



Image Classification of Brain Tumors Using Improved CNN Framework with Data Augmentation

Xin Ning¹, Zhanbo Li², and Haibo Pang¹(✉)

¹ School of Software Technology, Zhengzhou University, Zhengzhou 450002, China
innerpeacecx@163.com, panghbz@163.com

² Department of Network Management Center, Zhengzhou University, Zhengzhou
450001, China
iezbli@zzu.edu.cn

Abstract. At present, the problem of shortage of medical human resources can be solved through mobile medical equipment, the main method to improve the diagnostic performance of mobile medical equipment is to improve the accuracy of the algorithm. Brain tumor classification is to determine the tumor type of patients. The accurate brain tumor classification algorithm can improve the diagnostic performance of mobile medical equipment while assisting doctors in diagnosis. This paper proposes a multi-grade brain classification system using improved CNN framework with extensive data augmentation for differentiating among glioma, meningioma and pituitary tumors, which from three prominent types of brain tumor. First, we locate the tumor and extract the region of interest (ROI). Secondly, to solve the problem of insufficient data samples in the brain tumor classification, we use data augmentation techniques to augment the data samples. Thirdly, VGG-19 and Inception V3 model are improved, and the CNN model is optimized by Adam algorithm. Finally, the improved CNN framework is trained and classified with augmented dataset. Experimental results show that the system proposed in this paper based on data augmentation and improved CNN framework has better classification performance than traditional classifier, and this system can effectively solves the problem of low accuracy caused by insufficient data samples.

Keywords: Mobile medical equipment · Multi-grade brain tumor classification · Data augmentation · CNN · MRI · Adam

1 Introduction

Mobile medical equipment uses mobile communication technology and mobile internet to provide medical services and information. Mobile medical treatment can help improve medical problems. Accurately diagnosing patients through mobile medicine is one of the main problems at present. Brain tumor is one of the most dangerous cancers, the

World Health Organization (WHO) divides brain tumors into four grades I (benign) to IV (malignant). Accurately diagnosing the type of brain tumor can help doctors treat patients. The characteristics of malignant tumors are usually determined by histopathology, but different pathologists have different views on these characteristics, so magnetic resonance imaging (MRI) is introduced. MRI provides a higher contrast for imaging of brain soft tissue.

Deep learning is widely used in brain tumor segmentation, classification, and localization of lesions. Laukamp et al. [1] established a deep learning model that can perform full-automatic detection and segmentation of meningioma against two different MRI images (FLAIR and T1CE). Kamnitsas et al. [2] proposed a dual pathway, 11-layers deep, three-dimensional Convolutional Neural Network for the challenging task of brain lesion segmentation. The author devised an efficient and effective dense training scheme which joins the processing of adjacent image patches into one pass through the network. This method improved the computational efficiency of 3D medical scan processing. Kharrat et al. [3] proposed a hybrid scheme that uses genetic algorithms (GA) and support vector machines (SVM) to classify brain tumor into normal, benign and malignant. In order to solve the problem of insufficient and unbalanced clinical data of meningioma images, Zhu et al. [4] used the LeNet-5 network and oversampling technology to classify meningioma into three categories. Afshar et al. [5] proposed a CapsNets for brain tumor in addition. Zia et al. [6] proposed a generalized classification system for brain tumor based on rectangular window image clipping. This system utilized discrete wavelet transform and PCA, and used SVM as the classifier; the system had low computational complexity, and it was conducive to the development of a universal CAD system. It can be used in any clinical or institution to help radiologists to understand brain tumor MRI images.

Transfer learning is a common method for training large-scale models. The small amount of data can easily cause the model overfitting, and transfer learning is often used to solve such problems. Swati et al. [7] proposed a fine-tuning strategy based on transfer learning using a pre-trained deep convolutional neural network, which performed well on the CE-MRI dataset with an average accuracy of 94.82%. Deepak et al. [8] used pre-trained GoogLeNet to extract features from MRI images by using transfer learning to classify three brain tumors. The research in this paper showed that transfer learning was effective in solving the shortage of data.

Data augmentation is mainly divided into image processing techniques and generative adversarial networks (GAN) [9]. Chang et al. [10] used residual CNN to determine the status of Isocitrate Dehydrogenase (IDH) in low and high gliomas. To increase the size of the training set and prevent overfitting, the author augmented the training set images by introducing random rotations, translations, flips, shearing, and zooming. Han et al. [11] proposed a progressive growing of generative adversarial networks (PGGAN) to solve the problem of insufficient data in real image distribution.

Glioma, meningioma and pituitary tumor have the highest incidence among all brain tumors. For these three types of brain tumors with the highest incidence, how to classify them accurately is a hot topic of current research. The main problems now are as follows: 1) Large-scale labeled medical image datasets are difficult to obtain. 2) Training a large-scale deep learning model with a small amount of data easily leads to model overfitting.

3) Directly using classic deep learning networks for medical image processing may lead to low accuracy of the model, so it is necessary to improve the model.

