Special Section on Emerging Deep Learning Theories and Methods for Biomedical Engineering

Intervention of Hepatocellular Carcinoma in Fibrosis Staging Based on Multi-Parameter Magnetic Resonance Image Depth Learning Method

ZHENHUA TIAN1 AND XUDONG QU2

1Department of Interventional Therapy, Shangqiu First People’s Hospital, Shangqiu 476100, China
2Department of Interventional Therapy, Zhongshan Hospital, Fudan University, Shanghai 200032, China

ABSTRACT This paper mainly discusses the deep learning solution for non-invasive evaluation of the differentiation degree of hepatocellular carcinoma based on multi-parameter nuclear magnetic resonance images, combined with the clinical diagnosis experience of radiologists and the characteristics of nuclear magnetic resonance images. The method of multimodal data fusion is studied based on multi-parameter nuclear magnetic resonance imaging data. Multi-channel three-dimensional convolution neural network and multi-scale depth residual network are proposed to extract the features of three-dimensional medical image data and two-dimensional fusion medical image data, and to solve the problem of insufficient cases in clinical image data of hepatocellular carcinoma (HCC). We examine the role of transition learning and metric learning in medical image classification. In this study, we use a method of data fusion, transition learning and multi scale feature extraction to construct a deep learning model for medical image aided diagnosis. Multiple modal fusion decisions for finding complementary modal data fusion for the complementarity of multimodal images in diagnostic decisions can effectively improve diagnostic effects. Although there is a clear difference between natural and medical images through experiments, a model trained with a natural image dataset as an initialization of the network can ensure and converge the training. At the same time, improve the performance of the model on the test set. The multi-scale feature extraction model proposed in this paper enhances the robustness of the model and further improves the effect of medical image classification.

INDEX TERMS Computer-aided diagnosis, multi-parameter nuclear magnetic resonance imaging, convolution neural network, liver cancer.

I. INTRODUCTION

In recent years, artificial intelligence technology has made great progress in all aspects of practical application. With the development of medical imaging technology and the needs of clinical medicine, many scientific research institutions and hospitals have begun to deploy the application of artificial intelligence and cognitive computing in medical image-aided diagnosis in the diagnosis of some diseases [1]–[4]. The application of computer technology to clinical diagnosis has also become a research hotspot. At present, computer-aided diagnosis (Computer aided diagnosis, CAD) based on medical images has been applied to the automatic detection and diagnosis of brain, breast, prostate, lung and other lesions. Due to the large number of abdominal organs and complex structure, the process of image acquisition is easily disturbed by various factors [5]. At present, the research on computer-aided detection and diagnosis is rare [6]. The application of artificial intelligence technology to the auxiliary diagnosis of liver diseases is also one of the difficulties of computer-aided diagnosis [7]–[10].

Hepatocellular carcinoma ((Hepatocellular carcinoma, HCC)) is a kind of tumor that usually occurs in the epithelium of the liver [11]. It is the most common malignant tumor of the liver and the third deadliest cancer in the world. According to authoritative journals in the world, about 80% of patients with liver cirrhosis eventually converted to hepatocellular carcinoma. HCC generally does not show some symptoms...
different from the common liver disease before it becomes malignant and cannot be cured [12]. In other words, once the patient has the corresponding symptoms, it usually indicates a poor prognosis. Treatment plans for HCC patients with different degrees of differentiation are usually different [13]. If we can make a preliminary prediction of the degree of differentiation of HCC before surgery, we can make a better treatment plan for patients. In order to achieve the purpose of improving the survival rate of patients [14]. At present, pathological biopsy is the gold standard for the diagnosis of HCC differentiation at home and abroad, which is harmful to the human body and is very risky, so a non-invasive diagnosis method is urgently needed. Nuclear magnetic resonance imaging (MRI) examination is one of the possible ways for non-invasive evaluation of HCC [15]–[19].

There are many kinds of liver diseases, involving a wide range. In clinical diagnosis, hepatocellular carcinoma is a very important diagnostic node for most liver diseases. The prognosis of HCC and the formulation of a complete and appropriate treatment plan depend to a large extent on the degree of tumor differentiation of patients. At present, it is recognized that the most accurate and effective method to evaluate the differentiation degree of liver tumor is the pathological gold standard. In other words, the focus area is punctured before operation, but the tumor puncture biopsy has many shortcomings, such as invasiveness, needle tract metastasis, sampling error and so on [20]. These limitations limit the application of tumor biopsy in preoperative evaluation of HCC. Based on the non-invasive and non-local nature of imaging methods, compared with pathological examination, it provides a feasible solution for the grading and evaluation of diseases [21]. At present, most systems use the degree of imaging differentiation, that is, to distinguish the degree of HCC differentiation according to some symptoms shown by medical images. Nuclear magnetic resonance imaging (MRI) is a radiation-free imaging method for human body, and the scanning results are accurate, which makes it have unique advantages in the examination of some liver diseases and the quantitative analysis of various indexes [22]. In recent years, with the development of intelligent imaging technology, the deep learning method is applied to MRI images to distinguish the differentiation degree of HCC, and to establish a relatively objective and high-precision preoperative non-invasive evaluation, which is of great practical value for improving the current medical service quality, improving the clinical diagnosis effect and alleviating the work intensity of doctors [23].

