

DU-Net：基于改进型 U-Net 的视网膜血管分割模型

孟 瑶¹ 孔 劼¹
MENG Yao KONG Jie

摘 要

使用视网膜血管的图像可以识别各种疾病，很多眼部疾病需要通过眼球的毛细血管状况来进行判别，因此视网膜血管的准确分割对一些疾病的早期诊断非常重要。医疗影像数据量与 U-Net 模型在体量上相匹配，有效避免了过拟合。但 U-Net 网络通常包含大量的层和参数，特别是在处理大型图像数据集时训练和推理的时间较长。为更好地分割视网膜血管，提供一个基于改进 U-Net 的视网膜血管图像分割方式，用 Do 卷积代替 U-Net 路径上的卷积运算，在不增加网络推理运算的情况下加快收敛速度并改善网络性能。在 DRIVE 数据集上的实验结果表明，与原始 U-Net 相比，IoU 提升了 1.2%，Dice 提升了 1.7%，Global correct 提升了 0.1%。

关键词

视网膜血管分割；U-Net；DU-Net；Do 卷积；医学图像分割

doi: 10.3969/j.issn.1672-9528.2024.06.009

0 引言

视网膜血管是人体内唯一可以用肉眼直接观察到的血管。这些血管的形状、直径、规模、分支角度、增生和渗出的变化可以反映身体的疾病^[1]。因此，检测和提取视网膜图像中的血管对相关疾病的诊断和治疗具有重要的临床医学意义^[2]。随着最近医疗影像的激增，手动处理医疗影像需要花费大量的时间和精力，临床诊断和治疗面临很大困难。因此，迫切需要一种自动图像分割方法来帮助临床医生快速诊断并减轻医生的负担，这是研究人员试图解决的问题^[3]。此外，医学图像中经常存在许多异常，如目标物体的对比度低、边界模糊以及损伤区域存在大量噪声。这些问题使得医学图像的精确和有效分割变得极其困难和具有挑战性^[4]。近年来，深度卷积神经网络在计算机视觉领域得到了广泛应用。但是，由于上述问题，训练具有高泛化能力的深度网络来分割视网膜图像并不简单。将深度学习应用于该领域的两个主要策略是：（1）将每个图像分割成小块，最大限度地提高训练样本数；（2）结合现有的卷积层和上采样，同时学习本地和全局特征。在图像处理领域，深度学习已经显示出强大的优势，人工神经网络已经广泛用于图像分类、分割和识别。对于医学图像分割方法，许多学者进行了相关研究。2015 年，LONG J 等人^[5]首次提出了全卷积神经网络（fully convolutional

network, FCN）来进行语义图像分割，随后出现了各种不同的语义分割网络。其中，U-Net^[6]网络模型因其良好的分割性能，逐渐成为医学图像分割领域的焦点，之后出现了许多基于 U-Net 改进的网络模型。Shankaranarayana 等人^[7]将残差模块和 U-Net 结合提出 Res-UNet，残差模块加深了网络结构，增强浅层与深层特征的融合，加快了网络的收敛速度，同时避免了梯度的消失与爆炸问题。Oktay 等人^[8]将注意力机制应用于 U-Net 分割网络中，提出基于注意力的 U-Net，利用注意力门可以较好地捕捉显要特征，并对无关背景区域进行抑制。Rui 等人^[9]结合残差模块和 U-Net 提出了 Res-UNet，残差模块深化了网络结构，避免了梯度消失和爆炸的问题。Zhou 等人^[10]提出 U-Net++，通过跳跃连接整合各个层次的特征，并设计了一种剪枝策略，在保证性能的同时加快了推理速度，灵活的网络构造与深监督结合，大幅减少了参数量。Li 等人^[11]基于空间位置信息提出了可变形 U-Net，用可变卷积的空间变形能力来处理病灶部分的几何变换，同时改进了池化模块，可以更好地保留池化操作后的信息。Kitrungrotsakul 等人^[12]提出深度交互的优化网络 RefineNet，可自动优化分割效果。Zhao 等人^[13]提出了用于医学图像分割的小波残差注意网络（WRANet）。他们用离散小波变换取代了 CNN 中的标准降采样模块（如最大池化和平均池化），将特征分解为低频成分和高频成分，并去除高频成分以消除噪声。同时，通过引入注意力机制，可以有效解决特征丢失的问题。Mehreen M 等人^[14]提出了一种新的基于 U-Net 的医

