Brain Tumor Identification Based on AlexNet and VGG

Rongchen Zhao*

Department of Artificial Intelligence, South China University of Technology, Guangzhou, China

*Corresponding author: 202264640361@mail.scut.edu.cn

Abstract. Early diagnosis of brain tumor symptoms through Computer Tomography (CT) imaging is an important means of treating brain tumors. However, doctors who identify and diagnose brain tumors with naked eyes may cause misdiagnosis due to overwork or inexperience, which greatly affects the rehabilitation of brain tumor patients. In recent years, there is a lack of a clinical application tool for automatic identification and classification of brain tumor CT images to help doctors improve the rate of brain tumor treatment. This study proposes an approach for identifying and classifying brain tumors using a modified AlexNet and Visual Geometry Group (VGG). The methodology involves the comparative analysis of the performance of the modified AlexNet and VGG in identifying and classifying brain tumors. Our findings reveal that the modified AlexNet demonstrates superior performance, with an accuracy rate of 69.54%, in comparison to the VGG. The proposed approach holds tremendous potential for future clinical application as an auxiliary tool to aid medical professionals in accurately diagnosing brain tumors.

Keywords: Convolutional Neural Network; Brain Tumor detection; AlexNet; VGG.

1. Introduction

Malignant Gliomas (MGs) have been a problem that has plagued humans for nearly a century. As the normal cells are affected by pathogenic factors, they maybe undergo gene mutations which lead them to transform into cancer cells, thereby inducing tumors. According to the statistical report published by the Central Brain Tumor Registry of the United States (CBTRUS), brain tumors constitute a diverse category of neoplastic growths that may be either primary or metastatic in nature and occur within the central nervous system. These pathological conditions are considered to be highly dangerous due to their low survival rates [1]. For this problem, relevant medical studies have proved that early resection of brain tumors and adjuvant postoperative radiotherapy are one of the important means to improve the survival rate of patients [2, 3]. Therefore, finding the presence of brain tumors as early as possible based on the Computer Tomography (CT) images of the patient's brain has become one of the top priorities. However, nowadays, most of brain tumor identification methods still rely on doctors to visually identify the images with experience. This not only has low efficiency but also has limited accuracy. To address this issue, this paper proposes an automatic detection method based on neural network algorithms in artificial intelligence, which may assist doctors to improve the accuracy of brain tumors identification.

In August 1955, at the Dartmouth College in the United Kingdom, the Dartmouth Conference was held for two months, and the concept of artificial intelligence was first proposed at the conference. Since then, artificial intelligence has come out for the first time. In the following nearly 70 years of development, artificial intelligence has developed rapidly, and neural networks, as an important branch of artificial intelligence, have gradually matured, including the Convolutional Neural Networks (CNNs) model. In 2021, Dwivedi et al. have highly praised the accuracy of the CNN model in the field of recognition and prediction in their articles [4, 5]. After analysis, two classic CNN network architectures, Visual Geometry Group (VGG) and AlexNet, are mainly used for the recognition problem in this paper. In 2014, VGG was proposed [6], which can reduce the convolution kernel and increases the depth of network to extract more delicate features to attach the final goal of reducing errors. In the following year, Residual Networks was first proposed by Kaiming He et al., and won the first place in the classification and target detection tasks in the ImageNet competition that year [7]. The creator broke the simple basic mapping relationship between layers but added the relationship of residual mapping. This makes CNN model with a great number of layers possible and further improves the accuracy of the CNN model.

Based on the problem of brain tumor recognition, this paper uses a dataset of Computer Tomography (CT) images of brain cancer patients from the website Kaggle [8]. The experiment shows that the network built based on AlexNet has a good performance with an accuracy of 69.54%.

2. Method

2.1. Dataset preparation

2.1.1. Dataset introduction

The data set used in this paper collected from 3, 274 Computer Tomography (CT) images of the brains of brain cancer patients released on the kaggle website [8]. In the four categories (i.e. glioma, meningioma, pituitary, and no tumor), there are 928, 939, 903, and 502 images size of 512×512 respectively. A sample of each category is shown in Fig. 1 (a), (b), (c), and (d) respectively. In order to be suitable for the training of the convolutional neural network, this paper represents each image in the form of a 512×512 gray matrix, and 2, 875 samples are used for training and 398 samples for testing.

