

引用格式: GAO Weiwei, YANG Yile, FANG Yu, et al. Improved Faster-RCNN Based on Multi Feature Scale Fusion for Automatic Detection of Microaneurysms in Retina[J]. Acta Photonica Sinica, 2023, 52(4):0410002

高玮玮, 杨亦乐, 方宇, 等. 多特征尺度融合改进 Faster-RCNN 视网膜微动脉瘤自动检测算法[J]. 光子学报, 2023, 52(4): 0410002

多特征尺度融合改进 Faster-RCNN 视网膜 微动脉瘤自动检测算法

高玮玮¹, 杨亦乐¹, 方宇¹, 樊博¹, 宋楠²

(1 上海工程技术大学 机械与汽车工程学院, 上海 201620)

(2 复旦大学附属眼耳鼻喉医院 眼科, 上海 200031)

摘要: 视网膜微动脉瘤的检测对于早期发现糖尿病视网膜病变等重要疾病至关重要, 但该病灶尺寸相对较小, 属于眼底图像中的微小目标, 现有的微动脉瘤检测算法难以实现该病灶的精准检测, 为此提出了基于多特征尺度融合的改进 Faster-RCNN 微动脉瘤自动检测算法。该算法在 Faster-RCNN 网络模型的基础上, 首先采用多特征尺度融合对特征提取网络与 RPN 结构进行改进以提高网络对于微小目标特征的利用; 然后, 通过感兴趣区域齐平池化以消除感兴趣区域池化过程中引入的量化误差; 最后, 通过对损失函数中的 smooth L1 损失函数进行重新设计得到平衡 L1 损失函数以实现损失函数优化, 从而有效降低大梯度难学样本与小梯度易学样本间的不平衡问题, 进而使得模型能够得到更好地训练。针对眼底图像中微动脉瘤的自动检测, 将优化后的 Faster-RCNN 网络模型在 Kaggle 数据集上进行训练及测试, 并与其他方法进行对比。实验结果表明, 与其他各种结构的 Faster-RCNN 网络模型相比, 所提出的基于多特征尺度融合的改进 Faster-RCNN 算法能显著提高检测结果 (F-score 与原始 Faster-RCNN 相比提升了 9.36%); 与其他网络模型以及方法相比, 所提出的基于多特征尺度融合的改进 Faster-RCNN 的自动检测精度明显更优。故所提出的基于多特征尺度融合的改进 Faster-RCNN 算法性能较优, 能准确、有效地检测出眼底图像中的微动脉瘤。

关键词: 眼底图像; 微动脉瘤; Faster-RCNN; 多特征尺度融合; 深度学习

中图分类号: R318; TP391.41

文献标识码: A

doi: 10.3788/gzxb20235204.0410002

0 引言

视网膜微动脉瘤 (Microaneurysms, MAs) 是一种微小血管病变, 由血管壁变弱引起的毛细血管肿胀形成^[1], 因此 MAs 可能与不同的眼科和心血管疾病相关^[2]。例如, MAs 已被证实是中风的先兆因素^[3]。此外, MAs 更是糖尿病视网膜病变 (简称“糖网”, Diabetic Retinopathy, DR)^[4] 的最早期典型症状。糖网是一种由糖尿病引发的视网膜疾病, 是目前全球导致中老年人群失明的主要原因^[5]。对于糖网, 至关重要的是对该疾病的及早发现, 以阻止疾病的进展和预防潜在的视力损失。然而, MAs 尺寸较小, 相对于眼底图像而言属于微小目标; 且由于视觉条件不理想, MAs 可能相对于背景呈现低对比度, 或者可能受到图像中不均匀照明的影响; 此外, MAs 还可能与图像中的其他结构混淆, 例如微出血、色素沉着变化, 甚至是眼底照相机中的灰尘颗粒。故针对眼底图像中的 MAs 进行自动检测是一项极具挑战性的任务, 很多研究人员就此展开了相关研究。

MAs 自动检测算法通常可分为三类: 基于物理模型的方法、基于分类器的方法以及基于深度学习的方法。基于物理模型的方法主要根据 MAs 的灰度分布特征进行数学建模, 如: 赖小波等^[6] 提出了一种基于特

基金项目: 国家自然科学基金 (No. 61703268)

第一作者: 高玮玮, gww03020234@sina.com

收稿日期: 2022-08-11; 录用日期: 2022-11-21

征相互关系的视网膜MAs自动提取算法;郑绍华等^[7]提出了一种新的基于Radon变换的方法对眼底图像中的MAs进行识别;高玮玮等^[8]提出了一种基于数学形态学扩展极小值变换的MAs自动检测方法;ZHANG Xinpeng等^[9]提出了一种基于特征转移网络和局部背景抑制的MAs检测方法等。基于分类器的方法一般先获取MAs候选区,然后再通过所设计分类器对候选区进行分类,如ORLANDO JI等^[10]首先采用中值滤波去除眼底图像中的MAs,再经过形态学膨胀得到糖网背景图像,在此基础上将原图与糖网背景图像相减得到MAs的候选区,然后对候选区提取传统特征和深度特征用于MAs分类;WU B等^[11]利用从眼底图像中提取出的多个轮廓和局部特征训练KNN分类器以实现MAs的准确分类;YADAV D等^[12]首先通过分析直方图进行MAs候选区域分割,然后再根据MAs候选区域的形状、灰度、纹理等特征进行识别等。基于深度学习的MAs自动检测方法主要为搭建端到端的深度神经网络,如DAIL等^[13]提出了一种临床报告引导的多筛分卷积神经网络,该网络利用临床报告中的少量监督信息,通过特征空间中的图像到文本映射来识别潜在的MAs;ROHAN R A^[14]提出了一种基于YOLO(You Only Look Once)的MAs自动检测算法;赵学功等^[15]提出了基于卷积神经网络的MAs自动检测方法;XU Y等^[16]提出一种基于改进U-Net网络的MAs自动分割算法等;郭松等^[17]提出一种基于多任务学习的分割模型Red-Seg来获取眼底图像中的MAs。

