

基于深度学习分割掩膜的胸片图像配准技术及其应用

刘远明¹, 王小义², 郭琳¹, 夏丽¹, 权中文¹, 钱令军¹, 李宏军³

1. 深圳市智影医疗科技有限公司, 广东 深圳 518000; 2. 青海省第四人民医院, 青海 西宁 810007; 3. 首都医科大学附属北京佑安医院, 北京 100069

【摘要】医学图像配准技术对临床诊断治疗具有重要意义。相比传统的图像配准方法, 目前基于深度学习的配准方法提高了配准的精度和速度。为了将深度学习图像配准技术应用于胸片的配准以及减影分析, 本研究先采用深度学习掩膜对原始胸片进行预处理, 然后以掩膜图像作为输入, 以ResUNet网络作为配准框架来实现胸片图像配准, 最后评估配准效果。结果显示深度学习掩膜结合深度学习图像配准方法训练出的模型在胸片配准上具有良好的图像配准精度。这种基于掩膜的深度学习配准模型可以较好地应用于胸片的减影分析。

【关键词】胸片; 深度学习; 分割掩膜; 图像配准; 减影分析

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2022)10-1231-05

Chest radiograph registration technique based on segmentation mask obtained by deep learning and its application

LURE Fleming Yuanming¹, WANG Xiaoyi², GUO Lin¹, XIA Li¹, QUAN Shenwen¹, QIAN LingJun¹, LI Hongjun³

1. Shenzhen Zhiying Medical Imaging Co., Ltd. Shenzhen 518000, China; 2. The Fourth People's Hospital of Qinghai Province, Xi'ning 810007, China; 3. Beijing Youan Hospital, Capital Medical University, Beijing 100069, China

Abstract: Medical image registration technique is of vital importance for clinical diagnosis and treatment. The deep learning based registration method improves the accurate and speed of registration when compared with conventional registration methods. In order to apply deep learning algorithm to do chest radiograph registration and subsequent subtraction analysis, the original chest radiographs is preprocessed with the use of the mask obtained by deep learning, and the chest radiograph registration is achieved with mask images as input, ResUnet as registration structure. The evaluation of the registration results showed that the model developed by mask and registration technique based on deep learning has high image registration accuracy in chest radiograph registration. The proposed registration model can be well applied to the subtraction analysis of chest radiographs.

Keywords: chest radiograph; deep learning; segmentation mask; image registration; subtraction analysis

前言

可变形图像配准是医学图像分析领域的一项重要技术^[1], 在疾病诊断、疾病治疗、术前术后比对等应用中

有着至关重要的作用。医学背景下, 图像配准被用于不同患者、不同时间、不同模态的两个或多个图像的对齐, 这种对齐一般通过变形场实现, 先使源图像产生变形, 然后使变形后的图像与目标图像对齐, 配准过程中采用相似性度量源图像与目标图像的匹配程度。

目前深度学习在医学图像处理方面获得了出色表现^[2], 能够帮助我们在完全不同的模式下进行图像配准。深度卷积神经网络(Convolutional Neural Networks, CNN)应用于计算机视觉任务(如图像分割和图像分类)中的效果优于之前的算法^[3-5]。随着图像配准技术的发展, 其在配准精度以及计算时间上的性能不断得到优化。Balakrishnan等^[1]提出了VoxelMorph快速学习框架, 给定新的扫描图像对后, VoxelMorph能够直接评估并快速计算变形场, 加快

【收稿日期】2022-05-26

【基金项目】国家重点研发计划(2019YFE0121400); 深圳市科技计划项目(KQTD2017033110081833, JCYJ20190813153413160, JSGG20201102162802008)

【作者简介】刘远明, 博士, 教授, 研究方向: 机器学习与人工智能神经网络, E-mail: f.lure@hotmail.com; 王小义, 主任医师, 研究方向: 感染与炎症放射学, E-mail: qhssyfsk@126.com(刘远明和王小义共为第一作者)

【通信作者】李宏军, 博士, 主任医师, 研究方向: 感染与炎症, E-mail: lihongjun00113@126.com; 郭琳, 博士, 研究方向: 人工智能与医学影像, E-mail: guolin913@outlook.com

