

目标轮廓检测技术新进展

冯芙蓉 张兆功

黑龙江大学计算机科学技术学院 哈尔滨 150080

(crazyidog@163.com)

摘要 轮廓检测是计算机视觉研究领域中最基础、最重要、最具挑战的问题之一。随着近年来深度学习的发展,视觉领域的其他研究方向取得了突破,例如目标检测、实例分割,这些逐渐证明了轮廓检测与其他研究方向的密切关系,因此轮廓检测任务也受到了越来越广泛的关注。文中讨论了多个主体内容,不仅包括对现有轮廓检测算法的细致回顾,而且根据轮廓检测提取特征的特点将其分为3个阶段即低层、中层和高层来介绍,还包括对应用到的数据集、性能评估指标、模型结构和模型细节、轮廓检测的应用及结果的应用进行详细分析,对轮廓检测发展进行了深入介绍。最后,还对轮廓检测所面临的挑战和未来趋势进行了分析和预测,以期为该领域后续的研究提供新思路及参考。

关键词:轮廓检测;计算机视觉;深度学习;分割;聚类

中图法分类号 TP391

Recent Advances for Object Contour Detection Technology

FENG Fu-rong and ZHANG Zhao-gong

School of Computer Science and Technology, Heilongjiang University, Harbin 150080, China

Abstract Object contour detection is one of the most foundational, significant and challenging problems in the field of computer vision research. With the development of deep learning in recent years, breakthroughs have been made in other research directions in the field of vision, such as object detection and instance segmentation, which gradually prove the close relationship between contour detection and other research directions, so more and more attention has been paid in contour detection. This paper discusses several main contents, including not only a detailed review of the existing contour detection algorithms, but also three stages according to the features of contour detection and extraction: low-level, middle-level and high-level, and a detailed analysis of the applied datasets, performance evaluation indicators, model structure and model details, the application of contour detection and the application of its results, so as to make a deep understanding of the development of contour detection. Finally, the challenges and future trends of contour detection are analyzed and predicted. This paper provides new ideas and references for the follow-up research in this field.

Keywords Contour detection, Computer vision, Deep learning, Segmentation, Cluster

1 引言

简单说,轮廓^[1]就是图片中一个沿着物体外延并能包围物体展示出物体的大致形状或曲线,通过这些形状或曲线能识别出来物体的类别。而轮廓检测的任务就是从含有背景物体和前景物体的图片中抑制背景轮廓,忽视前景物体内部的纹理以及各类噪声干扰的影响,将这些代表物体的大致形状或曲线,采用一定的方法提取出来的过程,结果集合即为轮廓。在文献^[1]中有说明与轮廓密切联系的两个概念:边界^[2]和边缘^[3]。轮廓检测与边缘检测息息相关。边缘检测作为计算机视觉领域的关键组成部分,具有丰富的历史。与轮廓检测不同,边缘检测的结果既包括一张图片背景中物体的边界,又包括前景中物体的边界。轮廓检测则是筛去了背景中研究人员不感兴趣的轮廓。只提取感兴趣的轮廓,最近的研究更多开始关注目标的轮廓,也称之为目标轮廓检测^[4]。

轮廓检测提取图片中目标的形状信息,它是计算机视觉任务的基础。例如,它可以用于图像分割^[5-6]、目标检测^[7-8]、语义分割^[9-11]等。实际应用方面,最直接的也是最初的应用则是生成区域提议^[12-13],其实就是目标检测的前期候选区域的生成。再则就是目前研究取得重大成就的自动驾驶^[14-16],虽然自动驾驶是利用实例分割^[17-20]和语义分割,但其本质上仍是找到物体的轮廓范围。

(1)本文综述轮廓相关检测技术并介绍新进展

1)对轮廓检测算法的全面回顾:本文回顾了轮廓检测发展史上的各类方法,并对最新研究做出了分析。

2)讨论了轮廓检测各个部分的细节:本文分析了轮廓检测中模型结构,主要是深度学习模型^[21]。目前与轮廓密切相关的研究方向发展迅速,用于语义分割、图像分割的卷积网络有利于CEDN^[4]的发展,同时证明了图像分割、目标轮廓检测和边缘检测之间的密切关系,所以也对相关领域的模型进行

项目基金:黑龙江省自然科学基金(F2017024, F2017025)

This work was supported by the Natural Science Foundation of Heilongjiang Province, China(F2017024, F2017025).

通信作者:张兆功(2013010@hlju.edu.cn)

分析。本文还分析了当前应用于计算机视觉领域的各个数据集,并指出在数据集上一直存在的精确标注工作量大的问题,同时也提出了可行的解决办法——将弱监督学习应用到轮廓检测中。

3)加入了最新研究成果并结合本问题最新的研究状况和相关问题研究提出了新的可能发展的方向。比如,“最新研究解决了什么问题?”“最新研究的贡献和优缺点”等,同时针对其不足提出自己的看法和思路,这也是作者的下一研究方向。

(2) 轮廓检测的难点、目前与未来面临的挑战

轮廓检测发展的过程就像是一个上升的螺旋线,经历了曲折与坎坷,也看见过曙光并且不断走上其发展的光明大道。轮廓检测与其他研究任务都会有各式各样难点。首先轮廓检测自开始就有自身的困难,因为轮廓检测不像目标检测只需要一个矩形框框住目标并标上其置信度,也不像图像分类只需要分类器分辨出目标的类别,它需要精确把握住目标的边界像素点。对于人工标注的数据集,由于它的准确度不能得到保证,缺失或不闭合或不精准也是在所难免,故检测的难度加大。在实际检测中,还包括照片拍摄的角度变化、光照变化、目标遮挡等情况。这些将在第4节详细介绍。

本文第2节从轮廓检测提取特征的特点出发,将其分为3个不同阶段来分别介绍轮廓检测,第一个是低层次特征阶段,基于像素点,第二个是中层次特征阶段,基于部分边缘检测(非深度学习),第三个是高层次特征阶段,主要基于深度学习以及与轮廓检测密切相关的研究方向;第3节讨论了轮廓检测数据集、评估指标、模型相关的细节和轮廓检测结果的应用;第4节将讨论轮廓检测所面临的挑战和未来趋势;最后总结全文。

