

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2020)10-2171-11

论文引用格式: Yao F Z, Li Z, Wang L H, Cheng X Y and Zhang J. 2020. Deep iterative fusion network on skull removal of brain magnetic resonance images. Journal of Image and Graphics 25(10): 2171-2181(姚发展, 李智, 王丽会, 程欣宇, 张健. 2020. 深度迭代融合的脑部磁共振图像颅骨去除网络. 中国图象图形学报 25(10): 2171-2181 [DOI: 10.11834/jig.200218])

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

贵州大学计算机科学与技术学院智能医学影像分析与精准诊断重点实验室, 贵阳 550025

摘要: 目的 去除颅骨是脑部磁共振图像处理和分析中的重要环节。由于脑部组织结构复杂以及采集设备噪声的影响导致现有方法不能准确分割出脑部区域, 为此提出一种深度迭代融合的卷积神经网络模型实现颅骨的准确去除。方法 本文 DIFNet(deep iteration fusion net) 模型的主体结构由编码器和解码器组成, 中间的跳跃连接方式由多个上采样迭代融合构成。其中编码器由残差卷积组成, 以便浅层语义信息更容易流入深层网络, 避免出现梯度消失的现象。解码器网络由双路上采样模块构成, 通过具有不同感受野的反卷积操作, 将输出的特征图相加后作为模块输出, 有效还原更多细节上的特征。引入带有 L2 正则的 Dice 损失函数训练网络模型, 同时采用内部数据增强方法, 有效提高模型的鲁棒性和泛化能力。结果 为了验证本文模型的分割性能, 分别利用两组数据集与传统分割算法和主流的深度学习分割模型进行对比。在训练数据集同源的 NFBS(neurofeedback skull-stripped) 测试数据集上, 本文方法获得了最高的平均 Dice 值和灵敏度, 分别为 99.12% 和 99.22%。将在 NFBS 数据集上训练好的模型直接应用于 LPBA40(Ioni probabilistic brain atlas 40) 数据集, 本文模型的 Dice 值可达 98.16%。结论 本文提出的 DIFNet 模型可以快速、准确地去除颅骨, 相比于主流的颅骨分割模型, 精度有较高提升, 并且模型具有较好的鲁棒性和泛化能力。

关键词: 卷积神经网络(CNN); 颅骨分割; 脑部磁共振图像; 深度迭代融合; 数据增强

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

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

Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis of Guizhou Province,

School of Computer Science and Technology, Guizhou University, Guiyang 550025, China

Abstract: **Objective** Magnetic resonance imaging (MRI) is frequently used in clinical applications. It is a common means to detect lesions, injuries, and soft tissue variations in neural system diseases. Skull removal is an important preprocessing step for brain magnetic resonance (MR) image analysis. Its purpose is to remove nonbrain tissue from the brain MRI, thereby facilitating subsequent extraction and analysis of brain tissue. The MR images acquired using clinical scanners inevitably have blurring or noise characteristics due to the complexity of brain tissue structure and the effects of equipment noise and field offset. Differences also exist in the anatomical structure of the brain tissue for different individuals, which cause difficulties in the skull segmentation in brain MR images. Most traditional methods for skull segmentation are incompletely automatic and often require the operator to use the mouse and other tools to determine the center point of the region of interest and adjust the parameters manually. The current automatic skull segmentation method does not require human-computer

收稿日期: 2020-05-29; 修回日期: 2020-06-29; 预印本日期: 2020-07-05

interaction but has poor adaptability, and satisfactory segmentation results in different MR images are difficult to achieve. On the contrary, the deep learning-based method exhibits advanced performance in multiple segmentation tasks in the field of computer vision. Therefore, we propose a deep iterative fusion convolutional neural network model (DIFNet) in this work to realize skull segmentation. **Method** The main structure of DIFNet is composed of an encoder and a decoder. The skip connection between the encoder and decoder is realized by multiple upsampling iterative fusion, which means that the input information of one decoder layer comes from not only the same layer but also the deep layers of the encoder. The encoder consists of several residual convolution blocks, which allow the shallow semantic information to flow into deep networks to avoid gradient vanishment. The decoder is composed of double-way upsampling modules. The feature maps generated from the double-way upsampling modules are added as real outputs through deconvolution operations with different receptive field sizes. This process enables to restore the image details effectively by adding multiple scale information. The internal data enhancement method is adopted to enhance the generalization capability of the model. First, the image is randomly scaled, in which the interval of scaling factor sets is determined in accordance with the ratio of the original image size to the output block size. Then, a center point is randomly selected in the scaled image, and the cutting area is determined. Lastly, the cut image patches are fed into the network for training. The Dice loss function embedded with an L2 regularization item is used to optimize the model parameters and overcome the overfitting problem. We use two datasets in this work to evaluate the accuracy and robustness of the proposed model. Each dataset has a brain segmentation mask provided by a professional doctor as the gold standard of the model. One dataset is NFBS (neurofeed back skullstripped), from which a part of images are used for testing (the ratio of the training dataset to the test dataset is 4 : 1). The other dataset is LPBA40 (loni probabilistic brain atlas 40), which is used as an independent dataset for testing the generality of the models. For quantitative analysis, the Dice score, sensitivity, and specificity are used in this work. **Result** For the NFBS dataset, the method in this paper obtains the highest average Dice score and sensitivity of 99.12% and 99.22%, respectively, compared with U-Net, U-Net with residual block (Res-U-Net), and U-Net with double-way upsampling modules (UP-U-Net). The Dice score is increased by 1.88%, 1.81%, and 0.6%. The sensitivity and specificity are increased by at least 0.5% compared with the U-Net model. The segmentation results of the model are similar to the manual segmentation results of experts. The model trained with the NFBS dataset is applied directly to the LPBA40 dataset to verify the segmentation capability of the model. The Dice value obtained in the test experiment is up to 98.16%. By contrast, the Dice values of U-Net, UP-U-Net, and Res-U-Net are 81.69%, 77.34%, and 76.42%, respectively. Compared with these models, our proposed model is robust. **Conclusion** Experiments illustrate that the internal data augmentation and deep iterative fusion make the proposed model be easily trained and acquire the best segmentation results. The deep iterative feature fusion can guarantee the robustness of the segmentation model.