In the aspect of computer-aided diagnosis based on medical images, computer-aided diagnosis technology has been applied to the segmentation, detection and diagnosis of a variety of diseases [24]–[28]. For example, computer-aided detection and diagnosis of brain microhemorrhage based on MR images, computer-aided diagnosis of benign and malignant breast tumors, automatic detection and diagnosis of prostate cancer and detection of pulmonary nodules based on CT images, there is no related research on computer-aided grading diagnosis of hepatocellular carcinoma based on CT and MR images. Medical images such as CT and MR are mostly three-dimensional images. There are many computer-related methods to deal with similar three-dimensional data. Dou Q et al proposed a cascaded three-dimensional convolution neural network framework for the detection of cerebral microvascular hemorrhage (Cerebral Microbleeds, CMB) [29]. The three-dimensional full convolution network is used to retrieve the candidate regions of CMB with high probability, and then the trained model is used to distinguish the CMB from the non-CMB in the candidate regions. Jing Li et al compared the performance of two-dimensional and three-dimensional convolution neural networks to distinguish malignant and benign breast tumors on the basis of MR images, and proved that three-dimensional convolution neural networks can achieve higher accuracy [30]. However, they only consider the two time series of DCE-MRI, and perform a simple linear transformation between the plain scan image and the enhanced image to obtain the enhanced transformation samples. It is still not accurate to describe the dynamic contrast enhancement process of the lesions in this way. At present, in practical clinical application, only one kind of imaging examination is carried out, and obtaining a single type of medical image is not enough to provide complete diagnostic information. For radiologists to accurately diagnose the benign or malignant lesions or the malignant degree of the tumor, doctors usually need to make a comprehensive analysis combined with the size, shape, average grayscale of different modal images and the sections of different directions of three-dimensional images. In order to get a more accurate diagnosis or work out a more appropriate treatment plan [31]. Wu Zhou et al extracted texture features such as the average intensity of HCC DCE-MRI images to predict the histological grading of HCC, which is the only work available in the literature to analyze the degree of HCC differentiation by computer [32]. At present, there is no related work on the use of deep learning method to predict the degree of HCC differentiation [33]–[35]. In this paper, based on the progress and achievements of deep learning technology in the field of computer-aided diagnosis based on medical images, deep learning technology is applied to the non-invasive evaluation of hepatocellular carcinoma differentiation. Computer aided diagnosis technology based on medical images was always a hot spot in the field of computer applications. The urgent necessity of controllability and practical clinical diagnosis of acquiring medical image data brings breakthrough progress to the detailed learning landing. The paper mainly discusses non invasive evaluation of hepatocellular carcinoma differentiation based on multi parameter nuclear magnetic resonance imaging. Based on this, multichannel 3D convolution neural network and multiscale depth residual network are proposed to extract features of three-dimensional medical image data and two-dimensional fusion medical image data. The role of metastasis learning and metric learning in the classification of medical images is verified.
II. SPATIO-TEMPORAL FEATURE EXTRACTION OF MULTI-PARAMETER MAGNETIC RESONANCE IMAGES BASED ON DEPTH LEARNING ALGORITHM

A. REPRESENTATION AND DIMENSIONALITY REDUCTION OF HIGH-DIMENSIONAL MEDICAL IMAGE DATA

Medical images reflect the information of human tissues and organs, and it is usually necessary to retain the spatial structure information of multi-layer cross-sectional scanning, so it has three-dimensional spatial structure information. In addition, clinical diagnosis usually requires different imaging methods to scan diseased organs or tissues, and a number of three-dimensional image data will be obtained, so medical image data often have high-dimensional characteristics. DCE-MRI data, which have good clinical diagnostic effect on hepatocellular carcinoma, not only include three-dimensional image data, but also have time series, that is, three-dimensional image data are collected at several time points before and after injection of contrast agent, so DCE-MRI data can be regarded as four-dimensional data including three-dimensional space and one-dimensional time. How to establish an effective depth network to learn the discriminable feature representation of this kind of data. The realization of automatic evaluation of HCC differentiation degree based on DCE-MRI is an important problem to be solved in this chapter. Aiming at this research topic, this paper introduces the methods of feature extraction, dimensionality reduction and classification for high-dimensional data such as DCE-MRI.

This study uses DCE-MRI medical image data as input to realize the automatic evaluation of HCC differentiation degree. The data are from Beijing Friendship Hospital. The radiology department of Beijing Friendship Hospital uses GE 3.0T750 magnetic resonance imaging system to collect images for patients and obtain the original image data of the focus area. The data preprocessing process of the original image data is introduced below. The data used in this paper contains a series of five periods of DCE-MRI, each of which is a three-dimensional data with three dimensions (represented by coronal plane, sagittal plane and cross section in human anatomy). These six time series are shown in Figure 1. From the picture, we can see that the DCE-MRI image consists of six time series: S0, S1, S2, S3, S4 and SS, which represent the MRI images collected at 16s, 26s, 60s, 180s and 300s before and after contrast injection, respectively, including plain scan, early arterial phase, late arterial phase, portal phase, balance phase and delay phase, and the images collected in each phase are 3D images.

The background area of liver cirrhosis and the three-dimensional location of HCC lesions in six sequences of each case were marked one by one by the attending radiologist engaged in abdominal imaging diagnosis for 10 years, and then examined by the chief radiologist engaged in abdominal imaging diagnosis for 30 years. In order to adapt to the input of the neural network, when selecting the experimental samples, the HCC cross-sectional scale of each time series is further normalized to $32 \times 32$. On the longitudinal level, the image information of two levels on each side of the largest region is preserved uniformly, and the images of the first five periods are preserved in this chapter. Therefore, the (Region of Interest, ROI) size of each time series is $32 \times 32 \times 5$, and each case has five time series ROI. A fourth-order tensor with a size of $32 \times 32 \times 5 \times 5$ was used to represent the data of each case. The time-signal intensity curve of a region in the DCE-MRI image can be obtained by quantifying the MRI signal intensity change of the DCE-MRI image with time. In this paper, we also extract the $16 \times 16$ size region of the ROI center of the S0∼S5 mode of each sample, and calculate the average signal strength of this region as the strength of the current sequence. The TSIC curves of the HCC region and the background region of the three differentiation levels are shown in Figure 2.

