

一种基于编解码结构的多尺度边缘检测方法

田岸霖¹, 雷为民¹, 张鹏^{1,2}, 张伟¹

(1. 东北大学 计算机科学与工程学院, 辽宁 沈阳 110169; 2. 沈阳二一三电子科技有限公司, 辽宁 沈阳 110027)

摘要: 针对传统边缘检测算法易出现边缘线断裂、不连续的情况以及基于深度学习的方法无法保证边缘清晰性和准确性且容易丢失边缘细节信息的问题,提出了一种基于编解码结构的轻量级的边缘检测方法. 通过编码器提取图像的边缘特征,在解码端恢复编码器下采样时丢失的边缘信息,编解码器之间采用跳跃连接方法实现低层特征和高层特征之间的融合,采用具有注意力机制的深监督模块,进一步学习多尺度多层次的边缘特征生成精细的图像边缘. 该网络模型在BSDS500S数据集上进行训练,实验结果表明,本文方法 ODS(optimal dataset scale)与 OIS(optimal image scale)分别达到 0.808 和 0.830,在 GTX 1060 机器上帧率达到 60 帧/s,超过了基于卷积神经网络的主流边缘检测方法,具有较好的效果.

关键词: 边缘检测;编解码;跳跃连接;注意力机制;深监督

中图分类号: TP 391 文献标志码: A 文章编号: 1005-3026(2024)07-0936-08

A Multi-scale Edge Detection Method Based on Encoder-Decoder

TIAN An-lin¹, LEI Wei-min¹, ZHANG Peng^{1,2}, ZHANG Wei¹

(1. School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China; 2. Shenyang 213 Electronic Technology Co., Ltd., Shenyang 110027, China. Corresponding author: LEI Wei-min, E-mail: leiweimin@ise.neu.edu.cn)

Abstract: Traditional edge detection algorithms often suffer from fractured and discontinuous edge lines, while deep learning-based approaches fail to ensure edge clarity and accuracy, often leading to the loss of edge details. To address these issues, a lightweight edge detection method based on encoder-decoder structure is proposed. The encoder extracts the edge features of the image, recovers the edge information lost when sampling under the encoder at the decoding end, uses the jump connection method between the codecs to achieve the fusion between low-level features and high-level features, and uses the deep supervision module with attention mechanism to further learn multi-scale and multi-level edge features to generate fine image edges. The network model is trained on the BSDS500S dataset. Experimental results show that the ODS and OIS of the proposed method reach 0.808 and 0.830 respectively, and the frame rate on the GTX 1060 machine reaches 60 frames/second, which exceeds the mainstream edge detection methods based on convolutional neural networks, thus showing its effectiveness.

Key words: edge detection; encoder-decoder; skip connection; attention mechanism; deep supervision

在视频图像领域中,边缘特征的提取一直是研究的热点问题.一幅图像中包含各种各样的丰富信息,直观感受到的就是边缘,它象征着一个区域的开始与另一个区域的结束.边缘是识别物体主要依赖的特征,现有的研究工作证明,在一

些计算机视觉上游任务中,包括图像分类^[1]、目标检测^[2]、语义分割^[3]等,提取恰当的边缘特征可以帮助获取对象的结构信息,更好地解决边缘模糊和边界丢失问题,提高任务精度.因此,边缘检测具有重要的研究意义.

收稿日期: 2023-03-07

基金项目: 2022年辽宁省“揭榜挂帅”科技重大专项(2022JH1/10400025); 国家重点研发计划项目(2018YFB1702000); 中央高校基本科研业务费专项资金资助项目(N2216010).

作者简介: 田岸霖(1998-),男,山西吕梁人,东北大学硕士研究生; 雷为民(1969-),男,山西平遥人,东北大学教授,博士生导师.

边缘检测的目的是在图像中找到像素亮度发生强烈变化的点的集合.早期的边缘检测算法基于不同的微分算子来获得图像的边缘轮廓,例如Sobel算子^[4]、Prewitt算子^[5]以及广泛使用的Canny算子^[6],这些边缘检测算子使用简单,但仅仅考虑了图像的局部梯度特性,忽略了纹理变化,容易造成边缘识别不准确以及边缘模糊等问题.通过对特征的显示分析,研究者认为边缘不仅依赖梯度变化,还应该包含颜色、纹理等低维特征,因此提出了基于手工特征提取的边缘检测方法gPb(global pixel boundary)算法^[7]、SE(structured edge)算法^[8]等.这些方法都将亮度、颜色、纹理等低维特征结合起来获取边缘,并且取得了比较不错的结果.但手工提取的特征参数量少,用于表示高级语义信息能力有限,无法做到精细的边缘检测.

