文章编号:1674-2974(2024)08-0013-10 DOI:10.16339/j.cnki.hdxbzkb.2024274

复杂稠密网络下的并置多尺度融合边缘检测模型

党建武[1],张天胤[2],田彬[1†]
(1. 兰州交通大学 电子与信息工程学院,甘肃 兰州 730070;
2. 兰州交通大学 光电技术与智能控制教育部重点实验室,甘肃 兰州 730070)

摘 要:边缘检测是计算机视觉任务的基础.目前的技术主要依赖于深度学习,但是大多数的模型在评价阶段会借助非极大值抑制来提高预测边缘的准确率.该策略仅着重关注预测边缘的准确性,没有同时考虑边缘的粗细程度.针对这一问题,本文基于复杂稠密网络,提出了一种新的特征融合策略.该策略在传统深度学习边缘检测器侧输出的基础上,添加了自顶向下的并置多尺度融合架构.此架构可以更好地将高层的高语义特征与低层的高分辨率纹理特征结合,有效地抑制了边缘预测中存在的背景模糊和伪线条的现象.在特征连接处使用Concat block块代替单一的Concat操作,更好地融合了不同尺度的语义信息.最后使用一个简单的注意力融合块融合多个网络输出,并结合跟踪损失对输出的不同尺度的预测图进行深度监督.此方法不依赖于非极大值抑制,并且充分利用了目标的多尺度、多层次信息,在精细边缘图像的同时也提升了预测的准确性.实验结果显示,在未使用和使用形态学非极大值抑制方案的情况下,所提出的模型在BIPED数据集上的ODS、OIS、AP分别达到了0.891、0.895、0.900和0.894、0.899、0.931,优于所有比较算法.在MDBD的数据集上也取得了最优的结果.

关键词:边缘检测;卷积神经网络;注意力机制;多尺度融合
中图分类号:TP181 **文献标志码**:A

Multi-scale Fusion Edge Detection Model with Spatial Co-location Rule Based on Dense Extreme Inception Network

DANG Jianwu[1], ZHANG Tianyin[2], TIAN Bin[1†]
(1. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;
2. Key Lab of Opt-electronic Technology and Intelligent Control of Ministry of Education, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: Edge detection is the basis of many computer vision tasks. Current techniques mainly rely on deep learning, but most models improve the accuracy of predicted edges using Non-Maximum Suppression in the evaluation stage. These models only focus on the accuracy of predicted edges without considering the coarseness and

* **收稿日期**:2023-08-25
基金项目:国家自然科学基金资助项目(62067006,62367005),National Natural Science Foundation of China(62067006,62367005);甘肃省知识产权计划项目(21ZSCQ013),Gansu Provincial Intellectual Property Project(21ZSCQ013);2022年度中央引导地方科技发展资金项目(22ZY1QA002),Central Government Guides Local Funds Project for Science and Technology Development of Gansu Province in 2022(22ZY1QA002);教育部人文社会科学研究项目(21YJC880085),The Ministry of Education of Humanities and Social Science Project(21YJC880085)
作者简介:党建武(1963—),陕西富平人,兰州交通大学教授,博士
† 通信联系人,E-mail:tianbin@mail.lzjtu.cn

fineness of the edges. To address this issue, this paper proposes a new feature fusion strategy based on the dense extreme inception network. This strategy incorporates top-down multi-scale fusion edge detection with spatial co-location rule and retains the multi-network structure based on the traditional deep learning edge detection side output. The proposed strategy can better integrate the high semantic characteristic of high-layer information with the high-resolution texture characteristic of low-layer information, thereby suppressing pixel confusions in backgrounds and lines that are predicted incorrectly in edge detection. In the feature connection, Concat block is used to replace the single operation of Concat, to better fuse semantic information in different scales. Lastly, a simple attention fusion block is used to fuse outputs of multiple networks. Also, different output prediction maps at different scales are deeply supervised combining the tracing loss. This model is independent of Non-Maximum Suppression. By fully utilizing the multi-scale and multi-level information of the target image, this model improves the accuracy of prediction along with improving images' edges. The experimental results show that without the morphological Non-Maximum Suppression, on the BIPED data set, the proposed model on ODS, OIS, and AP are 0.891, 0.895, 0.900, respectively; with the morphological Non-Maximum Suppression, the proposed model on ODS, OIS, and AP are 0.894、0.899、0.931, respectively, which is superior to all comparison algorithms involved in this article. Also, on the MDBD data set, optimal results were also achieved.

Key words: edge detection; convolutional neural network; attention mechanism; multi-scale fusion

边缘检测是一种从图像中提取显著性边缘和边界的技术.几十年以来一直都是计算机视觉面临的主要挑战之一.边缘检测技术一直被认为是一种低级的下游任务,各种高级任务比如形状匹配[1-2]、显著性检测[3-4]、图像分割等[5-6],都得益于边缘检测技术的发展.经过了研究人员近几十年以来的研究,提出了很多方法.这些方法大致可以分为四类:早期的方法、手工特征、初级神经网络的方法和基于深度学习的方法.

在早期,边缘检测方法主要集中在使用颜色、纹理、梯度等局部底层特征信息.这些方法可以获得细致、边界清晰的边缘图,Huntsberger等[7]讨论了一种定量方法,用于选择颜色坐标以提取视觉上显著的边缘和边界,而Huang等[8]和Torre等[9]则基于零交叉理论提出了相应的算法.Sobel提出了著名的Sobel[10]算子,通过计算图像的梯度图并对其进行阈值处理,得到边缘信息.作为 Sobel的扩展版本,Canny [11]在预处理步骤中引入了高斯平滑,并采用双阈值的方法获得对噪声更具鲁棒性的边缘检测结果.随着方法不断改进,效果也在不断提升,然而其局限性也是非常明显的,由于这些方法使用的是局部、低级的特征信息,无法有效抑制高频纹理,在许多场景下无法实现良好的鲁棒性,因此难以适应现今的应用需求.

