

分类号\_\_\_\_\_

编 号\_\_\_\_\_

U D C\_\_\_\_\_

密 级\_\_\_\_\_



南方科技大学  
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

# 本科生毕业设计（论文）

题 目：\_\_\_\_\_ 基于超分辨的轻量 OCT 分割算法 \_\_\_\_\_

姓 名：\_\_\_\_\_ 张佳意 \_\_\_\_\_

学 号：\_\_\_\_\_ 11911537 \_\_\_\_\_

系 别：\_\_\_\_\_ 计算机科学与工程系 \_\_\_\_\_

专 业：\_\_\_\_\_ 智能科学与技术 \_\_\_\_\_

指导教师：\_\_\_\_\_ 刘江 \_\_\_\_\_

2023 年 5 月 8 日

# 诚信承诺书

1. 本人郑重承诺所呈交的毕业设计（论文），是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料均真实可靠。
2. 除文中已经注明引用的内容外，本论文不包含任何其他人或集体已经发表或撰写过的作品或成果。对本论文的研究作出重要贡献的个人和集体，均已在文中以明确的方式标明。
3. 本人承诺在毕业论文（设计）选题和研究内容过程中没有抄袭他人研究成果和伪造相关数据等行为。
4. 在毕业论文（设计）中对侵犯任何方面知识产权的行为，由本人承担相应的法律责任。

作者签名：张佳意

2023 年 5 月 8 日

# 基于超分辨的轻量 OCT 分割算法

张佳意

(计算机科学与工程系 指导教师：刘江)

**[摘要]**：一些基于深度学习的自动化 OCT 分割算法工作已经取得了较大的突破，极大提高了眼科疾病早期检测的辅助诊断效率。然而，这些突破和改进通常是以增加模型复杂性和降低计算效率为代价的，不适用于实时医疗应用场景及移动医疗设备的工业生产。因此，本文提出一个用于在边缘计算设备部署的 OCT 分割算法，该算法在 U-Net 的编码-解码的结构上通过分组卷积降低计算量，使得模型能够在准确率和速度上做出权衡，并通过引入超分辨辅助的方法在降低输入分辨率以减少计算量的同时保证模型的精度。在 AROI 数据集上的实验证明，本文提出的模型能够相对于 U-Net 的 3% 的计算量的情况下实现 OCT 分割，并取得了极具竞争力的结果。

**[关键词]**：OCT 分割；超分辨；语义分割；多任务

**[ABSTRACT]:** Current automated OCT segmentation algorithms based on deep learning have made great breakthroughs, which greatly improved the auxiliary diagnosis efficiency of early detection of ophthalmic diseases. However, these breakthroughs and improvements usually come at the cost of increasing model complexity and reducing computational efficiency, which are not suitable for real-time medical applications and industrial production of mobile medical devices. Therefore, this work proposes an OCT segmentation algorithm for deployment in edge computing devices, which reduces the computational amount by applying group convolution on the encoding-decoding structure of U-Net, so that the model can make trade-offs in accuracy and speed and ensures the accuracy of the model while reducing the input resolution by introducing super-resolution assistance to reduce the calculation amount. Experiments on the AROI dataset show that the proposed model can complete OCT segmentation with 3% of the computation amount of U-Net and achieves highly competitive results.

**[Key words]:** OCT segmentation; Super-resolution; Semantic Segmentation; Multi-task

# 目录

1. 引言 .....	1
2. 相关工作 .....	5
2.1 基于深度学习的自动 OCT 图像分割网络 .....	5
2.2 基于深度学习的超分辨率网络 .....	5
2.3 基于深度学习的轻量分割网络 .....	5
3. 方法 .....	6
3.1 分割网络 .....	6
3.1.1 3x3 卷积替换 .....	6
3.1.2 Axial-Attention .....	7
3.2 超分辨率网络模型引入 .....	8
3.2.1 超分辨率网络 .....	8
3.2.2 超分辨率网络引入方式 .....	9
3.2.3 融合模块 .....	9
3.2.4 损失函数 .....	11
4. 实验 .....	13
4.1 数据集 .....	13
4.2 模型评价标准 .....	13
4.3 实验结果 .....	14
4.3.1 分割网络对比实验 .....	14
4.3.2 消融实验 .....	15

4.3.3 与现有分割网络对比 ..... 16

5. 总结 ..... 18

参考文献 ..... 19

致谢 ..... 22

# 1. 引言

光学相干断层扫描技术（Optical Coherence Tomography, OCT）是一种广泛应用于临床视网膜诊断的非侵入性成像技术，能够获取眼底视网膜的高分辨率横截面图像（如图 1 所示），常用于辅助诊断年龄相关性黄斑变性（Age-related Macular Degeneration, AMD）、糖尿病性黄斑水肿（Diabetic Macular Edema, DME）和糖尿病性视网膜病变（Diabetic Retinopathy, DR）等眼科疾病。

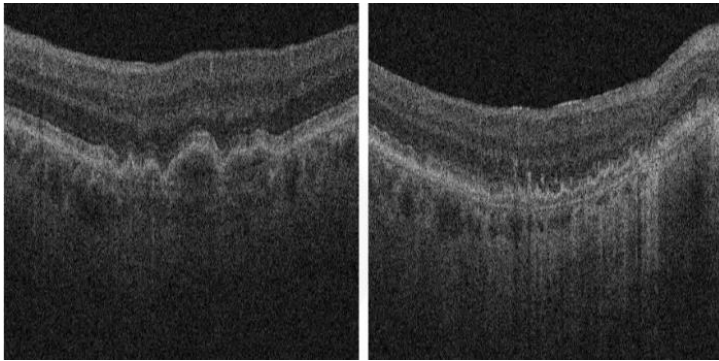


图 1 OCT 图像示意图

在眼科学中，由于不同的眼科疾病以不同的形式影响各视网膜层，使用特异性和灵敏的方法分析视网膜形态结构将有助于眼科疾病的早期检测，例如，糖尿病黄斑水肿（DME）通常会导致视网膜内积液（Intraretinal Fluid, IRF）（如图 2 所示）；图 3 展示 OCT 图像分层，其中视网膜纤维层（Retinal Fiber Layer, RNFL）变薄可能与多发性硬化症有关。因此，准确地分割 OCT 图像中的视网膜层有助于辅助眼科疾病诊断。

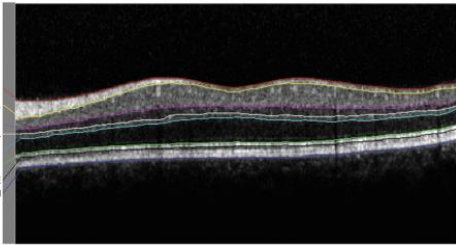
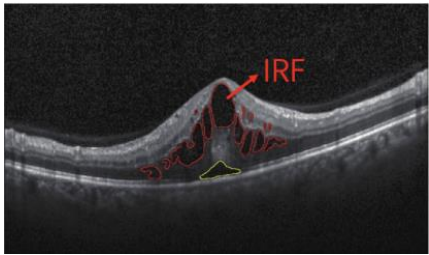


图 2 糖尿病黄斑水肿（DME）导致视网膜内积液（IRF）      图 3 OCT 图像分层示意图

然而，传统分割方法采用医学专家人工手动标注来进行，是非常耗时和低效的。而相比于人工手动 OCT 分割，自动化 OCT 分割算法极大的提高了辅助诊断的效率。自动化 OCT 分割方面的一些现有工作已经取得了较大的突破，尤其是基于深度学习的网络模型工作。经过调研，现有基于深度学习的自动化 OCT 分割算法大多根据 OCT 图像分割难点对经典分割网络进行改进，例如 U-Net<sup>[1]</sup>，SegNet<sup>[3]</sup>，FCN-

