

基于 CNN 跨层融合结构的边缘检测算法^①



李金迪, 张陶界, 周迪斌, 刘文浩

(杭州师范大学 信息科学与技术学院, 杭州 311121)

通信作者: 周迪斌, E-mail: dibinz@zju.edu.cn

摘要: 传统边缘检测算法难以处理复杂的图像, 而现有基于深度的边缘检测模型, 其检测结果往往存在边缘定位错误和信息丢失等现象. 针对此类问题, 提出一种基于 RCF 的高精度的边缘检测算法 RCF-CLF. 首先, 引入 HDC 结构设计用于避免因叠加相同膨胀卷积而引起的网格效应; 其次, 设计了一种特征增强结构, 旨在融合多尺度信息、扩大感受野; 然后, 设计了跨层融合结构, 将高层信息和低层信息融合, 用于提取准确的边缘信息; 最后, 引入注意力机制 *CBAM*, 通过聚焦物体边缘区域, 抑制非边缘区域, 从而提高网络对边缘信息的提取能力. 本文在 BSDS500 和 BIPED 数据集上评估所提出的方法, 与 RCF 算法相比, 在 BIPED 数据集上, 主要指标 ODS、OIS 和 AP 分别达到了 0.893、0.901 和 0.945, 提高了近 5 个百分点, 在 BSDS500 数据集上, 主要指标也有所提升. 此外, 与其他同类算法相比, 本文算法也具有一定的优势, 可以实现更加准确的边缘定位.

关键词: 边缘检测; 卷积神经网络; 特征增强; 跨层融合; 注意力机制

引用格式: 李金迪, 张陶界, 周迪斌, 刘文浩. 基于 CNN 跨层融合结构的边缘检测算法. 计算机系统应用, 2024, 33(2): 207-215. <http://www.c-s-a.org.cn/1003-3254/9398.html>

Edge Detection Algorithm Based on CNN Cross-layer Fusion Structure

LI Jin-Di, ZHANG Tao-Jie, ZHOU Di-Bin, LIU Wen-Hao

(School of Information Science and Technology, Hangzhou Normal University, Hangzhou 311121, China)

Abstract: Traditional edge detection algorithms are difficult to deal with complex images, and the existing depth-based edge detection models often have edge positioning errors and information loss in the detection results. Aiming at such problems, this study proposes a high-precision edge detection algorithm RCF-CLF based on RCF. First, the HDC structure is introduced to avoid the grid effect caused by superimposing the same dilated convolution. Second, a feature enhancement structure is designed to fuse multi-scale information and expand the receptive field. Then, a cross-layer fusion structure is designed, which integrates high-level and low-level information to extract accurate edge information. Finally, the attention mechanism *CBAM* is introduced to focus on the edge area of the object and suppress the non-edge area, thereby improving the ability of the network to extract edge information. This study evaluates the proposed method on the BSDS500 and BIPED datasets. Compared with the RCF algorithm, the main indicators ODS, OIS, and AP reached 0.893, 0.901, and 0.945, respectively, with an increase of nearly 5 percentage points on the BIPED dataset. On the BSDS500 dataset, the main indicators have also improved. In addition, compared with other similar algorithms, the proposed algorithm also has certain advantages, which can achieve more accurate edge positioning.

Key words: edge detection; convolutional neural network (CNN); feature enhancement; cross-layer fusion; attention mechanism

① 基金项目: 国家自然科学基金联合重点项目 (U21A20466)

收稿时间: 2023-08-17; 修改时间: 2023-09-15; 采用时间: 2023-09-26; csa 在线出版时间: 2023-12-27

CNKI 网络首发时间: 2023-12-28

图像是人类相互交流和认识客观世界的工具。科学研究表明,人们获取的信息大约有80%来源于视觉图像。图像边缘作为图像的基本特征,它包含了很多具有重要意义的信息,是一种低层次特征。图像边缘检测技术的实质是通过边界检测算法对图像中有意义的图像边缘进行提取,它是图像处理的基础。后期图像处理的过程容易受到边缘检测结果的影响。因此,边缘检测在图像分割^[1,2]、目标检测和识别^[3,4]任务中具有重要意义。通过边缘检测,可以获得有关物体形状、轮廓和几何结构的关键信息,为后续的图像处理和任务提供基础和指导。

边缘检测技术广泛应用于医学救助^[5,6]、路面检测^[7,8]、地质勘测^[9,10]、日常生活^[11-13]等方面。在医学领域,对通过红外线、CT、超声波等技术得到的医学图像进行边缘检测,可以辅助医生进行诊断,便于医生进行会诊工作。在路面检测领域,边缘检测技术可以帮助检测人员快速地定位到路面的裂缝,极大地提高了检测人员的工作效率。在日常生活中,校园的人脸识别系统,停车场的车牌识别系统,地铁的安检仪等都涉及到了边缘检测技术。边缘检测技术的广泛应用对各行各业都起到了重要的支持和推动作用。通过提取图像中的边缘信息,边缘检测技术能够为相关领域的研究和实践工作提供准确的视觉指导和信息基础。

边缘检测的主要目的是标识数字图像中亮度变化明显的点。图像属性中的显著变化通常反映了属性的重要事件和变化。然而,由于在成像过程中的投影、混合、畸变和噪声等因素影响,使得图像模糊和变形,进而使得这种变化难以被检测到,因此,边缘检测一直是图像处理中长期存在的一种挑战。

