IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON NEW TRENDS IN BRAIN SIGNAL PROCESSING AND ANALYSIS

Received January 19, 2019, accepted January 30, 2019, date of publication February 4, 2019, date of current version March 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897131

Research on Feature Extraction of Tumor Image Based on Convolutional Neural Network

AIMIN YANG[1], XIAOLEI YANG[1], WENRUI WU[2], HUIXIANG LIU[1], AND YUNXI ZHUANSUN[1]

[1]The Key Laboratory of Engineering Calculation in Tangshan City, North China University of Science and Technology, Tangshan 063000, China
[2]School of Mining Engineering, North China University of Science and Technology, Tangshan 063000, China

Corresponding author: Wenrui Wu (huali_wuwenrui@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51674121, in part by the Natural Science Foundation of Hebei under Grant E2017209178, and in part by the Outstanding Youth Fund Project of North China University of Science and Technology under Grant JQ201705.

ABSTRACT Medical images play a very important role in making the right diagnosis for the doctor and in the patient's treatment process. Using intelligent algorithms makes it possible to quickly distinguish the lesions of medical images, and it is especially important to extract features from images. Many studies have integrated various algorithms into medical images. For medical image feature extraction, a large amount of data is analyzed to obtain processing results, helping doctors to make more accurate case diagnosis. In view of this, this paper takes tumor images as the research object, and first performs local binary pattern feature extraction of the tumor image by rotation invariance. As the image shifts and the rotation changes, the image is stationary relative to the coordinate system. The method can accurately describe the texture features of the shallow layer of the tumor image, thereby enhancing the robustness of the image region description. Focusing on image feature extraction based on convolutional neural network (CNN), the basic framework of CNN is built. In order to break the limitations of machine vision and human vision, the research is extended to multi-channel input CNN for image feature extraction. Two convolution models of Xception and Dense Net are built to improve the accuracy of the CNN algorithm. It can be seen from the experimental results that the CNN algorithm shows high accuracy in tumor image feature extraction. In this paper, the CNN algorithm is compared with several classical algorithms in the local binary mode. The CNN algorithm has more accurate feature extraction ability for tumor CT images on a larger data basis. Furthermore, the advantages of CNN algorithms in this field are demonstrated.

INDEX TERMS Convolutional neural network, image feature extraction, local binary mode.

I. INTRODUCTION

Medical image is the main carrier and important carrier of modern medical alternative language description. It gradually plays an increasingly important role in the medical field with very intuitive expression and objective description. Medical image analysis is a modern image analysis technology that integrates mathematical modeling, artificial intelligence, integrated medical imaging, digital image processing and other multidisciplinary.

Defining whether a tumor is benign or malignant according to the image is also a subject that the medical community has not stopped. The feature extraction of the image is to first segment the image, and then extract the texture features in the image after segmentation. There are a lot of textures on trees, cloths, and clothes. These things that can be seen visually are taken from their textures. Of course, medically, for tumors, it is not as intuitive as trees or clothes. It can only be sliced into tumors, relying on imaging equipment to make its texture features into medical images such as CT or MRI. Texture is a basic and very useful information feature in an image. It is an important parameter for describing image content, and its academic research has gradually become a new topic.

In the concept of typing, the texture in the image has rotation, translation, and scale invariance, which means that regardless of the state of the image, the position, orientation and size of the texture are constant. For texture feature extraction, the commonly used methods are statistical analysis, model analysis, structural analysis and spectrum analysis. The first successful application of image feature extraction texture features was in 1972, when Sutton and Hall identified normal and abnormal lungs [1]. Taleb Ahmed uses the theory

The associate editor coordinating the review of this manuscript and approving it for publication was Victor Hugo Albuquerque.

of typing to extract the CT image of the bone, and obtains the image difference between the osteoporosis patient and the normal human bone. Further, the osteoporosis patient can be discriminated according to the image [2]. Ganeshan *et al.* analyzed the MRI image of the liver and achieved the detection of patients with cirrhosis [3]. In 2014, Sun *et al.* [4] developed a face recognition system through the feature extraction of the figure, with an error rate of only 3.55%, reaching the world's top effect. Similarly, Chen *et al.* [5] also studied face recognition.

For the local binary mode and the convolutional neural network to solve the problem of image feature extraction, Liu *et al.* established a three-level scale space pyramid, and used the canny detector to detect the edge and tested on 326 images, achieving an accuracy of 0.93. Good results [6]. Sidibé *et al.* [7] solved the image problem based on the semi-supervised learning Gaussian mixture model, reaching 93% and 80% of SE and SP. Lee *et al.* [8] first trained the CNN model and then performed image segmentation, but because of the small amount of data samples, the final segmentation accuracy and generalization ability of the model were limited. Kermany *et al.* [9] used the CNN model to classify and identify bacterial pneumonia and viral pneumonia with a final resolution of 96.6%. However, there are still some shortcomings in the process of processing images: the source of the training image samples is different, which may cause the image scale to change too much. The mark made during the image cutting process will affect the image feature extraction, and then the data sample size is thin. Not enough to train the CNN model.

This paper proposes to use the local binary mode and convolutional neural network based on image translation rotation invariance to cut and classify the image, and then extract the feature to achieve the purpose of classification. In this study, a large number of labeled tumor CT images are used to train the deep convolutional neural network model, and then the trained model is used to extract the texture features of medical images, and the texture features are identified and classified according to these texture features. The lesion image was pre-screened and later assisted.

Convolutional neural network (CNN) is a relatively advanced image processing algorithm. In addition to recognizing and classifying medical images, it can extract features of photos in surveillance videos in intelligent logistics, intelligent transportation and other aspects, so as to achieve the purpose of intelligent sorting express and distinguishing road violations.

II. LOCAL BINARY PATTERN FEATURE EXTRACTION BASED ON IMAGE ROTATION INVARIANCE

A. LOCAL BINARY PATTERN FEATURE

Local binary pattern is the most commonly used method for image recognition. Rotation invariance is an improved local binary model which can accurately describe the shallow texture features of medical images.