图像配准的分析、处理流程。基于变形先验模型的图像配准方法可应用于低信噪比的同一模态图像配准,适合用于常规配准方法无法实现精确配准的超声及MRI图像^[6]。为了解决当前基于深度学习的配准网络需要不断迭代的问题,Sokooti等^[7]提出采用大量人工生成的位移矢量场进行网络训练,最后得到一种一次完成非刚性配准的学习网络RegNet。此外,其他学者在对不同配准对象实现高精度配准这一问题上提出了各自的研究路线,如利用标志点集优化可变形图像配准的空间精度以及采用人工合成的变形图像对,来训练深度学习网络,从而对变形较大的图像实现高准确度的配准^[8-9]。在缩短图像配准计算时间方面,Woods等^[10]提出和验证了基于体素强度匹配的自动图像配准方法,通过平衡速度和准确性之间的关系,可让MRI图像配准的准确度达到75~150 μm ,在准确度损失很小的前提下大大减少了配准时间。Ashburner等^[11]提出了一种用于微形图像配准的算法,采用恒定欧拉速度值,实现快速的矩阵变换和计算,从而提高配准速度。

基于Lucas等^[12]的研究成果,许多可变形图像配准建模的研究陆续出现^[13],其中大部分图像配准问题被视为优化问题,一般可以表示为:

$$f^* = \arg \min_{S_{\chi}} (I_M \cdot f, I_F) + \lambda R(f, \chi) \quad (1)$$

其中, f^* 表示图像对(源图像与目标图像)的优化对比, I_M 表示源图像(移动), I_F 表示目标图像(固定), f 表示源图像到目标图像每一点进行空间转换的位移量, χ 表示一个规则网格中排列的 N 个点的集合, R 表示正则化项, λ 表示平衡因子, S_{χ} 表示源图像经过空间转换后生成图像与目标图像的相似度测量值。图像配准的过程就是计算源图像经过空间转换后与目标图像产生最大相似度的优化问题。这里采用均方误差(Mean Square Error, MSE)来度量相似度^[14-15]:

$$S_{\text{MSE}} = \frac{1}{|\chi|} \sum_{x \in \chi} (I_M(x + f(x)) - I_F(x))^2 \quad (2)$$

其中, S_{MSE} 表示均方误差。典型的用于平滑转换形变场的正则化项可以表示为:

$$R_{\text{diff}} = \frac{1}{|x|} \sum_{x \in \chi} \sum_{i=1}^d \|\nabla f_i(x)\|^2 \quad (3)$$

其中, R_{diff} 表示源图像与目标图像空间变换的正则化。

准确分析两张图像的差异非常重要,现有的技术大多对变形场施加平滑约束,或将像素级损失纳入目标函数中,以增强结构的合理性。然而,平滑的变形场不一定能产生真实客观的结果,当两张图像器官相对位置距离太大时,普通平滑变形场产生的

形变图像会在结构上失真,为了保护图像的形态结构,就需要对其做出合理的形变。客观量化医学图像形变的合理性,并将其融于基于深度学习的配准模型训练过程是一项复杂的任务,本研究针对这一问题采用深度学习掩膜来参与训练,即先用掩膜对图像进行预处理,然后基于掩膜图像去训练模型。本研究主要针对胸片进行配准模型的优化,从而将其应用于胸片的减影分析。

1 材料和方法

1.1 图像预处理

图像预处理过程一般可以分为图像去噪、图像增强、图像分割和图像增广。研究采用高斯滤波平滑以及自适应直方均衡化分别对图像进行去噪和增强。进行图像分割时,研究首先获取了胸部X光片影像,然后对影像中的肺区进行标记,并用ResUNet对标记后的图像进行分割模型训练以获取掩膜^[3, 16-18]。ResUNet由UNet和ResNet组成,其中UNet部分包括4个编码层和4个解码层,在UNet的编码层之前加一层ResNet卷积层作为编码层,在解码层之后加一层ResNet卷积层作为解码层。本研究中,3位高年资影像科医生对肺区进行平行标注,对医生之间标注结果Dice值 $\geq 98\%$ 的勾勒区域取平均值作为金标准,Dice值 $< 98\%$ 的影像由3位医生进行讨论,以讨论后的结果作为金标准。一共标注406例病例,年龄15~82岁,其中,男198例,女208例,正常样本112例,异常样本296例;其中326例作为肺区分割训练集(20例作为验证集),80例作为测试集。测试集在训练后模型上的Dice值为97.4%。最后在训练数据上进行水平翻转、旋转、缩放、亮度变化等一系列操作以实现图像增广。