In this paper, we propose a new framework to classify three pathologic types of brain tumors (glioma, meningioma and pituitary tumor). First, we segment the tumor ROI from the MRI image. Then we use data augmentation technique to augment the data samples. In this paper, we compare two data augmentation methods, one based on image processing technique and the other based on deep convolutional generative adversarial networks (DCGAN) [12]. Finally, we improve the performance of the model by adjusting the network framework and parameters of the model, and use the Adam [13] algorithm to optimize the improved model. The major contributions of this paper are listed below:

1. Due to the lack of public brain tumor datasets, we use two data augmentation methods to augment the data, One is using different image processing techniques to add noise and transformation to the data, and another is using DCGAN to generate new data, and we analyze the impact of the two methods on the model performance.
2. Two improved deep CNN models are proposed. The first model is the improved VGG-19 model, and we improve its network structure. The second model is the improved the Inception V3 model, we use inductive transfer learning to improve it. And using Adam algorithm to optimize the two models. The improved two models have a stronger ability to learn the features of brain tumor data.
3. The effects of different learning rate and batch size on model accuracy and convergence speed are studied. The influence of parameters in the Adam algorithm on the model convergence speed is studied.
4. The two improved models have high specificity, both exceeding 90%, which can effectively judge the authenticity of negative patients.
5. Our diagnostic system has a higher classification accuracy than other methods, so the system is suitable for mobile medical equipment, it can improve the diagnostic accuracy of doctors, and can be used for the positive diagnosis.

2 Method

The design framework of this paper mainly includes three parts: 1) ROI segmentation. 2) Data augmentation. 3) Training model and classification. The system framework is shown in Fig. 1.

The first step is to locate the tumor area of the sample data in the original dataset, segment the tumor, and then obtain the ROI image of the tumor. The second step is to augment the dataset. Expanding a image of the original dataset to 32 images. The size of the datasets after the two data augmentation methods are the same. The third step is to improve and optimize two classic deep learning CNN models, the VGG-19 model and the Inception V3 model. Then using the augmented datasets to train two improved models and classify brain tumor.

2.1 Brain Tumor ROI Segmentation

Accurately locating ROI or lesions in MRI images is a key part of a doctor's diagnosis. The experiments in this paper use the publicly available figshare dataset [14], this dataset

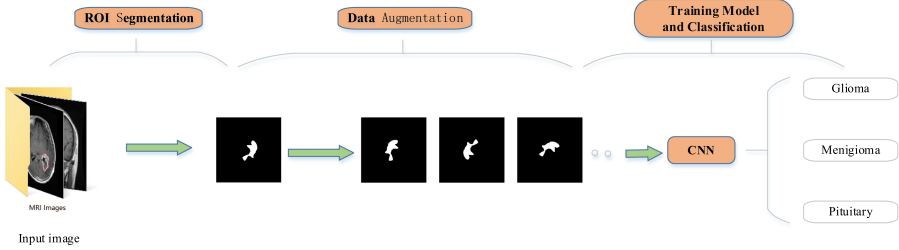


Fig. 1. System framework

contains the tumor boundary coordinates of each data sample, as well as a binary image labeled with the tumor area. We use the threshold segmentation technique proposed by Otsu et al. [15] to segment the ROI of 3 types of brain tumors, and the binary images are generated from the original MRI grayscale images, and white represents the tumor area, as shown in Fig. 2.

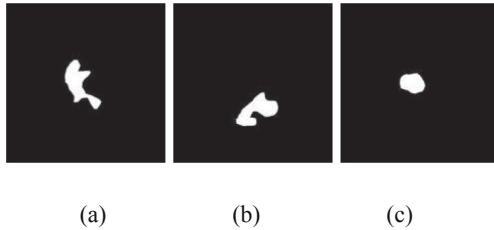


Fig. 2. Images of three brain tumors after ROI segmentation (a: Glioma. b: Meningioma. c: Pituitary tumor.)

2.2 Data Augmentation

The quality and quantity of training samples are important factors that affect the classification performance of deep learning models. In this paper, we use image processing techniques and DCGAN to augment the dataset.

Data Augmentation Based on Image Processing Technique

Xue et al. [16] used fixed scaling, rotation plus flip and shearing to augment the images. Training deep CNN requires a large-scale dataset, so we use a variety of image processing techniques to achieve data augmentation. A total of 9 image processing techniques, including 5 methods of data augmentation through geometric change techniques: flipping, rotation, shears, skewness and distortions. Considering that when the number of samples increases to a certain extent, it will affect the performance of the model and reduce the robustness of the model. We need to add noisy data samples to the dataset to verify the robustness of the model. So we add noise to the data samples through the four noise addition techniques of gaussian blur, sharpening, edges detection and median blur. Table 1 shows the above 9 data augmentation techniques and 32 parameters, the

Table 1. Different data augmentation techniques with their respective parameters

S. no.	Data augmentation technique	Parameters	
1	Rotation (angle)	90°	
		45°	
		-45°	
		-90°	
		180°	
2	Flip	Level	
		Vertical	
3	Shear	Left 15°	
		Right 15°	
4	Skew	Bottom	0.5
			1.0
		Right	0.5
			1.0
5	Gaussian blur (σ)	0.5	
		1.0	
		1.5	
		2.0	
6	Distortions (grid)	Width = 5.0 height = 10.0	
		Width = 10.0 height = 5.0	
7	Sharpen (lightness)	1.0	
		2.0	
8	Median blur (kernel)	5 × 5	
		7 × 7	
		11 × 11	
9	Edge detection	Direction = 0.5	$\alpha = 0.25$
			$\alpha = 0.50$
			$\alpha = 0.75$
			$\alpha = 1.0$
		Direction = 1.0	$\alpha = 0.25$
			$\alpha = 0.50$
			$\alpha = 0.75$
			$\alpha = 1.0$

second column refers to augmentation techniques and the third column is the parameters of each technique. Here are a total of 32 parameters, that is, one sample in the original dataset can be expanded to 32 samples.

