

引用格式: 张晓闻, 任勇峰. 基于改进 HED 的多尺度边缘检测方法[J]. 微电子学与计算机, 2021, 38(6): 1-6. [ZHANG X W, REN Y F. Improved multi-scale edge detection method based on HED[J]. Microelectronics & Computer, 2021, 38(6): 1-6.]

基于改进 HED 的多尺度边缘检测方法

张晓闻, 任勇峰

(中北大学 电子测试技术国家重点实验室, 山西 太原 030051)

摘 要: 针对基于深度学习的多尺度边缘检测不可避免出现自适应性低, 参数增加, 计算量大, 检测边缘不连续等问题, 本文提出一种基于改进整体嵌套的多尺度边缘检测方法. 该方法将多尺度检测与弱监督模型相结合, 解决参数多计算量大的问题. 为了充分利用卷积强大的特征表达能力, 在整体嵌套边缘检测的基础上, 提出了一种多尺度下深度学习结构, 一个相互独立的多网络多尺度结构, 由不同深度和输出的多个网络组合. 同时引用整体嵌套的权重混合层, 权重混合层将所有的弱监督预测结果连接到一起, 并在训练的过程中学习混合权重. 将所提方法在数据集 BSDS500 上通过评价指标进行性能评价. 实验结果表明, 所提方法能够在数据集 BSDS500 上达到很好的性能.

关键词: 边缘检测; 多尺度; 弱监督学习; 整体嵌套边缘检测

中图分类号: TP31

文献标识码: A

文章编号: 1000—7180(2021)06—0001—06

Improved multi-scale edge detection method based on HED

ZHANG Xiao-wen, REN Yong-feng

(Science and Technology on Electronic Test & Measurement Laboratory,
North University of China, Taiyuan 030051, Shanxi, China)

Abstract: Aiming at the problems of deep learning-based multi-scale edge detection inevitably low adaptability, increased parameters, large calculations, and discontinuous detection edges, this paper proposes a multi-scale edge detection method based on improved overall nesting. The method combines multi-scale detection with weak supervision model to solve the problem of large amount of parameter calculation. In order to make full use of the powerful feature expression ability of convolution, based on the whole nested edge detection, a multi-scale deep learning structure is proposed, which is an independent multi-network multi-scale structure, which is composed of multiple networks with different depths and outputs. At the same time, the whole nested weight mixing layer is referenced. The weight mixing layer connects all the weak supervised prediction results together and learns the mixed weight in the training process. The performance of the proposed method is evaluated by the evaluation index on the data set BSDS500. Experimental results show that the proposed method can achieve good performance on BSDS500.

Key words: Edge detection; Multi-scale; Weak supervised learning; Holistically-Nested Edge Detection

1 引言

边缘检测是图像处理中的基本步骤之一, 理想的边缘检测主要针对边缘的有无、真假和定位等问

题进行展开. 但由于环境、噪声分布及方差等未知因素的加入, 图像的边缘仅靠传统的方法并不能很好的解决. 随着人们对边缘检测任务的要求日益严格, 实际图像中的边缘问题常常发生在不同尺度下, 单

一尺度解决边缘检测问题已经不能满足,因此,多尺度边缘检测成为了理想的解决途径.多尺度边缘检测最初由 Rosenfeld^[1]提出的,结合 Marr^[2]的边缘检测理论,Witkin^[3]的尺度滤波和小波变换理论,经过一系列发展及完善,形成了一套完整的理论体系.但是多尺度边缘检测中始终存在 3 个问题:如何界定尺度范围,如何确定最佳滤波尺度及如何有效的利用多个尺度正确的检测边缘.

