



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

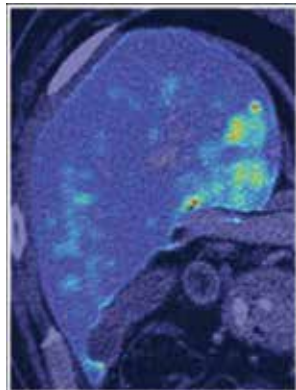
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

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

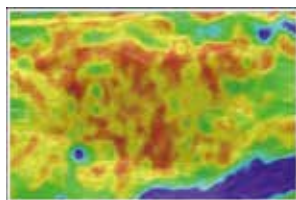
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



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

许正玺, 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

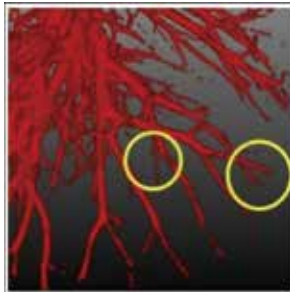
肖照林, 杨志林, 刘欢, 金海燕 0988

基于图像的介入器械前端变形建模与受力估计

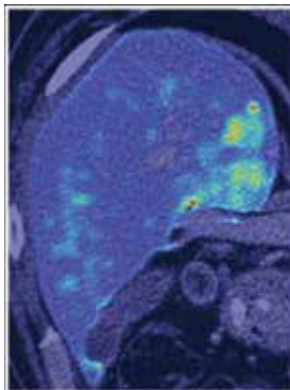
任龙飞, 孟偲, 刘博, 白相志, 周付根 1001

CONTENTS

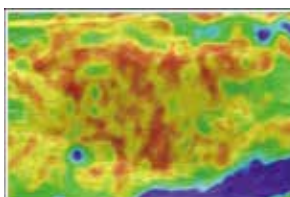
JOURNAL OF IMAGE AND GRAPHICS



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: TP391 文献标识码: A 文章编号: 1006-8961(2022)03-0862-11

论文引用格式: Xu J C and Xiao Z Y. 2022. A spatio-temporal multi-scale network for cardiac dynamic MRI image segmentation. Journal of Image and Graphics, 27(03):0862-0872(徐佳陈, 肖志勇. 2022. 心脏动态 MRI 图像分割的时空多尺度网络. 中国图象图形学报, 27(03):0862-0872) [DOI: 10.11834/jig.200783]

心脏动态 MRI 图像分割的时空多尺度网络

徐佳陈, 肖志勇*

江南大学人工智能与计算机学院, 无锡 214122

摘要: **目的** 心脏磁共振成像(magnetic resonance imaging, MRI)的自动分割技术有利于在临床诊断中评估心脏的功能参数。然而由于心脏磁共振成像技术产生的图像边界不清晰、各向分辨率异性等特性, 现有的大多数方法依旧存在类内不确定、类间不清晰问题。针对这一问题, 提出了一种利用时间信息进行特征增强, 并利用空间信息进行特征矫正的多输入、多分支和多任务的分割网络(spatio-temporal UNet, ST UNet)。**方法** 为充分获取动态心脏 MRI 图像的时间信息, 提出了全新的时间增强编码模块, 将需要进行分割的目标帧和一段包含了目标帧的连续时间片段作为关键序列一同输入网络。关键序列用于获取更丰富的时间特征, 目标帧提供更精准的边缘特征。为了聚集更多有益的特征, 更好地融合时域特征和边缘特征, 采用可变形全局连接代替传统的长连接, 为网络的解码部分提供更广泛的多维特征信息。在训练过程中额外学习空间方向场特征, 并使用该特征对原有的分割结果进行矫正。**结果** 在 ACDC(Automated Cardiac Diagnosis Challenge)心脏分割挑战中, 以 Dice 系数和 HD(Hausdorff distance)距离为评价指标, 该方法在左心室、右心室和左心肌分割的平均 Dice 系数分别为 95%、91.5% 和 91%, HD 距离的平均值分别为 6.77、11.39 和 8.54。**结论** 实验表明, 提出的新型网络能够充分地利用心脏 MRI 图像的时空信息, 有效地提升目标器官的分割效果, 更有助于医生对于心脏诊断。

关键词: 医学影像; 心脏磁共振成像(CMRI); 特征增强; 空间方向场; U 型网络(UNet)

A spatio-temporal multi-scale network for cardiac dynamic MRI image segmentation

Xu Jiachen, Xiao Zhiyong*

School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China

Abstract: Objective The segmentation of cardiac dynamic magnetic resonance imaging (MRI) is essential to evaluate cardiac functional parameters in clinical diagnosis. Based on the segmentation results, the qualified analysis can be obtained the myocardial mass and thickness, ejection fraction, ventricular volume and other diagnostic indicators effectively. Currently, the method of heart segmentation is still limited to manual segmentation. This method is time-consuming and human behavior-oriented. Therefore, the issue of automatic and accurate cardiac-MRI (CMRI) segmentation has been focused on. However, due to the non-uniform magnetic field intensity, artifacts resulted blurred boundaries in the process of imaging. Organs of different subjects vary greatly, especially the variable shape and size of the right ventricle, which tends to produce volume effect. In addition, dynamic magnetic resonance imaging is featured based on a small scale of short axis sec-

收稿日期: 2020-12-21; 修回日期: 2021-04-22; 预印本日期: 2021-04-29

* 通信作者: 肖志勇 zhiyong.xiao@jiangnan.edu.cn

基金项目: 江苏省自然科学基金优秀青年项目 (BK20190079)

Supported by: Natural Science Foundation of Jiangsu Province for Excellent Young Scholars (BK20190079)

