

# 基于多任务学习的糖尿病视网膜病变图像分割<sup>①</sup>



雷凯杰<sup>1</sup>, 崔永俊<sup>1</sup>, 马巧梅<sup>2</sup>

<sup>1</sup>(中北大学 仪器与电子学院, 太原 030051)

<sup>2</sup>(中北大学 软件学院, 太原 030051)

通信作者: 崔永俊, E-mail: annayucyj@nuc.edu.cn

**摘要:** 针对糖尿病视网膜病变 (DR) 图像, 提出了一种基于多任务学习的图像多分类分割方法. 首先, 通过 Otsu 阈值算法将大部分无病灶信息像素去除; 其次, 通过滑动窗口切割的方法将图像切分为若干小尺寸的图像, 以解决医学图像分辨率过大以及病灶在图像中占比较小的问题; 再次, 将不存在病灶的子图剔除, 以增大含病灶子图的比例; 最后, 利用 UNet++ 多任务学习属性, 并且用转置卷积代替传统上采样, 进行多输出多病灶的图像分割. 通过在国际公开 IDRiD 和 DDR 数据集上进行验证, 在 IDRiD 上取得 0.7131 的 mAUPR, 在 DDR 上取得 0.5691 的 mAUPR.

**关键词:** 深度学习; 糖尿病视网膜病变; 图像分割; 多任务学习; 多分类

引用格式: 雷凯杰, 崔永俊, 马巧梅. 基于多任务学习的糖尿病视网膜病变图像分割. 计算机系统应用, 2023, 32(1): 413-419. <http://www.c-s-a.org.cn/1003-3254/8961.html>

## Diabetic Retinopathy Image Segmentation Based on Multi-task Learning

LEI Kai-Jie<sup>1</sup>, CUI Yong-Jun<sup>1</sup>, MA Qiao-Mei<sup>2</sup>

<sup>1</sup>(School of Instrument and Electronics, North University of China, Taiyuan 030051, China)

<sup>2</sup>(School of Software, North University of China, Taiyuan 030051, China)

**Abstract:** A multi-class image segmentation method based on multi-task learning is proposed for diabetic retinopathy (DR) images. Specifically, the dominant pixels without lesion information are removed by the Otsu thresholding algorithm; subsequently, the image is segmented into several small-sized images by the method of sliding window segmentation to solve the problems that the resolution of medical images is too large and the proportion of lesions in the image is small; then, sub-images without lesions are eliminated to increase the proportion of those with lesions; finally, multi-output multi-lesion image segmentation is performed by leveraging the multi-task learning properties of UNet++ and replacing traditional upsampling with transposed convolution. When the proposed method is verified on the international public Indian diabetic retinopathy image dataset (IDRiD) and dataset for diabetic retinopathy (DDR), it achieves a mean area under precision-recall curve (mAUPR) of 0.7131 on IDRiD and an mAUPR of 0.5691 on DDR.

**Key words:** deep learning; diabetic retinopathy; image segmentation; multi-task learning; multi-class

根据世界卫生组织 2019 年 10 月发布的《世界展望报告》, 全球有 4.18 亿人患有糖尿病视网膜病变、青光眼、年龄相关性黄斑病变或者其他可能导致失明的眼部疾病. 眼病患者往往不知道无症状病情的恶化, 因此早期筛查和治疗眼病尤为重要.

随着近几年患病人数的增多, 医生的工作负担加

重, 繁重的阅片任务导致医生无法给予患者及时的反馈, 从而错过最佳治疗时机, 同时由于病灶在眼底图像中占比较小, 医生极易出现误诊、漏诊的情况<sup>[1]</sup>. 因此亟需开发计算机辅助系统来提高医生的筛查率, 从而降低患者的风险.

