

# 基于端到端深度神经网络和图搜索的 OCT 图像视网膜层边界分割方法\*



胡凯<sup>1,3</sup>, 蒋帅<sup>1</sup>, 刘冬<sup>1</sup>, 高协平<sup>2</sup>

<sup>1</sup>(智能计算与信息处理教育部重点实验室(湘潭大学), 湖南湘潭 411105)

<sup>2</sup>(智能计算与语言信息处理湖南省重点实验室(湖南师范大学), 湖南长沙 410081)

<sup>3</sup>(医学影像人工智能湖南省重点实验室(湘南学院), 湖南郴州 423000)

通信作者: 高协平, E-mail: [xpgao@xtu.edu.cn](mailto:xpgao@xtu.edu.cn)

**摘要:** 视网膜层边界的形态变化是眼部视网膜疾病出现的重要标志, 光学相干断层扫描 (optical coherence tomography, OCT) 图像可以捕捉其细微变化, 基于 OCT 图像的视网膜层边界分割能够辅助相关疾病的临床判断. 在 OCT 图像中, 由于视网膜层边界的形态变化多样, 其中与边界相关的关键信息如上下文信息和显著性边界信息等对层边界的判断和分割至关重要. 然而已有分割方法缺乏对以上信息的考虑, 导致边界不完整和不连续. 针对以上问题, 提出一种“由粗到细”的基于端到端深度神经网络和图搜索 (graph search, GS) 的 OCT 图像视网膜层边界分割方法, 避免了非端到端方法中普遍存在的“断层”现象. 在粗分割阶段, 提出一种端到端的深度神经网络—注意力全局残差网络 (attention global residual network, AGR-Net), 以更充分和有效的方式提取上述关键信息. 具体地, 首先设计一个全局特征模块 (global feature module, GFM), 通过从图像的 4 个方向扫描以捕获 OCT 图像的全局上下文信息; 其次, 进一步将通道注意力模块 (channel attention module, CAM) 与全局特征模块串行组合并嵌入到主干网络中, 以实现视网膜层及其边界的上下文信息的显著性建模, 有效解决 OCT 图像中由于视网膜层形变和信息提取不充分所导致的误分割问题. 在细分阶段, 采用图搜索算法去除 AGR-Net 粗分割结果中的孤立区域或和孔洞等, 保持边界的固定拓扑结构和连续平滑, 以实现整体分割结果的进一步优化, 为医学临床的诊断提供更完整的参考. 最后, 在两个公开数据集上从不同的角度对所提出的方法进行性能评估, 并与最新方法进行比较. 对比实验结果也表明所提方法在分割精度和稳定性方面均优于现有方法.

**关键词:** OCT 图像; 视网膜层边界分割; 残差神经网络; 注意力; 图搜索

**中图法分类号:** TP391

中文引用格式: 胡凯, 蒋帅, 刘冬, 高协平. 基于端到端深度神经网络和图搜索的 OCT 图像视网膜层边界分割方法. 软件学报, 2024, 35(6): 3036–3051. <http://www.jos.org.cn/1000-9825/6895.htm>

英文引用格式: Hu K, Jiang S, Liu D, Gao XP. Segmentation of Retinal Layer Boundary in OCT Images Based on End-to-end Deep Neural Network and Graph Search. Ruan Jian Xue Bao/Journal of Software, 2024, 35(6): 3036–3051 (in Chinese). <http://www.jos.org.cn/1000-9825/6895.htm>

## Segmentation of Retinal Layer Boundary in OCT Images Based on End-to-end Deep Neural Network and Graph Search

HU Kai<sup>1,3</sup>, JIANG Shuai<sup>1</sup>, LIU Dong<sup>1</sup>, GAO Xie-Ping<sup>2</sup>

<sup>1</sup>(Key Laboratory of Intelligent Computing and Information Processing of Ministry of Education (Xiangtan University), Xiangtan 411105, China)

\* 基金项目: 国家自然科学基金 (62272404, 61972333); 湖南省自然科学基金 (2022JJ30571); 湖南省教育厅优秀青年基金 (21B0172); 长沙市科技重大专项 (kh2202005)

收稿时间: 2022-06-30; 修改时间: 2022-09-23; 采用时间: 2022-12-11; jos 在线出版时间: 2023-06-28

CNKI 网络首发时间: 2023-06-29

<sup>2</sup>(Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing (Hunan Normal University), Changsha 410081, China)

<sup>3</sup>(Key Laboratory of Medical Imaging and Artificial Intelligence of Hunan Province (Xiangnan University), Chenzhou 423000, China)

**Abstract:** The morphological changes in retina boundaries are important indicators of retinal diseases, and the subtle changes can be captured by images obtained by optical coherence tomography (OCT). The retinal layer boundary segmentation based on OCT images can assist in the clinical judgment of related diseases. In OCT images, due to the diverse morphological changes in retina boundaries, the key boundary-related information, such as contexts and saliency boundaries, is crucial to the judgment and segmentation of layer boundaries. However, existing segmentation methods lack the consideration of the above information, which results in incomplete and discontinuous boundaries. To solve the above problems, this study proposes a coarse-to-fine method for the segmentation of retinal layer boundary in OCT images based on the end-to-end deep neural networks and graph search (GS), which avoids the phenomenon of “faults” common in non-end-to-end methods. In coarse segmentation, the attention global residual network (AGR-Net), an end-to-end deep neural network, is proposed to extract the above key information in a more sufficient and effective way. Specifically, a global feature module (GFM) is designed to capture the global context information of OCT images by scanning from four directions of the images. After that, the channel attention module (CAM) and GFM are sequentially combined and embedded in the backbone network to realize saliency modeling of context information of the retina and its boundaries. This effort effectively solves the problem of wrong segmentation caused by retina deformation and insufficient information extraction in OCT images. In fine segmentation, a GS algorithm is adopted to remove isolated areas or holes from the coarse segmentation results obtained by AGR-Net. In this way, the boundary keeps a fixed topology, and it is continuous and smooth, which further optimizes the overall segmentation results and provides a more complete reference for medical clinical diagnosis. Finally, the performance of the proposed method is evaluated from different perspectives on two public datasets, and the method is compared with the latest methods. The comparative experiments show that the proposed method outperforms the existing methods in terms of segmentation accuracy and stability.

**Key words:** optical coherence tomography (OCT) image; segmentation of retinal layer boundary; residual neural network; attention; graph search (GS)

## 1 引言

眼部疾病是人类常见疾病之一,尤其是视网膜疾病,对视功能损害巨大,是影响人类视觉健康的重要因素。随着疾病类型和患者数量的增多,临床诊断往往借助医学影像技术以更高的效率为患者提供诊断结果和治疗建议。光学相干断层扫描(optical coherence tomography, OCT)<sup>[1]</sup>是近年来发展迅速的新型层析成像技术,填补了传统超声和显微镜成像之间的空白<sup>[2]</sup>,且由于其对光具有高敏感度,成像也具有较高分辨率,因此被广泛应用于眼部结构成像及相关疾病的诊断及筛查等<sup>[3-5]</sup>。由于眼部疾病会导致视网膜层结构出现显著性变化,临床医师往往通过观察视网膜 OCT 图像中的层边界来帮助分析眼部疾病的病变程度,通过视网膜层的厚度变化进行定量或定性分析能够辅助诊断青光眼<sup>[6]</sup>、年龄性黄斑病变<sup>[7,8]</sup>和糖尿病视网膜病变<sup>[9]</sup>等视网膜疾病。从图 1 中可以看出,与正常眼部 OCT 图像相比,脉络膜新生血管(choroidal neovascularization, CNV)、糖尿病黄斑水肿(diabetic macular edema, DME)病变的 OCT 图像视网膜层边界发生了显著的形态变化,此类变化会使视网膜层边界变得模糊不清甚至消失,使得视网膜层边界的分割成为医学图像领域的挑战性难题之一。临床上,医师往往手动分割视网膜层边界以辅助眼部疾病的判断,尽管这种方式能获得较好的结果,但由于其受主观因素影响、对人工依赖性较大,导致诊断效率较低。而基于计算机算法的自动分割方法具有精度高、速度快等特点,近年来受到了研究人员的广泛关注<sup>[10,11]</sup>。

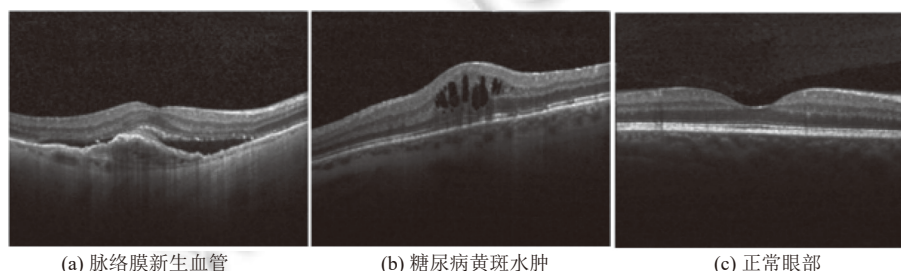


图 1 病变与正常眼部的 OCT 图像

视网膜层自动分割方法大致可以分为两类:传统处理方法和基于机器学习的方法.前者以需要人工提取并选择与任务强相关的图像特征为共同特性,分割结果往往取决于特征选取的恰当性,很大程度上限制了其在 OCT 图像分割领域的有效应用.具体包括基于阈值法<sup>[12-14]</sup>、水平集模型<sup>[15]</sup>、马尔可夫随机场模型<sup>[16]</sup>和基于图论<sup>[17-20]</sup>等传统图像处理方法,以及充分利用机器学习能较好逼近复杂非线性关系并对噪声有较强鲁棒性和容错能力特性而产生的诸如支持向量机<sup>[21]</sup>、随机森林<sup>[22,23]</sup>和 K 近邻<sup>[24]</sup>和神经网络<sup>[25-27]</sup>方法等传统机器学习方法.其中,图搜索(graph search, GS)分割算法近年来表现突出,它将图像映射为加权图,图的顶点对应其连接的像素点或区域,带权边上的权值代表其连接的相邻顶点间的非负相似度.图结构有效建立了图像中像素点或区域之间的相互关系,通过特定的建图方式以分割目标表面(surface)边界来实现目标的准确分割,从原理上保证了分割效果的固定拓扑结构和边界连续平滑,避免了视网膜层的过分割和欠分割现象,因此基于图搜索算法的视网膜分层具有先天的优势.但它存在两个方面的不足,一是由于 OCT 图像分辨率高、影像尺寸大,直接使用图搜索方法存在计算代价大、运算时间过长的缺点,加上 OCT 图像存在大量噪声且对比度低,手动提取的特征无法充分地表征图像信息,对视网膜分割带来挑战;二是图搜索方法大多基于图像强度和梯度来提取边界,受到图像本身视网膜层厚度的先验知识约束,而不同个体的视网膜层厚度或特征不同,利用同一组规则处理不同病例的 OCT 图像数据时可能会出现错误.

