



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
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

2020
04
VOL.25

ISSN1006-8961
CN11-3758/TB



海岛岸线自动分割 P768



第25卷第4期 (总第288期)

2020年4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 中国科学院遥感与数字地球研究所
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 顾行发
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@radi.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 Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Gu Xingfa
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@radi.ac.cn
Telephone 010-58887035
Website www.cjig.cn

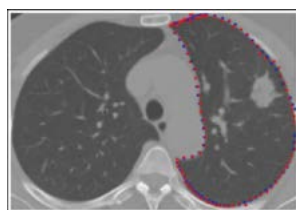
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

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



自适应在线判别外观学习的
分层关联多目标跟踪
(第0708页)



低秩和主动形状模型的肺区
分割(第0759页)



结合曲率滤波的HTM算法去
除遥感影像云雾(第0791页)

综述

深度学习目标检测方法综述

赵永强, 饶元, 董世鹏, 张君毅 0629

图像处理和编码

结合矩阵分解与差分隐私的人脸图像发布

张啸剑, 付聪聪, 孟小峰 0655

图像分析和识别

多尺度深度特征融合的变化检测

樊玮, 周末, 黄睿 0669

面向类内差距表情的深度学习识别

陈亮, 吴攀, 刘韵婷 0679

带后缀印刷体维吾尔文映射关系检索

伊克萨尼·普尔凯提, 阿布都萨拉木·达吾提, 艾斯卡尔·艾木都拉 0688

结合混沌鸟群算法的阴极铜板表面缺陷检测

王卓, 张长胜, 李伟, 钱俊兵, 唐都作, 蔡兵, 常以涛 0697

自适应在线判别外观学习的分层关联多目标跟踪

方岚, 于凤芹 0708

图像理解和计算机视觉

多方向Laplacian能量和与tetrolet变换的图像融合

沈瑜, 陈小鹏, 杨倩 0721

双重对偶生成对抗网络的跨年龄素描—照片转换

吴柳玮, 孙锐, 阚俊松, 高隽 0732

医学图像处理

残差密集相对平均CGAN的脑部图像配准

王丽芳, 张程程, 秦品乐, 蔺素珍, 高媛, 窦杰亮 0745

低秩和主动形状模型的肺区分割

孙申申, 田丹, 吴微, 康雁, 赵宏 0759

遥感图像处理

改进Deeplab网络的遥感影像海岸线分割

王振华, 钟元芾, 何婉雯, 曲念毅, 徐利智, 张文苹, 刘智翔 0768

卷积神经网络提取风云影像土壤湿度

韩颖娟, 周顺武, 王文文, 张学艺, 张承明, 栗佩琪 0779

结合曲率滤波的HTM算法去除遥感影像云雾

杨珂珂, 贾渊, 沈川 0791

ChinaMM 2019会议专栏

利用稳健非负矩阵分解实现无监督高光谱解混

宋晓瑞, 吴玲达, 孟祥利 0801

关键点深度特征驱动人脸表情识别

王善敏, 帅惠, 刘青山 0813

融合区域和边缘特征的水平集水下图像分割

孙杨, 陈哲, 王慧斌, 张振, 沈洁 0824

深度时空能量特征表示下的人体行为识别

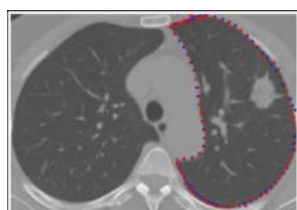
巢新, 侯振杰, 李兴, 梁久桢, 宦娟, 刘浩昱 0836

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Multi-object tracking based on adaptive online discriminative appearance learning and hierarchical association(P0708)



Lung segmentation by active shape model approach based on low rank theory(P0759)



Haze and cloud removal from remote sensing image using HTM algorithm based on curvature filtering(P0791)

Review

Survey on deep learning object detection

Zhao Yongqiang, Rao Yuan, Dong Shipeng, Zhang Junyi 0629

Image Processing and Coding

Private facial image publication through matrix decomposition

Zhang Xiaojian, Fu Congcong, Meng Xiaofeng 0655

Image Analysis and Recognition

Multiscale deep features fusion for change detection

Fan Wei, Zhou Mo, Huang Rui 0669

Depth learning recognition method for intra-class gap expression

Chen Liang, Wu Pan, Liu Yunting 0679

Mapping relationship retrieval of Uyghur-printed suffix word

Eksan Firkat, Abdusalam Dawut, Askar Hamdulla 0688

Cathodic copper plate surface defect detection based on bird swarm algorithm with chaotic theory

Wang Zhuo, Zhang Changsheng, Li Wei, Qian Junbing, Tang Duzuo, Cai Bing, Chang Yitao 0697

Multi-object tracking based on adaptive online discriminative appearance learning and hierarchical association

Fang Lan, Yu Fengqin 0708

Image Understanding and Computer Vision

Image fusion of multidirectional sum modified Laplacian and tetrolet transform

Shen Yu, Chen Xiaopeng, Yang Qian 0721

Double dual generative adversarial networks for cross-age sketch-to-photo translation

Wu Liuwei, Sun Rui, Kan Junsong, Gao Jun 0732

Medical Image Processing

Image registration method with residual dense relativistic average CGAN

Wang Lifang, Zhang Chengcheng, Qin Pinle, Lin Suzhen, Gao Yuan, Dou Jieliang 0745

Lung segmentation by active shape model approach based on low rank theory

Sun Shenshen, Tian Dan, Wu Wei, Kang Yan, Zhao Hong 0759

Remote Sensing Image Processing

Island shoreline segmentation in remote sensing image based on improved Deeplab network

Wang Zhenhua, Zhong Yuanfu, He Wanwen, Qu Nianyi, Xu Lizhi, Zhang Wenping, Liu Zhixiang 0768

Extracting soil moisture from Fengyun satellite images using a convolutional neural network

Han Yingjuan, Zhou Shunwu, Wang Wenwen, Zhang Xueyi, Zhang Chengming, Li Peiqi 0779

Haze and cloud removal from remote sensing image using HTM algorithm based on curvature filtering

Yang Keke, Jia Yuan, Shen Chuan 0791

Column of ChinaMM 2019

Unsupervised hyperspectral unmixing based on robust non-negative matrix factorization

Song Xiaorui, Wu Lingda, Meng Xiangli 0801

Facial expression recognition based on deep facial landmark features

Wang Shanmin, Shuai Hui, Liu Qingshan 0813

Level set method combining region and edge features for segmenting underwater images

Sun Yang, Chen Zhe, Wang Huibin, Zhang Zhen, Shen Jie 0824

Action recognition under depth spatial-temporal energy feature representation

Chao Xin, Hou Zhenjie, Li Xing, Liang Jiuzhen, Huan Juan, Liu Haoyu 0836

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2020)04-0745-14

论文引用格式: 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(04): 0745-0758 (王丽芳, 张程程, 秦品乐, 蔺素珍, 高媛, 窦杰亮. 2020. 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(04): 0745-0758) [DOI:10.11834/jig.190116]

残差密集相对平均 CGAN 的脑部图像配准

王丽芳, 张程程, 秦品乐, 蔺素珍, 高媛, 窦杰亮

中北大学大数据学院山西省生物医学成像与影像大数据重点实验室, 太原 030051

摘要: 目的 针对图像合成配准算法中鲁棒性差及合成图像特征信息不足导致配准精度不高的问题, 提出了基于残差密集相对平均条件生成对抗网络(residual dense-relativistic average conditional generative adversarial network, RD-RaCGAN)的多模态脑部图像配准方法。方法 相对平均生成对抗网络中的相对平均鉴别器能够增强模型稳定性, 条件生成对抗网络加入条件变量能提高生成数据质量, 结合两种网络特点, 利用残差密集块充分提取深层网络特征的能力, 构建 RD-RaCGAN 合成模型。然后, 待配准的参考 CT(computed tomography)和浮动 MR(magnetic resonance)图像通过已训练好的 RD-RaCGAN 合成模型双向合成对应的参考 MR 和浮动 CT 图像。采用区域自适应配准算法, 从参考 CT 和浮动 CT 图像中选取骨骼信息的关键点, 从浮动 MR 和参考 MR 图像中选取软组织信息的关键点, 通过提取的关键点指导形变场的估计。从浮动 CT 图像到参考 CT 图像估计一个形变场。类似地, 从浮动 MR 图像到参考 MR 图像估计一个形变场。另外, 采用分层对称的思想进一步优化两个形变场, 当两个形变场之间的差异达到最小时, 将两个形变场融合得到最终的形变场, 并将形变场作用于浮动图像完成配准。结果 实验结果表明, 与其他 6 种图像合成方法相比, 本文模型合成的目标图像在视觉效果和客观评价指标上均优于其他方法。对比 Powell 优化的 MI(mutual information)法、ANTs-SyN(advanced normalization toolbox-symmetric normalization)、D. Demons(diffeomorphic demons)、Cue-Aware Net(cue-aware deep regression network)和 I-SI(intensity and spatial information)的图像配准方法, 归一化互信息分别提高了 43.71%、12.87%、10.59%、0.47%、5.59%, 均方根误差均值分别下降了 39.80%、38.67%、15.68%、4.38%、2.61%。结论 本文提出的多模态脑部图像配准方法具有很强的鲁棒性, 能够稳定、准确地完成图像配准任务。

关键词: 医学图像配准; 图像合成; 相对平均生成对抗网络; 残差密集块; 最小二乘; 条件生成对抗网络(CGAN)