随后的研究者们利用低级信息,如强度、梯度和纹理,手动设计特征,并采用复杂的学习范式来区分边缘和非边缘像素[12-13].Dollar等[14]利用随机决策森林来表征局部图像块的结构,并输入颜色、梯度特征,以得到高质量的结构森林边缘检测结果.除了上述方法之外,还有一些其他的方法,如PMI[15]、GPB-OWT-UCM[16]等.然而,这些方法都基于手动设计的特征,无法充分表示语义上显著的边缘信息,因而存在一定的局限性.

随着神经网络在计算机视觉中的广泛应用,边缘检测也进入了神经网络时代.其中,图像变换的神经网络最近邻场(N^4-Fields)[17]是经典方法之一.该方法利用卷积神经网络(CNN)来提取每个图像块的特征,并在字典中寻找相似的边缘,最终将它们融合以得到边缘检测结果.DeepEdge算法[18]也采用图像块的方法,通过将图像块分类到不同边缘类型并融合它们,得到最终的边缘检测结果.尽管早期的神经网络方法在特定数据集上表现良好,但它们大多数是基于局部策略的,只关注单个图像块,未能充分利用高级信息,导致与真实场景存在较大差距.

近年来,随着卷积神经网络的发展,特别是在FCN[19]、AlexNet[20]、VGG[21]、ResNet[22]、Xception[23]、DenseNet[24]等深度卷积网络结构提出之后,学者们

应用这些网络结构,利用卷积层和全连接层获取初始的局部边缘特征图,最后使用非极大值抑制(NMS)等后处理方法获取边缘图像.其中,最广泛使用的主干网络为VGG模型.由Xie等[25]提出的HED就是在VGG网络[21]的基础上提出的第一个端到端的边缘检测方法.该方法使用多尺度策略,在卷积层的侧面插入一个侧输出层,并对侧面输出层进行深度监督,最终通过连接不同层的输出得到结果.在此基础上,Liu等提出了RCF方法[26-27].他认为不同卷积层之间的信息可以互补.该方法利用每一层之间的信息,但仅利用池化层前最后一个卷积层的信息.因此,RCF方法利用每个卷积层的信息进行边缘检测.DexiNed是在HED和Xception网络[23]的基础上提出的,其在BIPED[28]和MDBD[29]数据集上都取得了最优的结果.除了上述提到的两种算法外,还有许多优秀的检测方法,如CED[30]、BDCN[31]、CATS[32]、PiDiNet[33]等.

虽然从传统方法到基于卷积神经网络的方法,检测效果有了很大的提高,但仍存在一些有待改进的问题.例如大多数的模型,为了获得清晰的边缘图,得到更高的分数,在测试前对网络输出都应用了形态学非极大值抑制.该方案抑制了粗边缘周围的像素模糊,让评价结果看似很高,然而抑制后的边缘图并不美观,实际使用的却是网络原始输出的相对较粗并且有较多像素模糊的预测图.一个良好的边缘检测器必须在边缘的正确性(区分边缘和非边缘像素)和边缘的"脆度"(精确定位边缘像素)之间取得平衡.

为了改善这一问题,本文在DexiNed和CATS的启发下,提出了一种新的并置多尺度融合架构.我们首先保留了DexiNed的复杂稠密网络结构,然后在其原有的侧输出的基础上添加了空洞卷积[34],通过不同膨胀率的卷积来达到丰富多尺度的目的[35],紧接着引入一个简单的空间注意力结构,用于去除背景纹理信息.在特征融合时,设计了一个多尺度并置融合架构,最后引入CATS中的边缘追踪和纹理抑制损失函数[32].我们的方法在BIPED数据集上取得了很好的效果,相较于最先进的方法,我们的方法在未使用NMS的条件下,OIS、ODS和AP分别达到了0.891、0.895、0.900,提升了2%~4%,在使用NMS的条件下,OIS、ODS和AP分别为0.894、0.899、0.931,也达到了最优结果.本文的主要贡献如下.

1)本文提出了一种新的边缘检测方法,该方法可以在提升边缘预测准确性的同时精细边缘图像.

2)本文提出了一种自顶向下的并置多尺度融合策略,该策略更好地融合了不同尺度的语义信息,减少边缘预测中的背景和线条周围的像素模糊.

3)本文提出的边缘检测方法不依赖于NMS,生成的边缘预测图背景清晰、线条精细.

4)实验结果表明,在BIPED数据集上,相比于其他使用NMS的研究方法,本文方法在未使用NMS的情况下取得了更好的结果,在使用NMS的条件下,依旧优于其他对比算法.

1　本文模型

我们的模型不依赖于预训练权重,并且允许端到端进行训练直接生成边缘图像.如图1所示,下面通过几个部分来分别构建算法架构.

