

文章编号: 1005-0523(2020)04-0116-09

DOI: 10.16749/j.cnki.jecjtu.2020.04.018

低照度边缘增强的语义分割模型研究

罗 晖, 芦春雨

(华东交通大学信息与通信工程学院, 江西 南昌 330013)

摘要: 为提高对低照度图像的语义分割精度, 提出了一种基于 RPN 的边缘增强语义分割模型(EESN)。在该模型中, 首先利用深度残差网络提取图像的高阶语义特征, 并通过 RPN 快速生成待分割目标候选区域; 然后, 利用设计的融合算法对候选区域进行融合, 并剔除重复的候选区域; 最后, 对融合的目标候选区域做低照度边缘搜索, 并利用失真代价较小的局部增强算法对低照度边缘进行特征增强。将 EESN 用于 Pascal VOC12 和 Cityscapes 两个数据集的语义分割中, 分别获得了 81.2% 和 67.6% 的 mIoU, 该结果证明了 EESN 对具有低照度边缘的图像具有较好的分割性能。

关键词: 语义分割; 候选区域网络; 低照度; 边缘检测; 局部增强

中图分类号: TP391

文献标志码: A

本文引用格式: 罗晖, 芦春雨. 低照度边缘增强的语义分割模型研究[J]. 华东交通大学学报, 2020, 37(4): 116-124.

Citation format: LUO H, LU C Y. Research on semantic segmentation model of low-illumination edge enhancement[J]. Journal of East China Jiaotong University, 2020, 37(4): 116-124.

语义分割是将整个图像分割成多个像素组的机器视觉任务。基于深度学习的图像语义分割网络是近几年的研究热点, 其分割精度和处理效率都明显优于传统分割模型^[1]。文献[2]提出全卷积网络(FCN, fully convolutional network), 实现了对输入图像的密集语义分割预测, 证明了卷积神经网络(CNN, convolutional neural network)在语义分割中的可行性。相较此前的分割方法, FCN 分割效果更好, 但其分割图仍较粗略。此外, CNN 的空间不变性, 会造成特征图位置信息的丢失, 影响分割。为此, 文献[3]提出了 SegNet, 通过引入跳跃卷积, 并将由池化索引得到稀疏特征图进行解码得到分割结果, 提高了分割速率。文献[4]将 CNN 的最后一层输出与条件随机场相结合来改善分割精度, 并利用膨胀卷积(atrous convolution)增加卷积的特征感受野, 改善分割效果。随后, 为了进一步提升基于 CNN 的语义分割模型的分割精度和分割效率, 文献[5]引入空间金字塔池(ASPP, atrous spatial pyramid pooling)结构对 DeepLab V1 进行了改进; 文献[6]中, 将 Xception 中可分卷积结构引入 ASPP 中作为编码, 并用膨胀卷积的输出编码特征的分辨率来平衡编码-解码结构的精度和运行时间。文献[7]将条件独立的假设添加到全连接的 CRF 中, 并与卷积神经网络结合, 加快了网络的训练和推断速度。文献[8]在深度特征输出上施加空间金字塔注意力结构, 并结合全局池化策略学习出更好的特征表征。文献[9]将深度特征卷积之后, 通过上采样与浅层特征相乘, 获取融合特征图, 以改善分割效果。

以上文献介绍的语义分割模型虽然都在一定程度上对以往的模型进行了改进, 但针对光照较弱背景下的图像语义分割的研究并不多。为此, 在 FCN 的基础上, 从局部特征增强的角度, 引入一种局部增强算法来增强低照度边缘特征, 并提出低光照边缘增强的语义分割模型(EESN, semantic segmentation model with low-illumination edge enhancement), 以提高对低照度图像语义分割精度。

收稿日期: 2019-11-01

作者简介: 罗晖(1969—), 男, 教授, 研究方向为多媒体信息处理, 无线通信系统, 无线多媒体传感器网络, 压缩感知。

1 基本理论

在本节中更详细地回顾了用于语义分割的残差网络(ResNet,residual network)^[10]的结构,并讨论专为生成高质量区域建议而设计的区域提议网络(RPN,region proposal network)^[11]。

