

Received November 19, 2019, accepted November 30, 2019, date of publication December 9, 2019,
date of current version December 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2958370

Brain Tumor Detection Based on Multimodal Information Fusion and Convolutional Neural Network

MING LI^{ID1}, LISHAN KUANG^{ID2}, SHUHUA XU^{ID2}, AND ZHANGUO SHA^{ID2}

¹Department of Nuclear Medicine, Rizhao People's Hospital, Rizhao 276826, China

²Department of Radiology, Rizhao People's Hospital, Rizhao 276826, China

Corresponding author: Zhanguo Sha (shazhanguo@126.com)

ABSTRACT Aiming at the problem of low accuracy of traditional brain tumor detection, in this paper, a combination of multimodal information fusion and convolution neural network detection method of brain tumors, we call it a Multi-CNNs. First, this paper uses the extension of the 2D-CNNs to multimodal 3D-CNNs, and can obtain brain lesions under different modal characteristics of three-dimensional space. It can solve the 2D-CNNs raw input requires large neighborhood of faults, at the same time better to extract the modal of the differences between information. Then the real normalization layer is added between the convolution layers and pooling layer to improve the convergence speeds of the network and alleviate the problem of overfitting. In the end, the loss function was improved, and the weighted loss function was used to enhance the feature learning of the lesion area. The experimental results showed that the brain tumor detection method proposed in this paper could effectively locate tumor lesions, and better results were obtained in correlation coefficient, sensitivity, and specificity. Compared with two-dimensional detection network and single mode brain tumor detection methods, the detection accuracy is significantly improved.

INDEX TERMS Multimodal, fusion, convolutional neural network, brain tumor.

I. INTRODUCTION

With the improvement of modern medical standards, medical imaging technology plays an increasingly important role in daily medical diagnosis and medical research. Therefore, research on medical diagnostic image data is very important. As a tumor disease with frequent occurrence and complexity, brain tumor has become a key research topic in the medical field [1], [2]. The diagnosis of brain tumors is usually based on imaging data analysis of brain tumor images. Accurate analysis of brain tumor images is a key step in determining a patient's condition. However, the accumulation of doctors' personal medical knowledge, differences in experience levels, and visual fatigue can affect the correct analysis of image results. Therefore, how to accurately detect brain tumor images is very important [3], [4].

Magnetic Resonance Imaging (MRI) [5] can provide information on the shape, size, and position of human tissues and organs without high ionizing radiation. The images obtained are very clear and precise. MRI greatly improves the diagnostic efficiency, avoids the operation of thoracotomy or

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei^{ID}.

laparotomy exploration, and provides a good guide for lesion localization and surgical treatment. Brain tumor MRI uses three-dimensional multi-band imaging technology, and chest X-ray scanning, etc. Compared with 2D images, 3D multi-band MRI can provide the coordinate position of the lesion area to assist the doctor to accurately locate the lesion area. In addition, MRI imaging can also obtain different structures of the same tissue using the unused development sequence. That is, a multimodal MRI image. Brain MRI imaging can be divided into four modes according to the difference of imaging auxiliary conditions: T1 weighted mode, T1ce mode, T2 weighted mode and Flair mode. Different modes can display different brain tumor features [6].

The existing methods of brain tumor detection mainly include level set detection [7], fuzzy clustering [8], region growing [9], and machine learning [10]. Feizollah *et al.* [11] used fuzzy clustering algorithm to image pre-detection is performed to obtain the region of interest of the brain tumor MRI image, and the clustering detection result is taken as the initial contour of the level set evolution, and finally the clustering result is used to detect the tumor. Konur [12] trained the Support Vector Machine (SVM) model with known samples, and then processed other brain tumor images

with the trained SVM model. Currently, convolutional neural networks develop rapidly, in target detection [13], speech processing [14], target detection [15], super-resolution reconstruction [16] and other areas have made breakthroughs, and extended to the field of computer-aided medical diagnosis, such as the diagnosis of lung nodules and Alzheimer's disease, some achievements have been made in the diagnosis of brain tumors. Gnouna *et al.* [17] proposed an automatic detection framework based on stacked noise reduction auto-encoder and multi-modal brain tumor images. The best similarity coefficient dice of the experimental results can be up to 93%, the average dice is 86%.

However, the existing methods of brain tumor detection and methods based on convolutional neural networks generally have the following problems:

- 1) The method is single and the accuracy is low, which cannot provide valuable information for clinicians;
- 2) Strong dependence on manual intervention and data preprocessing;
- 3) Most of them are processed based on a single modality, and MRI data of different modalities are not utilized efficiently.

Combined with MRI images, 3D information of multiple modalities can be obtained. This paper proposes an algorithm of brain tumor detection based on multimodal information fusion and convolutional neural network. Firstly, the 2D-CNNs are extended to multimodal 3D-CNNs to obtain the three-dimensional features of brain tumor lesions under different modalities. Secondly, the problem of slow convergence and over-fitting is solved by normalization. Finally, a new weighted loss function was constructed to reduce the interference of non-focal region to the detection of brain tumors.

Specifically, the technical contributions of this paper can be concluded as follows:

This paper proposes an algorithm of brain tumor detection based on multimodal information fusion and convolutional neural network. Firstly, the 3D-CNNs network is introduced to fuse multiple modal information, which can better extract the difference information among the modalities. Secondly, the improvement of the loss function can reduce the interference of non-lesional areas on brain tumor detection, and thus improve the accuracy of detection.

II. RELATED WORKS

This section mainly introduces the method of brain tumor detection based on multimodal MRI images and the structure of the basic method used in this paper, such as the structure of convolutional neural network and the structure of 3D convolutional network.

A. BRAIN TUMOR DETECTION METHOD BASED ON MULTIMODAL FUSION

Brain tumor detection is a technique for dividing different tumor tissues, such as active tumor tissue, edema tissue, and necrotic tumor tissue, from normal tissues such as

gray matter, white matter, and cerebrospinal fluid. Due to the high clinical relevance and challenge of tumor detection itself, the problems of brain tumor detection have received widespread attention in the past 20 years.

According to the degree of human intervention, image detection of brain tumors can be mainly divided into three categories: manual detection based on manual, semi-automatic detection based on manual initialization, and fully automatic detection without human intervention. Manual detection involves manually depicting the contour of the tumor and the structure of interest, or depicting the area of the anatomy with a different label. Manual testing is a tedious, tedious, time-consuming task. Different detectors have different detection tendencies. However, manual detection is a relatively professional and relatively accurate detection method for obtaining tumor information from images. In the semi-automatic and automatic detection methods, the results of manual detection are often used as the object of our comparison. We call it ground truth, which provides us with a qualitative and quantitative analysis basis. For semi-automatic detection, its detection results are relatively dependent on human initialization, and the user can interact with it, not just the output. Compared with the semi-automatic detection method, the automatic detection can automatically locate and detect the tumor area without human intervention after setting the relevant parameters.

Detection, localization, diagnosis, and classification of brain tumors are very important in clinical and oncological research. Early detection and localization of tumor diseases, and accurate determination of the type and stage of the tumor have a crucial impact on the rehabilitation and life extension of patients. Image studies of brain tumors have been associated with the development of related technologies for a long time, and various excellent algorithms have been introduced into the image research of brain tumors. In the development of brain tumor detection technology, most excellent algorithms were originally developed to detect other structures or lesions, such as the detection of white matter and gray matter, and obtained quite high detection on these problems. Accuracy was later introduced to the problem of brain tumor detection.

In brain tumor detection studies, there is no one method that can obtain satisfactory results for all brain tumor images. Typically, the method of detection is optimized for specific image data or specific data modalities. The methods of tumor image detection can be roughly divided into:

1) THRESHOLD-BASED APPROACH

Threshold is a very effective method of area detection, tumor target and brain tissue background, characterized by the gray value of pixels or voxels. Jiji and Dehmeshki [18] proposed an unsupervised method for the detection of pixel grayscale for T1c images with tumors.

2) REGION-BASED APPROACH

The corresponding MRI brain tissue images are divided into different target regions by pre-defining the similarity criterion

and combining the neighboring pixels or voxels in the intersecting regions. Vijayakumar and Gharpure [19] developed a tumor detection method based on MRI images using region growth. Nasrulloh *et al.* [20] proposed an improved region growing method to eliminate the volume effect and integrate the gradient information, obtain more accurate boundary information, and have a certain filling of the leakage gap that may occur after the detection. Zhao *et al.* [21] proposed a multi-scale watershed transform detection method for the detection of tumors by means of watershed transformation. Shi *et al.* [22] constructed a manual-assisted semi-automated method of brain tumor detection using the hierarchical watershed method.

