Application of Medical Brain CT/MRI Image Fusion Algorithm based on Neural Network

Dan Yang*
School of Software and Big Data, Changzhou College of Information Technology, Changzhou, China

***Abstract*—In recent years, fused images have been developed for fast processing of medical images, which provide a more reliable basis for reducing the burden on physicians because they can contain multiple times the image information. In order to achieve fast and accurate recognition results in medical image recognition, avoid similar blocks and shadow fitting in CT/MR fusion images, and improve the entire medical system, in this study, CT/MRI image fusion of brain images is studied based on algorithms generated by Convolutional Neural Network (CNN). The study utilizes Rolling Guidance Filter (RGF) to divide medical CT/MRI images into two parts, one of which is used for model training and the other for image fusion. In the experiments, the results of all three experiments are compared with the Nonsub Sampled Contourlet Transform - Piecewise Convolutional Neural Network (NSCT - PCNN), and the CNN-RGF MI/ IE/SSIM/AG values of CNN-RGF are superior compared to the conventional algorithm of NSCT-RCNN with an average improvement of 10.0% and above, and the resulting CNN-RGF observed meningitis, hydrocephalus, and cerebral infarction with an average of 24.8% higher compared to NSCT-RCNN. The outcomes show that for brain image fusion and detection, the CNN-RGF approach put forward in the study performs better.**

Keywords*—Convolutional neural network; image; integration; CT; MRI

I. Introduction

In recent years, medical imaging has been rapidly developing as it has started to be involved in disease diagnosis and widely used in clinical treatment. As the amount of comprehensive information increases, medical images using a single modality mode gradually fail to meet the needs of physicians. The information provided by traditional medical images can be too one-sided, and there are many tissues and organs in the human body, so doctors using the naked eye to identify the images will inevitably produce eye fatigue, thus affecting the accuracy of diagnosis [1, 2]. Moreover, the development of medical devices is really backward, and the blurriness of imaging is sometimes even lower than the resolution of the human eye, so it is necessary to consider the fusion of multiple images together. Through a method to aggregate multiple medical images into one, not only can the useful information be concentrated into one, which improves the image utilization rate, but also can reduce the amount of images for doctors to see, which is convenient for doctors to locate the precise lesion and target medication to patients, and perhaps treat more difficult and complicated diseases [3]. Therefore, research on fusing multiple images together is imminent. Recently, a classical algorithm, Convolutional Neural Network (CNN), has come into the view of researchers and become popular in image fusion. The CNN has powerful feature extraction capability, and the Rolling Guidance Filter (RGF) is able to handle the similarity blocks and anaglyph of fused images well [4]. Based on the advantages of both, this study establishes a CNN-RGF algorithm to fuse medical images at pixel level considering four plain objective quantities after CT/MRI fusion images. The aim is to achieve fast and accurate recognition results in medical image recognition, and to avoid similar blocks and shadow-fitting CT/MRI fused images. This will reduce the processing burden of doctors, improve their efficiency, and thus improve the whole medical system. The main contribution of the research is to extract and fuse different image information of the same target from different angles, levels, or types of sensors. At the same time, the low transparency information in the image is processed through image denoising, enhancement, and other image processing techniques, thereby significantly improving the accuracy, restoration, and reliability of the image, and providing clearer and more accurate expression for the generation of target images a fused image with complete content and rich image information. The innovation of the research lies in the use of a pyramid based multi-scale image decomposition method, which enables fusion at each decomposition level. Each source image is decomposed through a regional Laplacian pyramid, making the image features more distinct. Therefore, this method plays an important role in medical image fusion.

The research structure is mainly divided into four parts. The first part is a summary of relevant research on medical image fusion at home and abroad. The second part is to build a brain CT/MRI image fusion model based on CNN, and introduce the specific improvement process and research of the algorithm. The third part is to analyze the performance of the constructed model, reflecting its performance through indicators such as accuracy and error. The fourth part is a summary and analysis of the research, discussing the achievements and shortcomings of the research, and proposing suggestions for future research directions.

II. Related Works

A very important branch of image processing technology, i.e., medical image fusion, has a very important role in doctors' rapid treatment of patients, targeted drug administration, etc. Wang et al. [5] studied multi-feature fusion in depth and proposed a medical brain image algorithm based on it. Texture information was obtained by feature extraction of CNNs, and morphological features were obtained by feature extraction of voxel information. These two types of features were concatenated and then the feature selection stage was

optimized using a heuristic search algorithm. They analyzed experimentally to select the optimal values of the parameters based on the heuristic search and extracted the optimal feature subset after determining the parameter values. Finally, the algorithm improved the accuracy and efficiency of brain image classification compared to similar algorithms. Polinati et al. proposed a new method for medical image fusion, incorporating content decomposition and sigmoid function [6]. They considered and implemented the use of empirical wavelet transform for content-based decomposition for preserving edges and corner points. They discovered that using detail layer fusion directly results in significant artefacts, so they used the sigmoid function to improve weight scaling. They tested their suggested method with previous fusion methods after fusing 24 pairs of MRI-PET and MRI-SPECT pictures, and they discovered that both the qualitative and quantitative outcomes had significantly improved. By first filtering the CT and MR image sets through a set of various scaled filter sets, different pairs of representations of CT and MR were obtained. Each pair of different representations was then used to train the corresponding CNN to obtain the final fused image, and it was compared with nine recent state-of-the-art multimodal fusion methods. Wang et al. [7] proposed a fusion method based on a multi-CNN combination of fuzzy neural networks. The experimental findings demonstrated that in objective evaluation and visual quality, the fusion approach greatly exceeded other comparative fusion methods. The method excelled in four measures, enhanced multimodal medical picture fusion quality, and helped doctors diagnose diseases more accurately. A brand-new picture fusion technique based on sparse representation was proposed by Yu et al. [8]. They studied that after merging all source images into a joint matrix and training it by an algorithm, an overcomplete coefficient would be obtained that can be used to represent this matrix. The obtained over-completeness was used as coefficients of the image features and combined with choose-max fusion rules. The fused images were reconstructed from the connected coefficients and the overcomplete dictionary and compared with the conventional algorithms. They found that the method had better fusion performance compared to three state-of-the-art algorithms.

