



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象 图形学报

2021
09
VOL.26

ISSN1006-8961
CN11-3758/TB



“深度学习+智慧医疗”专刊



第26卷第9期（总第305期）
2021年9月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@aircas.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

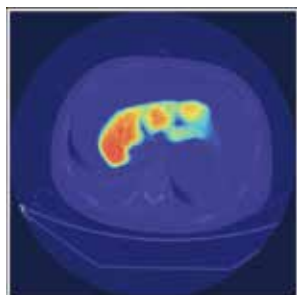
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

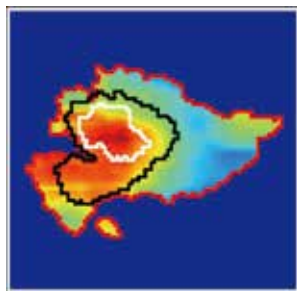
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



肺部影像解剖结构分割数据集及应用(第2111页)



LFSCA-UNet: 基于空间与通道注意力机制的肝纤维化区域分割网络(第2121页)



语义拉普拉斯金字塔多中心乳腺肿瘤分割网络(第2193页)

编者按 I

综述

深度医学图像配准研究进展: 迈向无监督学习

马露凡, 罗凤, 严江鹏, 徐哲, 罗捷, 李秀 2037

U-Net网络医学图像分割应用综述

周涛, 董雅丽, 霍兵强, 刘珊, 马宗军 2058

医学图像深度学习技术: 从卷积到图卷积的发展

唐朝生, 胡超超, 孙君顶, 司马海峰 2078

多尺度变换像素级医学图像融合: 研究进展、应用和挑战

周涛, 刘珊, 董雅丽, 霍兵强, 马宗军 2094

数据集论文

肺部影像解剖结构分割数据集及应用

覃文军, 李小硕, 周庆华, 刘盼, 杨金柱 2111

计算机断层扫描图像

LFSCA-UNet: 基于空间与通道注意力机制的肝纤维化区域分割网络

陈弘扬, 高敬阳, 赵地, 吴忌, 陈金军, 全显跃, 李欣明, 薛峰, 周沐瑶, 柏冰冰 2121

融合上下文和多尺度特征的胸部多器官分割

吉淑滢, 肖志勇 2135

注意力机制下密集空洞卷积的肺部图像分割

郭宁, 柏正尧 2146

融合注意力机制和特征金字塔网络的CT图像肺结节检测

张福玲, 张少敏, 支力佳, 周涛 2156

改进Faster R-CNN模型的CT图磨玻璃密度影目标检测

杨淑莹, 邓东升, 郑清春 2171

融合注意力机制与可变形卷积的多尺度骨病变检测

方成, 柏正尧 2181

磁共振图像

语义拉普拉斯金字塔多中心乳腺肿瘤分割网络

王黎, 曹颖, 郭顺超, 唐雷, 鄧子翔, 王荣品, 王丽会 2193

引入注意力机制和多视角融合的脑肿瘤MR图像U-Net分割模型

罗恺锴, 王婷, 叶芳芳 2208

集成注意力增强和双重相似性引导的多模态脑部图像配准

田梨梨, 程欣宇, 唐堃, 张健, 王丽会 2219

LRUNet: 轻量级脑肿瘤快速语义分割网络

何康辉, 肖志勇 2233

自适应多模态特征融合胶质瘤分级网络

王黎, 曹颖, 田梨梨, 陈祈剑, 郭顺超, 张健, 王丽会 2243

研究应用

面向运动想象脑电图识别的镜卷积神经网络

罗靖, 王耀杰, 刘光明, 王晓帆, 鲁晓峰, 黑新宏 2257

先验权重共享码本下内窥镜图像大肠病变分类

朱霆威, 李胜, 何熊熊 2270

融合空洞卷积与注意力的胃癌组织切片分割

陈颖锲, 李晗, 周雪婷, 万程 2281

视觉显著性的眼底图像视盘检测

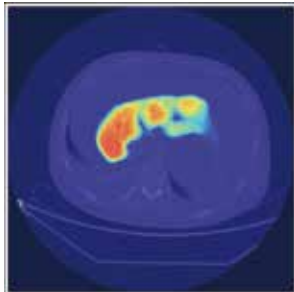
吕鹏飞, 王莹, 王思齐, 于晓升, 吴成东 2293

CONTENTS

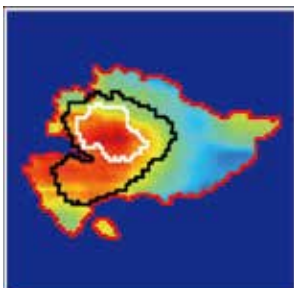
JOURNAL OF IMAGE AND GRAPHICS



Pulmonary image anatomical structure segmentation dataset and applications(P2111)



LFSCA-UNet: liver fibrosis region segmentation network based on spatial and channel attention mechanisms(P2121)



Semantic Laplacian pyramids network for multicenter breast tumor segmentation(P2193)

Review

- Deep-learning based medical image registration pathway: towards unsupervised learning
Ma Lufan, Luo Feng, Yan Jiangpeng, Xu Zhe, Luo Jie, Li Xiu 2037
- U-Net and its applications in medical image segmentation:a review
Zhou Tao, Dong Yali, Huo Bingqiang, Liu Shan, Ma Zongjun 2058
- Deep learning-based medical images analysis evolved from convolution to graph convolution
Tang Chaosheng, Hu Chaochao, Sun Junding, Sima Haifeng 2078
- Research on pixel-level image fusion based on multi-scale transformation: progress application and challenges
Zhou Tao, Liu Shan, Dong Yali, Huo Bingqiang, Ma Zongjun 2094

Dataset

- Pulmonary image anatomical structure segmentation dataset and applications
Tan Wenjun, Li Xiaoshuo, Zhou Qinghua, Liu Pan, Yang Jinzhu 2111

Computerd Tomography Image

- LFSCA-UNet: liver fibrosis region segmentation network based on spatial and channel attention mechanisms
Chen Hongyang, Gao Jingyang, Zhao Di, Wu Ji, Chen Jinjun, Quan Xian Yue, Li Xinming, Xue Feng, Zhou Muyao, Bai Bingbing 2121
- Integrated context and multi-scale features in thoracic organs segmentation
Ji Shuying, Xiao Zhiyong 2135
- The integration of attention mechanism and dense atrous convolution for lung image segmentation
Guo Ning, Bai Zhengyao 2146
- Detection of pulmonary nodules in CT images by combining an attention mechanism and a feature pyramid network
Zhang Fuling, Zhang Shaomin, Zhi Lijia, Zhou Tao 2156
- Ground-glass opacity target detection in CT scans based on improved Faster R-CNN model
Yang Shuying, Deng Dongsheng, Zheng Qingchun 2171
- Multi-scale bone lesion detection based on attention mechanism and deformable convolution
Fang Cheng, Bai Zhengyao 2181

Magnetic Resonance Image

- Semantic Laplacian pyramids network for multicenter breast tumor segmentation
WANG Li, CAO Ying, Guo Shunchao, Tang Lei, Kuai Zixiang, Wang Rongpin, Wang Lihui 2193
- U-Net segmentation model of brain tumor MR image based on attention mechanism and multi-view fusion
Luo Kaikai, Wang Ting, Ye Fangfang 2208
- Multimodal brain image registration with integrated attention augmentation and dual similarity guidance
Tian Lili, Cheng Xinyu, Tang Kun, Zhang Jian, Wang Lihui 2219
- LRUNet: a lightweight rapid semantic segmentation network for brain tumors
He Kanghui, Xiao Zhiyong 2233
- Adaptive multi-modality fusion network for glioma grading
Wang Li, Cao Ying, Tian Lili, Chen Qijian, Guo Shunchao, Zhang Jian, Wang Lihui 2243

Research and Application

- Mirror convolutional neural network for motor imagery electroencephalogram recognition
Luo Jing, Wang Yaojie, Liu Guangming, Wang Xiaofan, Lu Xiaofeng, Hei Xinhong 2257
- Colorectal lesion classification of endoscopic images based on priori weight shared codebook application
Zhu Tingwei, Li Sheng, He Xiongxiang 2270
- The fusing of dilated convolution and attention for segmentation of gastric cancer tissue sections
Chen Yingsi, Li Han, Zhou Xueting, Wan Cheng 2281
- Optic disc detection based on visual saliency in fundus image
Lyu Pengfei, Wang Ying, Wang Siqi, Yu Xiaosheng, Wu Chengdong 2293

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2021)09-2037-21

论文引用格式: Ma L F, Luo F, Yan J P, Xu Z, Luo J and Li X. 2021. Deep-learning based medical image registration pathway: towards unsupervised learning. *Journal of Image and Graphics*, 26(09): 2037-2057 (马露凡, 罗凤, 严江鹏, 徐哲, 罗捷, 李秀. 2021. 深度医学图像配准研究进展: 迈向无监督学习. *中国图象图形学报*, 26(09): 2037-2057) [DOI:10.11834/jig.200361]

深度医学图像配准研究进展: 迈向无监督学习

马露凡¹, 罗凤¹, 严江鹏¹, 徐哲^{1,2}, 罗捷², 李秀^{1*}

1. 清华大学深圳国际研究生院, 深圳 518055; 2. 哈佛医学院, 美国波士顿 02115

摘要: 在疾病诊断、手术引导及放射性治疗等图像辅助诊疗场景中, 将不同时间、不同模态或不同设备的图像通过合理的空间变换进行配准是必要的处理流程之一。随着深度学习的快速发展, 基于深度学习的医学图像配准研究以其耗时短、精度高的优势吸引了研究者的广泛关注。本文全面整理了 2015—2019 年深度医学图像配准方向的论文, 系统地分析了深度医学图像配准领域的最新研究进展, 展现了深度配准算法研究从迭代优化到一步预测、从有监督学习到无监督学习的总体发展趋势。具体来说, 本文在界定深度医学图像配准问题和介绍配准研究分类方法的基础上, 以相关算法的网络训练过程中所使用的监督信息多少作为分类标准, 将深度医学图像配准划分为全监督、双监督与弱监督、无监督医学图像配准方法。全监督配准方法通过采用随机变换、传统算法和模型生成等方式获取近似的金标准作为监督信息; 双监督、无监督配准方法通过引入图像相似度损失、标签相似度损失等其他监督信息以降低对金标准的依赖; 无监督配准方法则完全消除对标注数据的需要, 仅使用图像相似度损失和正则化损失监督网络训练。目前, 无监督医学图像算法已经成为医学图像配准领域的研究重点, 在无需获得代价高昂的标注信息下就能够取得与有监督 and 传统方法相当甚至更高的配准精度。在此基础上, 本文进一步讨论了医学图像配准研究后续可能的 4 个未来挑战, 希望能够为更高精度、更高效率的深度医学图像配准算法的研究提供方向, 并推动深度医学图像配准技术在临床诊疗中落地应用。

关键词: 医学图像配准; 深度学习 (DL); 全监督学习; 双监督学习; 弱监督学习; 无监督学习

Deep-learning based medical image registration pathway: towards unsupervised learning

Ma Lufan¹, Luo Feng¹, Yan Jiangpeng¹, Xu Zhe^{1,2}, Luo Jie², Li Xiu^{1*}

1. Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China;

2. Harvard Medical School, Boston 02115, USA

Abstract: Medical image registration (MIR) has aimed to implement the optimal transformation via aligning anatomical structures of a pair of medical images spatially. The crucial clinical applications like disease diagnosis, surgical guidance and radiation therapy have been involved. Scholars have categorized MIR into inter-/intra-patient registration, uni-/multi-modal registration and rigid/non-rigid registration. Image classification has been developing deep learning-based (DL-based) MIR methods. The DL-based MIR has demonstrated substantial improvement in computational efficiency and task-specific registration accuracy over traditional iterative registration approaches. A sophisticated literature review of DL-based

收稿日期: 2020-07-09; 修回日期: 2020-09-22; 预印本日期: 2020-09-29

* 通信作者: 李秀 li.xiu@sz.tsinghua.edu.cn

基金项目: 国家自然科学基金项目 (41876098); 清华大学深圳国际研究生院海外科研合作基金项目 (HW2018008)

Supported by: National Natural Science Foundation of China (41876098)

MIR have benefited to the disciplines. The current MIR has been analysed based on iterative optimization to one-step prediction and supervised learning to unsupervised learning. The DL-based MIR has been classified into fully supervised, dual supervised, weakly supervised and unsupervised approaches to train the DL network via the amount of supervision. Each category has been systematically reviewed. At the beginning, fully supervised methods have been reviewed in terms of the initial exploration to remove the time-consuming with low inference speed issues of deep iterative registration algorithms (deep similarity-based registration, reinforcement learning-based registration). One-step fully supervised registration has predicted the final transformation. The lack of training datasets with ground-truth transformations have barriered to train a fully supervised registration network. Most scholars have generated synthesized transformations with the following three approaches as below: 1) random augmentation-based generation; 2) traditional registration-based generation; 3) model-based generation. Next, the integration of dual-supervised and weak supervised registration have alleviated the reliance on ground truth compared with fully supervised approaches via the transition technologies between fully supervised and unsupervised methods. Dual-supervised registration frameworks have integrated image similarity metric to supervise the training. Weak supervised registration in the context of anatomical labels of interest (solid organs, vessels, ducts, structure boundaries and other subject-specific ad hoc landmarks) has replaced ground truth. The label similarity using label-driven supervised registration has facilitated the network to directly estimate the transformation for paired fixed image and moving image. The end-to-end unsupervision has been used to indicate the DL-based medical image registration evolved into the unsupervised field gradually. The unsupervision has avoided the acquisition of ground-truth transformations and segmentation labels for the supervised methods. Unsupervised registration frameworks have performed spatial data based on spatial transformer network (STN) to flat image similarity loss calculation during the training process with unknown transformations further. The latest developments and applications of DL-based unsupervised registration methods have been summarized from the aspects of loss functions and network architectures. DL-based unsupervised registration algorithms on liver CT (computed tomography) scan datasets have also been re-implemented. The demonstrated analyses have the priority to baseline model. At the end, the potentials and possibilities have been illustrated as following: 1) constructing more robust similarity metrics and more effective regularization terms to deal with multi-modality MIR; 2) quantifying registration result confidence of various DL-based models or integrating domain knowledge into current data-driven networks; 3) designing more qualified networks with fewer parameters (e. g., 3D convolution factorization, capsule network architecture).

