

利用可分离卷积和多级特征的实例分割*

王子愉¹, 袁春², 黎健成²

¹(清华大学 计算机科学与技术系, 北京 100084)

²(清华大学 深圳研究生院, 广东 深圳 518000)

通讯作者: 王子愉, E-mail: wangziyu17@mails.tsinghua.edu.cn



摘要: 实例分割是一项具有挑战性的任务,它不仅需要每个实例的边界框,而且需要精确的像素级分割掩码。最近提出的端到端的全卷积实例感知分割网络(FCIS)在检测与分割的结合方面做得很好。但是,FCIS 没有利用低层特征,而低层次的特征信息在检测和分割上都证明是有用的。在 FCIS 的基础上,提出了一种新的模型,充分利用了各层次的特征,并对实例分割模块进行了优化。该方法在检测分支中使用了具有大型卷积核的可分离卷积来获得更精确的边界框。同时,设计了一个包含边界细化操作的分割模块,以获得更精确的掩模。此外,将 Resnet-101 网络中的低级、中级和高级特征组合成 4 个不同级别的新特征,每个新特征都被用于生成实例的掩码。这些掩码被相加之后通过进一步细化以产生最终的最精确的掩模。通过这 3 项改进,实验结果表明,该方法明显优于基线方法 FCIS,相比于 FCIS,该方法在 PASCAL VOC 数据集上的评测指标 $mAP^r@0.5$ 和 $mAP^r@0.7$ 分别提高了 4.9% 和 5.8%。

关键词: 实例分割;可分离卷积;边界细化;多级特征

中图法分类号: TP391

中文引用格式: 王子愉,袁春,黎健成.利用可分离卷积和多级特征的实例分割.软件学报,2019,30(4):954-961. <http://www.jos.org.cn/1000-9825/5667.htm>

英文引用格式: Wang ZY, Yuan C, Li JC. Instance segmentation with separable convolutions and multi-level features. Ruan Jian Xue Bao/Journal of Software, 2019,30(4):954-961 (in Chinese). <http://www.jos.org.cn/1000-9825/5667.htm>

Instance Segmentation with Separable Convolutions and Multi-level Features

WANG Zi-Yu¹, YUAN Chun², LI Jian-Cheng²

¹(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

²(Graduate School at Shenzhen, Tsinghua University, Shenzhen 518000, China)

Abstract: Instance segmentation is a challenging task for it requires not only bounding-box of each instance but also precise segmentation mask of it. Recently proposed fully convolutional instance-aware semantic segmentation (FCIS) has done a good job in combining detection and segmentation. But FCIS cannot make use of low level features, which is proved useful in both detection and segmentation. Based on FCIS, a new model is proposed which refines the instance masks with features of all levels. In the proposed method, large kernel separable convolutions are employed in the detection branch to get more accurate bounding-boxes. Simultaneously, a segmentation module containing boundary refinement operation is designed to get more precise masks. Moreover, the low level, medium level, and high level features in Resnet-101 are combined into new features of four different levels, each of which is employed to generate a mask of an instance. These masks are added and refined to produce the final most accurate one. With the three improvements, the proposed approach significantly outperforms baseline FCIS as it provides 4.9% increase in $mAP^r@0.5$ and 5.8% increase in $mAP^r@0.7$

* 基金项目: 国家自然科学基金(U1833101); 深圳市基础研究资助项目(JCYJ20160428182137473); 腾讯清华大学联合研究中心资助项目

Foundation item: National Nature Science Foundation of China (U1833101); Shenzhen Foundational Research Project (JCYJ2016 0428182137473); The Joint Research Center of Tencent & Tsinghua University

本文由“多媒体数据的知识关联与理解专题”特约编辑蒋树强研究员、刘青山教授、孙立峰教授、李波教授推荐。

收稿时间: 2018-04-23; 修改时间: 2018-06-13, 2018-09-30; 采用时间: 2018-10-31

on PASCAL VOC.

