

doi: 10.3978/j.issn.1000-4432.2021.01.10

View this article at: <http://dx.doi.org/10.3978/j.issn.1000-4432.2021.01.10>

结合多尺度特征融合的扩张残差U-Net分割网络在视网膜自动分层中的应用

梁姗姗¹, 曾鸿为¹, 何杰¹, 张军¹, 袁进²

(1. 中山大学电子与信息工程学院, 广州 510006; 2. 中山大学中山眼科中心, 眼科学国家重点实验室, 广州 510000)

[摘要] **目的:** 对视网膜光学相干断层扫描图像中不同层和积液区域的分割。**方法:** 提出一种基于深度学习的轻量级的神经网络, 参考DRUNet体系、膨胀卷积和残差网络的架构, 通过连接不同深度网络处得到的上采样输出, 进行多尺度特征融合, 使网络能够更好地识别出图像中的边界信息。**结果:** 改进型DRUNet显著提升了视网膜分层的效果, 准确率较U-Net提高了1.25%, 同时能提前1~2次迭代达到传统U-Net的准确度。**结论:** 本文采用的网络结构提高了对视网膜光学相干断层扫描图像的分割性能, 同时降低了网络参数, 具有强大的应用潜力。

[关键词] 扩展残差U-Net; 眼底分割; 光学相干层析成像

Multie-scale hierarchical feature extraction combined with dilated-residual U-Net for retina automatic segmentation

LIANG Shanshan¹, ZENG Hongwei¹, HE Jie¹, ZHANG Jun¹, YUAN Jin²

(1. School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006;

2. State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou 510000, China)

Abstract **Objective:** To achieve the segmentation of different layers and fluid areas on the optical coherence tomography (OCT) image of the retina. **Methods:** A lightweight neural network based on deep learning was proposed. The network structure adopted in this study was designed based on the architecture of dilated-residual U-Net. By connecting the upsampling output obtained at different depth networks, multi-scale feature fusion was performed to enable the system to accurately identify the boundaries on the OCT image. **Results:** Compared with U-Net, this algorithm could achieve the same accuracy with 1–2 epochs less, and the accuracy was also improved by 1.25%. **Conclusion:** The proposed network improves the segmentation performance of retinal OCT images, and reduces the number of parameters, which demonstrates the network has great application potential.

Keywords dilated-residual U-Net; retina segmentation; optical coherence tomography

收稿日期 (Date of reception): 2020-08-17

通信作者 (Corresponding author): 梁姗姗, Email: liangshsh@mail.sysu.edu.cn

基金项目 (Foundation item): 国家重点研发计划 (2017YFC0112401); 国家自然科学基金 (61505267; 61975246)。This work was supported by the National Key R&D Program of China (2017YFC0112401), National Natural Science Foundation of China (61505267, 61975246).

视网膜是眼睛最重要的部分之一, 其病变会导致视力下降甚至失明。多数视网膜病变会导致视网膜形态学改变, 因此观测和分析其形态变化对临床诊断具有重要意义^[1-2]。视网膜是紧贴在脉络膜内面的薄膜。它由视觉细胞构成, 有感受光刺激的作用, 视信息在视网膜上形成视觉神经冲动, 沿视路将视信息传递到视中枢形成视觉, 在头脑中建立起图像。视网膜在组织学上共分为10层: 神经纤维层(nerve fiber layer, NFL)、神经节细胞层(ganglion cell layer, GCL)、内丛状层(inner plexiform layer, IPL)、内颗粒层(inner nuclear layer, INL)、外丛状层(outer plexiform layer, OPL)、外颗粒层(outer nuclear layer, ONL)、内节段髓(inner segmental medulla, ISM)、内节段椭圆体(inner segment ellipsoid, ISE)、外节段(outer segment, OS)、视网膜色素上皮层(retinal pigment epithelium, RPE)^[3]。视网膜形状的变化往往就表现为这十层膜结构的形态的变化。

