



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

中国图象学报 中国图形学报

2022
03
VOL.27

ISSN1006-8961
CN11-3758/TB



医学图像及临床应用



第27卷第3期（总第311期）
2022年3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

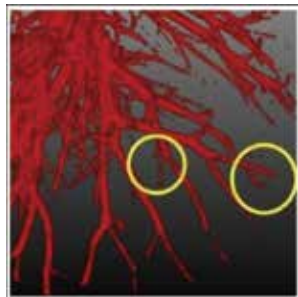
Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

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

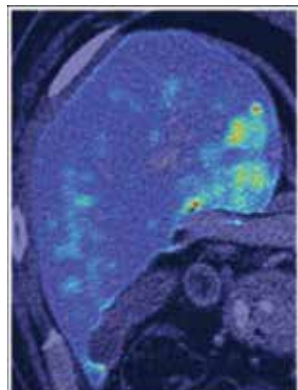
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

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

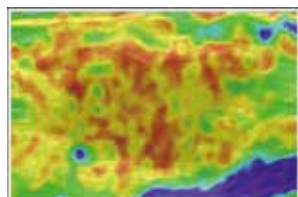
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



慢阻肺患者CT图像中肺内血管分割及定量分析(第0762页)



基于特征选择与残差融合的肝肿瘤分割模型(第0838页)



面向乳腺超声计算机辅助诊断的两阶段深度迁移学习(第0898页)

专刊封面展示了上海科技大学生物医学工程学院和上海联影智能医疗科技有限公司合作的脑影像智能分析方向研究成果。

《中国图象图形学报》医学图像及临床应用专刊简介

沈定刚, 刘天明, 周涛, 夏勇, 白相志, 王乾, 李刚, 徐寿平, 刘昊, 韩向娣 0653

综述

中国医学影像人工智能20年回顾和展望

蒋希, 袁奕宣, 王雅萍, 肖振祥, 朱美芦, 陈泽华, 刘天明, 沈定刚 0655

迁移学习在医学图像分类中的研究进展

黎英, 宋佩华 0672

生成对抗式网络及其医学影像应用研究综述

张颖麟, 胡衍, 东田理沙, 刘江 0687

基于深度学习的心脏磁共振影像超分辨率前沿进展

李书林, 冯朝路, 于鲲, 刘鑫, 江鑫, 赵大哲 0704

CT图像肺及肺病变区域分割方法综述

冯龙锋, 陈英, 周滔辉, 胡菲, 易珍 0722

计算机断层扫描图像

结构图注意力网络的新冠肺炎轻重症诊断

刘彦北, 李赫南, 张长青, 肖志涛, 张芳, 陈英, 高耀宗, 石峰, 单飞, 沈定刚 0750

慢阻肺患者CT图像中肺内血管分割及定量分析

赵宏, 李璋, 张杰华, 王琨, 孙家兴, 廖玺铭, 阎昱升, 钟正, 张鑫, 孙健, 于起峰, 葛俊辉 0762

基础信息保持和细节强化的胸部CT图像增强

张余, 张顺利, 白相志, 张利 0774

面向小样本股骨骨折分型的多视角注意力融合方法

张亚东, 汪玲, 兰海, 翟禹樵, 程洪 0784

三维多尺度嵌套U结构CT影像肺结节检测

许正奎, 张少敏, 支力佳, 周涛 0797

融合多尺度残差和注意力机制的特发性肺纤维化进展预测

陈舞, 孙军梅, 李秀梅 0812

COVID-19肺部CT图像多尺度编解码分割

陆倩杰, 柏正尧, 樊圣澜, 周雪, 许祝 0827

基于特征选择与残差融合的肝肿瘤分割模型

乔伟晨, 黄冕, 刘利军, 黄青松 0838

磁共振图像

面向多模态MRI脑胶质瘤区域三维分割与生存期预测的级联U-Net网络

余力, 刘霄雪, 闫朝阳, 李建瑞, 张志强, 黄轶枢, 徐军 0850

心脏动态MRI图像分割的时空多尺度网络

徐佳陈, 肖志勇 0862

融合跨阶段深度学习的脑肿瘤MRI图像分割

夏峰, 邵海见, 邓星 0873

融入平滑组稀疏化的脑部MRI图像分类

黄帅辉, 王金凤 0885

超声图像

面向乳腺超声计算机辅助诊断的两阶段深度迁移学习

贡荣麟, 施俊, 周玮珺, 汪程 0898

结合注意力机制的乳腺双模态超声分类网络

赵绪, 龚勋, 樊琳, 罗俊 0911

中医图像

视觉Transformer与多特征融合的脑卒中检测算法

赵琛琦, 王华虎, 赵涓涓, 冀伦文, 王麒达, 李慧芝, 赵紫娟 0923

多分支深度特征融合的中医脑卒中辅助诊断

王麒达, 冀伦文, 强彦, 王华虎, 赵琛琦, 李慧芝, 赵紫娟 0935

研究应用

多区域融合注意力网络模型下的核性白内障分类

章晓庆, 肖尊杰, Risa Higashita, 陈婉, 胡衍, 袁进, 刘江 0948

面向高度近视条纹损伤的深监督特征聚合网络

谭晓, 刁逸超, 陈新建, 石霏, 樊莹, 谢嘉受, 朱伟芳 0961

结合目标检测与匹配修正的手腕骨兴趣区域提取

毛科技, 汪敏豪, 陈立建, 陆伟, 武坤秀, 陈庆章, 赵小敏 0973

面向旷场实验视频分类的特征拼接矩阵学习方法

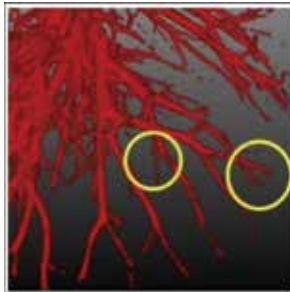
肖照林, 杨志林, 刘欢, 金海燕 0988

基于图像的介入器械前端变形建模与受力估计

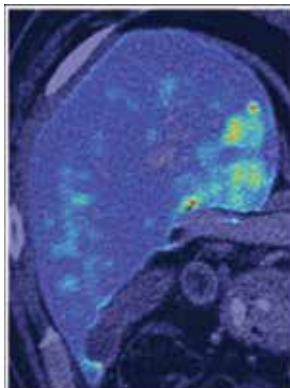
任龙飞, 孟偲, 刘博, 白相志, 周付根 1001

CONTENTS

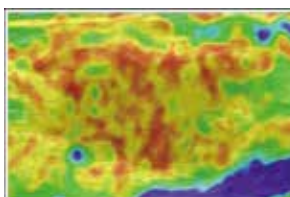
JOURNAL OF IMAGE AND GRAPHICS



Segmentation and quantitative analysis of intrapulmonary vasculature in CT images from COPD patients(P0762)



Feature selection and residual fusion segmentation network for liver tumor(P0838)



Two-stage deep transfer learning for human breast ultrasound computer-aided diagnosis(P0898)

Review

- A 20-year retrospect and prospect of medical imaging artificial intelligence in China
Jiang Xi, Yuan Yixuan, Wang Yaping, Xiao Zhenxiang, Zhu Meilu, Chen Zehua, Liu Tianming, Shen Dinggang 0655
- Review of transfer learning in medical image classification
Li Ying, Song Peihua 0672
- A review of generative adversarial networks and the application in medical image
Zhang Yinglin, Hu Yan, Risa Higashita, Liu Jiang 0687
- Critical review of human cardiac magnetic resonance image super resolution reconstruction based on deep learning method
Li Shulin, Feng Chaolu, Yu Kun, Liu Xin, Jiang Xin, Zhao Dazhe 0704
- Review of human lung and lung lesion regions segmentation methods based on CT images
Feng Longfeng, Chen Ying, Zhou Taohui, Hu Fei, Yi Zhen 0722

Computerd Tomography Image

- Diagnosis of COVID-19 by using structural attention graph neural network
Liu Yanbei, Li Henan, Zhang Changqing, Xiao Zhitao, Zhang Fang, Wei Ying, Gao Yaozong, Shi Feng, Shan Fei, Shen Dinggang 0750
- Segmentation and quantitative analysis of intrapulmonary vasculature in CT images from COPD patients
Zhao Hong, Li Zhang, Zhang Jiehua, Wang Kun, Sun Jiaxing, Liao Ximing, Yan Yusheng, Zhong Zheng, Zhang Xin, Sun Jian, Yu Qifeng, Ge Junhui 0762
- Human chest CT image enhancement based on basic information preservation and detail enhancement
Zhang Yu, Zhang Shunli, Bai Xiangzhi, Zhang Li 0774
- Multi-view attention fusion method for few-shot femoral fracture classification
Zhang Yadong, Wang Ling, Lan Hai, Zhai Yuqiao, Cheng Hong 0784
- Detection of pulmonary nodules in three-dimensional multiscale nested U-structure computed tomography images
Xu Zhengxi, Zhang Shaomin, Zhi Lijia, Zhou Tao 0797
- Multi-scale residual and attention mechanism fusion based prediction for the progression of idiopathic pulmonary fibrosis
Chen Wu, Sun Junmei, Li Xiumei 0812
- Multiscale codec network based CT image segmentation for human lung disease derived of COVID-19
Lu Qianjie, Bai Zhengyao, Fan Shenglan, Zhou Xue, Xu Zhu 0827
- Feature selection and residual fusion segmentation network for liver tumor
Qiao Weichen, Huang Mian, Liu Lijun, Huang Qingsong 0838

Magnetic Resonance Image

- Brain glioma tumor segmentation and survival prediction from multi-modality MRIs via cascaded U-Net
Yu Li, Liu Xiaoxue, Yan Chaoyang, Li Jianrui, Zhang Zhiqiang, Huang Yunzhi, Xu Jun 0850
- A spatio-temporal multi-scale network for cardiac dynamic MRI image segmentation
Xu Jiachen, Xiao Zhiyong 0862
- Cross-stage deep-learning-based MRI fused images of human brain tumor segmentation
Xia Feng, Shao Haijian, Deng Xing 0873
- Brain MRI image classification incorporating smooth group sparseness
Huang Shuaihui, Wang Jinfeng 0885

Ultrasound Image

- Two-stage deep transfer learning for human breast ultrasound computer-aided diagnosis
Gong Ronglin, Shi Jun, Zhou Weijun, Wang Cheng 0898
- Attention-based networks of human breast bimodal ultrasound imaging classification
Zhao Xu, Gong Xun, Fan Lin, Luo Jun 0911

Traditional Chinese Medical Image

- Cerebral stroke detection algorithm for visual Transformer and multi-feature fusion
Zhao Chenqi, Wang Huahu, Zhao Juanjuan, Ji Lunwen, Wang Qida, Li Huizhi, Zhao Zijuan 0923
- Multi-branch deep feature fusion method for traditional Chinese medicine intervened Human cerebral stroke aided diagnosis
Wang Qida, Ji Lunwen, Qiang Yan, Wang Huahu, Zhao Chenqi, Li Huizhi, Zhao Zijuan 0935