由于眼底图像会因为患者、环境、采集设备等因素的不同而出现不同的亮度和对比度,且还会存在类似于MAs的对象,如小血块、血管结构、背景噪声等,故基于物理模型的方法不仅难以建立广泛适用的算法且容易造成误检。基于分类器的方法通常先采用滤波、形态学、背景估计等传统方法进行候选区提取,而这些方法极易受到参数的制约,同时由于眼底图像内容复杂,故容易使对于微动脉瘤的检测精度不高。基于深度学习的方法相较于上述两种方法所存在的不足,有明显的优势,但针对眼底图像的微小目标——MAs的自动检测而言,现有深度学习模型仍存在问题,如:小目标的特征无法得到充分利用、回归框位置不准确、检测精度不够高等,故提出了多特征尺度融合的改进Faster-RCNN(Improved Faster-RCNN, Faster-RCNN-Pro)用于实现眼底图像中MAs的精准检测。该方法首先选用检测精度较优的Faster-RCNN作为基础网络,然后结合眼底图像中的MAs自动检测,对其进行优化,具体为:首先通过采用多特征尺度融合对特征提取网络与区域候选网络(Region Proposal Network, RPN)结构进行改进以提高网络对于微小目标特征的利用从而提高MAs的检测精度;然后,通过感兴趣区域齐平池化以消除感兴趣区域池化过程中引入的两次量化误差;最后,对损失函数中的smooth L1损失函数进行重新设计得到平衡L1损失函数,以实现通过损失函数优化有效降低大梯度难学样本与小梯度易学样本间的不平衡问题,进而使模型能够得到更好的训练。在完成Faster-RCNN网络模型优化的基础上,将该网络模型用于眼底图像中MAs的自动检测,并进行性能分析。

1 实验材料及设备

实验采用Kaggle DR数据集用以训练和测试所提出的Faster-RCNN-Pro算法对于眼底图像中MAs的自动检测性能。Kaggle DR数据集由EyePACS(视网膜病变筛查免费平台)提供,来源于44 351位患者(左右眼各一幅),共计有35 126幅训练图像和53 576幅测试图像,图像的分辨率在 433×289 像素~ $5 184 \times 3 456$ 像素之间^[18]。眼科医师基于国际标准对每幅眼底图像进行了0~4的糖网病患病程度的5级分级,但由于本文研究的是所提出算法对于MAs病灶的自动检测性能,故需对该数据集进行预处理以满足本研究的需求,具体为:利用文献^[8]所述方法并辅以眼科医师的指导对Kaggle DR数据集进行预处理以获取MAs图像及注释文件。由于Kaggle DR数据集的分布不平衡,经过该阶段预处理去除了大量图像(该数据集中几乎90%的图像不包含任何病变),最后只剩下5 519幅包含MAs的眼底图像。对于这5 519幅眼底图像,均生成了注释文件。生成注释文件后,整个数据集(5 519幅图像)按90/10的比例进行划分,5 000幅用于训练,519幅眼底图像用于测试。此外,为充分验证所提出改进网络的检测性能,对原本包含519幅眼底图像的测试集予以增加200幅健康眼底图像进行扩充,即实际测试时测试集包含719幅眼底图像。

实验采用的操作系统为ubuntu-18.04.5;CPU为Intel Core i5-10300H;RAM大小为32 GB;GPU为RTX2060,显存大小为8 GB。运行软件环境为Pycharm;python版本为3.7;Pytorch版本为1.3.0;OpenCV版本为4.5;此外,安装Cuda10.1和Cudnn7.51进行GPU加速运算。

2 网络设计

Faster-RCNN^[19-21]检测算法相较于YOLO^[22-24]检测算法,多增设了一个RPN结构,这不仅使其成为了两阶段的检测算法,而且提高了检测精度,故其检测准确性通常要优于YOLO算法,具体算法流程如图1。由于MAs尺寸较小(直径在10~100 μm 之间,具体形态如图1),属于眼底图像中的微小目标,因此利用原始Faster-RCNN对眼底图像中的MAs进行检测时会出现一些不足,如检测性能不够好、回归框位置不准确、数据分布不均匀等问题。针对以上问题,对Faster-RCNN网络提出了3点改进,分别为:多特征尺度融合、感兴趣区域齐平池化、损失函数优化(具体为对损失函数函数中的smooth L1损失函数进行重新设计得到平衡L1损失函数),改进后的算法流程如图2。