1.1　主干网络

主干网络是一个类似于Xception的架构,如图1(a)左侧所示,包含了6个块,每个块是由一组卷积层组成的.每个块生成的特征图被单独提取出来,经过一系列的侧处理输入USNet中生成侧输出边缘图.最后这些中间边缘特征图被融合成为一个单一的边缘预测.图中蓝色块为2个3×3卷积进行堆叠,在每一个卷积后加入了批量归一化函数和非线性单元ReLU.第一个块中的卷积步幅为2.红色矩形是最大池化操作,内核大小为3×3,步幅为2.随着大量的卷积操作和网络深度的增加,重要的边缘特征会消失,因此引入了并行跳跃连接以缓解这一问题.所有跳跃连接在融合策略上采用的是相加.从第二个块的最大池化输出的特征会与第三个块内部进行密集链接,在块内部的密集连接与主干相连时采用的是平均的操作.这是本文保留的DexiNed的主干特征提取的部分.

1.2　侧结构

1.2.1　膨胀卷积与空间注意力

如图1(c)所示,在主干特征提取网络之后得到了6个边缘特征,为了细化特征图,我们首先构建一个膨胀卷积块,通过不同卷积核的卷积最后相加,达到了丰富多尺度边缘信息的目的,我们采用的卷积

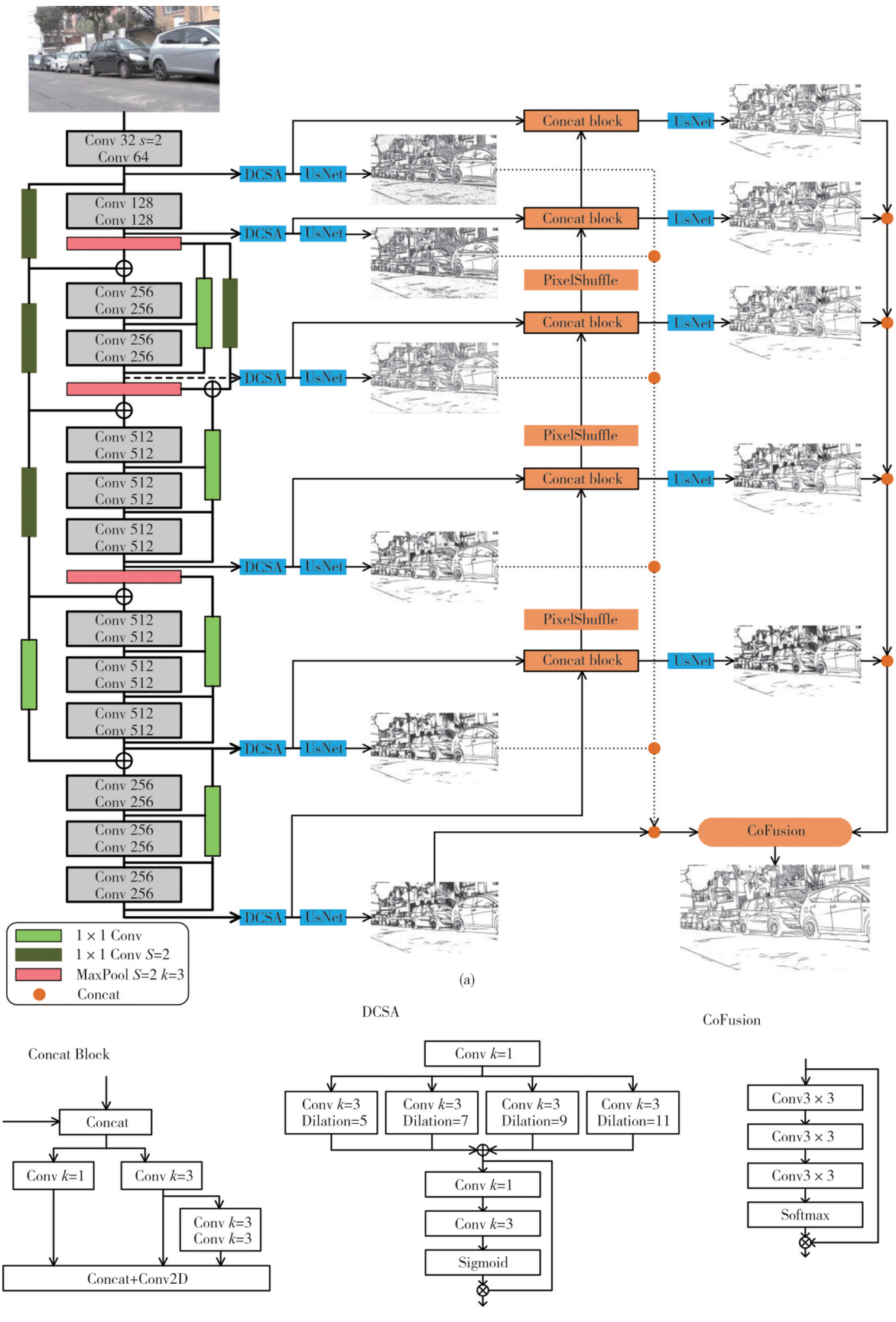


图 1 整体模型和各个主要块的结构

Fig.1 Overall model and structure of each major block

核分别为5,7,9,11,输入通道数为C,输出通道数为M(M<C),以减少后续融合阶段的计算开销.紧接着引入一个空间注意力块,用于消除背景纹理信息.