近年来, 以深度学习方法为基础的系统在 DR 疾

① 收稿时间: 2022-05-26; 修改时间: 2022-07-06; 采用时间: 2022-09-07; csa 在线出版时间: 2022-11-14

CNKI 网络首发时间: 2022-11-16

病检测中取得显著成果<sup>[2]</sup>,尤其是卷积神经网络及其相关算法的改进<sup>[3]</sup>.但传统的DR筛查系统通常首先识别感兴趣的区域进行特征提取,然后再使用提取的特征进行分类<sup>[4]</sup>.但这种深度学习系统在DR筛查中缺乏可解释性,这导致了语义分割的研究,该研究将像素分为各种视网膜特征,即视盘、微动脉瘤、出血、软硬渗出物等<sup>[5]</sup>.

用于DR病灶分割的方法主要有两种:第1种是编码器-解码器架构,该类架构中最具代表的网络是全卷积网络<sup>[6]</sup>、U-Net<sup>[7]</sup>等;第2种是基于空洞卷积的模型,该类架构中最具代表性的是DeepLab<sup>[8]</sup>系列模型.

本文用滑动窗口切割的方法,将大尺寸的图像切割成小尺寸的子图,并将不存在病灶的子图从子图数据集中剔除,只剩下含有病灶的子图,以保证每张图片的学习都能学到至少一种病灶.多任务学习可以多个任务同时学习,在一定程度上可以缓解模型的过拟合,提高模型的泛化能力.结合以上思路,本文提出了一种利用UNet++多任务学习属性结合滑动窗口切割处理数据集的方法.分别在两个公开数据集DDR和IDRiD进行实验验证.

## 1 病灶图像分割

### 1.1 图片预处理

通过观察数据集中的图像可知,眼底图像的边界处有面积较大的黑色部分,其中并不含有任何眼部图像相关的信息,所以首先需要做的就是将这部分冗余信息去除.在这里,选用Otsu阈值算法.Otsu阈值算法会根据图像的灰度值,计算出一个阈值,将图像分为前景和背景,然后获取前景边界的坐标,在原图上进行背景的切割.效果如图1所示,白色部分为前景,黑色部分为背景,最后一张为将部分背景切除后的效果.



图1 Otsu算法预处理图片

### 1.2 子图切割

由于医疗图像自身的高分辨率特性,以及病灶在整张图像中占比较小,对实验环境有较高要求,同时注意力网络聚集到大病灶导致局部细小的特征会被忽略<sup>[9]</sup>,通用的做法是对图像做出等比例的缩放,但这个过程

会导致小病灶的模糊甚至丢失,故使用滑动窗口切割图像的方法,将图像切割成一个个子图.由于小病灶的存在,所以切割以后,大量的子图中不含有任意一种病灶,我们需要将这分子图剔除,以提高含病灶子图的占比.子图切割如图2所示.

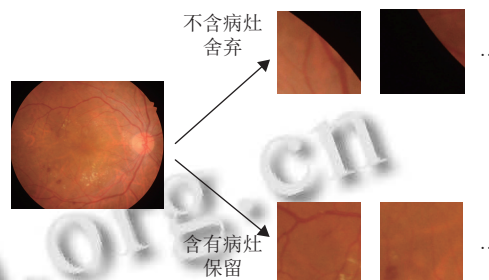


图2 子图切割

## 1.3 多任务学习

### 1.3.1 多任务学习形式

多任务学习的一种具体形式,主要包含两部分:一部分为共享层,另一部分为每个任务的特殊层,共享层是多个任务之间共用的一部分网络,特殊层是每个任务本身拥有的特性.多任务学习有效的原因主要是:多个任务一起学习时,会互相增加噪声从而提高模型的泛化能力,多个任务相互作用,逃离局部最优解,降低过拟合风险<sup>[10]</sup>.

### 1.3.2 UNet++多层次的特征提取网络

U-Net作为经典的医学图像骨干网络,在医学图像方向广泛使用.但其固定深度的下采样并不一定是最佳的下采样深度,而UNet++不倾向于选择网络深度,因为它在其体系结构中嵌入了不同深度的U-Net网络<sup>[11]</sup>.这样就可以一次训练多个层次不同深度的U-Net,然后从中选出最佳的深度.

UNet++中所有的U-Net共用一个编码器层,而解码器层相互交织,所有的U-Net会被同时训练,这种设计提高了整体的分割性能<sup>[11]</sup>.如图3所示,橙色的为共享层,其余为特殊层.