The horizontal axis of the line chart is the time series of DCE-MRI images, and the vertical axis is the average.
signal strength of the corresponding series. The TSIC curves of background (Background, BG) and HCC differentiation degree of well, moderately, poorly are calculated respectively. From the DCE-MRI TSIC curves of the three differentiated lesions, we can see that the overall trend of the three kinds of HCC differentiation degree data is similar, but there is a progressive relationship in the data signal strength, in which the lowest malignant degree of well (highly differentiated) grade is closer to the background. Thus it can be seen that the blood perfusion information reflected by DCE-MRI images plays an important role in the diagnosis of HCC differentiation.

B. FEATURE DIMENSIONALITY REDUCTION OF HIGH-DIMENSIONAL MULTI-PARAMETER MAGNETIC RESONANCE IMAGE DATA

Aiming at the multimodal image data (such as CT, MRI, etc.), the multimodal image data of M patients are spliced according to the image type and image data to construct a fourth-order tensor $X$. Combined with the results of pathological examination of the patients, that is, the degree of HCC differentiation, according to the evaluation criteria of WHO, the differentiation degree of HCC is divided into three grades: well, moderately and poorly, as $F1$ and $F2$, and $F3$ forms the differentiation degree of HCC. According to the above idea, the Tucker decomposition of the image. The constraint here means that the tensor decomposition results can better reflect the correlation between image data and pathology. We add correlation constraints to the Tucker decomposition factors that reflect different patient data, that is, the corresponding decomposition factor matrix of the fifth order. For other orders, we can carry out orthogonal constraints, sparse constraints, smooth constraints, non-negative constraints and so on. Based on the above idea, the Tucker decomposition of $T$ is to solve the following problems:

$$T_i = \left[ \frac{\text{vec}(X)}{U^1 \times U^2 \times U^3 \times U^4} \right] \times \text{vec}(G)$$

The above objective function is further simplified to get:

$$U = \frac{1}{2} \sum_{i=1}^{5} \xi_i \times [X_i \times U^{(1)} \times U^{(2)} \times U^{(3)} / U^{(4)} U^{(5)}]$$

According to the above formula, the factor matrix $U$ of tensor decomposition is obtained to project the untagged patient image data:

$$Y = X \times \sqrt{U^{(1)} \times U^{(2)} \times U^{(3)} / U^{(4)} U^{(5)}}$$

In formula (2), $G$ is the decomposition kernel tensor and $U$ is the corresponding factor matrix of each order after the constrained Tucker decomposition of the image. The constraint here means that the tensor decomposition results can better reflect the correlation between image data and pathology. We add correlation constraints to the Tucker decomposition factors that reflect different patient data, that is, the corresponding decomposition factor matrix of the fifth order. For other orders, we can carry out orthogonal constraints, sparse constraints, smooth constraints, non-negative constraints and so on. Based on the above idea, the Tucker decomposition of $T$ is to solve the following problems:

$$T_i = \frac{\text{vec}(X)}{U^1 \times U^2 \times U^3 \times U^4} \times \text{vec}(G)$$

The above objective function is further simplified to get:

$$U = \frac{1}{2} \sum_{i=1}^{5} \xi_i \times [X_i \times U^{(1)} \times U^{(2)} \times U^{(3)} / U^{(4)} U^{(5)}]$$

According to the above formula, the factor matrix $U$ of tensor decomposition is obtained to project the untagged patient image data:

$$Y = X \times \sqrt{U^{(1)} \times U^{(2)} \times U^{(3)} / U^{(4)} U^{(5)}}$$

In this case, the projection result after low-rank approximation is still a fourth-order tensor, the tensor characteristic $z$. These features and the corresponding diagnostic tags are used to matrixize the tensor to get the image data $y$ as the training data to train the classifier.

When performing the convolution operation, the three-dimensional convolution neural network has a number of independent three-dimensional convolution kernels and the local multiplication and summation of the third-order three-dimensional data of the fourth-order tensor. Suppose the input image is $I$, where $X$, $Y$ and $Z$ represent the spatial dimension of the input 3D data respectively, and $u$ is the fourth dimension of the fourth-order tensor (similar to the color channel in the natural image). The operation of 3D convolution layer can be expressed as:

$$u(x,y,z) = \sum h^0_{i-1}(x-i) \times \int W(i,j,k)$$

C. TENSOR REPRESENTATION AND SPATIO-TEMPORAL FEATURE EXTRACTION OF MULTI-PARAMETER MAGNETIC RESONANCE IMAGE DATA

From the above analysis, we can see that if the fourth-order tensor samples are sent into the three-dimensional convolution neural network as multi-channel three-dimensional data, the convolution kernel of three-dimensional convolution can only extract the third-order information of the fourth-order tensor at the same time, and carry out numerical summation in the fourth order, so the following two kinds of data tensor representations are designed from the point of view of extracting DCE-MRI image time series information and three-dimensional image space information. In this paper, when the tensor proposed in this paper is used to represent the DCE-MRI data as the input of 3DCNN, the three-dimensional convolution nuclear energy is convoluted with the data of five phases at the same time, which is more beneficial to extract the time series features of DCE-MRI. When the second data representation method is used, the three-dimensional convolution nuclear energy is used to extract the spatial structure features of each phase of DCE-MRI. As far as the data with the shape of $32 \times 32 \times 5$ collected in this paper is concerned, the parameters of the model required by the two data representation methods are exactly the same. The method used in processing the actual medical image data is the transposition of the tensor, which directly changes the order of the two orders of the exchange tensor, as shown in Figure 3.