边缘检测作为一种常见的图像处理技术,被广泛研究,目前其检测方法一般可以分为3类。第1类是传统的边缘检测方法。传统的边缘检测方法通常使用图像的一阶导数信息和二阶导数信息对图像的边缘进行研究。早期的经典传统算法有Sobel算子,Canny算子^[14]等。Sobel检测器就是采用一组 3×3 的滤波器直接计算梯度大小。而Canny算子则是从一个平滑过程开始,将图像与高斯核进行卷积,然后计算梯度,最后应用非最大抑制(NMS)算法来精确定位边缘位置。后来,Li等人^[15]在Canny算子的基础上进行了改进,他们使用新的滤波器代替高斯滤波器,并采用最大类间方差确定最佳阈值。这种改进有效地提高了Canny算子在噪声

环境下的检测效果。然而,传统方法的主要缺点是:它们只关注局部强度的变化,而不能识别和忽略那些不属于边缘的纹理。第2类是采用机器学习技术构造边缘检测器。例如Arbeláez等人^[16]和Dollár等人^[17]通过结合颜色、亮度和其他线索来构建纹理描述符、手工特征和稀疏表示。基于学习的边缘检测器在一定程度上克服了传统方法在纹理表示问题上的局限性。然而,这些方法仍需要设计针对特定任务的特征提取算法,而且这些算法在适用性和泛化性方面限制较大。近年来,随着卷积神经网络(CNN)的普及,第3类边缘检测器被提出。该类方法利用卷积神经网络具备的特征学习和分类能力,实现了更有效的边缘检测。2015年,Xie等人提出了一种端到端的边缘检测模型HED^[18]。HED使用VGGNet^[19]和神经网络分类器,融合了VGGNet中每层的最后一个输出作为最终的结果。然而,该方法仅利用了每个卷积阶段最后一层的卷积特征,没有充分利用CNN层级结构的丰富特性,导致一些边缘信息出现丢失的现象。为了解决这个问题,Liu等人改进了HED网络,提出了一种新的深层结构——更丰富的卷积特征(RCF)^[20]。RCF可以自动融合所有卷积层的信息,获得不同尺度和更精细的特征,从而取得了更高的边缘检测精度。Su等人^[21]提出了一种简单、轻量且高效的边缘检测网络架构,名为PiDiNet(pixel difference network),该网络采用了一种新颖的像素差卷积操作,将传统的边缘检测算子集成到流行的卷积神经网络(CNN)中,从而提高了边缘检测任务的性能。PiDiNet通过有效地结合像素差和卷积运算,能够更好地捕捉图像中的边缘信息,实现准确的边缘检测。

无论是HED网络还是RCF网络,其检测得到的预测结果中大多包含较多的噪声,存在边缘定位错误和部分细节信息丢失的情况。这是由于随着网络深度的加深,网络提取的特征损失了较多的细节信息,而高层特征缺少细节信息导致网络对边缘的定位不够准确。同时,受限于各个阶段的单层输出,该类网络不能很好地识别粗细不同的边缘。

针对以上问题,本文在RCF模型的基础上提出了一种自底向上的跨层融合模型RCF-CLF。首先,引入混合膨胀卷积(hybrid dilated convolution, HDC)^[22],用来避免网格效应^[23],较好的保持了图像中的细节信息,其次,设计了一种特征增强模块,用来增强特征,强化边缘;然后,设计了一种跨层融合模块用于将高层信息和

低层信息进行融合;最后,为了进一步提高网络对边缘信息的提取能力,又引入注意力机制 convolutional block attention module (CBAM)^[24]. 本文在 BSDS500^[16] 和 BIPED^[25]数据集上对该模型进行性能评估,实验结果证明,相较于原有模型 RCF,该模型能够获得更好的边缘检测精度.

1 RCF-CLF 网络结构

针对复杂图像边缘检测时容易出现漏检和细节丢失的问题,本文在 RCF 网络的基础上进行改进得到 RCF-CLF,其整体结构如图 1 所示.该网络由输入,特征提取和边缘预测 3 大模块组成.其中,输入模块用于对原始图像进行预处理,特征提取模块利用卷积神经网络对输入图像进行特征提取,边缘预测模块则利用

一个小型网络对边缘进行预测.边缘检测属于像素级预测任务,而相同膨胀率卷积的串联使用会使得像素之间出现一定的间隔,进而导致出现网格效应的问题.通过将原始网络第 5 阶段中的 3 个膨胀卷积改为 HDC 混合膨胀卷积,在增大感受野的同时,还能避免由间隔而引起的边界模糊或不准确问题.由于每阶段的特征经过降维后直接输出会存在大量的冗余信息和非边缘信息,通过设计 FEM 模块将不同膨胀率的并行卷积输出融合,增强特征的表达能力,减少噪声影响.随着网络结构的加深,高层特征逐渐抽象,从而缺乏了低层的细节信息.通过设计带有 CBAM 的跨层融合模块 CLF,可以将高层信息和低层信息进行融合,得到更具表征能力和判别能力的特征表示,从而提高边缘信息预测的准确性.