2 轮廓检测的发展历程

在介绍各个算法之前,先通过一个时间轴来了解轮廓检测发展的整体脉络,如图1所示。

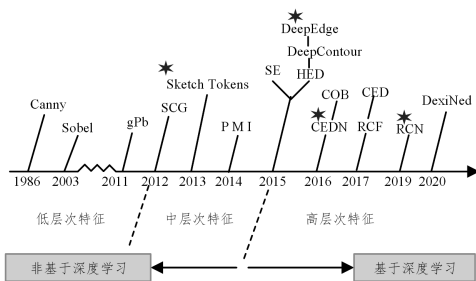


图1 轮廓检测发展历程

Fig. 1 Development historical map of contour detection

在这个发展过程中,由于技术的不断更新,轮廓检测这项任务在不同阶段考虑的因素不同,低层、中层、高层关注的特征不同,所以方法也不一样。低层关注像素、纹理,中层关注部分纹理和部分轮廓,高层关注中间细节信息和轮廓。

2.1 低层次特征

本文将更多关注点放在一张图片像素点的灰度值、颜色值、亮度值的变化以及一些图像纹理等单一的或者它们之间相结合的特征的轮廓检测方法视为低层次的。

轮廓检测的发展和边缘检测密不可分,最开始是使用传统的边缘检测算子作为基础提取物体轮廓。早期研究集中在设计简单的过滤器来检测在局部相邻的像素的最高梯度,这

种方法通常将图像与微分算子卷积,他们只看相邻像素之间的颜色或者亮度差异而不能在更大的感受野中分辨纹理差异,这也是他们没有考虑到高层的特征,比如借助于Sobe^[22]和Canny^[23]等方法,因此很难得到相对完整和封闭的目标轮廓。随着研究的发展、各类边缘检测的方法也在不断更新,尤其是纹理^[24]描述的进步。Martin等^[2]在边界检测器中结合了颜色、亮度和纹理梯度,提出了Pb算法。Arbelaez等^[6]又通过计算多尺度和SVD的局部线索改进了当时精度非常高的gPb,唯一不足的是其计算量不断增加。以上这些方法用于轮廓检测仅仅依赖低层次特征信息,缺乏较高层次的特征信息。

2.2 中层次特征

本文将部分基于非深度学习的边缘检测来提取轮廓作为中层次的特征阶段方法,因为这种方法在一定程度上关注了比之前更多、范围更大的特征,检测任务应用到更多的信息,从而检测结果也更加准确。

sketch token:Lim等提出了一个中层的学习轮廓和目标检测的表征——素描令牌^[25](sketch token)方法。他们提出了一个针对中级特征的新方法来学习与检测基于真实轮廓的模型表示。中级特征指的是在计算机视觉研究领域中,处于低级别像素信息与高级特征之间的过渡概念。首先,将人工标注好的二值图像按照 35×35 区域分成多个图像块,只有每个区域的中心像素正好在轮廓上才使用这个图像块,这个产生过程是通过有监督的完成的。然后就可获得很多含有轮廓形状的图像块,使用K-means对其进行聚类,那么每个图像块都有自己的类标,称为sketch token——代表图像中各种局部边缘结构。将这些作为数据集训练分类器自动识别图像块中所属的sketch token。故对于新的彩色图像,就需要检测出其存在的sketch token,将图像中所有的图像块(提取图像块中包含的特征——通道索引特征和自相似特征)输入分类器(结构化随机森林^[26])中,预测一个图像块隶属于sketch token或不属于的可能性。由于每个sketch token都有一个轮廓位于其中心,利用sketch token概率的总和计算轮廓在中心像素的概率,通过非极大值抑制得出最终的轮廓的响应值。该方法在轮廓的提取中将像素点进行二分类(是轮廓或不是轮廓)转化为多类分类,将轮廓信息作为多类分类的类标,降低了分类器分类的难度。

SE:把结构化学学习的方法应用在决策森林上,因此称之为结构化随机森林^[26](Structured Forests for Fast Edge Detection,SE)。该方法利用图像局部块固有的结构(如相邻像素之间的关系,这是之前方法中没有考虑的方面)将边缘检测问题表述为给定输入图像中图像块的局部分割掩模预测问题。它利用图像的局部特征,将这些图像块放到每个森林里预测每一块边缘像素的标签,再把这些标签聚合在图像上计算最终的边缘图。

中层次的轮廓检测考虑了图像的局部块特征,虽然无法看到图像的全局,且没有涉及到卷积神经网络(Convolutional Neural Networks,CNNs),但相比于传统低层次的方法,其性能已经提高很多,并且也逐步看到了特征之间的关系,而不是单独考虑某一个特征。

2.3 高层次特征

基于深度学习的轮廓检测因为更关注高层次所提取的抽

象特征,本文称之为高层次特征的轮廓检测,具体分为局部高层次特征和全局高层次特征。

2.3.1 局部高层次特征

N^4 -Fields^[27]:轮廓检测中一个非常直观的方法。针对图像中很多的块,通过 CNN 提取出每个块的特征,然后在字典里进行查找,找出与块特征相近的边缘,最后将所有的边缘信息集中在一起得到结果。

DeepEdge^[28]:第一个使用图像高层次特征来进行轮廓检测的算法(A Multi-Scale Bifurcated Deep Network for Top-Down Contour Detection)。该文提出轮廓检测和物体识别等任务相关性的概念,并提出以特征提取通用性非常强的 Knet^[29]作为对象级特征提取网络与一个分叉全连接的子网络共同组成的多尺度深层网络结构。模型借鉴了 Knet 的前 5 个卷积层,是为了保存在全连接前准确的位置信息;分叉全连接的子网络又包括两个部分:一个是学习轮廓分类(通过给定判断是否为轮廓),一个是学习轮廓回归。选择 4 种不同大小、不同程度的包含整个图像范围的局部块作为输入,然后调整它们的尺度以适应 Knet,并将这 4 个平行且相同的数据流连接到两个独立训练的分支组成的分叉子网络,再通过两个分支预测。它将一个轮廓检测问题转化成分类和回归综合问题,把它们的结果作为最后的轮廓检测的结果输出。在数据集上,它是直接从原始的图像上学习物体轮廓。

DeepContour^[30]:它结合了 sketch token^[25]优势和 CNN 来提取轮廓,相比素描令牌能更好地提取特征。首先利用聚类对真实标注的图像的局部轮廓进行分类;然后通过 CNN 来学习这些轮廓的深层特征,区分开物体本身的纹理和轮廓,还额外增加了轮廓错误分类到背景和背景错误分类到轮廓的损失函数,凸显轮廓和背景之间分类出错不可忽视;最后使用结构化随机森林将局部的轮廓特征图整合在一起,输出最终的轮廓图。