光学相干断层扫描技术(optical coherence tomography, OCT)是近年来发展较快的一种极具发展前途的新型层析成像技术, 由于其低损、高分辨率、非侵入性等特点, 已被广泛应用到眼科的临床诊断中, 成为继CT, MRI后又一重大技术突破^[4]。在临床分析中, 对视网膜OCT图像的分层处理一般采用5~9层的划分模式。视网膜积液往往预示视网膜发生了病变, 如糖尿病性黄斑水肿、视网膜炎等。医生可以通过分析视网膜积液区域的分布和积液块的大小来诊疗眼底病变情况, 分割视网膜中可能出现的积液区域有益于疾病诊断和治疗。传统的分割方法是靠医生或者专业人士手动分割, 这种方式耗时长、主观性强。随着计算机的普及以及图像处理算法的发展, 计算机辅助医生进行医疗诊断成为一种炙手可热的方法^[5]。

图像分割常见的方法有3类: 一种是基于灰度阈值的传统分割方法^[6-7], 它们通常只能进行简单的视网膜层分割, 分割的方法只是根据像素点的值, 缺失大量信息, 精度低; 还有基于图论的方法, 如采用稀疏重建和图搜索的方式^[8]对视网膜OCT图像中的黄斑区域进行分割, 这类方法有计算速度快、分层灵活的特点, 但同时也更易受到噪声的干扰; 另一类方法是基于机器学习的方法, 如利用随机森林^[9]、支持向量机^[10-11]和一些其他机器学习手段^[12]完成对视网膜层的分割, 这类方法有分割精度高、受噪声影响小的特点, 也有研究人员^[13]在使用机器学习方法的同时加入传统

的边界提取算法来更好地获得分割层的边界。

Devalla等^[14]结合残差结构可以提供更好的网络梯度信息流的优势和在DeepLab中提到的膨胀卷积概念, 提出了DRUNet体系架构。该网络由一条下采样路径和一条上采样路径组成, 每条路径包括一个标准卷积模块和两个残差卷积模块, 两条路径呈对称分布, 通过跳级连接实现互连。最后, 使用softmax激活函数应用于此输出层, 以获得各像素的类概率。在临床应用中, 分割的准确度和速度对医生的诊疗都有重要影响, 所以不断提高精度与速度一直是医学图像分割领域深度学习的着力点。DRUNet网络在自然图像分割任务上有很好的表现, 而本文旨在改进编码与解码的方式, 在减少参数数量的同时达到提高分割的准确率和加快处理速度的目的。

本文设计了一种基于DRUNet的改进型轻量级的适用于小样本医学图像分割任务的全卷积神经网络, 选择使用深度学习的方法来完成分割任务的原因在于视网膜OCT图像中各区域之间灰度差距小, 依靠传统阈值判决的方式难以获得较好的分割, 而卷积神经网络在提取图像高级语义和低级特征方面有着较传统算法独有的优势。本文搭建的网络是一个带有膨胀卷积^[15]和残差结构^[16]的网络, 并且在输出端采用了多尺度分层特征提取^[17]来帮助恢复视网膜各层的边界。经实验可知, 所设计的网络较普通的U-Net训练更快且对比发现分割视网膜各区域效果良好。

1 方法

1.1 网络结构

从不同下采样程度的特征图恢复到原图大小时所形成的分割结果其边界精细程度是不同的。在浅层的特征图中, 包含有更多的细节信息(局部特征); 在深层特征图中, 包含有更多的上下文信息(全局特征), 因此如果考虑将不同深浅层的特征进行多尺度融合, 得到的分割效果可能更好。本文提出的网络继承了DRUNet的结构, 并结合多尺度分层特征提取的优势, 对上采样网络进行了改进。在DRUNet中, 仅仅使用了残差模块和膨胀卷积, 残差模块可以加速训练时的梯度流动, 加快收敛; 而膨胀卷积可以在同样的数据规模中获得更大的感受野, 即考虑更多上下文的关联。但是在恢复原图尺寸的时候只是有类似U-Net架构中的跳级连接。本网络添加的多尺度特征提取, 将其

