

文章编号:1007-2780(2024)08-1090-13

跨尺度点匹配结合多尺度特征融合的图像配准

欧卓林¹, 吕晓琪^{1,2*}, 谷 宇¹

(1. 内蒙古科技大学 信息工程学院, 内蒙古 包头 014010;
2. 内蒙古工业大学 信息工程学院, 内蒙古 呼和浩特 010051)

摘要: 图像配准在脑部疾病的计算机辅助诊疗和远程手术等方面具有重要作用。U-Net 及其变体网络广泛应用于医学图像配准领域, 在配准精确度和配准时间上取得了较好效果。然而, 现有的配准模型在处理复杂图像形变时, 难以学习到图像中微小结构的边缘特征, 且忽视了不同尺度上下文信息的关联。针对上述问题, 本文提出了一种基于跨尺度点匹配结合多尺度特征融合的配准模型。首先, 在模型的编码结构中引入跨尺度点匹配模块, 增强对图像突出区域特征的表达以及对微小结构边缘细节特征的把握; 然后, 在解码结构中对多尺度特征进行融合, 形成更全面的特征描述; 最后, 在多尺度特征融合模块中融入注意力模块, 突出空间和通道的信息。在 3 个脑部核磁共振 (Magnetic Resonance, MR) 数据集上的实验结果表明, 以 OASIS-3 数据集为例, 本文方法的配准精确度相较于 Affine、SyN、VoxelMorph 以及 CycleMorph 等方法, 本文方法分别提升了 23.5%、12.4%、0.9% 和 2.1%; ASD 值相较于各方法分别降低了 1.074、0.434、0.043 和 0.076。本文提出的模型能更好地把握图像的特征信息, 提升配准的精确度, 对医学图像配准的发展具有重要意义。

关键词: 医学图像配准; 编码器-解码器结构; 特征加权; 特征匹配; 注意力机制

中图分类号: TP391 **文献标识码:** A **doi:** 10.37188/CJLCD.2023-0278

Image registration combining cross-scale point matching and multi-scale feature fusion

OU Zhuolin¹, LÜ Xiaoqi^{1,2*}, GU Yu¹

(1. School of Information Engineering, Inner Mongolia University of Science and Technology,
Baotou 014010, China;

2. School of Information Engineering, Inner Mongolia University of Technology, Hohhot 010051, China)

Abstract: Image registration plays an important role in computer-aided diagnosis of brain diseases and remote surgery. The U-Net and its variants have been widely used in the field of medical image registration, achieving good results in registration accuracy and time. However, existing registration models have difficulty in learning the edge features of small structures in complex image deformations and ignore the correlation of contextual information at different scales. To address these issues, a registration model is proposed based on cross-scale point matching combined with multi-scale feature fusion. Firstly, a cross-scale point matching module is introduced into encoding structure of the model to enhance the

收稿日期: 2023-08-24; 修订日期: 2023-09-15.

基金项目: 国家自然科学基金 (No.62001255, No.61841204, No.61771266)

Supported by National Natural Science Foundation of China (No.62001255, No.61841204, No.61771266)

*通信联系人, E-mail: lxiaoqi@imut.edu.cn

representation of prominent region features and grasp the edge details of small structure features. Then, multi-scale features are fused in the decoding structure to form a more comprehensive feature description. Finally, an attention module is integrated into the multi-scale feature fusion module to highlight spatial and channel information. The experimental results on three brain Magnetic Resonance (MR) datasets show that, taking the OASIS-3 dataset as an example, the registration accuracy has been improved by 23.5%, 12.4%, 0.9%, and 2.1% compared to methods such as Affine, SyN, VoxelMorph and CycleMorph, respectively. The corresponding ASD values for each method have decreased by 1.074, 0.434, 0.043, and 0.076. The proposed model can better grasp the feature information of images, which improves registration accuracy and has important implications for the development of medical image registration.

Key words: medical image registration; encoder decoder structure; feature weighting; feature matching; attention mechanism

1 引言

医学图像配准技术^[1]在疾病检测、临床诊断、手术导航等方面广泛应用,该技术是寻找两幅或多幅图像之间的空间对应关系,利用空间对应关系将浮动图像映射到固定图像的空间域并对齐的过程。传统配准方法^[2-3]通过迭代优化最小化图像对像素之间的相似性度量,寻求最佳的空间对应关系,具有较好的精确度和鲁棒性。此类方法迭代寻优的过程运算量大,导致配准时间较长。

随着深度学习的方法在配准领域的广泛应用,基于卷积神经网络(Convolutional Neural Networks, CNN)的医学图像配准方法取得了极大的进步。Eppenhof等人^[4]通过人工合成具有待配准图像代表性的变形场,在CNN的训练过程中多次迭代约束新生成的变形场,实现了有监督的3D肺部图像配准。Hu等人^[5]仅通过标记图像对之间对应的解剖结构,完成了弱监督的多模态图像配准。在前列腺数据上的验证结果表明,该方法的配准精确度与传统配准方法相当,配准时间极大缩短。Balakrishnan等人^[6]提出了名为VoxelMorph的配准网络,该网络采用稳定的U-Net结构,通过直接预测生成的密集变形场实现了无监督的配准^[7]。但VoxelMorph处理较大或者更复杂的图像时,配准结果欠佳。Wang等人^[8]利用级联U-Net网络配准具有大位移和大形变的图像,其配准结果的优越性在肝脏CT图像任务中得到体现。上述配准网络逐步从基于监督的方法迈向无监督的方法,并逐渐优化网络中既有的缺陷,但仍旧缺乏对网络中不同尺寸特征之间关联性的

利用。Yang等人^[9]利用全尺寸跳跃连接将U-Net编码结构中不同尺寸的特征结合起来以更准确地预测变形场。简单地融合不同尺寸的特征虽然能够获得更丰富的细节信息,却增加了冗余信息,且难以体现跨尺寸特征间的相互联系。

基于上述问题,本文在U-Net的基础上提出了一种跨尺度点匹配结合多尺度特征融合的图像配准模型CM-RegNet,用于3D脑部MR图像的配准。首先,在编码结构的下采样卷积块中巧妙地结合提出的跨尺度点匹配模块(Cross-scale Point Matching Module, CPM),在生成不同分辨率的特征图的同时,增强对图像突出区域特征以及解剖结构边缘细节特征的表达能力,削弱背景信息及冗余特征信息的干扰^[10];其次,在解码结构中,提出了多尺度特征融合模块(Multi-scale Feature Fusion Module, MFF),融合多个分支中不同尺寸的卷积块生成的描述不同细节的特征,形成更全面的特征描述;最后,在多尺度特征融合模块的分支中添加注意力模块(Concurrent Spatial and Channel Squeeze and Excitation, SCSE),对不同分支的特征在通道和空间维度的表达进行加权。

2 基本原理

2.1 配准网络

提出的CM-RegNet配准框架如图1所示,固定图像 F 和浮动图像 M 拼接后作为网络的输入。通过CM-RegNet学习输入图像对之间的相似性,产生空间映射的变形场 φ 。随后利用空间变

