

基于RCF的跨层融合特征的边缘检测

宋杰, 于裕*, 骆起峰

(安徽大学 计算机科学与技术学院, 合肥 230039)

(* 通信作者电子邮箱 yu_ah@163.com)

摘要:针对当前基于深度学习的边缘检测技术产生的边缘线条杂乱且模糊等问题,提出了一种基于RCF的端到端的跨层融合多尺度特征的边缘检测(CFF)模型。该模型使用RCF作为基线,在主干网络中加入CBAM,采用具有平移不变性的下采样技术,并且去除了主干网络中的部分下采样操作,以保留图像的细节信息,同时使用扩张卷积技术增大模型感受野。此外,采用跨层融合特征图的方式,使得高低层特征能够充分融合。为了平衡各阶段损失和融合损失之间的关系,以及避免出现多尺度特征融合之后低层细节过度丢失的现象,对每个损失添加了一个权重。在伯克利分割数据集(BSDS500)和PASCAL VOL Context数据集上进行了训练,在测试时使用图像金字塔技术提高边缘图像的质量。实验结果表明,CFF模型提取的轮廓比基线网络更加清晰,能够解决边缘模糊问题。在BSDS500基准上进行的评估表明,该模型将最佳数据集规模(ODS)和最佳图像比例(OIS)指标分别提高到0.818和0.839。

关键词:深度学习;边缘检测;注意力机制;平移不变性;跨层融合

中图分类号:TP391.41 **文献标志码:**A

Cross-layer fusion feature based on richer convolutional features for edge detection

SONG Jie, YU Yu*, LUO Qifeng

(School of Computer Science and Technology, Anhui University, Hefei Anhui 230039, China)

Abstract: Aiming at the problems such as chaotic and fuzzy edge lines caused by current deep learning based edge detection technology, an end-to-end Cross-layer Fusion Feature for edge detection (CFF) model based on RCF (Richer Convolutional Features) was proposed. In this model, RCF was used as a baseline, the CBAM (Convolutional Block Attention Module) was added to the backbone network, translation-invariant downsampling technology was adopted, and some downsampling operations in the backbone network were removed in order to preserve the image details information, dilated convolution technique was used to increase the model receptive field at the same time. In addition, the method of cross-layer fusion of feature maps was adopted to enable high-level and low-level features to be fully fused together. In order to balance the relationship between the loss in each stage and the fusion loss, and to avoid the phenomenon of excessive loss of low-level details after multi-scale feature fusion, the weight parameters were added to the losses. The model was trained on Berkeley Segmentation Data Set (BSDS500) and PASCAL VOL Context dataset, and the image pyramid technology was used in testing to improve the quality of edge images. Experimental results show that the contour extracted by CFF model is clearer than that extracted by the baseline network and can solve the edge blurring problem. The evaluation performed on the BSDS500 benchmark shows that, the Optimal Dataset Scale (ODS) and the Optimal Image Scale (OIS) are improved to 0.818 and 0.839 respectively by this model.

Key words: deep learning; edge detection; attention mechanism; translation invariance; cross-layer fusion

0 引言

边缘检测是图像处理和计算机视觉中的一个基本问题,目标是从自然图像中提取对象边界和感知上突出的边缘,从而保留图像的要点并忽略不重要的细节。它通常被认为是一种低层次的技术,各种高层次的任务^[1-2]从边缘检测的发展中受益匪浅,如对象检测^[3]、对象建议^[4]和图像分割^[5]。

早期的边缘检测利用图像的一阶和二阶梯度信息的原理来检测边缘。如:Sobel^[6]算子、Canny^[7]算子。这些算法尽管实时性较好,但抗干扰性差,不能有效克服噪声影响,定位方面欠佳。早期的边缘检测器是人工设计的,用于发现强度和

颜色上的不连续性。Martin等^[8]发现将亮度、颜色和纹理的变化转化为特征,并训练分类器将这些特征信息组合起来,可以显著提高性能。最近的工作探索基于学习的边缘检测方法。Dollár等^[9]使用随机决策森林来表示局部图像斑块中的结构,输入颜色和梯度特征,结构化森林输出高质量的边缘信息。然而,上述方法都是基于手工设计,成本高,设计繁琐,实用性不强,且对于语义上有意义的边缘检测来说,手工特征表示高级信息的能力是有限的。