Key words: instance segmentation; separable convolution; boundary refinement; multi-level feature

1 介绍

实例分割是计算机视觉中的一项具有挑战性的任务,它结合了目标检测^[1,2]和语义分割^[3-9]的任务.目标检测要求在实例层级上进行分类和定位,而语义分割在像素层级上进行分类.实例分割不仅需要每个实例的边界框,还需要精确的分割掩码.

实例分割领域的早期方法,如 Deepmask^[10]、Sharpmask^[11]和 InstanceFCN^[12],都使用了 Faster R-CNN 对实例进行分类和定位.它们在兴趣区域(region of interest)的基础上利用全卷积网络生成分割掩码.给定一个图像片段,Deepmask 做两件事:生成一个可以分辨每个像素是否属于某一个实例的掩码;获取表示此图像片段中心是否包含实例的置信度的分数.Deepmask 的关键是同时预测掩码和实例的得分.Sharpmask 通过生成具有不同深度的图像特征的更高质量掩模来改进 Deepmask.它首先从高层次特征生成实例的掩码,并使用较低层次的特征逐步加以细化.InstanceFCN^[12]提出了位置敏感评分图,以保证分割任务的平移可变性.位置敏感评分图被整合以生成每个实例的分割掩码.另一种称为多任务网络级联(MNC)的方法将实例划分为 3 个阶段:包围盒提议、掩码生成和实例分类.每个阶段都以前一阶段的结果作为输入,整个网络是端到端的.

上述方法的一个共同缺点是它们以类不可知的方式分割实例,这可能导致不准确的分割.为了解决上述问题,一种端到端的解决方案 FCIS^[13]被提了出来.FCIS 继承使用了位置敏感评分图的方法,将其扩展到内外位置敏感评分图,充分利用了上下文信息.另一方面,它为分割子任务中的每个类生成掩码候选项,以使最终掩码更准确.基于 FCIS,Biseg^[14]利用语义分割的结果作为实例分割的先验信息.多尺度的位置敏感评分图在 Biseg 中也被证明是有效的.通过这两项技巧,Biseg 比 FCIS 有了显著的提高.

最近,由何明凯提出的 Mask R-CNN^[15]是一种简单、灵活、通用的对象实例分割框架.作者在 Faster R-CNN 中加入了一个分割分支,在检测的同时对实例进行分割.Mask R-CNN 改进了分割的损失函数,由原来常用的基于单像素 softmax 的多项式交叉熵变为了基于单像素 sigmoid 二值交叉熵,二值交叉熵会使得每一类的掩码不相互竞争,而不是与其他类别的掩码比较.另外一点改进在于,作者加入了 RoIAlign 层,利用双线性插值使得兴趣区域的池化结果更加接近未池化前的特征,从而降低了误差.尽管结构简单,但在 FPN^[2]等一系列实用技术的帮助下,Mask R-CNN 取得了良好的效果.Mask R-CNN 需要在至少 8 个 GPU 的实验环境下才能复现出与论文结果相近的表现,GPU 的数量直接影响实验中批处理的数量,而大的批量能够带来稳定的批量归一化(batch normalization)效果,从而带来好的实验结果.受限于实验环境,我们的 GPU 数量目前无法满足要求,因此复现的 Mask R-CNN 效果与原论文相差较大,失去了比较的意义.出于这个原因,我们不能在 Mask R-CNN 上进行实验,只能作为参考,相关实验留待进一步研究.

本文继承使用了位置敏感评分图的思想,提出了一种比 FCIS 更复杂、更精确的模型.我们的方法充分利用了目标检测和语义分割的最新进展,在检测子任务中加入更有效的检测框架,设计了一个具有边界细化模块和多级特征的复杂分割分支.最后,在 PASCAL VOC 数据集上对该方法的性能进行了评价.我们在 $mAP^r@0.5$ 和 $mAP^r@0.7$ 中分别达到 69.1%和 54.9%,显著高于我们的基准线方法 FCIS.本文的主要贡献总结如下.

- (1) 利用具有大型卷积核的可分离卷积改进了检测分支,得到了更精确的包围盒.
- (2) 利用可分离卷积和边界细化操作改进了分割分支,得到了更精确的分割边界.
- (3) 我们的模型融合了 Resnet-101 的低、中、高 3 个层次的特征.实验验证了融合后特征的有效性.