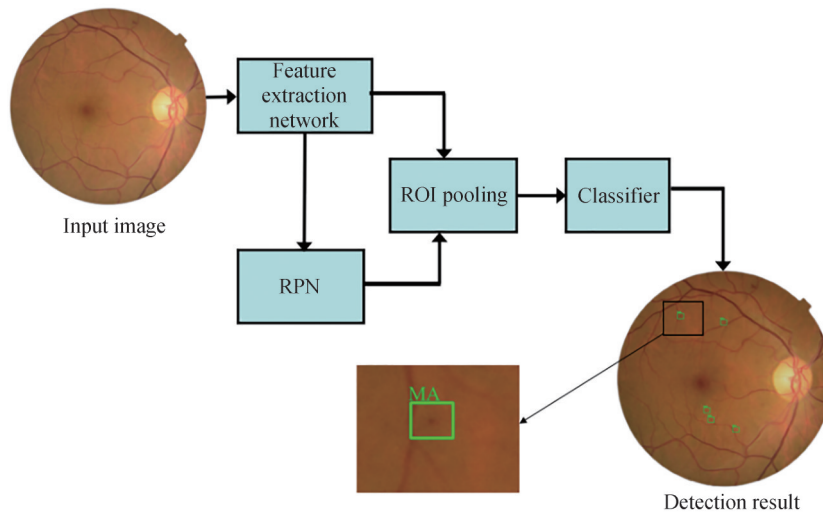


图1 Faster-RCNN算法结构
Fig. 1 Faster-RCNN algorithm structure

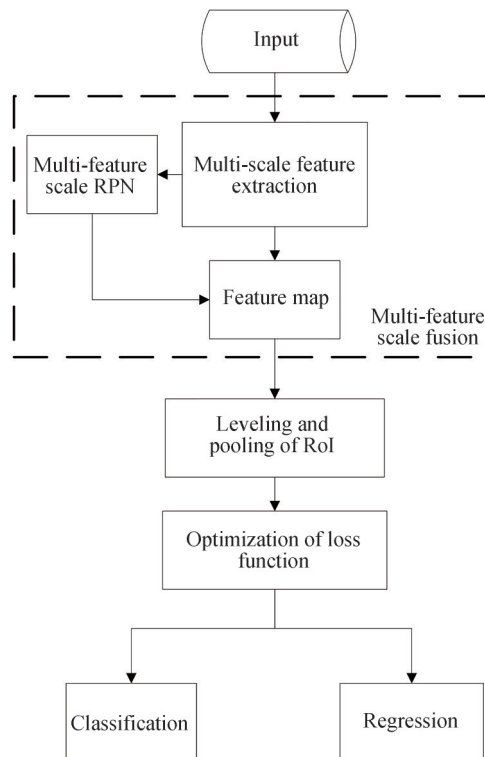


图2 改进后的算法流程
Fig. 2 Flow of improved algorithm

2.1 多特征尺度融合

在特征提取网络VGG16的结构中,特征图的尺寸在卷积的过程中越来越小,与之对应的待检测目标的特征信息也在不断变小。而RPN网络采用最后一层特征图卷积得到的信息作为特征图并进行后续的操作,这会导致小目标的特征无法得到充分利用,这很明显对于MAAs的自动检测是非常不利的,故采用多特征尺度融合的方法对特征提取网络与RPN结构进行改进,将该改进结构的Faster-RCNN称为Faster-RCNN-I。

多特征尺度融合是通过一定的方式整合具有高层语义信息的深层网络特征图的空间特征维度,使其与具有低层语义信息的浅层网络特征图相同后,按照相加的方式进行融合,从而使提取出的特征图能够具有更丰富的语义信息以供后续环节使用^[25]。多特征尺度融合后,需要在网络结构、检测和分类层、感兴趣池化层共3个层上进行改进。

在特征提取主干网络上,将卷积层每次降采样称为一个Stage,如图3所示,主干网络可分为5个Stage,分别为 C_1, C_2, C_3, C_4 和 C_5 。对于 $C_i (i=1, 2, 3, 4, 5)$,通过 1×1 的卷积操作与下采样,使得其与 C_{i-1} 的尺度相匹配并相加,最后通过 3×3 的卷积消除相加后由多维特征不连续产生的混叠效应,从而得到新的特征图 $P_i (i=1, 2, 3, 4, 5)$ 。

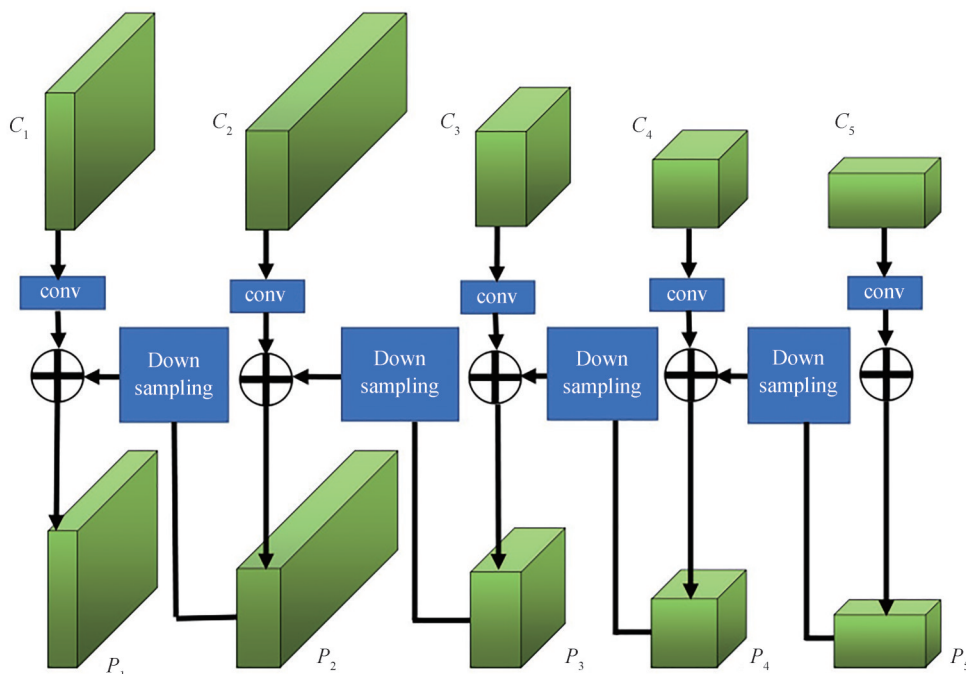


图3 多特征尺度融合

Fig. 3 Multi-feature scale fusion

RPN网络计算采用的是网络最后一层的特征图,对原图映射至特征图的 k 个预设大小的Anchor Box进行分类与回归。这 k 个Anchor Box预先设置为固定的长宽比,且在RPN中使用同一个Stage的特征图进行计算。为融合多特征尺度,Anchor Box在保持原有预设大小与比例尺度的情况下,在不同的特征图上匹配与之相应的Anchor Box。考虑到计算效率,将Anchor Box分别分布在 P_2, P_3, P_4 这三个特征层上面,每一个特征图上分配的Anchor Box均有3种比例尺度,分别为1:1、1:2、2:1。在RPN检测时,目标在哪个特征图上能够分类出则采用该层的Anchor Box,具体过程如图4。

在改变RPN生成的区域建议后,对于特征图的Pooling方式也需随之变化。原始Faster RCNN中RoI (Region of Interest) Pooling是把最后一层特征图作为输入,其中特征图与RPN产生的映射关系可表示为

$$P = 4 + \log_2 \frac{\sqrt{wh}}{k} \quad (1)$$

式中, P 表示特征图的层级, P 越大,对应特征图的尺度越小;4表示输入层级; k 是输入图像尺度; w, h 分别表示RoI区域的宽与长。由式(1)可知,RoI Pooling的输入单一,通过在最后一层特征图上使用RPN提取出的