与最终上采样结果融合再获得输出的方法便于网络进一步考虑分割目标的深层抽象特征和浅层边界信息,可以避免产生积液区域的分割空洞和形成更为精确的边界。

一个输入图像首先经过Standard Block(由两个 3×3 的卷积层)和3个Residual Block(由Standard Block在输入端和输出端跨接一个 1×1 的卷积层得到)进行编码,从每个Block输出都会做一次 2×2 的最大池化,然后再通过两个Residual Block和一个Standard Block进行解码,相同尺度的Block彼此跳接,对下采样中3个模块输出的上采样采用步长分别为1, 2和4的转置卷积,最后在深度方向拼接后送入输出层(图1)。这样做的好处是便于网络融合多层级的特征,有利于恢复扫描图像中的边界。

对比两个网络结构发现:本网络并不如U-Net深,仅进行3次下采样。因本网络采用了膨胀卷积,使得网络不用进行很多下采样也可获得足够的感受野。另外,网络中两个较深的层采用了残差结构来加速梯度的流动,并在跳接结构的基础上加入了多尺度特征提取。再对比两个网络的规模,经统计发现U-Net总共有58 825个可训练参数,而本网络总可训练参数仅为52 649个。可见添加残差结构和多尺度分层特征提取后,网络总参数还是比U-Net少,网络更加精简。

1.2 数据集

本实验数据来自杜克大学的一项关于视网膜OCT图像分割的项目^[18]。OCT图像大小不一,但基本都被裁剪为 496×768 ,标签为 8×768 的二维数组。经对图像和标签对应进行适当裁剪,最终获得110张尺寸为 288×560 的灰度图和对应的标签,其中包括32张不含积液的样本和78张含积液的样本。本文分配90张OCT图像供训练,其余20张用于测试,训练集和测试集中均包含带积液和不带积液的样本。同时对图像进行随机水平翻转、随机平移、随机旋转、错切变换和随机缩放以扩充数据集。

1.3 参数设置

在两个采样路径中,除最后一层输出外,其余层均进行Batch Normalization^[19],并采用ELU函数^[20]激活。使用带有Nesterov动量(动量=0.9)的随机梯度下降对整个网络进行端到端训练。初始学习率为0.1(当验证损失在两个连续的时期内未能改善时学习率缩小10倍),损失函数采用输出结果与标签的交叉熵。利用Keras的数据生成器,实时不断生成进行随机操作后的数据并馈入网络进行训练,受限于内存大小,每批次仅生成并馈入2张OCT图像,进行50轮迭代。整个训练及测试过程在2 GHz双核Inter[®] Core[™] i7-3667U CPU上进行。

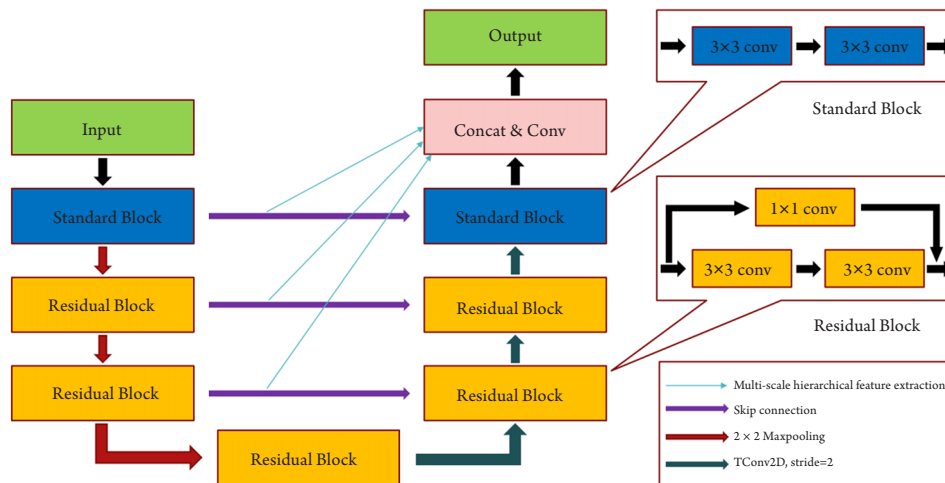


图1 改进的DRUNET网络结构,沿用了DRUNET的基础结构,同时将下采样路径中每一模块的输出分别上采样到原图像大小,对下采样中3个模块输出的上采样采用步长分别为1, 2和4的转置卷积,加入多尺度分层特征提取,在深度方向拼接后送入输出卷积层