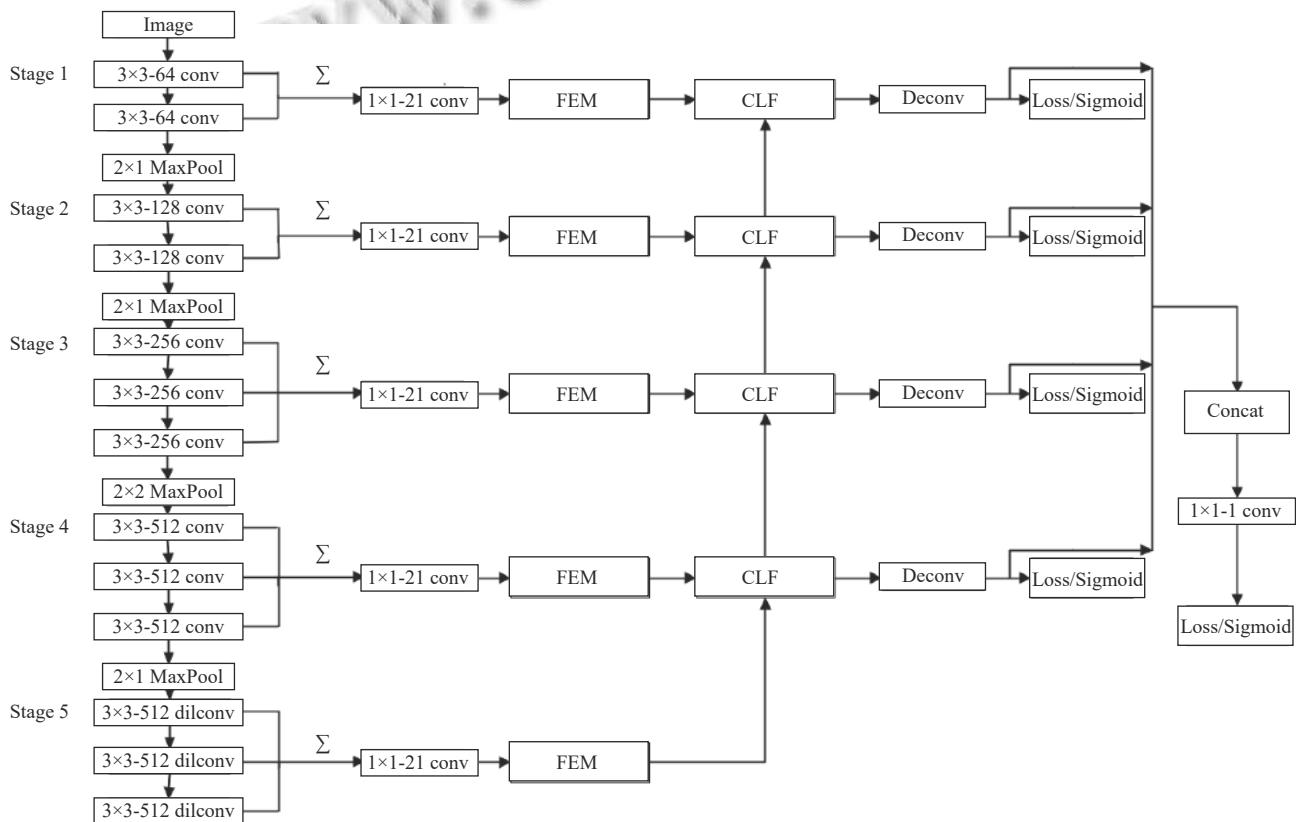


图 1 RCF-CLF 网络结构图

1.1 带有 HDC 结构的主干网络

RCF-CLF 的主干网络是去除 3 个全连接层的 VGGNet,其由 10 个内核大小为 3 的普通卷积、3 个不同膨胀系数的膨胀卷积和 4 个最大池化层组成.通过 4 个最大池化层将卷积操作划分为 5 个阶段,从而

使得每个阶段能够提取到不同尺度的边缘特征.例如第 1 阶段用于提取较低层次的边缘特征,第 5 阶段则用于提取更广泛,更抽象的边缘特征.通过这样的卷积操作和阶段划分,RCF-CLF 能够在不同阶段上提取到多尺度的边缘特征,并逐渐获得更广泛,更抽象的特征

信息,从而提高边缘检测的准确性.

为了提取到更准确,更丰富的边缘信息,本文在第5阶段引入HDC卷积结构,将原来3个膨胀系数为2的卷积替换为膨胀率分别为1、2、5的3×3大小的膨胀卷积.这样第5阶段不仅可以在不增加额外参数的情况下增大网络的感受野,而且还能避免相同膨胀卷积的串联造成的网格效应.

1.2 特征增强模块 FEM

在卷积神经网络中,每个阶段的输出包含着越来越抽象的特征表示.然而,由于卷积操作的局部感受野性质,每阶段输出可能无法捕捉到图像中不同尺度和粗细不同的边缘信息.同时每阶段仅进行降维处理还可能存在大量的冗余信息,影响最终的检测结果.

因此,为了解决这种问题,本文设计一种特征增强模块FEM,通过对每个阶段应用多个不同膨胀率的并行膨胀卷积,增加网络的感受野,提高网络对不同尺度边缘的感知能力.最后,通过将原始输入特征与学习到的多尺度边缘特征相加,可以实现特征的融合.相加操作有助于保留原始特征的细节,并抑制冗余噪声,进而使得网络可以更好地聚集于有意义的边缘特征,提高边缘提取的准确性和鲁棒性.

特征增强模块FEM的结构如图2所示.