The local binary pattern feature extraction is to compare the gradation difference values of the pixel points in the center portion, the quarter portion and the edge portion of the image to represent the local texture feature information of each portion. If the image features of the central part of the study are studied, then the quarter and the edge are the neighborhoods of the central part. The local gray matter of the central pixel and the neighborhood can be used to obtain the local texture features of the central pixel, which is:

$$LBP_{N,R} = \sum_{i=0}^{N-1} 2^i s(g_i - g_c);\ s(x) = \begin{cases} 1 & |x| \geq T_{LBP} \\ 0 & |x| < T_{LBP} \end{cases} \tag{2.1}$$

where: the gray value of the neighborhood pixel represented as the center; is the neighborhood radius, and the center point is the gray value of the first pixel in the center of the circle, which is a small positive value to enhance the robustness of the image region description.

B. COMPLETE LOCAL BINARY PATTERN FEATURE

The complete local binary mode describes the texture features of the pixel from three aspects: the gray value relationship feature, the gray value difference amplitude feature, and the pixel value and the average value of all pixel gray values, which is:

$$CLBP\_S_{N,R} = \sum_{i=0}^{N-1} 2^i s(g_i - g_c);\ s(x) = \begin{cases} 1 & |x| \geq T_{LBP} \\ 0 & |x| < T_{LBP} \end{cases} \tag{2.2}$$

$$CLBP\_M_{N,R} = \sum_{i=0}^{N-1} 2^i t(m_N - c);\ t(m_N - c) = \begin{cases} 1 & m_N \geq c \\ 0 & m_N < c \end{cases} \tag{2.3}$$

$$CLBP\_C_{N,R} = t(g_c, c_l);\ t(x, c) = \begin{cases} 1 & x \geq c \\ 0 & x < c \end{cases} \tag{2.4}$$

where: indicates the relationship between the magnitude of the gray value, the amplitude value of the gray value difference, and the average value relationship, which represents the central pixel point difference, which represents the mean value in the partial image, and represents the global gray mean value.

C. LOCAL BINARY PATTERN FEATURE EXTRACTION BASED ON IMAGE ROTATION INVARIANCE

According to the above-mentioned local binary pattern [10]–[12] sequence, the coordinate system is constructed. When the image changes in translation, rotation, etc., the coordinate system also changes. Compared with the coordinate system, the image is relatively static, so there will be an image. It is rotation invariance.

After the above process, a binary value string is obtained, which represents the value of the first digit on the binary value string ($i = 1, 2, \cdots, 8$), and then converted to decimal as the LBP value of the texture unit.

$$B_i = \begin{cases} 1 & c_l \geq m_N \\ 0 & c_l < m_N \end{cases} \tag{2.5}$$

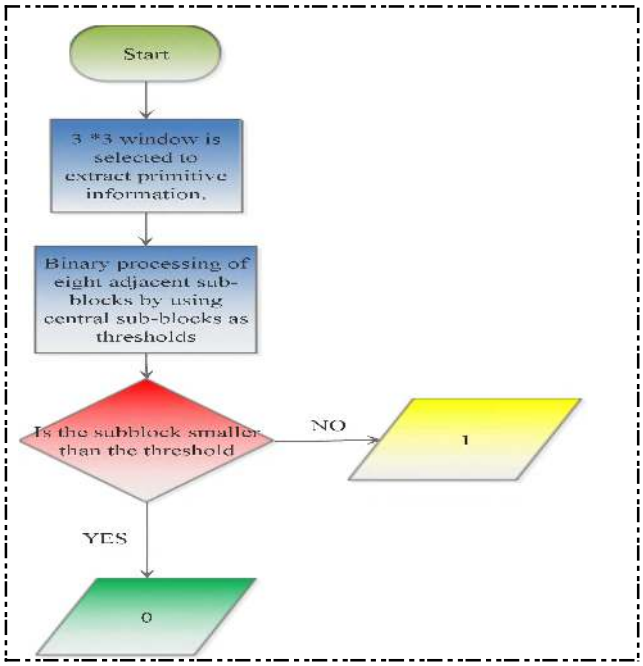


FIGURE 1. Flow chart of LBP algorithm based on image rotation invariance.

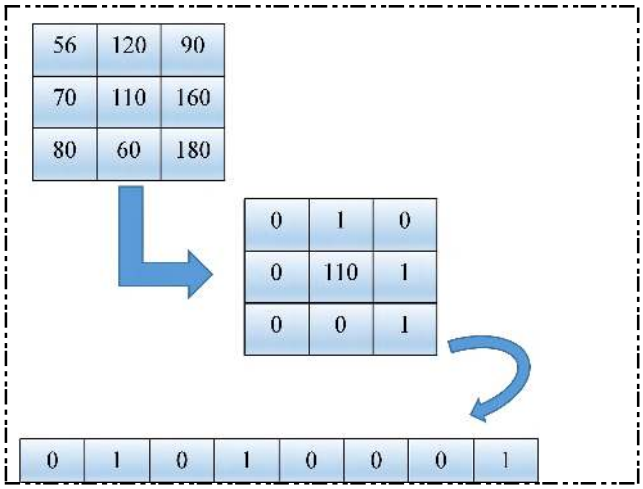


FIGURE 2. LBP algorithm analysis.

Figure 2 fully describes the texture features of the image and also preserves the shape and orientation of the texture.

III. CONSTRUCTING CNN IMAGE FEATURE EXTRACTION MODEL FOR MULTI-CHANNEL INPUT

A. THE BASIC PRINCIPLE OF CONVOLUTIONAL NEURAL NETWORKS

The convolutional neural network [13], [14] is a feed forward neural network that can perform image processing. The convolutional neural network consists of a convolutional layer, a pooled layer, and a fully connected layer. Convolutional neural networks are MLPs (multilayer perceptron's) inspired by biological thinking. They have different categories of levels, and each layer works differently and in different ways.