DCE-MRI ‘s five time series MRI images, each time series image has a three-dimensional structure, in other words, each sample we collected contains multiple three-dimensional data blocks. According to the analysis in the previous section, we can directly regard the original three-dimensional data blocks as the input of the three-dimensional convolution neural network, regard the time dimension of the data as a channel, exchange the spatial slice dimension and time series
dimension of the data, and convert the data into the following form. At this time, each three-dimensional data block contains all the sequences of DCE-MRI data. In addition, we can also see from the Figure that the overall shape of the data dimension remains unchanged before and after the exchange, which is $32 \times 32 \times 5$. The advantage of this is: for different data representation methods, the shape is the same, so as the input of the same neural network, the number of network parameters in the first layer of the network is exactly the same, so it is convenient for comparative experiments. Three-dimensional convolution neural network uses fourth-order tensor data as the input of three-dimensional convolution neural network, and then constructs a deep three-dimensional convolution neural network model by stacking convolution layer, pooling layer and full connection layer. The output layer uses softmax classifier to generate the prediction probability of HCC differentiation degree. The multi-channel three-dimensional convolution neural network architecture is shown in Figure 4.

**FIGURE 3.** Transposition of multi-sequence 3D medical image data.

**FIGURE 4.** The structure of three-dimensional convolution neural network.

Input, Output is input and output layer, C1 and C2 are convolution layer, M1 and M2 are maxpooling layer, FC 1 and FC2 are fully connected layer. The number of 3DCNN feature graph, the size of convolution core and the size of pool layer are all marked in the diagram. The number of convolution nuclei in the two convolution layers was only 6 and 8, respectively, and the number of neurons in the two fully connected layers was only set to 100 and 32, respectively. Because considering that most of the parameters of the convolution neural network are concentrated in the full connection layer, the purpose of this is to reduce the number of parameters and avoid the occurrence of overfitting. In the network implementation, the rectifier linear unit (Rectified linear unit, relu) is used as the nonlinear activation function of the convolution layer and the fully connected layer. Initializes the 3D convolution kernel from Gaussian distribution N, where f. The number entered for this convolution layer. Using the exponential decay rate estimated by the first moment, the exponential decay rate estimated by the second moment. The adaptive matrix estimation (Adaptive moment estimation, Adam) optimization algorithm with an initial learning rate of 0.001 adjusts the trainable parameters in the network to minimize the crossover loss. At the same time, the learning attenuation method is also used. At the same time, the dropout method with a scale of 0.5 is used to prevent overfitting. In order to make full use of the temporal and spatial structure information of DCE-MRI data, this paper proposes a 3DCNN fusion network, which uses independent convolution networks with the same structure to extract the image features of the third-order components of the fourth-order tensor, and then fuses the extracted features.

Due to the limitations of medical image data quality, data completeness and other factors, the data set is very small, and the number of differentiated lesions is not balanced enough. In order to prevent the model from overfitting, after dividing the data set, the ROI data of the training set and the test set are transposed, rotated and flipped respectively to increase the amount of data. The transpose operation can increase the amount of data of the original data set by 2 times, and rotate 90%. The amount of data of the original data set can be increased by 3 times, and the amount of data of the original data set can be increased by 2 times by horizontal and vertical flip operation. Therefore, after using the
data expansion method, the number of data samples can be increased to 8 times. The imbalance of data categories has a certain impact on the results of the convolution neural network model. In order to solve the problem of imbalance of all kinds of sample data in the convolution neural network training data set, the class sampling method proposed in the imagenet2015 competition overcomes the problem of category imbalance. Similarly, the champion of scene classification task in the imagenet2016 competition also uses a similar method to solve the problem of sample category imbalance.

In this paper, the eclectic sample resampling method is used to sample up and down multi-class samples, and the sampling method on a few categories of samples makes up for the lack of too few data samples in some degree of differentiation.

The experiment was repeated 10 times independently, and the resampling process was performed in each experiment to ensure that different samples were resampled randomly each time. The experimental results were evaluated by correct (Accuracy), sensitivity or recall (Sensitivity&Recall), accuracy (Precision) and F 1-score. The mean and standard deviation of the 10 experiments are shown in Figure 5 below.

In order to analyze and classify the results in more detail, we draw the confusion matrix of the results with the highest accuracy in the test set in dozens of experiments, in which the label 0 ∼ (1) and 2 represent three categories of poorly, moderately, well respectively. From the above confusion matrix, we can see that the ability of MCF-3D CNN to distinguish different pathological grades of HCC is different, and it has the strongest diagnostic ability for highly differentiated HCC. It can distinguish all highly differentiated samples from low and moderately differentiated samples, while low and moderately differentiated samples are easy to be misclassified. Therefore, the three classification problems of high, middle and low differentiation of HCC are further divided into three binary classification problems, that is, to distinguish between high differentiation HCC and middle and low differentiation HCC; C2, to distinguish high school differentiation HCC and low differentiation HCC; C3, to distinguish middle differentiation HCC and the other two differentiation HCC, to repeat the experiment ten times using MCF-3DCNN model, the experimental results are shown in Table 1 below.