Image registration method with residual dense relativistic average CGAN

Wang Lifang, Zhang Chengcheng, Qin Pinle, Lin Suzhen, Gao Yuan, Dou Jieliang

The Key Laboratory of Biomedical Imaging and Imaging on Big Data, North University of China, Taiyuan 030051, China

Abstract: **Objective** Multimodal medical image registration is a key step in medical image analysis and processing as it complements the information from different modality images and provides doctors with a variety of information about diseased tissues or organs. This method enables doctors to make accurate diagnosis and treatment plans. Image registration based on image synthesis is the main method for achieving high-precision registration. A high-quality composite image indicates good registration effect. However, current image-based registration algorithm has poor robustness in synthetic models and provides an insufficient representation of synthetic image feature information, resulting in low registration accuracy. In recent

收稿日期: 2019-04-12; 修回日期: 2019-09-16; 预印本日期: 2019-09-23

基金项目: 山西省自然科学基金项目(201701D121062)

Supported by: Natural Science Foundation of Shanxi Province, China (201701D121062)

years, owing to the success of deep learning in many fields, medical image registration based on deep learning has become a focus of research. The synthetic model is trained according to the modal type of the image to be registered, and the synthetic model bidirectional synthetic image is used to guide the subsequent registration. Anatomical information is employed to guide the registration and improve the accuracy of multimodal image registration. Therefore, a multimodal brain image registration method based on residual dense relative average conditional generative adversarial network (RD-RaCGAN) is proposed in this study. **Method** First, the RD-RaCGAN image synthesis model is constructed by combining the advantages of the relative average discriminator in the relativistic average generative adversarial network, which can enhance the model stability, and the advantages of the conditional generative adversarial network, which can improve the quality of the generated data, and also the ability of residual dense blocks to fully extract the characteristics of the deep network. Residual dense blocks are utilized as core components for building a generator. The purpose of this generator is to capture the law of sample distribution and generate a target image with specific significance, that is, to input a floating magnetic resonance (MR) or reference computed tomography (CT) image and generate the corresponding synthetic CT or synthetic MR image. The convolution neural network is used as a relative average discriminator, which correctly distinguishes an image generated by the generator from the real image. The generator and relative average discriminator perform confrontational training. First, the generator is fixed to train the relative average discriminator. Then, the relative average discriminator is fixed to train the generator, and the loop training is subsequently continued. During training, the least square function optimization generator and relative average discriminator, which are more stable and less saturated than the cross entropy function are selected. The ability of the generator and the relative average discriminator is enhanced, and the image generated by the generator can be falsified. At this point, the synthetic model training is completed. Subsequently, the CT image and MR image to be registered are bidirectionally synthesized into the corresponding reference MR image and floating CT image through the RD-RaCGAN synthesis model that has been trained. Four images obtained by bidirectional synthesis are registered by a region-adaptive registration algorithm. Specifically, the key points of the bone information are selected from the reference CT image and the floating CT image. The key points of the soft tissue information are selected from the floating MR image and the reference MR image, and the estimation of the deformation field is guided by the extracted key points. In other words, one deformation field is estimated from the floating CT image to the reference CT image, and a deformation field is estimated from the floating MR image to the reference MR image. At the same time, the idea of hierarchical symmetry is adopted to further guide the registration. The key points in the image are gradually increased when the reference and floating images are close to each other. Moreover, anatomical information is used to optimize the two deformation fields continuously until the difference between the two deformation fields reaches a minimum. The two deformation fields are fused to form the deformation field between the reference CT image and floating MR image. Finally, the deformation field is applied to the floating image to complete registration. Given that the synthesis of a target image from two images to be registered through the synthesis model requires time, the algorithm efficiency in this study is slightly lower than that of D. Demons (diffeomorphic demons) and ANTs-SyN (advanced normalization toolbox-symmetric normalization). **Result** Given that the quality of the synthesized image directly affects registration accuracy, three sets of contrast experiments are designed to verify the effect of the algorithm in this study. Different algorithms are by MR synthesis CT, and different algorithms are compared by CT synthesis MR, and comparison of the effect of different registration algorithms. The experimental results show that the target image synthesized by the synthesis model in this study is superior to those obtained by the other methods in terms of visual effect and objective evaluation index. The target image synthesized by RD-RaCGAN is similar to the real image and has less noise than the target images generated by the other synthetic methods. As can be seen from the bones of the synthesized brain image and the area near the air interface, the synthetic model in this work visually shows realistic texture details. Compared with the Powell-optimized MI (mutual information) method, ANTs-SyN, D. Demons, Cue-Aware Net (cue-aware deep regression network), and I-SI (intensity and spatial information) image registration methods, the normalized mutual information increased by 43.71%, 12.87%, 10.59%, 0.47%, and 5.59%, respectively. In addition, the mean square root error decreased by 39.80%, 38.67%, 15.68%, 4.38%, and 2.61%, respectively. The results obtained by the registration algorithm in this study are close to the reference image. The registration effect diagram that the difference between the registration image and the reference image obtained by the algorithm in this study is smaller than that obtained by the other

three methods. Small difference between the two images means good registration effect. **Conclusion** This study proposes a multi-modal brain image registration method based on RD-RaCGAN, which solves the problem of the poor robustness of the model synthesis algorithm based on image synthesis, leading to the inaccuracy of the synthetic image and the poor registration effect.

Key words: medical image registration; image synthesis; RaGAN (relativistic average generative adversarial network); residual dense blocks; least squares; CGAN (conditional generative adversarial network)

0 引言

在临床医学中,将不同模态的图像进行配准,可以给医生提供病灶部位的不同信息,提高医生诊断疾病的准确性。因此,多模态医学图像配准在图像融合、肿瘤生长监测、图像引导手术治疗及放疗计划制定等方面发挥着重要作用(Sotiras 等,2013)。

多模态脑部图像配准主要有基于特征、基于灰度和基于图像合成等3类方法。1)基于特征的方法是通过提取图像的显著特征建立图像间的映射关系完成配准,计算效率较高,但配准的精度取决于特征提取的准确性(Oliveira 和 Tavares, 2014),如RAMMS、Hammer 和 SURF (speeded up robust feature)。2)基于灰度的方法直接对不同模态图像的灰度特性进行全局最优化实现配准,配准精度高,但计算量大。如Powell 优化的MI(mutual information)法(Maes 等,1996)、ANTs-SyN(advanced normalization toolbox-symmetric normalization)(Avants 等,2008)和D. Demons(diffeomorphic demons)(Vercauteren 等,2009)。3)基于图像合成的方法是实现高精度配准的主要方法。由于不同模态的图像灰度强度不一致,通过一种模态合成另一种模态,可以将复杂的多模态配准转换为相对容易的单模态配准,当配准图像存在较大形变时比其他两类方法有更大优势。

用于医学图像合成的方法大致分为传统方法和基于深度学习的方法两类。1)传统方法。如基于图谱(Atlas)的方法(Burgos 等,2015)、基于稀疏表示(sparse representation, SR)的方法(Roy 等,2010; Ye 等,2013)和基于随机森林(random forest, RF)的方法(Huynh 等,2016)等。基于图谱的方法通过将图谱集中的所有图像与源图像做配准,根据形变场与图谱标记的映射关系合成目标图像。此类方法对配准的准确性非常敏感,且合成图像边缘轮廓比较模糊。基于稀疏表示的方法由于需要利用稀疏编码理论将源图像块和对应的目标图像块分别表示为相

应的字典,再通过学习字典之间的映射关系,生成给定源图像的目标图像,花费时间较长。基于随机森林的方法当样本数据存在较大噪声时会出现过拟合现象(Ye 等,2015)。2)基于深度学习的方法。近年来,深度学习在图像处理领域取得了显著成果,在大量任务中,深度学习得到的特征被证实比传统方法得到的特征具有更强的表达能力(Krizhevsky 等,2012)。Han(2017)利用深度卷积神经网络(deep convolutional neural network, DCNN)从MR(magnetic resonance)图像合成CT(computed tomography)图像,与基于回归的方法(Huynh 等,2016)相比计算量较小,但当输入图像尺寸较小时,容易忽略预测目标图像中的邻域信息。Xie 等人(2016)将可以保留结构信息的全卷积网络(fully convolutional networks, FCN)用于图像合成,然而要训练稳健且准确的模型需要大量的数据集而且训练耗费时间很长。之后,Nie 等人(2018)提出了基于生成对抗网络(generative adversarial network, GAN)(Goodfellow 等,2014)的合成算法,通过生成器和鉴别器的对抗训练构建合成模型。实验表明,相比其他深度学习算法(Han,2017;Xie 等,2016),结合对抗学习能够生成更逼真的图像且模型结构易于扩展。但生成对抗网络由于生成过程过于自由,当数据集规模较大且图像内容复杂时,会导致生成样本与目标不一致,且在训练过程中易出现模式崩塌、梯度消失的问题(Salimans 等,2016),另外,以上图像合成方法均在单一方向上进行,即只有一种图像的解剖信息指导配准,容易引入偏差。

针对上述问题,本文提出了基于残差密集相对平均条件生成对抗网络(residual dense-relativistic average conditional generative adversarial network, RD-RaCGAN)的多模态脑部图像配准方法,通过结合相对平均生成对抗网络(relativistic average GAN, RaGAN)(Jolicœur-Martineau,2018)和条件生成对抗网络(conditional GAN, CGAN)(Mirza 和 Osindero, 2014)来改进脑部图像的建模方法。其中,生成器

的结构对生成样本的质量有较大影响,Zhang 等人(2018)认为残差密集块能充分利用深层网络中各卷积层的分层特征,并有效缓解网络退化问题。因此,本文的生成器以残差密集块作为核心组件,卷积神经网络为相对平均鉴别器,通过对抗学习训练合成模型,完成 CT 到 MR 及 MR 到 CT 的双向图像合成。然后,采用区域自适应配准算法联合估计形变场,并将形变场作用于浮动图像完成配准。