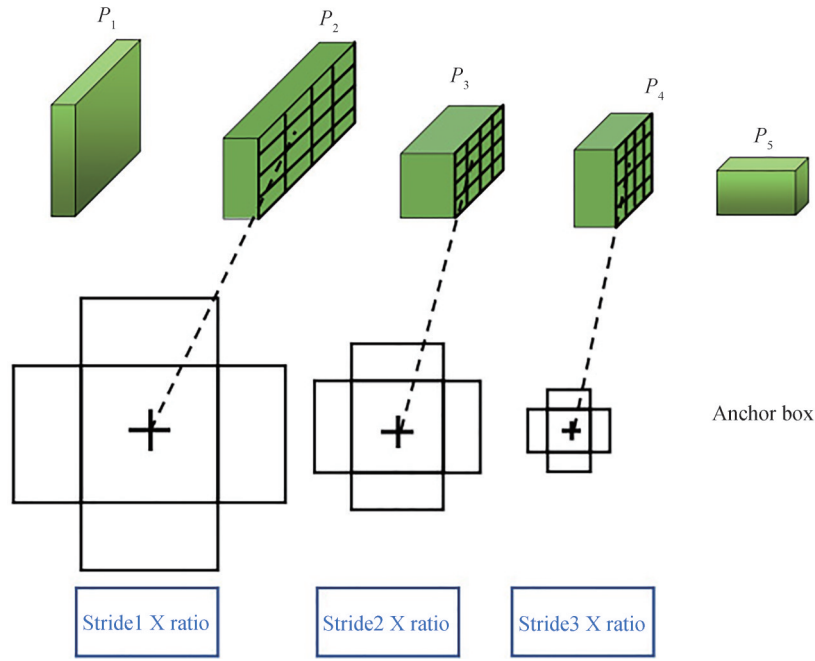


图4 多特征融合下的RPN

Fig. 4 RPN based on multi feature fusion

RoI区域作为输入,这无法有效利用多特征尺度RPN获取的RoI区域,故需对RoI Pooling进行改进,具体可表示为

$$\begin{cases} P = P_i + \log_2 \frac{\sqrt{wh}}{k} \\ P_i = \{P_{i+1}, P_i\} \end{cases} \quad (2)$$

式中,输入的特征图为当前层的特征图与深一层的小尺度特征图通过下采样的方式求和后的结果。通过将RPN在多特征尺度上得到的RoI区域对应的特征图作为RoI Pooling的输入,确保了目标特征能够映射到特征图上,从而提高了目标检测精度。

2.2 感兴趣区域齐平池

为进一步提升对MA的检测性能,还需对RoIPooling做出改进。原始RoI Pooling中,初始像素点与RPN得出的特征图上RoI区域像素点对应关系为^[20]

$$(x_{\text{feature}}, y_{\text{feature}}) = \left\lfloor \frac{x}{s}, \frac{y}{s} \right\rfloor \quad (3)$$

式中, $(x_{\text{feature}}, y_{\text{feature}})$ 为特征图上的像素坐标, (x, y) 为RoI区域在原始图像上的坐标, s 为下采样倍数, $\lfloor \cdot \rfloor$ 表示为取比括号内值小的最大整数,即向下取整。由于在计算 $(x_{\text{feature}}, y_{\text{feature}})$ 时,存在向下取整的计算,故此处引入了一次量化误差。此外,在RoI Pooling的过程中,需将RoI的输入直接固定成预设的大小,这带来了第二次量化误差。这两次量化误差使得原始特征区域与映射特征区域产生了偏差,这些偏差对数据集中的小目标——MA的检测影响很大,故提出对感兴趣区域进行齐平池化(如图5),图5中黑色矩形框表示RoI区域在特征图上的映射位置,具体关系可表示为

$$(x_{\text{feature}}, y_{\text{feature}}) = \frac{x}{s}, \frac{y}{s} \quad (4)$$

图5中绿色实线表示根据RoI Pooling的实际需求将区域分割为 n 个子区域,其中 n 的大小对应输出特征图尺度。图5中每一个子区域中设置4个点,每个点的像素值由特征图上最邻近区域的四个像素点通过双线性插值法获取^[26],如红色箭头所示;随后,池化的过程在每个子块中进行,最终得到尺寸齐平的输出。该感兴趣区池化过程与Faster-RCNN池化即RoI Pooling过程的对比如图6。将该改进结构的Faster-RCNN称为Faster-RCNN-II。

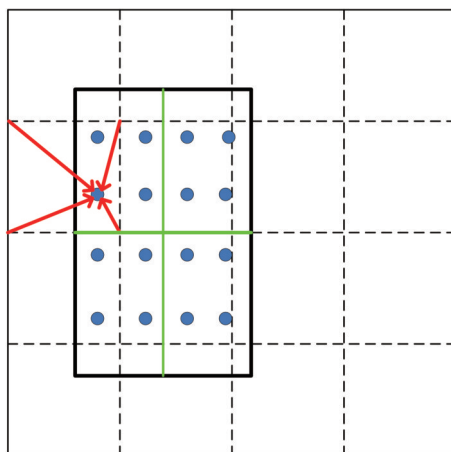


图5 感兴趣区域齐平池化
Fig. 5 Flush pooling of RoI

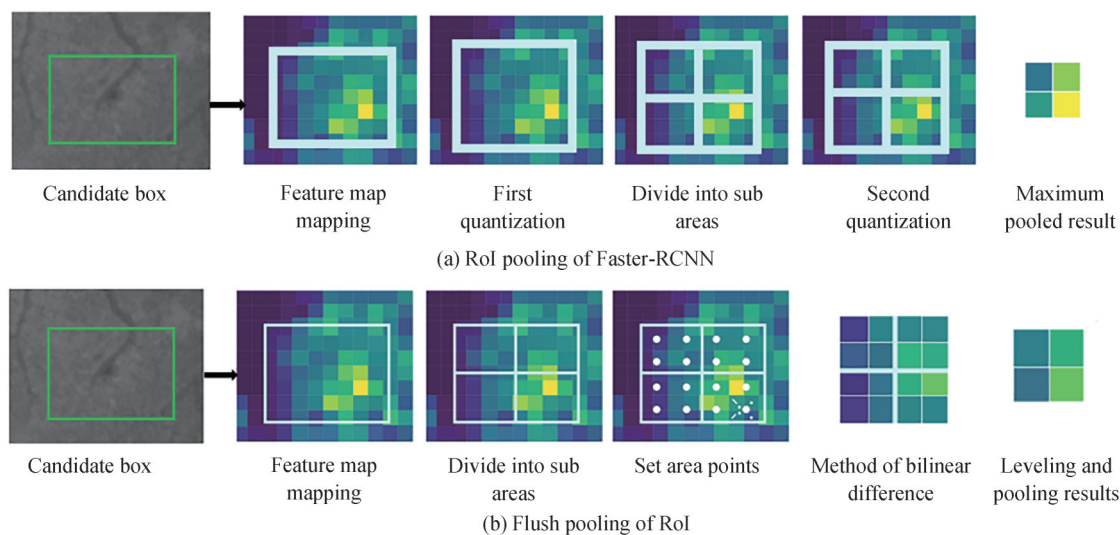


图6 感兴趣区池化过程对比
Fig.6 Comparison of pooling process in RoI

2.3 损失函数优化

Faster-RCNN的损失函数包括分类损失和回归损失两部分,可表示为^[20]