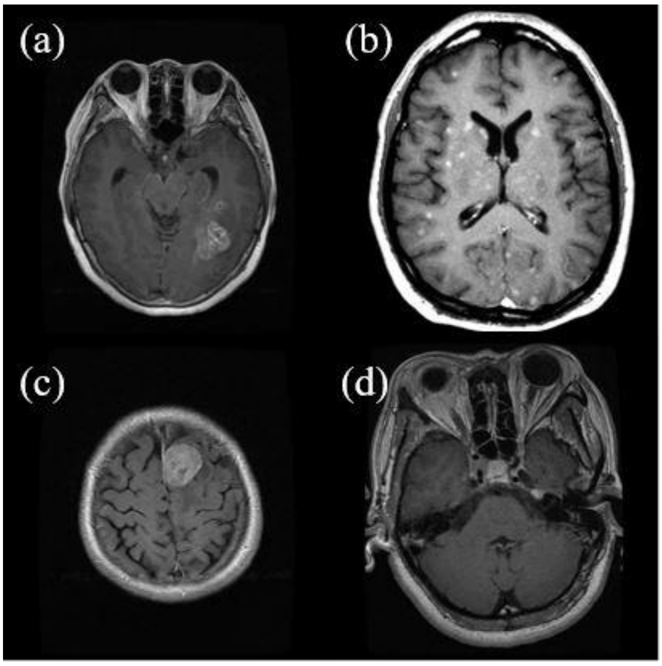

Fig. 1 The example of images on the collected dataset

2.1.2. Preprocessing process

For a large amount of data, some preprocessing of the data before training is essential. In this paper, the normalization and augmentation of the data are mainly carried out in the preprocessing stage. In order to improve the training speed, this paper adopts normalization processing on the gray value. That is to say, after normalization, each gray value (with an interval from 0 to 255 initially) is in the interval from 0 to 1. At same time the data augmentation is adopted. There have been relevant studies showing that data augmentation operations such as geometric transformation, and frame position transformation can greatly improve the accuracy and the generalization ability of the model [9, 10]. Therefore, this paper tries to use the ImageDataGenerator function in the keras library to achieve the data augmentation. Five basic data augmentation operations (i.e. rotation, horizontal translation, vertical translation, horizontal flip and vertical flip) are performed on each sample image of every batch. After building the Convolution Neural Network model, these batches will be used for training one by one in each epoch.

2.2. The proposed CNN model

A basic neural network mainly includes convolutional, pooling, and fully connected layers. The convolutional layer is mainly responsible for a preliminary extraction of image features, and the pooling layer is to filter the features extracted by the convolutional layer. At last, in the end of the network, the fully connected layer converts the feature matrix trained by each layer into a feature vector for output. In this paper, two classic convolutional neural network architectures, AlexNet and VGG, are adopted for tumor recognition. AlexNet is a new network architecture proposed by Alex.etc in the ImageNet competition in 2012 [11]. It mainly contains 8 learning layers, including 5 convolutional layers and 3 fully connected layers. The difference is that the Rectified Linear Units (ReLUs) nonlinear activation unit is used in this architecture, and the paper by Alex et al. shows that this activation function is faster than the training speed using the tanh unit as the activation function. Several times [11]. At the same time, in order to prevent over-fitting problems, the 'Dropout' method is used in this network architecture, which randomly sets the output of neurons in the hidden layer to 0 with a probability of 50%, which means that these parameters are not will participate in the feedforward and backpropagation of the network, greatly reducing the experimental error [12]. Then in 2015, the Very Deep Convolution Neural Network (VGG) proposed by Karen et al. further optimized the network architecture on the basis of AlexNet [6]. Compared with the 5×5 convolution kernel used in AlexNet, the VGG architecture only uses a convolution kernel with a step size of 1 and a size of 3×3. That is to say, VGG can be more delicate when extracting image features compared to AlexNet.

In this paper, two neural networks are built by referring to the two classic CNN architectures of ALexNet and VGG, and the respective fitting effects are compared and evaluated. The first Convolutional Neural Network is mainly based on the architecture of VGG16. It has 3×3 convolution kernels. At the same time, a Dropout layer is set to regularize the training data in order to prevent the problem of overfitting. Different from the classic VGG16, in order to speed up the training speed and reduce the cost of computing power, only 12 layers (9 convolutional layers, 4 pooling layers, 2 fully connected layers, and 1 output layer) are set. The second Convolutional Neural Network is built based on the basic algorithm of AlexNet with optimized adjustments. It consists of 8 layers (5 convolutional layers, 3 pooling layers, 1 fully connected layer, 1 output layer). The Dropout method is applied on after each pooling layer and fully connected layer to prevent overfitting. At the same time, Batch Normalization (BN) is introduced to further prevent the occurrence of overfitting.