1.2 图像配准网络结构

图像配准网络架构遵循类似于UNet的编码-解码器结构^[17],如图1所示,首先编码器采用ResNet网络结构对图像进行特征提取,解码器部分的每一个上采样块都引入了残差连接。然后向该网络输入一对源图像I与目标图像J,图像I和J作为一个单一的多通道图像被连接和反馈到网络,最后经过ResUNet网络的一系列卷积和池化层处理后输出一个预测的变形场 ϕ 。此外,网络使用了一个类似于空间变形网络的参数共享的可微变形模块(Spatial Transformer Networks, STN)^[19-20],使用 ϕ 对源图像I进行变形,会产生一个扭曲的图像 $\phi \cdot I$ 。

在基于优化策略的CNN回归配准方法中,可变形配准方法需要进行详尽的迭代优化,同时需要参

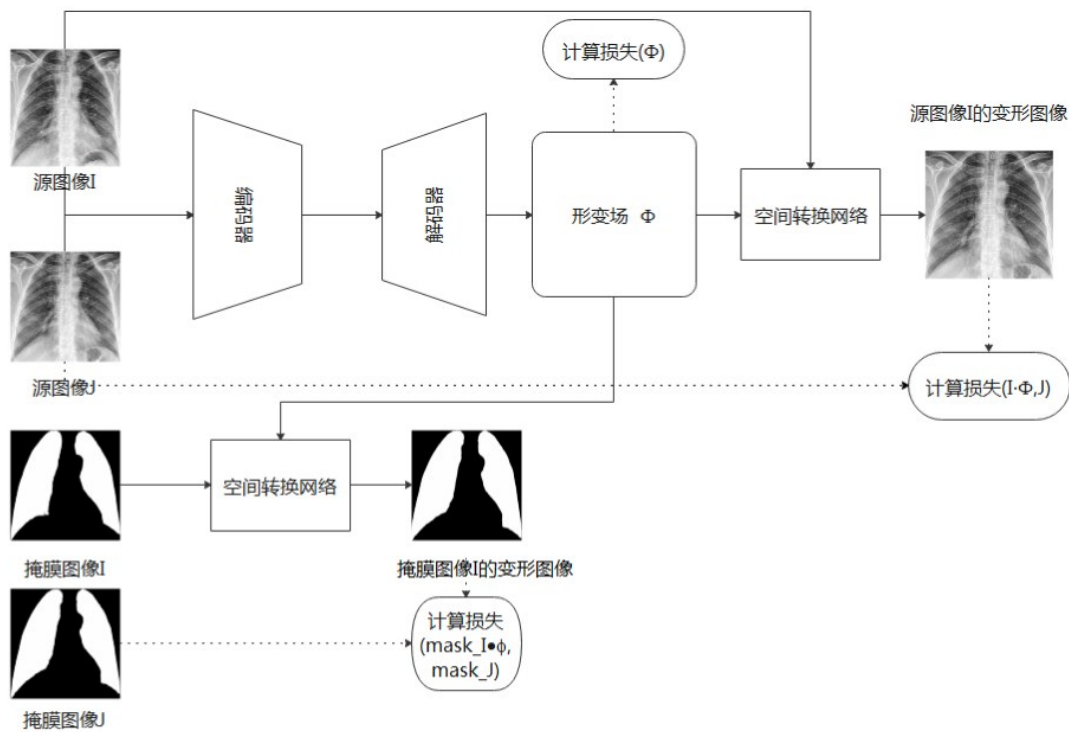


图1 网络结构
Figure 1 Network structure

数调整来估计图像间的变形场^[21]。因此,训练开始时,网络会产生一个随机变形场 ϕ ,经过训练不断调整参数,最后使源图像经过扭曲后产生的图像与目标图像相似度最大,本文采用归一化互相关性(Normalized Correlation Coefficient, NCC)作为损失函数来指导网络学习,同时还引入了一个简单的正则化项 $R(\phi)$,该正则化项通过计算变形场的总变化量得到,其作用在于防止平滑变形场过拟合。与此同时,用掩膜限制肺区,用Dice loss 计算目标掩膜与变形后的掩膜差异。本研究将损失函数定义为:

$$L(I, J, \phi) = M(I \cdot \phi, J) + \gamma R(\Phi) + \lambda L(SI \cdot \phi, SJ) \quad (4)$$

1.3 图像配准网络训练细节

