

## 基于改进 SOLO\_v2 的糖尿病黄斑水肿分割模型

郑宗生, 唐鹏飞, 王振华, 卢鹏  
上海海洋大学信息学院, 上海 201306

**【摘要】**糖尿病黄斑水肿(DME)是导致糖尿病患者视力损害的常见原因。光学相干断层扫描技术(OCT)有助于增强对糖尿病视网膜病变的早期检测和预防。目前,OCT图像中的DME区域存在大量散斑噪声及小目标区域,现有的实例分割方法存在漏分割等问题。针对上述问题,本文利用特征金字塔转换器(FPT)改进 SOLO\_v2 模型,提出了一种新的DME分割模型(SOLO-OCT),包括:(1)利用基于双域滤波去噪算法去除图像上存在的大量散斑噪声,提高输入图像质量;(2)引入FPT,提高模型对小目标的识别能力和学习能力;(3)改进非极大值抑制(NMS)算法,缓解对小目标区域的漏分割问题。将 SOLO-OCT 模型与其他实例分割模型(包括Mask R-CNN、SOLO和 SOLO\_v2)进行了比较,以评估其对DME区域的分割性能。与Mask R-CNN、SOLO和 SOLO\_v2 模型相比,SOLO-OCT模型对DME区域的分割精度(mAP)提高了3.1%,对小目标DME区域的分割精度(APs)提高了2.2%,而单幅图像的处理时间(Fps)只增加了0.009 9 s。本文提出的DME分割模型(SOLO-OCT)可用于大规模糖尿病视网膜病变筛查。

**【关键词】**糖尿病黄斑水肿;实例分割;特征增强;非极大值抑制

**【中图分类号】**R318;TP391.41

**【文献标志码】**A

**【文章编号】**1005-202X(2023)01-0024-07

### A novel model for diabetic macular edema segmentation based on improved SOLO\_v2

ZHENG Zongsheng, TANG Pengfei, WANG Zhenhua, LU Peng

College of Information Technology, Shanghai Ocean University, Shanghai 201306, China

**Abstract:** Diabetic macular edema (DME) is a common cause of visual impairment in diabetic patients. Optical coherence tomography (OCT) can enhance the early detection and prevention of diabetic retinopathy. At present, there are a lot of speckle noises and small target areas in the DME region in OCT images, and the existing instance segmentation methods have some problems such as missing segmentation. To address the above issues, SOLO\_v2 model is improved by feature pyramid transformer, and a novel model (SOLO-OCT model) is proposed for DME segmentation. The proposed method improves the quality of the input image by removing the speckle noises from the image using dual-domain filtering algorithm, and enhances the model's ability to recognize and learn small target areas by feature pyramid transformer, and alleviates the problem of missing segmentation for small target areas through improved non-maximum suppression. The SOLO-OCT model is compared with other instance segmentation models (Mask R-CNN, SOLO and SOLO\_v2) to evaluate its performance in DME segmentation. Compared with Mask R-CNN, SOLO and SOLO\_v2 models, SOLO-OCT model improves the segmentation accuracy of DME region (mAP) by about 3.1% and raises the segmentation accuracy of small-target DME region (APs) by about 2.2%, but the processing time of a single image (Fps) is increased by only about 0.009 9 s. The proposed SOLO-OCT model for DME segmentation can be used for large-scale screening for diabetic retinopathy.

**Keywords:** diabetic macular edema; instance segmentation; feature enhancement; non-maximum suppression

### 前言

糖尿病黄斑水肿(Diabetic Macular Edema, DME)是指由糖尿病引起在黄斑中心凹一个视盘直径范围内的细胞外液积聚所致的视网膜增厚或硬性渗出沉积<sup>[1-2]</sup>。光学相干断层扫描技术(Optical Coherence Tomography, OCT)是近年来发展较快的一种新型层析成像技术<sup>[3]</sup>,在生物组织活体检测和成

**【收稿日期】**2022-07-15

**【基金项目】**国家自然科学基金(41671431);国家海洋局数字海洋科学技术重点实验室开放基金(B201801034);上海市科委地方能力建设项目(19050502100);上海海洋大学科技发展专项基金(A2-2006-20-200211)

**【作者简介】**郑宗生,博士,副教授,研究方向:深度学习、遥感图像处理,  
E-mail: zszheng@shou.edu.cn

像方面具有极高的应用价值,尤其在视网膜疾病的诊断和视网膜内部结构的检测中发挥重大作用<sup>[4-7]</sup>。如何在OCT图像中快速、准确地提取出DME区域成为研究热点之一。

