

## 基于循环一致生成对抗网络的多模态影像刚性配准

郭翌<sup>1</sup>, 吴香奕<sup>1</sup>, 吴茜<sup>2</sup>, 陈志<sup>1</sup>, 徐榭<sup>1</sup>, 裴曦<sup>1</sup>

1. 中国科学技术大学物理学院工程与应用物理系, 安徽 合肥 230026; 2. 安徽医科大学, 安徽 合肥 230032

**【摘要】目的:**采用循环一致生成对抗网络(CycleGAN)方法在保证医学影像刚性配准精度的同时,降低图像配准的时间以及训练数据获取难度。**方法:**首先对训练数据进行标准化与归一化,同时对图像进行重采样与剪裁,去除多余的空气部分。其次采用阈值法与扫描线法获取图像的外轮廓信息,基于CycleGAN建立两个生成器与两个判别器,生成器输入配准图像对和输出配准结果,判别器输入配准图像对和输出配准程度。在原始CycleGAN损失函数基础上,增加轮廓损失项,以约束网络训练方向,提高收敛速度。**结果:**选取75例腹部病例,其中65例作为训练数据集,10例作为测试数据集,配准结果与配准软件Elastix对比。计算测试图像集外轮廓Dice系数,配准前图像对的平均Dice系数为0.858,Elastix配准后的平均Dice系数为0.926,本方法配准后的平均Dice系数为0.925。配准时间上Elastix的平均配准时间为12.1 s,本研究方法的平均配准时间为0.04 s,加速比达到302。**结论:**本方法在保证图像配准精度的同时极大降低了图像配准所需的时间,提高了配准流程工作效率。除此之外,与其他深度学习网络相比,本方法不需要真实配准结果以及传统相似性测度。

**【关键词】**医学影像;刚性配准;多模态;循环一致;生成对抗网络

**【中图分类号】**R318

**【文献标志码】**A

**【文章编号】**1005-202X(2021)02-0198-06

## Rigid registration of multimodal images based on CycleGAN

GUO Yi<sup>1</sup>, WU Xiangyi<sup>1</sup>, WU Qian<sup>2</sup>, CHEN Zhi<sup>1</sup>, XU Xie<sup>1</sup>, PEI Xi<sup>1</sup>

1. Department of Engineering and Applied Physics, University of Science and Technology of China, Hefei 230026, China; 2. Anhui Medical University, Hefei 230032, China

**Abstract: Objective** To reduce the time of image registration while ensuring the rigid registration accuracy of medical images. **Methods** Firstly, the training data were standardized and normalized. At the same time, the images were re-sampled and clipped to remove the redundant air. Secondly, threshold method and scanning line method were used to obtain the outline information of the images. Based on cycle-consistent generative adversarial networks (CycleGAN), two generator networks and two discriminator networks were established. The generator networks received multimodal image pairs and output the registration results. The discriminator networks received registration results and output registration accuracy. To improve the convergence speed of CycleGAN, the contour loss was added to restrict the training direction of the network. **Results** 75 abdominal cases were selected, of which 65 cases were used as training data set and 10 cases were used as test data set. Results of registration were compared with those of the registration software Elastix. The average Dice coefficient of the image pairs before registration was 0.858, that after Elastix registration was 0.926, and that after CycleGAN registration was 0.925. The average registration time of Elastix is 12.1 s. However, CycleGAN took 0.04 s to finish image registration, and the acceleration ratio was 302. **Conclusion** The proposed method not only ensures the accuracy of image registration, but also greatly reduces the time required for image registration and improves the efficiency of the registration process. In addition, compared with other deep learning methods, this method does not need real registration results and similarity measures.

