Research on Medical Image Registration Based on Graphic Neural Network Reinforcement Learning

Zhejia Dong
School of mathematics and statistics, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, P.R. China
dongzhejia@stu.xjtu.edu.cn

Abstract. On the basis of traditional medical image registration methods, aiming at the shortcomings of rigid body images, sequential rigid body images and non-rigid body images registration methods, this paper proposes a new method of medical image registration application based on graphic neural network reinforcement learning. Reinforcement learning adjusts the strategy from the actual system learning experience, and it is a process of gradually approaching the optimal strategy, mapping the image from high-dimensional gray space to low-dimensional feature category space; Then, the elastic registration of mutual information images is carried out in the feature class space. The advantage of this method is that it does not need extra supervision information, which effectively saves manpower and material resources. So as to make an accurate diagnosis or make a proper treatment plan.

1. Introduction
Image registration is a basic technology of digital image processing. It is the basis of using information fusion method to analyze the scene shape by using a variety of sensor images in the same scene. Using computer technology, we can seek one or a series of spatial transformations for a medical image, and make it consistent with the corresponding points on another medical image in space[1]. In recent years, medical imaging technology is experiencing a leap from plane to stereo, from static to dynamic, from structure to function. Modern medical image data presents higher complexity and diversity. Image registration is a key step in image processing and analysis, and a necessary prerequisite for image comparison, data fusion, change analysis and target recognition. It is the basis of other image processing algorithms and applications, such as image segmentation and image reconstruction [2].

The conventional medical image registration process mainly includes five steps: feature space extraction, spatial transformation, image interpolation, similarity measure and iterative optimization. When the image is decomposed into subgraphs for rigid registration, due to the small number of sample points, the statistics are prone to deviation, which leads to the maximum value of mutual information not necessarily corresponding to the best registration parameters; Theoretically, the research on the convergence analysis of tabular algorithm is mature [3], but the research results on the convergence of enhanced learning algorithm based on value function approximation are few. In this paper, the nonlinear registration algorithm of medical images based on graphical neural network reinforcement learning is deeply studied by designing different methods of graphical neural network reinforcement learning.

2. Specific Method
Extract features. Generally, the completion of medical image registration is based on either the original image or the feature image. Because the graph neural network is simple to calculate the main
eigenvectors, and the main eigenvectors usually converge into an integral vector, the calculation of rotation angle is simplified and the calculation amount can be reduced as a whole. It should be noted that under normal circumstances, the number of edge pixels of two images to be registered cannot be completely consistent, but we can take the number of edge pixels of template image as the standard. Supervised learning can be regarded as a function approximation process, that is, the output of graphical neural network approaches the given function value of sample points. Feature extraction is a difficult process. For many images, the accuracy of image registration is directly affected by the accuracy of feature selection.

The gray feature value of an image can be simply expressed as the gray value of the image pixel. Specifically, the gradient vector \( \nabla_x(\sigma) \) (\( x \) is the pixel of the image, \( \sigma \) is the scale parameter of the high Gaussian kernel function, and \( \sigma = 0.5, N = 100 \) is taken in the experiment) is obtained by convolution of the first-order differential between the image and the Gaussian kernel, and the direction cosine of the gradient vector is [4]:

\[
\alpha_x(\sigma) = \arccos \frac{\nabla_x(\sigma)}{|\nabla_x(\sigma)|}
\]

In this way, the gradient characteristic function is defined as:

\[
g_x(\alpha) = \left[ \cos(2\alpha) + \frac{1}{2} \times (N - 1) \right] + 1
\]

Stereo matching is a classic problem in computer vision research. Using optimization method to solve corresponding problems is a trend in stereo matching research at present. Reflected in the image gray histogram, the histogram shows obvious bimodal distribution, and there is no overlap between the gray levels of two kinds of objects. Based on the center voxel position, image blocks with the same size are extracted from the original resolution image and the middle and low resolution image, and then image features are extracted based on each image block. Most existing multimodal image registration methods based on pixel or voxel similarity are based on joint histogram. The normalized joint histogram is also called the joint probability distribution of gray values of two images. The purpose of image registration is to find a mapping relationship, so that the pixel points of the floating image after coordinate transformation are consistent with the anatomical structure positions pointed by the pixel points in the reference image in space. Finally, when the iterative optimization end condition is reached, the spatial transformation formula and the transformed floating image are output, that is, the registered image. So far, the registration process has been completed, which shows that medical image registration is an iterative optimization process to solve the transformation parameters.

