Journal of Intelligent & Fuzzy Systems 38 (2020) 127–137
DOI:10.3233/IFS-179387
IOS Press

Medical brain image classification based on multi-feature fusion of convolutional neural network

Dan Wang[a,b], Hongwei Zhao[a,c,e,*] and Qingliang Li[d,e]
[a]*Department of Computer Science and Technology, Jilin University, Changchun, China*
[b]*College of Information Technology and Media, Beihua University, Jilin, China*
[c]*State Key Laboratory of Applied Optics, Changchun, China*
[d]*Changchun University of Science and Technology, Changchun, China*
[e]*Department of Symbolic Computing and Knowledge Engineering, Key Laboratory of the Ministry of Education, Jilin University, Changchun, China*

Abstract. This paper presents a medical brain image algorithm based on multi-feature fusion. Feature extraction based on convolutional neural network was used as texture information, feature extraction based on voxel information was used as morphological feature, and then the two types of features were combined in series. Feature extraction based on convolutional neural network was used as texture information, feature extraction based on voxel information was used as morphological feature, and then the two types of features were combined in series. Then the heuristic search algorithm is used to optimize the feature selection stage. Based on the feature score table extracted by the recursive feature elimination method of support vector machine, the correlation between features is added. Moreover, through experimental analysis, the optimal value of the parameter K was selected according to the heuristic search, and the optimal feature subset was extracted after determining the value of the parameter K. Experiments show that compared with similar algorithms, this algorithm improves the accuracy and efficiency of the classification of brain images.

Keywords: Convolutional neural network, multi-feature fusion, heuristic search, medical image classification

1. Introduction

At present, under the inspiration of the national policy, the introduction of artificial intelligence technology in the medical field has become a hot trend in the medical field. That is to say, the computer can analyze and process all kinds of medical images intelligently, quickly and accurately so as to assist experts in medical diagnosis. Medical image processing and analysis technology has played an irreplaceable role in the medical industry at home and abroad [1]. It has been applied in assistant diagnosis, screening, classification, treatment effect evaluation, treatment decision-making and guidance of health-threatening diseases. Among them, the computer-aided diagnosis technology based on medical image classification has attracted the attention of scholars and researchers. It has achieved good results in the diagnosis of brain function and mental disorders, benign and malignant tumors of body organs, cardiovascular and cerebrovascular diseases, etc. In clinical work, it can realize the screening of medical images and the checking and classification of lesions. As the convolutional neural network (CNN) in the deep learning

*Corresponding author. Hongwei Zhao, Department of computer science and technology, Jilin University, Changchun, China. E-mail: wangdanjl_dx@163.com.

ISSN 1064-1246/20/$35.00

theory has applied to the field of medical image classification. It can directly input medical images into the network, thus avoiding the poor classification accuracy caused by the unsatisfactory feature selection link [9].

In this paper, the detection of Alzheimer's disease is carried out, and two advantages of texture analysis and quantitative analysis are taken into account. In order to better describe the two states of Alzheimer's disease, a medical brain image classification method based on multi-feature fusion of convolutional neural network was proposed.

Firstly, we extract visual features from all medical brain images in the image database. It contains the texture information and contour information described by the convolutional neural network. And also the medical features of medical brain images described by voxel-based morphometry (VBM) method. Then, after obtaining the two types of features, they need to be normalized and fused. hen the recursive feature elimination strategy based on support vector machine is applied to the fusion feature. Thus, the feature dimension with little correlation can be filtered to improve the feature discrimination and reduce the dimension. Finally, large-scale linear support vector machine (LLSVM) is introduced, and input dimension-reduced features into it, an effective and time-saving classification method for Alzheimer's disease is realized.

This paper is divided into three parts. The first part introduces the application status of convolutional neural network in the field of image processing and analysis. The second part is the construction process of multi-feature fusion brain image classification model based on CNN. The third part is to verify the effectiveness of the multi-feature fusion medical brain image classification model based on CNN and design the training process of the model. Experiments have proved that the proposed algorithm can reduce the dimension of fusion features and remove the interference of redundant features. In the task of medical brain image classification, this method can greatly improve the classification performance of Alzheimer's disease and greatly reduce the computational complexity.

2. Related work

So far, Convolution Neural Network (CNN) has been widely used in medical image classification and recognition. The medical image retrieval system designed by Wang Yuanyuan has tested 5000 CT and MRI images of four human body parts, belonging to the classification retrieval of relatively rough level [2]. Qin Zhiguang proposed a method of organizing and retrieving massive medical images based on hierarchical depth learning, which can classify medical images according to different medical image acquisition devices, human body parts and diseases [3]. For the first time, Li Jian et al. Used CNN method to segment brain tumors and achieved remarkable results. Compared with traditional supervised machine learning, depth-based learning does not rely on manual feature extraction, and can automatically learn features from image data to adapt to the complexity and individual differences of brain tumors [4]. Therefore, the CNN method for brain tumor segmentation has been gradually concerned by researchers in the field.