Data Augmentation Based on DCGAN

GAN is used to generate new data samples, while DCGAN applies the concept of CNN to GAN. In DCGAN, the discriminant network is a CNN. After several layers of convolution, the input image gets a feature, which is then inputted into the logistic function. The output of the function can be regarded as a probability. Compared with the GAN, DCGAN has the following advantages: 1) The network framework of DCGAN is relatively stable. 2) DCGAN can generate high-quality images. We use DCGAN to augment the dataset. During training, when it is found that some abnormal images are generated, these images will be deleted. The data samples generated after training are shown in Fig. 3.

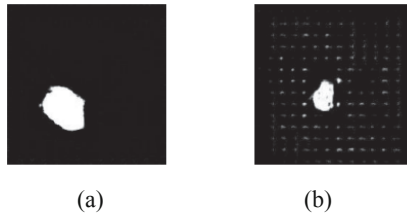


Fig. 3. The images generated by DCGAN (a: Normal image. b: Anormal image.)

2.3 Network Model Optimization

In this paper, we improve two deep learning networks and extract features and classification. First, the VGG-19 [17] network and the Inception V3 [18] network are improved respectively, and then adam is used to optimize the two models.

VGG-19 Network Improvement

VGG-19 network consists of 19 weighted layers, in which there are 16 convolutional layers (CONV) and 3 fully connected layers (FC), as shown in Fig. 4 (a). The improved network framework in this paper is 17 weighted layers, with 15 convolutional layers and 2 fully connected layers, as shown in Fig. 4 (b).

The specific improvement steps are as follows: 1) In the original VGG-19 network, a convolutional layer with 64 channels is added after the second convolutional layer, and a convolutional layer with 128 channels is added after the fourth convolutional layer. The purpose is that because the segmented ROI image is a binary grayscale image, the contour features of the tumor are very important, so adding a convolutional layer near the input layer can better extract low-order features. 2) Deleting the 11th, 12th, 15th, and 16th convolutional layers, which can speed up model convergence. 3) Deleting the last full connected layer and using a convolutional layer with output 3 as the new full connected

layer, the size of the convolutional kernel is 1×1 . And in order to prevent and reduce overfitting, dropout [19] optimization scheme is added after 2 fully connected layers and the last 1 convolutional layer. BP neural network algorithm is used to update and learn the parameters in the CNN framework. After the completion of one epoch, some neurons will be deleted randomly during the next epoch until the training is completed. Figure 4 shows the VGG-19 network framework before and after improvement.

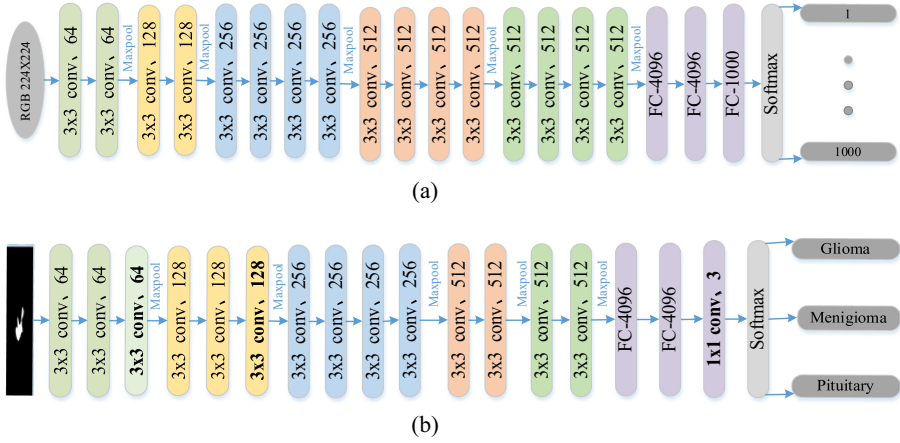


Fig. 4. VGG-19 network framework before and after improvement (a: Original VGG-19 network framework. b: Improved VGG-19 network framework)

Inception V3 Network Improvement

The improvement of Inception V3 in this paper uses inductive transfer learning. In inductive transfer learning, the learning tasks of the source domain and the target domain are different, and the feature space and the edge probability distribution of the source domain and target domain are the same. Figure 5 shows the principle of inductive transfer learning in this paper. We use the imagenet dataset to train the original Inception V3 model, and transfer the learned knowledge to the classification model in this paper. The source domain is the imagenet dataset, the target domain is the figshare dataset, the source task is 1000-class classification, and the target task is the 3-class brain tumor classification. The knowledge learned from the source domain is the feature of the natural images in the imagenet dataset, such as the color feature and shape feature in its sample.

The improvements to the Inception V3 model in this paper are as follows: 1) Deleting the original fully connected layer of the Inception V3 network and reinsert a new fully connected layer with an output size of three to adapt it to the target domain. Deleting the loss function in the original classification layer and use the new loss function. 2) Using the method of transfer learning, and low-level features of the data in the imagenet dataset are learned by the pre-trained layers from the original Inception V3 model, then using the segmented and preprocessed data samples of the figshare brain tumor dataset to train the improved Inception V3 model for the experiment. We set the learning factors for weights and bias at the new fully connected layer to 10, the purpose is to make the

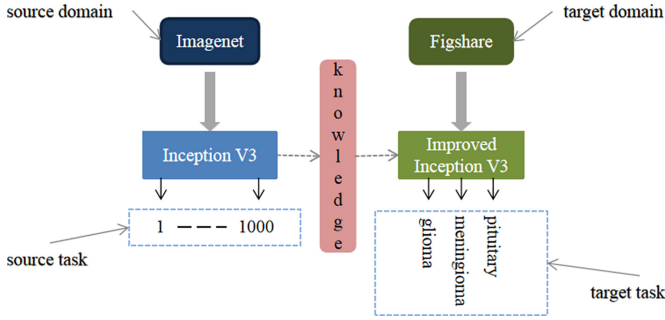


Fig. 5. Model for transfer learning

model learn specific high-level features of the target domain. The convolutional layer of the model is unchanged, only the structure after the FC layer is changed. The Inception V3 network framework before and after the improvement is shown in Fig. 6.

