



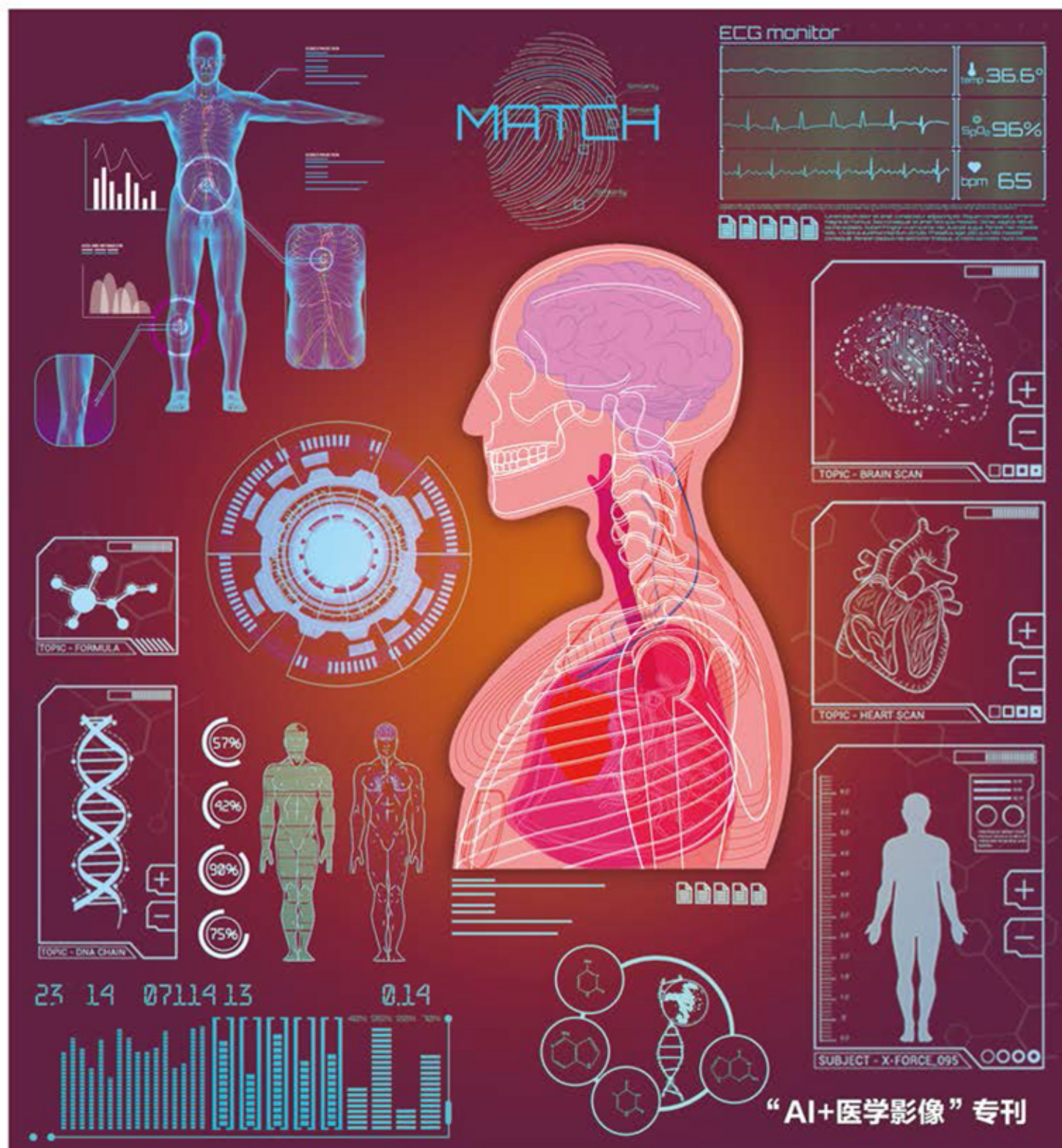
JOURNAL OF IMAGE AND GRAPHICS

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

# 中国图象图形学报

2020  
10  
VOL.25

ISSN1006-8961  
CN11-3758/TB





第25卷第10期（总第294期）

2020年10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 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** Aerospace Information Research Institute, 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 4<sup>th</sup> 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

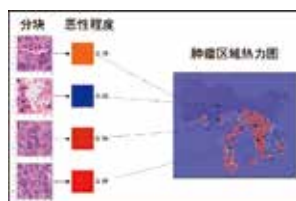
CODEN ZTTXFZ

国外发行代号 M1406

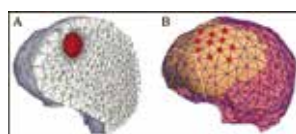
国内邮发代号 82-831

国内定价 60.00元

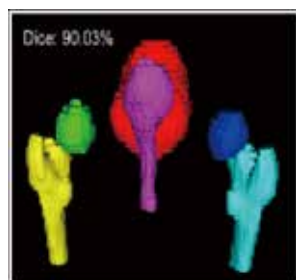




深度学习在组织病理学中的应用综述(第1982页)



机器学习在术中光学成像技术中的应用研究(第1994页)



2D级联CNN模型的放疗危及器官自动分割(第2110页)

序言 .....	安虹, 张云泉 I
编者按 .....	II

## 综述

### 深度学习在医学影像中的应用综述

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 .....	1953
---	------

### 深度学习在组织病理学中的应用综述

金旭, 文可, 吕国锋, 石军, 迟孟贤, 武铮, 安虹 .....	1982
------------------------------------	------

### 机器学习在术中光学成像技术中的应用研究

张崇, 王坤, 田捷 .....	1994
------------------	------

## 前沿进展

### 医学3D计算机视觉: 研究进展和挑战

杨健程, 倪冰冰 .....	2002
----------------	------

### 点云算法在医学领域的研究进展

李美佳, 于泽宽, 刘晓, 颜荣耀, 于媛媛, 王大明, 陈涓, 陆军, 祁鹏, 王俊杰, 刘杰 .....	2013
--	------

### 肝脏肿瘤CT图像深度学习分割方法综述

马金林, 邓媛媛, 马自萍 .....	2024
---------------------	------

### 磁共振影像深度学习在精神分裂症诊断中的应用综述

宋佩伦, 王雅萍, 耿秀娟, 张航, 宋学勤 .....	2047
------------------------------	------

### 新型冠状病毒肺炎(COVID-19)医学影像AI诊断研究进展

孟琢, 李铭辉 .....	2058
---------------	------

### 等强度婴儿脑MR图像分割的深度学习方法综述

张航, 王雅萍, 耿秀娟, 付鹏飞 .....	2068
-------------------------	------

### 医学影像疾病诊断的残差神经网络优化算法研究进展

周涛, 霍兵强, 陆惠玲, 师宏斌 .....	2079
-------------------------	------

### IVIM及纹理分析在术前预测宫颈癌类型和淋巴结转移研究进展

李翠平, 董江宁 .....	2093
----------------	------

### 4D时空纵向分析在生物医学领域中的应用现状与趋势

徐春园, 曾晓天, 宋泽雨, 唐晓英 .....	2100
--------------------------	------

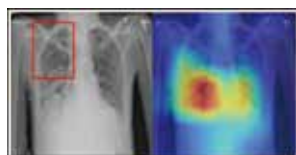
## 计算机断层扫描图像

### 2D级联CNN模型的放疗危及器官自动分割

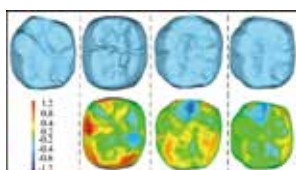
石军, 赵敏帆, 薛旭东, 郝晓宇, 金旭, 安虹, 张红雁 .....	2110
--------------------------------------	------

### 融合双注意力机制3D U-Net的肺肿瘤分割

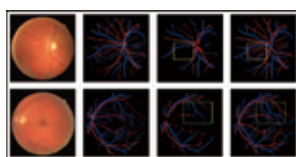
郝晓宇, 熊俊峰, 薛旭东, 石军, 文可, 韩文廷, 李晓扬, 赵俊, 傅小龙 .....	2119
--	------



密集挤压激励网络的多标签胸部X光片疾病分类  
(第2238页)



高分辨率深度生成网络的缺失牙体形态设计  
(第2249页)



语义融合眼底图像动静脉分类方法(第2259页)