2.3. Implementation details

In this paper, tensorflow is mainly used to build the neural network structure. The learning rate, loss function, optimizer, epochs and batch size were set to 0.01, categorical cross entropy, adam, 200, 32, respectively.

3. Results and discussion

3.1. Results

Table 1. The performance comparison of three CNN models

Model	Training Loss	Training Accuracy	Testing Loss	Testing Accuracy
Optimized AlexNet (epochs=100, batch size=32）	0.3999	0.8798	3.7455	0.5584
Optimized AlexNet (epochs=200, batch size=32）	0.2155	0.9310	2.8017	0.6954
Optimized VGG-16 (epochs=200, batch size=32）	0.7545	0.6561	4.1868	0.2513

By adjusting different epochs, three CNN networks were built based on the AlexNet and VGG-16 models respectively. The experimental results show that batch size=32, epochs=200, and the optimized AlexNet has the best performance as shown in Table 1, with an accuracy of 69.54%. The expected VGG model with a large number of parameters does not have a good performance on this issue, and the accuracy rate is only 25.13%.

3.2. Discussion

In comparison to the deeper VGG model, the AlexNet model has demonstrated better performance, which may be attributed to various factors. Overfitting is one possible reason for this discrepancy, as evidenced by the observation that the VGG model's training loss decreases but its test loss increases when trained for 100 epochs compared to 200 epochs. This trend suggests that the VGG model's complex architecture may be unsuitable for the problems presented in this study, thus providing an explanation for why the simpler architecture of AlexNet leads to better performance. Another contributing factor to the performance gap between the two models is the size of their respective convolution kernels. While the VGG model uses a smaller convolution kernel and deeper network layers to improve the depth of the network while maintaining the same receptive field as AlexNet, this design choice may not be optimal for the brain tumor recognition problem examined in this paper. Specifically, reducing the depth of the VGG network to prevent overfitting may also impact the size of its receptive field, while the smaller convolution kernel may not be able to effectively extract the necessary brain tumor feature matrix. This could lead to a noisy and ineffective training process, as the recognition of brain tumor CT images may not require excessively intricate feature matrices. In the future, some more advanced neural networks e.g. ResNet may be considered due to their excellent performance in the similar task [13, 14].

4. Conclusion

This article proposes the use of CNN to identify and classify brain tumors, using CT slice images of brain tumors published on the kaggle website to avoid privacy issues. At the same time, two classic convolutional neural network architectures, AlexNet and VGG, are used for reference, and corresponding optimizations are made for the problems in this paper to improve the performance. Experiments show that compared with the complex optimized VGG architecture, the optimized AlexNet with fewer parameters is more suitable for solving the problem of brain tumor CT image recognition and classification. At the same time, it is also showed that the problem solved with the method in this paper has a better performance. Future work aims to propose a more appropriate recognition network architecture for tumor cells and to extend the application of CNN to other medical image recognition problems, with the ultimate goal of improving patient treatment rates.

References

[1] Dolecek T A et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005–2009. Neuro. Oncol. 2012, 14: v1–v49.

[2] Stupp R et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N. Engl. J. Med. 2005, 352:987–996.

[3] Chaichana K L et al. Multiple resections for patients with glioblastoma: prolonging survival. J Neurosurg. 2013, 118:812–820.

[4] Dwivedi Y K et al. Artificial intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. Int J Inf Manag 57:101994, 2021.

[5] Salehi M et al. one-dimensional probabilistic convolutional neural network for prediction of breast cancer survivability. Comput J, 2021.

[6] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[7] Kaiming H et al. Deep Residual Learning for Image Recognition, 2015.

[8] Brain Tumor Classification. https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri, 2020.

[9] Barret Z et al. Learning Data Augmentation Strategies for Object Detection, 2019.

[10] Maximilian I Jakub M, Tomczak P F, Selecting Data Augmentation for Simulating Interventions, 2020.

[11] Alex K, Ilya S, Geoffrey E. Hinton ImageNet Classification with Deep Convolutional Neural Networks, 2012.

[12] Hinton G E et al. Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580, 2012.

[13] Yu Q, Chang C S, Yan J L, et al. Semantic segmentation of intracranial hemorrhages in head CT scans, 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS). IEEE, 2019: 112-115.

[14] Khan H A, Jue W, Mushtaq M, et al. Brain tumor classification in MRI image using convolutional neural network. Math. Biosci. Eng, 2020, 17(5): 6203-6216.