Research and Application

- Nuclear cataract classification based on multi-region fusion attention network model
Zhang Xiaoqing, Xiao Zunjie, Risa Higashita, Chen Wan, Hu Yan, Yuan Jin, Liu Jiang 0948
- Deep-supervision and feature-aggregation network for linear lesion segmentation of high myopia
Tan Xiao, Diao Yichao, Chen Xinjian, Shi Fei, Fan Ying, Xie Jiamin, Zhu Weifang 0959
- Hand-wrist region of interest extraction based on object detection and matching correction
Mao Keji, Wang Minhao, Chen Lijian, Lu Wei, Wu Kunxiu, Chen Qingzhang, Zhao Xiaomin 0973
- A spliced feature matrices learning method for open field test video classification
Xiao Zhaolin, Yang Zhilin, Liu Huan, Jin Haiyan 0988
- Image-based deformation modeling and force estimation of the front end of vascular interventional surgical apparatus
Ren Longfei, Meng Cai, Liu Bo, Bai Xiangzhi, Zhou Fugen 1001

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2022)03-0687-17

论文引用格式: Zhang Y L, Hu Y, Risa Higashita and Liu J. 2022. A review of generative adversarial networks and the application in medical image. Journal of Image and Graphics, 27(03): 0687-0703 (张颖麟, 胡衍, 东田理沙, 刘江. 2022. 生成对抗式网络及其医学影像应用研究综述. 中国图象图形学报, 27(03): 0687-0703 [DOI:10.11834/jig.210247])

生成对抗式网络及其医学影像应用研究综述

张颖麟¹, 胡衍¹, 东田理沙^{1,2}, 刘江^{1*}

1. 南方科技大学, 深圳 518055; 2. TOMEY 株式会社, 名古屋 451-0051, 日本

摘要: 生成对抗式网络(generative adversarial network, GAN)由负责学习数据分布的生成器和负责鉴别样本真伪的判别器构成,二者在相互对抗过程中互相学习逐渐变强。该网络模型使深度学习方法可以自动学习损失函数,减少了对专家知识的依赖,已经广泛应用于自然图像处理领域,对解决医学影像处理的相关瓶颈问题亦具有巨大应用前景。本文旨在找到生成对抗式网络与医学影像领域面临挑战的结合点,通过分析已有工作对未来研究方向进行展望,为该领域研究提供参考。1)阐述了生成对抗式网络的基本原理,从任务拆分、条件约束以及图像到图像的翻译等角度对其衍生模型进行分析回顾;2)对生成对抗式网络在医学影像领域中的数据增广、模态迁移、图像分割以及去噪等方面的应用进行回顾,分析各方法的优缺点与适用范围;3)对现有图像生成质量评估方法进行小结;4)总结生成对抗式网络在医学影像领域的研究进展,并结合该领域问题特性,指出现有理论应用存在的不足与改进方向。生成对抗式网络提出以来,理论不断完善,在医学影像的处理应用中也取得了长足发展,但仍然存在一些亟待解决的问题,包括3维数据合成、几何结构合理性保持、无标记和未配对数据使用以及多模态数据交叉应用等。

关键词: 生成对抗式网络(GAN);医学影像;深度学习;数据增广;模态迁移;图像分割;图像去噪

A review of generative adversarial networks and the application in medical image

Zhang Yinglin¹, Hu Yan¹, Risa Higashita^{1,2}, Liu Jiang^{1*}

1. Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China;

2. TOMEY Corporation, Nagoya 451-0051, Japan

Abstract: The generative adversarial network (GAN) consists of a generator based on the data distribution learning and an identified sample's authenticity discriminator. They learn from each other gradually in the process of confrontation. The network enables the deep learning method to learn the loss function automatically and reduces expertise dependence. It has been widely used in natural image processing, and it is also a promising solution for related problems in medical image processing field. This paper aims to bridge the gap between GAN and specific medical field problems and point out the future improvement directions. First, the basic principle of GAN is issued. Secondly, we review the latest medical images

收稿日期: 2021-04-06; 修回日期: 2021-05-26; 预印本日期: 2021-06-02

* 通信作者: 刘江 liuj@sustech.edu.cn

基金项目: 广东省教育厅普通高校重点领域专项项目(2020ZDZX3043); 深圳市基础研究项目(JCYJ20200109140820699); 深圳市稳定支持项目(20200925174052004); 广东省重点实验室项目(2020B12101001)

Supported by: Guangdong Provincial Department of Education (2020ZDZX3043); Shenzhen Natural Science Fund (JCYJ20200109140820699); Shenzhen Stable Support Plan Program (20200925174052004); Guangdong Provincial Key Laboratory (2020B12101001)

research on data augmentation, modality migration, image segmentation, and denoising; analyze the advantage and disadvantage of each method and the scope of application. Next, the current quality assessment is summarized. At the end, the research development, issue, and future improvement direction of GAN on medical image are summarized. GAN theoretical study focus on three aspects of task splitting, introducing conditional constraints and image-to-image translation, which effectively improved the quality of the synthesized image, increased the resolution, and allowed more manipulation across the image synthesis process. However, there are some challenges as mentioned below: 1) Generate high-quality, high-resolution, and diverse images under large-scale complex data sets. 2) The manipulation of synthesized image attributes at different levels and different granularities. 3) The lack of paired training data and the guarantee of image translation quality and diversity. GAN application study in data augmentation, modality migration, image segmentation, and denoising of medical images has been widely analyzed. 1) The network model based on the Pix2pix basic framework can synthesize additional high-quality and high-resolution samples and improve the segmentation and classification performance based on data augmentation effectively. However, there are still problems such as insufficient synthetic sample diversity, basic biological structures maintenance difficulty, and limited 3D image synthesis capabilities. 2) The network model based on the CycleGAN basic framework does not require paired training images. It has been extensively analyzed in modality migration, but may lose the basic structure information. The current research on structure preservation in modality migration limits in the fusion of information, such as edges and segmentation. 3) Both the generator and the discriminator can be fused with the current segmentation model to improve the performance of the segmentation model. The generator can synthesize additional data, and the discriminator can guide model training from a high-level semantic level and make full use of unlabeled data. However, current research mainly focuses on single-modality image segmentation. 4) GAN application in image denoising can reconstruct normal-dose images from low-dose images, reducing the radiation impact suffered by patients. The critical issues of GAN in medical image processing are presented as follows: 1) Most medical image data is three-dimensional, such as MRI (magnetic resonance imaging) and CT (computed tomography), etc. The improvement of the synthesis quality and resolution of the three-dimensional data is a critical issue. 2) The difficulty in ensuring the diversity of synthesized data while keeping its basic geometric structure's rationality. 3) The question on how to make full use of unlabeled and unpaired data to generate high-quality, high-resolution, and diverse images. 4) The improvement of algorithms' cross-modality generalization performance, and the effective migration of different modality data. Future research should focus on the issues as following: 1) To optimize network architecture, objective function, and training methods for 3D data synthesis, improving model training stability, quality, resolution, and diversity of 3D synthesized images. 2) To further promote the prior geometric knowledge integration with GAN. 3) To take full advantage of the GAN's weak supervision characteristics. 4) To extract invariant features via attribute decoupling for good generalization performance and achieve attribute control at different levels, granularities, and needs in the process of modality migration. To conclude, ever since GAN was proposed, its theory has been continuously improved. A considerable evolution in medical image applications has been sorted out, such as data augmentation, modality migration, image segmentation, and denoising. Some challenging issues are still waiting to be resolved, including three-dimensional data synthesis, geometric structure rationality maintenance, unlabeled and unpaired data usage, and multi-modality data application.

Key words: generative adversarial network (GAN); medical image; deep learning; data augmentation; modality migration; image segmentation; image denoising

0 引言

生成对抗式网络 (generative adversarial network, GAN) 已经广泛应用于图像合成 (Wan 等, 2019; Zhang 等, 2019a)、图像重建 (Li 等, 2018; Zhao 等, 2019)、风格迁移 (Liu 等, 2019; Mejjati 等, 2018) 和

图像修复 (Wang 等, 2017; Wang 等, 2020) 等自然图像处理任务, 并用于解决医学影像中的相关问题。深度学习算法在图像分类识别 (He 等, 2016; Huang 等, 2016; Krizhevsky 等, 2012)、语义分割 (Badrinathan 等, 2015; Shelhamer 等, 2017) 和语言处理 (Bahdanau 等, 2016; Sutskever 等, 2014) 等领域取得了显著成就, 这主要得益于计算性能的提升、网络结

构的优化以及大规模的数据集。但是由于数据标注昂贵、类别不平衡、域移以及隐私保护等问题,在医学影像中的应用十分有限。首先,医学影像的数据标注需要医学和医学成像方面的专业知识,通常由专业的医学专家操作,因此标注过程十分耗时且非常昂贵(Greenspan 等,2016)。其次,由于健康组织通常占主导地位且呈高度重复模式,通常会过度代表正常类别,这种类别不平衡问题会导致模型收敛缓慢或者过拟合(Ravi 等,2017)。此外,来自不同设备或者具有不同参数设置的同一设备通常具有不同的噪声分布、图像特征以及形状等。例如,在用于眼部成像的 OCT(optical coherence tomography)中,不同设备使用了不同波长的激光以及不同的扫描频率等(Schmitt 等,1999;Xiang 等,1998),因此,基于某一设备成像数据训练得到的模型难以很好地适应其他设备。最后,在建立和共享数据集的过程中需要充分考虑患者的隐私安全问题(Dysmorphology Subcommittee of the Clinical Practice Committee and American College of Medical Genetics,2000),这给创建大规模医疗数据集带来了很大挑战。生成对抗式网络的出现为解决上述问题提供了新的思路。Frid-Adar 等人(2018)使用 GAN 来扩展具有不同疾病的医学脑 CT(computed tomography)图像的数据集,有效提升了分类性能。同时,Frid-Adar 等人(2018)为了保证跨模态数据合成过程中的形状一致性,在 Zhu 等人(2017)的基础上引入在线分割任务,使得生成器与分割器的优化过程相辅相成。实验结果表明,对于 CT 和 MRI(magnetic resonance imaging)合成图像的 shape 一致性评价指标 S-score 的得分,从引入在线分割任务之前的 66.8% 和 67.5% 分别提升至引入后的 69.2% 和 69.6%。

国内已有一些综述性工作对生成对抗式网络的发展进行分析总结,但没有针对医学影像处理领域的相关工作进行回顾。曹仰杰等人(2018)分析了 GAN 模型的改进,重点介绍了其在计算机视觉方面的几类突出应用,如图像修复、图像合成以及风格迁移等。陈佛计等人(2021)从模型架构、目标函数和 GAN 在训练过程中存在的问题,以及如何应对模式崩溃等角度对 GAN 的相关研究与应用进行分析回顾。已有一些医学影像处理领域的深度学习综述,但没有讨论生成对抗式网络的相关工作。施俊等人(2020)介绍了深度学习在病灶检测、图像分割