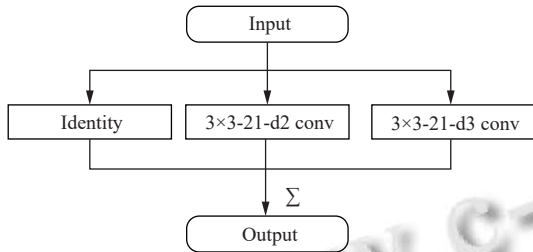


图2 FEM模块结构图

为了避免在模型中使用大尺度卷积核导致计算量和训练难度增加,采用扩张卷积代替大尺度卷积核.这种方法可以在不增加模型参数数量的前提下学习到更大尺度的特征信息.首先,针对降维融合后的特征 f ,使用扩张系数为2和3,卷积核大小为3×3的扩张卷积对图像进行特征提取,然后,通过ReLU激活函数得到两个特征图 f_1 和 f_2 .最后,将 f_1 、 f_2 和 f 进行特征融合,生成最终的输出特征图 f_3 ,并将其作为下一阶段的输入进行进一步处理.每层的结构类似,具体的计算公式如下:

$$f_1 = \text{ReLU}(\otimes_{3 \times 3}(f, d = 2)) \quad (1)$$

$$f_2 = \text{ReLU}(\otimes_{3 \times 3}(f, d = 3)) \quad (2)$$

$$f_3 = f + f_1 + f_2 \quad (3)$$

其中,ReLU表示修正线性单元激活函数; $\otimes_{3 \times 3}(\cdot, d)$ 表示卷积核大小为3×3,膨胀系数为 d 的卷积输出.

1.3 跨层融合模块 CLF

受U-Net^[26]模型中Decoder模块的启发,本文设计了一种自底向上的跨层融合模块cross-layer fusion (CLF),用来避免由于网络模型加深而导致的细节信息丢失.该设计通过将高层的抽象信息逐步反馈到低层,使得低层特征也能够关注到高层的抽象特征,从而达到减少信息丢失的目的.为了进一步获取更精细的特征信息,在CLF中引入了CBAM注意力机制,通过将注意力集中于物体边缘区域,以达到提高边缘检测精确度的目的.

如图3所示,CLF的结构由两个部分组成,包括边缘融合模块EFM和残差偏置模块RBM.其中在边缘融合模块中,首先,通过上采样将高层特征映射恢复到与低层特征相同的尺寸,其次为了获得高度压缩和重要性加权后的边缘特征,使用CBAM注意力块对低层特征进行操作,最后,将得到的低层特征与高层特征进行加和得到融合后的输出特征 f_o .边缘融合模块的整体过程可以被表示为式(4)–式(6):

$$f_{k+1} = \text{upsample}(f_{k+1}) \quad (4)$$

$$fk = \text{CBAM}(fk) \quad (5)$$

$$f_o = f_{k+1} + f_k \quad (6)$$

其中,upsample表示双线性插值上采样操作;CBAM表示注意力机制.

残差偏置模块的输入为边缘融合模块的输出 f_o ,首先,利用Canny算子提取输入图像 f_o 的边缘图 f_c ,其次,对 f_c 分别进行最大池化和平均池化,将生成的特征 f_{\max} 和 f_{avg} 进行通道维度拼接后执行特征卷积操作,最后,将经激活函数得到的偏置特征 f_p 与输入特征 f_o 进行特征相乘,得到输出 f_e .残差偏置模块的整体过程可以被表示为式(7)–式(9):

$$f_c = \text{Canny}(f_o) \quad (7)$$

$$f_{\max} = \max(f_c) \quad (8)$$

$$f_{\text{avg}} = \text{avg}(f_c) \quad (8)$$

$$f_e = P(\otimes_{1 \times 1}(\text{cat}(f_{\max}, f_{\text{avg}}))) \times f_o \quad (9)$$

其中, *Canny*表示传统算法 *Canny* 算子的; *max*和*avg*分别表示最大化池化操作和平均池化操作; *cat*表示通道维

度拼接; $\otimes_{1 \times 1}$ 表示使用 1×1 卷积进行特征升维; *P*表示 Sigmoid 激活函数

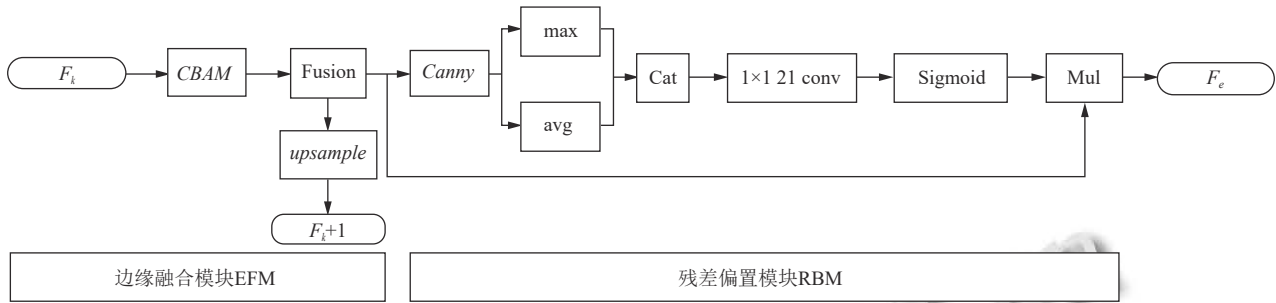


图3 CLF 模块结构图

1.4 损失函数

边缘检测数据集通常需要由多个标注者对边缘对象进行标注, 由于每个标注者对于边缘对象的认知存在差异, 因此对同一幅图像的边缘标记可能存在一定的不一致性. 为了解决这一问题, 本文采用平均值的方式生成一个边缘概率图, 该概率图的范围从 0 到 1, 其中 0 表示该像素点没有被任何标注者标记, 1 表示所有标注者都给该像素点打上了标记. 在进行样本分类时, 将边缘概率大于阈值 θ 的像素视为正样本, 将边缘概率等于 0 的像素视为负样本. 这样做的原因在于, 如果某个像素点被标记但其边缘概率较小, 则该像素点可能存在一定的争议, 将其视为正或负样本可能会混淆网络输出. 因此忽略了这些像素点并将阈值 θ 设置为 0.3, 以减少网络输出的含糊程度.