In the convolutional layer, there are usually several feature planes, each of which consists of a number of rectangular arranged neurons. The neurons on the same feature plane share weights and the shared weights are called convolution kernels. Weight sharing can reduce the number of connections between neurons and reduce the risk of over fitting. Down-sampling, also known as pooling, can be seen as a special convolution process that reduces computational dimensions and improves model generalization. Convolution and down sampling greatly simplifies model complexity and reduces model parameters. Convolutional neural network training process mainly includes forward propagation and backward propagation. During the training process, the convolutional layer and the down sampling layer alternately appear. Convolutional neural network multi-layer feature fusion structure is shown in Figure 3.

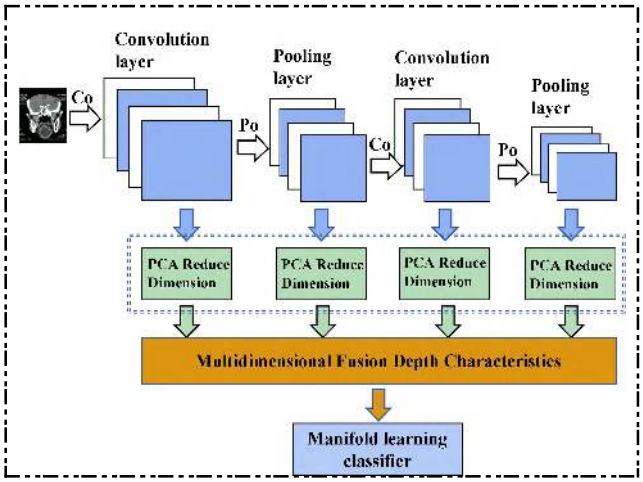


FIGURE 3. Convolutional neural network basic architecture.

B. IMAGE FEATURE EXTRACTION BASED ON MULTI-CHANNEL INPUT CONVOLUTIONAL NEURAL NETWORK

The traditional CNN directly inputs the picture as data, which theoretically conforms to the viewpoint that "pixel itself is the most redundant representation of image speech" in sparse representation theory [15], but machine vision has certain limitations compared with human vision. Humans can do a lot of pattern classification work at a time when looking at things with the naked eye. However, when we do research, we only study one or several specific pattern classification problems at a time. In order to avoid this limitation, when CNN is used for saliency detection, the image is first super pixel-divided, and then the segmented superpixel is input as a new network, and three channels are simultaneously input[16]–[18]. It shows Figure 4:

The segmented super pixel is input through three channels and then forwarded for feature extraction, mainly by convolution operation and down sampling operation. The image enters the convolution layer from the input layer and the

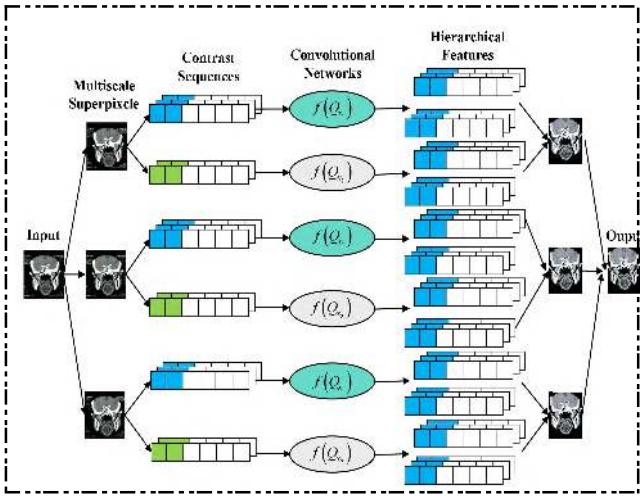


FIGURE 4. Convolutional neural network multi-channel input.

output value is obtained by the activation function.

$$x^l = f\left(W^l x^l + b^l\right) \tag{3.1}$$

Among them, the l representative number of layers, ω representing the weight, b representing the offset, f represents the activation function.

In the forward propagation process, several feature maps of the previous layer are convolved by a learnable convolution kernel to obtain a new feature map using the activation function.

$$x_j^l = f\left(\sum_{i \in M_j} X_i^{l-1} * k_{ij}^l + b_j^l\right) \tag{3.2}$$

where down represents is the down sampling function.

The current layer, l representing the previous layer, $l-1$ represents the first feature map of the current layer, and X_j^l represents the j convolution kernel corresponding to the first feature map of the current layer and the first feature map of the j previous layer, b_j^l which is an offset value.

After the down sampling layer is next to the convolutional layer, the relative positional changes of the target tilt and rotation can be neglected, the performance and robustness of the algorithm are improved, and the feature map dimensions are reduced, and over-fitting is avoided to some extent. The down sampling layer is calculated as follows:

$$x_j^l = f\left(\beta_j^l down\left(x_j^{l-1}\right) + b_j^l\right) \tag{3.3}$$

where down represents is the down sampling function.

The purpose of back propagation is to update the convolution kernel weight value, that is, to calculate the error value of the forward propagation output value and the predetermined sample label, and then obtain a gradient by convolution, pooling, etc. to update the convolution kernel. The error operation loss function generally adopts a square difference loss function. For multi-classification problems with categories and number of samples, the squared difference loss function table is not as follows:

$$E^N = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{c}\left(t_k^n - y_k^n\right)^2 \tag{3.4}$$

where E^N represents the total error of N samples, ct_k^n represents the n-dimensional label corresponding to the k sample, and f represents the n dimension output corresponding to the k sample.

After obtaining the image features through the convolutional layer, the following is the pooling layer, and the features extracted by the convolution layer are used for classification. On the one hand, the feature map is reduced, and the network computation complexity is simplified; on the other hand, feature compression and extraction are performed. The main feature is to reduce the feature dimension. That shows Figure 5:

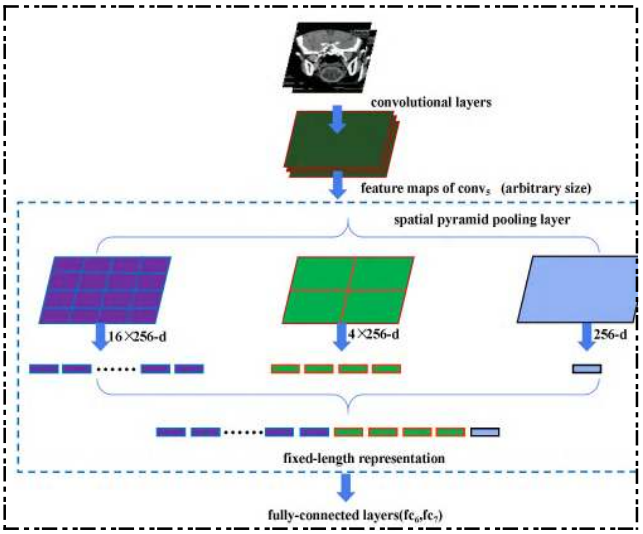


FIGURE 5. Empty pyramid pooling layer.