换网络(Spatial Transformer Network, STN)^[11]将 φ 作用于 M 生成形变图像 $M(\varphi)$ 。最后将 $M(\varphi)$ 与固定图像 F 进行相似性度量。变形场 φ 通过CM-RegNet对参数函数 $g_{\theta}(F, M) = \varphi$ 进行学习而得, θ 即是函数 g 的可学习的参数。通过最小化损失函数找到最优的相似性度量参数,此过程如公式(1)所示:

$$\theta = \arg_{\theta} \min \left(L_{\text{Sim}} \left(F, M(\varphi) \right) + \lambda L_{\text{Smooth}}(\varphi) \right), \quad (1)$$
式中,损失函数 L 由相似性损失函数 $L_{\text{Sim}}(\cdot, \cdot)$ 和正则化损失函数 $L_{\text{Smooth}}(\cdot)$ 组成,即最小化固定图像 F 和形变图像 $M(\varphi)$ 之间的相似性损失以及变形场 φ 的正则化损失。相似性度量的计算与变形场的正则化程度呈反比关系,超参数 λ 用于平衡两者^[12]。

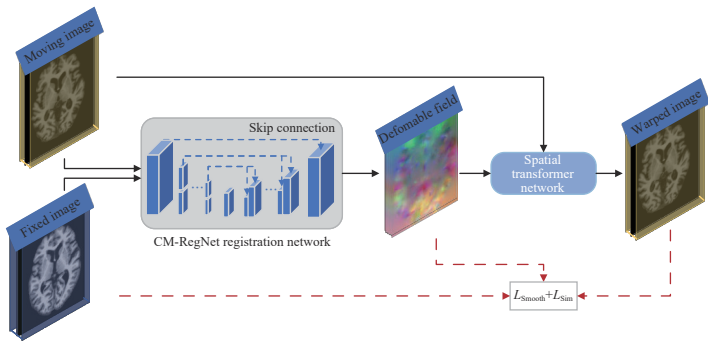


图1 配准网络的整体结构图

Fig. 1 Overall structure diagram of the registration network

2.2 CM-RegNet配准网络

本文提出的CM-RegNet网络结构如图2所示,采用与VoxelMorph类似的稳定且对称的编码器-解码器结构作为网络主干,并将所有池化层替换为步长为2的采样卷积层。

将两幅大小为 $160 \times 192 \times 224$ 组成的双通道3D图像对作为网络的输入,经过1个步长为2、卷积核为 $3 \times 3 \times 3$ 的采样卷积层和一个步长为1的不变卷积层,再经过LeakyReLU激活函数并利用Batch Normalization对图像特征进行批量归

一化处理,生成的特征图尺寸缩减为输入尺寸的1/2。将生成的特征图进行两步处理:其一是继续沿着网络结构进行图像缩减的下采样操作,生成尺寸为输入尺寸1/4的特征图,并将其作为跨尺度点匹配的第一个输入;其二是直接将1/2大小的特征图作为跨尺度点匹配的第二个输入,跨尺度点匹配模块利用不同尺寸的特征图,建立跨尺度特征图之间相似特征的关系,并输出1/4大小的特征图。利用相同的处理方式,生成尺寸为1/8和1/16的特征图。在解码阶段,将瓶颈处获

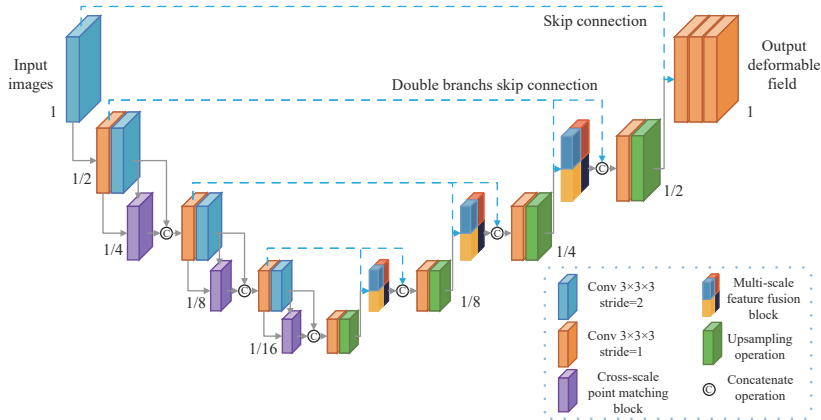


图2 跨尺度点匹配结合多尺度特征融合网络的结构图

Fig. 2 Structure diagram of cross-scale point matching combined with multi-scale feature fusion network

得的大小为 $1/16$ 的特征图通过一个步长为 1、卷积核为 $3 \times 3 \times 3$ 的不变卷积和一个最邻近插值的上采样层,生成 $1/8$ 大小的特征图。将特征图和跳跃连接传送的对应尺寸的特征图在通道维度拼接后送入多尺度特征融合模块,并在多尺度特征融合模块的内部再次拼接跳跃连接传送的对应尺寸的特征图。经过两次类似的操作可生成尺寸为 $1/2$ 的预测变形场。将 $1/2$ 的预测变形场上采样并结合利用跳跃连接传送的对应大小的特征图,最后经过 3 个步长为 1 卷积核为 $3 \times 3 \times 3$ 的不变卷积层逐渐缩减通道数,生成通道数为 3 的尺寸与输入图像相同的变形场 φ 。将 φ

通过 STN 作用于浮动图像 M ,即生成形变图像 $M(\varphi)$ 。

2.3 跨尺度点匹配模块

图像中不同尺度特征图之间存在一定的相似性^[13-14]。低分辨率图像中保留着高分辨率图像的细节信息,建立不同尺度特征图之间的联系可以获得更加真实、精确的特征细节^[15]。在 Li 等人^[16]的启发下,本文提出了一种双分支跨尺度点匹配模块,建立了相邻尺寸的特征图之间信息的相互联系,增强了对图像边界区域及结构内部较突出的特征,同时消除了冗余的信息。提出的跨尺度点匹配模块结构如图 3 所示。

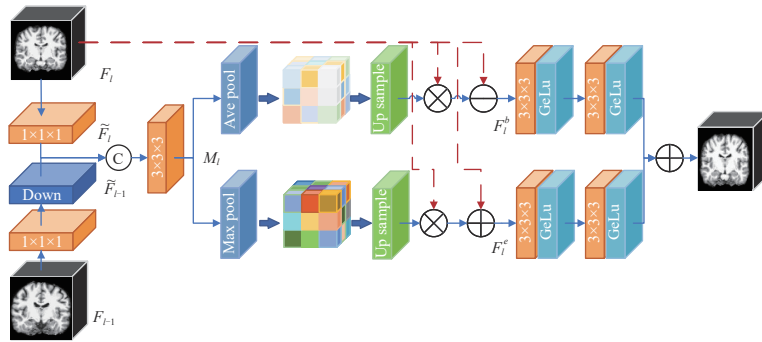


图3 跨尺度点匹配模块的结构图

Fig. 3 Structure diagram of cross-scale point matching module

将高分辨率特征 F_{l-1} 及相邻层次的低分辨率特征 F_l 送入模块。对 F_l 使用 $1 \times 1 \times 1$ 的卷积操作产生低分辨率、高语义的特征 $\tilde{F}_l$, 同时对 F_{l-1} 使用 $1 \times 1 \times 1$ 的卷积操作产生高分辨率、低语义的特征图。将此特征图进行下采样,生成与低分辨率特征 $\tilde{F}_l$ 相同尺寸的保留了高分辨率、低语义信息的特征 $\tilde{F}_{l-1}$ 。将 $\tilde{F}_l$ 和 $\tilde{F}_{l-1}$ 在通道维度进行拼接,利用一个 $3 \times 3 \times 3$ 的卷积操作计算局部区域内相邻点之间的相关性,并通过把 2 个通道缩减为 1 个通道的操作将两个相关性特征图中的点进行一对一匹配,再经过一个 $\text{Sigmoid}(\cdot)$ 激活函数生成显著性关联的特征映射 M_l ,该过程可由式(2)表示:

$$M_l = \text{Sigmoid}\left(\text{conv}_1\left(\text{Concat}\left(\tilde{F}_l, \tilde{F}_{l-1}\right)\right)\right). \quad (2)$$