1.1 残差网络

CNN 可被视为由浅层网络和深层网络共同组成,且深层网络是浅层的等价映射。浅层网络的输出 x 为深层网络的输入,且对应的深层网络的输出为 $H(x)$ 。因为深层网络是浅层网络的等价映射,则有

$$\lim(x)=H(x) \quad (1)$$

而直接学习该恒等映射比较困难。因而,将重点放在深层网络和浅层网络的差值即残差 $F(x)$ 分析

$$F(x)=H(x)-x \quad (2)$$

这样学习恒等映射式(1)则可以转换为式(3)的学习

$$\lim F(x)=0 \quad (3)$$

而对深度卷积神经网络而言, $F(x)$ 随 x 的变化幅度远远大于 $H(x)$ 随 x 的变化幅度。根据 $F(x)$ 的变化来调节训练中网络参数,能够有效地避免因神经网络层数过深而带来的梯度消失或梯度爆炸问题。

ResNet 的提出使得训练更加深层的卷积网络成为可能,进而使学习图像中更深层的语义信息成为可能。

1.2 区域建议网络

在 RPN 之前,常用的候选区域块生成算法包括选择性搜索(SS,selective search),Bing,Edge b-Boxes 等。这些算法首先生成一系列候选目标区域,然后利用 CNN 提取目标候选区域的特征,再根据学习到的特征做后续处理,这需要耗费大量的计算量。

为减少候选区域生成的耗时,提出了 RPN 来生成目标候选区域,具体的过程如图 1 所示。

在 RPN 中,首先根据设定的长宽比和尺度在特征图像以滑动生成不同尺度的锚点,然后根据对这些锚点的二分类结果去除负样本锚点,最后基于锚点与真实目标边框做初步回归,得到较高质量的建议区域。使用 RPN 作为 EESN 的区域建议网络能够减少计算量,同时保证其分割的效率。

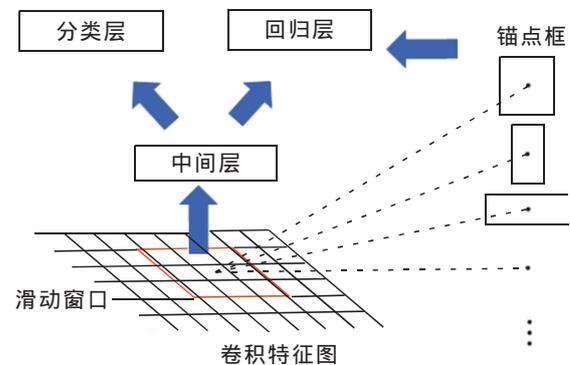


图 1 基于 RPN 目标候选区域生成示意图
Fig.1 Schematic diagram of candidate region generation based on RPN

2 低照度边缘增强语义分割

低照度场景下,光靠调节拍摄装置的曝光率仍旧不能完全解决某些区域出现的模糊和信息丢失的问题。为了克服图像中低照度边缘的识别困难,将一种局部增强算法引入语义分割网络,其结构模型如图 2 所示。

利用 EESN 模型实现语义分割的具体过程如下:首先利用卷积层提取输入图像的语义特征;然后通过 RPN 生成目标区域建议,再利用融合层将重叠区域融合;在融合形成的兴趣域基础上,通过搜索窗口检测是否包含低照度边缘,并根据增强算法对检测到的低照度边缘进行局部增强;增强之后的特征图输入联合池化层,再由上采样层进行上采样后,得到与原图等大的预测结果。