1 相关工作

1.1 相对平均生成对抗网络

为了解决 GAN 中存在的稳定性问题,相对平均

生成对抗网络(RaGAN)从鉴别器(discriminator) D 的角度对 GAN 做出了改进。在 GAN 中, D 判断输入数据是来自生成器(generator) G 生成的数据 x_f 还是真实数据 x_r 。如图 1 所示,若输入数据为 x_r ,则输出为 1,若输入为 x_f ,则 $D(x_f)$ 输出为 0。RaGAN 通过向 D 中加入梯度惩罚因子 $E_{x_r}[C(x_f)]$ 和 $E_{x_f}[C(x_r)]$ 得到相对平均鉴别器(relative average discriminator) D_{Ra} , $D_{Ra}(x_r, x_f)$ 表示 x_r 比平均给定的 x_f 更真实的概率, $D_{Ra}(x_f, x_r)$ 表示 x_f 比平均给定的 x_r 更假的概率。RaGAN 通过限定 D 的输出值,避免 D 过早饱和,增加网络的稳定性,实验表明,RaGAN 可以经过多次迭代训练而不造成梯度消失。

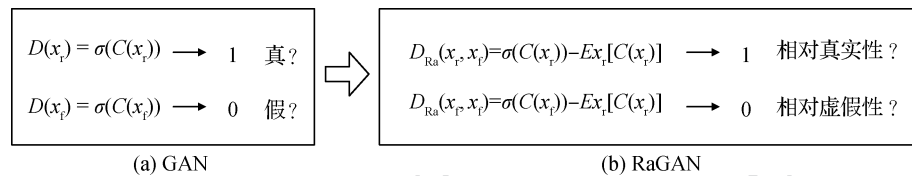


图 1 GAN 与 RaGAN 鉴别器的区别

Fig. 1 Difference between GAN discriminator and RaGAN discriminator ((a) GAN; (b) RaGAN)

1.2 条件生成对抗网络

GAN 作为一种无监督方法,当数据集中的图像内容复杂时,使用 GAN 模型很难控制生成的结果。条件生成对抗网络(CGAN)通过在 G 和 D 的建模中均引入条件变量 c ,将无监督 GAN 改进为有监督的生成模型。 c 可以基于多种信息,如类别标签、图像特征或来自不同模态的数据。 c 对模型增加了约束条件,用于指导数据生成过程,使得 CGAN 更关注对于阐释样本相关的特征,并忽略不

相关特征,有效提高了网络收敛速度和生成数据的质量。

1.3 残差密集块

残差密集块(residual dense block, RDB)将残差网络(He 等, 2016)中的跳跃连接与密集网络(Huang 等, 2017)中的密集连接相结合,使得网络不仅具有残差网络能缓解深层网络退化的效果,又具有密集网络增强特征表达的作用。基本的残差密集块如图 2 所示。

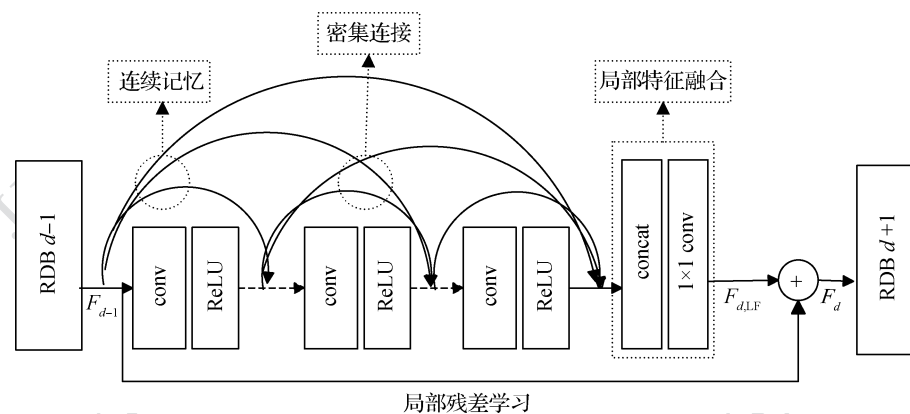


图 2 基本残差密集块

Fig. 2 Basic residual dense block

图2中,每个RDB包含3层卷积(conv),后接ReLU激活函数, F_{d-1} 和 F_d 分别表示第 $d-1$ 和第 d 个RDB的输出,RDB通过连续记忆将 F_{d-1} 的特征信息输入到第 d 个RDB的每个卷积层中。同时,密集连接使得层与层之间的特征信息充分利用,concat表示第 $d-1$ 和第 d 个RDB产生的特征图的级联,随着层数和特征数的增多,为了去除冗余信息,对有效的特征进行局部特征融合(local feature fusion, LF)得到 $F_{d,LF}$ 。 F_{d-1} 到 F_d 的局部残差学习使得网络能够顺畅训练。最终,第 d 个RDB的输出为 $F_d = F_{d-1} + F_{d,LF}$ 。

2 基于RD-RaCGAN的脑部图像配准算法

2.1 RD-RaCGAN 算法网络结构

综合CGAN中加入条件 c 能够提高生成图像质量和RaGAN中采用相对平均鉴别器能够增加模型稳定性的特点,使用有监督的相对平均条件生成对抗网络模型训练生成器 G 和相对平均鉴别器 D_{Ra} 。由于RDB具有在深层网络中充分利用所有卷积层的分层信息的优势, G 通过使用多层RDB来保证提取信息的充分性,并在训练阶段选用相对交叉熵更稳定的最小二乘损失函数优化 G 和 D_{Ra} ,构建残差密集相对平均条件对抗网络合成模型,具体框架如图3所示。

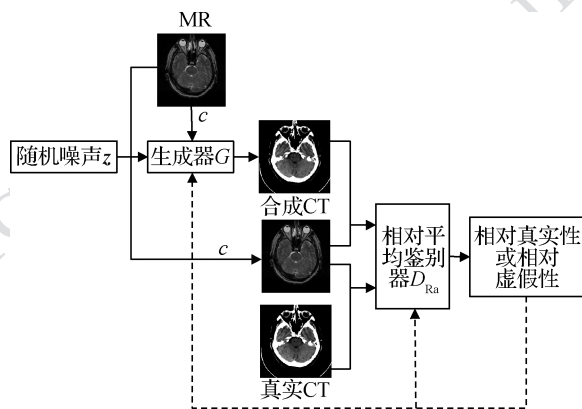


图3 RD-RaCGAN 的基本框架

Fig. 3 Basic framework of RD-RaCGAN

RD-RaCGAN (residual dense-relativistic average conditional generative adversarial network) 由生成器(G)和相对平均鉴别器(D_{Ra})构成,数据集中每幅MR图像都有1幅对应的CT图像,当MR合成CT

图像时,在生成器中,输入随机噪声 z 的同时,将MR图像作为条件 c ,这样通过图像可以直接调节 G ,每幅MR图像输入 G 都会生成1幅合成CT (synthesize CT, S-CT)图像,然后,CT、S-CT和MR图像输入 D_{Ra} 完成判别。 G 和 D_{Ra} 进行对抗训练,并用最小二乘损失函数不断更新网络参数,指导生成高质量的图像,最终 G 生成的S-CT图像能够以假乱真,模型训练完成。CT合成MR图像的过程与此相同,通过训练好的合成模型可以将待配准的CT/MR图像分别合成对应的S-CT图像和合成MR (synthesize MR, S-MR)图像,并将CT、S-CT、MR、S-MR图像用于区域自适应配准算法。

2.1.1 生成器

生成器 G 的目的是捕捉样本分布的规律,生成具有特定意义的目标图像,即输入MR或CT图像,生成相对应的S-CT或S-MR图像。本文通过增加网络的深度来提高网络的性能。由于过深的网络难以训练,将RDB设置为9层,并将浅层特征提取部分设计为3层以达到充分提取特征信息的目的,另外,采用可学习的反卷积层保证输入图像和生成图像尺寸一致,得到更细致的目标图像,并减少孔洞的出现。生成器的网络结构如图4所示,当MR合成CT图像时, I_{MR} 和 I_{CT} 表示输入和输出, k 、 n 和 s 分别表示内核(kernel)尺寸、滤波器个数(number)和步长(stride),并在对图像卷积操作时,进行了填充操作,填充适宜的0用以保持图像的尺寸。网络主要分为4个部分:第1部分为3个卷积层(conv),输入大小为 $256 \times 256 \times 1$ 的MR图像,通过3个卷积层的浅层特征提取生成较小尺寸的特征图像,并用于全局残差学习及RDB的输入;第2部分为9个RDB,每个RDB包括3层卷积,其中卷积层都为256个 3×3 大小滤波器的卷积层;第3部分为全局特征融合及全局残差学习,concat层将9个RDB的分层特征融合,经 1×1 卷积核进行降维及信息整合(Tai等,2017)后得到深层特征,通过跳跃连接的全局残差学习将浅层特征与深层特征结合在一起,得到全局密集特征;第4部分是反卷积(conv transposed)和重建,为了使生成器生成与输入图像大小一致的CT图像,通过两层反卷积操作使前一阶段获得的特征逐渐恢复到原图大小,之后是ReLU激活函数,最后通过 4×4 的卷积层后输出大小为 $256 \times 256 \times 1$ 的CT图像。