随着深度学习的发展,尤其是卷积神经网络(CNN)的出现,以上问题有了相对轻松的解决办法.2011 年开始,依靠 CNN 在自动学习自然图像的高级表示方面具有的优势,利用神经网络进行边缘检测已成为一种新的趋势.同年,Arbelaez 等^[4]将多尺度局部亮度、颜色和纹理结合应用到谱聚类框架中,获得了全局的边缘检测算法.实验结果表明,在数据集 BSDS500 上 F-measure 的值为 0.71.2012 年,Ren 等^[5]提出了一种使用正交匹配追踪算法^[6]和 K-means^[7]字典训练方法来学习,用线性支持向量机分类处理,以生成最终的轮廓.结果在数据集 BSDS500 上 F-measure 的值为 0.74.最近,越来越多的网络被提出用于多边界检测.2015 年,Xie 等^[8]提出了整体嵌套边缘检测(Holistically-Nested Edge Detection, HED),用于以嵌套方式检测和提取自然图像的边缘,而 HED 的优点是增大特征图的感受野,缓解特征图分辨率低的问题,HED 在训练中对注释噪声具有惊人的鲁棒性,缺点是参数增加,计算量大.同年,P. Dollár 等^[9]提出结构化边缘检测算法(Structured Edges, SE),该算法利用封闭轮廓作为训练输入,使用大量的手工设计特性,结构信息,融合多尺度响应.2019 年,Wang 等^[10]提出了一种精确的多类别边缘检测网络,即更丰富的类别感知语义边缘检测网络(R-CASENet),以 ResNet101 网络为骨干,得到最终的融合分类器.尽管如此,边缘检测依旧是一个很有挑战性的技术问题,原因在于:基于深度学习的边缘检测器的训练通常依赖于大量注释良好的图像,注释过程是耗时、昂贵和低效的.为了解决这个问题,研究学者们引入弱监督学习,在弱监督检测技术中,只使用图像级的标注或部分使用包围盒标注对检测器进行训练,对于降低劳动成本和提高检测灵活性具有重要意义.文献[11]中仅使用边框注释就足以实现高质量的对象边界估计.但由于弱监督模型以常规的方式进行训练,从而使用生成的(有噪声)训练数据作为输

入,而不是人工注释.2016 年,Khoreva 等^[12]提出了一种生成弱监督约束条件的方法,该方法使用了一个决策森林和一个对流边缘探测器.利用所提出的弱监督技术,实现了对目标边界检测任务的最优执行,并证明了仅使用边界框注释就可以达到高质量的对象边界,而不需要使用任何特定于对象的边界注释,在很大程度上超过了目前最先进的全监督方法.2016 年,Li 等^[13]提出了一种简单而有效的无监督边缘检测器来训练图像,证明了使用无监督方案训练的边缘检测器的性能接近于在完全监督下训练的方法.

本文从边缘检测及弱监督学习的角度出发,提出了一种基于弱监督学习的改进 HED 多尺度边缘检测方法.在 HED 的基础上,提出了一种网络结构,既解决了全监督学习的计算量问题,又避免了单一输出的不精确问题.同时,利用局部交叉熵及权重控制损失函数训练特征数据,使每个像素能够以不同的概率自适应的进行归类.

2 相关理论

2.1 深度学习网络架构

基于深度学习的多尺度边缘检测由于卷积神经网络具有层次学习的性质,多尺度、多层次学习的概念会因情况而异.而总的概括起来,共分五种结构,如图 1 所示.

(a)多数据流学习结构(Multi-stream learning).多个并行的网络流具有不同的参数值和感受野大小,对应于多个尺度.输入数据被同时输入到多个流中,然后由不同流产生连接的特征响应被输入到一个全局输出层中,从而产生最终的结果.

(b)跳跃式网络学习结构(Skip-layer network learning).该架构的拓扑以一个主流为中心,而不是训练多个并行流.添加链接以合并来自主网络流不同级别的特征响应,然后将这些响应合并到共享输出层中.(a)和(b)两种设置的一个共同点是,在两种架构中,只有一个输出损失函数和一个预测产生.然而,在边缘检测中,通常是获得多个预测值并求取平均值.

(c)多输入的单一网络模型(Single model on multiple inputs).在多尺度的输入图像上运行一个单一的网络或权重固定的网络.这种策略可以作为数据增加发生在训练阶段,也可以作为“集成测试”发生在测试阶段.这种方法在基于非深度学习的方法中也很常见.但是,集成测试会降低学习系统的预测效率,特别是使用更深层次的模型.

