Image Classification of Brain tumor based on Channel Attention Mechanism

Mengxin Liu*, Jun Yang*

School of Information Engineering, Ningxia University, Yinchuan, Ningxia, 750021, China

*author’s e-mail:229810803@qq.com

Abstract. The main method of traditional medicine to diagnose brain tumors is that expert doctors read the pictures with naked eyes. Meanwhile, the annotation of this type of medical image data set requires to be done manually by experienced experts, which makes the production cost of labeled medical image data set higher. In view of the above situation, this study firstly used unsupervised learning method Deep Convolution Generative Adversarial Networks (DCGAN) to generate data set expansion on the four types of MRI brain tumor images including glioma, meningioma, pituitary tumor and no tumor. In addition, an improved DenseNet-201 brain tumor classification network model is proposed, which uses the channel attention mechanism SENet as a bypass connecting the Convolutional layer and the Transition layer (or the Transition layer and the Transition layer), and interweaves the classification model on this basis. Optimized the problem of weak resolution of the original model to extract features, thereby accurately capturing the position, shape and texture information of the brain tissue in the MRI image. The model improves the accuracy of the brain MRI image classification, and enhances the generalization ability of the model.

1. Introduction

In recent years, medical imaging technology has developed rapidly and has become an indispensable tool for the study of various diseases. Common medical images include X-ray CT (X-CT), Magnetic Resonance Imaging (MRI), radionuclides, etc. These methods help doctors evaluate and diagnose symptoms of cancer, tumors, and cardiovascular diseases in the early stage[1]. In order to help doctors find abnormal symptoms in patients' brains early, detecting the type of brain tumors in MRI images has become the core of this work. Since the University of Toronto professor Geoffrey Hinton[2] proposed deep learning in the top international academic journal Science in 2006, deep learning technology has received widespread attention and has been applied in various fields of medicine, such as tumor screening and lung nodules, benign and malignant classification of lung nodules, and diagnosis of cardiovascular and cerebrovascular diseases. Malignant classification and diagnosis of cardiovascular and cerebrovascular diseases. At present, the application of deep learning in brain MRI medical images is also becoming popular. Adrien et al. [3] used 3D-CNN to classify brain MRI images to diagnose Alzheimer’s disease. Zikic et al. [4] used a shallow CNN model to classify brain tumors. The model has two convolutional layers and the maximum pooling step of the convolutional layer is set to 3, as well as a fully connected layer and a Softmax layer. Kumar et al. [5] used the integration method to classify medical images, and obtained various image features with strong resolution by integrating features of
different levels. Selvara et al. used the least square support vector machine method to classify normal and abnormal brain tissues and achieved very good results [6].

2. Materials and Methods

2.1. Deep convolution generative adversarial networks (DCGAN)
Generative Adversarial Network (GAN) is a deep learning network model proposed by Goodfellow et al. [7] in 2014. Its core is to achieve Nash equilibrium through mutual competition between two models. The principle of DCGAN is the same as that of GAN, except that the generator and discriminator in the GAN network are replaced with a convolutional neural network [8]. The generating network G processes the input noise through a series of deconvolution and upsampling to generate false samples, the discriminant network D judges whether the input image is a true sample through a series of subsamples. If it is a true sample, the output probability is 1; otherwise, the output probability is 0 [9]. The DCGAN generator (left) and the discriminator (right) are shown in Figure 1.

![Figure 1. DCGAN generator (left) and discriminator (right)](image)

2.2. Training and effect
Data came from Kaggle. The data set contains tumor images of 223 patients, a total of 4566 images, including 3,119 images in the training set, 1,025 in the verification set and 422 in the test set. Randomly selected 120 from each category for training GAN network. The training parameters are set as follows: the Batchsize is set to 75, the learning rate is 0.0001, and the optimal exponential decay rate $\beta$ is set to 0.5. Each step the discriminant network D is trained once for each step, and the generated network G is trained twice. As can be seen in Figure 2, as the number of epochs increases, the quality of the generated images is getting higher and higher. The samples generated by the model in the first few epochs (number of training iterations) are almost noisy images, the samples generated after the 27th epoch begin to have the approximate outline of the brain tumor; after about 650 epochs, the generated images are enough to pass off as real.

