

[文章编号] 1003-4684(2023)01-0038-05

基于多级残差多尺度的医学图像分割网络

李 伟, 吴 聪

(湖北工业大学计算机学院, 湖北 武汉 430068)

[摘 要] 针对目前多数 U 型网络存在编码阶段卷积核尺度单一难以提取变化较大特征以及深层网络难以训练优化的情况,提出一种新的基于多级残差和多尺度的神经网络,利用多级残差使神经网络更易学习,提高网络的深度,使它在模型不退化的情况下拥有更丰富的特征表达能力,提出了多尺度交叉融合模块,通过不同的感受尺度去提取特征,交叉融合也使得特征信息更加充分的交流和融合。网络在 CHASE_DB1 数据集上进行测试,并进行数据对比,性能表现优良,特别是 ACC 达到了 0.9744,SP 达到了 0.9876。提出的网络在增加深度的同时并不影响它的学习过程和表现效果。

[关键词] 血管分割; 深度学习; U-Net; 残差学习; 多尺度

[中图分类号] TP391 [文献标识码] A

很多疾病可通过视网膜血管的细微特征反映出来,专业医师可以根据视网膜血管的形态变化来进行病理分析和判断,制定诊疗计划。深度学习在医学图像处理领域的应用在近些年来取得了极大的进步,Ben-Cohen^[1]将全卷积神经网络(Fully Convolutional Networks, FCN)应用在分割肝脏和肿瘤的 CT 影像分割上。Dasgupta^[2]将 FCN 引入视网膜血管分割的领域,在 DRIVE 数据集上的实验证明了 FCN 的强大性能。Ronneber^[3]提出的经典的 U-Net 扩展了 FCN 使其效果更好并且仅仅需要更少的标注数据。

深度学习的发展伴随着网络深度的增加,Alex-Net^[4]仅仅只有 5 个卷积层,随后 VGG, GoogleNet, DenseNet 先后被提出,性能提升的同时结构更加复杂。神经网络的发展历程^[5-6]说明了网络的深度对于网络模型的表达能力非常重要。Wu Yan-Cheng^[7]进行的实验说明网络深度的增加可以决定网络是否可以取得良好效果。

1 相关

一定范围内,随着网络深度的增加,模型可以拟合更加复杂的函数,模型的性能也可以提升,但是在深度达到某种程度时,单纯增加网络深度,网络模型并不会得到优化。He^[8]等在 Highway 网络的基础上提出了残差网络,残差学习机制可以解决由于网络深度增加带来的退化问题,较深的网络可以更好

地训练。ZL Ni 等提出了 RAUNet^[9]用于语义分割,RAUNet 是在 U-Net 的基础上结合了残差学习机制和注意力机制,是对 U 型网络的成功改进。Zhang^[10]等人在实验的基础上提出设想:如果残差映射容易学习,那么残差映射中的残差映射更容易学习。他们在 ResNets 的基础上逐级加入 shortcut 支路,建立了 Residual network of Residual network(RoR),这就是多级残差,RoR 在 CIFAR-10, CIFAR-100 和 SVHN 等数据集上均取得了较 ResNets 更好的分类结果。LIAN 等^[11]构建了多尺度残差网络,在残差结构中,由级联的多尺度卷积层作为残差映射分支。Zheng^[12]等人在卷积的过程中引入多级残差来弥补 CNN 中缺失的特征从而提高识别精度。WU 等人提出了新颖的残差网络结构,也就是深度多级残差网络,他们在原有的残差网络结构上再加上多级捷径连接,用来挖掘残差网络的优化能力。

2 方法

单一尺度卷积核感受野固定,网络层数增加导致拟合能力退化,针对这些问题,可以将多尺度和多级残差机制相结合,前者增强特征的提取能力,后者让每一个多尺度模块更好地学习和优化,计划在 U 型网络的基础上设计一种多级残差多尺度网络(Multilevel residual Multi-Scale Net, MRMS-Net),在编码器中将多尺度机制集成在多级残差机制内。