近年兴起的深度学习方法<sup>[28]</sup>,摒弃了传统机器学习方法中人工提取图像特征的步骤,主要通过深度神经网络进行自动的特征提取和选择,在 OCT 图像视网膜层分割方面取得了很好进展,受到了专家学者们的青睐.根据模型输入的不同,基于深度学习的方法大致可分为基于图像块的方法<sup>[29,30]</sup>和端到端的方法<sup>[31-36]</sup>.对于基于图像块的方法,其实现方式是在将图像输入到神经网络前对其采集图像块,从而减少迭代过程中的复杂性和对计算资源的需求.相反,端到端的深度学习方法则无需考虑图像块采样,而是对输入的整图信息进行建模.在图像分割的研究中,端到端方法略去了裁剪图像的步骤,以更直接的方式处理输入图像,减少了图像信息丢失的可能.如 McDonough 等人<sup>[26]</sup>提出了端到端的 BP 神经网络使其专注于学习视网膜层的特征;Pekala 等人<sup>[33]</sup>提出了一种卷积神经网络与高斯过程结合的视网膜层分割方法;Kepp 等人<sup>[34]</sup>设计了基于卷积神经网络的形状回归方式以保留视网膜层的拓扑结构;Wei 等人<sup>[35]</sup>提出了一种新的基于深度最大池化的全卷积网络,并将不同视网膜层间的互斥关系引入到损失函数中.目前深度学习方法是 OCT 图像中视网膜层分割的主流算法,主要是基于卷积神经网络(convolutional neural networks, CNNs)或循环神经网络(recurrent neural networks, RNNs)的方法, CNN 和 RNN 本质上是特征提取器,对训练过程自动提取到的图像特征具备很强的表征能力,但由于它们只对各个像素点的结果进行判断,因此存在分割结果的连续性较差,会出现边界断裂、跳变或模糊等情况,也不能保证分割层的拓扑结构,分割结果存在过分割和欠分割等现象.

研究表明,人工设计的先验特征和自动学习获取的深度特征是相辅相成的,二者合理组合应用可以带来更好的分割结果,将传统处理算法与深度学习算法结合是图像分割的一种优化选择<sup>[37]</sup>,如在视网膜层分割任务中,将图搜索算法与深度学习结合则是已被证明的一种有效方式.目前常见的是基于图像块的深度学习和图搜索的结合,如 Fang 等人<sup>[25]</sup>以图像块输入利用 CNN 得到概率图,再使用图搜索算法对视网膜层边界进行搜索;Kugelman 等人<sup>[38]</sup>在 Fang 等人<sup>[25]</sup>所提方法的基础上使用 RNN 代替 CNN;Hu 等人<sup>[39]</sup>提出视网膜层多尺度信息提取的策略,并将其与图搜索算法结合以优化层分割效果;唐艳红等人<sup>[40]</sup>则保留了 CNN 部分,而通过改进图搜索算法以提升视网膜层的分割效果.尽管这些方法已经取得了良好的效果,但在采集图像块的过程中丢失了图像的全局信息和图像块间的位置信息,图搜索算法中施加的表面平滑约束仅考虑到各条边界本身的平滑性要求,而忽略了各边界间的关联性,类似于线间最大、最小距离,或形状相似性等信息.因此视网膜层边界的连续性在取块的过程中实际上被破坏,并且重叠采样也引入了大量冗余的图像信息从而在一定程度上影响分割性能.如对于糖尿病患者,通过这一类方法所得到的分割结果往往会存在如图 2(a)中明显的“断层”现象,即使经过了图搜索算法的平滑依然未得到改善,而对于黄斑变性患者,即使病例种类不同,如图 2(b)中也同样存在类似的现象.

将图搜索与深度学习结合应用于 OCT 分割还存在另外一种方式,即基于端到端深度神经网络和图搜索的

OCT 图像视网膜层边界分割方法. 如 Ben-Cohen 等人<sup>[41]</sup>提出了全卷积网络用于视网膜层分割, 并在后处理的过程中引入了 Sobel 边缘滤波器和图搜索算法; Mishra 等人<sup>[42]</sup>则将经典医学图像分割网络 U-Net<sup>[43]</sup>与最短路径算法结合用于视网膜层的自动分割. 然而, 这些方法所采用的通用医学分割网络由于缺少针对视网膜层的设计而表现出了明显不足的鲁棒性. 如图 2(c) 和图 2(d) 所示, 由于黄斑中心凹附近的视网膜层形变程度大, 呈现出两端宽、中心窄的特点, 处于中间层的边界分割往往会出现错误. 实际上, 对于中间层的边界像素而言, 边界两侧不同视网膜层间的差异性是由于辅助判断其类别的关键信息, 而这些信息往往蕴含于视网膜层边界间的上下文信息中, 因此如何充分提取和利用 OCT 图像中的上下文信息是获得视网膜层边界精准分割的关键. 已有方法缺乏对上述问题的考虑, 分割精度还存在较大提升空间. 综上所述, 本文将端到端的深度学习方法与图搜索算法结合, 提出了一种新颖的由粗到细的基于注意力全局残差网络 (attention global residual network, AGR-Net) 和图搜索的 OCT 图像视网膜层边界分割方法, 有效弥补了已有 OCT 图像视网膜层边界分割方法的不足. 本文主要贡献如下.

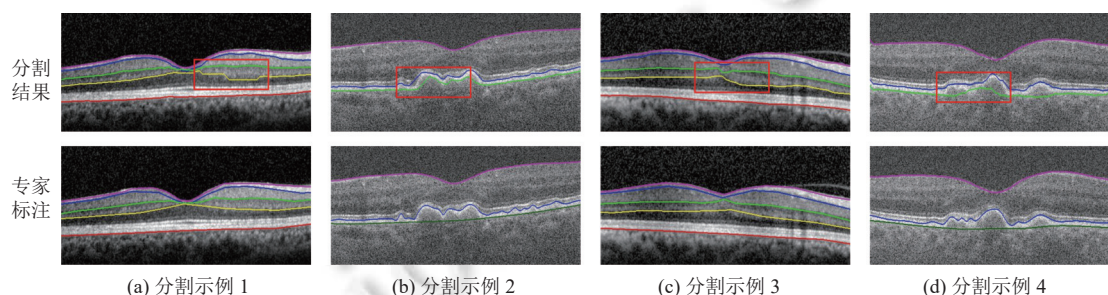


图 2 不同方法对不同患者类型的视网膜层边界分割结果示例

- 提出了一种新颖的“由粗到细”的端到端深度神经网络与图搜索结合的 OCT 图像视网膜层边界分割框架. 不同于以往的方法, 该框架既能充分利用深度学习自动提取复杂特征和有效表征 OCT 图像信息的长处, 又能充分利用图搜索方法从原理上保证了分割结果的固定拓扑结构和边界连续平滑, 从而避免了视网膜层的过分割和欠分割现象, 更能以端到端方式解决原来以图像块方式带来的“断层”现象等弊端.

- 提出了一种基于注意力全局残差网络的端到端深度学习模型, 并将其作为整体框架的粗分割阶段. 该模型以残差块为基础, 利用短连接能够很好避免底层信息的过度丢失, 保护特征提取的完整性. 通过串行组合全局特征模块 (global feature module, GFM) 和通道注意力模块 (channel attention module, CAM) 并将其嵌入到残差网络中, 能够充分提取 OCT 图像的全局上下文信息和图像特征之间的关联性信息, 并同时增强视网膜层边界的强相关信息, 有效地解决了视网膜 OCT 图像由于灰度值与背景十分接近、边界模糊难以定义的问题.

- 在两个公开数据集上开展消融实验验证了本文方法的有效性, 并与多种主流方法展开了对比实验. 实验结果表明本文方法对于 OCT 图像视网膜层边界分割在定量效果和视觉效果上都优于目前最先进的方法.

## 2 本文算法

如图 3 所示为本文提出的视网膜层边界分割算法整体流程图, 整个分割过程分为两个串行阶段: 粗分割阶段 (coarse segmentation stage) 和细分割阶段 (fine segmentation stage). 在粗分割阶段中, 通过利用训练集优化图 3 中 AGR-Net 的参数后将测试集输入网络即得到粗分割结果. 具体地, 本文在残差网络中嵌入了全局特征模块和通道注意力模块, 前者能够充分捕获图像的全局上下文信息, 后者则使得网络更关注于 OCT 图像中视网膜层边界相关的显著性信息. 值得注意的是, 神经网络得到的视网膜层分割结果容易出现层边界不连续的现象, 本文利用图搜索算法进一步对粗分割的结果进行细分割, 以获得连续的视网膜层边界分割结果. 具体地, 本文首先经端到端深度神经网络得到粗分割结果, 再利用粗分割结果生成的候选边界构造出图模型, 并对该生成图进行最短路径搜索, 弥补粗分割结果中存在的边界不连续的问题, 从而优化各视网膜层边界.

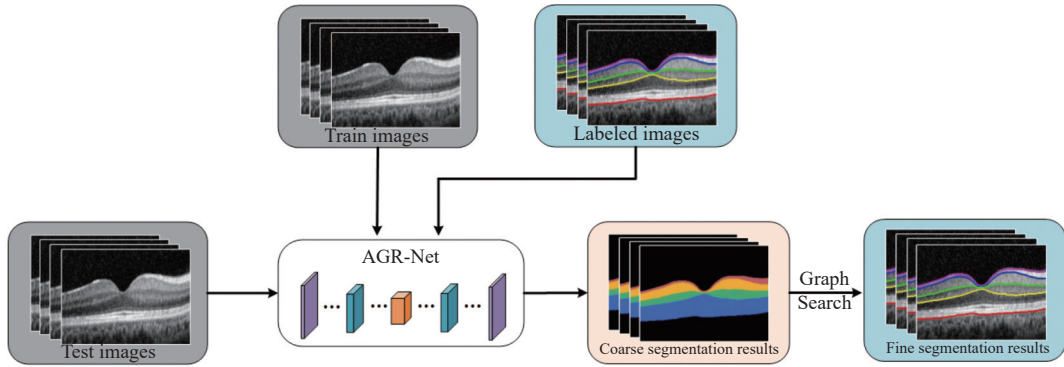


图 3 本文所提的视网膜层分割算法流程图

2.1 数据预处理

由于 OCT 图像之间存在的较大差异性, 在粗分割阶段往往会给神经网络的训练带来问题, 不仅会降低网络的学习速度, 还会影响对数据的拟合程度. 为了解决数据差异性带来的若干问题, 任意样本输入均被归一化处理, 通过归一化不仅避免了数据奇异性问题, 而且能够加快模型的求解速度和预测精度<sup>[44]</sup>.

