

文章编号 1004-924X(2023)21-3203-09

## 注意力感知的多尺度语义视杯盘分割

燕 杨\*, 曹娅迪, 黄文博

(长春师范大学 计算机科学与技术学院, 吉林 长春 130032)

**摘要:**为了解决编码器-解码器网络结构在目标提取中抑制无关语义、跨越语义鸿沟等问题,以获取更高精度,采用 U-Net 作为提取特征的主干网络;为了减轻浅层特征与深层特征语义的差异,设计一种融合注意力感知的多尺度语义池化模块(Channel-Spatial-Pyramid, CSP),替代早期层中的跳跃链接。CSP 模块从空间与通道两个层面强调更有意义的语义信息,通过 4 个不同池化核的并行分支提取不同尺度特征,聚合所有分支结果与后期层特征拼接。实验结果表明,CSP-Net 在彩色眼底图像视盘分割中的 Dice 指数可达 99.6%,视杯分割结果的 Dice 指数为 92.1%,相比现有算法均有提高。所提出的 CSP-Net 对于眼底图像中的微小目标提取的有效性及其抗干扰性较高,可为青光眼筛查与诊断临床提供借鉴。

**关键词:**多尺度语义;注意力感知;目标提取;U-Net

**中图分类号:**TP394.1;TH691.9 **文献标识码:**A **doi:**10.37188/OPE.20233121.3203

## Multi-scale semantic OD/OC segmentation method based on attention perception

YANG Yan\*, CAO Yadi, HUANG Wenbo

(School of Computer Science and Technology, Changchun Normal University, Changchun 130032, China)

\* Corresponding author, E-mail: yanyang2016@hotmail.com

**Abstract:** To suppress irrelevant semantics and cross semantic gaps in object extraction using an encoder-decoder network structure, thereby achieving higher accuracy. U-Net is used as the backbone network for feature extraction. To reduce semantic differences between shallow and deep features, a multi-scale semantic pooling module (CSP, Channel-Spatial-Pyramid) integrates attention perception and replaces skip links in early layers. The CSP module emphasizes more meaningful semantic information from two levels corresponding to space and channel, extracts features at different scales through parallel branches of four different pooling cores, and aggregates all branch results to splice with the features of later layers. The experimental results show that the Dice index of CSP-Net in color fundus image disc segmentation reaches 99.6%, whereas that of cup segmentation reaches 92.1%. Both results represent improvements over existing algorithms. CSP-Net exhibits a high effectiveness and anti-interference ability for extracting small targets in fundus images, making it appropriate for clinical reference in glaucoma screening and diagnosis.

**Key words:** multi scale semantics; attention perception; target extraction; U-Net

收稿日期:2023-05-22;修订日期:2023-06-15.

基金项目:吉林省科技厅自然科学基金(联合基金项目)(No. YDZJ202101ZYTS147)

## 1 引言

视盘(Optic Disc, OD)是视神经的起点,位于视网膜中心,视杯(Optic Cup, OC)即其中央凹陷处。由于青光眼导致的视神经损伤,OC区域会不断扩大,导致杯盘比(Cup to Disk Ratio, CDR)增加<sup>[1]</sup>。青光眼早期阶段并无明显症状,CDR即为诊断青光眼的重要指标。通过CDR诊断青光眼需要眼科医生对大量眼底图像中的OD与OC区域进行标注,耗时耗力,为此学者们对OD/OC自动提取算法展开研究。基于数字图像处理的OD和OC分割方法主要分为两种,一种是通过水平集、主动轮廓等算法寻找边缘确定OD范围;另一种是通过分析各类特征,使用分类器提取每个像素点分类,为了提高分割效率,这类方法多应用在彩色眼底图像的超像素分割中<sup>[2-5]</sup>。上述早期算法无法提取更抽象的特征,仅依赖低级特征,不仅易受干扰,而且精度过低。

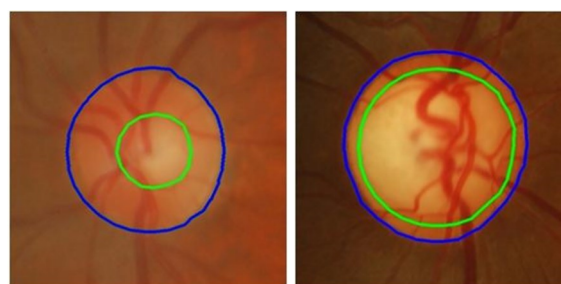
神经网络具有很强的特征提取能力,其中以M-Net<sup>[6]</sup>最为经典。M-Net采用金字塔多尺度输入,并通过side-out层为早期层生成伴随的局部输出图,最终将多个尺度的特征图聚合生成输出。针对OD像素与背景像素严重不均衡的问题,M-Net通过极坐标转换以及采用基于Dice系数的多标签损失函数替代常见的多类分割损失来缓解。目前,许多OD和OC分割方法都借鉴了M-Net的思想。Liu<sup>[7]</sup>等提出一种密集连接的深度可分离卷积网络,使用金字塔多尺度输入避免网络参数大幅增加,使用深度可分离卷积减少计算消耗,采用密集连接的形式,获取眼底图像中更丰富、更复杂的特征,提高OD和OC的提取精度。于舒扬等<sup>[8]</sup>基于Seg-Net框架提出了融合感受野模块的Seg-RFNet框架,其编码器采用ResNet-50结构,进行分支处理拓宽网络结构,获得了更抽象的语义信息;其解码器的输入来自不同级编码器,以更完整地恢复图像。