In the research work of using depth learning to solve the problem of image segmentation, Ye Jiayin et al. proposed a full convolution neural network (FCN) model for the first time. The skip frame structure of this method combines the high-level image representation with the low-level image appearance features to realize the fine segmentation task [5]. However, this method can easily lose context information. Zhang Jun et al. proposed a pyramid pool module to solve this problem [6]. Wu Xinjie extracted the features of the original input image, the deeper the network, the stronger the feature expression ability. Wu Xinjie's remaining blocks can train the deep network model, but it will lead to feature reuse attenuation [7]. Wen Mengfei et al. proposed the method of Wide Residual Network (WRN). The performance of the shallow network model can be similar to that of the depth model by expanding the coefficients [8]. In addition, Liu Lei et al. Found that once the anatomical information of the image is extracted, the final segmentation result will be greatly influenced by the intensity of a particular voxel signal (relative to texture features or other higher-order features). Therefore, they re-introduce the original image in the penultimate layer of the network, which significantly improves the segmentation performance of the network [10, 11].

Therefore, the convolution neural network method is applied to the classification and recognition of brain images, taking patients with Alzheimer's disease and mild cognitive impairment as examples, a brain image classification model based on multi-feature fusion was constructed, and the effects of feature dimension and feature correlation on classification rate and running time were discussed, and the validity of the model was verified [12–14].

3. Classification model of multi-feature fusion medical brain images based on CNN

3.1. Model framework

Based on CNN model, this paper designs a medical brain image classification model based on multi-feature fusion. As shown in Fig. 1.

The main research content of this paper is the selection, extraction, fusion methods of medical brain image features, the dimension reduction method and classification algorithm of feature set after optimization fusion. Next, this paper will expatiate on each link of medical brain image classification framework based on multi-feature fusion.

3.2. Texture features

Texture features can describe the overall gray distribution of medical brain images, By analyzing and comparing the texture features of medical brain images, we can characterize the texture structure of the brain images of Alzheimer's disease patients, detect other brain images similar to the texture structure, and then diagnose the patients as suspected Alzheimer's disease, and provide help for the attending physician in the treatment. The traditional feature extraction method is based on machine learning algorithm and training for image pixels. This method describes the features with low discrimination. With the rise of deep learning, many scholars and researchers gradually use convolution neural network for feature learning, reduce the dependence on expert knowledge, and improve the performance of recognition. With the development of convolution neural network, a lot of improved versions have emerged, such as new elements in convolution network architecture, invariance of convolution features, and so on. The main process of extracting CNN features as texture features is as follows: Firstly, the convolution neural network model is used to extract the features of all images in the image database. In this section, the output of the second all-connected layer in the network model is used as the CNN features of medical brain images, and the final texture feature library of medical images is generated. The CNN feature described in this section is the 1024-dimensional, 2048-dimensional, or 2096-dimensional feature vectors that result from many convolution and cisternization operations of medical brain images in the network layer:

$$fc_2 = \sigma(\omega_1 \times fc_1) \tag{1}$$

Where fc_1 denotes the output of the first full connection layer and ω_1 denotes the weight of the output of the first full connection layer.

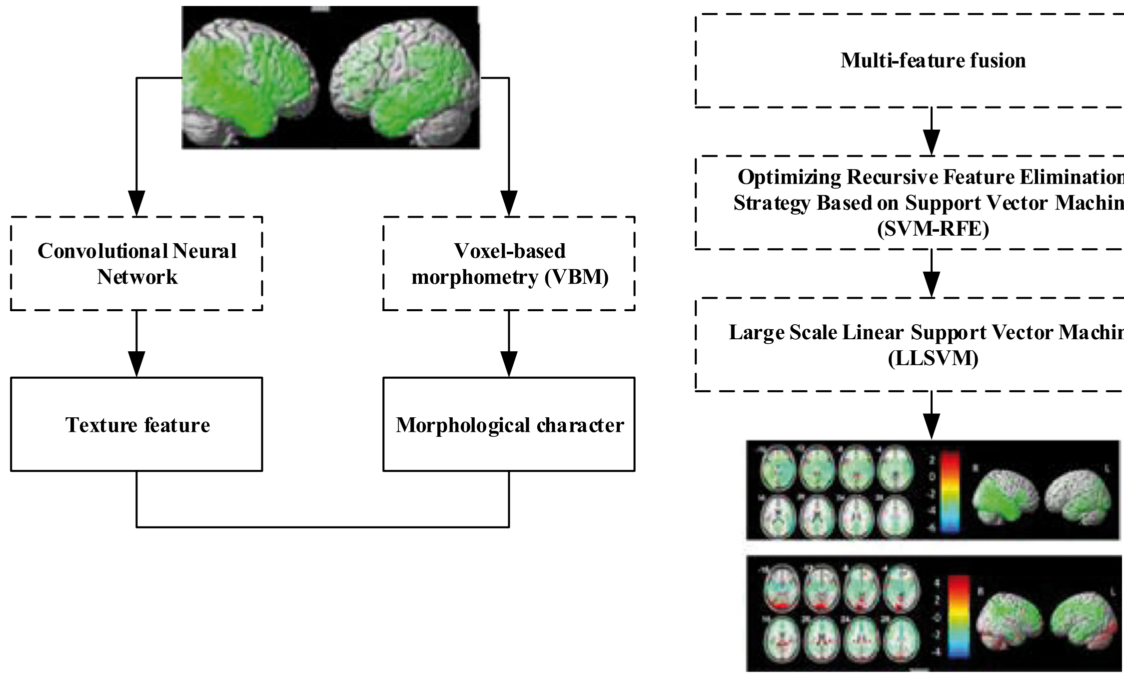

Fig. 1. Medical brain image classification framework based on multi-feature fusion.