3) METHOD OF PIXEL CLASSIFICATION

MRI data of brain tumors are generally stored in 2D slices or 3D volume. Slice-based brain tumor detection is no different from traditional image detection. A pixel is a basic unit of an image, and features such as gray level, local texture, color component, and topology information appear for each pixel in the feature space of the pixel using the attributes of the pixel. A brain tumor detection method based on pixel classification uses a supervised or unsupervised classifier to cluster related pixels in a brain tissue image. In the unsupervised classification, clustering is representative, and the correlation and non-correlation between tumor tissue and normal tissue in tumor images are mainly measured by distance [23]. The supervised classifier uses training data with tumor tags to train the parameters of the model to accommodate specific tumor detection problems [24].

4) MODEL-BASED APPROACH

The model-oriented tumor detection method has formed two major schools on the tumor detection problem: respectively, based on the detection model of the generation model and the discriminant model. The generated model utilizes the shape and spatiotemporal distribution of different tissues to generalize the relevant images of the invisible brain tissue [25]. Relative to the generation model, the discriminant model learns the difference in characteristics of brain tumors and other tissues directly from the training images [26]. Of course, the discriminant method generally requires a considerable amount of training samples. Through training, the discriminant model is greatly reduced in sensitivity to MRI image artifacts and grayscale and tumor shape transformation.

However, these methods are basically based on the single-modal information of MRI images to detect brain tumors. With the continuous development of technology, methods of brain tumor detection based on multi-modal MRI images have become a research hotspot of many researchers.

Multimodal MRI images, in a common sense, are different MRI images obtained from different MR sequences to the same tissue. Briefly, multimodal MRI images are images that are imaged differently under different contrasts from different MR imaging sequences. The introduction of

multimodal information has injected new vitality into brain tumor detection.

From a single-modal MRI image, we generally only get accurate information about one or two categories of tumors. For example, in the image of the glioma in the Flair mode, the free water in the brain tissue is converted to the bound water due to the normal brain tissue after the tumor lesion occurs. Free water was inhibited by Flair mode inversion restoration sequence. The entire tumor has a bright signal in Flair image, with high signal by binding water. At this time, only tumor-based and non-tumor structures can be obtained based on the separate Flair image detection. In order to have a more detailed tumor structure, multimodal information plays a crucial role.

Brain tumor detection combined with multimodal MRI information can be traced back to 2001. Fletcher-Heath *et al.* [27] proposed an automatic method of brain tumor detection using three modes of modality and clustering of T1, T2, and PD. After that, for the first time in the MICCAI, the competition BRATS for multimodal brain tumor segmentation [28] was established. This competition established the multimodal brain tumor detection to the world. Therefore, more and more outstanding researchers have also gathered to participate in the problem of multimodal brain tumor detection, which greatly promotes the development technology of brain tumor detection. Hu *et al.* [29] constructed a hybrid clustering and classification method of tumor detection by logistic regression. Wu *et al.* [7] studied an MRF model based on the detection clustering of super voxels. Pereira *et al.* [30] constructed a fully automated method of brain tumor detection by extracting neighborhood and local content features using a random forest classifier.

With the continuous breakthrough of convolutional neural networks in natural language processing, target detection and recognition, and image recognition and classification, methods of brain tumor detection have also begun to appear as a method based on convolutional neural networks. Since 2014, based on 2D and 3D convolution filters, many people use the ability of CNN of automatically learn image features. Pereira *et al.* [31] proposed a method by using multi-modal image block and 3D CNN structure of 3D modal brain tumor detection. García-Zapirain *et al.* [32] developed a method for predicting tumors by interpreting the size of the input four modal 3D patch to 2D-patch size using the traditional 2D-CNN structure. Havaei *et al.* [33] developed a two-way stacked CNN structure that uses both small size patches and large size patches. Lunga *et al.* [34] used a small 3x3 filter in the convolutional layer to construct a deeper CNN structure, while constructing different depths of network structures for different tumor types.

B. CONVOLUTIONAL NEURAL NETWORK STRUCTURE

In 1989, Lecun *et al.* [35] first proposed CNN, but it did not receive much attention until after the shocking results in the 2012 ImageNet competition [36]. The ImageNet dataset contains 1 000 objects, each containing 1 000 images,

and CNN almost cuts the error rate of the previous best method by half or even below the human recognition error level. Over time, the hardware and software for deep learning has been continuously improved, and CNN's architecture has become increasingly complex, even more than 100 layers, and the number of training weights required has reached millions.

A typical convolutional neural network consists of multiple convolutional layers, pooled layers, and fully connected layers. Usually, the convolutional layer and the pooled layer appear alternately in the front part of the network, while the latter part is the fully connected layer. The convolutional layer generates a feature map by convolving on the input image. The pooling layer is used to pass the value to the next layer by using the defined sliding window and taking its maximum or average value. The output of the convolutional layer before it is down sampled also needs to be delivered to a nonlinear function. This nonlinear function is called an activation function and can perform nonlinear transformation on data. To predict the input data, the output score of the last fully connected layer of the CNN is connected to the loss function, and the cross entropy loss function can normalize the score to the probability distribution of each label, i.e., the score large, the greater the probability that the input image belongs to the label. Finally, the Stochastic Gradient Descent (SGD) is used to optimize the parameters of the network by minimizing the loss function between the prediction and the truth tag. For example, the Stochastic Gradient Descent (SGD) is used to update the weights of the network in each iteration with back propagation until convergence. The training process of convolutional neural network is shown in Figure 1.

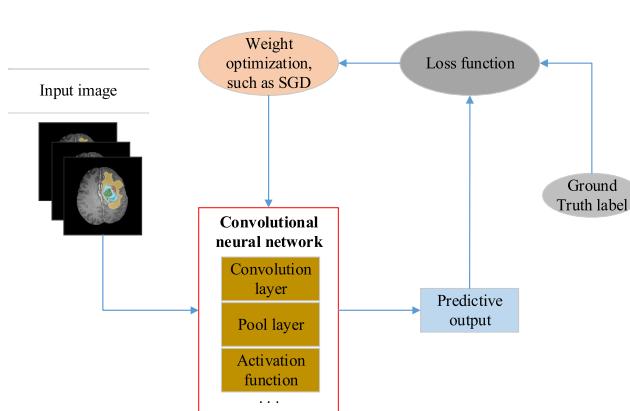


FIGURE 1. Convolutional neural network training process.

CNN is a feedforward artificial neural network developed from a multi-layer perceptron. Similar to the visual cortex of biology, the use of weight sharing reduces the complexity of the model inherited from the multi-layer perceptron and greatly reduces the number of weights. Very good results have been achieved in the fields of speech recognition, image classification and recognition. A typical CNN network structure consists of the following parts.

1) INPUT LAYER

The data input corresponding to the CNN model is usually a multi-channel image. Unlike traditional machine learning methods, for most problems, input can generally be directly input into the CNN network without pre-processing, and achieve better results.

2) CONVOLUTIONAL LAYER

Convolution is a very common linear operation in signal processing theory. The name of CNN is also derived from this special convolution operation for image signals. The purpose of convolution in CNN is to extract features from images. Convolution operations with convolution kernels of different levels and images can be used to obtain feature descriptions from general to abstract to more advanced, while retaining the spatial relationship between pixels during the extraction process.

By using the strategy of local perception, CNN can obtain the global information equivalent to the fully connected network through the fusion of feature information at a higher level. Each convolution kernel convolved with the image, you can get a set of images, also known as feature graph or feature map, different convolution kernel, convolved with the same image, you can get different feature graph, namely image features. Taking advantage of this, CNN adopts the method of multi-kernel convolution to extract richer image features and reduce the number of corresponding parameters. In practice, CNN obtains weights in convolution kernel through training. Figure 2 shows a schematic diagram of convolution operation.

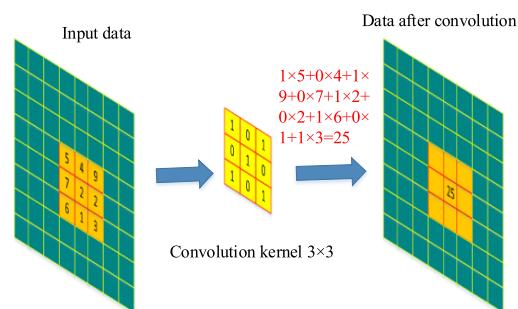


FIGURE 2. Schematic diagram of CNN convolution operation.