**Keywords:** medical images; rigid registration; multimodal; cycle-consistent generative adversarial networks

**【收稿日期】**2020-09-13

**【基金项目】**安徽省自然科学基金(1908085MA27);国家自然科学基金(11575180);安徽省重点研究与开发计划(1804a09020039);安徽省高校自然科学基金项目(KJ2019A0240)

**【作者简介】**郭翌,在读博士,研究方向:医学影像配准与可视化,E-mail: gyi@mail.ustc.edu.cn

**【通信作者】**裴曦,E-mail: xpei@ustc.edu.cn

## 前言

图像配准是大多数医学影像分析问题的一个基本步骤,广泛应用于医学诊断、手术放射治疗规划、图谱构建、增强现实、疗效分析评价和医学图像分割等<sup>[1-3]</sup>。医学影像配准分为刚性配准与非刚性配准,其中刚性配准方法通过寻找固定图像(Fixed Image)和浮动图像

(Moving Image)之间的旋转平移变换矩阵对齐两幅图像<sup>[4]</sup>。刚性配准不仅为进一步的非刚性配准提供前提基础、节省图像优化迭代的计算时间,同时能直观地显示不同模态之间图像的解剖结构差异,辅助医生进行准确勾画。传统的配准方法包括基于表面的方法、基于点的方法(通常基于解剖标记物)和基于体素的方法<sup>[5]</sup>。尽管基于表面的方法和基于点的方法在图像配准中仍占有一席之地,但基于体素的方法借助于计算机技术快速发展的优势而得到了广泛应用<sup>[5]</sup>。基于体素的方法的目标是通过计算两个输入图像之间的相似性获得几何变换参数,而不需要预先提取特征<sup>[6]</sup>。但是这些传统的配准方法往往需要迭代计算相似性度量如均方误差、互信息与归一化互信息(NMI)等,由于相似性度量在参数空间上的非凸性等问题,使得配准过程较为耗时且鲁棒性较差<sup>[1-2]</sup>。除此之外,其他方法如基于强度的特征选择算法通过提取出与强度对应的图像特征进行图像配准,然而提取出的特征在解剖学方面难以很好地对应<sup>[7-9]</sup>。

近年来,深度学习方法在图像分割、图像变换、图像分类和高分辨率图像重建等许多领域取得了巨大成功。已有文献报道深度学习方法在图像配准中的应用,证实该方法的可行性。Miao等<sup>[10]</sup>使用卷积回归神经网络(CNN)进行实时的数字重建放射影像和X-ray影像的刚性配准。Liao等<sup>[11]</sup>采用CNN和强化学习方法对计算机断层扫描影像(CT)和锥形束CT影像进行迭代刚性配准。Jaderberg等<sup>[12]</sup>提出空间变换网络进行输入图像的空间对齐。为了改进CNN网络鲁棒性不足的问题,Miao等<sup>[13]</sup>利用全卷积神经网络(FCN)代替CNN进行脊柱的2D/3D刚性配准。另外,Gutierrez-Becker等<sup>[14]</sup>采用有监督式的回归网络模型进行脑部核磁共振图像(MRI)的刚性配准。尽管上述网络模型比传统迭代计算配准方法提升了配准速度,但是这些网络模型大多是有监督网络(即需要预先配准好的真实结果)或者需要使用相似性度量参数,这从数据获取与配准结果精确度的角度上阻碍了网络模型的训练。

生成对抗网络(Generative Adversarial Networks, GANs)是一种深度学习网络模型,它至少由两个部分组成:生成器和判别器,使得生成的数据具有与真实数据相同的分布<sup>[15]</sup>。为了克服在某些应用中获取图像对的困难,Zhu等<sup>[16]</sup>和Isola<sup>[17]</sup>等提出了循环一致生成对抗网络(Cycle-consistent Generative Adversarial Networks, CycleGAN),该网络能够在没有成对训练数据的情况下,学习从输入到输出图像的映射。近年来,一些基于GANs的医学图像配准研究已经被报道。Tanner等<sup>[18]</sup>通过结合两种多模态图

像相似性测量NMI和模态独立邻域描述符(Modality Independent Neighborhood Descriptor, MIND)研究GANs在多模态医学图像配准中的应用,但是该研究仅将GANs用于生成中间图像,而非直接用于图像配准。Fan等<sup>[19]</sup>提出了一种对抗性相似网络自动学习相似性度量,用于训练脑部MRI配准。Mahapatra等<sup>[1]</sup>在成本函数中采用了新的约束条件,将GANs用于多模态医学图像配准,确保训练后的网络能够轻松生成真实、具有合理变形场的图像。尽管上述工作在一定程度上验证了GANs在图像配准方面应用的可行性,这些方法在神经网络训练过程中仍需要有真实配准结果对判别器进行训练,因此这些网络容易受到配准结果质量的影响。