tions with thick slice, which results in low short axis resolution and sparse information of the image. As the existing data sets only including ground truth of dynamic MRI images at two sites in the end of systole and end of diastole, the existing networks usually only take the images at these two sites as segmentation objects, which ignores the information of dynamic MRI images timescale. Hence, automatic segmentation of dynamic cardiac MRI images is challenged of the issues of intra-class uncertainty and inter-class imbalance. This illustration demonstrates a spatio-temporal multi-scale UNet that makes use of time information to conduct feature enhancement and spatial information to get feature correction. **Method** First, a new time-enhanced coding path is developed in order to fully obtain the time information of dynamic CMRI images, which consists of two branches are those of the target frame branch and the key sequence branch. The target frame is the image to be segmented, and the key sequence is the consistent time series containing the target frame. The key sequence is used to obtain richer time features, while the target frame provides more accurate edge features. Due to the beating frequency of the heart, the boundary information of the target organ would be in conflict with that of the time scaled images in the key sequence, and the edge information extracted from the target frame could make up for this problem. It is worth noting that the key sequence and target frame are cut or filled into 256×256 pixels due to the different sizes of images in the data set. Simultaneously, this analysis also makes the same data enhancement for the target frame and the key sequence and random affine changing and random rotation in order to extend the amount of data in the training set and improve the generalization ability and robustness of the model. Meanwhile, in order to aggregate more beneficial features and better integrate edge information and time features, this research illustrated a deformable global connection instead of the traditional long connection to provide more extensive multi-dimensional feature information for the decoding part of the network. Compared with the original network which only transfers the results of the encoding layer to the decoding layer, the features of the deformable global connection are the weighted sum of the global features. The deformable global connection is composed of the context modeling part, the feature transformation part and the deformable convolutional layer part. The part of modeling algorithm obtains the attention weight of the feature vector in the x - y plane, which represents the global context feature. The feature transformation captures the dependencies between channels to reduce the optimization difficulty via layer normalization. The deformable convolutional layer captures the geometry of the heart. In the process of decoding, the decoding layer can perform targeted aggregation to improve the compact and semantic consistency within the class based on the obtained weights. At the end, the spatial directional field features are learned and the original segmentation results are calibrated by these features. The spatial direction field, generated by ground truth, provides each pixel with a unique path from the boundary to the central region. The orientation relationship is revealed amongst the pixels on the target organ. The feature of central region is used to correct the segmentation map based on the requirement of spatial direction field. **Result** The online test results of the Automated Cardiac Diagnosis Challenge (ACDC) cardiac segmentation challenge illustrated that the average dice for the segmentation of left ventricle, right ventricle and left ventricle in this research analyses were 95%, 91.5% and 91% each, which is 2%, 1.5% and 2.5% higher than the initial network analysis. The average Hausdorff distance of the segmentation results is 6.77, 11.39 and 8.54, respectively, which is 2.22, 2.07 and 3.64 lower than the initial network. **Conclusion** This demonstrated experiments show that the proposed network can improve the segmentation effect of target organs effectively based on the spatio-temporal information of cardiac MRI images.

Key words: medical image; cardiac magnetic resonance imaging (CMRI); feature enhancement; spatial directional field; UNet

0 引言

心脏动态磁共振成像技术在临床诊断中具有重要的意义。根据其分割结果,医生可以有效地获取心肌质量和厚度、射血分数以及心室体积等诊断指标。然而准确地对其进行分割是一项具有挑战性的

任务,因为在成像的过程中由于不均匀的磁场强度,容易产生伪影,造成边界模糊。不同受试者的器官差异很大,尤其是右心室的形状大小多变,容易产生体积效应。此外动态磁共振成像技术本身就具有短轴切片数量少,切片厚度过厚等特点,导致图像的短轴分辨率过低、信息量稀疏等问题。现阶段的心脏分割方法依旧以手动分割方式为主。但这种手动分

割的方法耗时费力且对医生的专业素质要求严格。因此,如何自动对心脏磁共振图像(cardiac magnetic resonance imaging, CMRI)进行精确的分割已经成为一个重要的医学图像问题。

深度学习在医学图像分割领域大放异彩。与此同时,提出了许多基于深度学习的心脏分割算法。其中 Chang 等人(2018)提出了利用 YOLO(you only look once (Redmon 等, 2016))网络提取感兴趣区域,再输入到 U 型神经网络的方法,从而减少无关信息对分割效果的影响。Zotti 等人(2019)将心脏结构的先验信息嵌入网络,并为心脏分割提出了新的损失函数。Liu 等人(2020)提出了将残差模块(He 等, 2016)与双向长短记忆模块相结合的方法以减少不符合物理特征的分割异常点的出现。Hasan 和 Linte(2020)提出群卷积与权重剪枝技术相融合的 U 型神经网络,增加网络深度的同时减少了参数量。高强等人(2020)提出了一种基于组归一化与最近邻插值的分割算法,能快速准确地提取特征信息。刘畅等人(2021)为了在分割心肌内膜与外膜时保证其空间位置关系,提出了先用胶囊网络提取出相对位置信息和大小等信息,再输入到全连接网络进行预测的方法。然而现存的方法通常存在以下几个问题:1)针对动态心脏 MRI 图像本身短轴(z 轴)切片厚度过厚、信息稀疏的特点,现有方法为避免 3D 卷积带来的边界效应(Baumgartner 等, 2018)往往会使用 2 维网络来进行分割,使用单帧的 3 维图像在 z 轴的切片作为网络的输入。这样会丢失图像在 3 维层面的上下文信息,产生目标器官之间的类内不确定、类间不清晰的问题。2)由于现有数据集仅有动态 MRI 图像在收缩期末期和舒张期末期两个时间点的标签值(ground truth),现有网络通常也仅将这两个时间点的图像作为分割对象,这忽略了动态 MRI 图像在时间序列上的信息。3)由于心脏结构中细小器官的形状、大小多变的特性,现有方法往往会产生一些不符合目标器官相对物理特性的异常点(左心机包含了右心室的情况)。

为解决以上 3 个问题,提出了一种新的多输入、多分支和多任务的伪 3D 分割网络 ST UNet(spatio-temporal UNet)。该网络的主要创新点如下:1)为了充分利用心脏 MRI 影像的时间信息解决目标器官的类内不确定性问题,以连续时间内的多帧输入代替传统的单帧输入,让网络学习更为丰富的时域特

征。2)利用目标帧学习可变形偏移量,并将其融合到时域特征,使其具有更为精准边缘信息。3)为了减小细小器官的分割错误,更好地融合时域特征与边缘信息,提出可变全局注意力连接来代替传统的长连接,为网络提供更加广泛的全局上下文信息。4)针对由 2 维卷积造成的上下文信息丢失所引起的目标器官边界模糊、类间不清晰的问题,提出根据方向场特征的学习来进行特征向量的矫正。

1 本文方法

1.1 网络结构阐述

鉴于心脏动态 MRI 图像的切片过厚、层间信息稀疏的特点,提出了基于 U 型神经网络的伪 3D 分割网络(ST UNet)的结构如图 1(a)所示。该网络的基本骨架来源于 Ronneberger 等人(2015)提出的 UNet 结构。该结构由编码层、解码层与跳跃连接组成。其中编码层又称为下采样层,其作用为提取高级语义特征,并利用最大池化层减小特征图尺寸来扩大卷积的感受野。解码层又称为上采样路径,其作用是将高级语义特征图逐步还原成原始图片的大小,实现端到端的分割效果。跳跃连接则是将编码层学习到的高级特征图扩充到解码路径,从而避免特征提取过程中信息丢失的问题,使得解码层还原出来的结果更为精准。UNet 解码层的还原结果可表示为

$$\mathbf{x}^i = c([\boldsymbol{\alpha}^i, up(\mathbf{x}^{i+1})]) \quad (1)$$

式中, $\mathbf{x}$ 表示解码层的还原结果, $\boldsymbol{\alpha}$ 表示编码层的高级语义特征, i 表示当前所在的层数, $[\cdot]$ 表示通道连接操作, $up(\cdot)$ 表示上采样操作, $c(\cdot)$ 表示解码操作。

