

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

引入双编码器模型的OCT视网膜图像分割

陈明惠, 王腾, 袁媛, 柯舒婷

引用本文:

陈明惠, 王腾, 袁媛, 等. 引入双编码器模型的OCT视网膜图像分割[J]. *光电工程*, 2023, 50(10): 230146.

Chen M H, Wang T, Yuan Y, et al. Study on retinal OCT segmentation with dual-encoder[J]. *Opto-Electron Eng*, 2023, 50(10): 230146.

<https://doi.org/10.12086/oe.2023.230146>

收稿日期: 2023-06-26; 修改日期: 2023-11-10; 录用日期: 2023-11-14

相关论文

基于注意力机制多特征融合的视网膜病变分级算法

梁礼明, 董信, 李仁杰, 何安军

光电工程 2023, 50(1): 220199 doi: 10.12086/oe.2023.220199

边界注意力辅助的动态图卷积视网膜血管分割

吕佳, 王泽宇, 梁浩城

光电工程 2023, 50(1): 220116 doi: 10.12086/oe.2023.220116

光学相干层析技术在眼科手术导航方面的研究进展

李云耀, 樊金宇, 蒋天亮, 唐宁, 史国华

光电工程 2023, 50(1): 220027 doi: 10.12086/oe.2023.220027

更多相关论文见光电期刊集群网站 



<http://cn.ojournal.org/oe>



 OE_Journal



Website

DOI: 10.12086/oe.2023.230146

引入双编码器模型的 OCT 视网膜图像分割

陈明惠*, 王 腾, 袁 媛, 柯舒婷

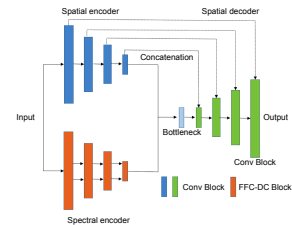
上海理工大学 健康科学与工程学院 上海介入医疗器械 工程技术研究中心
教育部医学光学工程中心, 上海 200082

摘要: OCT 视网膜图像中存在着噪声和散斑, 单一的提取空间特征往往容易遗漏一些重要信息, 导致不能准确地分割目标区域。而 OCT 图像本身存在光谱频域特征, 针对 OCT 图像的频域特征, 本文基于 U-Net 和快速傅立叶卷积提出一种新的双编码器模型以提高对 OCT 图像视网膜层、液体的分割性能, 提出的频域编码器可以提取图像频域信息并通过快速傅里叶卷积转换为空间信息, 将很好地弥补单一空间编码器遗漏特征信息的不足。经过与其他经典模型的对比和消融实验, 结果表明, 随着频域编码器的添加, 该模型能有效提升对视网膜层和液体的分割性能, 平均 Dice 系数和 mIoU 相较于 U-Net 均提高 2%, 相较于 ReLayNet 分别提高 8% 和 4%, 其中对液体的分割提升尤为明显, 相较于 U-Net 模型 Dice 系数提高了 10%。

关键词: 光学相干层析成像; 卷积神经网络; 图像分割; 双编码器

中图分类号: TP394.1; TH691.9

文献标志码: A



陈明惠, 王腾, 袁媛, 等. 引入双编码器模型的 OCT 视网膜图像分割 [J]. 光电工程, 2023, 50(10): 230146

Chen M H, Wang T, Yuan Y, et al. Study on retinal OCT segmentation with dual-encoder[J]. *Opto-Electron Eng*, 2023, 50(10): 230146

Study on retinal OCT segmentation with dual-encoder

Chen Minghui*, Wang Teng, Yuan Yuan, Ke Shuting

Shanghai Engineering Research Center of Interventional Medical, Shanghai Institute for Interventional Medical Devices, School of Health Science and Engineering College of Health Sciences and Engineering, University of Shanghai for Science and Technology, Shanghai 200082, China

Abstract: There are noises and speckles in OCT retinal images, and a single extraction of spatial features is often easy to miss some important information. Therefore, the target region cannot be accurately segmented. OCT images themselves have spectral frequency domain characteristics. Aiming at the frequency domain characteristics of OCT images, this paper proposes a new dual encoder model based on U-Net and fast Fourier convolution to improve the segmentation performance of the retinal layer and liquid in OCT images. The proposed frequency encoder can extract image frequency domain information and convert it into spatial information through fast Fourier convolution. The lack of feature information that can be omitted by a single space encoder will be well-complemented. After comparison with other classical models and ablation experiments, the results show that with

收稿日期: 2023-06-26; 修回日期: 2023-11-10; 录用日期: 2023-11-14

基金项目: 上海市科委产学研医项目 (15DZ1940400)

*通信作者: 陈明惠, cmhui.43@163.com。

版权所有©2023 中国科学院光电技术研究所

the addition of a frequency domain encoder, the model can effectively improve the segmentation performance of the retinal layer and liquid. Both average Dice coefficient and mIoU are increased by 2% compared with U-Net. They are increased by 8% and 4% compared with ReLayNet, respectively. Among them, the improvement of liquid segmentation is particularly obvious, and the Dice coefficient is increased by 10% compared with the U-Net model.