Figure 1 The network of optimized DRUNET, which follows the infrastructure of DRUNET. Output in down-sampling path is upsampled to the same scale of original images. Deconvolution with stride of 1, 2, 4 are respectively used in the three module outputs in down-sampling path, incorporating multi-scale feature extraction, the feature maps will reach output layer after being concatenated in channels

2 结果

观察网络的训练过程发现：与U-Net的对比，本工作中所提出网络的精确率达到95%时经过10轮训练，U-Net达到相同精度需要进行30轮以上，且最终稳定情况下精确率也较U-Net高1.25%(图2)。此外，Loss下降速度也明显快于单纯的U-Net，U-Net的Loss下限也不如改进型DRUNet。实验表明本文所提出的网络更加适用于训练，学习效率更高。

接下来分别以不带有积液的层分割和带有积液的分割为例，观察两个网络的分割结果。先讨论

不带有积液的层分割：经过本文所提出的改进型DRUNet处理后的图像较U-Net得到了更好的分割，待分割的各视网膜层的整体可见性也很好(图3)。再看带有积液的分割：U-Net对视网膜各层的分割不如改进型DRUNet；在积液区域，U-Net分割得到的结果出现了很多空洞(图4)，而这在积液区域通常是不会发生的，积液常为成块存在，因此改进型DRUNet分割的结果更加贴近于真实值。

