SFCN: Symmetric feature comparison network for detecting ischemic stroke lesions on CT images

Long Zhang1 | Chuang Zhu1 | YueWei Wu1 | Yang Yang1 | Yihao Luo1 | Ruoning Song1 | Lian Liu2 | Jie Yang1

1 School of Artificial Intelligence, Intelligent Perception and Computing Research Center, Beijing University of Posts and Telecommunications, Beijing, China
2 The interventional neuroradiology department, Beijing Tiantan Hospital, Beijing, China

Correspondence Lian Liu, Rm523, A1 Building, Beijing Tiantan Hospital, Beijing, 10070, China. Email: liuliandaxiang@126.com

Funding Information This work was supported in part by the National Natural Science Foundation of China Grant 61972248, in part by the Beijing Municipal Natural Science Foundation Grant 72010056, in part by the Beijing Natural Science Foundation under Grant 4182044, in part by the Beijing Laboratory of Advanced Information Networks of BUPT, in part by the Beijing Key Laboratory of Network System Architecture and Convergence of BUPT, in part by the National Key Research and Development Program of China under Grant 2017YFB0802701.

Abstract
Ischemic stroke is the most common stroke and the leading cause of disability and death in the world. Computed tomography (CT) is a popular and economical diagnostic device for the stroke. However, the ischemic stroke lesions are not evident on CT images and the diagnostic result relies on the visual observation of neurologists, which may vary from doctor to doctor. To facilitate the treatment, a computer-aided detection algorithm on CT images is proposed to help clinician for the ischemic stroke screening. In order to obtain accurate lesion annotation on CT images, novel automatic algorithms are developed to achieve image pairing, calibration, and registration. Then, a new framework with the symmetric feature extraction and comparison is proposed to identify and locate the ischemic stroke lesion. Experimental results show that this method achieves 75% of DICE in the detection of ischemic stroke lesions, which is higher than other methods by 4%. Its competitive results compared with seven latest methods is shown in terms of extensive qualitative and quantitative evaluation. This method can accurately detect the lesion in the CT images through the comparison of symmetric regional features, which has contributed to the clinical diagnosis of ischemic stroke.

1 INTRODUCTION

Cerebrovascular accident (CVA) is a leading cause of disability and death in the world. There are 15 million new CVA cases and 5 million deaths in 2018 [1]. Ischemic stroke is the most common type, accounting for 87% of all CVA. It can be caused by a blockage in an artery supplying the brain with blood [2]. If patients are not effectively treated within 4.5 h of onset, they may have a high probability of disability or even death [3]. Magnetic resonance imaging (MRI) and Computed Tomography (CT) are the most commonly used diagnostic equipment for ischemic stroke [4]. MRI is more sensitive to ischemic stroke than CT [5]. However, due to the high maintenance costs of MRI equipment and the need for more professionals, it is usually only available in large hospitals, and MRI examination also has many limitations [6]. For example, if there is metal in the brain, MRI cannot be used. Therefore, the identification and location of ischemic stroke lesions based on CT scans are still of great significance. But this is not an easy task. Figure 1 shows the DWI and CT images of the same location in the brain. The red rectangle box marks the lesion information. We can see that the ischemic stroke lesion on CT is not obvious enough, which leads to many classic deep learning models cannot achieve nice performance on this task.

In recent years, deep learning (DL) methods have been widely used in the fields of medical image analysis to help physicians in improving the accuracy and promptness of medical diagnosis. Deep convolution structure can extract a series of effective features without manual intervention [7]. Some methods have been proposed to solve the problem of ischemic stroke lesion detection on CT [8–10]. However, these methods have not answered two fundamental questions: (1) How to resolve the differences...
in CT labelling? (2) How to clean the data used in research? Reliable annotation is essential for all deep learning tasks, and the quality of annotation directly determines the performance of the trained model. In this task, it is difficult to directly obtain highly consistent annotations. The reason is shown in Figure 1. The intracranial structure is delicate and complex, and limited by the CT imaging mechanism, the grayscale difference between different tissues is small. Therefore, we use DWI as an intermediary to eliminate the inconsistencies in the labelling as much as possible and greatly improve the quality of the labelling. In addition, medical image data is different from traditional image data. It only has one channel of information, but this one channel contains rich information. Therefore, data preprocessing is very important and should be explained clearly.

The deep learning model also has a performance bottleneck, because the calculation form of the convolution kernel makes it unable to use some inherent features flexibly. For example, the human brain is symmetrical, and the difference between the left and right hemispheres can be used to determine whether there is a stroke lesion.

To solve the problems mentioned above, (1) we design a label transfer model (LTM) to obtain the CT ground truth. (2) We design SFCN using the symmetry feature to improve the feature extraction capabilities. In addition, brain tissue extraction, brain mid-axis recognition, and rotation methods are proposed to improve the overall performance. To sum up, the main contributions of our work are:

1. Solve the label conflict problem on CT. We design algorithms to pair MRI and CT automatically and register ground truth annotations of DWI to CT scans

2. Break through the limitations of traditional convolution kernels. We design a symmetric feature comparison network (SFCN), which can be applied to datasets with symmetric features

2 RELATED WORK

Due to the low price and high penetration rate of CT equipment, it has already become a commonly used diagnostic equipment for stroke. Early studies of ischemic stroke and haemorrhagic stroke are based on non-enhanced CT images [11]. Haemoglobin can absorb X-rays, so CT is very sensitive to haemorrhagic stroke. Many early studies used CT scans to detect or segment haemorrhagic stroke lesions [12–15]. The detection or segmentation of haemorrhagic stroke lesions on CT is an easy task, and many traditional algorithms or machine learning algorithms have achieved good performance.