此外,近年来有一种使用卷积神经网络(Convolutional Neural Network, CNN)的发展浪潮,强调自动分层特征学习的

收稿日期:2019-12-04;修回日期:2020-01-14;录用日期:2020-01-15。 基金项目:国家自然科学基金资助项目(61974001)。

作者简介:宋杰(1966—),男,安徽合肥人,副教授,博士,CCF会员,主要研究方向:智能计算、计算机体系结构; 于裕(1995—),男,安徽阜阳人,硕士研究生,主要研究方向:计算机视觉、数据挖掘; 骆起峰(1995—),男,安徽芜湖人,硕士研究生,主要研究方向:计算机视觉、智能计算。

重要性,极大地提高了边缘检测的性能。Ganin等^[10]提出将CNN与近邻搜索结合起来,用CNN算出图像中每个斑块的特征,然后在字典里面进行检索,查找与其相似的边缘,最后把这些相似的边缘信息集成起来,输出最终的结果。Xie等^[11]提出了第一个端到端的边缘检测模型HED(Holistically-Nested Edge Detection),该网络基于全卷积网络(Fully Convolutional Network, FCN)^[12]架构,采用多尺度和多层次的特征学习方式,在VGG16(Visual Geometry Group)^[13]网络的基础上显著提高了边缘检测的成绩。Liu等^[14]在HED模型基础上提出RCF(Richer Convolutional Features for edge detection),使用了更丰富的卷积特征,并提出了一个更鲁棒的损失函数,提升了检测性能。

在边缘检测领域,像HED和RCF这类经典的模型,虽然在边缘处理领域取得了不小的进展,但都是基于VGG16传统卷积(Convolution, conv)结构,不能提取图像的全局特征,且采用了过多的下采样,影响测试集的泛化能力。另外,在网络的较低阶段产生的特征图往往太过于杂乱,包含过多无关重点的细节纹理,虽然网络在最后做了一层融合,用于融合多尺度特征,但是简单地将5个stage使用 1×1 的卷积层融合,会丢失一些多尺度信息。为了解决这些问题,本文提出了一种跨层融合特征的边缘检测(Cross-layer Fusion Feature for edge detection, CFF)模型。该模型在VGG主干网络中引入了CBAM(Convolutional Block Attention Module)注意力机制^[15],并采用Zhang^[16]提出的让卷积网络具有移位不变性的方法,增强网络的特征提取能力。另外在主干网络的第5个stage中使用扩张卷积技术(dilation, dil)^[17]增加网络的感受野,以提取更多的语义信息。针对每个阶段产生的特征信息,使用了一种类似特征金字塔^[18]的方式充分融合多尺度特征,使得低层也能够关注高层的全局特征,并对每个阶段进行监督学习,最后再对这些多尺度特征使用 1×1 的卷积层融合,生成最终的边缘图。在伯克利分割数据集(Berkeley Segmentation Data Set, BSDS500)^[19]和PASCAL VOC Context^[20]数据集上训练。在测试时,使用图像金字塔技术进一步提高边缘的质量,实验结果表明,CFF模型能够有效地解决边缘图像线条模糊的问题。

1 跨层融合的检测模型

1.1 基线网络

RCF^[14]模型在HED^[11]的基础上,优化了网络的结构。RCF网络使用了更丰富的卷积特征,分别每个stage进行监督学习,提高了模型的收敛速度。最后对各层的多尺度特征使用 1×1 卷积层融合,并监督学习。RCF相比HED网络,利用了VGG16所有卷积层提取的特征信息,并且提出了一个更鲁棒的损失函数,使得检测结果有了大幅度的提升。

1.2 提出的模型

本文基于RCF网络,提出了一种跨层级融合多尺度特征的CFF模型,结构如图1所示。在下面小节中,将介绍各个模块的改进。

1.2.1 主干网络

如图1所示,使用VGG16作为CFF模型的主干网络,并去掉了全连接层,采用全卷积网络框架。在主干网络中引入了CBAM注意力机制^[15],并将下采样技术^[16]应用到主干网络的

前3个stage。在第4个stage之后使用 2×2 的最大池化(Maxpool),步长设为1,使得第5个stage的特征图分辨率保持不变,保留图像的细节信息。因此网络只有3次下采样操作,分辨率降低 $1/8$ 。另外,为了解决去除下采样操作后感受野被限制的问题,在第5个stage的卷积操作中使用了扩张卷积技术^[17],扩张参数设为2,保证网络参数不变的情况下增加模型的感受野。