ST UNet 保留了 UNet 中编码块、解码块与跳跃连接层的基本结构框架。但为充分利用心脏动态 MRI 图像的时空信息,ST UNet 对这 3 个结构做了不同改进,形成了一种多输入、多路径和多任务的新型分割网络。首先在编码阶段,ST UNet 提出了多路径、多输入的时间信息聚合编码块。将目标帧和关键序列分别输入到不同的路径。规范化卷积对关键序列提取丰富的时域特征,可变形卷积(Dai 等, 2017)提取目标帧中更为精准边缘信息。其次将特征通过可变形全局注意力连接扩充到解码路径。

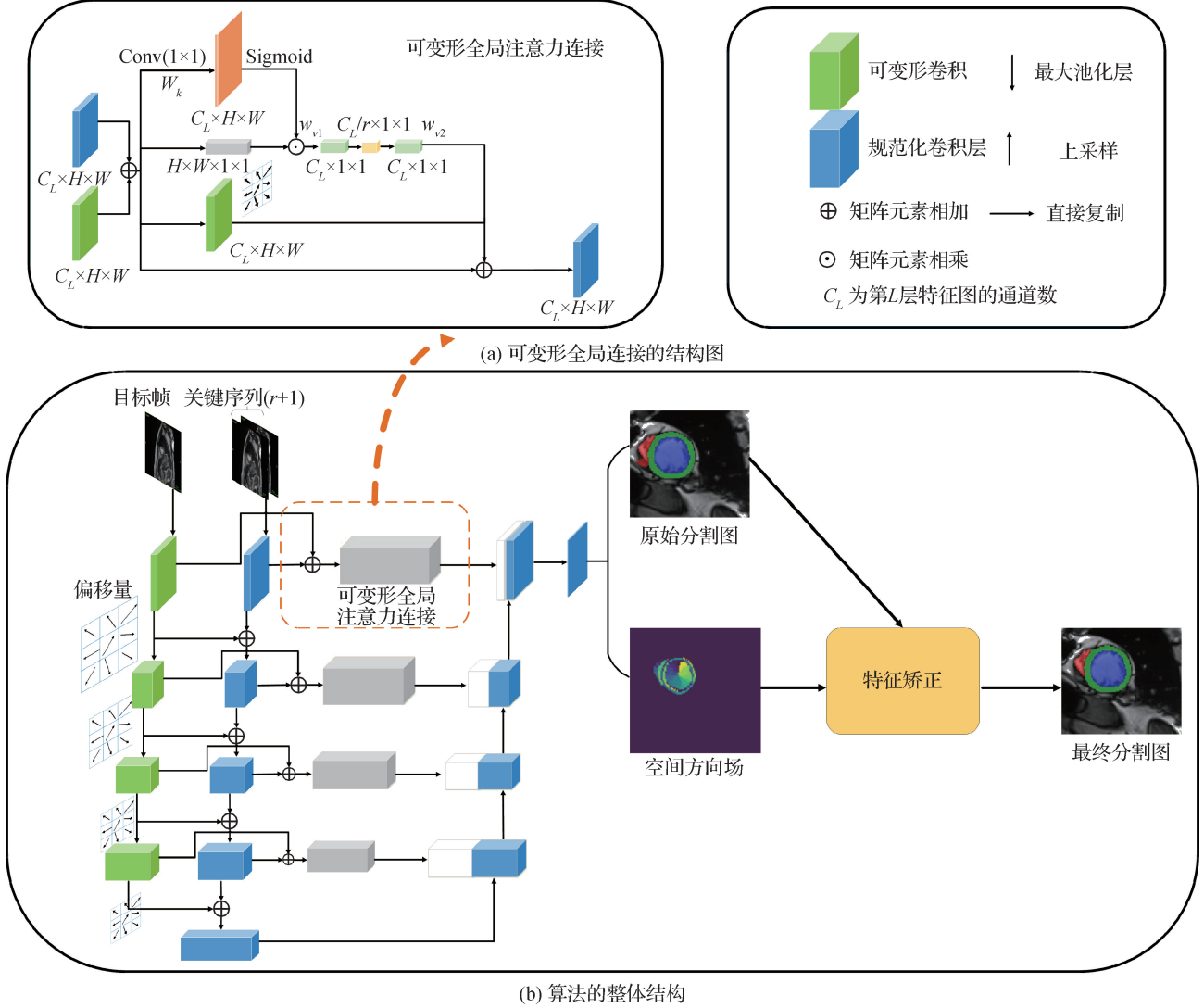


Fig. 1 Network structure diagram

((a) structure diagram of deformable global attention connection; (b) overall structure of the algorithm)

最后在网络解码过程中,规定 ST UNet 在预测原始分割图的同时还需额外预测分割图所对应的空间方向场,再利用预测到的空间方向场对原始的分割图进行特征矫正,得到改进后的分割结果。

1.2 时间增强编码块与可变形全局连接

现有的大多数方法都忽略了心脏动态 MRI 图像的时间信息,仅以收缩期末期和舒张期末期的图像作为单独的输入,从而导致分割性能下降。因此 ST UNet 选择将需要分割的图像作为目标帧,以包含目标帧的连续时间序列一同作为网络的输入,该连续序列称为关键序列。

心脏动态 MRI 图像的维度表示为 $[x, y, z, t]$,其中 x - y 为短轴平面, z 为短轴, t 为时间维度。考虑到

MRI 心脏图像在 z 轴上存在较大的层间间隙,选择在时间维度 t 上连续的 x - y 平面的图像组成关键序列。鉴于数据中心心脏 MRI 舒张期末期的时间维度 t 均为 0,选择从目标帧开始的后 r 帧组成关键序列。即假设 $c_{t_0} \in \mathbf{R}^{H \times W}$ 表示在 t_0 时刻的目标帧,则关键序列为 $\{c_{t_0}, \dots, c_{t_0+r}\}$ 。其中 H, W 分别代表了特征向量的高和宽。