随着深度学习在计算机视觉任务中取得重大进展,很多基于深度学习的边缘检测算法也相继被提出.Ganin等^[12]最先提出将CNN与最近邻搜索相结合的N4-Fields方法,首先利用卷积神经网络提取图像块的特征,然后使用最近邻搜索找出字典中相匹配的边缘,将这些边缘聚集起来就得到最后的结果.Shen等^[10]提出了Deep Contour边缘检测算法,将轮廓数据分割成紧凑的簇,然后学习不同的模型参数拟合不同形状的轮廓数据.Xie等^[13]提出了一种全新的整体嵌套边缘检测算法HED(holistically-nested edge detection),这种算法基于预训练的VGG16^[14]网络为主干,通过深度学习模型学习图像端到端的特征表示,并且采用深监督策略进行多尺度多层次的特征学习,大幅提高了边缘检测精度,为之后的边缘检测算法提供了全新的思路.Liu等^[15]在HED的基础上提出一种新的网络结构RCF(richer convolutional features for edge detection),充分利用CNN丰富特性的层次结构,将每个卷积层的不同层次信息组合到一起,进而得到更精细的多尺度特征.Deng等^[16]引入亚像素上采样方法,采用细化模块进一步得到精细的边缘.除了上述方法外,BDCN(bi-direction cascade network)^[17],LPCB(learning to predict crisp boundaries)^[18],PiDiNet^[19]等基于深度学习方法的检测结果都可以达到甚至超过人类的识别结果.

上述基于卷积神经网络的边缘检测算法大部分都是使用在大规模ImageNet数据集上预训练的VGG16^[14]网络作为主干,尽可能多使用每一个中间层特征进行端到端检测.通过分析发现上述这些模型大而复杂,具有内存消耗大、计算

成本高、低吞吐量等缺点.其次,这些模型都局限于编码器部分,只负责提取特征,为了获得更大的感受野,在每个阶段的卷积层后面加入下采样操作,丢失了大量的边缘信息,造成边缘模糊,边缘定位不准确,无法有效抑制背景纹理信息.

基于编解码结构的网络(SegNet^[20],Unet^[21]等)极大地提升了语义分割的性能,这种模型既可以在编码端获得图像中的物体结构信息以及位置信息,还可以在解码端对下采样的特征图进行恢复,弥补池化造成的边缘细节损失.因此本文提出了一种基于编解码结构的边缘检测方法,该方法编码器采用逐级递增卷积层方法在不同的感受野下提取不同的分层信息,利用解码器以及跳跃连接实现边缘特征的找回,获得更加完整精细的边缘.最后,加入深监督模块处理小目标物体的边缘以及进行多尺度多层次的特征学习,进一步提高边缘检测的精度.

1 基于编解码的边缘检测模型

传统的边缘检测算法基于手工设计特征,提取的边缘不准确或者边缘线出现断裂的情况,基于卷积神经网络的边缘检测方法局限于编码器部分,经过下采样操作后,特征的分辨率和粒度会丢失,损失了大量的边缘信息,存在线条较厚、边缘模糊问题.本文提出一种基于编解码结构的边缘检测方法,在一般边缘检测网络基础上加入解码部分,从而获取更加完整、精细的边缘.

1.1 总体结构

整体架构由编码器、解码器和深监督三个部分组成,如图1所示.

编码器通过神经网络学习输入图像的特征表示;解码器负责表达图像深层的、抽象的语义信息,将编码器处理后的特征进一步优化;跳跃连接将底层特征与深层特征拼接,减少编码端的边缘信息丢失,提高特征细粒度;最后,加入深监督模块进行多尺度多层次的特征学习.

1.2 编码器模块

在深度学习中,浅层卷积层提取的特征和输入相近,包含更多底层空间位置信息.通过卷积层数的不断叠加,图像分辨率降低,感受野增大,可以提取到更多抽象的语义信息.边缘检测的目的是提取图像中的边缘特征信息,减少信息的损失,因此编码器的设计遵循浅层阶段卷积数量少,提取更多的边缘信息,深层阶段卷积操作多,更加专注于抽象信息的提取的准则.编码器网络如图2所示.