The convolutional layer has two important characteristics: local perception and parameter sharing, which can reduce the number of network model training parameters. Local perception means that each neuron node only needs to respond to certain specific regions of the global image, because the local pixel connections are relatively close, and the spatially distant pixels are weakly connected. Parameter sharing means that each feature map corresponds to the same convolution kernel, and the number of channels of the convolution kernel depends on the number of channels outputted by the previous layer. At the same time, the convolutional layer uses multiple convolution kernels in order to fully extract features.

3) ACTIVATION LAYER

The activation layer is an activation function that is concatenated after the convolutional layer. It is precisely because of the features captured by the activation function CNN that the nonlinear description is more prominent. Common activation functions are mainly Sigmoid, Relu, tanh, and so on. The basic schematic is shown in Figure 3.

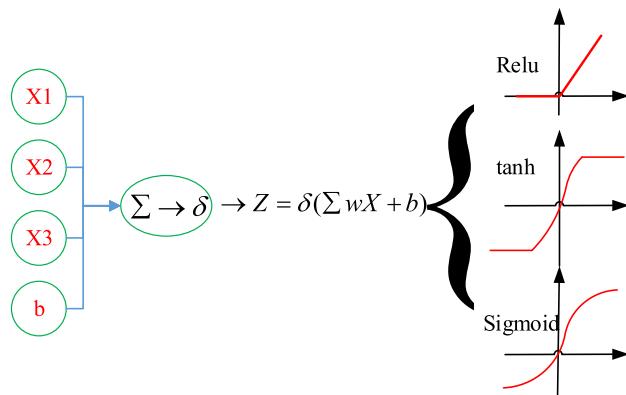


FIGURE 3. Schematic diagram of neuron activation and activation functions.

The Sigmoid function is used as the activation function. In the case of backpropagation, the division operation is involved. At the same time, for the deeper network structure, the gradient disappears easily, and the derivative gradually becomes zero in the process of Sigmoid back propagation. Moreover, it is unable to complete the training. As an activation function, Relu has a part of the neuron's output becoming zero due to the unilateral suppression of the function itself, thus reducing the interdependence of parameters, making the network have a certain sparsity, more in line with the human brain neurons working in response to the stimulus. The feature of sparsity, as an elemental operation, Relu can be applied to individual pixels. In addition, Relu itself is a piecewise linear function, and the derivation in the process of backward propagation still maintains piecewise linearity, which makes the network easier to optimize and learn. Therefore, most networks use Relu as an activation function.

$$\text{Sigmoid} = 1/(1 + \exp(-x)) \quad (1)$$

$$\text{Relu} = \max(0, x) \quad (2)$$

$$\tanh = (1 - \exp(-2x))/(1 + \exp(-2x)) \quad (3)$$

4) POOLING LAYER

Pooling is a convolutional neural network that is a means of reducing the amount of computation and is an aggregate statistical operation of images. The feature map obtained after the input image passes through the convolution layer, if directly expanded to do classification or other tasks, results in a large amount of calculation. The purpose of the pooling operation is to remove some redundant information and reduce the dimension of the feature map, which can reduce the feature, thereby reducing the amount of calculation, and effectively avoid over-fitting. In addition, the convolutional

layer obtains the local features of the upper layer output through local sensing, and the function of adding the pooling layer is to combine these similar local features into more advanced features.

Pooling is actually a down sampling operation. It is generally pooled on a 2×2 scale to reduce the size of the feature map, reducing feature redundancy and reducing the load on the next layer of computation. Commonly used pooling mainly includes maximum pooling, averaging pooling, and weighted pooling. Taking the 2×2 -space domain as an example, the maximum pooling is to select the largest element from the window to correct the elements in the feature map. If the average is the mean pooling, the elements in the window are summed and pooled. Figure 4 shows a schematic of the pooling of the maximum values.

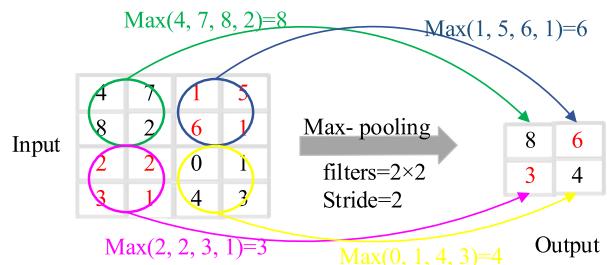


FIGURE 4. Schematic diagram of the maximum pool.

The pooling operation actually replaces the area information by calculating the average value or the maximum value in the local area of the feature map, thereby implementing the function of down sampling. As shown in Figure 4, the 4×4 feature graph is divided into $4 2 \times 2$ small regions, and then the maximum value or draw value of each region is calculated to obtain a 2×2 feature graph. Even if the target object in the image has a small translation or scaling, the pooling operation can still obtain the same pooling features as before the change. For example, when recognizing a handwritten number, moving it in any direction will cause the classifier to still classify it as the same number without errors. Therefore, the feature map information after pooling has certain rotation, translation, and telescopic invariance.

5) FULLY CONNECTED LAYER

The fully connected layer is actually a multi-layer perceptron network, which is fully connected to the features or activation data of the previous layer output. Like the general neural network, the activation of the fully connected layer is only the neuron weight and bias. After the convolution kernel is imaged, the image is imported into the softmax classifier through the fully connected layer to obtain the classification result. For the softmax classification, the sum of the output probabilities obtained from the fully connected layer is one, that is, softmax converts the input eigenvector into a numerical vector between zero and one, and the sum is one.

The basic training process of CNN can be summarized as follows:

1. Random or zero initialization of the parameters of the network convolution kernel.
2. Input training samples to the CNN network, through forward propagation process such as convolution, activation, pooling, and full connectivity layers, get the output probability of each class.
3. Calculate the error using the loss function at the output layer.
4. Using a back-propagation algorithm to modify the network parameters, moving the weight of the convolution kernel to a method of reducing the error gradient, thereby updating the parameter weights of all convolution kernels to minimize the error.
5. Repeat steps 1-4 for all image samples in the training data.

C. STRUCTURE OF 3D CONVOLUTIONAL NEURAL NETWORKS

Figure 5 is a schematic structural diagram of a conventional two-dimensional CNN, which is mainly composed of a convolutional layer, a pooled layer, and a fully connected layer. It can be seen that the data processed by the network is in a two-dimensional format.

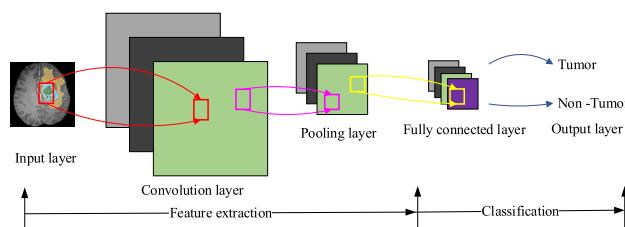


FIGURE 5. Schematic diagram of two-dimensional CNN structure.

The three-dimensional CNN is similar to the two-dimensional CNN. The difference is that the three-dimensional CNN performs convolution operations on three-dimensional data, and the convolution kernel and output data of network are all in three-dimensional format. The schematic diagram of the convolution is shown in Figure 6. As can be seen from Figure 6, the 3D convolution consists of stacking multiple consecutive images into one cube and then applying a 3D convolution kernel in the cube. In this structure,

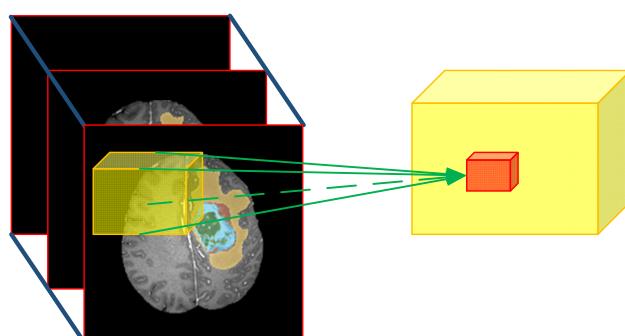


FIGURE 6. 3D convolution diagram.

each feature map in the convolutional layer is connected to a plurality of adjacent images in the upper layer, thereby capturing feature information of the image.