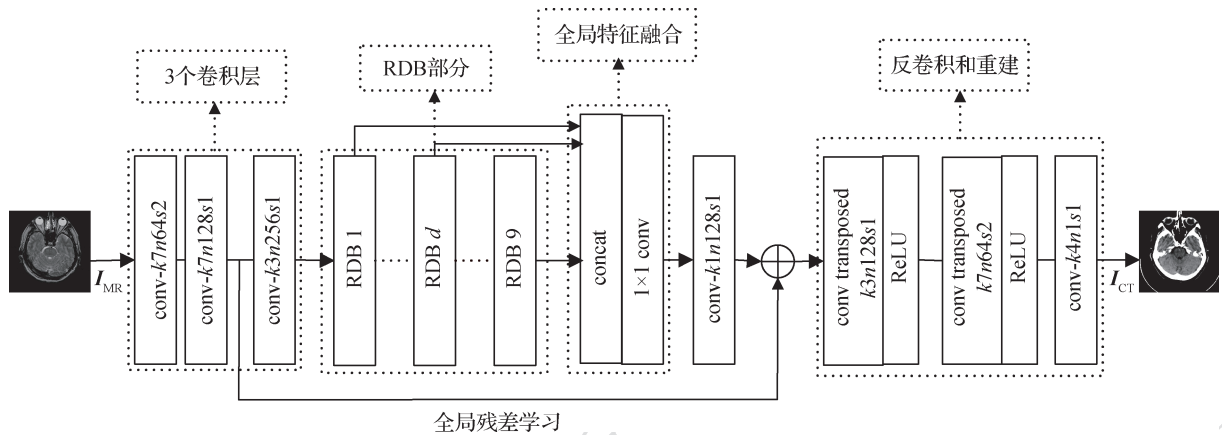


图4 生成器网络结构图

Fig. 4 Generator network structure diagram

2.1.2 相对平均鉴别器

相对平均鉴别器 (D_{Ra}) 的目的是正确区分出合成图像和真实图像,当MR由G生成S-CT时,将CT、S-CT以及MR图像输入 D_{Ra} 进行判断,通过最小二乘函数计算鉴别损失来衡量鉴别器的判别能力,并将信息反馈给 D_{Ra} ,使 D_{Ra} 继续增强学习,直到与G达到平衡状态,训练完成。本文中的 D_{Ra} 选用典型的卷积神经网络(convolutional neural network,

CNN)架构,如图5所示,输入待鉴别的CT图像,通过5层的conv层提取图像特征,为了增大局部感受野,采用 4×4 尺寸的卷积核,步长为2,然后用BN(batch normalization)层来提高训练速度和避免局部最优,后接ReLU激活函数。最后,经全连接层(fully connected layer, FC)后加一个sigmoid函数将输出数据归一化到0和1,然后根据阈值输出概率值作为判别结果。

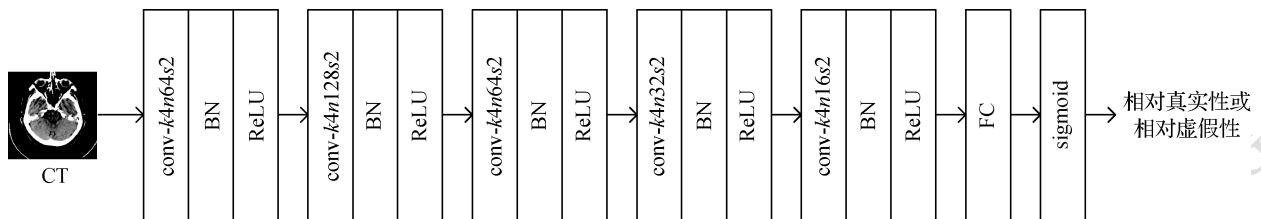


图5 相对平均鉴别器网络结构图

Fig. 5 Relativistic average discriminator network structure diagram

2.1.3 最小二乘损失函数

损失函数的选择是影响网络性能的重要因素, GAN、CGAN 和 RaGAN 均以交叉熵为损失函数。Arjovsky等人(2017)指出 GAN 训练不稳定是由于交叉熵不适合衡量不相交部分的分布造成的, Mao 等人(2016)指出对抗学习过程中的不稳定性部分是由鉴别器的饱和引起的。如图6所示,最小二乘函数的饱和区域远小于交叉熵函数的饱和区域,对于对抗学习,最小二乘损失函数比交叉熵损失函数更稳定,生成图像质量更高。因此,本文采用最小二乘作为损失函数,将最小二乘函数看作最优化问题,通过最小化目标函数值,达到合成图像与真实图像相似的目的。

以最小二乘为损失函数的 CGAN 中 G 和 D 的

损失分别为

$$L_G^{CGAN} = \frac{1}{2} E_{c \sim p(c)} [D(\mathbf{x}_f) - 1]^2 \quad (1)$$

$$L_D^{CGAN} = \frac{1}{2} E_{\mathbf{x}_r \sim p(\mathbf{x}_r)} [(D(\mathbf{x}_r) - 1)^2] + \frac{1}{2} E_{c \sim p(c)} [D(c, G(c, \mathbf{z}))^2] \quad (2)$$

式中, p 为数据的概率分布, c 为条件, $\mathbf{z}$ 为随机噪声,并用 L_1 形式的正则项约束生成图像 $\mathbf{x}_f$ 与真实图像 $\mathbf{x}_r$ 之间的偏差,具体为

$$L_1(G) = E_{\mathbf{x}_r, \mathbf{x}_f \sim p(\mathbf{x}_r, \mathbf{x}_f)} [\|\mathbf{x}_r - \mathbf{x}_f\|_1] \quad (3)$$

在式(1)和式(2)中加入相对平均鉴别器的梯度惩罚因子,得到 RD-RaCGAN 的损失函数为

$$L_{RD-RaCGAN} = \{\min(L_{D_{Ra}}^{RaCGAN}), \min(L_G^{RaCGAN})\} \quad (4)$$

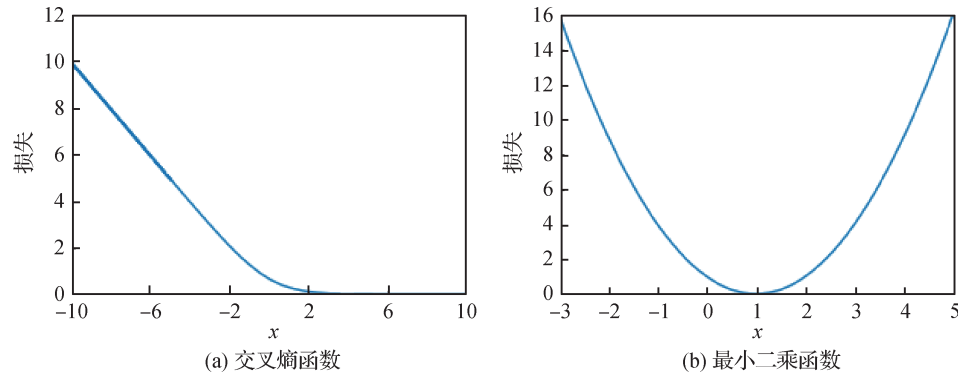


图6 交叉熵和最小二乘函数

Fig. 6 Cross entropy and least squares functions ((a) cross entropy; (b) least squares)

$$L_{D_{Ra}}^{RaCGAN} = \frac{1}{2} E_{x_r, x_f \sim p(x_r, x_f)} \times \{ [D_{Ra}(x_r, x_f) - 1]^2 + [D_{Ra}(x_f, x_r)]^2 \} \quad (5)$$

$$L_G^{RaCGAN} = \frac{1}{2} E_{x_r, x_f \sim p(x_r, x_f)} \times [D_{Ra}(x_r, x_f) - 1]^2 + \lambda L_1(G) \quad (6)$$

式中, $L_{RD-RaCGAN}$ 表示 RD-RaCGAN 算法的损失函数, $L_{D_{Ra}}^{RaCGAN}$ 、 L_G^{RaCGAN} 分别为相对平均鉴别器和生成器的损失函数, λ 为正则项系数, 网络训练的目的即最小化 $L_{D_{Ra}}^{RaCGAN}$ 和 L_G^{RaCGAN} 。损失函数值越小, 说明图像生成质量越高。在 G 和 D_{Ra} 的对抗训练过程中, D_{Ra} 的训练过程是最大化准确率、最小化损失函数的过程, 其对真实图像的输出 $D_{Ra}(x_r, x_f)$ 趋向 1, 对合成图像的输出 $D_{Ra}(x_f, x_r)$ 趋向 0。而训练 G 的目的是最小化 D_{Ra} 的准确率, 使 $D_{Ra}(x_f, x_r)$ 趋向 1。最终, 随着 G 和 D_{Ra} 的对抗训练, G 和 D_{Ra} 达到纳什均衡 (Ratliff 等, 2013), 此时, L_G^{RaCGAN} 和 $L_{D_{Ra}}^{RaCGAN}$ 达到最优值, 训练完成, 整体网络结构如图 7 所示。

2.2 区域自适应配准算法

以 CT 为参考图像, MR 为浮动图像, 经已训练好的 RD-RaCGAN 合成模型分别合成参考 S-MR 和浮动 S-CT 图像。为了避免单向合成的图像信息指导多模态图像配准引入偏差, 在 MR 与 CT 配准时, 同时利用 S-MR 和 S-CT 的解剖信息, 采用区域自适应配准算法 (Cao 等, 2018a) 进行配准, 主要过程分为 3 个阶段:

1) 关键点采样。为了得到 CT 和 MR 之间的形变场 φ , 在具有解剖信息的 CT、MR、S-CT 和 S-MR 图像域中进行关键点采样, 具体为

$$P = \frac{|\nabla_x| + |\nabla_y| + |\nabla_z|}{|\nabla G|} \quad (7)$$

式中, P 为采样概率, ∇_x, ∇_y 和 ∇_z 分别表示 3 个方向的图像梯度算子, $|\nabla G|$ 是图像中的最大梯度, 以将 P 归一化为 $[0, 1]$

2) 关键点的区域自适应选择。对式 (7) 采取截

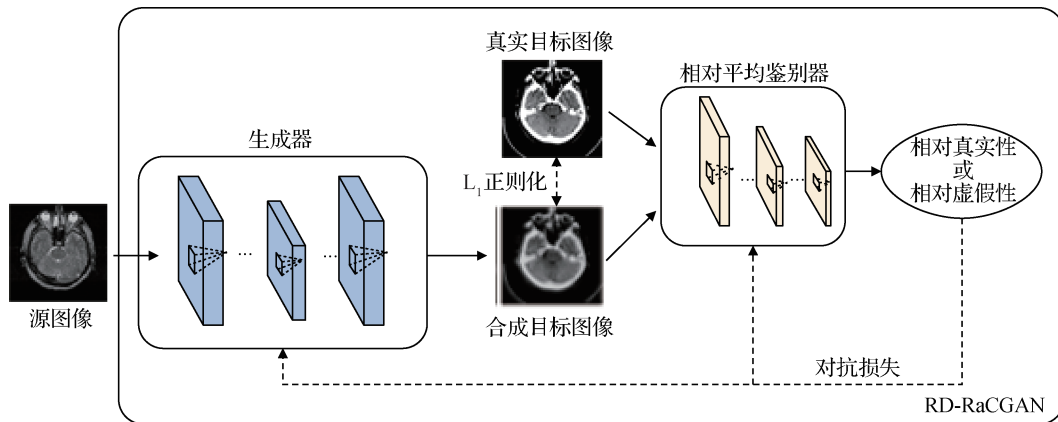


图7 整体网络结构

Fig. 7 Overall network structure

断阈值操作,从 CT 对(CT 和 S-CT)中选取骨骼信息的关键点,从 MR 对(MR 和 S-MR)中选取软组织信息的关键点,并通过归一化互相关(normalized cross correlation, NCC)检测关键点间的匹配关系。

3) 分层对称配准。通过双向图像合成,将 MR 和 CT 的配准转化为 CT 对和 MR 对的单模态配准,即

$$\varphi = \arg \min M(I_{CT}, D(I_{S-CT}, \varphi)) + M(I_{S-MR}, D(I_{MR}, \varphi)) + \alpha R(\varphi) \quad (8)$$

式中, I_{CT} , I_{MR} , I_{S-CT} 和 I_{S-MR} 分别表示 CT、MR、S-CT 和 S-MR 图像, $D(\cdot)$ 表示通过形变场 φ 变形图像的操作符。 R 为约束 φ 平滑度的正则项, M 为相似性测度, α 为权重衡量系数。为了求解式(8),采用分层对称配准的思想,通过最小化 CT 对和 MR 对之间的差异优化形变场 φ , 得

$$\{\varphi_1, \varphi_2\} = \arg \min_{\varphi_1, \varphi_2} M(D(T, \varphi_1), D(S, \varphi_2)) + \alpha(R(\varphi_1) + R(\varphi_2)) \quad (9)$$

将 CT、S-MR 作为模板空间 T , MR、S-CT 作为目标空间 S , R 为基于薄板样条(thin plate spline, TPS) (Bookstein, 1989) 的变换模型。如图 8 所示,通过关键点指导 T 和 S 中的图像同时扭曲得到两个形变场,随着 T 和 S 彼此靠近,逐渐增加关键点,并通过更丰富的信息改进两个形变场,直到它们在整个变形的中点彼此相似,得到两个形变场 φ_1 和 φ_2 , 此时, φ_1 和 φ_2 之间的差异已达到最小。最终得形变场 φ 为

$$\varphi = \varphi_2 \circ \varphi_1^{-1} \quad (10)$$

$$\varphi^{-1} = \varphi_1 \circ \varphi_2^{-1}$$

式中,“ $\circ$ ”表示变形成分, φ 、 φ^{-1} 表示前向和后向形变场, φ^{-1} 是 φ 的倒数。

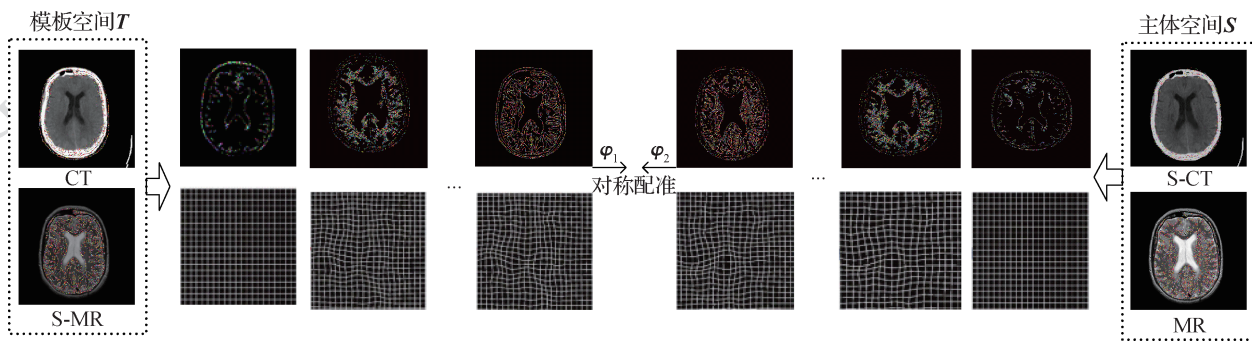


图 8 区域自适应配准算法

Fig. 8 Region adaptive registration algorithm

2.3 RD-RaCGAN 的脑部图像配准算法

本文提出的基于 RD-RaCGAN 的多模态脑部图像配准算法流程如图 9 所示,具体步骤为:

1) 构建 RD-RaCGAN 合成模型。CGAN 中条件 c 能直接指导生成器并提高生成图像质量、RaGAN 中相对平均鉴别器能提高网络稳定性、RDB 能在深层网络中充分提取信息,综合这 3 种能力,以 RDB 为核心组件构建生成器 G , CNN 为相对平均鉴别器 D_{Ra} , 对抗训练过程中,采用相对交叉熵函数更稳定、饱和度更小的最小二乘损失函数优化 G 和 D_{Ra} , 直到 G 、 D_{Ra} 达到平衡状态, RD-RaCGAN 合成模型训练完成。

2) 合成图像生成。输入两幅大小相同的待配准脑部图像,如参考 CT 和浮动 MR 图像,经 RD-RaCGAN 合成模型执行 CT 到 S-MR 和 MR 到 S-CT 的图像合成。

3) 区域自适应配准。双向合成后得到 4 幅图

像:CT 对(CT 和 S-CT)和 MR 对(MR 和 S-MR)。从 CT 对选取骨骼信息的关键点,从 MR 对选取软组织信息的关键点,采用区域自适应配准算法,在配准期间逐渐增加关键点的数量,分层对称地指导形变场的估计。

4) 将最终的形变场作用于浮动图像完成配准。

3 实验结果与分析

3.1 数据集与实验环境

本文的训练集来自数据库 RIRE (retrospective image registration evaluation) (<http://insight-journal.org/rire>) 和 Atlas (Johnson 和 Becker, 1999)。从 RIRE 数据库的 18 位患者及 Atlas 数据库中的急性脑中风、高血压性脑病变和颅内肿瘤患者中挑选了 518 对清晰度高、纹理丰富、细节复杂的 MR/CT 图像作为数据集,其中, RIRE 数据库中 CT 图像的体

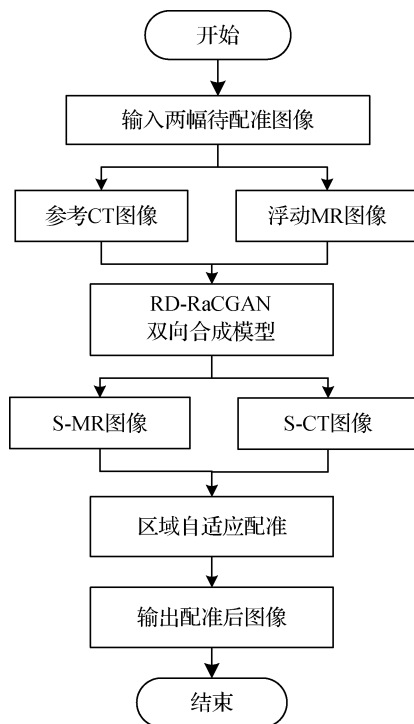


图9 本文算法流程图

Fig. 9 Algorithm flow chart in this paper

素大小在 $0.40 \times 0.40 \times 3.00 \text{ mm}^3 \sim 0.65 \times 0.65 \times 4.00 \text{ mm}^3$ 之间, MR 图像的体素大小在 $0.82 \times 0.82 \times 3.00 \text{ mm}^3 \sim 1.25 \times 1.25 \times 4.00 \text{ mm}^3$ 之间, Atlas 数据库中 CT 图像大小为 512×512 像素, MR 图像为 256×256 像素。首先, 将 RIRE 数据库中的 DICOM (digital imaging and communications in medicine) 格式的数据及 Atlas 数据库中的 GIF (graphics interchange format) 格式的数据均转换为 PNG (portable network graphic format) 格式。其次, 采用 N4 (Tustison 等, 2010) 偏差校正算法消除 MR 图像的灰度不均匀性, 最后, 将所有图像重新采样到 256×256 像素, 切片厚度为 1 mm 。数据集分为 3 部分, 包括训练集、测试集、验证集。训练集用于训练网络模型, 占数据的 80%, 共 414 对; 验证集用于对模型准确率进行初步评估, 以决定是否停止训练, 占 10%, 共 52 对; 测试集用于评估最终模型的泛化能力, 占 10%, 共 52 对; 由于深度模型通常受益于大数据, 为了充分利用数据集, 采用 Albumentations (Buslaev 等, 2018) 对训练集进行数据增强, 对数据集中的每幅图像分别进行 90° 、 180° 和 270° 旋转、水平和垂直翻转、转置、弹性变换、网格变形、光学畸变等操作, 获得 $4 \times 6 - 1 = 23$ 倍的数据, 即 9 936 对图