本文提出一种基于CycleGAN的CT与MR图像3D刚性配准网络,该网络与其他图像配准神经网络相比,具有如下特点:(1)完全无监督性,即不需要预先配准的真实结果;(2)端对端网络,即不产生中间图像,只需输入待配准的图像即可得到配准结果;(3)不需使用传统的相似性度量参数。

## 1 材料与方法

### 1.1 数据预处理

神经网络训练前需要对图像数据进行预处理。良好的数据预处理能够帮助神经网络加速收敛,同时得到更精确的结果。本研究采用75例腹部病例作为训练与测试数据集,每个病例包含1幅CT图像与1幅MR图像。所有图像数据(包括CT与MR数据)均进行图像标准化处理,同时将所有病例进行图像重采样并进行图像剪裁。重采样后的图像分辨率为1 mm×1 mm×5 mm,图像剪裁后大小为400 mm×400 mm,去除多余的空气部分。为使图像细节更为突出,设置CT图像像素值的窗宽窗位分别为350和40,即将图像原本的CT值范围限制在(-135, 215)范围内,MR图像不变。完成以上步骤后,对图像数据进行归一化,将所有图像像素值映射到(-1, 1)范围内。所有图像数据分为训练数据集与测试数据集,其中65例病例作为训练数据集,10例病例作为测试数据集,并与Elastix软件配准结果进行比较。

### 1.2 网络训练流程

本研究的神经网络结构基于CycleGAN网络模型<sup>[16-17]</sup>,网络主要分为4个部分:两个生成器 $G_{AB}$ 、 $G_{BA}$ 和两个判别器 $D_{AB}$ 、 $D_{BA}$ ,其中A为浮动图像,B为固定图像。刚性变换参数为 $F$ ,则 $G_{AB}$ 产生由A到B的变换 $F_{AB}$ , $G_{BA}$ 产生由B到A的变换 $F_{BA}$ , $D_{AB}$ 和 $D_{BA}$ 分别用于评价配准后两幅图像的配准程度。整体网络训练流程如图1所示。

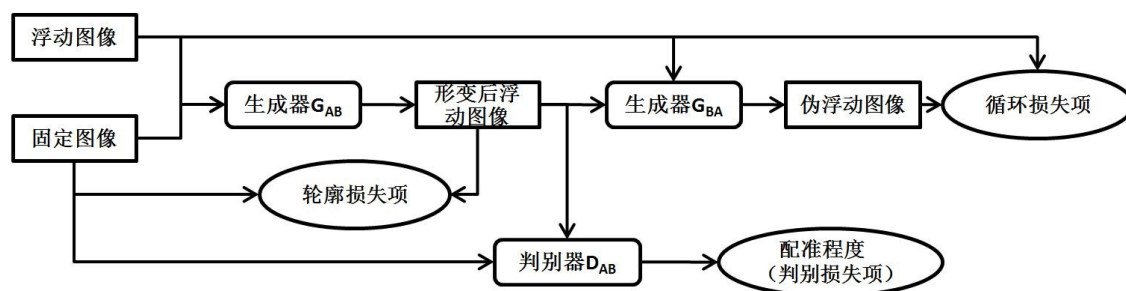


图1 网络训练流程示意图

**Fig.1 The flow chart of training process**

### 1.3 生成器网络

生成器的目的是根据输入的图像对得到刚性变换参数,具体结构如图2所示。网络输入为多模态3D图像对,输出为刚性变换参数与配准后的图像。

$G_{AB}$ 与 $G_{BA}$ 的网络结构一致,区别在于 $G_{AB}$ 中的固定图像为 $G_{BA}$ 中的浮动图像, $G_{AB}$ 中的浮动图像为 $G_{BA}$ 中的固定图像。生成器网络的输出结果作为判别器网络的输入之一。