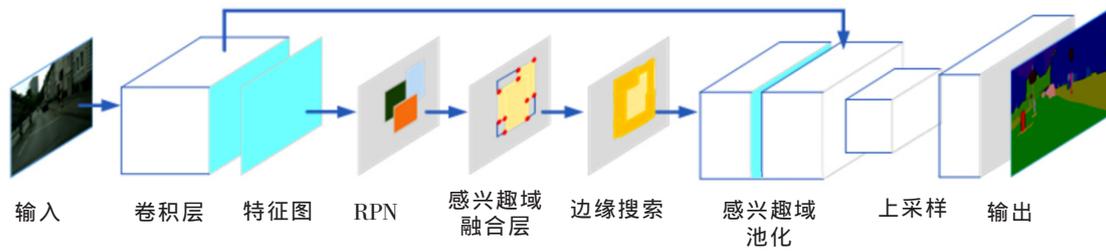


图2 EESN示意图
Fig.2 Schematic diagram of EESN

2.1 基于 RPN 的区域建议生成

EESN 采用 ResNet-101 作为主干网,对输入图像进行卷积特征提取,并将其最后卷积层的输出特征图输入到 RPN 中。RPN 在输入特征图上滑动产生锚点,并采用与 Faster R-CNN 相同的锚点生成方式,即面积有三种(128*128,256*256,512*512),长宽比有三种(1:1,1:2,2:1);用来采样的锚点的 IoU 阈值为 0.7;每张图最终得到约 300 个区域建议。通过 RPN 生成的区域建议具有较高质量,这些区域建议是由具有相似特征的像素聚合而成的同质块,它们反映了图像中重要的统计特征。与单个和孤立的像素相比,区域建议中可以提取更有意义的统计特征,同时能够保留目标的原始真实边界信息,进而更利于提高语义分割的鲁棒性。

2.2 区域建议融合

RPN 所生成的区域建议的集合 R 为

$$R = \{r_1, r_2, r_3, \dots, r_n\} \tag{4}$$

真实的语义分割区域集合 GR 为

$$GR = \{g_1, g_2, g_3, \dots, g_n\} \tag{5}$$

其中: n 为图像真实存在的语义类别的个数; r_i 为第 i 个语义类中实例的区域建议集合; g_i 为第 i 个语义类别的真实边界框。为了减少后续低照度边缘搜索的计算量,需要去除冗余的区域建议;同时,为提高边缘搜索的效率,将属于同一实例的区域建议进行融合,以保证被搜索的区域包含分割实例的完整轮廓信息。所提出的融合算法的概念图如图 3 所示。

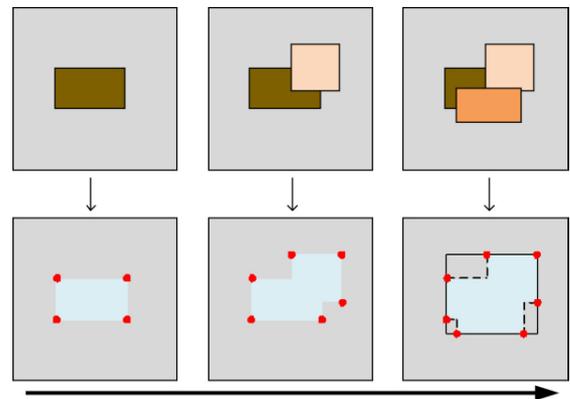


图3 融合算法概念图
Fig.3 Conceptual diagram of fusion algorithm

区域建议融合层的算法流程如下:

输入:区域建议集合 R 以及融合后区域集合 F

$$F = \{f_1, f_2, f_3, \dots, f_n\} \tag{6}$$

其中

$$f_i = \{x_i^1, y_i^1, x_i^2, y_i^2\} \tag{7}$$

f_i 为 r_i 中置信度最高的建议区域,其位为

$$r_{ic} = \{x_{ic}^1, y_{ic}^1, x_{ic}^2, y_{ic}^2\} \tag{8}$$

1) 初始化

$$i = 1 \tag{9}$$

2) 以原特征图右下角角点为原点,水平向右方向为 x 轴,垂直向上方向为 y 轴,将区域建议的坐标 r_{ik} 初始化为

$$r_{ik} = \{x_{ik}^1, y_{ik}^1, x_{ik}^2, y_{ik}^2\} \tag{10}$$

其中: $(x^1, y^1), (x^2, y^2)$ 分别对应区域建议的左上角和右下角;