## 深度学习结合影像组学的肝脏肿瘤CT分割

刘云鹏, 刘光品, 王仁芳, 金冉, 孙德超, 邱虹, 董晨, 李瑾, 洪国斌 ..... 2128

## 超级计算支撑的新冠肺炎CT影像综合分析辅助系统应用

康波, 郭佳, 王帅, 徐波, 孟祥飞 ..... 2142

## 鼻咽癌原发肿瘤放疗靶区的自动分割

薛旭东, 郝晓宇, 石军, 丁轶, 魏伟, 安虹 ..... 2151

## 磁共振图像

### 组卷积轻量级脑肿瘤分割网络

赵奕名, 李锦, 关欣 ..... 2159

### 深度迭代融合的脑部磁共振图像颅骨去除网络

姚发展, 李智, 王丽会, 程欣宇, 张健 ..... 2171

### 融合密集残差块和GAN变体的医学图像非刚性配准

张桂梅, 胡强, 龚磊 ..... 2182

## 超声图像

### 结合分段频域和局部注意力的超声甲状腺分割

胡屹杉, 秦品乐, 曾建潮, 柴锐, 王丽芳 ..... 2195

### 面向乳腺超声图像分割的混合监督双通道反馈U-Net

贡荣麟, 施俊, 王骏 ..... 2206

### 乳腺超声双模态数据的协同约束网络

杨子奇, 龚勋, 朱丹, 郭颖 ..... 2218

### 乳腺超声肿瘤动态噪声指数及分割算法

邹海鹏, 龚勋, 胡毕杰, 罗俊 ..... 2229

## 研究应用

### 密集挤压激励网络的多标签胸部X光片疾病分类

张智睿, 李锦, 关欣 ..... 2238

### 高分辨率深度生成网络的缺失牙体形态设计

郭闯, 戴宁, 田素坤, 孙玉春, 俞青, 刘浩, 程筱胜 ..... 2249

### 语义融合眼底图像动静脉分类方法

高颖琪, 郭松, 李宁, 王恺, 康宏, 李涛 ..... 2259

### AS-PANet:改进路径增强网络的重叠染色体实例分割

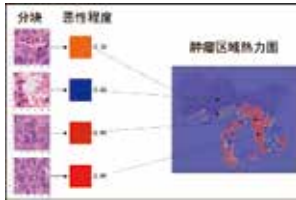
林成创, 赵淦森, 尹爱华, 丁笔超, 郭莉, 陈汉彪 ..... 2271

### TSCNN:面向可穿戴心电信号监测与分析的卷积神经网络

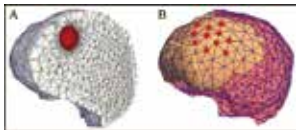
孟琨, 葛康, 宋阳, 杨东溟 ..... 2281

# CONTENTS

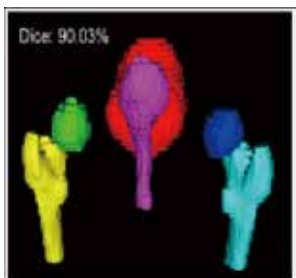
## JOURNAL OF IMAGE AND GRAPHICS



A Survey on Applications of Deep Learning in Histopathology(P1982)



Review: the application of machine learning in intraoperative optical imaging technologies(P1994)



Automatic segmentation of organs-at-risk in radiotherapy using 2D cascade-CNN model(P2110)

### Review

#### Applications of deep learning in medical imaging: a survey

Shi Jun, Wang Linlin, Wang Shanshan, Chen Yanxia, Wang Qian, Wei Dongming, Liang Shujun, Peng Jialin, Yi Jiajin, Liu Shengfeng, Ni Dong, Wang Mingliang, Zhang Daoqiang, Shen Dinggang ..... 1953

#### Survey on the applications of deep learning to Histopathology

Jin Xu, Wen Ke, Lyu Guofeng, Shi Jun, Chi Mengxian, Wu Zheng, An Hong ..... 1982

#### Review: the application of machine learning in intraoperative optical imaging technologies

Zhang Chong, Wang Kun, Tian Jie ..... 1994

### Frontier

#### Advances and challenges in medical 3D computer vision

Yang Jiancheng, Ni Bingbing ..... 2002

#### Progress of point cloud algorithm in medical field

Li Meijia, Yu Zekuan, Liu Xiao, Yan Rongyao, Yu Yuanyuan, Wang Daming, Chen Juan, Lu Jun, Qi Peng, Wang Junjie, Liu Jie ..... 2013

#### Review of deep learning segmentation methods for CT images of liver tumors

Ma Jinlin, Deng Yuanyuan, Ma Ziping ..... 2024

#### Survey of deep learning in MRI-based diagnosis of schizophrenia

Song Peilun, Wang Yaping, Geng Xiujuan, Zhang Hang, Song Xueqin ..... 2047

#### Progress of artificial intelligence diagnosis and Prognosis technology for COVID-19 medical imaging

Meng Lu, Li Ronghui ..... 2058

#### Review of deep learning methods for isointense infant brain MR image segmentation

Zhang Hang, Wang Yaping, Geng Xiujuan, Fu Pengfei ..... 2068

#### Progress of residual neural network optimization algorithm for medical imaging disease diagnosis

Zhou Tao, Huo Bingqiang, Lu Huiling, Shi Hongbin ..... 2079

#### Advances in preoperative identification of subtypes and lymph node metastasis of cervical cancer by IVIM and texture analysis

Li Cuiping, Dong Jiangning ..... 2093

#### Research status and trend of 4D spatiotemporal longitudinal analysis in biomedical field

Xu Chunyuan, Zeng Xiaotian, Song Zeyu, Tang Xiaoying ..... 2100

### Computerd Tomography Image

#### Automatic segmentation of organs-at-risk in radiotherapy using 2D cascade-CNN model

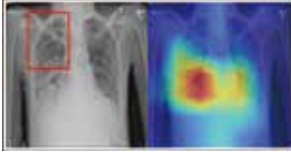
Shi Jun, Zhao Minfan, Xue Xudong, Hao Xiaoyu, Jin Xu, An Hong, Zhang Hongyan ..... 2110

#### 3D U-Net with dual attention mechanism for lung tumor segmentation

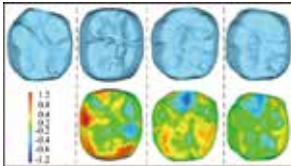
Hao Xiaoyu, Xiong Junfeng, Xue Xudong, Shi Jun, Wen Ke, Han Wenting, Li Xiaoyang, Zhao Jun, Fu Xiaolong ..... 2119

#### Accurate segmentation method of liver tumor CT based on the combination of deep learning and radiomics

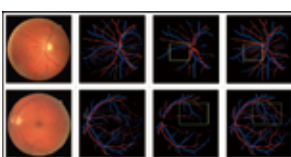
Liu Yunpeng, Liu Guangpin, Wang Renfang, Jin Ran, Sun Dechao, Qiu Hong, Dong Chen, Li Jin, Hong Guobin ..... 2128



Multilabel chest X-ray disease classification based on a dense squeeze-and-excitation network (P2238)



Morphological design of missing tooth driven by high-resolution deep generation network(P2249)



Arteriovenous classification method in fundus images based on semantic fusion(P2259)

## Supercomputing-supported COVID-19 CT image comprehensive analysis assistant system

Kang Bo, Guo Jia, Wang Shuai, Xu Bo, Meng Xiangfei ..... 2142

## Auto-segmentation of high-risk primary tumor gross target volume for the radiotherapy of nasopharyngeal carcinoma

Xue Xudong, Hao Xiaoyu, Shi Jun, Ding Yi, Wei Wei, An Hong ..... 2151

## Magnetic Resonance Image

### Lightweight brain tumor segmentation algorithm based on a group convolutional neural network

Zhao Yiming, Li Qiang, Guan Xin ..... 2159

### Deep iterative fusion network on skull removal of brain magnetic resonance images

Yao Fazhan, Li Zhi, Wang Lihui, Cheng Xinyu, Zhang Jian ..... 2171

### Non-rigid medical image registration based on residual-in-residual dense block and GAN

Zhang Guimei, Hu Qiang, Gong Lei ..... 2182

## Ultrasound Image

### Ultrasound thyroid segmentation based on segmented frequency domain and local attention

Hu Yishan, Qin Pinle, Zeng Jianchao, Chai Rui, Wang Lifang ..... 2195

### Hybrid supervised dual-channel feedback U-Net for segmentation of breast ultrasound images

Gong Ronglin, Shi Jun, Wang Jun ..... 2206

### Cooperative suppression network for bimodal data in breast cancer classification

Yang Ziqi, Gong Xun, Zhu Dan, Guo Ying ..... 2218

### Dynamic noise index and segmentation algorithm for breast ultrasound tumor

Zou Haipeng, Gong Xun, Hu Bijie, Luo Jun ..... 2229

## Research and Application

### Multilabel chest X-ray disease classification based on a dense squeeze-and-excitation network

Zhang Zhirui, Li Qiang, Guan Xin ..... 2238

### Morphological design of missing tooth driven by high-resolution deep generation network

Guo Chuang, Dai Ning, Tian Sukun, Sun Yuchun, Yu Qing, Liu Hao, Cheng Xiaosheng ..... 2249