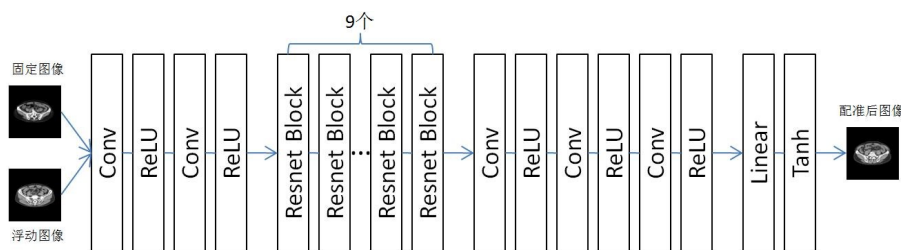


图2 生成器网络结构

Fig.2 The structure of the generator network

## 1.4 判别器网络

判别器网络的目的是根据输入的图像对判断两幅图像的配准程度,具体结构如图3所示。网络的输入为配准后的图像对,输出为图像对的配准程度。图像配准程度用0至1区间的数表示,输出结果越接

近1,表示图像配准程度越高,输出结果越接近0,表示图像配准程度越低。 $D_{AB}$ 与 $D_{BA}$ 的网络结构一致,区别在于 $D_{AB}$ 中的输入图像对为固定图像与配准后的浮动图像, $D_{BA}$ 中的输入图像对为浮动图像与配准后的固定图像。

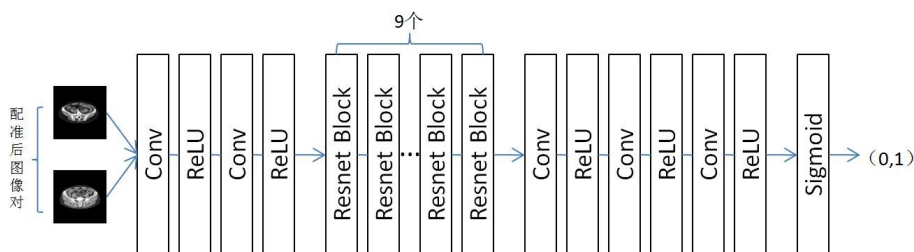


图3 判别器网络结构

Fig.3 The structure of the discriminator network

## 1.5 损失函数

神经网络训练涉及到损失函数的选择,不同的损失函数对网络训练结果具有较大影响。本研究在CycleGAN网络损失函数的基础上进行了改进,添加了新的损失项,能够对网络进行约束,使得网络更容易收敛。生成器网络的损失函数主要由3个部分组成:(1)判断损失项 $L_d$ 。该损失项与判别器相关,其

作用是使由生成器产生的配准图像对能够“瞒过”判别器,令判别器判定配准结果准确可靠。该损失项的具体公式如公式(1)所示,其中 $D_{AB}$ 为判别器,A表示固定图像,B表示配准后的浮动图像。(2)轮廓损失项 $L_2$ 。为使网络更易收敛,引入待配准图像的轮廓信息作为约束条件,使配准过程朝轮廓重叠的方向进行。其中轮廓信息为图像的外轮廓,损失项为配



准图像轮廓之间的 Dice 系数,如公式(2)所示,  $\text{Contour}_A$  和  $\text{Contour}_B$  分别为固定图像与配准后浮动图像的外轮廓信息。(3)循环损失项  $L_3$ 。大部分关于图像配准的神经网络只考虑前向传播过程而忽略反向过程,即只训练从浮动图像配准到固定图像,而忽略从固定图像配准到浮动图像,因此这样的网络不具备循环一致性,网络训练的结果不够可靠。为了克服这一缺点,CycleGAN 加入了循环损失项,保证了结果的一致性。如公式(3)所示,其中  $B'$  表示浮动图像,  $A$  表示固定图像。总的生成器网络损失函数  $L_G$  如公式(4)所示。除了生成器网络具有损失函数之外,判别器网络的损失函数如公式(5)所示,其中  $D$  代表判别器,公式的第1项表示由生成器产生的图像经判别器判断为假,公式第2项表示对于配准好的图像经判别器判断为真。