2.3.2 全局高层次特征

HED^[31]:一种基于 FCN^[32](Fully Convolutional Networks)和深度监督网络的新的边缘检测算法(Holistically-Nested Edge Detection, HED),与 DeepEdge^[28]和 DeepContour^[30]不同的是它是基于全局高层特征。它的思想源于多尺度和多级别学习的概念。该网络使用同一个深度学习网络的不同层级侧面输出(side output),再将 Deep supervision 应用在每个 side-output 层,指导 side-outputs 获取特征进行边缘预测,进一步在不同分辨率的图层上附加一个线性分类器,并使用双线性插值(反卷积)将其恢复到原图像的分辨率。由于 HED 的输出是多尺度和多层次的,并且 side-output 平面尺寸逐渐变小,感受野尺寸逐渐变大,因此加入了一个权重融合层去自动学习多尺度输出的结合,即训练一个权重融合函数得到最终的边缘输出。该网络是第一个实现了端到端训练的轮廓检测。

COB^[33]:在 HED 的基础上,使用一个类似 HED 的网络结构衍生出来的多尺度(低尺度、高尺度、多方向)面向对象和一个高效的层次图像分割算法。其性能优于大多数轮廓检测算法。

CEDN^[4]:之前的文献^[34]中对物体轮廓检测的探索相对较少,许多工作都致力于边缘检测,以响应前景物体和背景边界,基于全卷积编解码网络的目标轮廓检测(Object Con-

tour Detection with a Fully Convolutional Encoder-Decoder Network, CEDN)第一次将全卷积编解码网络应用在轮廓检测上,而且更关注目标的轮廓。与 HED 相比,它没有使用 Deep supervision,而是通过特征的逐层解码来构造解码网络。由于图像标记问题的不对称性,它还引入了一种轻量级解码器,打破了反卷积网络的对称结构。在数据集上,为了使数据更加准确,可以直接用于训练,提出 Dense CRF^[35]将轮廓对齐问题看作一个多类标记问题,然后用相邻的实例标签填充不确定的区域,从而得到标签边界处的精确轮廓。

RCF^[36]:在 HED 的基础上提出了基于更丰富的卷积特征(Richer Convolutional Features, RCF)的精确边缘检测算法,其核心思想依然是编解码网络,但与以往基于 CNN 不同的是它关注每阶段每次卷积后的结果而不仅仅是 CNN 最后一层的结果。该网络共分为 5 个阶段,每一阶段的每次卷积后的特征图都通过 1×1 的卷积核,与该阶段其他同样操作之后的特征图相融合,再进行解码,最后将所有阶段解码后的特征图融合在一起生成最终的轮廓。该方法大幅度提高了检测的准确性和性能。

CED:为了解决之前边缘检测中模糊的情况,Wang 等^[37]在线性的编码-解码结构^[4]的基础上提出了 CED(Deep Crisp Boundaries)网络,通过使用高效的子像素卷积^[38]代替了 HED 中的双线性插值,CED 能够生成更好的与图像边界对齐的边缘映射。

RCN^[39]:基于 RefineContourNet 的目标轮廓与边界检测(Object Contour and Edge Detection with RefineContourNet, RCN),它在 RefineNet^[40]的基础上提出了自己的概念,并对其改进:将原图第一次卷积操作作为第一层;在最后的融合操作后加入了 sigmoid 激活函数预测是否为轮廓的概率。最重要的是它融合了高、中、低层次的特点,使用具有最高水平特征的张量作为起点,将它逐层与较低抽象层的特征结合起来与远距离残差连接来做出较高分辨率的预测,相比于简单的跳层融合更有效。不足的是,对于目标轮廓检测任务来说,粗边缘预测对 RCN 的不利影响依然存在。图 2 给出了 RCN 的网络结构。

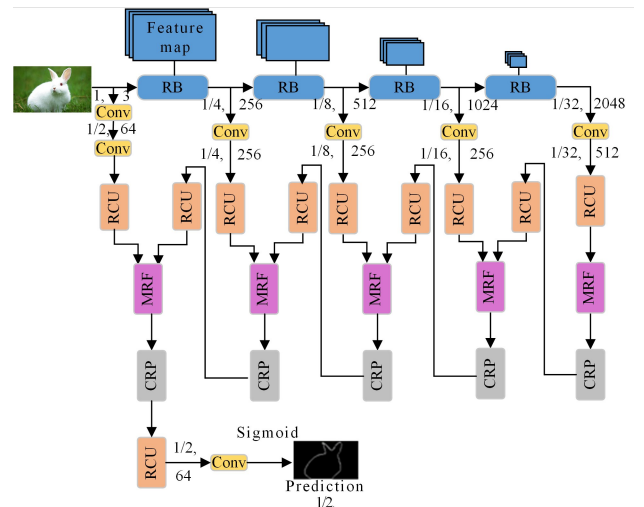


图 2 RCN 的网络结构

Fig. 2 Network structure of RCN

DexiNed:一个与之前基于 CNN 截然不同的检测器,不像 HED 等需要经过 ImageNet 预训练,它是从头开始训练

的,是一个新的 CNN 网络结构^[41],主要用于稠密的边缘检测或者轮廓检测。3.1 节有更具体的说明。

综上所述,基于深度学习的轮廓检测前期是通过将图像局部区域送入网络中进行了局部特征的提取,最后将其整合输出轮廓图。而后期都开始注重弱化或解决 CNN 的缺点,

表 1 各类轮廓检测算法在数据集 BSDS500 和 PASCAL 上的性能

Table 1 Performance of partial contour detection algorithm based on deep learning on BSDS500 and PASCAL

Algorithm	BSDS500			PASCAL		
	ODS	OIS	AP	ODS	OIS	AP
HED	0.778	0.808	0.840	0.587	0.592	0.443
COB	0.793	0.819	0.849	0.624	0.657	0.593
RCF	0.811	0.830	0.846	—	—	—
CED	0.822	0.840	0.895	0.726	0.750	0.778
RCN	0.823	0.838	0.853	0.752	0.773	0.641
RCN-VOC	0.824	0.839	0.837	0.721	0.746	0.613
DexiNed-fla	0.707 0.709	0.723 0.726	0.520 0.738	0.431 0.475	0.458 0.497	0.274 0.329

3 模型与评价指标

3.1 相关模型结构和模型细节分析

这一部分首先对最经典的基础网络做出了介绍,其次针对最近几年在基础网络上改进的变体网络做出分析,并分别说明它们的贡献和特点。

AlexNet:在 2012 年,AlexNet^[29]是第一个在 ImageNet 分类上表现出色的大规模卷积神经网络体系结构,大大超越了以前所有基于非深度学习的模型。AlexNet 网络结构更简单、更容易理解,共包含 5 个卷积层和 3 个全连接层。其中每个卷积层都添加了 RELU 激活函数和局部响应归一化(LRN)的操作,紧接着是常规的下采样和池化。