After the pooling layer reduces the image features and then transmits them to the full layer, the fully connected layer plays the role of "classifier" in the whole convolutional neural network, that is, after deep network through convolution, activation function, pooling, etc. Then, the results are identified and classified through the fully connected layer. First, the results of the convolution, activation function, and pooled deep network are strung together. The structure is shown in Figure 6.

C. TWO EXPERIMENTAL CONVOLUTION MODELS BASED ON MULTI-CHANNEL INPUT

1) XCEPTION MODEL

The Xception model is mainly based on the fact that the cross-channel correlation and spatial correlation process of the convolutional neural network processing feature map can be separated, and then the Inception module is replaced by the deep volume convolution module, that is, the input channel

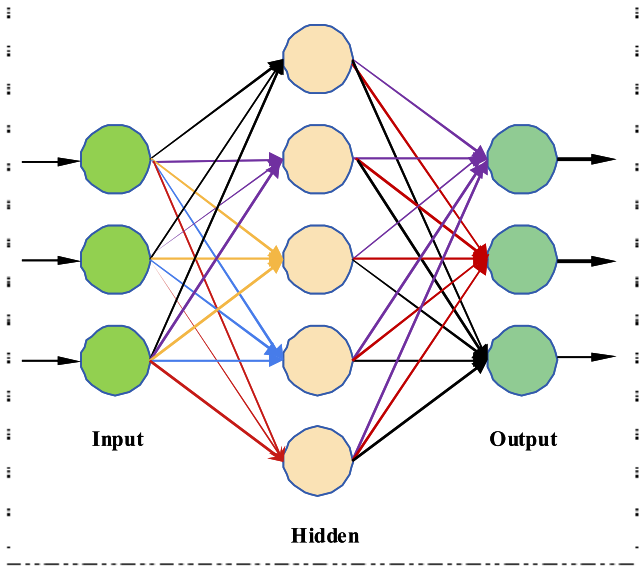


FIGURE 6. Full connection layer structure.

is independent for each input channel. Spatial convolution is performed, then a 1 × 1 size filter is used for point-by-point convolution, and finally the channel output is mapped to the new channel space as shown in Figure 7. The Xception model achieves better recognition than Inception with the same amount of parameters as the Inception model, which proves that this performance improvement is due to the improvement of the model.

As can be seen from Figure 7, multiple feature maps are output during the whole process, and each graph can be combined to convolve multiple values:

$$x_j^l = f(u_j^l) \tag{3.5}$$

$$u_j^l = \sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l \tag{3.6}$$

where x_i^l is the output of the l convolution channel of the j layer convolutional layer, and u_j^l is the net activation of the l convolutional channel of the j layer convolutional layer, which is convolved and offset by the previous G x_i^{l-1} owned.

Then use the above two-dimensional feature map as the input of the connection layer:

$$x^l = f(u^l) \tag{3.7}$$

$$u^l = w^l x^{l-1} + b^l \tag{3.8}$$

where w^l is the weight coefficient of the fully connected layer of layer l, and b^l is the threshold offset term.

Next, the gradient descent method is used to adjust the direction of the training error reduction:

$$\delta^l = \frac{\partial E}{\partial u^l} \tag{3.9}$$

where δ^l represents the function of the squared error with a change in u^l, and E represents the square of the difference between the desired output and the actual output.

2) DENSE NET MODEL

The main feature of the Dense Net model [19], [20] is that a densely concatenated convolutional neural network is proposed novelty, that is, the input of the current convolutional layer comes from the output of all previous convolutional layers. This mode makes convolution network trains deeper, more accurately, and more efficiently. As shown in Figure 8, this excellent structural design of the Dense Net model enhances feature reuse, which better solves the gradient dispersion problem and greatly reduces the number of network parameters.

From Figure 8, we can see that the Dense Net network structure model constructed based on CNN algorithm has several advantages, which reduces vanishing-gradient, enhances the feature transfer, and more effectively utilizes the feature. To a lesser extent, the parameters are less. Quantity, in the network of intelligent algorithm learning, as the depth of the network deepens, the problem of gradient disappearance becomes more and more obvious. At present, many researches have proposed solutions to such problems. This study constructs the core of the network structure to ensure the middle layer of the network. Directly connect all layers directly under the premise of maximum information transfer between layers.

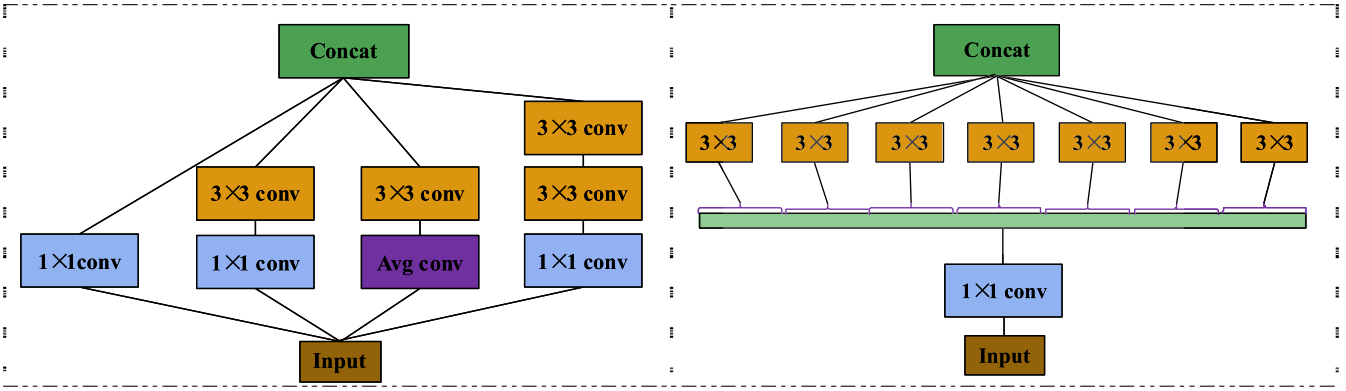


FIGURE 7. Inception module (left) and Xception module (right).