由于浅层特征与深层特征之间存在语义鸿沟,U型网络编码器-解码器通过跳跃链接将二者直接拼接的行为通常会导致噪声,增加干扰。为此,本文提出了注意力感知的多尺度语义目

标提取模块(Channel-Spatial-Pyramid, CSP)。该模块在多尺度池化方法中引入通道-空间联合注意力机制,在丰富上下文信息的同时,增大对目标提取有意义的特征权重。进一步提出基于CSP模块的OD/OC自动分割网络CSP-Net,将CSP模块应用于编码器-解码器网络,增强早期层特征,减轻语义鸿沟,提高算法性能和精度。

## 2 原理

正常情况下,OD和OC的形态如图1(a)所示,其中蓝色轮廓内区域为OD,绿色轮廓内区域为OC(彩图见期刊电子版)。由于青光眼导致的OC病变,如图1(b)所示,OD与OC特征接近,难以区分。本文设计了CSP模块加入网络,提高了网络特征提取能力,能够更精准地分割OD与OC。CSP-Net整体结构如图2所示。提取特征的主干网络在U-Net基础上减少了滤波器数量和训练参数,降低训练所需要的计算消耗。针对早期特征与后期特征之间存在的语义鸿沟问题,在主干网络下采样路径的早期层加入CSP模块,增强网络提取上下文信息的能力并抑制无关特征,减轻浅层特征与深层特征的语义差异。传统的金字塔池化模块(Pyramid Pooling Module, PPM)<sup>[9]</sup>通过使用不同尺寸的池化核,达到获取上下文信息的目的,但早期特征仍比较粗糙,含有大量与分割目标无关的干扰特征。因此,CSP模块在金字塔池化中引入通道-空间联合注意力模块,从通道和空间两个维度放大对识别目标更有意义特征。



(a) 未患病视杯盘 (b) 青光眼视杯盘  
(a) Normal OD and OC (b) Glaucoma OD and OC

图1 彩色眼底图像中的OD和OC区域

Fig. 1 OD and OC regions in color fundus images

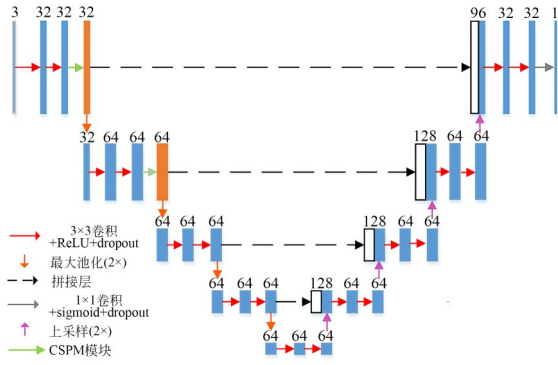


图 2 CSP-Net 网络模型

Fig. 2 CSP-Net model

## 2.1 轻量级 U-net

早期特征在卷积与池化过程中均有丢失,仅凭后期特征进行分割不够全面。U-Net<sup>[10]</sup>通过拼接层将早期特征与后期特征进行融合,解决了这一问题,因此成为最有效的特征提取网络之一。U-Net 主要由两部分组成,其中收缩路径用于获取上下文信息,扩张路径用于完成精准分割,整个网络结构呈对称状态。

为减少训练参数,提高训练效率,这里减少了主干网络卷积层上的滤波器,并且不增加用于降低分辨率的滤波器数目。这些改变并没有降低网络提取特征的能力,而是在参数数量和训练时间方面变得更加轻量级<sup>[11]</sup>。

## 2.2 损失函数

在 OD 分割任务中,由于 OD 区域仅占全部图像很小的部分,因此存在正负样本严重不均衡的问题,训练时损失函数会陷入局部最小值,产生更重视背景部分的网络,样本量较少的前景部分常会丢失或是仅被部分提取<sup>[12]</sup>。为了解决这一问题,将损失函数定义为:

$$l(A, B) = -\log(d(A, B)), \quad (1)$$

$$d(A, B) = \frac{2 \sum_{i,j} a_{ij} b_{ij}}{\sum_{i,j} a_{ij}^2 + b_{ij}^2}, \quad (2)$$

其中:  $d(A, B)$  是 Dice 损失函数,是基于一种反应正负样本相似度的评价函数;Dice 系数的取值为  $[0, 1]$ ;  $A$  是网络输出的概率图,  $B$  是专家手工标注的真实标签,包含每个像素及其所属的类;  $a_{ij}$  和  $b_{ij}$  分别代表  $A$  和  $B$  中的某一像素。Dice 损失函数评估真实标签与预测结果的相似程度,而非像素级的准确度,避免网络为追求更低损失将全部样本分类为负样本的情况。

