D Convolutional Neural Network(CNN)for MRI Gliomas Brain Tumor Classification[J]. *J Digit Imaging*, 2020, 33(4):

- 903-915.  
<https://doi.org/10.1007/s10278-020-00347-9>
- [11] Singh G, Manjila S, Sakla N, *et al.* Radiomics and radiogenomics in gliomas: a contemporary update[J]. *Br J Cancer*, 2021, 125(5): 641-657.  
<https://doi.org/10.1038/s41416-021-01387-w>
- [12] Zhuge Y, Ning H, Mathen P, *et al.* Automated glioma grading on conventional MRI images using deep convolutional neural networks[J]. *Med Phys*, 2020, 47(7): 3044-3053.  
<https://doi.org/10.1002/mp.14168>
- [13] Yang Y, Yan L F, Zhang X, *et al.* Glioma Grading on Conventional MR Images: A Deep Learning Study With Transfer Learning[J]. *Front Neurosci*, 2018, 12: 804.  
<https://doi.org/10.3389/fnins.2018.00804>
- [14] Li Y, Wei D, Liu X, *et al.* Molecular subtyping of diffuse gliomas using magnetic resonance imaging: comparison and correlation between radiomics and deep learning[J]. *Eur Radiol*, 2022, 32(2): 747-758.  
<https://doi.org/10.1007/s00330-021-08237-6>
- [15] Shin I, Kim H, Ahn S S, *et al.* Development and Validation of a Deep Learning-Based Model to Distinguish Glioblastoma from Solitary Brain Metastasis Using Conventional MR Images[J]. *AJNR Am J Neuroradiol*, 2021, 42(5): 838-844.  
<https://doi.org/10.3174/ajnr.A7003>
- [16] Bangalore Yogananda C G, Shah B R, Vajdani-Jahromi M, *et al.* A novel fully automated MRI-based deep-learning method for classification of IDH mutation status in brain gliomas[J]. *Neuro Oncol*, 2020, 22(3): 402-411.  
<https://doi.org/10.1093/neuonc/noz199>
- [17] YU J, Yang B, Wang J, *et al.* 2D CNN versus 3D CNN for false-positive reduction in lung cancer screening[J]. *J Med Imaging(Bellingham)*, 2020, 7(5): 051202.  
<https://doi.org/10.1117/1.JMI.7.5.051202>
- [18] Zeineldin, R.A., Karar, M.E., Elshaer, Z. *et al.* Explainable hybrid vision transformers and convolutional network for multimodal glioma segmentation in brain MRI. *Sci Rep* 14, 3713 (2024).  
<https://doi.org/10.1038/s41598-024-54186-7>
- [19] Babu Vimala, B., Srinivasan, S., Mathivanan, S.K. *et al.* Detection and classification of brain tumor using hybrid deep learning models. *Sci Rep* 13, 23029 (2023).  
<https://doi.org/10.1038/s41598-023-50505-6>
- [20] Ranjbarzadeh, R., Bagherian Kasgari, A., Jafarzadeh Ghouschi, S. *et al.* Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. *Sci Rep* 11: 10930 (2021).  
<https://doi.org/10.1038/s41598-021-90428-8>

---

Received on 14-04-2024

Accepted on 23-05-2024

Published on 30-05-2024

DOI: <https://doi.org/10.12974/2313-1047.2024.11.02>

© 2024 Yuanzhe *et al.*

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.