1.2.2　并置多尺度融合

为了获取更加丰富的多尺度信息,我们在原有的6个输出的基础上添加了自顶向下的路径,把高层的特征传下来,补充低层的语义信息,让输出可以获得高分辨率、强语义的特征,这一方法能够有效减少由低层纹理信息引起的最终边缘预测中出现背景模糊和伪线条的问题,整体传播路径如图1(a)所示.在横向特征融合路径上,以往的工作中使用的是单一的add或者Concat操作,本文引入Concat Block块,在Concat后引入一组多分支卷积,更好地融合了来自不同路径不同尺度的语义信息,如图1(b)所示,在自顶向下的路径中对于不同尺度的图像首先采用PixelShuffle上采样操作后再进行并置融合.对于所有的侧输出,我们均采用转置卷积进行循环上采样到与原图相同大小.由图1(a)可以看出所提网络有11个侧边缘预测图像,最终的结果来自这些预测的融合.由于受到CATS[13]的启发,本文采用CoFusion块来融合这些边缘图像.我们对该块在维度和归一化上进行了轻微的修改,以适应我们提出的多网络结构,如图1(d)所示,得到最终的融合图像加上11张侧输出图一共12张图分别对GT计算损失进行深度监督.

1.3　损失函数

本文的损失函数主要来源于CATS[32].

1.3.1　加权交叉熵

加权交叉熵函数为大多数边缘检测所使用的损失函数.给定边缘预测为$\hat{Y}=\{\hat{y}_i\}_{i=1}^{N}$,对应的真值标签$Y=\{y_i\}_{i=1}^{N}$,加权交叉熵函数如式(1)所示:

$$L_{\text{ce}}(\hat{Y},Y)=\lambda\alpha\sum_{i\in Y^+}\lg\hat{y}_i-(1-\alpha)\sum_{i\in Y^-}\lg(1-\hat{y}_i) \tag{1}$$

式中:$Y^+=\{i|y_i\in Y,y_i>\delta\}$、$Y^-=\{i|y_i\in Y,y_i=0\}$分别表示边缘和非边缘样本集,$\alpha$表示边缘和非边缘的整体样本集中负样本的比例.由于边缘真值中正负样本数量差距很大,所以引入λ用于平衡正负样本比例.δ是一个像素阈值,用来删除有争议的边缘像素.

加权交叉熵虽然可以有效地监督网络学习合理的边缘图,但其对于边缘样本和非边缘样本的注意力高度不平衡,导致了网络难以区分那些边缘具有相同特征的像素.所以,输出的边缘预测图像常常会出现预测错误的混淆像素,在真正的边缘线条周围出现模糊像素,从而产生粗边,在纹理区域得到假阳性预测,形成暗影.

1.3.2　边缘追踪

由于混淆像素通过共享特征与边缘交织在了一起堆积在边缘周围,使得边缘看起来是一道粗边,因此需要边缘追踪抑制来从混淆像素中分离出清晰的边缘.边缘追踪函数如式(2)所示:

$$L_{\text{bdry}}(\hat{Y},Y)=-\sum_{p\in E}\lg\Big(\sum_{i\in L_p}\hat{y}_i/\Big(\sum_{i\in R_p^e/L_p}\hat{y}_i+\sum_{i\in L_p}\hat{y}_i\Big)\Big) \tag{2}$$

式中:E是边缘标签y中所有边点的集合,R_p^e表示中心点在一个边缘点p上的一个7×7图像补丁,L_p表示R_p^e中所有边点的集合.为了使$L_{\text{bdry}}(\hat{Y},Y)$的值越小,$\sum_{i\in L_p}\hat{y}_i$要增大,$\sum_{i\in R_p^e/L_p}\hat{y}_i$要趋向于0.利用边缘感知函数可以抑制边缘周围的混淆像素并且有效地对网络进行监督.

1.3.3　纹理抑制

边缘追踪函数处理完边缘周围的混淆像素后,剩余的纹理区域的模糊可以通过定义纹理区域抑制函数进行抑制,如式(3)所示:

$$L_{\text{tex}}(\hat{Y},Y)=-\sum_{p\in Y/\hat{E}}\lg\Big(1-\sum_{i\in R_p^t}\hat{y}_i/|R_p^t|\Big) \tag{3}$$

式中:R_p^t表示中心点p不是边缘点的一个3×3图像补丁,$\hat{E}$表示边缘最终函数中用到的所有边缘和混淆像素的集合.要使得$L_{\text{tex}}(\hat{Y},Y)$减小,$\sum_{i\in R_p^t}\hat{y}_i/\left|R_p^t\right|$就要趋向于0.纹理抑制函数实际上与边界跟踪函数是互补的,它很好地抑制了除边界以外的纹理区域的像素模糊.

1.3.4　整体损失函数

利用边缘追踪和纹理抑制,最终损失函数表示为:

$$\text{FLoss}(\hat{Y},Y)=L_{\text{ce}}+\lambda_1L_{\text{bdry}}+\lambda_2L_{\text{tex}} \tag{4}$$

式中:$\hat{Y}$和Y分别表示边缘预测和边缘的真值标签,λ_1和λ_2为超参数用于平衡边缘追踪损失和纹理抑制损失.在模型训练中,L_{ce}用来对边缘进行粗学习,L_{bdry}通过边缘定位对边缘进行细化,L_{tex}对其余纹理区域

进行抑制.该损失函数相比单用加权交叉熵可以实现生成更加清晰细致的边缘预测图像.

2 实验与结果

本部分主要介绍实验与结果.在本文中采用了三个数据集对我们的模型进行定量评估,分别为BIPED、MDBD、BRIND.