## 2.3 CSP 模块

考虑到语义鸿沟问题, CSP-Net 在主干网络早期层中引入基于注意力机制的多尺度池化 CSP 模块。在编码器-解码器网络中,随着网络层数的加深,特征分辨率不断降低,这一过程使深层特征丢失了许多上采样操作无法恢复的细节信息。通过跳跃链接方法将浅层特征与深层特征进行拼接虽然可以解决这一问题,却因早期特征更粗糙、后期特征更抽象这一差异导致噪声,影响分割精度。因此,这里在主干网络早期层引入 CSP 模块,对早期特征进行进一步处理再与后期特征进行拼接,减少语义差异,如图 3(a) 所示。

U-Net 将不同卷积层的特征拼接实现了上下文信息融合,但提取特征时依旧使用固定尺寸的卷积核,缺少对每一层上下文信息的有效提取和利用。为充分提取早期层所包含的上下文信息, CSP 模块基于金字塔池化结构,依靠多个不同尺寸的感受野检测尺寸各异的目标,获得更丰富的多尺度特征。普通最大池化使用单个池内核,金字塔池化结构采用 4 条并行分支进行池化,每条分支池化核尺寸均不相同。4 个不同尺寸感受野收集特征图的上下文信息并进行编码,产生 4 个不同尺寸的特征图并将所有分支结果拼接。最后,对拼接的特征图使用  $1 \times 1$  卷积,将通道数恢复至原特征图尺寸。

传统的金字塔池化结构虽然提取到丰富的空间上下文信息,但无法阻止无关特征进入后期层,因此 CSP 在金字塔池化结构中引入通道-空间联合注意力模块。通道-空间联合注意力模块拥有两个顺序子模块,即通道注意力模块和空间注意力模块。对于一个给定的特征图  $F \in \mathbb{R}^{H \times W \times D}$ , 该模块首先通过沿着通道模块产生一个一维的通道注意力特征图  $A_c \in \mathbb{R}^{1 \times 1 \times D}$ , 将它乘以输入特征图进行自适应特征细化,产生细化的特征图  $F'$ 。  $F'$  作为空间注意力模块的输入,产生通道数为 1 的空间注意力图  $A_s \in \mathbb{R}^{H \times W \times 1}$ , 并与  $F'$  相乘,产生最终的注意力特征图  $F''$ 。该过程的数学形式如下:

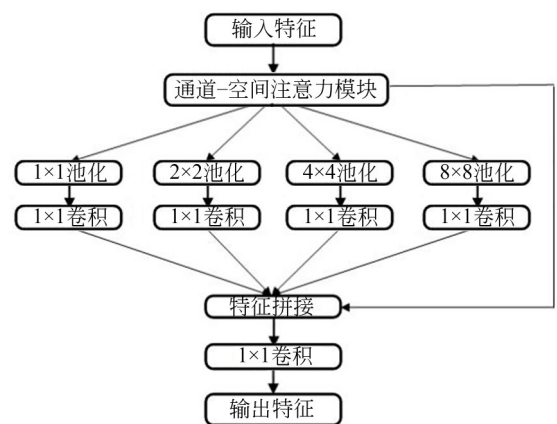
$$\begin{cases} F' = A_c \otimes F \\ F'' = A_s \otimes F' \end{cases}, \quad (3)$$

其中 $\otimes$ 表示对特征图进行逐元素相乘。在该过程中,通道注意力模块产生的注意值会继续在空间维度传播。通道注意力模块与空间注意力模块结构分别如图3(b)和3(c)所示。

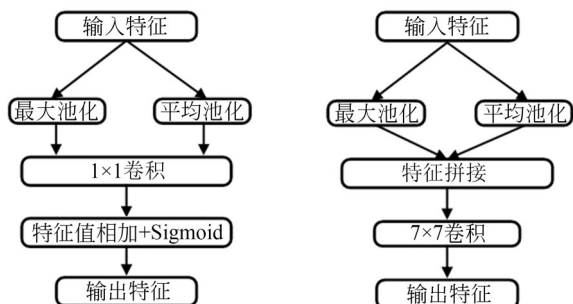
输入通道注意力模块的特征图,首先经过最大池化和平均池化聚合特征映射的空间信息,产生两组不同的特征信息——平均池化所聚合的空间信息与最大池化所收集的不同类别的特征,使网络可以有效学习 OD 和 OC 的范围,减少无关背景对识别目标的影响,同时细化特征,有效降低 OC 目标的提取难度。通道注意力模块对池化产生的两组特征分别进行  $1 \times 1$  卷积操作后,再进行求和,最后使用 sigmoid 函数进行激活,产生最终的通道特征图。通道注意力模块可表示为:

$$A_c = \sigma(\text{conv}_{1 \times 1}(\text{AvgPool}(F)) + \text{conv}_{1 \times 1}(\text{MaxPool}(F))), \quad (4)$$