本研究中配准模型共使用 500 个患者在不同时期拍摄的图像对,所有数据均由青海省各地方医院提供并进行脱敏处理,研究首先将图像缩放到统一尺寸 512×512,然后将数据集按照 8:1:1 的比例随机分为训练集,验证集和测试集。训练过程采用深度学习框架 Pytorch 以及优化器 Adam,学习率为 10^{-3} ,输入网络的批次为 16,公式(4)中的 γ 与 λ 分别设置为 10^{-2} 和 1。该模型在 NVIDIA GPU GTX 2080ti 16 GB 内存上训练和测试,共进行 200 次迭代。

1.4 评估指标

图像配准基于图像间相似性最大化原理来实现。配准后需要对其性能进行评估,不同类型的图像使用不同的评估标准,目前没有一个绝对的金标

准可以评估图像配准的质量。医学图像配准质量评估有两种经典方法:一种是单模图像配准常用的 NCC,一种是多模图像配准常用的互信息(Mutual Information, MI)^[7, 22-24]。本研究采用 NCC。

$$S_{NCC} = \frac{\sum I_F \cdot (I_M \cdot f) - \sum E(I_F) E(I_M \cdot f)}{|\chi| \cdot \sum \text{Var}(I_F) \text{Var}(I_M \cdot f)} \quad (5)$$

其中, S_{NCC} 表示归一化互相关性损失函数, E 表示期望值(均值), Var 表示相应图像的方差, χ 表示图像 N 个点的集合。本研究首先与两个传统配准模型进行对比,并引入了两个常用于评价分割精度的指标来评估模型的对齐效果,一个是平均表面距离(Average Symmetric Surface Distance, ASSD),一个是豪斯多夫距离(Hausdorff Distance, HD)。此外,本研究还与一个基于深度学习的胸片配准模型进行了比较^[25]。

2 结果与讨论

2.1 图像预处理与配准网络训练结果

研究共设置 4 组实验来训练网络模型,以寻找最佳的方式进行 X 射线胸片的配准,如图 2 所示,图 2a 表示原始图像训练,图 2b 表示原始图像截取肺部区域图像训练,图 2c 表示原始图像加肺区分割掩膜训练,图 2d 表示肺区图像加肺区分割掩膜训练。实验在保持网络其它参数不变的情况下进行对比训练。