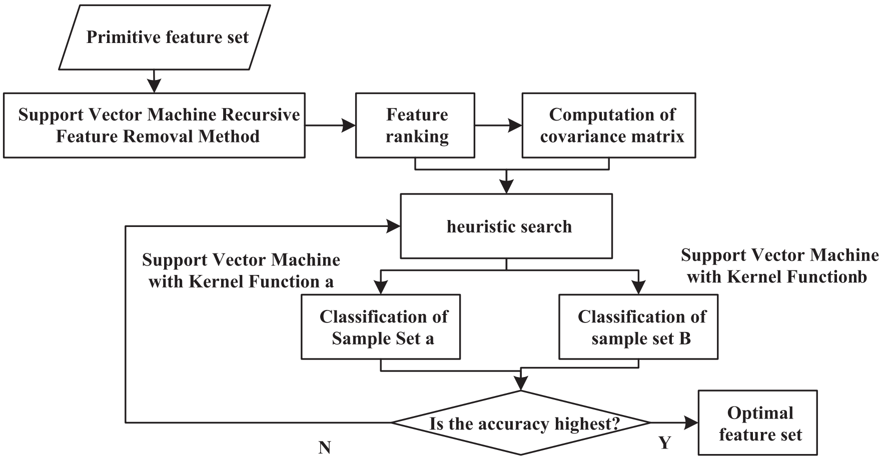

Fig. 2. Optimized support vector machine recursive feature removal framework.

3.3. Morphological characteristics

3.3.1. VBM feature extraction

In 2000, Ashburner et al. Proposed a voxel-based morphometry (VBM) to analyze brain tissue information. This method calculates the volume and probability density of brain tissue based on voxel information in three-dimensional images, and then compares the differences of brain tissue information among samples and their distribution areas. Therefore, this method is statistic and detection in the whole brain, which is sensitive to the abnormal information in different brain regions, so as to avoid misjudgment and missed judgment, and has strong practical value in the field of brain medical image application.

The process of morphological feature extraction based on VBM: Firstly, the brain images to be tested are registered into three-dimensional space in the same coordinate system by using standardized template, and then the brain gray matter, white matter, cerebrospinal fluid and so on are segmented according to different probability templates of brain tissues. Next, the segmented brain tissue information is preprocessed to improve the signal-to-noise ratio of the image, such as Gaussian smoothing. Finally, the image is analyzed to obtain the morphological features.

3.3.2. VBM statistical analysis

Three kinds of brain tissues, gray matter, white matter and cerebrospinal fluid, can be extracted by the above methods. The gray matter is mainly composed of neuronal cell bodies and is the main controller of the nervous system. The white matter is composed of nerve fibers, which is the signal transmission channel of nervous system. CSF is predominantly ventricular filling. Quantitative analysis of gray matter in medical images is an important method to detect functional diseases. Generalized Linear Modeling (GLM) is used in this section. Then the gray matter volume and probability density of two groups of images were detected according to the two samples in the region of specific significant difference.

The calculation process is as follows: Firstly, the GLM model is constructed. Assuming that the number of groups representing the sample and the number of J representing the sample is the gray value of the first random voxel of the brain tissue sample in the first group of samples. Applying the above data processing method, two sets of random samples with normal distribution and independent of each other can be obtained:

$$X_{ij} = \mu_i + \varepsilon_{ij} \tag{2}$$

Where:μ_i is the mean of the ith sample, ε_{ij} conforming to the normal distribution of variance σ^2. A sum of two variable is introduced in Equation (2–13) to indicate that that sample is from the first or second sample group:

$$\begin{aligned} X_{ij} &= x_{1ij}\mu_1 + x_{2ij}\mu_2 + \varepsilon_{ij}, x_{1ij} \\ &= \begin{cases} 1\,,\ i=1 \\ 0\,,\ i=2 \end{cases}, x_{1ij} = \begin{cases} 1\,,\ i=1 \\ 0\,,\ i=2 \end{cases} \end{aligned} \tag{3}$$

The matrix is in the form of:

$$X = Y\beta + \varepsilon, Y = \begin{bmatrix} 1\cdots1\ 0\cdots0 \\ 0\cdots0\ 1\cdots1 \end{bmatrix}^T \tag{4}$$

Where the parameter vector $\beta = (\mu_1, \mu_2)^T$. The zero hypothesis with equal sample mean is defined as follows:

$$H_0 : \mu_1 = \mu_2 \sim H_0 : c^T \beta = 0 \tag{5}$$

Substituted into matrix 2-15, the estimated values of the calculated parameters are:

$$\beta' = (Y^T Y)^{-1}(Y^T X) = (\overline{X_1}, \overline{X_2})^T \tag{6}$$

$$\sigma^{2'} = \frac{\varepsilon^T \varepsilon}{n_1 + n_2 - 2} \tag{7}$$

Where, respectively, β', $\sigma^{2'}$ are estimates of β and σ^2. Then the statistics are obtained and calculated as follows:

$$t = \frac{c^T \beta'}{\sqrt{\sigma^2 \left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} \tag{8}$$

A statistical model is established for each voxel, and the statistics and estimates are calculated separately. Then the threshold is designed according to the actual situation, and a comparison is made.