Key words: medical image registration; deep learning (DL); full supervised learning; dual supervised learning; weakly supervised learning; unsupervised learning

0 引言

图像配准是一种根据图像外观将浮动图像和参考图像对齐的优化过程,其优化目标是找到使输入图像中感兴趣区域对齐的最佳空间变换(罗述谦和吕维雪,1999)。作为图像辅助诊疗的关键技术,医学图像配准试图建立不同图像间的解剖学对应关系,在疾病诊断、手术引导和放射性治疗等多个临床场景中得到应用(王彩芳和姜明,2006)。医学图像配准是一个重要的研究主题,有各种各样的分类方式:按照配准区域可分为大脑、肺部、肝脏以及乳房等解剖结构配准;根据待配准图像类型可分为单模态配准和多模态配准;根据图像维度可分为2D-2D、2D-3D、3D-2D和3D-3D配准;根据空间变换类型可

分为仿射配准和可变形配准(Andrade等,2018)——其中仿射配准是针对全局图像的整体线性变换,适用于骨骼等坚硬结构的刚性配准,常作为复杂多阶段配准的预配准操作(Boveiri等,2020),可变形配准则允许图像发生弹性变换,能够捕获组织器官的局部不规则形变和多模态设备成像差异,实现图像的精确对齐(Oliveira等,2014)。更多关于配准研究的细分方式如图1所示。

医学图像配准研究由来已久,可追溯至20世纪70年代(Barnea和Silverman,1972)。传统配准方法通过迭代优化输入图像间相似度来指导参数更新,以求解最优变换(Auer等,2005;Shekhar和Zagrodsky,2002;Kabus等,2004;Tang等,2006),其步骤如图2所示。这类方法每配准一对图像都要从零开始优化目标函数,限制了最优参数的拟合速度,较低的

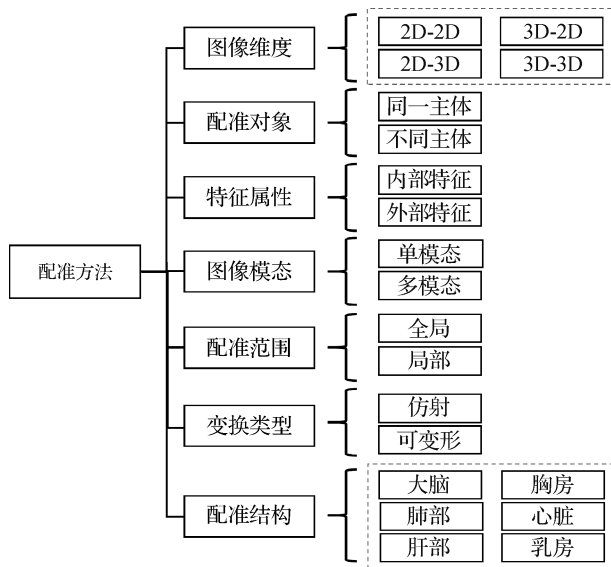


图1 医学图像配准分类

Fig. 1 Classification of medical image registration methods

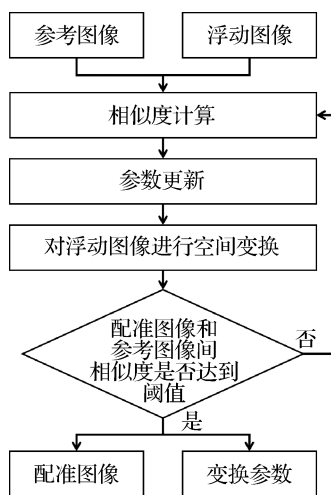


图2 传统的迭代优化配准算法

Fig. 2 Traditional iterative algorithm for optimization of registration

计算效率给临床实时应用带来困难。深度学习 (deep learning, DL) 技术在分类、检测和分割等计算机视觉任务中发挥关键作用, 研究者也验证了使用深度学习方法解决医学图像配准问题的可行性 (Kabus 等, 2004; Tang 等, 2006; Miao 等, 2016a), 取得了令人鼓舞的结果。

最初, 学者们在传统的迭代配准算法中嵌入神经网络来提取特征或学习相似度度量 (Wu 等, 2013; Cheng 等, 2018; Simonovsky 等, 2016), 提升了基于灰度和基于特征的配准算法性能。学者也利用深度强化学习将医学图像配准转化为迭代求解最优

动作序列的过程 (Liao 等, 2016; Krebs 等, 2017)。但这类深度迭代算法收敛速度缓慢, 快速配准的需求促使学者提出用于一步变换估计的全监督配准框架 (Yan 等, 2018; Eppenhof 等, 2018)。获取全监督方法所需的金标准 (ground truth) 十分困难, 为此学者逐渐迈向更少监督的配准方法。双监督配准及弱监督配准通过引入图像分割掩码等新的监督信息来降低对金标准的依赖。然而, 这类方法仍然需要大量标注数据, 高昂的标注成本驱动学者进一步开发端到端的无监督配准框架 (de Vos 等, 2019; Li 和 Fan, 2018)。为了能够全面地梳理深度医学图像配准领域的研究文献, 以“深度学习”、“神经网络”、“基于学习”、“卷积神经网络”、“医学图像”、“配准”等为关键词, 在谷歌学术和中国知网期刊数据库中检索到 122 篇在 2015—2019 年公开发表的学术文献, 并对其进行分类统计。如图 3 所示, 无监督配准研究的文献数量增长迅速, 显示出巨大的发展潜力。基于深度学习的医学图像配准研究呈现出从迭代优化到一步估计、从有监督到无监督的发展趋势, 如图 4 所示。一步变换估计的配准方法是研究主流, 因此下文提及的深度医学图像配准研究均属这一范畴。

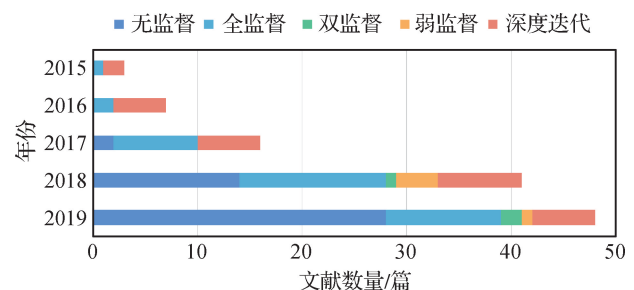


图3 基于深度学习的医学图像配准研究统计

Fig. 3 Number of publications in DL-based medical image registration

相较于通用计算机视觉任务, 深度医学图像处理研究起步较晚, 因此国内现有的医学图像配准综述大多停留在对传统迭代配准算法的梳理 (罗述谦和吕维雪, 1999; 王海南 等, 2005; 李雄飞 等, 2010), 基于深度学习的医学图像配准综述数量稀少, 仅有邹茂扬等人 (2019) 从配准网络架构的角度归纳已有工作。相对于他人的工作, 本文针对医学图像配准任务中标注数据匮乏这一普遍存在的问题, 按照网络训练过程中监督信息不断减少的研究

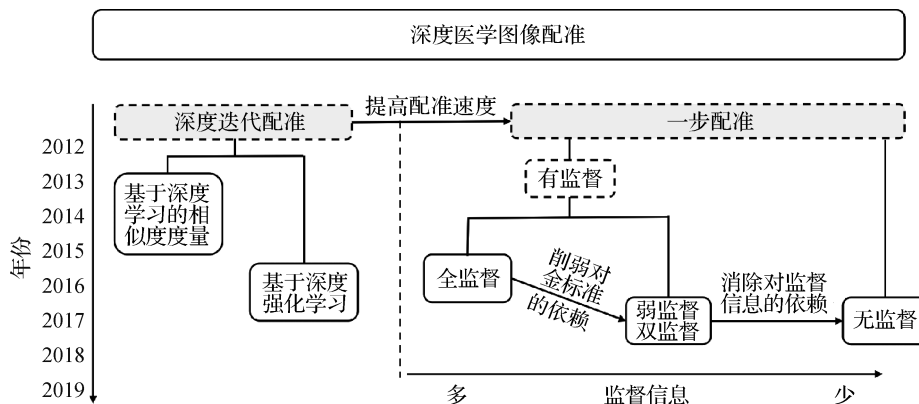


图4 基于深度学习的医学图像配准研究路线图

Fig. 4 Overview of DL-based medical image registration approaches in timeline

发展趋势,系统地总结了深度医学图像配准领域的最新研究进展,比较分析不同配准方法的优势与局限,并对未来研究趋势进行展望。

1 深度医学图像配准的界定及分类

深度医学图像配准是指利用深度卷积网络直接估计图像的空间映射关系,实现不同医学图像对齐的过程,其公式化表述如下:浮动图像 I_m 是 N 维实数向量空间 $\mathbf{R}^N$ 中子区域到 C 维实数向量空间 $\mathbf{R}^C$ 的映射,即 $I_m: \Omega \rightarrow \mathbf{R}^C$ 。同样地,参考图像 I_f 可表示为 $I_f: \Omega \rightarrow \mathbf{R}^C$ 。其中 Ω 表示 $\mathbf{R}^N$ 中的子区域, C 表示图像通道数量。由于医学图像配准一般使用通道数为1的2维或3维灰度图像,则上述表示可简化为 $I_m: \Omega \rightarrow \mathbf{R}, I_f: \Omega \rightarrow \mathbf{R}, \Omega \subseteq \{\mathbf{R}^2 \cup \mathbf{R}^3\}$ 。端到端的深度配准网络以 I_m 和 I_f 为输入,输出空间变换 $\Phi: \Omega \rightarrow \{\mathbf{R}^2 \cup \mathbf{R}^3\}$,使得扭曲变换后的 I_m 对齐至 I_f ,即

$$I_f(x) \approx I_m(x + \Phi(x)) \quad (1)$$

式中, I_m 和 I_f 分别表示浮动图像和参考图像, Φ 表示浮动图像 I_m 中的各像素或体素 $I_m(x)$ 对齐到参考图像 I_f 中对应像素或体素 $I_f(x)$ 所需的位移。一般地,经空间变换 Φ 扭曲后的 I_m 可定义为 I_m^{warped} ,表示变换后的浮动图像,即 $I_m^{\text{warped}} = I_m(x + \Phi(x))$ 。

根据上述定义,深度医学图像配准是一个端到端的变换预测过程。网络以待配准图像对为输入,在单次前向传播中完成预测,输出对应空间变换。根据实现细节的不同,深度医学图像配准可分为以下3类:

1) 按照网络结构划分。深度医学图像配准可分为基于卷积神经网络(convolutional neural net-

work, CNN)(Krizhevsky 等,2012)、基于全卷积网络(fully convolutional network, FCN)(Long 等,2015)、基于生成对抗网络(generative adversarial nets, GAN)(Goodfellow 等,2014)的图像配准。基于 CNN(Goodfellow 等,2014)的配准网络包含特征提取和参数回归两部分,常用于仿射配准。U-Net(Ronneberger 等,2015)是一种基于 FCN(Long 等,2015)的典型医学图像分割网络,但也常用于可变形配准之中。网络由降低图像分辨率的编码器和恢复图像分辨率的解码器构成,并在编码器和解码器之间加入了跳跃连接操作,有助于实现图像语义信息捕捉和像素级的定位。基于 GAN(Goodfellow 等,2014)的配准网络由估计变换参数的生成器和评估配准质量的鉴别器两部分组成,在对抗训练中提升配准效果。

2) 按照输入图像尺寸划分。可分为基于图像块和基于整幅图像的配准。基于图像块的配准可解决训练数据不足和计算资源紧张的问题。但这种方法在融合多个图像块的变形场时可能沿图像块边缘出现网格状伪影(Fu 等,2020a);基于整幅图像的配准则包含更多与配准相关的全局信息。

3) 按照网络输出的变换类型划分。可分为预测变换矩阵的仿射配准模型和预测位移向量场的可变形配准模型。仿射变换矩阵包含表示平移、旋转、缩放和剪切四类变换的参数。位移向量场则表示图像中稀疏或密集控制点的位移,插值后得到的变形场尺寸与输入图像一致。