Key words: convolutional neural network (CNN); skull segmentation; brain magnetic resonance image; deep iterative fusion; data augmentation

0 引言

脑部磁共振图像可以为医生提供丰富的脑组织信息以辅助临床诊断,去除颅骨是脑部医学图像分析中非常重要的图像预处理步骤,其结果的质量会影响后续分析的准确性,如影响图像配准、脑组织分割以及脑区体积分析等。通常,颅骨分割的金标准需要经验丰富的医生进行手动划分,费时费力。另外由于脑部组织结构的复杂性以及采集设备噪声、场偏移效应等的影响,脑部磁共振图像不可避免会

产生模糊、灰度不均匀等问题,采用传统的分割方法并不能准确地将颅骨与脑室分割开来,上述因素促进了自动颅骨分割方法的发展。

相关专家和学者已经提出了许多脑部自动提取方法。Archibald 等人(2003)基于 Gegenbauer 重建方法对采集的脑部数据进行分割预处理,来改进脑组织分割。Rex 等人(2004)提出了脑提取元算法(brain extraction meta-algorithm, BEMA)来解决不同方法在不同数据上分割结果具有差异性的问题,并行执行多个提取算法和运算过程,以智能方式组合并获得优于任何单个算法的改进结果。Huang 等人

(2006) 结合期望最大化(expectation maximization, EM) 算法与数学形态学,降低脑部皮质区域的分类失误率。Somasundaram 和 Kalaiselvi(2010) 针对 T2 加权 MRI(magnetic resonance imaging) 的脑部提取,加入低通滤波器进行数据预处理,将阈值分割与最大连接成分分析结合,提高提取大脑的准确度。基于这些方法,已经开发了许多脑颅骨自动分割软件,如 BrainSuite 18a 软件中的脑部区域提取的方法(brain surface extractor, BSE)(Shattuck, 2001)、FSL(the fmrib software library) 的脑提取方法(brain extraction tool, BET)(Smith, 2002) 和基于学习的鲁棒性大脑提取(robust brain extraction, ROBEX)(Iglesias 等, 2011) 等。在这些方法中,当特征未被识别或脑部区域和颅骨未正确分离时,就会发生分割异常。此外 BET 和 BSE 在进行颅骨分割的时候都需要手动设置参数,并没有像 ROBEX 达到全自动分割,以上方法的分割结果与金标准相比仍有一定差距(Sara 等, 2014)。

卷积神经网络(convolutional neural network, CNN) 和基于深度学习的方法在医学图像分割中取得了出色的表现,其分割准确度接近于专家手动标记。大量脑部 MRI 的颅骨分割方法相继提出。Kleesiek 等人(2016) 首次提出了一种 3D 卷积神经网络结构,用于不同模态脑部 MRI 的大脑提取,但是 Kleesiek 采用 3D 的网络结构较为简单,限制了网络学习更多特征的能力。Valvano 等人(2018) 提出了一种以 2D U-Net(Ronneberger 等, 2015) 为基线改进的卷积神经网络,同时加入了数据增强操作,分割精度比常用的颅骨分割软件高。Han 等人(2018) 提出了一种基于主成分分析(principal components analysis, PCA) 的模型,专门用于从脑部病理图像中提取大脑。Lucena 等人(2019) 针对在专家标记数据不多的情况下,使用银标准掩膜来训练卷积神经网络,以期达到金标准的效果。Liu 等人(2019) 考虑到 3 维网络模型具有较高的计算成本和 GPU 内存限制,将特征融合模块转换为 2D 编码模块和 3D 空间编码模块,以提取区域和空间特征,最后将二者结合,这样既减少了计算量又增强了网络学习能力。Hwang 等人(2019) 提出了一种基于 3D U-Net 的端到端深度学习的分割模型,分割精度有了一定提高。但是 3D 网络模型具有太多参数,训练时间较慢而且测试过程要加载过多参数。

以上网络模型虽然在公开数据集上都有了较好的分割效果,但仍存在一个共性问题,即泛化能力一般,也就是说在一个数据集上表现良好,而在另一个数据集上表现欠佳,并且分割精度仍需要提高。如何提升模型的泛化能力是亟待解决的一个关键问题。

为了提高模型的泛化能力和分割精度,研究学者们以 U-Net 模型为基础,对分割模型进行改进。Jégou 等人(2017) 用密集连接卷积块(Huang 等, 2017) 代替常规卷积操作,提高了网络模型的可训练性。Li 等人(2018) 提出了 DeepUNet 模型,将残差网络思想(Szegedy 等, 2017) 应用于网络模型的下采样块和上采样块中,比 U-Net 分割精度要高 2%。在改变卷积方法的同时,也有人在讨论如何设计模型的宽度和深度以实现高性能。Zhou 等人(2018) 提出了 U-Net++ 模型,其中编码器和解码器子网通过一系列嵌套的密集跳跃连接,重新设计了基线网络中的跳跃连接方式,减少编码器和解码器子网的特征图之间语义信息的差距,并在多个公共医学图像数据集上实现了比 U-Net 更好的精度。上述改进方法表明,改进卷积方法以及网络结构,可以提高模型的学习能力以及分割精度。

基于这一思想,本文提出了一种深度迭代融合的卷积神经网络模型(deep iteration fusion net, DIF-Net),以期提高脑部磁共振图像中颅骨分割的准确度以及模型的泛化能力。主要框架思想是将不同层之间的特征联系起来,跨通道和跨深度的语义信息融合可以提高对学习任务的认知,通过自下而上的特征迭代融合,丰富的深层特征会最大化还原。另外更深层次的融合可以获得不同层的信息,同时以迭代和分层方式合并特征,使模型具有更高的学习能力和更少的参数。此外,采用残差卷积模块替换传统卷积,以使得浅层语义信息更易流入深层网络中,提取更丰富的语义信息,增强网络模型的可训练性。同时,增设双路上采样模块,利用空洞卷积扩大感受野区域,以便还原出更多细节特征,从而有利于细节分割。