3) 对第 i 类语义区域建议中剩余的区域建议 t_i 进行遍历,其中

$$t_i = r_i - r_{ic} \tag{11}$$

比较 t_{ik} 与 f_i ,并对 f_i 更新,更新原则如下

$$\begin{cases} f_i \leftarrow f_i \cup \{x_{ik}^1, y_{ik}^1, x_{ik}^2, y_{ik}^2\} \\ x_i^1 \leftarrow \max(x_{ik}^1, x_i^1); y_i^1 \leftarrow \max(y_{ik}^1, y_i^1) \\ x_i^2 \leftarrow \max(x_{ik}^2, x_i^2); y_i^2 \leftarrow \max(y_{ik}^2, y_i^2) \end{cases} \tag{12}$$

4) 重复第二步,直到遍历完所有语义类;

5) 开始下一个语义兴趣域的融合,令

$$i = i + 1 \tag{13}$$

6) 重复 2) 到 5),直到 $i = n$;

输出:融合后兴趣域 F 。

通过以上过程,能够将存在重叠区域的区域建议融合成数量与图像中语义类别数量相同的兴趣区域。

2.3 低照度边缘搜索及局部增强

融合形成的兴趣域中可能存在低照度边缘,这些边缘会影响分割效果,因此需要对兴趣域进行低照度边缘搜索及特征增强。图像的边缘信息可以体现在特征图中梯度值的变化上,因此,这里用一个 3×3 窗口对融合后的兴趣域进行滑动搜索,搜索步长设为 2,滑动搜索窗口 S 如图 4。

分别从 -45°, 0°, 45°, 90° 四个方向对兴趣域 F 对应的特征图 I 进行滑动梯度检测,并对检测到的低照边缘特征进行局部增强,该过程的具体算法流程如下:

S ₁₁	S ₁₂	S ₁₃
S ₂₁	S ₂₂	S ₂₃
S ₃₁	S ₃₂	S ₃₃

图 4 滑动搜索窗口
Fig.4 Sliding search window

输入:融合后兴趣域 F 映射的特征图区域为

$$I = \{I_1, I_2, I_3, \dots, I_n\}, I_i = \{p_{11}, p_{12}, p_{13}, \dots, p_{mk}\} \quad (14)$$

其中: m, k 分别对应特征区域 I_i 中特征点的横、纵坐标;

1) 初始化

$$i=1 \quad (15)$$

2) 滑动窗口遍历第 i 个特征区域 I_i 中的特征点,即

$$S_{22}=p_{mk} \quad (16)$$

3) 计算特征点 p_{mk} 邻域的四个方向梯度

$$\begin{cases} 0 < \|s_{11}-s_{33}\| = s_{-1/2} < \delta \\ 0 < \|s_{21}-s_{23}\| = s_0 < \delta \\ 0 < \|s_{31}-s_{13}\| = s_{1/2} < \delta \\ 0 < \|s_{32}-s_{12}\| = s_{\perp} < \delta \end{cases}, \quad (17)$$

其中: δ 是设定的低照度边缘梯度边缘的最大值,超过 δ 则视为非低照度边缘; $s_{-1/2}, s_0, s_{1/2}, s_{\perp}$ 分别表示 S_{22} 领域(如图 5 所示)中

的 -45°, 0°, 45°, 90° 四个方向上的梯度值;

4) 将搜索到的低照度区域进行增强

$$\xi = \max(s_{-1/2}, s_0, s_{1/2}, s_{\perp}) \quad (18)$$

其中 ξ 用来记录该区域的四个方向上梯度变化最大的梯度值,即边缘信息最明显的特征差;

5) 对最大梯度变化方向的正交方向做强差变换,并将搜索框对应的中心值做等值处理,例如,当 $\xi = s_{-1/2}$ 时,该区域的梯度特征被变化为

$$s_{31}=s_{13}=s_{22}=\min(s_{11}, s_{33})+\delta \quad (19)$$

6) 更新到下一个特征区域的遍历,令

$$i=i+1 \quad (20)$$

7) 重复 2) 到 5), 直到 $i=n$;