DenseNet<sup>[4]</sup>，DeepLab<sup>[5]</sup>等经典分割网络。最常见的是以 U-Net 网络作为网络主体进行改进以应用于 OCT 图像分割，主要存在对 U-Net 网络结构进行微调、在 U-Net 结构中加入新的模块、在 U-Net 结构中加入新的模块三种改进方向。U-Net 网络结构如图 4 所示，结构中的跳跃连接融合深浅层的特征图，整合不同尺度的特征信息。上述提到的几种改进方式往往是在 U-Net 的基础上增加参数量和计算量以缓解 OCT 分割难点带来的问题，从而提升分割精度。

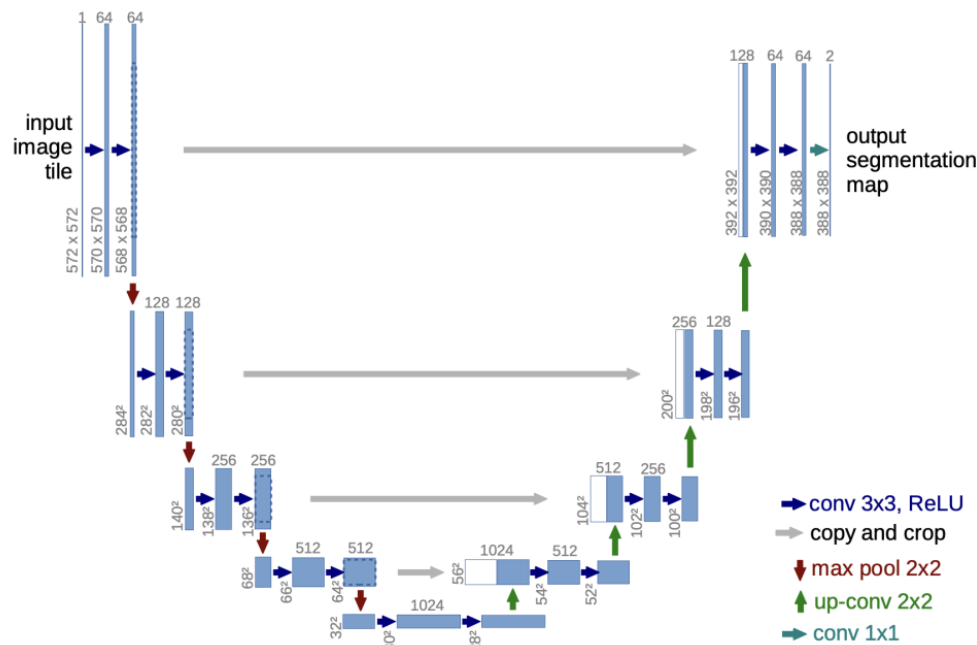


图 4 U-Net 网络模型示意图<sup>1</sup>

OCT 分割的难点主要为如图 5 所示的层与层之间边界模糊，对比不明显，层形状不规则，以及图像存在大量的背景噪声。参考文献<sup>[6]</sup>中提出的 WATNet 在 U-Net 结构基础上嵌入 Wavelet Attention 机制来考虑层间结构（如图 6 所示），参考文献<sup>[7]</sup>中提出的 LF-UNet 将 UNet 网络与 FCN<sup>[8]</sup>网络结合在一起以将层边缘跟全局信息结合考虑（如图 7 所示）。

<sup>1</sup> 图片来源：<https://arxiv.org/abs/1505.04597>



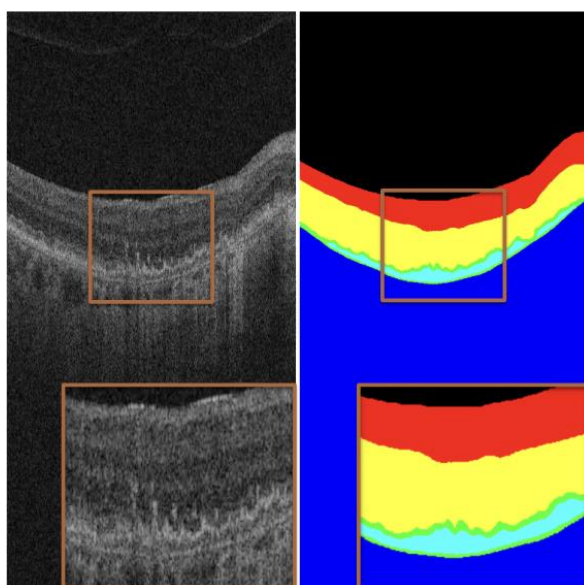


图 5 OCT 图像分割难点示意图

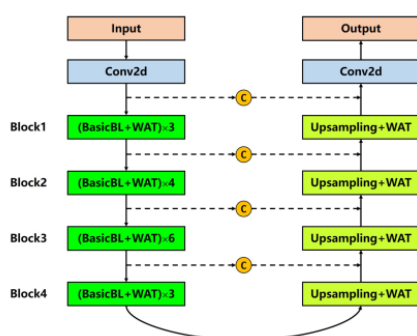


图 6 WATNet 示意图<sup>2</sup>

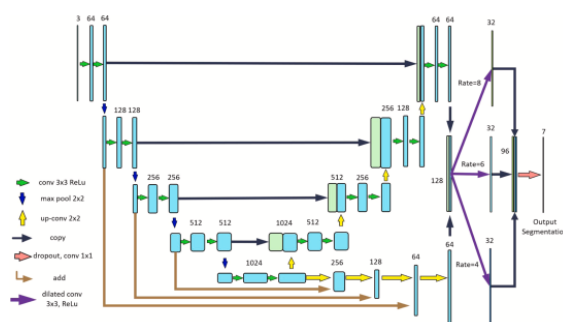


图 7 LF-UNet 示意图<sup>3</sup>

然而这种计算密集型分割网络在特定场景下并不适用：从应用场景来说，计算密集型分割网络由于推理速度慢、所需要的计算资源较多，不适用于一些实时医疗应用场景，例如影像引导手术、在线适应性放疗和实时监测。同时从移动医疗设备的工业生产角度考虑，计算密集型分割网络难以部署且所需硬件成本较高，难以有效大规模投入使用。因此，本文在不牺牲分割准确性的情况下提出一个轻量的能够部署在计算资源匮乏设备上的 OCT 分割算法，为提高自动化 OCT 分割对医疗诊断的有效辅助及实际应用的可行性做出贡献。

本文关注到降低输入图像分辨率能够降低网络计算量，但低分辨率图像中的细节丢失问题往往导致分割精度的下降。超分辨率技术是从观测到的低分辨率图像重建出相应的高分辨率图像的技术。其中重建的高分辨率图像可以作为分割网络的信息补充来源<sup>[9][10][11]</sup>。在超分辨率技术的辅助下，分割网络通过输入低分辨率图像可以获得有竞争力的分割结果，大大缓解了计算资源有限的困境。因此，本文根据 OCT

<sup>2</sup> 图片来源: <https://doi.org/10.1364/BOE.475272>

<sup>3</sup> 图片来源: <https://doi.org/10.1016/j.compmedimag.2021.101988>

图像分割的特点和难点引入超分辨分支以在不增加计算量和参数量的情况下提高分割的准确性，主要研究工作如下：

- 1) 基于 U-Net 网络模型进行轻量优化，在分割精度基本一致的情况下提出参数量、计算量较小的轻量 OCT 分割模型。
- 2) 在优化模型上引入超分辨网络 ESPCN<sup>[12]</sup>，设计融合模块提取超分辨网络中的有效信息对分割网络进行补充，在低分辨率输入及不增加计算量的情况下得到有竞争力的分割结果。
- 3) 在 AROI<sup>[2]</sup>数据集上验证所提出模型，并与现有经典分割网络及轻量分割网络进行对比，其在计算量和参数量较小的情况下超越一些现有分割网络模型。