在拉普拉斯金字塔中,将平滑处理后的图像与原始图像相减即可得到原始图像的边缘部分^[17]。在特征映射 M_l 上使用平均池化,并将池化后的

映射作用于 F_l 以达到平滑处理的效果,然后从 F_l 中减去平滑部分以生成锐化图像边缘的特征 F_l^b , F_l^b 可用公式(3)表述:

$$F_l^b = F_l - \text{AvgPool}(M_l) \times F_l, \quad (3)$$

同时,对特征映射 M_l 进行最大池化操作以获得最显著的特征映射,并将此映射作用于 F_l 。为了缓解因池化造成的特征损失,将获得的表征图像突出区域的特征添加在原始特征 F_l 上生成特征图 F_l^c ,此过程用式(4)表述:

$$F_l^c = F_l + \text{MaxPool}(M_l) \times F_l, \quad (4)$$

将 F_l^b 和 F_l^c 通过 2 个带有 GeLu 激活函数和批量归一化的 $3 \times 3 \times 3$ 卷积层,在加深网络结构的同时描述更丰富的语义特征信息。最后将两个描述不同重点的特征图中的特征点按照对应的空间位置关系进行求和,生成关联不同尺寸、能够强调边缘及组织结构内部较突出区域的特征图。跨尺度点匹配模块增强了网络对各解剖结构特征细节的学习能力,更有利于突出输入图像

对之间各个结构空间位置的对应关系,为变形场的形成提供更充分的信息。

2.4 多尺度特征融合模块

多尺度卷积通过多个不同的分支来处理相同或不同尺寸及层次的特征信息,大幅增强了感受野范围,在建立起局部信息相关性的同时捕捉全局图像的背景信息^[18]。在形成配准图像密集变形场的过程中,需要大量关注局部和全局的特征信息及图像的空间映射关系。因此,在文献^[18-19]的启发下,提出了多尺度特征融合模块,利用不同尺寸的卷积核生成不同感受野的特征信息,为变形场的形成提供充足的细节信息,且融合后的信息能够减少不同通道特征之间的语义差距。多尺度特征融合模块的结构如图 4(a)所示。依托网络结构,将上采样后的特征与跳跃连接传送的相同尺寸的特征相乘并通过 ReLU 激活函数,生成通道层面表征两个特征图关联性的特征映射 M_i ,并将 M_i 并行通过多尺度卷积的分支。多尺度卷积的分支分别由卷积核大小为 1、3、5、7 的 3D 卷积层构成,1×1×1 的卷积层能够增加网络的线性表达,使网络对特征的表达更丰富,生成对应的中间特征映射 M_{i1} ,尺寸大小分别为 3、5、7 的卷积层逐渐扩大网络的感受野,提高模型对全局特征的把控,同时生成对应的中间特征映

射 M_{i3} 、 M_{i5} 和 M_{i7} 。4 个中间特征映射叠加后通过 Sigmoid 激活函数,并重新作用到上采样后的特征以减少语义差距,最后再和跳跃连接传送的相同尺寸的特征拼接并通过一个 3×3×3 的不变卷积层将两个特征图融合并输出。

为了更好地让网络理解空间和通道的特征信息,在多尺度特征融合模块的 4 个分支位置(图 4(a)中的灰色虚线框),引入了同时关注空间和通道的 SCSE 模块^[20]。SCSE 模块的结构如图 4(b)所示,3D SCSE 模块由两个分支结构和一个跳跃连接组成。上分支特征图通过全局平均池化层产生 C 个通道的权重系数,通过一个带有 LeakyReLU 激活函数的非线性操作以及一个带有 Sigmoid 激活函数的非线性操作后产生关于通道的掩码。上分支计算过程如式(5)所示:

$$S = \sigma \left(\omega \left(\delta \left(\omega \left(\frac{1}{H \times W \times D} \sum_{i=1}^H \sum_{j=1}^W \sum_{k=1}^D u_c(i, j, k) \right) \right) \right) \right), \quad (5)$$

其中: $\omega(\cdot)$ 表示线性运算, $\delta(\cdot)$ 和 $\sigma(\cdot)$ 分别是 LeakyReLU 激活函数和 Sigmoid 激活函数, H 、 W 和 D 分别表示特征图的高度、宽度以及深度, $u_c(i, j, k)$ 表示对应高度、宽度和深度为 i 、 j 、 k 的特征点。下分支将特征图通过通道数为 1、卷积核

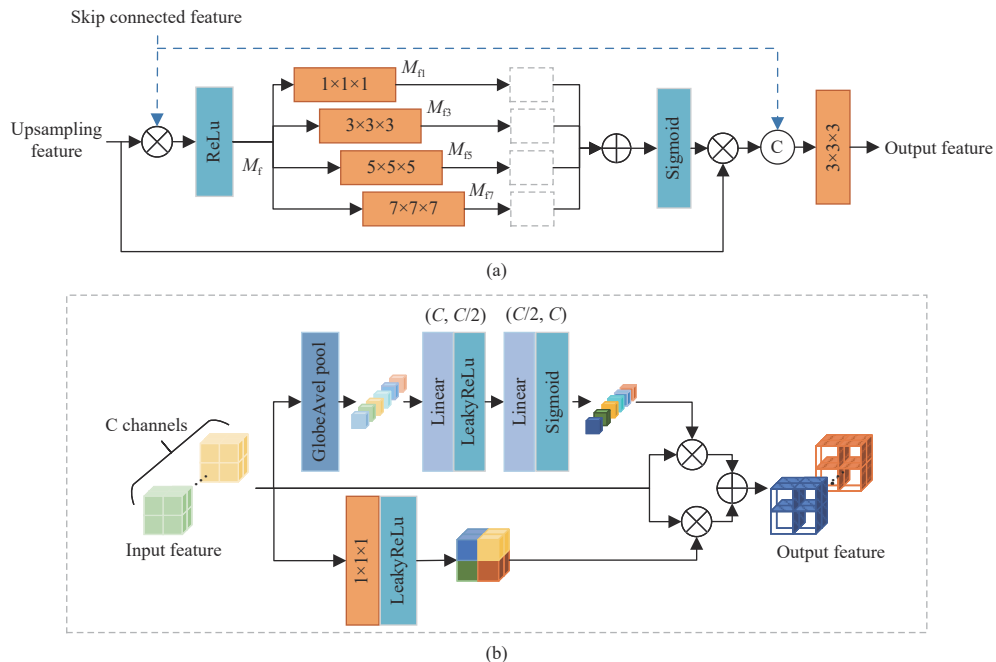


图 4 多尺度特征融合模块 (a) 及 SCSE 模块 (b) 的结构图

Fig. 4 Structure diagram of the multi-scale feature fusion module (a) and SCSE module (b)