表1和表2统计了不带积液样本和带积液样本中各类别的精确率、Dice系数、交并比(IoU)，这里统计了前8个类别。

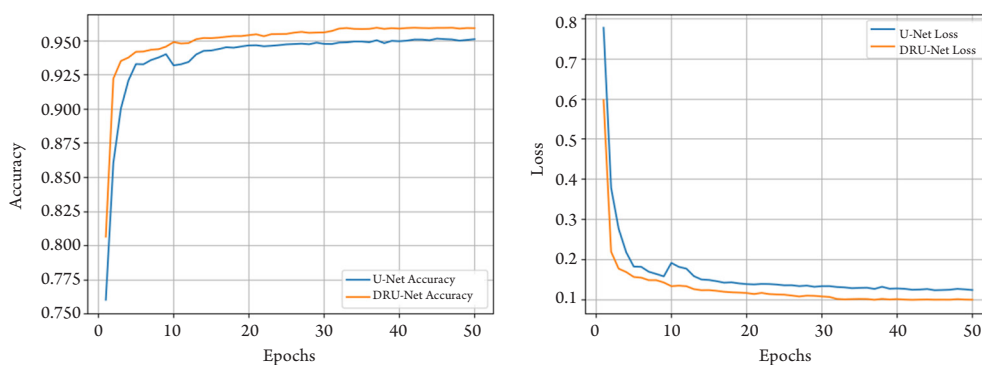


图2 改进型的DRUNet与U-Net训练过程准确度和损失曲线

Figure 2 Comparison of training accuracy and loss between DRUNet and U-Net

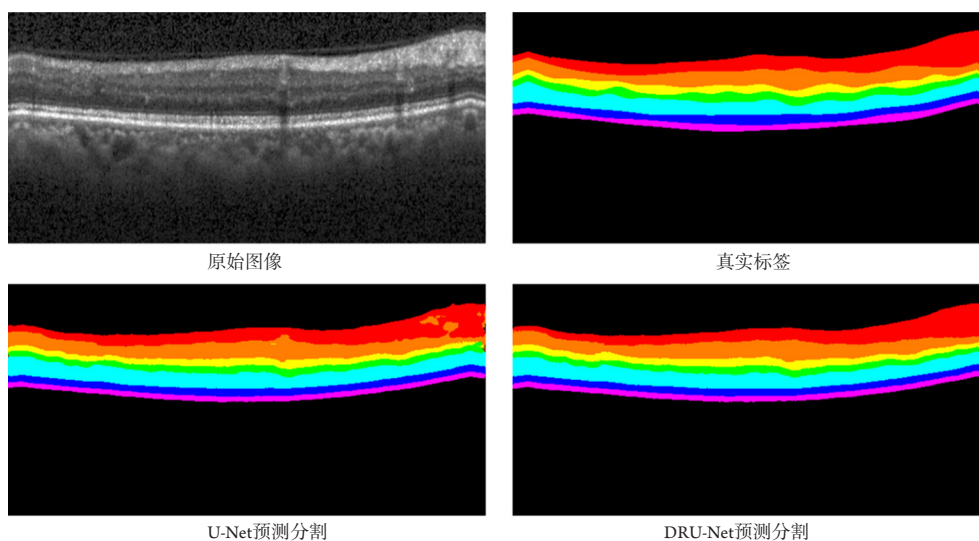


图3 不带积液的分割结果对比图，本文所提出的改进型DRUNet层次分割优于传统U-Net

Figure 3 Comparison of segmentation results on samples without effusion, the proposed optimized DRUNet outperforms original U-Net

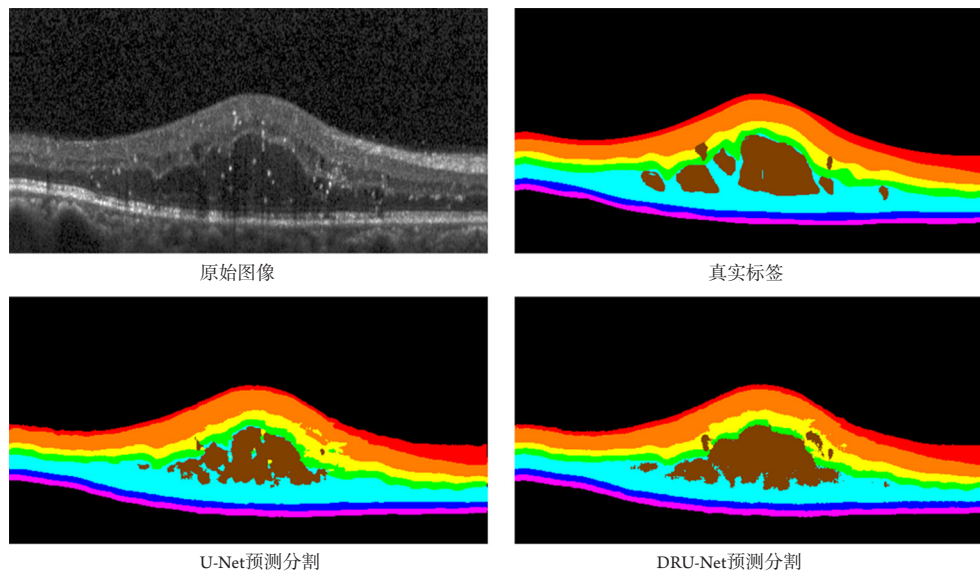


图4 带积液的分割结果对比图, 在积液区域, U-Net分割得到的结果出现了很多空洞, 而这对于积液区域通常是改进型DRUNet分割的结果更加贴于真实值

Figure 4 Comparison of segmentation results on samples with effusion, U-Net shows many cavities in effusion area, while the segmentation result of DRUNet is closer than true label

表1 不带积液样本的分割结果统计

Table 1 Comparison of segmentation results on samples without effusion

组别	U-Net			DRUNet		
	精确率	Dice系数	IoU	精确率	Dice系数	IoU
NFL	0.9165	0.8176	0.7110	0.8853	0.8755	0.7901
GCL-IPL	0.7458	0.7675	0.6702	0.8057	0.8041	0.7294
INL	0.6238	0.6447	0.5278	0.7294	0.7370	0.6166
OPL	0.7799	0.6387	0.5205	0.7617	0.6795	0.5601
ONL-ISM	0.9014	0.9105	0.8375	0.8249	0.8469	0.7660
ISE	0.8933	0.8724	0.7752	0.8898	0.8462	0.7386
OS-RPE	0.8801	0.8412	0.7265	0.8378	0.8076	0.6809