训练过程中 MSE 损失值变化如图 3 所示,根据网络训练情况以及相似性测量标准,研究发现肺区

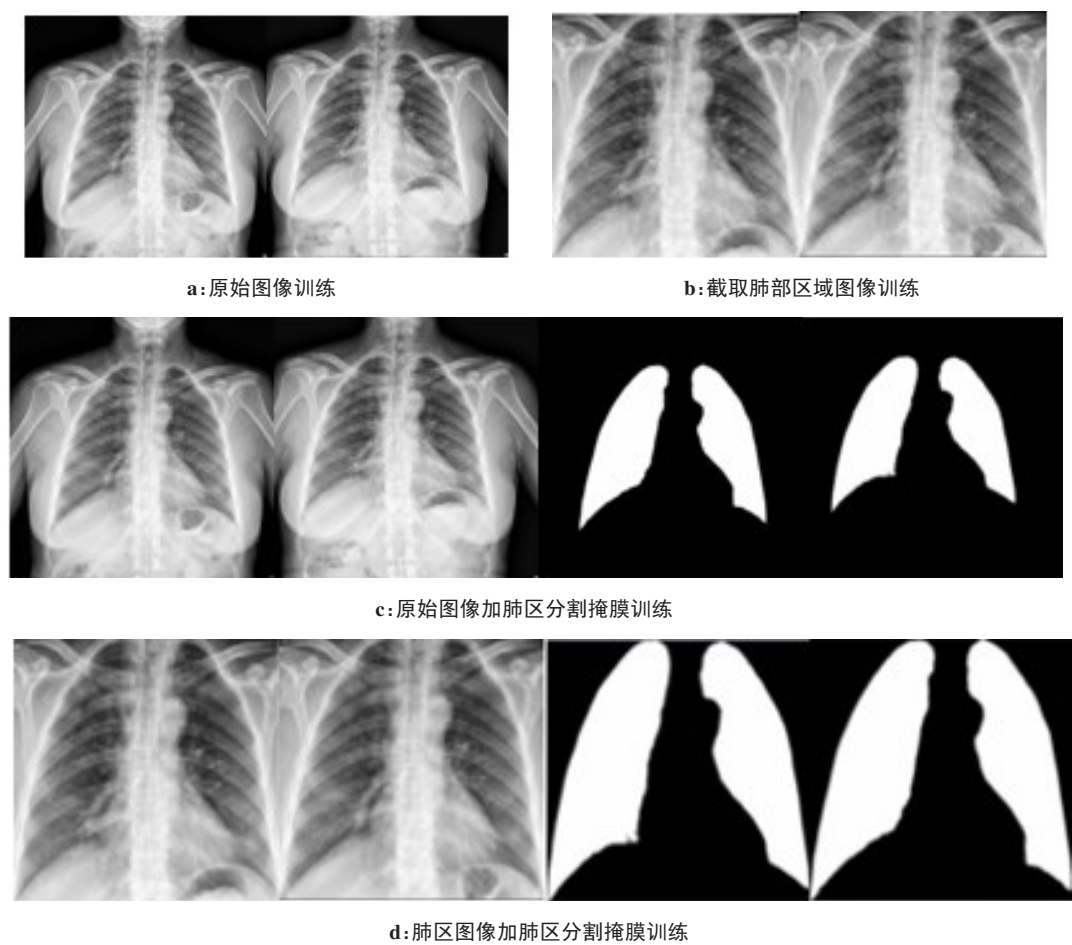


图2 不同网络模型训练

Figure 2 Training of different network models

图像加肺区分割掩膜训练的模型损失值更低,因此采用掩膜截取的肺区图像来训练网络会得到一个最佳的实验结果,这个结果也成为了本文模型的基准。

2.2 模型比较

为了评价本研究所提出模型的配准性能,研究一方面采用了两种传统的配准方法 Opencv ORB 和 SimpleElastix^[13]进行胸片配准以作为对照,另一方面还与一个基于深度学习的胸片配准模型进行了对比。两种传统模型之前均未用于胸片配准,因此研究将训练产生的模型与两个传统模型同时进行胸片配准。实验结果如表1所示,3个模型中,本研究提出的模型在配准后固定图像与浮动图像的相似性最大。Fang等^[25]采用基于深度学习的模型进行训练测试,根据局部互相关(Local Cross-Correlation, LCC)得到的相似度为0.350 3,本研究根据LCC得到的相似度为0.437 1。