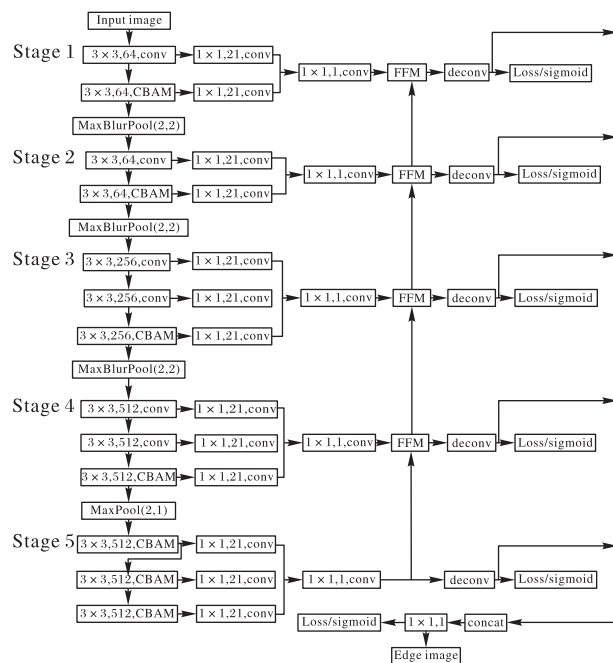


图1 CFF网络结构

Fig. 1 Structure of CFF network

1.2.2 基于CBAM注意力的主干网络

CFF模型在主干网络中引入了CBAM^[15]注意力模块。另一个相关工作是SENet(Squeeze-and-Excitation Networks)^[21],网络根据loss去学习特征权重,使得有效的特征图权重高,无效或效果小的特征图权重小的方式训练模型,以达到更好的结果。然而SENet的不足之处只考虑到不同通道的像素的重要性,忽略了不同位置的重要性。而CBAM模型比SENet多了空间注意力机制,这个空间注意力模块能够学习到每个特征图不同位置的重要性。CBAM模块结构如图2所示。

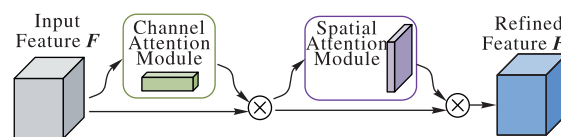


图2 CBAM模块结构

Fig. 2 Module structure of CBAM

CBAM模块可以分为通道注意力模块(图3)和空间注意力模块(图4)。细致来说,在通道注意力模块中,使用平均池化和最大池化对输入的特征图 F 在空间维度上进行压缩,得到两个不同的空间描述符 F_{avg}^c 和 F_{max}^c ,分别表示平均池化特征和最大池化特征,然后将两个描述符发送到多层感知器(Multi-Layer Perceptron, MLP)中,将MLP输出的特征进行基于element-wise的加和操作,再经过sigmoid激活操作,生成以生成通道关注图 M_c 。最后将 M_c 与 F 相乘得到具有通道关注的特征图 F' 。

在空间注意力模块中Stage1,首先对 F' 沿着通道维度应用平均池化和最大池化操作,得到 F'_{avg} 和 F'_{max} ,然后将这两个特征连接起来,并通过标准卷积层进行卷积,再经过sigmoid激活生成空间关注图 M_s 。最后将空间关注图 M_s 与 F' 相乘得到具有注意力机制的特征图 F'' 。

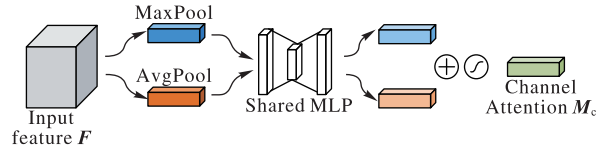


图3 通道注意力模块

Fig. 3 Channel attention module

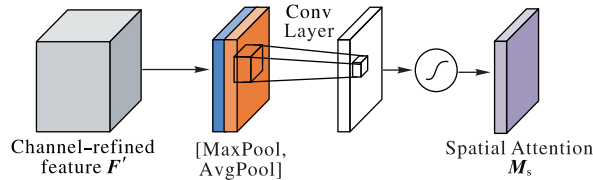


图4 空间注意力模块

Fig. 4 Spatial attention module

考虑到网络参数的问题,只在网络的部分卷积层添加了CBAM注意力机制,如图1中所示。在加入CBAM模块之后,网络能够学习到不同特征图和不同位置像素点的重要性,从

而增强模型的特征提取能力。

1.2.3 最大池化

现代卷积网络并不是移位不变的,是因为常用的下采样方法,如Maxpool,strided-Conv和Avgpool忽略了采样定理,所以小的输入移位或平移会导致输出发生剧烈变化。为解决这一问题,Zhang^[16]提出了一种模糊采样方法(BlurPool)。