本文选用训练过程中所使用的监督信息作为分类标准,将深度医学图像配准划分为全监督、双监督、弱监督和无监督配准方法。全监督配准以图像

对间的真实变换(金标准)作为监督信息,采用真实变换参数和估计变换参数间的误差损失监督网络训练。双监督配准在使用参数误差损失基础上又额外加入图像相似度损失,以降低对金标准的依赖。弱监督使用分割掩码、生物关键点等标签信息替代金标准。无监督配准则不再需要任何标注数据,仅使用图像相似度损失监督网络训练。后续的文献梳理也将按照监督信息递减的线索来组织。

2 全监督医学图像配准

一步估计的全监督配准框架以按通道拼接的待配准图像对 (I_m, I_f) 为输入,通过配准网络直接推断出最优变换,过程中使用参数误差损失监督网络训练,其优化目标可表示为

$$\Phi^* = \arg \min_{\Phi} L_{\text{dist}}(\Phi_{\text{gt}}, \Phi(I_m, I_f)) \quad (2)$$

式中, Φ_{gt} 表示实际的空间变换(即 ground truth, 金标准), Φ^* 表示最优空间变换, L_{dist} 表示真实变换 Φ_{gt} 与预测变换 Φ 之间的参数误差损失。

全监督配准中常用的损失函数为绝对误差和(sum of absolute error, SAE)、平均绝对误差(mean absolute error, MAE)、误差平方和(sum of squared error, SSE)、均方误差(mean squared error, MSE)等。全监督方法的最大挑战是式(2)中金标准 Φ_{gt} 的获取成本高昂,且高度依赖专业知识,仅有极少数研究使用由专家手动配准的训练数据集(Yan 等, 2018; Zheng 等, 2017, 2018)。所以,围绕全监督医学图像配准的研究问题往往转化为如何生成高质量的、含已知变换的训练数据集的问题。为解决这一问题,使用合成变换 Φ_{syn} 来代替 Φ_{gt} ,从而得到训练数据 $(I_m, I_f, \Phi_{\text{gt}})$ 。常用的变换合成方法分为3类:1)基于随机变换生成 Φ_{syn} ;2)采用传统配准算法求解 Φ_{syn} ;3)基于模型生成 Φ_{syn} 。相关的全监督配准文献归纳如表1所示。

随机变换是一种通过叠加平移、旋转和缩放等数据增强技巧得到空间变换 Φ_{syn} 的方法。已知图像 I_1 , 经 Φ_{syn} 空间变换后得到图像 I_2 作为参考图像,与 I_1 、 Φ_{syn} 共同构成一组已知变换的训练数据 $(I_1, I_2, \Phi_{\text{syn}})$ 。为使合成数据尽量真实,这类方法还常对参考图像 I_2 添加高斯随机噪声。Salehi 等人(2018)采用此方法,首先按照均匀分布生成旋转参

数矩阵、平移参数矩阵和缩放参数矩阵,然后结合3个矩阵合成仿射变换矩阵。在此基础上,采用由粗略到精细的配准策略,设计了3维位姿估计网络3D PoseNet。网络使用特征层提取图像特征,后输入回归层预测3维旋转和平移的角轴表示,并在MSE基础上创新性地引入测地距离损失,极大提升了配准精度。Chee 和 Wu(2018)采用类似的变换合成方法生成训练数据,提出融入 DenseNet(Iandola 等, 2014)特征提取模块的脑部MRI(magnetic resonance imaging)图像配准模型 AIRNet。林家庆等人(2020)通过随机旋转和缩放合成小样本训练数据集,提出一种串行回归的深度神经网络以实现CT图像方向的精准校正。基于成对的T1-T2 MR图像, Sloan 等人(2018)将此随机变换方法拓展到多模态仿射配准,额外引入可逆性损失,实现对估计变换的逆一致性约束。不同于上述合成仿射矩阵的解决方法,学者还基于随机变换合成变形场(Eppenhof 等, 2018; Sun 等, 2018; Sentker 等, 2018; Eppenhof 和 Pluim, 2018)。初始化与输入图像等大的位移向量场,从中全采样或随机采样若干点,按随机概率分布生成各点在不同方向上的位移,得到密集或稀疏的合成变形场。Sokooti 等人(2017)引入不同方差的高斯核函数合成不同位移范围的变形场,并提出多尺度分流卷积网络 RegNet,模型融入上下文信息,实现了端到端的3D胸部CT图像非刚性配准。为了更好地捕获呼吸运动带来的肺部形变, Eppenhof 和 Pluim(2019)多次叠加基于稀疏网格随机生成的大形变和基于密集网格随机生成的小形变,提高了 Φ_{syn} 的可靠性和模型的泛化性。

随机合成的 Φ_{syn} 过于简单、类型有限,无法准确刻画复杂人体器官的真实生理运动,对此学者使用 diffeomorphic demons (Vercauteren 等, 2009)、LD-DMM (large deformation diffeomorphic metric mapping) (Ashburner, 2007)、SyN (symmetric image normalization) (Avants 等, 2008) 等传统配准算法求解空间变换 Φ_{syn} 来代替 Φ_{gt} 。Miao 等人(2016a)在手动粗配准图像对后使用基于灰度的 Powell 迭代配准算法得到 Φ_{syn} (Miao 等, 2016b),并将其应用于端到端的有监督刚性配准。网络采用特征提取和分层回归机制,以局部感兴趣区域(region of interest, ROI)图像残差作为输入特征,以变换参数的欧氏距离作为损失函数,将6维的参数空间划分为3个区

表 1 全监督医学图像配准研究归纳
Table 1 Overview of fully-supervised registration methods

参考文献	配准区域	维度	模态	变换	模型	合成方法
Miao 等人(2016a) Miao 等人(2016b)	探头	2D-3D	Xray	仿射	CNN	传统算法
Yang 等人(2016)	脑部	2D-2D,3D-3D	MR, MR	可变形	FCN	传统算法
Yan 等人(2018)	前列腺	3D-3D	MR-TRUS	仿射	GAN	手动标注
Eppenhof 和 Pluim(2018)	肺部	3D-3D	CT	可变形	FCN	随机变换
Sun 等人(2018)	肝部	2D-2D	CT-US	可变形	DVFNet	随机变换
Uzunova 等人(2017)	脑部 心脏	2D-2D	MR	可变形	FlowNet	基于模型
Sentker 等人(2018)	肺部	3D-3D	CT	可变形	FCN	随机变换
Sokooti 等人(2019)	肺部	3D-3D	CT	可变形	U-Net	基于模型
Chee 和 Wu(2018)	脑部	3D-3D	MR	仿射	CNN	随机变换
Salehi 等人(2018)	脑部	3D-3D	T2	仿射	ResNetCNN	随机变换
Sloan 等人(2018)	脑部	2D-2D	T1, T1-T2	仿射	FCN	随机变换
Zheng 等人(2017)	探头	2D-2D	Xray	仿射	CNN	手动标注
Zheng 等人(2018)	脊椎	2D-3D	Xray-CT	仿射	CNN	手动标注
Sokooti 等人(2017)	肺部	3D-3D	CT	可变形	RegNet	随机变换
Eppenhof 和 Pluim(2018)	肺部	3D-3D	CT	可变形	CNN	随机变换
Eppenhof 和 Pluim(2019)	肺部	3D-3D	CT	可变形	U-Net	随机变换
Cao 等人(2017) Cao 等人(2018a)	脑部	3D-3D	MR	可变形	CNN	传统算法
Rohé 等人(2017)	心脏	3D-3D	MR	可变形	SVFNet	传统算法
Onieva 等人(2018)	肺部	3D-3D	CT	可变形	RegNet	传统算法
Yang 等人(2017)	脑部	3D-3D	T1, T1-T2	可变形	CNN	传统算法
沈延延和冯汉升(2020)	颅骨	2D-3D	X-DRR	仿射	FCN	传统算法
林家庆等人(2020)	肝脏	2D-2D	CT	仿射	CNN	随机变换

注:CT(computed tomography)为计算机断层扫描成像,MR(magnetic resonance)为磁共振成像,Xray 为 X 光图像,US(ultrasonography)为超声成像,TRUS(transrectal ultrasonography)为经直肠超声。数据合成方法中的随机变换指第 1 类变换合成方法:基于随机变换生成 Φ_{syn} ;传统算法指第 2 类变换合成方法:采用传统配准算法求解 Φ_{syn} ;基于模型指第 3 类变换合成方法:基于模型生成 Φ_{syn} ,手动标注指通过专家手动标注获取 Φ_{syn} 。

域,并分层训练基于面内、面外旋转和平面外平移的参数,揭示了从图像残差到参数残差的映射关系。该方法将复杂的回归任务分解为可独立学习的简单子任务,计算效率明显高于基于互信息(mutual information, MI)、互相关系数(cross correlation, CC)和梯度相关系数(gradient correlation, GC)等传统迭代配准方法(Gendrin 等, 2012; Schmid 和 Chênes, 2014; Hatt 等, 2015)。类似地,沈延延和冯汉升

(2020)将空间刚体变换参数分解到两个平面上,利用 FCN(Long 等, 2015)从 X-DRR(digitally reconstructed radiograph)图像对的残差回归出空间刚体变换参数,实现实时的 2D-3D 配准。不同于使用传统配准算法直接求解图像对间的空间变换(Miao 等, 2016a, b), Rohé 等人(2017)则基于迭代对数近似法(Arsigny 等, 2006)求解图像分割掩码对间的变换作为参考变换 Φ_{syn} ,充分利用感兴趣区域的轮廓

信息。在利用 SyN (Avants 等, 2008) 初配准脑部 MRI 灰度图像后 (Cao 等, 2017, 2018a), 使用 diffeomorphic demons (Vercauteren 等, 2009) 求解经初配准后的器官分割掩码的变形场, 叠加两阶段求解的变形场作为 Φ_{syn} 。合成数据后, Cao 等人 (2017) 设计了一个融入图像块相似度的 CNN 配准网络, 额外输入 (I_m, I_f) 的互相关图, 降低模板与被试者间的外观差异给网络带来的训练难度。采用微分同胚度量映射 (LDDMM) 算法求解变形场合成 Φ_{syn} (Yang 等, 2016, 2017; Onieva 等, 2018)。Yang 等人 (2016) 探索了一种基于 LDDMM 的非参数深度配准模型, 预测空间维度上的初始动量作为中间值, 通过贝叶斯概率网络估计出变形场, Yang 等人 (2017) 后续将该方法拓展应用到多模态配准中, 现已集成到 Quicksilver 软件中。

通过传统配准算法求解空间变换的方法提高了合成变换的真实性, 但会直接限制后续深度配准模型的精度上限, 对此学者采用另外一种基于模型的数据合成方法。基于 Wilms 等人 (2017) 提出的统计形状模型 (statistical shape models, SSM), Uzunova 等人 (2017) 提出具备高度表征能力的统计外观模型 (statistical appearance models, SAM), 通过少量的训练样本合成了大量 ($I_1, I_2, \Phi_{\text{syn}}$) 数据对。实验表明, 使用 SAM 合成数据训练的模型配准精度优于使用随机合成数据训练的模型。结合胸部横纵向最大扩张范围的相关生物学知识, Sokooti 等人 (2019) 使用呼吸运动模型合成的变形场 Φ_{syn} 训练 3D 胸部 CT 图像配准网络, 其精度优于基于随机合成数据训练的模型, 但仍未能超过传统配准算法 (Berendsen 等, 2014)。

基于随机变换的 Φ_{gt} 生成方法仅能合成有限类型的空间变换, 无法有效模拟复杂的人体组织器官的真实生理运动。基于传统配准算法合成 Φ_{gt} 的方法使配准网络在训练时的优化目标为预测一个无限接近 Φ_{syn} 的空间变换, 这为深度医学图像配准模型设置了性能上限。尽管基于模型合成 Φ_{gt} 的方法优于上述两种方法, 但组织器官的形变十分复杂, 具有较高的建模难度。总的来说, 3 类方法在不同程度上均有局限, 但这类 Φ_{syn} 合成方法不再需要专家手动配准图像获取金标准, 推动了有监督配准框架的发展。

与传统配准算法和深度迭代配准算法相比, 全

监督配准方法利用单次前向传播一步估计空间变换, 极大缩短了配准时间, 使深度医学图像配准向临床应用更进一步。然而 Sun 等人 (2018) 和 Sokooti 等人 (2019) 的工作均表明合成训练数据的全监督配准仍然很难超越传统配准算法, 仅能达到与之相当的配准精度。这是因为有监督方法的配准精度高度依赖于变形场 Φ_{gt} 的质量, 而上述基于 Φ_{syn} 合成的训练数据与临床数据存在较大偏差, 使得配准模型无法很好地适用于实际测试。针对此问题, 可以引入域适应模块实现合成数据与真实数据间的域转换, 提升训练模型的泛化性和鲁棒性。