### Arteriovenous classification method in fundus images based on semantic fusion

Gao Yingqi, Guo Song, Li Ning, Wang Kai, Kang Hong, Li Tao ..... 2259

### AS-PANet: a chromosome instance segmentation method based on improved path aggregation network architecture

Lin Chengchuang, Zhao Gansen, Yin Aihua, Ding Bichao, Guo Li, Chen Hanbiao ..... 2271

### TSCNN: a convolutional neural network for the monitoring and analysis of the electrical signals of the heart from wearable devices

Meng Lu, Ge Kang, Song Yang, Yang Dongming ..... 2281

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2020)10-2002-11

论文引用格式: Yang J C and Ni B B. 2020. Advances and challenges in medical 3D computer vision. Journal of Image and Graphics, 25(10): 2002-2012  
(杨健程, 倪冰冰. 2020. 医学 3D 计算机视觉: 研究进展和挑战. 中国图象图形学报, 25(10): 2002-2012) [DOI: 10.11834/jig.200244]

# 医学 3D 计算机视觉: 研究进展和挑战

杨健程, 倪冰冰

上海交通大学电子工程系, 上海 200240

**摘要:** 医学影像的诊断是许多临床决策的基础, 而医学影像的智能分析是医疗人工智能的重要组成部分。与此同时, 随着越来越多 3D 空间传感器的兴起和普及, 3D 计算机视觉正变得越发重要。本文关注医学影像分析和 3D 计算机的交叉领域, 即医学 3D 计算机视觉或医学 3D 视觉。本文将医学 3D 计算机视觉系统划分为任务、数据和表征 3 个层面, 并结合最新文献呈现这 3 个层面的研究进展。在任务层面, 介绍医学 3D 计算机视觉中的分类、分割、检测、配准和成像重建, 以及这些任务在临床诊断和医学影像分析中的作用和特点。在数据层面, 简要介绍了医学 3D 数据中最重要的数据模态: 包括计算机断层成像 (computed tomography, CT)、磁共振成像 (magnetic resonance imaging, MRI)、正电子发射断层成像 (positron emission tomography, PET) 等, 以及一些新兴研究提出的其他数据格式。在此基础上, 整理了医学 3D 计算机视觉中重要的研究数据集, 并标注其数据模态和主要视觉任务。在表征层面, 介绍并讨论了 2D 网络、3D 网络和混合网络在医学 3D 数据的表征学习上的优缺点。此外, 针对医学影像中普遍存在的小数据问题, 重点讨论了医学 3D 数据表征学习中的预训练问题。最后, 总结了目前医学 3D 计算机视觉的研究现状, 并指出目前尚待解决的研究挑战、问题和方向。

**关键词:** 医学影像分析; 3D 计算机视觉; 深度学习; 卷积神经网络 (CNN); 预训练

## Advances and challenges in medical 3D computer vision

Yang Jiancheng, Ni Bingbing

Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

**Abstract:** Medical imaging is an important tool used for medical diagnosis and clinical decision support that enables clinicians to view the internal of human bodies. Medical image analysis, as an important part of healthcare artificial intelligence, provides fast, smart, and accurate decision supports for clinicians and radiologists. 3D computer vision is an emerging research area with the rapid development and popularization of 3D sensors (e.g., light detection and ranging (LIDAR), RGB-D cameras) and computer-aided design in game industry and smart manufacturing. In particular, we focus on the interface of medical image analysis and 3D computer vision called medical 3D computer vision. We introduce the research advances and challenges in medical 3D computer vision in three levels, namely, tasks (medical 3D computer vision tasks), data (data modalities and datasets), and representation (efficient and effective representation learning for 3D images). First, we introduce classification, segmentation, detection, registration, and reconstruction in medical 3D computer vision at the task level. Classification, such as malignancy stratification and symptom estimation, is an everyday task for clinicians and radiologists. Segmentation denotes assigning each voxel (pixel) a semantic label. Detection refers to localizing key objects from medical images. Segmentation and detection include organ segmentation/detection and lesion segmentation/detection. Registration, that is, calculating the spatial transformation from one image to another, plays an impor-

收稿日期: 2020-05-31; 修回日期: 2020-07-07; 预印本日期: 2020-07-14

基金项目: 国家自然科学基金项目 (61976137, U1611461)

Supported by: National Natural Science Foundation of China (61976137, U1611461)

tant role in medical imaging scenarios, such as spatially aligning multiple images from serial examination of a follow-up patient. Reconstruction is also a key task in medical imaging that aims at fast and accurate imaging results to reduce patients' costs. Second, we introduce the important data modalities in medical 3D computer vision, such as computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET), at the data level. The principle and clinical scenario of each imaging modality are briefly discussed. We then depict a comprehensive list of medical 3D image research datasets that cover classification, segmentation, detection, registration, and reconstruction tasks in CT, MRI, and graphics format (mesh). Third, we discuss the representation learning for medical 3D computer vision. 2D convolutional neural networks, 3D convolutional neural networks, and hybrid approaches are the commonly used methods for 3D representation learning. 2D approaches can benefit from large-scale 2D pretraining, triplanar, and trislice 2D representation for 3D medical images, whereas they are generally weak in capturing large 3D contexts. 3D approaches are natively strong in 3D context. However, few publicly available 3D medical datasets are large and sufficiently diverse for universal 3D pretraining. For hybrid (2D + 3D) approaches, we introduce multistream and multistage approaches. Although they are empirically effective, the intrinsic disadvantages within the 2D/3D parts still exist. To address the small-data issues for medical 3D computer vision, we discuss the pretraining approaches for medical 3D images. Pretraining for 3D convolutional neural network (CNN) with videos is straightforward to implement. However, a significant domain gap is found between medical images and videos. Collecting massive medical datasets for pretraining is theoretically feasible. However, it only results in thousands of 3D medical image cases with tens of medical datasets, which is significantly smaller compared with natural 2D image datasets. Research efforts exploring unsupervised (self-supervised) learning to obtain the pretrained 3D models are reported. Although its results are extremely impressive, the model performance of up-to-date unsupervised learning is incomparable with that of fully supervised learning. The unsupervised representation learning from medical 3D images cannot leverage the power of massive 2D supervised learning datasets. We introduce several techniques for 2D-to-3D transfer learning, including inflated 3D (I3D), axial-coronal-sagittal (ACS) convolutions, and AlignShift. I3D enables 2D-to-3D transfer learning by inflating 2D convolution kernels into 3D, and ACS convolutions and AlignShift enable that by introducing novel operators that shuffle the features from 3D receptive fields into a 2D manner. Finally, we discuss several research challenges, problems, and directions for medical 3D computer vision. We first determine the anisotropy issue in medical 3D images, which can be a source of domain gap, that is, between thick- and thin-slice data. We then discuss the data privacy and information silos in medical images, which are important factors that lead to small-data issues in medical 3D computer vision. Federated learning is highlighted as a possible solution for information silos. However, numerous problems, such as how to develop efficient systems and algorithms for federated learning, how to deal with adversarial participants in federated learning, and how to deal with unaligned and missing data, are found. We determine the data imbalance and long tail issues in medical 3D computer vision. Efficient and effective learning of representation from the noisy, imbalanced, and long-tailed real-world data can be extremely challenging in practice because of the imbalanced and long-tailed distributions of real-world patients. We mention the automatic machine learning as a future direction of medical 3D computer vision. With end-to-end deep learning, the development and deployment of medical image application is inapplicable. However, excessive engineering staff need to be tuned for a new medical image task, such as design of deep neural networks, choices of data augmentation, how to perform data preprocessing, and how to tune the learning procedure. The tuning of these hyperparameters can be performed with a hand-crafted or intelligent system to reduce the research efforts by numerous researchers and engineers. Thus, medical 3D computer vision is an emerging research area. With increasing large-scale datasets, easy-to-use and reproducible methodology, and innovative tasks, medical 3D computer vision is an exciting research area that can facilitate healthcare into a novel level.