The features of ischemic stroke lesions on CT images are confusing. Image enhancement methods can magnify the difference between lesions and healthy tissue. Sajjadi et al. [16] proposed an image enhancement algorithm, and the task of this paper is to output the ASPECT score [17] of the patient, so it can be regarded as a classification task or a regression task. Usinskas et al. [18] uses the Ostu method to find the optimal segmentation threshold to segment the lesion, but it has assumed that the image contains healthy tissue and stroke lesions. Yahiaoui [9] uses the Laplace enhanced image algorithm to amplify the differences between pixels and then uses FCM to cluster the images. Since this method needs to determine the number of classifications in advance, it requires multiple attempts and is not suitable for clinical application.
Many researchers found that the symmetry of the brain is a useful feature [19]. Doctors often judge the location of the lesion by comparing the differences between the left and right hemispheres. Maldjian et al. [20] uses the brain’s atlas to divide the brain area, and calculates the grey value distribution of the corresponding partition, and determines whether the difference includes the stroke lesion in the distribution. The accuracy of this method is low. Task-oriented Deep Network [21] has two computing branches, one k-means for pixel clustering, and the other is to separate the left and right brains through Bezier curves, extract texture features and finally identify the lesions by SVM classifier. Tyan et al. [19] extracts target tissues through a region growth algorithm and then pairs them with symmetry to extract features to determine the location of the lesion. In recent years, due to the wide application of deep learning technology in medical imaging, it has become a trend to hand over the task of feature design to the deep learning model. Wang et al. [22] designed a deep CNN network to extract the abstract features and segment stroke lesions. Rubin [23] proposes a CT to MRI conversion model and then used the deep CNN to segment the stroke lesions on transferred images. These methods make full use of the advantages of deep learning models, but when the model depth reaches a certain extend, the model performance will no longer improve.

With the popularity of new diagnostic equipment, many ischemic stroke studies based on other images have emerged in recent years. MRI-based research is a major trend in recent years. The region of DWI abnormality can act as a gold standard for irreversible brain infarction in clinical practice. Multimodal MRI can be used to classify the onset time of ischemic stroke [24], and they may help to eliminate the artefacts in segmentation results [25]. For example, the method uses a fuzzy C-means algorithm to segment the lesion on DWI, then eliminate artefacts through T1, T2, and Flair [26]. The task of ischemic stroke based on MRI is relatively simple. However, there are few medical institutions equipped with MRI. CT perfusion can show the state of blood flow. The blood flow velocity at ischemic stroke lesions is very low, and low density features are present on CT image. [27] Is an ischemic stroke lesion segmentation method on CT perfusion image. However, perfusion CT has not been widely used due to its high cost and harm to the human body.

CT-based research still has research value. However, the existing research does not answer these two fundamental questions, and is obsessed with pursuing deeper models. The traditional model structure has reached the performance bottleneck. Therefore, we improve performance by designing a symmetric comparison network.

3 METHODS

In this section, we will introduce three modules, Label Transfer Model (LTM), which are used to normalize CT images and obtain ground truth. Symmetric Feature Comparison Network (SFCN) to identify and locate ischemic stroke lesions. And the image enhancement module magnifies the difference between diseased and healthy tissue. The relationship between the modules is shown in Figure 2.

3.1 Label transfer model (LTM)

As shown in Figure 1, it is difficult to obtain accurate labels of CT data. However, the quality of data annotation has a significant impact on model training. DWI is a gold indicator of ischemic stroke in clinical practice. Therefore, CT and DWI can be registered to obtain CT labelling information. To make the experiment more rigorous, the patients selected have the following requirements:

1. The medical record contains keywords: ischemic stroke
2. Patients should have CT and MRI images before surgery
3. Time interval between CT and MRI is less than two hours

Figure 3 shows the LTM workflow. The patient select module requires data to meet the above three constraints to ensure that the lesions of MRI and CT are similar. For the selected data, first, noise suppression; second, data calibration; and finally, the DWI image is registered to the CT image; Doctors and experts label the registered DWI to obtain the ground truth of the CT images.

3.1.1 Image denoising

In order to achieve sufficient disk storage, the original pixel values in DICOM undergo a linear transformation. Doctors commonly use pixel values in Hounsfield units (HU), so pixel transformation must be performed. The transformation equation is defined as:

\[ H(x) = \text{RescaleSlope} \times x + \text{RescaleIntercept} \] (1)

where \( x \) is the value of DICOM pixel. RescaleSlope and RescaleIntercept are determined by the hardware manufacturer.

The associated symptom of ischemic stroke is cerebral edema. Therefore, we are more interested in the information around the water’s HU. We designed a bandpass filter to extract the information of interest. It is defined as:

\[ E(x) = \begin{cases} 
- (k - 1) T, & x < \langle -T \rangle \\
\end{cases} \] (2)

\[ k x, & |x| \leq T \\
\end{cases} \] (2)

\[ x + (k - 1) T, & x > T \]

where \( x \) is the pixel value, \( k \) is the sensitivity coefficient, and \( T \) is the sensitive area. The original image and the image after our processing are shown in Figure 4.
Next, we use an anisotropic filter for blurring processing, which can reduce noise and increase the brightness of regional differences in the image. The filter is defined as:

$$I(x, y) = \frac{\sum_{i=-1}^{+1} \sum_{j=-1}^{+1} I(x + i, y + j) G(x, y)}{\sum_{i=-1}^{+1} \sum_{j=-1}^{+1} G(x, y)}$$  \hspace{1cm} (3)$$

$$G(x, y) = \exp \left( \frac{-\sqrt{G_x^2 + G_y^2}}{2k^2} \right)$$  \hspace{1cm} (4)$$

$$(G_x, G_y) = \left( \frac{\partial I(x, y)}{\partial x}, \frac{\partial I(x, y)}{\partial y} \right)$$  \hspace{1cm} (5)$$

where $I(x, y)$ represents an original image, $G(x, y)$ represents the Gaussian kernel function, and $k$ is the variance of the Gaussian mask.