关键序列的编码块由两个 3×3 的卷积核组成,并在每个卷积核后面添加批量归一化 (batch-normalization) 和 ReLU 激活函数。时间信息增强后的特征可以表示为

$$F_{\text{Key}}(p_0) = \sum_{t=t_0}^{t_0+r} \sum_{p_n \in \mathbf{R}} w(p_n) \times C_t(p_0 + p_n) \quad (2)$$

式中, $F_{\text{key}}(\cdot)$ 表示增强后的特征向量, p_0 表示特征向量中的任意位置, $p_n \in \mathbf{R}$ 表示卷积核中的任意位置, $w(p_n)$ 表示卷积核在位置 p_n 的权重值, 而 $C_i(p_0 + p_n)$ 表示图像 C_i 在 $(p_0 + p_n)$ 位置的值。

此外, 由于心脏随着时间变化的跳动, 目标器官的边缘信息也会随着时间的变化产生变化, 从而造成关键序列中不同时间点的边界信息冲突, 因此在对关键序列的卷积过程中可能会产生边缘语义信息的冲突问题。为解决此问题, 本文在编码过程中增加了一条针对目标帧 C_{t_0} 的可变形编码路径, 利用可变形卷积学习偏移量 δ 来增强卷积对于目标器官形状的适应能力, 更精准地提取目标器官的信息。

规范卷积中卷积核对于输入的特征图的采样位置都是固定的, 如 3×3 大小的卷积核的采样位置固定为 $\mathbf{R} = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}$, 其输出的特征图可以直接表示为

$$F(p_0) = \sum_{p_n \in \mathbf{R}} w(p_n) \times x(p_0 + p_n) \quad (3)$$

式中, $x(p_0 + p_n)$ 表示图像 x 在 $(p_0 + p_n)$ 位置的值。

固定的采样方式限制了卷积核对不同位置、不同尺寸或变形物体的特征提取能力, 可变形卷积通过对卷积核中每一个采样点增加一个偏移变量 δ , 改变卷积核采样点的位置。

由图2可以看出, 相较于常规的卷积, 可变形卷积更加适应几何形变, 其原理是让网络在训练过程中额外学习每一个卷积核的位置偏移矩阵 δ , 使得卷积的采样点发生偏移, 实现在当前位置附近的随意采样, 提高网络的灵活性。可变形卷积中输出的特征图表示为

$$F_{\text{tar}}(p_0) = \sum_{p_n \in \mathbf{R}} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (4)$$

式中, $\Delta p_n \in \delta$ 表示在 p_n 位置的权重的位置偏移量。

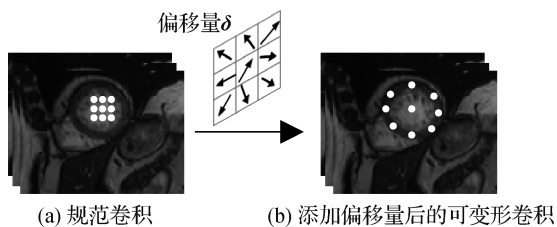


图2 可变形卷积与规范卷积

Fig. 2 Deformable convolution and standard convolution
((a) standard convolution; (b) deformable convolution)

再将由关键序列生成的特征与目标帧生成的特征进行融合, 生成当前层编码块的特征图, 最后将融

合后的特征图作为下一层关键序列编码块的输入, 该特征图表示为

$$F(p_0) = F_{\text{key}}(p_0) + F_{\text{tar}}(p_0) \quad (5)$$

融合后的特征图不仅包含了心脏动态 MRI 图像中丰富的时间信息, 也保留了目标器官在 t_0 时刻 (收缩期末期 ed 或舒张期末期 es) 的精准边缘信息。

考虑到目标器官之间的相对位置固定, 而实验中会产生一些不符合该相对位置的异常点, 因此本文受 GC block (global context block) (Cao 等, 2019) 的启发设计了可变形全局注意力连接, 为解码层提供一个全局的注意力矩阵, 让网络在解码还原的过程中获取图像全局上下文信息, 从而使得还原的结果更为准确。

可变形全局注意力连接的结构如图1(b)所示。该结构主要分为上下文建模部分、特征转化部分以及可变形卷积层部分。

上下文建模部分利用 1×1 的卷积和 sigmoid 激活函数获取特征向量在 x - y 平面的注意力权值, 然后将该权值与图像相乘得到全局的上下文特征。特征转化部分则通过两个 1×1 的卷积 w_{v1} 与 w_{v2} 捕获图像通道间的依赖, 并利用层归一化来降低优化难度。该过程可以表示为

$$\alpha_j = \frac{\exp(w_k x_j)}{\sum_{m=1}^{N_p} \exp(w_k x_m)} \quad (6)$$

$$z_i = x_i + w_{v2} \text{ReLU}(\text{LN}(w_{v1} \sum_{j=1}^{N_p} \alpha_j x_j)) \quad (7)$$

式中, N_p 为图像的位置数量, $i \in [1, N_p]$, x_i 表示第 i 个位置的输入, z_i 是对应的输出。 α_j 是采用了嵌入式高斯表达式后的注意力权值, LN 表示层归一化操作。

最后利用可变形卷积层捕获心脏的几何信息, 将其与 z_i 以及原始特征进行融合, 并将融合后的结果传给解码层, 进行解码还原。

与原始网络仅仅将编码层的结果传给解码层相比, 可变形全局连接后的特征是全局特征的加权和。在解码还原的过程中解码层可以根据权值进行有选择的聚合, 提高类内紧凑型和语义一致性。

1.3 空间方向场

受 Cheng 等人(2020)的启发, 为了进一步增强网络预测的结果的类间差异性与类内相似性, 提出在预测原始分割图的同时预测对应的空间方向场。

空间方向场是由真实标签图所生成,其生成规则为对每一个前景像素点 p ,找到目标器官边界最近的一点 b 。然后对向量 bp 以 b 到 p 的距离为 $|bp|$ 进行归一化,最终像素点 p 的方向场可以表示为

$$DF(p) = \begin{cases} bp/|bp| & p \in \text{目标器官} \\ (0,0) & p \notin \text{目标器官} \end{cases} \quad (8)$$

由式(8)可以看出,空间方向场为每一个像素提供了从边界指向中心区域的唯一路径,揭示了目标器官上像素与像素的方向关系。

用一个简单的方向场模块来学习空间方向场。该模块由一个 1×1 的卷积层组成,模块的输入为 UNet 最后一个解码层还原出来的 64 通道的特征,输出为双通道的方向场特征。其中一个通道反映像素点 x 轴之间的方向关系,另一个通道反映 y 轴之间的方向关系。