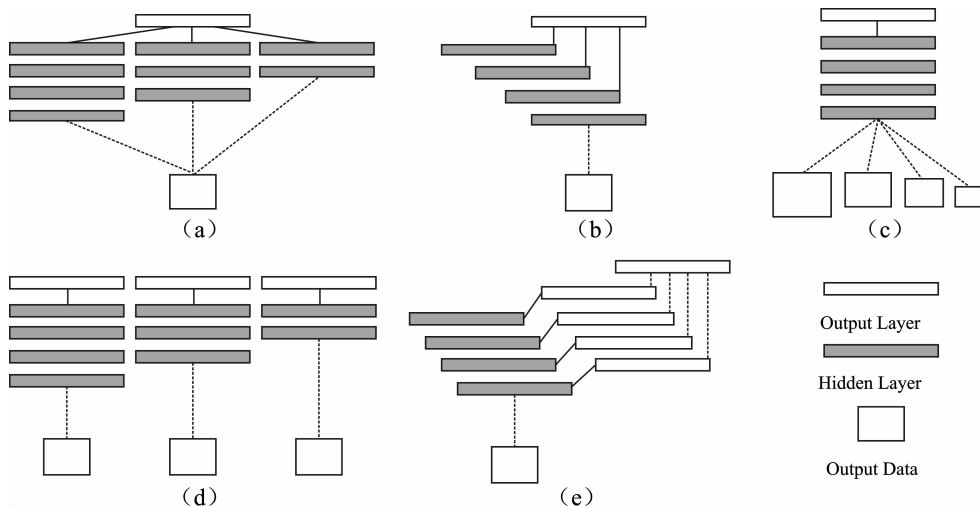


图1 几种网络架构

(d) 多网络独立结构 (Training independent networks). 该结构通过训练具有不同深度和不同输出损失层的多个独立网络来进行多尺度预测. 该结构执行起来可能具有实际的挑战性, 因为这种重复会增加训练所需的资源.

(e) 整体嵌套网络 (Holistically-nested networks). 该网络结构融合了(b)和(d)的优点, 在多个尺度中产生预测值, 它是一个具有多侧输出的单流深度网络.

2.2 HED

文献[8]中提出的 HED, 利用一种新的网络架构组合 ResNet 的复合块, 框架的层次结构主要考虑了架构的深度和步幅的不一致性, 以有效地产生感知多层次的特征和捕获边缘映射的固有尺度. 因此 HED 采用了 VGGNet 架构, 但做了一些修改: (a) 在每一阶段都将侧输出层连接到最后的卷积层, 每个卷积层的感受野大小与相应的侧输出层相同; (b) 去掉了 VGGNet 的最后一个阶段, 包括第五池化层和所有完全连接层.

损失函数作为神经网络中的关键因素, 可以很好得反映模型与实际数据差距的工具, HED 中的损失函数由两部分组成, 单边输出层的损失函数和混合权重的损失函数, 用数学公式表示如下:

$$(W, \omega, h)^* = \operatorname{argmin} (L_{\text{side}}(W, \omega) + L_{\text{fuse}}(W, \omega, h)) \quad (1)$$

式中, L_{side} 表示单边输出层的损失函数; L_{fuse} 表示混合权重的损失函数.

设 $S = \{(X_n, Y_n), n = 1, \dots, N\}$ 表示输入训

练数据集, 样本 $X_n = \{x_j^{(n)}, j = 1, \dots, |X_n|\}$ 表示原始输入图像, $Y_n = \{y_j^{(n)}, j = 1, \dots, |X_n|\}$, $y_j^{(n)} \in (0, 1)$ 表示图像 X_n 对应的实际的二值边缘图. 所有标准网络层参数集合为 W , 假设在网络中有 M 个单边输出层, 每个单边输出层都与分类器相关联, 其中响应权重表示为 $\omega = (\omega^{(1)}, \dots, \omega^{(M)})$. 目标函数为:

$$L_{\text{side}}(W, \omega) = \sum_{m=1}^M \alpha_m \ell_{\text{side}}^{(m)}(W, \omega^{(m)}) \quad (2)$$