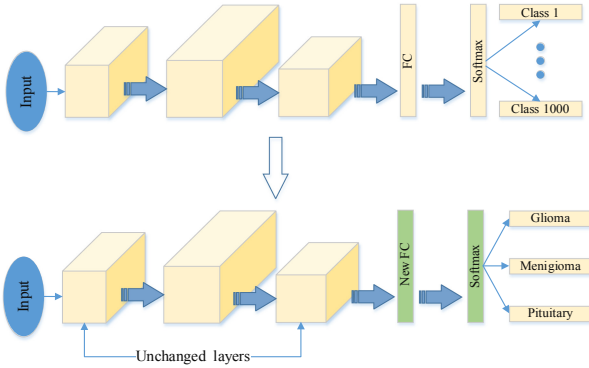


Fig. 6. Improved the inception V3 network framework

Figure 7 shows the improved Inception V3 model classification framework. When training, the model learns classification based on the learned training features and training labels. During testing, the features of the test features are input into the classifier, and the model predicts the test labels. Then calculating the classification accuracy of the test by the actual test label.

Optimizing CNN Networks with Adam

In this paper, we use Adam to optimize two improved models. Traditional CNN networks use stochastic gradient descent (SGD) in the BP framework to minimize the loss function. SGD randomly selects samples to update the gradient at each epoch. During the training process, it will maintain a single learning rate, which will cause each epoch to be not the optimal direction and will cause the loss function to oscillate severely, which will cause the accuracy of the model to decrease. We train the improved VGG-19 model on the original dataset with the SGD optimizer and the Adam optimizer, respectively. The

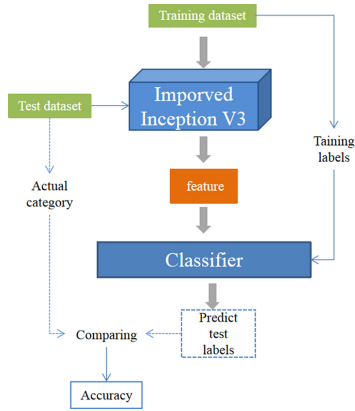


Fig. 7. Improved inception V3 classification framework

average loss function values of the first ten epochs are shown in Fig. 8. In the ten epochs, the changing trend of the Loss value of the SGD optimizer is not obvious, and the Loss value of the Adam optimizer drops rapidly at the third epoch and continues to decline thereafter.

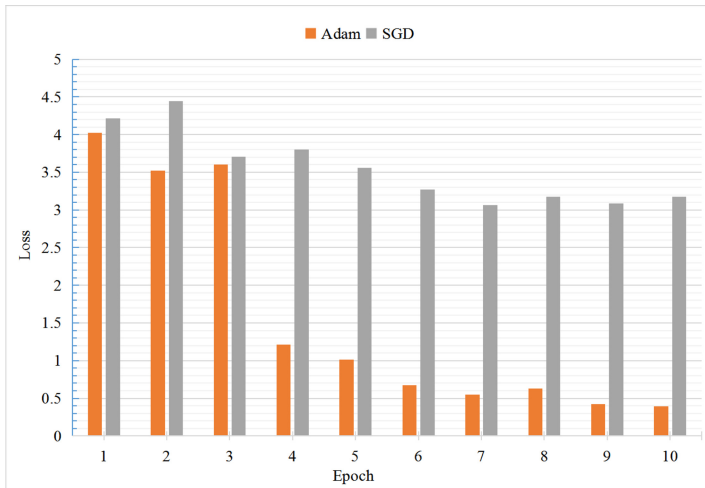


Fig. 8. Comparing the performance between the Adam optimizer and the SGD optimizer

According to Fig. 8, it can be seen that the SGD decline rate is slower than Adam, so we use Adam instead of SGD to minimize the loss function. Adam calculates the adaptive learning rate of each parameter. The process of network parameter update is as follows:

$$g_t = \frac{1}{n} \nabla_{\theta} \sum_i L(F(Y_i, \theta), X_i) \quad (1)$$

$$m_t = u \times m_{t-1} + (1 - u) \times g_t \quad (2)$$

$$n_t = v \times n_{t-1} + (1 - v) \times g_t^2 \quad (3)$$

First, the Adam initializes the parameter vector, the first moment vector and the second moment vector, and then it will iteratively update each part to make the parameter θ converge. In Eqs. (1), (2) and (3), $L(\theta)$ is the mean square error function, g_t is the gradient function of $L(\theta)$ against θ . m_t is the first moment estimate and n_t is the second raw moment estimate. u and v are exponential decay rates for the moment estimates. At time $t + 1$, updating biased first moment estimate and biased second raw moment estimate.