## 2. 相关工作

### 2.1 基于深度学习的自动 OCT 图像分割网络

现有基于深度学习的自动 OCT 图像分割工作大多数是基于全卷积网络的，主要是基于 U-Net 网络模型及 FCN-DenseNet 网络模型，并根据 OCT 图像分割难点进行优化。参考文献<sup>[13]</sup>中提出的网络使用混合卷积和基于图的方法来识别 OCT 边界层。参考文献<sup>[14]</sup>中的研究使用了级联 U-Net 模型结构，并将性能与基于随机森林的经典方法的性能进行了比较。在另一项现有的工作<sup>[15]</sup>中使用基于 U-Net 结构的变体来分割视神经头 OCT。参考文献<sup>[16]</sup>中的工作使用了 FCN-DenseNet 模型，并将其与后处理相结合，后者使用启发式方法或使用高斯过程回归和径向基函数内核<sup>[17]</sup>。但这些现有工作均为了提高分割准确性，在原有基础框架上增加了参数量和计算量，限制了模型部署及应用场景。

### 2.2 基于深度学习的超分辨率网络

单幅图像超分辨率技术受到研究人员的关注，尤其是在自然场景中。由于大多数医学图像受制于成像设备无法提供高分辨率图像，超分辨率相关算法在医学领域并未得到发展。直到最近，研究人员提出了一些基于深度学习的医学图像的超分辨率方法<sup>[18][19]</sup>。为了构建高质量的图像，参考文献<sup>[20]</sup>中提出了一种用于 2D 脑电子计算机断层扫描图像的改进 U-Net，而参考文献<sup>[21]</sup>中设计一个具有多尺度特征的融合模块。此外，参考文献<sup>[22]</sup>中提出了一种基于亚像素卷积层的快速医学图像超分辨率方法。在本文中，超分辨率分支以 U-Net 为主干，两个任务使用相同的编码特征进行任务相关解码，并使用高效亚像素卷积网络重构高分辨率图像。

### 2.3 基于深度学习的轻量分割网络

一些研究工作提出了用于语义分割的轻量级神经网络架构。ShelfNet<sup>[23]</sup>使用具有空间跳跃连接和减少通道数量的多个编码器-解码器分支来提高准确性和效率。LedNet<sup>[24]</sup>通过通道拆分和打乱（shuffle）操作修改残差块以减少计算量。FANet<sup>[25]</sup>引入了一种用于实时语义分割的轻量级 Attention 模块。BiSeNetv2<sup>[26]</sup>引入了双分支架构来解耦空间和语义信息，从而带来有竞争力的性能及进行快速推理。STDCNet<sup>[27]</sup>通过用更轻量的处理模块替换空间路径来改进 BiSeNet，性能损失可以忽略不计。DensePASS<sup>[28]</sup>提出了在跨域条件下进行高效全景分割的无监督域自适应轻量模型。

### 3. 方法

#### 3.1 分割网络

U-Net 网络结构的优点是整合信息能力强、在小数据集上效果较好，但网络参数量较大，需要较多计算资源支撑，因此本文期望能够在 U-Net 的基础上进行优化，减少计算量和参数量，同时保证分割准确性基本一致。

##### 3.1.1 3x3 卷积替换

由于 U-Net 网络中采用了两个 3x3 卷积放入编码器和解码器结构层中（如图 8 所示），在优化 U-Net 结构时主要采用卷积分解的方式，本文设计出轻量高效的 MS Block 替换 3x3 卷积。

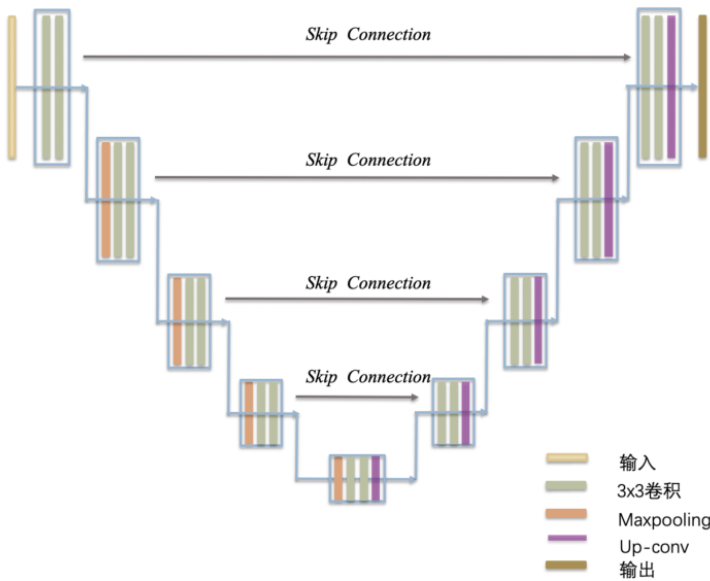


图 8 U-Net 模型示意图（强调卷积结构）

在 MS Block 结构中（如图 9 所示），首先使用一个 1x1 卷积将特征图通道数从  $C_{in}$  转换到  $C_{out}$ ，接着将特征图按照指定通道数  $C_{out}/K$  分成相等的  $K$  个特征，对每个部分采用可分离的空洞卷积，空洞卷积率为  $k$ ， $k=\{1, 2, \dots, K\}$ ，在保证感受野的同时更加高效的进行计算。接着对得到的相邻特征图进行叠加，增加通道上的信息。最后将特征图在通道上合并在一起，通过 1x1 卷积进行更好的信息融合得到输出。同时为了达到更加轻量的效果，减少了通道数及编码层，优化后的模型结构图如图 10 所示。