2.1 数据集

BIPED:该数据集是提出DexiNed模型的文章中给出的专门用于边缘检测的数据集.BIPED包含250张1 280像素×720像素的巴塞罗那户外图像,其中200张用于训练,50张用于测试.该数据集的边缘图仅有一个经过验证的注释.虽然之前有很多数据集例如BSDS500、NYDU、MDBD等,这些数据集不仅用于边缘检测任务,还用于边界检测轮廓检测等任务.BIPED数据集的提出,不仅弥补了边缘检测数据集的空白,还可以作为用于评估边缘检测算法的基准.因此本文主要采用此数据集对模型进行评估.此外,作者还提供了相应的图像增强功能,包括图像分割、旋转、剪切、反转以及相应的伽马校正等.BIPED有两个版本,本文采用的是拥有更多注释的BIPEDv2.

MDBD:该数据集有100张1 280像素×720像素高清图像.其中80张用于训练,20张用于评估模型.MDBD数据集每张图片都有多个标注,不仅有边界注释,还有边缘注释(5次边界注释,6次边缘注释),但尚未进行验证,因此有些边缘对应的不是正确的注释.在BIPED数据集提出之前,MDBD被普遍认为是与边缘检测最相关的数据集.为了公平起见,MDBD也使用了与BIPED数据集相同的数据增强.

BRIND:该数据集对BSDS500数据集中边缘部分重新标注,并将所有类型的边缘的注释混合在一起,并且进行了评估验证,最后仅保留了一个注释.在BRIND数据集中,300张用于训练,其余的200张用于测试.BRIND同样也使用了与BIPED数据集相同的数据增强.

2.2 训练细节

大多数的模型都依赖于ImageNet数据集的预训练权重,而本文提出的网络不依赖预训练权重,从头开始训练.训练使用Adam优化器,初始学习率设置为1e-4,权重衰退设置为1e-8,批量大小设置为24.对于BIPED、MDBD、BRIND这三个数据集,网络输入的彩色图像大小分别为352像素×352像素、480像素×480像素、320像素×320像素,网络收敛的轮数分别为9、4、10.BIPED和BRIND数据集的真值在所有大于0.2的值加0.6后被剪入0和1.MDBD数据集的真值在所有大于0.3的值加0.7然后被剪入0和1.训练时CATS损失函数的各参数如表1所示,训练使用的系统为Ubuntu 16.04.6 LTS,CPU为Intel(R) Core(TM) i9-9900KF CPU@3.6 GHz,显卡为TITAN RTX,我们的网络在BIPED和BRIND数据集上,平均在第9轮开始收敛,MDBD在第5轮开始收敛.

表1 损失函数的参数设置

Tab. 1 Parameter setups of loss function

数据集	λ_1/λ_2		
	L_1-L_3	L_4-L_{11}	Final
BIPED	2/0.05	1/0.1	4/0.01
MDBD	4/0.01	2/0.01	6/0.01
BRIND	2/0.05	2/0.1	6/0.05

2.3 评价指标

评估边缘检测的核心步骤是在指定最大允许距离的公差的基础上,将输出的边缘预测与地面真值进行像素匹配.由于边缘检测得到的边缘预测图像并不是二值矩阵,是经过Sigmoid后取值为0-1的矩阵.为了评估结果,需要将矩阵二值化,故需要设置阈值η.用于评估的指标有3个,分别为ODS(数据集最优阈值)、OIS(单场图片最优阈值)、AP(平均精度).

在对给定的边缘预测进行二值化之前,大部分边缘检测器为了获得细化的原边预测图,都会使用非极大值抑制(NMS)操作和数学形态学操作的后处理方案.虽然后处理去除了部分像素周围的混沌像素,但边缘检测器的原始输出是否具有良好的性能是一个更加值得探索的问题,因此我们去掉了后处理方案,直接在网络的原始输出与真值之间进行评估.

2.4 实验结果对比

首先将我们提出的方法与几种先进的方法在BIPED数据集上进行了比较.图2显示了这些检测方法的效果,其中,CATS和DexiNed的预测图一个看起来线条更加的精细,尤其是对于小目标的边缘提取比较精细,另一个线条较粗但是背景的模糊抑制效果较好,BDCN和PiDiNet网络的预测图都有不同程度上的漏检问题.从图中可以看出我们的方法与

其他方法比较,在做到了准确定位边缘的同时,得到了更加精细的边缘效果.如表2所示,我们的网络ODS分数达到了0.891,相较于其他方法有1%~5%的提升.

表2　不同方法在BIPED数据集上的评价结果

Tab. 2　Evaluation results of different methods on BIPED data set

方法	ODS ↑	OIS ↑	AP ↑
RCF[26-27]	0.850	0.860	0.730
BDCN[31]	0.878	0.886	0.827
CATS[32]	0.883	0.889	0.898
DexiNed[28]	0.872	0.877	0.798
PiDiNet[33]	0.854	0.860	0.786
本文算法	0.891	0.895	0.900