Keywords: optical coherence tomography; convolutional neural network; segmentation of image; dual-encoder

1 引言

视网膜病变是一种常见地眼科疾病, 大多数眼底疾病是由视网膜病变引起的^[1]。研究表明, 在大多数眼底疾病中, 视网膜形态结构的变化早于患者在视觉上感知到的变化, 检测视网膜各层的微小形态变化来分析视网膜形态结构对早期发现眼底视网膜疾病具有关键价值。近年来, 视网膜相关的各种医学图像被广泛应用于视网膜疾病的诊断、筛查和治疗中^[2-4]。光学相干层析成像 (Optical coherence tomography, OCT) 可以直接显示视网膜的分层结构状态并获取其相应的厚度, 目前已成为一项重要的眼科诊断技术, 并被常规用于监测视网膜状态变化^[5]。视网膜层厚度和液体分布可用于诊断眼科疾病, 自动分割视网膜 OCT 图像中的视网膜层可以帮助有效地诊断和监控视网膜疾病, 如糖尿病黄斑水肿、青光眼、多发性硬化症等。因此, 准确分割视网膜图像中的视网膜层边界是进行病理分析和诊断的关键步骤。然而, 视网膜各层的图像有时是复杂的, 难以分割。例如, 糖尿病可能导致黄斑水肿 (macular edema, ME), 在检查 ME 的视网膜 OCT 图像中可以看到部分液体区域, 该区域具有低反射强度, 是视网膜组织中的低对比度区域^[6]。液体区域使视网膜组织变形, 这可能导致分割算法失败。因此, 快速准确地分割视网膜各层及液体区域是具有挑战性的。

传统的视网膜层分割方法通常基于数学模型, 通过分析解剖结构和先前的临床知识来检测视网膜层。Monemian 等^[7]提出利用像素的纹理属性来提取层间边界像素的区分特征。Liu 等^[8-10]基于水平集的方法对视网膜各层进行自动分割, 该方法可以在训练随机森林的过程中获得边界概率图, 随着训练次数的不断增加可以将分割结果准确到亚像素精度。Hussain 等^[11]通过将问题建模为图并运用 Dijkstra 最短路径算法, 找到三个参考层的近似位置, 然后将搜索空间

限定为实际层。由于拓扑在边界确定中起到很好的作用, 因此可以克服扭曲的影响。这些方法在视网膜层分割中是有效的, 但鉴于不同患者 OCT 图像内的特征有所差异, 传统算法在不同数据集上的效果也都有明显差异, 这导致基于传统算法的分割任务鲁棒性不太理想, 并且一些算法的相关参数还需要有经验的实验人员去实时调整。

近些年, 深度学习的发展尤为迅速, 特别是在 OCT 视网膜分层中达到了比传统方法更准确的效果。卷积神经网络 (Convolutional neural network, CNN) 可自动提取和学习图像特征, 已被广泛应用于各种模式的图像处理中, 后续出现了许多关于 CNN 的变体结构^[12-16]。U-Net^[17]是最早且最著名的基于自动编码器和解码器的医学图像分割架构之一。杜克大学 OCT 实验室在 2017 年基于 U-Net 的编码器、解码器结构提出了一种新的全卷积深度模型, 称为 ReLayNet^[18], 成功地对 OCT 图像中各类视网膜层以及液体区域进行了分割。ReLayNet 通过编码器级联进行下采样使模型对特征图进行信息提取分类, 从而进行有效的对上下文特征的学习, 通过解码器的上采样路径进行信息的整合恢复以生成完整的分割图。后续也有相关工作采取解码器单级联多输出进行分割工作^[19]。MDAN-UNET^[20]使用多尺度特征和注意力机制来提高现有方法的分割性能, 可以基于空间和通道两方面很好地提取全局信息, 实验对 OCT 视网膜液体区域和视网膜层进行了分割并取得了理想的效果。Ngo 等^[21]提出了一种基于特征学习回归网络的自动分割方法, 该方法将图像片段的亮度、梯度和自适应归一化强度分数作为学习特征, 然后预测相应的视网膜边界像素。然而, 这些基于卷积神经网络的方法感受野范围较小, 对全局信息关注较少。Chi 等^[22]实验表明, 空间信息只关注局部信息, 而往往忽略了图像中所有像素的全局信息, 从而提出了一种新的卷积-快速傅里叶卷积 (fast Fourier convolution, FFC), 它不仅观察到非局部的特征区域, 还可以在卷积内部

进行跨尺度信息的处理融合。为了减少卷积网络的计算成本,人们探索了 U-Net^[17] 和快速傅立叶变换^[23] 的组合,其目的是进一步利用图像中不能被常规卷积层很好地提取的全局特征。