3 双监督及弱监督医学图像配准

为了降低全监督配准对金标准 Φ_{gt} 的依赖, 尝试使用更少的监督信息来指导医学图像配准, 提出了双监督和弱监督配准方法。双监督配准框架以待配准的 (I_m, I_f) 为输入, 通过深度卷积网络估计出从 I_m 到 I_f 的空间映射 Φ^* , 同时使用 Φ_{gt} 与 Φ^* 间的参数误差损失 L_{dist} 及 I_m^{warped} 和 I_f 间的相似度损失 L_{sim} 监督网络训练。与全监督方法相比, 双监督配准在金标准的基础上引入 I_m^{warped} 和 I_f 间图像相似度作为另一监督信息, 其整体优化目标变为

$$\Phi^* = \arg \min_{\Phi} (L_{\text{dist}} + L_{\text{sim}}) \quad (3)$$

式中, Φ^* 表示最优空间变换, L_{dist} 表示真实变换 Φ_{gt} 与预测变换 Φ 之间的参数误差损失, L_{sim} 表示变换后的浮动图像 I_m^{warped} 和参考图像间的相似度损失。现有的基于双监督方法的医学图像配准研究如表 2 所示。

Fan 等人 (2019a) 基于分层学习、多通道输入策略提出双监督配准网络 BIRNet, 利用合成变形场 Φ_{syn} 指导网络进行粗略配准后借助 I_m^{warped} 和 I_f 间相似度完成精细配准, 解决了只使用不准确的 Φ_{syn} 监督网络带来的偏差问题。Liu 等人 (2019) 则将双监督框架引入跨模态配准, 使用 Xception (Chollet, 2017) 学习 I_m 和 I_f 在公共潜空间的特征表示, 构建可微几何约束回归得到可变形变换参数 Φ^* , 过程中使用有监督的合成变换和无监督的图像相似度共同指导语义配准网络训练。实际上, 双监督配准并无严格定义, 在广义上通常指同时利用两种损失函数监督配准网络训练。Cao 等人 (2018b) 提出一种用模态内相似

表 2 双监督医学图像配准研究归纳
Table 2 Overview of dual-supervised registration methods

参考文献	配准区域	维度	模态	变换	模型	损失函数
Fan 等人(2019a)	脑部	3D-3D	MR	可变形	U-Net	L2
Liu 等人(2019)	骨骼	2D-2D	Xray	可变形	Xception	L2
Cao 等人(2018b)	盆骨	3D-3D	MR-CT	可变形	U-Net	NCC

注:L2 为 L2 范数,NCC 为归一化互相关(normalized cross correlation)。

度指导跨模态配准网络学习的方法,分别在 MR 和 CT 模态上计算成对图像的相似度损失,包括 MR-MR 损失和 CT-CT 损失。网络在训练过程中共同优化两个相似度损失,结果表明双模态相似度在 Dice 相似系数(Dice similarity coefficient, DSC)和平均表面距离(average surface distance, ASD)指标上显著优于 SyN(Avants 等,2008)和使用单模态相似度的方法。

相较于直接获取 Φ_{gt} ,基于感兴趣解剖结构的标签信息显得更加可靠和实用。所以不同于上述全监督和双监督方法,弱监督配准选择使用解剖标签

(如器官、血管或其他特殊结构的分割掩码及生物关键点)来指导网络一步估计出输入图像对 (I_m, I_f) 间的空间变换 Φ^* ,并将最小化浮动图像标签 $Label_m$ 和参考图像标签 $Label_f$ 间的相似度损失作为优化目标

$$\Phi^* = \arg \min_{\Phi} L_{sim}(Label_m, Label_f) \quad (4)$$

式中, Φ^* 表示最优空间变换, $Label_m$ 和 $Label_f$ 分别表示浮动图像和参考图像标签, L_{sim} 表示变换后的 $Label_m$ 和 $Label_f$ 间的相似度损失。相关文献归纳如表 3 所示。

表 3 弱监督医学图像配准研究归纳
Table 3 Overview of weakly-supervised registration methods

参考文献	配准区域	维度	模态	变换	模型	损失函数
Hu 等人(2018a)	前列腺	3D-3D	MR-TRUS	仿射、可变形	CNN	CE L2
Hu 等人(2018b)	前列腺	3D-3D	MR-TRUS	仿射、可变形	GAN	L2
Hu 等人(2018c)	前列腺	3D-3D	MR-TRUS	可变形	FCN	CEDice
Hering 等人(2019)	前列腺	2D-2D	cine-MR	可变形	U-Net GAN	L2

注:cine 为心血管快速磁共振成像,CE 为交叉熵损失(cross entropy),Dice 为分割骰子系数,用于度量图像分割掩码的相似度,取值范围在[0,1]。

Hu 等人(2018a)首次提出一种标签驱动的弱监督配准框架。网络由全局子网和局部子网分别拟合仿射和可变形变换,使用 I_m 和 I_f 标签对的相似度损失和正则化损失监督训练,实现 MR-TRUS 跨模态图像配准。后续从损失函数和网络结构两个方面对该弱监督框架进行优化。针对损失函数设置,Hu 等人(2018b)引入 GAN(Goodfellow 等,2014)网络,在标签相似度损失基础上新增对抗损失,实现对变形场的正则化约束,得到一个更加光滑的变形场。同时,Hu 等人(2018c)设置多尺度 Dice 指标衡量标签对的相似度损失,解决了使用二值分割掩码计算标签相似度时存在的非均衡权重问题。针对网络结构设计,Hu 等人(2018c)在弱监督框架中引入深监

督机制,有效融合了浅层预测变形场的全局信息和深层预测变形场的局部信息。Hering 等人(2019)提出一个结合全局弱标签语义信息和局部距离度量的弱监督模型,充分利用互补的全局信息和局部信息,实现了心脏 2D cine-MR 图像配准。网络遵循 U-Net(Ronneberger 等,2015)结构,使用分割标签相似度和基于边缘的归一化梯度设计损失函数,有效对齐血管等局部结构,配准效果优于 Ruhaak 等人(2013)提出的多级配准方法。

双监督及弱监督方法分别通过引入其他监督信息来削弱配准模型对于金标准的依赖。双监督配准方法在参数误差损失的基础上引入图像相似度损失,融合使用有监督损失和无监督损失指导网络优

化,降低了对 Φ_{gt} 的准确度要求。标签驱动的弱监督配准方法在网络训练阶段引入包含更少监督信息的分割标签,使用标签相似度替代图像相似度,消除了对 Φ_{gt} 的依赖并巧妙避开了跨模态图像相似度计算的困难。但是,这两种配准方法仍然需要大量的数据标注,并未完全消除网络训练对监督信息的依赖。

4 无监督医学图像配准

获取上述有监督方法所需的真实变换和分割标签极具挑战、成本高昂,所以越来越多的学者关注无监督配准框架。无监督配准方法进一步弱化了网络训练对监督信息的需要,仅需待配准的图像对(I_m, I_f)就能构建一个端到端的配准网络直接估计变换参数 Φ^* 。然而,在没有金标准 Φ_{gt} 的情况下,很难定义合适的网络损失函数。解决此问题的一项基础工作是Jaderberg等人(2015)提出的空间变换网络

(spatial transformer network, STN)。STN允许网络根据变形场(deformable fields, DFs)实现对 I_m 的空间变换,是一个完全可微的模块,能被插入到现有的卷积配准网络中。STN模块使得在训练过程中计算图像相似度损失成为可能,推动了无监督配准研究的发展。一个典型的无监督配准框架由卷积配准网络和空间变换网络(STN)组成。以图像对(I_m, I_f)为输入,利用卷积配准网络直接估计高维的变形场 Φ^* ,借助STN实现对 I_m 的扭曲变换得到 I_m^{warped} ,进而计算 I_f 与 I_m^{warped} 间的图像相似度损失,其优化目标为

$$\Phi^* = \arg \min_{\Phi} L_{\text{sim}}(I_m^{\text{warped}}, I_f) \quad (5)$$

式中, Φ^* 表示最优空间变换, I_m^{warped} 表示变换后的浮动图像, L_{sim} 表示变换后的 I_m^{warped} 和参考图像 I_f 间的相似度损失。针对无监督配准问题,从损失函数和配准框架两个方面给出了许多解决方案,相关文献归纳如表4所示。

表4 无监督医学图像配准研究归纳
Table 4 Overview of unsupervised registration methods

参考文献	配准区域	维度	模态	变换	模型	损失函数
Kori 和 Krishnamurthi (2019)	脑部	3D-3D	T1, T2, FLAIR	仿射	CNN MLP	SSD
de Vos 等人(2017)	心脏	2D-2D	MR	可变形	DIRNet	NCC
Yoo 等人(2017)	神经元组织	2D-2D	SSEM	可变形	CAE STN	MSE AE
Sheikhjafari 等人(2019)	心脏	2D-2D	MR	可变形	SAE	MSE
Balakrishnan 等人(2018)	脑部	3D-3D	MR	可变形	U-Net	LCC L2
Balakrishnan 等人(2019)	脑部	3D-3D	MR	可变形	U-Net	LCC Dice
Qin 等人(2018)	心脏	2D-2D	MR	可变形	FCN STN	MSE CE
Elmahdy 等人(2019)	前列腺	3D-3D	CT	可变形	GAN	Dice NCC
Mahapatra 等人(2018a)	骨骼	2D-2D	Xray	可变形	GAN	NMI SSIM BCE
Sun 和 Zhang (2018)	脑部	3D-3D	MR-US	可变形	FCN	L2 图像梯度
Mahapatra (2018)	眼底 心脏	2D-2D	FA, MR	可变形	GAN	NMI SSIM L2
Xu 等人(2020)	肾脏 腹腔	3D-3D	CT-MR	可变形	U-Net GAN	SSIM L2

续表 4 无监督医学图像配准研究归纳
Table 4 Overview of unsupervised registration methods(continued)

参考文献	配准区域	维度	模态	变换	模型	损失函数
Liu 等人(2017)	胸部 脑部	2D-2D	T1, T2, PD	可变形	PCANet	RMSE SSD MIND
Qin 等人(2019)	肺部 脑部	2D-2D	T1-T2, CT-MR	可变形	CNN GAN	L1
Shan 等人(2017)	脑部 肝脏	2D-2D	CT, MR	可变形	CNN	光度差 L1
Jiang 等人(2020)	肺部	3D-3D	CT	可变形	CNN	NCC L2
Yu 等人(2019)	腹部	3D-3D	CT-PET	可变形	CNN	NCC L1
Zhang(2018)	脑部	3D-3D	MR	可变形	ICNet	MSE L2
Kuang(2019)	脑部	2D-2D	T1	可变形	U-Net STN	NCC L2
Kim 等人(2019)	肝脏	3D-3D	CT	可变形	CNN	NCC L2
Fan 等人(2018)	脑部	3D-3D	CT	可变形	GAN	BCE
Fan 等人(2019b)	脑部 盆骨	3D-3D	MR, CT	可变形	GAN	SSD CC
Mahapatra 等人(2018b)	眼底 心脏	2D-2D	FA, MR	可变形	GAN	NMI SSIM L2
Fu 等人(2020b)	肺部	3D-3D	CT	可变形	GAN	NCC BCE L2
Lei 等人(2019)	腹部	3D-3D	CT	可变形	GAN	CC L2
de Vos 等人(2019)	心脏 肺部	3D-3D	MR, CT	仿射 可变形	CNN	NCC 弯曲能量
Zhao 等人(2020)	脑部 肝脏	3D-3D	MR, CT	仿射 可变形	U-Net	CC L2
Zhao 等人(2019)	脑部 肝脏	3D-3D	MR, CT	仿射 可变形	U-Net	CC L2

注:FLAIR(fluid attenuated inversion recovery)为磁共振成像液体衰减反转恢复序列,SSEM(serial-section electron microscopy)为切片扫描电镜成像,PET(positron emission computed tomography)为正电子发射型计算机断层显像,FA(fluorescein angiography)为眼底血管荧光造影,AE(absolute error)为绝对误差,LCC(local cross correlation)为局部互相关,CC(cross correlation)为互相关,Dice系数同表3,NMI(normalized mutual information)为归一化互信息,SSIM(structural similarity)为结构相似指数,BCE(binary cross entropy)为二元交叉熵,RMSE(root mean squard error)为均方根误差,MIND(modality independent neighbourhood descriptor)为模态独立领域描述子,L1为L1范数。

de Vos 等人(2017)提出了首个基于图像相似度的无监督配准网络 DIRNet,以 I_m^{warped} 和 I_f 间相似度为损失函数,使端到端的网络训练成为可能。Yoo 等人(2017)和 Sheikhfafari 等人(2019)采用卷

积自编码器将输入的 I_m 和 I_f 编码为特征向量后计算基于特征的相似度损失。结果表明,基于特征的相似度度量方法优于基于灰度的相似度度量方法。Balakrishnan 等人(2018)提出一种基于卷积神经网络