In this paper, the three-classification problem of high, middle and low differentiation of HCC is decomposed into three binary classification problems. The AUC value of multi-channel three-dimensional convolution neural network for distinguishing highly differentiated (low malignant) samples and the other two categories is 0.70 ± 0.03, 0.70 ± 0.03, 0.64 ± 0.04, 0.70 ± 0.03 and 0.64 ± 0.04, respectively. The experimental results show that the proposed model can more easily distinguish the highly differentiated samples from the other two samples. As shown in Figure 6, after analyzing the original data, it is found that the sample data of the two categories with higher malignant degree (low differentiation and middle differentiation) have diversity, great individual differences, and less data of low differentiation samples. These reasons make the model easy to confuse the samples of the two categories with higher malignant degree.

### III. COGNITIVE CALCULATION METHOD BASED ON MULTI-PARAMETER MAGNETIC RESONANCE IMAGING OF LIVER CANCER WITH FIBROSIS STAGING

#### A. INFORMATION FUSION OF LIVER CANCER IMAGES WITH MULTI-PARAMETER MRI FIBROSIS STAGING

In this study, a multi-parameter magnetic resonance imaging (Multi-parametric MRI, mp-MRI) data set was constructed to evaluate the differentiation of HCC. Here we call the images obtained by different imaging methods under different parameters as multimodal MRI image data, and the following multimodal images refer to different kinds of MRI images. One of the most valuable MRI sequence images for HCC diagnosis is dynamic contrast enhanced imaging. The DCE-MRI data consists of six phases, including the

| Types  | Accuracy  | Sensitivity | Specificity | AUC  |
|--------|-----------|-------------|-------------|------|
| Poorly | 0.770±0.0539 | 0.343±0.0449 | 0.942±0.0612 | 0.644±0.0469 |
| Moderately | 0.682±0.0221 | 0.703±0.0453 | 0.671±0.0341 | 0.706±0.0322 |
| Well   | 0.910±0.0303 | 0.968±0.0403 | 0.896±0.0593 | 0.964±0.0391 |

### FIGURE 6. The average ROC curve of MCF-3DCNN to distinguish highly differentiated HCC.
five periods used in the previous chapter, as well as the delay period. The acquisition of DCE-MRI images has been introduced, this time a total of six complete time series of DCE-MRI three-dimensional images are used. In the experimental analysis in the previous chapter, we draw the following conclusions: compared with the three-dimensional spatial structure information of tumors, the time series information of DCE-MRI is more helpful to distinguish the background of HCC and liver cirrhosis and to evaluate the differentiation degree of HCC. The main reason is that the two-dimensional texture information of the longitudinal section of the MRI image can reflect the spatial structure information of the tumor to a certain extent, and the time series information of the DCE-MRI image reflects the blood perfusion information of the tumor, which is obviously more important than the texture information of the third dimension of the tumor. In order to process DCE-MRI image data more conveniently and realize multi-modal MRI image fusion, one of the two-dimensional layers of DCE-MRI 3D image is extracted. In addition to the DCE-MRI with time series information, the same method is used to obtain the marked images of T2WI and T1WI in the same phase (T1WI/IN, T1WI/OUT). Because the scan thickness (6.0mm) of T2WI and T1WI in the same inverse phase is larger than that of DCE-MRI (2.2mm), the 2D plane of each tumor data is retained, and the 2D data information of three modes is obtained. Figure 7 is an example of data of 9 modes in some cases.

FIGURE 7. T2WI, T1WI same and inverse phase, DCE-MRI image and its fusion image.

As can be seen from the picture, the tumor details in the color images with multiple modal data fusion are clearer, and the images fused with different modal data are also different. We must comprehensively use the images with multiple parameters of MRI images in order to evaluate the degree of HCC differentiation more accurately. Through the experiment, we can choose the data with complementary characteristics for fusion, which has achieved a better classification effect. When we use two-dimensional tumor data for diagnosis, single-modal image blocks and multi-modal fused image blocks can correspond to grayscale images and color images in natural images, respectively. This allows us to pre-train the network on the natural image data set, and then fine-tune the pre-trained model on the collected medical image data set for medical image classification.

All the data blocks in the picture have been normalized to the same size, but in fact, the size of the tumor is different, and the tumor size information plays an important role in clinical imaging diagnosis. In order to use the tumor scale information and ensure the batch training of the model, it is necessary to learn from the methods of dealing with multi-scale targets in natural images. After several convolution layers, the spatial pyramid pool is used to extract the multi-scale features of the image. Based on the diagnosis experience of doctors, the key information that doctors pay attention to in the process of diagnosis is fully considered, including the characteristics of MRI images with various parameters and a priori knowledge about tumor scale, multimodal data are fused, and the complementarity of data is used to calculate the features of medical images.

B. DESIGN OF MULTI-SCALE DEEP RESIDUAL NETWORK MODEL BASED ON DEEP LEARNING METHOD

The deeper the network is, the more difficult it is to train, Kaiming He and others to propose a deep residual network, which reduces the classification error rate to 3.5700 in the ILS VRC 2015 task, even higher than the human visual cognitive level (the error rate is 5% to 10%) o Resnet uses a 152-layer deep network structure, which has achieved excellent performance in image recognition, target detection and other fields. The number of network layers of Resnet is 152 layers, which is much higher than that of VGG-Net. In a certain range, the network performance will be relatively improved with the deepening of network layers, but when the layers of ordinary neural networks continue to deepen, the correct rate will reach a zero point, and then it will begin to decline. It is obvious that the increase of error rate in the training phase is not caused by overfitting, and the model begins to degenerate with the simple deepening of depth, indicating that not all networks are easy to optimize. Suppose that when the newly added layer is an identical layer, it will at least have the same effect as the shallow network.

