

基于深度学习的医学图像配准研究进展

周 勤, 王远军

(上海理工大学 医疗器械与食品学院, 上海 200093)

摘要: 图像配准是图像引导手术、图像融合、器官图谱生成、肿瘤和骨骼生长监测等临床任务应用的关键技术, 也是一个极具挑战性的问题。近年来, 深度学习技术对医学图像处理方法的研究产生重要的影响, 在医学图像配准领域发展迅速。对使用深度学习技术实现医学图像配准的研究进行综述, 首先按照深度学习模型将医学图像配准方法分为 3 类, 包括监督、弱监督和无监督医学图像配准; 然后分别介绍国内外研究进展, 并总结这些研究方法的优缺点; 在此基础上, 阐述常用的深度学习配准框架以及评价标准, 并总结常用的开源医学影像数据集; 最后对深度学习技术在医学配准图像领域中存在的问题进行分析, 展望未来发展的方向。

关键词: 医学图像配准; 深度学习; 变形配准; 卷积神经网络

中图分类号: O 482.53 **文献标志码:** A

Advances in medical image registration based on deep learning

ZHOU Qin, WANG Yuanjun

(School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Image registration is a key technology for image-guided surgery, image fusion, organ map generation, tumor and bone growth monitoring and other clinical applications. It is also a very challenging problem. In recent years, deep learning technology has exerted an important influence on the research of medical image processing methods, and has developed rapidly in the field of medical image registration. Research on medical image registration using deep learning technology was reviewed. Firstly, according to the deep learning model, medical image registration methods were divided into three categories, including supervised, weakly supervised and unsupervised medical image registration. Then the research progress at home and abroad was introduced, and the advantages and disadvantages of these research methods were summarized. On this basis, the commonly used deep learning registration framework and evaluation criteria were described, and the commonly used open source medical image data sets were summarized. Finally, the existing problems of deep learning technology in the field of medical registration image were analyzed, and the future development direction was forecasted.

收稿日期: 2021-02-02

基金项目: 国家自然科学基金资助项目 (61201067); 上海市自然科学基金资助项目 (18ZR1426900)

第一作者: 周 勤 (1996-), 女, 硕士研究生. 研究方向: 图像配准. E-mail: 1764354931@qq.com

通信作者: 王远军 (1980-), 男, 副教授. 研究方向: 医学图像处理与分析、计算机辅助诊断治疗. E-mail: yjusst@126.com

Keywords: *medical image registration; deep learning; deformable registration; convolutional neural network*

图像配准是基于图像相似性度量标准，寻求将一幅或多幅图像映射到目标图像上的空间变换，使得参考图像和目标图像达到空间上的对应关系。通常情况下，需要映射变换的图像称为浮动图像，在配准过程中保持不变的图像称为固定图像或者参考图像。当从不同时间、不同角度或使用不同传感器分析获得的图像时，需要进行配准^[1]。医学图像配准在临床上有很多应用，如图像引导、图像融合、肿瘤生长监测等。

在深度学习技术兴起之前，已经有很多传统的配准方法，但深度学习的出现极大地影响图像配准研究的热点^[2]。自从 AlexNet 在 2012 年 ImageNet 挑战赛^[3]中获得成功以来，深度学习使许多计算机视觉任务的最新性能得以实现，例如：目标检测、特征提取、分割、图像分类、图像去噪和图像重建。最初，深度学习技术主要用于增强迭代、基于强度的医学图像配准，但配准速度较慢，对快速配准算法的需求推动基于端到端深度学习的监督配准方法的发展。由于监督学习获取真实形变场比较困难，促使许多学者开发用于端到端的无监督配准框架。目前，无监督配准框架中的最大挑战是图像相似性量化，近来基于信息论的相似性度量、弱监督框架和生成对抗性网络 (generative adversarial networks, GAN) 框架等的研究都是用来解决图像相似性量化的策略^[4]。

本文根据使用的深度学习技术进行分类，总结常用的医学图像配准框架及评价指标、常用的数据集和开源工具，分析深度学习技术在医学图像领域中存在的问题，并进行展望。

1 配准方法的分类

通过调研分析，近几年的相关文献主要分为 2 种^[5]：a. 使用深度学习网络来估计两幅图像的相似性度量，以驱动迭代优化；b. 使用深度学习网络直接估计变换参数。前者深度学习技术主要用于增强迭代，配准速度慢且耗时长；后者利用深度学习的端到端网络进行配准，速度快。因此，对后者的调研分析比较多，根据所选用的深度学习框架，将配准方法分为监督学习、弱监督学习