3.4. *Multi-feature fusion*

After obtaining the image texture features and VBM voxel features, we also need to fuse the two features. Fusion methods can be divided into many kinds, in which the traditional way is to cascade features into new features to increase the dimension of features. In practical applications, fusion methods and the optimization of features after fusion is of great significance.

3.4.1. *Principal component analysis*

This paper chooses PCA dimension reduction algorithm because the way of fusion feature is easy to lead to too high feature dimension of medical images, not only lead to too high feature dimension of computation, but also make the related description information redundant. The advantages of PCA algorithm just solve the drawbacks of the way of fusion feature introduced in this paper. Principal Component Analysis (PCA) is a classical dimension reduction algorithm, whose core idea is to map the original eigenvectors into the new eigenspace by orthogonal matrix computation. Then, a new feature combination is generated by merging in the new feature space. PCA is a classic and effective feature dimension reduction method, which not only reduces the dimension of feature, but also preserves the feature set with discrimination, eliminates the correlation between features, and improves the discrimination between features to a certain extent.

3.4.2. *Feature selection*

The PCA dimensionality reduction method effectively improves the computational speed, but the classification accuracy decreases. Feature fusion algorithm based on multi-kernel support vector machine can increase the accuracy of classification, but the computational efficiency is not ideal. Aiming at the shortcomings of the above two methods, the feature selection method is used to optimize the high-dimensional information generated by the fusion feature. Support vector machine recursive feature removal method is used to determine the relationship between the features, and the strong discriminant feature subset is extracted from the original feature set, and applied to the classification of medical brain images in this paper.

We use heuristic search method to construct the overall framework of feature selection, The specific implementation is as follows: First define a feature set, and set to blank, In the loop, the feature x with the highest score is selected from the scoring sequence of the recursive feature removal method of support vector machine into the feature set X, and then the K features with the highest correlation degree with the feature x are selected from the original feature set by using the correlation coefficient of the covariance matrix. These features do not take into account the individual discrimination ability. Then, according to the above rules, the corresponding features are selected from the original feature set and added into the feature set X, and the optimal evaluation index is calculated.

In order to make the feature set X as concise as possible and the feature discrimination ability in the set as strong as possible, in this paper, two groups of samples are selected randomly and tested by support vector machines with different kernel functions. When the classification accuracy of the two groups of samples is the highest, the feature selection is finished. The feature set is defined as feature set X. The calculation process is shown in the figure.

The feature selection process in this paper mainly involves two classification algorithms: The first classification algorithm uses support vector machine recursive feature removal method to calculate the discrimination ability evaluation criteria for the original feature set. Next, the feature is not selected according to the traditional support vector machine recursive feature removal method, and the feature is sorted only according to the score. The second classification algorithm is the traditional support vector machine algorithm. In this stage, the covariance matrix of principal component analysis is used to calculate the correlation between all features of the sorted feature set. If the redundancy of a pair of features is minimum, it is judged that they are related features, that is to say, the closer the relationship between the features described above is. Then, according to the heuristic search algorithm, the feature set is extracted from the score table based on the feature discrimination ability and the feature correlation information cycle, and a new feature subset is obtained until the optimal result is obtained.

4. Experimental design and analysis

4.1. Medical brain imaging database

The Medical Brain Imaging Database constructed in this paper is from the Public Database of Alzheimer's Disease Imaging Initiative (ADNI). 300 samples were selected as the Medical Brain Imaging Database in this public database. These samples were divided into 3 categories: the first 100 samples were patients with Alzheimer's disease; A sample of 100 patients in the second group showed mild cognitive impairment, which was not severe; The last 100 samples represent healthy elderly people. The samples in the Alzheimer's Disease Imaging Initiative (ADNI) public database are all T1-weighted MRI brain imaging data with flux 1.5. The medical brain imaging databases constructed in this paper have completed the necessary preprocessing, such as format conversion and background enhancement.

All experiments were carried out on a Core I7 computer with 64G memory, GTX1050Ti 4G stand-alone graphics card, Windows 7.0 system and MATLAB2012b programming environment, in which SPM8 and LibSVM expansion kits were loaded.

The specific information of the sample is shown in Table 1. The age of the samples in the three categories are all over 70 years old, and the number of samples in the three categories is set at 100. The ratio of male to female is basically balanced, and the educational level is basically similar. The value of 16 in the table indicates that most of the samples are with bachelor's degree.