2 相关工作

2.1 全卷积实例感知网络(FCIS)

这项工作是以 InstanceFCN 为基础的.InstanceFCN 引入了位置敏感评分图,在一定程度上保持了平移可变性.InstanceFCN 主要用于生成掩码,但也存在一些不足,如不能判断语义类别,需要后续的网络辅助判别,这意味

着它不是端到端的结构。

基于 InstanceFCN,全卷积实例感知网络(FCIS)被提了出来,InstanceFCN 只有单个对象输出,没有类别信息,需要单独的网络完成类别信息,FCIS 一个网络同时输出实例掩码和类别信息.FCIS 的两个子任务(目标检测和分割)不仅共享卷积特征,而且共享位置敏感评分图。

在实例分割中,相同像素可以是目标的前景或另一目标的背景.仅仅通过为每个类别输出单个得分图来区分这两种情况是不够的.因此,FCIS 的主要创新之一就是提出了内外位置敏感评分图.在 FCIS 中,对于一个类别,每个像素在每个分区中具有两个分数,表示在目标边界内部(或外部的可能性.通过这种方法,FCIS 中位置敏感得分图记录了更多的语义信息并生成了更精确的掩码.输入图像,使用卷积神经网络提取特征,用区域提议网络(RPN)提取候选区域矩形框,计算这些矩形框的部件内外位置敏感评分图,就是部件对象边界内的概率图和对对象边界外的概率图.通过整合得到完整对象的对象边界内的概率图和对对象边界外的概率图,通过逐像素的 softmax 操作得到实例掩码,通过逐像素的 max 操作得到类别概率。

2.2 可分离卷积(seperable convolution)

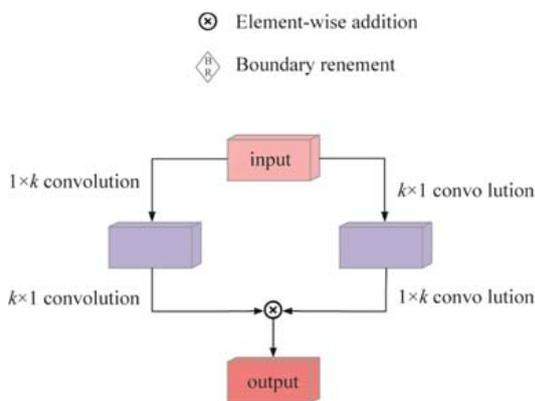


Fig.1 Separable convolution

图 1 可分离卷积

为了减小卷积运算中的参数,可分离卷积的概念被提了出来.利用可分离卷积,卷积运算可以分成多个步骤.假设卷积运算由 $y=conv(x,k)$ 表示,其中,输出图像是 y ,输入图像是 x ,卷积核大小是 k .接下来,假设 k 可以通过以下等式计算: $k=k1.dot(k2)$.这实现了可分离的卷积运算,因为二维卷积运算不是由卷积核大小为 k 的一维卷积运算执行的,而是由卷积核大小为 $k1$ 和 $k2$ 的两个一维卷积运算代替.这样就减少了大量的参数,因为通常 $k1 \times k2$ 比 $k1+k2$ 大得多.可分离卷积如图 1 所示。

在目标检测和语义分割任务中,具有大型卷积核的卷积运算被证明是有效的,因为大的卷积核带来了大而有效的感知域^[7,16,17].对于目标检测和语义分割任务,大的感知域能够使得大物体的信息被完整地传递到更深层

的网络,对于目标包围框的定位更加准确.例如,在后文所示的图 5 中,FCIS 方法检测出来的火车在车尾部分明显丢失了一小段,这就是因为感受野不够大而导致的.用具有大卷积核的可分离卷积扩大感受野后,目标包围框明显扩大,这部分缺失的内容被弥补了回来。