### 3.1.2 Image calibration

The non-standard lying position of the brain during CT imaging increases the difficulty of lesion detection. Our image calibration workflow is shown in Figure 5.

We use edge information to identify the brain midline, and perform image calibration based on it. There is a lot of salt and pepper noise in CT images, and it has a significant influence on edge extraction. Median filtering can solve this problem well. We take Robert Operator to extract the edges [28]. Robert Operator is defined as follow:

$$G(f(x, y)) = \left\{ f(x+1, y+1) - f(x, y) \right\}^2 + \left\{ f(x+1, y) - f(x, y+1) \right\}^{\frac{1}{2}}$$  \hspace{1cm} (6)$$

Since the calculation load of the above formula is too high, we use the following approximate formula:

$$G(f(x, y)) = |f(x+1, y) - f(x, y)| + |f(x, y+1) - f(x, y)|$$  \hspace{1cm} (7)$$
After edge extraction, we eliminate some interference points through connected areas. Commonly used are four connected areas and eight connected areas:

\[ N4(x, y) = (x, y) \cup (x + 1, y) \cup (x - 1, y) \cup (x, y + 1) \cup (x, y - 1) \quad (8) \]

\[ N8 (x, y) = N4(x, y) \cup (x + 1, y + 1) \cup (x + 1, y - 1) \cup (x - 1, y + 1) \cup (x - 1, y - 1) \quad (9) \]

So far, we have obtained many continuous line segments and scattered points, and we regard the continuous line segments as a collection of scattered points. We use Hough Transform to find the best-fit straight line [29].

First, we convert the scattered points in the Cartesian coordinate system to a polar coordinate system, such as \((x_0(i), y_0(i))\) to \((\rho(i), \theta(i))\)

\[ \rho = x_0 \cos \theta + y_0 \sin \theta \]

\[ = \sqrt{x_0^2 + y_0^2} \left( \frac{x_0 \cos \theta + y_0 \sin \theta}{\sqrt{x_0^2 + y_0^2}} \right) \]

\[ = \sqrt{x_0^2 + y_0^2} \sin (\theta + \varphi_0) \quad (10) \]

\[ \sin \varphi_0 = \frac{x_0}{\sqrt{x_0^2 + y_0^2}} \quad (11) \]

Counting the intersection of these curves, the parameter with the highest frequency is the result of the Hough transform. However, the Hough transform may generate multiple fitting straight lines to make the final brain axis result more accurate. We will perform a linear regression on the result of the Hough transform. For points:

\[ D = \{(x_1, x_2, \ldots, x_d, y_i)\}_{i=1}^{m} \quad (12) \]

where \(d\) is the dimension of the input space, \(m\) is the number of data. The form of the function we need to fit is defined as:

\[ f(x) = w_1 x_1 + w_2 x_2 + \cdots + w_d x_d + b \quad (13) \]

For any point \(x\) must satisfy the following relationship:

\[ f(x_i) = w_1 x_{i1} + w_2 x_{i2} + \cdots + w_d x_{id} + b \approx y_i \quad (14) \]

We use the mean square error to determine the parameters \(w\) and \(b\):

\[ (w^*, b^*) = \arg \min_{w,b} \sum_{i=1}^{m} (f(x_i) - y_i)^2 \]

\[ = \arg \min_{w,b} \sum_{i=1}^{m} (y_i - w x_i - b)^2 \quad (15) \]

We use the parameter estimation method to differentiate \(w\) and \(b\) respectively:

\[ \frac{\partial E_{(w,b)}}{\partial w} = 2 \left( w \sum_{i=1}^{m} x_i^2 - \sum_{i=1}^{m} (y_i - b) x_i \right) \quad (16) \]

\[ \frac{\partial E_{(w,b)}}{\partial b} = 2 \left( mb - \sum_{i=1}^{m} (y_i - w x_i) \right) \quad (17) \]

Let the above formula be equal to 0 to get the optimal solution

\[ w = \frac{\sum_{i=1}^{m} y_i (x_i - \bar{x})}{\sum_{i=1}^{m} x_i^2 - \frac{1}{m} (\sum_{i=1}^{m} x_i)^2} \quad (18) \]

\[ b = \frac{1}{m} \sum_{i=1}^{m} (y_i - w x_i) \quad (19) \]

Once the brain midline is obtained, it can be calibrated by image rotation.
3.1.3 Image registration

The region of DWI abnormality can act as a gold standard for irreversible brain infarction in clinical practice. Register the DWI to the calibrated CT, and then label the registered DWI to obtain the labelled data corresponding to the calibrated CT. Due to device parameter limitations, the DWI sequence has 24 slices, and the CT sequence has 32 slices. We screen out the slices containing the brain and pair the two sequences by linear mapping. Figure 6 shows the slice mapping method.