为 $1 \times 1 \times 1$ 的卷积层以及一个 LeakyReLU 激活函数产生不同通道空间位置的非线性注意力映射。将通道的掩码和空间的注意力映射利用跳跃连接对输入的特征图进行加权并累加得到最终更具优势的输出。

2.5 损失函数

实验使用相似性损失函数 $L_{\text{Sim}}(\cdot, \cdot)$ 和正则化损失函数 $L_{\text{Smooth}}(\cdot)$, 其中 $L_{\text{Sim}}(\cdot, \cdot)$ 约束固定图像 F 和形变图像 $M(\varphi)$ 之间外观和空间位置上的差异, $L_{\text{Smooth}}(\cdot)$ 约束变形场 φ 局部空间的变换。采用均方体素差 MSE 作为相似性度量函数, 具体表达形式如式(6)所示:

$$\text{MSE}(F, M(\varphi)) = \frac{1}{\Omega} \sum_{p \in \Omega} \|F(p) - M(\varphi(p))\|^2, \quad (6)$$

其中: Ω 为输入图像对的空间域, p 为空间域中的体素值, 即在整个空间域 Ω 中计算固定图像 F 和形变图像 $M(\varphi)$ 对应的每个体素 p 的均方差值。在变形场空间的梯度上使用扩散正则化函数, 具体表达形式如式(7)所示:

$$L_{\text{Smooth}}(\varphi) = \sum_{p \in \Omega} \|\nabla \varphi(p)\|^2, \quad (7)$$

其中: $\|\nabla \varphi(p)\|^2$ 用于计算变形场中空间体素 p 一阶导数的平方, 即对空间域 Ω 中所有体素 p 的一阶导数求平方和用于平滑变形场中较为突出的区域。

3 实验与结果分析

3.1 数据集与实验参数

实验在3个公开的脑部MR数据集 OASIS-3^[21]、LPBA40^[22] 和 ABIDE^[23] 上进行。OASIS-3 和 ABIDE 数据集通过 FreeSurfer^[24] 图像处理工具进行预处理, 即对刚性的颅骨进行剥离并去除多余的脖颈部分, 将剩余的非刚性脑部组织重采样到 1 mm 体素的网格中, 并对皮层下结构进行分割及仿射变换等, 得到大小为 $255 \times 255 \times 255$ 的 3D 图像。处理后的图集带有 43 个解剖结构的分割标签, 对图像中多余的边缘区域利用居中裁剪处理, 生成 $160 \times 192 \times 224$ 的图集。LPBA40 数据集是已经预处理的大小为 $160 \times 192 \times 160$ 的图集, 且带有 56 个解剖结构的分割标签。实验数据

集的划分依次为: OASIS-3 数据集使用 1 610 张图像用于训练, 50 张用于验证, 50 张用于测试; LPBA40 数据集中 30 张图像用于训练, 4 张用于验证, 5 张用于测试; ABIDE 数据集随机选择 50 张图像用于泛化实验。图 5 展示了 3 个数据集中随机挑选的一张样例图像, 各行分别对应图像横断面、矢状面和冠状面的二维剖面。

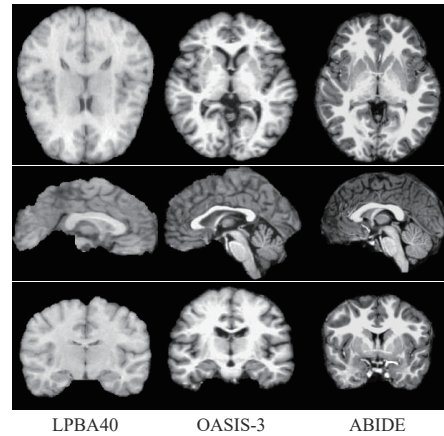


图5 3个数据集实验数据示意图

Fig. 5 Illustrates of experimental data of three datasets

实验利用 Python3.6 作为基础编程环境并使用 Pytorch 作为主体编程框架、使用 DGX NVIDIA v100 显卡进行深度学习, 学习率设置为 10^{-4} , 正则化损失函数 L_{Smooth} 的超参数 λ 设置为 0.1, 使用自适应矩估计优化器 (Adaptive Moment Estimation Optimizer, Adam) 对网络中的参数通过反向传播进行更新。采用基于 atlas 的训练方法, 即随机选取一张图像作为固定图像, 其他图像作为浮动图像与选取的固定图像进行配准。OASIS-3 数据集的训练次数设置为 100 epoch, 总迭代次数为 161 000; LPBA40 数据集的训练次数设置为 500 个 epoch, 总迭代次数为 15 000。每轮训练结束后都使用验证数据集对模型进行验证, 将验证集中精确度最好的模型作为最优模型。

3.2 评价指标

为评价固定图像和形变图像之间的相似性, 即图像配准的精确度, 实验采用 Dice 相似系数描述形变图像和固定图像空间位置的重叠程度, 该评价指标的计算过程如式(8)所示:

$$\text{Dice}(I_{\text{pre}}, I_{\text{fixed}}) = \frac{2 * |I_{\text{pre}} \cap I_{\text{fixed}}|}{|I_{\text{pre}}| + |I_{\text{fixed}}|}, \quad (8)$$

其中 I_{pre} 是形变图像的标签, I_{fixed} 是固定图像的标签。Dice 系数的取值介于 0~1 之间, 值为 1 表示两幅图像完全匹配, 为 0 则表示两幅图像完全不匹配。另外, 采用平均表面距离 (Average Surface Distance, ASD) 评估形变图像表面的光滑程度, 局部区域内异常突出的点会导致 ASD 数值变大, ASD 的数学公式如式 (9) 描述:

$$\text{ASD}(P_s, R_s) = \frac{1}{|P_s| + |R_s|} \times \left(\sum_{a \in P_s} d(a, R_s) + \sum_{b \in R_s} d(b, P_s) \right), \quad (9)$$

其中 a 和 b 分别表示固定图像 P_s 和形变图像 R_s 空间位置中的体素值, $d(x, y)$ 表示每个体素 x 到 y 的最小距离。ASD 数值越小, 表明形变图像中局部区域内各体素之间的位置差异越小, 形变图像表面越光滑。