随着深度学习在图像分割方面的不断发展,越来越多图像分割算法被提出并应用于医学图像分割之中。早期主要通过滑动图像块的方式实现<sup>[8]</sup>:在目标像素周围截取固定大小的图像块送入卷积神经网络(Convolutional Neural Networks, CNN),得到的分类结果即为当前像素点所属类别,图像块滑动方法存在大量的重复计算操作且运行效率低下,同时分割精度直接受图像块大小限制,具有一定局限性;Long等<sup>[9]</sup>提出全卷积网络(Fully Convolutional Networks, FCN),其将分类网改写为用于分割的像素点分类网,在一定程度上提高了分割精度,但对像素与像素之间的关系并没有考虑到,忽略了在基于像素分类的分割方法中使用的空间规整步骤,缺乏空间一致性;Chen等<sup>[10]</sup>提出的Deep-Lab V3+网络借鉴了特征图金字塔网络(Feature Pyramid Networks, FPN)等网络的Encoder-decoder架构,实现了特征匹配跨区域的融合并使用了分组卷积来加速,但是其是在降采样8倍的尺度上进行预测的,导致了边界效果不甚理想;Ronneberger等<sup>[11]</sup>提出U-Net网络,其在医学分割方面得到广泛应用,U-Net网络同样基于编码器-解码器思想构建,采用跳跃连接方式直接融合上下采样结构中的高层和底层信息,用于提升分割结果细节精度的同时使U-Net具有更好的可解释性,但其使用的有效卷积增加了模型设计的难度和普适性,且其对图像上存在的散斑噪声不能有效去除,对存在的小目标区域不能有效地进行识别和分割。

由于OCT图像存在散斑噪声,限制了OCT图像的对比度和信噪比,直接观察或使用分割算法分割DME会造成困难<sup>[12]</sup>;由此本文引入以ResNet-101作为骨干网络的SOLO\_v2模型为主体框架<sup>[13]</sup>,在预处理阶段,引入基于双域滤波算法去噪方法<sup>[14]</sup>,对OCT图像中的散斑噪声进行去除,以此提高训练图像的质量;同时,考虑到实例分割需要综合不同层的上下文进行预测且对小目标区域的检测分割需要借鉴特征图上的其他特征点进行当前特征点增强,本文引入特征金字塔转换器(Feature Pyramid Transformer, FPT)改进原有的特征金字塔结构以提高模型对小目标的识别能力和学习能力<sup>[15]</sup>;并改进了非极大值抑制(Non-Maximum Suppression, NMS)方法,以此缓解对小目标区域的漏分割问题且提高对DME检测的精度,为DME的检测提供信息支持。

## 1 相关工作

### 1.1 基于双域滤波去噪算法

基于双域滤波的OCT图像去噪算法由高乃珺等<sup>[14]</sup>于2019年提出,其基于Knaus等<sup>[16]</sup>提出的通过结合传统的空域与变换域去噪算法优点的双域去噪算法,并结合OCT图像存在的大量散斑噪声的乘性特性建立的乘性噪声模型,从而提出一种简单但高质量的双域滤波去噪算法。OCT图像中存在的大量散斑噪声主要是由光线发生干涉时引入的乘性噪声,所以根据乘性噪声的特点来建立模型<sup>[17-18]</sup>:

$$y = x * \rho_m + \rho_a \quad (1)$$

其中, $y$ 为带有噪声的图像; $x$ 为输入的原始噪声图像; $\rho_m$ 和 $\rho_a$ 分别是散斑噪声和加性噪声。由于加性噪声与乘性噪声相比非常小,因此可以在此忽略不计。从而产生适用于基于双域滤波算法的OCT图像去噪模型;基于此模型可以得到去噪效果良好的OCT图像,将其作为后续训练的输入图像。

### 1.2 SOLO\_v2模型

SOLO\_v2模型旨在建立一个性能强大且简单、直接、快速的实例分割框架,模型整体保持了SOLO\_v1的网络结构。如图1所示,SOLO\_v2与YOLO类似<sup>[19]</sup>,对于输入图像,从概念上将其均分为 $S \times S$ 的网格,通过FCN与特征金字塔网络预测分类与特征提取<sup>[20]</sup>,并输出特征图I的高度H、宽度W及通道数E,然后判别实例的中心是否落入某个网格中,满足条件的网格将进入语义分支与掩膜分支,分别对对应的实例进行语义以及掩膜大小及位置的判别。

SOLO\_v2保持了SOLO\_v1算法高效的分割速度,并做了系列优化。提出动态学习对象分割器的头部分支(Mask head),这样使Mask head依赖于位置,具体来说,将掩码分支分解为卷积核分支(Mask kernel)和特征分支(Mask feature),分别学习卷积核和卷积特征;提出非极大值抑制矩阵(Matrix Non Maximum Suppression, Matrix NMS)以显著减少由于掩码的NMS造成的推理时间开销,Matrix NMS一次性使用并行矩阵操作执行NMS,并产生了更好的结果。

### 1.3 FPT

FPT是由Zhang等<sup>[15]</sup>于2020提出的一种新颖的特征金字塔强化网络。结构如图2所示,不同的纹理模式表示不同的特征变换器,不同的颜色表示不同比例的特征贴图。“Conv”表示输出维数为256的 $3 \times 3$ 卷积。输入为金字塔特征,首先对每层特征分别进

