

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2020)10-2206-12

论文引用格式: Gong R L, Shi J and Wang J. 2020. Hybrid supervised dual-channel feedback U-Net for segmentation of breast ultrasound images. Journal of Image and Graphics 25(10): 2206-2217(贡荣麟, 施俊, 王骏. 2020. 面向乳腺超声图像分割的混合监督双通道反馈 U-Net. 中国图象图形学报, 25(10): 2206-2217 [DOI: 10.11834/jig.200240])

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

上海大学通信与信息工程学院, 上海 200444

摘要: 目的 基于超声图像的乳腺病灶分割是实现乳腺癌计算机辅助诊断和定量分析的基本预处理步骤。由于乳腺超声图像病灶边缘通常较为模糊, 而且缺乏大量已标注的分割图像, 增加了基于深度学习的乳腺超声图像分割难度。本文提出一种混合监督双通道反馈 U-Net(hybrid supervised dual-channel feedback U-Net, HSDF-U-Net) 算法, 提升乳腺超声图像分割的准确性。方法 HSDF-U-Net 通过融合自监督学习和有监督分割实现混合监督学习, 并且进一步通过设计双通道反馈 U-Net 网络提升图像分割准确性。为了改善标记数据有限的问题, 首先在自监督学习框架基础上结合标注分割图像中的标签信息, 设计一种边缘恢复的辅助任务, 以实现对比病灶边缘表征能力更强的预训练模型, 然后迁移至下游图像分割任务。为了提升模型在辅助边缘恢复任务和下游分割任务的表现, 将循环机制引入经典的 U-Net 网络, 通过将反馈的输出结果重新送入另一个通道, 构成双通道编码器, 然后解码输出更精确的分割结果。结果 在两个公开的乳腺超声图像分割数据集上评估 HSDF-U-Net 算法性能。HSDF-U-Net 对 Dataset B 数据集中的图像进行分割获得敏感度为 0.848 0、Dice 为 0.826 1、平均对称表面距离为 5.81 的结果, 在 Dataset BUSI(breast ultrasound images) 数据集上获得敏感度为 0.803 9、Dice 为 0.803 1、平均对称表面距离为 6.44 的结果。与多种典型的 U-Net 分割算法相比, 上述结果均有提升。结论 本文所提 HSDF-U-Net 算法提升了乳腺超声图像中的病灶分割的精度, 具备潜在的应用价值。

关键词: 乳腺超声图像分割; 深度学习; 自监督学习; 混合监督学习; 双通道反馈 U-Net

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

Gong Ronglin, Shi Jun, Wang Jun

School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China

Abstract: Objective In the clinical diagnosis and treatment of breast cancer, ultrasound imaging is widely used because of its real-time, non-radiation, and low cost. Automatic segmentation of breast lesions is a basic preprocessing step for computer-aided diagnosis and quantitative analysis of breast cancer. However, breast ultrasound segmentation is challenging. First, more noises and artifacts can be found in ultrasound images, and the boundary of the lesions is more ambiguous than the foreground in general segmentation tasks. Second, the size of the lesions in different sample images is different. In addition, benign and malignant lesions are quite different. The segmentation effect depends on the ability of the algorithm to understand the overall image. However, traditional methods rely on the characteristics of artificial design, which is difficult

收稿日期: 2020-05-31; 修回日期: 2020-07-08; 预印本日期: 2020-07-15

基金项目: 国家自然科学基金项目(81830058); 国家自然科学基金重大科研仪器研制项目(81627804); 上海市科委项目(17411953400, 18010500600, 18411967400)

Supported by: National Natural Science Foundation of China (81830058, 81627804)

to deal with the noise and image structures. In recent years, excellent segmentation models such as U-Net are identified in the field of medical image segmentation. Many algorithms are based on U-Net, such as Auto-U-Net. Auto-U-Net uses the idea of iterative training to form a new input from the probability graph of model output and the original graph and sends them to the new U-Net model for training. However, the number of models needed in Auto-U-Net is the same as the total number of iterations, which leads to a complex training process and inefficient parameter utilization. The segmentation algorithm based on deep learning has a certain demand for data scale and annotation quality, whereas the professional requirements for accurate annotation of medical image data are high. Therefore, the number of samples cannot be guaranteed, resulting in the limited performance of the deep learning model. For the above mentioned challenges, in addition to transfer learning methods, using self-supervised learning to assist the training process is also a feasible solution. Considering that self-supervised learning emphasizes self-learning, this feature can deal with the problem of high cost in medical image field. Compared with the common transfer learning method, the advantage of self-supervised learning in the field of medical image lies in the stronger correlation between pretext task and target task. At present, in the field of medical image, the focus of research on self-supervised learning method is self-monitoring learning, whereas semantic segmentation is only a downstream task to evaluate self-supervised learning to features. In this process, the assistant task lacks the effective use of label information. Facing these limitations, this paper proposes a hybrid supervised dual-channel feedback U-Net (HSDF-U-Net) to improve the accuracy of breast ultrasound image segmentation. **Method** HSDF-U-Net achieves hybrid supervised learning by integrating self-supervised learning and supervised segmentation and improves the accuracy of image segmentation by developing a dual channel feedback U-Net network. The algorithm designs the edge recovery task on the basis of the information in the segmentation label to enhance the correlation between the pretext task and the target task in self-supervised learning. The location information of contour pixels is extracted from the segmentation label. The images with ambiguous edge and the images with gray value close to the segmentation mask are obtained by using this information. They are used as input and label of deep learning to obtain the pre-training model with stronger ability to represent the edge of lesions and then transferred to the downstream image segmentation task. In addition, the feedback mechanism is introduced into U-Net to improve the performance of the model in the pretext edge restoration task and the downstream segmentation task. The mechanism is based on the general feed-forward convolutional neural network (CNN) and integrates the idea of weight sharing in the recurrent neural network. Through feeding back feature map, the prediction results are continuously refined. Therefore, we propose a dual channel feedback U-Net. The output probability map is fed back to the coding stage as the input of the encoder part probability channel. It forms a dual channel input together with ultrasonic image, which is encoded and fused separately before decoding. Consequently, the prediction results are continuously refined. **Result** The performance of HSDF-U-Net algorithm was evaluated on two open breast ultrasound image segmentation datasets. HSDF-U-Net segmented the image in Dataset B obtained sensitivity of 0.848 0, dice of 0.826 1, and the average symmetrical surface distance of 5.81. The sensitivity of 0.803 9, dice of 0.803 1, and the average symmetrical surface distance of 6.44 were obtained on Dataset breast ultrasound images (BUSI). The above mentioned results were improved compared with the typical deep learning segmentation algorithm. **Conclusion** In this study, the proposed HSDF-U-Net improves the accuracy of breast ultrasound image segmentation, indicating potential application value.