**Self-organizing map graph neural network classifier.** Adjacent neurons in the graph neural network can compete with each other through lateral interaction, and develop into special detectors to detect different signals adaptively [5]. It maps a straight line to a straight line and maintains parallelism. The concrete expression can be uniform scale transformation with uniform scale transformation coefficients in all directions or non-uniform scale transformation and shear transformation with inconsistent transformation coefficients. The behavior selection strategy should not only satisfy the condition of strategy iteration, that is, select the element with the largest behavior value function from the behavior set with a high probability, but also have a good continuity in the estimation of value function and the change of weight value. The system evaluates the output of the network, but the evaluation is not used to adjust the weight of the network immediately. After some delay, it provides several environmental evaluations to the graphical neural network in batches to adjust the weight. A pixel in a fixed image may or may not correspond to points in multiple floating images. In this case, if you want to calculate the gray value of the pixel in the image to be registered, you must perform interpolation, which is a fitting of unknown quantities.

Graph neural network is used as function estimation, which inputs the state of the system and outputs \( Q \) value or \( V \) value for reinforcement learning system. Its structure is shown in Figure 1:
Figure 1 The basic way of combining reinforcement learning with graphical neural network

Reinforcement learning using graphical neural network has two convergence processes at the same time:

- Strengthen the convergence of learning and graphical neural network, and remember that the function estimation operators $\hat{M}: \hat{Q} \rightarrow \hat{Q}$ and $\hat{Q}$ of graphical neural network are $Q$ functions generated by graphical neural network, and the sequence of value functions generated in this way is as follows:

$$Q_0, M(Q_0), M(T(M(Q_0))), M(T(M(T(M(Q_0))))), \ldots$$

Finally, the effect of reinforcement learning is related to these two convergence processes.

- Deformable transformation model can represent local deformation in medical images. Deformable image registration can be realized by various transformation models. It can be a foreign marker artificially added into the image, an anatomical structure point reflected in the image itself, or an inflection point with certain geometric features. Some are particularly effective in image registration between two modes, and the objective function is smooth, without local extremum, which can get the exact optimal solution. However, the image registration between other modes is poor, and even cannot be used at all. In the process of model training, the anatomical structures of two modes are considered comprehensively, and more accurate bidirectional image synthesis is realized. According to the relationship between gray level and threshold, pixels are judged as object points or background points, which is called image binarization. The segmentation effect of binary image can be obtained by further analysis [6].

**Registration method.** In digital image processing, the image is usually divided into grids with equal intervals, and the points at the center of the grids are called pixel points. With the fixed image and floating image registered by affine as input, the neural network layer and down-sampling layer appear alternately after connection. The point set of curves and surfaces extracted from one image contour is called hat, and the point set extracted from another image contour is called head, and then the hat and head are registered. This is because it can give consistent results and show excellent performance for different images. This is because it can give consistent results and show excellent performance for different images. For example, in image recognition, it is only necessary to input many different image samples and corresponding recognition results into the artificial graphic neural network, and the network will slowly learn to recognize similar images through self-learning function. The criterion is used to judge whether a point is an edge point, and the search criterion is used to guide how to search for the next edge point.

The joint entropy of images is a quantitative measure of the uncertainty of the relationship between two images, and the joint histogram is the way to statistics the uncertainty. A two-dimensional pixel strength state grid of $L_a \times L_b$ is established.

For images $A$ and $B$, their corresponding pixel pair is $(a, b)$. When making statistics of joint histogram, we only need to put this pixel pair $(a, b)$ in the established two-dimensional state grid.
according to the value of pixels, and accumulate the number $h_{ij}(a,b)$ of pixel pairs in each grid. The joint probability distribution $p_{ij}(a,b)$ is [7]:

$$p_{ij}(a,b) = \frac{h_{ij}(a,b)}{\sum_{i=1}^{L} \sum_{j=1}^{L} h_{ij}(a,b)}$$

(3)

Then the joint entropy of the image is:

$$H(A,B) = -\sum_{i=1}^{L} \sum_{j=1}^{L} p_{ij}(a,b) \log p_{ij}(a,b)$$

(4)

There is the following relationship between the joint probability distribution $p_{ij}(a,b)$ of the image and the marginal probability distribution $p_i(a)$ and $p_j(b)$:

$$p_i(a) = \sum_{j=1}^{L} p_{ij}(a,b)$$

(5)

$$p_j(b) = \sum_{i=1}^{L} p_{ij}(a,b)$$

(6)

The training process of the model itself also embodies the thought of reinforcement learning, which uses the graph neural network technology to calculate the possibility of the existence of lesions and complete the detection of lumps. Therefore, if an image is rotated, its rotation angle is equal to its rotation angle in the first main direction. By designing a deformation fusion model driven by dual-core, the registration information of two modes can be effectively fused in the registration process. Gradient term is not only based on the magnitude of gradient, but also based on the direction of gradient. Because the place with large gradient in the image is usually the place where the tissue jumps and contains the most mutual information. Because the structure of human brain and bone is relatively hard, these parts can be regarded as rigid bodies approximately, and images imaged in different directions can be registered by rigid body transformation, which has been widely studied [8].