1 本文方法

1.1 深度迭代融合模型框架

传统 U-Net 模型结构由编码器和解码器构成,

如图 1(a) 所示 ,浅层特征图通过长跳跃连接直接与深层特征图融合 ,深层特征图通过上采样操作加上相应的浅层特征图 ,弥补了编码器过程中池化操作带来的损失 ,加快网络学习目标特征。在网络深层结构中 ,感受野区域较大 ,能获取更多区域或者全局语义信息;而在网络浅层结构中 ,感受野区域较小 ,获取的特征多包含细节信息。为了加深深层特征与浅层特征之间的融合 ,本文拓扑了 U-Net 结构 ,提出了深度迭代融合模块 ,如图 1(b) 所示。为了保证深层的语义信息能最大化还原 ,解码器中增设一路 ,也就是采用两路上采样的方式 ,同时两路之间由短跳跃连接联系。不同大小的感受野 ,对图像中大小不一的分割目标敏感程度不同 ,通过自下而上地将浅层与深层特征进行迭代融合 ,可以获取不同层次的特征 ,增强模型对分割目标特征的学习。

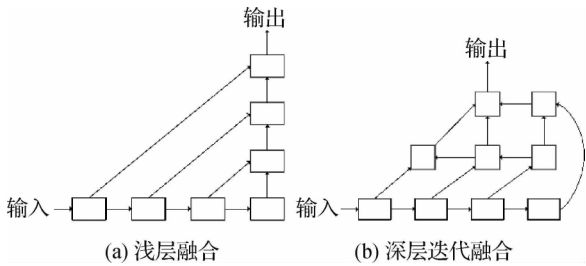


图 1 不同层次融合方式
Fig. 1 Different levels of integration
((a) shallow fusion (b) deep iteration fusion)

1.2 DIFNet 模型结构

DIFNet 网络的具体结构如图 2(c) 。DIFNet 具有编码器和解码器结构 ,整体骨干结构采用的是深度迭代融合思想。在编码器路径上由残差模块 (Res Block) 和步幅为 2 的最大池化层组成。在解

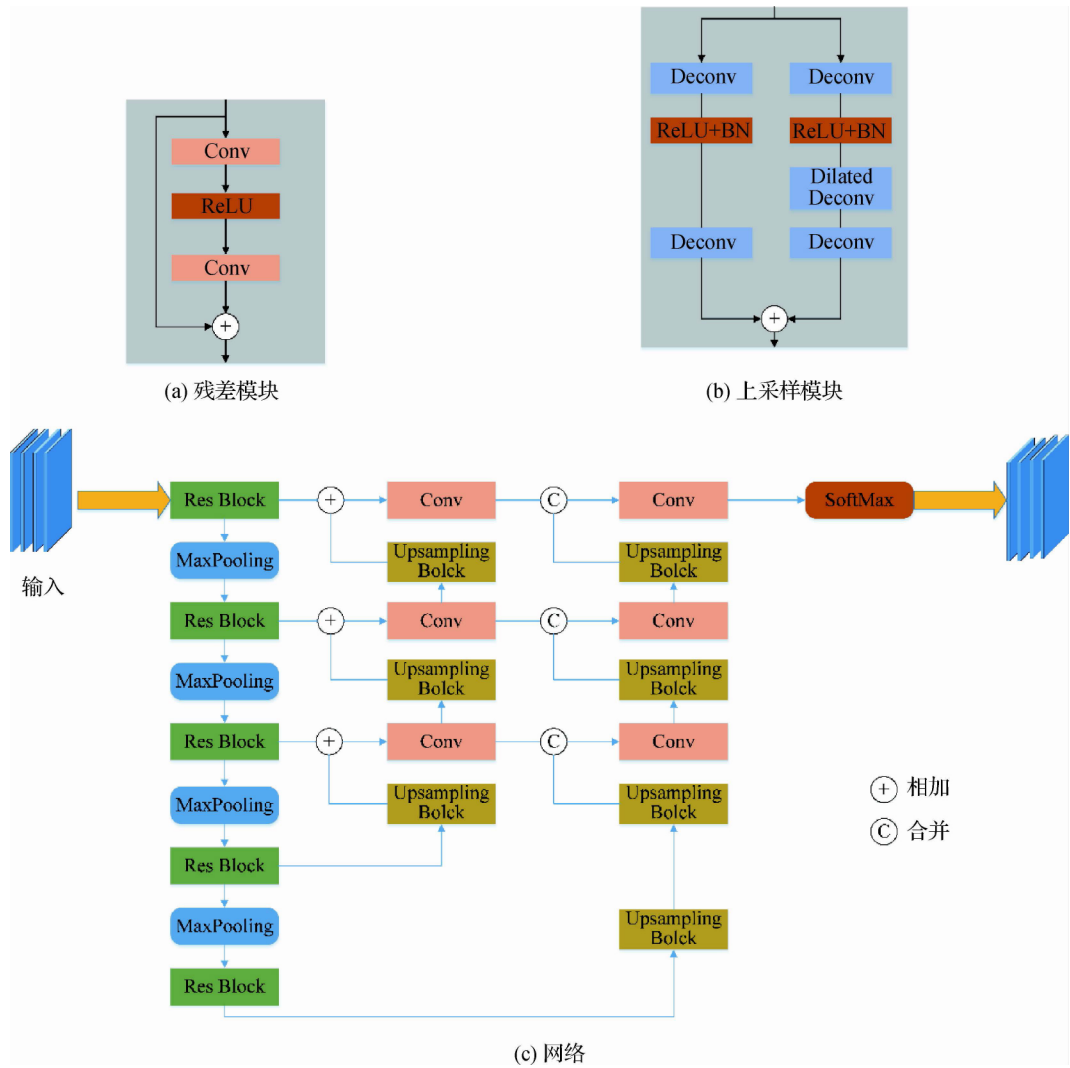


图 2 网络结构

Fig. 2 Architecture of the proposed convolutional network((a) Res Block; (b) Upsampling Block; (c) network)

码器路径中设置了上采样模块(Upsampling Block),采用双路上采样的方式,确保上采样操作能还原更多特征信息。输入的数据首先通过4次Res Block和池化操作,每次执行完池化操作,特征图的尺寸减半,通道数量增加一倍。最后一层采用Res Block操作,提取深层语义信息,然后采用Upsampling Block执行相同次数的上采样,每次执行完上采样操作,输出特征图的分辨率大小增加一倍。在网络结构的中间阶段,加入深度迭代融合的思想,取代U-Net中编码器的特征信息直接与相应层的解码器信息融合操作,将编码器输出的特征图通过上采样扩大一倍,并使用相同大小分辨率的编码器输出特征图进行加法运算,这种连接方式可以对池化操作后丢失的浅层特征信息进行补充,同时确保将特征信息完整地输入到深层阶段。深层特征融合将解码器的深层与中间层输出的特征图进行串联合并,并进行一个卷积操作,将不同层之间的特征进行融合,也增大了提取到的特征数量。最后,模型加入SoftMax,用于分类分割结果。