其中  $\sigma$  表示 sigmoid 激活函数。



(a) CSP模块  
(a) CSP module



(b) 通道注意力模块  
(b) Channel attention module

(c) 空间注意力模块  
(c) Spatial attention module

图3 CSP与其子模块结构

Fig. 3 CSP module and submodule structure

空间注意力模块作为通道注意力模块的补充,接收通道注意力特征图作为输入。经过全局最大池化和全局平均池化后,将结果聚合,经过  $7 \times 7$  卷积和 Sigmoid 函数,产生最终的空间注意力图。与通道注意力不同的是,空间注意力会聚焦特征图中的关键信息,即强调感兴趣目标,有助于解决前背景像素分布不均衡的问题。空间注意力模块还可表示为:

$$A_s = \sigma(f^{7 \times 7}([\text{AvgPool}(F') + \text{MaxPool}(F')])), \quad (5)$$

其中  $f^{7 \times 7}$  为卷积核尺寸为  $7 \times 7$  的卷积操作。

### 3 实验

#### 3.1 实验条件

本文分别在 RIM-ONE v. 3<sup>[13]</sup>, DRIONS-DB<sup>[14]</sup>, DRISHTI-GS<sup>[15]</sup> 公开彩色眼底数据集上进行实验,将本文方法与现有方法在 OD/OC 识别任务中进行对比,以验证本文方法的性能与泛化能力。在实验中,数据集被随机划分为训练集与验证集两部分,其中训练集占数据集总量的 80%,验证集占总量 20%。训练集与验证集中均包含正常眼底与青光眼眼底两类数据。

实验使用深度学习框架 Tensorflow 2.0 和 Keras 2.3.1,利用 GPU (graphic processing unit) 加速网络模型训练及测试,显卡型号为 NVIDIA Tesla K80。采用 Adam 优化器,Adam 算法可在训练时自适应调节学习率,且收敛速度更快。

训练前采用对比度受限自适应直方图均衡进行预处理。由于样本数量较少,为防止过拟合,在训练阶段对每幅图像分别进行 4 种数据增强操作,包括随机缩放、随机水平偏移、随机垂直偏移和随机旋转,并将图像统一缩放为  $256 \times 256$  像素进行训练与测试。为了充分利用先验知识,在 OC 识别前先根据图中 OD 的位置裁剪出尺寸为  $128 \times 128$  的区域,再输入网络中。训练时,批处理大小设置为 4,训练进行 500 个 epoch。

#### 3.2 评估函数

为客观评估本文算法性能,采用 Dice 系数、平均交并比 (Mean Intersection over Union, MIoU)、总体分割精度 (Accuracy, Acc),精确度 (Precision, PC) 和均方误差 (Mean Square Error,



MSE)作为评价标准,将本文算法结果与现有算法结果进行比较。评价函数公式如下:

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

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{TP + FP + FN}, \quad (7)$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \quad (8)$$

$$PC = \frac{TP}{TP + FP}, \quad (9)$$

其中:TP为预测正确的正样本像素数量,FP为预测错误的正样本像素数量,TN为预测正确的负样本像素数量,FN预测错误的负样本像素数量, $k+1$ 是包括背景在内的类的数量。

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N [y(i,j) - x(i,j)]^2}{M \times N}, \quad (10)$$

其中: $M$ 和 $N$ 分别代表图像的宽和高, $y(i,j)$ 代表原始图像中的某一像素, $x(i,j)$ 为预测图像中的对应像素。

### 3.3 对比实验及结果分析

在OD提取任务中,将本文方法与现有方法进行比较,结果如表1所示。在DRIONS-DB数据集上,本文方法的MIoU与Dice系数分别达到92.6%与96.9%,Acc达到99.8%,PC为97.0%。与未加入CSP模块的主干网络比较,各项指标均有所增加,且所有对比算法中,CSP-Net的Dice值最高。

图4展示了在DRIONS-DB数据集中,本文方法与其他方法预测结果的可视化结果。以第3

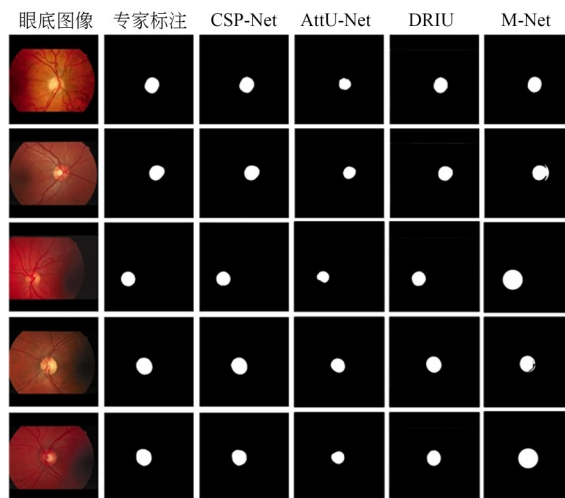


图4 DRIONS-DB数据集OD提取可视化结果对比

Fig. 4 Comparison of OD extraction results on DRIONS-DB dataset

行为例,AttU-Net由于血管干扰,视盘区域欠分割;M-Net的分割结果则远大于视盘区域,会影响CDR计算,最终影响青光眼筛查结果;DRIU方法的视盘位置与人工标注有偏差。相对而言,本文方法能够更准确地从病变区域中识别视盘,分割结果也更接近专家标注。