如图5所示,最大池化的第一步是先计算区域的最大值,然后进行下采样。而BlurPool则将抗锯齿操作嵌入到中间,通过引入模糊核来平滑输入信号,使得平移后的结果与未平移的结果相近。

以一维信号为例,如图6,当输入信号是[0,0,1,1,0,0,1,1,1],传统最大池化会得到[0,1,0,1],若将输入信号移位1步,最大池化会得到[1,1,1,1],两个结果变化较大,锯齿化明显。而MaxBlurPool操作先将输入信号做步长为1的求最大值操作得到[0,1,1,1,0,1,1,1],再对两端补位得到[1,0,1,1,1,0,1,1,0],然后引入模糊核kernel=[1,2,1],将模糊核与信号做点积运算,并除以模糊核的和值,得到[0.5,0.75,1,0.75,0.5,0.75,1,0.75],再做下采样后得到[0.5,1,0.5,1]。同样,若将输入信号移位1步之后,最后将会得到[0.75,0.75,0.75,0.75],结果相对来说比传统最大池化更平滑。需要注意的是,文献[16]提供了多个模糊核,且使用模糊核时需要手动选择。

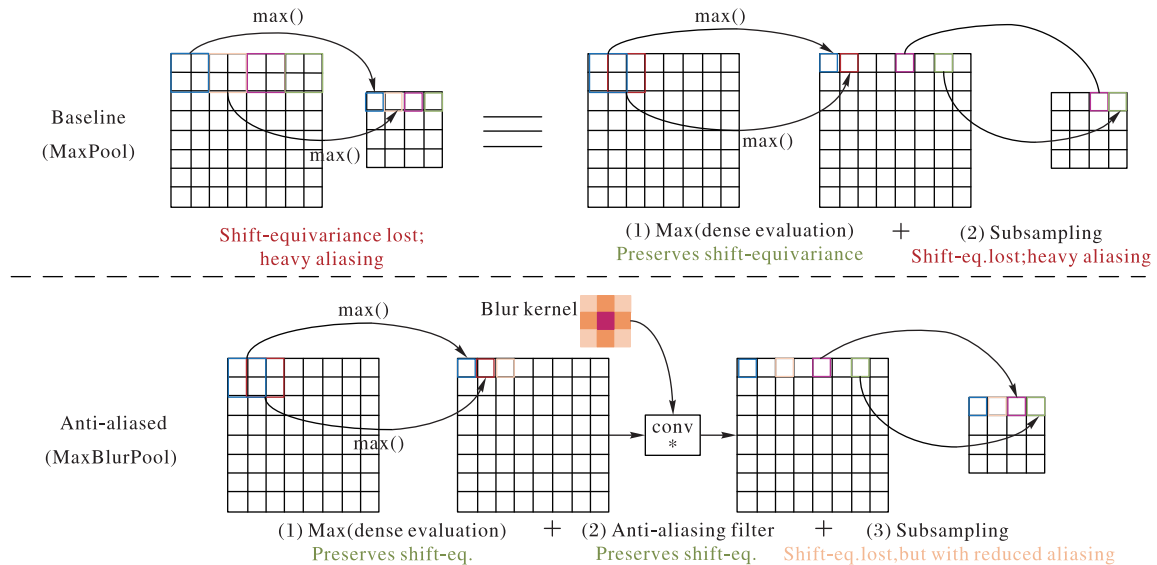


图5 抗锯齿的最大池化

Fig. 5 Anti-aliased max-pooling

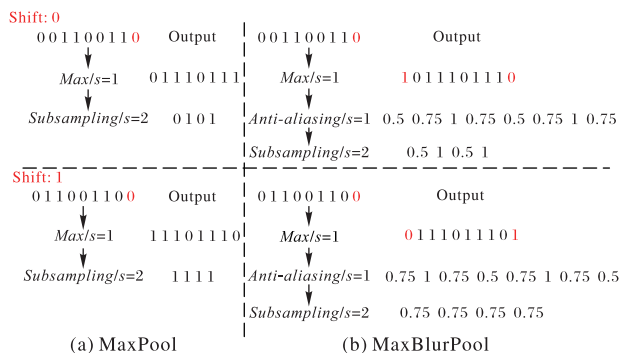


图6 一维信号的抗锯齿操作

Fig. 6 Anti-aliasing operations for one-dimensional signals

CFF模型在网络的前3个阶段后使用该下采样技术^[16],增强了模型的鲁棒性和泛化能力。

1.2.4 多尺度特征提取

在深度学习中,通常存在两种方式让网络学习到多尺度特征:第一种方法是在神经网络的内部,在网络中以越来越大的感受野和下采样层的方式,这样每层学到的特征自然是多尺度的;第二种方法是通过调整输入图像的尺寸。