If the three-dimensional data size of the network input is $W \times H \times D$, the size of the three-dimensional convolution kernel is $m \times n \times h$, and the convolution step is 1, the output size of the three-dimensional CNN is $(W - m + l) \times (H - n + l) \times (D - h + l)$. In the three-dimensional CNN, the formula for calculating the value of the voxel point of feature block (x, y, z) of i-th layer, like the formula (4).

$$g_{i,j}^{x,y,z} = f(\sum_m \sum_p \sum_q \sum_r w_{i,j,m}^{p,q,r} \times X_{(i-1)m}^{(x+p)(y+q)(z+r)} + b_{i,j}) \quad (4)$$

The variables of x , y and z represent the pixel values of pixel (x, y) in the three-dimensional MRI brain tumor. The variables $w_{i,j,m}^{p,q,r}$ represents the (i, j, r) between the i-th in the j-th feature map of i-th layer and the weight between the m-th feature map of the 1st layer. The variable $b_{i,j}$ represents the offset value; $f(\bullet)$ is the activation function.

After three-dimensional convolution, the network also needs to be pooled. The three-dimensional pooling is similar to the two-dimensional pooling. The three-dimensional pooling reduces the characteristic dimension of brain tumor lesions and improves the feature significance of tumor lesions. The area is $F_1 \times F_2 \times F_3$, and the calculation formula of three-dimensional pooling is equation (5).

$$Y_{l,m,n} = \max_{0 \leq i \leq F_1, 0 \leq j \leq F_2, 0 \leq k \leq F_3} (X_{l \times s+i, m \times t+j, n \times r+k}) \quad (5)$$

The variable $X_{l \times s+i, m \times t+j, n \times r+k}$ is the eigenvalue of the three-dimensional pixel point of the lesion ($l \times s + i, m \times t + j, n \times r + k$). The variables of s , t and r are the moving steps of the three-dimensional pooling. Moreover, the variable $Y_{l,m,n}$ is the output value after the three-dimensional pool.

D. MULTIMODAL AND IMAGE PREPROCESSING

Multimodality is a very broad concept, and images and speech are information about two different modes. Assuming that a person is now identified, using both image and voice information, recognition based on these two kinds of information is the application of a multimodal information, the image provides visual features, and the voice provides auditory features. Usually, with a kind of information, humans can basically know a person, but it is not so reliable for a machine. However, if both types of information are used at the same time, it is clear that the correct rate of recognition will be much higher. For MRI images, the concept of multimodality is very similar to the concept of common multimodality, as well as different descriptions of the same target, except that the multimodality in MRI images is obtained using different nuclear magnetic visualization sequences, including T1, T1c, T2, and Flair images of four modalities. The main difference between the four modal images is that the same tissue is presented by different development sequences at different grayscale contrasts.

In common image segmentation, recognition, or detection systems, image preprocessing is often a critical step influencing system performance. The same is true for brain tumor detection problems in this paper. Since registration and skull separation have been completed, the focus here is on grayscale normalization and contrast adjustment for multimodal MRI images. The gray level of MRI images is usually 16 bits or higher, the scattered gray scale distribution and large gray value have a certain negative impact on the improvement of brain tumor segmentation accuracy. To do this, perform grayscale normalization on the MRI data before performing the segmentation process:

$$g(x, y) = ((GWM - BWM)/(h_{\max} - h_{\min})) \\ (h(x, y) - h_{\min}) + BWM \quad (6)$$

where $h(x, y)$ is the original histogram of the image, and the variable h_{\max} and the variable h_{\min} are its maximum and minimum gray levels. The variable $f(x, y)$ is a histogram of the image after normalization. GWM and BWM are the maximum and minimum gray levels of the histogram of the image after normalization.

After normalization, usually MRI images still cannot meet the needs of brain tumor detection, because the contrast difference between different tissues may make the segmentation robustness greatly reduced. The original MRI image is adjusted using itk-snap [37], and the slices sequence of the multimodal MRI image is extracted, as shown in Figure 7.

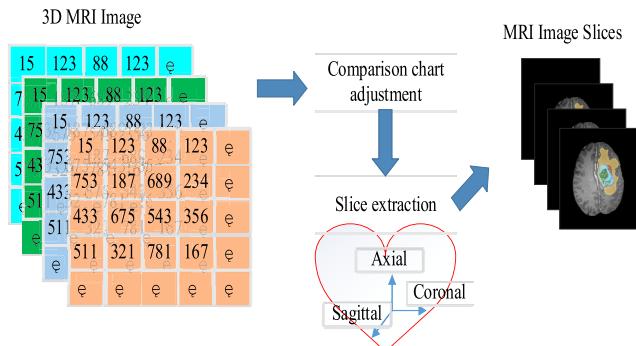


FIGURE 7. MRI image preprocessing process.

III. BRAIN TUMOR DETECTION ALGORITHM COMBINED WITH MULTIMODAL FUSION AND CNN

At present, most of the 3D medical image data is sliced into 2D data, and then the 2D slice data is used to train the network. This paper uses a 3D network of brain tumor detection to train 3D MRI images. The network structure is shown in Figure 8.

The whole network is divided into two parts: the contraction path and the expansion path. The purpose of the contraction path is feature extraction, which is composed of 3D-CNNs improved by the classic 2D-CNNs network; the expansion path can be used for brain tumor detection based on the features extracted from the contraction path, consists of

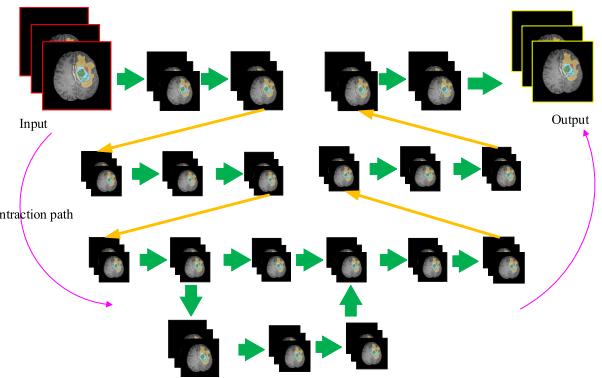


FIGURE 8. Schematic diagram of brain tumor detection network.

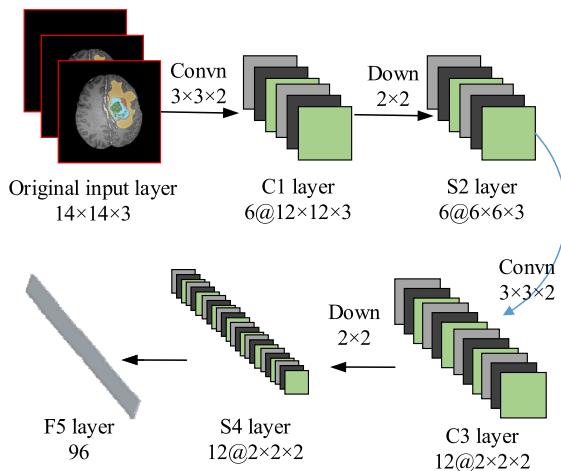
three-dimensional up sampling, feature cascading, and output layers.

A. MULTIMODAL 3D-CNNs FEATURE EXTRACTION NETWORK

The classic 2D-CNNs input image size is fixed, and then the feature is extracted from the whole image. Applying it to MRI brain tumor detection will have the following problems: First, to detect brain tumor images, individual pixel points must be classified. Therefore, the original input can only be the neighborhood of a single pixel, and the size of this neighborhood is difficult to grasp. Secondly, the size of brain tumors varies from patient to patient, and the size of brain tumors in different image layers of the same patient is different, even though the training layer. The neighborhood value of the original input layer is also difficult to ensure that this neighborhood is suitable for all tumor points of this patient. Third, how to make full use of MRI multimodal information to achieve higher precision classification is very important.

MRI images of different modal brain tumors can reflect the different characteristics of tumor and non-tumor regions. Therefore, if multiple information is fused and sent to the network for training, the accuracy of brain tumor detection will be improved. Therefore, it is necessary to A modal MRI three-dimensional angiography was used to construct high-dimensional brain tumor features. This paper improves the classical 2D-CNNs and builds a multi-modal 3D-CNNs network to achieve multi-modal information fusion.

The network of multimodal 3D-CNNs is shown in Figure 9. Small neighborhoods of the same position of four modalities that the specific neighborhood size is obtained according to the training data grid constitute the 3D original input layer. For example, if the size of small neighborhoods is 14×14 , the size of the 3D original input layer is $14 \times 14 \times 2$. A convolution template of $3 \times 3 \times 2$ size shared by six weights is used to convolve the original input layer to obtain six feature maps C1, and the size is $12 \times 12 \times 3$. The feature maps of the six C1 layers are respectively subjected to 2D average down sampling to obtain the S2 layer. After summing all the features of the S2 layer and passing through 12 convolution templates, the size is $3 \times 3 \times 2$. And we can

**FIGURE 9.** Multimodal 3D-CNNs.

obtain 12 feature map C3 layers, the size is $6 \times 6 \times 2$. The C3 layer is averaged down sampled to obtain the S4 layer. The S4 layer is normalized by column to obtain a 96-dimensional feature vector F5.