和无监督学习 3 大类。

1.1 监督学习的配准方法

监督学习的配准方式是要事先得到金标准，即形变场。这个形变场可以是刚性配准的 3 个参数，也可以是变形配准的位移场。然后神经网络训练出一个模型，输入待配准的两张图像，输出这个形变场。在神经网络训练的过程中，目标函数反向传播，从而更新各层的参数。训练时的损失函数常常是金标准和模型输出的形变场的均方误差 (mean squared error, MSE)。在测试阶段，对配准图像对进行采样并输入网络，利用预测的形变场对移动图像进行插值，即得配准图像。人工生成形变场有 3 种方式：a. 利用传统的经典配准方法进行配准，得到的形变场作为标签；b. 对原始图像进行模拟变形，将原始图像作为固定图像，变形图像作为移动图像，模拟形变场作为标签；c. 基于模型变换，图像变换是图像配准的前提，根据变换性质可分为刚体变换和非刚体变换。下面从刚性配准和非刚性配准两方面介绍近几年出现的一些成果。

1.1.1 刚性配准

Miao 等^[6]是第一个使用卷积神经网络 (convolutional neural networks, CNN) 来估计刚性变换参数的团队。他们用 CNN 估计 3D CT 与 2D X 线脊柱图像刚性配准的变换矩阵，提出将 6 个变换参数分为 3 组的层次回归方法。实验结果表明，该方法的性能优于互相关、互信息和基于梯度相关的迭代配准方法。Chee 等^[7]使用 CNN 估计用于 3D 脑磁共振 (magnetic resonance, MR) 图像的变换参数，该方法用仿射配准网络估计值和标准仿射变换之间的 MSE 来训练网络。在单模和多模情况下，它们都能优于基于迭代 MI 的配准。与前面两种使用随机生成的形变场不同的是，Zheng 等^[8]提出将对区域适应模块集成到预先训练的 CNN 中，该自适应模块可以使用少量成对的真实数据和合成数据进行训练，将训练后的模块接入网络，将真实的特征转移到网络中，使其接近于合成的特征。由于网络是在合成数据上训练的，所以网络在合成数据上表现良好。Sloan 等^[9]用 CNN 回归刚性变换参数，对 T1 和 T2 加权脑 MR 图

像进行配准, 研究了单模和多模配准。

1.1.2 非刚性配准

Yang 等^[10] 首先提出直接基于图像外观的变形模型逐块预测来实现脑MR 图像的快速配准。该方法用深度编解码网络将图像像素的初始动量值作为网络输入, 对这些值进行演化来获得估计的形变场。实验结果表明, 该方法能够对基于数值优化的配准结果进行准确且快速的预测。此外, Yang 等^[11] 利用变换参数的变分高斯分布的低秩 Hessian 近似量化 3D T1 和 T2 加权脑 MR 图像变形配准相关的不确定性。该方法在真实数据和合成数据上都得到了验证。Sokooti 等^[12] 提出用多尺度 CNN 来估计形变场的方法, 实验表明, 该方法性能优于基于 B 样条的配准。与使用随机变化或手工构建来生成形变场的方法不同, Ito 等^[13] 提出使用 CNN 学习形变场生成的合理变形, 通过在 ADNI 数据集上进行实验得出, 结果优于基于迭代 MI 的方法。

1.2 弱监督学习的配准方法

监督配准方法需要真实的形变场作为金标准, 而真实形变场的获取比较困难。弱监督方法的提出, 减轻了对金标准的依赖性。弱监督利用相应解剖结构的分段重叠来设计损失函数, 即采用标签的相似性度量作为目标函数, 反向迭代更新网络的权重。弱监督学习的配准成果, 相对前两种方法成果比较少, 多使用基于相似性测度的配准和非刚性配准。Hu 等^[14] 没有使用人工生成的形变场, 而是使用标签相似性来训练网络, 进行 3D MRI 和超声图像配准。同样地, Hering 等^[15] 结合分割标签和图像相似性的互补信息训练网络。实验结果表明, 该方法 Dice 分数高于仅使用图像相似性损失和分割标签损失。不同于以上方法, Liu 等^[16] 使用表示学习来学习具有置信水平概率图的特征描述符, 并使用监督合成变换和无监督描述符结合的图像相似性损失来训练网络。