像进行训练。

实验的硬件平台是一台 Intel Xeon 服务器, 搭载 2 块 NVIDIA Tesla M40 的 GPU, 每块显存 12 GB。软件平台是 64 位的 Windows 7 操作系统, Pycharm64, Tensorflow V1.2, MATLAB R2016a。

3.2 训练过程

生成器 G 和相对平均鉴别器 D_{Ra} 采用交替的方式进行对抗式训练。首先固定 G 训练 D_{Ra} , 然后固定 D_{Ra} 训练 G , 再继续循环训练, G 和 D_{Ra} 的能力都得到增强, 最终 G 生成的图像能够以假乱真, 此时训练结束, 当有新的图像输入时, G 生成的 x_f 就可以作为真实的图像结果使用。具体训练过程如下:

- 1) 固定 G , 用 x_r 训练 D_{Ra} , 训练 n 个 epoch (训练样本集中的每个样本训练 1 次);
- 2) 用 G 生成 x_f , 用 x_f 训练 D_{Ra} , 训练 n 个 epoch;
- 3) 固定 D_{Ra} , 用 D_{Ra} 的输出作为图像标签, 用 D_{Ra} 的训练损失训练 G , 训练 n 个 epoch;
- 4) 重复步骤 1)–3), 并用验证集对模型进行评估, 得到乱真的目标图像。
- 5) 选用测试集图像评估最终模型性能, 以及合成图像质量。

本文采用 Adam 优化算法 (Kingma 和 Ba, 2014) 在训练过程中促使鉴别损失和生成损失函数达到最小来不断更新网络的参数。网络初始学习率设置为 1×10^{-4} , 动量参数设定为 0.9, 权重衰减为 5×10^{-4} 。在初始赋值方面, 卷积层和反卷积层的赋值满足均值为 0、标准差为 0.01 的正态分布。受 GPU 显存的限制, 采用 mini-batch 的训练方式, batch-size 设置为 16, 由于 MR 图像相对 CT 图像解剖信息更丰富, MR 合成 S-CT, 属于由“复杂”到“简单”的合成, CT 合成 S-MR, 属于由“简单”到“复杂”的合成, 在训练过程中, 由 CT 到 S-MR 的合成模型训练时间相对较长, 经过 200 个 epoch 的训练后, L_C^{RaCGAN} 和 $L_{D_{Ra}}^{RaCGAN}$ 随着迭代次数的变化趋势如图 10 所示。

从图 10 可以看出, 随着迭代次数的增加, L_C^{RaCGAN} 整体呈下降趋势, 曲线后段逐渐稳定并接近 0.2, 说明生成器错误率已经小于 20%, 生成的目标图像与真实图像非常相似。另一方面, 随着迭代次数的增加, $L_{D_{Ra}}^{RaCGAN}$ 整体呈现缓慢上升并趋于稳定, 综合权衡生成损失和判别损失, 确定最终迭代次数为 $9\,936/16 \times 200 = 124\,200$ 次 (训练样本数:

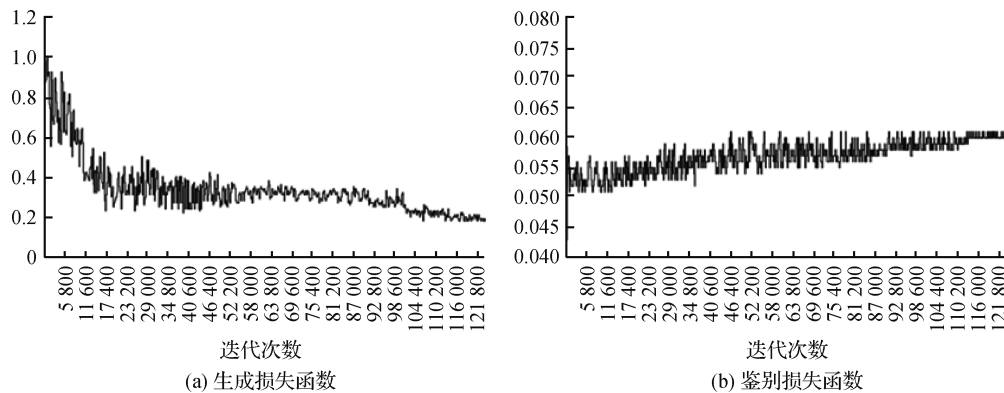


图 10 损失函数变化趋势图

Fig. 10 The change trend chart of loss ((a) generation loss function; (b) discriminator loss)

9 936, batch-size: 16, epoch 数量: 200)。另外,从损失函数的变化情况可以看出,在整个训练过程中, L_G^{RaCGAN} 和 $L_{D_{\text{Ra}}}^{\text{RaCGAN}}$ 均未出现大幅度的忽高忽低的情况,说明使用 RD-RaCGAN 训练模型较为稳定。

3.3 结果比较与分析

由于合成图像的质量直接影响配准精度,为了验证本文算法配准图像的效果,设计了3组对比实验。

1) 由 MR 合成 CT 不同算法效果比较。图 11 给出了测试集中的 3 组脑部 MR-T2 及对应的 CT 图像,本组实验选取 FCN-GAN (Nie 等, 2018)、Atlas、DCNN (Han, 2017)、SR 和 RF 等 5 种方法比较并分析合成效果。由图 11 可以看出, Atlas 算法合成的 CT 图像轮廓不清晰、不完整; SR 算法合成的图像整体看起来很模糊; DCNN 算法得到的图像有很多噪声; RF 和 FCN-GAN 算法在图像锐度和视觉效果方面都较好,但相比 RD-RaCGAN 在细节方面的处理仍显不足,例如,从第 1 行放大图可以看出耳道、更

细的骨骼和颅骨后侧这些复杂的细节部分合成不精确;而 RD-RaCGAN 算法合成的图像在视觉上显示出了更逼真的纹理细节,与真实 CT 图像更相似、噪声更小,从第 2 行可以看出骨骼和空气界面附近区域的合成较其他方法更准确。

为了客观评价合成效果,采用平均绝对误差 (mean absolute error, MAE) 和峰值信噪比 (peak signal-noise ratio, PSNR) 来衡量合成图像的精度,具体为

$$M_{\text{AE}} = \frac{\sum_{i=1}^H |\text{CT}_R(i) - \text{CT}_S(i)|}{H} \quad (11)$$

$$P_{\text{SNR}} = 10 \ln \left(\frac{Q^2}{M_{\text{SE}}} \right) \quad (12)$$

式中, CT_R 、 CT_S 分别为真实 CT 和合成 CT 图像, H 是每幅图像中像素的个数, Q 为 CT_S 和 CT_R 的最大强度值, M_{SE} 为均方误差。通常, M_{AE} 值越小, P_{SNR} 值越高,合成效果越好, CT_S 越接近 CT_R 。

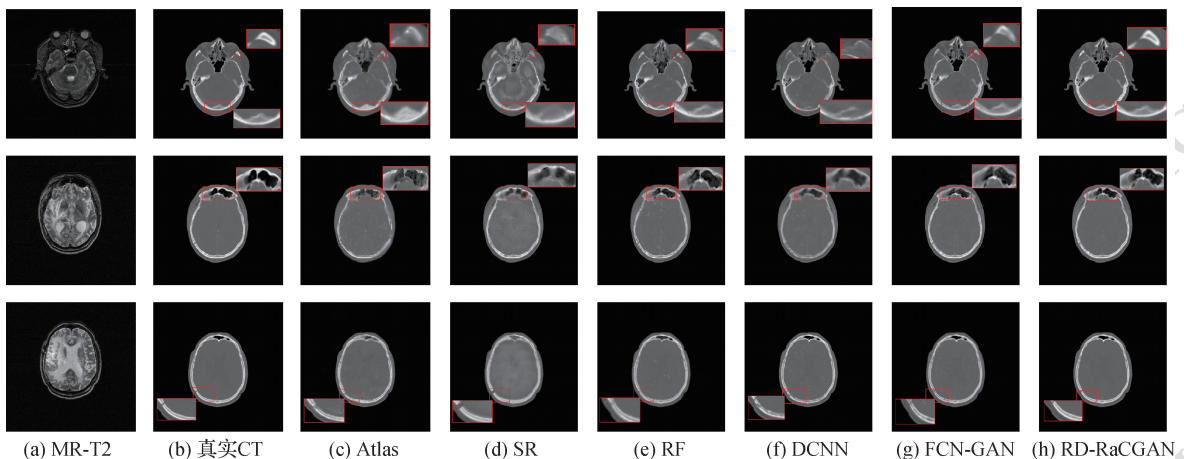


图 11 不同算法由 MR-T2 合成 CT 图像结果图

Fig. 11 Different algorithms from MR-T2 to synthesize CT image result

((a) MR-T2 image; (b) real CT image; (c) Atlas; (d) SR; (e) RF; (f) DCNN; (g) FCN-GAN; (h) RD-RaCGAN)