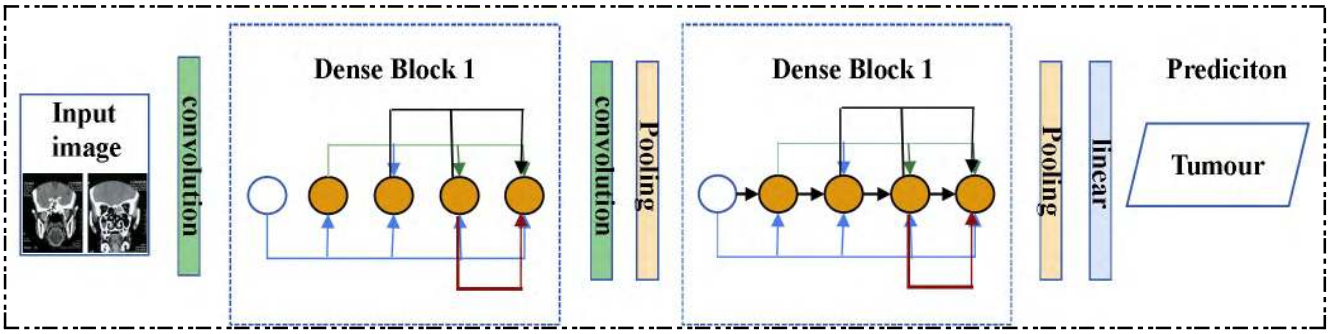


FIGURE 8. Dense Net network structure diagram.

One of the main points of the Dense Net network architecture is that the network is narrower and has fewer parameters.

$$X_l = H_l(X_{l-1}) + X_{l-1} \tag{3.10}$$

Here l represents the layer, X_l represents the output of the l layer, and H_l represents a nonlinear transformation. So for ResNet, the output of the l layer is the output of the $l-1$ layer plus the nonlinear transformation of the output of the $l-1$ layer.

$$X_l = H_l([X_0, X_1, \ldots, X_{l-1}]) \tag{3.11}$$

The above formula is DenseNet.

$[X, X_1, \ldots, X_{l-1}]$ indicates that the output feature map of the 0 to $l-1$ layer is concatenation. Concatenation is the merging of channels, just like Inception. The former resnet is the sum of the values, and the number of channels is constant. H_l includes convolutions of BN.

Based on the Dense Net model built by the convolutional neural network algorithm, Dense Net improves the back propagation of the gradient, which makes the network easier to train, the focus is on dense connections, and finally each layer can directly reach the error signal. An implicit channel is implemented. The model parameters under this algorithm are smaller and computationally efficient, and feature reuse is achieved, so that the feature map of each layer is relatively small.

IV. GUIDELINES FOR GRAPHICS PREPARATION AND SUBMISSION

The image formation process is like dividing a selected layer into a plurality of cuboids of the same volume, called voxels. The scanned information is calculated to obtain the X-ray attenuation coefficient or absorption coefficient of each voxel, and then arranged into a matrix, that is, a digital matrix. The digital matrix can be stored in a magnetic disk or an optical disk. By digital/analog converter, each digit in the digital matrix is converted into small squares of unequal gray scales from black to white, that is, pixels, and arranged in a matrix to form a CT image [21].

The data samples used in this paper are all from the network CANCER-CAPTAC-GBM database. 500 ultrasound images of 560 × 360 pixels were selected as experimental samples, of which 417 were malignant cases and 83 were benign cases, the ratio of the two training samples to the test samples was 7:3.

A. DATA PREPROCESSING

The preprocessing of ultrasound images [22]–[26] mainly includes the scale registration of images and the removal of two parts by artificial markers in the image.

Image scale registration: The texture detail scales of different image images are different. In order to obtain images of the training set with the same distance scale, the target images will be scaled. First, the threshold image method is used to obtain the binary image of the suspected scale position in the image [27]–[29], see equation (2.5). Again, calculate the autocorrelation function for all the columns in B. Find the position of the extreme points that meet the conditions in the correlation.

$$A_j = AR(B(:, j)), i \in [1, n] \tag{4.1}$$

$$S_j = A_j(i) > \alpha^* \max(A_j) \tag{4.2}$$

where $B(:, j)$ represents the j column of the obtained binary image, AR is an autocorrelation function, Aj is an autocorrelation coefficient, and $\max(A_j)$ is the maximum value of A_j. The position sequence is S_j.

$$\Delta S_j = S_j(i+1) - S(i), \quad i = 1, 2, \cdots N-1 \tag{4.3}$$

$$P = \begin{cases} \dfrac{\sum_{i=1}^{N} \Delta_i}{N} & \sigma(\Delta_{s_j}) \le T_\sigma \\ N/\alpha & \sigma(\Delta_{s_j}) > T_\sigma \end{cases} \tag{4.4}$$

where, ΔS_j is the distance between adjacent peak points in S_j. If the corresponding distance is less than the threshold T_σ, then the column is the column of the real scale ruler, and the pixel length of each scale is the peak of each adjacent peak point. It is average of the distance.

By calculating all the pixel lengths on the image and interpolating them with the length of the pixels in the image, a scaled image set with the same distance can be obtained.

$$F = s(f, \frac{P_f}{P_n}) \tag{4.5}$$

In the formula, the image is interpolated and scaled, and the ratio of the scale of the image to the standard image determines the scale of the zoom.

B. SAMPLE SET ALLOCATION

The collected image samples are divided into training samples and experimental samples, which are used to train neural network models and experimental samples, respectively.