Table 1
Sample specific information

Statistics	Alzheimer's disease (N = 100)	Mild disorder (N = 100)	Healthy elderly (N = 100)
Quantity (M/F)	68/32	55/45	52/48
Age (average)	74.56.8	75.54.8	74.85.4
Education (average)	16.52.1	16.12.6	16.32.3

The three major types of status were analyzed by Mini Mental State Inspection Scale and Clinical Dementia Rating Scale. The educational level in the scale was mainly based on the indicators given by the National Institute of Mental Illness and Stroke of the United States. The average educational level of 16 healthy elderly people was above 27, and the clinical dementia score was 0. There was no significant difference between the two groups. Patients with Alzheimer's usually score below 26 on the Mini-Mental State Inventory and 0.5 or 1.0 on the Clinical Dementia Rating Scale. Mild cognitive impairment is defined as mild cognitive impairment between Alzheimer's disease and health. Mild cognitive impairment can be classified into mild and severe cognitive impairment, but it is not strictly classified in this section, so long as the score of patients is between Alzheimer's disease and health, it is recognized as cognitive impairment.

4.2. Experiment flow and related parameter setting

Firstly, the REST component based on MATLAB tool is used for preprocessing, such as noise suppression, background enhancement and so on. Then, two kinds of features are extracted: texture feature (CNN feature extracted from medical brain image based on neural network structure) and morphological feature (morphological feature extracted by SPM8 component based on extracting element information). The two features are normalized in each feature extraction stage, and then they are fused in series. The feature set with high dimension after fusion is optimized. Firstly, linear kernel function is adopted for the feature set after fusion based on support vector machine recursive feature removal method, and the feature set with discriminant score can be calculated, and its feature set is sorted, and its penalty factor is defined as 100; Then the fusion feature set is divided into two kinds of feature subsets by heuristic search method, in which the extraction rule parameter K of heuristic search algorithm needs to be set, and K represents the feature with strong discrimination after recursive feature removal algorithm based on support vector machine, and the larger the number of features related to K, the more relevant features will be obtained. In this experiment, we divide the fusion feature into two kinds of feature subsets, and determine the best K value in the subset. Then, we use the best K value as input to generate the best feature subset in the heuristic search algorithm.

4.3. Evaluation criteria

In the medical brain image processing phase, the classification accuracy, sensitivity, specificity, positive and negative predictive value were used as the evaluation index.

Accuracy refers to whether the average value is consistent with the true value in a certain group of features, indicating the correct proportion of classification results. Generally speaking, the higher the classification accuracy is, the better the classification performance of the algorithm will be. The calculation is as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

Where TP, TN, FP and FN represent true positive, true negative, false positive and false negative, respectively.

Sensitivity refers to the ratio of the number of positive samples to the actual number of patient samples. Generally, the higher the sensitivity, the stronger the ability to detect patients, i.e., the detection rate of patient samples, is calculated as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{10}$$

Specificity versus sensitivity is the ratio of the number of negative samples to the actual number of patient samples. Usually, the higher the specificity, the stronger the ability to detect non-patients, i.e., the detection rate of patient samples, is calculated as follows:

$$Sensitivity = \frac{TN}{TN + FP} \tag{11}$$

Positive predictive value refers to the proportion of true positive samples in the samples whose test results are set as positive. It is mainly used to judge whether there is misjudgment or not, and other misjudgment rates are inversely proportional. The higher the value is, the lower the misjudgment rate is. Calculate as follows:

$$Positive\ predictive\ value = \frac{TP}{TP + FP} \tag{12}$$

Negative predictive value refers to the proportion of true positive samples in the samples whose test results are set as negative. It is mainly used to judge whether there is omission or not, which is inversely proportional to the omission rate. The higher the value, the lower the omission rate. Calculate as follows:

$$Negative\ predictive\ value = \frac{TN}{TN + FN} \tag{13}$$

Below we introduce the classification performance of medical brain images using the above five indicators for testing.

4.4. Optimal subset of fusion features

In this section, we mainly analyze the influence of the extraction rule parameter K on the classification results of medical brain images in the optimization stage (feature selection) of fusion features. Secondly, the influence of the feature subset after the introduction of correlation is analyzed. The K value represents the score table extracted according to the recursive feature removal method of support vector machine. After extracting the feature with the highest score from the table, the number of related features is extracted. In the above theoretical section, the significance of the K value has been explained, and in the experiment section, when $K = 0$, the experiment will not proceed to the feature screening stage, and the feature with the highest score will be extracted directly from the ranking table of the next process, that is, the correlation between the features will not be taken into account; When $K = 1$, the feature with the highest score is selected in each cycle, and then the feature with the lowest redundancy with the highest score is selected from the other feature sets. The larger the value of K, the stronger the correlation between the feature subsets and K in each cycle. The classification task of the experiment part is to distinguish the three states of Alzheimer's disease cognitive impairment and normal elderly people in the sample set. In each class, it is divided into two groups of samples, the ratio of which is 1:2, and the support vector machine method based on linear kernel and the support vector machine method based on radial basis function kernel are used to classify the samples. Each group was divided into two phases. In the first stage, according to the values of the extraction rule parameter K, which are 0, 1, 2, 3, 4, 5 respectively, the test is carried out in two sample groups after subdivision to test the influence of different types of feature sets on classification accuracy and the optimal value of K; In the second stage, after determining the optimal value of K, feature subsets can be screened, and then the classification performance of feature subsets under different cycles can be calculated for two groups of samples respectively to obtain the optimal feature subsets.