**Key words:** medical image analysis; 3D computer vision; deep learning; convolutional neural networks (CNN); pre-training

## 0 引言

深度学习在医学影像分析领域中已经取得了巨

大的成功 (Litjens 等, 2017; Shen 等, 2017)。在主流的医学 3D 数据的数据模态和视觉任务上, 如计算机断层成像 (computed tomography, CT)、磁共振成像 (magnetic resonance imaging, MRI) 上的分类、分割和



检测等任务,深度学习已经成为一种最主要的研究方式。深度学习的成功部分取决于大量来自于工业界和学术界贡献的医学数据集,这使得很多基于深度学习的数据驱动方法变成了可能。另外,在深度学习的帮助下,3D 计算机视觉也得到了长足的发展,产生了许多全新的处理 3D 数据的新方法,如体素 (voxel) (Wu 等, 2015)、点云 (point cloud) (Qi Charles 等, 2017)、3D 网格 (mesh) (Hanocka 等, 2019) 等。在医学 3D 数据的分析上,医学影像社区逐渐发展出了一系列 3D 计算机视觉方法。特别将这类研究称为“医学 3D 计算机视觉”或“医学 3D 视觉”。本文尝试对于医学 3D 视觉的研究进行较为全面地总结和分析。如图 1 所示,具体从如下几个方面进行讨论:

1) 任务。介绍医学 3D 视觉具体处理哪些任务,主要介绍分类、分割、检测、配准和成像重建等医学影像分析上最常见的几大任务。

2) 数据。介绍医学 3D 数据中主要的数据模式;同时,针对上述计算机视觉任务,介绍一些比较重要的数据集。

3) 表征。介绍和分析医学 3D 数据的表征学习 (representation learning) 方法,主要介绍和讨论 2D 网络、3D 网络及混合网络的特点。同时,针对医学影像普遍存在的小数据问题介绍一些预训练方法和思路。

最后,针对医学 3D 数据中不同于标准自然图像的特殊问题,介绍现存的研究挑战、问题和方向。

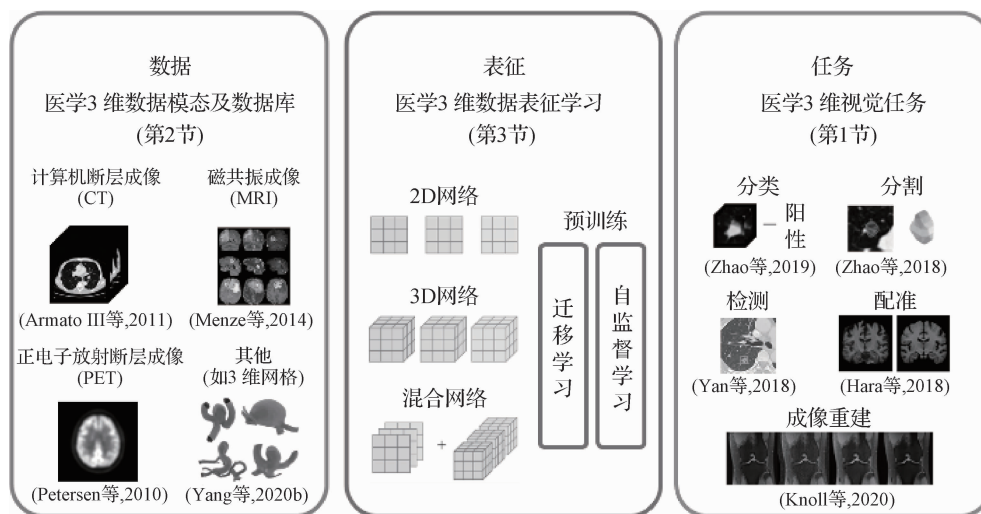


图 1 医学 3D 计算机视觉系统

Fig. 1 Medical 3D computer vision system

## 1 医学 3D 视觉任务

在人工智能的任务层次中,包含感知 (perception)、推理 (reasoning) 和因果 (causality)。在当前的发展阶段,计算机视觉的主要成果集中在感知层面,即分类 (Deng 等, 2009)、分割 (Zhou 等, 2017) 和检测 (Lin 等, 2014) 等。在医学 3D 视觉中,分类、分割和检测等任务都被赋予了一定特殊含义。此外,医学影像配准和成像重建也是医学 3D 视觉中的常见任务。

### 1.1 分类

在医学影像学诊断中,从良恶性风险诊断,到各

种征象的判断,分类通常是临床医生日常进行的高频任务。以肺癌诊断为例,医生通常需要判断相关病灶的风险,此外还需要根据病灶大小、病灶征象 (毛刺、磨玻璃等) 和病灶位置 (是否位于肺门旁) 等综合决定病人的治疗方案 (随访、手术或放化疗等)。因此,在医学影像的任务中,分类任务往往是多标签分类,即尝试对于同一个实体打上不同语义层面的标签。这些标签可以根据放射医生的手工标注 (Armato III 等, 2011), 如良恶性风险、征象; 这些标签也可以基于一些更加深入的医学检查, 如病理分析 (Zhao 等, 2018) 和基因检测 (Zhao 等, 2019) 等。

### 1.2 分割

医学影像的分割是指对医学影像中的每个体

素/像素赋予语义,通常分为器官分割(区分肝脏、气管和冠脉等)和异常分割(区分肝癌、肺结节和冠脉钙化等)。这些像素(或体素)级别的模型输出可以作为后续算法模型的重要构成;且一些场景下,分割结果本身就直接有用,如术前规划的可视化、放化疗的危及器官(organs at risk)分割(Tang等,2019)等。值得一提的是,由于医学影像中前景往往没有过多重叠,因此大部分医学影像分割采用的是语义分割(semantic segmentation)而非实例分割(instance segmentation)。

### 1.3 检测

医学影像检测指的是从较大的视觉范围内定位出单个或多个感兴趣区域(region of interest, ROI)。检测任务是放射科医生日常阅片的核心内容之一,如肺结节检测(Setio等,2017)。在实际的医学筛查和诊断中,只有检测往往不够,还需要伴随后续的分类、分割等任务才可能完成临床实践中的医学诊断报告。另外,针对小目标分割,先检测后分割(类似实例分割)也是一种行之有效的简单策略(Tang等,2019)。

### 1.4 配准

医学影像的配准是医学影像成像和展示中的常见问题,主要指的是将两个结构、纹理等相似的医学影像中各个像素相匹配。配准在很多医学场景是很实用的核心问题,例如针对随访的病人,如何将多次检查自动匹配以节约临床医生的时间,并便于后续的量化测量(体积、密度的变化)。医学影像配准的研究方法中,一个代表性的方法是VoxelMorph(Balakrishnan等,2018),将其将配准问题归约成利用深度神经网络预测两幅(2D或3D)图像之间形变的问题,并借鉴了空间形变网络(Jaderberg等,2015)的思路将之变成一个端到端可微分、无监督的学习框架。

### 1.5 成像重建

如何在降低成像质量的前提下减少病人检查的剂量和负担,如计算机断层成像的辐射量、磁共振所需的检查时间,一直是医学成像中的核心研究问题。由于深度学习等人工智能技术的兴起,数据驱动的方法可以利用数据先验,形成快速、准确的医学影像重建。这类医学影像成像的研究可以视

为计算摄影学在医学影像领域的延伸,如fastMRI(Knoll等,2020)结合人工智能技术,使用更少的成像时间获得更高质量的MRI成像,探索MRI成像中压缩感知(compressed sensing)以外的计算方法。更多研究可以参考Wang等人(2018)的相关成果。