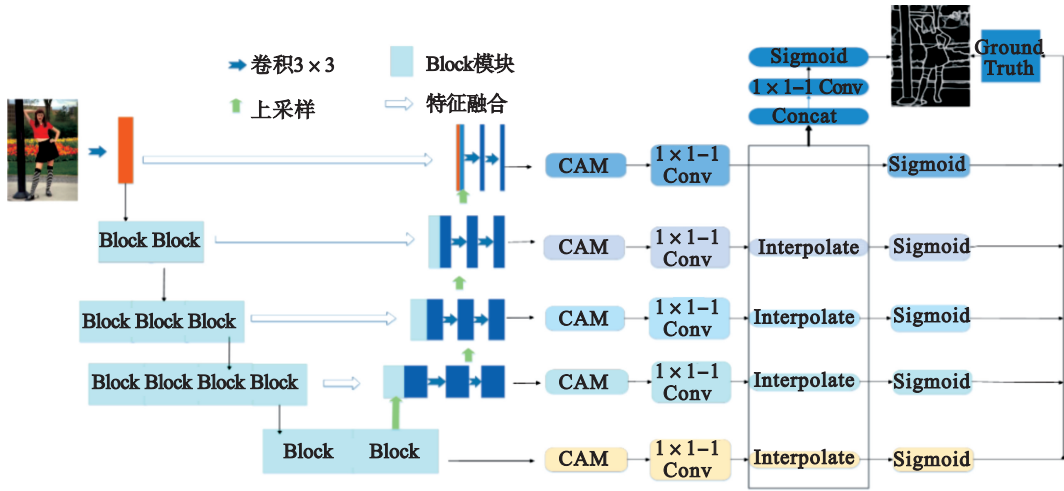


图1 基于编解码结构的边缘检测网络

Fig. 1 Edge detection network based on encoder-decoder

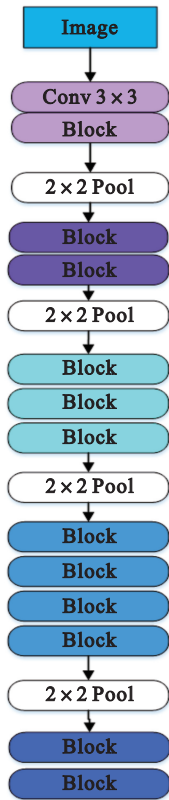


图2 编码器模块

Fig. 2 Encoder module

在第一个阶段,原始图像 I 经过一个普通卷积层和一个 Block 模块,如式(1)所示,当前卷积层数少,可以提取到尽可能多的边缘特征信息。

$$E_i = \text{Block}(\text{Conv}(I)). \quad (1)$$

其中: E_i 表示编码器第 i 阶段的输出特征; $\text{Conv}(\cdot)$ 表示普通 3×3 卷积运算; $\text{Block}(\cdot)$ 表示 Block 卷积运算。

接下来每个阶段采取递增策略逐级增加 Block 块,利用下采样操作使每个阶段获得不同

的感受野,这样在编码端能够提取不同尺度、不同层次的边缘特征.例如,在第1个阶段提取到更多的边缘位置信息,而在第5个阶段更多的是语义级别的高级特征。

受 MobileNetV3^[22] 结构启发,Block 模块采用深度可分离卷积结构,如图3所示,将普通卷积运算分解为 3×3 深度卷积,ReLU层和 1×1 逐点卷积,可以极大地降低参数量,当 stride 为 2 时,加入残差结构,防止梯度消失,提高模型训练的准确率.同时为了避免大模型,每个阶段的通道数量相当小,阶段 1~5 的通道数分别为 $C, 2 \times C, 4 \times C, 8 \times C$ 和 $16 \times C (C=20)$ 。

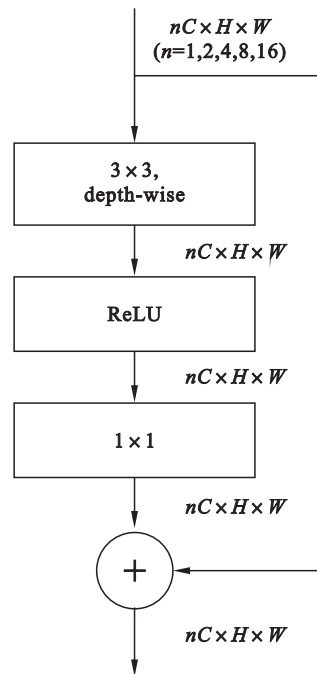


图3 Block 模块

Fig. 3 Block module

1.3 解码器模块

解码器的作用是将编码器主干提取的高级抽象特征逐步上采样到与原始图像相同的分辨率,实现对边缘像素的分类.为了能够充分利用编码端提取的特征表示,同时克服池化造成的重要边界信息丢失.解码端输入由两个部分构成,一个是编码端的输出,另一个是编码器中间阶段的特征输出,将左边编码器的低级特征与上采样后的高分辨率信息进行特征拼接,提供更细粒度的特征,实现对边界信息的找回.在编码端进行了 4 次下采样,获得 5 个不同分辨率的特征表示 $E = \{E_1, E_2, E_3, E_4, E_5\}$,解码端相应也需要 4 次解码器模块,逐步上采样,提升分辨率,解码器模块由一个上采样层、特征拼接和两个 3×3 卷积层构成.其每个阶段的输出特征如式(2)所示:

$$D_i = \text{DoubleConv}(\text{Concat}(E_i, \text{Up}(E_{i+1}))). \quad (2)$$

式中: $\text{Up}(\cdot)$ 表示上采样; $\text{Concat}(\cdot)$ 表示特征拼接; $\text{DoubleConv}(\cdot)$ 表示双卷积运算.