如表2所示,在RIM-ONE v. 3数据集上,本文方法的MIoU与Dice系数分别达到91.7与96.3%,Acc与PC分别为99.7%,95.7%,比主干网络分别提高了5.6%,3.2%,0.2%和1.5%。由表1与表2所展示的OD提取对比数据可见,本文方法有效提高了网络分割精度,与人

表1 DRIONS-DB数据库OD提取结果对比

Tab.1 Comparison of OD extraction results on DRIONS-DB database (%)

Method	Dice	MIoU	PC	MSE	ACC
M-Net <sup>[6]</sup>	76.5	64.1	64.9	1.40	97.3
U-Net <sup>[10]</sup>	81.4	68.9	86.3	1.76	94.8
DRIU <sup>[16]</sup>	90.9	86.9	94.5	0.55	98.4
AttU-Net <sup>[17]</sup>	87.9	82.1	78.1	0.64	98.1
R2U-Net <sup>[18]</sup>	82.1	70.0	74.5	3.17	94.1
Hyne <sup>[19]</sup>	93.7	—	—	—	—
SAM <sup>[20]</sup>	87.9	90.7	87.3	0.61	99.4
MedSAM <sup>[21]</sup>	92.4	93.8	92.7	0.32	99.5
轻量级U-Net	94.3	88.1	90.0	0.33	99.6
CSP-Net	96.9	92.6	97.0	0.14	99.8

表2 RIM-ONE v. 3数据库OD提取结果对比

Tab.2 Comparison of the results of OD extraction on RIM-ONE v. 3 database (%)

Method	Dice	MIoU	PC	MSE	ACC
M-Net <sup>[6]</sup>	79.35	67.0	66.9	2.19	97.1
DRIU <sup>[16]</sup>	93.4	90.2	91.5	0.70	97.4
AttU-Net <sup>[17]</sup>	93.4	87.7	93.3	0.46	97.9
MEAL <sup>[22]</sup>	91.2	—	—	—	—
U-Net <sup>[10]</sup>	87.6	82.5	90.8	4.90	91.7
R2U-Net <sup>[18]</sup>	90.3	82.8	85.3	4.35	95.2
SAM <sup>[20]</sup>	92.4	83.4	90.3	0.42	99.5
MedSAM <sup>[21]</sup>	93.0	85.2	94.1	0.35	99.4
轻量级U-Net	93.1	86.1	94.2	0.42	99.5
CSP-Net	96.3	91.7	95.7	0.23	99.7

工标注的结果最接近。

图 5 展示了 RIM-One v. 3 数据集 OD 提取的可视化结果。由图可见,在病变、血管结构及光照等影响下,本文方法仍然较为精准且稳定地提取 OD,在病变眼底中能够准确识别视盘。

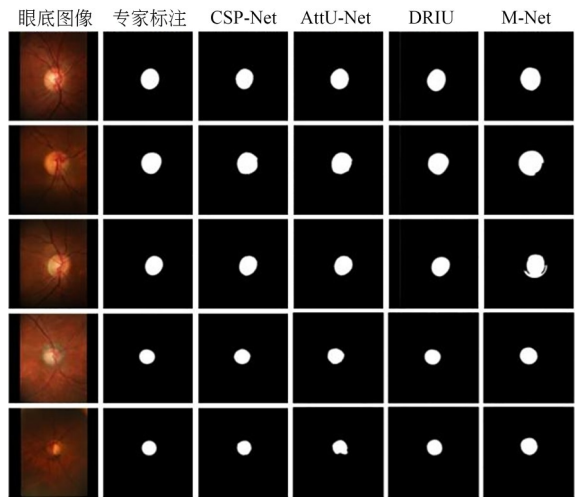


图 5 RIM-ONE v. 3 数据集 OD 提取的可视化结果对比  
Fig. 5 Comparison of OD extraction results on RIM-ONE v. 3 dataset

在 OC 提取任务中,本文方法在 DRISHTI-GS 和 RIM-ONE v. 3 数据集上与 M-Net,主干网络, MEAL<sup>[22]</sup>, pOSAL<sup>[23]</sup>, FCN<sup>[24]</sup>, SegNet<sup>[25]</sup>, ESS-Net<sup>[26]</sup>, Edge TPU<sup>[27]</sup>, SLSR-Net<sup>[28]</sup>, NAS-U<sup>2</sup>-Net<sup>[29]</sup>, SAM 和 MedSAM 方法进行了对比,结果如表 3 所示。由表可见,本文方法的 MIoU 和 Dice 点数分别为 85.4% 和 92.1%,显著优于其他方法。

由表 4 可见,在 RIM-ONE v. 3 数据集 OC 提取对比实验中,本文方法仍获得了最高的相似度,预测结果与真实标签最接近。以上结果表明,CPS-Net 可以准确地对 OD/OC 进行自动提取,为青光眼的诊断与筛查提供可靠依据。