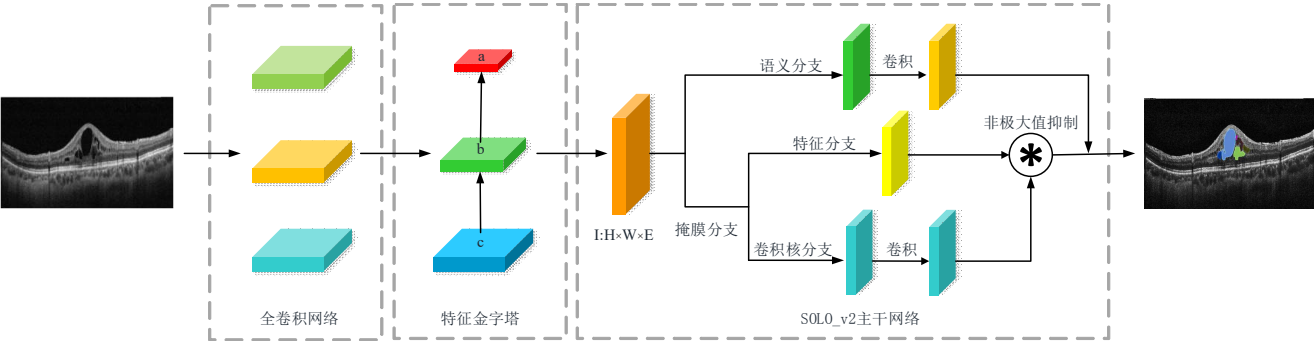


图 1 SOLO\_v2网络结构图  
Figure 1 Framework diagram of SOLO\_v2

行 ST、GT、RT 特征增强得到多个增强后的特征<sup>[21]</sup>，特征拼接到一起，通过卷积将增强后的特征维度恢复到输入时的相同维度，最后送入到头部网络。

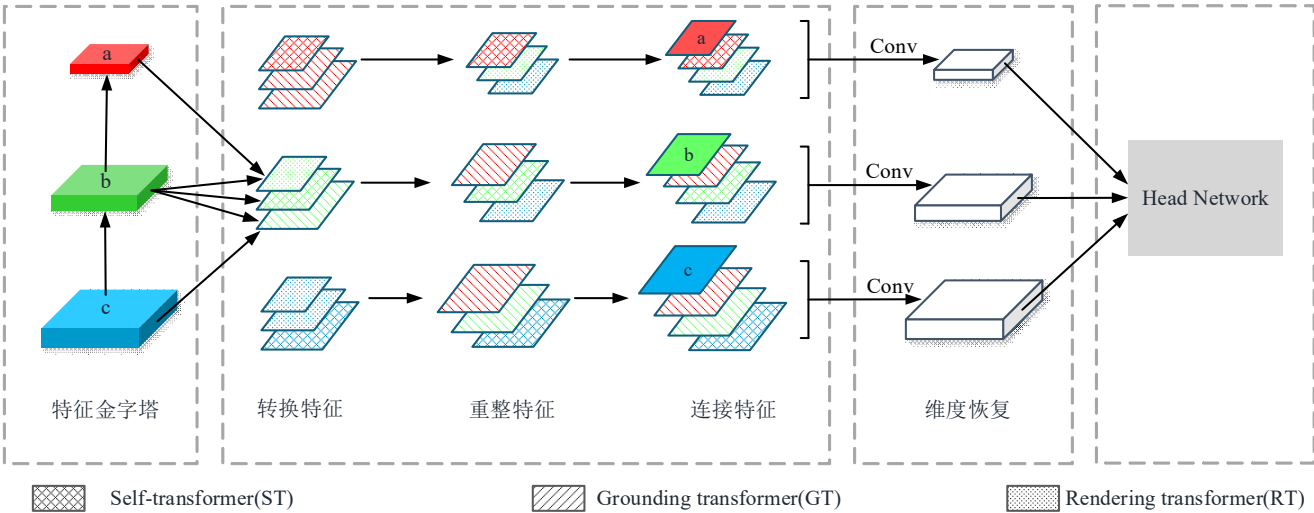


图 2 FPT网络结构图  
Figure 2 Framework diagram of FPT

ST:在同一张特征图上捕获共同发生的对象特征;GT:进行自上而下的非局部交互,借用高层的粗粒度特征来增强低层的细粒度特征,输出的特征图和低层的细粒度特征尺度相同;RT:以自下而上的方式工作,借用低层的细粒度特征来增强高层的粗粒度特征

FPT 核心思想是在特征金字塔进行自注意力增强,然后使用多层特征进行预测。其输入是一个特征金字塔,而输出是一个变换的金字塔,其中每个层级都是一个更丰富的特征图,它编码了跨空间和跨尺度的非局部交互作用。

## 2 方法

图 3 为改进的 SOLO\_v2 模型 (SOLO-OCT 模型),包括:(1)在预处理阶段加入基于双域滤波去噪算法来去除 OCT 图像上的散斑噪声,提高输入图像质量;(2)利用 FPT 提高对小目标的检测、分割能力;(3)改进原有的 NMS 算法,在后处理阶段有效地降低了小目标漏分割和被重复检测的问题,提高图像分割的精度。