VGGNet^[41]:相对于 AlexNet 更深的网络模型,而且使用更小的卷积核,减少了参数数量或者多个卷积核构成的卷积组,可以改变特征图的通道数,获取更多信息。该模型最常用的是 VGG16^[41]和 VGG19 等,其中 VGG16 是很多模型的主干网络,在实验中也表现出良好的效果。VGGNet 与 AlexNet 二者都是通过纵向增加网络的层数来达到很好的训练效果。由于网络越来越深,其反向传播参数调整过程中必然会出现问题,比如梯度消失、梯度爆炸等。

GoogleNet^[43]:2014 年 Szegedy 提出的一种全新的深度学习结构,其核心是 Inception 模块,模块内部由多个这样的 Inception 模块组成。该模块本质上是将一个卷积核拆分为 1×1 (可以降低维度,减少计算量)和 3×3 的卷积,其实就是将深的网络横向变宽了。实验发现这样做的效果比之前的好很多,且速度更快。它与 VGGNet 的分类模型在原理上与传统的 CNN 模型并没有太大差别。

Resnet:由微软研究院的 He 等^[44]4 名华人提出,其主要思想是在网络中加入了直连通道,即将前一层或前几层特征图信息直接加入后一层卷积后的特征图上,将它们融合,保护信息的完整性。整个网络只需要学习输入、输出差别的那一部分,简化学习目标和难度,同时一定程度上解决了深网络的弊端。尽管如此,但由于卷积神经网络固有的操作:不断地下采样和池化,直接使得特征图的分辨率越来越低。如今这个弊端愈演愈烈,检测的难度自然也就加大了。研究者针对这个问题想出了如下解决的办法。

(1)利用上采样、上池化的操作,将特征图恢复到原来图

并为了获取更多细节信息和丢失的信息,更加关注编码解码网络以及重视网络各层特征和融合各层特征对于轮廓提取的重要性。表 1 列出了各类轮廓检测算法在数据集 BSDS500 和 PASCAL 上的性能,其中性能指标 ODS, OIS, AP 在 3.3 节给出了详细说明。

像的尺寸。但由于上采样是在上一层操作后得到的特征图上将图上的每个像素点按照设定的步长复制该像素点进行填充,并且引入了更多参数去学习,所以即使得到了原来图像的大小,其分辨率也依然很低,细节和空间信息无法还原。上池化也是和上采样类似的操作,只不过上池化需要记录上一层操作后的路径,并对图像上的每个像素点周围按照步长进行 0 填充。当然,上采样还包括亚像素采样,同样不能达到预期的目的。

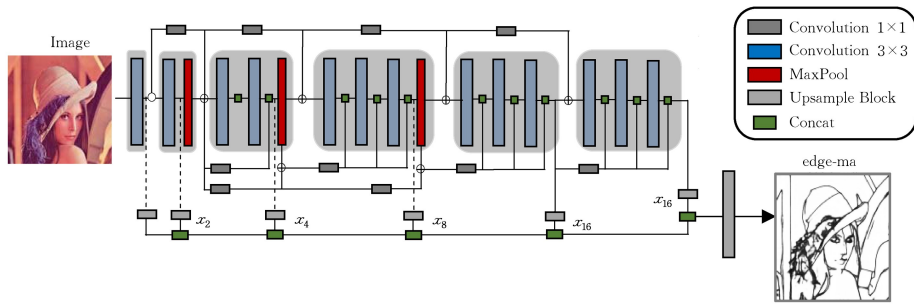
(2)利用空洞卷积^[45],在原来的像素点上能看到比之前更多的信息,而且具有更少的参数和更快的运行时间。但空洞卷积由于卷积核和卷积的区域较大,所以其计算量自然也就大;高层次的和高分辨率的特征图也被需要比较大的 GPU 存储。

RefineNet^[40]:针对 CNN 的缺陷,充分利用采样过程中的可用信息。该模型具体是将卷积后不同分辨率的特征图分成 4 个不同模块,模块与模块之间借助 Resnet 思想将其连接:将上一层模块或者上几层模块的信息直接输入下一层并与该层卷积后的特征图信息相融合,期望获得更多的细节信息和下采样、池化导致丢失的信息,再将每个模块除了第四个模块以外输入到 RefineNet 模块中。每一个 RefineNet 模块都有两个输入,一个是本阶段的特征图,另一个是低层产生的经过处理的特征图。这样随着上采样的进行,上下文语义信息也逐渐丰富,最终得到的得分图,经过上采样操作恢复原有图片大小。该模型在 PASCAL-VAL 数据集上进行评估时,达到了优异的性能和最佳的数据集规模(ODS,0.752)。用于边缘检测和轮廓检测的 RefineContourNet^[39]也是由 RefineNet 的思想而来。

DexiNed^[41]:该方法受到了 HED^[31]网络和 Xception^[46]网络的启发,是一个全新的网络结构(见图 3)。它将网络分为两个部分:Dexi(Dense Extreme Inception Network)和 UB(Upsampling Block)。Dexi 部分有 6 个主要的模块,类似 RefineNet,是根据不同分辨率划分。主块之间通过 1×1 的卷积连接,这其实就利用了 Resnet 思想,目的还是尽可能多地搜集信息。而主块内部各个卷积操作之间同样使用 1×1 的卷积连接,做到不丢失任何数据。即便如此,随着卷积次数和池化增多,特征图的分辨率依旧降低,为此就将每个主块产生的特征图作为 UB 部分的输入,试图在这个部分还原其分

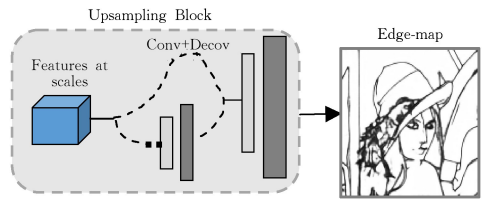
分辨率,事实证明是可以的。在 UB 内部有卷积和反卷积两种操作用以提取中间的边缘映射信息,为融合各个主块特征图

并恢复图像原来尺度做准备。该模型应用自己的数据集达到目前边缘检测的先进水平。



图中上方是 Dexi 部分,灰色区域块代表 6 个主模块。图中下方灰色小块是 UB 部分

(a) DexiNed 网络结构



(b) DexiNed 网络中的 UB(内部含有迭代卷积编码解码操作^[41])

图 3 DexiNed

Fig. 3 DexiNed

Polarmask^[47]:一种全新的实例分割方法。主要基于 FCOS^[48](Fully Convolutional One-Stage Object Detection)模型,借助 FPN^[49](Feature Pyramid Networks for Object Detection)的特点,再加上分割任务的需要而改进了 FCOS,把实例分割统一到 FCN 框架下,这样目标检测、实例分割和轮廓检测之间就可以在在一定程度上相互转化(模型上做到了通用的性能)。该文还提出了一种新的轮廓建模方式,对轮廓检测来说是一种更简单、更直接的方式,也是一种新的启发。这种建模方式与 Mask R-CNN^[20]不同。Mask R-CNN 是将目标检测和语义分割结合起来,通过先检测出来目标的位置,再在更小的范围内借助像素级的建模方式对每个像素点进行分类,比在大的范围内分类简单很多。