1. 西安石油大学计算机学院 陕西西安 710076

学图像分割架构 R2U++, 在所提出的架构中, 与普通 U-Net 相比, 所做的调整包括以下两方面: (1) 普通卷积骨干被更深的递归残差卷积块取代, 这些区块增加了视野, 有助于提取关键特征进行分割, 网络整体性能的提高也证明了这一点; (2) 通过密集的跳过路径, 缩小了编码器和解码器之间的语义差距。这些路径积累了来自多个尺度的特征, 并进行相应的连接。改进后的架构嵌入了多深度模型, 不同深度的输出集合提高了对图像中不同尺度前景物体的处理性能。全卷积性神经网络 (fully convolutional network, FCN)^[15-16] 是深度学习的一种方法, 是根据语义级别图像分割问题而给出的。将 GroundTruth 作为监督信息训练网络, 由网络进行图像等级的预测, 将图像等级的划分逐步扩展至图像等级的划分。

上述研究将对医学影像分割的研究和发展带来一些借鉴价值, 能增强网络的分割功能, 提高对影像中细节区的分离能力。同时, 针对眼底血管结构复杂、影像对比度降低的问题, 在 U-Net 编码过程中, 需要对输入图像进行特征提取, 利用传统卷积去处理数据参数量过大, 特征提取会有缺失, 本文提出了一种基于改进 U-Net 的分割网络 DU-Net, 利用 Do 卷积^[17-18] 替换传统 U-Net 编码结构中的卷积操作, 提高分割精度的同时, 增加了模型收敛速度。

1 方法

1.1 Do 卷积

利用传统的卷积操作, 本文使用的卷积核通道数要与输入特征图的通道数保持一致, 然后进行一次卷积运算之后得到一个通道的特征图, 如果要获得多通道的输出, 本文使用的卷积核的数量必须与输出通道保持一致。这种卷积操作的参数量较大, 因为每个卷积核都需要学习多个权重值, 且不同的卷积核之间没有共享权重, 参数量大, 计算复杂度高, 容易导致过拟合。因此, 选择使用 Do 卷积来代替 U-Net 编码器中的传统卷积操作, 在普通卷积层中加入额外的 depthwise 卷积操作, 构成一个 over-parameterized 的卷积层。在推理过程中, 可以通过 DO-Conv 转换为现有的卷积运算, 因此即使在网络上用 DO-Conv 代替现有的卷积, 计算要求也不会增加。Do 卷积实现可分为两种方式。

如图 1 所示, 对输入特征 P 进行 depthwise 卷积操作, 得到中间变量 P' ; 然后对 P' 进行传统卷积操作, 得到最终结果 O , 并且 $D_{mul} \geq M \times N$ 。

$$P' = D \cdot P : D_{mul} \times (M \times N) \times C_{in} \times (H \times W) \quad (1)$$

$$O = W * P' : C_{out} \times C_{in} \times H \times W \times D_{mul} \quad (2)$$