CFF采用和RCF网络相同的方式提取图像的多尺度特征。将主干的网络中的每一层侧输出通过 1×1 的卷积层进行特征压缩,并以stage为单位将所有侧输出相加。随后经过 1×1 的卷积层降维,输出一张单通道的特征图。

1.2.5 特征金字塔融合模块

为了充分融合各层级的多尺度特征,CFF模型采用特征金字塔的方式,通过将高层特征传递给低层,使得低层也能够关注全局特征。在充分融合多尺度特征的同时有效解决了低阶段生成的特征图中细节杂乱模糊的问题。

如图1中的特征融合模块(Feature Fusion Module, FFM),首先对高层的特征做上采样,然后与低层的特征连接,随后通过一层 1×1 的卷积进行特征压缩。通过这种方式使得低层能够很好地融合高层特征。另外,为避免模型过度忽略低层特征包含的重要细节,借鉴了残差网络^[22]的结构,在本模块中将原始特征图与输出特征图相加,并将结果作为FFM的最终输出。在第4个stage中的FFM结构如图7所示。其他层次中的结构类似。

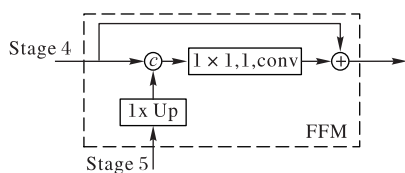


图7 特征融合模块

Fig. 7 Feature fusion module

对于特征融合模块的输出:一方面作为前一阶段的融合模块的输入;另一方面使用反卷积(Deconvolution, deconv)操作实现上采样,使每一个stage都输出一张边缘图,并且模型对每一个stage输出的边缘图都进行监督学习。

1.3 多尺度特征融合

经过特征金字塔模块后,多尺度特征已经充分融合,所以模型只使用一个 1×1 卷积层融合所有层级的多尺度特征,作为CFF模型最后的输出边缘图像,并进行监督学习。

1.4 损失函数

边缘检测数据集通常由多个注释者来标记。对于每个图像,平均所有注释者的标记以生成边缘概率图,范围从0到1。其中:0表示没有注释者在该像素上标记;1表示所有注释者都在此像素处标记。将边缘概率高于 η 的像素视为正样本,将边缘概率等于0的像素视为负样本;否则,如果像素被小于 η 的注释者标记,则该像素在语义上可能是有争议的边缘点,无论将其视为正样本还是负样本,都可能使网络混淆,所以模型忽略了这个类别中的像素。

由于边缘检测是对像素做分类,因此本模型使用交叉熵函数作为目标函数,与文献^[14]相同,计算每个像素相对于像素标签的损失为:

$$l(X_i; \mathbf{W}) = \begin{cases} \alpha \cdot \log(1 - P(X_i; \mathbf{W})), & y_i = 0 \\ 0, & 0 < y_i \leq \eta \\ \beta \cdot \log P(X_i; \mathbf{W}), & y_i > \eta \end{cases} \quad (1)$$

$$\begin{cases} \alpha = \lambda \cdot \frac{|Y^+|}{|Y^+| + |Y^-|} \\ \beta = \frac{|Y^-|}{|Y^+| + |Y^-|} \end{cases} \quad (2)$$

其中: $|Y^+|$ 和 $|Y^-|$ 分别代表正负样本的数量,超参数 λ 用来平衡正负样本数量差, X_i 表示神经网络的激活值, y_i 表示标签图中像素点 i 是边缘点的概率值, $\mathbf{W}$ 表示神经网络中可学习的参数。