1.4 深监督模块

为了进一步学习丰富的分层边缘特征,解决编解码网络收敛难的问题,在解码器不同阶段后面加入侧输出层,将标签图与各侧边输出特征图计算误差进行反向传播形成深监督,最后将不同尺度的侧输出层特征进行融合,在标签图的监督下输出最优的精确边缘图像,如图 4 所示.

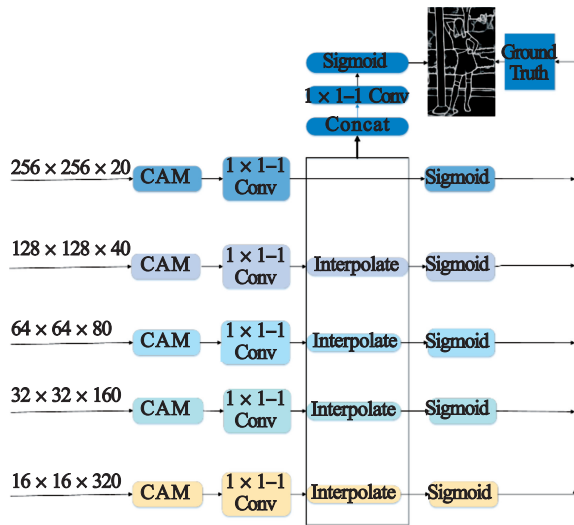


图 4 深监督模块

Fig. 4 Deep supervision module

不同阶段的输出通道数和所携带的信息量不同,所以先将解码端不同阶段的输出经过一个通道注意力模块 CAM(channel attention module),如图 5 所示.注意力模块原理是给输入的不同信

息分配不同的权重,从大量的信息中挑选出与当前任务有关的信息.因此,CAM 根据每个通道的重要程度得到更多有用的边缘信息,抑制无关的背景纹理信息. 1×1 卷积将 CAM 输出特征图减少为单通道的特征图,然后将其插值为与原始图像的大小相同,使用 Sigmoid 激活函数生成 5 个阶段的边缘概率图 $\text{Sides} = \{\text{side1}, \text{side2}, \text{side3}, \text{side4}, \text{side5}\}$.将 5 个阶段的单通道特征图拼接融合,通过 1×1 卷积转换为单通道特征图,最后利用 Sigmoid 函数得到最终边缘 $\text{Side}_{\text{fuse}}$.

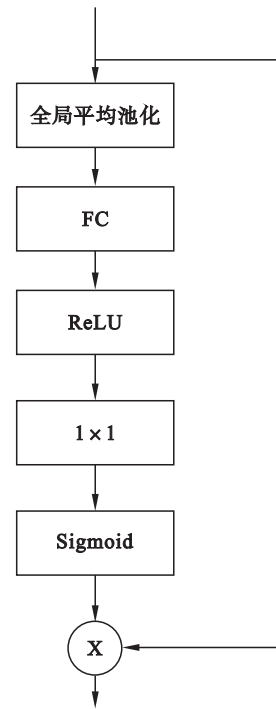


图 5 CAM 模块

Fig. 5 CAM module

1.5 损失函数

边缘检测任务的目的是将像素分为边缘像素和非边缘像素,归根结底是一个二分类任务.但与二分类任务不同的是,边缘检测存在一个类不平衡问题.在一幅图像中,边缘像素只占很少一部分,绝大部分都是非边缘像素,二者比例严重失衡,会导致最终检测出大部分非边缘像素,所以在边缘检测任务当中不能直接使用二进制交叉熵损失函数.另外,在边缘检测任务中还存在一个问题,边缘检测数据集中的标签图是由多个工作人员标记的,每个人的认知不同,例如一个人认为某个像素点是边缘,而其他人认为不是边缘.

为了解决以上两个问题,本文采用了文献[15]中提出的类别平衡损失函数.分别计算数据集中

边缘像素和非边缘像素的总数, α 是边缘像素占总像素的比例, β 是非边缘像素占总像素的比例.

$$\alpha = \lambda \frac{|Y^+|}{|Y^+| + |Y^-|}, \quad (3)$$

$$\beta = \lambda \frac{|Y^-|}{|Y^+| + |Y^-|}. \quad (4)$$

式中: λ 是超参数, 用来平衡正负样本的数量; Y^+ 和 Y^- 分别表示边缘和非边缘像素数量.