输出:边缘增强后的特征图 I 。

通过上述流程后,卷积特征图的局部边缘特征将被增强,同时,特征图中边缘噪声能够被很好的抑制。

3 实验及分析

该部分首先介绍了用于实验的数据集、实验设备和设备配置以及用于评价的语义分割模型性能的指标。然后,展示了 EESN 模型在两个典型语义分割数据集(Pascal VOC12 和 Cityscapes)上的分割实验结果。最后,将所提出的 EESN 同其他语义分割方法在以上几个数据集的基础上进行了性能对比实验,并给出实验数据。

3.1 数据集及硬件配置

Pascal VOC12 数据集包含了 20 个室内和室外目标类别以及一个背景类别。实验中采用了该数据集 中的 10 582 幅图像进行训练,1 449 幅图像进行验证,1 456 幅图像进行测试。

Cityscapes 数据集是一个通过车载摄像机采集到的大型城市街道场景数据集。它包含 5 000 张经过精细注释的图片。实验中使用该数据集中 19 种目标类别和一个背景类别的 2 975 张图像用于训练,500 张图像用于验证,1 525 张图像用于测试。

实验在配有 16 GB 内存、Intel i5-7600 处理器和两张 GT1080Ti GPU 显卡的图形工作站上进行。工作站同时安装了 CUDA 9.0 和 CuDNN 7.0。

3.2 评价指标

为了充分分析 EESN 模型的分割性能,除平均交并比(mIoU, mean intersection over union)之外,还引入

了标记精度(TA, tagging accuracy)、定位精度(LA, locating accuracy)和边界精度(BA, boundary accuracy)三个指标对 EESN 的分割结果进行评价, 其中:TA 用于评价预测的像素级标签与场景真实值标签之间的差异, 能够反映模型对包含多种语义类别图像的分类准确性;LA 定义为目标的预测边界框与真实边界框之间的 IoU, 用于估计模型对图像中目标定位的精度;BA 表示正确定位目标的预测语义边界与实际语义边界的差值, 它用于反映网络的语义分割精度。

3.3 实验结果

用 COCO 2014 对 EESN 进行预训练, 并选取 Pascal VOC12 中的 20 个类图像对 EESN 进行训练和测试, 预训练和训练的迭代次数分别为 150k 和 30k, mini-batch 的大小设置为 10, 学习率设置为 0.01。表 1 为 EESN 对 Pascal VOC12 中 20 类目标的分割结果的 TA, LA 和 BA 指标。

表1 EESN 模型在 Pascal VOC12 数据集上的分割评价
Tab.1 Evaluation of segmentation results of EESN on VOC12 dataset

数据类型	areo	bike	Bird	boat	bottle	bus	car	cat	chair	cow	table
TA	97.8	98.1	97.8	98.3	97.1	98.8	97.1	96.2	93.5	97.8	96.1
LA	82.6	78.5	79.1	72.3	62.1	89.5	75.5	87.2	55.8	81.3	62.5
BA	96.1	84.2	97.2	94.1	94.5	94.8	96.3	96.2	91.6	94.3	92.7
数据类型	dog	horse	mbike	person	plant	sheep	sofa	train	tv	Avg	
TA	96.5	96.7	98.3	93.5	96.5	98.3	93.1	97.8	96.6	96.8	
LA	86.7	82.5	82.3	78.5	58.5	81.9	72.6	78.3	62.8	75.6	
BA	95.9	95.7	93.8	95.4	92.1	95.6	92.3	95.9	97.2	94.3	

表 1 中, EESN 模型对 Pascal VOC12 数据集上图像分割结果的 TA, LA, BA 3 个评价的平均值分别为 96.8%, 75.6% 和 94.3%, 该结果表明 EESN 模型能够较好地对 Pascal VOC12 数据集中的图像进行语义分割, 即 EESN 模型能够实现图像语义分割。在此基础上, 为证明 EESN 模型能够提高语义分割精度, 进行了以下实验。