综合上述相关工作,在类别更多、更复杂的分割任务下,单一的设计增加各种机制模块无疑提升了模型的复杂度和硬件需求,同时一个大的有效感受野对于理解图像的整体结构和解决修复问题虽然是必不可少的,但在大掩模的情况下,大但有限的接受场也可能不足以访问能产生高质量修复所需的信息。因为 OCT 图像中存在频域特征,而这些特征在现有的空间卷积神经网络中可能会被遗漏。本文利用快速傅里叶卷积从光谱频域中提取和处理全局信息,提出一种基于 U-Net 的双编码器分割模型,用于分割 OCT 视网膜图像,分割任务包括内界膜(inner limiting membrane, ILM)、神经纤维层-内丛状层(nerve fiber ending to Inner plexiform layer, NFL-IPL)、内核层(inner nuclear layer, INL)、外丛状层(outer plexiform layer, OPL)、外核层-光感受器内节段髓系(outer nuclear layer to Inner segment myeloid, ONL-ISM)、内感光和外感光层联结(inner segment ellipsoid, ISE)、外节段-视网膜色素上皮细胞层(outer segment to retinal pigment epithelium, OS-RPE)以及视网膜图像存在的一些液体区域。利用频域编码器来提取频域特征,提出 FFC-DC 模块处理全局和局部信息,通过逆转换将频域特征转换为空间特征来补充空间编码器可能会忽略的信息。实验发现,基于快速傅里叶卷积的

神经网络的性能优于类似的卷积模型,并且表现出比其他分割模型更好的结果。

2 实验方法及原理

2.1 空间编码器、解码器

如图 1 所示,本文模型中的空间编码器和解码器与原始的 U-Net 相同,空间编码器包含四个卷积块。每个卷积块包含一个 3×3 卷积、BN 层、ReLU 层以及最大池化 MP 层,编码器通过最大池化层进行下采样四次,对应的解码器上采样也进行四次,其中左右两侧同级别的特征图中间会进行级联操作。空间解码器网络总共具有四个上卷积块,这四个上卷积块接收频域和空间特征,并在这四个上卷积块传递到瓶颈层之前将它们连接起来。然后,使用相等于空间编码器的卷积块来放大来自先前解码器块和跳过连接的特征,经过转置卷积层后解码器块生成最终分割图。

2.2 频域编码器

如图 2 所示,与空间编码器相同,在频域编码器中,实验选择应用 FFC 和扩张卷积(Dilated Convolution, DC)模块组成四个 FFC-DC 块,频域编码器接收与空间编码器相同的输入并从数据中提取频域特征。FFC-DC 块主要使用 DC 模块处理局部特征,使用频域 Transformer 来处理全局特征。输入图像 x 作为局部信息, x_1 被输入到第一个 FFC-DC 块。因为输入图像中没有全局特征,图像像素被认为是局部信息,所以输入进来的第一个 x_g 的值被设置为 0。

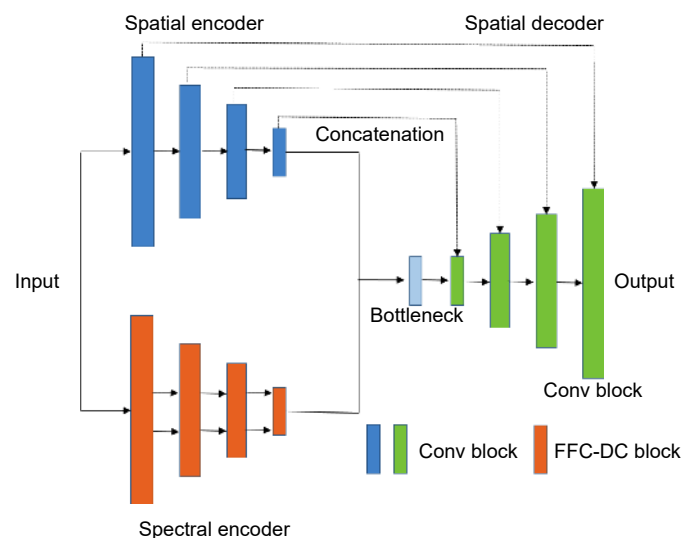


图 1 模型结构图

Fig. 1 Diagram of the model structure

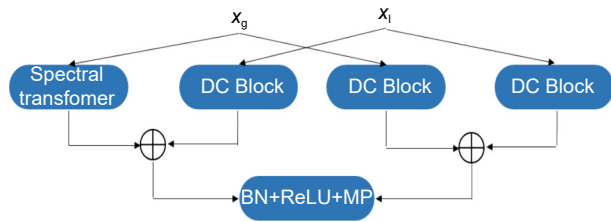


图 2 FFC-DC 模块

Fig. 2 Architecture of the FFC-DC block

在 FFC-DC 模块中特征图被四个分支所处理, 分别为局部到局部、局部到全局、全局到局部、全局到全局, 每条路径都可以捕捉到不同感受野的互补信息。这些路径之间的信息交换在内部执行。快速傅里叶卷积 FFC 块接收输入的全局和局部信息 x_g 、 x_l , 然后将 x_g 和 x_l 输入到三个 DC 模块和频谱 Transformer 中, 提取全局和局部空间特征以及频域范数, 从而进行频域特征提取。最后, 将批归一化、非线性激活函数和最大池化合并应用于特征, 以生成下一个 FFC 块的全局和局部特征 x'_g 、 x'_l 。