在 HED 中, 训练损失函数使用了图像 $X = (x_j, j = 1, \dots, |X|)$ 和图像 $Y = (y_j, j = 1, \dots, |X|)$, $y_j \in (0, 1)$ 中的所有像素. 在自然图像中, 90% 的像素都是非边缘的. HED 中为了自动平衡正样本和负样本之间的损失, 在每个像素项的基础上引入了一个类平衡权重 β . 因此, 使用类平衡交叉熵损失函数数学模型如下:

$$\begin{aligned} \ell_{\text{side}}^{(m)}(W, \omega^{(m)}) = & -\beta \sum_{j \in Y_+} \log P_r(y_j = 1 | X; W, \omega^{(m)}) - \\ & (1-\beta) \sum_{j \in Y_-} \log P_r(y_j = 0 | X; W, \omega^{(m)}) \end{aligned} \quad (3)$$

式中, $\beta = |Y_-| / |Y|$, $1-\beta = |Y_+| / |Y|$ 边缘和非边缘的正确测量数据集用 $|Y_+|$ 和 $|Y_-|$ 表示. $P_r(y_j = 1 | X; W, \omega^{(m)}) = \sigma(a_j^{(m)}) \in [0, 1]$ 使用激活函数 $\sigma(\cdot)$ 在像素 j 上的值进行计算. 在每一个单边输出层上, 进行边缘图预测 $\hat{Y}_{\text{side}}^{(m)} = \sigma(\hat{A}_{\text{side}}^{(m)})$, 其中, $\hat{A}_{\text{side}}^{(m)} \equiv \{a_j^{(m)}, j = 1, \dots, |Y|\}$ 表示层 m 上的单边输出.

为了直接利用单边输出预测, 添加了一个混合权重层到网络并且在训练期间学习融合权重. 混合

权重的数学表达如下:

$$L_{\text{fuse}}(W, w, h) = \text{Dist}(Y, \hat{Y}_{\text{fuse}}) \quad (4)$$

式中, $\hat{Y}_{\text{fuse}} \equiv \sigma(\sum_{m=1}^M h_m \hat{A}_{\text{side}}^{(m)})$, $h = (h_1, \dots, h_M)$ 是融合权重. $\text{Dist}(\cdot, \cdot)$ 是融合后的预测和真实标注之间的距离.

2.3 弱监督学习

文献[14]中将弱监督学习分为三种类型, 不完全监督, 不确切监督和不精确监督. 不完全监督是指在图像分类任务中, 训练数据的真实值标签由人类注释者给出; 虽然很容易从互联网上获取大量图像, 而由于人工成本, 只能注释一小部分图像. 该问题的解决方法, 主要利用主动学习和半监督学习两种技术. 不确切监督是指训练数据只给出了粗粒度标签, 并没有给出具体的示例为 YES 或 NO, 在重要目标检测中, 常常仅有图片级标签, 而没有对象级标签, 该问题的解决办法常常需要进行多示例学习; 不精确监督是指真实标注的值并不完全正确, YES 和 NO 被标记错误, 这将会导致结果错误, 因此利用带噪学习的方法解决不精确监督问题.

3 改进 HED 的边缘检测

文献[8]中, HED 结构是基于 VGG 网络, 将输出层与卷积层相连, 去掉了 VGG 的最后一个池化层和后面所有的连接层, 对整体图像的训练和预测, 可以增大特征图的感受野, 缓解特征图分辨率低的问题, 同时, HED 在训练中对注释噪声具有鲁棒性. 本文在文献的基础上, 通过一种新的网络结构结合弱监督学习进行多尺度边缘检测.

3.1 损失函数

在强监督的方法中, 利用深度学习进行边缘检测, 几乎所有的像素点都是带标注的, 并为每个像素点从 C 个类中预测其所属类别, 设 N 为像素点的个数, 其属于类别 i 的概率为 $\hat{y}_i(i)$, 传统用于多尺度分类的交叉熵损失函数形式为:

$$L_{\text{CE}}(\hat{y}_i, y_i) = - \sum_{c=1}^C y_i(c) \log \hat{y}_i(c) \quad (5)$$

式中, y 表示真实标注.