Using the non-subsampled shear wave transform (NSST), smooth wavelet transform, and impulsive coupled neural network, Singh and Gupta suggested a multilevel multimodal fusion model [9]. A weighted Laplace pyramid was used to extract structural features from the source image and apply them to an adaptive model that can map the feature weights used for low-band component fusion using absolute maxima and absolute differences, a rule that allows fusion of high-frequency NSST components to preserve complex directional details. The first step was to use NSST to decompose the source image into optimal sparse multi-resolution components. The strategy, when compared to previous methods, dramatically improved medical picture fusion with good visual quality and improved computational metrics, according to experimental results. The non-subsampled contour wave transform (NSCT) domain image fusion approach was proposed by Yu et al. and is based on pulse-output neural networks (PCNN) and hybrid frog-leaping algorithms (SFLA) [10]. First, the source image was decomposed into low-frequency and high-frequency subbands using NSCT, and secondly, different PCNN fusion rules were designed. Finally, the fused images were reconstructed by inverse NSCT. The fused image preserved more of the original image's information with strong edge retention, according to a visual and quantitative examination of the experimental results. Guan et al. proposed an image fusion algorithm based on multi-scale analysis coupled with approximate sparse representation to better deal with the singularity of high-dimensional features of images and to take into account the fusion of image target features and average intensity information [11]. The high-frequency and low-frequency information of the image was obtained by the scale analysis of the source image, and the specific target detail information was highlighted. The approximate sparse representation was designed to approximate the singular curve with the smallest coefficients. A decision mapping was constructed to analyze the activity and matching degree of all coefficients on the same subband and output the decision values, which were used to match and fuse the images. Then the final fused image was obtained by multi-scale inverse transform. The experimental results showed that better visual effects can be obtained with high robustness and wide application.

Multiple researchers have found that CT/MRI image fusion algorithms are very popular internationally and have achieved relatively successful data in experiments, with an overall success rate of over 80% for image fusion [12-16]. Although image fusion has made some progress, its effectiveness still has a significant room for improvement. Research has found that there is relatively little research on using CNN algorithms to form composite neural networks in image fusion. Therefore, combining CNN and RGF can leverage their respective advantages, compensate for the shortcomings of individual algorithms, and perfectly avoid their own shortcomings. The research aims to further improve the effectiveness of medical image fusion.

III. CNN-BASED BRAIN CT/MRI IMAGE FUSION STUDY

A medical image fusion method based on pyramid and CNN is proposed. By using multi-scale decomposition of pyramids that are more conducive to human visual perception, the fusion effect is improved. At the same time, the idea of support vector machine (SVM) is used to improve the CNN network, which does not rely on empirical initialization parameters and effectively extracts image features to obtain more suitable weight maps. The pooling and sampling layers in traditional CNN networks is removed to reduce the loss of image information.

A. Application of Multi-Scale Geometric Transform in Image Fusion Algorithm

The research on the overall framework of medical image fusion proposes a medical image fusion algorithm that can be summarized into the following four steps. The first step is to input the source image into an improved CNN and generate a weight map. The second step is pyramid decomposition, which uses the multi-scale image decomposition method of the pyramid to fuse at each decomposition level. Each source image is decomposed through the regional Laplace pyramid. The third step is coefficient fusion. The fourth step is the

reconstruction of the Laplace pyramid. The specific framework diagram is shown in Fig. 1.

Computed Tomography (CT) is an important tool for diagnosing lesions because of its rapid scanning capability, while Magnetic Resonance Imaging (MR) has the strongest resolution of soft tissue and can observe lesions without dead space [17-19]. By combining CT/MRI together, the accuracy of unimodal medical images can be greatly improved. For two images of the same target, the levels after fusion can be divided into three kinds, as in Fig. 2.

From Fig. 2, the fusion of images can be divided into three levels: pixel, feature and decision. Among them, direct fusion from the original image is called pixel fusion; feature extraction of the original image once and then fusion is called feature fusion; feature extraction of the image that has been extracted once and then fusion is called decision fusion, implying that a decision can be made directly from the decision fused image. The contribution of decision fusion to CT/MRI of the brain is many, including but not limited to the timely detection of lesions, saving the time of doctors and providing a possibility of cure for patients. In order that no one will suffer, this study investigates the method of brain CT/MRI image fusion, which will be discussed in detail next. Independent individual neurons have simple structures, but neural network systems composed of large numbers of neurons are rich in behavior. The relationships between neurons are intricate and complex. The expressions of neurons are shown in Equation (1) [20].

$$\begin{cases} y_l = f\left(\sum_{k=0}^{N_2-1} \omega''_{kl}\, x_k'' - \theta_l''\right) \\ x_k'' = f\left(\sum_{j=0}^{N_1-1} \omega'_{kj}\, x_j' - \theta_k'\right) \\ x_j' = f\left(\sum_{i=0}^{N-1} \omega_{ij} x_i - \theta_j\right) \end{cases} \tag{1}$$

In the above Equation (1), the nonlinear function is denoted as f; two neurons are defined as i, j; then the link between them is called ω_{ij} and its threshold is called θ. To reflect how much valid information is contained in the image, Mutual Information (MI) is chosen to judge the size of information data in the image. It supposes that there exist (X,Y) as random variables and their distribution is jointly located at $p(x,y)$, then $p(x)p(y)$ is called the edge of the distribution, then their relationship is as denoted in Equation (2).

$$I(X,Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \tag{2}$$

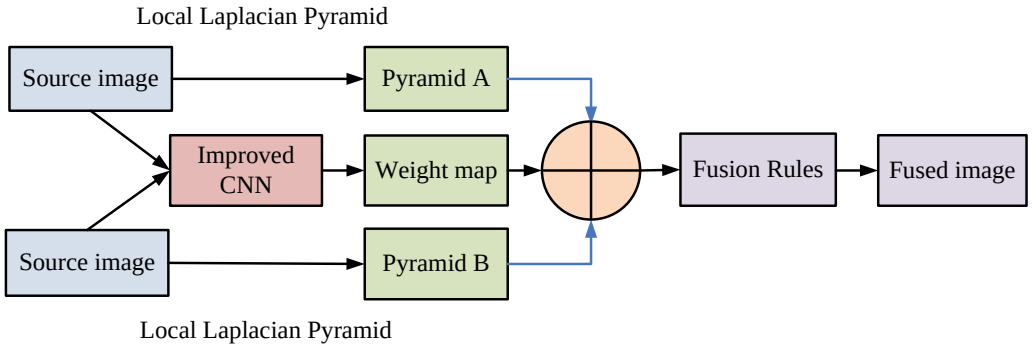

Fig. 1. Overall framework of medical image fusion process.

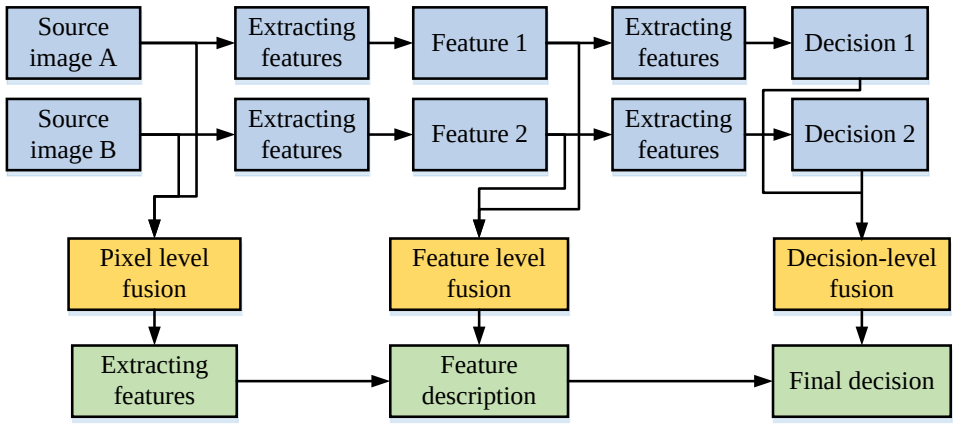

Fig. 2. Three different levels of image fusion.