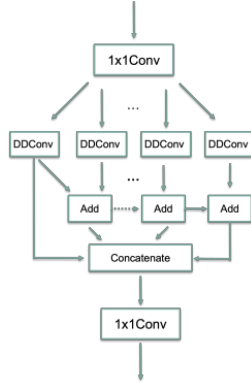


图 9 MS Block 示意图

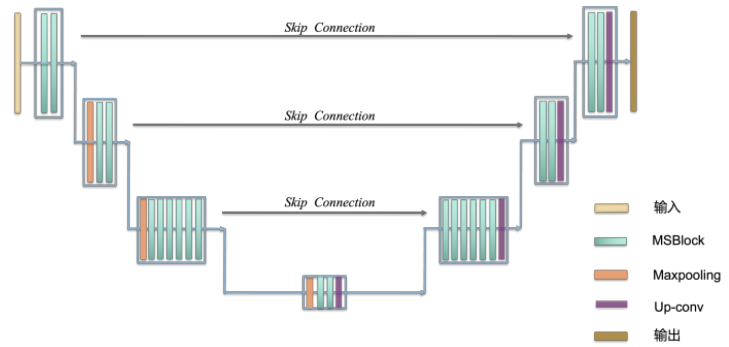


图 10 替换 3x3 卷积后模型示意图

### 3.1.2 Axial-Attention

由于卷积利用局部性来提高效率，但缺少远程上下文的联系，可以采用 self-attention<sup>[29]</sup>（如图 11（a）所示）以通过非局部交互来增加全局信息。self-attention 首先对输入采用不同编码得到 Query、Key、Value，用 Query 和 Key 相乘并通过 softmax 操作的方式得到当前位置与其他位置的关联度，相当于一个权重图，再跟 Value 相乘重新加权，得到全局关联信息。但由于 self-attention 的计算复杂度较高 ( $O(n^2)$ ， $n$  为序列长度)，对于输入较大的分割任务来说是不适用的。本文关注到 Axial-Deeplab<sup>[30]</sup>中提出的 Axial-Attention 思想，其在 2D self-attention 模块的基础上，通过分别计算长和宽的单维度 1D self-attention 以在保证全局连接的情况下减少计算复杂度。在减少计算量的同时，在 OCT 图像上，这种 X,Y 方向上各自的关联信息有助于疾病诊断（例如层间关系的捕捉）。因此，本文在 Axial-Attention 基础上，设计了 DAA Block（如图 11（d）所示）。其由 Scaled-Axial-Attention 模块和 Scaled-MLP 模块以及残差结构组成。Axial-Attention 模块如图 11（c）所示，包含 X 和 Y 方向上的 Multi-Head Attention。Multi-Attention 操作与 self-attention 基本一致，区别是其将模型分成多个头，形成多个子空间让模型关注不同方向的信息。相比于原始的 Multi-Head Attention, 本文增加了可学习参数控制 Key, Query, Value 的信息输入（如图 11（b）所示）。

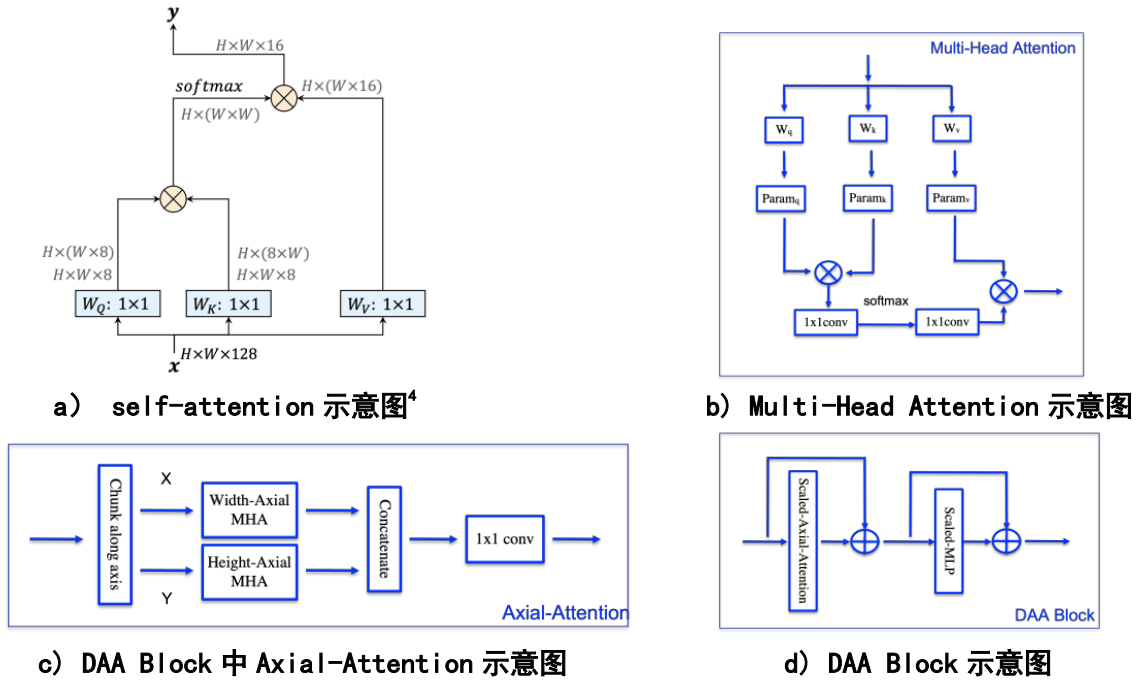
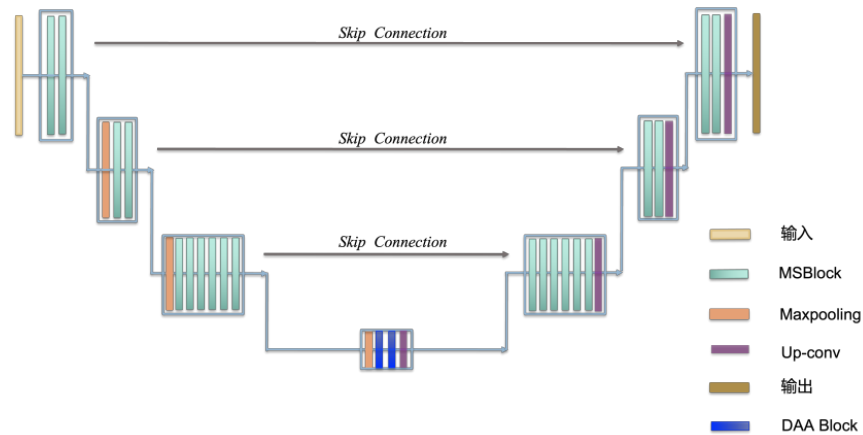


图 11 Axial-Attention 模块设计

为了能更有效地提取全局信息，DAA Block 添加在编码器的最后一层，改进后如图 12 所示。



## 3.2 超分辨率网络模型引入

### 3.2.1 超分辨率网络

超分辨率网络分支采用 ESPCN 网络模型，该工作直接输入低分辨率图像进行特征提取，后使用亚像素卷积层将没有插值的低分辨率图片映射到高分辨率图片（如图 13 所示）。亚像素卷积层有效融合通道信息，在没有额外计算量的同时将图像从  $H \times W$  映射到  $rH \times rW$  的大小，其中  $r$  是可设置的超参数。

<sup>4</sup> 图片来源：<https://arxiv.org/abs/2003.07853>

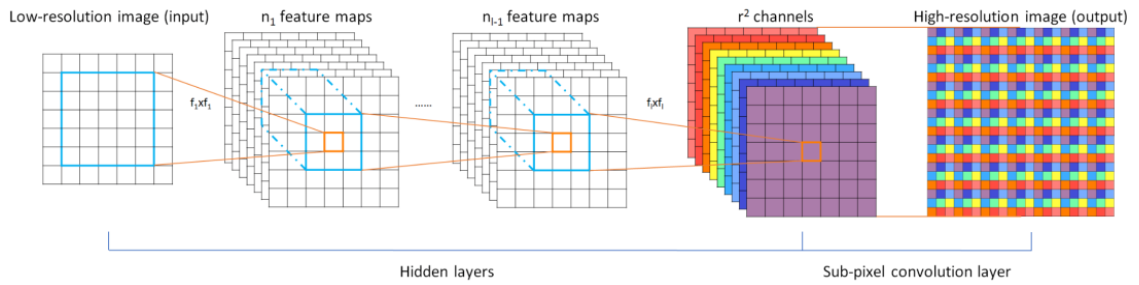


图 13 ESPCN 网络结构示意图<sup>5</sup>

### 3.2.2 超分辨率网络引入方式

超分辨率网络模型通过跟分割网络共用编码器的方式引入，在通过其独自的解码层后，通过亚像素卷积层得到高分辨率图像，同时采用损失函数进行引导。为了更有效更准确地提取两个分支中的有效信息进行相互补充，将两个分支解码器的最后一层输出作为融合模块的输入，具体结构图如图 14 所示。