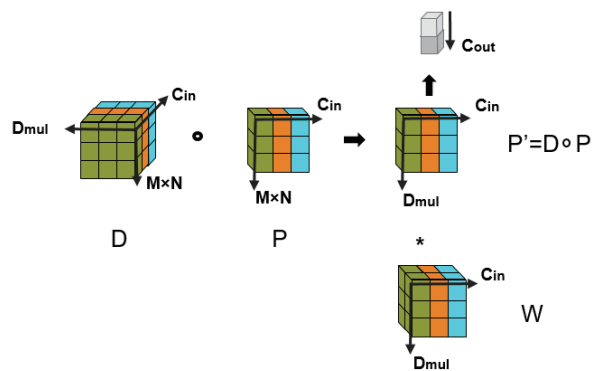


图 1 Do 卷积中的 Feature composition 示意图

如图 2 所示, 先计算两个权重的乘积, 得到新的权重 W' 。然后利用权重 W' 对输入特征 P 进行传统的卷积运算, 得到最终结果 O 。

$$W' = D^T \cdot W : D_{mul} \times (M \times N) \times C_{in} \times C_{out} \quad (3)$$

$$O = W' * P : C_{out} \times C_{in} \times H \times W \times (M \times N) \quad (4)$$

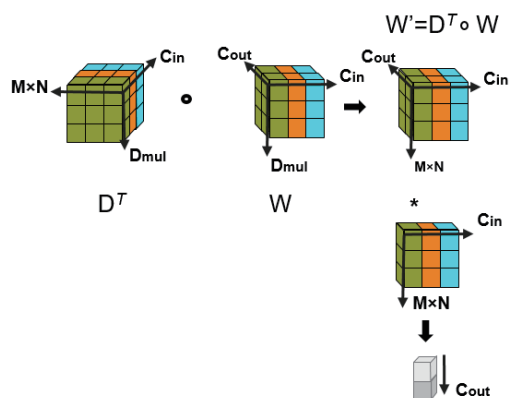


图 2 Do 卷积中的 Kernel composition 示意图

1.2 DU-Net 网络整体结构图

如图 3 所示, 将 U-Net 网络的编码部分所使用的 3×3 传统卷积, 替换为 Do 卷积, 但是在实验的过程中, 并不是一性替换所有的卷积层, 而是通过实验逐层替换, 以确定最好的实验效果。

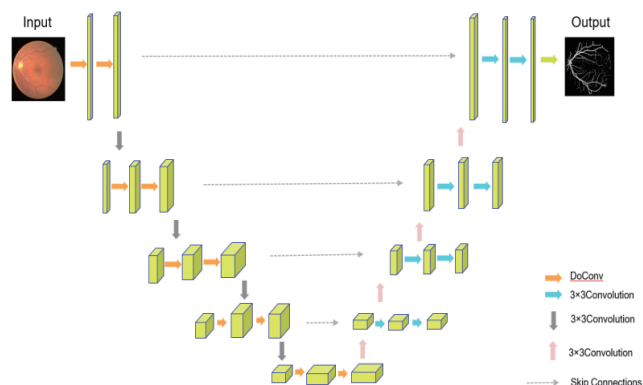


图 3 DU-Net 结构图

输入图像大小为 $572 \times 572 \times 1$ ，其中 1 是通道数，卷积核为 3 的两次 Do 卷积运算获得 $568 \times 568 \times 64$ 大小的特征图。之后通过最大池运算进一步缩小特征图的大小，获得 $284 \times 284 \times 64$ 大小的特征图。以此类推，不再赘述。在解码器部分利用双线性插值进行上采样，这里用的卷积为传统的 3×3 卷积。