而弱监督方法的目的是通过较少的像素获得更多的采样点, 只有部分像素点有标注, 会形成一个带有不完整结果的分割图 m_i , 由于无标注的像素点不能为学习过程提供有效信息, 因此, 称为局部交叉熵

损失函数, 表示为:

$$L_{\text{CE}}(\hat{y}, y) = \sum_{i=1}^N m_i L_{\text{CE}}(\hat{y}_i, y_i) / \sum_{i=1}^N m_i \quad (6)$$

为了在不引入任何先验知识的情况下减小弱监督方法和强监督方法的性能差距, 采用文献[15]中 CRF 控制损失函数. 定义类别 i 和 j 的位置 a 和 b 的能量项:

$$\phi_{a,b}(i, j) = \mu(i, j) G_{ab} K_{ab} \quad (7)$$

式中, $G_{ab} = m_a^{\text{dst}} m_b^{\text{src}}$, 其能够有选择的进行计算(只有当 b 是有效的‘源’位置且 a 是有效的‘目标’位置才能够参与计算), $m_a^{\text{src}} \in \{0, 1\}$ 表示与输入图像 x 大小相同的源掩码, $m^{\text{dst}} \in \{0, 1\}$ 表示与输入图像 x 尺寸相同的目标掩码. $\mu(i, h) = \begin{cases} 0, & \text{if } i = j \\ 1, & \text{otherwise} \end{cases}$;

$\phi_{a,b}$ 表示两个点对的势; μ 表示通用的类别矩阵, $K_{a,b} = \sum_{p=1}^P \omega^{(p)} k^{(p)}(f_a^{(p)}, f_b^{(p)})$; 在 $K_{a,b}$ 中, $\omega^{(p)}$ 表示权重; $f_a^{(p)}$ 表示第 p 个核的特征向量; $k^{(p)}(\cdot, \cdot)$ 表示核函数, 数学模型如下:

$$k^{(p)}(f_a^{(p)}, f_b^{(p)}) = \exp \left\{ -\frac{1}{2} \left| \frac{f_a^{(p)} - f_b^{(p)}}{\sigma^{(p)}} \right|^2 \right\} \quad (8)$$

因此预测值 $\hat{y}$ 的能量项表示为:

$$\phi_{a,b}(\hat{y}) = \sum_{i,j \in \{1,C\}} \phi_{a,b}(i, j) \hat{y}_a(i) \hat{y}_b(j) \quad (9)$$

$$\Psi(\hat{y}) = \sum_{a=1}^N \sum_{b=1, b \neq a}^N \phi(a, b)(\hat{y}) \quad (10)$$

在 CRF 中定义位置为 a , 半径为 r 的窗口, $\Omega_r(a) = [a_x - r, a_x + r] \times [a_y - r, a_y + r]$, 由于限制参与对位置 b 处像素点能量求和的范围, 使其只与邻近位置 $\Omega(a)$ 计算能量. 则能量和变为:

$$\Psi(\hat{y}) = \sum_{a=1}^N \sum_{b \in \Omega_r(a)} \phi_{a,b}(\hat{y}) \quad (11)$$

因此 CRF 损失函数表示为:

$$L_{\text{GCRF}}(\hat{y}) = \Psi(\hat{y}) \sum_{i=1}^N m_i^{\text{dst}} \quad (12)$$

为了利用输出预测, 添加了加权融合函数, 并在训练过程中学习融合权重

$$L_{\text{fuse}}(\hat{y}) = \sigma \left(\sum_{m=1}^M h_m y^{(m)} \right) \quad (13)$$

式中, $h = (h_1, \dots, h_m)$ 为融合权重; $y^{(m)}$ 为弱监督学习中的预测值.

因此, 对于最终损失函数为:

$$L(\hat{y}) = L_{\text{CE}}(\hat{y}, y) + \lambda L_{\text{GCRF}}(\hat{y}) + L_{\text{fuse}}(\hat{y}) \quad (14)$$

λ 为权重.

3.2 网络架构

根据文献,本文在五种网络结构的基础上,提出一种新的网络架构,该架构通过多个不同层实现多尺度边缘检测. 为了在弱监督学习的前提下,有效的产生感知多层次特征,该架构以 VGGNet 架构为基础,VGGNet 被认为在 ImageNet 挑战中达到了最先进的性能,也证明了预先训练在一般图像分类任务上对于低层次边缘检测任务是有用的.