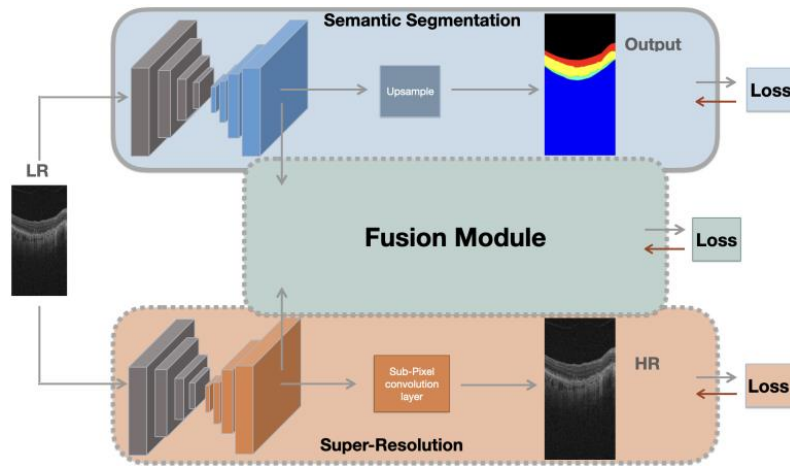


图 14 超分辨率网络引入方式示意图

### 3.2.3 融合模块

为了让超分辨率任务中的信息更加有效更加有针对性地补充进分割网络，本文设计了一个融合模块。融合模块的具体结构如图 15 所示，取两个任务分支的最后一层解码器特征作为融合模块输入，首先在通道上进行叠加并经过  $1 \times 1$  卷积进行融合。后经过 SplitFusion Block (SFB) 层，通过多尺度特征计算，更好地融合并提取层边缘信息。接着将融合特征分别通过两个  $1 \times 1$  卷积转化为分割融合特征图和超分辨率融合特征图，两种特征均经过一个 SelectFilter (SF) 模块，提取对方特征中的重要关联信息进行补充。再分别通过 Seg A 和 SR A 获得相应权重特征图，并通过与输入融合模块的原始特征相乘后相加的方式重新加权。最后同样分别通过一个上采样和亚像素卷积层得到分割融合输出以及超分融合输出。具体过程如公式(1)–(7)所示。

<sup>5</sup> 图片来源: <https://arxiv.org/abs/1609.05158>

$$F_{fusion} = SFB \left( Conv_{1 \times 1} \left( Concat(F_{seg}, F_{sr}) \right) \right) \quad (1)$$

$$Fusion_{seg} = Conv_{seg}(F_{fusion}) \quad (2)$$

$$Fusion_{sr} = Conv_{sr}(F_{fusion}) \quad (3)$$

$$M_{seg} = A_{seg} \left( SF_{seg}(Fusion_{seg}, Fusion_{sr}) \right) \quad (4)$$

$$M_{sr} = A_{sr} \left( SF_{sr}(Fusion_{sr}, Fusion_{seg}) \right) \quad (5)$$

$$O_{seg\_fusion} = F_{seg} * M_{seg} + F_{seg} \quad (6)$$

$$O_{sr\_fusion} = F_{sr} * M_{sr} + F_{sr} \quad (7)$$

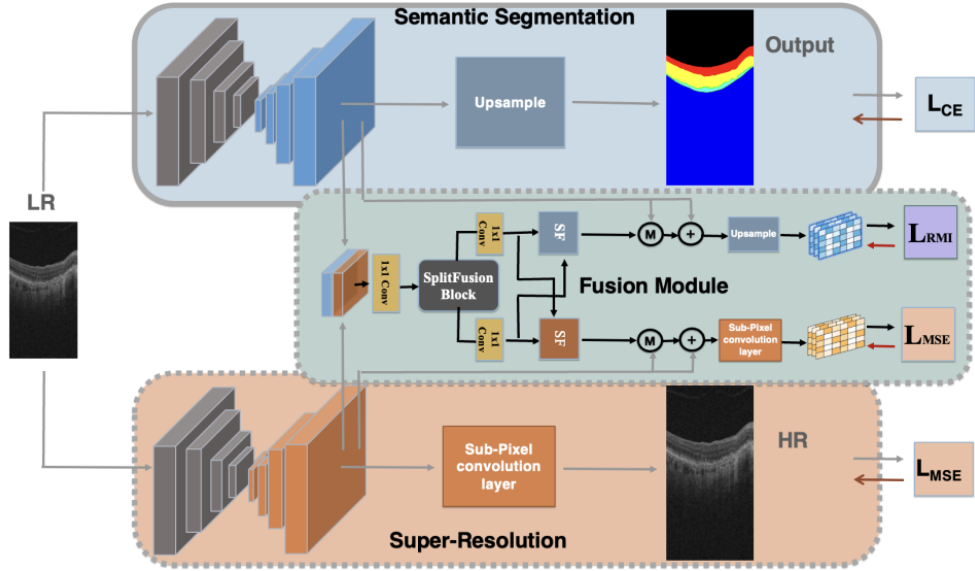


图 15 本文网络结构示意图

SFB 层主要是采用多尺度特征图融合的思想，具体结构如图 16 所示。首先是将  $n$  通道的输入按照通道分成  $j$  组有  $n/j$  通道的特征图。第一组特征图通过  $1 \times 1$  卷积进行浅层特征的提取，随后的  $j-1$  组采用空洞卷积率为  $i-1$  的  $3 \times 3$  空洞卷积进行感受野不变的多尺度高效特征计算，其中  $i$  为组编号， $i=1, 2 \dots j$ 。通过卷积多尺度计算后，前一层的输出特征图与下一层累积相加得到新的特征图，最后将  $j$  组特征输出图在通道上拼接在一起得到输出。



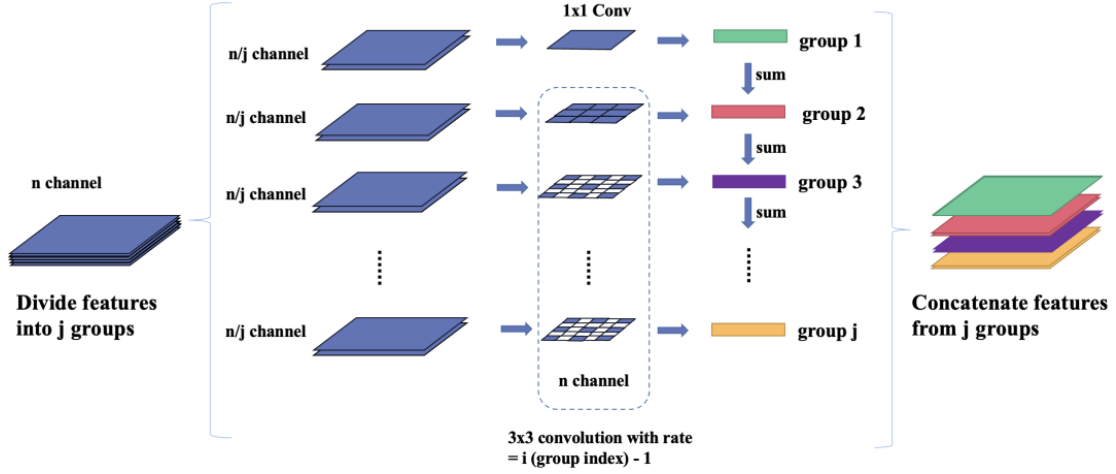


图 16 SFB 层结构示意图

SF 模块参考 Transformer<sup>[29]</sup> 的思想，计算分割融合特征和超分融合特征的相关性并筛选重要信息以高效对子任务进行补充。以提取超分融合特征中重要信息补充分割任务为例，SF 模块结构图如图 17 所示，首先将两个子任务输入融合特征进行下采样，后通过线性变换，由分割融合特征得到 Key 和 Value，超分融合特征得到 Query。Query 与 Key 相乘后进行 softmax 操作，得到超分融合特征图关于分割融合特征图的权重图。权重图与 Value 相乘重新加权，过滤出超分融合特征图中对分割特征图更有效的信息，实现高效信息筛选和补充。