当某个像素点被预测为边缘像素, 它的损失值乘以 β , 反之, 预测为非边缘像素的, 损失值乘以 α , 这样就实现了类别平衡. 将不同工作人员标记的标签图取平均值得到一个边缘概率图, 设置一个阈值 η , 如果一个点的概率值大于 η , 则认为该点是边缘; 若这个点的概率为 0, 则认为该点不是边缘; 如果概率值介于 0 到 η 之间, 则这个像素点存在争议, 将其忽略, 不计入损失函数. 综上, 每个像素点的损失函数如下:

$$l(X_i, W) = \begin{cases} \alpha \cdot \lg(1 - P(X_i, W)), & y_i = 0; \\ 0, & 0 < y_i < \eta; \\ \beta \cdot \lg P(X_i, W), & \text{其他.} \end{cases} \quad (5)$$

其中: y_i 表示真实边缘图像像素点是边缘的概率; X_i 是像素点 i 的 CNN 激活值; W 是网络中卷积操作所有的学习参数; $P(X_i, W)$ 是像素点 i 是边缘的概率值.

各个阶段侧输出层生成边缘特征图的损失函数如下:

$$L_{\text{side}} = \sum_{i=1}^{|I|} \left(\sum_{k=1}^5 l(X_i^k, W) \right). \quad (6)$$

其中: X_i^k 是第 k 阶段像素 i 点的 CNN 激活值; $|I|$ 代表图像 I 的像素个数.

融合层生成边缘特征图的损失函数为

$$L_{\text{fuse}} = \sum_{i=1}^{|I|} l(X_i^{\text{fuse}}, W). \quad (7)$$

其中, X_i^{fuse} 是融合层像素点 i 的 CNN 激活值.

总的损失函数是各个阶段侧输出层以及融合层的边缘图每个像素点的损失函数之和:

$$L(W) = L_{\text{side}} + L_{\text{fuse}}. \quad (8)$$

2 实验与分析

2.1 数据集

本实验使用了两个常见的边缘检测数据集, 分别是 BSDS 500 和 NYUD 数据集. 在 BSDS 500 数据集中, 共有 200 张训练集、100 张验证集和 200 张测试集, 其中每个图像都有人工标记的标

签图, 广泛应用于语义分割和边缘检测任务, 但是训练样本过少, 因此研究人员经过旋转、翻转、尺度缩放增强训练数据集, 最终得到了 28 800 张训练图片.

NYUD 是纽约大学创建的各种室内场景组成的数据集, 其中包含 1 499 个密集标记的 RGB 和深度图像. 本实验只使用 RGB 图像作为边缘检测数据集, 其中有 381 张训练图片, 414 张验证集, 654 张测试图片, 并且通过旋转、缩放和翻转进行了数据增强, 扩充了数据集.

2.2 训练参数

本文实验基于 Pytorch 深度学习框架实现, 可以利用 GPU 强大的算力加速网络训练. 在 Linux 环境下进行实验, 操作系统为 Ubuntu 18.04, 显卡为 GeForce RTX 3090.

在训练过程中, 本文使用 Adam 优化器更新神经网络权重, 初始学习率 lr 为 0.005, 以多步方式衰减, 学习率衰减因子 0.1, 权重衰减因子设置为 (weight_decay) 0.000 1, 最大训练 20 epoch.

2.3 评估与比较

提出的模型在 BSDS500 和 NYUDv2 两个数据集上进行了评估, 并且将结果与一些边缘检测方法进行定量评估和定性评估.

2.3.1 定量评估

网络参数. 本文提出了一种基于编解码结构的轻量级边缘检测方法, 与现有的基于 CNN 边缘检测方法相比, 本文方法参数量较少, 仅为 1.54×10^6 , 击败了目前的绝大多数模型, 例如 HED^[13], RCF^[15] 等, 如图 6 所示.

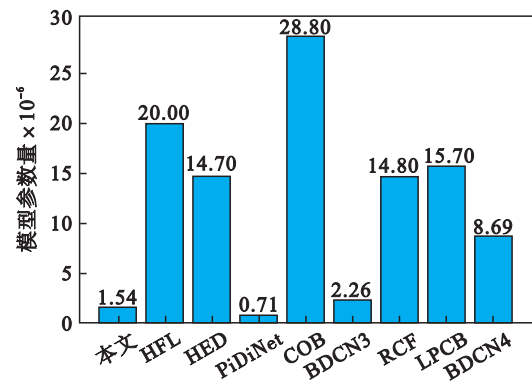


图 6 不同网络模型参数量

Fig. 6 Network model parameter quantity

评价指标. 基于深度学习的边缘检测方法得到的预测结果是一个边缘概率图, 需要先进行非极大值抑制处理得到用于评估的边缘细化图像. 非极大值抑制也就是在某个像素点的 8 邻域内