OCT 图像的获取较困难, 导致用于训练的 OCT 图像数据量相对较少, 而一般情况下, 神经网络参数较多, 从而需要大量的训练数据才能使网络参数正常工作. 由于 OCT 图像中视网膜层边界均为近似水平的走向, 考虑到这一水平特性, 本文通过水平翻转将训练数据进行扩增, 使得训练得到的网络模型具有更强的泛化性.

2.2 粗分割-注意力全局残差网络

如图 4 所示为本文所提出的注意力全局残差网络结构, 其由骨干网络和两个功能模块组成, 其中, 上方为网络的整体架构, (a) 和 (b) 分别为通道注意力模块和全局特征模块的结构细节. 骨干网络包含 4 个阶段, 且均由残差块构成, 两个功能模块分别为全局特征模块和通道注意力模块.

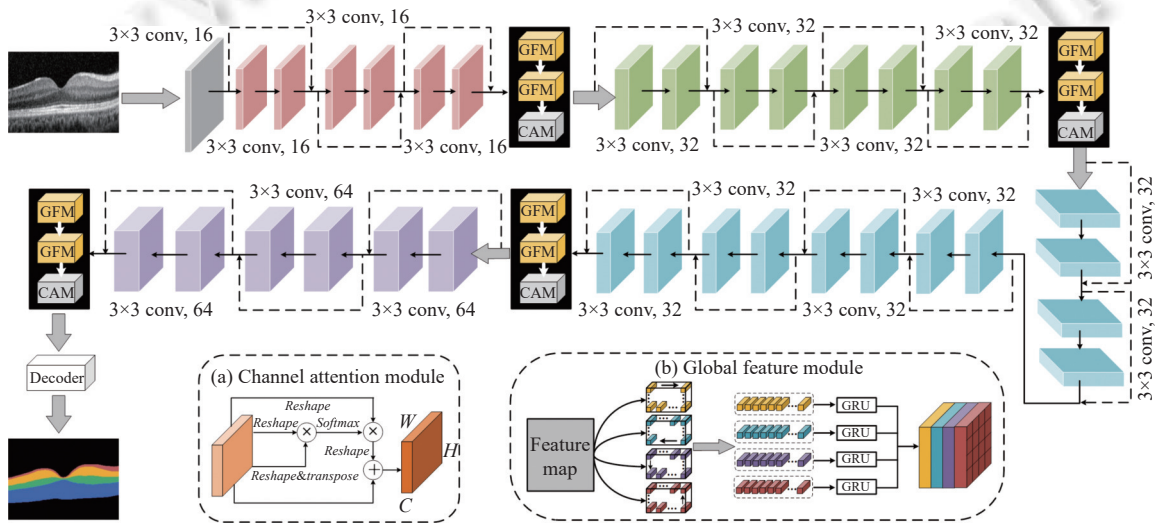


图 4 AGR-Net 的结构图

残差块: 在计算机视觉各种任务中, 为了更好地拟合大批量数据, 神经网络的深度不可避免的加深, 而当深度到达一定程度后, 网络前馈和反馈过程中梯度问题所造成的网络退化问题愈发明显, 导致网络对训练数据的拟合效率会大幅降低. 残差网络<sup>[45]</sup>的提出很好地解决了此类问题, 残差结构中跳跃连接的存在避免了避免梯度消失或

爆炸等问题, 有效地保护了网络的深层特征. 因此, 本文以连续的若干残差块为单个阶段, 并以其为基础构成 AGR-Net 的骨干结构, 能够避免梯度消失或爆炸等一系列由于网络加深所导致的网络退化问题, 且残差块中的恒等映射也极大程度地保护了特征提取的完整性.

**全局特征模块:** 图像信息的充分提取对于特定像素的分类具有关键性的作用, 尤其对于视网膜层边界而言, 判断其属于某一视网膜层结构时则需要结合若干方向临近像素的语义信息, 即上下文信息. 从 OCT 图像中可以发现, 视网膜层往往以水平的方式贯穿全图, 提取图像的全局上下文信息能够更好地建模视网膜层的水平连续性. 考虑到循环神经网络的特征输出具有全局化表示的特点, 其在进行特征提取的过程中, 通过保存每一步的信息以调整下一步, 从而保证均考虑了整个图像的激活. 基于循环神经网络的这一优势, 本文设计了用于提取全局上下文信息的全局特征模块 (GFM). 由于 OCT 图像中的视网膜层在一般情况下并非始终保持水平, 且循环神经网络在提取特征时具有时序性, 因此 GFM 由 4 个门控循环单元 (gated recurrent units, GRU) 构成, 它们分别在水平和垂直上以双向形式对图像进行扫描, 实现从 4 个方向以全局的方式获取视网膜 OCT 图像的上下文信息. 此外, GRU 能够较好地平衡内存使用和计算能力. 具体而言, GFM 中 4 个并行的 GRU 分别处理 4 个扫描方向, 输入的 OCT 图像被转化重组为 4 个不同的  $1 \times 1 \times n$  的序列, 并输入至对应的 GRU 中, 将得到的输出序列转化为 4 张特征图, 理论上每张特征图所包含的特征均不同, 将四者组合便得到了包含全局上下文信息的特征图序列.

**通道注意力模块:** 由于全局特征模块从不同方向对图像信息进行了表示, 经全局特征模块提取后的特征在通道维度必然存在不同的视网膜层相关性. 为了使网络更专注于挖掘视网膜 OCT 图像中与层边界相关的显著性信息, 即充分利用视网膜层结构的强相关特征, 通道注意力模块 (CAM) 通过在通道维度计算并生成通道注意力矩阵, 即生成对应的通道权重概率图, 从而建模任意通道之间的关系. 将该概率图加权至 CAM 的输入特征图上即可得到注意力图, 从而选择性地强调与层边界相关的通道信息并抑制无关的通道信息. 本文将通道注意力模块与连续的两个全局特征模块串行组合并集成到主干网络中以有效地实现视网膜层特征的提取和增强.

图 4 为本文提出的 AGR-Net 的具体细节. 对任意输入的 OCT 图像, 首先通过一个卷积层将其 4 倍下采样, 为了在网络提取图像特征的过程中保留空间信息, 该下采样是 AGR-Net 中唯一降低图像分辨率的步骤. 经下采样的图像再经过网络的 4 个阶段, 这 4 个阶段均由残差块 (每两个卷积层加跳跃连接组成一个残差层)、全局特征模块和通道注意力模块组成, 各个阶段的残差块的数目分别为 3、4、6 和 3, 而全局特征模块和通道注意力模块在 4 个阶段中的数目都被设置为 2 和 1. 值得注意的是, 第 4 阶段中的跳跃连接在每两个残差层之间, 且 GFM 和 CAM 被串行组合嵌入在 4 个阶段中以充分提取 OCT 图像中的层边界特征. 由于在上述特征提取的过程中图像尺寸仅缩小为原图的  $1/4$ , 采样倍数较小, 因此在网络的最后采用转置卷积进行 4 倍上采样后直接输出结果.

### 2.3 细分割-图搜索算法

一般情况下, OCT 图像中的视网膜层边界是近似水平连续的, 即使对于出现视网膜层形变的眼底疾病患者, 连续性也很难被改变. 这类连续性约束在神经网络提取 OCT 图像特征的过程中无法被考虑, 因此视网膜层分割结果中容易出现如图 5(a) 中的小的孤立区域或孔洞, 若直接使用神经网络得到的候选边界作为最终的视网膜层边界, 则会出现如图 5(b) 中边界不连续的情况. 视网膜层的不连续往往以个别孤立区域或孔洞的形式存在, 为了消除该现象, 本文通过图搜索算法对粗分割结果进一步处理, 从而得到层边界连续的细分割结果.

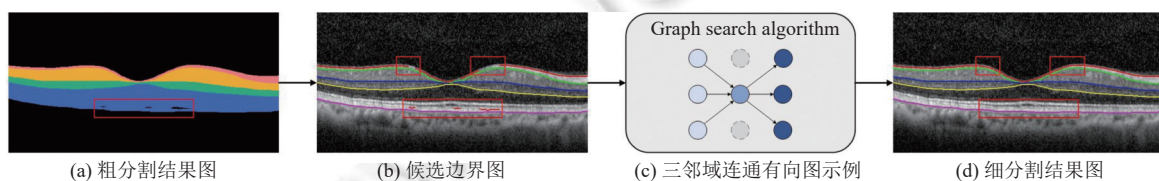


图 5 细分割处理流程

图论中, 有向图和无向图是描述实际问题的一种有效方式, 相比于无向图而言, 有向图的优势在于其连接任意节点的边都具有方向, 能够更有效地表示节点间的不对称关系. 本文在对视网膜层分割结果建立图模型的过程中,

令任意节点在水平方向上实现如图 5(c) 所示的三邻域有向连通, 在充分建立邻域关系的基础上能够极大地提升遍历图的效率. 本文将端到端分割网络与图搜索算法相结合, 通过神经网络得到概率图后, 采用改进的图搜索算法对层边界进行搜索, 使得整体分割框架在充分提取和利用视网膜 OCT 图像的全局特征和边界强相关信息的能够同时考虑连续性约束, 从而实现更精确的视网膜层边界分割.

图搜索算法整体流程如算法 1 所示, 算法主要包括以下 4 个步骤.

**Step 1.** 转换候选边界. 将粗分割结果中的边界像素定义为对应边界的候选像素, 如图 5(b) 所示. 对于粗分割结果中的任一像素, 若其与下方像素对应的视网膜层类别不一致时, 则该像素被认为是候选边界的组成部分.

**Step 2.** 构造有向图模型. 将转换候选边界后的图像构建为如图 5(c) 所示的由节点和连接节点的边组成的一个三邻域连通有向图模型, 其中 1 个节点即为 1 个像素, 每个节点都有其对应的权值.

**Step 3.** 计算有向图中各节点的权值和连接权值. 将 Step 1 中的候选边界像素 (节点) 的权值设为 1, 其余为 0. 有向图的节点的连接权值计算如下:

$$w_{ij} = 2 - (p_i + p_j) + w_{\min} \quad (1)$$

其中,  $p_i$  和  $p_j$  分别表示节点  $i$  和  $j$  的权值,  $w_{ij}$  表示节点  $i$  到  $j$  间的有向边的权重,  $w_{\min}$  为一个小正数, 以防止连接权值为 0. 对于有向图中的任意一条边, 其权重的大小与其连接的两个节点所对应的候选像素的多少成反比.