$$\hat{m} = \frac{m_t}{1 - u_t} \quad (4)$$

$$\hat{n} = \frac{n_t}{1 - v_t} \quad (5)$$

Next, computing bias-corrected first moment estimate and bias-corrected second raw moment estimate. In Eqs. (4) and (5), \hat{m}_t and \hat{n}_t are deviation corrections of m_t and n_t , respectively.

$$\Delta\theta_t = -\eta \times \frac{\hat{m}_t}{\sqrt{\hat{n}_t + \varepsilon}} \quad (6)$$

$$\theta_{t+1} = \theta_t + \Delta\theta_t \quad (7)$$

Finally, the gradient of the objective function for the parameter θ is updated at time $t + 1$, then updating the parameter θ of the model with the computed value above. In Eqs. (6) and (7), η is the step size, ε is the numerically stable small constant. $\Delta\theta_t$ is the updated value of θ_t at time t . At time $t + 1$, adding $\Delta\theta_t$ and θ_t to get the value of θ for θ_{t+1} .

3 Experiments

All experiments are evaluated using NVIDIA GeForce RTX-2080Ti with 64 GB onboard memory and deep learning framework Tensorflow that are installed on Ubuntu 18.04.

3.1 Dataset and Pre-processing

The figshare dataset contains 3064 brain MRI images of 233 patients, and the images belong to the T1-CE MRI. It contains 1,426 MRI images of glioma (89 patients), meningioma contains 708 images (82 patients), and pituitary tumor contains 930 images (62 patients). The resolution of the image is 512×512 . The gap between each slice is 1 mm, and the slice thickness is 6 mm. The data is augmented by the data augmentation techniques in Table 1 and DCGAN. In order to compare the impact of the two data augmentation techniques on model performance, the size of the datasets after the expansion of the two data augmentation techniques is the same. After data augmentation, the total number of samples in this dataset is increased from 3064 to 98048, of which glioma images is increased from 1426 to 45632, meningioma images is increased from 708 to 22656, and pituitary tumor images is increased from 930 to 29760.

3.2 Optimizer Setting

For the Adam optimizer, η is step size, the value of the η parameter affects the size of the update value $\Delta\theta_t$ of θ_t at time t , that is, the parameter η affects the descent direction of the loss function and the accuracy. When the value of η is larger, the model cannot converge, and the loss function will always oscillate. When it is small, the model needs to be iterated many times to reach the global optimum. According to the paper [13], we set the exponential decay rate of the moment estimation $u = 0.9$, $v = 0.999$, and the numerically stable small constant $\varepsilon = 10^{-8}$. In order to analyze the effect of η value on model performance, we set different η values, and then train improved network models on the original dataset. Finally, we compare the experimental results.

3.3 Classifier Setting

The VGG-19 model and the Inception V3 model were originally designed for RGB color images, with an input layer of size $224 \times 224 \times 3$. Because the images of the figshare dataset are grayscale images. First, we adjusted the image size, and then copied the grayscale values three times to create three channels. We used patient-level five-fold cross-validation on the dataset before and after the data augmentation to verify the prediction ability of the model. We used the cross-loss function as the loss function of the model.

Different parameter values will seriously affect the performance of the model. Too large learning rate will prevent the loss function from converging. Too small batch size will also make the model unable to converge and the training time will increase, too much batch size will adversely affect the quality of the model, and the computer hardware configuration cannot meet the calculation requirements.

The experimental process is as follows: First, we set different learning rate values and different batch size values. After the dataset was augmented by DCGAN, we trained two improved models separately, compared the accuracy of the models, and analyzed the impact of learning rate and batch size on model accuracy and loss. Secondly, we used the datasets before and after data augmentation to train the two improved models, then we compared the accuracy of the two improved models before and after data augmentation and analyzed the impact of data augmentation on brain tumor classification. Finally, we compared the performance of the two classifiers before and after data augmentation.

3.4 Evaluation Index

This paper conducted the experiments five times, and each experiment followed a patient-level five-fold cross-validation. The loss of the loss function is the average after five experiments. Based on the correct and incorrect classification of the three types of brain tumor, the precision, recall, accuracy and specificity of each brain tumor were calculated according to Eq. (8), where TP, FP, TN and FN are the number of classified cases of true positives, false positives, true negatives and false negatives, respectively.

$$\begin{aligned}
precision &= \frac{TP}{TP + FP} \\
recall &= \frac{TP}{TP + FN} \\
specificity &= \frac{TN}{TN + FP} \\
accuracy &= \frac{TN + TP}{TP + TN + FP + FN}
\end{aligned} \tag{8}$$

4 Analysis of Experimental Results

4.1 Effect of η Parameter on Adam Optimizer

When the two models use Adam optimizer with the original dataset, we set different η values. After 10 epochs, comparing the accuracy and loss of the two models, as shown in Table 2. The experiment result shows: 1) For the improved VGG-19 model, the accuracy is higher when η is 10^{-4} , the loss of the loss function is the smallest, and the convergence speed is the fastest. 2) When the improved Inception V3 model takes different values, the accuracies are not high after 10 epoch. Because the Inception V3 model has more network layers and fewer data samples in the original dataset, so fewer features are learned during training. When the number of epochs is small, the accuracy is low. 3) In the improved Inception V3 model, the accuracy and loss are close when η is 10^{-4} and η is 10^{-5} . When η is small, the model needs a longer training time, so we set η as 10^{-4} for comprehensive consideration.