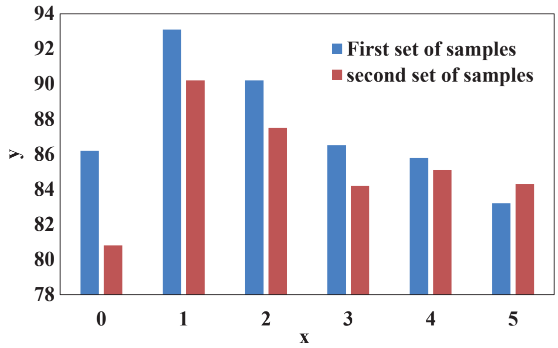

Fig. 3. Effect of K in Alzheimer's disease sample and healthy elderly sample experiment.

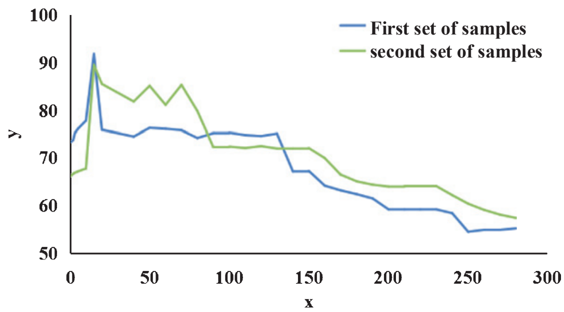

Fig. 4. Optimal feature subsets in Alzheimer's disease sample and healthy elderly sample experiment.

4.4.1. *Comparison of Alzheimer's disease and healthy elderly samples*

At first, that sample with Alzheimer's disease and the healthy elderly were classify. Figure 3 shows the classification result when six values of K were obtained. The first set of samples represents classification algorithms that take into account the interconnections between features, and the second set does not address the interconnections be tween features, but rather classifies features based on their individual discrimination capabilities.

When $K = 0$, the relationship between features is not considered, and the classification result based on recursive feature removal method is not ideal, and the difference between the two groups of classification results is large. When $K = 1$, i.e. Extracting one optimal feature from the fusion feature and selecting one related feature, the classification accuracy of the two groups of samples is the highest, and the optimal solution is obtained at this time. As K increases, the classification accuracy decreases gradually, and it should be noted here that the classification accuracy of the first set of samples is less than that of the second set when $K = 5$. It can be seen that feature correlation is helpful to improve classification accuracy to some extent.However, if feature correlation is introduced blindly, the impact of feature recognition ability on classification task will be impacted and classification accuracy will be affected easily.

According to the above experiments, in the classification of Alzheimer's disease samples and healthy elderly people, the best value of parameter K based on support vector machine recursive feature removal method is 1. Therefore, when $K = 1$, the feature subset is selected for the original feature set in a loop. Figure 4 indicates that the classification accuracy varies with the number of feature subsets.

In the initial state, the classification accuracy first increases with the number of feature subsets, When the number of feature subsets is 15, the classification accuracy of the two groups of samples is the highest. When the number of feature subsets is greater than 15, the classification accuracy begins to decrease. This is because too many redundant features will disturb the classifier to classify the feature sets. The experiments show that the algorithm can get better classification results by using smaller dimension features on the premise of larger dimension, which can not only improve the accuracy of classification, but also reduce the number of features involved in the calculation, and improve the computation speed.

4.4.2. *Comparison of mild cognitive impairment and healthy elderly samples*

As shown in Fig. 5, when $K = 5$, the classification accuracy of the two groups of samples reached the peak value and was above 90%, but the trend of the experimental results was different from that of the previous group of experiments. The values of parameter K were different in different databases, and the comparison of the classification results was also different.

When K is 1, the result of feature subset screening is shown in Fig. 6. When the dimension of feature set

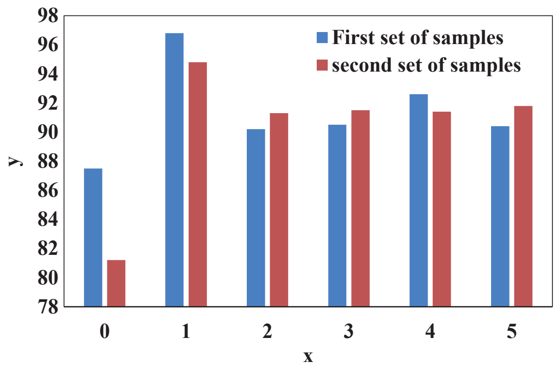

Fig. 5. Comparison of mild cognitive impairment and healthy elderly samples.

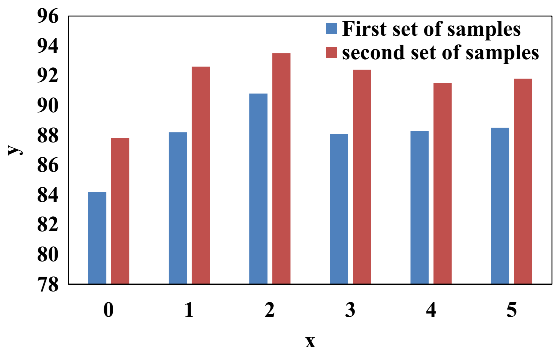

Fig. 7. Comparison of mild cognitive impairment and Alzheimer's disease samples.