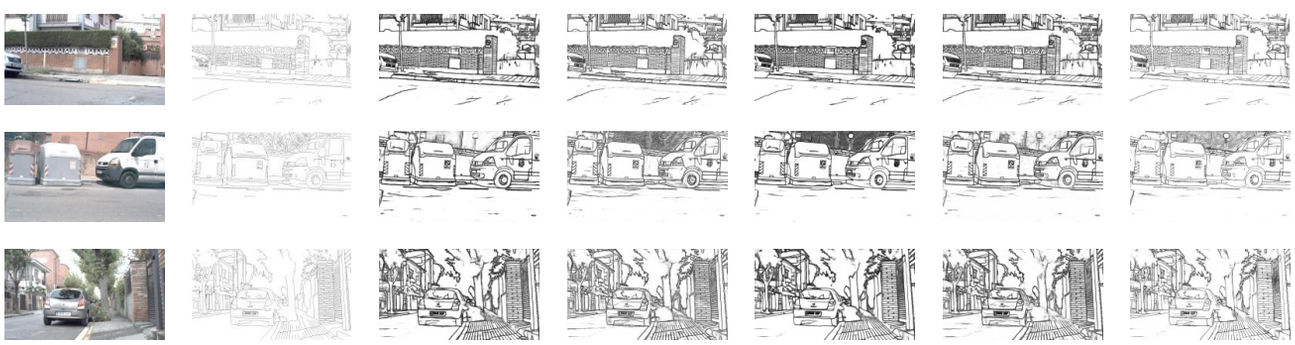
对于MDBD数据集,该数据集可以用作边缘检测,也可以用作边界检测.本文提出的方法主要是用作边缘检测,所以采用数据集中边缘注释部分进行实验.表3所示为我们的方法与其他先进的方法在MDBD数据集上的测试结果,可以看出三个指标中,我们的模型都取得了最优的结果,对于AP,我们的方法相对于其他的方法有较大的提升,说明我们的方法有更少的错误预测,边缘定位更加准确.从图3也可以看出,我们的方法输出的边缘预测图不仅背景的纹理模糊更少,而且线条更加清晰精细.

表3　不同方法在MDBD数据集上的评价结果

Tab. 3　Evaluation results of different methods on MDBD data set

方法	ODS ↑	OIS ↑	AP ↑
BDCN[31]	0.821	0.823	0.720
CATS[32]	0.854	0.856	0.680
DexiNed[28]	0.802	0.807	0.798
PiDiNet[33]	0.821	0.823	0.720
本文算法	0.864	0.871	0.876

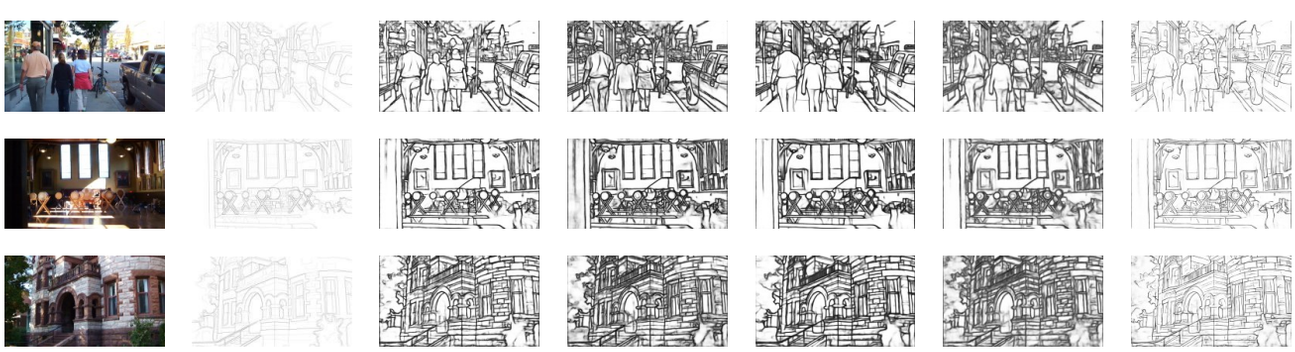
最后对于BRIND数据集,由于BRIND数据集来源于BSDS500数据集,它并不是一个专门进行边缘检测的数据集,即便进行了部分重新标注,但许多细节部分还是被丢弃了.我们的方法是专门用于边缘检测的,在训练时会尽可能多地捕捉那些边缘纹理细节,因此在用BRIND数据集进行测试时,会出现部分边缘预测和真值标签不符的现象.从表4可以看



(a)原图　(b)GT　(c)BDCN　(d)CATS　(e)DexiNed　(f)PiDiNet　(g)本文算法

图2　BIPED数据集上,各算法与本文算法进行定性比较

Fig. 2　Qualitative comparison of other algorithms with our algorithm on BIPED data set



(a)原图　(b)GT　(c)BDCN　(d)CATS　(e)DexiNed　(f)PiDiNet　(g)本文算法

图3　MDBD数据集上,各算法与本文算法进行定性比较

Fig. 3　Qualitative comparison of other algorithms with our algorithm on MDBD data set

出,在OIS和ODS两个指标上,我们的模型取得了次优的结果,在AP上取得了最优的结果.从图4中可以看出,我们的模型所预测的边缘图像在细节方面的处理更加精细,视觉效果上更好,总体上也更符合我们的审美规范.

表4 不同方法在BRIND数据集上的评价结果

Tab.4 Evaluation results of different methods on the BRIND data set

方法	ODS ↑	OIS ↑	AP ↑
BDCN[31]	0.778	0.790	0.411
CATS[32]	0.756	0.771	0.602
DexiNed[28]	0.747	0.759	0.663
PiDiNet[33]	0.748	0.762	0.644
本文算法	0.762	0.774	0.698

2.5 消融实验

为了探索网络中不同模块在边缘检测结果的有效性和在测试前采用形态学非极大值抑制(NMS)对最终测试结果的影响,我们使用BIPED数据集进行试验.

第一个实验探索不同模块的有效性.我们使用DexiNed算法的实验结果作为我们的基线,然后通过以下模型进行消融实验:1)为了探索CATS损失函数对预测的影响,模型A表示我们的模型使用加权交叉熵损失函数;2)为了探索所提的并置多尺度融合方法对预测的影响,模型B表示原DexiNed使用CATS损失函数但不使用并置多尺度融合方法;3)为了探索CDCM和CSAM对模型的影响,模型C表示不使用CDCM和CSAM的后处理,模型D表示在向上传播后的每个侧输出也都添加CDCM和CSAM.结果如表5所示.