## 2 医学3D数据模态及数据库

在医学诊断影像中,CT和MRI可以说是医学3D成像中两种最常见的影像模态。CT是利用X射线多次扫描(如螺旋扫描)后利用计算机重建算法构成立体成像,常见部位包括胸部CT、腹部CT和脑部CT等。使用显影剂可以拍摄增强CT,对于一些医学诊断更加有利。目前一些部位可以使用低剂量(low-dose)CT来获得足够好的成像,可以大幅降低病人的辐射暴露风险,因此肺癌早期筛查的推荐方式是低剂量CT(The National Lung Screening Trial Research Team,2011)。MRI通过射频脉冲激发人体内氢质子采集k-space原始数据,并通过计算方法复原出人眼可以理解的3D图像。直观地说,MRI不使用放射线,因此对人体的损伤较小,但通常需要较长的成像时间。MRI常用于软组织成像,如脑部、膝部等。另外,基于磁共振成像可以开发多种功能性磁共振成像(functional MRI)用来观测脑部活动。基于CT和MRI,都可以配合正电子放射断层成像(positron emission tomography, PET)(Petersen等,2010),开发出PET-CT和PET-MRI等技术,以观测全身代谢,是用来肿瘤诊断(如肿瘤转移)的一种常用成像手段。上述成像方法形成的原始医学3D数据是基于体素的,符合卷积的归纳偏置先验,因此常常使用卷积神经网络(convolutional neural network, CNN)来分析。

除了上述基于体素的医学3D数据外,Yang等人(2020b)通过半手工整理出了血栓解剖结构3D形状的3D网格(mesh),借此研究基于体素、网格和点云的深度学习方法在医学影像中的性能。

表1整理了部分重要的医学3D数据库,包含了从CT、MRI到图形学格式的分类、分割、检测、配准和成像重建问题。

表 1 部分医学 3D 数据库一览  
Table 1 An overview of 3D medical image datasets (partial)

数据库	数据模态	主要任务	数据量	数据库描述
LIDC-IDRI (the lung image database consortium image collection) (Armato III 等, 2011)	CT	分类 分割 检测	1 018 套标注数据	大规模肺结节数据集, 包含 4 位医生对胸部 CT 的疑似病灶标注, 每套 CT 包含 1~23 个结节标注, 包括病灶检测、病灶 3D 分割、良恶性判断、征象 (如毛刺、钙化等)。
BraTS (Menze 等, 2015)	MRI	分割	274 套标注数据 110 套测试数据	BraTS 2012-2019 挑战赛, 系列脑组织脑肿瘤分割数据集, 其中使用较广泛的是 BraTS 2015 数据集。每个病例包含 FLAIR, T1, T2 和 T1C 这 4 个模态, 以及对应的脑组织和脑部病灶 (包括肿瘤) 分割。江宗康等人 (2020) 提供了 BraTS 数据集及 MRI 脑肿瘤图像分割的研究综述。
LUNA16 (Setio 等, 2017)	CT	检测	888 套标注数据	大规模肺结节检测数据集, 是 LIDC-IDRI 的子集, 挑选了病灶较为显著、符合结节标准的胸部 CT 和病灶标注。
VoxelMorph (Balakrishnan 等, 2018)	MRI	配准	合计 7 829 套数据	大规模多中心数据集, 由作者搜集 8 个公开数据集整理而成, 其用于学习配准的配对数据由该数据集添加形变生成。
DeepLesion (Yan 等, 2018)	CT	检测 弱分割 分类	32 120 套标注数据 (关键帧)	大规模通用病灶检测数据集, 包含全身各种部位 CT 中带有标注的关键帧及其上下 30 mm 的相邻帧。除了病灶检测标注外, 提供了病灶的 RECIST (response evaluation criteria in solid tumors) 标注 (即病灶的长径短径), 可以用来生成弱标注的病灶分割。每张关键帧包含 1~3 个标注。后续数据集增添了部分分类标签。
LiTS (the liver tumor segmentation benchmark) (Bilic 等, 2019)	CT	分割	131 套标注数据 70 套测试数据	LiTS 2017 挑战赛, 肝脏肝癌分割数据集, 包含标注了肝脏和肝癌区域 3D 分割的腹部 CT。另有部分临床信息。
MSD (medical segmentation decathlon) (Simpson 等, 2019)	MRI CT	分割	合计 1 746 套标注数据 888 套测试数据	大规模医学 3D 数据分割数据集, 包含多种模态 (MRI 和 CT)、各种部位、各种数据规模的 10 个医学 3D 数据分割数据集。该数据集也是一个持续进行的挑战赛, 要求参赛者使用一套不需要人工调整参数的算法来完成多套数据的建模。
KiTS (Heller 等, 2019)	CT	分割	210 套标注数据 90 套测试数据	KiTS 2019 挑战赛, 肾脏肾癌分割数据集, 包含标注了肾脏和肾癌区域 3D 分割的腹部增强 CT。另有部分临床信息。
fastMRI (Knoll 等, 2020)	MRI	成像重建	1 594 套配对数据	大规模膝部 MRI 成像重构数据集, 将低质量的 MRI 图像 (降采样 k-space) 重构成高质量的 MRI (全采样 k-space)。除了已经重建好的配对 MRI 数据 (DICOM 格式) 外, 该数据集还包含了原始的 k-space 数据。该数据也是一个持续进行的挑战赛, 后续还增加了大规模的脑部 MRI 数据。
IntrA (Yang 等, 2020b)	3D 网格 (mesh)	分类 分割	1 909 套标注数据	颅内动脉瘤数据集, 整理了由医生手动分割的血管及动脉瘤的分割, 每个分割经过手工清洗后转换成了 3D 网格 (计算机图形学的基本格式), 因此该数据集可以使用基于点云、网格和体素的深度学习方法来处理。
RibFrac (RibFrac Team, 2020)	CT	检测 分割 分类	500 套标注数据 120 套测试数据	RibFrac 2020 挑战赛, 肋骨骨折检测、分割和检测数据集, 包含标注了肋骨骨折病灶像素级 3D 分割的胸部 CT, 并经由专家给出骨折类型。可以用来开发 3D 检测、实例分割和分类算法。挑战赛评估检测和分类性能。



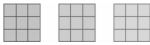

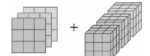
### 3 医学3D数据表征学习

基于深度学习的方法学习医学3D数据的表征(representation),具体讨论基于2D卷积神经网络、3D卷积神经网络和混合网络的方法,并分析

各自的优缺点(表2)。考虑到医学数据中普遍存在的小数据问题,介绍在医学3D数据的表征学习中引入预训练以提高学习效率,包含了基于时空数据/3D数据的直接预训练、基于医学3D的无监督预训练和部分2D到3D迁移学习方法。

表2 标准卷积神经网络在处理医学3D数据上的比较

Table 2 Comparison of standard convolutional neural networks for 3D medical image analysis

网络类型	优点	缺点
2D 网络 	直接利用大规模2D数据作为预训练。	2D表征,难以学习复杂的3D上下文信息。
3D 网络 	3D表征,可以学习复杂的3D上下文信息。缺少直接可用的大规模医学3D数据作为预训练。	
混合网络 (2D网络+3D网络) 	兼有2D表征和3D表征。	1) 2D网络部分仍然是2D表征; 2) 3D网络部分仍无合理的预训练; 3) 解决方案类似于模型集成; 其多阶段/多分支的网络通常比较冗余。

#### 3.1 2D网络

尽管医学3D数据是3D空间中的图像,直接使用2D网络逐帧处理也是一种行之有效的方法。好处是可以直接使用在大规模2D数据集(如ImageNet(Deng等,2009))上训练出的2D卷积神经网络。为了模仿自然图像上的RGB输入,可以将单幅灰度图像直接重复,也可以将3D医学影像以多平面或连续帧的2D图像拼接作为2D输入,后者有时也称为2.5D方法。早期的深度学习研究(Prasoon等,2013)使用多平面的表示方法,即医学3D数据以某个中心点展开横断位(axial)、冠状位(coronal)和矢状位(sagittal)3个平面输入2D卷积神经网络中。尽管这种方法在实践中有一定的效果,但是本质上并不完全符合卷积网络的归纳偏置,因为作为通道的3个平面在空间上并非对齐的。后续的大部分研究(Han,2017)采用连续帧的2D表示,即以中间帧并取上下连续帧作为2D输入的通道。

基于2D神经网络的方法虽然可以直接使用大规模的2D图像数据集作为预训练,但从本质上来说难以学习到复杂的3D上下文信息,尤其无法处理一些长范围的3D上下文信息(如血管等)。

#### 3.2 3D网络

除了将3D上下文信息作为通道输入的2D神经网络的方法,有很多研究直接使用3D卷积神经

网络来处理医学3D数据(Çiçek等,2016)。相较于2D网络方法,3D网络具有捕捉大范围3D空间上下文信息的优势,感受野随着卷积和池化在3D空间中增长,这对于区分小器官、小病灶和血管等需要长范围3D上下文信息的对象来说是至关重要的。