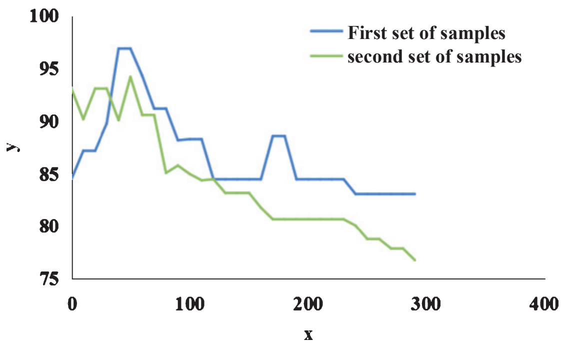

Fig. 6. Optimal feature subsets in Alzheimer's disease sample and healthy elderly sample experiment.

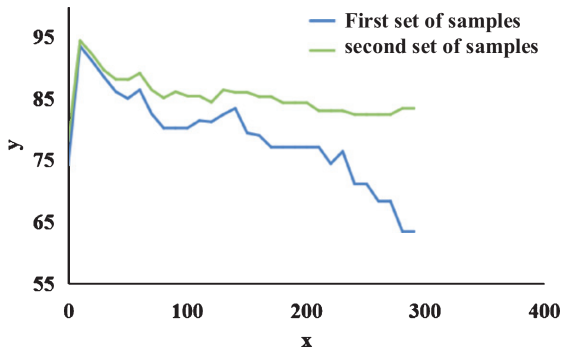

Fig. 8. Optimal feature subsets in the mild cognitive impairment and Alzheimer's disease sample experiment.

is 40, the classification results of the two classes are the best.

4.4.3. Comparison of mild cognitive impairment and Alzheimer's disease samples

This part is based on the mild cognitive impairment and Alzheimer's disease samples, the calculation process is the same as the previous two groups of experiments, but the results are very different. As shown in Fig. 7, when parameter $K = 0$, the classification accuracy is the lowest, and increases with the increase of parameter K. When parameter $K = 2$, the classification accuracy is the highest, but not as the first two groups of experiments, when $K = 1$, the classification accuracy is the highest. Up to $K = 3$, the accuracy of classification began to decrease.

In the task of comparing mild cognitive impairment with Alzheimer's disease samples, set the parameter $K = 2$, and test the dimension of the feature subset as shown in Fig. 8. The dimension of the feature subset is 10, and the optimal result can be obtained.

4.5. Experimental results and analysis

From the above experiments, we can see that the medical brain image retrieval method based on multi-feature fusion proposed in this paper is mainly improved in two stages. The first stage is to fusion multi-feature descriptors, and then extract the CNN of convolution neural network structure as the description of texture information, and then fuse the body shape features. In the second phase, support vector machine recursive feature removal method is used to improve the structure of this paper. Below, we compare the classification performance of these improved schemes. In this section, we compare the classification performance of the following cases: using a single texture feature descriptor, using a single morphological descriptor, and using the descriptor of fusion feature proposed in this paper, It should be noted that the descriptor classification of fused features still needs to pay attention to the feature selection stage in the classification stage. For

Table 2
Comparison of Alzheimer's Disease and Healthy Elderly Classification Performances

Classification algorithm	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
Texture feature	78.88	73.12	80.12	79.85	84.89
Morphological characteristics	80.32	72.98	83.95	82.34	84.53
Fusion Feature+PCA	86.24	78.65	91.91	90.23	80.55
Fusion Feature+Multi-kernel Support Vector Machine	89.42	86.16	92.32	91.24	84.32
Algorithm in this paper	93.9	88.6	98.56	97.56	88.65

Table 3
Comparison of Mild Cognitive Impairment and Healthy Elderly Classification Performances

Classification algorithm	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
Texture feature	82.11	76.32	87.28	86.12	79.12
Morphological characteristics	64.77	56.11	66.21	67.02	62.91
Fusion Feature+PCA	85.65	85.06	86.23	89.88	80.88
Fusion Feature+Multi-kernel Support Vector Machine	92.02	90.50	93.23	94.88	86.34
Algorithm in this paper	97.32	94.96	100	100	92.88

Table 4
Comparison of Mild Cognitive Impairment and Alzheimer's Disease Classification Performances

Classification algorithm	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
Texture feature	71.64	92.28	58.78	63.51	90.01
Morphological characteristics	68.84	63.45	71.78	70.21	65.56
Fusion Feature+PCA	72.52	80.21	64.28	67.45	78.56
Fusion Feature+Multi-kernel Support Vector Machine	78.21	86.24	71.11	73.02	85.22
Algorithm in this paper	92.15	100	82.15	83.65	100

example, in this stage, we need to use PCA dimensionality reduction as the feature selection stage, and multi-kernel support vector machines as the feature selection stage, which are compared with the algorithms proposed in my chapter respectively. The results of each group are as follows.

4.5.1. Comparison of Alzheimer's disease and healthy elderly samples

The accuracy of the proposed algorithm is 93.9%, the sensitivity is 88.6%, the specificity is 98.56%, the positive predictive value is 97.56%, the negative predictive value is 88.65%, and all the indexes are the highest.

4.5.2. Comparison of mild cognitive impairment and healthy elderly Samples

The classification performance of the proposed algorithm is shown in Table 3. The accuracy, sensitivity, specificity, positive predictive value and negative predictive value are 97.32%, 94.96%, 100%, 100% and 92.88% respectively, for the mild cognitive impairment and healthy elderly samples.