1.5 设置不同阶段的损失权重

网络中每个阶段输出的边缘图像之间差异较大,各阶段损失的量级可能不一致,且融合阶段的损失应该占主要地位。另外,在实验中发现,模型在训练至第20个epoch时,前2个阶段的特征图几乎不再包含任何细节纹理,猜想这可能是低层融合高层特征后带来的负面影响。这些问题对最终的预测都是不利的。为了抑制这种现象,本文降低了网络中5个阶段的损失比重,提高了融合阶段损失比重,以平衡各阶段损失和融合损失之间的关系。

将网络的5个stage的损失权重设置为 s^{side} ,最终融合层的损失权重为 s^{fuse} ,因此总损失函数可写为:

$$L(\mathbf{W}) = \sum_{i=1}^{|I|} \left(\sum_{k=1}^K S^k \cdot l(X_i^{(k)}; \mathbf{W}) + s^{\text{fuse}} \cdot l(X_i^{\text{fuse}}; \mathbf{W}) \right) \quad (3)$$

其中: S^k 表示第 k 个stage的损失权重, s^{fuse} 表示融合层的损失权重, $X_i^{(k)}$ 代表第 k 个stage输出图片中的第 i 个像素点的激励值, X_i^{fuse} 代表融合模块输出的图片中的第 i 个像素点的激励值, $|I|$ 代表每张图片像素点的总数, K 代表主干网络中stage的数量。

1.6 多尺度的边缘检测器

为了进一步提高边缘的质量,在测试时采用了图像金字塔技术。具体来说,在测试时调整图像的大小来构建一个图像金字塔,每个图像分别输入到已经训练好的单尺度检测器中;然后,使用双线性插值将所有得到的边缘概率图调整为原始图像的大小;最后,对这些结果进行加权平均,得到最终的预测边缘图。本模型使用了3个不同的尺度,分别为0.5、1.0和1.5。

2 模型训练

2.1 数据集介绍

BSDS500^[19]数据集和PASCAL VOC Context^[20]数据集是边缘检测中广泛使用的数据集。BSDS500数据集由200张训练图像,100张验证图像和200张测试图像组成,每个图像由4到9个注释者标记。为了防止模型出现过拟合现象,对BSDS500的训练集和验证集共计300张图片进行旋转、扩大、剪裁等操作,进行数据集增强。最后将BSDS500的增强数据集与PASCAL VOC Context数据集混合作为训练数据。

2.2 训练细节

CFF网络基于Python3编写,使用pytorch 1.0.1深度学习框架,以及其他一些库。在一台ubuntu服务器上实验,硬件包括E5-2678 v3 2.50 GHz的CPU和1块NVIDIA Tesla K40C的显卡,显存12 GB。

模型通过随机梯度下降算法训练30个epoch,batch size设为1,基准学习率learning rate设为 $1\text{E}-6$,为不同卷积层指定不同的学习率,momentum设置为0.9,weight decay设置为0.0002。在训练时,不使用任何预训练模型,网络参数使用Gaussian分布初始化。

3 实验

给定边缘概率图,需要阈值来产生边缘图像,设置此阈值有两种选择:第一个是最佳数据集规模(Optimal Dataset Scale, ODS),它对数据集中的所有图像采用固定阈值;第二种是最佳图像比例(Optimal Image Scale, OIS),它为每个图像选择最佳阈值。通常使用ODS和OIS作为边缘检测模型的指标。

3.1 实验分析

将非最大抑制技术^[9]应用于模型输出的边缘图像以获得用于评估的细化边缘图像,并用Edge Box^[4]工具包进行评估,图8显示了评估结果。RCF网络的边缘检测相对于人类已经取得了更好的结果,而CFF模型优化了RCF网络的短板,其多尺度策略将ODS评分提高到了0.818。