然而,由于直接使用纯3D神经网络,这些方法往往缺少像2D网络那样的预训练模型。由于数据稀缺、数据孤岛、标注困难和病种罕见等问题,3D图像的天生匮乏,并且从经济上、实操上都难以建立大规模的医学3D数据集。因此,很多研究重点放在了如何高效进行深度学习,包括深监督(Dou等,2017)和紧凑的网络设计(Zhou等,2018)等。但值得注意的是,这些方法往往并非专门针对3D网络设计,即并未直击3D网络方法的痛点。

#### 3.3 混合网络

为了解决2D网络和3D网络独立使用时无法调和的缺陷,部分研究将2D网络和3D网络混合使用,这类方法称为混合网络方法,通常包含一个预训练的2D卷积神经网络来处理医学3D数据连续帧的2D输入,还包括一个随机初始化的3D网络用来直接处理医学3D数据的体素输入。其中,2D网络部分和3D网络部分可以分别优化也可以联合优化,可以级联也可以并联。

级联方法也称为多阶段方法,通常包含多个2D



网络(甚至3D网络),之后紧接一个3D融合网络(Xia等,2018)。并联方法也可称为多分支方法,通常包含多个2D网络和3D网络进行相同的目标任务(或提取特征),之后再进行融合(Zheng等,2019)。

尽管在实践中,混合网络方法是有效的,但是这种简单的“2D+3D”的方法更像是一种简单的集成学习,并没有真正解决2D网络和3D网络各自的痛点。另外,在真实的工程环境中,往往也不允许使用这样的大规模混合网络来部署。

### 3.4 预训练方法

在计算机视觉中,针对目标任务数据量小的情况,往往会使用从大规模数据上预训练后迁移学习的方式提高小数据目标任务上的性能。然而,正如上文所述,能够直接使用2D图像数据集预训练的2D网络方法并不能很好地学习3D上下文信息,而可以学习3D上下文信息的3D网络缺少大规模的通用预训练。

前期研究讨论了针对3D网络的预训练,如使用视频序列(Hara等,2018)和集合多个医学3D数据的预训练(Chen等,2019)。考虑到视频数据的易得性和多样性,使用视频序列进行预训练对于3D网络的学习很有帮助。然而,其迁移到医学3D数据的任务上时性能往往并不如人意,这主要来源于两方面的差异:1)自然图像和医学影像的差异;2)时空的视频序列和3D空间的医学影像的差异。尽管医学3D数据中也存在较明显的各向异性,但并不显式存在视频序列中特殊的轴向(时间)。因此,Med3D(Chen等,2019)探索了集合多个医学影像数据集的预训练方式。即使集合了多个数据集,其关联的样本也仅有1 000多例,这和自然图像、视频序列动辄百万的数据集上的预训练是难以比较的,因此其预训练效果作用也比较有限。

除了上述的监督学习方式,Models Genesis(Zhou等,2019)探索了在医学3D数据上进行无监督学习的表征学习方式。通过一种类似于去噪自动编码器的学习方式,在多个数据集上验证了其无监督学习表征的有效性。然而,这种方法并不能显式地使用已有的大规模2D图像通用数据集。并且,监督学习的预训练往往比无监督学习的预训练更加有效。

另外,一些研究探索如何将2D卷积神经网络学习到的知识(权重),转换成能够学习3D表征的神经网络。I3D(inflated 3D)(Carreira和Zisserman,2017)提供了一种将2D卷积神经网络的权重直接转换成3D卷积神经网络权重的方法。其基本思路是利用卷积的线性,将2D的卷积核“扩充”成3D的卷积核(沿轴重复并等比例缩小数值)。除此之外,其他研究通过提出全新的卷积算子,包括ACS(axial-coronal-sagittal)卷积(Yang等,2019)和AlignShift(Yang等,2020a)来让2D到3D的迁移学习成为可能,其核心是通过不同卷积通道聚集3D卷积的感受野范围内部分位置的特征。

表3按照时间顺序总结了上述各种适用于医学3D数据的预训练方法。

## 4 结 语

### 4.1 总结

在基于深度学习的人工智能和计算机视觉技术的快速发展下,医学3D计算机视觉,主要指医学3D数据的分析和理解,得到了长足的发展,以至于深度学习已经成为医学3D计算机视觉中最核心的研究方式之一。本文重点介绍了医学3D计算机视觉系统的3个层面:任务、数据和表征。在任务层面,介绍了医学3D数据中的分类、分割、检测、配准和成像重建等任务。在数据层面,主要介绍了CT、MRI、PET等一系列医学3D数据模态,并整理一些重要的研究数据集。在表征层面,介绍了基于2D网络、3D网络和混合网络的表征学习方法各自的优缺点。同时,考虑到医学3D数据分析中广泛存在的小数据问题,介绍了在医学3D数据表征学习中可能的预训练方法。

### 4.2 未来挑战与研究机会

尽管目前医学3D计算机视觉已经获得了长足发展,甚至在某些领域已经达到了基本可用的程度,但仍然有很多的研究问题尚未充分解决。

1)各向异性。由于成像设备的限制,医学3D数据中往往存在比较明显的各向异性。具体而言,医学3D数据中横断面的间距和其他两个方向(冠状面和矢状面)的间距往往有较大的差别,这使得这些医学影像各个平面的分析存在一定区别。其中,横断面之前的间距通常称为层厚(thickness),根

表 3 医学 3D 数据的深度学习预训练方法比较  
Table 3 Comparison of pretraining in deep learning for 3D medical image analysis

方法简述 / 名称	卷积算子	预训练源	数据规模	数据多样性	是否监督预训练	是否特定轴向
不进行预训练	2D 卷积	—	—	—	—	是
不进行预训练	3D 卷积	—	—	—	—	否
2D 数据预训练直接迁移	2D 卷积	2D 图像	非常大	非常大	是	是
2D 数据预训练扩展权重 / I3D (Carreira 和 Zisserman, 2017)	3D 卷积	2D 图像	非常大	非常大	是	是
视频数据预训练直接迁移 / 3D CNNs (Hara 等, 2018)	3D 卷积	视频序列	较大	较大	是	是
医学数据预训练直接迁移 / Med3D (Chen 等, 2019)	3D 卷积	医学 3D 数据	中等	中等	是	否
3D 数据无监督直接迁移 / Models Genesis (Zhou 等, 2019)	3D 卷积	医学 3D 数据	较大	中等	否	否
2D 数据预训练修改算子 / ACS 卷积 (Yang 等, 2019)	特殊算子	2D 图像 (亦视频序列或医学 3D 数据)	非常大	非常大	是 (亦无监督训练)	否
2D 数据预训练添加算子 / AlignShift (Yang 等, 2020a)	特殊算子	2D 图像 (亦视频序列或医学 3D 数据)	非常大	非常大	是 (亦无监督训练)	是

注: 此处主要讨论提及的原始方法在处理医学 3D 数据时的情形, 不讨论其各自变种, “—”表示不适合进行分类。

据层厚可以将医学 3D 数据分为薄层图像和厚层图像。在实际的医学 3D 数据分析中, 薄层和厚层因其自身差异, 会导致同一套模型难以在不同层厚的图像上泛化, 这在实际部署中会严重影响模型的适用范围。目前已有一些研究 (Yang 等, 2020a) 考虑如何弥合模型在不同层厚的数据上的性能, 但各向异性的问题属于医学 3D 数据中的基本问题, 这个方向值得更多探索。

2) 数据孤岛。开展实际医学人工智能相关的研究时, 往往会面临数据孤岛问题。事实上, 医学数据非常丰富, 但研究者能够访问到的数据相比自然图像非常稀缺。重要原因是高质量的医学影像往往停留在一些高水平机构 (如三甲医院) 中, 而这些机构的数据考虑到数据隐私等问题, 彼此是孤立的, 从而导致数据孤岛问题进一步加剧了医学影像分析中小数据的问题。对此, 一个解决方法是联邦学习 (Bonawitz 等, 2019): 让数据停留在各个中心, 仅依靠通信来传输各个中心的数据上学习到的模型来学习全局的模型。然而, 目前联邦学习在真实使用中仍有非常多的挑战, 包括如何设计高效的联邦学习系统和算法, 如何处理恶意的联邦学习参与者, 如何考虑数据缺失和对齐问题, 如何解决不同中心的域差异 (domain gap) 等。