### 2.1 SOLO-OCT 模型

SOLO\_v2 模型在对图像做预处理时,并没有包含针对 OCT 图像上存在的散斑噪声方法,这导致输入的原始 OCT 图像上存在大量的散斑噪声,对图像的识别和分割造成影响。在预处理阶段引入基于双域滤波去噪算法,可以有效地去除图像上存在的大量散斑噪声,提高输入图像的质量,使 DME 区域的轮廓更加清晰,有利于后续深度学习模型的训练与学习。

为提升 SOLO\_v2 模型对小目标的识别能力,在 SOLO\_v2 模型的特征金字塔层后新增一个 FPT,改进后的网络模型结构如图 3 所示。图 3 中红色虚线框内为新增的 FPT。原有的特征金字塔对于一些像素级分割任务,比如实例分割,无法综合不同层的上



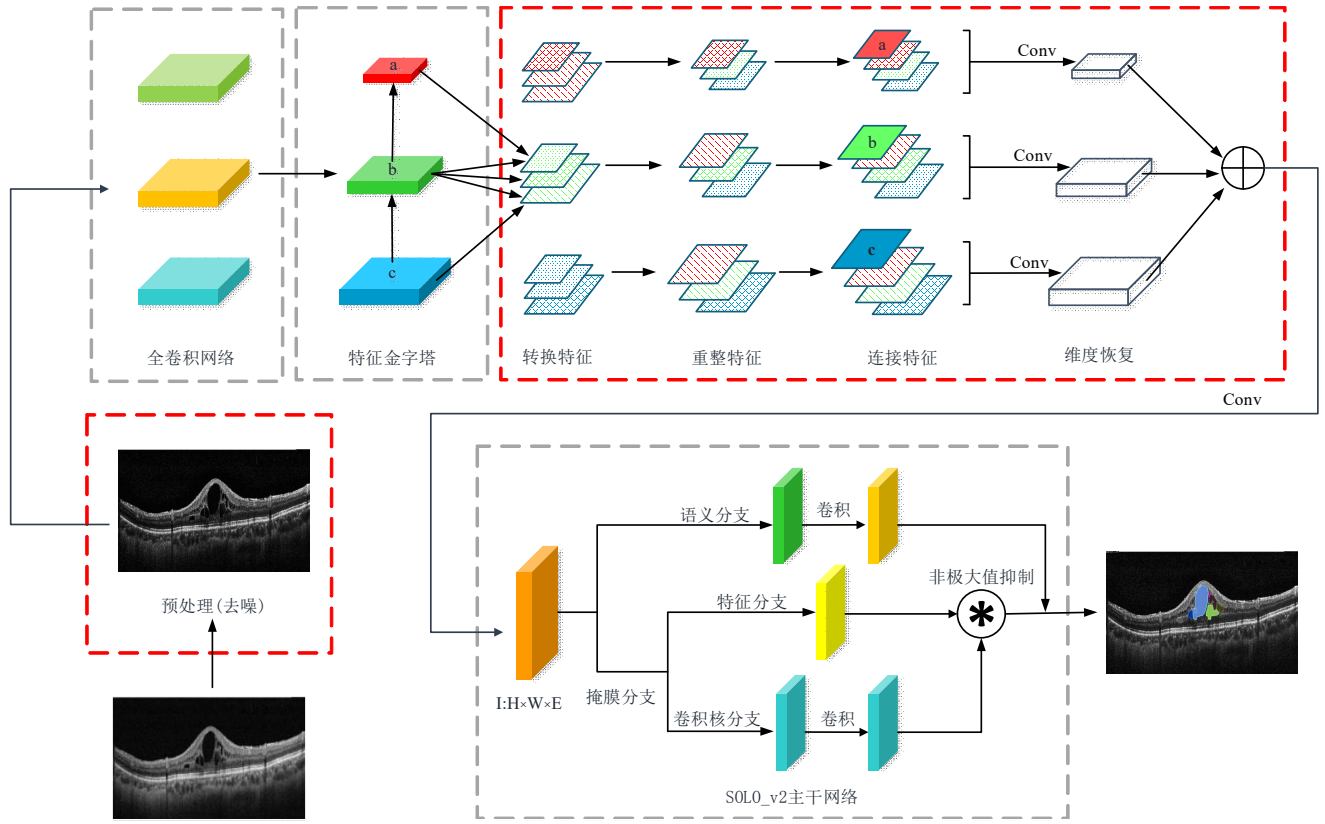


图3 改进的SOLO\_v2网络结构图

Figure 3 Framework diagram of improved SOLO\_v2

下文信息进行细致的预测<sup>[22-23]</sup>,导致图像在处理分割时会丢失部分细节信息,特别是针对小目标容易存在漏分割或错分割现象。SOLO-OCT模型利用FPT在特征金字塔进行自注意力增强,然后使用多层特征进行预测。在SOLO-OCT模型中,结合了不同尺度图像上的特征进行特征融合,以更明确和主动的方式对非局部上下文特征进行建模,并能够借鉴特征图上的其它特征点来对当前特征点进行增强。