以及辅助诊断方面的应用。张航等人(2020)针对深度学习方法在婴儿脑 MRI 图像分割中的最新进展进行了系统总结。

本文通过对医学影像处理相关瓶颈问题以及生成对抗式网络的深入分析,旨在找到两者的结合点,以及未来的改进方向,也为该领域相关研究人员提供参考。

1 GAN 概述

1.1 基本原理

生成对抗式网络包含生成器 G 和判别器 D 两部分,如图 1 所示。

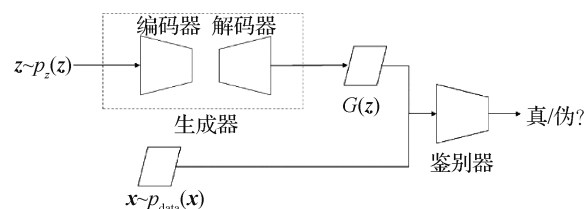


图 1 生成对抗式网络

Fig. 1 Generative adversarial network

生成器 G 用于学习数据分布,判别器 D 用于判别样本是来自真实数据还是生成器。二者进行值函数为 $V(D, G)$ 的二人博弈游戏,具体为

$$\min_G \max_D V(D, G) = E_{x \sim p_{\text{data}}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

式中, z 表示噪声, $p_z(z)$ 是噪声分布, x 是真实数据, $p_{\text{data}}(x)$ 是真实数据分布。判别器 D 的目标是尽可能区别真实数据与生成样本,即最大化 $E_{x \sim p_{\text{data}}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$ 。生成器 G 的目标是使生成数据 $G(z)$ 尽可能接近来自真实数据分布的样本,即最小化 $E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$ 。但是,在实际中由于在开始阶段 G 的合成能力非常差,因此会导致 $\log(1 - D(G(z)))$ 饱和,所以通常通过最大化 $E_{z \sim p_z(z)} [\log D(G(z))]$ 来优化生成器。在训练过程中,生成器 G 和判别器 D 通过交替优化的方式互相学习,最终,生成器会学习到数据的分布,而判别器则无法鉴别样本究竟来自真实数据还是生成器。二者在相互对抗的过程中互相学习逐渐变强。

1.2 衍生模型

生成对抗式网络模型提出(Goodfellow 等,

2014)以来,进行了许多改进与扩展,产生了多种衍生模型,大致可分为任务拆分模型、条件约束模型以及图像到图像的翻译模型。现有代表性的衍生模型生成器的计算流程与结构如图2所示,其中, $z \sim$

$p_z(z)$ 表示噪声分布, I_0 表示输入图像, I 表示输出图像, c 表示条件变量, G_i 表示第 i 个生成器, En_i 表示第 i 个编码器, De_i 表示第 i 个解码器, L_1 表示第一范数损失函数, D 表示判别器。

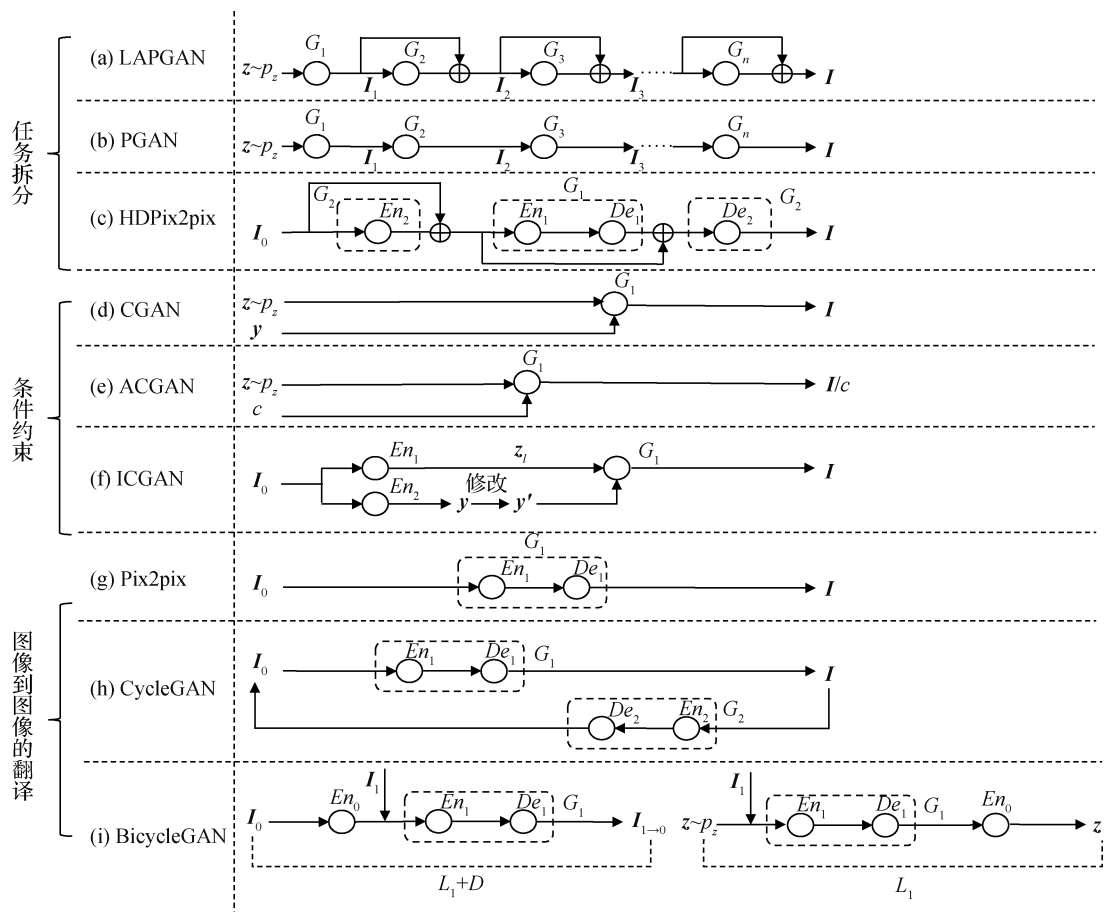


图2 GAN 衍生模型生成器的计算流程与结构

Fig. 2 The calculation process and structure of the GAN derivative generator

((a) LAPGAN; (b) PGAN; (c) HDPix2pix; (d) CGAN; (e) ACGAN; (f) ICGAN;
(g) Pix2pix; (h) CycleGAN; (i) BicycleGAN)

任务拆分指将一个困难任务拆分成多个相对简单的子任务,从而使 GAN 的训练变得更加稳定,提升模型性能。Denton 等人(2015)提出一种基于 Laplacian 金字塔的生成对抗式模型 LAPGAN (Laplacian GAN),将随机向量作为输入,在金字塔的每个层上训练一对生成器和判别器,每层上的生成器仅负责学习上一层到该层的差分图像 $I_n - I_{n-1}$,如图2(a)所示,该方法能够合成尺寸为 32×32 像素和 64×64 像素的图像。高分辨率图像的生成很困难,因为更高的分辨率使判别器更易于区分生成的图像和真实图像,从而极大放大了梯度问题。Karras 等人(2018)提出了一种生成对抗式网络

的训练方法 PGGAN (progressive GAN),关键思想是逐渐增加生成器和判别器,从低分辨率开始,随着训练的进行逐渐添加新层,对越来越精细的细节建模,将合成图像的尺寸提高到 1024×1024 像素,如图2(b)所示。Wang 等人(2018)提出一种多尺度网络架构 HDPix2pix,如图2(c)所示,它的生成器包括一个全局生成器 G_1 和一个局部增强生成器 G_2 ,在训练过程中,首先训练全局生成器,然后训练局部增强器,最后,对它们一起进行微调,该方法可生成 2048×1024 像素具有视觉吸引力的结果。

条件生成模型可通过基于条件信息指导数据生成过程。Mirza 和 Osindero(2014)通过简单地对生

成器和判别器馈送条件信息 y 首先构造了生成对抗式网络的条件版本 CGAN(condition GAN), 如图 2(d) 所示, 并证明了该模型可以生成以类别标签为条件的 MNIST(mixed national institute of standards and technology database) 数字。Odena 等人(2017) 提出 ACGAN(auxiliary classifier GAN) 以辅助分类器的形式引入了类条件信息 c , 如图 2(e) 所示, 尽管这种方法与已有模型没有很大区别, 但是这样的改进在 ImageNet 数据合成任务中产生了不错结果并且可以使训练更加稳定。Perarnau 等人(2016) 提出 ICGAN(invertible conditional GAN), 如图 2(f) 所示, 由编码器和 CGAN 两个关键部分组成, 可以将真实图像 x 压缩为潜在表示向量 z_i 和条件向量 y , 并且允许使用条件向量 y 控制生成样本的复杂属性。在 MNIST 和 CelebA 数据集上的实验表明, 该模型可以控制合成数字的内容和风格, 控制合成人脸图像的肤色、表情以及头发等更为复杂的属性。Olszewski 等人(2020) 提出一种交互式方法用于人脸图像中毛发的编辑与合成, 该研究工作通过一个交互界面使用户可以通过稀疏简单地引导笔触定义所需要的目标发式的结构和色彩, 但该方法的属性控制主要集中在像素层面。

图像到图像的翻译属于条件生成对抗式网络的一种特殊形式, 由于该方向有大量研究工作, 因此, 本文将其单列为一个子领域, 并对其相关工作进行回顾。图像到图像的翻译由 Isola 等人(2017) 首先提出, 类似语言翻译任务中可以将对同一事物的描述由英语翻译成法语一样, 他们认为在图像合成任务中, 可以将 RGB 图像、边缘图以及语义标签图等理解为对同一事物的不同描述, 因此这些不同类型的图像之间可以相互转换。他们提出了将条件对抗网络作为图像到图像翻译问题的通用解决方案 Pix2pix, 生成器负责学习从输入图像到输出图像的映射, 如图 2(g) 所示。此外, 该方法通过判别器学习损失函数来训练该映射, 并且证明了这种方法可以有效地从标签图合成照片等。但是该解决方案存在两个固有缺陷: 1) 需要成对地训练数据; 2) 合成数据多样性不足。Zhu 等人(2017) 提出了 CycleGAN, 如图 2(h) 所示, 通过引入循环一致性约束令 $G_2(G_1(I_0)) \approx I_0$, 缓解了 Pix2pix 中需要配对图像的问题, 实验结果证明了该方法在涉及颜色和纹

理变化翻译任务上的有效性, 但对需要更改几何结构的任务会发生失效, 此外作者指出未配对方法可达到的结果与配对训练数据可达到的结果之间仍然存在差距。为了解决图像翻译任务中模式崩溃问题, 并产生更多样化的结果, Zhu 等人(2018) 提出 BicycleGAN, 如图 2(i) 所示, 通过组合多个目标来鼓励潜在空间 $z \rightarrow I' \rightarrow z'$ 和输出空间 $I \rightarrow z \rightarrow I'$ 之间的双向映射, 防止训练期间从潜码到输出的多对一映射(也称模式崩溃问题), 实验结果表明, 该方法可以在广泛的图像到图像的翻译问题中产生视觉上有吸引力的结果, 且相比于其他基准更加多样化。