上述模型本质上都依然是 CNN。由于现在技术的不断进步和实际情况的日渐复杂,研究者们开始在 CNN 每层卷积或更细致的阶段做文章。从低层到高层,根据卷积特点和反卷积的特点,结合不同的研究任务努力提取更多细节、更符合真实情况的信息。这样就使得后续的研究在速度、精确度上得到更大的保证。

3.2 应用到的数据集

MDBD^[50]:从早期的视觉任务来看,边缘检测的多分辨率数据集主要用于自然场景中物体边缘检测的心理物理研究。数据集由短双目视频自然场景序列^[50]组成,包含 100 个高清晰度场景(1280 × 720)。每个场景有 5 个边界标注和 6 个边界标注。从给定的数据集中,选取 80 张图片用于训练,其余 20 张用于测试。

CID:一种受脑生物启发的边缘检测技术数据集^[51]。它的唯一不足是数据较少,只有 40 个图像和经过真实标注的图像。它关注的除了边缘的真实标注图像之外,还包括目标轮廓图像。

BSDS^[52]:主要用于图像分割和物体边缘检测。该数据集分为两个版本:BSDS300 和 BSCS500。目前常用的是 BS-

DS500,其中有 200 个用于训练,200 个用于测试,剩下的 100 个用于验证,并且数据集中每张图片都有 5 个对应的真值。在训练时可以通过对真值采用平均值的方式或者旋转、翻主、放大等操作来丰富数据集。

NYUD^[53]:该数据集共有 1449 张 RGB 图像,每张 640 × 480 像素,且被分成 3 个子集,即训练集、验证集和测试集。其中包含 464 个室内场景用于数据分割。除了 654 张图像用于测试,其他数据用来训练和验证模型。

Pascal-context^[54]:目前比较流行的数据集,它是 PASCAL2012 数据集的扩展,主要用于边缘检测和边界检测的深度学习模型训练和测试。它包括 11530 个注释图像,其中约 505 张图像已用于测试。

PASCAL VOC^[55]:这个数据集有两个不同的版本:PASCAL2007 和 PASCAL2012。它是视觉领域中对对象分类识别和检测的一个基准测试数据集,提供了检测算法和学习性能的标准图像注释数据集以及标准的评估系统。

BIPED(Barcelona Images for Perceptual Edge Detection)^[41]:最新的、被计算机视觉研究领域的专家经过仔细注解的、经过反复检查过、没有可能存在误差的边缘数据集。它包含 250 张外景图片,每张 1280 × 720 像素。因为在以前的边缘检测任务中缺乏这样的精准的数据集,所以这个数据集的出现也直接促使 DexiNed 网络能够在其他数据集中进行边缘检测评估,而不需要在这些数据集上进行训练,且效果达到先进水平。这个数据集是公开的,作为评估边缘检测算法的基准。

3.3 性能评估指标

一个好的性能评估指标对于模型检测的准确性评价是非常重要的,这一小节将讨论轮廓检测通用指标,这里 TP、FP 和 FN 分别代表轮廓像素点检测的正确样本个数、错误样本个数和遗漏样本个数。

1)P-R 曲线:以精确度为纵轴,召回率为横轴得到的图称

为 P-R 曲线或者 P-R 图,如图 4 所示。它可以直观地展示学习器在样本总体上的精确度和召回率,还可以观察到在同样的数值情况下,各类方法的精确度或召回率以及它们的平衡点位置也即是横轴、纵轴二者数值相等的情况。2004 年, Martin 等^[2]将其完善并应用到多标注的真实轮廓评价系统中。其定量指标下计算如式(1)所示:

$$F = \frac{P * R}{(1-\alpha)P + R} \quad (1)$$

其中, α 为权重,一般取 0.5; P 和 R 分别代表精确度和召回率。

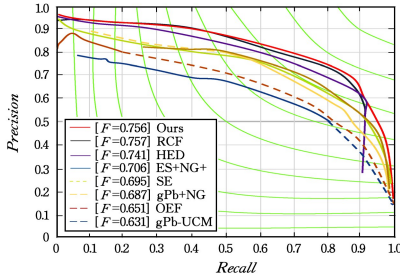


图 4 数据集 NYUDv2 下轮廓检测算法的 P-R 曲线^[56]

Fig. 4 P-R curve for contour detection on NYUDv2^[56]

(2)精确度(Precision):也称为查准率,是针对预测结果而言的,其含义是在所有被预测为正样本中实际为正样本的概率。计算公式如式(2)所示:

$$P = \frac{TP}{TP + FP} \quad (2)$$

(3)召回率(Recall):也被称为查全率,是指在原始样本中实际为正样本个数中在预测中仍然被预测为正样本的概率。其计算如式(3)所示:

$$R = \frac{TP}{TP + FN} \quad (3)$$

(4)最佳数据集规模(Optimal Dataset Scale, ODS):轮廓检测中最常应用的评价标准。对于给定的阈值 T ,整个数据集将会得到一个平均的最优 F 值。

(5)单图最佳阈值(Per-image Best Threshold, OIS):数据集中每一幅图像都单独给定一个最优阈值 T ,综合后的最优 F 值被称作最优的图像尺度。

3.4 目标轮廓检测的应用及结果的应用

轮廓检测的一个直接应用是生成候选区域。贝塔修斯等^[57]改进了卷积特征的语义轮廓检测,并展示了它在语义分割中的应用。目标轮廓检测器可以用来改进一个更具挑战性和实用性的实例级语义分割问题^[58]。

在目标检测任务中,最重要的也是区域的生成。传统的滑动窗口、BING^[58]和 Edge Boxes^[13]、SS^[12]等方法,都是为了对最后框出目标做的前期准备。随着深度学习的发展,各类深度学习模型的提出,不管是目标检测、轮廓检测还是边缘检测都得到了突破性发展。在目标检测 Faster R-CNN^[60]中引入了一个全卷积的、经过端到端的训练、可以生成高质量的区域提议网络(RPN)。在 CEDN 中提出将一种全卷积编码解码网络应用于轮廓检测,先产生轮廓,再结合多尺度组合分组算法^[11, 61-62]等方法,最终生成目标检测的候选区域。

上面介绍的是基于候选区域的产生的目标检测。随着近几年深度学习的发展,目标检测已经呈现出两个趋势:一种是 two-stage,就是借助生成区域提议的算法或者借助锚点,通过设置不同尺度、不同比例的参数映射到原图上生成大量的可