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (5)$$

式中,下标*i*表示第*i*个候选区域,*i*随同批次中真值数量变化; p_i 为检测网络中离散型概率分布,即判定第*i*个目标的类别; p_i^* 为第*i*个框的GT(Ground Truth)标定值,即真实值(当该框中存在该目标物时 $p_i^* = 1$;反之, $p_i^* = 0$);分类损失 L_{cls} 和回归损失 L_{reg} 分别由 N_{cls} 、 N_{reg} 以及权重 λ 归一化。 $t_i = \{t_x, t_y, t_w, t_h\}$ 表示目标物位置的预测值; $t_i^* = \{t_x^*, t_y^*, t_w^*, t_h^*\}$ 则是学习样本中人工给定的监督信息。

对于回归损失,由于参数回归时均近似为线性回归,且标注信息噪声较多,梯度回传不稳定^[27],为缓解这一情况,回归损失使用梯度平缓的 $S_{smooth L_1}$ 损失函数,可分别表示为

$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*) \quad (6)$$

$$S_{smooth L_1}(x) = \begin{cases} 0.5x^2 \times 1/\sigma^2, & |x| < 1/\sigma^2 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (7)$$

式(6)中的*R*即为式(7)中的Smooth L1损失函数。其中,参数 σ 用于控制平滑区域的范围,以避免过大的梯度对网络参数产生过多的影响。

对式(7)求梯度,即

$$\frac{\partial L_{\text{smoothL1}}}{\partial x} = \begin{cases} x, & |x| \leq 1 \\ 1, & \text{otherwise} \end{cases} \quad (8)$$

反向传播中损失较大的定义为困难样本,换言之容易样本损失较小。故由式(8)发现,在 smoothL1 函数中,困难样本相较于容易样本来说,提供了更多的梯度信息,这导致了难易样本间学习能力的不平衡。

为解决该问题,设计了 balanceL1 损失函数,即由式(8)得到 balanceL1 损失的梯度函数,表示为

$$\frac{\partial L_{\text{balanceL1}}}{\partial x} = \begin{cases} 0.3 \ln(b|x|+1), & |x| \leq 1 \\ \gamma, & \text{otherwise} \end{cases} \quad (9)$$

其中,反向传播时提供梯度高于 γ 的样本取 γ 作为梯度,以此平衡小梯度样本的梯度值与难学样本的梯度值。在此基础上,根据式(9)的梯度信息,通过积分求得平衡 L1 损失函数为

$$L_{\text{balanceL1}}(x) = \begin{cases} \frac{0.3c}{b}(b|x|+1)\ln(b|x|+1) - 0.3|x|, & |x| < 1 \\ \gamma|x| + d, & \text{otherwise} \end{cases} \quad (10)$$

式(10)中为了满足分段函数在分段处提供梯度,则有 $b = \exp \frac{\gamma}{0.3}$ 。在 Kaggle DR 数据集上 γ 取值为 1。通过平衡 L1 损失函数,可有效降低大梯度难学样本与小梯度易学样本间的不平衡问题,使模型能够更好地训练,将该策略优化后的 Faster-RCNN 称为 Faster-RCNN-III。其中,两种损失函数的梯度变化如图 7。

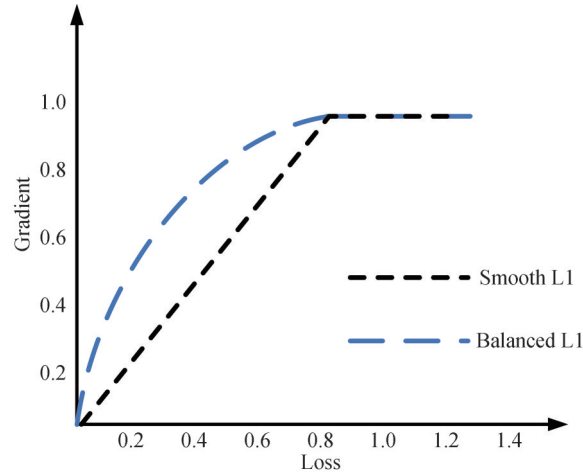


图7 损失函数梯度
Fig.7 Loss function gradient

3 结果与分析

3.1 评价指标

在目标检测任务中,通常使用召回率(Recall, R)和精确率(Precision, P)以及综合考虑召回率与精确率的指标 F-score 对网络性能进行评价。因此为评价所提出网络模型对于 MAs 的检测性能,基于病灶区域水平定义了上述指标,具体计算表达式分别为^[28]

$$R = TP / (TP + FN) \quad (11)$$

$$P = TP / (TP + FP) \quad (12)$$

$$F\text{-score} = 2 \times P \times R / (P + R) \quad (13)$$

式中,TP(True Positive)为成功预测的正例,TN(True Negative)为成功预测的负例,FP(False Positive)为被误判为正例的负例,FN(False Negative)为被错误预测为负例的正例。

3.2 实验结果及分析

利用 Kaggle DR 数据集对所提出的 Faster-RCNN-Pro 网络模型进行训练时,超参数设置为:输入图像

为彩色图像,尺寸经适当裁减去掉背景区域并进行压缩后为 $1\,024\times 768$,批处理大小设置为2,迭代次数共为31 000次,基础学习率为0.001 25(当第23 000次迭代后学习率衰减至0.000 125),优化方法采用带动量因子的小批量随机梯度下降法,特征提取网络部分采用在COCO数据集上的模型作为预训练模型。

Faster-RCNN-Pro深度学习检测模型训练的损失函数在20个轮次的训练周期内,呈现收敛趋势。在训练之初,损失函数快速下降,在迭代到22 000次后,损失函数收敛放缓,但存在波动;在20 000次迭代后,学习率下降到十分之一,损失函数也随之下降,波动减少;直至30 000次迭代,损失函数平稳收敛,识别性能较好,故取第31 000次迭代后的模型作为最终检测模型。