### 1.4 转置卷积代替传统上采样

为了将低分辨率的图像还原成高分辨率图像,会进行多次上采样,常用的上采样方式包括:最近邻插值、双线性插值等.转置卷积也可以作为一种上采样方式,因为其包含可学习的参数,所以经过训练的转置卷积可以更好地实现上采样.首先引入卷积输入和输出关系的计算公式<sup>[12]</sup>:

$$W_{out} = \frac{W_{in} + 2P - F}{S} + 1 \quad (1)$$

其中,  $W_{in}$  为输入的特征图宽度或高度,  $W_{out}$  为输出的特征图宽度或高度,  $P$  为 padding 的取值,  $F$  为卷积核的大小 (一般卷积核形状为方形),  $S$  为卷积核移动步幅. 将  $W_{in}$  和  $W_{out}$  交换位置公式就成为转置卷积的计算公式. 假设我们需要完成  $n$  倍的上采样操作, 设  $W_{out} = nW_{in}$ , 则这些参数应该满足如下关系:

$$W = \frac{nW + 2P - F}{S} + 1 \quad (2)$$

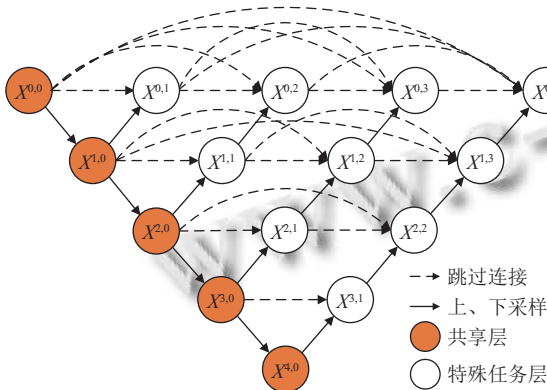


图3 UNet++多任务学习形式

然后可以推算出:

$$\begin{cases} S = n \\ F - S - 2P = 0 \end{cases} \quad (3)$$

上采样倍数一般是 2 的倍数, 于是有:

$$\begin{cases} S = 2k \\ P = t \\ F = 2(k + t) \end{cases} \quad (4)$$

其中,  $k, t$  取 1 到正无穷的整数.

当卷积核大小  $F$  无法被步长  $S$  整除时, 容易出现棋盘状伪影, 所以需要保证  $F/S$  为整数.

从上述推导可以找到满足条件的一组解为:  $F=2, P=0, S=2$ .

### 1.5 损失函数

图像分割任务输入尺寸为:  $h \times w \times c$ , 预测输出  $S$  的尺寸为:  $h \times w \times n$ , 其中  $h$  和  $w$  为长和宽,  $c$  为通道数,  $n$  为分割的种类. 然后将向量  $S$  和真值分割掩膜  $Y$  进行比较. 为此我们使用多分类的交叉熵损失函数, 如式 (5) 所示:

$$L(S, Y) = - \sum_{i=1}^h \sum_{j=1}^w \sum_{k=1}^n Y_{i,j,k} \log S_{i,j,k} \quad (5)$$

由于多任务学习可以进行多输出, 而且每个输出都进行相同的多分类任务, 所以赋予每个输出相同的损失函数, 最终的损失函数为所有的损失函数和求平均, 如式 (6) 所示:

$$L = \frac{1}{N} \sum_{i=1}^N L_i(S, Y) \quad (6)$$

### 1.6 网络具体结构

如图 4 所示为网络的具体结构, 用 VGG16 替换传统网络的编码器部分, 解码器部分为 VGG16 的镜像, 然后输出 4 个结果. 其中  $K$  表示卷积核的个数.

## 2 实验验证和结果分析

### 2.1 数据集和评价指标

糖尿病视网膜病变图像数据集来自于 IDRiD 和 DDR 两个公开数据集. IDRiD 即印度糖尿病视网膜病变图像数据集, 是印度人口眼底图像数据集<sup>[13]</sup>; IDRiD 是一个高质量的数据集, 其中的 81 幅图像提供了典型的 DR 相关病变和视盘的像素级标注. DDR 为来自中国的公开数据集, 其包含数据较多, 采用国际标准标注, 所有图像都有像素级的病变标注, 包括了所有的病灶种类<sup>[14]</sup>. 两个数据集标注情况见表 1.