3.3.1 Pascal VOC12 数据集上实验结果

为进一步验证 EESN 模型对提高语义分割精度的有效性, 将 EESN 模型与其他多种分割模型 (包括 Zoom-out^[12]、DeepLab V2^[5]、EdgeNet^[13]、BoxSup^[14]和 Higher-order CRF^[15]) 的分割性能进行了比较。不同模型的分割结果的评价结果如表 2 所示。

由表 2 可知, EESN 模型在 Pascal VOC12 测试集上的分割结果的 mIoU 为 80.5%, 高于其他网络。此外, EESN 对交通场景中目标的语义分割性能良好 (如 boat, bus, car, mbike, train 等), 而这些目标极易受光照不均或光照过低影响, 与场景中其他目标混淆。该实验证明了 EESN 模型对低照度图像的语义分割性能良好。

3.3.2 Cityscapes 数据集上实验结果

为避免实验结果的偶然性, 将 EESN 模型与上个实验中用到的几种分割模型对 Cityscapes 数据集中图像进行了语义分割, 并统计出不同模型的分割结果的 mIoU 指标, 结果如表 3 所示。

表 3 中, EESN 模型的分割结果的 mIoU 达到了 67.6% 的, 高于其他几种模型。同时也能发现 EESN 模型对 Cityscapes 数据集中 bus, car, road, train 和 truck 等目标的分割性能较好, 而这些目标同样具有易受光照影响而与周围目标混淆的特性。

表 2 Pascal VOC12 数据集上的分割结果比较
 Tab.2 Comparison of segmentation results evaluation on Pascal VOC12 dataset

数据类型	Zoom-out	EdgeNet	Box-Sup	Higer-order CRF	DeepLab V2	EESN
areo	85.6	86.1	89.8	90.3	91.2	87.8
bike	37.3	38.9	38.1	39.3	55.3	58.7
bird	83.2	78.8	89.2	89.1	88.6	86.8
boat	62.5	60.1	68.9	67.8	70.6	71.3
bottle	66.1	67.2	67.9	69.3	75.3	72.2
bus	85.1	87.2	89.6	90.1	90.8	91.6
car	80.7	81.2	82.9	83.3	85.8	87.6
cat	84.9	82.6	97.7	84.6	86.3	85.1
chair	27.3	37.1	34.4	39.5	35.1	34.8
cow	73.2	78.5	83.6	80.1	86.7	81.2
table	57.5	62.4	67.1	68.7	70.2	72.3
dog	78.1	80.6	81.5	81.7	83.2	82.1
horse	79.2	80.3	83.7	85.4	86.9	86.8
mbike	81.1	83.6	85.2	87.3	87.8	89.6
person	77.1	80.2	83.5	87.2	89.3	85.8
plant	53.6	56.3	58.6	65.7	66.8	68.7
sheep	74	79.7	84.9	85.1	88.6	85.3
sofa	49.3	52.9	55.8	62.3	62.8	61.1
train	71.6	78.4	81.2	84.2	83.3	85.2
tv	63.3	68.5	70.7	75.3	73.8	70.2
mIoU	69.6	73.6	75.2	77.9	79.7	80.5

表 3 Cityscapes 数据集上分割结果的比较
 Tab.3 Comparison of segmentation results evaluation on Cityscapes dataset