Key words: breast ultrasound image segmentation; deep learning; self-supervised learning; hybrid supervised learning; dual-channel feedback U-Net

0 引言

乳腺癌是女性最常见的癌症,严重危害女性身体健康。在乳腺癌的临床检测与诊治中,超声成像以其实时性、无辐射和检查费用低等优点而被广泛应用(Sahiner等,2007)。

基于超声图像的乳腺病灶分割是乳腺癌的计算机辅助诊断、定量分析的常用基本预处理步骤(Kim等,2013)。然而,由于超声图像斑点噪声、超声伪影等因素的影响,导致病灶边缘通常较为模糊(Noble和Boukerroui,2006);不同个体的病灶大小不一,良恶性病灶边缘差异较大,增加了病灶分割的难度。虽然针对乳腺超声图像分割的各种算法不断提

出(Huang 等 2017) ,但仍然是一项困难的任务。

基于深度学习的图像分割方法广泛应用于各种图像分割任务(Litjens 等 2017) 。在医学图像分割领域 ,针对不同影像模态、不同组织器官 ,基于深度学习的图像分割算法表现出较传统分割方法更为优异的性能(Hesamian 等 2019; 边子健 等 2018; 江宗康 等 2020) 。其中 ,U-Net 是基于卷积神经网络(convolutional neural network ,CNN) 的图像分割方法的典型代表(Ronneberger 等 2015) 。该网络由高度对称的编码和解码结构组成 ,编码和解码结构之间的跳跃连接有助于多尺度特征的复用 ,从而有效改善训练中梯度消失的问题。

U-Net 在各种医学图像分割任务中获得了广泛应用 ,但由于医学图像的多样性、复杂性 ,针对不同的分割任务 ,各种改进的 U-Net 算法也不断提出。例如 ,Salehi 等人(2017) 基于 Auto-Context 的迭代训练思想 ,提出了 Auto-U-Net 算法 ,将模型输出的概率图和原图组成新的输入 ,送入级联的下一个新的 U-Net 进行训练 ,通过不断迭代训练来不断提高分割精度。但是 ,Auto-U-Net 算法中需要的模型数量与迭代总次数相同 ,这导致整个网络的训练流程非

常复杂 ,网络结构非常庞大 ,并且堆叠端到端的网络模型会产生大量的参数 ,但并未获得有效的利用。

值得注意的是 ,前馈 CNN 模型在一次训练完成以后 ,即使参数优化结果陷入局部最小值 ,但是模型无进一步拟合的可能。因此 ,有学者提出了反馈机制 ,即在一般的前馈 CNN 基础上 ,融入循环神经网络中权重共享的思想 ,通过将高层次信息重新输入网络中较浅的层 ,进一步修正整体学习状态 ,提升了图像超分辨率的重建性能(Li 等 2019) 。因此 ,将反馈机制嵌入 U-Net ,在控制网络结构复杂性的前提下 ,具有提升图像分割精度的可行性。

此外 ,基于深度学习的图像分割方法(包括 U-Net) 对数据规模和标注质量有较高的要求(Tan 等 2018) ,而医学影像数据的精确标注对于专业性的要求非常高 ,导致实现大批量精准标注医学图像非常困难 ,从而影响了基于深度学习的分割模型的性能提升 ,训练过程中容易发生过拟合、不稳定等情况。针对这一问题 ,可以直接通过改进网络架构提升模型性能 ,有针对性地在监督方式或训练策略上进行调整也是一个可行的思路。两类方法的侧重点如表 1 所示。

表 1 不同算法侧重点比较
Table 1 Comparison of different algorithms

	方法	描述	基本网络架构	监督方式	总结
网络架构改进	内嵌 U-Net(Zhou 等 2018)	注意力门机制	U-Net	监督学习	通过各种拓扑连接方式和特征提取模块或机制 ,对网络架构的整体或局部进行改进
	注意力 U-Net(Oktay 等 2018)	多解码输出	U-Net	监督学习	
监督方式改进	上下文恢复(Chen 等 2019)	图像上下文恢复	编解码结构	自监督学习	针对数据标注不足的问题 ,对数据集的使用方式进一步挖掘 ,从而提升模型的泛化性能
	通用自主学习(Zhou 等 2019)	多变换混合恢复	编解码结构	自监督学习	

为了提升面向小样本医学图像的深度学习分割算法的性能 ,研究人员不断研究提出了各种改进方法 ,其中较为典型的机器学习方法包括半监督学习、弱监督学习和迁移学习等(Cheplygina 等 2019) 。

自监督学习吸引了越来越多研究兴趣。自监督学习属于无监督学习的一种 ,基于无标注数据设计自主生成伪标签的辅助任务 ,从而帮助提升网络性能(Kolesnikov 等 2019) 。在基于 CNN 的自监督学习方法中 ,一般通过训练辅助任务得到一个预训练模型 ,然后将其迁移至下游其他的机器学习新任务。该方法只需进行网络参数微调 ,即可提升模型在该任务上的性能(Jing 和 Tian ,2020) 。

由于自监督学习强调自主学习 ,这种特性适合应用于有限标注样本的医学图像分割任务 ,所以逐步开始应用于医学图像分割任务。例如 ,Chen 等人(2019) 提出了一种基于图像上下文恢复的自监督学习任务 ,通过编解码结构训练图像恢复模型 ,然后将其迁移至医学图像分割任务 ,提升了分割性能。已有研究表明: 对于 2 维医学图像分割任务 ,通过自监督学习预训练的深度神经网络模型可以达到与 ImageNet 预训练模型相似的效果(Zhou 等 2019) 。利用自监督学习的编解码模型可以同时迁移编解码两个部分的权重 ,信息量更为全面。

值得注意的是,由于自监督学习是属于无监督学习的一种方法,重在强调有效利用无标签数据来自我生成学习任务,以获得泛化性较强的预训练网络模型。而在医学图像分割任务中,常存在一定数量的标注图像。而这些标签信息能有效引导网络模型的训练。因此,在自监督学习的框架中嵌入标签信息,能够引导网络训练,增强与下游分割任务的相关性,从而形成新的混合监督学习模型,提升医学图像分割模型的性能。

因此,针对乳腺超声图像分割任务中存在的已标注样本有限的问题,本文提出一种混合监督双通道反馈 U-Net (hybrid supervised dual-channel feedback U-Net, HSDF-U-Net) 分割算法。首先设计了一种新的双通道反馈 U-Net (DF-U-Net) 网络以提升图像分割准确性;进一步通过融合自监督学习框架和已分割标注的图像,以混合监督的方式设计一种新的辅助学习任务,加强 DF-U-Net 模型在小样本学习中的表征能力和泛化性能。

本文工作的主要创新点包括:

1) 将循环机制引入经典的 U-Net 网络,提出一种新的 DF-U-Net,将前一次网络输出的概率图和原始超声图像分别作为下一次网络循环的两个通道输入,在编码器部分独立进行网络学习之后再进行特征融合,由一个解码器完成图像分割,通过循环训练不断学习优化分割图像。

2) 提出了一种基于自监督学习的混合监督学习模型,以自监督学习为基础框架,结合已分割标注图像中的标签信息,设计一种新的图像边缘恢复辅助任务,以混合监督模式提升预训练网络的泛化能力,增强与下游分割任务的相关性,提升 DF-U-Net 模型对病灶边缘的表征能力。

1 方法

1.1 总体流程

本文提出的 HSDF-U-Net 的算法流程如图 1 所示。该算法分为两大模块:结合标注图像的混合监督辅助任务模块和下游图像分割任务模块。

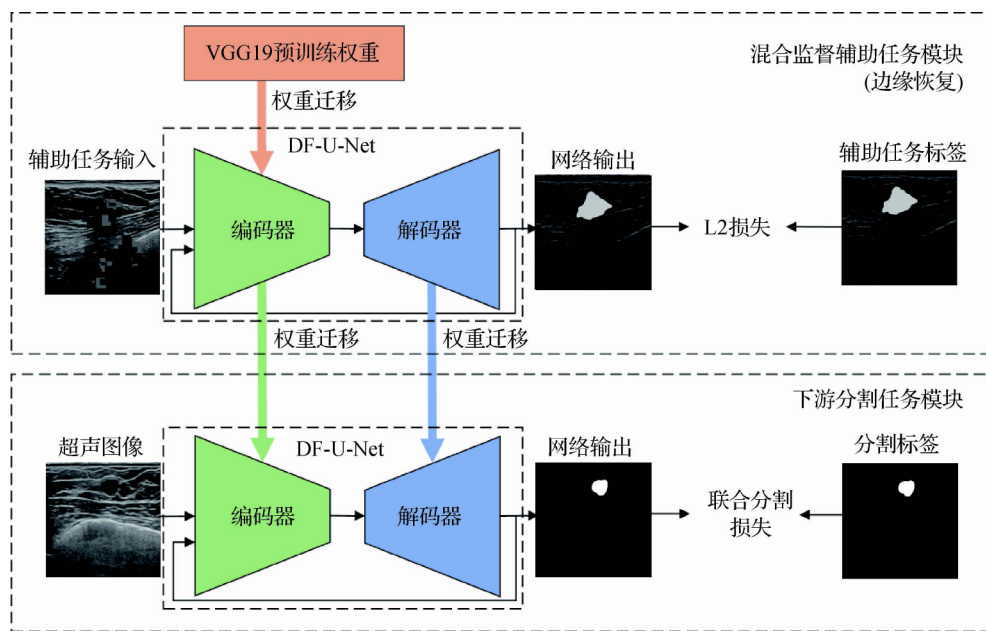


图1 HSDF-U-Net 算法框架

Fig. 1 The algorithm framework of HSDF-U-Net

1) 基于 DF-U-Net 的混合监督辅助任务模块: 从分割标签中获取轮廓像素点的位置信息,利用这些信息对原始超声图像进行变换,得到病灶边缘模糊的图像,前景和背景灰度值接近分割掩码的图像,以此作为所设计的边缘恢复辅助任务的输入和标签。

利用 L2 损失函数指导 DF-U-Net 网络模型学习两类图像的映射关系,完成边缘恢复。

2) 基于 DF-U-Net 的下游图像分割任务模块: 将边缘恢复辅助任务中训练获得的 DF-U-Net 作为下游图像分割任务的预训练模型进行迁移,通过

实际分割任务的训练数据对网络模型进行微调。在分割任务训练过程中,该网络将输出的概率图反馈至编码阶段,与原始超声图像形成两个网络通道的输入,通过多次反馈得到更精确的分割预测。该过程利用联合分割损失函数进行训练微调。

1.2 混合监督边缘恢复任务

在分割任务中,模型的表现很大程度取决于对病灶边缘的表征能力,而分割标签中含有前景区域的轮廓信息,可以对应到原始超声图像中病灶边缘像素点位置。考虑到既然最终目的是提升模型的分割性能,那么在设计辅助任务时,则不应限于自监督学习这个无监督学习模式。因此,本文结合原数据集中成对的超声图像和分割标注掩码图像,分别生

成混合监督学习中的输入和标签,设计了一个混合监督边缘恢复任务。

如图2所示,利用标注图像中的标签信息,定位得到原图中的病灶边缘像素的位置集合。随机选取集合中的部分像素点,在这些像素点上以子块为单位进行随机尺寸大小、随机灰度值的遮盖,并对该区域进行高斯模糊得到边缘恢复后的图像。以该图像作为DF-U-Net网络模型的输入,可以促使模型在边缘恢复的过程中学习高分辨率特征,细化对病灶边缘形态的理解。在图像的任意位置随机生成边缘模糊的伪病灶。模型要识别出这些伪病灶完成正确的映射,则需要结合超声影像中病灶周围的背景信息和图像的整体结构,从而促使模型更好地学习到深层语义特征。

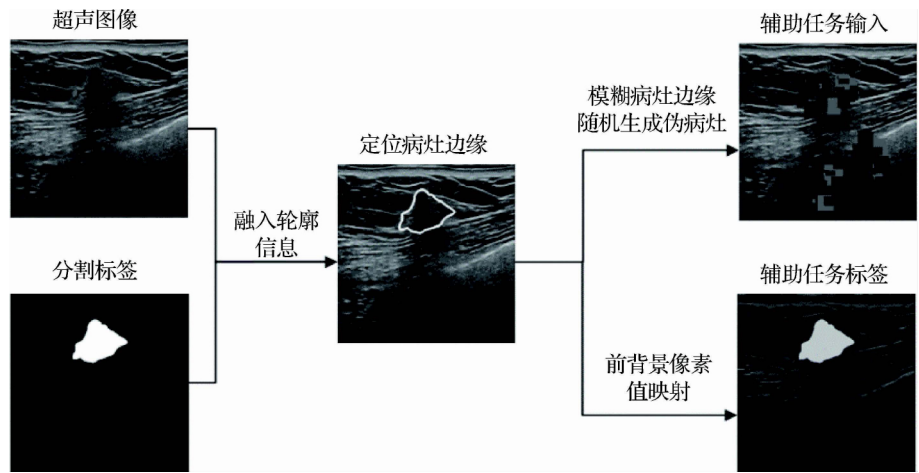


图2 混合监督边缘恢复任务设计

Fig.2 Design of hybrid supervised edge restoration task

本文将这些操作组合作为混合监督的输入。同时,考虑到预训练模型泛化性能的重要性,在制作混合监督数据集的输入图像时,不同的变换操作是以一定概率进行的。输入图像包含以下可能:同时含有边缘模糊的病灶和伪病灶的图像、只含有边缘模糊的病灶但不含有伪病灶的图像、保留边缘但含有伪病灶的图像以及原图。模型能够同时识别这些不同变换的输入图像并且完成映射,则可以认为具备了较好的泛化性能。