2.3 配准应用于胸片的对比减影

将本文提出的配准模型用于胸片的对比减影分析,如图4所示,减影图像可以清晰地反映出两个胸片的不同之处,让医生可以直观分析病灶变化的趋势。本文

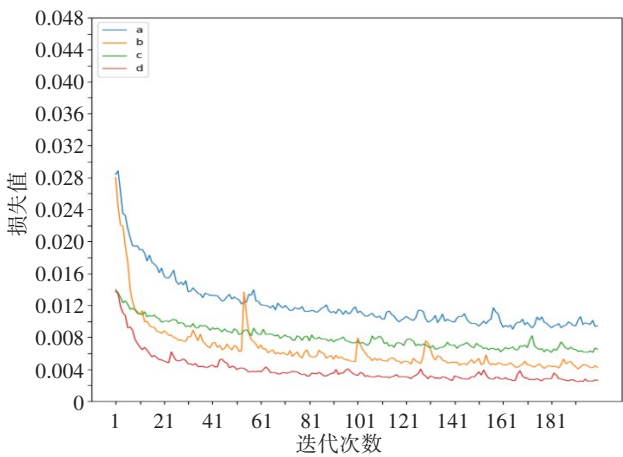


图3 训练过程中MSE损失值的变化

Figure 3 MSE loss value changes during training

a:原始图像训练;b:截取肺部区域图像训练;c:原始图像加肺区分割掩膜训练;d:肺部图像加肺区分割掩膜训练

通过获取两张图像的绝对差值对图像进行减影,然后对配准后的两张图像做病灶分割,通过病灶面积的占比量化两张图像的差异值,从而有利于图像配准在医学领域的运用。

表1 配准模型比较

Table 1 Comparison of registration methods

配准方法	NCC	HD	ASSD
Opencv ORB	0.886±0.174	25.713±9.563	10.22±5.035
SimpleElastix	0.936±0.083	17.432±8.441	4.128±2.051
本文模型	0.953±0.017	15.573±7.356	3.340±1.210

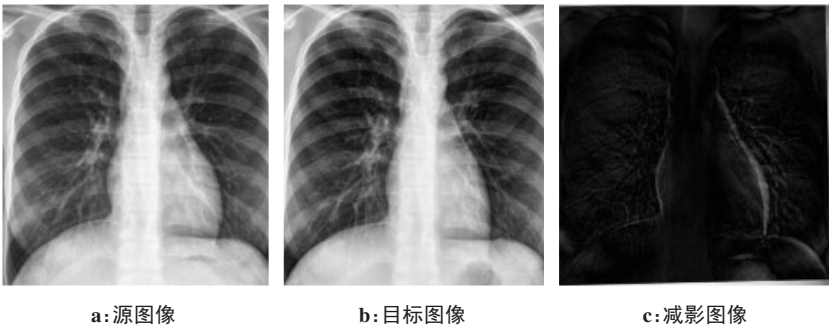


图4 配准在胸片对比减影分析中的应用

Figure 4 Application of registration in the subtraction analysis of chest radiographs

了可用于二维图像配准,也能探究其在三维图像上的应用效果,开发出基于掩膜的快速、精确的多模态下的图像配准。

【参考文献】

[1] Balakrishnan G, Zhao A, Sabuncu MR, et al. VoxelMorph: a learning framework for deformable medical image registration[J]. IEEE Trans Med Imaging, 2019, 38(8): 1788-1800.

[2] 刘飞, 张俊然, 杨豪. 基于深度学习的医学图像识别研究进展[J]. 中国生物医学工程学报, 2018, 37(1): 86-94.

[3] Liu F, Zhang JR, Yang H. Research progress of medical image recognition based on deep learning[J]. Chinese Journal of Biomedical Engineering, 2018, 37(1): 86-94.

[4] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proc of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015: 3431-3440.

[5] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks [C]//Proc of the Neural Information Processing Systems. New York, USA: Association for Computing Machinery, 2012: 1097-1105.

[6] Horn BK, Schunck BG. Determining optical flow[J]. Artif Intell, 1980, 17(1): 185-203.

[7] Glocker B, Komodakis N, Navab N, et al. Dense registration with deformation priors[C]//Proc of the Information Processing in Medical Imaging. Williamsburg, VA, USA: Springer-Verlag Berlin Heidelberg, 2009: 540-551.

[8] Sokooti H, De Vos B, Berendsen F, et al. Nonrigid image registration using multi-scale 3D convolutional neural networks[C]//Proc of the International Conference on Medical Image Computing and Computer Assisted Intervention. Cham: Springer International Publishing, 2017: 232-239.