3.3 对比实验分析

为了验证提出的 CM-RegNet 模型在脑部 MR 图像配准的有效性和准确性, 将 CM-RegNet 应用于 OASIS-3 和 LPBA40 数据集进行验证实验, 并与仿射配准 Affine、SyN^[2]、VoxelMorph^[6] 以及 CycleMorph^[25] 等 4 种方法进行对比实验, 其中 Affine 仅对图像进行仿射变换, SyN 是传统非线性配准方法中表现最优的, VoxelMorph 是 U 形卷积神经网络配准方法的代表, CycleMorph 是拓扑保持的且同样具有 U 形结构的多尺度配准模型。图 6 展示了几种配准方法在 OASIS-3 和 LPBA40 数据集上配准的可视化结果, 每种数据集中第一、第三和第四行分别对应图像的横断面、冠状面和矢状面, 第二行使用 HSV (Hue, Saturation, Value) 图像标记了横断面的局部结构, 便于更直观、更清晰地观测图像局部的变形。

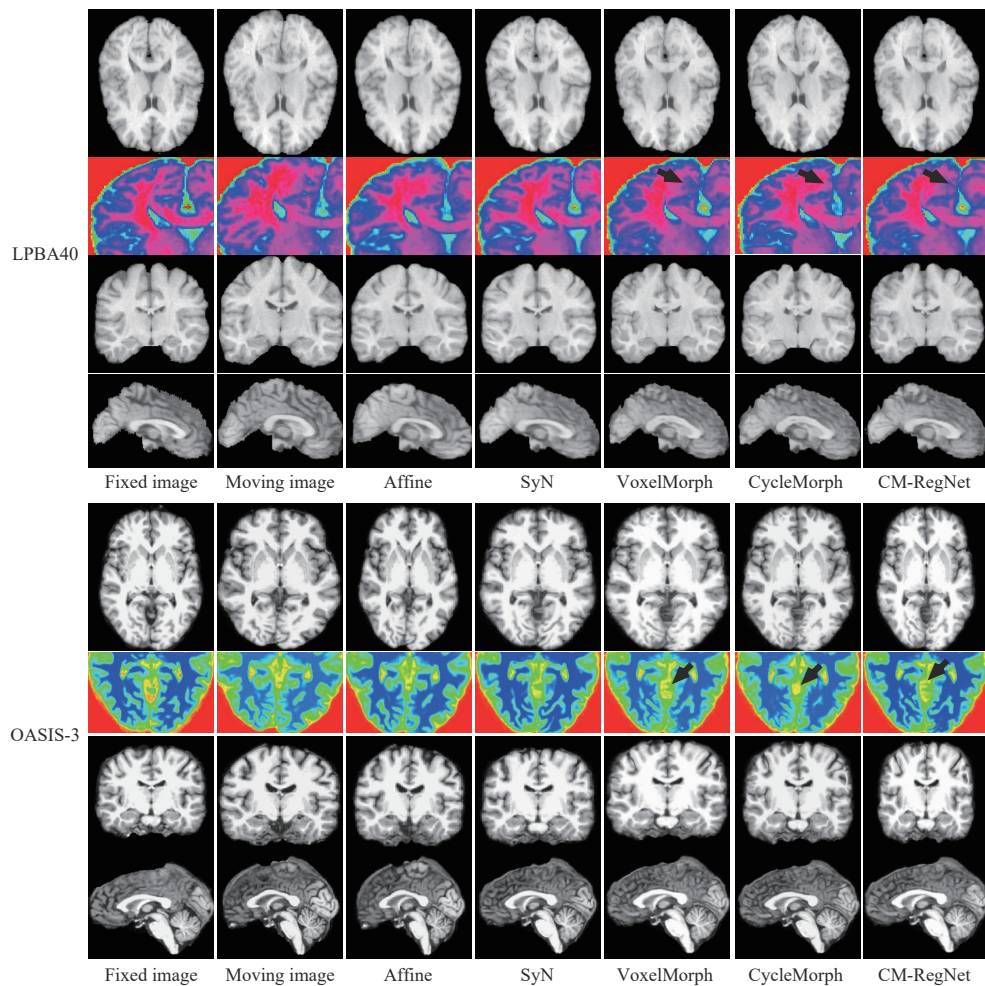


图 6 两个数据集上 5 种配准方法的结果展示

Fig. 6 Comparison of results from 5 registration methods on two datasets

从图6可以看出,仅使用Affine仿射配准方法生成的形变图像,其左右脑结构及大脑纵裂等与浮动图像更为相似。使用SyN生成形变图像的整体轮廓结构及沟壑等浅显明了,能大致体现脑部主要的功能区域,难以描述脑组织中各解剖结构精细的纹理。VoxelMorph、CycleMorph和提出的CM-RegNet在对配准图像纹理细节的描述上不同程度地优于前述的两种方法,如图6中箭头所示,CM-RegNet描述的细节信息相较于前两者更加丰富,更加精细。在OASIS-3数据集中,提出的网络能根据丰富的边缘细节特征信息大致描绘出小脑蚓及其周边的结构。表1给出了几种方法在2个数据集上所有解剖结构的平均Dice值和ASD值。在OASIS-3数据集上,提出的CM-RegNet在两个评价指标上都达到了5种方法的最优值,其中Dice值相较于Affine、SyN、VoxelMorph以及CycleMorph分别提升了23.5%、12.4%、0.9%和2.1%;ASD值对应于各方法分别降低了1.074、0.434、0.043和0.076。在LPBA40

数据集上,CM-RegNet仅次于SyN方法排在第二位,相较于其余3种方法同样有不错的提升,其Dice值相较于Affine、VoxelMorph以及CycleMorph分别提升了5.9%、1.3%和2.4%,ASD值对应于各方法分别降低了0.35、0.046和0.065。可视化和定量的结果都证实了CM-RegNet可以更好地把握脑部MR图像配准的细节信息,配准精确度有一定的提升。

表2展示了5种配准方法在两个数据集上的平均配准时间。值得注意的是,Affine和SyN方法基于CPU实现;其余配准方法基于GPU实现。从表中可以发现,基于CPU实现的Affine和SyN方法的配准时间较长,难以满足实时配准的需求;其余3种基于GPU实现的配准方法中,VoxelMorph在OASIS-3和LPBA40数据集上分别以0.15 s和0.14 s排名第一。提出的CM-RegNet平均配准时间略高于其他方法,在两个数据集上的平均配准时间分别为0.61 s和0.43 s,配准时间都低于1 s,具有在临床应用中实时配准的潜力。

表1 5种配准方法在两个不同数据集的比较(平均±标准差)

Tab. 1 Comparison of 5 registration methods on two different datasets (average±standard deviation)

Method	OASIS-3		LPBA40	
	Dice	ASD	Dice	ASD
Affine	0.481±0.021	1.870±0.110	0.594±0.012	1.620±0.146
SyN	0.592±0.014	1.230±0.085	0.664±0.011	1.220±0.098
VoxelMorph	0.707±0.012	0.839±0.062	0.640±0.028	1.316±0.126
CycleMorph	0.695±0.012	0.872±0.060	0.629±0.027	1.335±0.127
CM-RegNet	0.716±0.012	0.796±0.064	0.653±0.027	1.270±0.113

表2 5种配准方法在两个数据集的平均配准时间比较

Tab. 2 Comparison of the average registration time of 5 registration methods on two different datasets

Method	s	
	OASIS-3	LPBA40
Affine	5.39	3.58
SyN	23.50	15.69
VoxelMorph	0.15	0.14
CycleMorph	0.16	0.14
CM-RegNet	0.61	0.43