无约束的形变场估计会导致浮动图像的扭曲, 从而产生不真实的器官外观, 一种常用的方法是将估计形变场的 L2 范数加入损失函数。另一个有价值的贡献是将 GANs 用于弱监督估计变换。Lei 等^[17] 使用鉴别器来判断扭曲的图像对于原始图像是否足够逼真。Fu 等^[18] 使用了类似的想法, 并表明, 在 3D-CT 肺 DIR 中包括对抗损失可以提高配准精度。上述基于 GAN 的方法试图从图

像或变换外观的角度引入正则化。不同的是, Hu 等^[19] 使用 GAN 框架来产生逼真的变形。Hu 团队主要想法是通过引入判别器对局部形变场进行约束, 代替形变场的正则项。总体框架是通过生成器生成仿射变换的形变场和局部形变场, 通过传统的方法生成一个真实的局部形变, 然后判别器判别预测的局部形变场和仿真生成的形变场。Fan 等^[20] 研究使用 GAN 网络区分对齐和未对齐的图像对。但是, GANs 本身存在着难以训练的缺点, 因为训练 GANs 需要达到纳什均衡, 目前还没有很好达到纳什均衡的技术。此外, GANs 在配准上的应用方法不一样, 各个方法也有自己的缺点, 如训练不稳定、梯度消失和模式崩溃等问题。

1.3 无监督学习的配准方法

尽管监督和弱监督学习的图像配准取得很大成果, 但标签的获取很难, 没有标签, 又很难定义适当的损失函数, 这促使许多学者开始探索无监督学习的方法。2015 年, Jaderberg 等^[21] 提出一种空间变压器网络 (spatial transformer networks, STN), 它可以对网络内的数据进行空间操作。因为 STN 可以在训练过程中计算图像的相似性损失, 使无监督学习的图像配准技术成为可能。一种典型的变形无监督变换估计网络具体做法是以图像对作为输入, 直接输出预测的随机位移矢量场 (displacement vector field, DVF)。STN 利用该网络对运动图像进行扭曲, 生成扭曲图像; 然后计算扭曲图像与固定图像的相似性损失, DVF 平滑度约束通常用于对预测的 DVF 进行正则化。以下分类介绍这方面出现的文献。通过调研发现, 无监督学习的配准方法多用于非刚性配准。基于图像相似性的无监督图像配准方法不需要标签, 然而在多模配准应用中, 图像相似度的量化仍然是一个挑战。因此, 基于图像相似性的无监督学习的配准多用于单模配准。鉴于许多临床应用中经常需要多模配准, 研究者们开始研究无监督学习来学习特征表示以确定最佳空间对齐方式^[4], 并取得一些成果。下面分类介绍这两方面的文献。

1.3.1 基于灰度相似性度量的配准

配准的关键是选择合适的相似性度量^[22], 常见的做法是估计给定图像对之间的相似性度量, 将其作为代价函数进行反向更新迭代。Balakrishnan 等^[23] 提出一种基于 CNN 的无监督配准方法, 他们使用了一种类似于 U-Net 的架构, 并将其命名为

VoxelMorph。后来，他们扩展了该方法来利用训练数据中的辅助分段。结果表明，分割损失有助于提高 *Dice* 分数，性能可与 ANT 和 NifTYG 相比，但比 ANT 快 150 倍，比 NifTYG 快 40 倍。类似地，Qin 等^[24] 也使用心脏 MRI 分割信息作为心脏 MRI 配准的补充信息。他们发现通过配准 CNN 学习的特征也可以用于分割。预测的 DVF 用于运动图像的掩模变形以生成固定图像的掩模。他们训练了一个联合分割和配准模型，并证明了在分割和配准任务中，联合模式可以比单独的模型产生更好的结果。Zhang 等^[25] 提出了一种基于 CNN 端到端的配准模型。他们研究了变换的微分同胚映射，并提出使用逆一致正则化项来约束来自相应逆映射的两个变换之间的差异。损失函数由图像相似性损失、变换平滑性损失、逆一致损失和反折叠损失组成。在 *Dice*、灵敏度和平均表面距离方面，他们的方法优于 Demons 和 Syn。不同于上述 2 种方法，Jiang 等^[26] 提出了一种基于 CNN 的多尺度配准框架。他们级联 3 个 CNN 模型，使用图像块来训练网络，以优化图像相似性损失和 DVF 平滑度损失。他们的工作表明，在标准人工神经网络数据集上训练的网络可以推广到不同的人工神经网络数据集。