为了进一步加强与下游分割任务的关联性,本文利用分割标签中前景区域的轮廓信息,将前景区域直接映射为一个较高的灰度值,而背景区域则以一定比例映射降低亮度。这一策略可以很好保留原始超声图像中的细节信息,同时增强病灶边缘的显

著性。将该图像作为边缘恢复任务的标签,使得边缘恢复任务与分割有较高相似性,便于迁移微调。

1.3 双通道反馈 U-Net

虽然DF-U-Net最终目的是进行分割,但是为了得到泛化性能较强的预训练模型,会利用该网络进行基于混合监督的边缘恢复任务,然后再迁移至下游分割任务进行训练。

如图3所示为本文提出的DF-U-Net,该算法源于经典的U-Net模型,以此为基础架构引入了循环机制。模型的核心思想在于对反馈输出的结果再次编码学习,和超声通道共同构成双通道编码器,然后融合两个通道的特征进行解码输出。编码器阶段有超声和概率两个通道,分别输入样本的超声图像和反馈的概率图。在边缘恢复任务中,超声通道 U 负

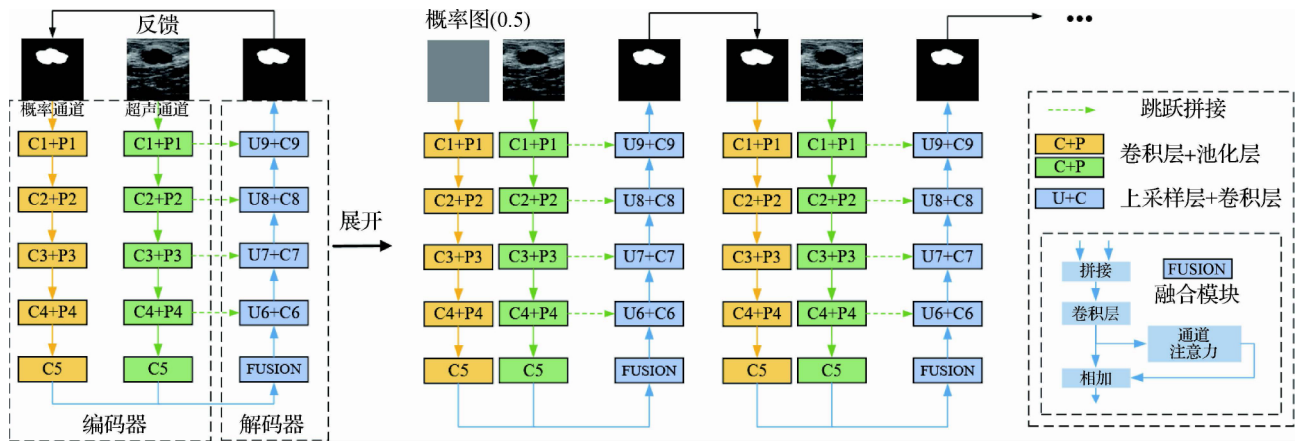


图3 双通道反馈 U-Net(DF-U-Net)流程图

Fig. 3 Flowchart of dual-channel feedback U-Net

责理解变换后图像中存在的散斑噪声和组织纹理,初步分辨边缘模糊的病灶和伪病灶。概率通道 P 根据上一个循环输出的结果,进一步消除伪病灶,并细化前背景区域不同的映射关系。而在病灶分割任务中,超声通道 U 负责学习原图的低维细节信息和高层语义信息,以及对病灶大小、位置和形态有一个初步认识。概率通道 P 可以学习上一个循环输出的概率图,尤其是病灶边缘和伪影区域的数值大小,辅助网络进一步分辨伪影和细化模糊的边缘形态。

编码器部分双通道的每一层,均由若干卷积和一个池化层构成(在图3中用黄色矩形和绿色矩形表示)。每经过一层,图像尺寸减为原来的 $1/2$,通道数变为原来的 2 倍。超声通道 U 和概率通道 P 编码输出的特征侧重点不同,是相互辅助的关系。所以需要有一个模块来融合两部分特征。如图3所示,本文简单设计了一个融合模块,在其中使用注意力机制,使网络自主学习两组特征的权重。双通道融合之后是解码阶段,每层都由一个上采样和若干卷积层构成,在图3中用蓝色矩形表示,网络要在这里完成编码器部分超声通道 U 同尺寸特征的跳跃。每经过一层,图像边长变为原来的 2 倍,通道数减为原来的 $1/2$ 。

如图3所示,DF-U-Net 在展开后的第1次循环,将每个像素数值为 0.5 的概率图 $p_{0.5}^1$ 作为概率通道 P 的输入,经过网络运算得到 p^2 。以此类推,将每个循环输出的概率图,送入下一次循环,同时计算一次损失,达到深监督的效果(Lee 等 2014)。整

个流程可表示为

$$p^{k+1} = D[U(i) P(p^k)] \quad (1)$$

式中, $U(\cdot)$ 表示超声通道编码, $P(\cdot)$ 表示概率通道编码, $D[\cdot]$ 表示将融合的特征解码, i 表示原图输入, p^k 表示第 k 次反馈的概率图 p 。

算法不断精细边缘恢复任务和病灶分割任务的结果,自动辅助模型理解病灶区域在整幅图像中的位置,解决边缘模糊和超声伪影的问题。由于所有循环可以理解成一个大网络,后面的每个循环的梯度回传会慢慢叠加,同样辅助训练之前的循环。每次循环都计算了损失,实现了深监督,改善了梯度消失(Lee 等 2014),而深监督的设计普遍被认为对于分割模型是有效的(Dou 等 2016; Zhao 等 2018)。该设计使算法在操作便捷性和参数利用率上,都优于类似 Auto-U-Net 的分步迭代算法。

高度对称的 U-Net 模型也已经被包括在所提出的结构中,即编码器的超声通道 U 和解码器。不同于一般反馈机制的运用,本文独立设置了概率通道 P ,不改变超声通道 U 和解码器之间原本简单高效的信息流通方式,保证了模型的鲁棒性。概率通道 P 独立对概率图进行学习之后,通过特征融合,对原 U-Net 进行补充更新。这种设计也有利于本文在混合监督学习过程中进一步使用迁移学习,即通过加载 VGG(visual geometry group) 19 基于 ImageNet 数据集的预训练模型,将可适配的预训练权重作为所提出算法编码器部分两个通道的初始化权重。