In the doctor’s clinical diagnosis, the size of the tumor is of reference significance for the doctor’s clinical diagnosis. However, in order to ensure the stability of the training in the training process of the deep learning model, the specific training method is usually adopted, which requires that the scale of the input data should be the same, but in practice, the size of the HCC tumor has different scales, in order to adapt to different sizes of tumors. First of all, we normalize the tumor data of all patients to the same size, which brings the deficiency of losing the original scale information of tumor
data to some extent. In order to adapt to different scales of HCC data and extract multi-scale features, combined with depth residual network and multi-scale feature expression, a multi-scale depth residual network is proposed to extract multi-scale features of medical images. Because the use of feature pyramid network or multi-branch network will greatly increase the number of network parameters, this paper uses spatial pyramid pooling to obtain different scale features, and the basic network is Resnet20, in order to extract multi-scale features of data. In order to extract multi-scale features of data, the global average pooling layer of the last layer of Resnet20 is replaced by spatial pyramid pooling layer, which is used to extract data features of different scales. After the input image passes through Resnet20+SPP, a 320D feature vector is obtained. On the far left side of the image is the input data, which is sent to Resnet20 to extract image features, and then the features are sent to the spatial pyramid pool layer to get 320D multi-scale features. The 320D multi-scale features extracted by the above network are sent to the single-layer fully connected neural network for re-coding, and the coding dimensions are set to 2 (for visualization), 16, 32 and 64. Finally, the features are sent to the softmax classifier to get the probability, and the training network is trained by minimizing cross-entropy loss. The network is pre-trained on MNIST and CIFAR 10 respectively to extract the features of 2D single-mode and mixed-mode data blocks.

![FIGURE 8. The result of multi-parameter magnetic resonance image feature mapping to two-dimensional plane.](image)

Figure 8 shows the two-dimensional features of the MNIST test set and the CIFAR10 test set image extracted from the original Resnet20 after training the original MNIST test set using the crossover loss function. After SPP Layer replaces GAP Layer in Resnet20, multi-scale Resnet is trained with cross-scale loss function, and then two-dimensional features of MNIST test set and CIFAR 10 test set are extracted. It can be seen from the Figure that the features obtained by the network using multi-scale features and the original features show a state of divergence centered on the origin in space. Compared with the original features, multi-scale features are not concentrated near the origin, which is more beneficial to classification. This is consistent with the classification results in Table 1.

For the task of distinguishing between HCC and non-tumor areas, radiologists can make judgments based on medical images, but usually need to carefully observe a variety of modal data images, such as DCE-MRI, T2WI, DWI, etc., to make a comprehensive judgment. The accuracy of the results is related to the experience of radiologists, and the judgment process usually takes a long time. In this section, we use clinical MRI images to train the discriminant model between HCC and non-tumor regions. In order to verify the effect of transfer learning on medical image diagnosis, we first ill Resnet20 and Resnet20+SPP models on CIFAR 10 data, in which the output features of Resnet20+SPP are encoded into 32-dimensional feature vectors and the accuracy reaches 90.19% on the verification set of CIFAR 10, so as to simulate human cognitive ability to natural images and initialize the model.

In this section, we use S0Query S2 and S4 as the three channels of the image to train the Resnet20 and Resnet20+SPP models that have been trained on the CIFAR10 data set. At the same time, we set up two control experiments, namely ab initio training Resnet20 and ab initio training Resnet20+SPP, in which the learning rate of pre-training experiment is set to 0.0005 to fine-tune the whole network. Figure 9 shows the average accuracy curve in the training process of the four models and the average ROC curve on the test set in the repeated experiment.

![FIGURE 9. Average ROC curve on magnetic resonance image test set.](image)

**C. BASED ON DEPTH LEARNING METHOD TO JUDGE THE DEGREE OF HCC DIFFERENTIATION OF MULTI-PARAMETER MRI IMAGES**

In the actual process of clinical diagnostics, even experienced radiologists cannot accurately distinguish the HCC differentiation from MRI images alone. Doctors can determine whether or not it is a tumor, and only the likelihood of benign and malignancy by medical images. Cannot determine the specific degree of differentiation of HCC. Currently, the evaluation of HCC differentiation can only be achieved by
pathological puncture, but this method has a great risk. In the clinical laboratory experience of physicians, doctors examine that merging multiple modal data characteristics for computer aided diagnosis using the recognition experience of a physician that gives liver tissue, comprehensive analysis, and the most rational diagnostic decision in various imaging methods of medical images. In order to verify the role of multimodal data fusion, we use clinical MR images to train the model and evaluate the degree of differentiation of HCC. First of all, the original Resnet20 was trained to judge the differentiation degree of HCC by using the data of T2WI, T1Magee S0, S1, S2, S3, S4 and S5 in Figure 10.

**FIGURE 10.** The effect of each mode to evaluate the differentiation degree of HCC.

Based on the analysis of the indexes of a single mode from the diagram, T2WI, T1 * and S4 (equilibrium phase) of DCE-MRI are better than other data in distinguishing the degree of HCC differentiation. However, the late phases of DCE-MRI, such as S2 (late arterial phase), S3 (portal phase), S4 (equilibrium phase) and SS (delayed phase), have a higher recall rate (recall) in judging the differentiation degree of HCC. In order to combine the advantages of multiple sets of MRI data, make use of the complementarity of different modal data, and find the best data fusion scheme to distinguish the differentiation degree of HCC, it is necessary to compare multiple groups of experiments. There are 56 different combination ways to choose 3 modes from 8 modes, and it is too complicated to carry out experiments under all combination. Based on the results and doctor’s experience, we listed the experimental results of T211 * (fusion of T2WI, T1WI/ IN, T1WI/OUT), T2S24 (fusion of T2WI, S2, S4), 5024 (fusion of S0, S2, S4) and T21 * S4 (fusion of T2WI, T1WI/OUT, S4) to distinguish the degree of differentiation of HCC.