保留局部最大值,采用该方法是因为基于深度学习的边缘检测方法得到的边缘概率图太密集,其中仍有许多非边缘点,使用非极大值抑制可以尽可能过滤掉非边缘点,得到用于评估的边缘细化图像.接着采用边缘检测中最常用的评价指标在不同的数据集上进行分析,分别是 ODS-F, OIS-F, AP(average precision)以及 P-R 曲线(P-R 指的是精确率-召回率). AP 是 P-R 曲线的积分. ODS-F 是整个数据集中的所有图像共用同一个阈值 η ,使其在整个数据集上的 F-score 最大; OIS-F 指的是在每一张图片上都有一个阈值 η ,使得 F-score 最大^[9].

BSDS500 数据集评估.将本文方法与现有的边缘检测算法在 BSDS500 数据集上比较,如表 1 所示, P-R 曲线如图 7 所示.

表 1 不同方法在 BSDS500 数据集上的比较
Table 1 Comparison of different methods on BSDS500 dataset

方法	ODS	OIS	AP
Human	0.803	0.803	—
Canny ^[6]	0.611	0.676	0.520
Sobel ^[4]	0.539	0.575	0.498
gPb-UCM ^[7]	0.729	0.755	0.745
SE ^[8]	0.743	0.764	0.800
OEF(oriented edge forests) ^[23]	0.746	0.770	0.815
DeepContour ^[10]	0.757	0.776	0.790
DeepEdge ^[11]	0.753	0.772	0.807
HED ^[13]	0.788	0.808	0.815
RCF ^[15]	0.806	0.823	0.834
本文	0.808	0.830	0.843

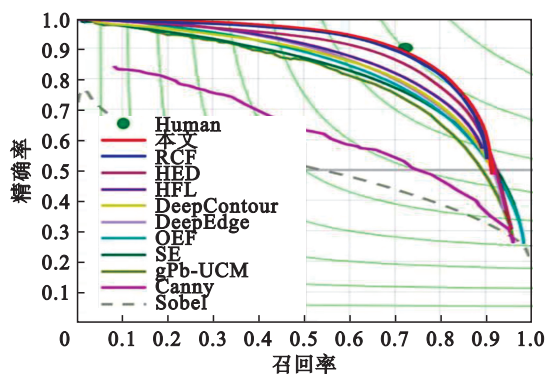


图 7 BSDS500 数据集上的 P-R 曲线
Fig. 7 P-R curves on BSDS500 dataset

从表 1 可以看出,本文提出的算法与其他方法相比,在 ODS, OIS 与 AP 上都取得了更高的得

分,例如与传统的 Canny 算法^[6]相比,本文的 ODS 与 OIS 分别提高了 32.2% 和 22.8%;相比经典的卷积神经网络方法 HED^[13], ODS 与 OIS 分别提高 2.5% 和 2.7%.这说明本文提出的方法性能更优越,可以进一步提高边缘检测的精细度.在 P-R 曲线图中,本文的 P-R 曲线相较其他一些方法比较短,表明本文的模型只会保留非常精确的边缘.

NYUDv2 数据集评估.同 BSDS500 数据集评估一样,将本文的模型与一些其他方法在 NYUD 数据集上进行定量评估,结果如表 2 所示, P-R 曲线见图 8.通过分析,表明该方法在 NYUD 数据集上 ODS 与 OIS 也都取得了较高的分数.

表 2 不同方法在 NYUD 数据集上的比较
Table 2 Comparison of different methods on NYUD dataset

方法	ODS	OIS	AP
gPb-UCM ^[7]	0.631	0.661	0.562
OEF ^[23]	0.651	0.667	0.653
SE ^[8]	0.695	0.708	0.719
HED ^[13]	0.741	0.757	0.749
本文	0.753	0.769	0.763

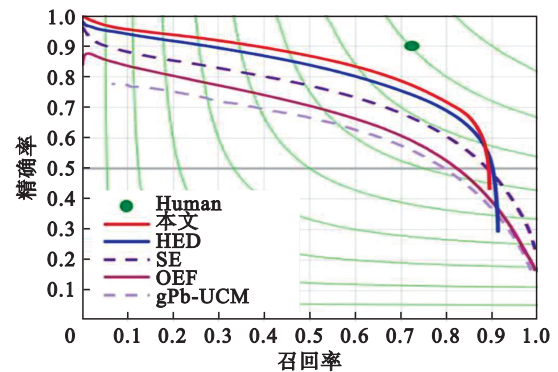


图 8 NYUD 数据集上的 P-R 曲线
Fig. 8 P-R curves on NYUD dataset

消融实验.为了进一步验证本文中 CAM 模块的有效性,以及确定它在本文网络中的最优配置,在 BSDS500 数据集上进行了消融实验.通过 3 次实验分别验证无 CAM 模块、CAM 模块位于编码器(encoder)后面以及 CAM 模块位于解码器(decoder)后面的性能, P-R 曲线如图 9 所示,从图中可以看到 CAM 模块位于解码器后面的 P-R 曲线将另外的 P-R 曲线完全包住,说明 CAM 模块位于解码器后面的性能优于另外两者.