2 实验结果与分析

2.1 数据集

本文使用的是 DRIVE (digital retinal images for vessel extraction) 数据集，此数据集用于视网膜血管分割。共有 40 个 JPEG 格式的彩色眼底图像，其中包括 7 个异常的病理案例。这些图像是在荷兰的糖尿病视网膜病检查项目中收集的。图像是使用 Canon CR5 非散瞳 3CCD 相机拍摄的，视场角 (FOV) 等于 45° 。每个图像的分辨率为 584 像素 $\times$ 565 像素，每个颜色通道为 8 位。这 40 个图像被等分为 20 个训练集图像和 20 个测试集图像。在这两个集合中，每张图像都有一个直径约为 540 像素的圆形视场掩码 (FOV)。在训练集中，每张图像都有一位眼科专家进行的手动分割。在测试集中，每个图像由两个不同的观察者手动分割，并将第一个观察者的分割作为性能评估的标准。图 4 显示了 DU-Net 对视网膜的分割结果，从左到右依次为原图、U-Net 分割结果和 DU-Net 分割结果。可以看出，本文提出的方法分割出的血管更加清晰，而且分割出的小血管数量更多。这些结果充分体现了本文提出的方法在特征提取方面的优势。



图 4 DRIVE 数据集的分割结果示例

2.2 评价指标

分别使用 IoU、Dice 和 Global correct 作为网络的评价指标。

(1) IoU 是将预测分割和标签之间的重叠区域除以预测分割和标签之间的结合区域的值，该指标的范围为 $0 \sim 1$ ($0 \sim 100\%$)，其中 0 表示没有重叠，1 表示完全重叠分割。如公式所示，TP (true positive) 表示模型正确地预测为正类的像素数量，即模型预测为正类，并且真实标签也为正类的像素数量；FP (false positive) 表示模型错误地预测为正类的像素数量，即模型预测为正类，但真实标签为负类的像素数

量；FN (false negative) 表示模型错误地预测为负类的像素数量，即模型预测为负类，但真实标签为正类的像素数量；TN (true negative) 表示模型正确地预测为负类的像素数量，即模型预测为负类，并且真实标签也为负类的像素数量。

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FN} + \text{FP}} \quad (5)$$

(2) Dice 系数的取值范围在 $0 \sim 1$ 之间，其中 1 表示完美预测，0 表示完全不匹配。较高的 Dice 系数表示模型的预测结果与真实标签有较好的重叠，即模型的性能更好。

$$\text{Dice} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (6)$$

(3) Global correct 用来衡量模型在整个数据集或测试集上的整体准确性。

$$\text{Global correct} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (7)$$

2.3 实验结果

利用 Do 卷积替换 U-Net 不同层次上的卷积，并测试每一层的效果。如表 2 所示，不断增加替换的层数，通过比对，本文发现将编码器所使用的传统卷积全部替换成 Do 卷积时，效果最好。

表 1 各层添加 Do 卷积的实验结果

	IoU	Dice	Global correct
U-Net	80.4	79.6	95.1
第一层	81.0	80.4	95.2
第二层	80.5	79.9	94.9
第三层	81.1	80.3	95.2
第四层	81.1	80.5	95.1
瓶颈层	81	80.4	95.3

表 2 不断增加替换层数后的实验结果

	IoU	Dice	Global correct
U-Net	80.4	79.6	95.1
前两层	80.9	80.2	95.1
前三层	81.3	80.8	95.3
前四层	81.3	80.7	95.3
DU-Net	81.6	81.3	95.2

3 总结

本文基于 U-Net 网络，设计了一种新的视网膜血管分割方案。在现有卷积运算中添加额外参数以形成 DO-Conv，使用 DO-Conv 代替现有卷积可以加快收敛速度并改善网络性能，而无需增加网络推论的计算量。本文所提出的 DU-Net 网络分割视网膜下血管精度更高，更加利于辅助医生进行病

情的诊断与治疗。在公共数据集 DRIVE 上进行测试,与基线模型相比,DU-Net 的三项评价指标都有所提升。但是该方法也有需要改进的地方,只做了在编码路径上的实验。在后期的研究中,将进一步测试在解码路径上利用 Do 卷积的分割效果。

参考文献:

- [1]BENSMAIL I, MESSADI M, FEROU I A, et al.New methodology based on images processing for the diabetic retinopathy disease classification[J].International journal of biomedical engineering and technology,2022,39(2):170-187.
- [2]LIU X, SONG L, LIU S,et al.A review of deep-learning-based medical image segmentation methods [J].Sustainability, 2021, 13(3): 1224.
- [3]VALANARASU J M J, VISHAL M P.UNeXt:MLP-based rapid medical image segmentation network[C]//Medical Image Computing and Computer Assisted Intervention- MICCAI 2022, Part V.Cham:Springer,2022:23-33.
- [4]ZHONG S, GUO X, ZHENG Y.Improved UNet network for lung nodule segmentation[J].Computer engineering and applications, 2020,56(17):203-209.
- [5]LONG J, SHELHAMER E, DARRELL T.Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference On Computer Vision And Pattern Recognition,[v.4].Piscataway:IEEE,2015:3431-3440.
- [6]RONNEBERGER O, FISCHER P,BROX T.U-net:convolutional networks for biomedical image segmentation[C]//Medical Image Computing And Computer-Assisted Intervention--MICCAI 2015,Part III.Cham: Springer, 2015: 234-241.
- [7]SHANKARANARAYANA S M, RAM K, MITRA K, et al.Joint optic disc and cup segmentation using fully convolutional and adversarial networks[C]//Fetal, Infant and Ophthalmic Medical Image Analysis. Cham: Springer, 2017:168-176.
- [8]OKTAY O, SCHLEMPER J, FOLGOC L L, et al.Attention U-Net:learning where to look for the pancreas[EB/OL].(2018-04-11)[2024-02-22]. <https://doi.org/10.48550/arXiv.1804.03999>.
- [9]RUI L, YING H, XIAO Y, et al.Ocular biomarkers of cognitive decline based on deep-learning retinal vessel segmentation[J]. BMC geriatrics, 2024,24(1):1471-2318.
- [10]ZHOU Z, SIDDIQUEE M M R, NIMA T, et al. UNet++: redesigning skip connections to exploit multiscale features in image segmentation[J].IEEE transactions on medical imaging, 2020, 39(6): 1856-1867.
- [11]LI Z, PAN H, ZHU Y, et al. PGD-UNet: a positionguided deformable network for simultaneous segmentation of organs and tumors[C]//2020 International Joint Conference on Neural Networks (IJCNN). Piscataway: IEEE,2020:1-8.
- [12]KITRUNGROTSAKUL T, YUTARO I, LIN L, et al. Interactive deep refinement network for medical image segmentation[EB/OL].(2020-06-27)[2024-02-28].<https://doi.org/10.48550/arXiv.2006.15320>.
- [13]ZHAO Y, WANG S, ZHANG Y, et al. WRANet:wavelet integrated residual attention U-Net network for medical image segmentation.[J].Complex intelligent systems, 2023, 9(6): 11-13.
- [14]MEHREEN M, HAZRAT A, CHRISTER G, et al.R2U++:a multiscale recurrent residual U-Net with dense skip connections for medical image segmentation[J].Neural computing applications, 2022,34(20):11-17.
- [15]JAGANNATH A, BIPUL N.Multi-Scale feature map aggregation and supervised domain adaptation of fully convolutional networks for urban building footprint extraction[J]. Remote sensing,2023,15(2):488.
- [16]JIN M, YANG Y, MING C, et al. DO-Conv: depthwise over-parameterized convolutional layer[C]//IEEE Transactions on Image Processing(Volume:31).Piscataway: IEEE, 2020: 3726-3736.
- [17]CAO J M, LI Y Y, SUN M C, et al. DO-Conv: depthwise over-parameterized convolutional layer[J]. IEEE transactions on image processing, 2022, 31:3726-3736.
- [18] WANG H, JIA H, LU L, et al. Thorax-net: an attention regularized deep neural network for classification of thoracic diseases on chest radiography[J]. IEEE journal of biomedical and health informatics, 2019, 24(2): 475-485.

【作者简介】

孟瑶(1998—),女,山东东营人,本科,研究方向:计算机技术。

(收稿日期:2024-03-17)