2.1 多级残差

图 1a 显示了基本的残差结构,在普通卷积块的基础上增加了一条 shortcut,这条没有权重的越层连接成为一条从输入到输出的通路可以避免特征图丢失。

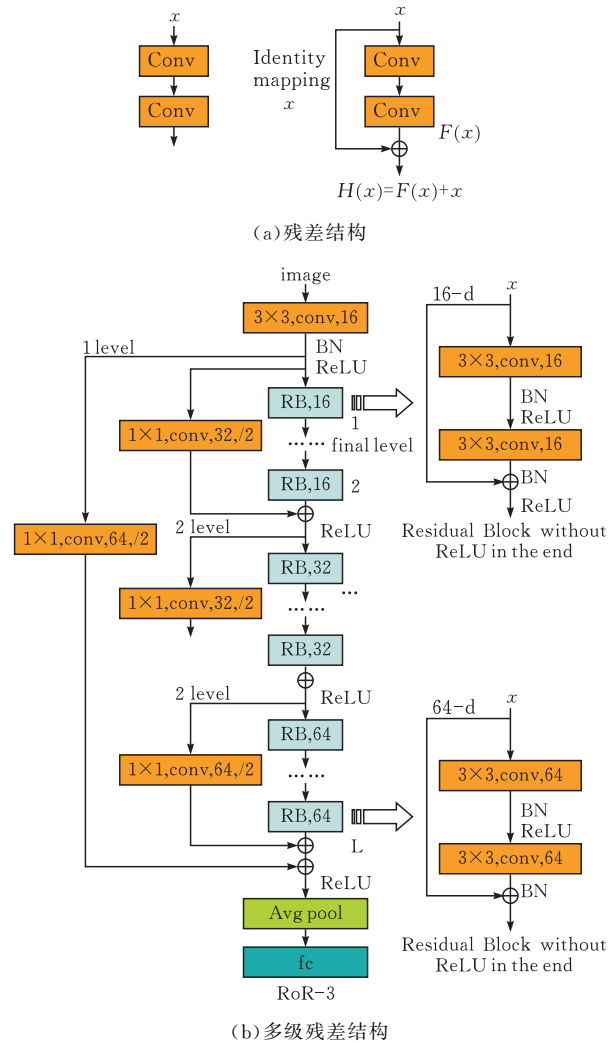


图 1 残差和多级残差说明

多级残差 RoR(Residual of Residual)是通过添加逐级快捷连接来实现对残差映射的优化,以此方式构建出的基于残差网络的 RoR,它的快捷连接较多,但等级分明。如图 1b 所示,这是一个拥有 L 个原始残差块的 RoR-3 网络,因为有 root-level shortcut,middle-level shortcut,final-level shortcut 这三级快捷连接而得名。这里存在 L 个 final-level shortcut, $L/2$ 个 middle-level short,1 个 root-level shortcut,最基本的残差块的 shortcut 是 final-level shortcut。设 m 为快捷连接级数, $m=1,2,3,\cdots$,当 $m=1$ 时,RoR 是一个基本的残差网络,当 $m=2$ 时,RoR 只有 root-level shortcut 和 final-level shortcut。

2.2 多尺度交叉融合模块

对于特征图,多尺度模块按照不同的感受野去

提取特征。使用空洞卷积^[13]的方法既能够提升卷积核的感受野又能够不增加参数量,可以保留更多的细节信息。多尺度特征提取会产生若干个特征图,一般意义的多尺度模块会直接对它们进行融合处理,如图 2a 所示,将输入特征图标记为 Input。