在编码器阶段加入残差卷积模块,如图2(a)所示。残差卷积模块中加入了ReLU激活函数,同时在模块最后的输出中也加入了ReLU激活函数。模型中的卷积核大小统一设置为 3×3 ,步长为 1×1 。在设置卷积核大小的时候,选择较小的而不是较大的卷积核,因为两个连续的 3×3 卷积层的感受野与 5×5 卷积层的感受野相同,并且前者的计算参数要远少于后者。

对于残差卷积模块,均为

$$x_{l+1} = f_c(f_r(x_l)) + x_l \quad (1)$$

式中, $f_r(\cdot)$ 为卷积操作后加入ReLU激活函数, $f_c(\cdot)$ 为卷积操作, x_l 为输入的特征图, x_{l+1} 为残差卷积模块的输出。

本文提出的上采样模块如图2(b)所示,以对编码器输出的特征图进行上采样。上采样模块具有双路结构,其中反卷积层包含不同大小的卷积核,即 1×1 和 3×3 。左侧路径包含多个反卷积层,并且每个层具有反卷积后的批标准化(batch normalization, BN)和ReLU激活函数。为了增加感受野的大小,在右侧路径中添加了一个卷积核为 3×3 的空洞卷积。最后使用步长为2的反卷积层对左侧和右侧生成的特征图扩大两倍,并通过相加操作输出为最终特征图,即

$$x_{l+1} = f_d^2(f_d(x_l)) + f_d^2(f_k(f_d(x_l))) \quad (2)$$

式中, $f_d(\cdot)$ 代表反卷积操作加BN和ReLU, $f_k(\cdot)$ 为空洞卷积, $f_d^2(\cdot)$ 为步长为2的反卷积操作。

在上采样模块中的卷积核数量逐渐增加。模块中采用不同的卷积核大小可以获取不同感受野区域的特征信息。设计双路上采样结构,主要优点是不同上采样操作将输出的特征图融合。如果仅使用反卷积层对输入的特征图进行上采样不能准确地还原细节信息,也就导致分割结果精度达不到最优。本文上采样模块,使用不同卷积核大小的反卷积层能获取到不同的特征,将最后的结果进行融合,一定程度上减少深层语义特征信息损失。

为了避免过度拟合,在本文模型中,每个Res Block和Upsampling Block之后添加Dropout操作(Srivastava等,2014),利于网络模型训练。

1.3 损失函数

在图像分割任务中,常用交叉熵作为损失函数,交叉熵损失函数将每个像素点的类预测概率与相应的热编码标签数据进行比较,最后将结果求平均值,也就是这个图像分割的效果是由平均值决定的。但是,对于医学图像常有类别不均衡的问题,使用交叉熵作为损失函数导致训练结果会被非感兴趣区域类主导,降低网络的分割性能。因此本文采用Dice损失函数训练网络。Dice通过计算网络的预测图像与对应标记图像的重合程度来评价分割效果。重合区域越大,Dice损失函数值则会越小。同时为了避免过拟合,在损失函数中引入L2正则,即总体损失L为

$$L = 1 - \frac{2 \sum_{x \in N_L} p(x) g(x)}{\sum_{x \in N_L} p(x) + \sum_{x \in N_L} g(x) + e} + \frac{\lambda}{2n} \sum w^2 \quad (3)$$

式中, $p(x)$ 是模型预测图, $g(x)$ 是专家手动分割掩膜, N_L 是每个类别的像素, e 为公差, λ 是正则项系数, n 是所有参数的总数, w 是模型中参数。

2 实验结果与分析

2.1 数据集

使用公开数据集NFBS(neurofeedback skull-stripped)(Puccio等,2016)对本文方法进行验证。该数据集包含来自21—45岁的125位参与者的T1

加权脑部 MRI 图像分辨率为 $256 \times 256 \times 192$ 像素, 每位参与者数据都有专家手动去除颅骨的分割掩膜, 如图 3 所示。本文使用该数据集进行训练、验证和初步评估。此外将在 NFBS 数据集上训练好的模型应用于另一个数据集 LPBA40 (loni probabilistic brain atlas 40) (Shattuck 等, 2008) 进行进一步评估。该数据集由 40 位参与者的 T1 加权核磁共振图像及相应的专家手动标记的去除头骨数据组成。

2.2 数据增强

为了解决过拟合问题, 在训练过程中使用了数据增强技术, 同时增加模型的鲁棒性和泛化能力。采用外部数据增强技术提前将数据进行平移、旋转、缩放等图像增强然后再利用增强后的数据训练模型, 这样产生的问题是训练一个 epoch 时间过长, 增加了网络训练时间。本文采用的数据增强方

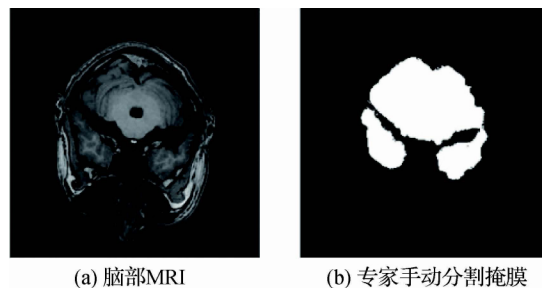


图 3 NFBS 数据展示

Fig. 3 NFBS data illustration

((a) brain MRI; (b) experts manually segmentation)

法为内部增强, 即网络输入数据后, 进行数据增强操作, 如图 4 所示。首先, 给定一个数值区间, 每个输入数据在这个区间内随机选择一个数值作为缩放因子, 然后随机选择缩放图像上的点作为中心点裁剪固定大小, 最后将裁剪后的数据送入网络进行训练。