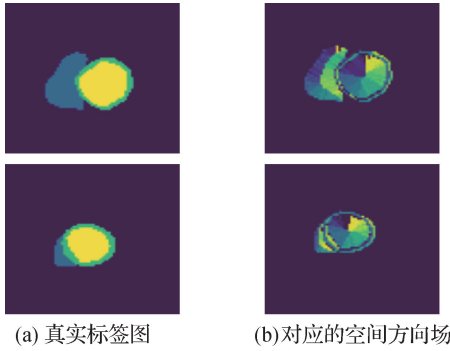


图 3 真实标签图与其对应的空间方向场
Fig. 3 Ground truth and spatial directional field
((a) ground truth; (b) spatial directional field)

根据空间方向场的指导,本文利用中心区域的特征对原始分割图进行矫正。如图 4 所示,最终特征图 F_{final} 是由原始特征图 F_0 和预测的方向场结合得到的,矫正方式为根据方向场提供的位置信息对原始特征图进行双线性插值。换句话说,让原始特征图 F_0 上的像素点 p_0 根据学习到的方向场进行修改,其整个优化过程为

$$\forall p \in \Omega, F(p) = F(p_x + DF(p)_x, p_y + DF(p)_y) \quad (9)$$

式中, Ω 表示所有的前景像素点集, p_x 表示像素点 p 的 x 轴坐标, p_y 表示像素点 p 的 y 轴坐标, $F(p)$ 表示经过修正后像素点 p 的值。由式(9)可以看出,像素点 p 的矫正后的特征值是根据方向场位置偏移

后的像素值。

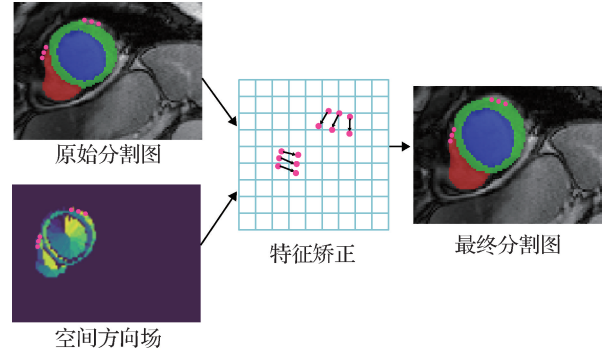


图 4 特征矫正过程图

Fig. 4 The process of feature correction

1.4 损失函数

本文的损失函数涉及分割图的损失函数 L_{seg} 以及空间方向场的损失函数 L_{DF} 。

分割图的损失函数为交叉熵损失函数与 Dice 损失函数的联合函数。其中交叉熵是将预测结果与真实标签图的每一个像素进行对比,然后对所有的对比结果求平均值。而 Dice 损失函数是评价预测图与真实标签值的重叠程度,重叠程度越高,分割效果越好。联合损失函数可以表示为

$$L_{\text{seg}} = 0.5 \times L_{\text{cross}} + L_{\text{dice}} \quad (10)$$

$$L_{\text{cross}} = - \sum_{i=1}^N (y^i \log m^i + (1 - y^i) \log (1 - \log m^i)) \quad (11)$$

$$L_{\text{dice}} = \sum_{i=1}^N 1 - \frac{2 \times y^i \times m^i}{y^i + m^i} \quad (12)$$

式中, L_{cross} 表示交叉熵损失函数, L_{dice} 表示 Dice 损失函数。 N 表示样品的总数, y^i 表示第 i 个样本的预测结果, m^i 表示第 i 个样本的真实标签值。

对于空间方向场的学习,以 L2 范式距离和角度之间的距离作为方向场的损失函数,即

$$L_{\text{DF}} = \sum_{p \in \Omega} w(p) (\|DF(p) - \widehat{DF}(p)\|_2 + \|\cos \langle DF(p), \widehat{DF}(p) \rangle\|^2) \quad (13)$$

式中, DF 和 $\widehat{DF}$ 分别表示网络预测的方向场与真实标签的方向场。 $\|\cdot\|_2$ 表示 L2 范式距离。 Ω 表示所有的前景像素点集, $w(\cdot)$ 表示像素点 p 的权重值, $w(p)$ 计算式为

$$w(p) = \begin{cases} \frac{\sum_{i=1}^N |C_i|}{N \cdot |C_i|} & p \in C_i \\ 1 & p \notin C_i \end{cases} \quad (14)$$

式中, $|C_i|$ 表示当前标签的所有像素的个数, N 表示所有标签的个数。

2 实验结果与讨论

2.1 实验数据与预处理

使用 ACDC (Automated Cardiac Diagnosis Challenge) (Bernard 等, 2018) 自动心脏分割挑战数据集, 该数据集是法国 Dijon 大学医院根据当地委员会制定的规定进行采集的。采集设备为两台磁场强度不同的 MRI 扫描仪: 1.5 T 磁场与 3.0 T 磁场。该数据集包含 150 个志愿者的心脏核磁共振图像, 150 名志愿者又分为 5 个亚组, 分别为: 1) 健康的志愿者; 2) 既往有心肌梗塞的志愿者; 3) 扩张型心肌病的志愿者; 4) 肥厚型心肌病的志愿者; 5) 右心室异常的志愿者。每个志愿者的数据包含心脏整个动态电影 MRI 图像, 以及图像中舒张期末期以及收缩期末期的真实标签值。该数据集的真实标签值是由两名资深医生手动绘制的。动态 MRI 图像的空间分辨率为 $1.37 \sim 1.68 \text{ mm}^2/\text{像素}$ 之间, 并且动态图像由 $28 \sim 40$ 幅图像组成, 完全或部分覆盖了心跳周期。该数据集将 100 名志愿者划分为训练集, 50 名志愿者划分为测试集, 训练集与测试集各亚组的分布比例相同并且并未公布测试集的真实标签值。本文所有实验结果都来自于 ACDC 的官方在线测试平台 (<https://acdc.creatis.insa-lyon.fr/>)。

对 ACDC 数据集进行如下预处理: 1) 本文中关键序列参数 r 取值为 1, 因此从短轴动态心脏 MRI 视频中获取收缩期末期 t_{es} 与舒张期末期 t_{ed} 以及其各自后 1 帧的时间点 t_{ed+1} 与 t_{es+1} 时间点的 3 维图像。2) 将 3 维图像沿 z 轴进行 2 维切片, 单独将 t_{ed} 或 t_{es} 时间点所得的切片作为目标帧, 再将目标帧对应的 t_{ed+1} 或 t_{es+1} 时间点所得的切片与目标帧组成关键序列。3) 为增加训练集的数据量, 提升模型的泛化能力与鲁棒性, 本文同时对目标帧与关键序列做相同的数据增强, 数据增强的方式有随机仿射变化与随机旋转。4) 由于数据集中图像的尺寸各不相同, 将所有的切片都剪切或者填补成 256×256 像素。5) 对图像进行归一化操作。