For problems involving brain images, it is not enough to discriminate the amount of data by MI alone, but also requires the complement of Information Entropy (IE). The smaller the IE, the more residual the image is, as indicated in Equation (3).

$$H(x)=E(I(x_i))=E\left(\log\left(2,\frac{1}{P(x_i)}\right)\right)$$
$$=-\sum P(x_i)\log(2,P(x_i))(i=1,2,\ldots,n) \quad (3)$$

CNN is the most commonly used network model in medical field. A complete CNN mainly consists of input, convolutional, pooling, fully connected and output layers. Among them, the convolutional layer is the core of CNN and is also the source of the name of CNN. It assumes that ω_i represents the full-valued vector, the convolution operation is noted as $\otimes$, and the activation function is called $\otimes$, the operations in the convolution layer are as expressed in Equation (4).

$$H_i=f(H_{i-1}\otimes\omega_i+b_i) \quad (4)$$

To construct multi-resolution images, the concept of Local Laplacian Pyramid (LLP) is also introduced. LLP has the power to not only accurately distinguish between the edges and textures of an image, but even to fuse the image with the tower layer. Due to many updates, LLP has never had any artifact problems. Before performing LLP operation on an image, a Gaussian Pyramid (GP) decomposition is performed, as in Equation (5).

$$\begin{cases} G_l^*\left(\frac{i+m}{2},\frac{j+n}{2}\right)=\begin{cases} G_1\left(\frac{i+m}{2},\frac{j+n}{2}\right) & \frac{i+m}{2},\frac{j+n}{2}\in Z \\ 0 & Elses \end{cases} \\ G_l^*(i,j)=4\sum_{m=-2}^{2}\sum_{n=-2}^{2}w(m,n)G_1\left(\frac{i+m}{2},\frac{j+n}{2}\right) \end{cases} \quad (5)$$

In Equation (5), G_l^* is the image after GP processing, and i,j means the current GP layer number. The image after undergoing GP decomposition cannot obtain the risk in the evolutionary process, such as the empirical risk, but also the information is lost [21]. To avoid information loss and at the same time deepen the impression of information in the network, the residual learning module is built according to Equations (3) to (5) as in Fig. 3.

In Fig. 3, the input primitive image can finally obtain an optimized image of size 6*6*256 after first undergoing C-level evolution. Then two steps of weighting are performed, and after passing it, it can be input among LLP, which is calculated as Equation (6).

$$\begin{cases} O=collapsywhane\left(S_i\left[I'(v)\right]\right) \\ I'(v)=\begin{cases} g+sign(v-g)\sigma_r\left(|v-g|/\sigma_r\right)^{\alpha} & |v-g|\le\sigma_r \\ g+sign(v-g)\left(\beta\left(|v-g|-\sigma_r\right)+\sigma_r\right) & else \end{cases} \end{cases} \quad (6)$$

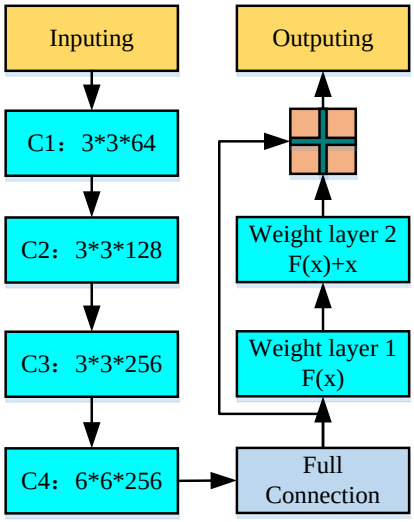

Fig. 3. Residual learning module.

In Equation (6), the image output by LLP is called O, and the image of each layer is noted as S_i. After a set of processes in LLP, the coefficients due to the decomposition are obtained, which are denoted as v, then $I'(v)$ is also called the standard function based on the decomposition coefficients. The reconstruction operator is then denoted as *collapsywhane* and the pixel value obtained from the LLP can be called g. In addition, α,β,σ_r are three variable parameters of LLP, which are intensity threshold, detail factor and ranging factor. The intensity threshold serves as a boundary to distinguish edges from details; the detail factor and the range factor control the enhancement and reduction, respectively, one controlling the details and the other controlling the range. After the GP and LP, the signal analysis is performed. The signal analysis tool is the well-known Multiscale Geometric Analysis (MGA), whose flow is shown in Fig. 4.

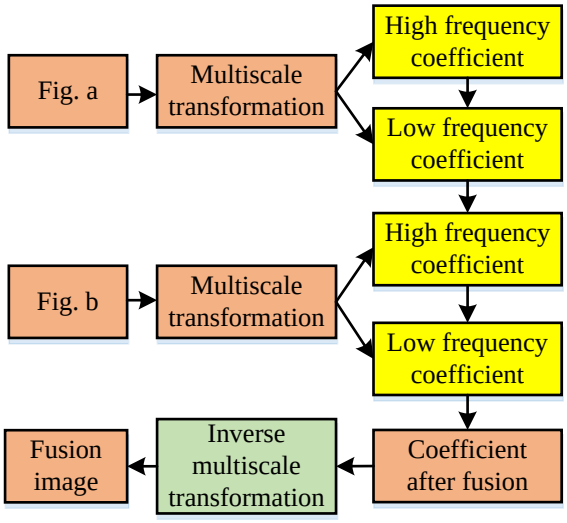

Fig. 4. Multi-scale geometric transformation image fusion.

In Fig. 4, the original image is first multi-scale transformed, which will result in a decomposed source image, and the fusion rule is applied to the decomposed image to fuse high-frequency coefficients or low-frequency coefficients, as appropriate. Then the fused coefficients are inverse multi-scale transformed, and finally the fused image is recombined to form [22]. For the fused image, the gap between the distortion and the original image is contacted, i.e., the Fidelity of Visual Information (VIFF) is calculated, as in Equation (7).

$$VIFF_k(A,B,F)=\frac{\sum_b FVID_{k,b}(A,B,F)}{\sum_b FVIND_{k,b}(A,B,F)} \quad (7)$$

In Equation (7), the visual information degree of distortion or not is expressed by $FVID, FVIND$, respectively. A,B,F denotes the degree of information of the pixel and b denotes the value of the coordinate at which it is located.