4.5.3. Mild cognitive impairment compared with Alzheimer's disease samples

Compared with the mild cognitive impairment and Alzheimer's disease samples, the classification performance is shown in Table 4. The accuracy, sensitivity, specificity, positive predictive value and negative predictive value of the proposed algorithm are 92.15%, 100%, 82.15%, 83.65% and 100% respectively. As a result of that above three group, In the classification of multi-feature fusion brain images, PCA dimensionality reduction algorithm mainly improves the efficiency of computer, but the accuracy of classification decreases, the final classification accuracy of multi-kernel support vector machine improves a little when optimizing the weights, the algorithm proposed in this section has significant improvement in computational efficiency and accuracy.

5. Conclusion

In this paper, a multi-feature fusion algorithm for medical brain images is proposed. CNN features are extracted as texture information by convolution neural network structure, and voxel-based features are extracted as morphological features, and the two types of features are fused in series. However, the feature dimension after fusion is often high and the computation is complex, and the redundant information of the feature after fusion will disturb the classification accuracy. For this reason, this paper utilizes the feature selection scheme to optimize the fusion features: A support vector machine-based recursive feature removal method incorporate a covariance matrix to calculate that correlation between features, And according to the heuristic search optimization feature selection stage, the feature correlation is added to the feature score table extracted based on the recursive feature removal method of support vector machine, and the optimal value of the parameter K is selected according to the heuristic search through the experimental analysis, and the parameter K value is determined, then the optimal feature subset is extracted.

Finally, through the experimental design and analysis, we can see that the proposed algorithm can reduce the dimension of fusion features, and remove the interference of redundant features, to some extent, improve the classification performance. In the task of medical brain image classification, the classification performance of Alzheimer's disease is greatly improved and the computational complexity is greatly reduced.

Acknowledgments

This project is supported by The Education Department of Jilin Province, "13th Five-Year" science and technology project "based on behavior recognition method adaptive multiple features fusion research" (Project No. JJKH20170037kj); The National Natural Science Foundation of China Youth Project (61101155); Jilin provincial development and Reform Commission industrial innovation special project (2016C035); The Key Laboratory of symbolic computation and Knowledge Engineering Ministry of education of Jilin University in 2017.

References

[1] J.E. Bibault, P. Giraud and A. Burgun, Big Data and machine learning in radiation oncology: State of the art and future prospects[J], *Cancer Letters* **382** (2016), 110–117.
[2] W. Yuanyuan, Z. Tao and W. Cuiying, Research on PET/CT multimodal image recognition based on convolutional neural network [J], *Television Technology* **41** (2017), 88–94.
[3] Q. Zhiguang, C. Hao, D. Yi, et al., Study on the method of cerebral vascular extraction based on multimodal convolution neural network [J], *Journal of University of Electronic Science and Technology* **45** (2016), 573–581.
[4] L. Jian, L. Man, L. Xiao, Y. Jiayin, Z. Wenming, L. Yang, et al., Study on brain tumor segmentation based on multi-scale convolution neural network [J], *China Medical Equipment* **2** (2016), 25–28.
[5] Y. Jiayin, Z. Wenming, L. Yang, et al., Multi-modal emotion recognition based on deep neural network [J]. Journal of Southeast University (English edition) **33** (2017), 444–447.
[6] S. Shukla and Dr. Neeraj Shukla, Smart Waste Collection System based on IoT (Internet of Things), *International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)* **5**(9) 2017.
[7] W. Xinjie, L. Hongyu and L. Nannan, *Application of convolutional neural network in ECT image reconstruction [J]. Journal of Liaoning University (Natural Science Edition)* **1** (2018).
[8] W. Mengfei, L. Weirong and C. Hu, Heterogeneous Multimodal Target Recognition Strategy for Large Data Stream in Network Media [J], *Computer Research and Development* **54** (2017) 71–79.
[9] A. Paul, A. Ahmad, M.M. Rathore and S. Jabbar, Smartbuddy: Defining human behaviors using big data analytics in social internet of things, *IEEE Wireless communications* **23** (5) (2016), 68–74.
[10] P. Sujatha and N. Hyrunnisha Mca., M.Phil, Leach- Low Energy Adaptive Over Data Communication in WSN, *International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)* **5**(10), 2017.
[11] N. Sharma, N. Kumar Upadhyay and N. Kumar Saraswat, An Adaptive Study for Different Methodology for Information Hiding, *International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)* **5**(8) 2017.
[12] Z Jinhan, Z. Yongdong, et al., Automatic image quality analysis based on convolutional neural network CT/CBCT [J], *Chinese Journal of Medical Physics* **5** (2018).
[13] Z. Jun, C. Venkatesan, P. Karthigaikumar, A. Paul, S. Satheeskumaran and R. Kumar, ECG signal preprocessing and SVM classifier-based abnormality detection in remote healthcare applications, *IEEE Access* **6** (2018), 9767–9773.
[14] L. Lei, Online visual target tracking method based on improved convolution neural network [J], *Inner Mongolia Normal University Journal (Natural Chinese Edition)* **6** (2017), 878–883.