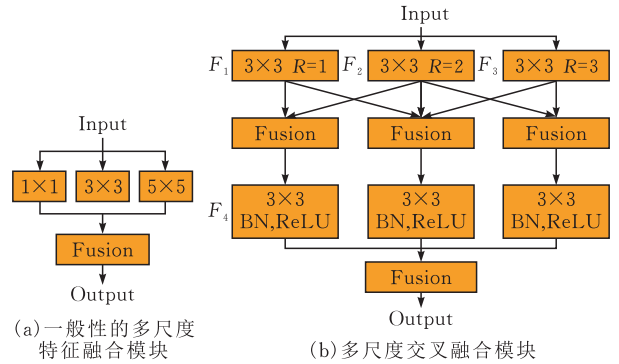


图 2 多尺度结构

本文提出了新颖的多尺度交叉融合模块,如图 2b 所示,感受野设置为 $3\times 3,5\times 5,7\times 7$,而且特征融合的方式也变成了交叉融合,交叉融合之后特征图再经历一次卷积操作,最后融合特征得到输出,最后的卷积层中包含了 BN 和 ReLU 操作。矩形表示特征操作,这里设输入特征图为 x , x 首先被尺寸大小分别为 $3\times 3,5\times 5,7\times 7$ 的卷积核同时提取特征,这三种卷积操作分别标记为 F_1,F_2,F_3 ,对应的输出标记为 $F_1(x),F_2(x),F_3(x)$,交叉融合的方式如图所示, $F_1(x)$ 和 $F_2(x)$ 进行融合, $F_1(x),F_2(x),F_3(x)$ 进行融合, $F_2(x)$ 和 $F_3(x)$ 进行融合,生成的三个特征图都传递到 F_4 操作中,最后再经历一次融合得到模块的输出。那么该模块的输出 x' 可以表示为:

$$F_4(F_1(x) + F_2(x)) + F_4(F_1(x) + F_2(x) + F_3(x)) + F_4(F_2(x) + F_3(x)) = x'$$

和一般性的多尺度模块存在一些区别,这里应用了交叉融合,然后又增加了 Conv-BN-ReLU 操作。交叉融合得到的特征包涵信息量更多,多尺度融合拥有一个融合结果,而交叉融合具有三个不同的融合结果,不同的融合结果包涵的信息丰富层次不同,尺度描述也不同,这是由于交叉路径使得信息的流动更加充分,不同尺度不同层次的特征信息可以相互结合生成更加丰富的表示。之所以再增加一层 Conv-BN-ReLU 操作,考虑有两点:一是因为经过交叉融合生成的特征图包含的信息存在冗余,这种不必要的冗余会影响后面的特征提取,所以额外设置一层 Conv-BN-ReLU 自适应的学习消除冗余,使特征的表示保持在合理的范围内,方便后续特征学习。二是由于残差结构的残差分支要求至少两层卷积, F_1,F_2,F_3 算作第一层卷积,这里使用 F_4 充当第二层。在实验部分将设计一组对比实验,保证总

体框架不变的情况下比较多尺度交叉融合模块和一般性的多尺度融合模块的性能作用。

2.3 总体结构

在编码器中堆叠多尺度交叉融合模块,这样的操作有两个好处:1)多尺度交叉融合机制可以极大地促进特征信息的传播流动;2)更多的卷积层能够提升网络的拟合能力。这样的设计存在深度增加网络退化的问题,这里应用了多级残差机制,遵循这样的思路:残差映射容易学习,如果让恒等映射也成为残差映射的一部分,那么这样的安排对于残差机制性能的挖掘是可观的。

本文提出了新颖的多级残差多尺度模块,结构见图 3。将这种模块嵌入每一个编码器中,对解码器的改动很轻微,仅仅把普通卷积换成深度可分离卷积,这样的操作是为了平衡整体网络模型的参数量,使模型容易训练。对于多级残差多尺度模块,设计它的 root-level 残差中包含有两个 middle-level 残差,每一个 middle-level 残差包含两个 final-level 残差,final-level 残差块的残差部分为多尺度交叉融合模块。