为验证所提出的3点改进策略对于Faster-RCNN算法检测性能的影响,采用消融实验比分析不同改进结构的Faster-RCNN算法分别在Kaggle DR数据集上对MAs的自动检测性能,具体检测结果见表1。由表1可知,每种网络改进策略均在不同程度上提高了原始Faster-RCNN的检测性能,具体为:由于MAs为尺寸微小目标,通过多特征尺度融合可以有效提高对该病灶的检测性能,故Faster-RCNN-I的F-score较Baseline(Faster-RCNN)提升了6.82%;通过感兴趣区域齐平池化,先验框中更准确的特征信息优化了模型表现,故Faster-RCNN-II的F-score较Baseline(Faster-RCNN)提升了4.68%;采用优化损失函数平衡了难学样本和易学样本对模型提供的梯度,降低了噪声影响,故Faster-RCNN-III的F-score较Baseline(Faster-RCNN)提升了2.73%;尤其是3种改进策略均采用的Faster-RCNN-Pro的检测性能提升更为明显,其F-score较Baseline(Faster-RCNN)提升了9.36%。

表1 不同改进Faster-RCNN模型在测试集上的性能评估

Table 1 Performance evaluation of different improved Faster-RCNN models on the test set

Network structure	Multi feature scale fusion	Leveling and pooling of RoI	Balanced L1 loss function	F-score/%
Faster-RCNN	×	×	×	83.21
Faster-RCNN-I	√	×	×	90.03
Faster-RCNN-II	×	√	×	87.89
Faster-RCNN-III	×	×	√	85.94
Faster-RCNN-Pro	√	√	√	92.57

由以上实验结果发现,所提出的对Faster-RCNN的3点改进在Kaggle DR数据集上针对MAs自动检测,对于网络性能提升是有效的。该网络模型对MAs的检测效果如图8。由图8可知,该网络模型可以实现对眼底图像中MAs的准确检测。

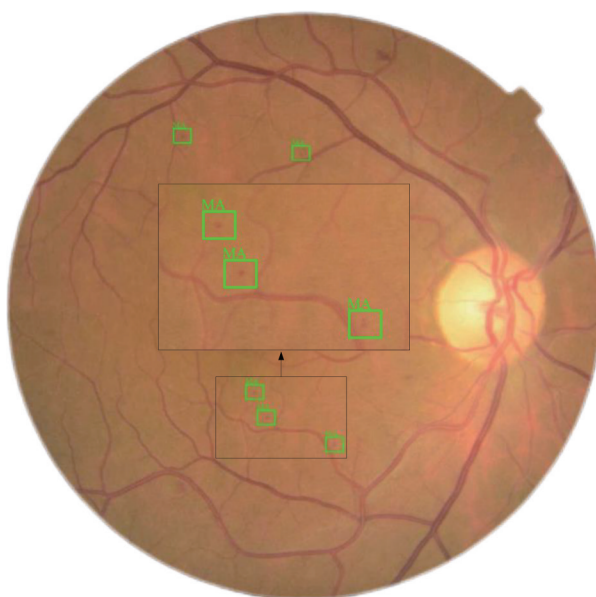


图8 Faster-RCNN-Pro网络检测MAs结果

Fig. 8 Detection results of MAs by Faster-RCNN-Pro

此外,为进一步验证所提出的基于多特征尺度融合的改进Faster-RCNN对眼底图像中MA的自动检测性能,将该方法与文献[8](基于图像处理的方法)、文献[12](基于分类器的方法)、文献[13](采用多筛分卷积神经网络检测MA)、文献[14](采用YOLO网络自动检测MA)、文献[15](利用卷积神经网络检测MA)、文献[16](采用改进U-Net网络自动分割MA)对MA的自动检测性能进行了对比,结果如表2。由表2可知,基于深度学习的方法即文献[13-16]以及所提出的Faster-RCNN-Pro算法的F-score均优于基于图像处理的方法以及基于分类器的方法。传统算法(基于图像处理的方法以及基于分类器的方法)的F-score低,是因为传统算法易受到参数限制,在眼底图像这样的复杂背景中提取的MA病灶候选区易存在较多干扰且无法排除,最终会成为FP,从而导致 P 值较低进而影响F-score。结合表1、表2可知,基于深度学习的方法即深度神经网络由于具有更强的拟合性,对于MA的自动检测结果更为准确,且两阶段的Faster-RCNN因为RPN区域建议网络的存在使得检测性能参数F-score明显更优,而经过优化后的Faster-RCNN-Pro的检测精度更是得到了明显提升。但同时也应注意到,深度神经网络也可能存在过拟合现象,从而导致部分MA会检测不出,因此文献[13-16]的召回率 R 会低于传统算法,而Faster-RCNN-Pro则很好地克服了这个问题。由此可见,所提出的Faster-RCNN-Pro算法能准确、有效地检测出眼底图像中的MA,即检测性能更优。

表2 不同算法检测性能对比

Table 2 Comparison for detection performance of different algorithms

Algorithm	$R/\%$	$P/\%$	F-score/ $\%$
Ref. [8]	88.30	81.12	84.57
Ref. [12]	87.56	82.15	84.77
Ref. [13]	84.18	88.79	86.42
Ref. [14]	83.91	86.75	85.31
Ref. [15]	85.12	87.41	86.25
Ref. [16]	87.05	90.91	88.94
Faster-RCNN-Pro	90.89	94.32	92.57

最后,将所提出的Faster-RCNN-Pro对于眼底图像中MA的自动检测性能与YOLOv5进行了对比,结果如表3。结合表2、表3可知,YOLOv5算法通过使用Focus与CSPDarknet53特征提取网络获取多尺度图像特征,相比之前的YOLO网络进一步消除网络敏感性,优化了小目标特征信息差异,从而提高识别精度^[29],但是与Faster-RCNN-Pro相比,其对于MA的检测精度仍逊色不少。因此,就检测精度而言,所提出的Faster-RCNN-Pro明显更优,能精准检测出眼底图像中的MA。

表3 与YOLOv5的检测性能对比

Table 3 Comparison of detection performance with YOLOv5

Algorithm	$R/\%$	$P/\%$	F-score/ $\%$
YOLOv5	88.75	90.03	89.39
Faster-RCNN-Pro	90.89	94.32	92.57