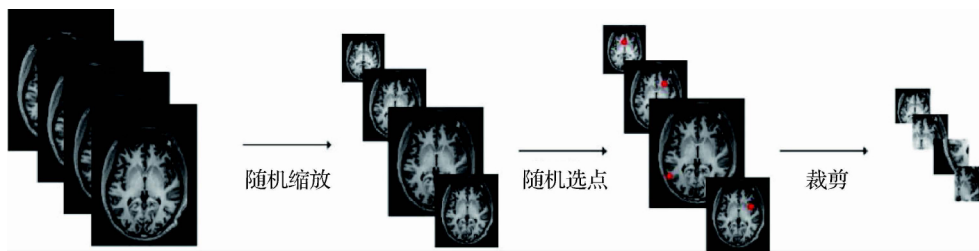


图 4 数据增强过程

Fig. 4 Data augmentation process

2.3 评价指标

采用 3 个指标对分割结果进行评价, 分别为: Dice 值、灵敏度 (S)、特异度 (P)。通过将预测输出图像 P 与专家手动分割的掩膜 R 重叠来计算 Dice 值, 具体定义为

$$Dice = \frac{2|P \cap R|}{|P| + |R|} = \frac{2T_p}{2T_p + F_p + F_N} \quad (4)$$

$$S = \frac{T_p}{T_p + F_N} \quad (5)$$

$$P = \frac{T_N}{T_N + F_p} \quad (6)$$

式中, T_p 表示正确预测的脑组织部分, T_N 表示正确预测的非脑组织部分, F_p 表示将非脑组织部分预测为脑组织, F_N 表示将脑组织部分预测为非脑组织。Dice 值越接近 1, 表示预测结果与手动分割掩膜越接近, 分割结果越准确。灵敏度是正确分类脑组织标签的度量, 其大小可以确定网络模型分割脑室的

效果。特异度是用来计算输出类别的正确分类标签的度量, 其值越高说明网络模型将颅骨与脑室分割的效果越好。

2.4 实验参数设置

本文所有网络模型均在 TensorFlow 平台上实现。采用 Momentum 优化器, 参数 $\beta = 0.9$, 每组实验都训练 100 个 epoch, 初始学习率设置为 $1E-3$, 训练 10 个 epoch 后学习率减少为 $6E-4$, 50 个 epoch 后学习率设置为 $1E-4$ 。本文网络参数的权重采用 L2 正则化, 系数 λ 为 $1E-4$ 。上采样模块中均加入了 BN 操作, 加速训练收敛。使用 NFBS 数据集进行训练模型时, 按照 4:1 的比例, 将数据集随机划分为训练集和验证集。使用 LPBA40 数据集来测试模型。

2.5 实验结果

2.5.1 训练过程中不同模型的对比

为了验证本文模型框架在脑图像分割中的有效

性,分别在原有的2D U-Net模型中加入本文的Res Block和Upsampling Block,形成Res-U-Net和Up-U-Net模型,并对比分析了本文模型与传统方法以及U-Net、Res-U-Net和Up-U-Net的分割结果。

图5显示了训练过程中不同模型在验证集上Dice值的变化。在前10个epoch内,Up-U-Net的收敛速度比Res-U-Net和U-Net快,尽管中期波动较大,但在后期会逐渐收敛,与Res-U-Net验证精度相差不多,比基准网络高一些,说明加入这些模块在后的基准模型精度会有所提升。而本文方法在训练的早期阶段模型收敛最快,学习特征能力较强,后期阶段Dice值并没有波动,趋于稳定,说明在颅骨分割任务上比基准网络具有更好的性能。

2.5.2 分割的可视化结果

为了验证模型颅骨分割的性能,使用两个数据集进行实验。在NFBS验证数据集上分别与BET、BSE、ROBEX、U-Net、Up-U-Net和Res-U-Net进行比较。然后使用在NFBS数据集训练好的DIFNet模型,直接分割LPBA40数据集,不做任何相关训练,以此验证模型的泛化能力。

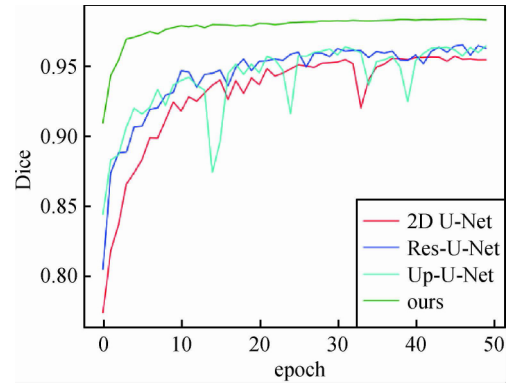


图5 不同模型在验证集上Dice值的变化

Fig.5 Change of Dice score on the validation set of different models

图6是在NFBS数据集上模型预测的结果,其中红色矩形表示分割失败的区域,未将颅骨组织剔除或者过度分割。图中BET和BSE的结果是手动调参3次以上获得的,尽管如此在部分数据上还会出现过度分割的现象。常用的颅骨分割方法并不能完全将颅骨与脑室分割开。ROBEX方法虽然不需要手动调参,但是分割结果不稳定,也会出现过度分割的现象。用深度学习的方法,把整幅图像输入到