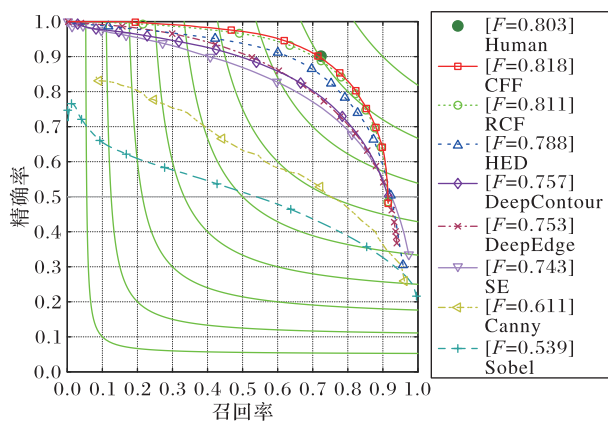


图8 BSDS500评估结果

Fig. 8 Evaluation results on BSDS500

将CFF与其他相关算法进行了比较,结果如表1所示。从表中的各项指标可以看出,CFF模型的ODS和OIS分别比RCF模型提高了0.7个百分点和0.9个百分点。

表1 CFF模型和其他算法的比较

Tab. 1 Comparison between CFF model and other algorithms

算法	ODS	OIS
Canny ^[7]	0.611	0.676
gPb-ucm ^[23]	0.726	0.757
SE ^[9]	0.743	0.763
N4-Fields ^[10]	0.753	0.769
DeepEdge ^[24]	0.753	0.772
DeepContour ^[25]	0.756	0.773
HED ^[11]	0.788	0.804
RCF ^[14]	0.811	0.830
CFF	0.818	0.839

CFF与RCF网络输出的边缘图像的对比结果如图9所示。通过对比可以看出,RCF模型产生边缘图像中有一些线条比较模糊,而CFF模型能够清晰地将图像中的边缘检测出来,且对一些细节模糊的问题处理得较好。

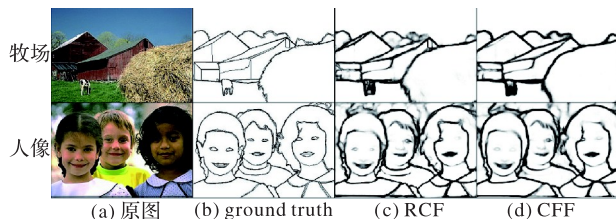
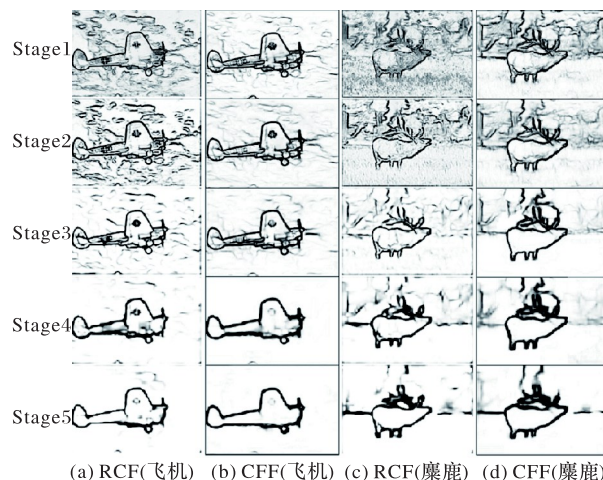


图9 CFF与RCF结果对比

Fig. 9 Comparison of the results of CFF and RCF

为进一步展示CFF模型的优化细节,在图10中给出了CFF模型在各个阶段输出的边缘图像与RCF网络的对比。在图中每一列从上到下分别为阶段1~5生成的边缘图像,可以看出RCF网络的各个阶段对一些无关的细节处理能力较差,每个阶段中都包含了一些模糊的线条。而CFF模型通过跨层

融合不同层次的特征,使得模型在低层能够关注一些全局的轮廓信息,帮助多尺度特征充分融合。从图中可以看出,CFF输出的边缘图像相较于RCF只包含了很少的无细节,尤其是在第1,2阶段,没有过多的杂乱纹理。



(a) RCF(飞机) (b) CFF(飞机) (c) RCF(麋鹿) (d) CFF(麋鹿)

图10 CFF与RCF各阶段输出图像的对比

Fig. 10 Comparison of output images of CFF and RCF at each stage

3.2 网络内部结构对比

在本节中,对CFF模型的内部结构进行性能分析。如表2所示,在主干网络中引入CBAM注意力模块和抗锯齿的下采样技术后,模型的ODS分数提高了0.4%,OIS提高了0.5%,证明了本模型的主干网络提取的特征信息更加丰富有效。另外,通过跨层融合不同阶段的输出特征,ODS和OIS都进一步提高了0.2%,说明通过将高层特征传递给低层后,能够充分融合多尺度特征。为了平衡各阶段损失之间的关系,抑制低层细节丢失的问题,通过设置不同阶段的损失权重后,模型的ODS和OIS分别提高了0.1%和0.2%。

表2 各模块改进效果的比较

Tab. 2 Comparison of improvement effect of different modules

Rank	CBAM+			ODS	OIS
	Dilation+	FFM	Weight loss		
	MaxBlurPool				
1				0.811	0.830
2	√			0.815	0.835
3	√	√		0.817	0.837
4	√	√	√	0.818	0.839