1.4 上下游模型训练

如图1所示,混合监督辅助任务的训练过程是

通过一个编解码模型,学习输入灰度图与标签灰度图之间的映射关系。训练开始前,首先将大型有标签自然图像数据集 ImageNet 的预训练权重加载到编解码结构的编码器部分,使得模型在初始阶段具有较强的提取特征的能力,也让整个混合监督学习过程包含一个监督学习的起点。此外,同一般的图像重建任务一样,本文采用了 L2 损失函数,通过计算反向梯度优化参数,指导边缘恢复图像的训练过程。最后,将训练得到的权重参数作为预训练模型,迁移到下游分割任务上作为初始化参数微调训练。

在对下游分割任务进行训练时,由于乳腺超声数据集的样本中,有一部分图像病灶区域较小,存在前景区域面积不均衡的情况。不同于多数分割问题中仅采用交叉熵作为损失函数,需要同时考虑训练的稳定性与类不均衡问题。所以,本文采用交叉熵损失函数与 DICE 损失函数的加权和作为损失函数。

二元交叉熵损失(binary cross entropy)对每个像素和类别关注度相同,稳定性好,也容易受到类不平衡问题的影响。其定义为

$$Loss_{bce} = - \sum y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i) \quad (2)$$

式中, y_i 表示第*i*个像素的标签值, \hat{y}_i 表示第*i*个像素的预测值。

Dice 是评价两块区域的重叠程度的指标,Dice 损失函数在训练过程中更关注前景区域,针对性地解决前景区域小的问题,但训练并不稳定。其定义为

$$Loss_{dice} = 1 - [2(y \cap \hat{y}) + \varepsilon] / (y + \hat{y} + \varepsilon) \quad (3)$$

式中, y 表示标签矩阵, \hat{y} 表示预测矩阵, ε 表示一个避免被零除的数字模糊常量。

由于引入了深监督,算法对每个循环的输出都要计算损失。将以上两种损失函数加权,联合*N*次循环得到最终损失函数为

$$Loss = \sum_{t=1}^N \beta^{N-t} [\alpha Loss_{bce} + (1 - \alpha) Loss_{dice}] \quad (4)$$

式中, α 决定了两种损失函数的系数,而 β 的值决定了每一次循环输出结果的权重。在这项工作中,本文针对乳腺超声数据集设置 α 为0.5, β 为0.5。

上下游模型均采用优化方法 Adam 更新训练过

程中的权重参数。

2 实验和结果

2.1 实验数据与预处理

本文在 Breast Ultrasound Dataset B 和 Dataset BUSI(breast ultrasound images)两个乳腺超声分割的公开数据集上进行实验,以验证所提算法的有效性。Dataset B 数据集由来自不同女患者的 163 幅超声影像组成,平均图像大小为 760 × 570 像素,其中 53 幅是恶性病变,110 幅是良性病变(Yap 等,2018)。Dataset BUSI 数据集共收集了 780 幅 20 ~ 75 岁女性的乳腺超声影像,平均图像大小为 500 × 500 像素。图像分为正常、良性和恶性 3 类,其中有 133 幅不含病灶的正常图像,437 幅良性病变和 210 幅恶性病变,良性样本和恶性样本中可能含有两个病灶(Al-Dhabhani 等,2020)。两个数据集中所有样本都有手工标注的分割掩码。为了测试算法在小样本情况下的表现,本文使用 Dataset BUSI 数据集时,仅从该数据集良性和恶性类别中随机选取 125 个样本进行实验。

为了进行数据增广,本文以 50% 概率进行水平翻转、放大裁剪以及缩小补零等操作。同时对训练集和测试集进行尺寸统一为 384 × 384 像素,并统一转成灰度图。最后,对灰度图做 Z-score 标准化预处理。

2.2 实验设计

为了验证本文所提出的 DF-U-Net 算法的有效性,本文选择了如下 5 个网络结构或者监督方式与 DF-U-Net 相关,并且是基于 U-Net 的图像分割算法进行对比:

1) U-Net(Ronneberger 等,2015):对称的编解码结构,内部的跳跃拼接补充了下采样过程中损失的高分辨率特征,同时改善训练中的梯度消失问题。该结构在医学图像分割领域广泛应用。

2) Attention U-Net(Oktay 等,2018):在 U-Net 的跳跃拼接上加入了注意力门机制,使得拼接的信息更集中于关键区域。

3) Nested U-Net(Zhou 等,2018):内嵌多解码输出的 U-Net 变体。为加强特征复用,在内部引入了稠密连接。

4) Auto-U-Net(Salehi 等,2017):迭代训练多个 U-Net 模型,将一个模型输出的概率图 and 原图组成新的输入,送入后一个模型进行训练,直到两次预测

结果之差小于阈值。

5) SS-U-Net (Chen 等, 2019): 自监督 U-Net (self-supervised U-Net) 是一种基于图像上下文恢复的自监督学习方法, 通过编解码结构训练图像恢复模型, 然后将其迁移至下游任务。

此外, 本文还对 DF-U-Net 算法本身进行了消融实验, 分别验证混合监督学习和双通道反馈 U-Net 结构的有效性。以下是 2 个消融实验的对比算法:

1) HS-U-Net: 混合监督 U-Net (hybrid supervised U-Net), 在 U-Net 上使用本文提出的混合监督学习的算法框架;

2) DF-U-Net: 本文提出的双通道反馈 U-Net。

本文使用 5 折交叉验证评估所提出算法的性能。分别对两个数据集的样本做 5 折划分, 并且在划分数据时考虑良恶性样本不平衡的问题, 尽可能使它们在每一折分布均匀。

本文采用敏感度 (sensitivity)、Dice 和平均对称表面距离 (average symmetric surface distance, ASSD) 作为分割结果的评价指标。其中, 敏感度表示实际前景像素中预测为前景区域的比例; Dice 表示预测前景区域和真实前景区域的重叠程度; 平均对称表面距离表示前景区域和真实前景区域的形状相似程度。所有评价指标的结果将以均值 \pm 方差的形式

呈现。

2.3 网络实现细节

由于参与对比的网络框架均为编解码结构, 同时也为了对比公平, 所有对比的算法均加载深度学习算法 VGG19 基于 ImageNet 数据集的预训练模型。去除全连接层, 将剩下的预训练权重作为所提出算法编码器部分的初始化权重, 这部分权重将在训练阶段微调。而在解码阶段, 网络与 U-Net 的解码阶段相同。

本文使用深度学习 Keras 框架构建模型并完成实验。批量大小设置为 1, epoch 设置为 15, 优化器设置为 Adam, 学习率 lr 设置为 $1E-5$ 。另外, 对网络层数和其余各项超参数的设置也尽可能相同。