从表 1 可以看出 DDR 相对于 IDRiD 有更多的标注图像. 在实际训练过程中, IDRiD 数据集特征更为明显, 同时所有图像的分辨率均为  $4288 \times 2848$ ; 而 DDR 数据集的图像分辨率大小不固定, 但 DDR 包含更多不同程度的病灶图像, 多样性更好. 同时这也是先切割成子图的原因之一, 为了统一训练图像的分辨率.

### 2.2 多病灶图像分割

图像分割关注 4 种病灶, 分别是: 微动脉瘤 (micro-aneurysm, MA)、出血 (hemorrhage, HE)、硬性渗出 (hard exudate, EX) 和软性渗出 (soft exudate, SE). 病灶图像分割在训练时可以采取多分类或者多个二分类两种方法.

多分类就是将多张标签最终合并成一张标签, 里面包含背景、4 种病灶等多个类别; 多个二分类是指打算分几个类别就有几张标签, 每张标签由病灶和背景组成. 本文选用多分类的方法, 因为通过初始实验对实际模型的训练比较, 这种方法分割效果更好, 各病灶的特征更加明显. 其中 IDRiD 数据集类别为: MA、HE、EX、SE 以及黑色背景, 同时将除前面 5 类以外的像素

都归为其他类, DDR 数据集与 IDRiD 一致. 这样做的原因是在实验过程中观察到有将别的类别识别成这 4 类的情况出现, 为了尽可能地减少这种情况, 同时为了提升分割效果, 所以在这里分割了多于 4 类. 虽然有多个类别, 但最终结果只计算 MA、HE、EX 以及 SE 这 4 类的指标.

在本文中, 我们使用精确召回曲线下的面积 (area under precision recall curve, AUPR) 作为我们的评估指标, 这与 IDRiD 挑战中使用的指标相同. 阳性预测值 (positive predictive value, PPV) 和真阳性率 (true

positive rate, TPR) 定义如下:

$$PPV = \frac{TP}{TP + FP} \tag{7}$$

$$TPR = \frac{TP}{TP + FN} \tag{8}$$

其中, 真阳性 (true positives, TP) 是正确分类的病变像素, 假阳性 (false positives, FP) 是错误分类为病变像素的非病变像素, 假阴性 (false negatives, FN) 是错误分类为非病变的病变像素. 通过绘制给定不同阈值的精确召回对, 获得精确召回曲线, 然后计算出面积.