After getting the image pair, we use image mutual information registration method, and the entropy of an image is defined as:

\[ p_i = \frac{h_i}{\sum_{i=1}^{N} h_i} \]  

where \( h_i \) represents the total number of pixels in the image whose gray value is \( i \), \( N \) represents the number of gray levels of the image. \( P_i \) represents the probability of gray level \( i \) appearing.

\[ H(Y) = - \sum_{i=0}^{N} p_i \log p_i \]  

where \( H(Y) \) is the entropy of the image \( Y \). We define the mutual information of the two images as:

\[ MI(R,F) = H(R) + H(F) - H(R,F) \]  

where image \( R \) and image \( F \) are the image to be registered and the floating image. The mutual information between image \( R \) and image \( F \) is defined as:

\[ H(R,F) = - \sum_{r,f} p_{RF}(r,f) \log p_{RF}(r,f) \]  

where \( r, f \) represent pixels from images \( R \) and \( F \) respectively. When the similarity of two images is higher, the correlation is more significant, the joint entropy is smaller, that is, the mutual information is greater. The best registration of the image can be calculated by maximizing the mutual information of the image. Figure 7 shows a set of CT-DWI pairs selected by the algorithm. DWI is calibrated on the basis of CT. After calibration, the two images are superimposed to observe the registration effect.

3.2 Image enhancement

There will be differences between healthy brain tissue and stroke lesions in CT, but this difference is fragile. We use an image enhancement algorithm based on the Laplacian pyramid to increase the difference between normal tissue and the diseased area. The Laplacian pyramid is the most common multiscale change and is usually used for image compression. It is a sequence of error images \( L_0, L_1, \ldots, L_N \). each is the difference between the two levels of the Gaussian pyramid. As a bandpass filter, pyramid construction tends to enhance image features, which are important for interpretation.

The zero levels of the pyramid \( A_0 \) is equal to the original image. This is low-pass filtered and sub-sampled by a factor of two to obtain the next pyramid level, \( A_1 \) is then filtered in the same way and sub-sampled to obtain \( A_2 \) and so on. If \( 0 < l < N \), the levels of the pyramid are obtained iteratively by:

\[ A_l(i,j) = \sum_{m} \sum_{n} w(m,n) A_{l-1}(2i+m,2j+n) \]  

where \( w(m,n) \) is Gaussian generating kernel. And it is convenient to refer to this process as standard REDUCE operation, and is given by:

\[ A_l = \text{REDUCE}[A_{l-1}] \]  

On the other hand, we need to assume that we only have the data after REDUCE, and interpolate the REDUCE data.
our model, the classification of image blocks into normal and abnormal images depends on three key factors, the image block features, the symmetric image block features, and entire image features. Symmetry is a remarkable characteristic of the human brain. The left and right brains are relatively independent. From the perspective of the pathogenesis of ischemic stroke, it is difficult to have both left and right brains sick at the same time. Therefore, when the lesions appear in the left or right hemispheres, the characteristics will be different from the symmetrical hemispheres. We also need to consider the global features of the image. On the one hand, it can be used to detect the rare occurrence of left and right hemispheres at the same time. On the other hand, we hope to optimize the classification results by introducing global features. Therefore, we design a lesion identification and classification network SFCN, based on symmetry features, as shown in Figure 8.

Each CT slice of the brain requires two types of processing. One is to extract local features through Block Feature Extraction (BFE). First, CT slices need to be divided into image blocks symmetrically. For each image block, we design a multi-scale sampling network model to extract features. We use the same convolution kernel to calculate the characteristics of each image block, and quantify the difference between the symmetric image blocks through the difference in the feature maps of the image blocks (It will be explained in detail later). For the difference features map of each block, we first sample to a fixed size, and then combine the image block features into the size of the global feature map. The local feature map and the global feature map will be concatenated (increase the number of channels), and the concatenated feature maps are sent to output an \( m \times m \) result, where \( m \times m \) represents the number of symmetrical image blocks.

For the convenience of description. We use a matrix to introduce the difference feature’s calculation process, as shown in Figure 9. The other processing is to extract global features through Global Feature Extraction(GFE). The image first passes through three convolutional layers. To obtain global features, we adopt three kinds of dilated convolution kernels to extract information on different scales, then concatenating the global feature and the differential feature. Finally, we use a convolutional network to obtain the result matrix.

4 | DATA

Our data comes from the CT and MRI data of the patients in Beijing Tiantan Hospital from 2018 to 2020. This study has been approved by the Institutional Review Board of the Beijing Tiantan Hospital (IRB No. KYSQ 2019-124-01), Permission data (12 February 2019). The selected data are all taken before surgery, and the time interval is within 2 h.

4.1 | Dataset description

In this paper, we have performed a series of comparison experiments on our CT dataset to confirm the validity and practicality.
FIGURE 8  The architecture of the symmetric feature comparison network (SFCN). Block Feature Extraction (BFE) is used to extract the features of image blocks. Differential Feature (DF) is used to calculate feature differences. Global Feature Extraction (GFE) is used to extract global features.

FIGURE 9  From left to right are the original matrix, flipped matrix and differential matrix.

of our method. These files are all Digital Imaging and Communications in Medicine (DICOM) format (Siemens) with the following acquisition parameters: image Dimensions: 512 × 512 × 32; voxel Dimensions: 0.479 mm × 0.479 mm × 0.479 mm;

For each patient's data, we have undergone LTM and image enhancement processing, and use this as a basis to generate a dataset. The acute ischemic lesions were annotated by at least three experienced experts. We use the data of 40 patients (1280 images) to train and validate the model, and the remaining 15 patients (480 images) are only used for testing.