2.4 结果定性分析

如图 4 所示, 第 1 和 2 行是不同算法在 Dataset B 上的分割结果, 第 3 和 4 行是不同算法在 Dataset BUSI 上的分割结果。由于散斑噪声和超声伪影等问题, Dataset B 中各个算法预测结果可能会出现不止一个前景连通区域。相比之下, 本文提出的 HS-DF-U-Net 算法较好地改善了这个问题, 模型基本能在影像复杂的背景中找到唯一的病灶区域。此外, HSDF-U-Net 在预测较大的、较为复杂的病灶时, 边缘更加精细, 对于不平滑的边缘也有不错的理解能力。

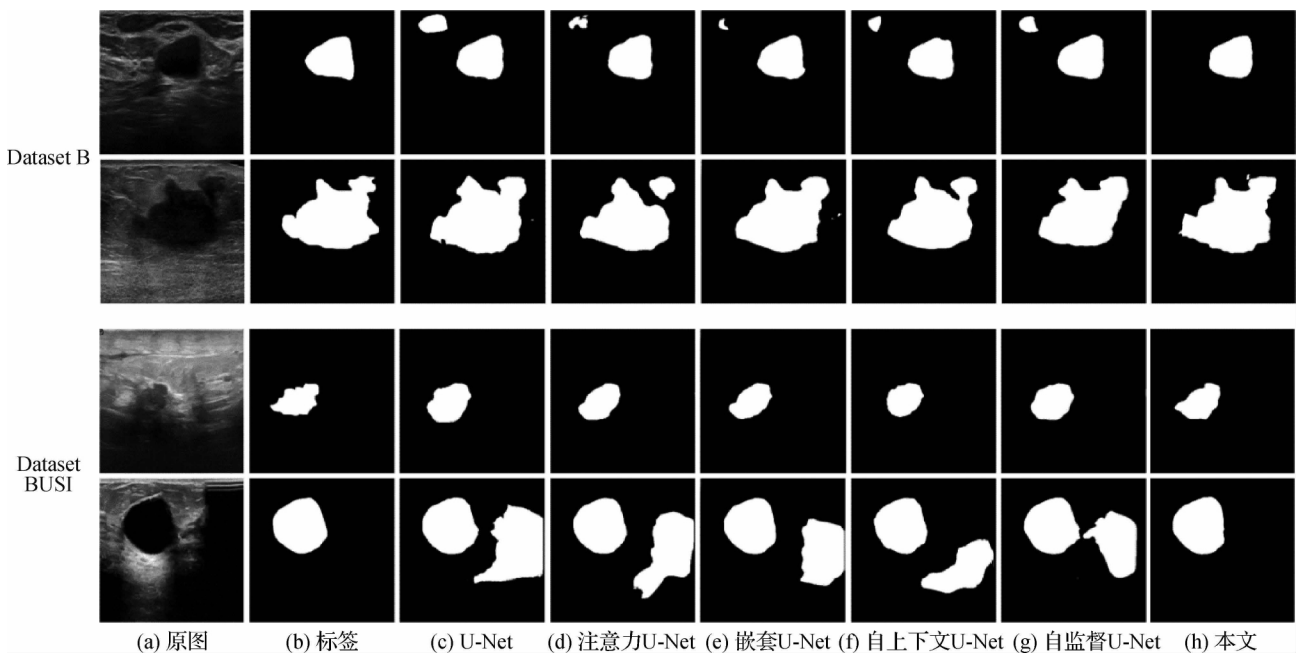


图 4 不同算法在 Dataset B 和 Dataset BUSI 上的分割结果

Fig. 4 The segmentation results of different algorithms in Dataset B and Dataset BUSI

((a) original images; (b) label; (c) U-Net; (d) Attention U-Net; (e) Nested U-Net; (f) Auto-U-Net; (g) SS-U-Net; (h) ours)

而 Dataset BUSI 中样本的噪声严重,病灶边缘非常模糊,总体相对于第 1 个数据集分割难度更大。如图 4 第 3 和 4 行所示,本文提出的 HSDF-U-Net 算法对图像上病灶面积大小的把握更加接近医生手工分割的结果。此外,HSDF-U-Net 在分割边缘较复杂的恶性病灶时,相对于其他算法在边缘形状处理上的精细度更好。综合来说,分割性能相对于之前的经典算法框架取得了一定进步。

2.5 结果定量分析

表 2 所示为不同算法在 Dataset B 上的指标结果。相较于其他算法,HSDF-U-Net 算法在 3 项指标中均取得了最好结果。与同类主流分割算法及自监督学习方法相比,HSDF-U-Net 敏感度至少提升了 0.036 5,Dice 至少提升了 0.035 8,ASSD 至少下降了 5.33。3 项指标均取得了明显提升。

表 2 不同医学图像分割算法在 Dataset B 上的结果
Table 2 Results of different medical image segmentation algorithms on Dataset B

方法	sensitivity	Dice	ASSD
U-Net	0.797 9 ± 0.040 4	0.777 9 ± 0.031 3	11.15 ± 5.11
Attention U-Net	0.779 8 ± 0.045 2	0.761 9 ± 0.028 7	13.60 ± 4.86
Nested U-Net	0.795 7 ± 0.051 0	0.780 3 ± 0.035 3	13.28 ± 8.41
Auto-U-Net	0.811 5 ± 0.040 0	0.790 3 ± 0.029 1	11.14 ± 5.64
SS-U-Net	0.804 9 ± 0.049 1	0.773 8 ± 0.040 0	13.01 ± 6.51
HSDF-U-Net	0.848 0 ± 0.044 1	0.826 1 ± 0.029 8	5.81 ± 2.20

注:加粗字体为每列最优值。

如表 3 所示,HSDF-U-Net 在 3 项指标上均取得最好的结果,而采用双通道反馈机制算法的敏感度和 Dice 仅次于 HSDF-U-Net,取得了第 2 好的结果,且相对于 U-Net 提升显著(敏感度提升了 0.041 3,Dice 提升了 0.042 3),表明了此反馈结构的有效

表 3 Dataset B 上的消融实验结果

Table 3 Ablation experiment results on Dataset B

方法	sensitivity	Dice	ASSD
U-Net	0.797 9 ± 0.040 4	0.777 9 ± 0.031 3	11.15 ± 5.11
SS-U-Net	0.804 9 ± 0.049 1	0.773 8 ± 0.040 0	13.01 ± 6.51
HS-U-Net	0.827 3 ± 0.036 8	0.793 3 ± 0.031 4	9.70 ± 4.42
DF-U-Net	0.839 2 ± 0.036 0	0.820 2 ± 0.029 3	10.11 ± 4.70
HSDF-U-Net	0.848 0 ± 0.044 1	0.826 1 ± 0.029 8	5.81 ± 2.20