3) 长尾/不平衡问题。由于医学疾病本身的长

尾效应, 医学影像的标注往往存在极端的类别不平衡现象 (Yan 等, 2019)。在实践中, 直接在极端不平衡的数据上进行学习, 会导致模型的表征学习或长尾数据的分类学习存在较大的偏差。随着医学 3D 计算机视觉的发展和广泛部署, 该问题会变得越来越重要。

4) 自动机器学习。在医学 3D 计算机视觉的算法任务中, 尽管依靠自动端到端学习的深度学习方法, 但仍有很大一部分工作是需要研究者和工程师人为指定的, 如网络架构选择、优化策略、预处理和数据增强等。事实上, 很多工程设计完全可以让机器自行决定。如 Isensee 等人 (2018) 提出了一套自动机器学习架构, 完全根据交叉验证来自行决定在模型上使用的工程参数。该方法取得了令人瞩目的表现, 但存在计算过大、人为设计过多等问题, 仍有较大的发展空间。

参考文献 (References)

Armato III S G, McLennan G, Bidaut L, McNitt-Gray M F, Meyer C R, Reeves A P, Zhao B S, Aberle D R, Henschke C I, Hoffman E A, Kazerooni E A, MacMahon H, Van Beek E J R, Yankelevitz D, Biancardi A M, Bland P H, Brown M S, Engelmann R M, Laderach G E, Max D, Pais R C, Qing D P Y, Roberts R Y, Smith A R, Starkey A, Batra P, Caligiuri P, Farooqi A, Gladish G W, Jude

- C M, Munden R F, Petkovska I, Quint L E, Schwartz L H, Sundaram B, Dodd L E, Fenimore C, Gur D, Petrick N, Freymann J, Kirby J, Hughes B, Vande Casteele A, Gupte S, Sallam M, Heath M D, Kuhn M H, Dharaiya E, Burns R, Fryd D S, Salganicoff M, Anand V, Shreter U, Vastagh S, Croft B Y and Clarke L P. 2011. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical Physics*, 38 (2): 915-931 [DOI: 10.1118/1.3528204]
- Balakrishnan G, Zhao A, Sabuncu M R, Dalca A V and Gutttag J. 2018. An unsupervised learning model for deformable medical image registration//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 9252-9260 [DOI: 10.1109/CVPR.2018.00964]
- Bilic P, Christ P F, Vorontsov E, Chlebus G, Chen H, Dou Q, Fu C W, Han X, Heng P A, Hesser J, Kadoury S, Konopczynski T, Le M, Li C M, Li X M, Lipková J, Lowengrub J, Meine H, Moltz J H, Pal C, Piraud M, Qi X J, Qi J, Rempfler M, Roth K, Schenk A, Sekuboyina A, Vorontsov E, Zhou P, Hülsemeyer C, Beetz M, Ettlinger F, Gruen F, Kaissis G, Lohöfer F, Braren R, Holch J, Hofmann F, Sommer W, Heinemann V, Jacobs C, Mamani G E H, Van Ginneken B, Chartrand G, Tang A, Drozdal M, Cohen A B, Klang E, Amitai M M, Konen E, Greenspan H, Moreau J, Hostettler A, Soler L, Vivanti R, Szeskin A, Lev-Cohain N, Sosna J, Joskowicz L and Menze B H. 2019. The liver tumor segmentation benchmark (LITS)[EB/OL]. [2020-05-01]. <https://arxiv.org/pdf/1901.04056v1.pdf>
- Bonawitz K, Eichner H, Grieskamp W, Huba D, Ingerman A, Ivanov V, Kiddon C, Konečný J, Mazzocchi S, McMahan B, Van Overveldt T, Petrou D, Ramage D and Roselander J. 2019. Towards federated learning at scale: system design[EB/OL]. [2020-05-01]. <https://arxiv.org/pdf/1902.01046.pdf>
- Carreira J and Zisserman A. 2017. Quo vadis, action recognition? A new model and the kinetics dataset//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE: 4724-4733 [DOI: 10.1109/CVPR.2017.502]
- Chen S H, Ma K and Zheng Y F. 2019. Med3D: transfer learning for 3D medical image analysis[EB/OL]. [2020-05-01]. <https://arxiv.org/pdf/1904.00625.pdf>
- Çiçek Ö, Abdulkadir A, Lienkamp S S, Brox T and Ronneberger O. 2016. 3D U-Net: learning dense volumetric segmentation from sparse annotation//Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens: Springer: 424-432 [DOI: 10.1007/978-3-319-46723-8\_49]
- Deng J, Dong W, Socher R, Li L J, Li K and Li F F. 2009. ImageNet: a large-scale hierarchical image database//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE: 248-255 [DOI: 10.1109/CVPR.2009.5206848]
- Dou Q, Yu L Q, Chen H, Jin Y M, Yang X, Qin J and Heng P A. 2017. 3D deeply supervised network for automated segmentation of volumetric medical images. *Medical Image Analysis*, 41: 40-54 [DOI: 10.1016/j.media.2017.05.001]
- Han X. 2017. Automatic liver lesion segmentation using a deep convolutional neural network method[EB/OL]. [2020-05-31]. <https://arxiv.org/pdf/1704.07239.pdf>
- Hanocka R, Hertz A, Fish N, Giryas R, Fleishman S and Cohen-Or D. 2019. MeshCNN: a network with an edge. *ACM Transactions on Graphics*, 38(4): #90 [DOI: 10.1145/3306346.3322959]
- Hara K, Kataoka H and Satoh Y. 2018. Can spatiotemporal 3D CNNs retrace the history of 2D CNNs and ImageNet? //Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 6546-6555 [DOI: 10.1109/CVPR.2018.00685]
- Heller N, Sathianathan N, Kalapara A, Walczak E, Moore K, Kaluzniak H, Rosenberg J, Blake P, Rengel Z, Oestreich M, Dean J, Tradedwell M, Shah A, Tejpal R, Edgerton Z, Peterson M, Raza S, Regmi S, Papanikolopoulos N and Weight C. 2019. The KiTS19 challenge data: 300 kidney tumor cases with clinical context, CT semantic segmentations, and surgical outcomes[EB/OL]. [2020-05-01]. <https://arxiv.org/pdf/1904.00445.pdf>
- Isensee F, Petersen J, Klein A, Zimmerer D, Jaeger P F, Kohl S, Wasserthal J, Koehler G, Norajitra T, Wirkert S and Maier-Hein K H. 2018. nnU-Net: self-adapting framework for U-Net-based medical image segmentation[EB/OL]. [2020-05-01]. <https://arxiv.org/pdf/1809.10486.pdf>
- Jaderberg M, Simonyan K and Zisserman A. 2015. Spatial transformer networks//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: ACM: 2017-2025
- Jiang Z K, Lyu X G, Zhang J X, Zhang Q and Wei X P. 2020. Review of deep learning methods for MRI brain tumor image segmentation. *Journal of Image and Graphics*, 25(2): 215-228 (江宗康, 吕晓钢, 张建新, 张强, 魏小鹏. 2020. MRI 脑肿瘤图像分割的深度学习方法综述. *中国图象图形学报*, 25(2): 215-228) [DOI: 10.11834/jig.190173]
- Knoll F, Zbontar J, Sriram A, Muckley M J, Bruno M, Defazio A, Pariente M, Geras K J, Katsnelson J, Chandarana H, Zhang Z Z, Drozdal M, Romero A, Rabbat M, Vincent P, Pinkerton J, Wang D, Yakubova N, Owens E, Zitnick C L, Recht M P, Sodickson D K and Lui Y W. 2020. fastMRI: a publicly available raw k-space and DICOM dataset of knee images for accelerated MR image reconstruction using machine learning. *Radiology: Artificial Intelligence*, 2(1): #190007 [DOI: 10.1148/ryai.2020190007]
- Lin T Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P and Zitnick C L. 2014. Microsoft COCO: common objects in context//Proceedings of the 13th European Conference on Computer Vision. Zurich: Springer: 740-755 [DOI: 10.1007/978-3-319-10602-1\_48]