本文采用式 (10) 来计算每个像素相对于其标签的损失:

$$l(X_i; W) = \begin{cases} \partial \times \log(1 - P(X_i; W)), & y_i = 0 \\ 0, & 0 < y_i \leq \theta \\ \beta \times \log(P(X_i; W)), & \text{otherwise} \end{cases} \quad (10)$$

其中,

$$\begin{cases} \partial = \gamma \frac{|Y^+|}{|Y^- + Y^+|} \\ \beta = \gamma \frac{|Y^-|}{|Y^- + Y^+|} \end{cases} \quad (11)$$

其中, Y^+ 和 Y^- 分别表示正样本集和负样本集. 超参数 γ 用于平衡正、负样本的数量. 激活值和像素 i 处的真值概率分别用 X_i 和 y_i 表示. W 表示网络结构中要学习到的所有参数. 因此, 整体损失函数可以表示为:

$$l(W) = \sum_{i=1}^{|I|} \left(\sum_{k=1}^K l(X_i^k; W) + l(X_i^{\text{fuse}}; W) \right) \quad (12)$$

其中, X_i^k 表示第 k 阶段的激活值, X_i^{fuse} 表示融合层的激活值, $|I|$ 表示图像 I 中的像素数, K 是阶段数, 在本文的算法中 K 等于 4.

2 实验设计与结果分析

2.1 数据集和数据处理

本文以 BSDS500 和 BIPED 两个数据集来评估 RCF-CLF 网络的性能. 其中, BSDS500 数据集被广泛用于边缘检测领域. 该数据集包含 200 张训练图像、100 张验证图像和 200 张测试图像. 每个图像都由 4-9 个标注者进行标记. 为了增加模型的鲁棒性和泛化能力, 本文还采取了一系列的数据增强操作, 包括随机翻转、缩放和旋转等. 通过这些操作, 训练集和验证集的数据得到了扩充, 使得模型可以更好地适应不同的场景和变化.

BIPED 数据集由 DexiNed 论文作者提供. 该数据集包含了 250 张户外图像, 分辨率为 1280×720 . 与 BSDS500 数据集相比, BIPED 数据集是边缘检测领域中的一个相对较新的数据集. 为了评估算法在不同场景下的准确性和鲁棒性, BIPED 数据集的提供者还进行了额外的增强处理. 这些增强方法包括图像分割、旋转、裁剪、翻转和伽马校正等. 因此, 本文直接采用原作者提供的数据增强方法对 BIPED 数据集进行数据扩充, 并使用扩充后的数据集来评估本文方法在边缘检测上的表现.

2.2 实验环境

本文中的所有实验均在 PyTorch 框架环境下进行, 使用单一的 NVIDIA GeForce RTX3060 服务器 (12 GB 显存) 进行训练和测试. 针对两个数据集均采用与 RCF

网络相同的训练策略,即在训练过程中,每个批次(batch)的大小设置为1,用SGD优化器进行权重更新,初始学习率为0.00001,并进行了30轮(epochs)的迭代训练.在测试阶段,批次大小为1,以评估网络的性能表现.

2.3 评价指标

基于深度学习的边缘检测算法输出一般都是一张概率图.该图需要经过一系列的处理,才能生成最终的边缘图像.为了实现这一目标,通常需要使用一个阈值对输出的边缘概率图进行操作,用于将其转化为边缘图像.常用的设置边缘检测阈值的方法包括最优数据集尺度(optimal dataset scale, ODS)和最优图像尺度(optimal image scale, OIS).前者是对所有的图像都采用一个固定的阈值,使得整个数据集上的 F 值(F -score/ F -measure)取得最大化,后者则是为每一张不同的图像样本设置一个不同的阈值,使得该图像上的 F 值最大.因此,为了在最大程度上评价一种方法的泛化能力和鲁棒性,边缘检测算法一般采用ODS, OIS和平均精度(average precision, AP)作为性能评价指标.

2.4 消融实验

为了评估所提出的不同改进方法对原始RCF算法性能的影响,进行了消融实验.实验旨在验证不同组合策略在两个数据集上的效果.这样可以对改进方法的效果进行全面评估,并确定其对算法性能的影响程度.实验汇总如表1所示.