B. MR/CT Medical Image Fusion based on the Combination of CNN and Rolling Guide Filtering

RGF has the ability to guarantee smooth details even under scale measurement. RGF works iteratively and converges particularly fast; RGF works over a wide range, and small structure removal and edge restoration are its features, which are well suited for medical image studies. If it assumes that I is the input image, then G represents the output image, the standard deviation of the Gaussian filter is noted as σ_S^2 and the pixel index can be expressed by p,q as in Equation (8).

$$G(p)=\frac{1}{K_p}\sum_{q\in N(p)}\exp\left(-\frac{\|p-q\|^2}{2\sigma_r^2}\right)I(q) \quad (8)$$

In Equation (8), when $K_p=\sum_{q\in N(p)}\exp\left(-\frac{\|p-q\|^2}{2\sigma_S^2}\right)$ is used, it means that it is available for normalization. When the RGF takes another approach to recover the edges, that must be the joint RGF iteration, at which time the source output of the filter is represented by J^t. When the filter iterates to t times, the output at that time is represented by J^{t+1}, and the relationship between them is as in Equation (9).

$$J^{t+1}(p)=\frac{1}{K_p}\sum_{q\in N(p)}\exp\left(-\frac{\|p-q\|^2}{2\sigma_S^2}-\frac{\|J^t(p)-J^t(q)\|^2}{2\sigma_S^2}\right)I(q) \quad (9)$$

In Equation (9), the weight range is controlled jointly using I,σ_r^2. The CNN model is testing the activity level metric while using a huge number of photos to train its data and create adaptive fusion rules. CNN is able to greatly reduce the difficulty of designing fusion rules because image fusion with CNN is more efficient than manual design, as shown in Fig. 5.

From Fig. 5, a normal image block size is 16*16, and after a nonlinear mapping, it gets 64 feature maps of size 16*16, which should be processed with special care to prevent information loss. 64 images undergo another nonlinear mapping, and then compression in the image, which can get 128 refined images of 8*8. The refined image can also undergo a final expansion to obtain 256 8*8 results. This is already the limit of convolutional kernel, if it is too large, it is easy to cause information loss; if it is too small, the feature extraction is not obvious enough [23-25]. For medical class images, fluctuations in other local regions can be better characterized, so a new parametric max-min filtering algorithm is introduced, calculated as in Equation (10).

$$\bar{I}(i,j)=\max_{(m,n)\in\Omega}(I(m,n))-\min_{(m,n)\in\Omega}(I(m,n)) \quad (10)$$

In Equation (10), the original image is recorded with I, the center of I is noted as Ω and (i,j) is any point. $\max_{(m,n)\in\Omega}(I(m,n))$ and $\min_{(m,n)\in\Omega}(I(m,n))$ are the max and the mini filters. Images sometimes have similar shared blocks, which can be established in Equation (11).

$$\eta_q=\|P_q-P_r\|_{E_w} \quad (11)$$

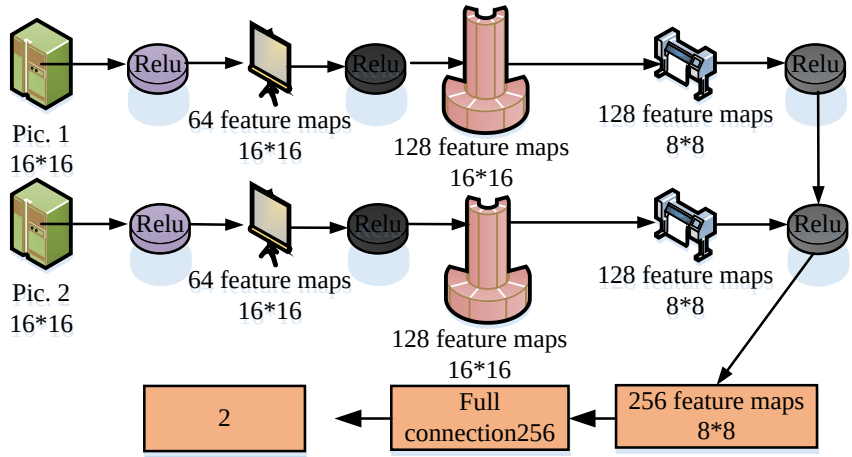

Fig. 5. CNN structure.

In Equation (11), η_q represents the Euclidean distance of the similar parts between the shared blocks. The given reference block is noted as P_q, and the candidate block is defined as P_r. The candidate blocks are to avoid gradient reciprocity and gradient destruction, and also to be adaptive to the weights that appear. The fusion of the input graph is set according to the known judging criteria, as in Fig. 6.

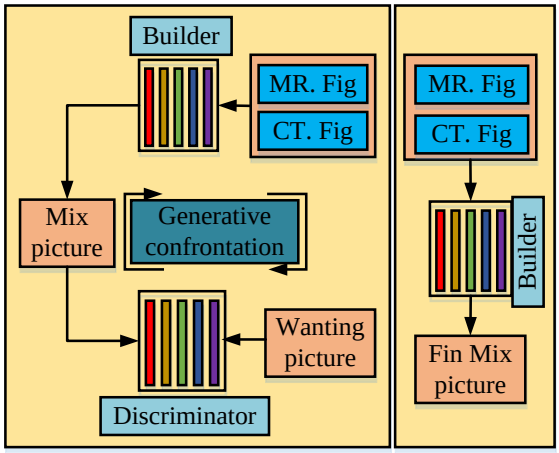

Fig. 6. Image fusion process.

From Fig. 6, if the detail image has the feature of local similarity, then the block can be cut into a large number of square blocks of equal size, called Shared Similar Block (SSB), which is defined by Equation (12).

$$L_W^S(r) = L_{W^A}(r) \cap L_{W^B}(r) \tag{12}$$

In Equation (12), the image A,B has k similar blocks to the image L_{W^A}, L_{W^B}, respectively, and the SSB to be calculated is $L_W^S(r)$. The SSB is somewhat different from the traditional perceptron, as reflected in the optimization of the fixed denominator, and the optimization function is as in Equation (13).

$$\min \frac{\|W\|^2}{2}, y_i\left(W^T x_i + b\right) \geq 1 (i = 1 - N) \tag{13}$$

In Equation (13), $\min \frac{\|W\|^2}{2}$ is defined as the large distance from all points to the shared plane, and y_i implies a regression analysis that can seek the optimal solution for learning ability and distance. For human vision, the color and state of the image are very sensitive, so a small deficiency in a key location can cause a large change. For medical matters, the larger the area of the image is the more informative it is. Medical imaging characteristics can prove that the largest and smallest pixel difference reflects important information. So for the study of images, the first step is to perform a pyramid decomposition so that the fusion applies to each level, as in Fig. 7.

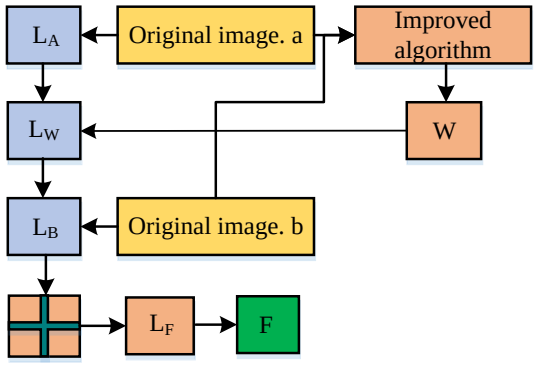

Fig. 7. Improved algorithm for image fusion.