2.2.1 DC 模块

以往的工作通常采取常规卷积来处理局部特征, 与传统卷积相比, 扩张卷积具有特殊的优点^[24]。它可以随着扩张参数的改变而随意生成不同的感受野, 并获得更丰富的上下文信息且不增加模型复杂度。扩张卷积产生的特征图可以与输入的特征图具有相同的尺度, 同时每个输出卷积会具有更大的感受野, 因此它可以获取更高级别的语义信息。DC 块的结构如图 3 所示, 在 DC 块中, 特征分别被输入 1×1 和 3×3 两个常规卷积, 所提取的特征信息会输入到 3×3 扩张卷积中。通过这种方式, DC 块可以获得不同大小的感受野。最后, 将所有特征送入 1×1 卷积, 以降低信道的维度。该方法可以使提出的模型更有效地进行局部的特征提取。

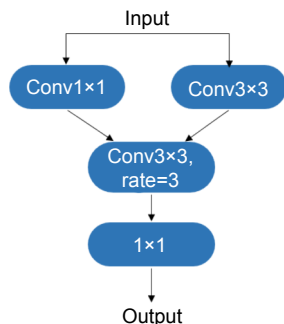


图 3 DC 卷积模块

Fig. 3 Architecture of the DC block

2.2.2 频域 Transformer

如图 4, 频域 Transformer 首先将核大小为 1 的卷积块应用于产生 x'' 的 x_g 。然后基于预定值 α 将 x'' 的通道分成两部分, 其中 α 百分比的通道被设定为全局信息, 而 $1-\alpha$ 百分比的通道被设定为局部信息。分割的全局特征和局部特征分别输送到傅里叶单元 GFU (global Fourier unit)、LFU (local Fourier unit) 以提取频域特征 x''_g 、 x''_l 。

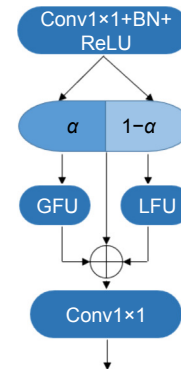


图 4 频域 Transformer

Fig. 4 Architecture of the spectral Transformer

如图 5 所示, 其中 GFU 和 LFU 结构相同, LFU 主要目的是捕捉半全局信息, 将输入图分为四份处理, 然后进行傅里叶变换, 这样可以处理四分之一空间内的关联特征。傅里叶单元接收 x'' 的一部分作为输入, 然后对这些特征应用傅里叶变换以获得实部和虚部 $a+bi \in \mathbb{C}$ 。将实部和虚部 a 、 b 叠加, 然后输入到核大小为 1 的卷积层, 将激活层和批归一化层应用于卷积层的输出。输出分成两部分, 即实部和虚部 a' 、 b' , 随后其会经过傅里叶逆变换将提取的频域特征转换回空间特征。最后, x 与全局和局部傅里叶单元 x''_g 、 x''_l 的输出相加。

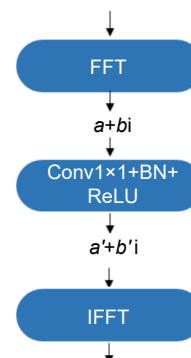


图 5 傅里叶单元

Fig. 5 Architecture of the Fourier unit

2.3 损失函数

本文提出的框架是一种端到端的深度学习方法。交叉熵损失函数被经常应用到深度学习图像分割中。然而, 医学方面的图像分割目标通常占据图像很小一部分的区域, 交叉熵损失函数对于这类任务并不是最好最理想的。经过实验验证, 选择结合了 Dice 系数损失和加权交叉熵损失的新损失函数来代替交叉熵损失函数, 损失函数公式如:

$$L_{\text{dice}} = 1 - \frac{\sum_k \sum_i 2\omega_k p_{(k,i)} g_{(k,i)}}{\sum_k \sum_i p_{(k,i)}^2 + \sum_k \sum_i g_{(k,i)}^2}, \quad (1)$$

$$L_{\text{ce}} = - \sum_k \sum_i g_{(k,i)} \log(p_{(k,i)}). \quad (2)$$

使用 λ_{dice} 、 λ_{ce} 作为每个损失期的加权系数时, 总损失额将变为:

$$L_{\text{all}} = \lambda_{\text{dice}} L_{\text{dice}} + \lambda_{\text{ce}} L_{\text{ce}}, \quad (3)$$

其中: N 是像素数, $p_{(k,i)} \in [0,1]$ 和 $g_{(k,i)} \in \{0,1\}$ 分别表示第 k 个类别的预测概率和真实标签; k 是类别数, $\sum_k \omega_k = 1$ 是类权重, 在本文中, 根据以往实验经验地设定了 $\omega_k = 1/k$, 表示为类别数的倒数。

3 实验与结果

3.1 数据集和实验设置

实验选择在杜克大学^[25]和温州医科大学的混合临床视网膜 OCT 数据集上训练和评估上节提出的方法。数据来自谱域 OCT 系统, 系统的参数为中心波长 1300 nm, 带宽 135 nm, 纵向分辨率为 5.5 μm , 横向分辨率为 13 μm , 信噪比 109 dB。该 OCT 数据集包括来自 11 个 OCT 扫描子集, 这些扫描对应的所需标签由两名专家注释。过程中选取前 7 个子集的扫描用于训练, 后选取两个子集用于验证, 最后两个子集的扫描作为测试集, 并将其与之前的相关工作进行了比较。实验基于公共 PyTorch 平台。在训练过程中, 经验地将原始图像裁剪到 224 \times 224 输入进网络以加快训练阶段, 因为收敛速度快的优点, 则选取 ADAM 作为优化器, 批大小为 10, 学习率设置为 0.0005, 权值衰减为 0.0001, 训练的最大周期为 100, 其中训练周期数是根据所有模型的最佳验证精度选择的。经过实验验证将 λ_{dice} 和 λ_{ce} 的值都设置为 1。