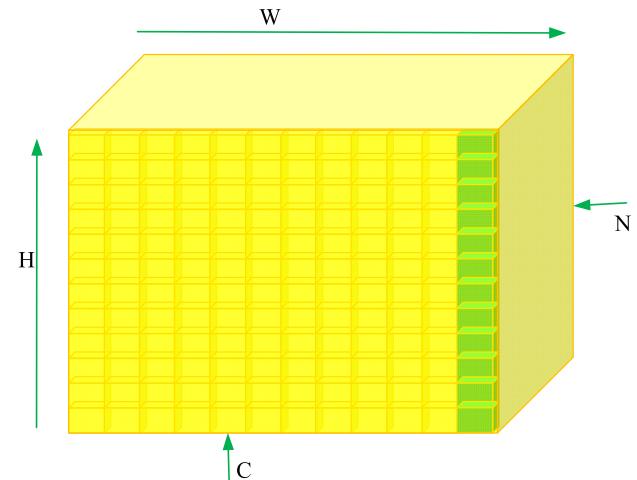
Multi-modal 3D-CNNs the original input layer is composed of four modes. The 3D convolution automatically extracts the difference information between the modes. The supervised learning method realizes different classification features for different patient difference information. Down sampling allows feature extraction to contain more structural edge information while eliminating redundant information and noise; multimodal common input makes the original input require less neighborhood information to accommodate tumor points in different image layers and improve brain tumors detection accuracy.

B. INSTANCE NORMALIZATION

In the deep neural network, the number of layers in the network is deep, and the neurons between different layers interact with each other. If the data is jittered at a certain layer of the network, the problem of data jitter will become stronger as the network deepens. The training optimization process affecting the model, normalization can effectively alleviate the problem of data oscillation [38].

Currently used data normalization methods include Batch Normalization (BN), Group Normalization (GN), Weight Normalization (WN), and Instance Normalization(IN), and so on [39]. In this experiment, the batch size of the training set and the test set are set to one, and BN, LN, and WN are all for the case where the batch value is large, and the IN can handle the smaller batch value. In this case, the normalization of brain tumor data by IN is used to speed up the convergence of the brain tumor detection network while solving the data jitter. The schematic diagram of the IN method is shown in Figure 10.

From the N direction of Figure 10, it is a picture of brain tumor, while from the C direction, it is the pixel of the characteristic map of brain tumor. Every 11 to vertical arrangement

**FIGURE 10.** Schematic diagram of the instance normalization method.

of small cubes rectangle represents a feature of a brain tumor image figure. The dark block is the part that is normalized together, so it can be clearly seen that IN refers to the single channel of a single brain tumor image that is normalized separately.

The main steps of the IN algorithms are:

- 1) Calculate the mean value of each brain tumor picture along the channel μ

$$\mu_{ij} = (1/HW) \sum_{l=1}^W \sum_{m=1}^H x_{ijlm} \quad (7)$$

where, variable $\sum_{l=1}^W \sum_{m=1}^H x_{ijlm}$ is the data that needs to be normalized.

- 2) Calculate the variance σ of each brain tumor picture along the channel

$$\sigma_{ij}^2 = (1/HW) \sum_{l=1}^W \sum_{m=1}^H (x_{ijlm} - m\mu_{ij})^2 \quad (8)$$

- 3) Combining equations (7) and (8), normalize the input brain tumor image to obtain normalized data y

$$y_{ijkg} = (x_{ijlm} - \mu_{ij}) / \sqrt{\sigma_{ij}^2 + \zeta} \quad (9)$$

The variable y_{ijkg} is the output after instance normalization. The advantage of adding the IN layer is that layer-by-layer normalization avoids gradient disappearance and gradient overflow, reduces the network's dependence on initialization such as weights, and accelerates network convergence. It can serve as a regularization approach that reduces the need for a network for dropout.

C. LOSS FUNCTION CONSTRUCTION

In the final output layer of the network of brain tumor detection, the brain tumor classification probability map with the same dimension as the input brain tumor data is output, indicating whether each individual prime point is a brain

tumor probability. The network of brain tumor detection calculates the output value through the loss function. In addition, the error value between the tag value and the reverse parameter correction of the network according to the error value to obtain the best weight and offset value of the network of brain tumor detection. In the target detection and target detection tasks of natural images, commonly used the loss function is loss, and its calculation formula is

$$\text{loss} = 1 - \frac{2 \sum_{i=0}^N p_i g_i}{\sum_{i=0}^N p_i^2 + \sum_{i=0}^N g_i^2} \quad (10)$$

The variable p_i represents the value of i-th pixel point. In addition, g_i represents the ground truth, respectively.

In the three-dimensional MRI brain tumor imaging, due to the particularity of medical images, compared with natural images, in the whole brain tumor image, the brain tumor lesions account for a relatively small proportion, and the non-lesion area accounts for a larger proportion. Dice correlation coefficient as a loss function, in the network training process, the network tends to learn the characteristics of non-brain tumor lesions, cannot effectively extract the characteristics of brain tumor lesions, resulting in false detection and missed detection. Therefore, improve the learning ability of the network on the brain tumor area, and improve the traditional loss function. The improved loss function formula is

$$\text{loss} = 1 - \frac{2 \times \sum_{i=0}^N ((1/4)p_i \times (3/4)g_i)}{\sum_{i=0}^N p_i^2 + \sum_{i=0}^N g_i^2} \quad (11)$$

According to equation (10), since the g_i part corresponds to the lesion area of brain tumor, weight weighting was carried out on the g_i part, and the ratio of predicted results and true values in the loss function was 1:3. The loss function has a larger loss coefficient for the true value distribution. It can strengthen the network's characteristic learning of the lesion area of brain tumor, weaken the distribution of the loss value of the network to the non-tumor area, and reduce the interference of the brain MRI background image on the characteristic learning of the lesion area, thereby improving the detection accuracy of the network.

IV. EXPERIMENTS AND RESULTS

A. EXPERIMENTAL DATA SET

The experimental data set is the actual patient data provided by MICCAI BraTS 2018 [21], of which 220 have advanced gliomas and 54 have lower grade gliomas. In the experimental data, each patient has four modal MRI images, namely FLAIR, T1, T1c, and T2. The dimension of each MRI image is $155 \times 240 \times 240$. All images are stored as signed 16-bit integers, but only non-negative values are used. Each data is labeled as five categories at the pixel level, which are normal

tissue (or background), necrosis, edema, non-enhanced and enhanced tumors.

Due to the limited source of medical image acquisition, the sample is augmented based on the existing data set. Firstly, the image with the size of 240×240 is cropped, and only the area containing the brain is retained, and the surrounding 0 pixel point is removed. Then use the nearest neighbor interpolation method to enlarge the brain image after removing the background, highlighting the image characteristics of the brain; finally, the processed sample is rotated, mirrored, and flipped. The schematic diagram of sample augmentation of brain tumor is shown in Figure 11. As can be seen from Figure 11, the ROI image is first acquired by cropping the original image, and the size of the ROI image is enlarged to the same size as the original image by the nearest neighbor interpolation method. The new sample image is then obtained by rotating different angles and flipping.

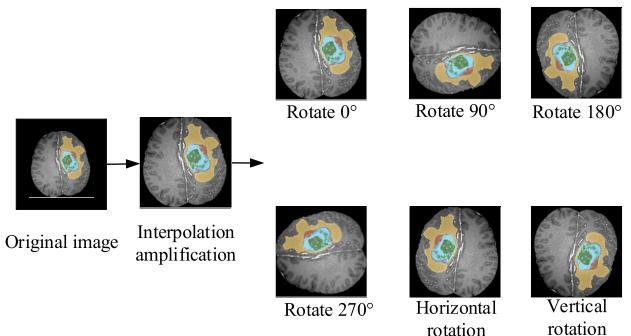


FIGURE 11. The schematic diagram of sample augmentation of brain tumor.

B. EVALUATION STANDARD

In order to verify the effectiveness of the algorithm of brain tumor detection, dice similarity coefficient, sensitivity (SN) and specificity (SE) were used to quantify the brain tumor detection results. Dice reflects the detection of network of brain tumor detection. The degree of similarity between the result and the true value, the larger the value indicates the higher the accuracy of the detection result and the formula is (12).

$$\text{dice}(A, B) = 2|A \cap B|/(|A| + |B|) \quad (12)$$

where, A is the lesion detection result of the brain tumor network, and B is the true value of the brain tumor lesion area.