In Fig. 7, the original images A, B and W are input to the algorithm based on the combination of rolling filter and CNN to obtain three initial images L_A, L_B and L_W. The obtained three images are input into the Laplace pyramid and then feature fusion operation is performed to be able to output the final image F. The risk in this is mainly in two aspects, containing empirical risk and structural risk, to be considered simultaneously. The empirical risk is to be controlled by the generalization of the CNN, and the study is to control both time and geographic location because it is most influenced by environmental factors and other factors are negligible [26-28]. Structural risk reduction is to be achieved by dimensionality reduction of the CNN, as reflected in the histogram of the probability distribution of the input. The function mapping is then performed by hidden neurons, and assuming that the filter convolves the image support values and the standard convolution kernel of step size is maximally pooled, Equation (14) can be obtained.

$$c_{i,j,o}(\psi) = f\left(\sum_{h=-0.5k}^{0.5k}\sum_{w=-0.5k}^{0.5k}\sum_{u=1}^{N}\theta_{h,w,u,o} \bullet \psi_{g(h,w,i,j,u)}\right) \tag{14}$$

In Equation (14), the kernel weights are represented by θ; u denotes the scale of the maximum pool operation; (i, j) is the coordinate; $f(\bullet)$ represents the activation function of Relu, which is chosen to represent the elements of the previous layer of feedback for this operation.

IV. Analysis of Medical MR/CT Image Fusion Model Based on CNN

A. Determination of Model Parameters for Fused CNN-RGF

For this study, the dataset from The First Affiliated Hospital of Harbin Medical University was used, and images of normal brains as well as common disease brains, such as brain atrophy, were selected. The experimental environment, i.e., the parameters, is shown in Table I.

For the image processing, the image was first normalized and then compared with NSCT - PCNN for comparison, as shown in Fig. 8 [29, 30].

TABLE I. EXPERIMENTAL PARAMETERS

GPU	Internal storage	Operating system	Channel output
NVID Tesla M60	256GB*2	128Ubnutu 21.02.20	64; 128; 256
Flash memory	Operator	Input	Filter parameters
CUDA Too kit 23.0	Python	3.0	3*3; Floor2
Step length	Number residual blocks	Display card	CPU
1	32	Tensor flow 2.40	Windows X10

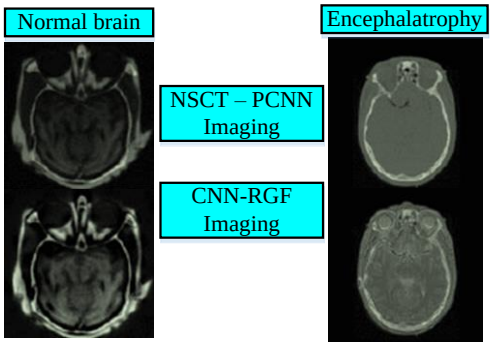

Fig. 8. Comparison chart of CT/MRI fusion effect.

From Fig. 8, the brain luminosity of NSCT - PCNN imaging was not as bright as CNN-RGF, and even in the processing of details, it was obvious that CNN-RGF was more carefully discriminated. Even though NSCT - PCNN was more economically cost effective, it was more important to be medically rigorous. Then the dataset had to be trained iteratively as well as with error training, as in Fig. 9.

From Fig. 9, the accuracy of CNN-RGF was lower than that of NSCT-PCNN until 200 iterations, but the error rate was higher than that of NSCT-RCNN with the increase of the iteration times. However, when the iteration times reached 200 or more, the accuracy of CNN-RGF was unmatched by NSCT-RCNN. Although the increase in the iteration times decreased the operational efficiency of the algorithm, the accuracy also increased with the iteration times. Since the recognition of brain images was important for brain diseases, the accuracy of image recognition was more important, and based on this, the CNN-RGF algorithm proposed in the study had more significant advantages. The corresponding results in Fig. 9 are shown in Table II.

In Table II, the highest accuracy value of CNN-RGF was 0.94, the average value was 0.85, and the error was 0.13; The highest accuracy of NSCT-PCNN was 0.81, with an average of 0.69 and a deviation of 0.18. The minimum error value of CNN-RGF was 0.52, the average value was 0.64, and the deviation was 0.11. The highest accuracy of NSCT-PCNN was 0.71, with an average of 0.86 and a deviation of 0.13. The results indicated that the proposed CNN-RGF model had higher accuracy and stability.

B. Experimental Data Validation based on CNN-RGF Model

To make the results more generalizable and applicable to all hospitals, this study provided a high-level evaluation of the results based on several common metrics. These were MI, IE, Structural Similarity (SSIM) and Average Grads (AG). The values obtained from the above four box indicators belonged to dimensionless values and were mainly used for comparison. Three common brain diseases were studied based on the four metrics, and the generated results were compared using NSCT-RCNN with CNN-RGF, as shown in Fig. 10.

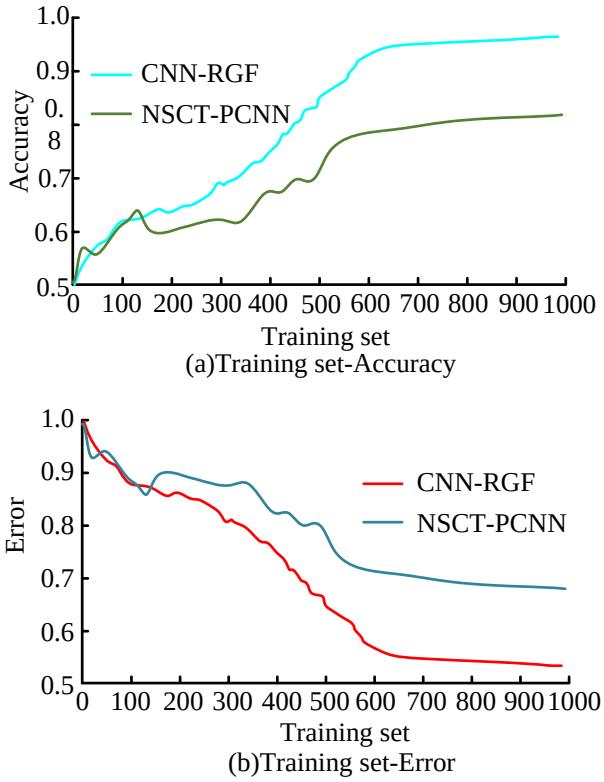

Fig. 9. Iterative training and error training of datasets.