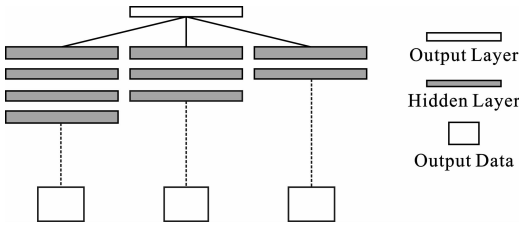


图 2 网络结构

因此,结合图 1(a)和图 1(d)的优点,在网络架构上做如下修改:

(1)多个并行网络流具有不同的参数值和感受野大小,对应于多个尺度,输入数据同时输入多个流中,不同流产生的连接特征响应被输入不同的输出层,获得多个预测值并求取平均值.

(2)由于弱监督模型使用弱注释生成的数据来训练边界检测器,不使用任何实例的对象边界标注或通用边界标注,所以对 VGGNet 中的池化层和完全连接层进行保留. 边缘检测与目标检测不同,在利用神经网络进行训练时,只需要生成外边界标注,这个工作利用弱监督学习即可完成.

(3)由于弱监督学习中仅有一部分的训练集,其余工作都是在训练的过程中调整适应,因此加入混合权重函数.

4 实验

4.1 数据集及实验细节

本文使用公共可用的 Caffe 库来实现,整个网络通过预训练的 VGG-16 网络模型进行初始化微调. 在伯克利大学提供的数据集 BSDS500 (Berkeley Segmentation Sata Set)上进行测试,该数据集包含 200 张训练图,100 张验证图,200 张测试图,每幅图像都标有真实轮廓. 利用三个测试指标:最优数据集规模 (ODS),每幅图像最佳阈值(OIS),平均精度 (AP)对 Canny,HED,SE 及本文方法评估边缘检测精度.

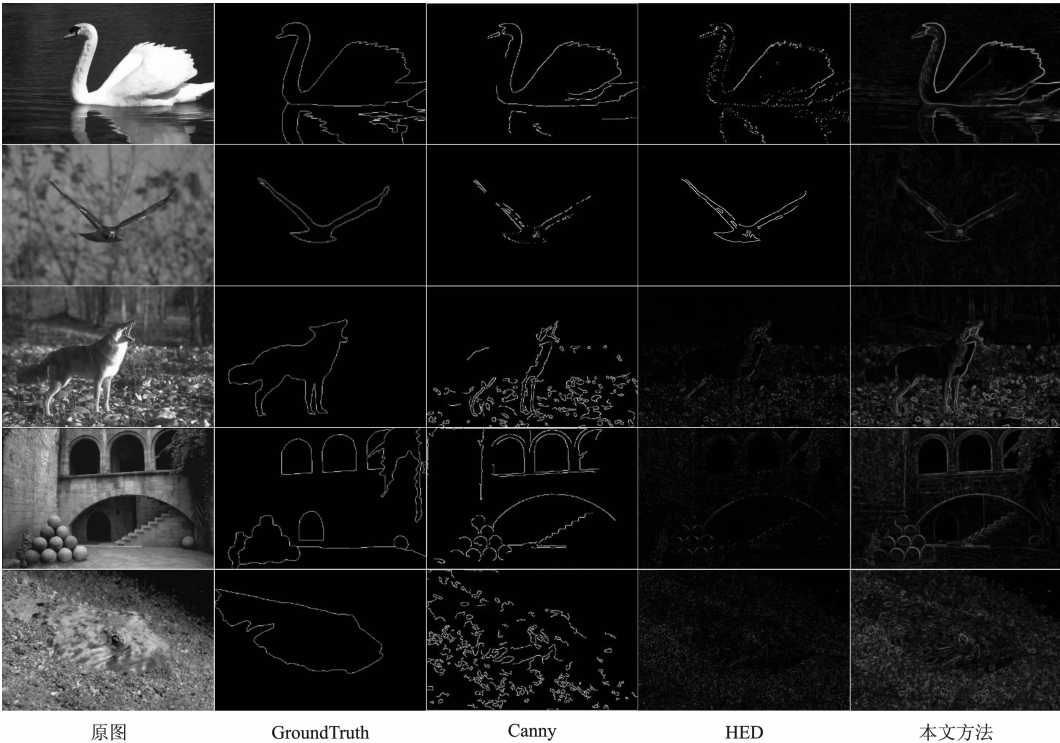


图 3 几种方法的效果对比图

4.2 实验结果

为进一步验证算法的有效性,具体测试数据及

结果见图 3 和表 1. 通过图 3 可以看出对于弱监督学习的多尺度边缘检测也可以达到很好的效果.