能包含待检测目标的区域框,再通过一个 classifier(SVM 或者 softmax)识别每个区域框里是否真的包含目标物体,以及物体的类别概率。例如 R-CNN^[63], SPP-Net^[64], Fast R-CNN^[65], Faster R-CNN, R-FCN^[66]。另一种是 one-stage,基于端到端(end to end)的,不需要产生候选区域,如 YOLO^[67], SSD^[68]等。二者发展都很迅速:区域提议准确率较好,但是实时性差些;端到端的方法速度较快而且实时性能好,但不能准确将物体检测到。如 YOLO,对于较小的物体是不能检测到的,尽管 SSD 针对 YOLO 不足进行了改善,实时性和准确性都得到了保证,但这些方法本质上依然没有做到完全的 proposal region-free, anchor-free。作者根据文献^[47]中的建模方式,提出一种由目标轮廓直接产生预测框的设想,这也是作者另一篇论文研究的点。因为这种建模方式可以知道目标轮廓的中心点和确定中心点过程得到的多个中心点以及中心点到轮廓点的距离,将这些距离降序排序,首先找出最长距离的轮廓点,以它为基点画相互垂直的射线包围轮廓,然后用与两条射线同侧的轮廓点做出到射线的距离,找出距离最大的轮廓点再分别向两条射线作垂线,那么射线交互点整体构成一个矩形框,此时所得到的矩形框可能不完全准确,接着可以根据以往目标检测预测框和目标之间的间隙作为一个先验值,将刚刚得到的矩形框进行变换。具体操作过程如图 5 所示。

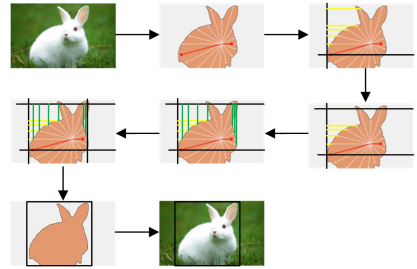


图 5 基于 Polarmask 的极坐标建模方式下根据目标轮廓直接产生目标检测的预测框过程

Fig. 5 Design process diagram of directly generating the final prediction boxes of object detection from the object contour, which based on polar modeling of polarmask

在分割任务中,可以由已知目标轮廓来确定目标的分割范围,再根据具体的分割要求实现分割。不管是实例分割还是语义分割,轮廓的检测显得非常重要。Mask-RCNN 中像素级的建模方式和 Polarmask 中极坐标的轮廓建模方式对轮廓检测有很大作用。

4 轮廓检测所面临的挑战和未来趋势

虽然现在的科技、技术不断进步,但是在数据集上,很多任务依然需要监督式的数据进行不断训练、验证、测试模型,以保证模型在未知的数据上测试也能达到如期的效果。目前在数据集上面临的挑战如下:

(1)大数据时代,数据多、杂乱,许多任务都需要特定的数据集;

(2)当前监督学习技术已经取得了巨大的成功,但人工标注数据成本高、耗时耗力;

(3)无监督学习由于学习过程太过困难,发展缓慢。

就有关数据集的问题,本文提出的一个新的思路:如果采

用弱监督学习方式标注数据集,不仅在训练数据标注上能够减少成本,而且扩大了训练所需要的特定数据的范围。文献[69]将弱监督技术应用在目标检测任务,该文中是指不需要任何框注释,仅仅用图像级标签,通过同时训练一个使用图像级数据的弱监督模型和一个监督模型就能实现目标检测,结果显示效果很好。IPRnet^[70] (Inter-pixel Relations Network)使用分类标注做弱监督信息,在整个网络中有两个分支:一个图片分类网络,因为CNN本身就能获得位置信息的class attention map,另外一支则是Inter-pix Relation网络,用来预测物体边界。Deepcluster^[71]则是使用图像级别标签来做语义分割任务,采用区域聚类的方法对这些数据进行聚类。最后将这些聚类完的区域与图片级的标签做对应,得到伪标签。由此可以看到弱监督技术在不断发展。文献[72]提出了一种生成弱监督注释数据的技术,并证明了仅仅使用 bounding box 注释就足以在不使用任何特定对象的边界注释的情况下获得高质量的对象边界。这篇文献主要是针对边缘检测的。文中弱注释的边缘产生思想,即仅仅从随意的 bounding box 注释和各类边缘检测方法的结合取其所共识的部分:在随意 bounding box 注释下的图像中,分别可与 F&H^[73], GrabCut, MCG^[74] 算法或者它们组合的算法结合。例如使用 F&H,在 bounding box 和 F&H 有交叠部分画出边界线,如果没有交叠部分,则这一部分的边界将会被忽视,而得到的数据集是用于边缘检测。在这个思想下,我们只需要来抑制得到的弱注释中背景边界即可得到轮廓的弱注释,但不完全的、不确切的、不精确的监督方式势必造成数据集的不精确,从而也会使得轮廓检测结果不理想。文献[75]提出一种新的衡量股价在某个时间区域内是否出现跳跃性波动的指标——波动次数。在弱监督技术下产生的弱轮廓注释,可以引入波动次数来优化弱注释边界。同时,针对照片拍摄的角度变化、光照变化这些非本质变化,可以利用波动次数的仿射不变性来进行处理,以加强弱注释轮廓数据的信息。这是现在研究中没有涉及的一个部分。

遮挡问题一直是计算机视觉的难题。在目标检测中,行人遮挡检测有了突破。其中文献[76]是对目标候选框分成不同部分逐一处理、区别对待再加以特征融合;而文献[77]从损失函数的角度使得目标候选框对相互遮挡的情况更具判别性。不同于以上两种方法的基于双盒回归的行人检测和遮挡评估^[78] (Bi-box Regression for Pedestrian Detection and Occlusion Estimation, Bi-box)是借助 Fast-RCNN 框架,让网络并行两个分支,分别输出两个框:一个是目标整个部分的候选框,另一个是可见部分的候选框。通过回归这两个框,分别定位行人全身和行人可见部分。这无疑是在给遮挡轮廓检测启示。那么如何很好地利用、借鉴这些相关方向的资源来促进遮挡轮廓检测的进步,也是继续要钻研的课题。

上述与轮廓检测相关的研究方向的各种方法也都显示了它们之间的密切关系。不论是目标检测、实例分割还是语义分割等都可以转化成轮廓检测的问题,而且与之相关的目标轮廓检测也将成为新的发展趋势。根据 3.1 部分的介绍,模型也将由不同的任务要求进行不同程度和不同形式的融合或者分工检测最终融合,以达到更好的效果。