TABLE II. PERFORMANCE MEAN AND STANDARD DEVIATION OF MODEL TRAINING

Algorithm	Precision			Deviation		
	Highest Value	Average Value	Standard Deviation	Minimum	Average Value	Standard Deviation
CNN-RGF	0.94	0.85	0.13	0.52	0.64	0.11
NSCT-PCNN	0.81	0.69	0.18	0.71	0.86	0.13

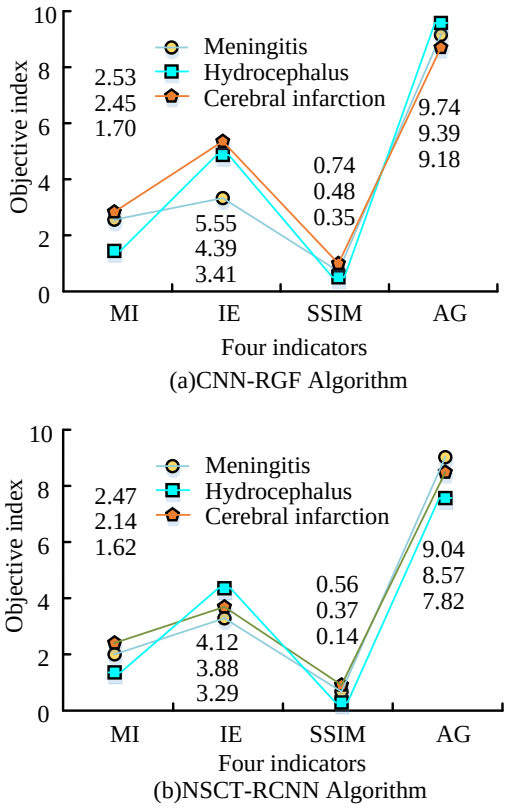

Fig. 10. Comparison of three kinds of encephalopathy and four indexes by using CNN-RGF (a) and NSCT-RCNN (b).

From Fig. 10, firstly three brain diseases, including meningitis, hydrocephalus and cerebral infarction, were characterized in CNN-RGF which was higher than NSCT-RCNN. In other words, the image quality obtained by fusion was better when CNN-RGF was used to characterize brain diseases. The CNN-RGF algorithm reached the optimum in all three groups of experiments for meningitis, hydrocephalus and cerebral infarction, and the enhancements for MI reached 25%, 2.5% and 20.5%, respectively, which were the plain objective constants with the largest enhancements. The smallest improvement was SSIM, but it also reached 16.6%, 1.6%, and 15.0%, respectively. Among them, MI could reflect the rate of change of image brightness, and SSIM could consider both brightness and contrast of the image. The MI and SSIM values were considered together, i.e., the higher their values, the clearer the image. The corresponding results in Fig. 10 are shown in Table III.

In Table III, both CNN-RGF and NSCT-PCNN had higher evaluation values for various indicators in the same disease. In the standard deviation, CNN-RGF also exhibited better stability. To make the test results more comprehensive, their CT/MRI objective indexes were also evaluated, as shown in Fig. 11.

In Fig. 11, CNN-RGF outperformed the NSCT-RCNN algorithm in all objective metrics when observing CT/MRI maps of the brain. The observation of MI in meningitis reached the optimal value, the observation of SSIM in cerebral infarction belonged to the suboptimal value, and the four objective values based on CT/MRI were improved by 22.5% on average compared with the NSCT-RCNN algorithm, which could provide a large number of medical CT/MRI quality images. Since medical images represent personal privacy, the First Hospital of Harbin Medical University did not keep some data. The dataset for this study required some pioneering, which largely limited the performance of the constructed model. The corresponding results in Fig. 11 are shown in Table IV.

In Table IV, both CNN-RGF and NSCT-PCNN had higher evaluation values for various indicators in the same disease; In the standard deviation, CNN-RGF also exhibited better stability. Taking into account the implications, a complementary experiment was designed and implemented, i.e., based on MR/SPECTION metric observations, as shown in Fig. 12.

TABLE III. DETECTION INDEX RESULTS OF DIFFERENT ALGORITHMS IN ENCEPHALOPATHY

Index		Meningitis	Hydrocephalus	Cerebral infarction
CNN-RGF	MI	1.70±0.13	2.45±0.22	2.53±0.23
	IE	3..41±0.33	4.39±0.36	5.55±0.42
	SSIM	0.35±0.06	0.48±0.09	0.74±0.11
	AG	9.18±0.59	9.39±0.61	9.74±0.62
NSCT-PCNN	MI	1.62±0.12	2.14±0.15	2.47±0.19
	IE	3.29±0.28	3.88±0.31	4.12±0.34
	SSIM	0.14±0.02	0.37±0.05	0.56±0.07
	AG	7.82±0.31	8.57±0.42	9.04±0.47

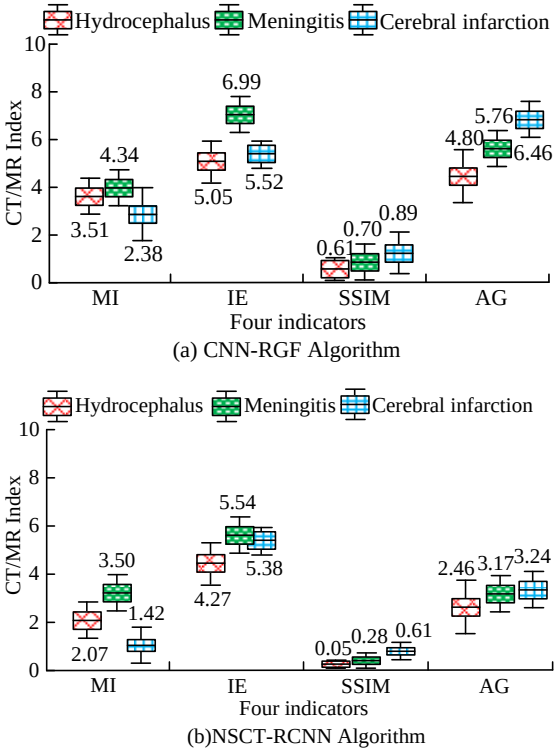

Fig. 11. Comparison of three kinds of encephalopathy and CT/MRI indexes by using CNN-RGF and NSCT-RCNN.

TABLE IV. OBSERVATION RESULTS OF CT/MRI INDICATORS

Index		Meningitis	Hydrocephalus	Cerebral infarction
CNN-RGF	MI	4.34±0.24	3.51±0.33	2.38±0.34
	IE	6.99±0.44	5.05±0.32	5.52±0.31
	SSIM	0.70±0.07	0.61±0.04	0.89±0.07
	AG	5.76±0.70	4.80±0.13	6.46±0.68
NSCT-PCNN	MI	3.50±0.50	2.07±0.34	1.42±0.58
	IE	3.29±0.28	3.88±0.31	4.12±0.34
	SSIM	0.14±0.02	0.37±0.05	0.56±0.07
	AG	7.82±0.31	8.57±0.42	9.04±0.47

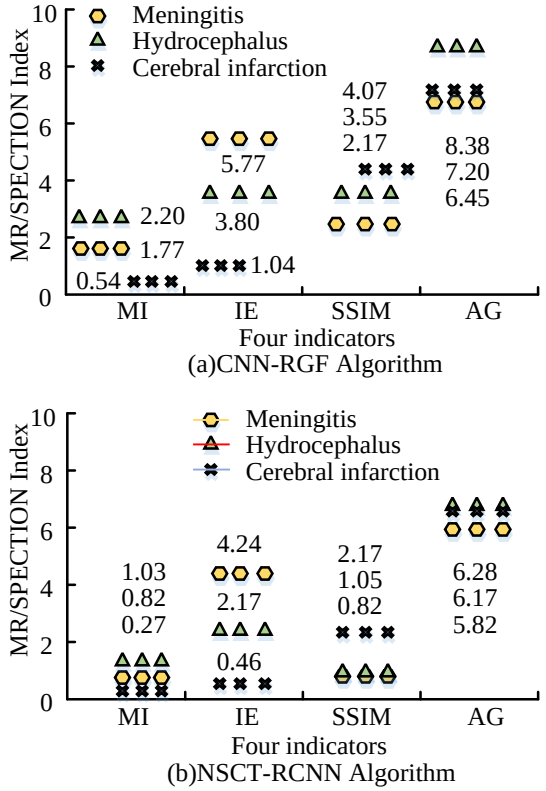

Fig. 12. Comparison of three kinds of encephalopathy and MR/SPECTION indexes by using CNN-RGF (a) and NSCT-RCNN (b).