1.3.2 基于特征的配准

基于特征的配准利用 CNN 学习图像特征表示，并用学习到的特征表示来训练网络以确定最优变换。Yoo 等^[27] 使用卷积自动编码器提取单模结构特征，并将其输入到 STN 中进行最终的可变形配准，他们发现这种基于特征的相似性度量优于基于灰度的相似性度量^[28]。Liu 等^[29] 提出了一种基于张量的方法，该方法使用主成分分析网络^[30] 进行单模和多模配准，在吸气-呼气两对胸部 CT 容积和多模脑 MR 图像的实验中得到验证。与前面两种方法不同，Kori 等^[31] 使用预先训练的网络，对二维 T1 和 T2 加权脑 MR 图像进行特征提取和仿射变换参数回归，对图像进行二值化，然后把运动图像和固定图像之间的 *Dice* 得分作为代价函数。

1.4 存在的问题

在监督学习配准中，真实的形变场在网络训练中发挥重要作用，但使用人工生成的形变场也存在一些问题，包括：生成的变换可能无法反映真实的生理运动；生成的变换可能无法捕捉实际图像配准场景的大范围变化。为解决这些问题，

常使用各种转换生成模型^[13] 和进行适当的数据扩充^[12]。弱监督配准技术可以对多模态配准进行相似性量化，而且弱监督变换估计方法减轻对手工标注的依赖性，但仍需要手动标注数据（如金标准或分段）。同监督变换估计相比，无监督的方法有效地缓解了数据集不足的问题。然而在无监督变换估计的情况下对多模图像配准比单模图像配准要困难得多，因为使用手工标注的相似性度量来量化两个图像之间的相似性是很困难的。近年来，利用无监督学习来学习特征表示以确定最佳转换，引起了研究界的极大兴趣。

1.5 常用基于深度学习的配准框架及评价指标

深度学习是用多层的神经网络来学习数据的表示^[27]，属于机器学习中的一类。近年来，深度学习被广泛用于图像处理。通过分析近几年深度学习对图像配准上的应用，总结 4 种常用的基于深度学习的配准框架，具体如下：

a. CNN 是一种前馈神经网络，具有局部连接、权重共享以及空间或时间上的次采样特性。这些特性使得卷积神经网络具有一定程度上的平移、缩放和扭曲不变性^[32]。在网络的每一层中，许多卷积滤波器在前一层的特征映射上“滑动”，输出是另一组特征映射。由这些卷积得到的特征映射被叠加并输入到网络的下一层，并允许对图像进行分层特征提取。此外，这些操作可以逐块执行，对许多图像配准任务有用^[33]。

b. RNN 用于处理和预测序列，适合处理时间信号。利用 RNN 的内部状态来模拟和记忆先前处理过的信息，因此 RNN 的输出即依赖它的即时输入，又依赖它以前的输入历史。长短期记忆 (LSTM) 是一种用于图像处理任务的 RNN。

c. 堆叠自动编码器 (stacked auto-encoders, SAE) 是网络隐藏层学习输入数据的隐含特征并重构为输出。在医学图像配准的文献中，该网络仅用于提供图像中最重要的、最有鉴别能力的特征，以作为替代配准方法的补充，而不是使用手工制作的特征^[10]。

d. GANs 由生成器和判别器两个相互竞争的神经网络组成。生成器将数据从一个域映射到另一个域。判别器的任务是区分来自所述域的真实数据和由生成器生成的数据。训练 GANs 的目标是收敛到一个可微的纳什均衡^[34]，此时生成的数据和实际数据是不可区分的。在医学图像配准中，

GANs 通常用于正则化或将多模态配准转换为单模态配准。方法是用判别器来代替传统的相似性度量, 多用于多模态配准。需要指出的是, 对于单模态而言, 还是手工设计的相似性测度更好, 深度学习学到的只是一种补充。