表1为使用6种算法合成CT图像所得MAE和PSNR的平均值(mean)和标准差(std)值。为了更好地评估各算法的合成性能,表1总结了大脑不同区域的MAE值,包括空气、骨骼和软组织区域。依据对实验结果中配准后图像以及表1的各项指标来看,本文使用的合成模型具有更好的合成效果。通过计算其他区域的MAE值可以得出各合成方法在骨骼和空气区域的产生误差较大,其余组织区域MAE值较小的结论。

表1 不同合成算法下合成CT图像的MAE和PSNR值
Table 1 MAE and PSNR values of CT images synthesized by different synthesis algorithms results

算法		全脑		空气	骨骼	组织
		MAE	PSNR/dB	MAE	MAE	MAE
Atlas	mean	169.5	20.9	294.21	271.39	55.6
	std	35.6	1.6	65.5	64.95	23.34
SR	mean	166	21.2	302.48	283.96	52.18
	std	37.6	1.7	68.99	60.5	22.84
RF	mean	99.9	26.3	240.94	264.1	42.61
	std	14.2	1.4	57.47	55.7	14.32
DCNN	mean	102.4	26.2	256.5	255.22	43.23
	std	11.3	1.22	57.32	55.32	16.9
FCN-GAN	mean	93.5	27.9	238.5	245.3	39.09
	std	13.9	1.18	52.5	48.97	11.2
本文	mean	88.3	28.15	235.44	240.2	38.65
	std	10.25	1.17	50.21	47.74	8.58

2)由CT合成MR不同算法效果比较。图12给出了从测试集中选出的两组脑部MR-T2及其对应的CT图像,并分别通过MT-RF(multi-target regression forest)(Cao等,2018a)算法与本文中的RD-RaCGAN算法从CT图像合成MR图像所得结果。从图12可以看出,MT-RF算法比RD-RaCGAN算法对丰富纹理处理仍然显得不够细腻,MT-RF算法合成的图像比较模糊且视觉过于平滑,RD-RaCGAN算法在视觉上显示出了更逼真的纹理细节。

表2列出了使用两种合成算法各项评价指标的计算结果,结果表明本文算法的各项评价指标均优于MT-RF算法。通过观察发现,表2中的MAE值略大于表1中的数值,而PSNR值略小于表1中的

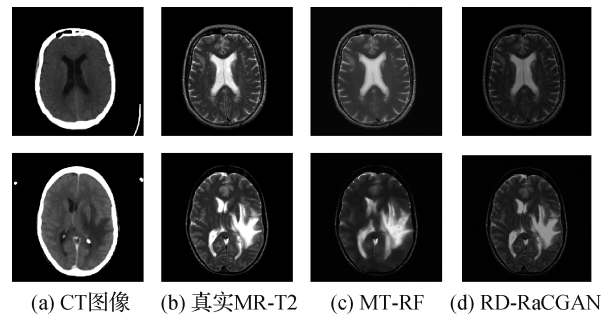


图12 不同算法由CT合成MR-T2图像结果图

Fig. 12 MR-T2 image synthesized from CT of different algorithms((a) CT image; (b) real MR-T2; (c) MT-RF; (d) RD-RaCGAN)

表2 不同合成算法下合成MR图像的MAE和PSNR值
Table 2 MAE and PSNR values of MR images synthesized by different synthesis algorithms results

算法		全脑		空气	骨骼	组织
		MAE	PSNR/dB	MAE	MAE	MAE
MT-RF	mean	124.53	26.52	238.45	246.96	40.52
	std	7.88	0.78	53.24	52.37	14.33
本文	mean	120.55	27.21	225.93	241.78	39.02
	std	6.43	0.82	51.85	48.02	7.96

数值,这是因为由解剖信息简单的CT合成复杂的MR图像比由MR合成CT的计算复杂度更高,合成产生的误差也相应的略高。

3)不同配准算法的效果比较。为了评估基于RD-RaCGAN算法的配准性能,选取Atlas数据库中的急性脑卒中患者CT/MR图像,图像大小均为 256×256 像素,将实验分为两组,每组从24对切片中随机抽取8对进行实验,以CT为参考图像,MR为浮动图像,并对两组图像中的MR图像分别随机施加 ± 20 像素的平移、 $\pm 10^\circ$ 的旋转和20%的纹波变形。实验分别使用常用于多模态医学配准Powell优化的MI法、D. Demons、ANTs-SyN、Cue-Aware Net(Cao等,2018b)、I-SI(Öfverstedt等,2019)以及本文方法进行比较。图12展示了采用上述6种方法配准的效果和配准后的图像差。

本组实验采用归一化互信息(normalized mutual information,NMI)、均方根误差(root mean square error,RMSE)评价图像的配准效果。NMI度量描述的是两幅图像之间的统计相关性,配准后图像与参考

图像之间的 NMI 值越大,两幅图像越相似,配准越精确;RMSE 表示配准后图像与参考图像的偏差, RMSE 值越小,则图像配准效果越好。NMI 和 RMSE 的表达式为

$$N_{MI} = \frac{H(R) + H(M(\varphi(n)))}{H(R, M(\varphi(n)))} \quad (13)$$

$$R_{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (M_i - R_i)^2} \quad (14)$$

式中, R 、 M 分别表示参考图像和浮动图像上的灰度, $H(R)$ 和 $H(M(\varphi(n)))$ 分别为参考图像和配准后的浮动图像的熵, $H(R, M(\varphi(n)))$ 为联合熵,且 NMI 的范围为 $[1, 2]$, n 表示图像中的像素点的数量。

表 3 列出了本文算法及对比算法的平均 NMI、RMSE 值及平均配准时间。从表 3 可以看出,本文算法与 Powell-MI、ANTs-SyN、D. Demons、Cue-Aware Net 和 I-SI 的图像配准的方法相比, NMI 均值分别提高了 43.71%、12.87%、10.59%、0.47%、5.59%, RMSE 均值分别下降了 39.80%、38.67%、15.68%、4.38%、2.61%,故基于 RD-RaCGAN 的配准算法效果更好,精度更高。但从配准时间上来看,本文算法的效率略低于 D. Demons 和 ANTs-SyN,主要是因为

表 3 不同算法配准图像的评价指标值及时间

Table 3 The average evaluation index value and time of images registered by different algorithms

配准方法	NMI	RMSE	配准时间/s
Powell-MI	1.025	1.304	208.86
ANTs-SyN	1.305	1.280	148.52
D. Demons	1.332	0.931	136.48
Cue-Aware Net	1.466	0.821	300.80
I-SI	1.395	0.806	209.40
本文	1.473	0.785	181.32

本文方法是基于图像合成且合成模型是通过 RD-RaCGAN 训练得到的,并在此基础上对待配准的 CT/MR 图像双向合成,因此,通过合成模型合成相应的 S-CT/S-MR 图像需要花费一定时间。

图 13 是不同算法的配准结果和配准后图像差,从图中箭头所指的纹理和边缘区域可以看出:1) 经过扭曲变形后的浮动图像,通过本文配准算法得到的结果更接近参考图像;2) 本文算法得到的配准图像与参考图像的图像差比其他 5 种方法得到的图像差的差异性小。配准后两幅图像的差异性越小,配准效果越好。

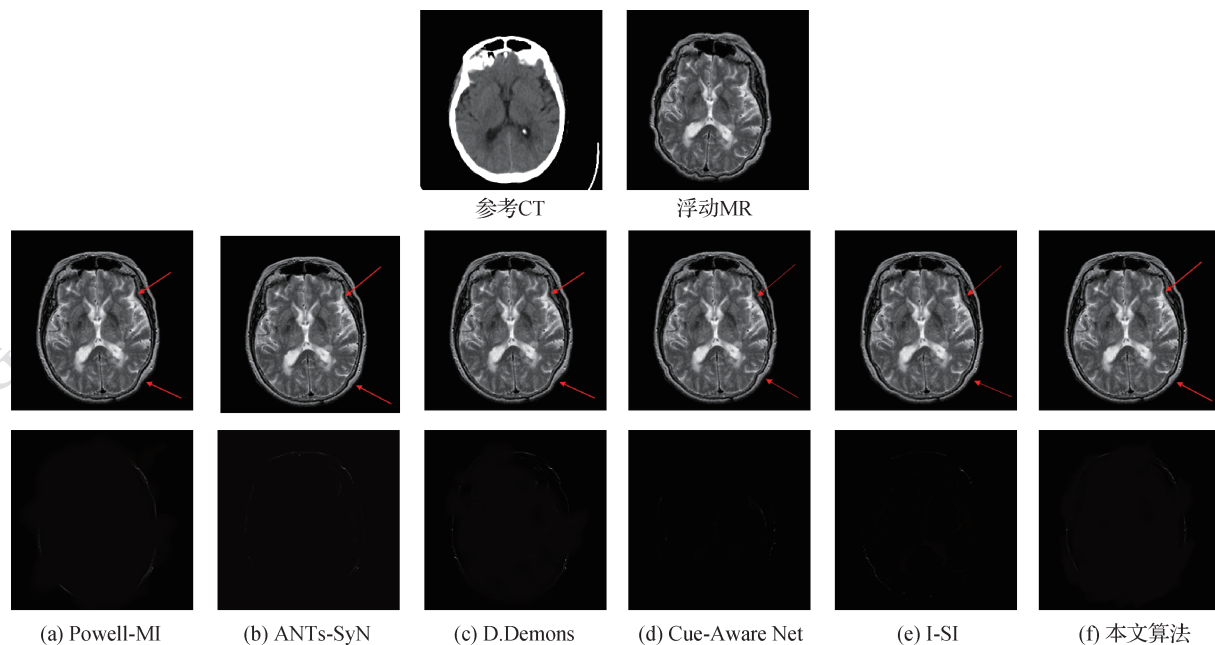


图 13 不同算法的配准结果和配准后图像差

Fig. 13 Registration results of different algorithms and post-registration image difference