**Step 4.** 最短路径算法搜索有向图. 利用 Dijkstra 最短路径算法从所确定的有向图中选定顶点出发, 遍历整个图并更新层边界的路径, 最后得到细分割结果.

图搜索算法本身具有最短路径算法的连续性, 能够弥补粗分割结果中存在的视网膜层边界不连续的问题. 本文的图搜索方法的优点在于对边界的优化过程中, 令最短路径尽可能多地包含候选像素, 从而在有效继承神经网络的分割结果的同时使得最终得到的视网膜层边界与真值图更接近.

---

#### 算法 1. 图搜索算法.

---

输入: 粗分割阶段输出的像素集  $V$  及边界点集  $C$  (候选像素).

初始化:

1. 图像像素集合  $V = \{v_1, v_2, \dots, v_k\}_{n \times m}$
2. 候选像素集合  $C = \{c_1, c_2, \dots, c_s\}$
3. 构建三邻域连通有向图  $G = (V, E), E = \{w_{ij}, i, j \in \{1, 2, \dots, k\}\}$
4. 节点权值集合  $P = \{p_1, p_2, \dots, p_k\}$ , 其中  $p_1 = p_2 = \dots = p_k = 0$

图搜索:

5.  $Initial(P) : P = \{p_i = 1 \mid 1 \leq i \leq k, V_i \in C\}$
6. For  $i = 1 \rightarrow k$  and  $j = 1 \rightarrow k$  Do
7.  $w_{ij} \leftarrow 2 - (p_i + p_j) + w_{\min}$
8. End For
9.  $S = \emptyset$  // 已经确定路径的节点集合
10. While  $V \neq \emptyset$
11.  $u \leftarrow \min Distance(E, V)$  // 从边集中选择出与节点  $V$  最小距离的顶点  $u$
12.  $S \leftarrow S \cup \{u\}$  // 将顶点  $u$  加入  $S$
13. For each vertex  $v \in Neighbor(u)$  Do
14.  $Update(u, v, e)$  // 对每一个与顶点  $u$  相邻的节点更新一次路径, 更新最短路径
15. End For
16. End While

输出:  $Path(G, S)$ .

---

### 3 实验与分析

#### 3.1 数据集介绍

数据集 1: U. of Miami OCT 数据集<sup>[46]</sup>, 由迈阿密巴斯科姆帕尔默眼科研究所采集, 包含 10 例轻度非增殖性糖尿病视网膜病变病例的 50 幅 OCT 图像, 其中患者年龄区间为 45–59 岁. 对于每个病例, 采集了患者视网膜黄斑中央凹中心位置图像 1 幅、中央凹周围图像 2 幅和中央凹旁图像 2 幅, 共 5 幅分辨率大小为 496×768 的 OCT 图像, 每幅图像的垂直分辨率为 11.11 μm/像素, 水平分辨率为 3.867 μm/像素. 所有采集到的视网膜 OCT 图像中的 5 层边界 (Surface1、Surface2、Surface4、Surface6 和 Surface11) 由两位眼科专家同时进行标注.

数据集 2: 该数据集来自于 4 家不同的眼科研究中心, 由 Chiu 等人<sup>[47]</sup>所在机构的 SD-OCT 成像系统采集, 包含 20 例年龄相关性黄斑变性 (AMD) 病例的 220 幅 OCT 图像. 其中每个病例包含 11 幅包含中心凹的 OCT 图像 (编号分别为  $F \pm n$ , 其中  $n$  等于 0、2、5、10、15 和 20). 所有采集到的视网膜 OCT 图像中 3 层边界 (脉络膜的外边界 (RBC)、视网膜色素上皮的内边界 (IZ-RPE) 和内界膜的内边界 (ILM-NFL)) 同样由两位眼科专家同时进行标注.

#### 3.2 实验方案

本文通过消融实验和对比实验的方式, 将从 4 个方面验证所提注意力全局残差网络及其组件的有效性, 以及其与图搜索算法结合的分割性能, 具体包括: (1) 全局特征模块的消融实验. 通过对比是否加入全局特征模块的实验结果, 以验证本文设计的全局特征模块是否有助于视网膜层边界分割. (2) 通道注意力的消融实验. 通过验证在网络框架中嵌入通道注意力模块能否进一步优化模型的分割性能. (3) 图搜索算法的消融实验. 通过比较粗分割和细分割的结果以验证图搜索算法的有效性. (4) 与已有方法的对比实验. 通过在两个数据集上从数值结果和视觉效果方面与主流方法进行对比, 以验证本文方法的优越性. 此外, 针对不同病例的 OCT 图像, 本文对其分割结果进行了分析.

在实验中, 本文采用水平翻转进行了数据增强, 并采用 5 折交叉验证进行性能评估. 其中数据集 1 选择 8 个病例 (40 幅图像) 作为训练集, 其余 2 个病例 (10 幅图像) 为测试集; 数据集 2 使用 16 个病例 (176 幅图像) 进行训练, 其余 4 个病例 (44 幅图像) 作为测试集. 数据集 1 和 2 均有两位专家的手动标注, 本文以各自的第 1 位专家标注为金标准, 并将两位专家的标注差异 (Inter-Observer) 作为判断分割精度的基本标准.

#### 3.3 评价指标

为了更直接公平地对比, 本文在数据集 1 和 2 上采用其最常用的指标作为模型分割性能的评价标准, 即在数据集 1 上采用如公式 (2) 所示的平均无符号误差 (mean unsigned error,  $MUE$ ), 在数据集 2 上采用如公式 (3) 所示的均方根误差 (root mean square error,  $RMSE$ ), 以上评价指标计算模型的预测边界与真实边界的误差的方式如下:

$$MUE_i = \frac{\sum_x |P_i(x) - G_i(x)|}{\|A\text{-scan}\|} \quad (2)$$

$$RMSE_i = \sqrt{\frac{\sum_x (P_i(x) - G_i(x))^2}{\|A\text{-scan}\|}} \quad (3)$$

其中,  $x$  表示  $x$  轴向扫描 ( $A\text{-scan}$ ),  $i$  表示第  $i$  个边界,  $P(x)$  和  $G(x)$  分别表示第  $x$  个  $A\text{-scan}$  图像的模型预测图和真值图.  $\|A\text{-scan}\|$  表示  $A\text{-scan}$  的图像总数.

#### 3.4 对比实验

##### 3.4.1 全局特征模块的消融实验

由第 3.2 节可知, 全局特征模块由 4 个并行的门控循环单元构成, 为了验证全局特征模块的有效性, 本文在两个数据集上进行了该模块的消融实验. 实验结果如表 1 和表 2 所示, 在使用全局特征模块时, 除数据集 2 上的 RBC 层边界误差稍微升高了 0.03 个像素以外, 其他边界的误差指标都有大幅度的下降. 总之, 当不集成全局特征



模块时, 视网膜层边界分割结果存在较大的误差, 而将全局特征模块嵌入网络后时, 视网膜层边界误差大幅减少. 全局特征模块通过利用门控循环单元, 一方面在水平方向上提取各个视网膜层自身的特征以建模层的水平连续性, 另一方面在垂直方向上提取不同视网膜层间的判别性特征, 通过融合两类特征以有效地进行视网膜层特征提取. 全局特征模块从水平和垂直两个方向对图像的全局上下文信息进行整合, 而不仅受限于局部特征的提取, 充分利用了图像各像素之间的关联性. 在两个数据集上的数值结果均表明本文设计的全局特征模块能够有效提取和建模 OCT 图像的全局信息, 从而提高视网膜层边界分割的分割性能.

表 1 在数据集 1 上对 GFM 和 CAM 的消融实验结果

Network	Mean	Std	Max	Surface1	Surface2	Surface4	Surface6	Surface11
Baseline	1.021	0.137	1.243	0.897	1.014	1.088	1.243	0.864
Baseline+GFM	0.914	<b>0.103</b>	<b>1.080</b>	0.769	0.956	<b>0.901</b>	<b>1.080</b>	0.862
Baseline+GFM+CAM	<b>0.911</b>	0.112	1.084	<b>0.768</b>	<b>0.955</b>	0.940	1.084	<b>0.811</b>

表 2 在数据集 2 上对 GFM 和 CAM 的消融实验结果

Network	Overall	RBC	IZ-RPE	ILM-NFL
Baseline	1.93±0.98	<b>1.66±0.84</b>	2.73±1.02	1.39±0.38
Baseline+GFM	1.73±0.96	1.69±1.19	<b>2.28±0.80</b>	1.22±0.40
Baseline+GFM+CAM	<b>1.72±0.97</b>	1.68±1.20	2.31±0.79	<b>1.16±0.29</b>

### 3.4.2 通道注意力模块的消融实验

为了验证通道注意力的引入是否有利于提升分割结果, 本文同样进行了有无通道注意力模块的对比实验. 在数据集 1 上对通道注意力模块的消融实验结果如表 1 所示, 当嵌入 CAM 后, 在 *MUE* 指标方面, 除 Surface4 和 Surface6 的误差略微上升外, 其他边界上的误差及平均误差均有不同幅度的降低. 从表 2 中对通道注意力模块的消融实验结果可以看到, 除 IZ-RPE 层边界的误差上升了 0.03 个像素外, 其他的边界误差均有不同程度的下降, 尤其是整体方面有了更小的误差. 综上所述, 通道注意力的引入能够使网络有选择性地强调信息特征, 过滤掉经 GFM 提取到的特征可能包含的少量冗余和无关信息, 从而有效抑制与层边界无关的通道信息. 在两个数据集上的消融实验结果也显示了通道注意力模块的嵌入对于边界分割有一定的提升效果, 而 GFM 和 CAM 二者的组合使用可以使得全局特征的建模更充分和有效.

### 3.4.3 图搜索算法的消融实验

由于粗分割结果得到的边界可能存在不连续的情况, 本文采用 Dijkstra 算法对设计的三连通有向图进行搜索以进行细分割. 表 3 和表 4 分别为两个数据集上有无图搜索算法时的分割结果, 可以看出, 图搜索算法有效地降低了粗分割结果中的误差. 图 6 和图 7 分别展示了在两个数据集上引入图搜索算法前后的可视化结果对比. 从中可以发现, 视网膜层不连续的现象被有效地消除, 尤其如图 6(b) 和图 7(c) 所示, 其中多处边界不连续在经过图搜索算法后均被平滑处理. 本文从以上两个角度验证了细分割阶段的图搜索算法对视网膜层边界分割具有显著的正向作用.