3.4 消融实验

为了验证CM-RegNet模型中提出的各模块的

效果,在OASIS-3和LPBA40两个数据集上进行了消融实验。将拥有稳定U形结构的VoxelMorph网络作为基础主干网络(Baseline),分别添加跨尺度点匹配模块(CPM)和多尺度特征融合(MFF)以及SCSE模块,得到两个数据集上的Dice值如表3所示。单独添加CPM对OASIS-3和LPBA40测试数据集图像的Dice值分别提升了0.002和0.011,验证了该模块可以更好地结合不同尺度之间特征的相关性,更好地捕捉突出区域和边缘轮廓的特征信息;提出的MFF在生成变形场的过程中,建立局部信息之间相关性的同时捕捉全局图像的图像信息,使得Dice值在两个数据集上分别提升了0.002和0.008,继续引入关注空间

和通道的 SCSE 模块后,特征细节在空间和通道维度上优势部分的加权得到了体现^[26],实现了跨通道信息交互,模块整体对 Dice 值分别提升了 0.006 和 0.010,整合所有模块后的 CM-RegNet 模型在两个数据集上的 Dice 值分别提升了 0.009 和 0.013,整个网络模型使得配准精确度达到了最优。

表 3 CM-RegNet模型消融实验结果比较(平均±标准差)

Tab. 3 Comparison of ablation experiment results of CM-RegNet model (average±standard deviation)

Baseline	CPM	MFF	SCSE	OASIS-3	LPBA40
✓				0.707±0.012	0.640±0.028
✓	✓			0.709±0.012	0.651±0.026
✓		✓		0.709±0.016	0.648±0.019
✓		✓	✓	0.713±0.011	0.650±0.021
✓	✓	✓	✓	0.716±0.012	0.653±0.027

实验还对两个数据集上添加各个模块后生成的配准图像进行展示,配准图像的可视化结果如图 7 所示,两个数据集中各行分别对应图像的冠状面、横断面和矢状面。从图中可以看出,随着各个模块的添加,形变图像相较于 Baseline 生成的图像在回、沟、脑白质轮廓及脑组织中细小的分支上都有不同程度的优化;蓝色方框和箭头标记了添加 CPM 模型后,图像在外部整体轮廓和边缘沟裂区域的形变效果提升较为明显,沟、裂区域的颜色更加突出;黄色方框和箭头描述了逐步添加 MFF 和 SCSE 模块后,由于图像对深层次语义细节特征能更好地学习,使得配准后图

像对细节轮廓的描述更加丰富、细致。结合所有模块后的 CM-RegNet 能够更具体地突出图像轮廓、表达更多的纹理细节。

3.5 解剖结构分析

为了探究 CM-RegNet 在配准时对各个解剖结构中细节特征的理解能力,对 OASIS-3 测试数据集中各图像的 43 个解剖结构配准后的 Dice 值进行计算。为了更简洁地描述相同的结构,将左右脑中相同的解剖结构进行合并,最终得到 27 个解剖结构的 Dice 值。图 8 展示了 27 个解剖结构 Dice 值的箱线图,可直观地发现, Affine 和 SyN 的箱体较宽,配准精确度的离散程度较大,且异常值较多,表现出一定的不稳定性;VoxelMorph、CycleMorph 和提出的 CM-RegNet 的配准精确度相对集中,且 CM-RegNet 在多数的解剖结构中拥有更少的异常值和更高的配准极值。在 27 个解剖结构中,CM-RegNet 在其中的 23 个结构都表现出最优的配准效果,其余 4 个最优的分布分别为: SyN 在 Pallidum 结构中表现最优, VoxelMorph 在 Optic Chiasm 和 CC Central 中表现出最好的配准精确度, CycleMorph 在 CC Mid Posterior 结构中展示出最佳的配准效果。由于数据较多,表 4 仅展示了 5 个常见解剖结构的平均 Dice 值,最佳的配准效果都由提出的 CM-RegNet 获得。值得注意的是,在区域最大的 White Matter 中, CM-RegNet 以 0.842 的准确度获得最优效果,相较于 Affine、SyN、VoxelMorph 以及 CycleMorph 分别提升了 0.244、0.185、0.011 和 0.017。即使在配准效果欠佳的 Vessel 区域,CM-RegNet 也以 0.308 的 Dice 值位列第一,而 SyN 配准方法在该区域配准的 Dice 值只有最低的 0.029。

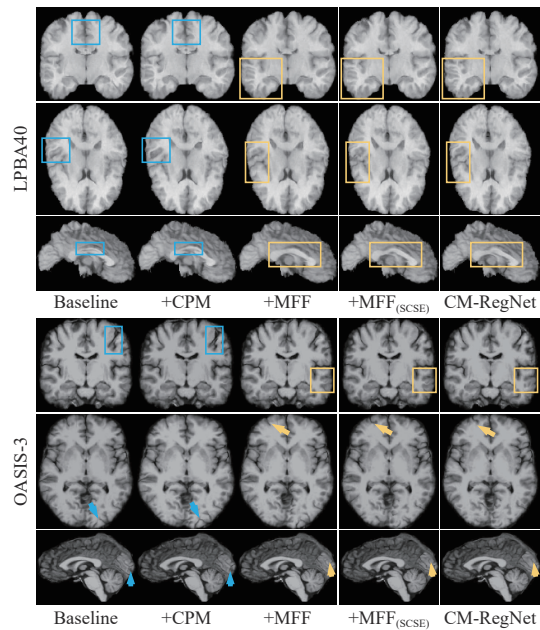


图 7 两个数据集上消融实验的可视化结果
Fig. 7 Visualization results of ablation experiments on two datasets