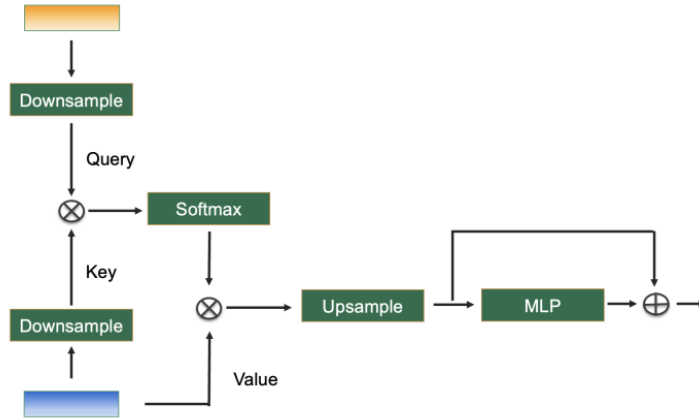


图 17 SF 模块结构示意图

#### 3.2.4 损失函数

损失函数主要对超分辨分支，分割分支以及融合模块这三个部分进行引导。对于分割分支和超分辨分支，分别采用两类任务中常用的交叉熵（CE）以及均方误差（MSE），分别如公式（8）和公式（9）所示。其中  $O_{seg}$  为分割网络的输出， $i$  为第  $i$  类， $y_i$  为第  $i$  类的原始标注图像。而在 MSE 损失公式中  $O_{sf}$  为超分辨网络的输出， $N$  为图像中的像素数， $j$  为像素索引，HR 为高分辨目标图像。

$$L_{CE} = \frac{1}{n} \sum_{i=0}^n y_i \log O_{seg_i} \quad (8)$$

$$L_{MSE} = \frac{1}{N} \sum_{j=0}^N y_i \left( O_{sr_j} - HR_j \right)^2 \quad (9)$$

对于融合模块中的损失函数，如公式（10）所示，主要分为两部分。对于分割融合输出，采用区域互感信息损失函数（Region Mutual Information, RMI）<sup>[31]</sup>进行引导，以更好关注到图形的局部区域。对于超分融合输出，则依旧采用 MSE 损失。整个网络模型所用的损失函数如公式（11）所示。

$$L_{fusion} = L_{RMI}(O_{seg\_fusion}, y) + L_{MSE}(O_{sr\_fusion}, HR) \quad (10)$$

$$L_{total} = L_{MSE} + L_{CE} + L_{fusion} \quad (11)$$

## 4. 实验

### 4.1 数据集

实验所用到的数据集为 AROI dataset，该公开数据集含有来自 24 位病人的 1136 张有标注的 OCT 图像，图像分辨率为 512x1024。AROI 数据集中的图像分为 8 个类（如图 18 所示），视网膜空间上方和下方，三个视网膜表面和三个病变。视网膜表面位于内限制膜（Internal Limiting Membrane, ILM）、内丛状层（Inner Plexiform Layer, IPL）/内核层（Inner Nuclear Layer, INL）、视网膜色素上皮（Retinal Pigment Epithelium, RPE）和布鲁赫膜（Bruch's Membrane, BM）的边界之间。注释的液体是色素上皮脱离（Retinal Pigment Epithelial Detachment, PED）、视网膜下液（Subretinal Fluid, SRF）和视网膜内液（Intraretinal Fluid, IRF）。由于数据集本身没有划分测试集，实验采用 K-fold 验证法（六折）进行实验。

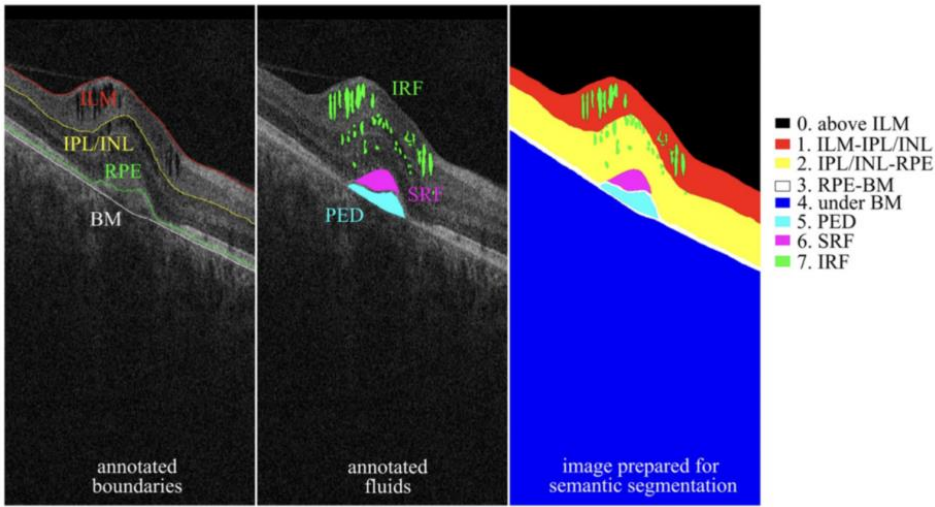


图 18 AROI 数据集分类示意图<sup>6</sup>

### 4.2 模型评价标准

根据设计轻量准确的 OCT 分层分割网络的目标，分别确定了分割准确性和模型是否轻量的判断标准：分割准确性方面，分别采用 IoU（Intersection of Union）、Recall（召回率）、Dice（Dice Similarity Coefficient）来进行衡量，公式如公式（12）-（14）所示；模型是否轻量则根据测试时模型的 FLOPs（浮点运算次数），Params（模型参数总数）来衡量。

<sup>6</sup> 图片来源：<https://doi.org/10.1080/00051144.2021.1973298>

$$IoU = \frac{TP}{TP+FN+FP} \quad (12)$$

$$Recall = \frac{TP}{TP+FN} \quad (13)$$

$$Dice = \frac{2*TP}{(TP+FN)+(TP+FP)} \quad (14)$$

### 4.3 实验结果

#### 4.3.1 分割网络对比实验

首先是本文提出的分割模型在以上两种评价标准下与 U-Net 网络以及分割网络 Deeplabv3+ 的对比实验，实验结果如表 1、表 2 所示。可以看到在相同输入和输出图像分辨率时，在分割准确性与 U-Net 基本一致且超过 Deeplabv3+ 的情况下，本文提出的分割网络在计算量、网络参数量上大大减少了。可视化结果如图 19 所示，可以看到虽然细节上跟 Ground Truth 依然有差距，但跟 U-Net 分割效果基本一致。