络(CNN)的无监督配准框架 VoxelMorph, 通过级联 U-Net(Çiçek 等, 2016)和 STN 结构一步估计稠密的变形场 Φ^* , 实现 3 维脑部 MRI 图像的可变形配准。网络借助 STN 模块在训练过程中计算图像相似度损失, 对 I_m^{warped} 和 I_f 间出现的外观差异进行惩罚, 最终模型在 Dice 指标上取得与 SyN(Avants 等, 2011)配准算法相当的精度。后来, Balakrishnan 等人(2019)在 VoxelMorph 中进一步引入解剖分割标签, 在原有的图像相似度损失基础上增加标签相似度损失监督训练。分割图的引入为配准网络提供了更多辅助信息, 新增的标签相似度约束能使网络收敛到更优的可变形变换参数, 有助于提高配准精度。类似地, Qin 等人(2018)在网络中引入分割标签, 建立了一个心脏 MR 图像的联合分割配准模型, 效果优于单任务模型。学者还将配准和分割任务相结合(Elmahdy 等, 2019; Mahapatra 等, 2018a), 使用图像相似度和标签相似度损失联合训练以提高配准性能。

跨模态图像间巨大的外观差异给图像相似度损失的计算带来困难, 传统的单模态图像相似度度量大多不再适用。对此, 一种直接的解决方案是改进相似度度量方式(Sedghi 等, 2019), 设计有效的跨模态图像相似度损失指导网络估计变形场 Φ^* 。Sun 和 Zhang(2018)基于图像灰度和梯度信息度量 I_m^{warped} 和 I_f 间相似度。学者还基于图像自相似性概念提出一种模态独立领域描述子 MIND(Liu 和 Leung, 2017; Heinrich 等, 2012)。MIND 对不同模态间的显著差异具有较高鲁棒性, 能够作为损失函数来衡量跨模态图像相似度。另一种间接的解决方案则是将跨模态图像映射到同一域空间, 统一 I_m 和 I_f 的数据分布。学者引入 GAN(Goodfellow 等, 2014)实现 I_m 从源域到目标域的跨模态转换(Choi 等, 2018; Isola 等, 2017; Liu 等, 2017; Yi 等, 2017; Singh 和 Raza, 2020), Mahapatra(2018)基于 Cycle-GAN(Zhu 等, 2017)将复杂的多模态配准转化为更简单的单模态配准任务, 进而继续使用 SSD(single shot multibox detector)、CC 等单模态相似度损失进行监督学习。Xu 等人(2020)提出一种基于双流配准场融合的新型跨模态图像配准方法, 借助改进的 Cycle-GAN 网络实现 CT-MR 跨模态图像转换, 通过双流配准网络有效融合原始跨模态流和合成单模态流估计的变形场, 缓解了基于 GAN(Goodfellow 等,

2014)的图像转换技术在生成图像中引入非真实的人工特征, 导致变形场失真的问题。王丽芳等人(2020)结合平均生成对抗网络和条件生成对抗网络的特点, 构建 RD-RaCGAN 合成模型, 实现输入图像对的双向模态转换。在 CT 模态和 MR 模态下分别估计单模态配准变形场, 并将二者融合得到最终变形场, 效果优于 Powell 优化的 MI 算法 SyN(Avants 等, 2008)、diffeomorphic demons(Vercauteren 等, 2009)、Cue-Aware Net(Cao 等, 2018a)和 I-SI(Öfverstedt 等, 2019)图像配准方法。Qin 等人(2019)则先将 I_m 和 I_f 映射到共同的语义空间以抽取特征表示, 然后再进行配准。王松伟等人(2020)采用 PCANet-SR 提取并融合跨模态输入图像的多级特征, 将不同模态的图像转换为同一模态, 减少跨模态图像之间的差异, 得到更好的配准性能。

然而, 仅靠图像相似度损失的约束并不能保证网络输出一个足够光滑的变形场, 所以学者们对不同的正则化损失进行探索。Shan 等人(2017)利用 FlowNet(Dosovitskiy 等, 2015)结构完成图像特征的学习和空间变换的预测, 设置 L1 正则化损失对网络估计的变形场 Φ^* 进行约束。Jiang 等人(2020)提出一种无监督的多尺度配准网络, 利用 NCC 图像相似度损失和 L2 形式的 DVF(displacement vectors-fully convolutional network)平滑度损失监督网络训练, 实现由粗略到精细的 3D 肺部 CT 图像配准。类似地, Yu 等人(2019)在 NCC 相似性损失基础上新增 L1 正则项约束来指导网络的端到端训练, Balakrishnan 等人(2018)则在损失函数中引入 L2 正则化损失。为提高估计变形场 Φ^* 的光滑度, 学者提出逆一致性损失和抗折叠损失(Zhang, 2018; Kuang, 2019; Kim 等, 2019), 在这些正则化损失的监督训练下模型取得比 Demons(Vercauteren 等, 2009)和 SyN(Avants 等, 2008)更优的配准性能。同时, 不少学者引入 GAN(Goodfellow 等, 2014), 给出了一种新的正则化方法。Fan 等人(2018)提出一个对抗性相似网络: 使用配准子网估计变形场 Φ^* 、判别子网区分 I_m^{warped} 和 I_f , 配准子网利用来自判别子网的对抗损失对变形场进行正则化约束。后续学者将此方法用于跨模态医学图像配准(Fan 等, 2019b; Mahapatra 等, 2018b)和 3D 肺部 CT 图像配准(Fu 等, 2020b)。Lei 等人(2019)使用扩张 inception(Szegedy 等, 2015)模块提取 I_m 和 I_f 的多尺度特征, 并在此基础上直接估

计变形场 Φ^* 。除图像相似度损失外,设置平滑度损失和对抗损失对 Φ^* 进行正则化约束,使得网络可以学习到一个更加真实可靠的光滑变形场。

除上述针对损失函数的改进外,学者不断探索更为高效的无监督配准框架。为了能够同时捕获输入图像 I_m 和 I_f 的全局变换和局部形变,学者将仿射变换和可变形变换集成到一个端到端的无监督配准框架中,解决了以往可变形配准之前需要进行仿射预配准的问题。de Vos 等人(2019)提出第一个无监督的仿射和可变形配准网络,通过叠加多个卷积配准网络实现由粗略到精细的 3D 医学图像配准。类似地,Zhao 等人(2020)基于 U-Net (Çiçek 等, 2016) 结构设计无监督配准框架 VTN (volume tweening network),在可变形配准网络中新增仿射预配准模块。学者在此基础上提出一个端到端的递归级联网络 RCVTN (recursive cascaded VTN) (Zhao 等, 2019),级联多个子网渐进学习 I_m 和 I_f 间的空间映射 Φ^* ,实现仿射和可变形子网参数的联合优化。级联策略有效处理了 3D 医学图像的较大位移形变,取得了优异的配准效果。

无监督医学图像配准发展迅猛,受到学者的广泛关注。现有的无监督配准研究表明,使用图像相似度损失和正则化损失代替有监督方法中的误差损失和标签相似度损失是可行有效的。从配准性能来看,无监督配准能够取得比有监督和传统配准方法更高的配准精度,且端到端配准网络显著提高计算效率。但是,目前的无监督配准研究主要集中在单模态图像,针对多模态图像配准的无监督研究较少,未来应有更多研究关注到这一问题。同时,现有的无监督配准框架结构单一,更多针对配准任务的高效网络结构有待提出。

5 配准实验比较

为直观展示端到端的无监督配准方法带来的优异性能,本文复现了两个代表性的基于深度学习的无监督可变形配准算法 VoxelMorph (Balakrishnan 等,2018)和 VTN (Zhao 等,2020)。由于获取开源的带有金标准和分割标注的配准数据集十分困难,并且现有的全监督配准算法 (Sun 等,2018;Sokooti 等, 2019)在配准精度上仍然无法超越 SyN (Avants 等, 2008)等传统配准算法,所以本文选择传统配准算

法 SyN (Avants 等,2008)作为基准对比。

5.1 数据集

上述两种算法的实现以及 ANTs SyN (Avants 等,2008)的应用均基于 Zhao 等人(2020)使用的肝脏 3D CT 扫描数据集。输入数据通过在现有肝脏数据集中随机选择两幅来组成待配准图像对,图 5 展示了一对待配准的肝脏 3D CT 切片图。

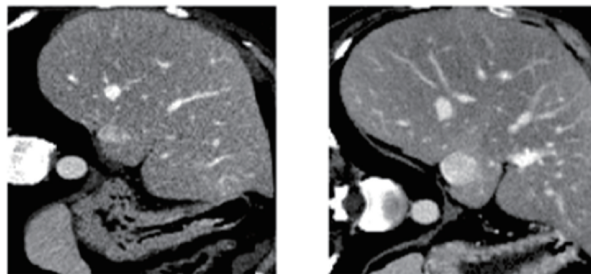


图 5 待配准的肝脏 3D CT 切片图

Fig. 5 Two slices from two liver CT scans for registration

训练数据集由 MSD 和 BFH (Beijing Friendship Hospital) (Zhao 等,2020)两个数据集构成。由于原始的 MSD 数据集中包含不同器官的医学图像,本文仅选择其中 933 幅包含肝脏的 CT 扫描图像,BFH 数据集则由 92 幅肝脏 CT 扫描图像构成。

测试数据集由 SLIVER、LiTS (liver tumor segmentation) 和 LSPG (Zhao 等,2020)3 个数据集构成。其中,SLIVER 数据集中包含 20 幅肝脏 3D CT 扫描图像,LiTS 数据集则包含 131 幅肝脏 3D CT 扫描图像,且均有对应的分割掩码。LSPG 是一个具有挑战性的数据集,包含 17 对猪的肝脏 3D CT 扫描图像,每对图像取自同一只猪在 13 mmHg 气腹压的手术前和手术中两个阶段。

5.2 实验细节

实验的细节设置严格遵从 VTN (Zhao 等,2020) 论文公布的源代码,具体网络结构及预处理方法均参照原文。模型使用 2 卡 12 GB NVIDIA TITAN Xp GPU 进行训练,所有实验 batch size 均为 4,每个实验的 epoch 数为 5,每个 epoch 迭代 20 000 步,总训练步数为 10^5 步,使用 Adam 优化器,初始学习率为 10^{-4} , 6×10^4 步后学习率减半, 8×10^4 步后学习率为以前的 0.25。

5.3 实验结果

选用 Dice 系数作为模型评估指标,在 3 个验证集上对上述两种无监督可变形配准算法和传统配准算法 SyN (Avants 等,2008)进行比较,不同方法配准

精度的度量结果和可视化结果分别如表5和图6所示。从配准精度和可视化效果图来看, VoxelMorph (Balakrishnan 等, 2018) 和 VTN (Zhao 等, 2020) 在无

监督条件下能够取得明显优于 SyN (Avants 等, 2008) 的配准精度, 印证了无监督配准方法相较于有监督和传统配准方法的优越性。

表5 不同配准方法精度度量结果(均值 ± 标准差)

Table 5 The quantitative results of different methods (average ± standard deviation)

配准算法	数据集		
	SLIVER	LiTS	LSPIG
ANTs SyN	0.895 ± 0.037	0.862 ± 0.055	0.825 ± 0.063
Affine + VoxelMorph	0.915 ± 0.025	0.869 ± 0.050	0.836 ± 0.056
Affine + VTN	0.915 ± 0.024	0.871 ± 0.048	0.829 ± 0.057

注: 表中数值为不同配准方法在验证集上对应的 Dice 值, Affine + VoxelMorph 为仿射配准子网和 VoxelMorph 可变形配准子网级联架构, Affine + VTN 为仿射配准子网和 VTN 可变形配准子网级联架构。

6 结 语

基于深度学习的医学图像配准方法在近年来取得了长足的进步和发展。本文按照网络训练过程中监督信息不断减少的研究发展趋势, 系统地总结了全监督、双监督及弱监督、无监督深度医学图像配准算法的最新研究进展, 比较分析了不同配准方法的优势与局限, 如表6所示。

其中, 基于无监督学习的配准方法消除了深度网络训练对于金标准和分割标签的依赖, 更加契合临床应用场景中缺乏数据标注的现状。从计算效率来看, 端到端的配准网络以无监督方式一步估计输入图像的空间映射, 能够取得比传统迭代配准算法更快的配准速率, 满足实时配准的需求。从配准精度来看, 无监督配准方法在图像相似度和正则项的约束下, 能够取得优于传统配准算法和有监督配准方法的性能表现。这些因素导致基于深度学习的医学图像配准研究将逐步迈向无监督领域。然而, 无监督的深度医学图像配准方法仍然面临着一些共性的挑战, 将其总结为下述4个主题: 多模态挑战、泛化性挑战、可解释性挑战和过参数化挑战。

1) 多模态挑战。对于无监督方法, 不同模态图像间巨大的外观差异使图像相似度损失的计算变得困难, 传统的单模态相似度度量往往不再适用, 亟需找到一个更加鲁棒的相似度指标。除相似度损失外, 学者也尝试使用正则化损失来约束变换估计, 但

常用的 L1、L2 等正则项在网络训练过程中并不能很好地约束估计变形场的光滑度。针对特定的配准任务, 学者应探索更为有效的正则化方法, 以确保网络输出一个真实可靠的光滑变形场。

2) 泛化性挑战。现有的配准模型泛化性不足, 大多针对于特定器官或模态, 当将其迁移到另一不可见域时, 模型性能出现明显退化, 如何使模型适用不同域成为亟待解决的问题。对此, 研究人员尝试在不同数据集上训练配准网络, 使用迁移学习方法来克服缺乏足够训练数据的问题, 让配准模型在脑部、肺部和肝脏等不同解剖部位的配准任务上表现出较强的泛化性。