上述衍生模型生成器的改进、优缺点及适用场景如表 1 所示。

1.3 面临的挑战

生成对抗式网络在许多领域表现了出色性能, 但是仍然存在一些亟待解决的挑战。

1) 生成高质量、高分辨率并具有多样性的图像。生成对抗式网络的训练过程是动态的, 并且对于参数优化以及模型体系结构等各方面都是十分敏感的。已有的大量研究提升了模型训练的稳定性, 改善了图像合成的质量和分辨率, 但是从复杂的数据集(例如 ImageNet) 生成高质量、多样的图像仍然十分困难。在大规模复杂 ImageNet 数据集上以 128×128 像素分辨率进行训练条件下, 将最新的先进模型进行性能比较, Wu 等人(2020) 提出的 LOGAN(latent optimisation for generative adversarial networks) 取得了 IS(inception score) = 148 的最佳成绩, Brock 等人(2019) 提出的最新先进图像生成模型取得的 IS 为 124.5。然而, 相比于实际数据 IS = 233 的结果仍然存在较大差距。

2) 根据实际需求、结合用户意图以控制图像的生成过程。控制生成图像的特定属性在实际应用中存在广泛需求, 例如, 改变人脸图像中的脸部表情、鸟类图像中鸟喙的长度, 以及嫌疑人面部毛发。这要求在图像生成过程中, 按照用户意图仅更改特定属性, 而保持其他属性不变, 然而, 已有研究在属性解耦以及控制方面仍然存在许多不足。Karras 等人(2019) 使用的特征解耦主要是基于不同尺度层级下的特征操作, 以实现在保留高层级特征的同时更改低层级特征, 但是难以单独更改特定属性。

表 1 生成对抗式网络的衍生模型
Table 1 Derivative model for generating adversarial networks

类型	方法	改进	优点	缺点	应用场景
任务拆分模型	LAPGAN	基于 Laplacian 金字塔框架, 改善了合成图像质量, 以从粗到细方式逐步生成图像, 每层仅负责学习上一层到该层的差分图像。	能够合成尺寸为 32×32 像素和 64×64 像素的图像。	无法稳定地合成高质量、高分辨率的图像。	由随机噪声合成图像。
	PGGAN	逐渐增加生成器和判别器, 从低分辨率开始, 随着训练的进行逐渐添加新层, 对越来越精细的细节建模。	将合成图像尺寸提高至 $1\,024 \times 1\,024$ 像素。	合成质量尤其语义信息无法达到真实图像效果, 且无法实现端到端的训练。	由随机噪声合成高分辨率图像。
	HDPix2pix	提出一种由粗到细的局部和全局相结合的架构, 首先训练全局生成器, 然后训练局部增强器, 最后联合微调。	可生成 $2\,048 \times 1\,024$ 像素的具有视觉吸引力的图像。	需要带有语义分割标签的成对数据进行训练。	由语义分割标签合成高质量、高分辨率图像。
条件生成模型	CGAN	对生成器和判别器馈送类别条件信息 y 。	可对 MINIST 数字生成结果的类别进行控制。	研究内容过于简单。	图像合成。
	ACGAN	以辅助分类器的形式引入类别条件信息。	提高了训练稳定性。	相对于 CGAN 没有太大改进。	图像合成。
	ICGAN	将真实图像 x 压缩为潜在表示向量 z 和条件向量 y , 并允许使用 y 控制生成样本的复杂属性。	允许根据任意属性重建和修改人脸的真实图像。	缺少人机交互, 无法按照用户意图进行特定修改。	基于特定属性控制人脸图像合成。
	Olszewski 等人(2020)	提出一种交互式图像合成方法。	允许用户显式地修改合成图像以符合目标需求。	无法从语义层面控制图像生成。	交互式图像合成。
图像到图像的翻译模型	Pix2pix	提出一种基于条件生成对抗式网络的图像到图像翻译问题的通用解决方案。	成功用于图像重建、分割、图像着色和图像合成等多种任务。	需要成对的数据进行模型训练。	图像到图像翻译的各类应用场景。
	CycleGAN	提出一种在没有成对数据情况下将图像从源域转换到目标域的方法。	无需成对训练数据。	缺少对图像内容的显式监督, 在与几何结构相关任务上会失效。	无成对数据的图像到图像翻译任务。
	BicycleGAN	通过组合多个目标来鼓励潜在空间和输出空间之间的双向映射。	避免从潜码到输出的多对一映射, 并产生更多样化的结果。	缺少基于语义和属性信息对潜在空间的深入研究, 且需要成对训练数据。	图像到图像翻译任务。

Nguyen-Phuoc 等人(2019)将姿态从身份属性中解耦出来,但是需要一个特定的几何架构,而这并不适合其他的属性。Alharbi 和 Wonka(2020)将空间不变编码与空间可变编码分开,并基于网格的结构控制特定区域的合成,然而该方法对于不同粒度层面的控制,需要进行不同的调整,也无法实现前景与背景的解耦。

3) 在无成对训练数据情况下,图像到图像翻译的解决方案。在实际应用中,成对的训练数据通常难以获得,例如,艺术作品的风格迁移、季节转换以及物体变形等,因此,基于无成对训练数据集的图像到图像翻译的解决方案具有重要研究意义。Zhu 等

人(2017)通过引入循环一致性损失,减少了可能的映射空间,实现无成对数据下的模型训练,但是该方法仅能在涉及颜色和纹理变化的翻译任务上取得不错的性能,无法应对需要更多变化和极端的转换的场景,尤其是几何结构的变化。Almahairi 等人(2018)提出了 Augment CycleGAN 用于从未配对的数据中学习域之间的多对多映射,以应对更多变化的场景,但是其合成图像的保真度较低。

1.4 小结

生成对抗式网络基于两人博弈原理,无需显式地指定学习目标,间接地对未知分布进行建模,在图像处理领域得到广泛应用。许多研究者在任务拆

分、引入条件约束以及图像到图像翻译3个方面进行了大量研究,有效改善了合成图像的质量,提高了分辨率,允许对图像的合成过程进行更多的控制,尤其是图像到图像的翻译作为一种通用框架广泛应用。但是,仍然存在一些挑战:1)在大规模复杂数据集下生成高质量、高分辨率并具有多样性的图像;2)根据实际需求、结合用户意图对图像生成过程在不同层面、不同粒度上对不同属性进行控制;3)在无成对训练数据的情况下,训练图像到图像翻译模型,并保证图像的翻译质量与多样性。

2 GAN 在医疗图像领域的应用

生成对抗式网络在医学影像领域的各类应用按应用目的可以归纳为数据增广、模态迁移、图像分割以及图像去噪等。数据增广可以通过学习训练数据分布以生成额外的样本,缓解由于标注困难、隐私保护以及类别不平衡等导致的医学影像数据匮乏问题。模态迁移可以通过学习图像到图像的映射关系,将源域的模态迁移到目标域,缓解由于域移导致的分割或分类模型泛化性能差的问题,并且可为疾病的诊断提供多视角信息。此外,生成对抗式网络的生成器和判别器对于提高基于深度学习的分割模型的性能均有重要意义。最后,生成对抗式网络还可以应用于去噪,将由较低剂量、较低分辨率或者较短的曝光时间等导致的有噪声的图像重建到近似理想的图像,同时保留图像中的重要细节。

2.1 数据增广

数据增广指通过学习训练数据的分布以生成额外的样本,从而扩大训练数据集规模,提升分类、分割等网络的性能。生成对抗式网络能够基于随机噪声或者是相关的条件先验合成新的数据,扩展原有的训练数据集。由于医疗数据的标注成本高昂、健康组织与病灶区域类别不平衡、并且医疗数据的使用涉及病患的隐私保护,因此,利用生成对抗式网络进行数据增广是一种有前景的解决方案。

1) 基于随机噪声的无条件 GAN 合成方法。Frid-Adar 等人(2018)利用生成对抗式网络变体 DCGAN 隐式地学习肝脏病变数据分布以合成尺寸为 64×64 像素的图像数据,然后利用其合成的数据扩展原始训练集,实验结果表明,通过添加合成数据扩展训练数据集,可以将仅使用传统数据增强的

评价指标灵敏度和特异性从 78.6% 和 88.4% 提升至 85.7% 和 92.4%。Kwon 等人(2019)结合自动编码器和 GAN 的优势提出一种新颖的 3D 生成对抗式模型,可以由随机向量生成尺寸为 $64 \times 64 \times 64$ 的 3D 脑 MRI 数据,缓解了模式崩溃、图像模糊以及训练不稳定的问题,结果表明,该方法可以生成多种类型和模态的逼真 3D 全脑图像。但是基于随机噪声的图像合成方法,一个主要缺点就是其合成图像的分辨率过小,而在很多疾病诊断中,通常需要观察组织的细微变化,这要求合成图像具有更高的分辨率。

2) 基于图像条件先验的 GAN 合成方法。图像到图像的翻译生成对抗式网络主要包括基于 Pix2pix 基础框架和基于 CycleGAN 的生成对抗式网络两大类。Pix2pix 的优点是合成图像质量高,基础生物结构保留完整,缺点是需要配准的图像对,如图 3 所示,其中, (x, y) 表示成对的血管结构图和真实的眼底图像, G_x 表示基于血管结构图 x 合成的眼底图像;CycleGAN 的优点是不需要成对的配准图像,但是其成像质量不如 Pix2pix,生物结构容易发生丢失。然而,医学影像处理领域中对生物结构的合理性有着严格要求,因此,在数据增广方面, Pix2pix 解决方案得到了广泛研究。

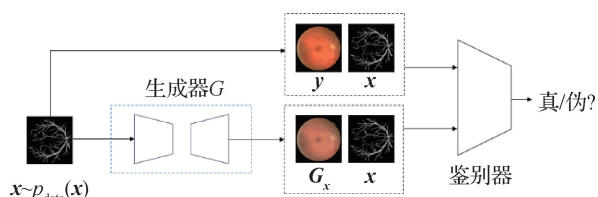


图3 图像到图像翻译的生成对抗式网络 Pix2pix

Fig. 3 Generative adversarial network

Pix2pix for image-to-image translation

Costa 等人(2018)将眼底图像的合成任务分解为二值血管树的分割与视网膜图像合成两个步骤,首先对输入视网膜图像进行分割得到二值血管树图像,然后以该二值血管树图像为条件先验,基于 Pix2pix 框架得到合成的视网膜图像。但是,该方法依赖于分割得到的二值血管树图像以生成新图像,限制了合成数据的多样性,并且合成算法会继承分割算法的潜在弱点。Zhang 等人(2019b)提出一种素描—绘制生成对抗式网络 SkrGAN (sketching-rendering unconditional GAN),该网络结构主要包括两个生成对抗式子模块,首先,第1个生成对抗式子模块负