4 结语

本文基于RCF网络,提出了一种关注全局的边缘检测网络。CFF模型在VGG16主干网络中加入CBAM模块,并采用具有平移不变性的下采样技术,提高了网络的特征提取能力。去除部分下采样层,防止图像分辨率过低,影响模型精度。并在第5个stage采用空洞卷积技术提升网络的感受野。另外,模型采用了一种由深到浅的特征融合方式,使得网络能够更加注重全局信息,并在计算损失时,给不同阶段设置不同的损失权重以平衡各阶段损失,防止模型过度忽略低层的细节。实验表明,CFF模型能够生成更清晰的边缘图像。

参考文献(References)

- [1] CHENG M, HOU Q, ZHANG S, et al. Intelligent visual media processing: when graphics meets vision[J]. Journal of Computer Sci-

- ence and Technology, 2017, 32(1):110-121.
- [2] HU S, CHEN T, XU K, et al. Internet visual media processing: a survey with graphics and vision applications[J]. The Visual Computer, 2013, 29(5):393-405.
 - [3] FERRARI V, FEVRIER L, JURIE F, et al. Groups of adjacent contour segments for object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(1): 36-51.
 - [4] ZITNICK C L, DOLLÁR P. Edge boxes: locating object proposals from edges[C]// Proceedings of the 13th European Conference on Computer Vision, LNCS 8693. Cham; Springer, 2014: 391-405.
 - [5] ARBELÁEZ P, PONT-TUSET J, BARRON J, et al. Multiscale combinatorial grouping[C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 328-335.
 - [6] FELDMAN J A, FELDMAN G M, FALK G, et al. The Stanford hand-eye project[C]// Proceedings of the 1st International Joint Conference on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann Publishers Inc., 1969: 521-526, 526a.
 - [7] CANNY J. A computational approach to edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, PAMI-8(6): 679-698.
 - [8] MARTIN D R, FOWLKES C C, MALIK J. Learning to detect natural image boundaries using local brightness, color, and texture cues [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(5): 530-549.
 - [9] DOLLÁR P, ZITNICK C L. Fast edge detection using structured forests[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(8): 1558-1570.
 - [10] GANIN Y, LEMPITSKY V. N^4 -fields: neural network nearest neighbor fields for image transforms[C]// Proceedings of the 12th Asian Conference on Computer Vision, LNCS 9004. Cham: Springer, 2014: 536-551.
 - [11] XIE S, TU Z. Holistically-nested edge detection[C]// Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 1395-1403.
 - [12] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3431-3440.
 - [13] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2019-03-02]. <https://arxiv.org/pdf/1409.1556.pdf>.
 - [14] LIU Y, CHENG M, HU X, et al. Richer convolutional features for edge detection[C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 5872-5881.
 - [15] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]// Proceedings of the 15th European Conference on Computer Vision, LNCS 11211. Cham: Springer, 2018: 3-19.
 - [16] ZHANG R. Making convolutional networks shift-invariant again [C]// Proceedings of the 36th International Conference on Machine Learning. New York: International Machine Learning Society, 2019: 7324-7334.
 - [17] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [EB/OL]. [2019-03-02]. <https://arxiv.org/pdf/1511.07122.pdf>.
 - [18] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 936-944.
 - [19] MARTIN D, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]// Proceedings of the 8th IEEE International Conference on Computer Vision. Piscataway: IEEE, 2001: 416-423.
 - [20] MOTTAGHI R, CHEN X, LIU X, et al. The role of context for object detection and semantic segmentation in the wild[C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 891-898.
 - [21] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.
 - [22] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.
 - [23] ARBELÁEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.
 - [24] BERTASIOUS G, SHI J, TORRESANI L. DeepEdge: a multi-scale bifurcated deep network for top-down contour detection[C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 4380-4389.
 - [25] SHEN W, WANG X, WANG Y, et al. DeepContour: a deep convolutional feature learned by positive-sharing loss for contour detection[C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3982-3991.
- This work is partially supported by the National Natural Science Foundation of China (61974001).
- SONG Jie**, born in 1966, Ph. D., associate professor. His research interests include intelligent computing, computer architecture.
- YU Yu**, born in 1995, M. S. candidate. His research interests include computer vision, data mining.
- LUO Qifeng**, born in 1995, M. S. candidate. His research interests include computer vision, intelligent computing.