$$L_1 = \ln(|D_{AB}(A, B) - 1|) \quad (1)$$

$$L_2 = \ln(|\text{Dice}(\text{Contour}_A, \text{Contour}_B) - 1|) \quad (2)$$

$$L_3 = \ln(|G_{BA}(G_{AB}(A, B'), B') - B'|) \quad (3)$$

$$L_G = L_{\text{judge}} + L_{\text{contour}} + L_{\text{cycle}} \quad (4)$$

$$L_D = \ln(|D(A, B')|) + \ln(|D(A, A') - 1|) \quad (5)$$

在判别器训练过程中,需要为判别器提供判断依据,即输入参考结果。大部分基于 GANs 网络的配准工作均采用了预先配准好的医学图像对参考结果进行训练<sup>[1,19]</sup>。这些方法的缺点在于:(1)需要采用其他传统迭代方法对图像对进行预先配准,无法保证配准结果的准确性与稳定性。(2)网络训练的精度随预先配准结果的精度而变,且不能超过预先配准方法的配准精度。为了避免采用传统迭代方法进行预先配准的缺点,考虑在图像配准过程中影响配准结果的主要因素为图像对之间的像素分布情况和轮廓形状等,次要因素为图像对之间对应位置像素值之间的差异,因此构造出与固定图像具有相同像素分布的图像代替采用传统迭代方法得到的配准结果图像。公式(5)中  $A'$  表示与固定图像具有相同的像素分布,同时像素值不同的构造图像。

### 1.6 硬件平台

本研究中所采用的硬件设备参数如下:操作系统为 64 位 Windows10 操作系统,CPU 为 Intel i7-8700K,主频为 3.70 GHz,内存为 16 GB, GPU 为 NVIDIA GeForce GTX 1070 Ti,显存为 8 GB。由于计算机显存限制,在网络训练过程中需要限制输入图像大小。本研究对原始输入图像进行了重采样与剪裁,生成器与判别器的输入图像尺寸为  $128 \times 128 \times 24$ 。

## 2 结果

本研究共应用 75 例腹部病例作为训练与测试数据集,其中 65 例病例作为训练数据集,10 例病例作为测试数据集。采用 Python 与 Pytorch 作为神经网络搭建语言与框架,使用 Adam 作为网络的优化器<sup>[20]</sup>。我们将提出的方法与 Elastix<sup>[21-22]</sup>配准结果进行对比,详细比较了传统迭代配准方法与本方法之间的配准精度与配准时间。

### 2.1 网络训练结果

网络训练的学习率前 400 次循环为 0.000 1,400~600 次循环为 0.000 05,600~1 000 次循环为 0.000 01,总循环次数为 1 000。神经网络训练总共耗时 3 h。图 4 为网络训练结束后生成器损失函数曲线图,其中黄线代表轮廓损失项,绿线代表判断损失项,红线代表循环损失项,蓝线代表生成器总损失项。曲线图的横坐标表示训练迭代次数,纵坐标表示损失值。所有损失项均取对数。

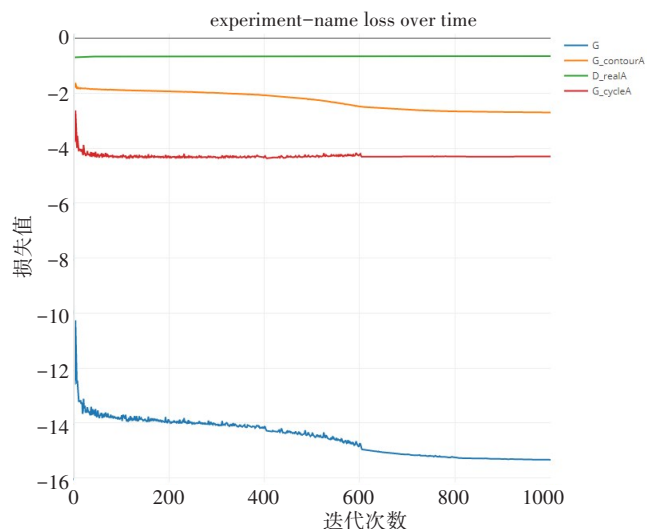


图4 神经网络生成器损失函数曲线图

Fig.4 Loss function curve of the network

可以看出,网络训练在前 20 次循环训练时损失项下降较为迅速,20~600 次循环训练之间下降速度变缓,600 次循环训练后损失函数趋于平缓。其中,轮廓损失项下降梯度较小,循环损失项对生成器总损失函数下降的贡献率较大,保证了网络训练结果是循环一致的。

### 2.2 结果对比

分别采用本研究提出的方法与 Elastix 对测试集图像数据进行 3D 刚性配准,分别计算未配准前与配准后图像对之间的 Dice 系数进行比较。10 例测试病例的配准前 CT 与 MR 图像之间外轮廓 Dice 系数与采