Considering the data set $\{X_i | X_i = [x_i, y_i]^T, 1 \leq i \leq N\}$, according to physical knowledge, its centroid can be obtained by the following equation

$$X_C = \frac{1}{N} \sum_i x_i$$

$$Y_C = \frac{1}{N} \sum_i y_i$$

(7)

The centroid $[X_C^f, Y_C^f]$ of the floating image can be obtained by equation (7), and the centroid $[X_C^r, Y_C^r]$ of the reference feature image can be obtained. The displacement is obtained by subtracting the centroid of the floating image from the centroid of the reference feature image, and its expression is as follows

$$delX = X_C^f - X_C^r$$

$$delY = Y_C^f - Y_C^r$$

(8)

delX and delY are displacements along $X$ axis and $Y$ axis, respectively.

Although the nearest neighbor interpolation method is simple and has a small amount of computation, its effect is the worst. The enlarged image is prone to mosaic and the reduced image is prone to distortion,
so bilinear interpolation algorithm is usually chosen in practical application. A typical anatomical landmark can be a punctate anatomical structure, for example, at the inflection point of the cochlea tip; The intersection of two linear structures; Bifurcation or intersection of blood vessels; Particular topological attribute on a surface; An identifiable part of a trench, etc. The final loss of the graphical neural network can not only be returned to the coordinates, but also be transmitted to the feature map, so that the spatial transformation parameters can be updated. Larger grid spacing means fewer control points, and more downsampling layers and neural network layers are needed. After the last downsampling layer, two additional neural network layers are needed. If a certain transformation is only aimed at a special sub-region of the image, then the transformation is local, such as nonlinear transformation based on spline function, elastic deformation and so on. Especially, using radial basis function to simulate the deformation of objects can be used to describe both global transformation and local transformation.

3. Registration Analysis
In registration, each image is associated with a coordinate system, and the external points are much easier to identify than the internal points in medical images. It is easy to visually detect the registration results by comparing the positions of the marks in the images. Because the signals will influence each other, the closer the time is, the greater the influence is. This operation has the characteristics of sparse interaction, equal representation and weight sharing. Each control point only affects a specific area in the image, that is, the change of the position of a control point only affects the change of several adjacent points. Gray level co-occurrence matrix not only contains the gray level statistical information of the image, but also reflects the spatial and directional information of gray level distribution, that is, the probability of a pair of adjacent pixels in different directions. The deformation field is updated iteratively by the estimated velocity field, which is the calculated displacement of each voxel under a specific similarity measure. However, after geometric transformation, the intensity points of the transformed image can not correspond to the integer grid points of the reference image one by one, so interpolation is needed to estimate the intensity values of pixels at grid points in the transformed image.

Figure 2 below shows the influence of image synthesis quality on the accuracy of the registration algorithm proposed in this paper. Among them, the implementation of image synthesis algorithm includes the single-objective random forest algorithm (ST-RF) for comparison, the algorithm proposed in this paper and the image synthesis algorithm optimized by one-layer self-iterative enhancement model (ACM1). Figure 2 shows the results of multi-modal registration experiments in the group.

![Figure 2](image.png)

Figure 2 The registration accuracy of CT and MRI in pelvic organs varies with the image synthesis accuracy
It can be seen from Figure 2 that compared with ST-RF, the algorithm proposed in this paper can synthesize and predict higher quality images, so it can obviously improve the registration accuracy. ACM model further improves the accuracy of image synthesis, so the performance of registration algorithm has been further improved. The experimental results can fully demonstrate that high-quality image synthesis results are helpful to achieve more accurate multi-modal image registration.