The SN reflects the ratio of the voxel point of the correct brain tumor area to the sum of the true voxel points, which can be calculated according to the true positive (TP) and the false negative (FN), and the formula is (13).

$$SN = TP/(TP + FN) \quad (13)$$

SE reflects the probability that non-tumor tissue is correctly determined to be part of normal tissue, and the formula is (14).

$$SE = TN/(TN + FP) \quad (14)$$

In the formula, TN is the number of voxel points that are correctly recognized in the non-tumor region; FP is the number of voxel points in which the non-tumor region is misjudged as the tumor region.

C. COMPARISON OF SINGLE-MODAL AND MULTI-MODAL DETECTION RESULTS

In order to verify the effectiveness of the multimodal fusion method used in this paper, the single-mode Flair and multimodal fusion MRI images were compared experimentally. The pre-assigned 228 samples were used as training samples and 57 samples were used as test samples. The network convergence map of the training set of the network of brain tumor detection and test set is shown in Figure 12 and Figure 13.

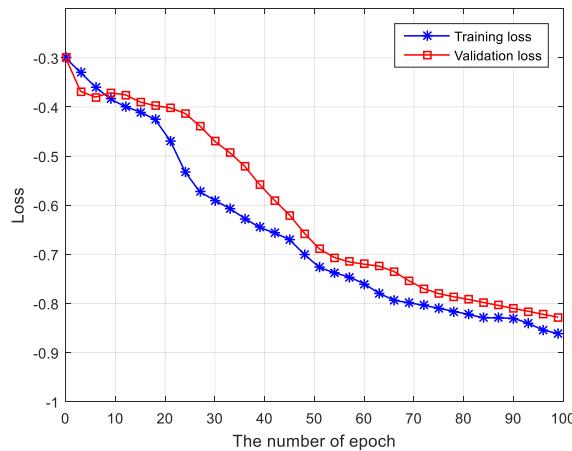


FIGURE 12. Schematic diagram of network convergence for single-mode brain tumor detection.

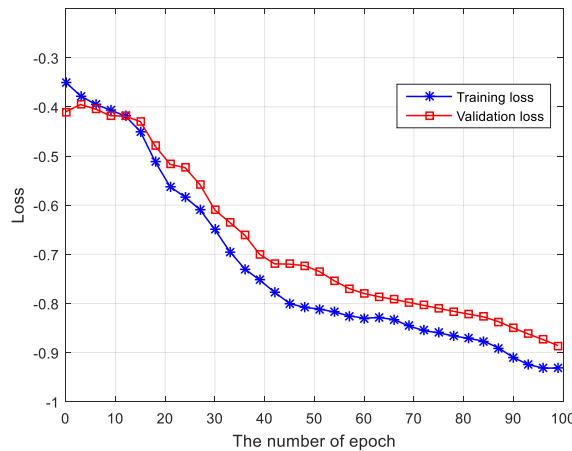


FIGURE 13. Schematic diagram of network convergence for multimodal brain tumor detection.

As can be seen from Figure 12 and Figure 13, the loss values of the training network and test set of the single-mode network of brain tumor detection are -0.883 and -0.830 , respectively, and the loss values of the multimodal network of brain tumor detection can reach -0.912 and -0.858 . Compared with the single-mode network, the loss value of the multi-modal network is smaller, and the training loss value

and the test loss value are increased by 0.029 and 0.027 , respectively, indicating that the performance of the multimodal brain tumor detection network is better, but the convergence from the loss function. From the curve, the multimodal network of brain tumor detection has a problem of overfitting.

Figure 14 and Figure 15 show the dice distribution of the evaluation index of 57 samples in the test set. Compared with Figure 14 and Figure 15, there are 4 samples in the single-mode network away from the cluster point, and only 2 samples in the multi-modality are far away from the cluster point.

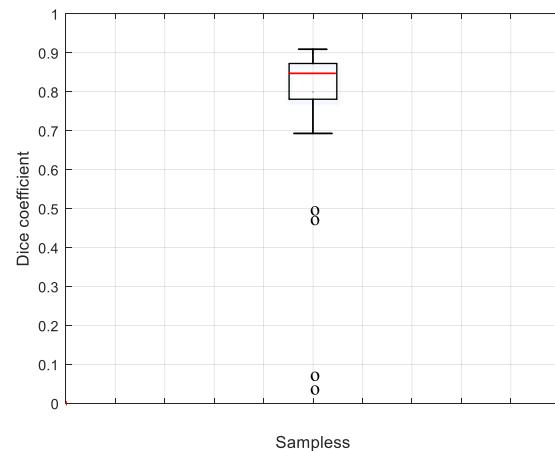


FIGURE 14. Dice distribution of single-mode test set of brain tumor.

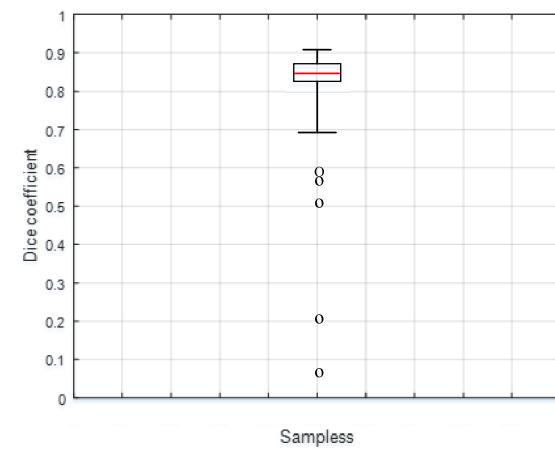


FIGURE 15. Dice distribution of multimodal test set of brain tumor.

Table 1 compares the three evaluation indexes of 57 samples of different modal training sets of the network of brain

TABLE 1. Comparison of three evaluation indicators for different modal brain tumors.

Method	Single modal	Multi modal
Dice	0.795	0.843
SN	0.824	0.825
SE	0.996	0.997

tumor detection. It can be seen that the three evaluation indicators dice, SN and SE of the multimodal brain tumor detection network increased by 4.7%, 0.13%, and 0.07%, respectively. The quantitative results from the test set indicate that the multimodal brain tumor detection method is better.

D. INSTANCE NORMALIZED TEST RESULTS COMPARISON

Schematic diagram of the normalized network convergence is shown as Figure 16 and Figure 17. In addition, the dice index distribution of the brain tumor test set. Comparing Figure 13 and Figure 16, the network of brain tumor detection combined with the example normalization has a faster convergence rate. The loss-value curve of the set and the test set is well fitted, indicating that the method of instance normalization not only accelerates the convergence speed of the network, but also alleviates the over-fitting problem in the multi-modal brain tumor detection network.

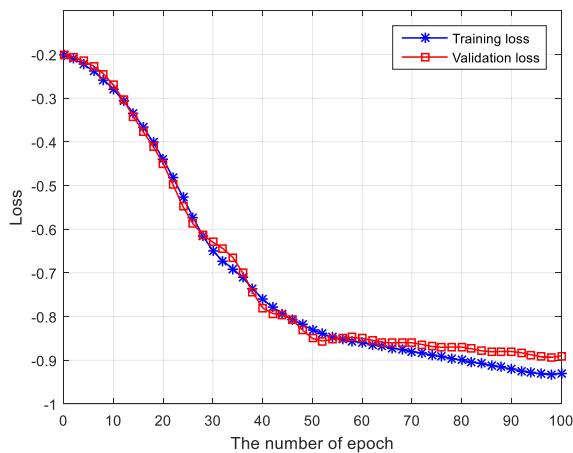


FIGURE 16. Schematic diagram of network convergence based on real-world normalization.

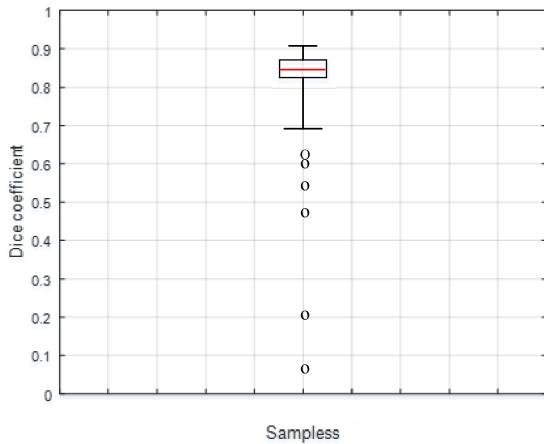


FIGURE 17. Dice distribution of brain tumor test set.