With the increase of the coding dimension of features, the capacity of the model is also increasing, so the evaluation indexes such as the accuracy of the model are also increasing, but the overall performance of the model tends to decline when it reaches a certain value. This may be related to the current lack of data, and the over-fitting phenomenon appears in the classification layer in the process of increasing the number of feature codes. When the feature coding dimension is 32, the model has the best effect on judging the differentiation degree of HCC. Furthermore, in order to more intuitively observe the effect of different networks on extracting multimodal MRI image features, we can directly draw the visual results of the network trained on the multimodal MRI data set encoding the test set data into two-dimensional features.

**D. DESIGN OF MULTI-SCALE DEEP RESIDUAL NETWORK MODEL BASED ON DEEP LEARNING METHOD**

The whole network consists of two deep learning modules: feature learning module and similarity comparison module, and two processing modules: data preprocessing and feature splicing. Among them, the data preprocessing module takes the doctor’s experience of reading films as a reference, preprocesses all the original data, and then sends the processed data into the feature learning module. After coding the data, the feature learning module splices the query samples (unknown categories) with the (poorly, moderately, well) features of the training samples with known differentiation categories in the form of sample pairs. Finally, the relational comparison network is used to score the similarity of the spliced data, and the differentiation category of the samples with unknown differentiation category is determined according to the scoring results.

In the conventional training of supervised feature representation learning network, there are a variety of loss functions to choose, such as using softmax classifier to get the class probability, then using cross loss to calculate the loss, training feature extraction network, cross loss calculation method. In the metric learning model, based on the constructed image classification model based on metric learning, the feature extraction network is used to calculate the features of sample pairs. Taking into account the metric learning, only considering the consistency of the same kind of sample pairs, and does not bring more influence on the increase of the difference between classes, this paper further uses the metric learning method of establishing the same kind and different class triple, on the basis of feature extraction, the Triplet loss function guides the optimization learning process of the feature representation network, which is guaranteed under the condition of small samples. The extracted features not only have a good ability to depict similar images, but also have a good ability to distinguish different kinds of images. Computer aided diagnosis technology based on medical images was always a hot spot in the field of computer applications. The urgent necessity of controllability and practical clinical diagnosis of acquiring medical image data brings breakthrough progress to the detailed learning landing. So it shows non invasive evaluation of hepatocellular carcinoma differentiation based on multi parameter nuclear magnetic resonance imaging. Based on this, multichannel 3D convolution neural network and multiscale depth residual network are proposed to extract features of three-dimensional medical image data
and two-dimensional fusion medical image data. The role of metastasis learning and metric learning in the classification of medical images is verified.

In order to verify the effectiveness of the metric learning method and its application ability in the feature representation learning of medical images, three different loss functions are used to train the feature extraction module. Visual comparative analysis experiments are carried out on the distribution of the features extracted by the model in the feature space. The method of difficult sample mining is widely used in machine learning and deep learning, such as support vector machine training, only support vector can be the final interface of the training effect. From the graph, we can see that softmax extracts training effect.

The work. From the graph, we can see that the features extracted by softmax also diverge around with the origin as the center, and the distribution is basically similar. Compared with the distance between the centers, the samples distributed on the boundaries between categories are more sparse, which can produce certain results in using metric learning methods to constrain the feature extraction process. However, due to the great differences between individuals with real clinical data, the use of the framework to achieve medical image classification has not achieved a better classification effect, and it is necessary to further collect data and adjust training strategies.

IV. DISCUSSION

Computer aided diagnosis technology based on medical images was always a hot spot in the field of computer applications. The urgent necessity of controllability and practical clinical diagnosis of acquiring medical image data brings breakthrough progress to the detailed learning landing. The paper mainly discusses non invasive evaluation of hepatocellular carcinoma differentiation based on multi parameter nuclear magnetic resonance imaging. Based on this, multichannel 3D convolution neural network and multiscale depth residual network are proposed to extract features of three-dimensional medical image data and two-dimensional fusion medical image data. The role of metastasis learning and metric learning in the classification of medical images is verified.

In this paper, we constructed a multi-parameter MRI medical image data set for judging the differentiation degree of hepatocellular carcinoma. First of all, according to the research results of scholars at home and abroad and the clinical diagnostic experience of doctors, this paper decided to take multi-parameter magnetic resonance imaging as the research object and cooperate with many radiologists in the radiology department of the hospital. The clinical image data of multi-parameter magnetic resonance imaging which has been pathologically confirmed as hepatocellular carcinoma were collected and labeled, and the medical image data set for training machine learning model was constructed. The specific data include: T1WI inverse phase, T2WI, dynamic contrast enhanced imaging (DCE-MRI) and other modal data, in which the data of each mode is 3D stereoscopic data. According to the experimental feedback results, a more
effective inspection scheme and image acquisition method for deep learning diagnosis performance improvement are proposed. In the aspect of data preprocessing, the methods of data enhancement and data sample resampling are used to overcome the problems of data shortage and data category imbalance.