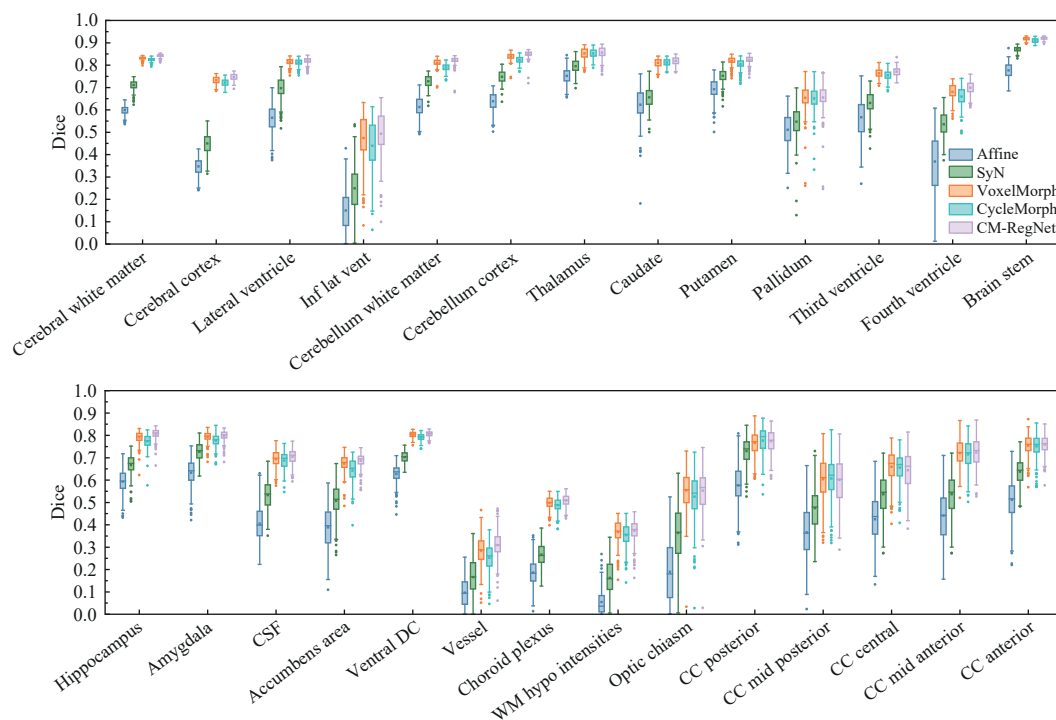


图 8 OASIS-3数据集中 27 个解剖结构 Dice 值的箱线图

Fig. 8 Box plot of Dice values for 27 anatomical structures in the OASIS-3 dataset

表 4 几种配准方法在 5 个解剖结构的精确度比较

Tab. 4 Comparison of accuracy of several registration methods in five anatomical structures

Method	Accuracy/%				
	White Matter	Inf. Lat. Vent.	Brain Stem	CSF	Vessel
Affine	59.8	14.9	77.6	40.5	9.6
SyN	65.7	13.8	85.6	51.0	2.9
VoxelMorph	83.1	47.4	91.7	69.3	28.4
CycleMorph	82.5	43.9	90.9	68.6	25.2
CM-RegNet	84.2	49.2	92.0	70.3	30.8

3.6 泛化实验

为探究 CM-RegNet 模型对没有学习过的脑部 MR 数据集的适应能力,进行了泛化实验,将几种方法在 OASIS-3 数据集获得的模型直接对尺寸相同且同样具有 43 个解剖结构的 ABIDE 数据集的 50 张脑部 MR 图像进行配准,并计算相应的 Dice 值和 ASD 值。几种方法的配准结果如表 5 所示,CM-RegNet 在 Dice 值评估中获得最好的效果,相较于 Affine、SyN、VoxelMorph 和 CycleMorph, Dice 精确度分别提升了 11.2%、0.4%、1% 和 2.8%。在 ASD 评价指标中,CM-RegNet 仅次于 SyN,位列第二;相较于 Affine、VoxelMorph 和 CycleMorph,获得了 0.397、

0.068 和 0.095 的提升。图 9 展示了配准后 ABIDE 数据集 43 个解剖结构的平均 Dice 值的

表 5 几种方法在 ABIDE 数据集的泛化性结果比较(平均±标准差)

Tab. 5 Comparison of generalization results of several methods on ABIDE dataset (average±standard deviation)

Method	Dice	ASD
Affine	0.462±0.036	1.698±0.121
SyN	0.570±0.020	1.054±0.074
VoxelMorph	0.564±0.126	1.369±0.444
CycleMorph	0.546±0.121	1.396±0.398
CM-RegNet	0.574±0.132	1.301±0.369

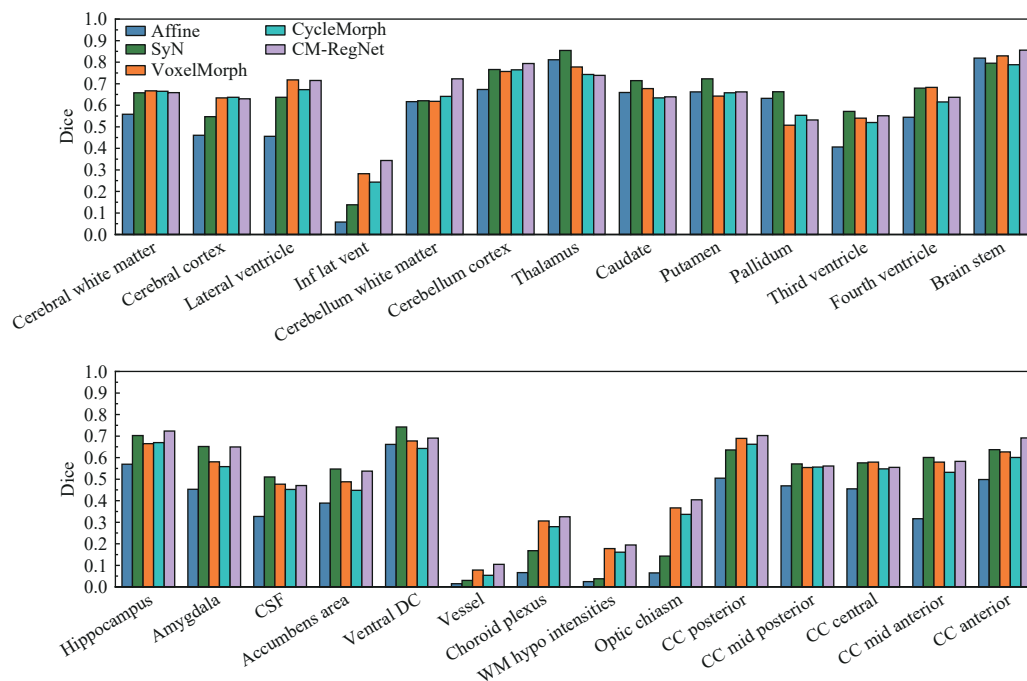


图9 ABIDE数据集中27个解剖结构平均Dice值的柱状图

Fig. 9 Bar chart of average Dice values for 27 anatomical structures in the ABIDE dataset

柱状图,同样将左右大脑中相同的结构合并,最终生成27个解剖结构。从柱状图可以直观地看出,CM-RegNet和SyN分别在11个解剖结构中优于其他方法,VoxelMorph在4个解剖结构上达到最优,CycleMorph仅在Cerebral Cortex结构中获得最优配准效果,即CM-RegNet模型具有较好的适应性、稳定性,对未知图像的学习能力具有明显优势。

4 结 论

三维医学图像配准在临床疾病诊疗、医学图像处理领域都具有重要研究意义。本文针对现有的基于编码器-解码器配准模型难以将不同尺度的上下文信息相结合并寻找其相互关系、不能充分利用采样的特征等问题,提出了一个关联不同尺度特征间的相关性并充分融合特征的CM-RegNet配准模型,实现了3D脑部图

像端到端的配准。本文提出的基于跨尺度点匹配结合多尺度特征融合的网络模型,利用相邻层次特征信息之间特征点的相关性进行点匹配,增强了各个解剖结构中特征的自我表达能力,在突出解剖结构边缘细节的同时增强了特征对纹理结构的描述能力。多尺度特征融合模块则通过跳跃连接、多尺度感受野以及SCSE模块对空间和通道加权等操作,丰富非线性特征的形成,突出网络中学习效果更好的通道,极大地促进更细节化的变形场的形成。提出的CM-RegNet在OASIS-3和LPBA40数据集上的配准精确度分别达到了71.6%和65.3%,在ABIDE数据集上泛化实验的配准精确度达到57.4%,所有的配准结果都比相同类型的主流配准方法更优。综合性实验的结果验证了CM-RegNet在医学图像配准领域对配准性能的提升,为该领域的脑部MRI图像配准提供了可行的方案。