3) 可解释性挑战。现有的深度配准网络均是数据驱动的, 公开医学图像数据集的质量参差不齐, 其中一些错误的人工标注会对配准方法的评估带来偏差 (Luo 等, 2020a)。完全基于数据的配准算法落地应用时, 如何让医生有效利用配准模型完成诊疗是需要考虑的问题。对此, 一种新的策略是通过给医生提供配准结果的不确定性, 使他们能根据配准结果和算法置信度来合理评估配准误差 (Luo 等, 2018, 2019)。所以, 针对配准有效性和配准不确定性 (Luo 等, 2020b) 的研究具有极大的参考价值, 是一个全新的研究方向。在此基础上, 未来研究也可考虑将领域专家知识融入到深度配准模型当中, 进一步加强模型的可解释性并改善总体配准效果。

4) 过参数化挑战。与 2D 配准网络相比, 现有的 3D 医学图像配准网络可训练参数随卷积层数呈指数增长, 极高的复杂度导致计算效率变低。在训

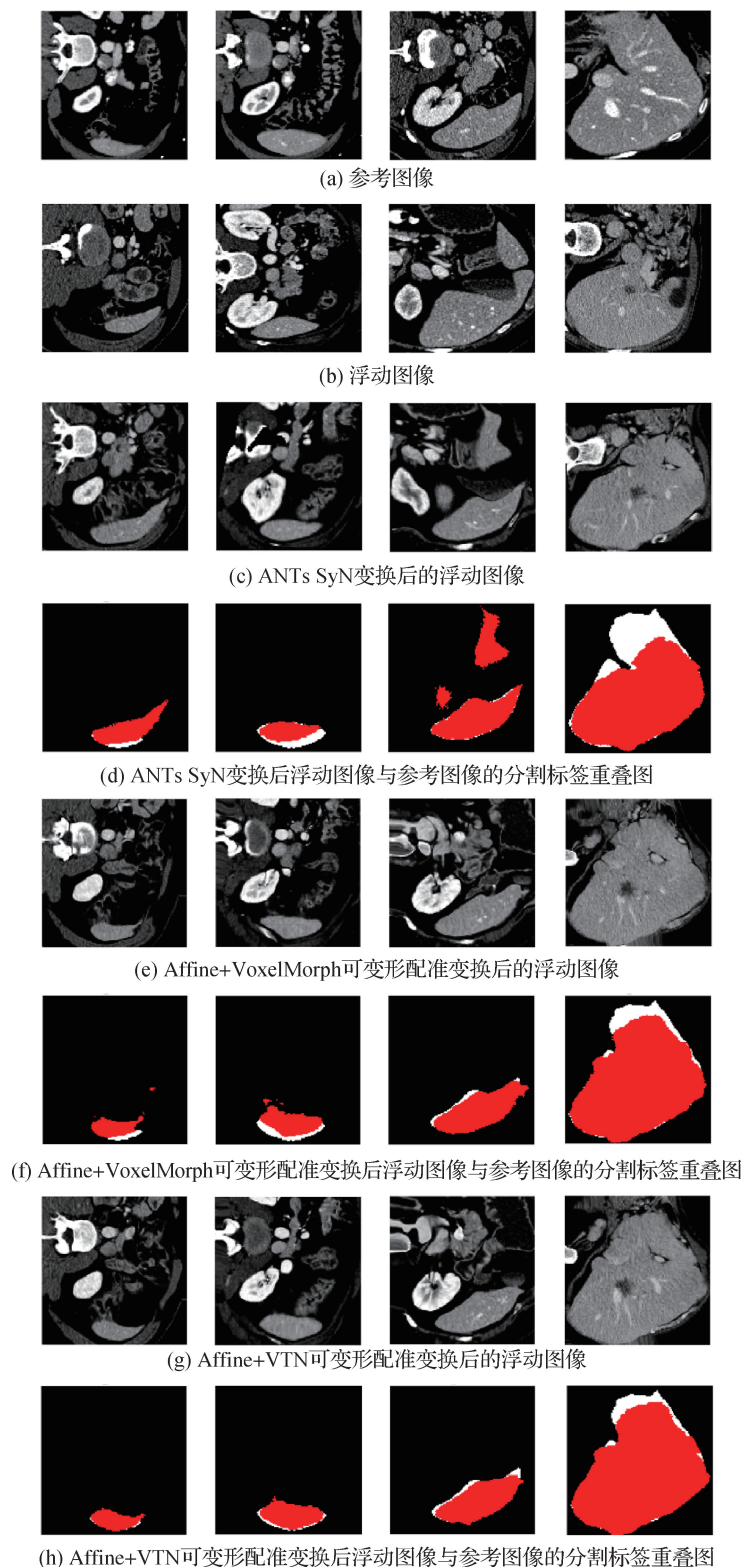


图6 不同配准方法可视化效果图

Fig. 6 Visualization results of different registration methods

((a) fixed images; (b) moving images; (c) warped images from ANTs SyN; (d) overlapped segmentation maps from ANTs SyN;
 (e) warped images from Affine + VoxelMorph registration; (f) overlapped segmentation maps from
 Affine + VoxelMorph registration; (g) warped images from Affine + VTN registration;
 (h) overlapped segmentation maps from Affine + VTN registration)

表6 4类深度医学图像配准方法比较

Table 6 Comparison of four types of deep medical image registration methods

类别	监督信息	优点	缺点
全监督	金标准	可以直接使用估计变换与真实变换间的参数误差损失监督网络训练;适用于低维的线性仿射配准。	缺乏带有已知真实变换的训练数据;合成变换类型有限,合成的训练数据与临床数据存在偏差;配准精度高度依赖于金标准的质量。
双监督	金标准、图像相似度	引入图像相似度损失,降低网络训练对金标准的依赖;融合使用有监督损失和无监督损失指导网络优化。	模型训练需要使用金标准,未完全消除配准网络对监督信息的依赖。
弱监督	分割标签相似度	引入包含更少监督信息的分割标签,比直接获取金标准更加可靠;使用标签相似度替代图像相似度,避开了跨模态图像相似度度量问题。	模型训练需要使用解剖分割标签等标注,未完全消除配准网络对监督信息的依赖。
无监督	图像相似度	完全去除网络训练对金标准和其他标注信息的依赖,更加契合实际应用场景;能够取得比有监督 and 传统配准方法更优的配准精度。适用于可变形配准。	传统单模态图像相似度大多不再适用于跨模态图像相似度损失计算;网络容易估计出折叠、体素漂移的非光滑变形场,导致变形场失真。

练数据和硬件资源有限的情况下,3D 配准模型的训练难度加大,在测试阶段也需耗费较长时间,不利于临床的快速诊断。为了有效平衡配准精度和计算效率,未来研究应提出更加高效的深度配准网络,可采用卷积分解(Qiu 等,2017)、胶囊网络(Sabour 等,2017)和神经网络结构搜索(Yan 等,2020)等方法来解决3D 配准网络过参数化、复杂度高的问题。

在深度医学图像配准研究迈向无监督学习领域的主流趋势下,上述挑战将成为该领域的下一个研究焦点。随着研究的不断深入,未来将出现更多高精度、高效率的无监督深度医学图像配准算法,进一步推动深度医学图像配准技术在临床诊疗中的应用。

参考文献(References)

- Andrade N, Faria F A and Cappabianco F A M. 2018. A practical review on medical image registration: from rigid to deep learning based approaches//Proceedings of the 31st SIBGRAPI Conference on Graphics, Patterns and Images. Parana, Brazil: IEEE: 463-470 [DOI: 10.1109/sibgrapi.2018.00066]
- Arsigny V, Commowick O, Pennec X and Ayache N. 2006. A log-Euclidean framework for statistics on diffeomorphisms//Proceedings of the 9th International Conference on Medical Image Computing and Computer-Assisted Intervention. Copenhagen, Denmark: Springer: 924-931 [DOI: 10.1007/11866565_113]
- Ashburner J. 2007. A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1): 95-113 [DOI: 10.1016/j.neuroimage.2007.07.007]
- Auer M, Regitnig P and Holzapfel G A. 2005. An automatic nonrigid registration for stained histological sections. *IEEE Transactions on Image Processing*, 14(4): 475-486 [DOI: 10.1109/TIP.2005.843756]
- Avants B B, Epstein C L, Grossman M and Gee J C. 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1): 26-41 [DOI: 10.1016/j.media.2007.06.004]
- Avants B B, Tustison N J, Song G, Cook P A, Klein A and Gee J C. 2011. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, 54(3): 2033-2044 [DOI: 10.1016/j.neuroimage.2010.09.025]
- Balakrishnan G, Zhao A, Sabuncu M R, Dalca A V and Guttag J V. 2018. An unsupervised learning model for deformable medical image registration//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 9252-9260 [DOI: 10.1109/cvpr.2018.00964]
- Balakrishnan G, Zhao A, Sabuncu M R, Guttag J and Dalca A V. 2019. VoxelMorph: a learning framework for deformable medical image registration. *IEEE Transactions on Medical Imaging*, 38(8): 1788-1801 [DOI: 10.1109/TMI.2019.2897538]
- Barnea D I and Silverman H F. 1972. A class of algorithms for fast digital image registration. *IEEE Transactions on Computers*, C-21(2): 179-186 [DOI: 10.1109/TC.1972.5008923]
- Berendsen F F, Kotte A N T J, Viergever M A and Pluim J P W. 2014. Registration of organs with sliding interfaces and changing topologies//Proceedings of SPIE 9034, Medical Imaging 2014: Image Processing. San Diego, USA: SPIE: 3877-2877 [DOI: 10.1117/12.2043447]
- Boveiri H R, Khayami R, Javidan R and Mehdizadeh A R. 2020. Medical image registration using deep neural networks: a comprehensive

- review [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/2002.03401.pdf>
- Cao X H, Yang J H, Wang L, Xue Z, Wang Q and Shen D G. 2018b. Deep learning based inter-modality image registration supervised by intra-modality similarity//Proceedings of the 9th International Workshop on Machine Learning in Medical Imaging. Granada, Spain; Springer: 55-63 [DOI: 10.1007/978-3-030-00919-9_7]
- Cao X H, Yang J H, Zhang J, Nie D, Kim M, Wang Q and Shen D G. 2017. Deformable image registration based on similarity-steered CNN regression//Proceedings of the 20th Medical Image Computing and Computer-Assisted Intervention. Quebec City, Canada; Springer: 300-308 [DOI: 10.1007/978-3-319-66182-7_35]
- Cao X H, Yang J H, Zhang J, Wang Q, Yap P T and Shen D G. 2018a. Deformable image registration using a cue-aware deep regression network. *IEEE Transactions on Biomedical Engineering*, 65(9): 1900-1911 [DOI: 10.1109/TBME.2018.2822826]
- Chee E and Wu Z Z. 2018. AIRNet: self-supervised affine registration for 3D medical images using neural networks [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1810.02583.pdf>
- Cheng X, Zhang L and Zheng Y F. 2018. Deep similarity learning for multimodal medical images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, 6(3): 248-252 [DOI: 10.1080/21681163.2015.1135299]
- Choi Y, Choi M, Kim M, Ha J W, Kim S and Choo J. 2018. StarGAN: unified generative adversarial networks for multi-domain image-to-image translation//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 8789-8797 [DOI: 10.1109/cvpr.2018.00916]
- Chollet F. 2017. Xception: deep learning with depthwise separable convolutions//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA; IEEE: 1800-1807 [DOI: 10.1109/CVPR.2017.195]
- Çiçek O, Abdulkadir A, Lienkamp S S, Brox T and Ronneberger O. 2016. 3D U-Net: learning dense volumetric segmentation from sparse annotation//Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens, Greece; Springer: 424-432 [DOI: 10.1007/978-3-319-46723-8_49]
- de Vos B D, Berendsen F F, Viergever M A, Sokooti H, Staring M and Išgum I. 2019. A deep learning framework for unsupervised affine and deformable image registration. *Medical Image Analysis*, 52: 128-143 [DOI: 10.1016/j.media.2018.11.010]
- de Vos B D, Berendsen F F, Viergever M A, Staring M and Išgum I. 2017. End-to-end unsupervised deformable image registration with a convolutional neural network//Proceedings of the 3rd International Workshop on Multimodal Learning for Clinical Decision Support. Québec City, Canada; Springer: 204-212 [DOI: 10.1007/978-3-319-67558-9_24]
- Dosovitskiy A, Fischer P, Ilg E, Häusser P, Hazirbas C, Golkov V, van der Smagt P, Cremers D and Brox T. 2015. FlowNet: learning optical flow with convolutional networks//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile; IEEE: 2758-2766 [DOI: 10.1109/ICCV.2015.316]
- Elmahdy M S, Wolterink J M, Sokooti H, Išgum I and Staring M. 2019. Adversarial optimization for joint registration and segmentation in prostate CT radiotherapy//Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen, China; Springer: 366-374 [DOI: 10.1007/978-3-030-32226-7_41]
- Eppenhof K A J, Lafarge M W, Moeskops P, Veta M and Pluim J P W. 2018. Deformable image registration using convolutional neural networks//Proceedings of SPIE 10574, Medical Imaging 2018: Image Processing. Houston, USA; SPIE: #105740S [DOI: 10.1117/12.2292443]
- Eppenhof K A J and Pluim J P W. 2018. Error estimation of deformable image registration of pulmonary CT scans using convolutional neural networks. *Journal of Medical Imaging*, 5(2): #024003 [DOI: 10.1117/1.JMI.5.2.024003]
- Eppenhof K A J and Pluim J P W. 2019. Pulmonary CT registration through supervised learning with convolutional neural networks. *IEEE Transactions on Medical Imaging*, 38(5): 1097-1105 [DOI: 10.1109/TMI.2018.2878316]
- Fan J F, Cao X H, Wang Q, Yap P T and Shen D G. 2019b. Adversarial learning for mono- or multi-modal registration. *Medical Image Analysis*, 58: #101545 [DOI: 10.1016/j.media.2019.101545]
- Fan J F, Cao X H, Xue Z, Yap P T and Shen D G. 2018. Adversarial similarity network for evaluating image alignment in deep learning based registration//Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention. Granada, Spain; Springer: 739-747 [DOI: 10.1007/978-3-030-00928-1_83]
- Fan J F, Cao X H, Yap P T and Shen D G. 2019a. BIRNet: brain image registration using dual-supervised fully convolutional networks. *Medical Image Analysis*, 54: 193-206 [DOI: 10.1016/j.media.2019.03.006]
- Fu Y B, Lei Y, Wang T H, Curran W J, Liu T and Yang X F. 2020a. Deep learning in medical image registration: a review. *Physics in Medicine and Biology*, 62(20): #20TR01 [DOI: 10.1088/1361-6560/ab843e]
- Fu Y B, Lei Y, Wang T H, Higgins K, Bradley J D, Curran W J, Liu T and Yang X. 2020b. LungRegNet: an unsupervised deformable image registration method for 4D-CT lung. *Medical Physics*, 47(4): 1763-1774 [DOI: 10.1002/mp.14065]
- Gendrin C, Furtado H, Weber C, Bloch C, Figl M, Pawiro S A, Bergmann H, Stock M, Fichtinger G, Georg D and Birkfellner W. 2012. Monitoring tumor motion by real time 2D/3D registration during radiotherapy. *Radiotherapy and Oncology*, 102(2): 274-280 [DOI: 10.1016/j.radonc.2011.07.031]

- Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y. 2014. Generative adversarial nets//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montréal, Canada: NIPS: 2672-2680 [DOI: 10.5555/2969033.2969125]
- Hatt C R, Speidel M A and Raval A N. 2015. Robust 5DOF transesophageal echo probe tracking at fluoroscopic frame rates//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer: 290-297 [DOI: 10.1007/978-3-319-24553-9_36]
- Heinrich M P, Jenkinson M, Bhushan M, Matin T, Gleeson F V, Brady S M and Schnabel J A. 2012. MIND: modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis*, 16 (7): 1423-1435 [DOI: 10.1016/j.media.2012.05.008]
- Hering A, Kuckertz S, Heldmann S and Heinrich M P. 2019. Enhancing label-driven deep deformable image registration with local distance metrics for state-of-the-art cardiac motion tracking//Proceedings of Bildverarbeitung für die Medizin. Lubeck, Germany: Springer: 309-314 [DOI: 10.1007/978-3-658-25326-4_69]
- Hu Y P, Gibson E, Ghavami N, Bonmati E, Moore C M, Emberton M, Vercauteren T, Noble J A and Barratt D C. 2018b. Adversarial deformation regularization for training image registration neural networks//Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention. Granada, Spain: Springer: 774-782 [DOI: 10.1007/978-3-030-00928-1_87]
- Hu Y P, Modat M, Gibson E, Ghavami N, Bonmati E, Moore C M, Emberton M, Noble J A, Barratt D C and Vercauteren T. 2018a. Label-driven weakly-supervised learning for multimodal deformable image registration//The 15th IEEE International Symposium on Biomedical Imaging. Washington, USA: IEEE: 1070-1074 [DOI: 10.1109/ISBI.2018.8363756]
- Hu Y P, Modat M, Gibson E, Li W Q, Ghavami N, Bonmati E, Wang G T, Bandula S, Moore C M, Emberton M, Ourselin S, N J A, Barratt D C and Vercauteren T. 2018c. Weakly-supervised convolutional neural networks for multimodal image registration. *Medical Image Analysis*, 49: 1-13 [DOI: 10.1016/j.media.2018.07.002]
- Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T and Keutzer K. 2014. DenseNet: implementing efficient convnet descriptor pyramids [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1404.1869.pdf>
- Isola P, Zhu J Y, Zhou T H and Efros A A. 2017. Image-to-image translation with conditional adversarial networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 5967-5976 [DOI: 10.1109/CVPR.2017.632]
- Jaderberg M, Simonyan K, Zisserman A and Kavukcuoglu K. 2015. Spatial transformer networks//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: NIPS: 2017-2025 [DOI: 10.5555/2969442.2969465]
- Jiang Z R, Yin F F, Ge Y and Ren L. 2020. A multi-scale framework with unsupervised joint training of convolutional neural networks for pulmonary deformable image registration. *Physics in Medicine and Biology*, 65(1): #015011 [DOI: 10.1088/1361-6560/ab5da0]
- Kabus S, Netsch T, Fischer B and Modersitzki J. 2004. B-spline registration of 3D images with Levenberg-Marquardt optimization//Proceedings of SPIE 5370, Medical Imaging 2004: Image Processing. San Diego, USA: SPIE: 304-313 [DOI: 10.1117/12.533976]
- Kim B, Kim J, Lee J G, Kim D H, Park S H and Ye J C. 2019. Unsupervised deformable image registration using cycle-consistent CNN//Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen, China: Springer: 166-174 [DOI: 10.1007/978-3-030-32226-7_19]
- Kori A and Krishnamurthi G. 2019. Zero shot learning for multi-modal real time image registration [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1908.06213.pdf>
- Krebs J, Mansi T, Delingette H, Zhang L, Ghesu F C, Miao S, Maier A K, Ayache N, Liao R and Kamen A. 2017. Robust non-rigid registration through agent-based action learning//Proceedings of the 20th International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec City, Canada: Springer: 344-352 [DOI: 10.1007/978-3-319-66182-7_40]
- Krizhevsky A, Sutskever I and Hinton G E. 2012. ImageNet classification with deep convolutional neural networks//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: NIPS: 1097-1105 [DOI: 10.5555/2999134.2999257]
- Kuang D Y. 2019. On reducing negative Jacobian determinant of the deformation predicted by deep registration networks [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1907.00068.pdf>
- Lei Y, Fu Y B, Harms J, Wang T H, Curran J W, Liu T, Higgins X and Yang X F. 2019. 4D-CT deformable image registration using an unsupervised deep convolutional neural network//Proceedings of the 1st Workshop on Artificial Intelligence in Radiation Therapy. Shenzhen, China: Springer: 26-33 [DOI: 10.1007/978-3-030-32486-5_4]
- Li H M and Fan Y. 2018. Non-rigid image registration using self-supervised fully convolutional networks without training data//The 15th IEEE International Symposium on Biomedical Imaging. Washington, USA: IEEE: 1075-1078 [DOI: 10.1109/isbi.2018.8363757]
- Li X F, Zhang C L, Li H P and Zang X B. 2010. Development of medical image registration technology. *Computer Science*, 37(7): 27-33 (李雄飞, 张存利, 李鸿鹏, 臧雪柏. 2010. 医学图像配准技术进展. *计算机科学*, 37(7): 27-33) [DOI: 10.3969/j.issn.1002-137X.2010.07.006]
- Liao R, Miao S, De Tournemire P, Grbic S, Kamen A, Mansi T and

- Comaniciu D. 2016. An artificial agent for robust image registration [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1611.10336.pdf>
- Lin J Q, Han J, Yuan Z M and Peng J L. 2020. Orientation correction for CT images via multitask deep convolutional network. *Journal of Huaqiao University (Natural Science)*, 41(3): 366-373 (林家庆, 韩娟, 袁直敏, 彭佳林. 2020. 多任务深度卷积网络的CT图像方向校正. *华侨大学学报(自然科学版)*, 41(3): 366-373) [DOI: 10.11830/ISSN.1000-5013.201911039]
- Liu C, Ma L H, Lu Z M, Jin X C and Xu J Y. 2019. Multimodal medical image registration via common representations learning and differentiable geometric constraints. *Electronics Letters*, 55(6): 316-318 [DOI: 10.1049/el.2018.6713]
- Liu M Y, Breuel T and Kautz J. 2017. Unsupervised image-to-image translation networks//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, USA: NIPS; 700-708 [DOI: 10.5555/3294771.3294838]
- Liu Q G and Leung H. 2017. Tensor-based descriptor for image registration via unsupervised network//*Proceedings of the 20th IEEE International Conference on Information Fusion*. Xi'an, China; IEEE; 1-7 [DOI: 10.23919/ICIF.2017.8009689]
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, USA; IEEE; 3431-3440 [DOI: 10.1109/CVPR.2015.7298965]
- Luo J, Frisken S, Wang D, Golby A J, Sugiyama M and Wells III W. 2020a. Are registration uncertainty and error monotonically associated?//*Proceedings of the 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention*. Lima, Peru; Springer; 264-274 [DOI: 10.1007/978-3-030-59716-0_26]
- Luo J, Ma G S, Frisken S, Juvekar P, Haouchine N, Xu Z, Xiao Y M, Golby A, Codd P, Sugiyama M and Wells III W. 2020b. Do public datasets assure unbiased comparisons for registration evaluation? [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/2003.09483.pdf>
- Luo J, Sedghi A, Popuri K, Cobzas D, Zhang M M, Preiswerk F, Toews M, Golby A, Sugiyama M, Wells III W M and Frisken S. 2019. On the applicability of registration uncertainty//*Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention*. Shenzhen, China; Springer; 410-419 [DOI: 10.1007/978-3-030-32245-8_46]
- Luo J, Toews M, Machado I, Frisken S F, Zhang M M, Preiswerk F, Sedghi A, Ding H Y, Pieper S, Golland P, Golby A, Sugiyama M and Wells III W M. 2018. A feature-driven active framework for ultrasound-based brain shift compensation//*Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention*. Granada, Spain; Springer; 30-38 [DOI: 10.1007/978-3-030-00937-3_4]
- Luo S Q and Lyu W X. 1999. Medical image registration technology. *Foreign medical science*. *International Journal of Biomedical Engineering*, (1): 1-8 (罗述谦, 吕维雪. 1999. 医学图像配准技术: 国外医学. 生物医学工程分册, (1): 1-8)
- Mahapatra D. 2018. Elastic registration of medical images with GANs [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1805.02369v1.pdf>
- Mahapatra D, Antony B, Sedai S and Garnavi R. 2018b. Deformable medical image registration using generative adversarial networks//*The 15th IEEE International Symposium on Biomedical Imaging*. Washington, USA; IEEE; 1449-1453 [DOI: 10.1109/isbi.2018.8363845]
- Mahapatra D, Ge Z Y, Sedai S and Chakravorty R. 2018a. Joint registration and segmentation of xray images using generative adversarial networks//*Proceedings of the 9th International Workshop on Machine Learning in Medical Imaging*. Granada, Spain; Springer; 73-80 [DOI: 10.1007/978-3-030-00919-9_9]
- Miao S, Wang Z J and Liao R. 2016a. A CNN regression approach for real-time 2D/3D registration. *IEEE Transactions on Medical Imaging*, 35(5): 1352-1363 [DOI: 10.1109/TMI.2016.2521800]
- Miao S, Wang Z J, Zheng Y F and Liao R. 2016b. Real-time 2D/3D registration via CNN regression//*The 13th IEEE International Symposium on Biomedical Imaging*. Prague, Czech; IEEE; 1430-1434 [DOI: 10.1109/ISBI.2016.7493536]
- Öfverstedt J, Lindblad J and Sladoje N. 2019. Fast and robust symmetric image registration based on distances combining intensity and spatial information. *IEEE Transactions on Image Processing*, 28(7): 3584-3597 [DOI: 10.1109/TIP.2019.2899947]
- Oliveira F P M and Tavares J M R S. 2014. Medical image registration: a review. *Computer Methods in Biomechanics and Biomedical Engineering*, 17(2): 73-93 [DOI: 10.1080/10255842.2012.670855]
- Onieva J O, Marti-Fuster B, de la Puente M P and Estépar R S J. 2018. Diffeomorphic lung registration using deep cnns and reinforced learning//*Proceedings of the 3rd International Workshop on Reconstruction and Analysis of Moving Body Organs*. Granada, Spain; Springer; 284-294 [DOI: 10.1007/978-3-030-00946-5_28]
- Qin C, Bai W J, Schlemper J, Petersen S E, Piechnik S K, Neubauer S and Rueckert D. 2018. Joint learning of motion estimation and segmentation for cardiac MR image sequences//*Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention*. Granada, Spain; Springer; 472-480 [DOI: 10.1007/978-3-030-00934-2_53]
- Qin C, Shi B B, Liao R, Mansi T, Rueckert D and Kamen A. 2019. Unsupervised deformable registration for multi-modal images via disentangled representations//*Proceedings of the 29th International Conference on Information Processing in Medical Imaging*. Hong Kong, China; Springer; 249-261 [DOI: 10.1007/978-3-030-20351-1_19]
- Qiu Z F, Yao T and Mei T. 2017. Learning spatio-temporal representation with pseudo-3D residual networks//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy; IEEE;