4.2 Data augmentation

The training of deep learning models requires many data, in order to add more data to the training set and make the data set more representative, performing data augmentation is necessary. Besides, data augmentation applies a variety of transformations to the dataset, which can improve the generalization ability of the model and correct the distorted label data to a certain extent. Since our model requires symmetrical data, our augmentation methods will be different from other models. For our proposed model, we adopt the following data enhancement method: (1) Using origin image; (2) randomly horizontally flip; (3) randomly vertically flip;

Since there is no symmetry requirement for the training of general models, more data augmentation methods can be used to improve the generalization ability of the models. The training data is randomly sampled, and one or more of the following transformations are randomly applied: (1) Using origin image; (2) randomly crop and resize at 512 × 512; (3) randomly horizontally flip; (4) randomly vertically flip; (5) randomly affine.

5 EXPERIMENTS AND RESULT

Table 1 shows the results of ablation experiments under different data processing methods. Sending data directly into the network model will cause the model's performance to drop sharply. This is because CT images are single-channel images and have higher accuracy than traditional images. Therefore, the numerical changes caused by lesions are not obvious enough. We need to extract the information of interest and suppress other information. When we denoise the image, most models have obvious performance improvements. After image calibration, all models obtain high performance gains, and symmetry is the most important feature. After we train the model with
TABLE 1  Denoising, calibration and enhancement ablation experiment

| Data processing | Model          |       |       |       |       |       |
|-----------------|----------------|-------|-------|-------|-------|-------|
|                 | Metric         | Denoising | Calibration | Enhancement | Linknet | FPN  | PAN  | PSPNet | Unet  | DeepLabV3 | DeepLabV3+ | Ours |
|                 |                | ×      | ×     | ×     | 0.23  | 0.37 | 0.33 | 0.23  | 0.31  | 0.36  | 0.40  | 0.42 |
|                 | F1-score       | ×      | ×     | ×     | 0.28  | 0.38 | 0.39 | 0.20  | 0.39  | 0.46  | 0.40  | 0.41 |
|                 |                | √      | ×     | ×     | 0.30  | 0.50 | 0.62 | 0.35  | 0.63  | 0.63  | 0.59  | 0.68 |
|                 |                | √      | √     | √     | 0.35  | 0.56 | 0.60 | 0.48  | 0.67  | 0.70  | 0.70  | 0.75 |

FIGURE 10  Each column from left to right represents the original image, the denoised image, edge detection results, ROI extraction, midline recognition, CT calibration enhanced images, most of the models have gained performance gains. The effects of different data processing methods will be shown in the next few sections.

5.1  Image calibration

We test our method on our dataset. Figure 10 shows the correction results of our approach. It can be seen that the original images [Figure 10(a)] become smoother after brain tissue extraction and median filtering [Figure 10(b)]. Figure 10(c) is the result of our edge detection method. We extract the region of interest (ROI) in the image through the centroid and gray distribution. It can effectively speed up the algorithm and reduce the recognition errors of the midline. Figure 10(e) is the result of our midline line detection algorithm. We calculated the slope of the midlines and adjusted these midlines to be parallel to the y-axis. [Figure 10(f)].

According to the details shown in Figure 10, we can see that the brain tissue extraction algorithm can remove the skull and remove the skin tissue outside the skull. In the follow-up experiments, we will demonstrate the excellent effect of the tissue extraction algorithm on model performance. After median filtering, the image becomes smoother, and the subsequent edge
detection of such an image is more accurate. Our ROI extraction algorithm can accurately extract the image part containing the brain midline. Since the previous operation filtered much noise, our midline extraction algorithm achieved excellent performance. The image calibration algorithm is fully automatic, which can significantly reduce the doctor’s burden and reduce the learning cost.

5.2 | Image registration

Image registration is a crucial step in obtaining ground truth of the data. Because the CT lesions are challenging to distinguish with the naked eye, the DWI sequence in MRI is a golden indicator of ischemic stroke and is very sensitive to the lesion. To make the ground truth in DWI closer to CT’s actual situation, we strictly screen patients, requiring patients to take CT and MRI before surgery, and the time interval between the two must not exceed 2 h. Such regulations can prevent patients from spreading lesions or absorbing thrombus during this time interval, resulting in inconsistent lesions in MRI and CT. MRI and CT are two different devices, so patients will have different lying positions when shooting. So we need to register the two. Considering the characteristics of our task, we will use the corrected CT as a fixed image, and the left and right floating images of MRI to fit the CT. We use pixel features to determine the starting and ending serial numbers of the two sequences. Image pairing is performed by linear mapping. Due to equipment limitations, the resolution of MRI sequences is lower than CT. The DWI sequence needs to be linearly interpolated before being registered with CT. When we get the registration adjustment, DWI can invite experts to mark it. Finally, transfer the annotation results to CT to obtain the ground truth of CT.

According to the details shown in Figure 11. Linear mapping can accurately match CT slices and MRI slices. The image registration algorithm is not affected by the skull skin in the MRI slices. It can be seen that the brain tissue in the two slices has been well registered. In actual clinical operations, doctors have many differences in labelling CT slices. It is difficult to obtain accurate labelling information for CT slices, so it is difficult to conduct related research. Existing research does not mention this difficulty. Through this method, we can solve this difficult problem well.

5.3 | Image enhancement

The image enhancement algorithm can be used to aid diagnosis, because it can magnify the slight gradient changes in the image, so that the doctor can more easily see the blood vessel veins and find the location of the lesion more conveniently. We take the image enhancement algorithm as a feature engineering in this paper and let it directly participate in the training and testing of the model. We tested the image enhancement algorithm under different parameters on our dataset, and the results are shown in Figure 12.