4 结论

本文提出了多特征尺度融合的改进Faster-RCNN算法用以实现眼底图像中MA的精准检测。该算法在Faster-RCNN网络模型的基础上,分别通过采用多特征尺度融合以提高网络对于微小目标特征的利用,感兴趣区域齐平池化以消除池化过程中引入的量化误差,以及优化损失函数有效降低大梯度难学样本与小梯度易学样本间的不平衡问题,从而实现针对眼底图像中MA精准检测的Faster-RCNN网络优化。在以Kaggle DR数据集为测试对象的基础上,采用消融实验分析验证了所提出的多特征尺度融合的改进Faster-RCNN算法能够有效提升MA的自动检测性能。此外,还将该方法与基于YOLO、基于深度卷积神经网络以及传统方法对MA的自动检测性能进行了对比,发现所提出的多特征尺度融合的改进Faster-RCNN算法在MA的自动检测精度方面明显更优。故所提出的多特征尺度融合的改进Faster-RCNN算法性能较

优,能精准检测出眼底图像中的MAs。但改进后的Faster-RCNN-Pro仍为二阶段目标检测算法,这使其与实时检测之间还存在一定距离,未来可在检测实时性上做进一步深入研究,以期达到更好的检测效果。

参考文献

- [1] SILVA P S, EIRAMI H, BARHAM R, et al. Hemorrhage and/or microaneurysm severity and count in ultrawide field images and early treatment diabetic retinopathy study photography[J]. *Ophthalmology*, 2017, 124(7): 970-975.
- [2] FLEMING A D, PHILIP S, GOATMAN K A, et al. Automated microaneurysm detection using local contrast normalization and local vessel detection[J]. *IEEE Transactions on Medical Imaging*, 2006, 25(9): 1223-1245.
- [3] WONG T Y, SUN J, KAWASAKI R, et al. Guidelines on diabetic eye care: the Int'l council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource setting[J]. *Ophthalmology*, 2018, 125(10): 1608-1622.
- [4] SHAO Yi, ZHOU Zhao, GE Qianmin. Consensus interpretation on the diagnosis and treatment of diabetic retinopathy and diabetic macular edema by the The Royal College of Ophthalmologists[J]. *Recent Advances in Ophthalmology*, 2021, (7): 601-607.
邵毅, 周召, 葛倩敏. 糖尿病视网膜病变及黄斑水肿诊疗规范: 英国皇家眼科医师学会指南解读[J]. *眼科新进展*, 2021, (7): 601-607.
- [5] CHEN Yixuan, WEN Li, PEI Cunwen, et al. Changes of microvascular diameter in non-proliferative diabetic retinopathy [J]. *International Eye Science*, 2021, 21(9): 1632-1636.
陈奕璇, 温力, 裴存文, 等. 非增殖期糖尿病视网膜病变微血管直径的变化[J]. *国际眼科杂志*, 2021, 21(9): 1632-1636.
- [6] LAI Xiaobo, LIU Huashan, FANG Chunjie. Retinal microaneurysm extraction by fusing relationship among features[J]. *Optics and Precision Engineering*, 2013, 21(8): 2187-2194.
赖小波, 刘华山, 方纯洁. 融合特征相互关系的视网膜微动脉瘤提取[J]. *光学精密工程*, 2013, 21(8): 2187-2194.
- [7] ZHENG Shaohua, PAN Lin, CHEN Jian, et al. Grading method for non-proliferative diabetic retinopathy images based on microaneurysms and hemorrhages automatic detection[J]. *Chinese Journal of Scientific Instrument*, 2014, 35(1): 59-67.
郑绍华, 潘林, 陈健, 等. 微动脉瘤与出血自动检测的NPDR图像分型方法[J]. *仪器仪表学报*, 2014, 35(1): 59-67.
- [8] GAO Weiwei, SHEN Jianxin, WANG Yuliang, et al. Efficient and automated detection of microaneurysms from non-dilated fundus images[J]. *Chinese Journal of Biomedical Engineering*, 2013, 31(6): 839-845.
高玮玮, 沈建新, 王玉亮, 等. 免散瞳眼底图像中微动脉瘤的高效自动检测[J]. *中国生物医学工程学报*, 2013, 31(6): 839-845.
- [9] ZHANG Xinpeng, WU Jigang, MENG Min, et al. Feature-transfer network and local background suppression formicroaneurysm detection[J]. *Machine Vision and Applications*, 2020, 32(1): 1-15.
- [10] ORLANDO J I, PROKOFYEVA E, FRESNO M D, et al. An ensemble deep learning based approach for red lesion detection in fundus images[J]. *Computer Methods and Programs in Biomedicine*, 2017, 153: 115-127.
- [11] WU B, ZHU W, SHI F, et al. Automatic detection of microaneurysms in retinal fundus images[J]. *Computerized Medical Imaging and Graphics*, 2017, 55: 106-112.
- [12] YADAV D, KARNA K, GIDDALUR A, et al. Microaneurysm detection using color locus detection method [J]. *Measurement*, 2021, 176(1): 109084.
- [13] DAI L, FANG R, LI H, et al. Clinical report guided retinal microaneurysm detection with multi-sieving deep learning[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(5): 1149-1161.
- [14] ROHAN R A. Film: finding the location of microaneurysms on the retina [J]. *Biomedical Engineering Letters*, 2019, 9: 497-506.
- [15] ZHAO Xuegong, DENG Jiakun, WEI Haoran, et al. CNN-based microaneurysm detection in fundus images[J]. *Journal of University of Electronic Science and Technology of China*, 2021, 50(6): 915-920.
赵学功, 邓佳坤, 魏浩然, 等. 基于卷积神经网络的眼底图像微血管瘤检测方法[J]. *电子科技大学学报*, 2021, 50(6): 915-920.
- [16] XU Y, ZHOU Z, LI X, et al. FFU-net: feature fusion-net for lesion segmentation of diabetic retinopathy[J]. *Biomed Research International*, 2021, 2021(3): 1-12.
- [17] GUO Song, LI Tao, LI Ning, et al. Red lesion segmentation of fundus image with multi-task learning[J]. *Journal of Software*, 2021, 32(11): 3646-3658.
郭松, 李涛, 李宁, 等. 基于多任务学习的眼底图像红色病变点分割[J]. *软件学报*, 2021, 32(11): 3646-3658.
- [18] FAN Jiawei, ZHANG Ruru, LU Meng, et al. Applications of deep learning techniques for diabetic retinal diagnosis[J]. *Acta Automatica Sinica*, 2021, 47(5): 985-1004.
范家伟, 张如如, 陆萌等. 深度学习方法在糖尿病视网膜病变诊断中的应用[J]. *自动化学报*, 2021, 47(5): 985-1004.
- [19] GIRSHICK R. Fast R-CNN [C]. *Proceedings of the 2015 IEEE International Conference on Computer Vision*, 2015:

- 1440-1448.
- [20] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards realtime object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [21] CAI Z, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection [C]. *Proceedings of the 2018 IEEE International Conference on Computer Vision*, 2018: 6154-6162.
- [22] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. *Proceedings of the 2017 IEEE International Conference on Computer Vision*, 2017: 2980-2988.
- [23] TIAN Z, SHEN C, CHEN H, et al. Fcos: fully convolutional onestage object detection [C]. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*, 2019: 9627-9636.
- [24] LUO Huilan, CHEN Hongkun. Survey of object detection based on deep learning[J]. *Acta Electronica Sinica*, 2020, 48(6): 1230-1239.
罗会兰, 陈鸿坤. 基于深度学习的目标检测研究综述[J]. *电子学报*, 2020, 48(6): 1230-1239.
- [25] CHEN Keqi, ZHU Zhiliang, DENG Xiaoming, et al. Deep learning for multi-scale object detection: a survey[J]. *Journal of Software*, 2021, 32(4): 1201-1227.
陈科圻, 朱志亮, 邓小明, 等. 多尺度目标检测的深度学习研究综述[J]. *软件学报*, 2021, 32(4): 1201-1227.
- [26] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2020, 42(2): 386 - 397.
- [27] ZHANG Cuiwen, ZHANG Changlun, HE Qiang, et al. Research on loss function of box regression in object detection [J]. *Computer Engineering and Applications*, 2021, 57(20): 97-103.
张翠文, 张长伦, 何强, 等. 目标检测中框回归损失函数的研究[J]. *计算机工程与应用*, 2021, 57(20): 97-103.
- [28] HUANG Hong, WANG Tao, LI Yuan, et al. Cancer pathological segmentation network based on depth feature fusion [J]. *Acta Photonica Sinica*, 2022, 51(3): 0310001.
黄鸿, 王涛, 李远, 等. 基于深度特征融合的癌症病理图像分割网络[J]. *光子学报*, 2022, 51(3): 0310001.
- [29] HE Bin, ZHANG Yibo, GONG Jianlin, et al. Fast recognition of tomato fruit in greenhouse at night based on improved YOLOv5[J]. *Transactions of the Chinese Society of Agricultural Machinery*, 2022, 53(5): 201-208.
何斌, 张亦博, 龚健林, 等. 基于改进 YOLOv5 的夜间温室番茄果实快速识别[J]. *农业机械学报*, 2022, 53(5): 201-208.

Improved Faster-RCNN Based on Multi Feature Scale Fusion for Automatic Detection of Microaneurysms in Retina

GAO Weiwei¹, YANG Yile¹, FANG Yu¹, FAN Bo¹, SONG Nan²

(1 *Institute of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China*)

(2 *Department of Ophthalmology, Eye, Ear, Nose and Throat Hospital of Fudan University, Shanghai 200031, China*)

Abstract: The presence of Microaneurysms (MAs) is the earliest detectable small abnormality of Diabetes Retinopathy (DR), a retinal disease that represents the leading cause of blindness among the middle-aged population globally. So the detection of MAs in fundus images is an important and challenging step for early diagnosis and prevention of critical health conditions. In particular, MAs are small vascular lesions consisting of swollen capillaries due to weakened vascular walls. Thus, retinal MAs can be associated with various ophthalmic and cardiovascular conditions. For instance, retinal MAs have been demonstrated as a risk factor for strokes. Therefore, it is crucial to detect the disease at its earliest stages to prevent its progression and consequent potential vision loss. However, MAs are a small target relative to the fundus image. Because the visual conditions are not ideal, MAs may present a low contrast with the background or may be affected by uneven illumination in the image. In addition, MAs may also be confused with other structures in the image, such as microbleeds, pigmentation changes, and dust particles in the fundus camera. Therefore, the automatic detection of MAs in fundus images is a significantly challenging task. The existing MA detection algorithm is difficult to achieve accurate detection of the lesion. Therefore, an improved faster-RCNN (Faster-RCNN-Pro) detection algorithm based on multi feature scale fusion is proposed. Firstly, the structure of feature extraction network and Region Proposal Network (RPN) are

improved by using multi feature scale fusion to increase the utilization of micro target features; then, the quantization error introduced in the process of pooling the region of interest is eliminated by homogenizing and pooling the region of interest; finally, by redesigning the smooth L1 loss function in the loss function, a balanced L1 loss function is obtained to realize the optimization of the loss function, so as to effectively reduce the imbalance between large gradient difficult samples and small gradient easy samples, so that the model can be better trained. For the automatic detection of MAs in fundus images, the optimized Faster-RCNN network model is trained and tested on the Kaggle DR dataset, and compared with other methods. Based on the Kaggle DR dataset, the ablation experiment analysis verifies that the proposed improved Faster-RCNN-Pro based on multi-feature scale fusion can effectively improve the automatic detection performance of MAs. Specifically F-score is increased by 9.36% compared to that of Faster-RCNN. In addition, the performance of this method is compared to the automatic detection performance of MAs based on YOLO, CNN, and traditional methods including image processing and classifiers. The results demonstrates that the F-scores of the methods based on deep learning, including YOLO, CNN and the Faster-RCNN-Pro proposed in this study, are superior to the methods based on image processing and classifiers. F-score of traditional methods such as image processing and classifiers is low because traditional algorithms are easily limited by parameters. The lesion candidate regions of MAs extracted from a complex background, such as the fundus images, are more prone to interference and can not be excluded, and will eventually become FP, resulting in a low P value and affecting the F-score. Moreover, the two-stage model of Faster-RCNN presents a significantly better detection performance, such as F-score, due to the existence of RPN; the detection accuracy of the Faster-RCNN-Pro has been significantly improved. However, the deep neural network may also present the overfitting phenomenon, which may lead to some MAs not being detected. Therefore, the R values based on deep neural networks in previous studies are lower than those obtained by using traditional algorithms; however, Faster-RCNN-Pro can overcome this problem. So the proposed Faster-RCNN-Pro can accurately and effectively detect MAs in fundus images, demonstrating a better detection performance.

Key words: Fundus image; Microaneurysms; Faster-RCNN; Multi feature scale fusion; Deep learning

OCIS Codes: 100.4996; 100.4994; 170.4580