Table 2. The accuracy and loss of two improved models with different η after 10 epochs

Improved VGG-19 model			Improved Inception V3 model		
η	Loss	Accuracy	η	Loss	Accuracy
10^{-2}	3.604	0.5964	10^{-2}	7.351	0.2666
10^{-3}	1.174	0.6812	10^{-3}	3.638	0.2667
10^{-4}	0.392	0.7911	10^{-4}	1.110	0.5667
10^{-5}	0.403	0.7829	10^{-5}	1.021	0.5625

4.2 The Impact of Learning Rate and Batch Size on Model Accuracy and Loss

Training the model on the dataset augmented by DCGAN, setting different parameters and different epochs, and compare the accuracy and loss of the two models. As shown in

Table 3, according to the results in the table: 1) when the learning rate of the two models are 1×10^{-4} and the batch size of the two models are 64, the accuracy of the two models are the highest and the loss of the loss functions are the smallest. 2) For the VGG-19 model, when the batch size is 64, the learning rate is 1×10^{-4} , and the accuracy of the model differs greatly, indicating that the VGG-19 model is more sensitive to the learning rate. 3) When the learning rate of the two models is 1×10^{-4} and the batch size is 48, 64 and 128 respectively, the difference in model accuracy is small, and the difference in loss value is large, indicating that batch size affects the direction of gradient descent. 4) Comparing when the learning rate is 1×10^{-5} and 1×10^{-4} , it is found that the smaller learning rate will significantly affect the accuracy of the two models.

Table 3. Comparing the accuracy and loss of two improved models with different parameters (where LR, BZ, and E refer to learning rate, batch size and epoch, respectively)

Improved VGG-19 model					Improved inception V3 model				
LR	BZ	E	Acc (%)	Loss	LR	BZ	E	Acc (%)	Loss
1×10^{-3}	32	200	88.69	1.071	1×10^{-3}	32	200	86.43	0.821
1×10^{-4}	48	80	90.68	0.107	1×10^{-4}	48	80	88.09	0.255
1×10^{-4}	64	50	91.73	0.061	1×10^{-4}	64	50	88.96	0.103
1×10^{-4}	128	50	91.12	0.079	1×10^{-4}	128	50	88.77	0.109
3×10^{-4}	64	50	90.35	0.109	3×10^{-4}	64	50	87.33	0.169
1×10^{-5}	64	50	90.30	0.072	1×10^{-5}	64	50	88.82	0.103

4.3 Accuracy Analysis of the Models Before and After Data Augmentation

In the histograms of Figs. 9 and 10, we can see the accuracy of the two models in the original dataset, the dataset augmented with DCGAN and the dataset augmented with 9 augmentation techniques. In the figures, the abscissa is the number of epochs, a total of 50 epochs, taking 10 epochs out of 50 epochs for comparison, and the ordinate is the accuracy. After 50 epochs, the accuracy of the improved VGG-19 on the original dataset is 0.8710, the accuracy of using DCGAN to augment the dataset is 0.9173, and the accuracy of using 9 data augmentation techniques is 0.9109. And the accuracy of the improved Inception V3 model on the original dataset is 0.8531, the accuracy of using DCGAN to augment the dataset is 0.8896, and the accuracy of using 9 data augmentation techniques is 0.8781. At the 20th epoch, the accuracies of the improved VGG-19 model trained with three datasets all reach above 80%, while the Inception V3 model has only the accuracy of dataset augmented by 9 data augmentation techniques reaches above 80% at the 20th epoch. According to the results, data augmentation is effective for training models, and the accuracy of two models using DCGAN is higher than the accuracy of two models using 9 augmentation techniques.

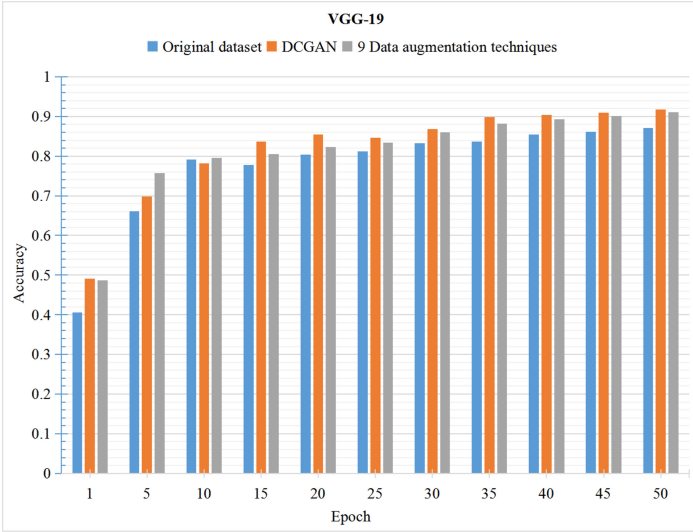


Fig. 9. The accuracy of the improved VGG-19 model before and after data augmentation

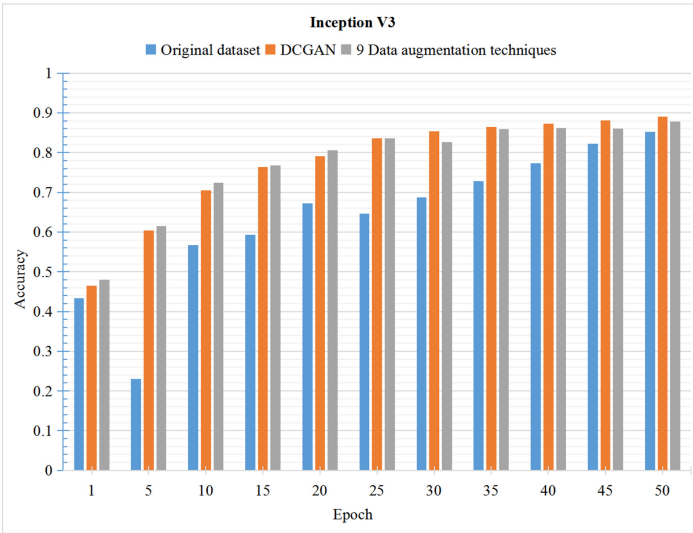


Fig. 10. The accuracy of the improved Inception V3 model before and after data augmentation