3.2 比较方法与评价指标

实验将提出的模型与 U-Net 和专门用于视网膜分层的 RelayNet 进行了比较评估, 所有实验都在同一数据集上进行训练和验证。为了评估网络分割性能, 本实验采用了 Dice 系数^[26]和平均交并比 (mean intersection over union, mIoU)^[27]来评估模型对 OCT 图像分割的性能, 二者都是衡量分割结果和原始标签相似度的度量指标。Dice 相似系数表示预测和真实标准两个样本之间的相似度, 取值范围在 [0,1], Dice 系数越接近 1, 表示模型性能越好。平均交并比 (mIoU) 表示多类分割中每个类别分割正确区域与真实标签、预测值两个集合之比。Dice、mIoU 的计算公式如:

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN+FP+TP}, \quad (4)$$

$$Dice = \frac{2TP}{FP+2TP+FN}, \quad (5)$$

其中: TP (true positive) 表示分割结果中正确分割的目标区域的像素个数; FP (false positive) 表示实际为背景像素但在分割结果中被错误分割为目标区域的像素个数; FN (false negative) 表示目标区域被错误分割为非目标区域的像素个数; k 表示类别的总数。

3.3 实验结果

表 1 显示了三种模型分割结果, 表示不同视网膜层和液体经过实验所获得的分割指标, 整体上本文模型在 Dice 和 mIoU 上都要优于 U-Net、ReLayNet, 就视网膜层而言, 除内界膜的提升较为明显外, 其他指标上的差别并不突出, 这是因为每一层的视网膜面积狭小, 在没有因病变或外力因素所导致特别大的形变下分割出的效果并没有太大差别。而在液体分布范围广且分散的情况下, 本文模型液体分割性能指标 Dice 系数相较于 U-Net, ReLayNet 模型分别提高了 10% 和 34%。在同等硬件条件下, 双编码模型用时比 U-Net 慢了 2 s, 但和 ReLayNet 相比基本差不多。

图 6 是黄斑水肿 (macular edema, ME) OCT 图像, 主要比较了本文模型和 U-Net 的分割生成图, 从视觉上验证了表中数据的真实性, 在本身视网膜结构复杂狭窄的情况下, 双编码器模型分割出的结果都近乎相似, 在区域范围较大、较分散的液体区域, 双编码器又能很好地利用图像的光谱频域信息进行转换, 成功捕捉到遗漏的空间信息, 识别并预测出其所包含的区

表 1 不同方法在数据集上的分割结果
Table 1 Results of each method on dataset

Model	Dice									mIoU
	ILM	NFL-IPL	INL	OPL	ONL-ISM	ISE	OS-RPE	Fluid	Mean	
U-Net	0.84	0.89	0.77	0.76	0.89	0.89	0.89	0.79	0.84	0.74
ReLayNet	0.84	0.85	0.70	0.71	0.87	0.88	0.84	0.55	0.78	0.72
Dual-encoder	0.87	0.89	0.77	0.75	0.90	0.90	0.90	0.89	0.86	0.76

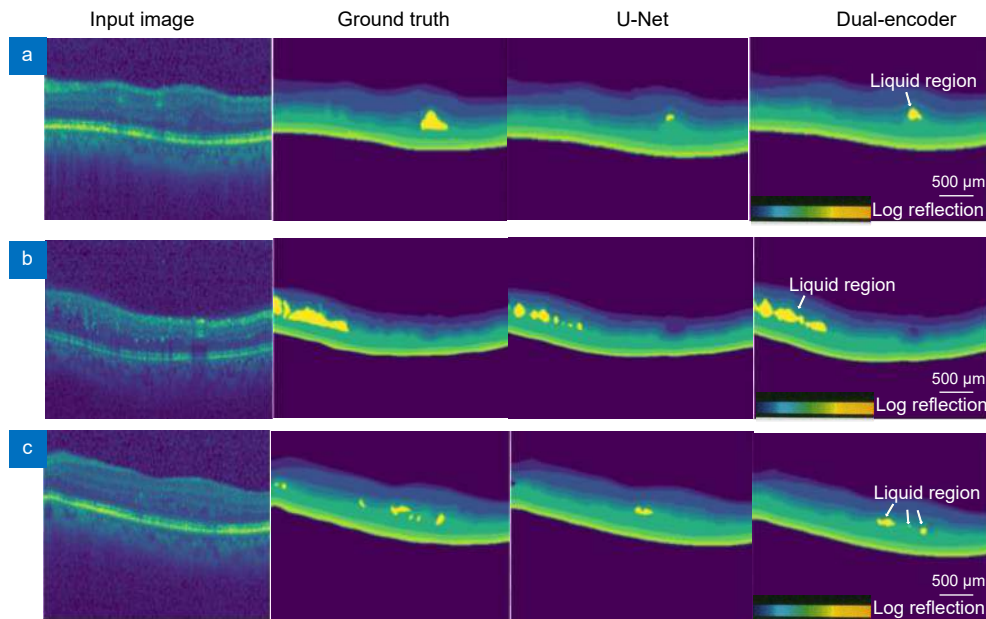


图 6 本文模型与 U-Net 的分割结果比较。(a) 小部分聚集液体区域的分割对比; (b) 长形相连接液体区域的分割对比; (c) 分散无规则分布液体区域的分割对比