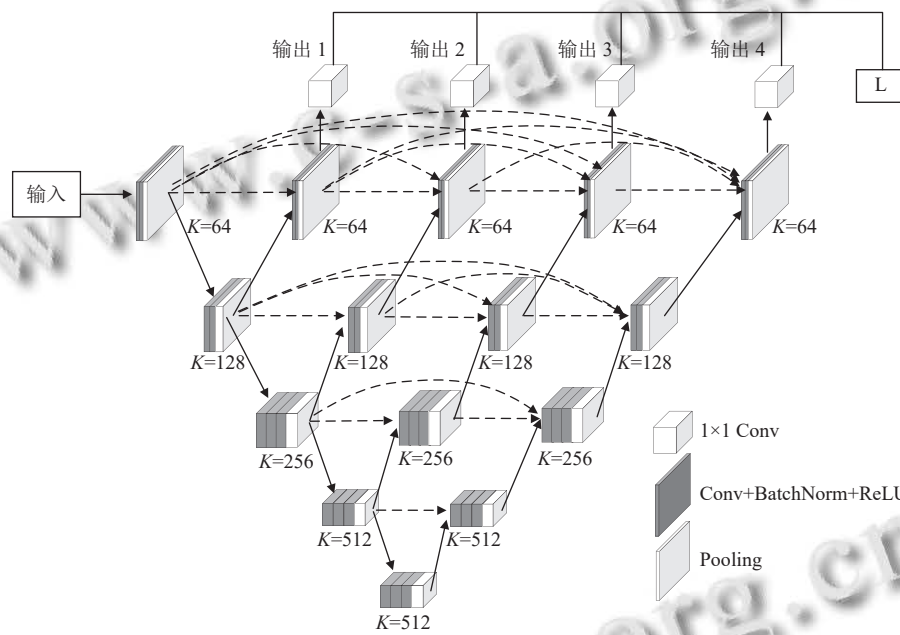


图 4 本文方法

表 1 糖网数据集分布

数据集	样本数			
	训练集	验证集	测试集	总计
DDR	383	149	225	757
IDRiD	40	14	27	81

此外, 我们还利用不同类型病变的平均精确召回曲线下的面积 (mean area under precision recall curve, mAUPR) 来评估该模型在多病变分割中的整体性能.

### 2.3 结果分析

为了探究不同分辨率对最终分割效果会产生什么影响, 同时满足网络所要求的子图分辨率为 32 的倍数, 分别使用多个 32 整数倍的大小进行对比实验, 实验结果如表 2、表 3 所示.

表 2 IDRiD 数据集不同分辨率 AUPR 对比

分辨率	AUPR				mAUPR
	EX	HE	MA	SE	
64×64	<b>0.8869</b>	<b>0.7017</b>	<b>0.6639</b>	<b>0.7651</b>	<b>0.7554</b>
128×128	0.8827	0.6438	0.6036	0.7453	0.7197
256×256	0.7553	0.4604	0.4724	0.7395	0.6069

表 3 DDR 数据集不同分辨率 AUPR 对比

分辨率	AUPR				mAUPR
	EX	HE	MA	SE	
64×64	<b>0.6704</b>	<b>0.4867</b>	<b>0.5034</b>	0.7456	<b>0.6018</b>
128×128	0.6405	0.4068	0.4221	0.7081	0.5444
256×256	0.5089	0.3472	0.3066	<b>0.7690</b>	0.4829

表 2、表 3 中的子图尺寸都为 32 的倍数, 为了使子图的分辨率变化更为明显, 从 64 开始每次增加一倍,

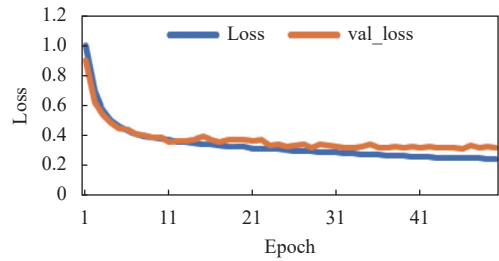
同时为了尽可能地提高病灶在图像中的占比, 最终将 256 作为最后一个分辨率. 从表中可以看出, 在两个数据集中, 64×64 是使得 AUPR 指标综合表现最好的分辨率, 虽然 64×64 分辨率没有在所有类别上做到最好, 但其综合效果最佳, 所以选择 64×64 作为子图切割的分辨率. 从上述分析可知不同的分辨率大小确实对图像分割的结果产生了影响.

本文利用了 UNet++ 的多任务学习属性, 实质上是多个完全相同的任务, 不同于一般定义中的多个相关任务, 或者说是多个不完全相同的任务. 为了验证该方法能否像一般多任务学习一样缓解学习中的过拟合, 从损失函数曲线变化角度进行分析, 如图 5 所示.

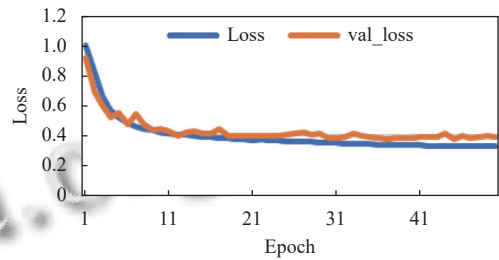
训练集损失不断下降, 验证集损失趋于不变或者上升, 说明网络开始进入过拟合. 从图 5(a) 和图 5(b) 可以看出在 IDRiD 数据集上, U-Net 与本文方法经过一段时间的学习以后训练集和验证集的损失开始较为明显的差值, 说明网络进入过拟合, 同时两种方法进入过拟合的训练周期 (epoch) 也较为相近; 图 5(c) 和图 5(d) 为 DDR 数据集上的损失曲线, 两种方法损失曲线趋势相近, 看不到较为明显的过拟合点, 到训练的后半段学习速度也明显放缓. 抑制过拟合的一种常用方式为减少模型参数, 本文方法参数远多于 U-Net, 但在学习趋势上与 U-Net 相近, 说明本文方法有更强的抑制过拟合的能力, 能够缓解在网络规模更大的情况下的过拟合.