2.2 实验环境与参数设置

实验硬件环境为 Intel Core i7 处理器与 NVIDIA GTX1080ti 显卡。编码语言为 python, 深度学习框架

选用的是 pytorch 1.4.0 版本。权重的初始化方式为 pytorch 的默认初始化, 优化器选用了 Adam 优化器。学习率 α 初始化为 0.001, 权重衰减率为 $1E-5$ 。网络的通道数从上到下分别为 64, 128, 256, 512, 1024。

2.3 评判指标

本文选用 ACDC 在线测试平台给出的两种评判标准来定量评估 ST UNet 的性能。

这两个评判标准分别为 Dice 相似性系数 (Dice similarity coefficient, DSC) 与豪斯多夫距离 (Hausdorff distance, HD), 其中 Dice 相似性系数 DSC 表示预测结果与真实标签图的重叠程度, 其值在 $[0, 1]$ 之间。而 HD 系数解释了预测结果与真实标签图的最大不匹配程度。DSC 的数学表达式为

$$DSC = \frac{2|Y \cap P|}{|Y| + |P|} \quad (15)$$

HD 的数学表达式为

$$d_{py} = \max_{p \in P} \min_{y \in Y} (p, y) \quad (16)$$

$$d_{yp} = \max_{y \in Y} \min_{p \in P} (y, p)$$

$$HD = \max\{d_{py}, d_{yp}\} \quad (17)$$

式中, P 与 Y 分别表示预测结果与真实标签图, p, y 为预测结果与真实标签图上的点。

2.4 消融实验

为探究本文算法的可行性, 做了 4 组消融实验来说明空间方向场、时间增强编码块和可变形全局连接对于实验结果的影响, 实验结果如表 1 和表 2 所示, LV 表示左心室, RV 表示右心室, MYO 表示左心肌。图 5 为 ST UNet 模型在测试集上的 DSC 箱型图。图 6 是 4 种方法的分割效果对比图, 分别选取了收缩期末期与舒张期末期 4 个不同患者进行对比。在分割图中, 红色、蓝色和绿色分别表示右心室、左心肌和左心室。

造成类内不确定、类间不清晰的主要原因是心脏动态 MRI 本身短轴切片数量少, 分辨率过低, 上下文信息量稀疏。稀疏的信息量造成以往的网络难以适应目标器官复杂的形状和成像时不同的磁场强度。而相邻时间的切片之间的形变较小, 尤其是背景与目标器官中心位置的信息并未因心脏跳动产生较大的变化, 该信息依旧有效, 可用于分割网络。时间增强编码块中的关键序列编码路径正是利用同一切片在相邻时间点的时间信息来弥补稀疏的 3 维上下文信息, 使得网络可以学习到的信息量倍增, 增强

表 1 消融实验 DSC 的对比结果
Table 1 DSC contrast results of ablation experiment

模型	LV DSC		RV DSC		MYO DSC	
	舒张期末期	收缩期末期	舒张期末期	收缩期末期	舒张期末期	收缩期末期
UNet	0.96	0.90	0.93	0.87	0.88	0.89
UNet + 空间方向场	0.96	0.92	0.94	0.88	0.90	0.91
UNet + 时间增强 + 空间方向场	0.97	0.92	0.95	0.89	0.90	0.91
ST UNet	0.97	0.93	0.94	0.89	0.90	0.92

注:加粗字体表示每列最优结果。

表 2 消融实验 HD 的对比结果
Table 2 HD contrast results of ablation experiment

模型	LV HD		RV HD		MYO HD	
	舒张期末期	收缩期末期	舒张期末期	收缩期末期	舒张期末期	收缩期末期
UNet	6.93	11.04	12.19	14.73	11.06	13.29
UNet + 空间方向场	8.43	10.09	11.36	13.74	9.94	11.01
UNet + 时间增强 + 空间方向场	6.99	8.65	10.83	13.25	8.52	8.03
ST UNet	5.52	8.02	10.50	12.29	8.24	8.84

注:加粗字体表示每列最优结果。

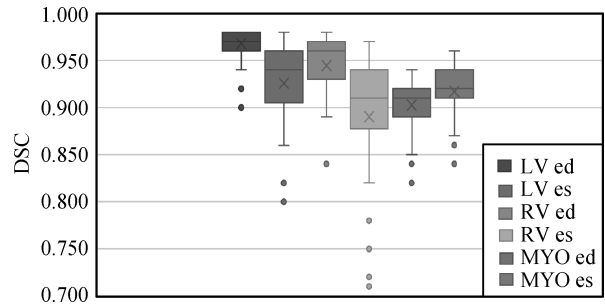


图 5 ST UNet 测试集 DSC 箱型图
Fig. 5 Test set DSC box diagram of ST UNet

了网络的适应能力,以此来解决类内不确定的问题。

然而相邻时间切片之间,目标器官的边界信息由于心脏跳动发生了变化,不同时间点的边界信息是相互冲突的。针对这一问题,时间增强编码块中的目标帧编码路径利用可变形卷积提取出目标器官在 ed 时刻或 es 时刻精准的边界信息,并通过可变形全局连接将精确的边缘信息与丰富的时间信息融合后扩散到解码路径,使得网络解码还原过程中可以根据权值进行有效的聚合,还原出更为精准的边缘信息,从而解决目标器官与目标器官之间、目标器官与背景之间的类间不清晰的问题。

对比表 1 和表 2 中 ST UNet 的实验结果和没有

添加时间增强编码块与可变形全局连接结构的实验结果可以看出,添加后的实验结果中 3 个目标器官的 DICE 值和 HD 距离都有了明显的改善。结合图 6 的分割结果图,可以发现,ST UNet 的分割结果更加接近真实标签,心肌不连贯、心室误分割等情况也都得到了改善。这充分体现了提出的利用时间信息来弥补稀疏的 3 维上下文信息的特征增强方式是可行的。

同时从表 1,表 2 与图 6 可以看出原始 UNet 分割容易出现边界模糊、将无关背景误分割成目标器官以及相邻器官分割错误等问题。通过加入额外的空间方向场训练与矫正后,边界模糊的现象有所改善,3 个目标器官的分割精度也有所提升,但是依旧有边缘不平滑以及误分割的现象。在此基础上再添加多输入的时间分支,其分割精度有了进一步的提升,这样更加证明了利用关键序列来进行特征增强能提高网络学习特征的能力,改善网络的分割效果。虽然此时的分割效果相较于前两个模型,边界平滑,没有明显的误分割现象,但是由于关键序列中不同时间点的图像的目标器官的边缘冲突问题,此时的分割结果的边缘出现了欠分割的情况。而加入全局注意力连接后的模型,利用目标帧的全局信息弥补