表 3 在数据集 1 上对图搜索算法的消融实验结果

Method	Mean	Std	Max	Surface1	Surface2	Surface4	Surface6	Surface11
AGR-Net (Ours)	0.97	0.17	1.14	0.78	1.09	1.10	1.14	<b>0.76</b>
AGR-Net-GS (Ours)	<b>0.91</b>	<b>0.11</b>	<b>1.08</b>	<b>0.77</b>	<b>0.96</b>	<b>0.94</b>	<b>1.08</b>	0.81

表 4 在数据集 2 上对图搜索算法的消融实验结果

Method	Overall	RBC	IZ-RPE	ILM-NFL
AGR-Net (Ours)	1.76±0.94	1.77±0.94	<b>2.26±0.76</b>	1.24±0.41
AGR-Net-GS (Ours)	<b>1.72±0.97</b>	<b>1.68±1.20</b>	2.31±0.79	<b>1.16±0.29</b>

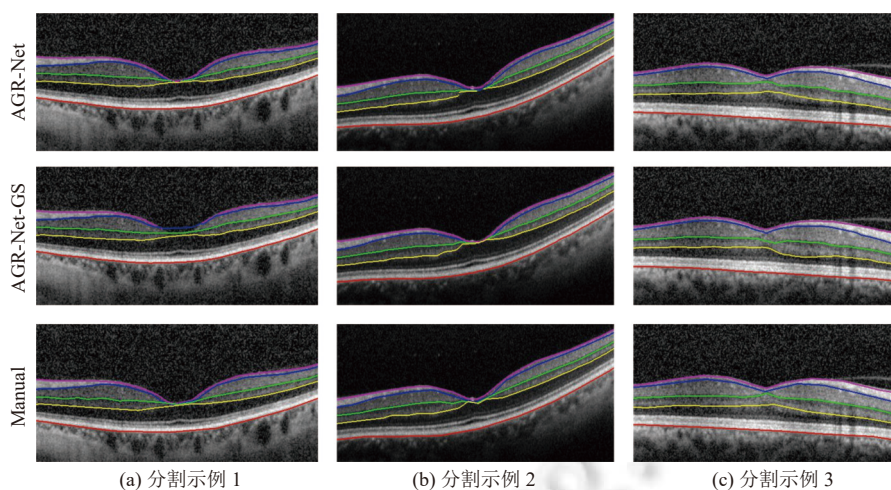


图 6 在数据集 1 上对图搜索算法的消融实验可视化结果

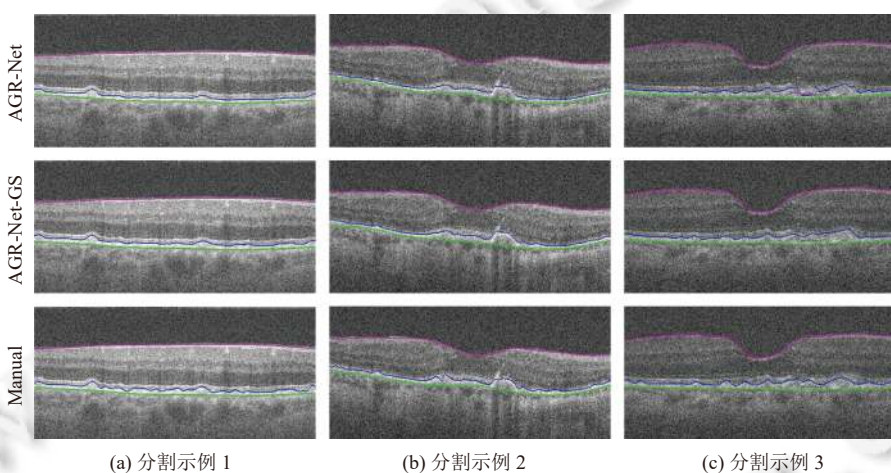


图 7 在数据集 2 上对图搜索算法的消融实验可视化结果

#### 3.4.4 与已有方法的对比实验

为了对比各种分割方法的分割性能,本文在两个数据集上与现有已有方法进行了对比试验,实验结果分别见表 5 和表 6。除分割结果外,为了衡量模型的推理速度,本文还计算了各模型的计算操作总数,即浮点运算数 (floating point operations, FLOPs),模型的 FLOPs 越小,计算量越低,则表示模型的推理速度越快。

表 5 展示了本文的方法与数据集 1 上的 14 种主流方法的对比结果,其中 Inter-Observer 表示两位专家标注的差异。除 IOWA<sup>[48]</sup>, Spectralis<sup>[49]</sup>和 AURA<sup>[41]</sup>的结果由对应开发的软件获得之外的对比方法的结果均经过实验得到,其中 Bern<sup>[50]</sup>, OCTRIMA<sup>[17]</sup>和 Pekala 等人<sup>[33]</sup>的结果摘自对应的论文原文,其余方法的结果为本文复现。为了保证比较的公平性,所有对比方法在实验过程中均使用了和本文方法相同的预处理和数据增强方法。从表 5 中可以看出,本文方法在 5 个边界指标中的 4 个都实现了最小误差,仅在 Surface4 指标上比最好结果略微差了 0.02 个像素,而其他边界误差均有大幅度的下降,平均误差也下降至 0.91 个像素。图 8 展示了本文方法与已有方法的可视化分割效果对比,可以看到,本文方法的分割结果更接近专家人工标注的结果。特别是与 DCNN-GS<sup>[25]</sup>、MCNN-GS<sup>[39]</sup>这类基于图像块的深度学习模型和图搜索算法结合的方法对比,对于糖尿病患者视网膜层存在的轻微形变,AGR-GS 的分割结果更加平滑,不存在明显的“断层”现象,而与 U-Net-GS<sup>[42]</sup>这类端到端方法的分割结果相比,AGR-GS 在处理形变的边界时也更加精确。

表 5 在数据集 1 上与已有方法的 *MUE* 数值结果比较

Method (Year)	Mean	Std	Max	Surface1	Surface2	Surface4	Surface6	Surface11	FLOPs (G)
Inter-Observer	1.10	0.15	1.29	0.87	1.14	1.10	1.29	1.12	—
Bern (2013) <sup>[50]</sup>	1.81	0.57	2.77	1.71	2.77	1.60	1.72	1.24	—
OCTRIMA (2015) <sup>[17]</sup>	1.17	0.23	1.52	0.95	1.18	0.99	1.52	1.20	—
IOWA (2015) <sup>[48]</sup>	1.66	0.30	2.03	2.03	1.74	1.79	1.51	1.22	—
DFCN-GS (2015) <sup>[51]</sup>	1.44	0.24	1.76	1.11	1.65	1.76	1.42	1.25	362.16
Spectralis (2016) <sup>[49]</sup>	1.33	0.37	1.92	1.09	1.45	1.92	1.19	0.99	—
AURA (2017) <sup>[41]</sup>	1.23	0.22	1.54	1.35	1.19	1.12	1.54	0.96	—
DCNN-GS (2017) <sup>[25]</sup>	1.12	0.20	1.45	0.85	1.45	1.05	1.20	1.02	148.33
MCNN-GS (2019) <sup>[39]</sup>	0.99	<b>0.10</b>	1.13	0.82	1.05	1.00	1.13	0.96	145.25
Pekala 等人 (2019) <sup>[33]</sup>	1.06	<b>0.10</b>	1.19	1.10	1.06	<b>0.92</b>	1.19	1.02	—
U-Net-GS (2020) <sup>[42]</sup>	0.97	0.12	1.16	0.80	1.04	<b>0.92</b>	1.16	0.90	364.39
LFU-Net (2021) <sup>[52]</sup>	1.48	0.63	2.41	0.85	1.59	2.41	1.61	0.98	253.74
CS <sup>2</sup> -Net (2021) <sup>[53]</sup>	1.34	0.37	1.66	0.96	1.66	1.41	1.55	1.13	68.94
MAG-Net (2022) <sup>[54]</sup>	1.49	0.59	2.12	1.05	1.48	2.12	1.89	0.93	<b>19.95</b>
GLFR-Net (2022) <sup>[55]</sup>	1.20	0.37	1.41	0.82	1.36	1.41	1.28	1.13	60.17
AGR-Net (Ours)	0.97	0.17	1.14	0.78	1.09	1.10	1.14	<b>0.76</b>	87.76
AGR-Net-GS (Ours)	<b>0.91</b>	0.11	<b>1.08</b>	<b>0.77</b>	<b>0.96</b>	0.94	<b>1.08</b>	0.81	232.87

表 6 在数据集 2 上与已有方法的 *RMSE* 数值结果比较

Method (Year)	Overall	RBC	IZ-RPE	ILM-NFL	FLOPs (G)
Inter-Observer	2.19±1.00	1.93±0.76	2.96±1.04	1.66±0.64	—
Chiu 等人 (2012) <sup>[47]</sup>	1.93±0.98	1.66±0.84	2.73±1.02	1.39±0.38	—
DFCN-GS (2015) <sup>[51]</sup>	2.56±2.05	2.94±1.83	2.63±0.91	2.12±2.84	311.42
DCNN-GS (2017) <sup>[25]</sup>	2.33±1.88	3.12±2.76	2.61±1.02	1.27±0.40	94.90
Hussain 等人 (2017) <sup>[18]</sup>	2.37±1.65	2.07±1.81	3.49±1.46	1.54±0.83	—
MCNN-GS (2019) <sup>[39]</sup>	2.12±1.46	2.64±2.05	2.48±0.93	1.25±0.35	94.51
U-Net-GS (2020) <sup>[42]</sup>	1.76±1.91	1.90±1.07	2.21±0.71	1.18±0.34	313.65
LFU-Net (2021) <sup>[52]</sup>	2.19±2.18	2.84±3.50	<b>2.08±0.60</b>	1.66±1.01	253.74
CS <sup>2</sup> -Net (2021) <sup>[53]</sup>	2.09±1.57	1.92±2.30	2.23±0.64	2.10±1.11	56.36
MAG-Net (2022) <sup>[54]</sup>	2.39±1.55	2.65±1.37	2.52±0.53	2.01±2.20	<b>16.44</b>
GLFR-Net (2022) <sup>[55]</sup>	2.04±0.60	<b>1.60±1.83</b>	2.37±0.87	2.16±1.67	60.17
AGR-Net (Ours)	1.76±0.94	1.77±0.94	2.26±0.76	1.24±0.41	72.30
AGR-Net-GS (Ours)	<b>1.72±0.97</b>	1.68±1.20	2.31±0.79	<b>1.16±0.29</b>	166.67