From the distribution of dice index of 57 cases of brain tumor test set shown in Figure 17, it can be seen that compared with Figure 15, the dice value with the lowest normalized brain tumor test result can reach 0.41, and only two serious points away from the cluster.

It can be seen from Figure 18 that the three evaluation indexes dice, SN and SE of the brain tumor detection network can reach 0.882, 0.894, and 0.997 combined with the real-normalization method, and more modal brain tumor detection models, dice and SN. The indicators increased by 3.91% and 6.82%, while the SE indicator decreased by 0.03%.

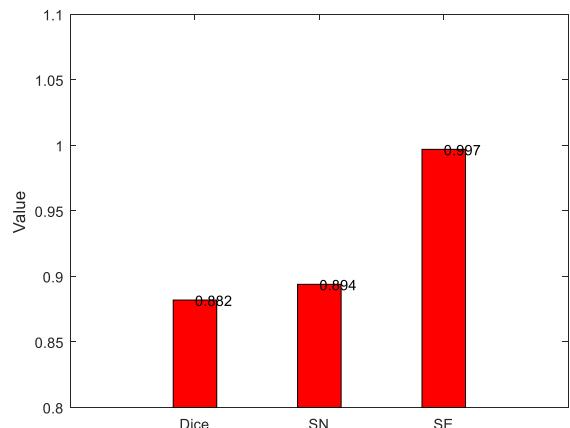


FIGURE 18. Evaluation index of brain tumor detection model combined with real column.

E. COMPARISON OF WEIGHTED LOSS FUNCTION DETECTION RESULTS WITH 3D-CNNs

To verify the effectiveness of the improved weighted loss function, the convergence graph of loss function and the test set dice index distribution of the network training are shown in Figure 19 and Figure 20, respectively. As can be seen from Figure 19 and Figure 20, compared with Figure 16 and Figure 17, this paper uses the weighted loss function is used to train the brain tumor detection network, strengthen the characteristic learning of the brain tumor lesion area, and weaken the characteristic expression of the non-focal area. The final training and test loss value of the network can reach -0.875 and -0.912 , the lowest. The dice value can also reach 0.07.

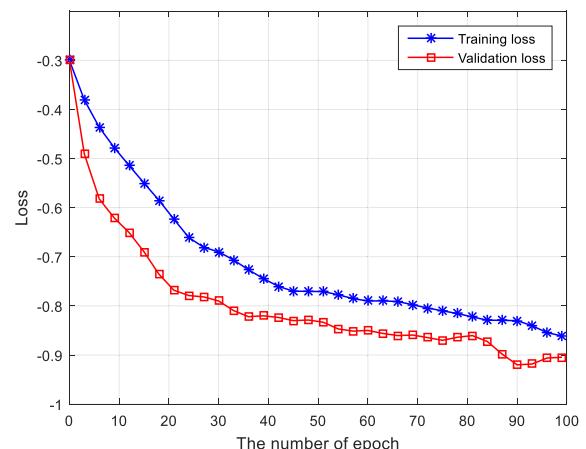


FIGURE 19. Schematic diagram of network convergence combined with weighted loss function.

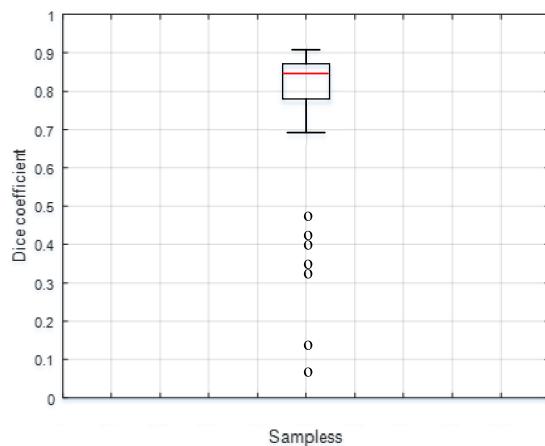


FIGURE 20. Brain tumor test set dice distribution.

F. COMPARISON OF 2D AND 3D INSPECTION NETWORK RESULTS

In order to verify the three-dimensional CNN, the experimental results of the two-dimensional detection network and the three-dimensional detection network are compared. Figure 21 is a schematic diagram of the convergence of the two-dimensional detection network.

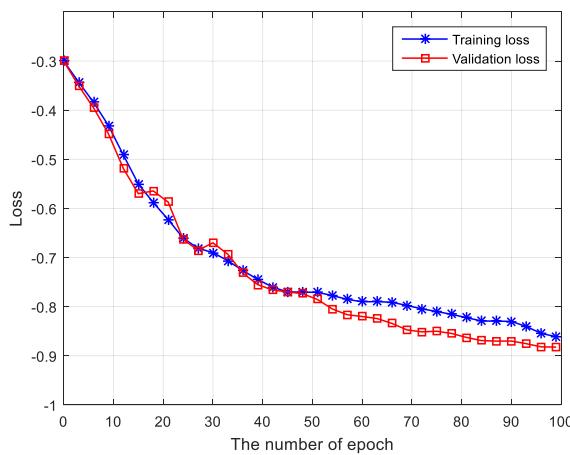


FIGURE 21. Schematic diagram of two-dimensional detection network convergence.

Comparing Figure 21 and Figure 19 it can be seen that the loss function of the three-dimensional detection network is better fitted, the training and test loss values of the two-dimensional detection network are -0.874 and -0.891 , and the three-dimensional detection network can reach -0.875 and -0.912 , respectively increased by 0.1% and 2.3% .

The two-dimensional detection network and the three-dimensional detection network evaluation of brain tumors are shown in Table 2. Compared with the two-dimensional detection network, the results of the three-dimensional detection networks dice, SN and SE are improved by 4.5% , 3.9% and 0.04% , respectively. Then the reliability of 3D detection network is verified.

TABLE 2. Comparison of 2D and 3D network evaluation indicators.

Method	2D	3D
Dice	0.881	0.927
SN	0.886	0.928
SE	0.997	0.998

G. COMPARED WITH OTHER METHODS

In order to verify the advantages of brain tumor detection methods combined with multimodal information fusion and CNNs, the methods of this paper were compared with those of literature [31], literature [32], literature [33], and literature [34].

The experimental results are shown in Table 3. It can be clearly seen that the algorithm has obvious advantages in three evaluation indicators. This is because, first, brain tumors are generally close to spherical, and the tumor size of different layers in the same patient is different. The 2D-CNNs model trained on the training layer is difficult to adapt to the image of the entire patient's tumor layer; secondly, the four models are respectively the 2D-CNNs feature extraction can theoretically obtain more rich information between different modes, but too much feature information increases the linear indivisibility of each pixel, making the detection result worse. Multimode state 3D-CNNs not only overcomes the shortcomings of the above classic 2D-CNNs, but also facilitates the combination of four modes, which is more conducive to the combination of different information between modes, and removes redundant information to promote the realization of effective classification. Finally, by constructing a new weighted loss function, weakening the interference of non-focal areas on brain tumor detection can reduce the interference of non-focal areas on brain tumor detection, and thus improve the detection accuracy.

TABLE 3. Comparison results of different algorithms.

Method	Dice	SN	SE
Literature [31]	0.812	0.811	0.823
Literature [32]	0.835	0.836	0.851
Literature [33]	0.799	0.801	0.825
Literature [34]	0.891	0.898	0.911
Multi-CNNs	0.927	0.928	0.998

V. CONCLUSION

In this paper, a method with three-dimensional MRI brain tumor detection combining multimodal information fusion and CNN is proposed. Firstly, it is better to use the improved multimodal 3D-CNNs to obtain the three-dimensional features of brain tumor lesions under different modalities. Extract the difference information between the various modes. Secondly, in order to solve the problem that the network convergence is slow and the over-fitting is serious, the brain tumor characteristic data is normalized. Then, according to the small volume of the lesion area and the large volume of the non-focal area, a new weighted loss function

is constructed to weaken the interference of the non-focal area on the detection of brain tumors, and the loss function can be improved to reduce the detection of brain tumors in non-focal areas. The experimental results show that the three evaluation indexes of dice, SN and SE are optimized respectively, and the two-dimensional brain tumor detection network and the original single-mode brain tumor detection method are compared. There has been a big improvement.