Fig. 6 Some qualitative results of ours compared to U-Net. (a) Segmentation and comparison of small areas of liquid accumulation; (b) Segmentation and comparison of long forms and connected liquid regions; (c) Segmentation and comparison of randomly distributed liquid regions

域。本文选取了小部分聚集的液体区域的情况 (对应于图 6(a))、原始标签液体区域呈现为一连串长条形的情况 (对应于图 6(b))、液体区域呈分散且不规则状的情况 (对应于图 6(c)) 作分析, 在图 6(a) OCT 样张中, 对于小部分聚集的液体区域, U-Net 只能识别到上方一角, 而本文模型识别出来的区域要远大于 U-Net。图 6(b) 样张中原始标签液体区域呈现为一连串的长条形, U-Net 分割结果呈不相连的块状区域, 没有将目标区域信息很好地连接, 而本文模型的分割结果没有出现断连, 将液体整体区域都分割出来, 使其更接近原始标签。图 6(c) 样张中的液体区域呈分散、不规则状, 这对于分割任务难度更大, 本文模型较 U-Net 分割出更多的液体区域, 证明双编码器对于细节局部特征的处理更加优秀, 也验证了该方法对于单编码器信息遗漏补全的可行性。结合数据和生成图, 本文模型

的实验结果最接近于专家注释的真实标签, 而 U-Net、ReLayNet 整体效果都不如双编码器的分割效果好, 尤其对于液体分割的结果不太理想, 在某些区域无法正确地分割液体。

3.4 消融实验及不同 α 的对比

如表 2 消融实验所示, 分析了频域 Transformer 中预定值 α 和 FFC-DC 模块对实验结果的影响, 其中表中第一行表示在频域解码器中只存在常规卷积块时的实验分割结果。数据表明, 随着 FFC-DC 模块的加入, 实验结果的 Dice 系数获得了明显提高, 特别是对于视网膜中液体分割的准确度, 这表明频域编码器的设计的确对于空间解码器所遗漏的部分空间信息有着良好的补全效果, 对于分散形状不一的液体区域具有很好的识别。因此有无 FFC-DC 模块, 在液体分割上效果明显。

表 2 关于 FFC-DC 和 α 的消融实验
Table 2 Ablation study on the FFC-DC blocks and α

Method	α	Dice								
		ILM	NFL-IPL	INL	OPL	ONL-ISM	ISE	OS-RPE	Fluid	Mean
-	-	0.86	0.88	0.76	0.77	0.89	0.89	0.89	0.80	0.843
+FFC-DC	0.25	0.86	0.89	0.77	0.76	0.89	0.89	0.89	0.90	0.856
+FFC-DC	0.5	0.87	0.89	0.77	0.75	0.90	0.90	0.90	0.89	0.860
+FFC-DC	0.75	0.86	0.89	0.77	0.76	0.87	0.89	0.89	0.87	0.850
+FFC-DC	1	0.84	0.88	0.76	0.76	0.88	0.89	0.88	0.88	0.846

表 2 探索了实验在不同 α 值下的分割性能, 结果显示, 当 α 值为 0.25 时, 液体分割性能最佳, Dice 系数达到了 90%; 当 α 值为 0.5 时, 整体性能最佳, 模型在液体分割上的 Dice 系数达到了 89%, 而在视网膜层分割上与其他实验组不相上下。实验表明, 局部和全局特征都可通过本文模型学习到有价值的信息。通过设置 α 的值, 所提出的模型能够将全局和局部特征关联起来, 以实现最佳性能。

4 结论与讨论

本文提出了一种基于 U-Net 的端到端双编码器分割模型, 用于分割 OCT 图像中的视网膜层和液体区域。区别于 U-Net 的单一编码器, 主要利用傅里叶卷积的提取和转换特点, 提出频域编码器作为第二个编码器分支, 进而从 OCT 图像中提取光谱频域特征, 包括具有依赖于组织和视网膜层的高频非均匀斑点, 频域中的特征提取使本文的网络能够对每一层中的斑点分布进行建模和学习。频域编码器作为本文模型的第二编码器提取 OCT 光谱特征, 比单编码器的 U-Net、ReLayNet 多了频域特征, 所以在视网膜层的内界膜层得到较好的分割效果, 因为内界膜层面积稍微大一点, 病变影响显著。而液体区域体现了其高光谱频域特征, 特别是液体区域呈分散、不规则状的图像中更加明显, 这在图 6(c) 也得到了很好的验证。