扩大感受野的效果是由进行大卷积核卷积带来的,而大卷积核的卷积操作的问题在于参数会变得非常多,通常计算资源不能满足.可分离卷积的目的在于大幅度减少参数,从而使得大卷积核的卷积操作能够可行,利用可分离卷积,可以用较少的参数和较快的速度进行大型卷积核卷积运算。

2.3 Light-Head RCNN

在流行的两阶段检测框架(如 Faster R-CNN 和 R-FCN)中,第 1 步是生成足够的候选对象,称其为主体;第 2 步是识别候选框,即头部.为了获得最佳精度,头部通常被设计得非常繁重,这需要大量的参数和复杂的计算.为了解决这个问题,使用具有大型卷积核的可分离卷积来生成一个较薄的分数图,称为轻头部.为了减少信息丢失,Light-Head RCNN 删除了 Faster R-CNN 或 R-FCN 中的全局平均池化层.在这之后,就没有办法直接得到最终的结果.然而,通过添加一个全连接层则可以获得分类和回归的最终结果.实验中,我们用 10 代替了原本的类别数,因此降低了位置敏感的兴趣区域池化层和全连接层的计算量.此外,我们使用了具有大卷积核的分离卷积代替 1×1 卷积,将 $k \times k$ 的卷积转化为 $1 \times k$ 和 $k \times 1$,最后加上了一个全连接层作通道变换,再进行分类和回归.通过实验我们还发现,在通道数少的特征图谱上作兴趣区域的池化,不仅准确率提高,还会节约内存及计算量资源.通过这种方法,Light-Head R-CNN 提高了检测的速度和精度。

2.4 边界细化(boundary refinement)

通常大量的物体会有复杂且不规则的边界形状,边界形状的精确预测对于整个实例的分割影响重大.在最近的语义分割工作中,为了提高边缘的分割精度,有研究者提出一种带残差结构的边界细化^[7]模块.这种边界细化模块在初步得到实例的分割掩码之后使用,以进一步拟合实例边界.由于实例分割的分割子分支中使用的方法与语义分割有大量的相似之处,这种方法在实例分割中也是有用的.该模型采用边界细化模块对生成的掩模进行细化.详情如图 2 和图 3 所示.

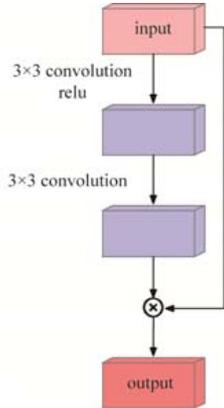


Fig.2 Boundary refinement
图 2 边界细化

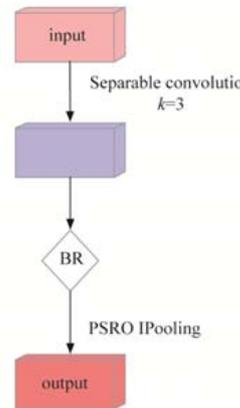


Fig.3 Segmentation module
图 3 分割模块

3 提出的方法

在该模型中,网络输入任意大小的图像并输出实例分割结果.该网络具有多个子网络,它们分享共同的卷积特征:区域提议、边界框回归、实例分割.区域建议子网络根据区域建议网络(RPN)^[1]的工作生成兴趣区域.边界框回归子网络对初始输入的兴趣区域进行细化,以产生更精确的检测结果.实例分割子网络估计实例掩码.我们的模型如图 4 所示.