[9] Castillo R, Castillo EM, Guerra R, et al. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets[J]. Phys Med Biol, 2009, 54(7): 1849-1870.

[10] Krebs J, Mansi T, Delingette H, et al. Robust non-rigid registration through agent-based action learning[C]//Proc of the Medical Image Computing and Computer Assisted Intervention. Cham: Springer International Publishing, 2017: 344-352.

[11] Woods RP, Grafton ST, Holmes CJ, et al. Automated image registration: I. general methods and intrasubject, intramodality validation[J]. J Comput Assist Tomogr, 1998, 22(1): 139-152.

[12] Ashburner J. A fast diffeomorphic image registration algorithm[J]. NeuroImage, 2007, 38(1): 95-113.

[13] Lucas BD, Kanade T. An iterative image registration technique with an application to stereo vision[C]//Proc of the 7th International Joint Conference on Artificial Intelligence. Vancouver, BC, Canada: Morgan Kaufmann Publishers, 1981: 674-679.

3 结 论

本研究采用深度学习掩膜对原始胸片进行预处理,使用掩膜图像训练 ResUNet 网络,然后将它与其他已报道的配准模型进行比较。研究结果显示深度学习掩膜结合深度学习图像配准方法开发出的模型在胸片配准上具有良好的配准精度。提出的模型除

[13] Marstal K, Berendsen FF, Staring M, et al. SimpleElastix: a user-friendly, multi-lingual library for medical image registration[C]//Proc of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2016: 574-582.

[14] Cachier P, Pennec X. 3D non-rigid registration by gradient descent on a gaussian-windowed similarity measure using convolutions[C]//Proc of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis. IEEE, 2011: 182-189.

[15] Avants BB, Tustison NJ, Song G, et al. A reproducible evaluation of ANTs similarity metric performance in brain image registration[J]. NeuroImage, 2011, 54(3): 2033-2044.

[16] Candemir S, Antani S. A review on lung boundary detection in chest X-rays[J]. Int J Comput Assist Radiol Surg, 2019, 14(4): 563-576.

[17] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [C]//Proc of the International Conference on Medical Image Computing and Computer Assisted Intervention. Cham: Springer International Publishing, 2015: 234-241.

[18] Oktay O, Ferrante E, Kamnitsas K, et al. Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation[J]. IEEE Trans Med Imaging, 2018, 37(2): 384-395.

[19] Jaderberg M, Simonyan K, Zisserman A, et al. Spatial transformer networks[C]// Proc of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2015: 2017-2025.

[20] Yoo I, Hildebrand DG, Tobin WF, et al. ssEMnet: serial-section electron microscopy image registration using a spatial transformer network with learned features[C]//Proc of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision. Cham: Springer International Publishing, 2017: 249-257.

[21] 邹茂扬, 杨昊, 潘光晖, 等. 深度学习在医学图像配准上的研究进展与挑战[J]. 生物医学工程学报, 2019, 36(4): 677-683.

[22] Zou MY, Yang H, Pan GH, et al. Research progress and challenges of deep learning in medical image registration[J]. Journal of Biomedical Engineering, 2019, 36(4): 677-683.

[23] Rohé MM, Datar M, Heimann T, et al. SVF-Net: learning deformable image registration using shape matching[C]//Proc of the Medical Image Computing and Computer Assisted Intervention. Cham: Springer International Publishing, 2017: 266-274.

[24] Yang X, Kwitt R, Styner M, et al. Quicksilver: fast predictive image registration-a deep learning approach[J]. Neuro Image, 2017, 158: 378-396.

[25] Wu GR, Kim M, Wang Q, et al. Scalable high-performance image registration framework by unsupervised deep feature representations learning[J]. IEEE Trans Biomed Eng, 2016, 63(7): 1505-1516.

[26] Fang Q, Yan J, Gu X, et al. Unsupervised learning-based deformable registration of temporal chest radiographs to detect interval change [C]//Proc of the Medical Imaging 2020: Image Processing. Houston, Texas, USA: SPIE Press, 2020.

(编辑:薛泽玲)