责从随机噪声中生成高质量的结构草图,然后,第2个生成对抗式子模块基于该结构草图的指导渲染进一步生成具有逼真前景结构的图像,该方法在视网膜彩色眼底、胸部 X 光、肺部 CT 和脑部 MRI 等医疗影像的合成任务中取得了先进表现。Jin 等人(2018)开发了一种 3D 生成对抗式网络用于学习 3D 空间中的肺结节特性分布,以周围肺组织为条件来引导肺结节的生成,并通过多任务重建损失以生成减轻边界不连续伪影的高质量逼真结节,该方法可用于帮助克服获取医学图像中“边缘病例”数据不平衡的困难。Bissoto 等人(2018)基于 Pix2pixHD 框架,利用语义分割与实例分割结果作为先验知识,以合成皮肤病损图像,利用合成的图像扩展真实图像训练集,将皮肤癌分类网络的性能提升了 1%。Li 等人(2020)基于生成对

抗式网络模型,由非 COVID-19 (corona virus disease 2019) 数据合成 COVID-19 CT 扫描,并通过合成的数据提高了分类模型的识别和检测能力。

在医学图像数据增广任务中,生成对抗式网络通过合成数据有效提高了分割、分类等模型的性能。无条件数据合成可生成多种类型、逼真的图像,但训练难度大,并且难以生成高分辨率、高质量的图像。基于 Pix2pix 框架的生成对抗式网络在医学影像领域广泛使用,合成的 2 维图像质量以及分辨率较高,但受限于合成样本的多样性不足,尤其在基础结构方面,目前无法通过几何约束在有效保证生物结构合理性的同时,合成具有多样性的图像。此外,已有模型能合成的 3 维影像分辨率低、成像质量较差。上述数据增广模型的改进、优缺点及适用场景如表 2 所示。

表 2 数据增广应用
Table 2 Data augmentation application

类型	模型	改进	优点	缺点	应用场景
无条件合成	Frid-Adar 等人 (2018)	利用生成对抗式网络变体 DCGAN 隐式地学习肝脏病变数据分布。	相比传统的数据增强方法,提高了分类模型性能。	仅局限于 2D 影像数据,合成图像尺寸过小。	2D 影像数据合成。
	Kwon 等人 (2019)	结合自动编码器和 GAN 的优势提出一种 3D 生成对抗式模型。	可由随机向量生成 64×64×64 的 3D 脑 MRI 数据。	图像分辨率过小。	3D 影像数据合成。
条件合成	Costa 等人 (2018)	将眼底图像的合成任务分解成二值血管树的分割和图像渲染两阶段。	可保持合成图像基础几何结构的合理性。	限制了合成数据的多样性,且会继承分割算法的潜在弱点。	图像到图像的 2D 数据合成。
	Zhang 等人 (2019b)	将图像合成任务分解为草图和渲染两阶段,分别由两个 GAN 子模块负责。	从随机噪声中生成草图提高了图像合成的多样性。	需要对训练数据,且不能充分保证合成草图结构的合理性。	图像到图像的 2D 数据合成。
	Jin 等人 (2018)	提出一种 3D 生成对抗式网络以学习 3D 空间中的肺结节特性分布。	缓解了边缘肺结节数据过少的数据不平衡问题。	局限于小区域数据合成,依赖较多周围先验信息,分辨率低。	3D 影像数据合成。
	Bissoto 等人 (2018)	利用语义分割与实例分割结果作为先验知识,以合成皮肤病损图像,基于合成数据提高了皮肤癌分类性能。	能够合成包含细粒度细节的高分辨率(最大 1 024×512 像素)的皮肤病变图像。	需要含分割标签的数据,而实际应用中这类数据通常难以获得。	2D 高分辨率影像数据合成。

2.2 模态迁移

模态迁移解决的是一类视觉和图形问题,目标是通过包含来自不同域的配准或者未配准的成对图像训练集来学习输入图像与输出图像间的映射关系。多模态图像可以对目标组织进行全面检查,包括目标组织的确切定位和代谢活性、周围组织内的组织流动和功能变化以及导致最终疾病的病理组织

学变化,有助于更加准确地诊断疾病。然而,由于高昂的成本和复杂的采集过程,不同模态的相应图像元组的采集非常稀缺且昂贵,模态迁移为该问题提供了一个有前景的解决方案。

对于跨模态图像迁移任务,由于成对训练数据通常难以获得,因此,模态迁移领域中基于 CycleGAN 的解决方案得到广泛研究,如图 4 所示,其

中, X 表示源域, Y 表示目标域, $G_{X \rightarrow Y}$ 表示源域到目标域的生成器, $G_{Y \rightarrow X}$ 表示目标域到源域的生成器, D_X 表示源域的判别器, D_Y 表示目标域的判别器, L_{cycle} 表示循环一致性损失。

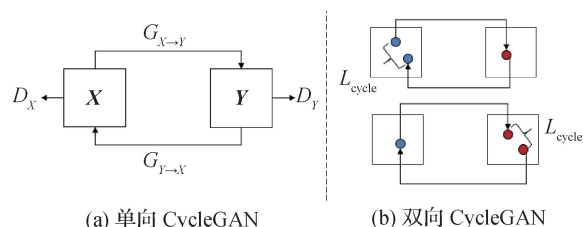


图 4 图像到图像翻译的生成对抗式网络 CycleGAN

Fig. 4 Generative adversarial network CycleGAN for image-to-image translation((a) unidirectional CycleGAN;(b) bidirectional CycleGAN)

Hiasa 等人 (2018) 通过添加梯度一致性损失扩展了 CycleGAN 方法, 实现了 MRI 到 CT 的跨模态图像合成, 以将 MRI 与 CT 一起用于识别肌肉结构并诊断骨坏死, 避免了 CT 图像拍摄过程中的放射性照射。为了解决域移问题, Chen 等人 (2018) 提出一种无监督域自适应的语义感知生成对抗式网络 SeUDA (semantic-aware GAN for unsupervised domain adaptation), 在语义标签空间中嵌入嵌套的对抗性学习, 通过语义感知损失来改善 CycleGAN。为了将 CT 数字重建射线照片 (digitally reconstructed radiographs, DRR) 上训练的图像分割模型推广到 X 射线域, Zhang 等人 (2018a) 提出一种任务驱动的生成对抗式网络 TD-GAN (task driven GAN), 用于将 X 射线图像迁移到 DRR 模态下。该模型包含一个经过修改的 CycleGAN 子结构, 用于在 DRR 和 X 射线图像之间进行像素到像素的转换, 以及一个附加模块, 利用预训练的 DI2I (dense image-to-image network) 分割网络来保证分割一致性。为了应对不同设备或不同参数配置下同一设备成像结果的噪声分布存在差异的问题, Zhang 等人 (2020) 提出了一种新颖的噪声自适应生成对抗式网络 NAGAN (noise adaptation GAN), 其中包含一个生成器和两个判别器, 两个判别器分别用于内容和噪声分布的训练引导, 并通过特征级的损失 Frobenius 范数进一步对内容与噪声模式两方面进行约束, 实验结果证明该方法可以有效迁移噪声分布。为避免模态迁移过程中目标内容的丢失, Xie 等人 (2020) 提出 MI2GAN (mutual

information constraint GAN), 从图像中解耦出内容特征, 并通过最大化内容特征间的互信息引导模态迁移过程。Yu 等人 (2019) 提出一种用于跨模态 3 维 MR 图像合成的边缘感知生成对抗式网络 Ea-GANs (edge-aware GAN), 优化了合成图像内容结构的纹理细节, 缓解了现有 2 维 cGAN 模型在切片间的不连续合成问题, 在更大范围内捕获图像上下文。Yang 等人 (2020) 提出一种基于顺序生成对抗式网络和半监督学习的双模态医学图像合成方法, 优先合成低复杂度的模态图像, 然后再合成高复杂度的图像。

模态迁移对于解决域移问题、提高深度学习模型泛化能力具有重要意义, 并且多模态数据的融合往往有助于疾病得到更加准确的诊断。基于无成对训练图像的生成对抗式网络在该领域得到广泛研究, 但是在模态迁移过程中, 基础几何结构信息容易丢失, 已有的相关研究对结构信息保持的研究仅局限在边缘以及语义分割等信息的融合。如何更深层次地将几何先验知识与生成对抗式网络结合起来是需要深入研究的方向。现有各模态迁移模型的改进内容、优缺点及适用场景如表 3 所示。

2.3 图像分割

图像分割即将图像分成若干特定的、具有独特性质的区域并提出感兴趣目标的技术和过程。在医学影像处理领域, 图像的精确分割是临床参数量化以及疾病分类的关键步骤, 对临床诊断、治疗决策以及手术预后等方面具有重要意义。基于深度学习技术的图像分割方法得到迅速发展, 但是由于医学领域的数据瓶颈问题, 其数据驱动的优势未能充分利用。

生成对抗式网络的生成器和判别器对提高基于深度学习的分割模型的性能均有重要意义。生成器可以合成额外数据, 判别器可以从高级语义层面对分割模型训练过程进行监督, 并充分利用无标签数据。Tan 等人 (2021) 提出一种基于深度学习的通用对抗式网络分割方案 LGAN (GAN-based lung segmentation), 通过生成器实现图像到分割结果的翻译, 并结合判别器的 EM (earth-mover) 损失和分割结果的交叉熵损失对模型进行优化。Lei 等人 (2020) 提出一种结合密集卷积 U-Net 和双重判别器的分割网络, 其中一个判别器检查输入图像中目标对象的上下文环境, 另一个判别器则检查分割结果的边缘

表 3 模态迁移应用
Table 3 Modality migration application

模型	改进	优点	缺点	应用场景
Hiasa 等人 (2018)	添加梯度一致性损失,扩展了 CycleGAN 方法。	加强了模态迁移过程中对于几何结构的约束。	迁移后的图像应用于分割任务的精度与其具体模态相关。	识别肌肉结构并诊断骨坏死,避免 CT 图像拍摄过程中的放射性照射。
NAGAN	分别引导内容和噪声的训练,通过特征级的损失 Frobenius 范数进一步对内容与噪声模式进行约束。	引入了特征损失函数,并对特征进行解耦。	需要针对每种不同目标域数据集训练模态迁移模型,无法根本解决需要重新训练模型的问题。	噪声分布迁移,以解决同一设备不同参数设置或不同设备上噪声分布不一致问题,如 OCT 和超声影像。
TD-GAN	基于优化后的 Cycle-GAN 子结构在 DRR 和 X-ray 影像间进行模态迁移,并利用分割网络保证结构一致性。	可同时进行模态迁移与分割的端到端框架。	未考虑拍摄过程中移动导致伪影等情形下的器官分割。	将 X 射线图像迁移到 DRR 模态进行分割,利用 DRR 数据训练的模型对 X 射线图像进行分析。
SeUDA	在语义标签空间中嵌入嵌套的对抗性学习,通过语义感知损失改善 CycleGAN。	缓解了域移问题。	未与其他基于生成对抗式网络的模态迁移方法进行充分比较。	通过将目标域的图像迁移到源域,利用源域下训练的模型进行医学图像分割。
M2GAN	从图像中解耦出内容特征,并通过最大化内容特征间的互信息引导模态迁移过程。	有效保留了源域图像的内容信息。	未与迁移学习的方法进行比较。	通过域适应提高模型泛化能力,将某一域上训练的模型推广到多中心数据的其他域。
Ea-GANs	提出一种用于跨模态 3 维 MR 图像合成的边缘感知生成对抗式网络。	在生成器与判别器中均引入了边缘感知信息,并实现了 3 维跨模态图像合成。	仅通过边缘检测任务实现结构信息保留。	优化内容结构,并缓解 2D 切片间不连续的问题。
Yang 等人 (2020)	基于复杂度顺序生成对抗式网络和半监督学习的双模态医学图像合成方法。	基于 3 种类型的评估指标和一项用户研究进行了非常全面的质量评估。	仅限于两种模态的图像合成,并且需要成对数据。	T1w-T2w 脑部双模态图像合成。