其他实验指标结果如表 3 所示.从表 3 可以看到在网络结构中加入 CAM 模块后各项指标都得到了提高.另外将 CAM 模块放到解码器后面

比放到编码器后面 ODS 及 OIS 分别提高了 1% 和 1.8%。以上结果表明,加入 CAM 模块并且将其放置于解码器后面是本文网络的最优配置,可以抑制对于边缘检测任务不重要的特征通道,使模型专注于边缘信息,提取更精确的边缘特征。

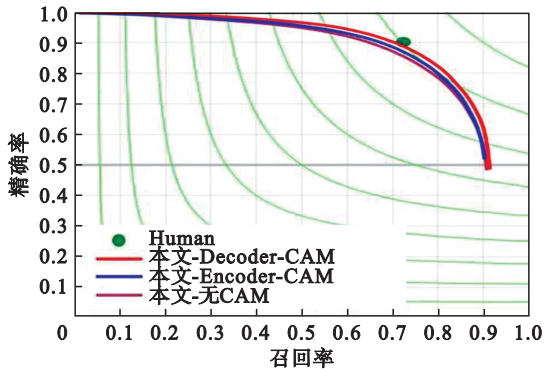


图 9 CAM 模块在不同配置下的 P-R 曲线
Fig. 9 P-R curves of CAM modules in different configurations

表 3 CAM 模块在不同配置下的 ODS 与 OIS 对比
Table 3 Comparison of ODS and OIS of CAM modules in different configurations

配置	ODS	OIS
无 CAM	0.792	0.809
Encoder-CAM	0.800	0.815
Decoder-CAM	0.808	0.830

2.3.2 定性评估

图 10 展示了本文方法提取的边缘图的直观效果,将其与传统的 Canny 方法、基于神经网络的 HED 方法对比,其中 Canny 方法的超参数 $\sigma=2.0$ 。从图中可以看到 Canny 算子检测到的边缘较多且杂乱,将本文方法与 HED 方法相比较,本文方法可以抑制纹理细节信息,有效地消除大多数噪声和模糊的边界,得到更加清晰、细化的边缘。

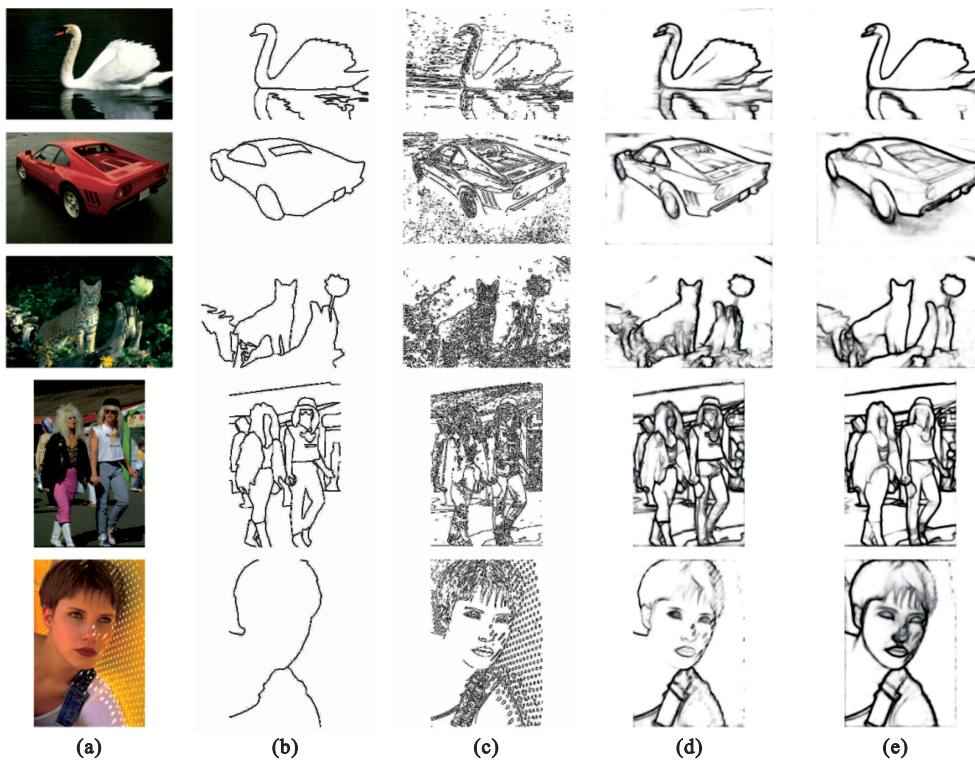


图 10 不同方法的边缘检测效果对比
Fig. 10 Comparison of edge detection effects of different methods
(a)—原图; (b)—Ground Truth; (c)—Canny 方法; (d)—HED 方法; (e)—本文方法。