一直以来图像配准的评价指标没有金标准, 由于不同类型图像差别很大, 而且研究的重点也各有不同, 所以配准效果很难用一个评价指标界定, 通常需要多个评价指标来评定。通过调研发现, *Dice* 和靶点定位误差 (target registration error, *TRE*) 是最常用的评价指标, 公式见表 1。

表 1 常用的评价指标
Tab.1 Common evaluation indicators

评价指标	公式	备注
<i>Dice</i>	$Dice = \frac{2 A \cap B }{ A + B }$	<i>A</i> 和 <i>B</i> 分别为金标准和配准后的金标准, $ A \cap B $ 表示 <i>A</i> 和 <i>B</i> 的交集, $ A $ 和 $ B $ 分别表示 <i>A</i> 和 <i>B</i> 的元素个数。0表示没有重叠, 1表示完全重叠
<i>TRE</i>	$TRE = \frac{1}{n} \sum_{i=1}^n l_i^A - l_i^B $	<i>A</i> 和 <i>B</i> 为输入图像, l_i^A 和 l_i^B 是 <i>A</i> 和 <i>B</i> 中第 <i>i</i> 个相应地标, <i>n</i> 为标记总数

1.6 常用的开源数据和开源工具

医学图像配准任务是复杂的, 配准的部位不同选用的数据集也不同。通过查阅文献, 发现近年来, 基于深度学习的医学图像配准多用来实现脑部和肺部的配准, 对用于图像配准和评估研究的常用脑部和肺部数据集进行归纳总结, 结果见表 2。其中, DIRLAB 是 4D CT 胸部图像配准研究中引用最多的公共数据集之一, 该数据集经常用于 4D CT 肺部配准的基准测试。与肺部图像配准相比, 脑图像配准在数据库中有更广泛的选择, LPBA40 和 IBSR18 是较为常用的两个。

表 2 常用的数据集
Tab.2 Common data sets

数据集	数量	模态	区域	格式
LPBA40	40	3D MRI	T1 Brain	.img+.hdr+.nii
IBSR18	18	3D MRI	T1 Brain	.img+.hdr
CUMCU12	12	3D MRI	Brain	.img+.hdr
MGH10	10	3D MRI	Brain	.img+.hdr
DIRLAB	10	4D CT	Lung	.img
EMPIRE	30	4D CT	Lung	.mhd+.raw

随着计算机技术的发展, 已经有很多开源工具可用于构建图像配准的网络, 如: ITK、ANTs、

Elastix、SimpleElastix 和 3D slicer 等。近年来, 已有不少学者将自己基于深度学习的医学图像配准研究成果公开分享到 GitHub 上, 这些工作加快了医学图像配准的发展。目前 GitHub 已经有许多公共的存储库可用于构建本文中描述的网络, 如 Voxelmorph^[23], RegNet3D^[12], Weakly-supervised CNN^[14] 等。

2 挑战与对策

通过对医学图像配准中应用深度学习方法的研究, 研究者们碰到了一些常见的挑战, 其中一些常见的挑战和解决方案如下所述。

2.1 缺少已知转化的训练数据集

将监督学习技术应用到图像配准之中, 经常会因为缺少已知转化的训练数据集而导致图像配准效果不好。在深度学习中, 常用以下 4 种方法解决这类问题:

a. 数据扩充技术。简单的数据扩充技术即图像的几何变换可直接应用于图像领域, 如图像的旋转、翻转、尺度变换、随机抠取、噪声扰动和颜色变换。根据近些年学者在样本增强技术方面的探索, Chen 等^[35]总结了 6 种复杂的样本增强技术, 具体如下: (a)在弱监督环境下, 利用大量未标记数据的信息学习标记样本^[36]; (b)从已有的训练模型中自适应地学习样本^[37]; (c)借用相关类别的实例或语义词汇扩充样本的量^[38]; (d)通过扭曲现有的训练实例来合成新的标记训练数据^[39]; (e)使用 GANs 在原有数据集的基础上生成新的数据^[40]; (f)根据期望值或强度, 利用属性引导增强 (attribute-guided augmentation) 技术合成样本^[41]。