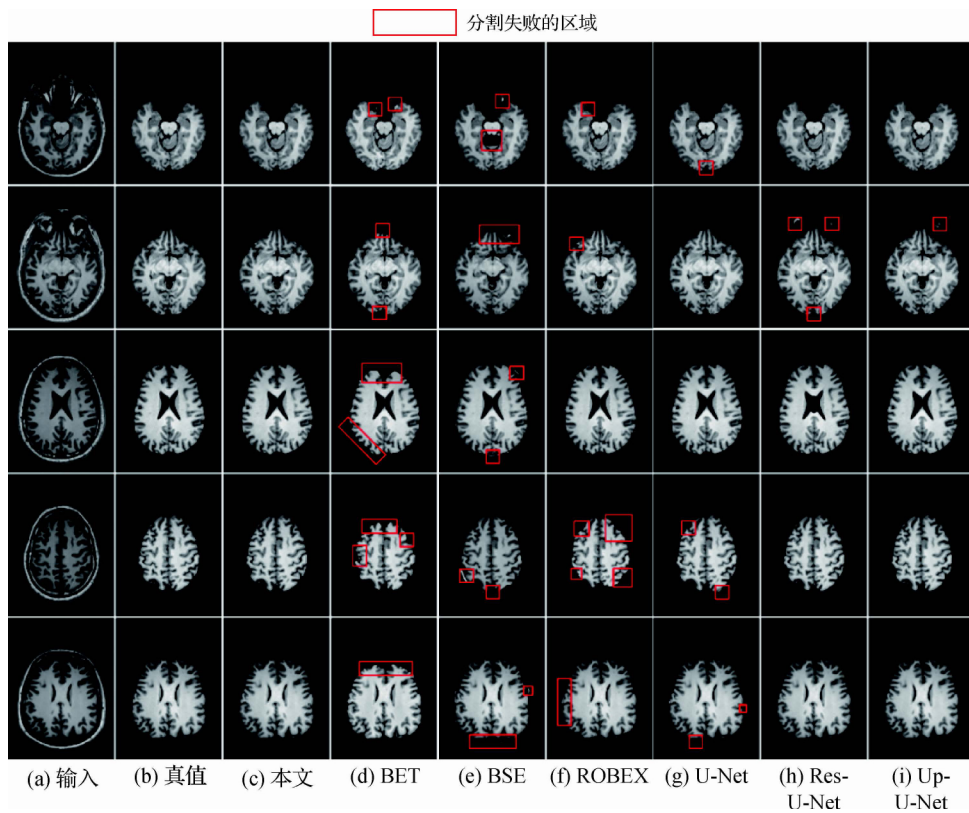


图6 多种方法的分割结果比较

Fig.6 Comparison of segmentation results of multiple methods

((a) input; (b) ground truth; (c) ours; (d) BET; (e) BSE; (f) ROBEX; (g) U-Net; (h) Res-U-Net; (i) Up-U-Net)

网络进行测试。U-Net 网络的分割结果在细节上比 Res-U-Net、Up-U-Net 表现较差,分割目标较小时,往往无法准确分割出脑室。本文方法的分割结果与金标准相近,在脑实质边缘处分割结果较为平滑,没有出现过度分割的现象。

图 7 是在 LPBA40 数据集上从 3 个方向展示测试的结果。从图中可以看出,BSE 和 ROBEX 都会将脑室边缘区域错误分割,出现过度分割的现象,BET 则不能完全将颅骨分割出去。为了进一步验证 DIFNet 模型的泛化能力,将同时在 NFBS 数据集训练好的 U-Net、Res-U-Net 和 Up-U-Net 模型去测试 LPBA40 数据集,这 3 种模型都出现了过度分割的现象,分割结果中也出现了噪点,导致整个图像的灰度值范围与其他方法差异较大,从图中可以看出这 3 种模型分割结果亮度偏暗。从矢状方向和冠状方向,可以看出 Res-U-Net 和 Up-U-Net 模型并不能完全将颅骨与脑部区域分割开,图像中存在不连续的颅骨组织,并且脑部下的颈部区域也不能很好剔除。这 3 种模型的分割结果与常用的 3 种软件相比,相差较大。而本文方法在 NFBS 数据集上学习到的知识应用于 LPBA40 上,并没有因为采集设备、图像灰度、图像分辨率不同而分割较差,从 3 个方向相比都优于常用的去除颅骨软件和基线模型的分割结果。

2.5.3 定量分析

为了更直观地体现本文方法在颅骨分割任务上的性能优势,从 Dice 值、灵敏度、特异度和模型参数上与常用的方法和模型进行比较。表 1 是在 NFBS 数据集上多种方法的评价指标对比。可以看出,本文模型在 Dice 值、灵敏度和特异度上均为最高。常用的颅骨分割方法的分割精度比深度学习的方法都要低,表明用深度学习的方法更能准确分割出脑室与颅骨。Up-U-Net 与 2D U-Net 相比有了 0.1 的提升,说明使用上采样模块替代简单的反卷积操作可以还原更多的特征信息。但是 Res-U-Net 的 Dice 值并没有多大提升,可以看出仅增加模型的深度并不能带来更好的效果。

图 8 给出了不同方法对 NFBS 数据测试集上分割结果的盒形图,更加直观地表现出多种方法对 25 个脑部 MRI 数据的统计分割结果。可以看出,BET 和 BSE 在 3 个评价指标的盒形图中数值波动区间较大,说明这两个分割方法对不同数据分割结

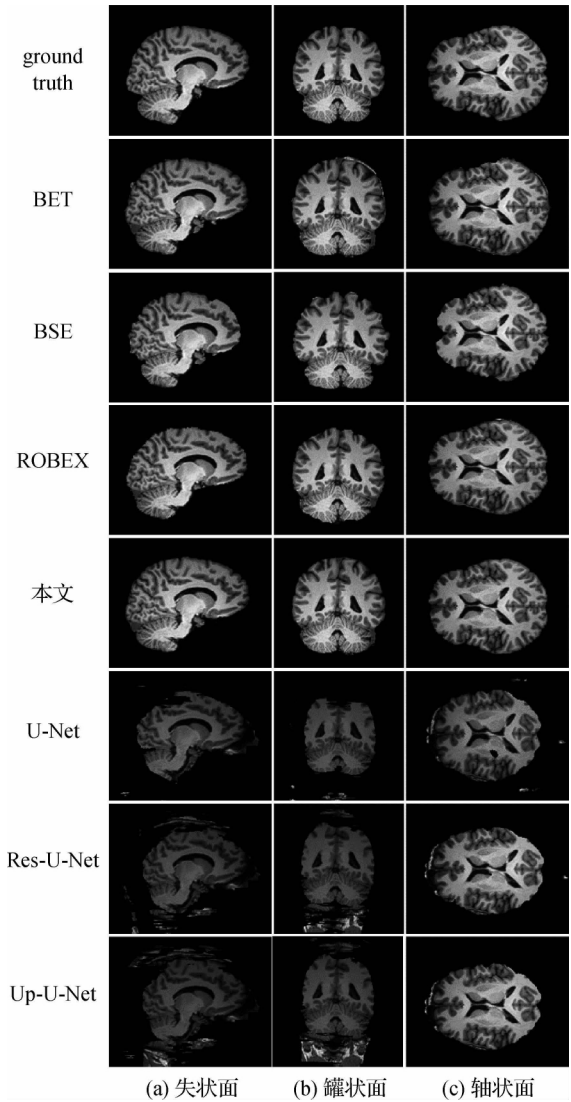


图 7 LPBA40 数据集上的分割结果

Fig. 7 Segmentation results on LPBA40 dataset
((a) sagittal; (b) coronal; (c) axial)