In this study, a three-dimensional convolution neural network architecture based on multi-channel fusion is proposed, which is used to extract the spatio-temporal features of DCE-MRI images. The multi-channel fusion 3D convolution neural network is based on 3 DCNN. Aiming at DCE-MR images, a tensor-based data representation model and a data fusion model are proposed to explore the effects of extracting time series information and spatial texture information from DCE-MRI images with different data representations and network structures. At the same time, two kinds of network structures are designed: the 3DCNN with multi-channel input and the fusion network using feature splicing strategy to extract the temporal and spatial features of DCE-MRI images to the maximum extent and improve the effect of auxiliary diagnosis. Through comparative experiments, it is found that the extraction of D-CE-MRI time series information is more helpful to distinguish the background of HCC and liver cirrhosis, and to evaluate the degree of HCC differentiation. Compared with 3 DCNN and ordinary data representation methods, using multi-channel converged network architecture and tensor data representation method can improve the accuracy of evaluating HCC differentiation by 7.3%.

This paper explores the role of data fusion, transfer learning and multi-scale features in medical image assistant diagnosis, and verifies the effects of multi-modal data fusion, multi-scale feature extraction and transfer learning on the constructed multi-parameter MRI image data sets. The experimental results show that when using multiple modal fusion decisions, it is necessary to select and find complementary modal data for fusion can get better results; although there are great differences between natural images and medical images, the pre-trained model on the natural image data set as the initialization of the network can ensure and accelerate the convergence of training and improve the performance of the model on the test set. Because the size of the batch training data needs to be adjusted before training the convolutional neural network, yes, the data size is the same in this batch of data, so the original scale information of the data will be lost. However, the scale information has a certain reference significance for tumor diagnosis, and the model extraction of multi-scale features is helpful to improve the effect of medical image classification. Combining the methods of multi-modal data fusion, multi-scale feature extraction and transfer learning, compared with the multi-channel fusion three-dimensional convolution neural network, the differentiation degree of HCC is judged, and the performance of the evaluation is further improved, in which the accuracy is improved by more than ten percentage points.

With the help of the idea of measurement learning, the problem of medical image classification is transformed into the problem of measuring the degree of similarity between medical image samples. To explore the effect of measurement learning method in the case of lack of medical image training data. Here are also some shortcomings for this method. The measurement learning method can not be used to guide the feature extraction process of medical images and reduce the inter individual differences in clinical data and overcome insufficient sample size problems.

V. CONCLUSION

In this paper, there are still many deficiencies in the non-invasive evaluation of the differentiation degree of hepatocellular carcinoma based on multi-parameter MRI images, which need to be further supplemented and improved in the future research, mainly reflected in the following aspects. In the construction of medical image data set, it is necessary to continuously mark more image data to train the auxiliary diagnosis model, which has achieved better results. In the aspect of extracting the multi-scale features of the tumor region, the spatial pyramid pool without increasing the number of parameters is used to extract the features. This method only uses the high-level semantic features of the network and has some limitations. In the future, we can consider using the feature pyramid network which is effective in target detection and pedestrian re-recognition to extract multi-scale features under high-level and low-level fusion. In the aspect of eliminating the differences between individuals of clinical data samples, we can consider the use of data preprocessing methods to reduce the difficulty of measurement learning.

REFERENCES

[1] A. Hughes, A. Moreno, and T. H. My, “Which destination is smarter? Application of (SA6) frameworks to establish a ranking of smart tourist destinations,” Int. J. Inf. Syst. Tourism, vol. 4, no. 1, pp. 19–28, May 2019.
[2] X. Yang, L. Lin, P.-Y. Cheng, X. Yang, Y. Ren, and Y.-M. Huang, “Examining creativity through a virtual reality support system,” Educ. Technol. Res. Develop., vol. 66, no. 5, pp. 1231–1254, Oct. 2018.
[3] J. D. Hemanth, U. Kose, O. Deperlioglu, and V. H. C. de Albuquerque, “An augmented reality-supported mobile application for diagnosis of heart diseases,” J. Supercomput., vol. 76, no. 2, pp. 1242–1267, Feb. 2020.
[4] C. E. Mora, J. Martín-Gutiérrez, B. Añorbe-Díaz, and A. González-Marrero, “Virtual technologies trends in education,” EURASIA J. Math., Sci. Technol. Educ., vol. 13, no. 2, pp. 469–486, Jan. 2017.
[5] N. Chung, H. Lee, J.-Y. Kim, and C. Koo, “The role of augmented reality for experience-influenced environments: The case of cultural heritage tourism in Korea,” J. Travel Res., vol. 57, no. 5, pp. 627–643, May 2018.
[6] A. Milosavljević, D. Rančić, A. Dimitrijević, B. Predić, and V. Mihailović, “Integration of GIS and video surveillance,” Int. J. Geograph. Inf. Sci., vol. 30, no. 10, pp. 2089–2107, Oct. 2016.
[7] S. C. H. Li, P. Robinson, and A. Oriade, “Destination marketing: The use of technology since the millennium,” J. Destination Marketing Manage., vol. 6, no. 2, pp. 95–102, Jun. 2017.
[8] S. Liu, C. I. Seresinhe, Y. Shen, and M. Gutierrez-Rojo, “Street-frontagenet: Urban image classification using deep convolutional neural networks,” Int. J. Geograph. Inf. Sci., vol. 34, no. 4, pp. 681–707, Apr. 2020.
[9] J. M. Haut, M. E. Paolotti, J. Plaza, A. Plaza, and J. Li, “Visual attention-driven hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 10, pp. 8065–8080, Oct. 2019.
[10] X. Liu, R. Zhang, Z. Meng, R. Hong, and G. Liu, “On fusing the latent deep CNN feature for image classification,” World Wide Web, vol. 22, no. 2, pp. 423–436, Mar. 2019.