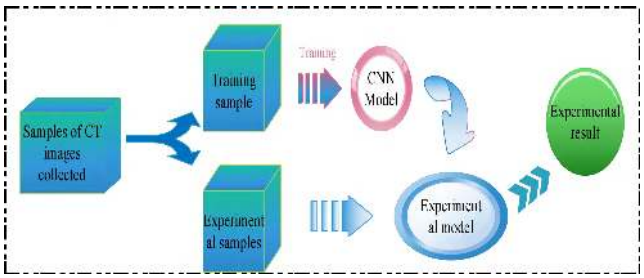


FIGURE 9. Training and experimental calculation flow chart.

The specific allocation is taken as a case of a benign case. First, 10% of the cases are taken from the total number, and then the same number is randomly taken from the second group, and sequentially taken down, and the last remaining cases are individually placed in one group. Malignant cases are taken in the same way as this method. Randomly taken out a group in the test group, all the remaining groups are placed in the training group, and finally cross-validation is used to obtain a more stable model. The specific number of each grouping of the data set is shown in Table 1 below:

TABLE 1. Distribution of training and test groups under cross-validation.

Group	Training group (example)		Test group (example)		Total (column)
	Benign Case	Malignant Case	Benign Case	Malignant Case	
1	240	210	27	22	500
2	240	210	27	22	500
3	240	210	27	22	500
4	240	210	27	22	500
5	240	210	27	22	500
6	240	210	27	22	500
7	240	210	27	22	500
8	241	209	26	23	500
9	241	209	26	23	500
10	241	209	26	23	500

C. BLOCK EXTRACTION

In this study, two-dimensional data blocks are extracted from each pixel in the image and used as training data. After binarization, the effective area can be cut out, then the value is 255 is the effective area. The pixel points further derive the distribution of such pixel points. Referring to the standard of tumor segmentation gold in medicine, the extracted data is divided into positive and negative samples by using the central pixel as a standard, wherein the corresponding portion of the tumor region is a positive sample, and the corresponding region of the negative sample is a normal tissue region. Avoid the difference between the positive sample and the negative sample. Take the random sampling method to make the number of positive and negative samples input in the training model equal. In this study, four different sizes of data blocks were extracted, which were 13×13, 15×15, 17×17, and 19×19. Finally, the four different sizes of data were used to conduct experiments. The experimental results are shown in the figure. Show: It can be obtained that the method is more suitable for 17×17 data blocks.

$$\text{Specificity } SPE = \frac{TN}{TN + FP} \tag{4.6}$$

$$\text{Accuracy } ACC = \frac{TP + FN}{FP + FN + TN + FP} \tag{4.7}$$

$$\text{True positive rate } APR = \frac{TP}{TP + FN} \tag{4.8}$$

$$\text{False positive } rate\ FPR = \frac{FP}{FP + TN} \tag{4.9}$$

A set of specificities and sensitivities will vary with the decision threshold. If only SEN and SPE are used, only a specific threshold can be evaluated [30]–[32].

TABLE 2. Recognition rate of three data samples in the same database by different LBP algorithm and CNN algorithm (%).

Data sample	LBP	LBP_C	LBP_M	LBP_S	Xception	DenseNet
A	77.32	85.32	96.95	96.62	97.67	97.56
B	78.90	82.52	96.23	96.53	97.90	98.03
C	63.25	73.22	82.31	87.12	91.6	90.14

D. ANALYSIS OF RESULTS

In order to verify the feasibility and effectiveness of the algorithm, this paper selects the image features provided by the database. The LBP mode algorithm is the rotation invariant mode and the CNN is the convolutional neural network algorithm. To avoid the contingency of training results, training samples were randomly selected from database samples and repeated 150 times for each experiment. It can be seen from Table 2 that the recognition rate of the data samples by the two convolutional neural network algorithms is significantly higher than that of the local binary mode algorithm. In order to exclude the influence of a single database on the experimental results, this study experiments the two algorithms into three different databases. The training results are shown in Table 3. It is available that the recognition rate of the database based on the two algorithms constructed by CNN still has Advantage.

The three samples A, B and C in Tables 2 and 3 are CT images under different conditions. TC10 and TC12 represent different databases respectively. LBP, LBP_C, CLBP_M and

TABLE 3. Recognition rate of different databases of Outex (TCl0, TCl2) by different LBP algorithm and CNN algorithm (%).

Algorithm	TC10	TC12		Average Value
		Tl84	horizon	
LBP	83.91	85.70	79.02	82.88
LBP_C	87.93	82.36	82.63	84.31
LBP_M	98.92	98.25	97.25	98.14
LBP_S	99.33	99.26	98.13	98.91
Xception	99.60	99.41	98.90	99.30
Dense Net	99.59	99.65	99.15	99.46

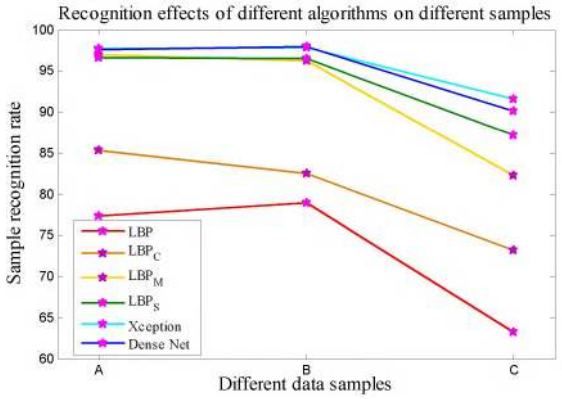


FIGURE 10. Recognition effect of different algorithms on different samples.

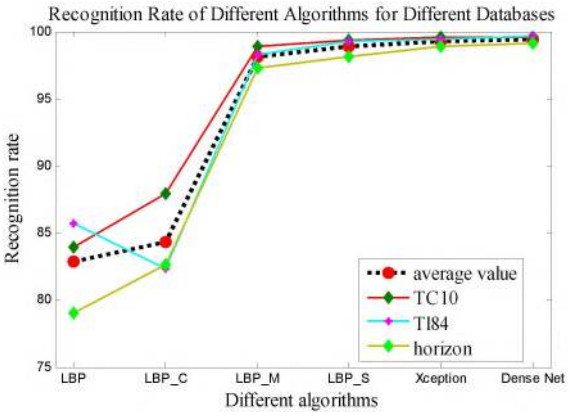


FIGURE 11. Recognition effect of different algorithms on different databases.