表 1 评估指标

方法	ODS	OIS	AP
Canny	0.600	0.640	0.580
SE	0.746	0.767	0.803
HED	0.782	0.804	0.833
本文方法	0.763	0.779	0.817

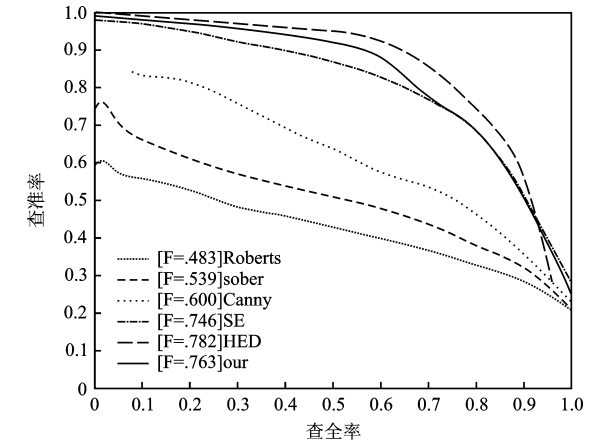


图 4 BSDS500 数据集下边缘检测 P-R 曲线

表 1 中对比了几种算法的三个测试指标,虽然本文方法没有达到传统 HED 的精度范围,但是利用弱监督进行多尺度边缘检测也达到了相当高的精度标准,本文方法的结果是在多个输出后取平均的结果,既验证了弱监督学习的可用性,又验证了网络结构的有效性.图 4 是几种对比方法的 P-R (Precisions-Recall) 曲线,P-R 曲线是对分类器的预测结果按照置信度进行降序排列,然后按此顺序对样例进行预测,计算出当前的查全率作为 x 轴,查准率作为 y 轴,通过 P-R 曲线进一步的验证了算法的有效性.

5 结束语

本文提出一种基于改进的 HED 的多尺度边缘检测方法,将多尺度与弱监督模型相结合,解决参数多计算量大的问题.同时为了充分利用卷积强大的特征表达能力,在 HED 的基础上,提出了一种多尺度下深度学习结构,一个相互独立的多网络多尺度结构,由不同深度和输出的多个网络组合.同时引用 HED 的权重混合层,权重混合层将所有的预测结果连接到一起并且在训练的过程中学习混合权重.将所提方法在 SBD 数据集上通过评价指标进行评价.实验结果表明,所提方法能够在大 BSDS500 数据集上达到很好的性能.

参考文献:

[1] ROSENFELD A, THURSTON M. Edge and curve detection for visual scene analysis[J]. IEEE Transactions on Computers, 1971, C-20(5): 562-569. DOI: 10.1109/T-C.1971.223290.

[2] MARR D, HILDRETH E. Theory of edge detection [J]. Proceedings of the Royal Society of London. Series B, Biological Sciences, 1980, 207 (1167): 187-217. DOI: 10.1098/rspb.1980.0020.

[3] WITKIN A P. Scale-space filtering[C]//Proceedings of the Eighth International Joint Conference on Artificial Intelligence. San Francisco: Morgan Kaufmann Publishers Inc. , 1983: 1019-1022.

[4] ARBELAEZ 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. DOI: 10.1109/TPAMI.2010.161.

[5] REN X F, BO L F. Discriminatively trained sparse code gradients for contour detection[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc. , 2012, 1: 584-592.

[6] PATI Y C, REZAIIFAR R, KRISHNA P S. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition[C]// Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers. Pacific Grove: IEEE, 1993: 1412-1416. DOI: 10.1109/ACSSC.1993.342465.

[7] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, 54 (11): 4311-4322. DOI: 10.1109/TSP.2006.881199.

[8] XIE S N, TU Z W. Holistically-nested edge detection [C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 1395-1403. DOI: 10.1109/ICCV.2015.164.

[9] 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. DOI: 10.1109/TPAMI.2014.2377715.

[10] WANG L Y, SHEN Y, LIU H D, et al. An accurate and efficient multi-category edge detection method[J]. Cognitive Systems Research, 2019, 58: 160-172. DOI: 10.1016/j.cogsys.2019.06.002.