For medical image registration, the threshold segmentation image is taken as the characteristic image. Relying on these methods, many tumors with remarkable characteristics can be accurately detected [9]. In contrast, the internal point method is friendly to the subjects and is a full retrospective registration; The loss function is calculated by propagating the output results forward, and then the weights are updated by gradient descent method. The pixel values at grid points are estimated by weighting the intensity of selected pixels. The weight depends on the distance between adjacent points and grid points. It reflects the degree of randomness in the region, and contains comprehensive information about the change of direction, adjacent elements and amplitude of image gray scale. Combining with case-based reasoning technology can improve retrieval efficiency; Combined with rough set theory, it can select important symptoms before training network and improve training efficiency. It provides more abundant anatomical structure information for image registration, so that the multi-modal nonlinear registration algorithm driven by dual-core can be realized.

Figure 3 shows the registration results obtained by reusing the trained registration model in the actual registration problem. When the model is reused for the second time, the registration accuracy is obviously improved, which shows that the algorithm has the ability to estimate the displacement vector of large deformation. Subsequently, the registration accuracy gradually converges, and the local deformation in the image to be registered has been fully estimated.

Considering that the images to be registered often represent the same object in different imaging modes, two columns of random signals describing its edge contour should have strong correlation. Based on the estimated displacement vector of key points, the final complete dense deformation field is obtained through the trained interpolation network of deformation field. In this way, it is easy to get the rotation angle of the image, that is, the angle between the two directions. The matching result is not the global optimal solution, which is related to the matching strategy and network topology. The iterative process of this hierarchical multi-stage registration strategy can avoid the optimization falling into local optimum. The flatness and parallelism of the image can be preserved, that is, after affine transformation, the originally parallel lines remain parallel. In addition to rotation and translation, affine transformation allows cutting and scaling. Moreover, with the increase of image noise content and the decrease of image sample space, mutual information also appears multi-extremum and extreme deviation.

In addition, sometimes it is necessary to compare the image of the subject with the image of the same part of a typical normal person to determine whether the subject is normal or not; If it is abnormal, it may
be necessary to compare with the typical images of some diseases to determine whether the patients belong to the same kind. Comparing the difference between the registered image and the fixed image, the registration effect can be judged. Although visualization does not mean good results, they can be used to find bad registration results.

4. Summary
Medical image plays an immeasurable role in the diagnosis and treatment of modern medicine. Modern medical imaging system, which is integrated with advanced computer technology, started late, but its development prospect is immeasurable. This paper proposes a mutual information image registration method based on graphic neural network reinforcement learning, which is unique in that the calculation of mutual information is no longer based on the joint gray distribution space of registered images, but on the joint feature category distribution space of images. Under the condition of low spatial resolution, noise and partial image defect, the algorithm has the characteristics of fast calculation speed, high precision and strong robustness. With the further development of graphic neural network reinforcement learning technology, it will be widely used in the field of medical image processing and analysis.

References
[1] Liu Zhaohui, Li Minghao, Xiao Yanli, et al. Medical image registration based on deep learning [J]. Electronic Production, 2019, 000(018): 52-53.
[2] Zhang Jiagang, Li Daping, Yang Xiaodong. Deformation medical image registration algorithm based on deep convolution feature optical flow [J]. Journal of Computer Applications, 2020, 040(006):1799-1805.
[3] Shen Yanyan, Feng Hansheng. A 2D-3D registration algorithm for dual X-ray images based on neural network [J]. Chinese Journal of Medical Physics, 2020, 037(003):293-298.
[4] Dai Shenghong, Li Zhibin. Research on Ramp Control Deep Reinforcement Learning Algorithm Based on Image Convolutional Neural Network [J]. Road Traffic and Safety, 2019, 019(004):1-6.
[5] Li Qiong, Gan Yongjin, Ning Weilian. Medical image segmentation based on Kohonen neural network algorithm [J]. Electronic Test, 2020, 000(005): 65-67.
[6] Wang Lingjiao, Li Qian, Guo Hua. Research on Deep Learning Model of Face Emotion Recognition Based on Swish Activation Function [J]. Image and Signal Processing, 2019, 008(003): P.110-120.
[7] Chen L, Wu H, Cui X, et al. Convolutional neural network SAR image target recognition based on transfer learning Chen Lifu[J]. Chinese Space ence and Technology, 2018, 38(6):46-51.
[8] Li Haowen, Hu Fangxu, Bai Yanan, et al. Development of medical image registration software based on ITK and VTK[J]. Electronic Production, 2019, 000(017):46-47,30.
[9] Lin Jinchao, Pang Yu, Xu Liming, et al. Research progress of medical image processing based on deep learning[J]. Life Science Instruments, 2018, 16(Z1):47-56.