为了证明多任务学习在糖尿病视网膜病变中的有效性, 设计了一组实验, 将本文方法多个输出的最好结果和经过相同改进的 U-Net 网络进行对比, 表 4、表 5 为实验结果.

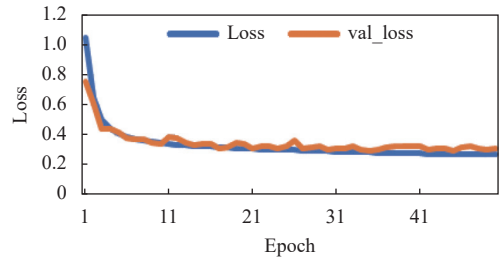
表 4 和表 5 中的输出整合为将各个输出最好值整合到一起. 从表 4 可以看出, 在 IDRiD 数据集上, 输出 3 在 MA 上, 输出 4 在 EX、HA 以及 mAUPR 上效果比 U-Net 效果更好, 在 SE 上与 U-Net 相差 3.9%; 从表 5 可以看出, 在 DDR 数据集上, 输出 4 在 EX、MA、SE 上比 U-Net 效果更好, 在 HE 上与 U-Net 相差 4.6%, 最终的 mAUPR 比 U-Net 表现更好. 从上述分析可知, 多任务学习可以取得更好的多分类效果, 即使在一个输出表现不佳的情况下, 也可由其他输出来补足. 在 IDRiD 数据集上相比于单任务的 U-Net 提升明显; 在 DDR 数据集上有提升, 但提升不是很大. 图 6、图 7 为部分子图分割结果.



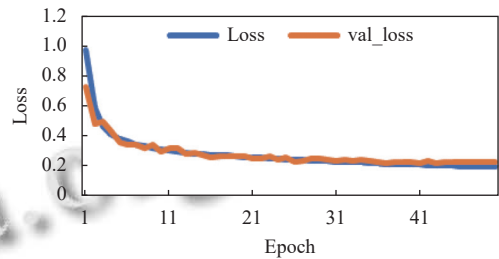
(a) IDRiD U-Net loss 曲线



(b) IDRiD 本文 loss 曲线



(c) DDR U-Net loss 曲线



(d) DDR 本文 loss 曲线

图 5 Loss 曲线

表 4 多任务学习在 IDRiD 数据集效果对比

模型	AUPR				mAUPR
	EX	HE	MA	SE	
U-Net	0.8646	0.6729	0.6025	<b>0.8041</b>	0.7360
输出1	0.7626	0.5684	0.2982	0.7434	0.5931
输出2	0.8741	0.6493	0.6345	0.7552	0.7283
输出3	0.8843	0.6971	<b>0.6639</b>	0.7636	0.7522
输出4	<b>0.8869</b>	<b>0.7017</b>	0.6591	0.7651	0.7532
输出整合	<b>0.8869</b>	<b>0.7017</b>	<b>0.6639</b>	0.7651	<b>0.7554</b>

表 6、表 7 中将本文方法和不同模型之间的 AUPR 值对比, 可以看出本文的方法, 在多个病灶上都取得较好的效果.