注:加粗字体为每列最优值。

性。此外,使用混合监督学习范式的 U-Net 在 ASSD 上也取得了第 2 好的结果,相比使用典型的自监督学习 SS-U-Net 算法,ASSD 缩小了 3.31,证明了所提出的混合监督学习框架的有效性。

表 4 给出了不同算法在 Dataset BUSI 上的指标结果。HSDF-U-Net 算法敏感度为 0.803 9 ± 0.044 0,Dice 为 0.803 1 ± 0.036 8,ASSD 为 6.44 ± 2.80。与 U-Net 相比,该算法的敏感度提升了 0.010 5。在医学图像分割领域更重要的后两项指标上,提升则更为显著。与其他主流医学图像分割算法相比,HSDF-U-Net 的 Dice 至少提升了 0.045 1,ASSD 的均值至少下降了 13.61。

表 4 不同医学图像分割算法在 Dataset BUSI 上的结果

Table 4 Results of different medical image segmentation algorithms on Dataset BUSI

方法	sensitivity	Dice	ASSD
U-Net	0.793 4 ± 0.025 7	0.755 3 ± 0.020 3	20.05 ± 7.59
Attention U-Net	0.772 8 ± 0.031 0	0.756 6 ± 0.024 6	20.20 ± 9.50
Nested U-Net	0.788 4 ± 0.022 2	0.758 0 ± 0.026 7	21.70 ± 10.16
Auto-U-Net	0.787 3 ± 0.038 9	0.753 7 ± 0.037 5	21.10 ± 8.97
SS-U-Net	0.798 9 ± 0.029 7	0.742 1 ± 0.029 9	19.40 ± 5.55
HSDF-U-Net	0.803 9 ± 0.044 0	0.803 1 ± 0.036 8	6.44 ± 2.80

注:加粗字体为每列最优值。

如表 5 所示,HSDF-U-Net 在医学图像分割领域较重要的 Dice 和 ASSD 两项指标上,均为最优。DF-U-Net 的这两项指标为次优,该算法在敏感度上的提升最为明显,相比 U-Net 提升了 0.043 7,达到了最优结果。而相比使用 ImageNet 预训练模型的 U-Net 和使用自监督学习的 SS-U-Net,使用混合监督的 HS-U-Net 虽然在更重要的 Dice 和 ASSD 两项指标上表现更优,但是提升幅度相比 DF-U-Net 的提升幅度并不显著。因此,在该数据集上,网络结构

表 5 Dataset BUSI 上的消融实验结果

Table 5 Ablation experiment results on Dataset BUSI

方法	sensitivity	Dice	ASSD
U-Net	0.793 4 ± 0.025 7	0.755 3 ± 0.020 3	20.05 ± 7.59
SS-U-Net	0.798 9 ± 0.029 7	0.742 1 ± 0.029 9	19.40 ± 5.55
HS-U-Net	0.791 4 ± 0.035 0	0.756 9 ± 0.022 9	19.12 ± 11.16
DF-U-Net	0.837 1 ± 0.037 6	0.785 5 ± 0.043 6	9.58 ± 4.27
HSDF-U-Net	0.803 9 ± 0.044 0	0.803 1 ± 0.036 8	6.44 ± 2.80

注:加粗字体为每列最优值。

的改进对算法性能提升的贡献更大。

此外, 结合表 2 和表 4 的各项对比算法的结果, 从网络内部改进拓扑结构的主流变体算法, 如 Attention U-Net 和 Nested U-Net, 相对于原始 U-Net 在乳腺超声分割问题上没有明显优势, 反而可能在指标上低于 U-Net 的结果。而需要多次迭代的 Auto-U-Net 算法相比 U-Net 在第 1 个数据集上有一定提升, 在第 2 个数据集上的指标反而有所下降。由此可见, 多模型迭代的复杂步骤和高资源消耗, 并未带来相应的好结果。虽然 Auto-U-Net 本身继承了 U-Net 的高鲁棒性和泛化性能, 但是直接将前一模型输出的预测结果输入新的模型, 可能导致新的模型陷入局部最小值, 只能输出与前一结果相近的值, 总体来说反而制约了后面模型的训练。而本文提出的双通道反馈结构, 则利用深监督避免了此问题。

此外, 结合表 3 和表 5 的各项对比算法结果, 使用自监督学习的 SS-U-Net 算法未能取得理想的提升, 只能实现与加载 ImageNet 预训练模型的 U-Net 较为相近的性能。这在一定程度上表明, 目前在医学图像领域, 通过自监督学习方法得到的预训练模型, 还无法超越监督学习得到的预训练模型。而本文在监督方式上进行探索, 提出的混合监督学习框架, 相比前两种方法取得了更好的综合性能。

2.6 模型空间占用与计算消耗分析

为了进一步评估本文算法的执行性能, 还对比了参数量 (Params) 和浮点运算数 (floating point operations, FLOPs)。前者是衡量算法模型大小的指标, 后者则表示完成一次预测的计算量, 是衡量算法计算复杂度的指标。表 6 为不同模型在预测时的参数量以及完成一幅图像预测所需的浮点运算数。其中, 原始 U-Net 的模型最小, 计算复杂度最低, 因此在网络结构上对 U-Net 进行改进的算法, 如 Attention U-Net 和 Nested U-Net, 会在计算消耗上有不同程度的提高。而级联模型 Auto-U-Net 的计算消耗会随级联 U-Net 数量的增加而成倍增加。相比之下, 虽然使用自监督方法或者本文提出的混合监督方法需要两阶段的训练, 但是预测时资源消耗直接取决于模型本身, 实时性不受影响。此外, 使用双通道反馈机制的网络虽然计算复杂度最高, 但是其端到端网络在训练和测试的流程较 Auto-U-Net 更简易, 且参数利用率较好。

表 6 不同医学图像分割算法的模型大小与计算复杂度

Table 6 Model size and computational complexity of different medical image segmentation algorithms

方法	Params/M	FLOPs/G
U-Net	31.16	278.66
Attention U-Net	31.95	281.99
Nested U-Net	37.40	713.69
Auto-U-Net	93.48	836.49
SS-U-Net	31.16	278.66
HSDF-U-Net	62.72	966.36

3 结 论

针对乳腺超声影像中病灶难以分割的问题, 本文提出了一种混合监督双通道反馈 U-Net 算法。该算法创新性地结合分割标签的信息制作自监督数据集的输入和标签, 指导模型在混合监督学习中更加关注对病灶边缘的理解, 以此来增强辅助任务和下游分割任务的关联性。在网络结构设计方面, 本文将循环机制引入经典的 U-Net 网络, 反馈输出的概率图和原 U-Net 网路的编码器共同组成双通道编码器, 输出结果的精度能够随反馈次数增加而提升。该结构以原 U-Net 为基础模块, 参数利用高效、鲁棒性好。本文在两个公开数据集上进行了实验, 验证了混合监督的学习框架和双通道反馈结构的有效性。实验结果表明, 与其他同类的基于 U-Net 代表性分割算法相比, 本文算法具有更强的泛化性能, 且不论在定性视觉结果还是定量评价指标上均取得一定程度的提升。