- Litjens G, Kooi T, Bejnordi B E, Setio A A A, Ciompi F, Ghafoorian M, Van Der Laak J A W M, Van Ginneken B and Sánchez C I. 2017. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42: 60-88 [DOI: 10.1016/j.media.2017.07.005]
- Menze B H, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R, Lanczi L, Gerstner E, Weber M A, Arbel T, Avants B B, Ayache N, Buendia P, Collins D L, Cordier N, Corso J J, Criminisi A, Das T, Delingette H, Demiralp Ç, Durst C R, Dojat M, Doyle S, Festa J, Forbes F, Geremia E, Glocker P, Golland P, Guo X T, Hamamci A, Iftekharruddin K M, Jena R, John N M, Konukoglu E, Lashkari D, Mariz J A, Meier R, Pereira S, Precup D, Price S J, Raviv T R, Reza S M S, Ryan M, Sarikaya D, Schwartz L, Shin H C, Shotton J, Silva C A, Sousa N, Subbanna N K, Szekely G, Taylor T J, Thomas O M, Tustison N J, Unal G, Vasseur F, Wintermark M, Ye D H, Zhao L, Zhao B S, Zikic D, Prastawa M, Reyes M and Van Leemput K. 2015. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10): 1993-2024 [DOI: 10.1109/TMI.2014.2377694]
- Petersen R C, Aisen P S, Beckett L A, Donohue M C, Gamst A C, Harvey D J, Jack C R, Jagust W J, Shaw L M, Toga A W, Trojanowski J Q and Weiner M W. 2010. Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology*, 74(3): 201-209 [DOI: 10.1212/WNL.0b013e3181cb3e25]
- Prasoon A, Petersen K, Igel C, Lauze F, Dam E and Nielsen M. 2013. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network//Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention. Nagoya: Springer: 246-253 [DOI: 10.1007/978-3-642-40763-5\_31]
- Qi Charles R, Su H, Mo K C and Guibas L J. 2017. PointNet: deep learning on point sets for 3D classification and segmentation//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE: 77-85 [DOI: 10.1109/CVPR.2017.16]
- RibFrac Team. 2020. MICCAI 2020 RibFrac challenge: rib fracture detection and classification[EB/OL]. [2020-05-01]. <https://ribfrac.grand-challenge.org>
- Setio A A A, Traverso A, De Bel T, Berens M S N, Van Den Bogaard C, Cerello P, Chen H, Dou Q, Fantacci M E, Geurts B, Van Der Gugten R, Heng P A, Jansen B, De Kaste M M J, Kotov V, Lin J Y H, Manders J T M C, Sónora-Mengana A, García-Naranjo J C, Papavasileiou E, Prokop M, Saletta M, Schaefer-Prokop C M, Scholten E T, Scholten L, Snoeren M M, Torres E L, Vandemeulebroucke J, Walasek N, Zuidhof G C A, Van Ginneken B and Jacobs C. 2017. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images; the LUNA16 challenge. *Medical Image Analysis*, 42: 1-13 [DOI: 10.1016/j.media.2017.06.015]
- Shen D G, Wu G R and Suk H I. 2017. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19: 221-248 [DOI: 10.1146/annurev-bioeng-071516-044442]
- Simpson A L, Antonelli M, Bakas S, Bilello M, Farahani K, Van Ginneken B, Kopp-Schneider A, Landman B A, Litjens G, Menze B, Ronneberger O, Summers R M, Bilic P, Christ P F, Do R K G, Gollub M, Golia-Pernicka J, Heckers S H, Jarnagin W R, McHugo M K, Napel S, Vorontsov E, Maier-Hein L and Cardoso M J. 2019. A large annotated medical image dataset for the development and evaluation of segmentation algorithms[EB/OL]. [2020-05-01]. <https://arxiv.org/pdf/1902.09063.pdf>
- Tang H, Chen X M, Liu Y, Lu Z P, You J H, Yang M Z, Yao S Y, Zhao G Q, Xu Y, Chen T F, Liu Y and Xie X H. 2019. Clinically applicable deep learning framework for organs at risk delineation in CT images. *Nature Machine Intelligence*, 1(10): 480-491 [DOI: 10.1038/s42256-019-0099-z]
- The National Lung Screening Trial Research Team. 2011. Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine*, 365(5): 395-409 [DOI: 10.1056/NEJMoa1102873]
- Wang G, Ye J C, Mueller K and Fessler J A. 2018. Image reconstruction is a new frontier of machine learning. *IEEE Transactions on Medical Imaging*, 37(6): 1289-1296 [DOI: 10.1109/TMI.2018.2833635]
- Wu Z R, Song S R, Khosla A, Yu F, Zhang L G, Tang X O and Xiao J X. 2015. 3D ShapeNets: a deep representation for volumetric shapes//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE: 1912-1920 [DOI: 10.1109/CVPR.2015.7298801]
- Xia Y D, Xie L X, Liu F Z, Zhu Z T, Fishman E K and Yuille A L. 2018. Bridging the gap between 2D and 3D organ segmentation with volumetric fusion net//Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention. Granada: Springer: 445-453 [DOI: 10.1007/978-3-030-00937-3\_51]
- Yan K, Peng Y F, Sandfort V, Bagheri M, Lu Z Y and Summers R M. 2019. Holistic and comprehensive annotation of clinically significant findings on diverse CT images: learning from radiology reports and label ontology//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE: 8515-8524 [DOI: 10.1109/CVPR.2019.00872]
- Yan K, Wang X S, Lu L, Zhang L, Harrison A P, Bagheri M and Summers R M. 2018. Deep lesion graphs in the wild: relationship learning and organization of significant radiology image findings in a diverse large-scale lesion database//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 9261-9270 [DOI: 10.1109/CVPR.2018.00965]
- Yang J C, He Y, Huang X Y, Xu J W, Ye X D, Tao G Y and Ni B B.



- 2020a. AlignShift: bridging the gap of imaging thickness in 3D anisotropic volumes [EB/OL]. [2020-05-01]. <https://arxiv.org/pdf/2005.01969/pdf>
- Yang J C, Huang X Y, Ni B B, Xu J W, Yang C Q and Xu G Z. 2019. Reinventing 2D convolutions for 3D images [EB/OL]. [2020-05-01]. <https://arxiv.org/pdf/1911.10477/pdf>
- Yang X, Xia D, Kin T and Igarashi T. 2020b. IntraA: 3D intracranial aneurysm dataset for deep learning [EB/OL]. [2020-05-01]. <https://arxiv.org/pdf/2003.02920v1.pdf>
- Zhao W, Yang J C, Ni B B, Bi D X, Sun Y L, Xu M D, Zhu X X, Li C, Jin L, Gao P, Wang P J, Hua Y Q and Li M. 2019. Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning. *Cancer Medicine*, 8(7): 3532-3543 [DOI: 10.1002/cam4.2233]
- Zhao W, Yang J C, Sun Y L, Li C, Wu W L, Jin L, Yang Z M, Ni B B, Gao P, Wang P J, Hua Y Q and Li M. 2018. 3D deep learning from CT scans predicts tumor invasiveness of subcentimeter pulmonary adenocarcinomas. *Cancer Research*, 78(24): 6881-6889 [DOI: 10.1158/0008-5472.CAN-18-0696]
- Zheng H, Zhang Y Z, Yang L, Liang P X, Zhao Z, Wang C L and Chen D Z. 2019. A new ensemble learning framework for 3D biomedical image segmentation//Proceedings of 2019 AAAI Conference on Artificial Intelligence. Honolulu: AAAI: 5909-5916 [DOI: 10.1609/aaai.v33i01.33015909]
- Zhou B L, Zhao H, Puig X, Fidler S, Barriuso A and Torralba A. 2017. Scene parsing through ADE20K dataset//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE: 5122-5130 [DOI: 10.1109/CVPR.2017.544]
- Zhou Z W, Siddiquee M M R, Tajbakhsh N and Liang J M. 2018. UNet++: a nested U-Net architecture for medical image segmentation//Proceedings of the 4th Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Granada: Springer: 3-11 [DOI: 10.1007/978-3-030-00889-5\_1]
- Zhou Z W, Sodha V, Siddiquee M M R, Feng R B, Tajbakhsh N, Gotway M B and Liang J M. 2019. Models genesis: generic autodidactic models for 3D medical image analysis//Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen: Springer: 384-393 [DOI: 10.1007/978-3-030-32251-9\_42]

### 作者简介



杨健程, 1993年生, 男, 博士研究生, 主要研究方向为医学影像分析、3D 计算机视觉、可信人工智能。

E-mail: jekyll4168@sjtu.edu.cn



倪冰冰, 通信作者, 男, 教授, 博士生导师, 主要研究方向为计算机视觉、机器学习与多媒体计算。

E-mail: nibingbing@sjtu.edu.cn