![Figure 2. DCGAN training process](image)

![Figure 3. Dense block structure](image)
2.3. Brain tumor image classification model based on channel attention mechanism

Convolutional neural network is one of the typical representatives of deep learning. The deep features of medical images are learned through deep convolutional neural networks, which has better robustness in expressing images [10]. The following is a detailed description of the network model SE-DenseNet-201 based on the channel attention mechanism proposed in this paper.

2.3.1. SE-DenseNet-201

In recent years, deep convolutional neural networks have performed in full swing in image feature extraction. However, due to the continuous deepening of the convolutional neural network model, the phenomenon of overfitting and gradient disappearance (explosion) are very likely to occur, which leads to a decrease in the accuracy of the model [11]. In order to fully capture the brain MRI image information and improve the accuracy of model classification, this paper proposes the SE-DenseNet-201 model to solve this problem. The model SE-DenseNet-201 is shown in Figure 4.

![SE-DenseNet-201 model](image)

2.3.2. Dense block and transition layer

Dense Block is an important part of DenseNet, which is used to improve the characteristic information flow between layers. It is composed of BN [12], Relu [13], 1×1Conv and 3×3Conv. Since the input of the following layer will be very large, Dense Block internal uses the bottleneck [14] layer to reduce the amount of calculation. The details are shown in formula 1.

\[ x_l = H_l([x_0, x_1, \cdots, x_{l-1}]) \]  

(1)

Among them, \([x_0, x_1, \cdots, x_{l-1}]\) refers to the feature map generated in the 0, 1, \cdots, 1 - 1 layer, and \(H_l(\cdot)\) is defined as the composite function of inputting three consecutive operations in the \(l^{th}\) layer. The dense block structure is shown in Figure 3.

The transition layer is located between two dense blocks and is used to change the size of the feature map. It is composed of BN, Relu, 1×1Conv and 2×2 average pooling. The convolution is responsible for extracting features from the output of the previous layer, and sharing a group weight when extracting features [15]. The convolution process is shown in formula 2.

\[ z^l = W^l \cdot f_l(z^{(l-1)}) + b^l \]  

(2)

Where \(z^l\) is the neuron state of the \(l^{th}\) layer, \(f_l(\cdot)\) represents the activation function, \(W^l\) and \(b^l\) represent the weight matrix and the deviation from the \((l - 1)^{th}\) layer to the \(l^{th}\) layer respectively.

2.3.3. SENet structure

SENet is composed of five parts: global average pooling, fully connected layer, Relu, fully connected layer and sigmoid. This "compression and excitation" network (as shown in Figure 5) is a sub-structure...
that can be used with any model. By modeling the interdependence between feature channels, beneficial feature channels can be enhanced and useless feature channels can be suppressed, so that feature channel adaptive calibration can be realized [16].

Figure 5. SENet structure

SENet implements the sequeeze operation through global average pooling to generate channel statistics \( Z \in \mathbb{R}^{C} \), and converts the input of \( H \times W \times C \) into the output of \( 1 \times 1 \times C \), as shown in formula 3. Excitation can help capture the dependency of the channel. It greatly reduces the parameters and calculation amount, and is mainly composed of two fully connected layers and two activation functions, as shown in formula 4.

\[
z_c = F_{sq}(U_c) = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} u_c(i, j)
\]

\[
S = F_{ex}(Z, W) = \sigma(g(Z, W)) = \sigma(W_2 \delta(W_1 z))
\]

Among them, \( F_{sq} \) represents the compression operation, and \( U_c \) represents the \( c^{th} \) feature map of the spatial dimension \( H \times W \times C \). \( F \) represents the compression operation, \( W_1 \in \mathbb{R}^{c \times C} \), \( W_2 \in \mathbb{R}^{C \times c} \), \( r \) is a scaling parameter. \( W_1 z \) represents the first fully connected layer, and then passes through a Relu layer, \( \delta(\cdot) \) represents Relu. Then it is multiplied by \( W_2 \) to form the second fully connected layer, and then passes through the sigmoid function, \( \sigma(\cdot) \) represents sigmoid. Finally, the input channels are multiplied by their respective weights to obtain \( \tilde{x}_k \), as shown in formula 5.