We found that when $N = 3,4$, the image enhancement result is not much different from the original image, that is, the details of the image are not significantly enlarged. When $N = 5$, the image enhancement result is different from the original image.
When $N > 5$, the results are not significantly improved. Considering the computational efficiency, we use the Laplace pyramid enhancement method, where $N = 5$.

### 5.4 Lesion identification and localization

In this section, we compare the segmentation results of the proposed method with other seven segmentation methods on the CT ischemic stroke dataset. The experiment is conducted on Pytorch-1.2.0 in python3.6 with NVIDIA Tesla V100-DGXS-32GB GPU. We adopt seven widely used segmentation methods as comparison methods, which include LinkNet [30], FPN [31], PAN [32], PSPNet [33], Unet [34], DeepLabV3 [35], Deeplabv3Plus [36].

#### 5.4.1 Typical scene classification

The dataset on which the proposed algorithm and the other comparison methods are tested contains several typical scenes, which involve significant issues in ischemic stroke lesion identification and localization. The major issues include scenes without lesions, large lesion targets, small lesion targets, and symmetric lesion targets, as shown below. Example images are given in Figure 13.

Without lesions: Under this situation, the difference between CT tissues is usually minimal. The brain tissue has gray matter and white matter, and the gray values of the two are different. These differences would bring difficulties to a precise result. Especially for the CNN-based algorithm, the difference in grayscale affected by the lesion may be confused with the difference in gray and white matter.

Large lesion targets: Patients with stroke usually have a more extensive lesion area, which is also the most common situation. Although large lesions represent more opportunities to detect lesions, improving recall is difficult for deep learning models.

Small lesion targets: Under this situation, the ischemic stroke lesion target is usually melted into the background. The ambiguous contour of the lesion would bring difficulties to precise results. Especially for the multi-scale based algorithm, the global feature maps of the image are very similar to those without lesions.

Symmetric lesion targets: Symmetrical lesions are relatively rare in clinic, because the symmetry features may be offset, which will lead our model to degenerate into the traditional CNN based deep learning model.

The model we proposed and the model used for comparison are both trained with the same loss function. DICE loss is a commonly used loss function, which is defined as:

$$\mathcal{L} = 1 - DICE$$

where $DICE$ is dice coefficient.

$$DICE = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

Therefore, the loss function is defined as

$$\mathcal{L} = \frac{FP + FN}{2 \cdot TP + FP + FN}$$

#### 5.4.2 Evaluation methods

The output of our model is to judge whether the image block contains a lesion. The size of the image block in the experiment is $32 \times 32$. Considering the characteristics of the task, we use the semantic segmentation model as a benchmark model for comparison. We divide the pixels of the segmentation result into
blocks, and the segmentation results of all pixels in the block jointly determine the label of the block. Since the small size of lesions, it is impossible to determine the label of the image block based on the proportion of diseased pixels, which is easy to miss the image block containing the small lesion. Therefore, we stipulate that there are lesion pixels in the image block, then the entire image block is regarded as a lesion.

We use many criteria to evaluate our method. The Dice coefficient is used to compare the agreement with neurologists’ segmentation results. It measures the overlap between our image block segmentation $X$ and the image block ground truth $Y$ and is defined as:

$$\text{Dice}(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|}$$  \hspace{1cm} (34)

IOU is similar to the Dice coefficient, which is defined as

$$\text{IOU}(X, Y) = \frac{|X \cap Y|}{|X \cup Y|}$$  \hspace{1cm} (35)

To make a comprehensive comparison of model performance, we used five additional pixel level indicators: accuracy, sensitivity, specificity, recall and F-Measure as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$  \hspace{1cm} (36)

$$\text{Sensitivity} = \frac{TP}{TP + FP}$$  \hspace{1cm} (37)

$$\text{Specificity} = \frac{TN}{TN + FP}$$  \hspace{1cm} (38)

$$\text{Recall} = \frac{TP}{TP + FN}$$  \hspace{1cm} (39)

$$F\text{-measure} = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta^2 \times (\text{Precision} + \text{Recall})}$$  \hspace{1cm} (40)

where $\beta$ is a coefficient that adjusts the proportion of precision and recall.

To prove the necessity of data processing, we used three kinds of data for training and testing: (1) Unprocessed original data. (2) Brain tissue data after tissue extraction. (3) Enhanced image data. Enhanced image data is to use the enhanced image as an image channel to participate in model training and testing. The experimental results of the three data are shown in Table 2.

According to the performance indicators shown in Table 2, we can see that compared with the unprocessed data directly involved in training, the F1 score of the model trained by tissue extraction data has been significantly improved, which shows that the skull and skin will damage the model Performance. We can also see that after adding the enhanced image channel, the F1 score part of the model has been improved. The global feature branch of SFN we proposed is similar to the structure of DeeplabV3, but the performance is much better than other models, mainly because the model structure uses symmetric characteristics.