结束语 本文讨论和回顾了目前已有的各种轮廓检测算法,从提取特征特点的角度分了低层、中层和高层 3 个阶段来

介绍,理清了轮廓检测的发展和演变。又从其他的几个方面讨论了轮廓检测和应用,以加深对轮廓检测的理解。然后对轮廓检测的发展趋势和面临的挑战进行分析和预测,最后针对最新轮廓检测算法的缺点和不足提出了自己的解决思路和方法。

参考文献

- [1] MING Y, LI H, HE X, et al. Contour Completion Without Region Segmentation[J]. IEEE Trans Image Process, 2016, 25(8): 3597-3611.
- [2] MARTIN D R, FOWLKESC C, MALIK J, et al. 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.
- [3] ZIOU D, TABBONE S. EDGE Detection Techniques-An Overview[J]. International Journal of Pattern Recognition & Image Analysis, 1998, 8(4): 537-559.
- [4] YANG J, PRICE B, COHEN S, et al. Object Contour Detection with a Fully Convolutional Encoder-Decoder Network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.
- [5] ROTHER C, KOLMOGOROV V, BLAKE A, et al. GrabCut: Interactive foreground extraction using iterated graph cuts[J]. ACM Trans. Graph., 2004, 23(3): 309-314.
- [6] ARBELAEZ P, MAIRE M, FOWLKESC, et al. Contour Detection and Hierarchical Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.
- [7] 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, 2007, 30(1): 36-51.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]//CVPR. IEEE, 2014.
- [9] ZHANG H, JIANG K, ZHANG Y, et al. Discriminative Feature Learning for Video Semantic Segmentation[C]//2014 International Conference on Virtual Reality and Visualization (ICVRV). IEEE Computer Society, 2014.
- [10] WANG Y, ZHAO X, LI Y, et al. Deep Crisp Boundaries: From Boundaries to Higher-Level Tasks[J]. IEEE transactions on image processing: a publication of the IEEE Signal Processing Society, 2019, 28(3): 1285.
- [11] BERTASIUS G, SHI J, TORRESANI L, et al. Semantic Segmentation with Boundary Neural Fields[J]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 3602-3610.
- [12] UIJLINGS J R R, SANDE K E A V D, GEVERST, et al. Selective Search for Object Recognition[J]. International Journal of Computer Vision(IJCV), 2013, 104(2): 154-171.
- [13] ZITNICK C L, DOLLAR P. Edge Boxes: Locating Object Proposals from Edges[C]//European Conference on Computer Vision. Cham; Springer, 2014: 391-405.
- [14] CHEN C, SEFF A, KORNHAUSER A, et al. DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving[C]//ICCV. 2015: 2722-2730.

- [15] LIU H, YAO Y, SUNZ, et al. Road Segmentation with Image-LiDAR Data Fusion[C]//ECCV 2020. 2020.
- [16] ZHANG M, HU X, ZHAOL, et al. Learning Dual Multi-Scale Manifold Ranking for Semantic Segmentation of High-Resolution Images[J]. *Remote Sensing*, 2017, 9(9):500.
- [17] HARIHARAN B, ARBELÁEZ P, GIRSHICK R, et al. Simultaneous Detection and Segmentation[C]//European Conference on Computer Vision. Springer, 2014;297-312.
- [18] HARIHARAN B, ARBELÁEZ P, GIRSHICK R, et al. Hypercolumns for Object Segmentation and Fine-grained Localization [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). 2015:447-456.
- [19] DAI J, HE K, SUN J, et al. Instance-aware Semantic Segmentation via Multi-task Network Cascades[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016;3150-3158.
- [20] KAIMING H, GEORGIA G, PIOTR D, et al. Mask R-CNN[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017;1-1.
- [21] LE C Y, BEN G Y, HINTON G, et al. Deep learning[EB/OL]. (2015-05-28) [2020-09-26]. <https://doi.org/10.1038/nature14539>.
- [22] DAVID A, FORSYTH, JEAN P, et al. *Computer Vision: A Modern Approach*, 2/E[M]//Computer Vision: A Modern Approach. Prentice Hall Professional Technical Reference, 2002.
- [23] CANNY J. A computational approach to edge detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, 8(6):679-698.
- [24] MALIK J, BELONGIE S, LEUNG T, et al. Contour and Texture Analysis for Image Segmentation[J]. *International Journal of Computer Vision*, 2001, 43(1):7-27.
- [25] LIM J J, ZITNICK C L, DOLLAR P, et al. Sketch Tokens: A Learned Mid-level Representation for Contour and Object Detection [C]//2013 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2013.
- [26] PIOTR D, ZITNICK L C. Structured Forests for Fast Edge Detection[C]//IEEE International Conference on Computer Vision (CVPR). IEEE Computer Society, 2013.
- [27] GANIN Y, LEMPITSKY V. N^1 -Fields: Neural Network Nearest Neighbor Fields for Image Transforms[C]//Asian Conference on Computer Vision. Cham; Springer, 2014.
- [28] BERTASIUS G, SHI J, TORRESANI L. DeepEdge: A Multi-Scale Bifurcated Deep Network for Top-Down Contour Detection[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR). 2015;4380-4389.
- [29] KRIZHEVSKY A, SUTSKEVER I, HINTON G, et al. ImageNet Classification with Deep Convolutional Neural Networks [J]. *Advances in Neural Information Processing Systems*, 2012, 25(2).
- [30] SHEN W, WANG X, WANG Y, et al. DeepContour: A deep convolutional feature learned by positive-sharing loss for contour detection[C]//Computer Vision & Pattern Recognition. IEEE, 2015;3982-3991.
- [31] XIE S, TU Z. Holistically-Nested Edge Detection [C]//IEEE International Conference on Computer Vision (ICCV). 2015: 1395-1403.
- [32] EVAN S, JONATHAN L, TREVORD, et al. Fully Convolutional Networks for Semantic Segmentation[M]. *IEEE Computer Society*, 2017.
- [33] MANINIS K K, PONT-TUSET J, PABLO ARBELÁEZ, et al. Convolutional Oriented Boundaries[C]//European Conference on Computer Vision. Springer, 2016;580-596.
- [34] UIJLINGS J, FERRARV. Situational object boundary detection [C]//CVPR. 2015;4712-4721.
- [35] PHILIPP K, KOLTUN V. Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials[C]//Advances in Neural Information Processing Systems. 2011.
- [36] LIU Y, CHENG M M, HU X, et al. Richer Convolutional Features for Edge Detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017;5872-5881.
- [37] WANG Y, ZHAO X, HUANG K, et al. Deep Crisp Boundaries [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2017;3892-3900.
- [38] AKBARINIA A, PARRAGA C A. Feedback and surround modulated boundary detection [J]. *International Journal of Computer Vision*, 2018, 126(12):1367-1380.
- [39] KELM A P, RAO V S, ZOELZERU, et al. Object Contour and Edge Detection with RefineContourNet[J]. *arXiv*:1904.13353, 2019.
- [40] LIN G, MILAN A, SHENC, et al. RefineNet: Multi-path Refinement Networks for High-Resolution Semantic Segmentation [C]//CVPR. 2017;5168-5177.
- [41] POMA X S, RIBA E, SAPPAA D, et al. Dense Extreme Inception Network: Towards a Robust CNN Model for Edge Detection[C]//WACV. 2020;1912-1921.
- [42] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. *arXiv*:1409.1556, 2014.
- [43] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//CVPR. 2015;1-9.
- [44] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]//IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2016;770-778.
- [45] YU F, KOLTUN V. Multi-Scale Context Aggregation by Dilated Convolutions [C]//International Conference on Learning Representations (ICLR). 2016.
- [46] CHOLLET F. Xception: Deep Learning with Depthwise Separable Convolutions[J]. *arXiv*:1610.02357, 2016.
- [47] XIE E, SUN P, SONG X, et al. PolarMask: Single Shot Instance Segmentation With Polar Representation[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.
- [48] TIAN Z, SHEN C, CHEN H, et al. FCOS: Fully Convolutional One-Stage Object Detection[J]. *arXiv*:1904.01355, 2019.
- [49] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2017.
- [50] MÉLY D A, KIM J, MCGILLM, et al. A systematic comparison between visual cues for boundary detection [J]. *Vision Research*. 2016, 120:93-107.
- [51] GRIGORESCU C, PETKOV N, WESTENBERGM A, et al.