实验将提出的模型与经典模型进行了比较, 结果表明, 本文模型都取得了最优分割效果, 整体的 Dice 系数和 mIoU 相较于 U-Net 均提高 2%, 相较于 ReLayNet 分别提高 8% 和 4%, 尤其在液体分割上其分割 Dice 系数比 U-Net 高出 10%, 比 ReLayNet 高出 34%, Dice 系数达到了 89%, 证明了频域编码器的可行性以及双编码器对于视网膜图像分割的优越性。

参考文献

- [1] Hayreh S S, Zimmerman M B. Fundus changes in branch retinal arteriolar occlusion[J]. *Retina*, 2015, 35(10): 2060–2066.
- [2] Fraz M M, Remagnino P, Hoppe A, et al. Blood vessel segmentation methodologies in retinal images—A survey[J]. *Comput Methods Programs Biomed*, 2012, 108(1): 407–433.
- [3] Li Y Y, Fan J Y, Jiang T L, et al. Review of the development of optical coherence tomography imaging navigation technology in ophthalmic surgery[J]. *Opto-Electron Eng*, 2023, 50(1): 220027.
李云耀, 樊金宇, 蒋天亮, 等. 光学相干层析技术在眼科手术导航方面的研究进展[J]. *光电工程*, 2023, 50(1): 220027.
- [4] Jiang P Q, Wang P H. Research on dispersion compensation technology for SDOCT system[J]. *Opto-Electron Eng*, 2021, 48(10): 210184.
姜盼秋, 汪平河. 谱域光学相干层析系统的色散补偿技术研究[J]. *光电工程*, 2021, 48(10): 210184.
- [5] Huang D, Swanson E A, Lin C P, et al. Optical coherence tomography[J]. *Science*, 1991, 254(5035): 1178–1181.
- [6] Tranos P G, Wickremasinghe S S, Stangos N T, et al. Macular edema[J]. *Surv Ophthalmol*, 2004, 49(5): 470–490.
- [7] Monemian M, Rabbani H. Mathematical analysis of texture indicators for the segmentation of optical coherence tomography images[J]. *Optik*, 2020, 219: 165227.
- [8] Liu Y H, Carass A, He Y F, et al. Layer boundary evolution method for macular OCT layer segmentation[J]. *Biomed Opt Express*, 2019, 10(3): 1064–1080.
- [9] Yang P, Peng Q, Liu W P, et al. A segmentation algorithm of OCT image for macula edema[J]. *J Biomed Eng*, 2011, 28(5): 1001–1006.
杨平, 彭清, 刘维平, 等. 一种眼底黄斑水肿 OCT 图像分割方法[J]. *生物医学工程学杂志*, 2011, 28(5): 1001–1006.
- [10] He J T, Chen M H, Jia W Y, et al. Segmentation of diabetic macular edema in OCT retinal images[J]. *Opto-Electron Eng*, 2018, 45(7): 170605.
何锦涛, 陈明惠, 贾文字, 等. 眼底 OCT 图像中糖尿病性黄斑水肿的分割[J]. *光电工程*, 2018, 45(7): 170605.
- [11] Hussain A, Bhuiyan A, Turpin A, et al. Automatic identification of pathology-distorted retinal layer boundaries using SD-OCT imaging[J]. *IEEE Trans Biomed Eng*, 2017, 64(7): 1638–1649.
- [12] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 2016: 770–778. <https://doi.org/10.1109/CVPR.2016.90>.
- [13] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-

- ResNet and the impact of residual connections on learning[C]//*Proceedings of the 31st AAAI Conference on Artificial Intelligence*, San Francisco, 2017. <https://doi.org/10.1609/aaai.v31i1.11231>.
- [14] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]//*Proceedings of the 3rd International Conference on Learning Representations*, San Diego, 2015.
- [15] Wu Y C, Lin L, Wang J X. Application of semantic segmentation based on convolutional neural network in medical images[J]. *J Biomed Eng*, 2020, 37(3): 533–540.
吴玉超, 林岚, 王婧璇. 基于卷积神经网络的语义分割在医学图像中的应用[J]. *生物医学工程学杂志*, 2020, 37(3): 533–540.
- [16] Qin C B, Song Z Y, Zeng J Y, et al. Deeply supervised breast cancer segmentation combined with multi-scale and attention-residuals[J]. *Opt Precis Eng*, 2021, 29(4): 877–895.
秦传波, 宋子玉, 曾军英, 等. 联合多尺度和注意力-残差的深度监督乳腺癌分割[J]. *光学精密工程*, 2021, 29(4): 877–895.
- [17] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, 2015: 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
- [18] Roy A G, Conjeti S, Karri S P K, et al. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks[J]. *Biomed Opt Express*, 2017, 8(8): 3627–3642.
- [19] Yuan K, Huo L. Multiple-scale inpainting convolutional neural network for retinal OCT image segmentation[J]. *Chin J Lasers*, 2021, 48(15): 1507004.
袁坤, 霍力. 用于视网膜OCT图像分层的多尺度补全卷积神经网络[J]. *中国激光*, 2021, 48(15): 1507004.
- [20] Liu W, Sun Y K, Ji Q G. MDAN-UNet: multi-scale and dual attention enhanced nested U-net architecture for segmentation of optical coherence tomography images[J]. *Algorithms*, 2020, 13(3): 60.
- [21] Ngo L, Cha J, Han J H. Deep neural network regression for automated retinal layer segmentation in optical coherence tomography images[J]. *IEEE Trans Image Process*, 2020, 29: 303–312.
- [22] Chi L, Jiang B R, Mu Y D. Fast Fourier convolution[C]//*Proceedings of the 34th International Conference on Neural Information Processing System*, Vancouver, 2020: 376. <https://dl.acm.org/doi/abs/10.5555/3495724.3496100>.
- [23] Nair V, Chatterjee M, Tavakoli N, et al. Optimizing CNN using fast Fourier transformation for object recognition[C]//*Proceedings of the 19th IEEE International Conference on Machine Learning and Applications*, Miami, 2020: 234–239. <https://doi.org/10.1109/ICMLA51294.2020.00046>.
- [24] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[Z]. arXiv: 1511.07122, 2016. <https://arxiv.org/abs/1511.07122>.
- [25] Chiu S J, Allingham M J, Mettu P S, et al. Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema[J]. *Biomed Opt Express*, 2015, 6(4): 1172–1194.
- [26] Mondal P K, Prodhan U K, Al Mamun S, et al. Segmentation of white blood cells using fuzzy C means segmentation algorithm[J]. *IOSR J Comput Eng*, 2014, 16(3): 1–5.
- [27] Singh V K, Romani S, Rashwan H A, et al. Conditional generative adversarial and convolutional networks for X-ray breast mass segmentation and shape classification[C]//*Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, 2018: 833–840. https://doi.org/10.1007/978-3-030-00934-2_92.