Some segmentation results of typical scenes summarized above have been presented in Page 10. Figure 14(a) shows the case where there are no lesions on CT, one of the most common scenarios. Because the deep learning model has a strong ability to deal with details. On CT, the boundary between ischemic stroke lesions and healthy tissues is blurred, and the intracranial structure is complex, and there are differences in gray

FIGURE 13  Typical scene of ischemic lesion distribution in brain CT images
TABLE 2  Performance comparison of segmentation methods on our task

| Method       | Accuracy | Precision | Specificity | Recall | DICE  | IOU  |
|--------------|----------|-----------|-------------|--------|-------|------|
| Linknet      | 0.90     | 0.29      | 0.93        | 0.45   | 0.35  | 0.22 |
| FPN          | 0.81     | 0.81      | 0.99        | 0.43   | 0.56  | 0.39 |
| PAN          | 0.94     | 0.72      | 0.98        | 0.54   | 0.62  | 0.45 |
| PSPNet       | 0.91     | 0.41      | 0.94        | 0.56   | 0.48  | 0.31 |
| Unet         | 0.95     | 0.66      | 0.97        | 0.67   | 0.67  | 0.50 |
| DeepLabV3    | 0.95     | 0.75      | 0.98        | 0.67   | 0.71  | 0.55 |
| DeepLabV3+   | 0.95     | 0.69      | 0.97        | 0.71   | 0.70  | 0.54 |
| ours         | 0.96     | 0.72      | 0.97        | 0.77   | 0.75  | 0.59 |

FIGURE 14  Test results of LinkNet, FPN, PAN, PSPNet, Unet, DeepLabV3, DeepLabV3+ in typical scenarios features of some image blocks are similar to the lesions, PSPNet and Unet will make wrong judgments. In this general situation, SFCN can correct the edge part's recognition result by the characteristics of symmetric blocks. DeepLabV3 and DeepLabV3Plus have identified too large lesions, and FPN and PAN have not identified the lesions at the edge.
values. The traditional deep learning model may be constrained by details, resulting in misjudgement. As shown in the figure, LinkNet, FPN, PAN, PSPNet, Unet, and other networks have detected lesions at the same location. It shows that the features of this image block are similar to the lesions, and our proposed model effectively avoids this error.

CT images contain large-area lesions, which is another common situation. Due to the large proportion of lesions, the image difference is more obvious, so in this case, all models’ performance is better. The result is shown in Figure 14(b). However, because the local Table 2: Performance comparison of segmentation methods on our task.

Figure 14(c) shows the scene containing small lesion objects. This is a severe problem in the task of ischemic stroke localization detection. Because the lesion is too small, the grayscale change caused by the lesion is likely to be recognized by the model as a regular grayscale value change in the brain. At this time, the symmetric feature has an advantage over the network. It can amplify the difference through the difference feature, and the model is easy to identify the lesion. In this situation, our model works much better than other models.

In clinical practice, symmetric stroke lesions are rare. Nevertheless, this is also a situation we need to consider. We know that symmetrical lesions will cause our model to degenerate into a traditional model. Figure 14(d) shows a scene containing symmetrical lesions. The results show that the lesions identified in our model are tiny, mainly because symmetric lesions offset some symmetrical features. Although we can see that PSPNet works better, observations show that PSPNet works poorly in other more common situations.

From the above scenario analysis, we can see that SFCN performs better than other models in common scenarios. For rare symmetrical lesions, SFCN is not effective, which is a limitation of its model structure.

6 | DISCUSSION AND CONCLUSION

In this paper, we propose a practical symmetric feature comparison framework to Identify and locate ischemic stroke lesions on CT. It is tough for doctors to mark ischemic stroke lesions directly on CT. In our data processing framework, we first do some necessary processing to make CT slices have symmetrical characteristics. Second, we designed a CT nuclear magnetic mapping algorithm to construct CT-MRI image pairing and designed an image registration algorithm. Finally, labeling is performed on the registered MRI to obtain labelling information on CT data. CT ischemic stroke lesions can be divided into the following categories: without lesions, big lesion objects, small lesion objects, and symmetric lesion objects. In our method, we use the designed symmetric feature comparison network to extract image block-level features and differentiate them from the fast features of symmetric objects, thereby amplifying the difference between the lesion and healthy tissue. At the same time, the differential feature can avoid the interference of intracranial gray value differences. In order to evaluate the method’s performance, seven latest methods with comparative significance were selected for comparison. As shown in the experiment, these seven comparison methods cannot effectively distinguish the lesion from healthy tissue in the case of small lesions and no lesions, and the corresponding results are also poor. From these experimental results, it can be seen that the simple method is challenging to process challenging clinical ischemic stroke lesion identification and location on CT. According to visual experiments, our method is superior to other methods in small lesion objects and non-lesion images. Besides, the F-score indicates that our method is valid and reliable. The experimental results show that our method can help clinicians make a diagnosis.

ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China Grant 81972248, in part by the Beijing Municipal Natural Science Foundation Grant 7202056, in part by the Beijing Natural Science Foundation under Grant 4182044, in part by the Beijing Laboratory of Advanced Information Networks of BUPT, in part by the Beijing Key Laboratory of Network System Architecture and Convergence of BUPT, in part by the National Key Research and Development Program of China under Grant 2017YFB0802701.