## 2.2 改进的NMS算法

SOLO\_v2模型采用Matrix NMS来筛除冗余检测框,其以Soft NMS为基础<sup>[24]</sup>,实现了交并比(Intersection over Union, IoU)矩阵的并行化运算<sup>[25]</sup>,大大提高了算法运行的效率。其抑制函数如式(2)所示:

$$s_f = \begin{cases} s_i, & \text{IoU}(M, b_i) < N_t \\ s_i \times (1 - \text{IoU}(M, b_i)), & \text{IoU}(M, b_i) \geq N_t \end{cases} \quad (2)$$

其中,  $b_i$  为第  $i$  个待检测框,  $s_i$  为  $b_i$  原始得分,  $s_f$  为  $b_i$  的最后得分,  $M$  为每一轮得分最高的检测框,  $N_t$  为预设阈值,  $\text{IoU}(M, b_i)$  为  $b_i$  和  $M$  的交并比。可以看出, IoU 是其唯一考量因素;然而在实际应用场景中,当两个不同物体挨得很近时,由于 IoU 值比较大,往往经过 NMS 处理后,只剩下一个检测框,导致漏检的错误情

况发生。

基于此,本文在处理 NMS 算法时:(1)加入计算真实框与预测框中心点之间的距离;(2)使用 log 函数对检测框的得分  $s_i$  进行非线性的权值惩罚。改进后的抑制函数如式(3)所示:

$$s_f = \begin{cases} s_i, & \text{DIOU}(M, b_i) < N_t \\ s_i \times (1 - \log \text{DIOU}(M, b_i)), & \text{DIOU}(M, b_i) \geq N_t \end{cases} \quad (3)$$

其中:

$$\text{DIOU} = \text{IoU} - \frac{\rho^2(b, b^{gt})}{c^2} = \text{IoU} - \frac{d^2}{c^2} \quad (4)$$

其中,  $b$  为预测框的中心,  $b^{gt}$  为真实框的中心,  $\rho^2(b, b^{gt})$  代表真实框与预测框中心距离的平方  $d^2$ ,  $c$  表示两个框的最小闭包区域的对角线长度。因此,在计算检测框得分时,如果两个框之间 IoU 比较大,但是两个框的中心距离也比较大时,可能会认为这是两个物体的框而不会被过滤掉,从而避免漏检情况的产生;并且每个检测框的得分呈非线性变化,目标区分度更高,保留下来的检测框更准确,从而提高后续图像分割的精度。

## 3 实验设计

### 3.1 数据集

实验数据选取于带有 DME 的原始 OCT 图像。

如图4所示。原始数据为1 583张图像,大小为941像素×425像素,其中绝大部分图像中含有小目标水肿区域。为提升模型泛化性能和防止模型出现过拟合现象,将原始图像进行增强,目标扩充和旋转等操作

作,将实验数据集扩充至6 125张图像。用LabelMe软件对水肿区域进行标注,在6 125张图像中标注水肿区域24 341个。

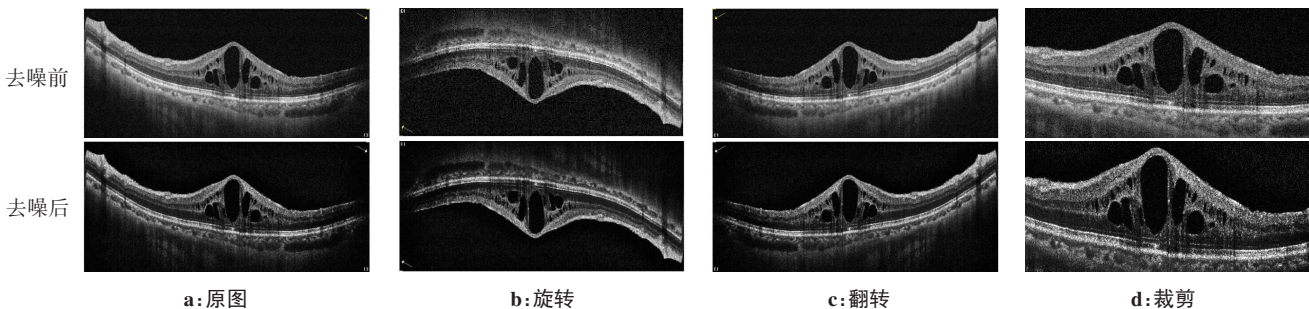


图4 数据集图像示例

Figure 4 Image examples from data set

a:原始OCT图像,未进行旋转、翻转、裁剪等操作;b:经过旋转后的OCT图像;c:经过翻转后的OCT图像;d:经过裁剪后的OCT图像

3.2 实验环境

实验硬件配置为Inter Core i7-9700 CPU处理器,显卡为GeForce GTX 1660Ti 6 GB显存的服务器。操作系统为64位Ubuntu16.04。深度学习框架为pytorch。