参 考 文 献:

- [1] XIAO H N, TENG X Z, LIU C Y, *et al.* A review of deep learning-based three dimensional medical image registration methods [J]. *Quantitative Imaging in Medicine and Surgery*, 2021, 11(12): 4895-4916.
- [2] AVANTS B B, EPSTEIN C L, GROSSMAN M, *et al.* Symmetric diffeomorphic image registration with cross-

- correlation: evaluating automated labeling of elderly and neurodegenerative brain [J]. *Medical Image Analysis*, 2008, 12(1): 26-41.
- [3] FU Y B, LEI Y, WANG T H, *et al.* Deep learning in medical image registration: a review [J]. *Physics in Medicine & Biology*, 2020, 65(20): 20TR01.
- [4] EPPENHOF K A J, PLUIM J P W. Pulmonary CT registration through supervised learning with convolutional neural networks [J]. *IEEE Transactions on Medical Imaging*, 2019, 38(5): 1097-1105.
- [5] HU Y P, MODAT M, GIBSON E, *et al.* Weakly-supervised convolutional neural networks for multimodal image registration [J]. *Medical Image Analysis*, 2018, 49: 1-13.
- [6] BALAKRISHNAN G, ZHAO A, SABUNCU M R, *et al.* An unsupervised learning model for deformable medical image registration [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 9252-9260.
- [7] 林立昊, 易见兵, 曹锋, 等. 多尺度并行全卷积神经网络的肺计算机断层扫描图像非刚性配准算法[J]. 激光与光电子学进展, 2022, 59(16): 1617004.
- LIN L H, YI J B, CAO F, *et al.* Non-rigid registration algorithm of lung computed tomography image based on multi-scale parallel fully convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2022, 59(16): 1617004. (in Chinese)
- [8] 王鹏, 严赞琦, 钱黎俊, 等. 基于纹理滤波的无监督配准方法及其在肝脏电子计算机断层扫描中的应用[J]. 生物医学工程学杂志, 2021, 38(5): 819-827.
- WANG P, YAN Y Q, QIAN L J, *et al.* Texture filtering based unsupervised registration methods and its application in liver computed tomography images [J]. *Journal of Biomedical Engineering*, 2021, 38(5): 819-827. (in Chinese)
- [9] YANG Y F, WU H H. Deformable medical image registration based on CNN [J]. *Journal of X-Ray Science and Technology*, 2023, 31(1): 85-94.
- [10] 石甜甜, 郭中华, 闫翔, 等. 基于多尺度融合注意力改进UNet的遥感图像水体分割[J]. 液晶与显示, 2023, 38(3): 397-408.
- SHI T T, GUO Z H, YAN X, *et al.* Water body segmentation in remote sensing images based on multi-scale fusion attention module improved UNet [J]. *Chinese Journal of Liquid Crystals and Displays*, 2023, 38(3): 397-408. (in Chinese)
- [11] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks [C]//*Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal: MIT Press, 2015: 2017-2025.
- [12] 张明娜, 吕晓琪, 谷宇. 残差混合注意力结合多分辨率约束的图像配准[J]. 光学精密工程, 2022, 30(10): 1203-1216.
- ZHANG M N, LÜ X Q, GU Y. Image registration based on residual mixed attention and multi-resolution constraints [J]. *Optics and Precision Engineering*, 2022, 30(10): 1203-1216. (in Chinese)
- [13] GLASNER D, BAGON S, IRANI M. Super-resolution from a single image [C]//*Proceedings of the IEEE 12th International Conference on Computer Vision*. Kyoto: IEEE, 2009: 349-356.
- [14] ZONTAK M, IRANI M. Internal statistics of a single natural image [C]. The 2011 IEEE/CVF Computer Vision and Pattern Recognition Conference, CVPR 2011. Colorado Springs: IEEE, 2011: 977-984.
- [15] MEI Y Q, FAN Y C, ZHOU Y Q, *et al.* Image super-resolution with cross-scale non-local attention and exhaustive self-exemplars mining [C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 5689-5698.
- [16] LI X T, HE H, LI X, *et al.* PointFlow: flowing semantics through points for aerial image segmentation [C]//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021: 4215-4224.
- [17] ADELSON E H, ANDERSON C H, BERGEN J R, *et al.* Pyramid methods in image processing [J]. *RCA Engineer*, 1984, 29(6): 33-41.
- [18] FAN J H, BOCUS M J, HOSKING B, *et al.* Multi-scale feature fusion: learning better semantic segmentation for road pothole detection [C]. 2021 IEEE International Conference on Autonomous Systems (ICAS). Montreal: IEEE, 2021: 1-5.
- [19] YANG T J, BAI X H, CUI X J, *et al.* DAU-Net: An unsupervised 3D brain MRI registration model with dual-

- attention mechanism [J]. *International Journal of Imaging Systems and Technology*, 2023, 33(1): 217-229.
- [20] ROY A G, NAVAB N, WACHINGER C. Concurrent spatial and channel ‘squeeze & excitation’ in fully convolutional networks [C]. 21st International Conference on Medical Image Computing and Computer Assisted Intervention. Granada: Springer, 2018: 421-429.
- [21] LAMONTAGNE P J, KEEFE S, LAUREN W, *et al.* OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease [J]. *Alzheimers & Dementia the Journal of the Alzheimers Association*, 2018, DOI:10.1016/j.jalz.2018.06.1439.
- [22] SHATTUCK D W, MIRZA M, ADISETIYO V, *et al.* Construction of a 3D probabilistic atlas of human cortical structures [J]. *NeuroImage*, 2008, 39(3): 1064-1080.
- [23] DI MARTINO A, YAN C G, LI Q, *et al.* The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism [J]. *Molecular Psychiatry*, 2014, 19(6): 659-667.
- [24] FISCHL B. FreeSurfer [J]. *NeuroImage*, 2012, 62(2): 774-781.
- [25] KIM B, KIM D H, PARK S H, *et al.* CycleMorph: cycle consistent unsupervised deformable image registration [J]. *Medical Image Analysis*, 2021, 71: 102036.
- [26] 杨云,周瑶,陈佳宁. 基于多尺度混合卷积网络的高光谱图像分类[J]. 液晶与显示, 2023, 38(3): 368-377.
- YANG Y, ZHOU Y, CHEN J N. Hyperspectral image classification based on multi-scale hybrid convolutional network [J]. *Chinese Journal of Liquid Crystals and Displays*, 2023, 38(3): 368-377. (in Chinese).

作者简介:



欧卓林,男,硕士研究生,2020年于辽宁石油化工大学获得学士学位,主要从事图像处理、医学图像配准等方面的研究。E-mail: o2236554402@163.com



吕晓琪,男,博士,教授,2003年于北京科技大学获得博士学位,主要研究方向为医学图像处理、配准与融合技术、模式识别等。E-mail: lxiaoqi@imut.edu.cn