ORCID

Long Zhang https://orcid.org/0000-0003-4124-9868

REFERENCES

1. Wikipedia, Stroke, https://en.wikipedia.org/wiki/Stroke, Accessed December 26, 2019
2. Stroke Severity and Mortality: Types, Treatments, and Symptoms, https://www.healthline.com/health/can-you-die-from-a-stroke, Accessed December 26, 2019
3. Liao, X.-L., et al.: T. Implementation, and M. of acute ischemic Stroke in China (TIMSCChina) Investigators, Implementation and outcome of thrombolysis with alteplase 3 to 4.5 h after acute stroke in Chinese patients. CNS Neurosci. Ther. 19, 43–47 (2013)
4. Lansberg, M.G., et al.: Comparison of diffusionweighted MRI and CT in acute stroke. Neurology 54, 1557–1561 (2000)
5. Smašljar, D., Šimn, O.: Sensitivity of the neuroimaging techniques in ischemic stroke. Medicinski Arhiv. 58, 282–284 (2004)
6. Zhang, L., et al.: Ischemic stroke lesion segmentation using multi-plane information fusion. IEEE Access 8, 45715–45725 (2020)
7. Shen, Y., et al.: Learning semantic representations using convolutional neural networks for web search. In Proceedings of the 23rd International Conference on World Wide Web, Seoul, Korea, pp. 373–374 (2014)
8. Crinion, J., et al.: Neuroimaging in aphasia treatment research: Quantifying brain lesions after stroke. Neuroimage 73, 208–214 (2013)
9. Ashton, E.A., et al.: Accuracy and reproducibility of manual and semi-automated quantification of MS lesions by MRI. J. Magn. Reson. Imag. 17, 300–308 (2003)
10. Yahiaoui, A.F.Z., Bessaid, A.: Segmentation of ischemic stroke area from CT brain images. In 2016 International Symposium on Signal, Image, Video and Communications (ISIVC), Tunis, Tunisia, pp. 13–17 (2016)
11. Gillebert, C.R., Humphreys, G.W., Mantini, D.: Automated delineation of stroke lesions using brain CT images. NeuroImage: Clin. 4, 540–548 (2014)
12. Lončarić, S.; Kočevčič, D.: A method for segmentation of CT head images. In: A. Del Bimbo (eds.), Image Analysis and Processing, Springer, Berlin, pp. 388–395 (1997)
13. Ruttmann, U., et al.: Fully automated segmentation of cerebrospinal fluid in computed tomography. Psychiatry Res. 50, 101–119 (1993)
14. Rebouças, E.d.S, et al.: New level set approach based on Parzen estimation for stroke segmentation in skull CT images. Soft Comput. 23(12), 9265–9286 (2018)
15. Randhawa, R.S., et al.: Improving boundary classification for brain tumor segmentation and longitudinal disease progression. In: International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Springer, Berlin, pp. 65–74 (2016)
16. Sajjadi, M., et al.: A promising method of enhancement for early detection of ischemic stroke. J. Res. Med. Sci. 17, 843–9 (2012)
17. Jung, S.-M., Whangbo, T.K.: Estimating aspect score for stroke from brain CT images based-on deep-learning. In: Advanced Science and Technology 2018, Ottawa, Canada, pp. 342–346 (2018)
18. Usinskas, A., Dobrovolskis, R., Tomandl, B: Ischemic stroke segmentation on CT images using joint features, informatica. Lith. Acad. Sci. 15, 283–290 (2004)
19. Tyan, Y.-S., et al.: Ischemic stroke detection system with a computer-aided diagnostic ability using an unsupervised feature perception enhancement method. J. Biomed. Imag. 2014, 19 (2014)
20. Maldjian, J., et al.: Automated CT segmentation and analysis for acute middle cerebral artery stroke. AJNR. Am. J. Neuroradiol. 22, 1050–5 (2001)
21. Rajini, N., Bhavani, R.: Computer aided detection of ischemic stroke using segmentation and texture features. Measurement 46, 1865–1874 (2013)
22. Wang, L., et al.: Taskoriented deep network for ischemic stroke segmentation in unenhanced CT imaging. In 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NNS/MIC), Manchester, pp. 1–3 (2019)
23. Rubin, J., Abulnaga, S.: CT-To-MR conditional generative adversarial networks for ischemic stroke lesion segmentation. In 2019 IEEE International Conference on Healthcare Informatics (ICHI), Xian, China, pp. 1–7 (2019)
24. Ho, K.C., et al.: Classifying acute ischemic stroke onset time using deep imaging features. AMIA Ann. Symp. Proc. 2017, 892 (2018)
25. Kabir, Y., et al.: Multimodal MRI segmentation of ischemic stroke lesions. In: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, pp. 1595–1598 (2007)
26. Khotanlou, H., et al.: 3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models. Fuzzy Sets Syst. 160, 1457–1473 (2009)
27. Clérisque, A., et al.: Acute ischemic stroke lesion core segmentation in CT perfusion images using fully convolutional neural networks. Comput. Biol. Med. 115, 103487 (2019)
28. Cai, W., et al.: The research of edge detection algorithm for Fingerprint images. In 2008 World Automation Congress (pp. 1-5). IEEE, (2008)
29. Duda, R., Hart, P.: Use of the hough transform to detect lines and curves in pictures. Commun. ACM 15(1), 11–15 (1975)
30. Chaurasia, A., Galarceillo, E.: LinkNet: Exploiting encoder representations for efficient semantic segmentation. In 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, Florida, pp. 1–4 (2017)
31. Lin, T.-Y., et al.: Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, Hawaii, pp. 2117–2125 (2017)
32. Li, H., et al.: Pyramid attention network for semantic segmentation, arXiv preprint arXiv:1805.10180 (2018)
33. Zhao, H., et al.: Pyramid scene parsing network. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, Hawaii, pp. 6230–6239 (2017)
34. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, Munich, Germany, pp. 234–241 (2015)
35. Chen, L.-C., et al.: Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. 40, 834–848 (2017)
36. Choudhury, A.R., et al.: Segmentation of brain tumors using DeepLabv3+. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Springer, Berlin, pp. 154–167 (2018)

How to cite this article: Zhang L, Zhu C, Wu YW, Yang Y, Liao Y, Song R, Liu L, Yang J. SCFCN: Symmetric feature comparison network for detecting ischemic stroke lesions on CT images. IET Image Process. 2021;1–15.
https://doi.org/10.1049/ipt2.12267.