3.3 评价指标

选用精准率(AP)、单幅图像处理时间(Fps)和平均精确率(mAP)对实例分割模型的性能进行评价。AP<sub>0.5</sub>与AP<sub>0.75</sub>表示IoU阈值分别取0.5与0.75时的AP指标;mAP表示IoU阈值从0.5到0.95(步长为0.05)所对应AP的平均值;APs表示对小目标物体(分割掩码中像素数量少于322)分割后的AP。其计算公式如式(5)、(6)所示:

AP = TP / (TP + FP) (5)

mAP = (sum from k=1 to C of AP(k)) / C (6)

其中,TP表示模型将负样本预测为正样本的个体数量;FP表示模型将正样本预测为正样本的个体数量;C为数据集中的类别总数,AP(k)为第k个类的平均精准率。

4 实验对比与分析

本文使用SOLO-OCT模型对OCT图像中的DME区域进行分割。图5给出了对OCT图像分割的代表性结果。首先,在预处理阶段使用基于双域滤波去噪算法对原始输入OCT图像(图5a)进行去噪,得到噪声较小及轮廓清晰的OCT图像(图5b);然后送入后续主干网络将DME区域完整分割出来(图5c)。

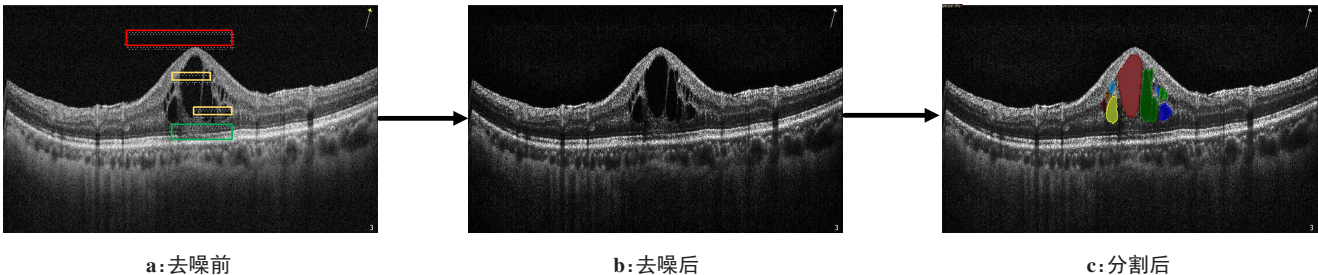


图5 对原始的OCT图像进行分割,得到最后分割的DME区域

Figure 5 Obtaining the final segmented DME region by the original OCT image segmentation

a:原始OCT图像,不同颜色的方框展示了图像上存在于不同位置的散斑噪声;b:预处理阶段使用基于双域滤波去噪算法对原始OCT图像去噪;c:SOLO-OCT模型对DME区域进行分割,不同的颜色表示对识别到的不同的DME区域进行分割

4.1 去噪前后分割结果对比

对OCT数据集去噪前后分割性能比较。如表1

所示,两个实验皆以ResNet-101-FPT为骨干网络;当OCT数据集进行去噪后,mAP提高了0.8%,而Fps基

本相同;可以看出 OCT 图像上存在的散斑噪声对 DME 区域的分割存在着一定影响,合适的去噪方法可以提高对 OCT 图像分割的精度。

表 1 相同骨干网络去噪前后的性能实验结果对比  
Table 1 Comparison of the performance of the same backbone network before and after noise reduction

是否去噪	AP <sub>0.5</sub> /%	AP <sub>0.75</sub> /%	mAP/%	APs/%	Fps/s
去噪前	70.013	45.387	42.432	18.037	0.131 6
去噪后	71.167	46.263	43.232	18.051	0.132 1

表 2 相同骨干网络改进NMS前后的性能实验结果对比  
Table 2 Comparison of performance experiment results after improving NMS on the same backbone network

NMS算法	骨干网络	AP <sub>0.5</sub> /%	AP <sub>0.75</sub> /%	mAP/%	APs/%	Fps/s
改进前	ResNet-101-FPT	70.813	44.987	42.132	16.751	0.129 6
改进后	ResNet-101-FPT	71.167	46.263	43.232	18.051	0.132 1

4.3 不同网络模型的性能实验结果对比

本文比较了 SOLO-OCT 模型与其他传统实例分割模型(包括 Mask R-CNN、SOLO、SOLO\_v2)的分割性能。如图 6 所示,Mask R-CNN 模型整体分割效果较差。SOLO 模型对 DME 区域边缘的分割无法得到清晰的结果。而传统的 SOLO\_v2 模型对小目标的分