表1 不同组合策略在不同数据集上的实验结果

数据集	组合策略	ODS	OIS	AP
BSDS500	—	0.797	0.815	0.819
	HDC	0.801	0.819	0.804
	HDC+FEM	0.804	0.817	0.810
	HDC+FEM+EFM	0.809	0.820	0.815
	HDC+FEM+FEM+RBM	0.812	0.826	0.827
BIPED	—	0.849	0.861	0.482
	HDC	0.854	0.869	0.582
	HDC+FEM	0.879	0.882	0.877
	HDC+FEM+EFM	0.881	0.887	0.912
	HDC+FEM+FEM+RBM	0.893	0.901	0.945

通过对表1的结果分析,可以发现本文所提出的模块在BSDS500数据集上的性能提升在1.5个百分点,这表明,这些模块能够提升模型的整体性能.值得注意的是,这些模块在BIPED数据集上展现了较好的性能提升.具体而言,在加入HDC模块后,ODS和OIS均增加了1个百分点左右,AP值更是增加了10个百

分点,说明HDC结构设计相较于叠加相同膨胀系数确实能够产生更高的精度,此外,引入特征增强模块FEM后,ODS和OIS指标分别提高了2个百分点和1.3个百分点,其中AP值更是提高了近30个百分点.表明,以并行的不同膨胀系数的卷积结构不仅能够增大感受野,而且有效地降低了图像中的噪声.最后,加入CLF模块,取得了最佳的性能.一方面是CLF模块中的EFM模块通过融合不同尺度下的特征信息,增强了特征的多样性,进而使得网络能够更好地捕获边缘信息.另一方面则是CLF模块中的RBM模块通过利用Canny算子和池化操作构建一种特征偏置系数,进一步强化边缘,弱化非边缘信息.

根据表1中的消融实验结果,可以观察到本文算法中的各组件在BSDS500数据集和BIPED数据集上表现出较大的差异.这种差异主要源于数据集的特点.图4展示了两个数据集的差异,图4(a)是BSDS500数据集中的原始图像和标签图像,图4(b)是BIPED数据集中的原始图像和标签图像.可以看出,即使是描述相同的轿车边缘,两个数据集的边缘描述之间存在巨大差异.在BSDS500数据集中,主要关注轿车的整体轮廓边缘,忽略了其他部分的边缘描述;而在BIPED数据集中,对轿车的整体细节进行了详细描述,包含了大量的边缘信息.

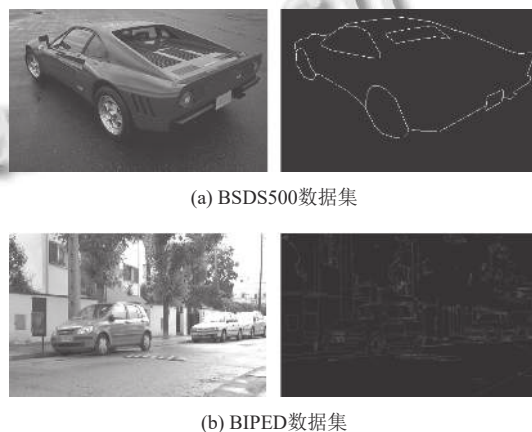


图4 BSDS500数据集与BIPED数据集的差异对比

因此,BSDS500数据集通常被广泛用于图像轮廓/边界和语义分割任务,但相对于专门用于边缘检测任务的BIPED数据集而言,BSDS500数据集中包含的图像边缘细节较少.在这种情况下,BSDS500数据集的标签图仅包括轮廓边缘,缺乏对细节边缘的充分描述,这使得对RCF-CLF模型检测到的细节边缘进行全面评

估和验证变得困难. 因此, 在 BSDS500 数据集上, 本文算法中的各组件仅能够提升 1.5 个百分点的性能, 而在 BIPED 数据集上, 则能够取得良好的性能表现.

2.5 对比实验与分析

为了验证本文方法在边缘检测任务中的性能, 将本文算法与近几年主流的边缘检测算法 RCF、CED^[27]、DexiNed 和 CHRNet^[28] 分别在 BSDS500 和 BIPED 数据集上进行训练和测试, 实验结果显示在表 2 和图 5 中.

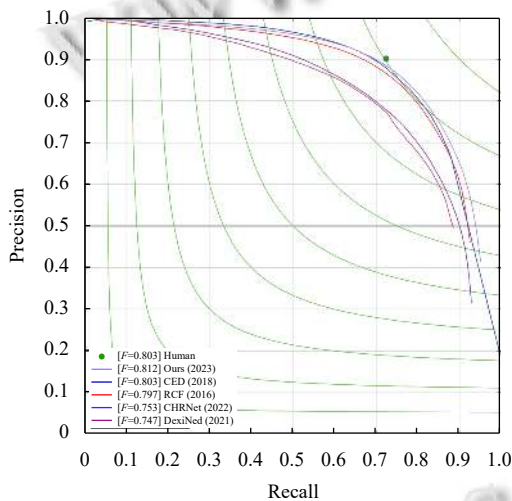
基于 BSDS500 数据集的实验结果表明, RCF-CLF 方法相对于原始的 RCF 方法, 在整体性能上实现了一定的提升. 具体而言, ODS 和 OIS 指标分别提高了 1.5 个百分点和 1.1 个百分点.

为了更直观地展示本文方法与其他方法之间的差异,

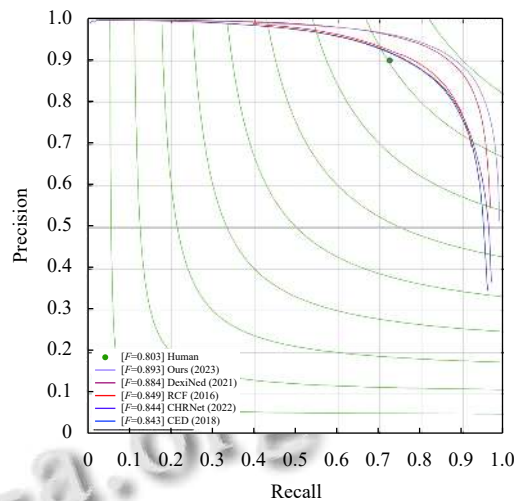
图 5(a) 展示了本文方法以及其他方法的 P-R (precision-recall) 曲线图. 通过观察该曲线图, 可以看出本文方法的 P-R 曲线位于其他方法之上, 说明本文方法在边缘检测过程中能够取得更高的精确率和召回率. 进一步验证了本文方法在边缘检测任务中的良好性能.