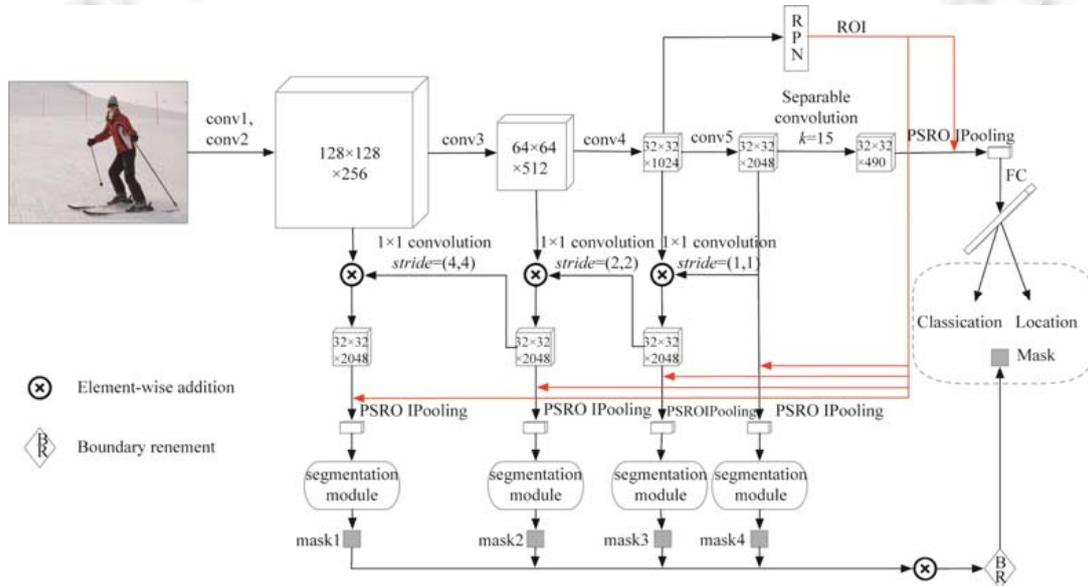


Fig.4 Network architecture
图 4 网络结构

3.1 Resnet-101特征融合

我们使用 ResNet-101^[18]模型作为卷积主干结构进行特征提取。Resnet-101 模型可分为 5 个阶段,每个阶段产生不同深度的特征。特别地,在基准线方法 FCIS 的实现过程中,Resnet-101 的第 5 级特征与第 4 级特征具有相同的大小输出。在我们的模型中,第 2 层、第 3 层和第 4 层的特征通过 1×1 的卷积操作被下采样到与第 5 层相同的大小,然后这些特征被组合成新的 4 层特征,具体的操作方法为,从第 5 层到第 2 层的特征逐层进行对应位置上的元素加法,4 次操作得到 4 层新的特征,这些特征图有着相同的大小和深度,但是特征融合的信息由少到多。每个新的特征图将作为实例分割子网络的输入。

3.2 检测子分支

如上所述,采用 Light Head R-CNN 的结构作为我们的子网络。在 Resnet-101 的 conv5 层特征图上进行大型卷积核的分离卷积,得到轻头部。我们将内核大小设置为 15,这与原始 Light Head R-CNN 中的大小相同。在轻头部中对于每个候选框进行位置敏感的兴趣区域池化,这样,每个候选框的特征就被提取出来。将结果通过一个全连接层转化为 4 096 维的向量,通过全连接层进行分类和边界框回归。

3.3 实例分割子分支

对于第 3.1 节中描述的每个新特征,我们都执行相同的一系列操作,这些操作可以归纳为一个分割模块。分割模块的细节如图 4 所示。与语义分割不同,我们发现具有大型卷积核的可分离卷积在实例分割中并不那么有效,原因是实例分割中每个实例在图像中的面积不一定很大,这导致大的感受野只对大的物体具有积极效应,而对于小的物体无效甚至有负面效果。在我们的实验中,通过多次的实验验证,最终把卷积核大小定为 3,得到了最好的结果。

3.4 损失函数

实验使用的损失函数为实例分割任务中常用的联合训练损失函数:

$$L = L_{cls} + L_{det} + L_{mask},$$

其中, L_{cls} 为目标检测任务中对于兴趣区域目标类别预测的交叉熵, L_{det} 为目标检测中对于兴趣区域的位置修正损失函数, L_{mask} 为语义分割任务中预测的掩码与实际掩码的逐像素交叉熵之和。由于实验采用的是端到端的训练方式,因此这 3 种损失函数同步计算并回传梯度,3 种损失函数直接相加,不进行加权。

4 实验

我们在 VOC 2012 训练集(5 623 张图片)上对模型进行了训练,并对 VOC 2012 验证集进行了评估(5 732 张图片)。我们使用的评估度量是区域上的平均精度,称为 mAP^r 。使用 0.5 和 0.7 的交并比阈值下的 mAP^r 指标,在预测区域和真实区域上计算交并比。