\[
\tilde{x}_k = F_{scale}(U_c, S_c) = S_c \cdot U_c
\]

Where \( \tilde{x}_k \in \mathbb{R}^{H \times W} \), \( F_{scale}(U_c, S_c) \) refers to the channel-level multiplication between the channel weight and the feature map \( U_c \).

2.3.4. Classification model

The classification network consists of global average pooling, full connection layer, BN and Softmax function. It is mainly used to reduce parameters and to differentiate brain tumors, as shown in Figure 6. Global average pooling summarizes the spatial information and reduces each the dimension of feature map[17]. Softmax can increase or decrease the signal exponentially to highlight the information to be enhanced, as shown in formula 6.

\[
\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{c=1}^{C} e^{z_c}}
\]
3. Experimental results and analysis

3.1. Sample comprehensive data set and performance evaluation index

1200 images of each type of tumor image (original sample + generated sample) are selected, a total of 3600 images are used as the training set. The original tumor image is used for the test set. Part of the brain MRI image is shown in Figure 7.

In order to verify the performance of the model, the effect of the model is evaluated after the model is constructed. The evaluation method of the confusion matrix of the two classification results was adopted in this paper:

Accuracy: proportion that was predicted correctly.

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{7}
\]

Precision: correctly predicted to be positive accounted for the proportion of all actually predicted to be positive.

\[
\text{Precision} = \frac{TP}{TP + FP} \tag{8}
\]

Recall: correctly predicted to be positive accounted for the proportion of all that should be predicted to be positive.

\[
\text{Recall} = \frac{TP}{TP + FN} \tag{9}
\]

Among them, TP is correctly predicted as a positive example; TN refers to a correct prediction as a negative example; FP refers to the false forecast as a positive example; FN refers to a wrong prediction as a negative example. Accuracy and recall are a pair of contradictory measures. Generally speaking, when the accuracy rate is high, the recall rate tends to be low; and when the recall rate is high, the precision rate tends to be low. Usually only in some simple tasks, it is possible to make both of them very high [18].

3.2. Experimental results

In the experiment, the network input size is set to 224×224×3, and the hyperparameters required for training the model are set as follows: The model is trained using a small batch of samples, the size is set to 42. The Adam [19] algorithm is used to optimize the loss function, the learning rate is set to 0.0001, the penalty factor is set to 0.0001, and L2 regularization is used to prevent overfitting. The dropout of the training process is set to 0.3, and the Relu activation function is used. And use random rotation, flip, zoom, and adjust brightness to enhance the training set.

In order to select a model that is more suitable for brain tumor image classification, DenseNet-121, DenseNet-161, DenseNet-169, and DenseNet-201 are used for training respectively, and the
classification results obtained are shown in Table 1. Separately merge the SENet and the classification network with the above model to extract image features, and obtain the SE-DenseNet-121, SE-DenseNet-161, SE-DenseNet-169 and SE-DenseNet-201 network models. The classification results obtained are shown in Table 2.

| Model       | Accuracy (%) |
|-------------|--------------|
| DenseNet-121| 87.43 ± 1.5  |
| DenseNet-161| 89.25 ± 0.8  |
| DenseNet-169| 90.68 ± 0.3  |
| DenseNet-201| 91.85 ± 0.9  |

Table 2 shows that in the face of complex and high resolution, deep convolutional neural networks have more advantages in classification effects. Therefore, SE-DenseNet-201 is a better match for MRI brain tumor image classification. When training the model, the loss change curve and accuracy change curve of the validation set are shown in Figure 8 and Figure 9.