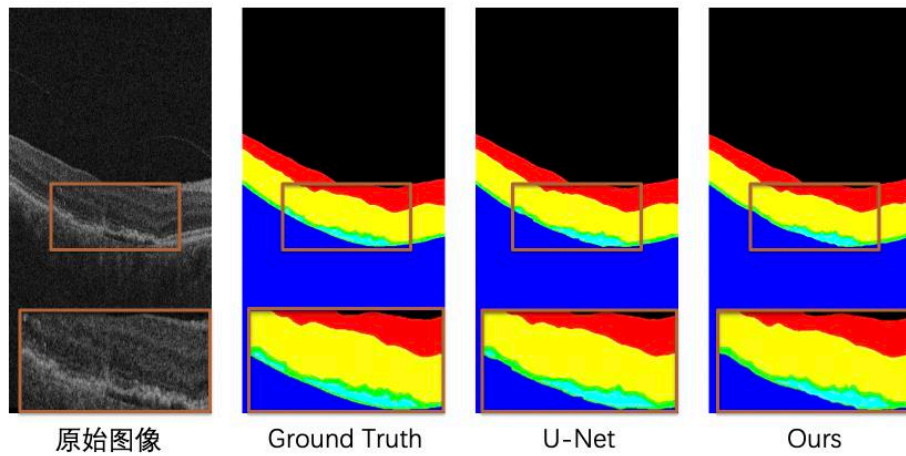


图 19 优化网络与 U-Net 网络可视化结果对比图

表 1 分割网络模型模型是否轻量衡量标准对比表

模型	输入大小	输出大小	计算量/G	参数量/M
Deeplabv3+	512x1024	512x1024	1467.82	40.36
U-Net	512x1024	512x1024	1544.86	28.95
Ours	512x1024	512x1024	39.05	0.97

表 2 分割网络模型分割准确度衡量标准对比表

模型	输入大小	输出大小	IoU	Dice	Recall
Deeplabv3+	512x1024	512x1024	64.25±3.33	72.53±4.21	73.08±4.43
U-Net	512x1024	512x1024	70.48±3.25	79.43±3.69	80.89±1.91
Ours	512x1024	512x1024	69.02±3.22	77.65±3.95	79.41±3.24

4.3.2 消融实验

为进一步降低计算量，降低输入图像分辨率进行实验，实验结果如表 3 所示。这里关注到，由于输入图像分辨率降低，虽然模型计算量降低到原来的四分之一，但在分割精度上跟 U-Net 模型差距变大。降低输入图像分辨率前后两个模型的分割准确性对比如表 4 和表 5 所示，关注到相比于 U-Net 模型来说，由于网络模型结构更加轻量，优化模型受降低输入图像分辨率的影响更大。因此，后续在优化模型上引入超分辨网络以提高分割精度是必要的。

表 3 降低输入图像分辨率后实验结果对比表

模型	输入大小	输出大小	计算量/G	参数量/M	IoU	Dice	Recall
U-Net	256x512	256x512	386.85	28.95	69.34±2.77	77.84±3.60	78.50±3.96
Ours	256x512	256x512	9.74	0.97	67.01±3.05	75.92±4.05	76.17±3.27

表 4 降低输入图像分辨率前后本文分割网络分割精度对比表

模型	输入大小	输出大小	IoU	Dice	Recall
Ours	512x1024	512x1024	69.02±3.22	77.65±3.95	79.41±3.24
Ours	256x512	256x512	67.01±3.05	75.92±4.05	76.17±3.27

表 5 降低输入图像分辨率前后 U-Net 模型分割精度对比

模型	输入大小	输出大小	IoU	Dice	Recall
U-Net	512x1024	512x1024	70.48±3.25	79.43±3.69	80.89±1.91
U-Net	256x512	256x512	69.34±2.77	77.84±3.60	78.50±3.96

针对只引入超分辨分支的网络模型设计以及加入融合模块的网络模型设计进行了实验，与不加入超分辨分支的分割网络模型进行对比，实验结果如表 6 所示。在测试时均不加超分分支以及融合模块，没有增加额外的测试计算量及参数量。从表 6 中可以看出，直接引入超分辨网络并用损失函数引导的方式反而在三个分割精度评判维度上都使网络分割精度下降，原因可能是这种引导方式无法有效提取超分辨网络中有效信息对分割网络进行补充，反而加入干扰信号导致分割精度下降。在加上融合模块后，分割精度在三个衡量标准下有较大提升，验证了融合模块高效提取超分辨分支中重要信息对分割模型进行补充的有效性。

表 6 消融实验对比结果表

模型	输入大小	输出大小	IoU	Dice	Recall
分割模型	256x512	256x512	67.01±3.05	75.92±4.05	76.17±3.27
引入超分辨	256x512	512x1024	66.02±2.38	74.98±3.28	75.13±3.77
加入融合模块	256x512	512x1024	<b>70.97±3.30</b>	<b>79.36±3.96</b>	<b>79.83±3.32</b>

#### 4.3.3 与现有分割网络对比

为更准确的评判本文所提出网络的轻量程度与分割网络精度，将本文所提出 OCT 分割网络模型分别与经典模型 U-Net，Deeplabv3+，U-Net++以及轻量模型 BiSeNetv2，STDCNet 进行对比。实验设置均保持一致，实验结果如表 7 所示。从表中可以看出，与计算量、参数量较大的经典分割模型相比，本文所提出分割模型在更为轻量的同时，保证了有竞争力的 OCT 分割结果。与轻量分割网络模型相比，本文所提出分割模型针对 OCT 分割进行设计，取得更好的分割结果。可视化结果如图 20 所示，可以看到本文所提出的 OCT 分割模型在全局以及局部区域最为贴近原始标注和原始图像，虽然在一些区域仍有误判，但相较于其他对比网络来说较为完整和准确。

表 7 对比实验结果表

模型	输入大小	输出大小	计算量/G	参数量/M	IoU	Dice	Recall
U-Net	512x1024	512x1024	1544.86	28.95	70.48±3.25	79.43±3.69	<b>80.89±1.91</b>
Deeplabv3+	512x1024	512x1024	1467.82	40.36	64.25±3.33	72.53±4.21	73.08±4.43
U-Net++	512x1024	512x1024	1118.99	9.16	69.87±2.99	78.53±3.76	79.19±2.90
BiSeNetv2	512x1024	512x1024	158.64	28.67	69.95±3.80	78.56±4.48	78.95±2.97
STDCNet	512x1024	512x1024	110.08	14.43	66.23±6.10	75.35±6.20	75.82±6.50
Ours	256x512	512x1024	9.74	0.97	<b>70.97±3.30</b>	<b>79.46±3.96</b>	79.83±3.32