CLBP_S represent different local binary algorithms. Xception and Dense Net are two different convolutional neural network models. From the data in the table, the recognition rate of the convolutional neural network is higher than the local binary mode.

Figure 10 and Figure 11 correspond to Table 2 and Table 3, respectively. For each image sample, the recognition effect of these six methods is shown in 4.2, which indicates the recognition degree of each sample by different algorithms. Figure 11 shows the database for different databases. The recognition effect of different algorithms, the dotted line in the figure is the average value.

In this paper, the data set is separately tested and the index value is calculated. Because there are too many types of indicators, the average value is used to represent the classification accuracy rate, sensitivity and specificity of the test index classification algorithm of each algorithm. The following table shows that the algorithm provided in this study can ensure the sensitivity and high specificity of the algorithm at the same time. In medical treatment, it can accurately distinguish between benign and malignant tumors. AUC is a commonly used evaluation index for evaluating the effect of classifiers [30]. As shown in Table 4.

TABLE 4. Comparison of classification effects of various classification algorithms in open source test data sets.

Algorithm	Accuracy /%	Sensitivity /%	Specificity /%	AUC /%
LBP	97.12	97.35	92.15	98.20
LBP_C	98.20	99.12	93.60	99.52
LBP_M	93.21	93.84	91.16	94.12
LBP_S	92.01	92.15	84.53	97.51
Xception	99.01	96.44	96.80	99.70
Dense Net	99.16	99.75	95.86	99.71

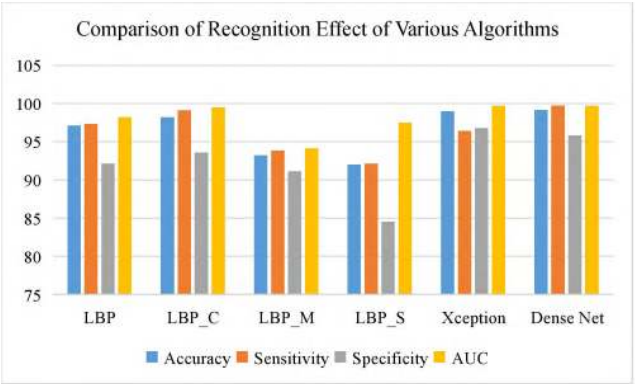


FIGURE 12. Classification effect diagram of each classification algorithm in open source test data set.

Figure 12 shows the effect of different methods on different indicators of different samples. From the perspective of AUC, the image feature extraction effect of convolutional neural network is good, and the average evaluation index value is 99.7%.

It can be seen that convolutional neural networks exhibit very high accuracy, sensitivity and specificity in image feature extraction, pattern recognition, and classification [31], [32]. A direct comparison between the convolutional neural network algorithm and the local binary model can reflect that the algorithm mentioned in this study can perform more accurate cancer risk assessment on tumor CT images on larger data.

V. CONCLUSION

In this paper, local binary model algorithm and convolutional neural network algorithm are used to image and extract features of tumor CT images in medical field. The local binary mode is an image recognition method based on image-based translational rotation invariance. The paper also introduces two models of convolutional neural networks. For the same data samples and sample sizes, the convolutional neural network model has a recognition rate of 99.7% for medical images, which is provided for the follow-up development expert diagnosis system and self-checking system, which is Strong technical support.

The algorithm and model described in this paper, as an algorithm for continuous optimization, use the extended data set to train a new tumor image-sensitive deep convolutional neural network, and strive to add other different types of cost factors to the classification process to continue the algorithm. Improve.

REFERENCES

[1] K. S. C. Chao *et al.*, "A novel approach to overcome hypoxic tumor resistance: CU-ATSM-guided intensity-modulated radiation therapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 49, no. 4, pp. 1171–1182, Mar. 2001.

[2] D. L. Schwartz *et al.*, "FDG-PET prediction of head and neck squamous cell cancer outcomes," *Arch. Otolaryngol. Head Neck Surg.*, vol. 130, no. 12, pp. 1361–1367, Dec. 2004.

[3] B. Ganeshan, K. Skogen, I. Pressney, D. Coutroubis, and K. Miles, "Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: Preliminary evidence of an association with tumour metabolism, stage, and survival," *Clin. Radiol.*, vol. 67, no. 2, pp. 157–164, Feb. 2011.

[4] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Sep. 2014, pp. 1891–1898.

[5] L. Chen, S. Wang, W. Fan, J. Sun, and S. Naoi, "Beyond human recognition: A CNN-based framework for handwritten character recognition," in *Proc. IEEE Asian Conf. Pattern Recognit. (IAPR)*, Nov. 2015, pp. 695–699.

[6] Y. Y. Liu, M. Chen, H. Ishikawa, G. Wollstein, J. S. Schuman, and J. M. Rehg, "Automated macular pathology diagnosis in retinal OCT images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding," *Med. Image Anal.*, vol. 15, no. 5, pp. 748–759, Oct. 2011.

[7] D. Sidibé *et al.*, "An anomaly detection approach for the identification of DME patients using spectral domain optical coherence tomography images," *Comput. Methods Programs Biomed.*, vol. 139, pp. 109–117, Feb. 2017.

[8] C. S. Lee, A. J. Tyring, Y. Wu, A. Rokem, A. Y. Lee, and N. P. Deruyter, "Deep-learning based, automated segmentation of macular edema in optical coherence tomography," *Biomed. Opt. Express*, vol. 8, no. 7, pp. 3440–3448, Jun. 2017.

[9] D. S. Kermany *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, Jun. 2018.

[10] Z. Ji and L. H. Nie, "Texture image classification with noise-tolerant local binary pattern," *J. Comput. Res. Develop.*, vol. 53, no. 5, pp. 1128–1135, May 2016.