实验环境:显卡为 2×Titan X(Pascal),训练时的学习率设置为 0.000 5,共训练 22 次迭代,训练时的批处理数量为 1,其他实验设置与官方开源的代码相同。

FCIS*:此基准线方法基于 FCIS 的开源实现。由于实验使用的代码是在基于 FCIS 作者开源的官方代码上修改实现的,在运行官方代码时,原代码使用了 4 个 GPU 运行,而我们的实验环境只有 2 个 GPU,前文在介绍中已经解释了 GPU 数量对于实验结果的影响,使用较少的 GPU 数量通常都会带来结果上的略微下降。由于使用的 VOC 2012 数据集图片数量较少,因此,虽然复现的结果略低于论文结果,但是差距较小,有一定的比较意义。实际上,在 Biseg^[14]中,作者复现的实验结果同样略低于论文结果,这并不是由于实验方法或者实验设置上的错误造成的。参考 Biseg^[14],我们将运行 FCIS 的结果标记为 FCIS*,这个结果是由 5 次运行官方代码平均得到的结果,偶然性较小。我们的工作基于此。

我们的方法:与作为对照实验的 FCIS*相同,我们使用在 ImageNet 上预先训练的 ResNet-101 模型,同时并不在我们的模型上应用在线硬示例挖掘(OHEM)。将模型中的 ROI Pooling 和位置敏感 ROI Pooling(PSROI Pooling)

改为 ROIAlign 方法,取得了较好的效果,分别为 64.8%和 51.1%.多层次的特征帮助我们达到了 65.9%和 52.1%的效果.将 Light Head R-CNN 的结构改进到我们的模型中之后,结果分别为 68.6%和 53.8%.最后对网络结构进行了调整,得到了最好的效果,分别为 69.1%和 54.9%.我们将本文模型与最近的方法进行了比较,结果显示在表 1 中.这证明了我们模型的有效性和高效性.

图 5 展示了我们的一些结果.通过边界细化和融合特征,我们的方法在边界线处产生了高质量的掩码.有了更好的检测器,从而避免了一些丢失或冗余多余的部件.

Table 1 Comparison of results on PASCAL VOC2012

表 1 PASCAL VOC2012 实例分割结果比较

Method	$mAP^r@0.5$ (%)	$mAP^r@0.7$ (%)
SDS ^[19]	49.7	25.3
Hypercolumn ^[20]	60.0	40.4
CFM ^[21]	60.7	39.6
MPA ^[22]	61.8	-
MNC ^[23]	63.5	41.5
InstanceFCN ^[12]	61.5	43.0
IIS ^[24]	63.6	43.3
CRF ^[25]	62.0	44.8
FCIS ^[13]	65.7	52.1
Biseg ^[14]	67.3	54.4
FCIS*(baseline)	64.2	49.0
Ours(with roialign)	64.8	51.1
+multi-level features	65.9	52.1
+Light Head R-CNN	68.6	53.8
+boundary refinements	69.1	54.9



Fig.5 FCIS(top) vs. Our Model(bottom). FCIS shows missing, redundancy and rough boundary

图 5 FCIS(上)vs.我们的模型(下).FCIS 表现出缺失、冗余和粗糙的边界

5 结论

实例分割是物体检测+语义分割的综合体.相对物体检测的边界框,实例分割可精确到物体的边缘;相对语义分割,实例分割可以标注出图上同一物体的不同个体.我们提出了一种有效且复杂的端到端全卷积解决方案以解决实例分割问题.该方法利用 Light Head R-CNN 方法对边界盒进行检测,并利用边界细化模块和多级特征来生成分割掩模.在 PASCAL VOC 数据集上取得了 69.0%和 54.9%的最佳结果,显著高于 FCIS 上的结果.在以后的工作中,我们打算在检测分支和分割分支之间建立一个直接连接,以加强它们之间的交流,使共同的特征充分共享.