表 2 不同方法在 BSDS500 和 BIPED 数据集上的实验结果

数据集	方法	ODS	OIS	AP
BSDS500	RCF	0.797	0.815	0.819
	CED	0.803	0.820	0.871
	DexiNed	0.747	0.770	0.739
	CHRNet	0.753	0.773	0.778
	Ours	0.812	0.826	0.827
BIPED	RCF	0.849	0.861	0.482
	CED	0.843	0.853	0.891
	DexiNed	0.884	0.892	0.926
	CHRNet	0.844	0.854	0.899
	Ours	0.893	0.901	0.945



(a) 不同方法在 BSDS500 数据集上的 P-R 曲线图



(b) 不同方法在 BIPED 数据集上的 P-R 曲线图

图 5 不同方法的 P-R 曲线

在 BIPED 数据集上的实验结果显示, 本文所提出的方法在性能上取得了较好的改进. ODS、OIS 和 AP 分别为 0.893、0.901 和 0.945, 相较于先前的 RCF 方法, 本文的方法表现出明显的提升, 尤其是在 AP 指标上实现了近 1 倍的增长. 图 6 展示了本文方法与其他方法在边缘检测任务上的效果对比图. 图 6 的前两列中, 展示了原始图像和其对应的基准真实边缘. 接下来的 5 列分别展示了使用 RCF、CED、DexiNed、CHRNet 以及本文方法所得到的原始边缘图. 通过对比这些边缘图, 可以观察到本文方法相对于其他方法在边缘检测方面的优势. 首先, 本文方法在边缘检测的准

确性方面表现出更高的水平. 通过观察第 1 行图片中的三角形标识区域, 我们可以发现其他方法存在边缘丢失或错误检测的问题, 而本文方法则能够更好地贴合真实边缘. 其次, 在结果的清晰性方面, 本文方法也表现出较大的优势. 观察第 2 行图片中红色框标识的大树区域, 我们可以注意到本文方法的检测结果更加干净、线条也更加清晰连贯, 相比之下, 其他方法存在边缘断裂或模糊等不足之处. 最后, 通过观察图 5(b) 中的 P-R 曲线位置, 可以发现本文方法具有更高的精确率和召回率. P-R 曲线的优势进一步验证了本文方法在正例检测和负例排除方面的良好表现. 综上所述, 本文

方法在边缘检测方面相对于其他方法具有较好的优势,包括更高的准确性、更清晰的结果以及更高的精确率

和召回率. 这些优势充分证明了本文方法在边缘检测任务中的出色性能.

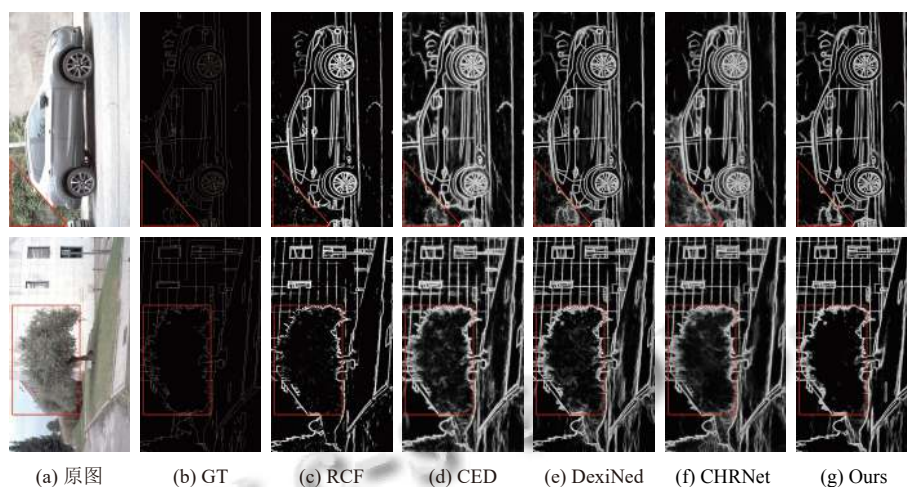


图6 在 BIPED 数据集上不同方法的边缘检测结果的可视化

3 结论

图像边缘检测是图像处理领域中较为基础的研究问题,也是很多计算机视觉任务的技术基础.已有的许多研究工作表明边缘检测在图像特征描述、高级特征提取、目标识别以及图像分割等领域都具有重要意义.本文提出的方法在边缘检测中利用了 RCF 的网络结构,通过使用 HDC 结构设计、特征增强模块和跨层融合模块更大程度上发挥了网络性能.因此,该方法不仅能够正确地检测边缘,而且能够更加注重细节信息.在使用专门用于边缘检测的数据集进行测试后,实验结果表明,相对于原始的 RCF 网络,本文提出的方法在 ODS、OIS 和 AP 这 3 个方面均得到提升,展现出更高的边缘检测性能.