从结果对比可以看到,CATS损失函数可以较大幅度提高ODS和OIS指标的结果.并置多尺度融合方法可以进一步提升性能,在AP指标上有较大提升,说明该方法可以使网络预测的边缘定位更加精准.CDCM和CSAM的添加可以让网络学习到更加丰富的多尺度信息,提升网络的性能.但是在向上传播阶段,添加过多的空洞卷积和空间注意力会适得其反.综上所述,网络中设计的每个模块都在最终的性能中起着至关重要的作用.

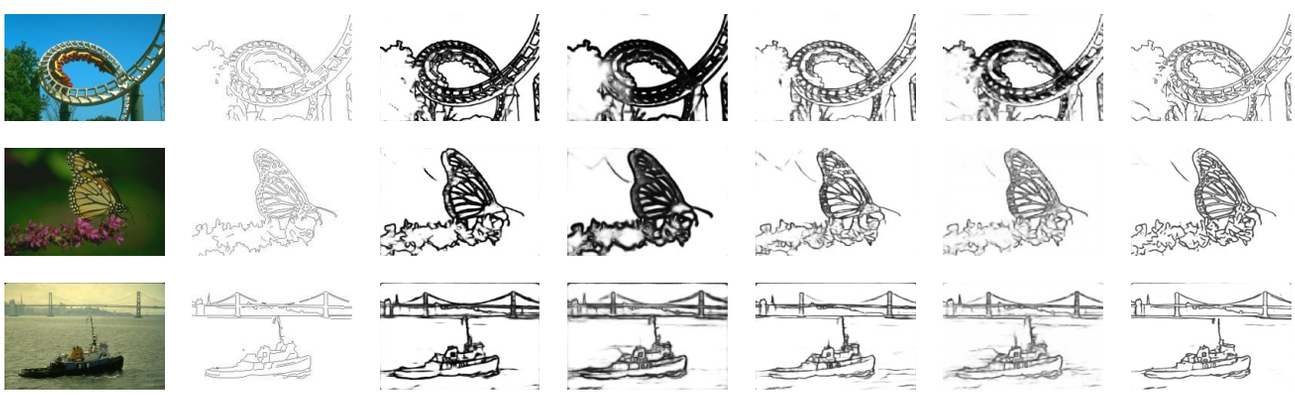
表5 消融实验的评价结果

Tab.5 Evaluation results of ablation study

方法	ODS ↑	OIS ↑	AP ↑
DexiNed[28]	0.872	0.877	0.798
模型A	0.881	0.886	0.827
模型B	0.887	0.892	0.886
模型C	0.890	0.895	0.887
模型D	0.887	0.893	0.906
所提模型	0.891	0.895	0.900

第二个实验试探索大多数边缘检测器使用的形态学非极大值抑制的后处理方法对最终性能的影响.我们使用目前在BIPED数据集上的最优模型DexiNed来进行对比.从图5中可以明显看出,使用NMS后的预测图要比原图更加精细.

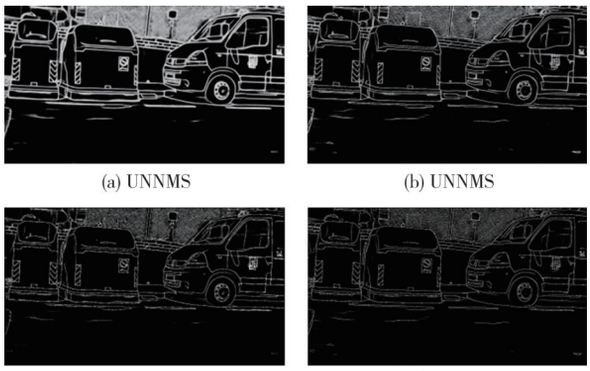
从表6可以看出,由于我们输出的边缘图已经很细了,NMS对我们的输出在性能上的增益并不高,不过即便是使用了NMS,我们的模型的性能依旧是超过了目前BIPED数据集上的最优模型,ODS、OIS、AP分别达到了0.894、0.899、0.930,与其他先进方法相比达到了最优.



(a)原图 (b)GT (c)BDCN (d)CATS (e)DexiNed (f)PiDiNet (g)本文算法

图4 BRIND数据集上,各算法与本文算法进行定性比较

Fig.4 Qualitative comparison of other algorithms with our algorithm on BRIND data set



(a) UNNMS　(b) UNNMS

(c) NMS　(d) NMS

(Ⅰ)DexiNed　(Ⅱ)本文算法

图5　是否使用非极大值抑制的定性比较

Fig.5　Qualitative comparison of whether using non-maximum suppression(NMS)

表6　是否使用非极大值抑制的评价结果

Tab.6　Evaluation results of whether using non-maximum suppression(NMS)

模型	UNNMS			NMS		
	ODS↑	OIS↑	AP↑	ODS↑	OIS↑	AP↑
RCF[26-27]	0.850	0.860	0.730	0.884	0.892	0.930
BDCN[31]	0.878	0.886	0.827	0.888	0.896	0.918
CATS[32]	0.883	0.889	0.898	0.884	0.891	0.927
DexiNed[28]	0.872	0.877	0.798	0.891	0.896	0.919
PiDiNet[33]	0.854	0.860	0.786	0.885	0.891	0.918
本文算法	0.891	0.895	0.900	0.894	0.899	0.931

3　总结

本文提出的基于复杂稠密网络的并置多尺度融合边缘检测算法,用于解决大多数边缘检测器提取边缘过粗的问题,并且可以实现端到端的边缘提取,最后通过实验在BIPED、MDBD和BRIND数据集上证明了所提算法可以兼顾边缘的准确性和精细程度,输出的边缘更加美观,并且拥有较好的泛化性.算法的不足在于对于某些特定颜色的边缘会异常的敏感,致使网络输出错误的预测.未来的算法会尝试加入结构相似度有关的损失函数或者寻找归一化的方法来提高模型对于不同颜色边缘的识别能力,以达到更好的效果.