细节。Yuan 等人(2020)提出一种 3 维的具有良好泛化性能的生成对抗式分割网络,将任意模态图像的翻译和分割结合在一起,其关键在于通过生成对抗式网络提取模态不变性结构特征,以实现任意模态图像的良好分割性能。Wu 等人(2021)将多阶段姿态估计网络 MSPN (multi-stage pose estimation network) 和判别器集成在一起,使用带有多尺度膨胀卷积模块的 MSPN 扩大深层特征感受野,并且在 GAN 训练过程中,交替提供标记的和未标记的 CMRI (cardiac magnetic resonance images) 数据,以充分利用无标记数据。Ruan 等人(2020)提出了多分支特征共享生成对抗式网络,用于在 CT 上同时进行肾脏肿瘤的分割和量化,该网络由多尺度特征提取

器 MSFE (multi-scale feature extractor)、关注区域定位器 LROI (localization region of interest) 和特征共享生成对抗式网络 FSGAN (feature share GAN) 组成。MSFE 在不同尺度的特征图上提供强大的语义信息,这在检测小肿瘤目标时特别有效。LROI 提取了肿瘤的目标区域,大幅降低了网络的时间复杂度。FSGAN 通过联合学习和对抗学习正确分割和量化了肾脏肿瘤,有效利用了两个相关任务之间的共性和差异。Zhang 等人(2018b)提出一种由生成器和分割器组成的互为有益的端到端的解决方案,通过循环一致性损失以及分割器监督下的形状一致性损失保持合成图像一致的解剖结构。现有各图像分割模型的改进内容、优缺点及适用场景如表 4 所示。

表 4 图像分割应用
Table 4 Image segmentation application

模型	改进	优点	缺点	应用场景
LGAN	通过生成器实现图像到分割结果的翻译,并结合判别器的 EM 损失和分割结果的交叉熵损失对模型优化。	可以从高级语义层面检查生成结果的合理性。	仅适用于单一模态影像的分割。	2 维单一模态医疗影像分割。
Lei 等人 (2020)	通过两个判别器辅助分割训练监督,一个判别器负责语义上下文检查,另一个负责纹理细节检查。	在皮肤病变数据集上获得了更好的分割性能,且具有良好扩展性。	同 LGAN	同 LGAN
Wu 等人 (2021)	使用多尺度膨胀卷积模块扩大深层特征感受野,并基于标记的和未标记的数据交替训练。	捕获更多上下文语义信息,充分利用无标记数据,提高模型性能。	同 LGAN	同 LGAN
Ruan 等人 (2020)	提出了多分支特征共享生成对抗式网络,用于在 CT 上同时进行肾脏肿瘤的分割和量化。	提高小目标分割精度,降低模型的时间复杂度,利用了多任务间的共性和差异。	同 LGAN	同 LGAN
Zhang 等人 (2018b)	提出一种由生成器和分割器组成的互为有益的端到端的解决方案。	通过循环一致性损失以及分割器监督下的形状一致性损失保持合成图像的解剖结构。	同 LGAN	同 LGAN
Yuan 等人 (2020)	提出一种 3 维的生成对抗式分割网络,可提取模态不变性结构特征,实现对任意模态图像的良好分割性能。	泛化性能好,可以实现任意模态图像的鲁棒分割。	GPU 内存资源需求大,多层特征融合会导致边界预测产生偏差。	3 维多模态医疗影像分割。

2.4 图像去噪

图像去噪过程中,需要将由于较低剂量、较低分辨率或者较短曝光时间等导致的有噪声图像重建到近似理想的图像,同时保留图像中的重要细节。Wolterink 等人 (2017) 提出一种基于体素损失和对抗性损失的生成对抗式网络用于从低剂量 CT 图像估计常规剂量 CT 图像,同时通过实验证明了该网络对低剂量 CT 图像进行降噪是可行的,并且允许在含有较高噪声的低剂量患者 CT 图像中对冠状动脉钙化进行评分。Liu 等人 (2020) 提出 TomoGAN, 用于在低剂量成像条件下提高重建 X-ray 图像的质量,并且在两种条件下评估了该方法:1) 足够数量的低剂量投影;2) 数量有限的高剂量投影。实验结果表明该方法可以显著减少重建图像中的噪声。Ran 等人 (2019) 提出一种基于残差编码器—解码器 Wasserstein 生成对抗式网络 RED-WGAN (residual encoder-decoder Wasserstein GAN) 的 MRI 去噪方法,通过计算预训练的 VGG-19 (Visual Geometry Group 19-layer network) 网络在特征空间中的距离来实现

感知相似性,并将其与均方误差 (mean square error, MSE) 以及对抗损失结合在一起以形成新的损失函数,在模拟和实际临床数据中,该方法表现了出色的噪声抑制能力和结构保持能力,并且运行速度得到大幅提高,平均执行时间为 0.16 s。You 等人 (2020) 提出一种以生成对抗式网络为基础的半监督方法 GAN-CIRCLE,以从低分辨率图像中恢复高分辨率 CT 图像,该方法基于 Wasserstein 距离强制执行循环一致性,并在损失函数中引入联合约束以实现几何结构保持,此外,基于几种基本的设计原则优化了网络,减少了参数量,缓解了过拟合问题。You 等人 (2018) 提出一种 3D 降噪方法,称为结构敏感的多尺度生成对抗式网络 SMGAN (structurally-sensitive multi-scale GAN),该方法结合了对抗性损失、结构损失以及 L1 损失进行模型优化,有效提高了低剂量 CT 的图像质量。

2.5 小结

生成对抗式网络已经广泛应用于医疗影像处理领域。1) 基于 Pix2pix 基础框架的网络模型能够合

成额外的高质量、高分辨率样本,以数据增广方式有效提高了分割和分类模型的性能。但是仍然存在合成样本多样性不足、基础生物结构保持困难以及3维影像合成能力有限等问题。2) 基于 CycleGAN 基础框架的网络模型无需成对训练图像,在模态迁移应用中得到广泛研究,但是基础结构信息容易丢失,已有的针对模态迁移过程中结构保持的研究仅局限在边缘和分割等信息的融合,未来需要研究更深层次的几何先验知识的使用。3) 生成器和判别器均可以与已有的分割模型融合以提升分割模型性能,生成器可以合成额外数据,判别器可以从高级语义层面指导模型训练,并能够充分利用无标签数据,然而已有研究主要集中在单一模态的图像分割,研究如何提取多种模态图像的不变性特征以实现具有更好泛化性能的分割模型,是临床应用的必要条件。4) 生成对抗式网络应用于图像去噪领域,可以实现由低剂量图像重建正常剂量图像,减少患者受到的放射影响。

3 合成图像质量评估

3.1 常用评价指标

常用评价指标主要有 IS、FID、SWD 以及 GAN-train 和 GAN-test。

1) IS (inception score) 是评估 GAN 的最常见方法之一 (Salimans 等, 2016)。IS 使用在 ImageNet 上经过预训练的 Inception 网络 (Shmelkov 等, 2018) 计算生成图像的 logits。具体为

$$IS(G) = \exp(E_{x \sim p_g} [D_{KL}(p(y|\mathbf{x}) \| p(y))]) \quad (2)$$

式中, $\mathbf{x}$ 是从生成器分布 p_g 采样得到的合成图像, E 是整个合成图像集合的期望, D_{KL} 是条件类分布 $p(y|\mathbf{x})$ 与标签 y 边缘类别分布 $p(y) = E_{x \sim p_g} [p(y|\mathbf{x})]$ 之间的 KL (Kullback-Leibler) 散度。但是, IS 没有考虑真实图像数据, 并且不能度量生成器分布与真实分布之间的接近程度 (Barratt 和 Sharma, 2018)。

2) FID (Fréchet inception distance) 比较了真实图像 $p_r(x)$ 与生成图像 $p_g(x)$ 在 Inception 网络中倒数第 2 层激活的分布 (Szegedy 等, 2014)。其分布都建模为多维高斯分布, 并由它们各自的均值和协方

差参数化。这两个高斯分布之间的距离度量为

$$d^2((m_r, C_r), (m_g, C_g)) = \|m_r - m_g\|^2 + \text{tr}(C_r + C_g - 2(C_r C_g)^{\frac{1}{2}}) \quad (3)$$

式中, $(m_r, C_r), (m_g, C_g)$ 分别表示真实图像与合成图像分布的均值和协方差。FID 引入了与真实数据分布之间的比较, 较低的 FID 意味着两个分布之间更接近, 也就意味着生成图像的质量较高、多样性较好。但是, FID 和 IS 都是基于特征提取, 也就是依赖于某些特征的出现或者不出现。但是它们都无法描述这些特征的空间关系。此外, IS 和 FID 都无法将图像质量与图像多样性区分开。

3) SWD (sliced Wasserstein distance) 是基于 Wasserstein 的距离度量 (Rabin 等, 2012), 它从真实图像 (P_r) 和生成图像 (P_g) 的拉普拉斯金字塔表示中提取局部图像 patch, 并据此计算出多尺度统计相似性。具体为

$$W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} E_{(x,y) \sim \gamma} [\|x - y\|] \quad (4)$$

式中, $\Pi(P_r, P_g)$ 表示 P_r 和 P_g 分布组合起来的所有可能的联合分布的集合。对于每一个可能的联合分布 γ , 可以从中采样得到 $(x, y) \sim \gamma$, 并计算出这对样本的距离 $\|x - y\|$, 所以可以计算该联合分布下样本对距离的期望 $E_{(x,y) \sim \gamma} [\|x - y\|]$ 。在所有可能的联合分布中能够对这个期望值取到的下界 $\inf_{\gamma \in \Pi(P_r, P_g)} E_{(x,y) \sim \gamma} [\|x - y\|]$ 就是 Wasserstein 距离。Wasserstein 距离相对于 KL 散度的优势在于, 即使两个分布的支撑集没有重叠或重叠非常小, 它仍然能够反映两个分布的远近, 而在此情况下 KL 散度则可能没有意义。SWD 使用真实图像和生成图像之间 Wasserstein-1 距离的随机投影 (Heusel 等, 2017) 进行了有效的近似, 适用于评估高分辨率 GAN, 但是并非所有模型都能生成高分辨率图像。