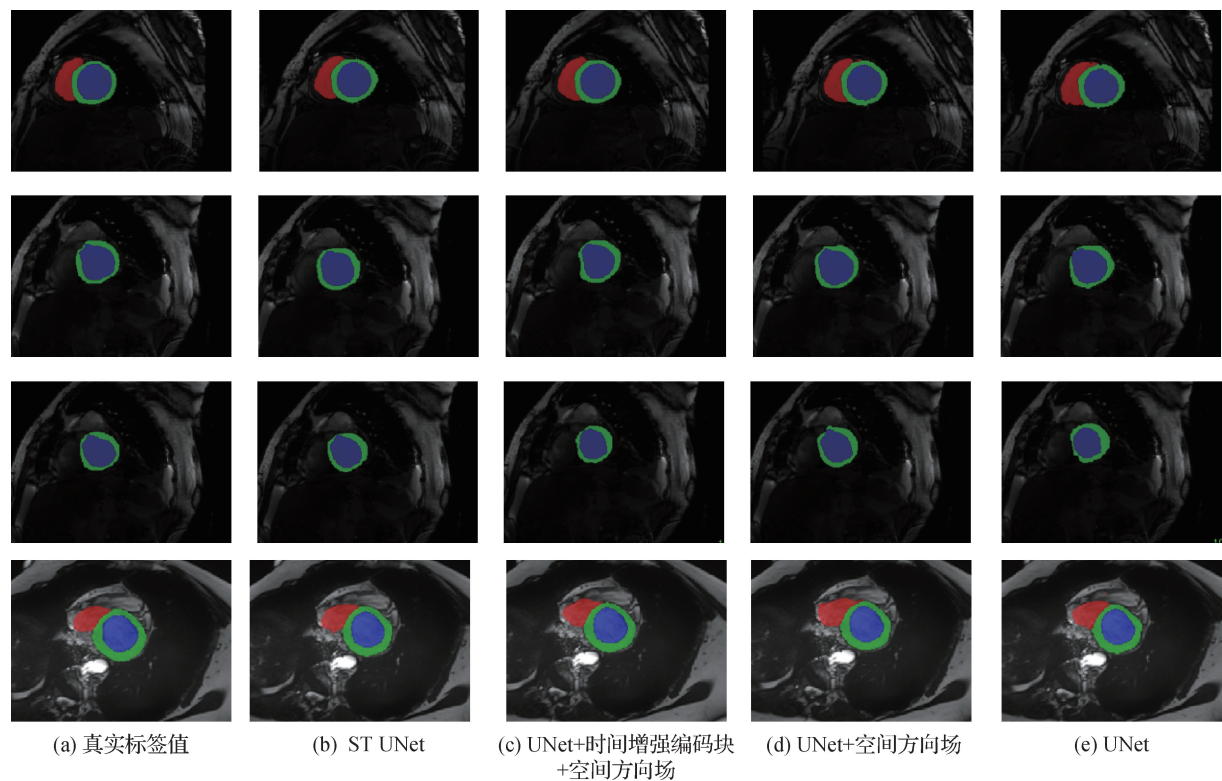


图6 不同模型的分割结果

Fig. 6 Segmentation results of different models((a)ground truth;(b)ST UNet;
(c)UNet + time branch + spatial directional field;(d)UNet + spatial directional field;(e)UNet)

了关键序列边缘冲突的问题,从而使得分割结果达到了更好的效果,更加接近真实标签。

2.5 与其他算法的对比

不同算法对比结果如表3和表4所示。其中Liu等人(2020)是利用残差网络与双向记忆网络来提取相邻切片的3维空间特征。Painchaud等人(2019)是在对抗性编码器(aVAE)中添加了一个潜在的心脏形态。Sun等人(2020)使用二次形流来捕获心脏的形态特征。Simantiris和Tziritas(2020)提出了新扩展卷积网络和自定义的损失

函数。

由表3和表4可以看出,本文算法的DSC与HD值都要优于其他方法,这是因为这些算法都仅仅以收缩期末期与舒张期末期的图像作为网络的训练对象,忽视了心脏视频在时间序列上的信息。而HD值的降低也从侧面表示利用空间方向场来修正边界的可行性。因此提出利用心脏动态MRI图像的时间信息来进行特征增强,再利用空间方向场提取出的空间信息来进行特征矫正的方法是可行的。截止2020年12月7日,与ACDC排行榜上的算法

表3 不同算法的DSC比对结果
Table 3 DSC comparison results of different algorithms

模型	LV DSC			RV DSC			MYO DSC		
	舒张期末期	收缩期末期	平均	舒张期末期	收缩期末期	平均	舒张期末期	收缩期末期	平均
Liu 等人(2020)	0.95	0.87	0.920	0.92	0.86	0.890	0.86	0.87	0.870
Painchaud 等人(2019)	0.96	0.91	0.935	0.93	0.88	0.910	0.88	0.90	0.890
Sun 等人(2020)	—	—	0.938	—	—	0.914	—	—	0.887
Simantiris 和 Tziritas(2020)	0.96	0.93	0.945	0.94	0.89	0.915	0.89	0.90	0.895
ST UNet	0.97	0.93	0.950	0.94	0.89	0.915	0.90	0.92	0.910

注:“—”表示论文中作者并未给出相应数值,加粗字体表示每列最优结果。

表4 不同算法的HD对比结果

Table 4 HD comparison results of different algorithms

模型	LV HD			RV HD			MYO HD		
	舒张期末期	收缩期末期	平均	舒张期末期	收缩期末期	平均	舒张期末期	收缩期末期	平均
Liu 等人(2020)	11.91	15.78	13.85	17.90	24.72	21.30	19.18	20.69	19.93
Painchaud 等人(2019)	6.15	8.29	7.22	13.72	13.32	13.52	8.65	9.60	8.83
Sun 等人(2020)	—	—	—	—	—	—	—	—	—
Simantiris 和 Tziritas(2020)	6.40	7.60	7.00	13.30	14.40	13.85	8.30	9.60	8.95
ST UNet	5.52	8.02	6.77	10.50	12.29	11.39	8.24	8.84	8.54