b. 迁移学习。在少量数据集上训练较复杂的网络容易过拟合, 可通过迁移学习的策略解决, 通常做法^[42]为: 对应样本量较少, 可以先用 Imagenet 等数据集预训练 CNN, 将预训练的 CNN 参数作为网络的初始值, 然后用自己的数据集对 CNN 中可学习的参数根据任务调优。在数据量较少的情况下, 可以从框架的开源数据库中下载相关的预训练模型。Cao 等^[43]使用 CNN 直接回归单模可变形的配准参数, 他们团队通过迁移学习的方法增加样本量, 该项研究表明, 迁移学习能减少因数据缺少而带来的影响。

c. 正则化。数据量较小会导致模型过拟合, 在损失函数后面加上正则项, 可以抑制过拟合问题

的产生。Srivastava等^[44]提出 dropout 技术,通过在训练过程中随机地从每一层中剔除一些输入,以减少过度拟合的影响。

d. 无监督预训练。用 Auto-Encoder 和受限玻尔兹曼机(restricted Boltzmann machine)的卷积形式实现无监督预训练,最后加上分类层作监督的微调。在这种情况下,预训练是一种正则化机制,有助于在泛化性能方面找到更好的训练误差。无监督的预训练在后来的监督训练中起着主要的规范化作用。Wang等^[37]在无监督预训练的工作表明,预先训练的网络是独一无二的,因为无监督的部分纯粹用于初始化,而且当进行监督的训练时,它的有益效果似乎不会减弱。

2.2 多模医学图像配准

将深度学习技术用于多模图像配准大致可分为两类:基于图像强度相似性度量的方法和基于特征的方法。前者利用增强学习对图像的灰度空间进行多模态图像配准,最大化像素强度之间的相似度来确定两幅图像之间的对齐程度;后者提取图像的显著不变特征,并利用这些特征之间的相关性来确定最佳对齐方式。基于深度学习的医学图像配准主要困难在于图像在不同的物理原理和环境具有很大的差异,缺乏鲁棒的相似性测度且特征提取较难。针对图像相似性的量化和特征提取问题,已有许多学者在这方面作了探索,如 Hu等^[14]提出了使用标签相似性来训练他们的网络,进行多模态医学影像配准,主要是用标签的相似性测度来代替图像之间的相似性测度,因为多模态图像之间的相似性测度不好设计,但是标签之间的相似性测度就好设计一些;Wu等^[45]将 CNN 与独立子空间分析相结合,利用学习到的图像特征来代替 HAMMER 配准模型中的手工特征;Kori等^[31]使用 CNN 网络对图像进行特征提取来实现多模图像配准。Liu等^[46]使用 10 层的 CNN 进行图像合成,不使用 GAN 将多模配准转换为单模配准。然后,他们使用基于 SSD 相似性度量来配准大脑模型和 IXI 数据集。实验结果表明,他们的算法性能优于基于 MI 的图像配准。但是,在多模配准应用中,图像相似度的测度和特征提取仍然是一个挑战。

3 结 论

总结了一种根据深度学习方法,把医学图像

配准方法分 3 类:监督医学图像配准、弱监督医学图像配准、无监督医学图像配准。每一类方法都有优缺点。因为监督预测变换和无监督预测变换各有利弊,未来更多的研究将集中在监督和无监督相结合的基础上。除此之外,弱监督和无监督的方法对人工标注的样本依赖性较小,这方面的研究具有现实意义。由于 GANs 不仅可以引入额外的正则化,还可以进行图像域转换,实现多模态到单模态的图像配准,因此基于 GANs 的方法逐渐成为学者研究的热点。

参考文献:

- [1] HILL D L G, BATCHELOR P G, HOLDEN M, et al. Medical image registration[J]. *Physics in Medicine & Biology*, 2001, 46(3): R1-R45.
- [2] AMBINDER E P. A history of the shift toward full computerization of medicine[J]. *Journal of Oncology Practice*, 2005, 1(2): 54-56.
- [3] ALOM M Z, TAHA T M, YAKOPCIC C, et al. The history began from alexnet: a comprehensive survey on deep learning approaches[EB/OL]. (2018-08-12)[2021-02-02]. <https://arxiv.org/abs/1803.01164v1>.
- [4] HASKINS G, KRUGER U, YAN P K. Deep learning in medical image registration: a survey[J]. *Machine Vision and Applications*, 2020, 31(1/2): 8.
- [5] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. *Medical Image Analysis*, 2017, 42: 60-88.
- [6] MIAO S, WANG Z J, LIAO R. A CNN regression approach for real-time 2D/3D registration[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1352-1363.
- [7] CHEE E, WU Z. Airnet: self-supervised affine registration for 3d medical images using neural networks[EB/OL]. (2018-10-12)[2021-02-02]<https://arxiv.org/abs/1810.02583>.
- [8] ZHENG J N, MIAO S, WANG Z J, et al. Pairwise domain adaptation module for CNN-based 2-D/3-D registration[J]. *Journal of Medical Imaging*, 2018, 5(2): 021204.
- [9] SLOAN J M, GOATMAN K A, SIEBERT J P. Learning rigid image registration-utilizing convolutional neural networks for medical image registration[C]//Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies. Funchal, Madeira, Portugal: SCITEPRESS-Science and Technology Publications, 2018.
- [10] YANG X, KWITT R, NIETHAMMER M. Fast predictive image registration[M]//CARNEIRO G, MATEUS D, PETER L, et al. *Deep Learning and Data Labeling for Medical Applications*. Cham: Springer, 2016.

- [11] YANG X. Uncertainty quantification, image synthesis and deformation prediction for image registration[D]. North Carolina: The University of North Carolina at Chapel Hill, 2017.
- [12] SOKOOTI H, DE VOS B, BERENDSEN F, et al. Nonrigid image registration using multi-scale 3D convolutional neural networks[C]//20th International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec City: Springer, 2017.
- [13] ITO M, INO F. An automated method for generating training sets for deep learning based image registration[C]//Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies. Funchal, Madeira, Portugal: SciTePress, 2018.
- [14] HU Y P, MODAT M, GIBSON E, et al. Label-driven weakly-supervised learning for multimodal deformable image registration[C]//2018 IEEE 15th International Symposium on Biomedical Imaging. Washington DC, USA: IEEE, 2018.
- [15] HERING A, KUCKERTZ S, HELDMANN S, et al. Enhancing label-driven deep deformable image registration with local distance metrics for state-of-the-art cardiac motion tracking[M]//HANDELS H, DESERNO T M, MAIER A, et al. Bildverarbeitung für die Medizin 2019. Wiesbaden: Springer, 2019, doi: [10.1007/978-3-658-25326-4_69](https://doi.org/10.1007/978-3-658-25326-4_69).
- [16] LIU C, MA L H, LU Z M, et al. Multimodal medical image registration via common representations learning and differentiable geometric constraints[J]. *Electronics Letters*, 2019, 55(6): 316–318.
- [17] LEI Y, FU Y, HARMS J, et al. 4D-CT deformable image registration using an unsupervised deep convolutional neural network[M]//NGUYEN D, XING L, JIANG S. Artificial Intelligence in Radiation Therapy. Cham: Springer, 2019.
- [18] FU Y B, LEI Y, WANG T H, et al. LungRegNet: an unsupervised deformable image registration method for 4D-CT lung[J]. *Medical Physics*, 2020, 47(4): 1763–1774.
- [19] HU Y P, GIBSON E, GHAVAMI N, et al. Adversarial deformation regularization for training image registration neural networks[M]//FRANGI A F, SCHNABEL J A, DAVATZIKOS C, et al. Medical Image Computing and Computer Assisted Intervention. Cham: Springer, 2018.
- [20] FAN J F, GAO X H, XUE Z, et al. Adversarial similarity network for evaluating image alignment in deep learning based registration[C]//21st International Conference on Medical Image Computing and Computer-Assisted Intervention. Granada: Springer, 2018.
- [21] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Quebec, Canada: NIPS, 2015: 2017–2025.
- [22] BOVEIRI H R, KHAYAMI R, JAVIDAN R, et al. Medical image registration using deep neural networks: a comprehensive review[J]. *Computers & Electrical Engineering*, 2020, 87: 106767.
- [23] BALAKRISHNAN G, ZHAO A, SABUNCU M R, et al. Voxelmorph: a learning framework for deformable medical image registration[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(8): 1788–1800.
- [24] QIN C, SHI B B, LIAO R, et al. Unsupervised deformable registration for multi-modal images via disentangled representations[M]//CHUNG A C S, GEE J C, YUSHKEVICH P A, et al. Information Processing in Medical Imaging. Cham: Springer, 2019.
- [25] ZHANG J. Inverse-consistent deep networks for unsupervised deformable image registration[EB/OL]. (2018-09-10)[2021-02-02]. <https://arxiv.org/abs/1809.03443v1>.
- [26] JIANG Z R, YIN F F, GE Y, et al. A multi-scale framework with unsupervised joint training of convolutional neural networks for pulmonary deformable image registration[J]. *Physics in Medicine & Biology*, 2020, 65(1): 015011.
- [27] YOO I, HILDEBRAND D G C, TOBIN W F, et al. ssEMnet: serial-section electron microscopy image registration using a spatial transformer network with learned features[M]//CARDOSO M J, ARBEL T, CARNEIRO G, et al. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 2017, doi: [10.1007/978-3-319-67558-9_29](https://doi.org/10.1007/978-3-319-67558-9_29).
- [28] SAALFELD S, FETTER R, CARDONA A, et al. Elastic volume reconstruction from series of ultra-thin microscopy sections[J]. *Nature Methods*, 2012, 9(7): 717–720.
- [29] LIU Q G, LEUNG H. Tensor-based descriptor for image registration via unsupervised network[C]//2017 20th International Conference on Information Fusion (Fusion). Xi'an, China: IEEE, 2017.
- [30] CHAN T H, JIA K, GAO S H, et al. PCANet: a simple deep learning baseline for image classification[J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 5017–5032.
- [31] KORI A, KRISHNAMURTHI G. Zero shot learning for multi-modal real time image registration[EB/OL]. (2019-08-17)[2021-02-02]. <https://arxiv.org/abs/1908.06213>.
- [32] HAO X, ZHANG G G, MA S. Deep learning[J].