From Fig. 12, it was indeed very necessary to test the complementation experiment, i.e., it was very different from the previous images. However, the difference was only in the shape, and the specific values of the four indexes of the MR/SPECTION images were still better for CNN-RGF than for NSCT-RCNN. Testing MR/SPECTION images, the MI/IE/SSIM/AG values of CNN-RGF were obtained as 1.77/5.77/2.17/6.45 for meningitis. The average values of MI/IE/SSIM/AG for hydrocephalus were 2.20/3.80/3.55/6.28; MI/IE/SSIM/AG for cerebral infarction were 0.54/1.04/4.07/7.20. The average values were 24.8% higher than those of NSCT-RCNN, and the CT/MRI images produced by the CNN-RGF algorithm were clearer. It is proved that CNN-RGF can retain the details of CT/MRI images well and can provide better quality CT/MRI images. The corresponding results in Fig. 12 are shown in Table V.

TABLE V. OBSERVATION RESULTS OF MRI/SPECTION INDICATORS

Index		Meningitis	Hydrocephalus	Cerebral infarction
CNN-RGF	MI	1.77±0.35	2.20±0.41	0.54±0.08
	IE	3.80±0.47	5.77±0.69	1.04±0.13
	SSIM	2.17±0.08	3.55±0.25	4.07±0.38
	AG	6.45±0.81	8.38±0.76	7.20±0.48
NSCT-PCNN	MI	0.82±0.04	1.03±0.05	0.27±0.03
	IE	4.24±0.39	2.17±0.11	0.46±0.07
	SSIM	0.82±0.16	1.05±0.13	2.17±0.28
	AG	5.82±0.68	6.28±0.75	6.17±0.47

In Table IV, both CNN-RGF and NSCT-PCNN had higher evaluation values for various indicators in the same disease. In the standard deviation, CNN-RGF also exhibited better stability. To ensure the without loss of generality of the proposed method, NWPU VHR-10 data set was used to verify the model performance. The NWPU VHR-10 dataset was an aerial photography dataset, and aerial images were also processed using methods such as cropping and stitching. Therefore, CNN-RGF was studied for this dataset, and the performance results were reflected through accuracy and recall indicators. The specific results are shown in Table VI.

TABLE VI. THE APPLICATION EFFECT OF CNN-RGF ALGORITHM IN NWPU VHR-10 DATASET

Comparison algorithm	Precision	Recall
Neural Network	0.673	0.616
K-Nearest Neighbor	0.785	0.690
SVM	0.820	0.714
CNN-RGF	0.913	0.873

In Table VI, the accuracy of CNN-RGF in the NWPU VHR-10 dataset was 0.913, and the recall rate was 0.873. Compared to other models, the performance proposed in the study has significantly improved. Although its accuracy was slightly lower, which might be due to the complex environment of aerial images, the results demonstrated the effectiveness of the proposed method.

V. CONCLUSION

This research used the brain image data of the First Affiliated Hospital of Harbin Medical University, and used CNN to extract its features. The extracted image used RGF to process the image's similar blocks and artifacts. Using the CNN-RGF method to fuse CT/MRI images, after the fusion was completed, it was first tested for naive objective quantities. The four simple objective indicators of CNN-RGF in meningitis, MI/IE/SSIM/AG, were 2.45/3.41/0.48/9.39; The MI/IE/SSIM/AG in Hydrocephalus was 1.70/4.39/0.35/9.74; The MI/IE/SSIM/AG in cerebral infarction was 2.53/5.55/0.74/9.18. At the same time, NSCT-RCNN algorithm was selected for comparative experiments. The measured MI/IE/SSIM/AG values of NSCT-RCNN in meningitis, Hydrocephalus, and cerebral infarction were 2.14/3.29/0.37/9.04, 1.62/4.12/0.14/7.82, 2.47/3.88/0.56/8.57, respectively. Among the three common brain diseases, the naive objective values obtained by CNN-RGF were higher than

those obtained by NSCT-RCNN, with an average improvement level of over 15%, indicating that the CT/MRI image quality obtained through CNN-RGF fusion was higher. To make the results universal, objective data testing of CT/MRI was supplemented. The CT/MRI objective data of CNN-RGF in meningitis, Hydrocephalus and cerebral infarction were 4.34/6.99/0.70/5.76, 3.51/5.05/0.61/4.80, 2.38/5.52/0.70/6.46 respectively; In NSCT-RCNN, they were 3.50/5.54/0.28/3.17, 2.07/4.27/0.05/2.46, 1.42/5.38/0.28/3.24, respectively. The values obtained from the above results were all dimensionless and were mainly used for comparing the effectiveness of algorithms. Therefore, from the extensive testing, the CT/MRI objective data values obtained from CNN-RGF were higher than those from NSCT-RCNN, with an average improvement level of over 10%. This proved that the CT/MRI obtained through CNN-RGF fusion was more suitable for major hospitals.

Because the first affiliated Hospital of Harbin Medical University lacked some data, the performance of the constructed model was limited. Due to scientific rigor, supplementary experiments were designed and implemented to observe MR/SPECTION indicators. The MR/SPECTION observations of meningitis in CNN-RGF were 1.77/5.77/2.17/6.45, respectively. The observed value of Hydrocephalus on MR/SPECTION was 2.20/3.80/3.55/6.28; The observed values of MR/SPECTION for cerebral infarction were 0.54/1.04/4.07/7.20, with an average of 24.8% higher than NSCT-RCNN. Taking the human eye resolution MI=8.00 as a reference, it could reach MI=2.40 as clear, indicating that for NSCT-RCNN fusion images, the characterization of meningitis and cerebral infarction cannot even reach the minimum standard. Compared with the traditional algorithm NSCT-RCNN, CNN-RGF was more suitable for application in hospitals. But there are not many medical images studied, because medical images have privacy and are not suitable for widespread dissemination. With the increase of volunteers, it is believed that future research can be improved.

REFERENCES

[1] X. Gao, M. Shi, X. Song, C. Zhang, and H. Zhang, "Recurrent neural networks for real-time prediction of TBM operating parameters," Autom. Constr., vol. 15, pp. 130-140, February 2019.

[2] J. H. Jung, H. Chung, Y. S. Kwon, and I. M. Lee, "An ANN to predict ground condition ahead of tunnel face using TBM operational data," KSCE J. Civ. Eng., vol. 23, pp. 5-6, May 2019.