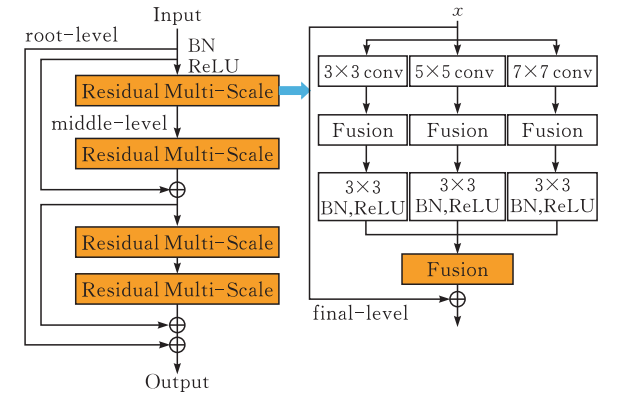


图 3 多级残差多尺度模块

模型总体架构如图 4 所示,除第一个编码器外,其余编码器中的第一个卷积层负责接收处理经过池化层后的特征图,并调整通道数,第二个卷积层仅负责提取特征并进行激活处理和归一化,在这个网络中,多级残差多尺度模块并不改变特征图的维度,仅仅起到提取特征的作用。网络模型的底部和第一个编码器类似,也是采用普通卷积加残差块的结构,希望能够规避过拟合。解码器结构顺序为特征拼接、深度可分离卷积,之所以大量采用深度可分离卷积,是为了一定程度降低参数量。

3 实验和结果

这项工作的实验使用了 CHASE_DB1 数据集。划分 20 个样本用于网络训练,另外 8 个样本用于测试。经过数据增强后的训练集容量为 600,每一张

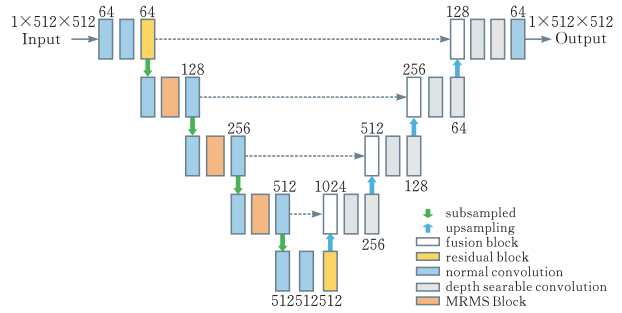


图 4 MRMS-Net 总体结构

训练图像裁减为 512×512 像素,并转换为灰度图。在训练阶段选择 RMSProp 优化算法,学习率为 0.0001,weight decay 为 e^{-8} ,momentum 为 0.9。

在图像分割领域存在着若干个衡量分割效果的指标,其中有 SE,SP,ACC。SE 指的是敏感性,正确分割的血管像素占真实血管像素的百分比,SP 指特异性,正确分割的背景像素占真实背景像素的百分比,ACC 指准确度,正确分割血管像素和背景像素在整个图像中的百分比。图 5 的内容显示了网络的工作效果,表 1 展示了 MRMS-Net 的性能表现以及和其它的网络效果的比较。从表 1 中可知,MRMS-Net 对视网膜血管的分割特异性和准确度表现较好,但是敏感性方面略有不足。

在这项工作中,还存在一组对照试验,MRMS-Net 作为标准网络,把 MRMS-Net 中的多尺度交叉融合模块替换为一般性多尺度融合模块,这种网络模型定义为 MRMS-Compare-Net,作为对照网络,这两种多尺度模块在前文中已经得到阐述,这里需要用实验结果来说明两个网络的差异(表 1)。

表 1 若干网络模型在 CHASE_DB1 数据集上的分割效果

Methods	Years	ACC	SE	SP
Recurrent U-Net ^[14]	2018	0.9622	0.7459	0.9836
Multi-scale Net ^[15]	2018	0.9657	0.7736	0.9871
Jiang, et al. ^[16]	2018	0.9668	0.8640	0.9745
LadderNet ^[17]	2018	0.9656	0.7978	0.9818
VesselNet ^[18]	2019	0.9661	0.8132	0.9814
Li X, et al. ^[19]	2020	0.9696	0.8452	0.9807
CSU-Net ^[20]	2020	0.9706	0.8427	0.9836
MRMS-Net	2021	0.9744	0.7753	0.9876
MRMS-Compare-Net	2021	0.9697	0.7832	0.9720

从表 1 中可以看到,标准网络的 ACC 和 SP 均高于对照网络,而对照网络的 SE 更高,有理由相信多尺度融合模块的不同在其中发挥了作用,虽然特征信息在标准网络中得到更充分的流动和融合,也采取了一些手段对冗余的信息加以抑制,但对照网络的多尺度融合模块毫无疑问在结构上更加简单直接,没有那么多的冗余堆叠。但也应该看到,这种简单模块并没有使残差机制发挥应有的作用,这一点在 ACC,SP 两项指标的差距上可以得到证明。