表 6 展示了本文的方法与数据集 2 上的 10 种主流方法以及两位专家标注的对比结果。其中 Chiu 等人<sup>[47]</sup>和 Hussain 等人<sup>[18]</sup>所提方法的结果直接摘抄自对应已发表的论文中。与数据集 1 一致, 为了更为充分地验证本文方法的有效性, 本文在数据集 2 上同样实现了其他的对比方法。由表 6 可以发现, 在整体误差评价指标 Overall 上的结果是所有对比方法中最小的, 在其中的 3 条边界上都取得了较小的误差, 尤其是 ILM-NFL 层边界。整体而言, 本文的注意力全局残差网络与图搜索联合的方法对于数据集 2 有比较好的分割效果。图 9 为不同分割方法的视觉效果对比。从图 9 中可以看到, 在黄斑中心凹位置附近, DCNN-GS<sup>[25]</sup>和 MCNN-GS<sup>[39]</sup>方法由于其存在对图像取块的原因导致 RBC 层边界出现了变形, 在真实边界和预测边界之间存在较大的差异。类似地, U-Net-GS<sup>[42]</sup>在 RBC 层边界上也有一定的错误, 但偏差略小于 DCNN-GS<sup>[25]</sup>和 MCNN-GS<sup>[39]</sup>等, 且不同于上述两种方法的是, 该方法出现边界弯曲变形的原因是边界像素的类别判断所依赖的图像上下文信息提取不够充分, 从而导致位于不同视网膜层间的像素类别判断错误。相反, 如图 9 所示, 本文方法在充分利用所提取的上下文信息并建模视网膜层间差异性的情

况下, 对于层边界的分割更加准确, 特别是 RBC 层边界. 从表 5 和表 6 中可以看出, GLFR-Net<sup>[55]</sup>和 MAG-Net<sup>[54]</sup>等基于深度神经网络的单阶段方法普遍具有较大的边界分割误差, 其中 GLFR-Net<sup>[55]</sup>通过重建全局和局部特征实现特征提取的方案在没有图搜索算法的情况下实现了小幅度的误差下降. 本文所提出的 AGR-Net 利用 GRU 提取全局特征的方式考虑到了视网膜层自身水平特性, 在推理速度稍下降的情况下极大地提高了分割精度, 且在两个数据集上较基于深度神经网络的单阶段方法均具有更小的误差. 更进一步地, 当引入图搜索算法后, 虽计算量稍有增加, 但分割误差被进一步降低.

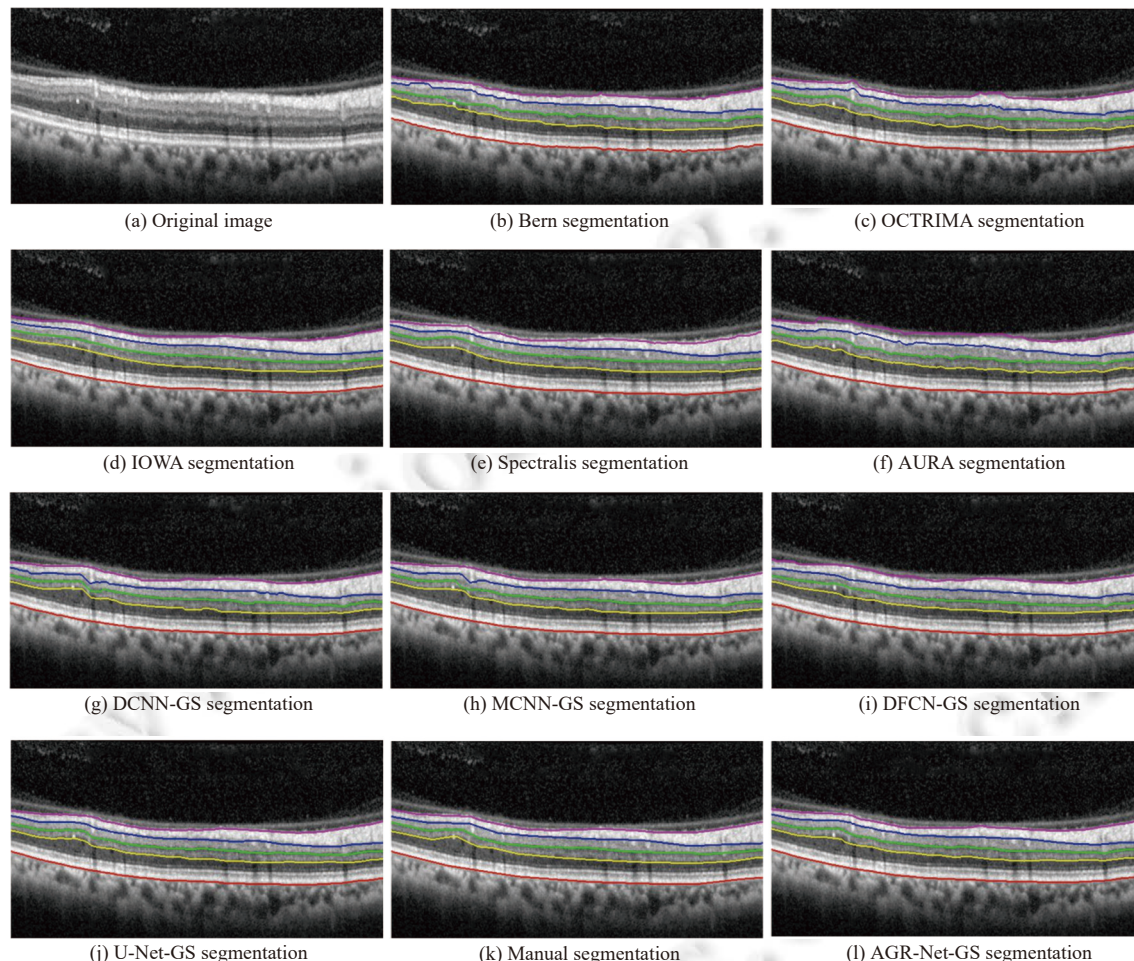


图 8 在数据集 1 上与已有方法的分割效果的比较

### 3.4.5 病例分析

不同疾病情形给算法在层边界分割方面带来了不同程度的困难与挑战, 为了本文对两个数据集进行了案例分析, 其可视化结果如图 10 所示. 由第 3.1 节可知, 数据集 1 和 2 分别包含不同疾病患者的视网膜 OCT 图像, 其中数据集 1 中均为轻度非增殖性糖尿病视网膜病变的患者, 而数据集 2 中均为年龄相关性黄斑变性的患者. 由图 10 可知, 黄斑变性患者的视网膜层变形程度与糖尿病患者相比更加严重. 如图 10(a) 所示, 对于糖尿病病变造成的视网膜层轻微变形, 导致了 AGR-Net-GS 的分割结果与真实边界存在较小的差异, 而对于如图 10(b) 所示的 AMD 造成的严重视网膜病变情形, 由于病变引起的层边界严重变形和不均匀, AGR-Net-GS 在提取全局特征时受到了更多的病变影响, 从而导致分割结果会存在些微的偏差, 但依然能够非常接近真实边界. 综合实验数据分析, 本文方法在这两个具有不同病变情况的数据集上均有着不错的表现, 这也表明本文所提方法对不同疾病所造成的视网膜病变均有一定的适用性.

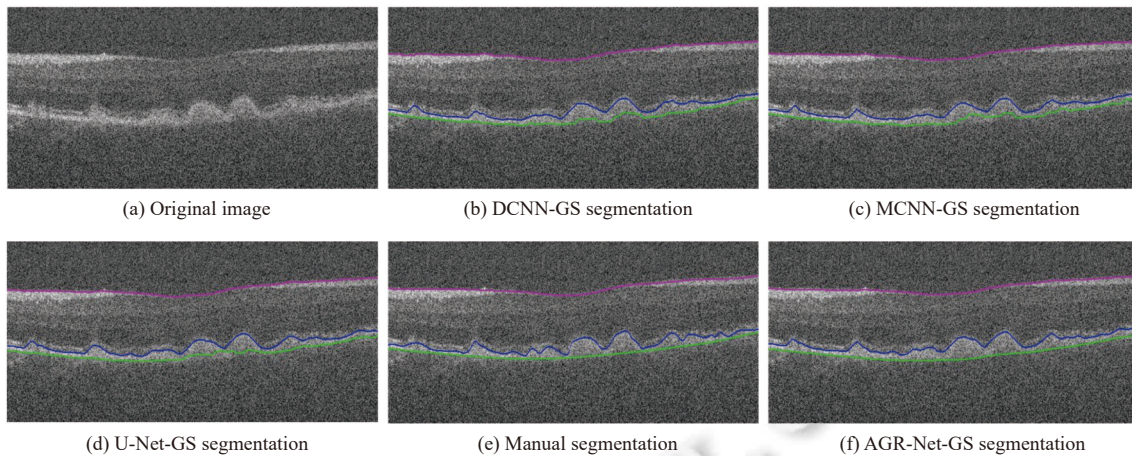


图 9 在数据集 2 上与已有方法的分割效果的比较

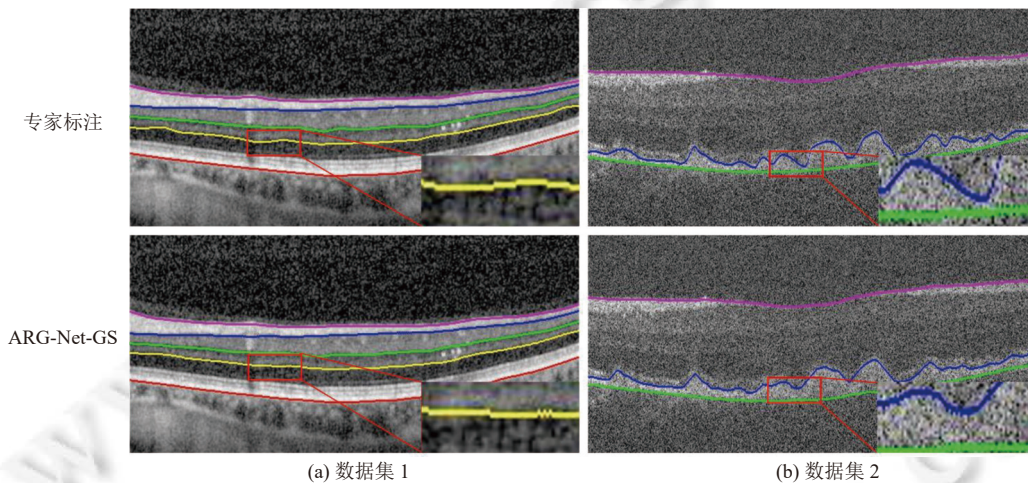


图 10 数据集 1 和数据集 2 中视网膜病变病例的分割结果图

## 4 总结

针对 OCT 视网膜层边界分割问题, 本文提出了一种新颖的“由粗到细”的基于注意力全局残差网络和图搜索的分割方法. 该方法包括基于端到端深度神经网络的粗分割和基于图搜索的细分割两个阶段. 具体地, 首先设计了一个全局特征模块并将其嵌入残差网络中以获取全局特征信息. 其次, 在主干网络中融入注意力机制捕获 OCT 图像中的目标信息. 最后, 采用图搜索方法去除网络粗分割结果其中的孤立区域和填充孔洞, 得到更精细的分割结果. 实验结果表明本文方法在两个主流的公开数据集上都取得了最佳的分割效果, 数值结果和视觉效果均优于目前的 OCT 视网膜层边界分割方法.