4.2 改进的NMS算法结果对比

改进的 NMS 算法前后分割性能比较,如表 2 所示,两个实验皆以 ResNet-101-FPT 为骨干网络;输入的数据均为经预处理去噪后的 OCT 图像。在 SOLO-OCT 模型中使用改进的 NMS 算法后,mAP 提高了 1.1%,APs 提高了 1.3%,而 Fps 只增加了 0.002 5 s;改进后的 NMS 算法在保持整个模型处理数据性能基本不变的情况下,不仅提高了分割精度,而且缓解了对小目标区域漏检的问题,能够更好地分割图像中的小目标。

割能力较差,容易造成漏分割且无法有效去除图像上的散斑噪声。相比之下,SOLO-OCT 模型不仅可以得到噪声较小且清晰的边缘分割效果,还可以成功对图像上较小的 DME 区域进行识别分割。这些结果表明 SOLO-OCT 模型适用于对 OCT 图像中的 DME 区域进行分割。

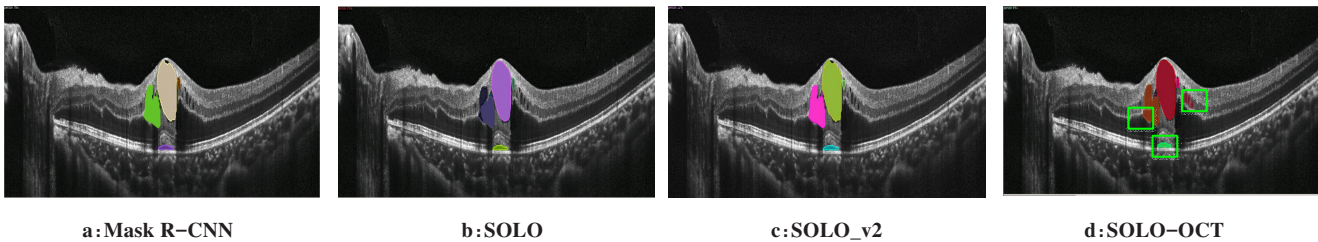


图 6 SOLO-OCT 模型与 Mask R-CNN、SOLO、SOLO\_v2 模型的分割性能比较  
Figure 6 Comparison of the segmentation performance of SOLO-OCT model with Mask R-CNN, SOLO and SOLO\_v2 models  
同一副图像中的不同颜色代表了对不同 DME 区域分割的效果;图 d 中绿色的方框可以看出 SOLO-OCT 模型优化了边缘分割的能力以及增强了对小目标的分割能力

SOLO-OCT 模型与其他实例分割模型的分割性能比较如表 3 所示,选用 ResNet-101 作为骨干网络,使用 SOLO\_v2 模型与 Mask R-CNN 和 SOLO 模型进行对比发现:相比 Mask R-CNN 模型,SOLO\_v2 模型的 AP<sub>0.5</sub>、AP<sub>0.75</sub>、mAP、APs 分别增加了 2.594%、4.866%、2.977%、3.301%;相比 SOLO,SOLO\_v2 的 AP<sub>0.5</sub>、AP<sub>0.75</sub>、mAP、APs 分别增加了 1.751%、3.451%、1.619%、3.250%。同时,SOLO\_v2 仅使用 0.122 2 s 即完成对 1 幅图像的识别与分割,单幅图像处理时间相比 Mask R-CNN 与 SOLO 模型分别减少了 47.0% 和

19%,极大地提升了模型的运行效率。综上所述,无论是从精度或是运行效率上,SOLO\_v2 模型都具有较好的效果。  
与 SOLO\_v2 模型相比,SOLO-OCT 模型的 AP<sub>0.5</sub>、AP<sub>0.75</sub>、mAP、APs 分别增加了 3.6%、2.2%、3.1%、2.2%,而在单幅图像处理时间上,SOLO-OCT 模型只增加了 0.009 9 s,虽然相较于原模型处理时间有所增加,但是增加的幅度极小,而整体的精度却有较大提升,所以 SOLO-OCT 模型对 OCT 图像上存在的 DME 区域有更好的分割效果。



表 3 不同网络模型下黄斑水肿分割性能的比较

Table 3 Performance comparison among different network models for macular edema segmentation

网络模型	骨干网络	AP <sub>0.5</sub> /%	AP <sub>0.75</sub> /%	mAP/%	APs/%	Fps/s
Mask R-CNN	ResNet-101-FPN	64.973	39.197	37.155	12.550	0.231 6
SOLO	ResNet-101-FPN	65.816	40.612	38.513	12.601	0.150 7
SOLO_v2	ResNet-50-FPN	66.073	43.508	38.755	13.367	0.091 3
	ResNet-101-FPN	67.567	44.063	40.132	15.851	0.122 2
SOLO-OCT	ResNet-101-FPT	71.167	46.263	43.232	18.051	0.132 1