References:

- [1] Ren XQ, He KM, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems. 2015. 91-99.

- [2] Lin TY, Dollar P, Girshick R, He KM, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proc. of the Computer Vision and Pattern Recognition. 2017,1–4.
- [3] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2015. 3431–3440.
- [4] Chen LC, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. arXiv Preprint arXiv: 1706.05587, 2017.
- [5] Yu F, Koltun V, Funkhouser T. Dilated residual networks. In: Proc. of the Computer Vision and Pattern Recognition. 2017. 1.
- [6] Lin GS, Milan A, Shen C, Reid I. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2017.
- [7] Peng C, Zhang XY, Yu G, Luo GM, Sun J. Large kernel matters-improve semantic segmentation by global convolutional network. arXiv Preprint arXiv: 1703.02719, 2017.
- [8] Wang W, Shen J, Shao L, *et al.* Correspondence driven saliency transfer. IEEE Trans. on Image Processing, 2016,25(11):5025–5034.
- [9] Wang W, Shen J, Shao L. Video salient object detection via fully convolutional networks. IEEE Trans. on Image Processing, 2017, 27(1):38–49.
- [10] Pinheiro PO, Collobert R, Dollar P. Learning to segment object candidates. In: Advances in Neural Information Processing Systems. 2015. 1990–1998.
- [11] Pinheiro PO, Lin TY, Collobert R, Dollar P. Learning to refine object segments. In: Proc. of the European Conf. on Computer Vision. Springer-Verlag, 2016. 75–91.
- [12] Dai JF, He KM, Li Y, Ren SQ, Sun J. Instance-sensitive fully convolutional networks. In: Proc. of the European Conf. on Computer Vision. Springer-Verlag, 2016. 534–549.
- [13] Li Y, Qi HZ, Dai JF, Ji XY, Wei YC. Fully convolutional instance-aware semantic segmentation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2017. 2359–2367.
- [14] Pham VQ, Ito S, Kozakaya T. Biseq: Simultaneous instance segmentation and semantic segmentation with fully convolutional networks. arXiv Preprint arXiv: 1706.02135, 2017.
- [15] He KM, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. In: Proc. of the 2017 IEEE Int'l Conf. on Computer Vision (ICCV). IEEE, 2017. 2980–2988.
- [16] Chollet F. Xception: Deep learning with depth-wise separable convolutions. arXiv Preprint, 2016.
- [17] Li ZM, Peng C, Yu G, Zhang XY, Deng YD, Sun J. Light-head R-CNN: In defense of two-stage object detector. arXiv Preprint arXiv: 1711.07264, 2017.
- [18] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 770–778.
- [19] Hariharan B, Arbelaez P, Girshick R, Malik J. Simultaneous detection and segmentation. In: Proc. of the European Conf. on Computer Vision. Springer-Verlag, 2014. 297–312.
- [20] Hariharan B, Arbelaez P, Girshick R, Malik J. Hypercolumns for object segmentation and fine-grained localization. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2015. 447–456.
- [21] Dai JF, He KM, Sun J. Convolutional feature masking for joint object and stuff segmentation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2015. 3992–4000.
- [22] Liu S, Qi XJ, Shi JP, Zhang H, Jia JY. Multi-scale patch aggregation (MPA) for simultaneous detection and segmentation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 3141–3149.
- [23] Dai JF, He KM, Sun J. Instance-aware semantic segmentation via multi-task network cascades. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 3150–3158.
- [24] Li K, Hariharan B, Malik J. Iterative instance segmentation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 3659–3667.
- [25] Arnab A, Torr PHS. Pixelwise instance segmentation with a dynamically instantiated network. arXiv: 1704.02386, 2017.



王子愉(1996—),男,湖北黄冈人,学士,主要研究领域为计算机视觉.



黎健成(1994—),男,硕士,CCF 学生会员,主要研究领域为计算机视觉.



袁春(1969—),男,博士,副研究员,博士生导师,CCF 高级会员,主要研究领域为机器学习,计算机视觉.

www.jos.org.cn

www.jos.org.cn