本文算法仍需要进一步研究。在混合监督学习和双通道反馈结构的消融实验中, 两者结合的方法在部分指标上提升不明显, 或并未取得最优结果。这是由于从辅助任务中得到的预训练模型迁移至下游任务的过程中含有不确定性。在本文的两个任务中, 使用双通道反馈机制分别能取得不错的结果, 但是循环结构会使得网络参数在不同任务的优化过程中差异性加大。未来的工作将探索如何利用损失函数来消除这种差异性。

参考文献 (References)

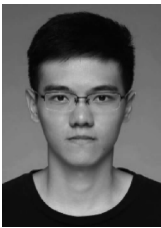
Al-Dhabyani W, Gomaa M, Khaled H and Fahmy A. 2020. Dataset of

- breast ultrasound images. *Data in Brief*, 28: #104863 [DOI: 10.1016/j.dib.2019.104863]
- Bian Z J, Qin W J, Liu J R and Zhao D Z. 2018. Review of anatomic segmentation methods in thoracic CT images. *Journal of Image and Graphics*, 23(10): 1450-1471 (边子健, 覃文军, 刘积仁, 赵大哲. 2018. 肺部 CT 图像中的解剖结构分割方法综述. *中国图象图形学报*, 23(10): 1450-1471) [DOI: 10.11834/jig.180067]
- Chen L, Bentley P, Mori K, Misawa K, Fujiwara M and Rueckert D. 2019. Self-supervised learning for medical image analysis using image context restoration. *Medical Image Analysis*, 58: #101539 [DOI: 10.1016/j.media.2019.101539]
- Cheplygina V, de Bruijne M and Pluim J P W. 2019. Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. *Medical Image Analysis*, 54: 280-296 [DOI: 10.1016/j.media.2019.03.009]
- Dou Q, Chen H, Jin Y M, Yu L Q, Qin J and Heng P A. 2016. 3D deeply supervised network for automatic liver segmentation from CT volumes//*Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece: Springer: 149-157 [DOI: 10.1007/978-3-319-46723-8_18]
- Hesamian M H, Jia W J, He X J and Kennedy P. 2019. Deep learning techniques for medical image segmentation: achievements and challenges. *Journal of Digital Imaging*, 32(4): 582-596 [DOI: 10.1007/s10278-019-00227-x]
- Huang Q H, Luo Y Z and Zhang Q Z. 2017. Breast ultrasound image segmentation: a survey. *International Journal of Computer Assisted Radiology and Surgery*, 12(3): 493-507 [DOI: 10.1007/s11548-016-4513-1]
- 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]
- Jing L L and Tian Y L. 2020. Self-supervised visual feature learning with deep neural networks: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1-1 [DOI: 10.1109/tpami.2020.2992393]
- Kim W H, Moon W K, Kim S J, Yi A, Yun B L, Cho N, Chang J M, Koo H R, Kim M Y, Bae M S, Lee S H, Kim J Y and Lee E H. 2013. Ultrasonographic assessment of breast density. *Breast Cancer Research and Treatment*, 138(3): 851-859 [DOI: 10.1007/s10549-013-2506-1]
- Kolesnikov A, Zhai X H and Beyer L. 2019. Revisiting self-supervised visual representation learning//*Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, USA: IEEE: 1920-1929 [DOI: 10.1109/cvpr.2019.00202]
- Lee C Y, Xie S N, Gallagher P, Zhang Z Y and Tu Z W. 2014. Deeply-supervised nets[EB/OL]. [2020-05-25]. <https://arxiv.org/pdf/1409.5185.pdf>
- Li Z, Yang J L, Liu Z, Yang X M, Jeon G and Wu W. 2019. Feedback network for image super-resolution//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, USA: IEEE: 3867-3876 [DOI: 10.1109/CVPR.2019.00399]
- 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]
- Noble J A and Boukerroui D. 2006. Ultrasound image segmentation: a survey. *IEEE Transactions on Medical Imaging*, 25(8): 987-1010 [DOI: 10.1109/tmi.2006.877092]
- Oktay O, Schlemper J, Le Folgoc L, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla N Y, Kainz B, Glocker B and Rueckert D. 2018. Attention U-net: learning where to look for the pancreas[EB/OL]. [2020-05-25]. <https://arxiv.org/abs/1804.03999.pdf>
- 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]
- Sahiner B, Chan H P, Roubidoux M A, Hadjiiski L M, Helvie M A, Paramagul C, Bailey J, Nees V A and Blane C. 2007. Malignant and benign breast masses on 3D US volumetric images: effect of computer-aided diagnosis on radiologist accuracy. *Radiology*, 242(3): 716-724 [DOI: 10.1148/radiol.2423051464]
- Salehi S S M, Erdogmus D and Gholipour A. 2017. Auto-context convolutional neural network (Auto-Net) for brain extraction in magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 36(11): 2319-2330 [DOI: 10.1109/tmi.2017.2721362]
- Tan C Q, Sun F C, Kong T, Zhang W C, Yang C and Liu C F. 2018. A survey on deep transfer learning//*Proceedings of the 27th International Conference on Artificial Neural Networks and Machine Learning*. Rhodes, USA: Springer: 270-279 [DOI: 10.1007/978-3-030-01424-7_27]
- Yap M H, Pons G, Martí J, Ganau S, Sentís M, Zwiggelaar R, Davison A K and Martí R. 2018. Automated breast ultrasound lesions detection using convolutional neural networks. *IEEE Journal of Biomedical and Health Informatics*, 22(4): 1218-1226 [DOI: 10.1109/jbhi.2017.2731873]
- 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]
- 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 International Workshop on Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Granada, Spain: Springer: 3-11 [DOI: 10.1007/978-3-030-00889-5_1]

Zhou Z W, Sodha V, Rahman Siddiquee M M, 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, China: Springer: 384-393 [DOI: 10.1007/978-3-030-32251-9_42]

作者简介



贡荣麟, 1996 年生, 男, 硕士研究生, 主要研究方向为深度学习、医学图像分析与处理。

E-mail: gongronglin@shu.edu.cn



施俊, 通信作者, 男, 教授, 主要研究方向为医学图像分析与处理、模式识别。

E-mail: junshi@shu.edu.cn

王骏, 男, 副教授, 主要研究方向为机器学习、医学图像分析与处理。E-mail: wangjun_shu@shu.edu.cn