5 结 论

本文针对 OCT 图像中的 DME 区域筛查存在繁琐且效率低下等现实需求问题,提出了一种改进的 SOLO\_v2 模型(SOLO-OCT)用于对 DME 的检测与分割,包括:(1)利用基于双域滤波去噪算法去除图像上存在的大量散斑噪声,提高输入图像质量;(2)引入 FPT,提高模型对小目标的识别能力和学习能力;(3)改进 NMS 算法,缓解对小目标区域的漏分割问题。实验结果表明,改进的 SOLO\_v2 模型对 DME 分割具有更高的精度和更细致的分割效果,同时提高了对小目标 DME 区域的分割能力。但是该模型对 DME 区域的边缘分割仍然不够平滑完整,下一步将对此进行改进。此外,还将研究如何减少网络参数并简化网络模型。

【参考文献】

[1] Țălu Ș, Nicoara SD. Malfunction of outer retinal barrier and choroid in the occurrence and progression of diabetic macular edema[J]. World J Diabetes, 2021, 12(4): 437-452.

[2] Ixamey M, Palma C. Diabetic macular edema[J]. DM-Dis Mon, 2021, 67(3): 101138.

[3] Yong WK, Lee J, Kim JS, et al. Diagnostic accuracy of wide-field map from swept-source optical coherence tomography for primary open-angle glaucoma in myopic eyes[J]. Am J Ophthalmol, 2020, 218: 182-191.

[4] Tayal A, Gupta J, Solanki A, et al. DL-CNN-based approach with image processing techniques for diagnosis of retinal diseases[J]. Multimedia Syst, 2022, 28: 1417-1438.

[5] Varela MD, Esener B, Hashem SA, et al. Structural evaluation in inherited retinal diseases[J]. Br J Ophthalmol, 2021, 105(12): 1623-1631.

[6] Robbins CB, Thompson AC, Bhullar PK, et al. Characterization of retinal microvascular and choroidal structural changes in Parkinson disease[J]. JAMA Ophthalmol, 2021, 139(2): 182-188.

[7] Pan L, Chen X. Retinal OCT image registration: methods and applications[J]. IEEE Rev Biomed Eng, 2021. Doi: 10.1109/RBME.2021.3110958.

[8] Ciresan D, Giusti A, Gambardella LM, et al. Deep neural networks segment neuronal membranes in electron microscopy images[C]//Advances in Neural Information Processing Systems, 2012: 2843-2851.

[9] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference On

Computer Vision and Pattern Recognition. IEEE, 2015: 3431-3440.

[10] Chen LC, Zhu Y, Papandreou G, et al. Encoder-decoder with at-rous separable convolution for semantic image segmentation [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 801-818.

[11] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015: 234-241.

[12] Kim J, Park SM, Kim GH, et al. Full-field optical coherence tomography using tunable-path-difference source for three-dimensional multi-layered imaging with enhanced signal-to-noise ratio [J]. Opt Laser Eng, 2021, 147: 106727.

[13] Wang XL, Zhang RF, Kong T, et al. SOLOv2: dynamic and fast instance segmentation[J]. arXiv preprint arXiv: 2003.10152, 2020.

[14] 高乃珺, 陈明惠, 张晨曦, 等. 基于双域滤波算法的 OCT 图像去噪 [J]. 光学技术, 2019, 45(3): 336-342.

Gao NJ, Chen MH, Zhang CX, et al. OCT image denoising based on double domain filtering algorithm[J]. Optical Technique, 2019, 45(3): 336-342.

[15] Zhang D, Zhang HW, Tang JH, et al. Feature pyramid transformer[C]//Computer Vision-ECCV 2020, 2020: 323-329.

[16] Knaus C, Zwicker M. Dual-domain image denoising [C]// IEEE International Conference on Image Processing. IEEE, 2014.

[17] Artyushenko VM, Volovach VI. Determination of the probability density function of the mixture signal and additive noise under influence of multiplicative noise[J]. JPCS, 2021, 2096(1): 012154.

[18] Wu Y, Zhao D, Liu S, et al. Fault detection for linear discrete time-varying systems with multiplicative noise based on parity space method[J]. ISA Trans, 2022, 121: 156-170.

[19] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C]//Conference on Computer Vision and Pattern Recognition. IEEE, 2016.

[20] Lin TY, Dollar P, Girshick R, et al. Feature pyramid networks for object detection [C]//Conference on Computer Vision and Pattern Recognition. IEEE, 2017.

[21] Carion, N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers [C]//European Conference on Computer Vision, 2020.

[22] Huang Y, Wang QQ, Jia WJ, et al. See more than once: kernel-sharing atrous convolution for semantic segmentation[J]. Neurocomputing, 2021, 443: 26-34.

[23] Long X, Zhang W, Zhao B. PSPNet-SLAM: a semantic SLAM detect dynamic object by pyramid scene parsing network[J]. IEEE Access, 2020, 8: 214685-214695.

[24] Bodla N, Singh B, Chellappa R, et al. Soft-NMS-improving object detection with one line of code[C]//IEEE International Conference on Computer Vision. IEEE, 2017.

[25] Nowozin S. Optimal decisions from probabilistic models: the intersection-over-union case [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2014.

(编辑:薛泽玲)