[3] R. Hasanpour, J. Rostami, J. Schmitt, Y. Ozcelik, and B. Sohrabian, "Prediction of TBM jamming risk in squeezing grounds using Bayesian and artificial neural networks," J. Rock Mech. Geotech. Eng., vol. 12, pp. 21–31, February 2020.

[4] L. Liu, W. Zhou, and M. Gutierrez, "Effectiveness of predicting tunneling-induced ground settlements using machine learning methods with small datasets," J. Rock Mech. Geotech. Eng., vol. 14, pp. 1028-1041, August 2022.

[5] D. Wang, H. Zhao, and Q. Li, "Medical brain image classification based on multi-feature fusion of convolutional neural network," J. Intell. Fuzzy Syst., vol. 38, pp. 127-137, January 2020.

[6] S. Polinati, D. P. Bavirisetti, K. N. V. P. S. Rajesh, and R. Dhuli, "Multimodal medical image fusion based on content-based and PCA-sigmoid," Cur. Med. Imag., vol. 18, pp. 546-562, Number 2022.

[7] L. Wang, J. Zhang, Y. Liu, J. Mi, and J. Zhang, "Multimodal medical image fusion based on Gabor representation combination of multi-CNN and fuzzy neural network," IEEE Access, vol. 9, pp. 67634-67647, April 2021.

[8] N. N. Yu, T. S. Qiu, and W. H. Liu, "Medical image fusion based on sparse representation with KSVD," Chin. J. Biomed. Eng., vol. 28, pp. 168-172, May 2019.

[9] S. Singh and D. Gupta, "Multistage multimodal medical image fusion model using feature-adaptive pulse coupled neural network," Int. J. Imag. Syst. Technol., vol. 31, pp. 981-1001, November 2020.

[10] M. Yu, C. Ning, and Y. Xue, "Brain medical image fusion scheme based on shuffled frog EAPING algorithm and adaptive pulse oupled neural network," Image Process., vol. 6, pp. 1203-1209, December 2020.

[11] J. S. Guan, S. B. Kang, and Y. Sun, "Medical image fusion algorithm based on multi-resolution analysis coupling approximate spare representation," Future Gener. Comput. Syst., vol. 98, pp. 201-207, September 2019.

[12] G. Wang, W. Li, X. Gao, B. Xiao, and J. Du, "Multimodal medical image fusion based on multichannel coupled neural P systems and max-cloud models in spectral total variation domain," Neurocomputing, vol. 480, pp. 61-75, April 2022.

[13] C. Wang, R. Nie, J. Cao, X. Wang, and Y. Zhang, "IGNFusion: An unsupervised information gate network for multimodal medical image fusion," IEEE J. Sel. Top. Sig. Proc., vol. 16, pp. 854-868, June 2022.

[14] K. P. Das and J. Chandra, "Multimodal classification on PET/CT image fusion for lung cancer: A comprehensive survey," ECS Trans., vol. 107, pp. 3649-3673, 2022.

[15] S. Akbar, S. A. Hassan, A. Shoukat, J. Alyami, and S. A. Bahaj, "Detection of microscopic glaucoma through fundus images using deep transfer learning approach," Microsc. Res. Tech, vol. 85, pp. 2259-2276, February 2022.

[16] M. T. Vo, A. H. Vo, T. Le, "A robust framework for shoulder implant X-ray image classification," Data Technol. Appl., vol. 56, pp. 447-460, 2022.

[17] G. Xiao, "Problems of railway tunnel construction under some special geological conditions in China and their countermeasures," Tunnel Constr., vol. 39, pp. 1748-1758, December 2019.

[18] J. Li, W. Zhang, W. Diao, Y. C. Feng, X. Sun, and K. Fu, "CSF-Net: Color spectrum fusion network for semantic labeling of airborne laser scanning point cloud," IEEE J. Sel. Top. Appl. Earth Observations Remote Sens., vol. 15, pp. 339-352, December 2022.

[19] T. Feng, C. Wang, J. Zhang, B. Wang, and Y. Jin, "An improved artificial bee colony-random forest (IABC-RF) model for predicting the tunnel deformation due to an adjacent foundation pit excavation," Undergr. Space, vol. 7, pp. 514-527, August 2022.

[20] R. Chen, P. Zhang, H. N. Wu, Z. T. Wang, and Z. Q. Zhong, "Prediction of shield tunneling-induced ground settlement using machine learning techniques," Front. Struct. Civil Eng., vol. 13, pp. 1363-1378, September 2019.

[21] R. Zhu, J. Fang, S. Li, Q. Wang, H. Xu, J. Xue, and H. Yu, "Vehicle re-identification in tunnel scenes via synergistically cascade forests," Neurocomputing, vol. 381, pp. 227-239, March 2019.

[22] G. X. Xu, X. Y. Xu, L. Wang, G. Q. Fu, P. Zhao, and A. Ding, "Sand-fix effects of Haloxylon ammodendron forests under the different densities and patterns under wind tunnel test," J. Arid Land Res. Environ., vol. 33, pp. 189-195, September 2019.

[23] L. Q. Yang, "Real-time prediction of rock mass classification based on TBM operation big data and stacking technique of ensemble learning," J. Rock Mech. Geotech. Eng., vol. 14, pp. 123-143, February 2022.

[24] B. Ramosaj and M. Pauly, "Consistent estimation of residual variance with random forest out-of-bag errors," Stat. Probab. Lett., vol. 151, pp. 49-57, August 2019.

[25] J. Yang, S. Yagiz, Y. J. Liu, and F. Laouafa, "Comprehensive evaluation of machine learning algorithms applied to TBM performance prediction," Undergr. Space, vol. 7, pp. 37-49, February 2022.

[26] C. Zhang, W. Z. Wang, C. Zhang, B. Fan, J. Wang, F. Gu, and X. Yu, "Extraction of local and global features by a convolutional neural network–long short-term memory network for diagnosing bearing faults," Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci., vol. 236, pp. 1877-1887, April 2022.

[27] M. S. Tamber, K. A. Scott, and L. T. Pedersen, "Accounting for label errors when training a convolutional neural network to estimate sea ice

concentration using operational ice charts," IEEE J. Sel. Top. Appl. Earth Observations Remote Sens., vol. 15, pp. 1502-1513, January 2022.

[28] Z. Zhao, J. Gui, A. Yao, N. Le, and M. Chua, "Improved prediction model of protein and peptide toxicity by integrating channel attention into a convolutional neural network and gated recurrent units," ACS Omega, vol. 7, pp. 40569-40577, October 2022.

[29] F. Masood, J. Masood, H. Zahir, K. Driss, N. Mehmood, H. Farooq, "Novel approach to evaluate classification algorithms and feature selection filter algorithms using medical data," J. Comput. Cogn. Eng., vol. 2, pp. 57-67, May 2023.

[30] S. Wang and Y. Shen, "Multi-modal image fusion based on saliency guided in NSCT domain," IET Image Process., vol. 14, pp. 3188-3201, November 2020.