4) GAN-train 和 GAN-test。GAN-train 是在合成图像上训练并在真实图像验证集上进行测试的分类器的准确性 (Theis 等, 2016)。GAN-test 是在真实图像训练集上训练但在合成图像上测试的分类器的准确性。GAN-train 类似于召回率, 良好的 GAN-train 性能表明生成的样本足够多样化。第 2 个度量值 GAN-test 与精度相似, 若 GAN-test 很高, 则表示生成的样本是真实图像分布逼真的近似。相对于 IS 和 FID, GAN-train 和 GAN-test 不依赖于 ImageNet 预

训练的 Inception 网络,并且考虑了与真实图像分布之间的相似度。此外,它们还可以将图像质量与图像多样性区分开。

除此之外,还有一些定性评估方法,如 AMT (Amazon Mechanical Turk) perceptual studies (Isola 等,2017),但是一些研究表明,对于这种基于人类感知的评估研究,生成图像质量必须有显著差异 (Heusel 等,2017)。

Theis 等人 (2016) 研究对比了生成模型的优化与评估,不同的度量标准可能导致不同的权衡取舍,因此,相对于某一指标可以获得良好的表现并不意味着对另一标准也可以获得相同的表现。

3.2 小结

数据增广、模态迁移、图像分割以及图像去噪等任务对于合成图像的质量以及多样性要求各有不同。数据增广侧重于数据合成的逼真质量以及丰富的多样性;模态迁移侧重于源域图像与目标域图像风格的接近以及内容的不变性;图像分割侧重于分割性能的提升;图像去噪侧重于成像质量的改善以及对于特定任务的有效性。因此,各任务在进行评价时,应充分考虑任务需求以选择合适的度量指标。

此外,已有的各类评价指标并不全面。如,如何有效评价图像去噪后的质量,获得良好的图像质量得分,并不能确保其去噪后的图像能提升下游分割和分类等任务的性能。

最后,对于不同任务应该统一其评价指标,以确保不同研究工作之间结果的参考性。

4 结 语

生成对抗式网络对于解决医学影像领域的瓶颈问题,充分发挥深度学习方法的数据驱动优势具有重要意义。

自生成对抗式网络提出以来,其理论不断完善,实际应用也取得了长足发展,但是在医学影像处理领域仍然存在一些亟待解决的问题:1) 医学影像数据大部分是3维的,如MRI、CT等,如何提高3维数据的合成质量和分辨率;2) 如何确保合成数据在获得多样性的同时保证其基础几何结构的合理性;3) 存在大量无标记和未配对的数据,如何充分利用这类数据以生成高质量、高分辨率并具有多样性的图像,提升分割、分类模型性能;4) 医学领域存在大量

多模态数据交叉使用的需求,如何实现提升算法的跨模态泛化性能,以及不同模态数据的有效迁移。

针对上述问题,未来的研究工作应当聚焦于:1) 针对3维数据合成优化网络架构、目标函数以及训练方式等,改善模型训练稳定性,提高3维合成图像质量、分辨率以及多样性;2) 更深层次地将几何先验知识与生成对抗式网络结合;3) 充分利用生成对抗式网络的弱监督特性;4) 通过属性解耦以提取不变性特征,以实现具有更好泛化性能的分割和分类模型,并且在模态迁移过程中实现不同层面、不同粒度以及不同需求的属性控制。

参考文献 (References)

- Alharbi Y and Wonka P. 2020. Disentangled image generation through structured noise injection//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 5133-5141 [DOI: 10.1109/CVPR42600.2020.00518]
- Almahairi A, Rajeshwar S, Sordoni A, Bachman P and Courville A. 2018. Augmented CycleGAN: learning many-to-many mappings from unpaired data//Proceedings of the 35th International Conference on Machine Learning. Stockholm, Sweden: PMLR: 195-204
- Badrinarayanan V, Handa A and Cipolla R. 2015. SegNet: a deep convolutional encoder-decoder architecture for robust semantic pixel-wise labeling [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1505.07293.pdf>
- Bahdanau D, Cho K and Bengio Y. 2016. Neural machine translation by jointly learning to align and translate [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1409.0473.pdf>
- Barratt S and Sharma R. 2018. A note on the inception score [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1801.01973.pdf>
- Bissoto A, Perez F, Valle E and Avila S. 2018. Skin lesion synthesis with generative adversarial networks//OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis. Switzerland: Springer: 294-302 [DOI: 10.1007/978-3-030-01201-4_32]
- Brock A, Donahue J and Simonyan K. 2019. Large scale GAN training for high fidelity natural image synthesis [EB/OL]. [2021-08-07] <https://arxiv.org/pdf/1809.11096v1.pdf>
- Cao Y J, Jia L L, Chen Y X, Lin N and Li X X. 2018. Review of computer vision based on generative adversarial networks. Journal of Image and Graphics, 23(10): 1433-1449 (曹仰杰, 贾丽丽, 陈永霞, 林楠, 李学相. 2018. 生成式对抗网络及其计算机视觉应用研究综述. 中国图象图形学报, 23(10): 1433-1449) [DOI: 10.11834/jig.180103]
- Chen C, Dou Q, Chen H and Heng P A. 2018. Semantic-aware genera-

- tive adversarial nets for unsupervised domain adaptation in chest X-ray segmentation//Proceedings of the 9th International Workshop on Machine Learning in Medical Imaging. Granada, Spain; Springer; 143-151 [DOI: 10.1007/978-3-030-00919-9_17]
- Chen F J, Zhu F, Wu Q X, Hao Y M, Wang E D and Cui Y G. 2021. A survey about image generation with generative adversarial nets. Chinese Journal of Computers, 44(2): 347-369 (陈佛计, 朱枫, 吴清潇, 郝颖明, 王恩德, 崔芸阁. 2021. 生成对抗网络及其在图像生成中的应用研究综述. 计算机学报, 44(2): 347-369)
- Costa P, Galdran A, Meyer M I, Niemeijer M, Abràmoff M, Mendonça A M and Campilho A. 2018. End-to-end adversarial retinal image synthesis. IEEE Transactions on Medical Imaging, 37(3): 781-791 [DOI: 10.1109/TMI.2017.2759102]
- Denton E, Chintala S, Szlam A and Fergus R. 2015. Deep generative image models using a Laplacian pyramid of adversarial networks//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada; MIT Press; 1486-1494
- Dysmorphology Subcommittee of the Clinical Practice Committee and American College of Medical Genetics. 2000. Informed consent for medical photographs. Genetics in Medicine, 2(6): 353-355 [DOI: 10.1097/00125817-200011000-00010]
- Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J and Greenspan H. 2018. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing, 321: 321-331 [DOI: 10.1016/j.neucom.2018.09.013]
- 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, Canada; MIT Press; 2672-2680
- Greenspan H, Van Ginneken B and Summers R M. 2016. Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. IEEE Transactions on Medical Imaging, 35(5): 1153-1159 [DOI: 10.1109/TMI.2016.2553401]
- 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, USA; IEEE; 770-778 [DOI: 10.1109/CVPR.2016.90]
- Heusel M, Ramsauer H, Unterthiner T, Nessler B and Hochreiter S. 2017. GANs trained by a two time-scale update rule converge to a local nash equilibrium//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA; Curran Associates Inc.; 6629-6640
- Hiasa Y, Otake Y, Takao M, Matsuoka T, Takashima K, Prince J L, Sugano N and Sato Y. 2018. Cross-modality image synthesis from unpaired data using CycleGAN; effects of gradient consistency loss and training data size//Proceedings of the 3rd International Workshop on Simulation and Synthesis in Medical Imaging. Granada, Spain; Springer; 31-41 [DOI: 10.1007/978-3-030-00536-8_4]
- Huang G, Liu Z, van der Maaten L and Weinberger K Q. 2016. Densely connected convolutional networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA; IEEE; 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- 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]
- Jin D K, Xu Z Y, Tang Y B, Harrison A P and Mollura D J. 2018. CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation//Proceedings of the 21st International Conference on Medical Image Computing and Computer Assisted Intervention. Granada, Spain; Springer; 732-740 [DOI: 10.1007/978-3-030-00934-2_81]
- Karras T, Aila T, Laine S and Lehtinen J. 2018. Progressive growing of GANs for improved quality, stability, and variation [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1710.10196.pdf>
- Karras T, Laine S and Aila T. 2019. A style-based generator architecture for generative adversarial networks//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA; IEEE; 4396-4405 [DOI: 10.1109/CVPR.2019.00453]
- 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; Curran Associates Inc.; 1097-1105
- Kwon G, Han C and Kim D S. 2019. Generation of 3D brain MRI using auto-encoding generative adversarial networks//Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention. Shenzhen, China; Springer; 118-126 [DOI: 10.1007/978-3-030-32248-9_14]
- Lei B Y, Xia Z M, Jiang F, Jiang X D, Ge Z Y, Xu Y W, Qin J, Chen S P, Wang T F and Wang S Q. 2020. Skin lesion segmentation via generative adversarial networks with dual discriminators. Medical Image Analysis, 64: #101716 [DOI: 10.1016/j.media.2020.101716]
- Li H, Hu Y, Li S, Lin W, Lin P, Higashita R and Liu J. 2020. CT scan synthesis for promoting computer-aided diagnosis capacity of COVID-19//Proceedings of International Conference on Intelligent Computing. Springer, Cham, 2020; 413-422
- Li R D, Pan J S, Li Z C and Tang J H. 2018. Single image dehazing via conditional generative adversarial network//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE; 8202-8211 [DOI: 10.1109/CVPR.2018.00856]
- Liu M Y, Huang X, Mallya A, Karras T, Aila T, Lehtinen J and Kautz J. 2019. Few-shot unsupervised image-to-image translation//Proceedings of 2019 IEEE/CVF International Conference on Computer