表 1 多种方法在 NFBS 数据集上的定量对比

Table 1 Quantitative comparison of multiple methods on the NFBS dataset

方法	Dice 值	灵敏度	特异度	参数/M
BSE	0.922 9	0.883 9	0.988 6	—
BET	0.893 1	0.909 8	0.978 1	—
ROBEX	0.952 1	0.963 6	0.993 1	—
2D U-Net	0.972 4	0.985 6	0.983 6	8.63
Res-U-Net	0.973 1	0.987 1	0.982 8	9.75
Up-U-Net	0.985 2	0.985 3	0.994 4	8.81
本文(DIFNet)	0.991 2	0.992 2	0.995 0	11.39

注:加粗字体为每列最优值,“—”表示无参数。

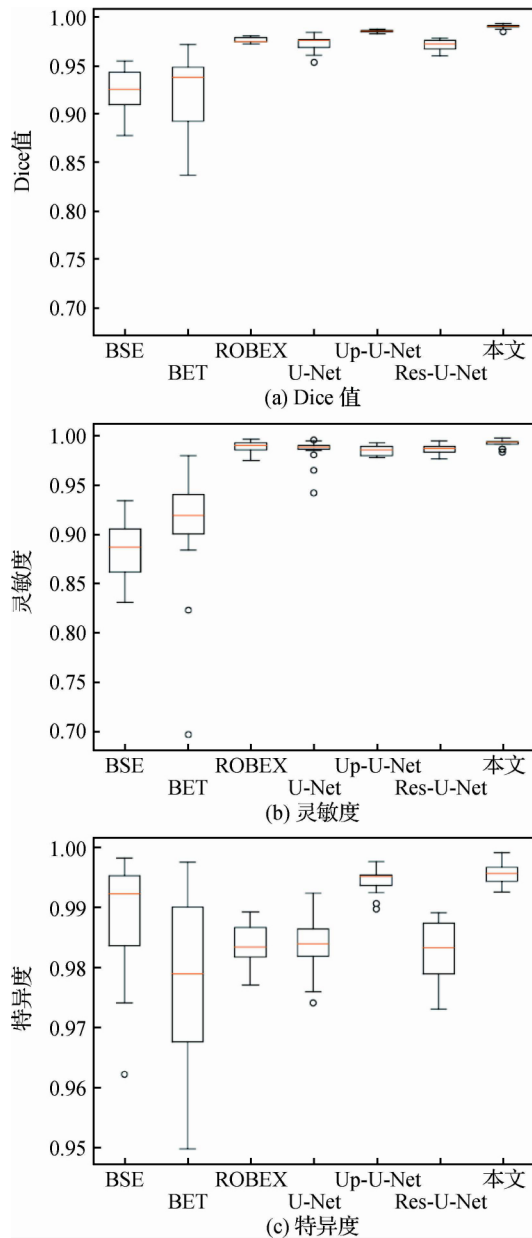


图8 在NFBS数据集上的定量结果盒形图

Fig. 8 Boxplot of quantitative results on the NFBS dataset
((a) Dice score (b) sensitivity (c) specificity)

果差异性较大,对样本较为敏感。

ROBEX 分割结果较为稳定,U-Net 的灵敏度区间相比其他方法有较多异常值,说明分割效果不稳定,在部分数据中分割脑室结果较差。Up-U-Net 的 Dice 值分布区间比 U-Net 稳定,同时在特异度区间中前者都高于后者,说明增强上采样操作,能促使模型性能更加稳定。在特异度盒形图中,Res-U-Net 并没有表现得很好,中位线与 ROBEX 相近,区分脑室与颅骨组织结果较差。相对本文方法,Dice 值盒形图和灵敏度盒形图,最大值与最小值之间的差值最

小,比其他方法都稳定。

为了验证本文方法的泛化能力,表2定量对比了不同模型在采用 NFBS 数据集训练后,直接用于 LPBA40 数据集测试分割得出的定量分析结果。可以看出,传统方法并没有取得良好的结果,也进一步表明 LPBA40 数据集的复杂性高于 NFBS 数据集,使用传统分割软件无法取得较好的分割结果。而采用深度学习的方法,U-Net、Up-U-Net 和 Res-U-Net 并没有像之前在 NFBS 上表现良好的分割精度,本文方法在几种方法对比中性能都是最高的,也验证了 DIFNet 的泛化能力优于其他模型。

表2 在 LPBA40 数据集上的定量对比

Table 2 Quantitative comparison on the LPBA40 dataset

方法	Dice 值	灵敏度	特异度
BSE	0.830 5	0.831 9	0.902
BET	0.927 7	0.889 2	0.985 8
ROBEX	0.917 7	0.979 3	0.895 1
U-Net	0.816 9	0.837 6	0.943 3
Up-U-Net	0.773 4	0.792 1	0.867 8
Res-U-Net	0.764 2	0.781 9	0.889 7
本文(DIFNet)	0.981 6	0.981 4	0.992 8

注:加粗字体为每列最优值。

3 结 论

本文提出了一种深度迭代融合的卷积神经网络模型 DIFNet,用于脑部 MRI 图像去颅骨分割。该模型采用编码器和解码器结构,中间加入短跳跃连接,加强不同层之间特征的融合。在编码器路径上加入了残差卷积模块,优化模型学习特征能力;在解码器中加入了双路上采样模块,还原出更加细节的特征;中间跳跃连接层由多个上采样迭代融合形成,加快了浅层语义信息输入到深层网络中,从而提高分割的精度。为了验证模型分割的准确性和所提模型分割的泛化能力,采用部分 NFBS 数据集进行训练,并利用剩余数据集进行测试,同时采用独立数据集 LPBA40 测试,实验结果表明,与现在流行的传统分割方法相比,本文模型在 Dice 值上提高了至少 3.91%,同时与深度学习分割方法相比,本文模型在 3 个评价指标上都是最高的,尤其是在独立数据集

LPBA40 上, Dice 值至少较其他模型高出 16.47%, 其鲁棒性明显高于其他模型。虽然本文模型可以较好地去除颅骨, 在后续工作中, 如何进一步优化网络模型、调节网络参数, 使模型适合于脑组织区域的细化分割是未来的主要研究内容。

参考文献 (References)