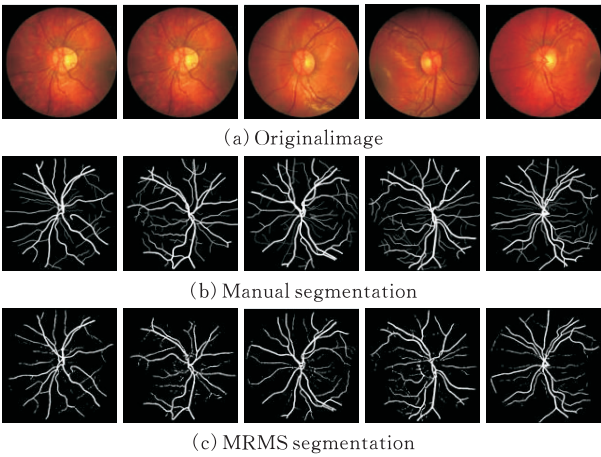


图 5 分割效果

4 结论

针对大部分 U 型网络编码器阶段卷积核的尺寸过于单一,网络拟合能力因深度增加而退化的问题,本文提出了一种多尺度交叉融合模块,其拥有较强的特征提取能力,并将多尺度交叉融合模块和多级残差机制相结合,形成了一种新的多级残差多尺度网络。一方面该网络同时兼顾良好的拟合能力和可收敛性,通过在公开数据集上进行实验比较,这种新的多尺度交叉融合相较于普通的多尺度融合在 ACC 和 SP 这两个指标上表现更优,这也体现在 MRMS-Net 和其它的网络的比较上。但是另一方面多级残差机制的引入也增加了网络的复杂度,复杂度的提升客观上使得网络参数量剧增,训练速度更慢。本论文所提出的网络在 SE 指标上亦存在不可忽视的问题,这种不平衡反映出网络模型在设计上存在一些问题,有待进一步研究解决。

[参 考 文 献]

[1] BEN-COHEN A,DIAMANT I,KLANG E,et al.Fully convolutional network for liver segmentation and lesions detection[C].// Athens, Greece: Springer Verlag,2016:77-85.

[2] DASGUPTA A,SINGH S.A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation[C].// Melbourne, VIC, Australia:IEEE Computer Society,2017:248-251.

[3] RONNEBERGER O,FISCHER P,BROX T.U-net: convolutional networks for biomedical image segmentation[C].// Munich, Germany: Springer Verlag, 2015:234-241.

[4] KRIZHEVSKY A,SUTSKEVER I,HINTON G.ImageNet classification with deep convolutional neural networks[J].Communications of the ACM,2017,60(6): 84-90.

[5] SIMONYAN K,ZISSERMAN A.Very deep convolutional networks for large-scale image recognition[EB/OL].(2014-9-4).[2021-9-20].<https://arxiv.org/abs/1409.1556>.

[6] SZEGEDY C,WEI LIU,YANGQING JIA,et al.Going deeper with convolutions[C].// Boston, MA, United States: IEEE Computer Society,2015:1-9.

[7] YAN CHENG WU,CHEN H C,SHAO MEI L I,et al.Person re-identification using attribute priori distribution[J].Acta Automatica Sinica,2019,45(5):953-964.

[8] HE KAIMING,ZHANG XIANGYU,REN SHAO-QING,et al.Deep residual learning for image recognition[C].// Las Vegas, NV, United states: IEEE Computer Society,2016:770-778.

[9] NI Z L,BIAN G B,ZHOU X H,et al.RAunet: residual attention U-net for semantic segmentation of cataract surgical instruments[C].// Sydney, NSW, Australia: Springer Science and Business Media Deutschland GmbH,2019:139-149.

[10] ZHANG K,SUN M,HAN X,et al.Residual Networks of residual networks: multilevel residual Networks[J]. IEEE Transactions on Circuits and Systems for Video Technology,2018,28(6):1303-1314.

[11] 练秋生,富利鹏,陈书贞,等.基于多尺度残差网络的压缩感知重构算法[J].自动化学报,2019,45(11):2082-2091.