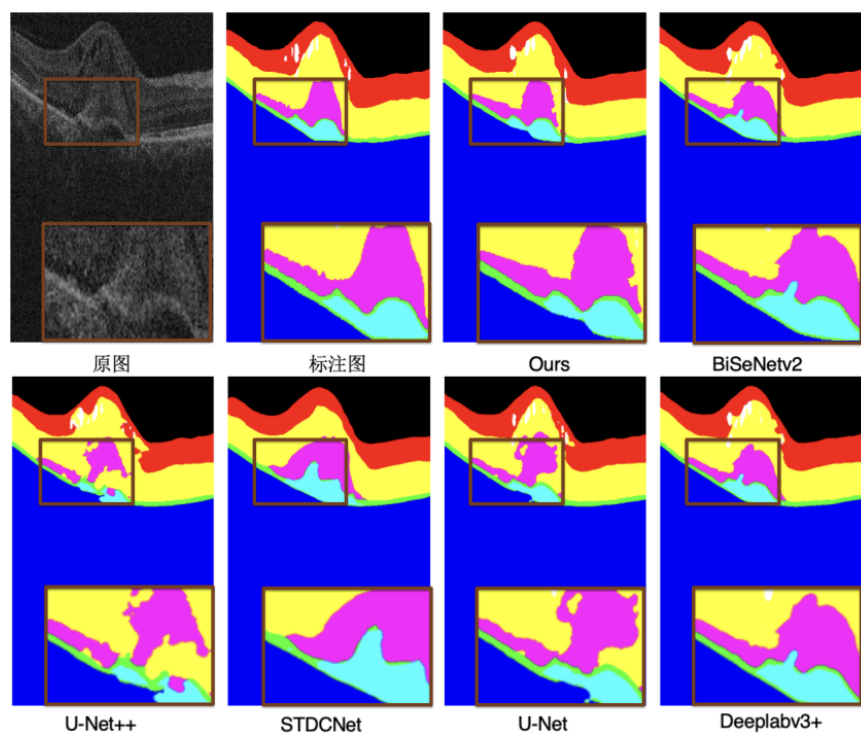


图 20 本文提出网络与现有工作可视化结果对比图

## 5. 总结

本文提出了一个基于深度学习的轻量 OCT 分割网络模型，主要是在 U-Net 网络基础上保证分割精度的同时进行轻量级优化，达到在计算量为 U-Net 模型 3% 的同时得到更好的分割准确性。后设计融合模块引入超分辨率网络以在降低输入图像分辨率以及不引入额外计算量的同时提高分割网络准确性。该模型在 AROI 数据集上进行验证，分割结果在大大减少网络的参数量和计算量的基础上超过一些现有经典分割网络及轻量分割网络，为自动化 OCT 分割在边缘医疗移动设备上的部署提供了新的思路。总而言之，本文提出的分割-超分辨双流 OCT 分割算法提高了自动 OCT 分割网络模型在实时医疗应用场景中应用的可行性，为 OCT 自动化分割进一步高效辅助医疗诊断做出贡献。虽然本文提出的轻量 OCT 分割模型在计算量、参数量均较小的情况下，分割效果相对对比网络较优，但其仍然存在一些缺陷。例如由于网络大大减小参数量和计算量，其稳定性相比于 U-Net 网络来说较为逊色，值得后续进一步探索和研究。



## 参考文献

- [1] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015: 234-241.
- [2] Melinščak M, Radmilović M, Vatauvuk Z, et al. Annotated retinal optical coherence tomography images (AROI) database for joint retinal layer and fluid segmentation[J]. Automatika, 2021, 62(3-4): 375-385.
- [3] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(12): 2481-2495.
- [4] Jégou S, Drozdal M, Vazquez D, et al. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017: 11-19.
- [5] 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, 2017, 40(4): 834-848.
- [6] Wang C, Gan M. Wavelet attention network for the segmentation of layer structures on OCT images[J]. Biomedical Optics Express, 2022, 13(12): 6167-6181.
- [7] Ma D, Lu D, Chen S, et al. LF-UNet–A novel anatomical-aware dual-branch cascaded deep neural network for segmentation of retinal layers and fluid from optical coherence tomography images[J]. Computerized Medical Imaging and Graphics, 2021, 94: 101988.
- [8] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.
- [9] DAI, Dengxin, et al. Is image super-resolution helpful for other vision tasks? [C]// 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2016. p. 1-9.
- [10] S. Lei, Z. Shi, X. Wu, et al. Simultaneous super-resolution and segmentation for remote sensing images[C]//IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium.

IEEE, 2019, pp. 3121–3124.

- [11] Tang Z, Pan B, Liu E, et al. Srda-net: super-resolution domain adaptation networks for semantic segmentation[J]. arXiv e-prints, 2020: arXiv: 2005.06382.
- [12] Shi W, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1874-1883.
- [13] Fang L, Cunefare D, Wang C, et al. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search[J]. Biomedical optics express, 2017, 8(5): 2732-2744.
- [14] He Y, Carass A, Yun Y, et al. Towards topological correct segmentation of macular OCT from cascaded FCNs[C]//Fetal, Infant and Ophthalmic Medical Image Analysis: International Workshop, FIFI 2017, and 4th International Workshop, OMIA 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 4. Springer International Publishing, 2017: 202-209.
- [15] Devalla S K, Chin K S, Mari J M, et al. A deep learning approach to digitally stain optical coherence tomography images of the optic nerve head[J]. Investigative ophthalmology & visual science, 2018, 59(1): 63-74.
- [16] Pekala M, Joshi N, Liu T Y A, et al. Deep learning based retinal OCT segmentation[J]. Computers in biology and medicine, 2019, 114: 103445.
- [17] Seeger M. Gaussian processes for machine learning[J]. International journal of neural systems, 2004, 14(02): 69-106.
- [18] Zhao X, Zhang Y, Zhang T, et al. Channel splitting network for single MR image super-resolution[J]. IEEE transactions on image processing, 2019, 28(11): 5649-5662.
- [19] Shi J, Liu Q, Wang C, et al. Super-resolution reconstruction of MR image with a novel residual learning network algorithm[J]. Physics in Medicine & Biology, 2018, 63(8): 085011.
- [20] Park J, Hwang D, Kim K Y, et al. Computed tomography super-resolution using deep convolutional neural network[J]. Physics in Medicine & Biology, 2018, 63(14): 145011.
- [21] Liu C, Wu X, Yu X, et al. Fusing multi-scale information in convolution network for MR image super-resolution reconstruction[J]. Biomedical engineering online, 2018, 17(1): 1-23.
- [22] Zhang S, Liang G, Pan S, et al. A fast medical image super resolution method based on deep learning

- network[J]. IEEE Access, 2018, 7: 12319-12327.
- [23] Zhuang J, Yang J, Gu L, et al. Shelfnet for fast semantic segmentation[C]//Proceedings of the IEEE/CVF international conference on computer vision workshops. 2019: 0-0.
- [24] Wang Y, Zhou Q, Liu J, et al. Lednet: A lightweight encoder-decoder network for real-time semantic segmentation[C]//2019 IEEE international conference on image processing (ICIP). IEEE, 2019: 1860-1864.
- [25] Hu P, Perazzi F, Heilbron F C, et al. Real-time semantic segmentation with fast attention[J]. IEEE Robotics and Automation Letters, 2020, 6(1): 263-270.
- [26] Yu C, Gao C, Wang J, et al. Bisenet v2: Bilateral network with guided aggregation for real-time semantic segmentation[J]. International Journal of Computer Vision, 2021, 129: 3051-3068.
- [27] Fan M, Lai S, Huang J, et al. Rethinking bisenet for real-time semantic segmentation[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 9716-9725.
- [28] Ma C, Zhang J, Yang K, et al. Densepass: Dense panoramic semantic segmentation via unsupervised domain adaptation with attention-augmented context exchange[C]//2021 IEEE International Intelligent Transportation Systems Conference (ITSC). IEEE, 2021: 2766-2772.
- [29] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
- [30] Wang H, Zhu Y, Green B, et al. Axial-deeplab: Stand-alone axial-attention for panoptic segmentation[C]//Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IV. Cham: Springer International Publishing, 2020: 108-126.
- [31] Zhao S, Wang Y, Yang Z, et al. Region mutual information loss for semantic segmentation[J]. Advances in Neural Information Processing Systems, 2019, 32.

## 致谢

四年本科生涯转眼就接近了尾声。在此，我首先要对我的学术导师刘江老师表示衷心的感谢。感谢您在我自大二进入实验室以后一直以来的关心和指导，让我从什么都不懂的“科研小白”，成长为有一定基础的科研入门者。感谢您在学习生活方面给予我的关怀和建议，让我在每个纠结迷茫的时刻做出相对最优的选择。

我还要感谢组里的邱忠喜师兄和其他为本工作提出指导意见的师兄师姐们。正是通过跟师兄师姐们的一次次探讨和修改，本工作才能顺利完成。

最后，我还要感谢我的家人和朋友们。感谢他们在本工作完成期间倾听我的烦恼，分享我的喜悦，给予我情感上强有力的支持。在此，我要向所有给予我帮助和支持人致以最真挚的感谢！