- [International Journal of Semantic Computing](#), 2016, 10(3): 417–439.
- [33] SOKOOTI H, SAYGILI G, GLOCKER B, et al. Accuracy Estimation for Medical Image Registration Using Regression Forests[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2016.
- [34] LV J, YANG M, ZHANG J, et al. Respiratory motion correction for free-breathing 3D abdominal MRI using CNN-based image registration: a feasibility study[J]. *The British Journal of Radiology*, 2018, 91(1083): 20170788.
- [35] CHEN Z T, FU Y W, ZHANG Y D, et al. Semantic feature augmentation in few-shot learning[EB/OL]. (2018-4-15) [2021-02-02]. <https://arxiv.org/abs/1804.05298v1>.
- [36] FU Y W, HOSPEDALES T M, XIANG T, et al. Transductive multi-view zero-shot learning[J]. [IEEE Transactions on Pattern Analysis and Machine Intelligence](#), 2015, 37(11): 2332–2345.
- [37] WANG Y X, HEBERT M H. Learning from small sample sets by combining unsupervised meta-training with CNNs[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2016.
- [38] FU Y W, SIGAL L. Semi-supervised vocabulary-informed learning[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016.
- [39] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2012.
- [40] MAO X D, LI Q, XIE H R, et al. Least squares generative adversarial networks[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017.
- [41] KWITT R, HEGENBART S, NIETHAMMER M. One-shot learning of scene locations via feature trajectory transfer[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 78–86.
- [42] ZOU M Y, ZHONG Y. Transfer learning for classification of optical satellite image[J]. [Sensing and Imaging](#), 2018, 19(1): 6.
- [43] CAO X H, YANG J H, ZHANG J, et al. Deformable image registration based on similarity-steered CNN regression[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec City, QC, Canada: Springer, 2017.
- [44] SRIVASTAVA N. Improving neural networks with dropout[D]. Toronto: University of Toronto, 2013.
- [45] WU G R, KIM M, WANG Q, et al. Unsupervised deep feature learning for deformable registration of MR brain images[C]//International Conference on Medical Image Computing and Computer-assisted Intervention. Nagoya: Springer, 2013.
- [46] LIU X L, JIANG D S, WANG M N, et al. Image synthesis-based multi-modal image registration framework by using deep fully convolutional networks[J]. *Medical & Biological Engineering & Computing*, 2019, 57(5): 1037–1048.

(编辑: 丁红艺)