[11] K. Zhou, L. M. Yang, and H. Song, "Face recognition method based on dynamic threshold local binary pattern," *Comput. Eng.*, vol. 35, no. 17, pp. 167–169, Sep. 2009.

[12] A. Yang, S. Li, H. Lin, and D. Jin, "Edge extraction of mineralogical phase based on fractal theory," *Chaos, Solitons Fractals*, vol. 117, pp. 215–221, Nov. 2018.

[13] H.-C. Shin *et al.*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.

[14] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[15] Y. Wu, S.-Y. Wang, and Y.-B. Hou, "A study on spectral super-resolution of hyperspectral imagery based on redundant dictionary," *Inf. Eng.*, vol. 3, no. 5, pp. 67–71, Oct. 2014.

[16] S. Li, R. Fan, G. Yue, C. Hou, and G. Lei, "A two-channel convolutional neural network for image super-resolution," *Neurocomputing*, vol. 275, no. 31, pp. 267–277, Jan. 2018.

[17] T.-Y. Lin, A. R. Chowdhury, and S. Maji, "Bilinear CNN models for fine-grained visual recognition," in *Proc. IEEE Int. Conf. Comput. Vis.*, Apr. 2015, pp. 1449–1457.

[18] M. Simon and E. Rodner, "Neural activation constellations: Unsupervised part model discovery with convolutional networks," in *Proc. 15th IEEE Int. Conf. Comput. Vis. (ICCV)*. vol. 3, Apr. 2015, pp. 1143–1151.

[19] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, and Y. Pan, "A survey of MRI-based brain tumor segmentation methods," *Tsinghua Sci. Technol.*, vol. 19, no. 6, pp. 578–595, Dec. 2014.

[20] Y. J. Huang and Q. J. Feng, "Segmentation of brain tumor on magnetic resonance images using 3D full-convolutional densely connect," *J. Southern Med. Univ.*, vol. 38, no. 6, pp. 661–668, Jun. 2018.

[21] C. Chaux, L. Duval, and J.-C. Pesquet, "Image analysis using a dual-tree M-band wavelet transform," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2397–2412, Nov. 2018.

[22] T. Fujiwara *et al.*, "The utility of sonographic features during endobronchial ultrasound-guided transbronchial needle aspiration for lymph node staging in patients with lung cancer: A standard endobronchial ultrasound image classification system," *Chest*, vol. 138, no. 3, pp. 641–647, Jan. 2010.

[23] S. Gupta, R. C. Chauhan, and S. C. Sexana, "Wavelet-based statistical approach for speckle reduction in medical ultrasound images," *Med. Biol. Eng. Comput.*, vol. 42, no. 2, pp. 189–192, Mar. 2004.

[24] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, "Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach," *Magn. Reson. Med.*, vol. 78, no. 6, pp. 2439–2448, Dec. 2017.

[25] J. X. Sun and Y. Z. Chen, "Novel speckle reduction for medical ultrasound images based on edge preservation," *Opt. Precis. Eng.*, vol. 10, no. 5, pp. 429–433, May 2002.

[26] Y. Zhou, "Ultrasound diagnosis of breast cancer," *J. Med. Imag. Health Inform.*, vol. 3, no. 2, pp. 157–170, Mar. 2013.

[27] J. Ringenberg, M. Deo, O. Berenfeld, P. Boyers, J. Gold, and V. Devabhaktuni, "Fast, accurate, and fully automatic segmentation of the right ventricle in short-axis cardiac MRI," *Comput. Med. Imag. Graph.*, vol. 38, no. 3, pp. 190–201, Apr. 2014.

[28] L. Chen, Y. Yang, X. Li, and Z. Cen, "Region labeling algorithm based on boundary tracking for binary image," *Proc. SPIE*, vol. 7850, no. 1, p. 78500K, Nov. 2010.

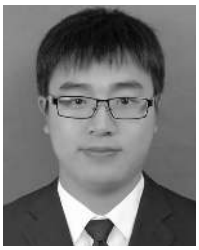
[29] G. Liu, B. Chen, S. Jiang, H. Fu, L. Wang, and W. Jiang, "Double entropy joint distribution function and its application in calculation of design wave height," *Entropy*, vol. 21, no. 6, p. 64, Jan. 2019.

[30] H. X. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* Jun. 2015, pp. 5325–5334.

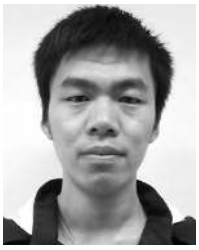
[31] W. Liu, "A image coding method based on CL multiwavelet transform," *Acta Scientiarum Natralium Univ. Sunyatseni*, vol. 49, no. 4, p. 012, Apr. 2011.

[32] L. R. Bachega, S. Hariharan, N. B. Shroff, and C. A. Bouman, "Distributed signal decorrelation and detection in multi view camera networks using the vector sparse matrix transform," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6011–6024, Dec. 2015.

AIMIN YANG received the B.Sc., M.Sc., and Ph.D. degrees from Yanshan University, in 2002, 2004, and 2015, respectively. He is currently a Professor with the North China University of Science and Technology. His current research interests include numerical computation, iron and steel big data, and intelligent computation.



XIAOLEI YANG was born in Cangzhou, Hebei, China, in 1992. He received the degree from the Qinggong College, North China University of Science and Technology, where he is currently pursuing the degree in mathematics with the Graduate School. His research interest includes applied mathematics.



WENRUI WU was born in Handan, Hebei, China, in 1995. He received the degree from the North China University of Science and Technology, where he is currently pursuing the degree in mining engineering with the Graduate School. His research interests include rock mechanics and rock engineering.

HUIXIANG LIU is currently pursuing the master's degree in mathematics majoring in numerical calculation with the North China University of Science and Technology, Tangshan, China. Her research interests include numerical calculation, big data, and intelligent computation.

YUNXI ZHUANSUN was born in Jining, Shandong, China, in 1995. He received the degree from Binzhou University. He is currently pursuing the degree in mathematics with the Graduate School, North China University of Science and Technology. His research interest includes numerical calculation and its application.