图 6 为不同方法在 RIM-ONE v. 3 数据集上的 OC 提取可视化结果对比。可以看出,pOSLA 算法在 RIM-ONE V3 数据集上出现了过分割的情况;M-Net 则由于可迁移性不强,实验结果也不理想。视杯提取的难点在于视杯包含在视盘中,且二者特征相似难以区分,而通道注意力使网络更专注与关键信息相关的通道,增强了 CPS-Net 的目标提取能力。以图 6 第 4 行为例,

表 3 DRISHTI-GS 数据库 OC 提取结果对比

Tab. 3 Comparison of OC extraction results in DRISHTI-GS database (%)

Method	Dice	MIoU	PC	MSE	ACC
M-Net <sup>[6]</sup>	81.9	71.2	89.3	0.49	98.6
SAM <sup>[20]</sup>	84.1	62.8	86.6	5.58	87.4
MedSAM <sup>[21]</sup>	86.2	73.7	88.4	1.58	97.8
MEAL <sup>[22]</sup>	87.0	—	—	—	—
pOSLA <sup>[23]</sup>	77.9	63.7	65.8	0.38	97.4
FCN-8 <sup>[24]</sup>	81.3	69.3	73.2	1.49	83.4
FCN-16 <sup>[24]</sup>	84.1	73.7	75.8	1.12	94.7
FCN-32 <sup>[24]</sup>	85.9	75.6	77.6	<b>0.86</b>	95.4
SegNet <sup>[25]</sup>	80.5	68.1	72.9	1.54	93.7
ESS-Net <sup>[26]</sup>	90.9	—	—	—	<b>99.7</b>
SLSR-Net <sup>[28]</sup>	89.5	—	—	—	<b>99.7</b>
NAS-U <sup>2</sup> -NET <sup>[29]</sup>	89.6	—	—	—	—
轻量级 U-Net	88.1	79.6	<b>92.1</b>	5.32	94.6
CSP-Net	<b>92.1</b>	<b>85.4</b>	90.3	3.05	96.9

表 4 RIM-ONE v. 3 数据库 OC 提取结果对比

Tab. 4 Comparison of OC extraction results on RIM-ONE v. 3 database (%)

Method	Dice	MIoU	PC	MSE	ACC
M-Net <sup>[6]</sup>	70.7	57.5	61.1	0.66	98.6
pOSLA <sup>[23]</sup>	77.4	63.2	70.7	<b>0.23</b>	98.1
MEAL <sup>[22]</sup>	82.1	—	—	—	—
SegNet <sup>[25]</sup>	76.2	62.4	82.9	3.72	90.9
FCN-8 <sup>[24]</sup>	84.3	73.6	<b>88.4</b>	2.43	93.9
FCN-16 <sup>[24]</sup>	84.5	74.0	85.9	2.35	93.9
FCN-32 <sup>[24]</sup>	80.6	68.7	82.6	5.85	92.3
ESS-Net <sup>[26]</sup>	85.63	—	—	—	<b>99.6</b>
Edge TPU <sup>[27]</sup>	84.0	—	—	—	—
SLSR-Net <sup>[28]</sup>	85.8	—	—	—	99.5
轻量级 U-Net	81.5	68.3	82.8	2.46	97.3
CSP-Net	<b>86.0</b>	<b>74.4</b>	86.6	2.20	97.5

虽然该图 OC 区域极小且特征不明显,但本文方法仍更接近专家标注,最大程度地减少 CDR 的计算误差。图 7 为不同方法在 DRISHIT-GS 数据集可视化 OC 提取对比。由图可见,本文方法未受血管网络影响,实现了精准分割。

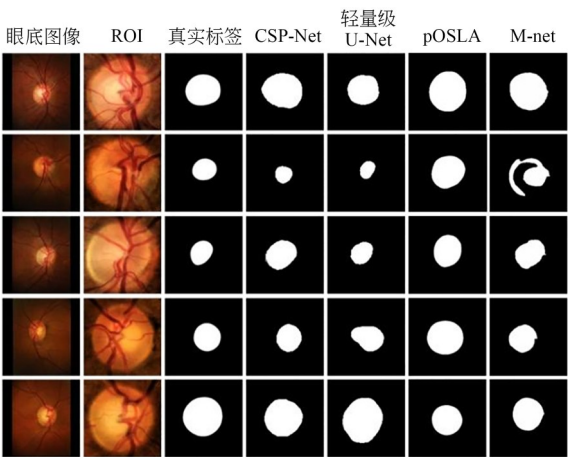


图6 RIM-ONE v. 3数据集的OC可视化结果对比  
Fig.6 Comparison of OC extraction results on RIM-ONE v. 3 dataset

3.4 消融实验及结果分析

本文方法在不同数据集上进行了消融实验,以验证CSP模块中所有模块的作用。实验在主干网络的基础上,分别加入了通道-空间联合注意力、金字塔池化结构以及CSP模块。实验环境、评估指标以及超参数均相同。