4.4 Performance Analysis of Improved VGG-19 Model

Table 4 shows the confusion matrix of the improved VGG-19 model in the original dataset and the dataset augmented using DCGAN, as well as the accuracy. Table 5 shows the classification performance of the improved VGG-19 model before and after data augmentation, including the precision, recall, and specificity of three brain tumors. In Table 4, the accuracy of the improved VGG-19 model before data augmentation is

87.10%. After data augmentation, the accuracy is 91.73%. According to the results in Tables 4 and 5: 1) Before and after data augmentation, the accuracy of the model is increased by about 4.63%, this result indicates that data augmentation has a greater impact on the performance of the improved VGG-19 model. 2) The specificities of the improved VGG-19 model before and after data augmentation are relatively high, both of which are more than 90%, which can effectively judge true negative patients and can be used for the positive diagnosis. 3) Before and after the data augmentation, the precisions of the model for glioma are higher, while the precisions of meningioma and pituitary tumor are lower, which may be caused by the imbalance of the data of the three brain tumors. 4) The precision and recall of meningioma are low, and the proportion of meningioma misclassified is higher than other brain tumors.

Table 4. Confusion matrix of improved VGG-19 model before and after data augmentation (where M, G, and P refer to meningioma, glioma and pituitary, respectively)

Confusion matrix of improved VGG-19 before data augmentation				Confusion matrix of improved VGG-19 after DCGAN					
		Predicted					Predicted		
		G	M	P			G	M	P
Actual	G	1273	81	72	Actual	G	43243	952	1437
	M	60	599	49		M	1369	19484	1803
	P	51	82	797		P	507	2040	27213
Accuracy		87.10%			Accuracy		91.73%		

In the experiment, we compared the accuracies of the model before and after the improvement on the dataset augmented using DCGAN. The accuracy of the VGG-19 model before improvement is 89.81%, and it is increased to 91.73% after improvement. According to the comparison results, the accuracy of the improved model is increased, indicating that the improvement of the VGG-19 model is effective on this dataset. After improvement, the ability to learn features of the model becomes stronger.

Table 5. Evaluation of improved VGG-19 model before and after data augmentation

Evaluation of improved VGG-19 before data augmentation				Evaluation of improved VGG-19 after DCGAN			
Type	Precision	Recall	Specificity	Type	Precision	Recall	Specificity
G	91.98%	89.27%	93.22%	G	95.84%	94.76%	96.42%
M	78.60%	84.60%	93.08%	M	86.69%	86.00%	96.03%
P	86.82%	85.70%	94.33%	P	89.36%	91.44%	95.26%

4.5 Performance Analysis of Improved Inception V3 Model

Table 6 shows the confusion matrix of the improved Inception V3 model in the original dataset and the dataset augmented using DCGAN, as well as the accuracy. Before and after data augmentation, the accuracy of the improved Inception V3 model increased from 85.31% to 88.96%. Table 7 shows the classification performance of the improved Inception V3 model before and after data augmentation. According to the results in Tables 6 and 7: 1) After the data augmentation, the performance of the improved Inception V3 model has been significantly improved, indicating that using data augmentation can improve the performance of the classifier, but the precision and specificity of glioma are reduced after data augmentation. 2) The accuracy of the improved Inception V3 model before and after data augmentation is increased by 3.65%, compared with the improved VGG-19 model, the increased of accuracy is relatively small. 3) The accuracy and specificity of meningioma and pituitary tumor were less than 90% before and after data augmentation. Comparing the results of the VGG-19 model, it can be seen that the data imbalance affects the CNN model to extract features, and the quality and quantity of the dataset seriously affect the classification performance of the model.

Table 6. Confusion matrix of improved Inception V3 model before and after data augmentation (where M, G, and P refer to meningioma, glioma and pituitary, respectively)

Confusion matrix of improved inception V3 before data augmentation					Confusion matrix of improved inception V3 after DCGAN				
		Predicted					Predicted		
		G	M	P			G	M	P
Actual	G	1229	101	96	Actual	G	41613	2335	1684
	M	55	594	59		M	1884	19353	1419
	P	20	119	791		P	842	2663	26255
Accuracy		85.31%			Accuracy		88.96%		

The accuracy of the Inception V3 model before improvement is 86.79%, and it is increased to 88.96% after improvement. According to the comparison results, after the improvement, the accuracy of the model is increased, this can indirectly explain that transfer learning is conducive to training a deeper network model.

4.6 Comparison with Other Methods

This section compares our method with other methods which use the figshare dataset to classify the brain tumor. As shown in Table 8, the last two lines are the accuracy of the improved VGG-19 model and the improved Inception V3 model based on DCGAN data augmentation proposed in this paper. We can see that the improved VGG-19 model proposed in this paper has the highest accuracy. After data augmentation, the accuracy of the improved VGG-19 model is higher than the improved Inception V3 model.