## References:

- [1] Fercher AF, Hitzenberger C, Juchem M. Measurement of intraocular optical distances using partially coherent laser light. *Journal of Modern Optics*, 1991, 38(7): 1327–1333. [doi: 10.1080/09500349114551491]
- [2] Li M, Landahl S, East AR, Verboven P, Terry LA. Optical coherence tomography—A review of the opportunities and challenges for postharvest quality evaluation. *Postharvest Biology and Technology*, 2019, 150: 9–18. [doi: 10.1016/j.postharvbio.2018.12.005]
- [3] Tearney GJ, Brezinski ME, Bouma BE, Boppart SA, Pitris C, Southern JF, Fujimoto JG. *in vivo* endoscopic optical biopsy with optical coherence tomography. *Science*, 1997, 276(5321): 2037–2039. [doi: 10.1126/science.276.5321.2037]

- [4] Drexler W, Fujimoto JG. Optical Coherence Tomography: Technology and Applications. Cham: Springer Int'l Publishing, 2015. 3–64. [doi: [10.1007/978-3-319-06419-2](https://doi.org/10.1007/978-3-319-06419-2)]
- [5] Fujimoto J, Swanson E. The development, commercialization, and impact of optical coherence tomography. *Investigative Ophthalmology & Visual Science*, 2016, 57(9): OCT1–OCT13. [doi: [10.1167/iovs.16-19963](https://doi.org/10.1167/iovs.16-19963)]
- [6] Puliafito CA, Hee MR, Lin CP, Reichel E, Schuman JS, Duker JS, Izatt JA, Swanson EA, Fujimoto JG. Imaging of macular diseases with optical coherence tomography. *Ophthalmology*, 1995, 102(2): 217–229. [doi: [10.1016/S0161-6420\(95\)31032-9](https://doi.org/10.1016/S0161-6420(95)31032-9)]
- [7] Keane PA, Liakopoulos S, Jivrajka RV, Chang KT, Alasil T, Walsh AC, Sadda SR. Evaluation of optical coherence tomography retinal thickness parameters for use in clinical trials for neovascular age-related macular degeneration. *Investigative Ophthalmology & Visual Science*, 2009, 50(7): 3378–3385. [doi: [10.1167/iovs.08-2728](https://doi.org/10.1167/iovs.08-2728)]
- [8] Malamos P, Ahlers C, Mylonas G, Schütze C, Deak G, Ritter M, Sacu S, Schmidt-Erfurth U. Evaluation of segmentation procedures using spectral domain optical coherence tomography in exudative age-related macular degeneration. *Retina*, 2011, 31(3): 453–463. [doi: [10.1097/IAE.0b013e3181eef031](https://doi.org/10.1097/IAE.0b013e3181eef031)]
- [9] Bavinger JC, Dunbar GE, Stem MS, Blachley TS, Kwark L, Farsiu S, Jackson GR, Gardner TW. The effects of diabetic retinopathy and pan-retinal photocoagulation on photoreceptor cell function as assessed by dark adaptometry. *Investigative Ophthalmology & Visual Science*, 2016, 57(1): 208–217. [doi: [10.1167/iovs.15-17281](https://doi.org/10.1167/iovs.15-17281)]
- [10] Sarhan MH, Nasser MA, Zapp D, Maier M, Lohmann CP, Navab N, Eslami A. Machine learning techniques for ophthalmic data processing: A review. *IEEE Journal of Biomedical and Health Informatics*, 2020, 24(12): 3338–3350. [doi: [10.1109/JBHI.2020.3012134](https://doi.org/10.1109/JBHI.2020.3012134)]
- [11] Ngo L, Yih G, Ji S, Han JH. A study on automated segmentation of retinal layers in optical coherence tomography images. In: Proc. of the 4th Int'l Winter Conf. on Brain-computer Interface (BCI). Gangwon: IEEE, 2016. 1–2. [doi: [10.1109/IWW-BCI.2016.7457465](https://doi.org/10.1109/IWW-BCI.2016.7457465)]
- [12] Ishikawa H, Stein DM, Wollstein G, Beaton S, Fujimoto JG, Schuman JS. Macular segmentation with optical coherence tomography. *Investigative Ophthalmology & Visual Science*, 2005, 46(6): 2012–2017. [doi: [10.1167/iovs.04-0335](https://doi.org/10.1167/iovs.04-0335)]
- [13] He QY, Li ZL, Wang XC, Nan N, Lu Y. Automated retinal layer segmentation based on optical coherence tomographic images. *Acta Optica Sinica*, 2016, 36(10): 1011003 (in Chinese with English abstract). [doi: [10.3788/AOS201636.1011003](https://doi.org/10.3788/AOS201636.1011003)]
- [14] Naz S, Akram MU, Khan SA. Automated segmentation of retinal layers from OCT images using structure tensor and kernel regression+GTD approach. In: Proc. of the 1st Int'l Conf. on Next Generation Computing Applications (NextComp). Mauritius: IEEE, 2017. 98–102. [doi: [10.1109/NEXTCOMP.2017.8016182](https://doi.org/10.1109/NEXTCOMP.2017.8016182)]
- [15] Liu YH, Carass A, He YF, Antony BJ, Filippatou A, Saidha S, Solomon SD, Calabresi PA, Prince JL. Layer boundary evolution method for macular OCT layer segmentation. *Biomedical Optics Express*, 2019, 10(3): 1064–1080. [doi: [10.1364/BOE.10.001064](https://doi.org/10.1364/BOE.10.001064)]
- [16] El Tanboly A, Ismail M, Switala A, Mahmoud M, Soliman A, Neyer T, Palacio A, Hadayer A, El-Azab M, Schaal S, El-Baz A. A novel automatic segmentation of healthy and diseased retinal layers from OCT scans. In: Proc. of the 2016 IEEE Int'l Conf. on Image Processing (ICIP). Phoenix: IEEE, 2016. 116–120. [doi: [10.1109/ICIP.2016.7532330](https://doi.org/10.1109/ICIP.2016.7532330)]
- [17] Tian J, Varga B, Somjai GM, Lee WH, Smiddy WE, DeBuc DC. Real-time automatic segmentation of optical coherence tomography volume data of the macular region. *PLoS One*, 2015, 10(8): e0133908. [doi: [10.1371/journal.pone.0133908](https://doi.org/10.1371/journal.pone.0133908)]
- [18] Hussain MA, Bhuiyan A, Turpin A, Luu CD, Smith RT, Guymer RH, Kotagiri R. Automatic identification of pathology-distorted retinal layer boundaries using SD-OCT imaging. *IEEE Trans. on Biomedical Engineering*, 2017, 64(7): 1638–1649. [doi: [10.1109/TBME.2016.2619120](https://doi.org/10.1109/TBME.2016.2619120)]
- [19] Niu SJ, Chen Q, Lu ST, Shen HL. SD-OCT image layer segmentation using multi-scale 3-D graph search method. *Computer Science*, 2015, 42(9): 272–277 (in Chinese with English abstract). [doi: [10.11896/j.issn.1002-137X.2015.9.053](https://doi.org/10.11896/j.issn.1002-137X.2015.9.053)]
- [20] Stankiewicz A, Marciniak T, Dabrowski A, Stopa M, Rakowicz P, Marciniak E. Novel full-automatic approach for segmentation of epiretinal membrane from 3D OCT images. In: Proc. of the 2017 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA). Poznan: IEEE, 2017. 100–105. [doi: [10.23919/SPA.2017.8166846](https://doi.org/10.23919/SPA.2017.8166846)]
- [21] Vermeer KA, Van der Schoot J, Lemij HG, De Boer JF. Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images. *Biomedical Optics Express*, 2011, 2(6): 1743–1756. [doi: [10.1364/BOE.2.001743](https://doi.org/10.1364/BOE.2.001743)]
- [22] Lang A, Carass A, Hauser M, Sotirchos ES, Calabresi PA, Ying HS, Prince JL. Retinal layer segmentation of macular OCT images using boundary classification. *Biomedical Optics Express*, 2013, 4(7): 1133–1152. [doi: [10.1364/BOE.4.001133](https://doi.org/10.1364/BOE.4.001133)]
- [23] Lang A, Carass A, Bittner AK, Ying HS, Prince JL. Improving graph-based OCT segmentation for severe pathology in Retinitis Pigmentosa patients. In: Proc. of the 2017 SPIE Medical Imaging Conf. on Biomedical Applications in Molecular, Structural, and Functional Imaging. Orlando: SPIE, 2017. 101371M. [doi: [10.1117/12.2254849](https://doi.org/10.1117/12.2254849)]
- [24] Nath SS, Anoop BK, Sankar P. Classification of outer retinal layers based on KNN-classifier. In: Proc. of the 2018 Int'l Conf. on Emerging Trends and Innovations in Engineering and Technological Research (ICETIETR). Ernakulam: IEEE, 2018. 1–4. [doi: [10.1109/](https://doi.org/10.1109/)]