- Archibald R, Chen K W, Gelb A and Renaut R. 2003. Improving tissue segmentation of human brain MRI through preprocessing by the Gegenbauer reconstruction method. *NeuroImage*, 20 (1): 489-502 [DOI: 10.1016/s1053-8119(03)00260-x]
- Han X, Kwitt R, Aylward S, Bakas S, Menze B, Asturias A, Vespa P, Van Horn J and Niethammer M. 2018. Brain extraction from normal and pathological images: a joint PCA/image-reconstruction approach. *NeuroImage*, 176: 431-445 [DOI: 10.1016/j.neuroimage.2018.04.073]
- Huang A, Abugharbieh R, Tam R and Traboulsee A. 2006. MRI brain extraction with combined expectation maximization and geodesic active contours // 2006 IEEE International Symposium on Signal Processing and Information Technology. Vancouver: IEEE: 107-111 [DOI: 10.1109/ISSPIT.2006.270779]
- Huang G, Liu Z, Van Der Maaten L and Weinberger K Q. 2017. Deeply connected convolutional networks // Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE: 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Hwang H, Rehman H Z U and Lee S. 2019. 3D U-Net for skull stripping in brain MRI. *Applied Sciences*, 9 (3): 569 [DOI: 10.3390/app9030569]
- Iglesias J E, Liu C Y, Thompson P M and Tu Z W. 2011. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Transactions on Medical Imaging*, 30 (9): 1617-1634 [DOI: 10.1109/TMI.2011.2138152]
- Jégou S, Drozdal M, Vazquez D, Romero A and Bengio Y. 2017. The one hundred layers tiramisu: fully convolutional denseNets for semantic segmentation // Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu: IEEE: 1175-1183 [DOI: 10.1109/CVPRW.2017.156]
- Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K, Bendszus M and Armin B. 2016. Deep MRI brain extraction: a 3D convolutional neural network for skull stripping. *NeuroImage*, 129: 460-469 [DOI: 10.1016/j.neuroimage.2016.01.024]
- Li R R, Liu W J, Yang L, Sun S L, Hu W, Zhang F and Li W. 2018. DeepUNet: a deep fully convolutional network for pixel-level sea-land segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11 (11): 3954-3962 [DOI: 10.1109/JSTARS.2018.2833382]
- Liu Z, Xiao B, Li Y and Fan Y. 2019. Context-encoding for neural network based skull stripping in magnetic resonance imaging [EB/OL]. [2020-05-29]. <https://arxiv.org/pdf/1910.10798.pdf>
- Lucena O, Souza R, Rittner L, Frayne R and Lotufo R. 2019. Convolutional neural networks for skull-stripping in Brain MR imaging using silver standard masks. *Artificial Intelligence in Medicine*, 98: 48-58 [DOI: 10.1016/j.artmed.2019.06.008]
- Puccio B, Pooley J P, Pellman J S, Taverna E C and Craddock R C. 2016. The preprocessed connectomes project repository of manually corrected skull-stripped T1-weighted anatomical MRI data. *GigaScience*, 5 (1): #45 [DOI: 10.1186/s13742-016-0150-5]
- Rex D E, Shattuck D W, Woods R P, Narr K L, Luders E, Rehm K, Stolzner S E, Rottenberg D A and Toga A W. 2004. A meta-algorithm for brain extraction in MRI. *NeuroImage*, 23 (2): 625-637 [DOI: 10.1016/j.neuroimage.2004.06.019]
- 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: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Shattuck D W, Mirza M, Adisetiyo V, Hojatkashani C, Salamon G, Narr K L, Poldrack R A, Bilder R M and Toga A W. 2008. Construction of a 3D probabilistic atlas of human cortical structures. *NeuroImage*, 39 (3): 1064-1080 [DOI: 10.1016/j.neuroimage.2007.09.031]
- Sara S, Samir B, Ahmed H and Bouchaib C. 2014. A robust comparative study of five brain extraction algorithms (BET; BSE; McStrip; SPM2; TMBE) // Proceedings of the 2nd World Conference on Complex Systems. Agadir: IEEE: 632-636 [DOI: 10.1109/ICoCS.2014.7060986]
- Shattuck D W, Sandor-Leahy S R, Schaper K A, Rottenberg D A and Leahy R M. 2001. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13 (5): 856-876 [DOI: 10.1006/nimg.2000.0730]
- Smith S M. 2002. Fast robust automated brain extraction. *Human Brain Mapping*, 17 (3): 143-155 [DOI: 10.1002/hbm.10062]
- Somasundaram K and Kalaiselvi T. 2010. Fully automatic brain extraction algorithm for axial T2-weighted magnetic resonance images. *Computers in Biology and Medicine*, 40 (10): 811-822 [DOI: 10.1016/j.combiomed.2010.08.004]
- Srivastava N, Geoffrey H, Alex K, Ilya S and Ruslan S. 2014. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15 (1): 1929-1958
- Szegedy C, Ioffe S, Vanhoucke V and Alemi A. 2017. Inception-v4, inception-ResNet and the impact of residual connections on learning // Proceedings of the 31st AAAI Conference on Artificial Intelligence. USA: AAAI Press: 4278-4284
- Valvano G, Martini N, Leo A, Santini G, Latta D D, Ricciardi E and Chiappino D. 2018. Training of a skull-stripping neural network with efficient data augmentation [EB/OL]. [2020-05-29]. <https://arxiv.org/pdf/1810.10853v1.pdf>

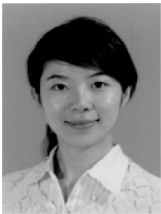
Zhou Z W , Siddiquee M M R , Tajbakhsh N and Liang J M. 2018.
Unet ++: a nested U-Net architecture for medical image segmenta-
tion//Proceedings of the 4th International Workshop on Deep Learn-
ing in Medical Image Analysis and Multimodal Learning for Clinical
Decision Support. Granada: Springer: 3-11 [DOI: 10.1007/978-
3-030-00889-5_1]

作者简介



姚发展,1994年生,男,硕士研究生,主要研究方向为医学图像处理及分割、深度学习。

E-mail: gs_fzyao18@gzu.edu.cn



李智,通信作者,女,教授,主要研究方向为医学成像、图像处理、模式识别与机器学习。

E-mail: lizhi@163.com

王丽会,女,副教授,主要研究方向为医学图像、图像处理、模式识别与机器学习。E-mail: lhwang@gzu.edu.cn

程欣宇,男,副教授,主要研究方向为机器学习、机器视觉、软件工程与网络通信。E-mail: ai_xycheng@qq.com

张健,男,讲师,主要研究方向为图像处理、3维重建和计算机视觉。E-mail: jzhang7@gzuu.edu.cn