Table 7. Evaluation of improved Inception V3 model before and after data augmentation

Evaluation of improved inception V3 before data augmentation				Evaluation of improved inception V3 after DCGAN			
Type	Precision	Recall	Specificity	Type	Precision	Recall	Specificity
G	94.25%	86.19%	95.42%	G	93.85%	91.19%	94.80%
M	72.97%	83.90%	90.66%	M	79.48%	85.42%	93.37%
P	83.62%	85.05%	92.74%	P	89.43%	88.22%	95.46%

Table 8. Comparison with other methods using the figshare dataset

Method	Accuracy (%)
Afshar [5]	86.56
Zia [6]	85.69
Afshar [20]	90.89
Cheng [14]	91.28
Improved inception V3	88.96
Improved VGG-19	91.73

5 Conclusion

In this paper, we proposed a 3-class brain tumor classification system based on data augmentation and improved CNN framework, our system has high accuracy and specificity, and the system can be applied to mobile medical equipment for the positive diagnosis. In this paper, we solved the problem of insufficient data samples in medical image classification tasks and analyzed the impact of data augmentation based on image processing technique and data augmentation based on DCGAN on model performance. We improved the VGG-19 model and the Inception V3 model, used Adam algorithm to optimize two CNN models, and discussed the impact of different parameters on the accuracy of the model. Comparing the accuracy of the two models with the augmented dataset before and after improvement, it showed that the improvement of the two models is effective. Comparing with other methods, the classification accuracy of the proposed system was the highest, and our algorithm has better performance, which provides the possibility for migration to mobile medical equipment and can improve the diagnostic performance of mobile medical equipment. But this paper still needs several improvements: 1) Too many data samples may cause the model to overfit. 2) When using transfer learning, the impact of different size datasets on model performance is not discussed. 3) Did not solve the problem of data imbalance.

Acknowledgments. This work is supported by National Natural Science Foundation of China (Grant nos. 81772009), Science and Technology Project of Henan Province Science and Technology Department (Grant nos. 182102310162), Key scientific research projects of Henan universities

in 2017 (Grant nos. 17A520014), 2017 Henan Province Science and Technology Project (Grant nos. 172102310496), 2019 Zhengzhou University offline excellent course construction project (Grant nos. 2019XXJPKC023), Research on Video Tracking Moving Object Modeling (Grant nos. 2019ZDGGJS029).

References

1. Laukamp, K.R., Thiele, F., Shakirin, G., et al.: Fully automated detection and segmentation of meningiomas using deep learning on routine multiparametric MRI. *Eur. Radiol.* **29**(1), 124–132 (2019)
2. Kamnitsas, K., Ledig, C., Newcombe, V.F.J., et al.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **36**, 61–78 (2017)
3. Kharrat, A., Gasmi, K., Messaoud, M.B., et al.: A hybrid approach for automatic classification of brain MRI using genetic algorithm and support vector machine. *Leonardo J. Sci.* **17**(1), 71–82 (2010)
4. Zhu, H., Fang, Q., He, H., et al.: Automatic prediction of meningioma grade image based on data amplification and improved convolutional neural network. *Comput. Math. Methods Med.* **2019** (2019)
5. Afshar, P., Mohammadi, A., Plataniotis, K.N.: Brain tumor type classification via capsule networks. In: 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 3129–3133. IEEE (2018)
6. Zia, R., Akhtar, P., Aziz, A.: A new rectangular window based image cropping method for generalization of brain neoplasm classification systems. *Int. J. Imaging Syst. Technol.* **28**(3), 153–162 (2018)
7. Swati, Z.N.K., Zhao, Q., Kabir, M., et al.: Brain tumor classification for MR images using transfer learning and fine-tuning. *Comput. Med. Imaging Graph.* **75**, 34–46 (2019)
8. Deepak, S., Ameer, P.M.: Brain tumor classification using deep CNN features via transfer learning. *Comput. Biol. Med.* **111**, 103345 (2019)
9. Goodfellow, I.: NIPS 2016 tutorial: generative adversarial networks. arXiv preprint [arXiv:1701.00160](https://arxiv.org/abs/1701.00160) (2016)
10. Chang, K., Bai, H.X., Zhou, H., et al.: Residual convolutional neural network for the determination of IDH status in low-and high-grade gliomas from MR imaging. *Clin. Cancer Res.* **24**(5), 1073–1081 (2018)
11. Han, C.: Infinite brain MR images: PGGAN-based data augmentation for tumor detection. In: Esposito, A., Faundez-Zanuy, M., Morabito, F.C., Pasero, E. (eds.) *Neural Approaches to Dynamics of Signal Exchanges*. SIST, vol. 151, pp. 291–303. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-8950-4_27
12. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint [arXiv:1511.06434](https://arxiv.org/abs/1511.06434) (2015)
13. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014)
14. Cheng, J., Huang, W., Cao, S., et al.: Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PloS ONE* **10**(10) (2015)
15. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979)
16. Xue, D.X., Zhang, R., Feng, H., et al.: CNN-SVM for microvascular morphological type recognition with data augmentation. *J. Med. Biol. Eng.* **36**(6), 755–764 (2016)
17. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014)

18. Szegedy, C., Vanhoucke, V., Ioffe, S., et al.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826 (2016)
19. Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. *Science* **313**(5786), 504–507 (2006)
20. Afshar, P., Plataniotis, K.N., Mohammadi, A.: Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries. In: ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1368–1372. IEEE (2019)