参考文献

- 1 Al-Amri SS, Kalyankar NV, Khamitkar SD. Image segmentation by using edge detection. *International Journal on Computer Science and Engineering*, 2010, 2(3): 804–807.
- 2 Rother C, Kolmogorov V, Blake A. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, 2004, 23(3): 309–314. [doi: 10.1145/1015706.1015720]
- 3 Ferrari V, Fevrier L, Jurie F, *et al.* Groups of adjacent contour segments for object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(1): 36–51. [doi: 10.1109/TPAMI.2007.1144]
- 4 Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014. 580–587.
- 5 Hsieh CJ, Huang TK, Hsieh TH, *et al.* Compressed sensing based CT reconstruction algorithm combined with modified Canny edge detection. *Physics in Medicine & Biology*, 2018, 63(15): 155011.
- 6 Nikolic M, Tuba E, Tuba M. Edge detection in medical ultrasound images using adjusted Canny edge detection algorithm. *Proceedings of the 24th Telecommunications Forum (TELFOR)*. Belgrade: IEEE, 2016. 1–4.
- 7 Ghandorh H, Boulila W, Masood S, *et al.* Semantic segmentation and edge detection—Approach to road detection in very high resolution satellite images. *Remote Sensing*, 2022, 14(3): 613. [doi: 10.3390/rs14030613]
- 8 Lu XY, Zhong YF, Zheng Z, *et al.* Cascaded multi-task road extraction network for road surface, centerline, and edge extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5621414.
- 9 Manzi MSD, Durrheim RJ, Hein KKA, *et al.* 3D edge detection seismic attributes used to map potential conduits for water and methane in deep gold mines in the Witwatersrand basin, South Africa. *Geophysics*, 2012, 77(5): WC133–WC147. [doi: 10.1190/geo2012-0135.1]
- 10 Prasad KND, Pham LT, Singh AP. A novel filter “ImpTAHG” for edge detection and a case study from

- cambay rift basin, India. *Pure and Applied Geophysics*, 2022, 179(6-7): 2351–2364. [doi: [10.1007/s00024-022-03059-z](https://doi.org/10.1007/s00024-022-03059-z)]
- 11 Rani S, Ghai D, Kumar S. Object detection and recognition using contour based edge detection and fast R-CNN. *Multimedia Tools and Applications*, 2022, 81(29): 42183–42207. [doi: [10.1007/s11042-021-11446-2](https://doi.org/10.1007/s11042-021-11446-2)]
- 12 Hussain BA, Hathal MS. Developing arabic license plate recognition system using artificial neural network and canny edge detection. *Baghdad Science Journal*, 2020, 17(3): 0909. [doi: [10.21123/bsj.2020.17.3.0909](https://doi.org/10.21123/bsj.2020.17.3.0909)]
- 13 Gao X, Ram S, Rodríguez JJ. Exploiting bilinear interpolation and predictive particle swarm optimisation for tilt correction of vehicle license plate images. *International Journal of Image Mining*, 2023, 4(2): 177–192. [doi: [10.1504/IJIM.2023.132346](https://doi.org/10.1504/IJIM.2023.132346)]
- 14 Canny J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, PAMI-8(6): 679–698. [doi: [10.1109/TPAMI.1986.4767851](https://doi.org/10.1109/TPAMI.1986.4767851)]
- 15 Li YB, Liu BL. Improved edge detection algorithm for canny operator. *Proceedings of the 10th IEEE Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*. Chongqing: IEEE, 2022. 1–5.
- 16 Arbeláez P, Maire M, Fowlkes C, *et al.* Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(5): 898–916. [doi: [10.1109/TPAMI.2010.161](https://doi.org/10.1109/TPAMI.2010.161)]
- 17 Dollár P, Zitnick CL. Fast edge detection using structured forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(8): 1558–1570. [doi: [10.1109/TPAMI.2014.2377715](https://doi.org/10.1109/TPAMI.2014.2377715)]
- 18 Xie SN, Tu ZW. Holistically-nested edge detection. *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*. Santiago: IEEE, 2015. 1395–1403.
- 19 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *Proceedings of the 3rd International Conference on Learning Representations*. San Diego: ICLR, 2015.
- 20 Liu Y, Cheng MM, Hu XW, *et al.* Richer convolutional features for edge detection. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 5872–5881.
- 21 Su Z, Liu WZ, Yu ZT, *et al.* Pixel difference networks for efficient edge detection. *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021. 5097–5107.
- 22 Wang PQ, Chen PF, Yuan Y, *et al.* Understanding convolution for semantic segmentation. *Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. Lake Tahoe: IEEE, 2018. 1451–1460. [doi: [10.1109/WACV.2018.00163](https://doi.org/10.1109/WACV.2018.00163)]
- 23 Odena A, Dumoulin V, Olah C. Deconvolution and checkerboard artifacts. *Distill*, 2016, 1(10): e3.
- 24 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 3–19.
- 25 Soria X, Riba E, Sappa A. Dense extreme inception network: Towards a robust cnn model for edge detection. *Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision*. 2020. 1912–1921.
- 26 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.
- 27 Wang YP, Zhao X, Li Y, *et al.* Deep crisp boundaries: From boundaries to higher-level tasks. *IEEE Transactions on Image Processing*, 2019, 28(3): 1285–1298. [doi: [10.1109/TIP.2018.2874279](https://doi.org/10.1109/TIP.2018.2874279)]
- 28 Elharrouss O, Hmamouche Y, Idrissi AK, *et al.* Refined edge detection with cascaded and high-resolution convolutional network. *Pattern Recognition*, 2023, 138: 109361. [doi: [10.1016/j.patcog.2023.109361](https://doi.org/10.1016/j.patcog.2023.109361)]

(校对责编:牛欣悦)