对于一个特定的类别, 属于该类别的像素点称为正样本, 否则为负样本, 假设TP(True Positive)表示实际为正样本被模型正确预测为正的样本数; TN(True Negative)表示实际为负样本, 被模型正确预测为负的样本数; FP(False Positive)表示实际为负样本, 被模型错误预测为正的样本数; FN(False Negative)表示实际为正样本, 被模型错误预测为负的样本数。精确率定义为 $Precision = (TP+TN)/(TP+TN+FN+FP)$; Dice系数定义为 $Dice = 2 \times TP / (2 \times TP + FP + FN)$; 交并比定义为

$IOU = 1/2 \times [TP/(FP+FN+TP) + TN/(FN+FP+TN)]$ 。

NFL, GCL-IPL, INL, OPL分割情况U-Net不如改进型DRUNet, 尤其在INL和OPL这类较难的分割对象中, U-Net表现也不如改进型DRUNet(表1)。对带积液的样本进行分割, 发现除ISE和OS, RPE外, 其余各类别改进型DRUNet的性能均较U-Net好(表2)。总体来看, 带积液的样本分割时, 改进型DRUNet各指标均优于传统U-Net, 而层分割时改进型DRUNet整体性能也更优于U-Net。

表2 带积液样本分割结果统计

Table 2 Comparison of segmentation results on samples with effusion

组别	U-Net			DRUNet		
	精确率	Dice系数	IoU	精确率	Dice系数	IoU
NFL	0.8361	0.8373	0.7241	0.8720	0.8493	0.7437
GCL-IPL	0.8915	0.8818	0.7915	0.8961	0.9044	0.8270
INL	0.7470	0.6990	0.5441	0.8256	0.7583	0.6149
OPL	0.6755	0.6918	0.5322	0.6570	0.7280	0.5738
ONL-ISM	0.8433	0.8342	0.7252	0.8363	0.8477	0.7443
ISE	0.8507	0.8598	0.7576	0.8410	0.8558	0.7529
OS-RPE	0.8282	0.8393	0.7240	0.8313	0.8320	0.7159
积液	0.4953	0.5155	0.3790	0.7098	0.6258	0.4768

3 讨论

本文提出了一种改进型DRUNet结构, 它是一个带有膨胀卷积和残差结构的网络, 并且在输出端采用了多尺度分层特征提取来帮助恢复视网膜各层的边界, 通过跟U-Net的对比可以发现, 该网络处理后的图像得到很好的分割, 待分割的各视网膜层的整体可见度较好, 而U-Net中出现了层次分割不准确, 某些层中有错误的散斑分割现象。更主要的是, 除了分割精度, 本文提出的网络结构参数量小, 交叉熵损失下降速度均明显优于U-Net, 具有广泛的应用前景, 可以实现对视网膜OCT图像较好的分割, 具有强大的临床应用潜力。

但通过结果可以看到, 进行层分割时, 不同层的分割结果往往差异很大, 如何设计约束使得在保持原有较好的层分割的同时, 完成对更加模糊的层的分割, 是另一个亟待解决的问题, 同时在有积液的OCT图像中, 积液区域出现的不确定性、不同程度积液在OCT图像中表现出的差异等问题仍然不能被很好地解决, 所以这个网络仍然有一定的优化空间。且由于样本数据量较少, 模型的表现受到一定的限制, 尚缺乏足够的外部验证测试模型的泛化能力。将进一步与相关眼科医生展开合作获取更多数据, 完善模型同时推进临床实验, 开展外部测试。

参考文献

1. 吴彩云. 高度近视黄斑病变的形态学分类及其影响因素[D]. 兰州: 兰州大学, 2014.
WU Caiyun. The morphological classification of myopia maculopathy and influence factors[D]. Lanzhou: Lanzhou University, 2014.
2. 滕岩, 刘英伟, 杨明明, 等. 糖尿病视网膜病变患者全视网膜激光光凝术后黄斑区功能与形态变化[J]. 中华眼底病杂志, 2010, 26(2): 120-123.
TENG Yan, LIU Yingwei, YANG Mingming, et al. The functional and morphological changes of macular after panretinal photocoagulation in the patients with diabetic retinopathy[J]. Chinese Journal of Ocular Fundus Diseases, 2010, 26(2): 120-123.
3. Chakravarty A, Sivaswamy J. A supervised joint multi-layer segmentation framework for retinal optical coherence tomography images using conditional random field[J]. Comput Methods Programs Biomed, 2018, 165: 235-250.
4. 陈玉平. 光学相干层析成像综述[J]. 价值工程, 2014, 33(32): 255-256.
CHEN Yuping. Review on optical coherence tomography[J]. Value Engineering, 2014, 33(32): 255-256.
5. Rogers W, 祝庆麟. 计算机辅助医学诊断文献回顾[J]. 国外医学. 生物医学工程分册, 1981(4): 26-31.
ROGERS W, ZHU Qinglin. Review of computer aided medical diagnosis[J]. Journal of Biomedical Engineering Foreign Medical Sciences Biomedical Engineering, 1981(4): 26-33.