- Contour detection based on nonclassical receptive field inhibition [J]. *IEEE Trans Image Process*, 2003, 12(7): 729-739.
- [52] DAVID R M, CHARLESS F, DORON T, et al. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics[C]//*ICCV*. 2001:416-425.
- [53] SILBERMAN N, HOIEM D, KOHLI P, et al. Indoor Segmentation and Support Inference from RGBD Images[C]//*Proceedings of the 12th European conference on Computer Vision-Volume Part V*. Berlin: Springer, 2012.
- [54] MOTTAGHI R, CHEN X, LIU X, et al. The Role of Context for Object Detection and Semantic Segmentation in the Wild [C]//*Computer Vision & Pattern Recognition*. IEEE, 2014.
- [55] EVERINGHAM M, ESLAMI S M A, VAN GOOL L, et al. The Pascal Visual Object Classes Challenge: A Retrospective[J]. *International Journal of Computer Vision*, 2015, 111(1): 98-136.
- [56] HE J Z, ZHANG S L, YANG M, et al. Bi-Directional Cascade Network for Perceptual Edge Detection[C]//*2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2020.
- [57] BERTASIUS G, SHI J, TORRESANI L, et al. High-for-Low and Low-for-High: Efficient Boundary Detection from Deep Object Features and Its Applications to High-Level Vision[C]//*ICCV*. IEEE, 2015.
- [58] HARIHARAN B, ARBELÁEZ P, GIRSHICK R, et al. Simultaneous Detection and Segmentation[J]. *Lecture Notes in Computer Science*, 2014, 8695: 297-312.
- [59] MING M C, YUN L, WENY L, et al. BING: Binarized normed gradients for objectness estimation at 300 fps[J]. *Computational Visual Media*, 2019, 5(1): 4-21.
- [60] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 39(6).
- [61] CARREIRA J, SMINCHISESCUC. Constrained parametric min-cuts for automatic object segmentation[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2011, 34(7): 3241-3248.
- [62] ARBELAEZ P, PONT T J, BARRON J, et al. Multiscale Combinatorial Grouping[C]//*2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2014.
- [63] ROSS B G, JEFF D, TREVOR D, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]//*CVPR*. 2014: 580-587.
- [64] HE K, ZHANG X, REN S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2014, 37(9): 1904-16.
- [65] GIRSHICK R. Fast R-CNN[J]. *Computer Science*, 2015.
- [66] JI F D, YI L, KAIM H, et al. R-FCN: Object Detection via Region-based Fully Convolutional Networks [J]. *arXiv*: 1605.06409, 2016.
- [67] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//*CVPR*. 2016: 779-788.
- [68] LIU W, ANGUELOV D, ERHAND, et al. SSD: Single Shot MultiBox Detector[C]//*European Conference on Computer Vision*. Cham: Springer, 2016.
- [69] WANG J, YAO J, ZHANG Y, et al. Collaborative Learning for Weakly Supervised Object Detection[C]//*Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*. 2018.
- [70] AHN J, CHO S, KWAK S, et al. Weakly Supervised Learning of Instance Segmentation with Inter-pixel Relations[C]//*IEEE/CVF Conference on Computer Vision & Pattern Recognition*. IEEE, 2019.
- [71] B X W A, C H M, D S Y, et al. Deep clustering for weakly-supervised semantic segmentation in autonomous driving scenes [J]. *Neurocomputing*, 2020, 381: 20-28.
- [72] KHOREVA A, BENENSON R, OMRAN M, et al. Weakly Supervised Object Boundaries [C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2016.
- [73] PEDRO F F, DANIEL P H. Efficient graph-based image segmentation[J]. *IJCV*, 2004, 59(2): 167-181.
- [74] PONT T J, ARBELAEZ P, BARRON J, et al. Multiscale Combinatorial Grouping for Image Segmentation and Object Proposal Generation[J]. *arXiv*: 1503.00848, 2015.
- [75] ZHANG Y T, ZHANG Z G. Research on the Correlation between Hot News and Stock Index Volatility Based on Deep Learning. [C]//*The 5th International Conference on Electronics Engineering and Informatics*. ICEEI, 2020.
- [76] ZHANG S, WEN L, BIAN X, et al. Occlusion-aware R-CNN: Detecting Pedestrians in a Crowd[C]//*European Conference on Computer Vision*. Cham: Springer, 2018.
- [77] WANG X L, XIAO T T, JIANG Y N, et al. Repulsion Loss: Detecting Pedestrians in a Crowd[C]//*CVPR*. 2018: 7774-7783.
- [78] ZHOU C L, YUAN J S. Bi-box Regression for Pedestrian Detection and Occlusion Estimation[C]//*ECCV*. 2018: 138-154.



FENG Fu-rong, born in 1986, postgraduate. Her main research interests include object contour detection and object detection.



ZHANG Zhao-gong, born in 1963, Ph.D., professor. His main research interests include bioinformatics, data mining, statistical genetics, big data, cloud computing etc.