- 5534-5542 [DOI: 10.1109/icc.2017.590]
- Rohé M M, Datar M, Heimann T, Sermesant M and Pennec X. 2017. SVF-Net: learning deformable image registration using shape matching//Proceedings of the 20th International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec City, Canada; Springer: 266-274 [DOI: 10.1007/978-3-319-66182-7_31]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany; Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Ruhaak J, Heldmann S, Kipshagen T and Fischer B. 2013. Highly accurate fast lung CT registration//Proceedings of SPIE 8669, Medical Imaging 2013: Image Processing. Lake Buena Vista, USA; SPIE: #86690Y [DOI: 10.1117/12.2006035]
- Sabour S, Frosst N and Hinton G E. 2017. Dynamic routing between capsules [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1710.09829.pdf>
- Salehi S S M, Khan S, Erdogmus D and Gholipour A. 2018. Real-time deep registration with geodesic loss [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1803.05982v1.pdf>
- Schmid J and Chênes C. 2014. Segmentation of X-ray images by 3D-2D registration based on multibody physics//Proceedings of the 12th Asian Conference on Computer Vision. Singapore, Singapore; Springer: 674-687 [DOI: 10.1007/978-3-319-16808-1_45]
- Sedghi A, Luo J, Mehrtash A, Pieper S, Tempany C M, Kapur T, Mousavi P and Wells III W M. 2019. Semi-supervised image registration using deep learning//Proceedings of SPIE 10951, Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling. San Diego, USA; SPIE: #109511G [DOI: 10.1117/12.2513020]
- Sentker T, Madesta F and Werner R. 2018. GDL-FIRE^{4D}: deep learning-based fast 4D CT image registration//Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention. Granada, Spain; Springer: 765-773 [DOI: 10.1007/978-3-030-00928-1_86]
- Shan S Y, Yan W, Guo X Q, Chang E I C, Fan Y B and Xu Y. 2017. Unsupervised end-to-end learning for deformable medical image registration [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1711.08608.pdf>
- Sheikhjafari A, Noga M, Punithakumar K and Ray N. 2019. Unsupervised deformable image registration with fully connected generative neural network//Proceedings of the Medical Imaging with Deep Learning. Amsterdam, the Netherlands; [s. n.]
- Shekhar R and Zagrodsky V. 2002. Mutual information-based rigid and nonrigid registration of ultrasound volumes. IEEE Transactions on Medical Imaging, 21(1): 9-22 [DOI: 10.1109/42.981230]
- Shen Y Y and Feng H S. 2020. 2D-3D double X-ray image registration method based on neural network. Chinese Journal of Medical Physics, 37(3): 293-298 (沈延延, 冯汉升. 2020. 基于神经网络的双X射线影像2D-3D配准算法. 中国医学物理学杂志, 37(3): 293-298) [DOI: 10.3969/j.issn.1005-202X.2020.03.007]
- Simonovsky M, Gutiérrez-Becker B, Mateus D, Navab N and Komodakis N. 2016. A deep metric for multimodal registration//Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens, Greece; Springer: 10-18 [DOI: 10.1007/978-3-319-46726-9_2]
- Singh N K and Raza K. 2020. Medical image generation using generative adversarial networks [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/2005.10687.pdf>
- Sloan J M, Goatman K A and Siebert J P. 2018. Learning rigid image registration – utilizing convolutional neural networks for medical image registration//Proceedings of the 11th Biomedical Engineering Systems and Technologies. Madeira, Portugal; Springer: 89-99 [DOI: 10.5220/0006543700890099]
- Sokooti H, de Vos B, Berendsen F, Ghafoorian M, Yousefi S, Lelieveldt B P, Išgum I and Staring M. 2019. 3D convolutional neural networks image registration based on efficient supervised learning from artificial deformations [EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1908.10235.pdf>
- Sokooti H, de Vos B, Berendsen F, Lelieveldt B P F, Išgum I and Staring M. 2017. Nonrigid image registration using multi-scale 3D convolutional neural networks//Proceedings of the 20th International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec City, Canada; Springer: 232-239 [DOI: 10.1007/978-3-319-66182-7_27]
- Sun L and Zhang S T. 2018. Deformable mri-ultrasound registration using 3D convolutional neural network//Proceedings of the 1st International Workshop on Computational Precision Medicine. Granada, Spain; Springer: 152-158 [DOI: 10.1007/978-3-030-01045-4_18]
- Sun Y Y, Moelker A, Niessen W J and van Walsum T. 2018. Towards robust CT-ultrasound registration using deep learning methods//Proceedings of the 1st International Workshop on Interpretability of Machine Intelligence in Medical Image Computing. Granada, Spain; Springer: 43-51 [DOI: 10.1007/978-3-030-02628-8_5]
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2015. Going deeper with convolutions//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA; IEEE: 1-9 [DOI: 10.1109/CVPR.2015.7298594]
- Tang L S, Hamarneh G and Celler A. 2006. Co-registration of bone CT and SPECT images using mutual information//IEEE International Symposium on Signal Processing and Information Technology. Vancouver, Canada; IEEE: 116-121 [DOI: 10.1109/ISSPIT.2006.270781]
- Uzunova H, Wilms M, Handels H and Ehrhardt J. 2017. Training CNNs

- for image registration from few samples with model-based data augmentation//Proceedings of the 20th International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec City, Canada: Springer; 223-231 [DOI: 10.1007/978-3-319-66182-7_26]
- Vercauteren T, Pennec X, Perchant A and Ayache N. 2009. Diffeomorphic demons: efficient non-parametric image registration. *NeuroImage*, 45(1): 61-72 [DOI: 10.1016/j.neuroimage.2008.10.040]
- Wang C F and Jiang M. 2006. Review of image registration methods for medical images. *CT Theory and Application*, 15(2): 74-80 (王彩芳, 姜明. 2006. 医学图像配准综述. *CT理论与应用研究*, 15(2): 74-80) [DOI: 10.3969/j.issn.1004-4140.2006.02.015]
- Wang H N, He C Y, Lei F Y and Zhang X Y. 2005. A survey of non-rigid medical image registration. *Computer Engineering and Applications*, 41(11): 180-184 (王海南, 郝重阳, 雷方元, 张先勇. 2005. 非刚性医学图像配准研究综述. *计算机工程与应用*, 41(11): 180-184) [DOI: 10.3321/j.issn:1002-8331.2005.11.059]
- Wang L F, Zhang C C, Qin P L, Lin S Z, Gao Y and Dou J L. 2020. Image registration method with residual dense relativistic average CGAN. *Journal of Image and Graphics*, 25(4): 745-758 (王丽芳, 张程程, 秦品乐, 蔺素珍, 高媛, 窦杰亮. 2020a. 残差密集相对平均CGAN的脑部图像配准. *中国图象图形学报*, 25(4): 745-758) [DOI: 10.11834/jig.190116]
- Wang S W, Zhao Q Y, Wang X Y and Rao X P. 2020. Region division method of brain slice image based on deep learning. *Journal of Computer Applications*, 40(4): 1202-1208 (王松伟, 赵秋阳, 王宇航, 饶小平. 2020. 基于深度学习的脑片图像区域划分方法. *计算机应用*, 40(4): 1202-1208) [DOI: 10.11772/j.issn.1001-9081.2019091521]
- Wilms M, Handels H and Ehrhardt J. 2017. Multi-resolution multi-object statistical shape models based on the locality assumption. *Medical Image Analysis*, 38: 17-29 [DOI: 10.1016/j.media.2017.02.003]
- Wu G R, Kim M, Wang Q, Gao Y Z, Liao S and Shen D G. 2013. Unsupervised deep feature learning for deformable registration of MR brain images//Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention. Nagoya, Japan: Springer; 649-656 [DOI: 10.1007/978-3-642-40763-5_80]
- Xu Z, Luo J, Yan J P, Pulya R, Li X, Wells III W and Jagadeesan J. 2020. Adversarial uni- and multi-modal stream networks for multi-modal image registration//Proceedings of the 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention. Lima, Peru: Springer; 222-232
- Yan J P, Chen S, Zhang Y B and Li X. 2020. Neural architecture search for compressed sensing magnetic resonance image reconstruction. *Computerized Medical Imaging and Graphics*, 85: #101784 [DOI: 10.1016/j.compmedimag.2020.101784]
- Yan P K, Xu S, Rastinehad A R and Wood B J. 2018. Adversarial image registration with application for MR and TRUS image fusion//Proceedings of the 9th International Workshop on Machine Learning in Medical Imaging. Granada, Spain: Springer; 197-204 [DOI: 10.1007/978-3-030-00919-9_23]
- Yang X, Kwitt R and Niethammer M. 2016. Fast predictive image registration//Proceedings of the 1st International Workshop on Deep Learning in Medical Image Analysis. Athens, Greece: Springer; 48-57 [DOI: 10.1007/978-3-319-46976-8_6]
- Yang X, Kwitt R, Styner M and Niethammer M. 2017. Quicksilver: fast predictive image registration – a deep learning approach. *NeuroImage*, 158: 378-396 [DOI: 10.1016/j.neuroimage.2017.07.008]
- Yi Z L, Zhang H, Tan P and Gong M L. 2017. DualGAN: unsupervised dual learning for image-to-image translation//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE; 2868-2876 [DOI: 10.1109/ICCV.2017.310]
- Yoo I, Hildebrand D G C, Tobin W F, Lee W C A and Jeong W K. 2017. ssEMnet: serial-section electron microscopy image registration using a spatial transformer network with learned features//Proceedings of the 3rd International Workshop on Multimodal Learning for Clinical Decision Support. Québec City, Canada: Springer; 249-257 [DOI: 10.1007/978-3-319-67558-9_29]
- Yu H J, Zhou X R, Jiang H Y, Kang H J, Wang Z G, Hara T and Fujita H. 2019. Learning 3D non-rigid deformation based on an unsupervised deep learning for PET/CT image registration//Proceedings of SPIE 10953, Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging. San Diego, USA: SPIE; #109531X [DOI: 10.1117/12.2512698]
- Zhang J. 2018. Inverse-consistent deep networks for unsupervised deformable image registration[EB/OL]. [2020-07-04]. <https://arxiv.org/pdf/1809.03443.pdf>
- Zhao S Y, Dong Y, Chang E and Xu Y. 2019. Recursive cascaded networks for unsupervised medical image registration//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea(South): IEEE; 10599-10609 [DOI: 10.1109/ICCV.2019.01070]
- Zhao S Y, Lau T, Luo J, Chang E I C and Xu Y. 2020. Unsupervised 3 d end-to-end medical image registration with volume tweening network. *IEEE Journal of Biomedical and Health Informatics*, 24(5): 1394-1404 [DOI: 10.1109/JBHI.2019.2951024]
- Zheng J N, Miao S and Liao R. 2017. Learning CNNs with pairwise domain adaption for real-time 6 DoF ultrasound transducer detection and tracking from x-ray images//Proceedings of the 20th International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec City, Canada: Springer; 646-654 [DOI: 10.1007/978-3-319-66185-8_73]
- Zheng J N, Miao S, Wang Z J and Liao R. 2018. Pairwise domain adaptation module for CNN-based 2-D/3-D registration. *Journal of Medi-*

cal Imaging, 5 (2): #021204 [DOI: 10.1117/1.JMI.5.2.021204]

Zhu J Y, Park T, Isola P and Efros A A. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2242-2251 [DOI: 10.1109/ICCV.2017.244]

Zou M Y, Yang H, Pang G H and Zhong Y. 2019. Research progress and challenges of deep learning in medical image registration. Journal of Biomedical Engineering, 36(4): 677-683 (邹茂扬, 杨昊, 潘光辉, 钟勇. 2019. 深度学习在医学图像配准上的研究进展与挑战. 生物医学工程学杂志, 36(4): 677-683) [DOI: 10.7507/1001-5515.201810004]

作者简介



马露凡, 1997年生, 女, 硕士研究生, 主要研究方向为深度学习和医学图像分析。

E-mail: malf19@mails.tsinghua.edu.cn



李秀, 通信作者, 女, 教授, 博士生导师, 主要研究方向为智能系统、深度学习与模式识别。

E-mail: li.xiu@sz.tsinghua.edu.cn

罗凤, 女, 硕士研究生, 主要研究方向为深度学习和医学图像分析。E-mail: luof19@mails.tsinghua.edu.cn

严江鹏, 男, 博士研究生, 主要研究方向为深度学习和医学图像分析。E-mail: yanjp17@mails.tsinghua.edu.cn

徐哲, 男, 硕士研究生, 主要研究方向为深度学习和医学图像分析。E-mail: xu-z18@mails.tsinghua.edu.cn

罗捷, 男, 博士研究生, 主要研究方向为医学图像配准。

E-mail: jluc5@bwh.harvard.edu