作者简介



【通信作者】陈明惠(1981-), 女, 博士, 副教授, 硕士生导师, 主要从事光学相干层析成像方面的研究。

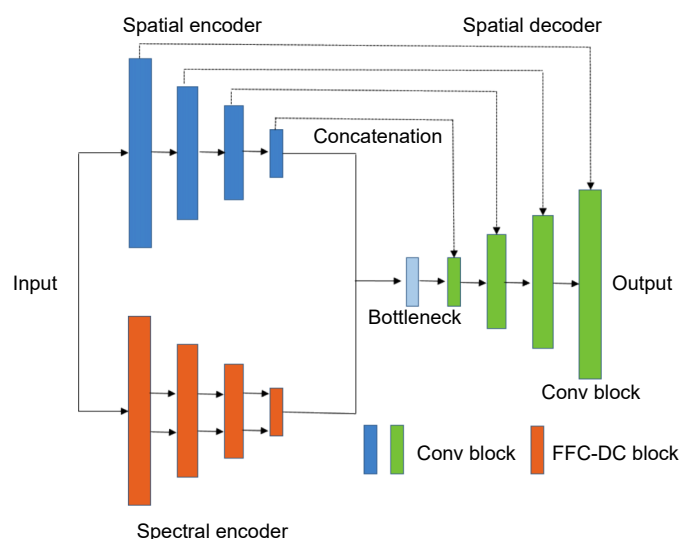
E-mail: cmhui.43@163.com



扫描二维码, 获取PDF全文

Study on retinal OCT segmentation with dual-encoder

Chen Minghui*, Wang Teng, Yuan Yuan, Ke Shuting



The figure of model structure

Overview: Deep learning methods have already had a profound impact on medical image processing. However, some noises and speckles contained in OCT images affect the quality of the images, coupled with the elongated and complex retinal layer and the irregular distribution of pathological fluid in it, which brings great challenges to the automatic segmentation task. At the same time, depending on the limited manpower and time, it is also difficult to sketch a large number of existing images by relying on the professional knowledge of doctors. For the above reasons, automatic medical image segmentation, a scientific medical auxiliary support, is of great clinical significance.

The research content of this paper mainly focuses on the analysis and processing of retinal OCT images. When it comes to monitoring the state of the patient's retina layer, due to the noise and speckle in OCT images and the subtle and complex structure of the retina layer itself, the performance of the model is limited by a single extraction space feature, and the target region cannot be accurately segmented. Aiming at the frequency domain characteristics of OCT images, this paper proposes a dual encoder model based on U-Net and fast Fourier convolution to improve the segmentation performance of the retinal layer and liquid in OCT images. The frequency domain encoder extracts the image frequency domain information and converts it into spatial information by fast Fourier convolution to complete the feature extraction of a single spatial encoder. The experimental results show that the model can effectively improve the segmentation performance of the retinal layer and liquid, both average Dice coefficient and mIoU are increased by 2% compared with U-Net, and the Dice coefficient of liquid is increased by 10%.

Chen M H, Wang T, Yuan Y, et al. Study on retinal OCT segmentation with dual-encoder[J]. *Opto-Electron Eng*, 2023, 50(10): 230146; DOI: [10.12086/oee.2023.230146](https://doi.org/10.12086/oee.2023.230146)

Foundation item: Project supported by Shanghai Science and Technology Commission Industry-university-research Medical Project (15DZ1940400)

Shanghai Engineering Research Center of Interventional Medical, Shanghai Institute for Interventional Medical Devices, School of Health Science and Engineering College of Health Sciences and Engineering, University of Shanghai for Science and Technology, Shanghai 200082, China

* E-mail: cmhui.43@163.com