用不同方法配准后的图像对 Dice 系数:配准前 10 例测试病例的 Dice 系数为 0.858, Elastix 配准结果平均 Dice 系数为 0.926, 本研究方法配准结果平均 Dice 系数为 0.925。可以看出采用 Elastix 配准软件与本研究方法进行刚性配准后图像对的 Dice 系数均比配准前有所提高, 同时本研究方法的配准精度与 Elastix 配准结果相当。在配准耗时上, Elastix 软件配准平均耗时 12.1 s, 而本方法为 0.04 s, 平均加速比达到了 302。

由于医学影像往往在刚性配准后还需要进行形变配准, 因此快速有效的刚性配准方法能够缩短图像配准流程的时间, 提高效率。

图 5 显示配准前图像对以及采用不同方法配准后图像对之间的差异。其中差异图像中红色部分代表浮动图像, 绿色部分代表固定图像。对比不同刚性配准方法得到的结果, 可以看出 Elastix 配准结果与本研究方法配准结果基本一致。

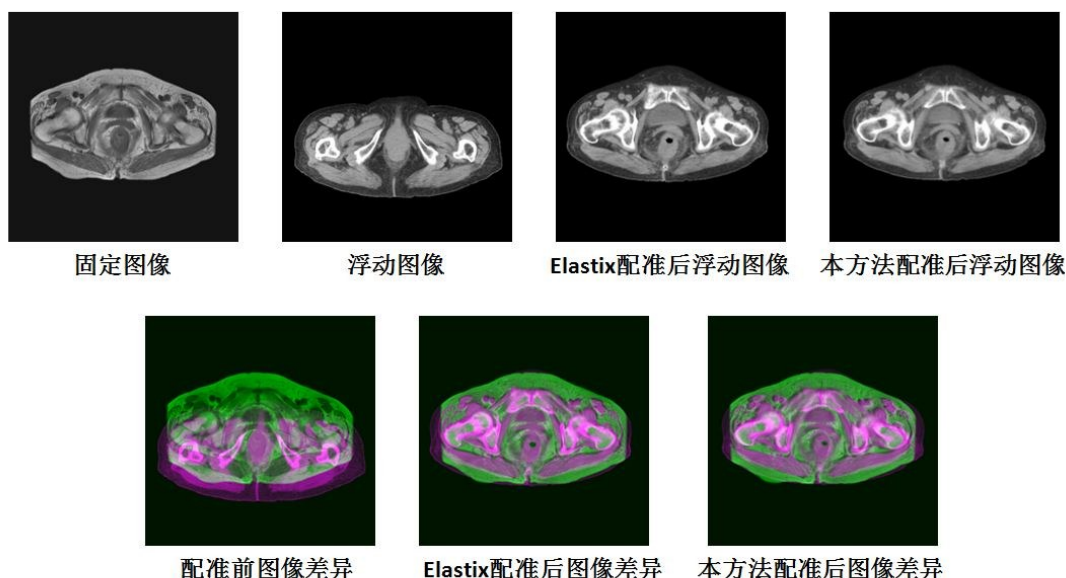


图5 配准前后图像以及图像差异

Fig.5 Image and image difference before and after registration

## 2.3 结果分析

上述结果表明采用 CycleGAN 方法进行腹部 3D 多模态影像刚性配准具有良好的应用性, 既保证了配准结果精度, 又极大缩短了配准时间。由于网络训练完全无监督, 不要求真实配准结果, 该网络易于扩展到其他部位多模态影像配准(如头颈部病例), 同时本研究将在刚性配准基础上进一步将网络拓展至多模态影像弹性配准。

## 3 讨论

传统的医学图像配准大多采用迭代计算方法, 相对耗时。同时迭代计算的相似性测度在参数空间上可能存在有非凸性问题, 因此在配准过程中往往需要根据不同情况选择合适的相似性度量以及优化方法才能达到满意的配准需求。本文提出了一种基于 CycleGAN 的 3D 多模态影像刚性配准方法, 相比于其他已有的图像配准网络, 该网络是端对端的无监督网络, 不需要预先配准的真实结果, 同时在网络训练过程中不需采用任何的相似性测量。在原有