For the improved model SE-DenseNet-201, the comparison curve of the training effect and the verification effect on the data set is shown in Figure 10 and Figure 11.
4. Conclusions
In order to enable doctors to detect abnormalities in the patient’s brain early and quickly diagnose tumor types, this article proposes a SE-DenseNet-201 model to classify four types of MRI brain tumors, including meningioma, glioma, pituitary tumor and no tumor. The channel attention mechanism is added to the model, which can not only pay attention to the position of the brain tissue in the image, but also effectively extract tumor-related feature information and improve the accuracy of model classification. In addition, based on the principle of generating pictures based on DCGAN, a series of MRI brain tumor images are generated from a small amount of original images to expand the data set, and use the random rotate, flip, scaling and adjust the brightness to enhancement of the training set. The image is classified based on the data set combined with the generated samples, and join the Dropout, which improves the generalization ability of the model and alleviates the overfitting phenomenon. In the future work, we will continue to study pathological information such as the location, type, and size of brain tumors in the brain, committed to different imaging images, such as X-ray, CT and ultrasound, continue to add other types of data sets for classification, and improve the robustness of the model.

Acknowledgments
1. Ningxia Natural Science Foundation, "Research on key technologies of large-scale wireless sensor network based on edge computing and its application in characteristic agriculture" (2020AAC03036).
2. Graduate Innovation Project of Ningxia University (GIP2021056).

References
[1] S. Bauer et al. A survey of MRI-based medical image analysis for brain tumor studies[J]. Phys. Med. Biol. 2013, 58(13): 97-129.
[2] Hinton, G. E. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(5786): 504-507.
[3] Payan A, Montana G. Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks[J]. arXiv preprint arXiv:1502.02506, 2015.
[4] D. Zikic et al. Segmentation of brain tumor tissues with convolutional neural networks[J]. MICCAI Multimodal Brain Tumor Segmentation Challenge(BraTS), 2014: 36-39.
[5] Kumar A, Kim J, Lyndon D, et al. An ensemble of fine-tuned convolutional neural networks for medical image classification[J]. IEEE Journal of Biomedical and Health Informatics, 2017, 21(1): 31-40.
[6] Selvaraj H, Thamarai SS, Selvathi D, et al. MRI slices classification using least squares supports vector machine [J]. International Journal of Intelligent Computing in Medical Sciences and Image Processing (IC-MED), 2007, 1(2): 21-33.
[7] Goodfellow Ian, Pouget Abadie J, Mirza M, et al. Generative adversarial nets[C]. Advances in Neural Information Processing Systems. 2014: 2672-2680.
[8] Radford A, Metz L, Chintala S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks [J]. Computer Science, 2015.

[9] T. Salimans, I. Goodfellow, W. Zaremba, et al. Improved techniques for training gans [c]. NIPS, 2016, 2234-2242.

[10] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [J]. Advances in neural information processing systems, 2012, 25: 1097-1105.

[11] Rumelhart D, Hinton G, Williams R J. Learning Representations by Back Propagating Errors [J]. Nature, 1986, 323(6088): 533-536.

[12] Abadi M, Agarwal A, Barham P, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems [J]. arXiv preprint arXiv:1603.04467, 2016.

[13] Maas A, Hannun A, Ng A. Rectifier nonlinearities improve neural network acoustic models [C]. In Proc. ICML, volume 30, 2013: 1152-1160

[14] Huang G, Liu Z, Maaten L, et al. Densely Connected Convolutional Networks [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017, 4700-4708.

[15] Lin M, Chen Q, Yan S. Network in network [J]. arXiv preprint arXiv:1312.4400, 2013.

[16] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE conference on computer vision and pattern recognition, 2018: 7132-7141.

[17] Lin M, Chen Q, Yan S. Network in network [J]. arXiv preprint arXiv:1312.4400, 2013.

[18] Gao W X, Hui R, Tian Z. Classification of CT brain images based on deep learning networks [J]. Computer Methods and Programs in Biomedicine, 2017, 138: 49-56.

[19] Kingma D, Ba J. Adam: A method for stochastic optimization [J]. arXiv Preprint, 2015, arXiv:1412.6980.