注:“—”表示论文中作者并未给出相应数值,加粗字体表示每列最优结果。

对比,ST UNet 的3个目标器官的平均DSC与平均HD值分别达到了92.5%与8.90 mm的成绩,仅次于第1名。

3 结 论

针对心脏磁共振图像自动分割的类内不确定、类间模糊等问题,本文提出了利用时间信息来进行特征增强,再利用空间信息来进行特征矫正的ST UNet。ST UNet的主要特点如下:1)相较于现有算法,本文将需要分割的3维图像作为目标帧,将包含目标帧的连续时间序列作为关键序列一起输入网络。2)利用关键序列学习丰富的类内特征,再利用可变形路径学习目标帧上更为精准的边缘信息,通过编码路径与可变形全局注意力连接将两者充分融合。3)网络额外学习空间方向场信息,并利用学习到的空间方向场对原始的预测结果进行矫正。实验结果表明,相较于其他算法,ST UNet能够得到更为精准的分割结果,各器官的分割精度均取得了提升。

尽管本文方法取得了一定进展,但依旧存在如下问题:收缩期末期的分割效果普遍差于舒张期末期的分割效果,原因是在收缩期末期,目标器官排血后体积变小,尤其是心尖与心室基层的形变较大,从而导致目标器官与背景的标签不平衡。后续工作将针对这一问题展开。

参考文献(References)

Baumgartner C F, Koch L M, Pollefeys M and Konukoglu E. 2018. An exploration of 2D and 3D deep learning techniques for cardiac MR image segmentation//Proceedings of the 8th International Workshop

on Statistical Atlases and Computational Models of the Heart. Quebec City, Canada: Springer: 111-119 [DOI: 10.1007/978-3-319-75541-0_12]

Bernard O, Lalande A, Zotti C, Cervenansky F, Yang X, Heng P A, Cetin I, Lekadir K, Camara O, Gonzalez Ballester M A, Sanroma G, Napel S, Petersen S, Tziritas G, Grinias E, Khened M, Kollerathu V A, Krishnamurthi G, Rohe M M, Pennec X, Sermesant M, Isensee F, Jager P, Maier-Hein K H, Full P M, Wolf I, Engelhardt S, Baumgartner C F, Koch L M, Wolterink J M, Isgum I, Jang Y, Hong Y O N M, Patravali J, Jain S, Humbert O and Jodoin P M. 2018. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE Transactions on Medical Imaging, 37(11): 2514-2525 [DOI: 10.1109/TMI.2018.2837502]

Cao Y, Xu J R, Lin S, Wei F Y and Hu H. 2019. GCNet: non-local networks meet squeeze-excitation networks and beyond//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshops. Seoul, Korea (South): IEEE: 1971-1980 [DOI: 10.1109/ICCVW.2019.00246]

Chang Y K, Song B Y, Jung C and Huang L Y. 2018. Automatic segmentation and cardiopathy classification in cardiac MRI images based on deep neural networks//Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary, Canada: IEEE: 1020-1024 [DOI: 10.1109/ICASSP.2018.8461261]

Cheng F, Chen C, Wang Y K, Shi H, Cao Y K, Tu D D, Zhang C Z and Xu Y C. 2020. Learning directional feature maps for cardiac MRI segmentation//Proceedings of the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention. Lima, Peru: Springer: 108-117 [DOI: 10.1007/978-3-030-59719-1_11]

Dai J F, Qi H Z, Xiong Y W, Li Y, Zhang G D, Hu H and Wei Y C. 2017. Deformable convolutional networks//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 764-773 [DOI: 10.1109/ICCV.2017.89]

Gao Q, Gao J Y and Zhao D. 2020. GNNUGnet: Precise Segmentation

- Neural Network of Left Ventricular Contours for MRI Images Based on Group Normalization and Nearest Interpolation. *Computer Science*, 47(08):213-220(高强,高敬阳,赵地. 2020. GNNI U-net:基于组归一化与最近邻插值的MRI左心室轮廓精准分割网络. *计算机科学*, 47(08):213-220) [DOI:10.11896/jsjx.190600026]
- Hasan S M K and Linte C A. 2020. CondenseUNet: a memory-efficient condensely-connected architecture for bi-ventricular blood pool and myocardium segmentation//*Proceedings of SPIE 11315, Medical Imaging 2020: Image-Guided Procedures, Robotic Interventions, and Modeling*. Houston, USA; SPIE: 113151 J [DOI: 10.1117/12.2550640]
- 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]
- Liu C, Lin N, Cao Y J and Yang C. 2021. Seg-CapNet: neural network model for the cardiac MRI segmentation. *Journal of Image and Graphics*, 26(02):0452-0463(刘畅,林楠,曹仰杰,杨聪. 2021. Seg-CapNet:心脏MRI图像分割神经网络模型. *中国图象图形学报*, 26(02):0452-0463) [DOI:10.11834/jig.190626]
- Liu T, Tian Y, Zhao S F, Huang X Y and Wang Q J. 2020. Residual Convolutional neural network for cardiac image segmentation and heart disease diagnosis. *IEEE Access*, 8: 82153-82161 [DOI: 10.1109/ACCESS.2020.2991424]
- Painchaud N, Skandarani Y, Judge T, Bernard O, Lalande A and Jodoin P M. 2019. Cardiac MRI segmentation with strong anatomical guarantees//*Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention*. Shenzhen; Springer: 632-640 [DOI: 10.1007/978-3-030-32245-8_70]
- Redmon J, Divvala S, Girshick R and Farhadi A. 2016. You only look once: unified, real-time object detection//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA; IEEE: 779-788 [DOI: 10.1109/CVPR.2016.91]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany; Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Simantiris G and Tziritas G. 2020. Cardiac MRI segmentation with a dilated CNN incorporating domain-specific constraints. *IEEE Journal of Selected Topics in Signal Processing*, 14(6): 1235-1243 [DOI: 10.1109/JSTSP.2020.3013351]
- Sun J, Darbeha F, Zaidi M and Wang B. 2020. SAUNet: shape attentive U-Net for interpretable medical image segmentation//*Proceedings of the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention*. Lima, Peru; Springer: 797-806
- Zotti C, Luo Z M, Lalande A and Jodoin P M. 2019. Convolutional neural network with shape prior applied to cardiac MRI segmentation. *IEEE Journal of Biomedical and Health Informatics*, 23(3): 1119-1128 [DOI: 10.1109/JBHI.2018.2865450]

作者简介



徐佳陈, 1995年生, 男, 硕士研究生, 主要研究方向为医学影像分析。

E-mail: xujosh@foxmail.com



肖志勇, 通信作者, 男, 副教授, 主要研究方向为人工智能、医学影像分析、图像处理。

E-mail: zhiyong.xiao@jiangnan.edu.cn