参考文献

[1] RADENOVIĆ F, TOLIAS G, CHUM O. Deep shape matching [M]//Computer Vision-ECCV 2018. Cham:Springer International Publishing, 2018: 774-791.

[2] SU Y Q, LIU Y H, CUAN B N, et al. Contour guided hierarchical model for shape matching [C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile.IEEE, 2015:1609-1617.

[3] CHEN C, LI S, WANG Y G, et al. Video saliency detection via spatial-temporal fusion and low-rank coherency diffusion[J]. IEEE Transactions on Image Processing, 2017, 26(7): 3156-3170.

[4] WANG X H, LI S, CHEN C, et al. Data-level recombination and lightweight fusion scheme for RGB-D salient object detection[J]. IEEE Transactions on Image Processing, 2021, 30:458-471.

[5] CHENG M, LIU Y, HOU Q, et al. HFS: hierarchical feature selection for efficient image segmentation[C]//Computer Vision-ECCV 2016. Amsterdam, the Netherlands: Springer, 2016: 867-882.

[6] WEI Y C, LIANG X D, CHEN Y P, et al. STC: a simple to complex framework for weakly-supervised semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(11): 2314-2320.

[7] HUNTSBERGER T L, DESCALZI M F. Color edge detection[J]. Pattern Recognition Letters, 1985, 3(3): 205-209.

[8] HUANG J S, TSENG D H. Statistical theory of edge detection [J]. Computer Vision, Graphics, and Image Processing, 1988, 43(3):337-346.

[9] TORRE V, POGGIO T A. On edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, PAMI-8(2):147-163.

[10] ZHANG J Y, CHEN Y, HUANG X X. Edge detection of images based on improved Sobel operator and genetic algorithms [C]// 2009 International Conference on Image Analysis and Signal Processing. Linhai, China. IEEE, 2009:31-35.

[11] CANNY J. A computational approach to edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, PAMI-8(6):679-698.

[12] DOLLAR P, TU Z W, BELONGIE S. Supervised learning of edges and object boundaries [C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). New York, NY, USA. IEEE, 2006:1964-1971.

[13] REN X F. Multi-scale improves boundary detection in natural images [C]//ECCV'08: Proceedings of the 10th European Conference on Computer Vision: Part III. Marseille, France: Springer, 2008: 533-545.

[14] DOLLAR P, ZITNICK C L. Fast edge detection using structured forests[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(8):1558-1570.

[15] ISOLA P, ZORAN D, KRISHNAN D, et al. Crisp boundary detection using pointwise mutual information [C]// Computer Vision-ECCV 2014 . Zurich, Switzerland: Springer, 2014: 799-814.

[16] SHI J B, MALIK J. Normalized cuts and image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888-905.

[17] GANIN Y, LEMPITSKY V. N^4-fields: neural network nearest neighbor fields for image transforms[M]//Computer Vision-ACCV 2014. Cham: Springer International Publishing, 2015: 536-551.

[18] BERTASIUS G, SHI J B, TORRESANI L. DeepEdge: a multi-scale bifurcated deep network for top-down contour detection[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA. IEEE, 2015: 4380-4389.

[19] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA. IEEE, 2015: 3431-3440.

[20] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.

[21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. http://arxiv.org/abs/1409.1556.

[22] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016: 770-778.

[23] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 1800-1807.

[24] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 2261-2269.

[25] XIE S N, TU Z W. Holistically-nested edge detection [J]. International Journal of Computer Vision, 2017, 125(1): 3-18.

[26] LIU Y, CHENG M M, HU X W, et al. Richer convolutional features for edge detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 5872-5881.

[27] LIU Y, CHENG M M, HU X W, et al. Richer convolutional features for edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(8): 1939-1946.

[28] SORIA X, RIBA E, SAPPA A. Dense extreme inception network: towards a robust CNN model for edge detection[C]//2020 IEEE Winter Conference on Applications of Computer Vision (WACV). Snowmass, CO, USA. IEEE, 2020: 1912-1921.

[29] MÉLY D A, KIM J, MCGILL M, et al. A systematic comparison between visual cues for boundary detection[J]. Vision Research, 2016, 120: 93-107.

[30] WANG Y P, ZHAO X, HUANG K Q. Deep crisp boundaries[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 1724-1732.

[31] HE J Z, ZHANG S L, YANG M, et al. BDCN: Bi-directional cascade network for perceptual edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(1): 100-113.

[32] HUAN L X, XUE N, ZHENG X W, et al. Unmixing convolutional features for crisp edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(10): 6602-6609.

[33] SU Z, LIU W Z, YU Z T, et al. Pixel difference networks for efficient edge detection [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada. IEEE, 2021: 5097-5107.

[34] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[EB/OL]. http://arxiv.org/abs/1511.07122.

[35] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//European Conference on Computer Vision. Cham: Springer, 2018: 833-851.