- Vision. Seoul, Korea (South); IEEE: 10550-10559 [DOI: 10.1109/ICCV.2019.01065]
- Liu Z C, Bicer T, Kettimuthu R, Gürsoy D, de Carlo F and Foster I. 2020. TomoGAN: low-dose synchrotron X-ray tomography with generative adversarial networks; discussion. *Journal of the Optical Society of America A*, 37(3): 422-434 [DOI: 10.1364/JOSAA.375595]
- Mejjati Y A, Richardt C, Tompkin J, Cosker D and Kim K I. 2018. Unsupervised attention-guided image to image translation [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1806.02311.pdf>
- Mirza M and Osindero S. 2014. Conditional generative adversarial nets [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1411.1784v1.pdf>
- Nguyen-Phuoc T, Li C, Theis L, Richardt C and Yang Y L. 2019. Hol-oGAN: unsupervised learning of 3D representations from natural images//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Seoul, Korea (South); IEEE: 2037-2040 [DOI: 10.1109/ICCVW.2019.00255]
- Odena A, Olah C and Shlens J. 2017. Conditional image synthesis with auxiliary classifier GANs [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1610.09585v4.pdf>
- Office for Civil Rights (OCR). 2022. Your rights under HIPAA [EB/OL]. US Department of Health and Human Services. [2022-02-24]. <https://www.hhs.gov/hipaa/for-individuals/guidance-materials-for-consumers/index.html>
- Olsewski K, Ceylan D, Xing J, Echevarria J, Chen Z L, Chen W K and Li H. 2020. Intuitive, interactive beard and hair synthesis with generative models//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA; IEEE: 7444-7454 [DOI: 10.1109/CVPR42600.2020.00747]
- Perarnau G, van de Weijer J, Raducanu B and Álvarez J M. 2016. Invertible Conditional GANs for image editing [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1611.06355.pdf>
- Rabin J, Peyré G, Delon J and Bernot M. 2012. Wasserstein barycenter and its application to texture mixing// Proceedings of the 3rd International Conference on Scale Space and Variational Methods in Computer Vision. Ein-Gedi, Israel; Springer: 435-446 [DOI: 10.1007/978-3-642-24785-9_37]
- Ran M S, Hu J R, Chen Y, Chen H, Sun H Q, Zhou J L and Zhang Y. 2019. Denoising of 3D magnetic resonance images using a residual encoder-decoder Wasserstein generative adversarial network. *Medical Image Analysis*, 55: 165-180 [DOI: 10.1016/j.media.2019.05.001]
- Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B and Yang G Z. 2017. Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*, 21(1): 4-21 [DOI: 10.1109/JBHI.2016.2636665]
- Ruan Y, Li D W, Marshall H, Miao T, Cossetto T, Chan I, Daher O, Accorsi F, Goela A and Li S. 2020. MB-FSGAN: joint segmentation and quantification of kidney tumor on CT by the multi-branch feature sharing generative adversarial network. *Medical Image Analysis*, 64: #101721 [DOI: 10.1016/j.media.2020.101721]
- Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A and Chen X. 2016. Improved techniques for training GANs//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain; Curran Associates Inc.: 2234-2242
- Schmitt J M, Xiang S H and Yung K M. 1999. Speckle in optical coherence tomography. *Journal of Biomedical Optics*, 4(1): 95-105 [DOI: 10.1117/1.429925]
- Shelhamer E, Long J and Darrell T. 2017. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4): 640-651 [DOI: 10.1109/TPAMI.2016.2572683]
- Shi J, Wang L L, Wang S S, Chen Y X, Wang Q, Wei D M, Liang S J, Peng J L, Yi J J, Liu S F, Ni D, Wang M L, Zhang D Q and Shen D G. 2020. Applications of deep learning in medical imaging: a survey. *Journal of Image and Graphics*, 25(10): 1953-1981 (施俊, 汪琳琳, 王珊珊, 陈艳霞, 王乾, 魏冬铭, 梁淑君, 彭佳林, 易佳锦, 刘盛锋, 倪东, 王明亮, 张道强, 沈定刚. 2020. 深度学习在医学影像中的应用综述. *中国图象图形学报*, 25(10): 1953-1981) [DOI: 10.11834/jig.200255]
- Shmelkov K, Schmid C and Alahari K. 2018. How good is my GAN? // Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany; Springer: 218-234 [DOI: 10.1007/978-3-030-01216-8_14]
- Sutskever I, Vinyals O and Le Q V. 2014. Sequence to sequence learning with neural networks//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada; MIT Press: 3104-3112
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2014. Going deeper with convolutions//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA; IEEE: 1-9 [DOI: 10.1109/CVPR.2015.7298594]
- Tan J X, Jing L L, Huo Y M, Li L H, Akin O and Tian Y L. 2021. LGAN: lung segmentation in CT scans using generative adversarial network. *Computerized Medical Imaging and Graphics*, 87: #101817
- Theis L, van den Oord A and Bethge M. 2016. A note on the evaluation of generative models [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1511.01844.pdf>
- Wan C H, Chuang S P and Lee H Y. 2019. Towards audio to scene image synthesis using generative adversarial network//Proceedings of 2019 IEEE International Conference on Acoustics, Speech and Signal Processing. Brighton, UK; IEEE: 496-500 [DOI: 10.1109/ICASSP.2019.8682383]
- Wang Q, Fan H J, Sun G, Ren W H and Tang Y D. 2020. Recurrent generative adversarial network for face completion. *IEEE Transac-*

- tions on Multimedia, 23: 429-442 [DOI: 10.1109/TMM.2020.2978633]
- Wang T C, Liu M Y, Zhu J Y, Tao A, Kautz J and Catanzaro B. 2018. High-resolution image synthesis and semantic manipulation with conditional GANs//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 8798-8807 [DOI: 10.1109/CVPR.2018.00917]
- Wang W Y, Huang Q G, You S Y, Yang C and Neumann U. 2017. Shape inpainting using 3D generative adversarial network and recurrent convolutional networks//Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy; IEEE: 2317-2325 [DOI: 10.1109/ICCV.2017.252]
- Wolterink J M, Leiner T, Viergever M A and Išgum I. 2017. Generative adversarial networks for noise reduction in low-dose CT. IEEE Transactions on Medical Imaging, 36(12): 2536-2545 [DOI: 10.1109/TMI.2017.2708987]
- Wu H S, Lu X H, Lei B Y and Wen Z K. 2021. Automated left ventricular segmentation from cardiac magnetic resonance images via adversarial learning with multi-stage pose estimation network and co-discriminator. Medical Image Analysis, 68: #101891 [DOI: 10.1016/j.media.2020.101891]
- Wu Y, Donahue J, Balduzzi D, Simonyan K and Lillicrap T. 2020. LOGAN: latent optimisation for generative adversarial networks [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1912.00953.pdf>
- Xiang S H, Zhou L and Schmitt J M. 1998. Speckle noise reduction for optical coherence tomography//Proceedings Volume 3196, Optical and Imaging Techniques for Biomonitoring III. San Remo, Italy; SPIE: 79-88 [DOI: 10.1117/12.297921]
- Xie X P, Chen J W, Li Y X, Shen L L, Ma K and Zheng Y F. 2020. MI²GAN: generative adversarial network for medical image domain adaptation using mutual information constraint//Proceedings of the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention. Lima, Peru; Springer: 516-525 [DOI: 10.1007/978-3-030-59713-9_50]
- Yang X, Lin Y, Wang Z W, Li X and Cheng K T. 2020. Bi-modality medical image synthesis using semi-supervised sequential generative adversarial networks. IEEE Journal of Biomedical and Health Informatics, 24(3): 855-865 [DOI: 10.1109/JBHI.2019.2922986]
- You C Y, Li G, Zhang Y, Zhang X L, Shan H M, Li M Z, Ju S H, Zhao Z, Zhang Z Y, Cong W X, Vannier M W, Saha P K, Hoffman E A and Wang G. 2020. CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE). IEEE Transactions on Medical Imaging, 39(1): 188-203 [DOI: 10.1109/TMI.2019.2922960]
- You C Y, Yang Q S, Shan H M, Gjestebj L, Li G, Ju S H, Zhang Z Y, Zhao Z, Zhang Y, Cong W X and Wang G. 2018. Structurally-sensitive multi-scale deep neural network for low-dose CT denoising. IEEE Access, 6: 41839-41855 [DOI: 10.1109/ACCESS.2018.2858196]
- Yu B T, Zhou L P, Wang L, Shi Y H, Fripp J and Bourgeat P. 2019. Ea-GANs: edge-aware generative adversarial networks for cross-modality MR image synthesis. IEEE Transactions on Medical Imaging, 38(7): 1750-1762 [DOI: 10.1109/TMI.2019.2895894]
- Yuan W G, Wei J, Wang J B, Ma Q L and Tasdizen T. 2020. Unified generative adversarial networks for multimodal segmentation from unpaired 3D medical images. Medical Image Analysis, 64: #101731 [DOI: 10.1016/j.media.2020.101731]
- Zhang H, Goodfellow I, Metaxas D and Odena A. 2019a. Self-attention generative adversarial networks [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1805.08318.pdf>
- Zhang H, Wang Y P, Geng X J and Fu P F. 2020. Review of deep learning methods for isointense infant brain MR image segmentation. Journal of Image and Graphics, 25(10): 2068-2078 (张航, 王雅萍, 耿秀娟, 付鹏飞. 2020. 等强度婴儿脑MR图像分割的深度学习方法综述. 中国图象图形学报, 25(10): 2068-2078) [DOI: 10.11834/jig.200285]
- Zhang T Y, Cheng J, Fu H Z, Gu Z W, Xiao Y T, Zhou K, Gao S H, Zheng R and Liu J. 2020. Noise adaptation generative adversarial network for medical image analysis. IEEE Transactions on Medical Imaging, 39(4): 1149-1159 [DOI: 10.1109/TMI.2019.2944488]
- Zhang T Y, Fu H Z, Zhao Y T, Cheng J, Guo M J, Gu Z W, Yang B, Xiao Y T, Gao S H and Liu J. 2019b. SkrGAN: sketching-rendering unconditional generative adversarial networks for medical image synthesis//Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention Society (MICCAI). Shenzhen, China; Springer: 777-785 [DOI: 10.1007/978-3-030-32251-9_85]
- Zhang Y, Miao S, Mansi T and Liao R. 2018a. Task driven generative modeling for unsupervised domain adaptation: application to X-ray image segmentation//Proceedings of the 21st International Conference on Medical Image Computing and Computer Assisted Intervention. Granada, Spain; Springer: 599-607 [DOI: 10.1007/978-3-030-00934-2_67]
- Zhang Z Z, Yang L and Zheng Y F. 2018b. Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA; IEEE: 9242-9251 [DOI: 10.1109/CVPR.2018.00963]
- Zhao L J, Bai H H, Liang J, Zeng B, Wang A H and Zhao Y. 2019. Simultaneous color-depth super-resolution with conditional generative adversarial networks. Pattern Recognition, 88: 356-369 [DOI: 10.1016/j.patcog.2018.11.028]
- 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

(ICCV). Venice, Italy: IEEE; 2242-2251 [DOI: 10. 1109/ICCV. 2017. 244]

Zhu J Y, Zhang R, Pathak D, Darrell T, Efros A A, Wang O and Shechtman E. 2018. Toward multimodal image-to-image translation [EB/OL]. [2021-08-07]. <https://arxiv.org/pdf/1711.11586.pdf>

作者简介



张颖麟, 1994 年生, 女, 硕士研究生, 主要研究方向为基于人工智能算法的医学图像处理、计算机视觉与深度学习技术。

E-mail: zhangyl7@mail.sustech.edu.cn



刘江, 通信作者, 男, 教授, 博士生导师, 主要研究方向为眼科人工智能、眼脑联动、医疗影像、手术机器人。

E-mail: liuj@sustech.edu.cn

胡衍, 女, 副研究员, 硕士生导师, 主要研究方向为手术辅助、术中导航、医学图像处理。E-mail: huy3@sustech.edu.cn

东田理沙, 女, 高级工程师, 主要研究方向为图像及信号处理。E-mail: k-chen@tomey.co.jp