REFERENCES

- [1] P. M. Shakeel, T. E. El Tobely, H. Al-Feel, G. Manogaran, and S. Baskar, "Neural network based brain Tumor detection using wireless infrared imaging sensor," *IEEE Access*, vol. 7, pp. 5577–5588, 2019.
- [2] S. Khajanchi and S. Banerjee, "A strategy of optimal efficacy of T11 target structure in the treatment of brain Tumor," *J. Biol. Syst.*, vol. 27, no. 1, pp. 1–31, 2019.
- [3] M. A. Khan, I. U. Lali, and A. Rehman, "Brain Tumor detection and classification: A framework of marker-based watershed algorithm and multilevel priority features selection," *Microsc. Res. Technique*, vol. 82, no. 6, pp. 909–922, 2019.
- [4] T. Cao, W. Wang, S. Tighe, and S. Wang, "Crack image detection based on fractional differential and fractal dimension," *IET Comput. Vis.*, vol. 13, no. 1, pp. 79–85, Feb. 2019.
- [5] M. Ramalho, A. P. Matos, and M. Aloabdy, "Magnetic resonance imaging of the cirrhotic liver: Diagnosis of hepatocellular carcinoma and evaluation of response to treatment—Part 1," *Radiologia Brasileira*, vol. 50, no. 1, pp. 38–47, 2017.
- [6] Q. Li, Z. Gao, and Q. Wang, "Glioma segmentation with a unified algorithm in multimodal MRI images," *IEEE Access*, vol. 6, pp. 9543–9553, 2018.
- [7] W. Wu, A. Y. C. Chen, and L. Zhao, "Brain tumor detection and segmentation in a CRF framework with pixel-pairwise affinity and super pixel-level features," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 9, no. 2, pp. 241–253, 2014.
- [8] V. P. Ananthi, P. Balasubramaniam, and T. Kalaiselvi, "A new fuzzy clustering algorithm for the segmentation of brain Tumor," *Soft Comput.*, vol. 20, no. 12, pp. 4859–4879, 2016.
- [9] I. Elaff, "Comparative study between spatio-temporal models for brain Tumor growth," *Biochem. Biophys. Res. Commun.*, vol. 496, no. 4, pp. 1263–1268, 2018.
- [10] S. Bonte, I. Goethals, and R. H. Van, "Machine learning based brain Tumour segmentation on limited data using local texture and abnormality," *Comput. Biol. Med.*, vol. 98, pp. 39–47, 2018.
- [11] A. Feizollah, N. B. Anuar, and R. Salleh, "Evaluation of network traffic analysis using fuzzy c-means clustering algorithm in mobile malware detection," *Adv. Sci. Lett.*, vol. 24, no. 2, pp. 929–932, 2018.
- [12] U. Konur, "Computerized detection of spina bifida using SVM with Zernike moments of fetal skulls in ultrasound screening," *Biomed. Signal Process. Control*, vol. 43, pp. 18–30, 2018.
- [13] C. Liu, S.-Y. Chen, C.-C. Chen, and C.-H. Tai, "Detecting newly grown tree leaves from unmanned-aerial-vehicle images using hyperspectral target detection techniques," *ISPRS J. Photogramm. Remote Sens.*, vol. 142, pp. 174–189, May 2018.
- [14] T. Ogunfunmi, R. P. Ramachandran, and R. Tognoni, "A primer on deep learning architectures and applications in speech processing," *Circuits Syst. Signal Process.*, vol. 38, no. 10, pp. 3406–3432, 2019.
- [15] J. Jeong, T. S. Yoon, and J. B. Park, "MOSnet: Moving object segmentation with convolutional networks," *Electron. Lett.*, vol. 54, no. 3, pp. 136–138, 2018.
- [16] S. Zhao, L. Zehua, and H. Liang, "A mixed non-local prior model for image super-resolution reconstruction," *Chin. J. Electron.*, vol. 26, no. 4, pp. 778–783, 2018.
- [17] M. Gnouma, A. Ladjalia, and R. Ejbali, "Stacked sparse autoencoder and history of binary motion image for human activity recognition," *Multimedia Tools Appl.*, vol. 78, no. 5, pp. 1–23, 2018.
- [18] G. M. Jiji and J. Dehmeshki, "Automatic CT brain image segmentation using two level multiresolution mixture model of EM," *J. Inst. Eng. India B*, vol. 95, no. 2, pp. 91–99, 2014.
- [19] C. Vijayakumar and D. C. Gharpure, "Development of image-processing software for automatic segmentation of brain Tumors in MR images," *J. Med. Phys.*, vol. 36, no. 3, pp. 147–158, 2011.
- [20] A. V. Nasrullah, C. G. Willcocks, P. T. G. Jackson, C. Geenen, M. S. Habib, D. H. W. Steel, and B. Obara, "Multi-scale segmentation and surface fitting for measuring 3-D macular holes," *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 580–589, Feb. 2018.
- [21] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, "A deep learning model integrating FCNNs and CRFs for brain Tumor segmentation," *Med. Image Anal.*, vol. 43, pp. 98–111, Jan. 2017.
- [22] G. Shi, Y. He, B. Li, L. Zuo, B. Yin, W. Zeng, and F. Ali, "Analysis and modeling of wireless channel characteristics for Internet of Things scene based on geometric features," *Future Gener. Comput. Syst.*, vol. 101, pp. 492–501, 2019.
- [23] S. Yan, H. Bo, and Z. Ying, "Segmentation of sidescan sonar imagery using Markov random fields and extreme learning machine," *IEEE J. Ocean. Eng.*, vol. 44, no. 2, pp. 502–513, Apr. 2019.
- [24] M. Safa, S. Samarasinghe, and M. Nejat, "Prediction of wheat production using artificial neural networks and investigating indirect factors affecting it: Case study in Canterbury province, New Zealand," *J. Agricult. Sci. Technol.*, vol. 17, no. 4, pp. 791–803, 2018.
- [25] V. Govindan, R. S. Chakraborty, and P. Santikellur, "A hardware trojan attack on FPGA-based cryptographic key generation: Impact and detection," *J. Hardw. Syst. Secur.*, vol. 2, no. 3, pp. 225–239, 2018.
- [26] Z. Zhai, Z. Jin, and R. Zhang, "Information integration of force sensing and machine vision for in-shell shrivelled walnut detection based on the golden-section search optimal discrimination threshold," *J. Sci. Food Agricul.*, vol. 99, no. 8, pp. 3941–3949, 2019.
- [27] L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, and F. R. Murtagh, "Automatic segmentation of non-enhancing brain Tumors in magnetic resonance images," *Artif. Intell. Med.*, vol. 21, nos. 1–3, pp. 43–63, 2001.
- [28] B. H. Menze *et al.*, "The multimodal brain Tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.
- [29] Y. Hu, M. Grossberg, and G. Mageras, "SU-D-BRA-03: What image features are useful for Tumor segmentation in multi-modal images?" *Med. Phys.*, vol. 42, no. 6, p. 3213, 2015.
- [30] S. Pereira, A. Pinto, J. Oliveira, A. M. Mendrik, J. H. Correia, and C. A. Silva, "Automatic brain tissue segmentation in MR images using random forests and conditional random fields," *J. Neurosci. Methods*, vol. 270, pp. 111–123, Sep. 2016.
- [31] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1240–1251, May 2016.
- [32] B. García-Zapirain, M. Elmogy, and A. El-Baz, "Classification of pressure ulcer tissues with 3D convolutional neural network," *Med. Biol. Eng. Comput.*, vol. 56, no. 6, pp. 1–14, 2018.
- [33] M. Havaei, A. Davy, and D. Wardle-Farley, "Brain Tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, 2017.
- [34] D. Lunga, H. L. Yang, A. Reith, J. Weaver, J. Yuan, and B. Bhaduri, "Domain-adapted convolutional networks for satellite image classification: A large-scale interactive learning workflow," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 962–977, Mar. 2018.
- [35] Y. Lecun, B. Boser, and J. S. Denker, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 2014.
- [36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [37] J. I. Rojas, L. Patrucco, and C. Besada, "Sex-related differences in atrophy and lesion load in multiple sclerosis patients," *Neurología*, vol. 28, no. 7, pp. 389–393, 2013.
- [38] M. B. Pereira, M. Wallroth, and V. Jonsson, "Comparison of normalization methods for the analysis of metagenomic gene abundance data," *BMC Genomics*, vol. 19, no. 1, pp. 274–291, 2018.
- [39] J. Wang, X. Chen, and W. Zhong, "Sensitivity study on quantitative methods by analysis of pressure fluctuations in a fluidized bed," *Ind. Eng. Chem. Res.*, vol. 57, no. 37, pp. 12551–12564, 2018.