表5 多任务学习在DDR数据集效果对比

模型	AUPR				mAUPR
	EX	HE	MA	SE	
U-Net	0.6545	<b>0.5332</b>	0.4784	0.7314	0.5994
输出1	0.4877	0.2701	0.2931	0.4528	0.3759
输出2	0.6391	0.4666	0.4948	0.6949	0.5739
输出3	0.6651	0.4867	0.5026	0.7455	0.5955
输出4	<b>0.6704</b>	0.4846	<b>0.5034</b>	<b>0.7456</b>	0.6000
输出整合	<b>0.6704</b>	0.4867	<b>0.5034</b>	<b>0.7456</b>	<b>0.6018</b>

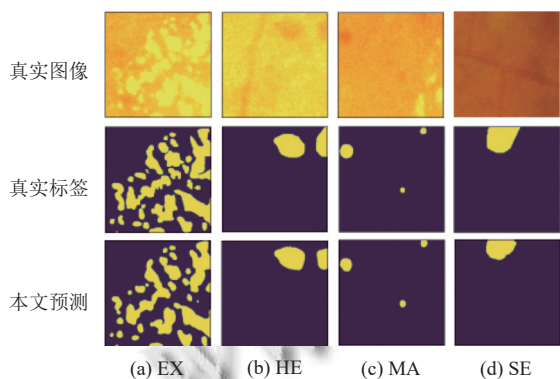


图6 IDRID 部分分子图分割结果

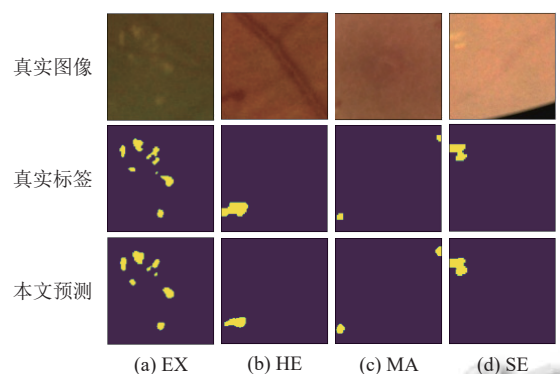


图7 DDR 部分分子图分割结果

表6 IDRID数据集上不同模型结果比较

模型	AUPR				mAUPR
	EX	HE	MA	SE	
HED <sup>[15]</sup>	0.7723	0.5084	0.4271	0.6673	0.5938
FCRN <sup>[16]</sup>	0.5472	0.4200	0.3383	0.5154	0.4552
CASENet <sup>[17]</sup>	0.7564	0.4462	0.3992	0.3275	0.4823
DeepLab V3+ <sup>[8]</sup>	0.7118	0.4772	0.1514	0.5912	0.4829
L-Seg <sup>[18]</sup>	0.7945	0.6374	0.4627	0.7113	0.6515
本文方法	<b>0.8869</b>	<b>0.7017</b>	<b>0.6639</b>	<b>0.7651</b>	<b>0.7554</b>

### 3 结语

针对糖尿病视网膜病变的图像分割问题进行了研

究, 本文提出了基于多任务学习结合滑窗切割的方法来进行多病灶图像分割. 该方法一方面考虑到医学图像高分辨率导致的病灶占比较小对分割结果产生的影响, 同时也考虑到样本不均衡导致学习效果不好的问题, 并通过增加非病灶种类以及提升含病灶子图占比的方法来解决. 实验结果表明了本文方法能够提升多种病灶的分割效果.

表7 DDR数据集上不同模型结果比较

模型	AUPR				mAUPR
	EX	HE	MA	SE	
HED <sup>[15]</sup>	0.4241	0.1919	0.0496	0.0972	0.1907
FCRN <sup>[16]</sup>	0.1996	0.1255	0.0134	0.0454	0.0960
CASENet <sup>[17]</sup>	0.2777	0.2625	0.1005	0.1304	0.1928
DeepLab V3+ <sup>[8]</sup>	0.5349	0.3894	0.0223	0.2120	0.2897
L-Seg <sup>[18]</sup>	0.5546	0.3586	0.1052	0.2684	0.3208
本文方法	<b>0.6418</b>	<b>0.4994</b>	<b>0.5034</b>	<b>0.7488</b>	<b>0.5983</b>