CycleGAN 损失函数基础上, 本研究提出了新的函数损失项, 有助于神经网络朝预定方向训练, 同时增强了网络的收敛性。测试结果表明, 与传统迭代配准方法相比, 本研究提出的方法极大降低了图像配准时间, 保持了图像配准精度。

## 【参考文献】

- [1] MAHAPATRA D, SEDAI S, GARNAVI R. Elastic registration of medical images with GANs [J]. arXiv preprint arXiv, 2018: 1805.02369.
- [2] MAHAPATRA D, ANTONY B, SEDAI S, et al. Deformable medical image registration using generative adversarial networks[C]. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, 2018: 1449-1453.
- [3] SLOAN J M, GOATMAN K A, SIEBERT J P. Learning rigid image registration-utilizing convolutional neural networks for medical image registration[C]. 11th International Joint Conference on Biomedical Engineering Systems and Technologies. 2018: 89-99.
- [4] 皮一飞, 吴茜, 裴曦, 等. 基于掩膜优化的多模态医学图像刚性配准[J]. 中国医学物理学杂志, 2018, 35(9): 1022-1029.

PI Y F, WU Q, PEI X, et al. Rigid registration of multimodal medical images based on mask optimization[J]. Chinese Journal of Medical Physics, 2018, 35(9): 1022-1029.

- [5] VIERGEVER M A, MAINTZ J B A, KLEIN S, et al. A survey of medical image registration-under review[J]. *Med Image Anal*, 2016, 33: 140-144.
- [6] OLIVEIRA F P, TAVARES J M. Medical image registration: a review [J]. *Comput Methods Biomech Biomed Engin*, 2014, 17(2): 73-93.
- [7] KEARNEY V, HAAF S, SUDHYADHOM A, et al. An unsupervised convolutional neural network-based algorithm for deformable image registration[J]. *Phys Med Biol*, 2018, 63(18): 185017.
- [8] ROHLFING T. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable[J]. *IEEE Trans Med Imaging*, 2012, 31(2): 153-163.
- [9] WU G, KIM M, WANG Q, et al. Hierarchical attribute-guided symmetric diffeomorphic registration for MR brain images [C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Heidelberg: Springer, 2012: 90-97.
- [10] MIAO S, WANG Z J, ZHENG Y, et al. Real-time 2D/3D registration via CNN regression[C]. *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI 2016)*. IEEE, 2016: 1430-1434.
- [11] LIAO R, MIAO S, DE TOURNEMIRE P, et al. An artificial agent for robust image registration [C]. *Thirty-First AAAI Conference on Artificial Intelligence*. 2017.
- [12] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks[J]. *Adv Neural Inf Process Syst*, 2015: 2017-2025.
- [13] MIAO S, PIAT S, FISCHER P, et al. Dilated FCN for multi-agent 2D/3D medical image registration [C]. *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
- [14] GUTIÉRREZ-BECKER B, MATEUS D, PETER L, et al. Learning optimization updates for multimodal registration[C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2016: 19-27.
- [15] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. *Adv Neural Inf Process Syst*, 2014: 2672-2680.
- [16] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]. *Proceedings of the IEEE international conference on computer vision*. 2017: 2223-2232.
- [17] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks [C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, 2017: 1125-1134.
- [18] TANNER C, OZDEMIR F, PROFANTER R, et al. Generative adversarial networks for MR-CT deformable image registration[J]. *arXiv preprint arXiv*, 2018: 1807.07349.
- [19] FAN J, CAO X, XUE Z, et al. Adversarial similarity network for evaluating image alignment in deep learning based registration[C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2018: 739-746.
- [20] KINGMA D P, BA J. Adam: a method for stochastic optimization[J]. *arXiv preprint arXiv*, 2014: 1412.6980.
- [21] SHAMONIN D P, BRON E E, LELIEVELDT B P F, et al. Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease[J]. *Front Neuroinform*, 2014, 7: 50.
- [22] KLEIN S, STARING M, MURPHY K, et al. Elastix: a toolbox for intensity-based medical image registration [J]. *IEEE Trans Med Imaging*, 2009, 29(1): 196-205.

(编辑:黄开颜)