[12] ZHENG K,XIA Z,ZHANG Y,et al.Speech emotion recognition based on multi-level residual convolutional neural networks[J].Engineering Letters,2020,28(2): 559-565.

[13] CHEN L C,PAPANDREOU G,KOKKINOS I,et al. DeepLab: semantic image segmentation with deep convolutional nets,atrous convolution,and fully connected CRFs[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2018,40(4):834-848.

[14] WANG W,YU K,HUGONOT J,et al.Recurrent U-net for resource-constrained segmentation[C].// Seoul, Korea: Institute of Electrical and Electronics Engineers Inc.,United States,2019:2142-2151.

[15] ZHANG B,HUANG S,HU S.Multi-scale neural networks for retinal blood vessels segmentation[EB/OL].(2018-4-11).[2021-9-20].<https://arxiv.org/abs/1804.04206>.

[16] JIANG Z,ZHANG H,WANG Y,et al.Retinal blood vessel segmentation using fully convolutional network with transfer learning[J].Computerized Medical Imaging and Graphics,2018,68(09):1-15.

[17] ZHUANG J.LadderNet: Multi-path networks based on U-Net for medical image segmentation[EB/OL].(2018-10-17).[2021-9-20].<https://arxiv.org/abs/1810.07810>.

[18] WU Y,XIA Y,SONG Y,et al.Vessel-net: retinal vessel

segmentation under multi-path supervision [C]. // Shenzhen, China: Springer Science and Business Media Deutschland GmbH, 2019: 264-272.

[19] LI X, JIANG Y, LI M, et al. Lightweight attention convolutional neural network for retinal vessel segmentation[J]. IEEE Transactions on Industrial Informatics, 2020, 17(03): 1958-1967.

[20] WANG B, WANG S, QIU S, et al. CSU-Net: a context spatial U-net for accurate blood vessel segmentation in fundus images [J]. IEEE Journal of Biomedical and Health Informatics, 2020, 25(04): 1128-1138.

Medical Image Segmentation Network based on Multilevel Residuals and Multi-scales

LI Wei, WU Cong

(School of Computer Science, Hubei Univ. of Tech., Wuhan 430068, China)

Abstract: At present, in most U-shaped networks, it is difficult to extract features with a single convolution kernel scale in the encoding stage, and it is also difficult to train and optimize the deep network. A new neural network based on multi-level residuals and multi-scale is proposed, which makes the neural network easier to learn, improves the depth of the network, and enables it to have richer feature expression ability, without reducing the performance of the model. In this paper, a multi-scale cross fusion module is proposed, which extracts features through different sensory scales. Cross-fusion also enables feature information to be more fully exchanged and fused. After testing on the data set CHASE_DB 1, the performance is excellent, especially with ACC being 0.9744, and SP being 0.982. The depth of the proposed network is increased without affecting its learning process and performance.

Keywords: blood vessel segmentation; deep learning; u-net; residual learning; multi-scale

[责任编辑: 张岩芳]

(上接第 37 页)

Fast Charging Method based on Improved Cuckoo Algorithm

ZHANG Yu, XU Si

(Hubei Key Laboratory of Solar Energy Efficient Utilization and Energy Storage Operation Control, Hubei Univ. of Tech., Wuhan 430068, China)

Abstract: In spite of the current rapid development of electric vehicles, the problem of rapid charging of lithium-ion batteries has become a bottleneck in the application. This paper presents a fast charging method based on an improved cuckoo algorithm. Firstly, the equivalent circuit model and thermal model of lithium-ion battery are established. Then a multi-step charging method is used with charging time and temperature rise as the optimization objectives. Finally, the improved cuckoo algorithm is used to search for the charging current at each stage. The influence of the weighting factors of charging time and temperature rise on the battery charging performance is discussed. Compared to the conventional cuckoo algorithm, the optimised cuckoo algorithm charging method is able to reduce the charging time by around 5% with almost the same temperature rise, balancing the charging speed and temperature rise and extending the life of the Li-ion battery.

Keywords: Lithium-ion battery; improve the cuckoo algorithm; charging time; temperature rise

[责任编辑: 张岩芳]