- [ICETIETR.2018.8529098]
- [25] Fang LY, Cunefare D, Wang C, Guymer RH, Li ST, Farsiu S. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. *Biomedical Optics Express*, 2017, 8(5): 2732–2744. [doi: [10.1364/BOE.8.002732](https://doi.org/10.1364/BOE.8.002732)]
- [26] McDonough K, Kolmanovsky I, Glybina IV. A neural network approach to retinal layer boundary identification from optical coherence tomography images. In: *Proc. of the 2015 IEEE Conf. on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. Niagara Falls: IEEE, 2015. 1–8. [doi: [10.1109/CIBCB.2015.7300299](https://doi.org/10.1109/CIBCB.2015.7300299)]
- [27] Yang JD, Tao YH, Xu QZ, Zhang YH, Ma X, Yuan ST, Chen Q. Self-supervised sequence recovery for semi-supervised retinal layer segmentation. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26(8): 3872–3883. [doi: [10.1109/JBHI.2022.3166778](https://doi.org/10.1109/JBHI.2022.3166778)]
- [28] Yin XH, Wang YC, Li DY. Suvery of medical image segmentation technology based on U-Net structure improvement. *Ruan Jian Xue Bao/Journal of Software*, 2021, 32(2): 519–550 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6104.htm> [doi: [10.13328/j.cnki.jos.006104](https://doi.org/10.13328/j.cnki.jos.006104)]
- [29] Hamwood J, Alonso-Caneiro D, Read SA, Vincent SJ, Collins MJ. Effect of patch size and network architecture on a convolutional neural network approach for automatic segmentation of OCT retinal layers. *Biomedical Optics Express*, 2018, 9(7): 3049–3066. [doi: [10.1364/BOE.9.003049](https://doi.org/10.1364/BOE.9.003049)]
- [30] Hassan T, Usman A, Akram MU, Masood MF, Yasin U. Deep learning based automated extraction of intra-retinal layers for analyzing retinal abnormalities. In: *Proc. of the 20th IEEE Int'l Conf. on e-Health Networking, Applications and Services (Healthcom)*. Ostrava: IEEE, 2018. 1–5. [doi: [10.1109/HealthCom.2018.8531198](https://doi.org/10.1109/HealthCom.2018.8531198)]
- [31] Liu YY, Chen M, Ishikawa H, Wollstein G, Schuman JS, Rehg JM. Automated macular pathology diagnosis in retinal OCT images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding. *Medical Image Analysis*, 2011, 15(5): 748–759. [doi: [10.1016/j.media.2011.06.005](https://doi.org/10.1016/j.media.2011.06.005)]
- [32] Loo J, Fang LY, Cunefare D, Jaffe GJ, Farsiu S. Deep longitudinal transfer learning-based automatic segmentation of photoreceptor ellipsoid zone defects on optical coherence tomography images of macular telangiectasia type 2. *Biomedical Optics Express*, 2018, 9(6): 2681–2698. [doi: [10.1364/BOE.9.002681](https://doi.org/10.1364/BOE.9.002681)]
- [33] Pekala M, Joshi N, Liu TYA, Bressler NM, DeBuc DC, Burlina P. Deep learning based retinal OCT segmentation. *Computers in Biology and Medicine*, 2019, 114: 103445. [doi: [10.1016/j.combiomed.2019.103445](https://doi.org/10.1016/j.combiomed.2019.103445)]
- [34] Kepp T, Ehrhardt J, Heinrich MP, Hüttmann G, Handels H. Topology-preserving shape-based regression of retinal layers in oct image data using convolutional neural networks. In: *Proc. of the 16th IEEE Int'l Symp. on Biomedical Imaging (ISBI 2019)*. Venice: IEEE, 2019. 1437–1440. [doi: [10.1109/ISBI.2019.8759261](https://doi.org/10.1109/ISBI.2019.8759261)]
- [35] Wei H, Peng P. The segmentation of retinal layer and fluid in SD-OCT images using mutex Dice loss based fully convolutional networks. *IEEE Access*, 2020, 8: 60929–60939. [doi: [10.1109/ACCESS.2020.2983818](https://doi.org/10.1109/ACCESS.2020.2983818)]
- [36] Wang B, Wei W, Qiu S, Wang SP, Li D, He HG. Boundary aware U-Net for retinal layers segmentation in optical coherence tomography images. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(8): 3029–3040. [doi: [10.1109/JBHI.2021.3066208](https://doi.org/10.1109/JBHI.2021.3066208)]
- [37] Wang D, Shan SG, Zhang HM, Zeng W, Chen XL. Coarse-to-fine hair segmentation method. *Ruan Jian Xue Bao/Journal of Software*, 2013, 24(10): 2391–2404 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4423.htm> [doi: [10.3724/SP.J.1001.2013.04423](https://doi.org/10.3724/SP.J.1001.2013.04423)]
- [38] Kugelman J, Alonso-Caneiro D, Read SA, Vincent SJ, Collins MJ. Automatic segmentation of OCT retinal boundaries using recurrent neural networks and graph search. *Biomedical Optics Express*, 2018, 9(11): 5759–5777. [doi: [10.1364/BOE.9.005759](https://doi.org/10.1364/BOE.9.005759)]
- [39] Hu K, Shen BW, Zhang Y, Cao CH, Xiao F, Gao XP. Automatic segmentation of retinal layer boundaries in OCT images using multiscale convolutional neural network and graph search. *Neurocomputing*, 2019, 365: 302–313. [doi: [10.1016/j.neucom.2019.07.079](https://doi.org/10.1016/j.neucom.2019.07.079)]
- [40] Tang YH, Chen YZ, Liu MD, Zeng YG, Zhou YX. Segmentation of retinal layers in OCT images based on CNN and improved graph search. *Laser & Optoelectronics Progress*, 2020, 57(24): 241702 (in Chinese with English abstract). [doi: [10.3788/LOP57.241702](https://doi.org/10.3788/LOP57.241702)]
- [41] Ben-Cohen A, Mark D, Kovler I, Zur D, Barak A, Iglicki M, Soferman R. Retinal layers segmentation using fully convolutional network in OCT images. 2017. <https://www.rsipvision.com/wp-content/uploads/2017/06/Retinal-Layers-Segmentation.pdf>
- [42] Mishra Z, Ganegoda A, Selicha J, Wang ZY, Satta SR, Hu ZH. Automated retinal layer segmentation using graph-based algorithm incorporating deep-learning-derived information. *Scientific Reports*, 2020, 10(1): 9541. [doi: [10.1038/s41598-020-66355-5](https://doi.org/10.1038/s41598-020-66355-5)]
- [43] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Proc. of the 18th Int'l Conf. on Medical Image Computing and Computer-assisted Intervention*. Munich: Springer, 2015. 234–241. [doi: [10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)]
- [44] Sola J, Sevilla J. Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Trans. on Nuclear Science*, 1997, 44(3): 1464–1468. [doi: [10.1109/23.589532](https://doi.org/10.1109/23.589532)]

- [45] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016. 770–778. [doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90)]
- [46] Tian J, Varga B, Tatrai E, Fanni P, Somfai GM, Smiddy WE, Debuc DC. Performance evaluation of automated segmentation software on optical coherence tomography volume data. Journal of Biophotonics, 2016, 9(5): 478–489. [doi: [10.1002/jbio.201500239](https://doi.org/10.1002/jbio.201500239)]
- [47] Chiu SJ, Izatt JA, O’Connell RV, Winter KP, Toth CA, Farsiu S. Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. Investigative Ophthalmology & Visual Science, 2012, 53(1): 53–61. [doi: [10.1167/iops.11-7640](https://doi.org/10.1167/iops.11-7640)]
- [48] Lee K, Abramoff M, Garvin M, Sonka M. The IOWA reference algorithms. 2014. <https://iibi.uiowa.edu/oct-reference>
- [49] Heidelberg Engineering. SPECTRALIS HRA+OCT user manual software version 6.0. 2015. <https://www.heidelbergengineering.com/>
- [50] Dufour PA, Ceklic L, Abdillahi H, Schroder S, De Dzanet S, Wolf-Schnurrbusch U, Kowal J. Graph-based multi-surface segmentation of OCT data using trained hard and soft constraints. IEEE Trans. on Medical Imaging, 2013, 32(3): 531–543. [doi: [10.1109/TMI.2012.2225152](https://doi.org/10.1109/TMI.2012.2225152)]
- [51] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015. 3431–3440. [doi: [10.1109/CVPR.2015.7298965](https://doi.org/10.1109/CVPR.2015.7298965)]
- [52] Ma D, Lu DH, Chen S, Heisler M, Dabiri S, Lee S, Lee H, Ding GW, Sarunic MV, Beg MF. LF-UNet—A novel anatomical-aware dual-branch cascaded deep neural network for segmentation of retinal layers and fluid from optical coherence tomography images. Computerized Medical Imaging and Graphics, 2021, 94: 101988. [doi: [10.1016/j.compmedimag.2021.101988](https://doi.org/10.1016/j.compmedimag.2021.101988)]
- [53] Mou L, Zhao YT, Fu HZ, Liu YH, Cheng J, Zheng YL, Su P, Yang JL, Chen L, Frangi AF, Akiba M, Liu J. CS<sup>2</sup>-Net: Deep learning segmentation of curvilinear structures in medical imaging. Medical Image Analysis, 2021, 67: 101874. [doi: [10.1016/j.media.2020.101874](https://doi.org/10.1016/j.media.2020.101874)]
- [54] Cazañas-Gordón A, da Silva Cruz LA. Multiscale Attention Gated Network (MAGNet) for retinal layer and macular cystoid edema segmentation. IEEE Access, 2022, 10: 85905–85917. [doi: [10.1109/ACCESS.2022.3198657](https://doi.org/10.1109/ACCESS.2022.3198657)]
- [55] Song JH, Chen XJ, Zhu QL, Shi F, Xiang DH, Chen ZY, Fan Y, Pan LJ, Zhu WF. Global and local feature reconstruction for medical image segmentation. IEEE Trans. on Medical Imaging, 2022, 41(9): 2273–2284. [doi: [10.1109/TMI.2022.3162111](https://doi.org/10.1109/TMI.2022.3162111)]

#### 附中文参考文献:

- [13] 贺琪欲, 李中梁, 王向朝, 南楠, 卢宇. 基于光学相干层析成像的视网膜图像自动分层方法. 光学学报, 2016, 36(10): 1011003. [doi: [10.3788/AOS201636.1011003](https://doi.org/10.3788/AOS201636.1011003)]
- [19] 牛四杰, 陈强, 陆圣陶, 沈宏烈. 应用多尺度三维图搜索的SD-OCT图像层分割方法. 计算机科学, 2015, 42(9): 272–277. [doi: [10.11896/j.issn.1002-137X.2015.9.053](https://doi.org/10.11896/j.issn.1002-137X.2015.9.053)]
- [28] 殷晓航, 王永才, 李德英. 基于U-Net结构改进的医学影像分割技术综述. 软件学报, 2021, 32(2): 519–550. <http://www.jos.org.cn/1000-9825/6104.htm> [doi: [10.13328/j.cnki.jos.006104](https://doi.org/10.13328/j.cnki.jos.006104)]
- [37] 王丹, 山世光, 张洪明, 曾炜, 陈熙霖. 一种由粗到细的头发分割方法. 软件学报, 2013, 24(10): 2391–2404. <http://www.jos.org.cn/1000-9825/4423.htm> [doi: [10.3724/SP.J.1001.2013.04423](https://doi.org/10.3724/SP.J.1001.2013.04423)]
- [40] 唐艳红, 陈允照, 刘明迪, 曾亚光, 周月霞. 基于CNN和改进的图搜索分割OCT图像中的视网膜层. 激光与光电子学进展, 2020, 57(24): 241702. [doi: [10.3788/LOP57.241702](https://doi.org/10.3788/LOP57.241702)]



胡凯(1984—), 男, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为机器学习, 模式识别, 生物信息学, 医学图像处理.



刘冬(1996—), 女, 硕士生, 主要研究领域为深度学习, 医学图像处理.



蒋帅(1999—), 男, 硕士生, 主要研究领域为深度学习, 医学图像处理.



高协平(1965—), 男, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为小波分析, 神经网络, 生物信息学, 图像处理, 计算机网络.