### 参考文献

- 1 Bellemo V, Lim G, Rim TH, *et al.* Artificial intelligence screening for diabetic retinopathy: The real-world emerging application. *Current Diabetes Reports*, 2019, 19(9): 72. [doi: 10.1007/s11892-019-1189-3]
- 2 Asiri N, Hussain M, Al Adel F, *et al.* Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey. *Artificial Intelligence in Medicine*, 2019, 99: 101701. [doi: 10.1016/j.artmed.2019.07.009]
- 3 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, 60(6): 84–90. [doi: 10.1145/3065386]
- 4 Lim G, Lee ML, Hsu W, *et al.* Transformed representations for convolutional neural networks in diabetic retinopathy screening. *Proceedings of the Workshops at the 28th AAAI Conference on Artificial Intelligence*. Québec City: AAAI, 2014.
- 5 Saha O, Sathish R, Sheet D. Fully convolutional neural network for semantic segmentation of anatomical structure and pathologies in colour fundus images associated with diabetic retinopathy. *arXiv:1902.03122*, 2019. [doi: 10.48550/arXiv.1902.03122.]
- 6 Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640–651. [doi: 10.1109/TPAMI.2016.2572683]

- 7 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference 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)]
- 8 Chen LC, Zhu YK, Papandreou G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 833–851. [doi: [10.1007/978-3-030-01234-2\\_49](https://doi.org/10.1007/978-3-030-01234-2_49)]
- 9 Wang Z, Yin YX, Shi JP, *et al.* Zoom-in-Net: Deep mining lesions for diabetic retinopathy detection. Proceedings of the 20th International Conference on Medical Image Computing and Computer-assisted Intervention. Québec City: Springer, 2017. 267–275. [doi: [10.1007/978-3-319-66179-7\\_31](https://doi.org/10.1007/978-3-319-66179-7_31)]
- 10 Ruder S. An overview of multi-task learning in deep neural networks. arXiv:1706.05098, 2017. [doi: [10.48550/arXiv.1706.05098](https://doi.org/10.48550/arXiv.1706.05098).]
- 11 Zhou ZW, Siddiquee MR, Tajbakhsh N, *et al.* UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE Transactions on Medical Imaging, 2020, 39(6): 1856–1867. [doi: [10.1109/TMI.2019.2959609](https://doi.org/10.1109/TMI.2019.2959609)]
- 12 Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning. arXiv:1603.07285, 2016. [doi: [10.48550/arXiv.1603.07285](https://doi.org/10.48550/arXiv.1603.07285).]
- 13 Porwal P, Pachade S, Kokare M, *et al.* IDRiD: Diabetic retinopathy—Segmentation and grading challenge. Medical Image Analysis, 2020, 59: 101561. [doi: [10.1016/j.media.2019.101561](https://doi.org/10.1016/j.media.2019.101561)]
- 14 Li T, Gao YQ, Wang K, *et al.* Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. Information Sciences, 2019, 501: 511–522. [doi: [10.1016/j.ins.2019.06.011](https://doi.org/10.1016/j.ins.2019.06.011)]
- 15 Xie SN, Tu ZW. Holistically-nested edge detection. Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015. 1395–1403. [doi: [10.1109/ICCV.2015.164](https://doi.org/10.1109/ICCV.2015.164)]
- 16 Mo J, Zhang L, Feng YQ. Exudate-based diabetic macular edema recognition in retinal images using cascaded deep residual networks. Neurocomputing, 2018, 290: 161–171. [doi: [10.1016/j.neucom.2018.02.035](https://doi.org/10.1016/j.neucom.2018.02.035)]
- 17 Yu ZD, Feng C, Liu MY, *et al.* CASENet: Deep category-aware semantic edge detection. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 1761–1770. [doi: [10.1109/CVPR.2017.191](https://doi.org/10.1109/CVPR.2017.191)]
- 18 Guo S, Li T, Kang H, *et al.* L-Seg: An end-to-end unified framework for multi-lesion segmentation of fundus images. Neurocomputing, 2019, 349: 52–63. [doi: [10.1016/j.neucom.2019.04.019](https://doi.org/10.1016/j.neucom.2019.04.019)]

(校对责编: 孙君艳)