如表5和表6所示,将仅加入注意力机制的网络记为+Attn,仅加入金字塔池化结构的网络记为+PPM。由表可知,以RIM-ONE v. 3数据

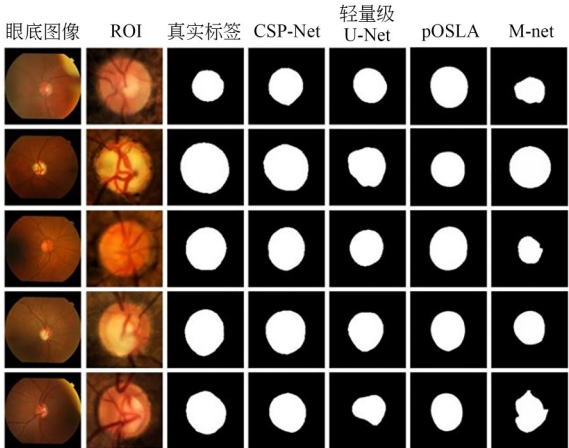


图7 DRISHTI-GS数据集的OC可视化结果对比  
Fig.7 Comparison of OC extraction results on DRISHTI-GS dataset

集的实验结果为例,使用主干网络提取OD的MIoU为86.1%,Dice系数为93.1%,OC提取的MIoU为68.3%,Dice为81.5%。将通道-空间联合注意力(Attn)与金字塔池化结构中的任何一个集成,都可以提高性能指标。将注意力机制引入金字塔池化结构后,OD的提取性能达到最高,MIoU为91.7%,比主干网络高5.6%;Dice为96.3%,比主干网络高3.2%。OC提取的MIoU可达74.4%,比主干网络高6.1%;Dice系数为86.0%,比主干网络高4.5%。

表5 OD提取消融实验结果对比

Tab. 5 Comparison of ablation experiment results for OD extraction task (%)						
模 块	DRIONS-DB			RIM-ONE v. 3		
	MIoU	Dice	MSE	MIoU	Dice	MSE
轻量级 U-net	88.1	94.3	0.33	86.1	93.1	0.42
t+Attn	91.2	95.4	0.16	90.5	95.7	0.26
t+PPM	89.5	95.1	0.21	90.8	95.8	0.28
CSP-Net	92.6	96.6	0.14	91.7	96.3	0.23

表6 OC提取消融实验结果对比

Tab. 6 Comparison of the ablation experiment results for OC extraction task (%)						
模 块	DRISHTI-GS			RIM-ONE v. 3		
	MIoU	Dice	MSE	MIoU	Dice	MSE
轻量级 U-net	77.5	85.2	6.32	68.3	81.5	2.46
t+Attn	80.1	88.3	4.31	70.6	83.9	3.22
t+PPM	82.1	89.1	3.80	70.0	83.6	3.16
CSP-Net	85.4	92.1	3.05	74.4	86.0	2.20

## 4 结 论

本文以更轻量级的编码器-解码器模型作为主干网络,加入基于注意力机制的多尺度语义目标提取 CSP 模块,对早期层特征进行进一步处理,并与上采样路径中的后期层特征进行拼接,在减轻语义鸿沟的同时,更加充分地利用来自每

一层的多尺度信息,有效提升了目标提取的精度。针对提取视杯盘任务,在 DRIONS-DB, RIM-ONE v. 3 及 DRISHTI-GS 3 个公开数据集上与现有方法进行对比。实验结果表明,本文提出的 CSP-Net 在 OD 和 OC 的提取任务中精度均为最优,且在处理病变区域、视杯特征不明显等区域具有很强的竞争力。

## 参考文献:

- [1] 于洋,蒋沁,曹国凡. 高度近视合并青光眼的临床诊断研究进展[J]. 国际眼科杂志, 2021, 21(6): 1008-1011.  
YU Y, JIANG Q, CAO G F. Progress in the clinical diagnosis of high myopia combined with glaucoma [J]. *International Eye Science*, 2021, 21 (6): 1008-1011. (in Chinese)
- [2] SIDDIQUEE M S, PATHAN N S. Optic disc segmentation using superpixel based features and random forest classifier[C]. 2019 4th International Conference on Electrical Information and Communication Technology. Khulna, Bangladesh EICT, 2019: 1-5.
- [3] REHMAN Z U, NAQVI S S, KHAN T M, et al. Multi-parametric optic disc segmentation using superpixel based feature classification [J]. *Expert Systems With Applications*, 2019, 120: 461-473.
- [4] SINGH R U, GUJRAL S. Assessment of disc damage likelihood scale (DDLS) for automated glaucoma diagnosis [J]. *Procedia Computer Science*, 2014, 36: 490-497.
- [5] YU T, MA Y, LI W. Automatic localization and segmentation of optic disc in fundus image using morphology and level set[C]. 9th International Symposium on Medical Information and Communication Technology, Kamakura, Japan. ISMICT, 2015: 195-199.
- [6] FU H Z, CHENG J, XU Y W, et al. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation[J]. *IEEE Transactions on Medical Imaging*, 2018, 37 (7): 1597-1605.
- [7] LIU B Y, PAN D R, SONG H. Joint optic disc and cup segmentation based on densely connected depthwise separable convolution deep network [J]. *BMC Medical Imaging*, 2021, 21(1): 1-12.
- [8] 于舒扬,袁鑫,郑秀娟. 融合感受野模块的卷积神经网络视杯视盘联合分割[J]. 中国生物医学工程学报, 2022, 41(2): 167-176.  
YU SH Y, YUAN X, ZHENG X J. Joint optic cup and disc segmentation using convolutional neural network with receptive field module [J]. *Chinese Journal of Biomedical Engineering*, 2022, 41 (2): 167-176. (in Chinese)
- [9] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017. Honolulu, HI. IEEE, 2017: 6230-6239.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional Networks for Biomedical Image Segmentation[M]. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234-241.
- [11] SEVASTOPOLSKY A. Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network [J]. *Pattern Recognition and Image Analysis*, 2017, 27(3): 618-624.
- [12] 黄文博,黄钰翔,姚远,等. 融合注意力的 ConvNeXt 视网膜病变自动分级[J]. 光学精密工程, 2022, 30(17): 2147-2154.  
HUANG W B, HUANG Y X, YAO Y, et al. Automatic classification of retinopathy with attention ConvNeXt [J]. *Optics and Precision Engineering*, 2022, 30(17): 2147-2154. (in Chinese)
- [13] CARMONA E J, RINCÓN M, GARCÍA-FEIJÓO J, et al. Identification of the optic nerve head with genetic algorithms [J]. *Artificial Intelligence in Medicine*, 2008, 43(3): 243-259.
- [14] FUMERO F, ALAYON S, SANCHEZ J L, et al. RIM-ONE: an open retinal image database for optic nerve evaluation [C]. 2011 24th International Symposium on Computer-Based Medical Systems