3 结 语

本文实现了一种基于编解码结构的轻量级边缘检测方法,在一般边缘检测网络的基础上加入了解码器部分,既可以充分利用上下文信息,又一定程度上弥补了下采样操作造成的边缘信

息丢失,同时加入深监督模块更好地识别小物体边缘.实验结果表明,本文方法在常见的边缘检测指标上都取得了较高的分数,可以获得更加清晰准确的边缘图像.但是本文没有加入语义信息的指引,后续的研究将进一步优化来实现语义边缘的检测。

参考文献:

- [1] Zhang M H, Luo H L, Song W, et al. Spectral-spatial offset graph convolutional networks for hyperspectral image classification [J]. *Remote Sensing*, 2021, 13(21): 4342.
- [2] 王玲敏, 段军, 辛立伟. 引入注意力机制的YOLOv5安全帽佩戴检测方法 [J]. *计算机工程与应用*, 2022, 58(9): 303-312.
(Wang Ling-min, Duan Jun, Xin Li-wei. Detection method of YOLOv5 helmet wearing with attention mechanism [J]. *Computer Engineering and Application*, 2022, 58(9): 303-312.)
- [3] Wang R, Qiu K. Fine-grained remote sensing semantic segmentation method under multi-stage supervision learning [C]//14th International Conference on Measuring Technology and Mechatronics Automation (ICMTA). Changsha, 2022: 361-365.
- [4] Sobel I, Feldman G. A 3x3 isotropic gradient operator for image processing [EB/OL]. (2015-06-14) [2022-11-10]. <https://www.researchgate.net/publication/239398674>.
- [5] Prewitt J M S. Object enhancement and extraction [J]. *Picture Processing and Psychopictorics*, 1970, 10(1): 15-19.
- [6] Canny J. A computational approach to edge detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, 8(6): 679-698.
- [7] Arbelaez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation [J]. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 2010, 33(5): 898-916.
- [8] Dollár P, Zitnick C L. Fast edge detection using structured forests [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 37(8): 1558-1570.
- [9] 刘胜男, 宁纪锋. 基于超像素的点互信息边界检测算法 [J]. *计算机应用*, 2016, 36(8): 2296-2300.
(Liu Sheng-nan, Ning Ji-feng. Point mutual information boundary detection algorithm based on super pixel [J]. *Computer Applications*, 2016, 36(8): 2296-2300.)
- [10] Shen W, Wang X G, Wang Y, et al. Deepcontour: a deep convolutional feature learned by positive-sharing loss for contour detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, 2015: 3982-3991.
- [11] Bertasius G, Shi J, Torresani L. Deepedge: a multi-scale bifurcated deep network for top-down contour detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, 2015: 4380-4389.
- [12] Ganin Y, Lempitsky V. N4-Fields: neural network nearest neighbor fields for image transforms [C]//Asian Conference on Computer Vision. Cham: Springer, 2014: 536-551.
- [13] Xie S N, Tu Z W. Holistically-nested edge detection [C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, 2015: 1395-1403.
- [14] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2015-04-10) [2022-11-10]. <https://arxiv.org/abs/1409.1556v2>.
- [15] Liu Y, Cheng M M, Hu X, et al. Richer convolutional features for edge detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 3000-3009.
- [16] Wang Y P, Zhao X, Huang K Q. Deep crisp boundaries [C]//Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition. Honolulu, 2017: 3892-3900.
- [17] He J Z, Zhang S L, Yang M, et al. Bi-directional cascade network for perceptual edge detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2019: 3828-3837.
- [18] Deng R X, Shen C H, Liu S J, et al. Learning to predict crisp boundaries [C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer, 2018: 562-578.
- [19] Su Z, Liu W Z, Yu Z T, et al. Pixel difference networks for efficient edge detection [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal, 2021: 5117-5127.
- [20] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 39: 2481-2495.
- [21] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer, 2015: 234-241.
- [22] Howard A, Sandler M, Chu G, et al. Searching for MobileNetV3 [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, 2019: 1314-1324.
- [23] Hallman S, Fowlkes C C. Oriented edge forests for boundary detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, 2015: 1732-1740.