6. 蔡怀宇. 眼科光学相干层析成像的图像处理方法[J]. 中国光学, 2019, 12(4): 731-740.
CAI Huaiyu. Image processing method for ophthalmic optical coherence tomography[J]. Chinese Optics, 2019, 12(4): 731-740.
7. 贺琪欲. 基于光学相干层析成像的视网膜图像自动分层方法[J]. 光学学报, 2016, 36(10): 309-318.
HE Qiyu. Automated retinal layer segmentation based on optical coherence tomographic images[J]. Acta Optica Sinica, 2016, 36(10): 309-318.
8. Fang L, Li S, Cunefare D, et al. Segmentation based sparse reconstruction of optical coherence tomography images[J]. IEEE Trans Med Imaging, 2017, 36(2): 407-421.
9. Lang A, Carass A, Sotirchos E, et al. Segmentation of retinal OCT images using a random forest classifier[C]//Medical Imaging 2013: Image Processing. International Society for Optics and Photonics, 2013: 145-155.
10. Vermeer KA, van der Schoot J, Lemij HG, et al. Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images[J]. Biomedical Optics Express, 2011, 2(6): 1743-1756.
11. 杨云. 集成支持向量机在OCT血管内斑块分割中的应用与研究[J]. 计算机应用与软件, 2019, 36(4): 103-107.
YANG Yun. Application and research of adaboost-SVM in OCT intravascular patch segmentation[J]. Computer Applications and Software, 2019, 36(4): 103-107.
12. 许毓鹏. 视网膜血管性疾病光学相干断层成像图像的自动分层研究[J]. 上海交通大学学报: 医学版, 2019, 39(6): 613-621.
XU Yupeng. Automatic layer segmentation of optical coherence tomography images in retinal vascular diseases[J]. Journal of Shanghai Jiaotong University. Medical Science, 2019, 39(6): 613-621.
13. Lang A, Carass A, Hauser M, et al. Retinal layer segmentation of macular OCT images using boundary classification[J]. Biomedical Optics Express, 2013, 4(7): 1133-1152.
14. Devalla SK, Renukanand PK, Sreedhar BK, et al. DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images[J]. Biomedical Optics Express, 2018, 9(7): 3244-3265.
15. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[J]. arXiv preprint arXiv: 1511.07122, 2015.
16. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.
17. Liu Y, Cheng MM, Hu X, et al. Richer convolutional features for edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(8): 1939-1946.
18. Chiu SJ, Allingham MJ, Mettu PS, et al. Kernel regression-based segmentation of optical coherence tomography images with diabetic macular edema[J]. Biomedical Optics Express, 2015, 6(4): 1172-1194.
19. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on Machine Learning, 2015: 448-456.
20. Clevert DA, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (ELUs)[C]//International Conference on Learning Representations, 2016: 1-14.

本文引用: 梁姗姗, 曾鸿为, 何杰, 张军, 袁进. 结合多尺度特征融合的扩张残差U-Net分割网络在视网膜自动分层中的应用[J]. 眼科学报, 2021, 36(1): 10-16. doi: 10.3978/j.issn.1000-4432.2021.01.10

Cite this article as: LIANG Shanshan, ZENG Hongwei, HE Jie, ZHANG Jun, YUAN Jin. Multie-scale hierarchical feature extraction combined with dilated-residual U-Net for retina automatic segmentation[J]. Yan Ke Xue Bao, 2021, 36(1): 10-16. doi: 10.3978/j.issn.1000-4432.2021.01.10