- (CBMS). June 27-30, 2011. Bristol, United Kingdom. IEEE, 2011: 1-6.
- [15] SIVASWAMY J, KRISHNADAS S R, DATT JOSHI G, *et al.* Drishti-GS: Retinal image dataset for optic nerve head (ONH) segmentation [C]. 2014 *IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. April 29-May 2, 2014. Beijing, China. IEEE, 2014: 53-56, .
- [16] MANINIS K K, PONT-TUSET J, ARBELÁEZ P, *et al.* *Deep Retinal Image Understanding* [M]. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016. Cham: Springer International Publishing, 2016: 140-148.
- [17] OKTAY O. Attention U-Net: learning where to look for the pancreas [C/OL]. <https://openreview.net/forum?id=Skft7cijM>.
- [18] ALOM M Z, YAKOPCIC C, TAHA T M, *et al.* Nuclei segmentation with recurrent residual convolutional neural networks based U-net (R2U-net) [C]. *NAECON 2018 - IEEE National Aerospace and Electronics Conference*. July 23-26, 2018. Dayton, OH, USA. IEEE, 2018: 228-233.
- [19] WEN Y, CHEN L T, QIAO L F, *et al.* An efficient weakly-supervised learning method for optic disc segmentation [C]. 2020 *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. December 16-19, 2020. Seoul, Korea (South). IEEE, 2020: 835-842.
- [20] KIRILLOV A, MINTUN E. Segment anything [J]. *arXiv e-prints*. abs/2304.02643.
- [21] MA J, HE Y, LI F, *et al.* *Segment Anything in Medical Images* [EB/OL]. 2023: *arXiv*: 2304.12306. <https://arxiv.org/abs/2304.12306.pdf>.
- [22] MA B Q, YANG Q, CUI H, *et al.* MEAL: me-  
ta enhanced entropy-driven adversarial learning for optic disc and cup segmentation [C]. 2021 *43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. November 1-5, 2021. Mexico. IEEE, 2021: 3273-3276.
- [23] WANG S J, YU L Q, YANG X, *et al.* Patch-based output space adversarial learning for joint optic disc and cup segmentation [J]. *IEEE Transactions on Medical Imaging*, 2019, 38(11): 2485-2495.
- [24] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.
- [25] 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.
- [26] HAIDER A, ARSALAN M, PARK C, *et al.* Exploring deep feature-blending capabilities to assist glaucoma screening [J]. *Applied Soft Computing*, 2023, 133: 109918.
- [27] CIVIT-MASOT J, LUNA-PEREJÓNF. A study on the use of Edge TPUs for eye fundus image segmentation [J]. *Engineering Applications of Artificial Intelligence*, 2021, 104: 104384.
- [28] HAIDERA, ARSALANM. Artificial Intelligence-based computer-aided diagnosis of glaucoma using retinal fundus images [J]. *Expert Systems With Applications*, 2022, 207: 117968.
- [29] SUN J D, YAO C, LIU J, *et al.* GNAS-U<sup>2</sup>Net: a new optic cup and optic disc segmentation architecture with genetic neural architecture search [J]. *IEEE Signal Processing Letters*, 2022, 29: 697-701.

## 作者简介:



燕 杨(1981—),女,汉族,博士,副教授,从事机器学习与模式识别方向的研究。E-mail: yanyang2016@hotmail.com



曹娅迪(1997—),女,汉族,硕士研究生,主要从事医学影像自动分割算法的研究。E-mail: AntOnia@163.com