((a) Powell-MI; (b) ANTs-SyN; (c) D. Demons; (d) Cue-Aware Net; (e) I-SI; (f) ours)

4 结 论

本文提出了基于残差密集相对平均条件生成对抗网络(RD-RaCGAN)的多模态脑部图像配准方法,解决了基于图像合成的配准算法合成模型鲁棒性差导致合成图像不准确和配准效果欠佳的问题。实验结果表明,本文方法能够将待配准的MR和CT图像通过RD-RaCGAN合成模型合成S-CT及S-MR图像,并用于联合估计形变场。本文算法有以下特点:1)RD-RaCGAN合成模型综合CGAN中条件变量能直接指导生成器并提高生成图像质量及RaGAN中相对平均鉴别器能提高网络稳定性的优点,具有鲁棒性强、合成图像效果好的特点;2)利用双向合成的图像信息指导配准,降低了多模态配准的难度,提高了配准精度。

但本文算法也存在局限性,尚有较大提升空间,具体表现在:1)本文算法需要将两幅待配准图像通过预先训练好的合成模型合成对应图像再进行相应的配准,相比传统方法需要更多的时间。后续工作将致力于有效降低算法执行时间的研究。2)本文算法中的图像合成模型仅适用于单图像域的合成任务,当进行多个域间的合成时,需要为每对域训练一个单独的生成器,计算成本很高。如何设计多图像域合成模型以适用于多模态图像配准任务将有助于提升配准效率。

参考文献(References)

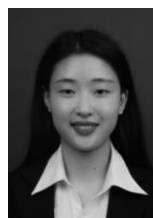
- Arjovsky M, Chintala S and Bottou L. 2017. Wasserstein GAN [EB/OL]. [2019-04-10]. <https://arxiv.org/pdf/1701.07875.pdf>
- Avants B B, Epstein C L, Grossman M and Gee J C. 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12 (1): 26-41 [DOI: 10.1016/j.media.2007.06.004]
- Bookstein F L. 1989. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11 (6): 567-585 [DOI: 10.1109/34.24792]
- Burgos N, Cardoso M J, Guerreiro F, Veiga C, Modat M, McClelland J, Knopf A C, Punwani S, Atkinson D, Arridge S R, Hutton B F and Ourselin S. 2015. Robust CT synthesis for radiotherapy planning: application to the head and neck region//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich: Springer: 476-484 [DOI: 10.1007/978-3-319-24571-3_57]
- Buslaev A, Parinov A, Khvedchenya E, Iglovikov V I and Kalinin A A. 2018. Alumentations: fast and flexible image augmentations [EB/OL]. [2019-04-10]. <https://arxiv.org/pdf/1809.06839v1.pdf>
- Cao X H, Yang J H, Gao Y Z, Wang Q and Shen D G. 2018a. Region-adaptive deformable registration of CT/MRI pelvic images via learning-based image synthesis. *IEEE Transactions on Image Processing*, 27 (7): 3500-3512 [DOI: 10.1109/TIP.2018.2820424]
- Cao X H, Yang J H, Zhang J, Wang Q, Yap P T and Shen D G. 2018b. 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]
- Goodfellow I J, 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*. Montreal: MIT Press: 2672-2680
- Han X. 2017. MR-based synthetic CT generation using a deep convolutional neural network method. *Medical Physics*, 44 (4): 1408-1419 [DOI: 10.1002/mp.12155]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Huang G, Liu Z, van der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE: 4700-4708 [DOI: 10.1109/CVPR.2017.243]
- Huynh T, Gao Y Z, Kang J Y, Wang L, Zhang P, Lian J and Shen D G. 2016. Estimating CT image from MRI data using structured random forest and auto-context model. *IEEE Transactions on Medical Imaging*, 35 (1): 174-183 [DOI: 10.1109/TMI.2015.2461533]
- Johnson K A and Becker J A. 1999. The whole brain atlas [EB/OL]. [2019-04-10]. <http://www.med.harvard.edu/aanlib/home.html>
- Jolicœur-Martineau A. 2018. The relativistic discriminator: a key element missing from standard GAN [EB/OL]. [2019-04-10]. <https://arxiv.org/pdf/1807.00734.pdf>
- Kingma D P and Ba J. 2014. Adam: a method for stochastic optimization [EB/OL]. [2019-04-10]. <https://arxiv.org/pdf/1412.6980.pdf>
- 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: Curran Associates Inc: 84-90 [DOI: 10.1145/3065386]
- Maes F, Collignon A, Vandermeulen D, Marchal G and Suetens P. 1996. Multi-modality image registration by maximization of mutual information//*Proceedings of 1996 Workshop on Mathematical Methods in Biomedical Image Analysis*. San Francisco: IEEE: 14-22

- [DOI: 10.1109/MMBIA.1996.534053]
- Mao X D, Li Q, Xie H R, Lau R Y K and Wang Z. 2016. Multi-class generative adversarial networks with the L2 loss function[EB/OL]. [2019-04-10]. <https://arxiv.org/pdf/1611.04076v1.pdf>
- Mirza M and Osindero S. 2014. Conditional generative adversarial nets [EB/OL]. [2019-04-10]. <https://arxiv.org/pdf/1411.1784.pdf>
- Nie D, Trullo R, Lian J, Wang L, Petitjean C, Ruan S, Wang Q and Shen D G. 2018. Medical image synthesis with deep convolutional adversarial networks. *IEEE Transactions on Biomedical Engineering*, 65(12): 2720-2730 [DOI: 10.1109/TBME.2018.2814538]
- 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]
- Ratliff L J, Burden S A and Sastry S S. 2013. Characterization and computation of local Nash equilibria in continuous games//*Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. Monticello: IEEE: 917-924 [DOI: 10.1109/Allerton.2013.6736623]
- Roy S, Carass A, Shiee N, Pham D L and Prince J L. 2010. MR contrast synthesis for lesion segmentation//*Proceedings of 2010 IEEE International Symposium on Biomedical Imaging: from Nano to Macro*. Rotterdam: IEEE: 932-935 [DOI: 10.1109/ISBI.2010.5490140]
- Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A and Chen X. 2016. Improved techniques for training GANs [EB/OL]. [2019-04-10]. <https://arxiv.org/pdf/1606.03498.pdf>
- Sotiras A, Davatzikos C and Paragios N. 2013. Deformable medical image registration: a survey. *IEEE Transactions on Medical Imaging*, 32(7): 1153-1190 [DOI: 10.1109/TMI.2013.2265603]
- Tai Y, Yang J, Liu X M and Xu C Y. 2017. MemNet: a persistent memory network for image restoration//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice: IEEE: 4549-4557 [DOI: 10.1109/ICCV.2017.486]
- Tustison N J, Avants B B, Cook P A, Zheng Y J, Egan A, Yushkevich P A and Gee J C. 2010. N4ITK: improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6): 1310-1320 [DOI: 10.1109/TMI.2010.2046908]
- Vercauteren T, Pennec X, Perchant A and Ayache N. 2009. Diffeomorphic demons: efficient non-parametric image registration. *NeuroImage*, 45(S1): S61-S72 [DOI: 10.1016/j.neuroimage.2008.10.040]
- Xie D, Cao X H, Gao Y Z, Wang L and Shen D G. 2016. Estimating CT image from MRI data using 3D fully convolutional networks//*Proceedings of the 1st International Workshop, LABELS 2016, and Second International Workshop on Deep Learning and Data Labeling for Medical Applications*. Athens: Springer: 170-178 [DOI: 10.1007/978-3-319-46976-8_18]
- Ye C Y, Yang Z, Ying S H and Prince J L. 2015. Segmentation of the cerebellar peduncles using a random forest classifier and a multi-object geometric deformable model: application to spinocerebellar ataxia type 6. *Neuroinformatics*, 13(3): 367-381 [DOI: 10.1007/s12021-015-9264-7]
- Ye D H, Zikic D, Glocker B, Criminisi A and Konukoglu E. 2013. Modality propagation: coherent synthesis of subject-specific scans with data-driven regularization//*Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Nagoya: Springer: 606-613 [DOI: 10.1007/978-3-642-40811-3_76]
- Zhang Y L, Tian Y P, Kong Y, Zhong B N and Fu Y. 2018. Residual dense network for image super-resolution//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE: 2472-2481 [DOI: 10.1109/CVPR.2018.00262]

作者简介



王丽芳, 1977年生, 女, 副教授, 主要研究方向为机器视觉、大数据处理、医学图像处理。
E-mail: wsm2004@nuc.edu.cn



张程程, 通信作者, 女, 硕士研究生, 主要研究方向为医学图像配准与机器学习。
E-mail: 1738490350@qq.com

秦品乐, 男, 副教授, 主要研究方向为机器视觉、大数据处理、3维重建。E-mail: QPL@nuc.edu.cn

蔺素珍, 女, 教授, 主要研究方向为图像处理、文物虚拟修复。E-mail: lsz@nuc.edu.cn

高媛, 女, 副教授, 主要研究方向为机器视觉、大数据处理、3维重建。E-mail: yuan_g@126.com

窦杰亮, 男, 硕士研究生, 主要研究方向为医学图像融合与机器学习。E-mail: liangdjsha@163.com