数据类型	Zoom-out	EdgeNet	BoxSup	Higer-order CRF	DeepLab V2	EESN
bike	61.3	62.5	66.8	65.9	64.2	63.1
build	85.2	86.7	89.2	92.3	91.2	90.1
bus	42.1	43.5	48.6	50.9	50.5	51.2
car	82.3	87.8	92.6	91.7	92.8	93.6
fence	40.1	42.6	44.2	48.9	46.7	48.1
mbike	48.2	55.3	51.6	54.8	54.3	53.9
person	72.3	78.9	79.2	78.7	77.1	78.1
pole	41.5	42.7	47.4	53.8	55.9	56.1
rider	48.7	50.3	55.3	53.2	51.4	54.6
road	89.8	92.1	93.3	97.1	96.8	97.4
sign	58.9	60.4	65.1	68.3	69.1	67.9
swalk	70.3	73.5	78.4	79.3	79.1	78.7
sky	86.3	88.6	91.3	92.6	93.9	90.1
terrain	62.5	65.4	72.1	70.3	69.3	70.8
tlight	53.5	56.7	61.3	62.9	60.1	61.8
train	42.6	48.3	48.9	54.3	56.5	56.2
turck	32.3	33.8	38.5	41.5	35.6	42.3
veg	83.9	87.6	89.4	90.7	91.4	88.3
wall	32.5	41.8	36.7	41.3	34.9	40.9
mIoU	62.5	63.1	64.8	66.4	67.1	67.6

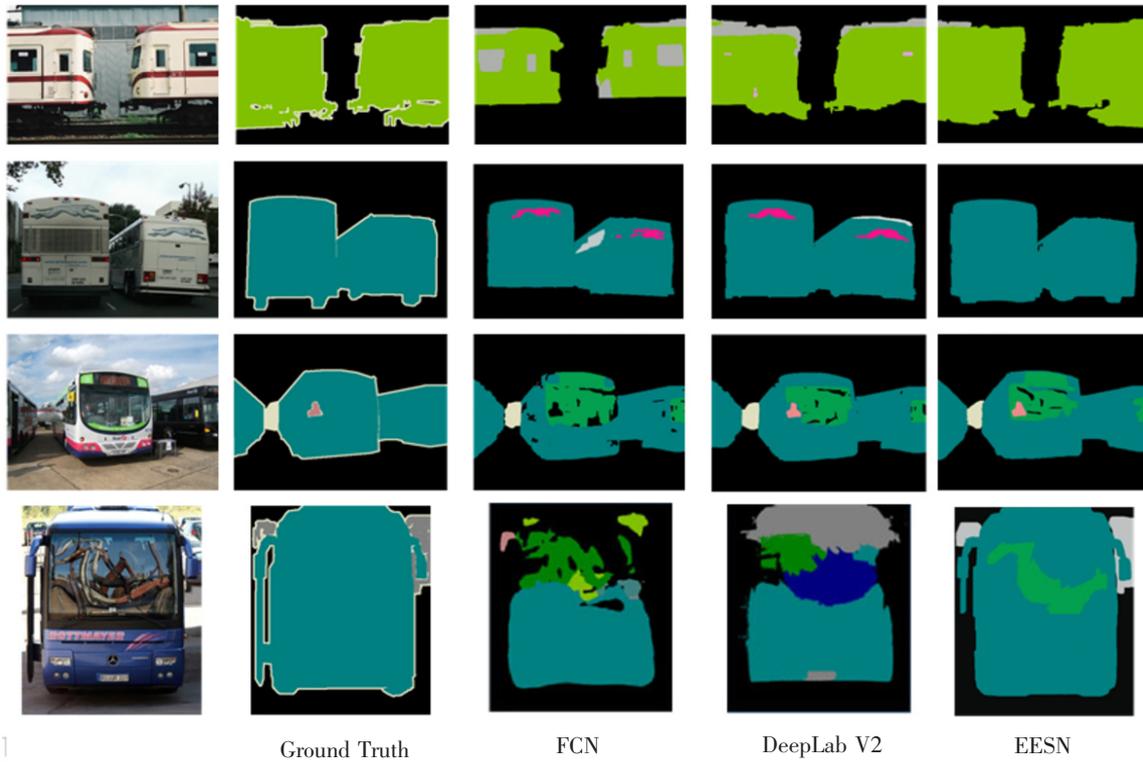


图5 Pascal VOC12数据集上几种语义分割模型的分割视觉质量图对比

Fig.5 Comparison of visual quality maps of several semantic segmentation models on Pascal VOC12 dataset

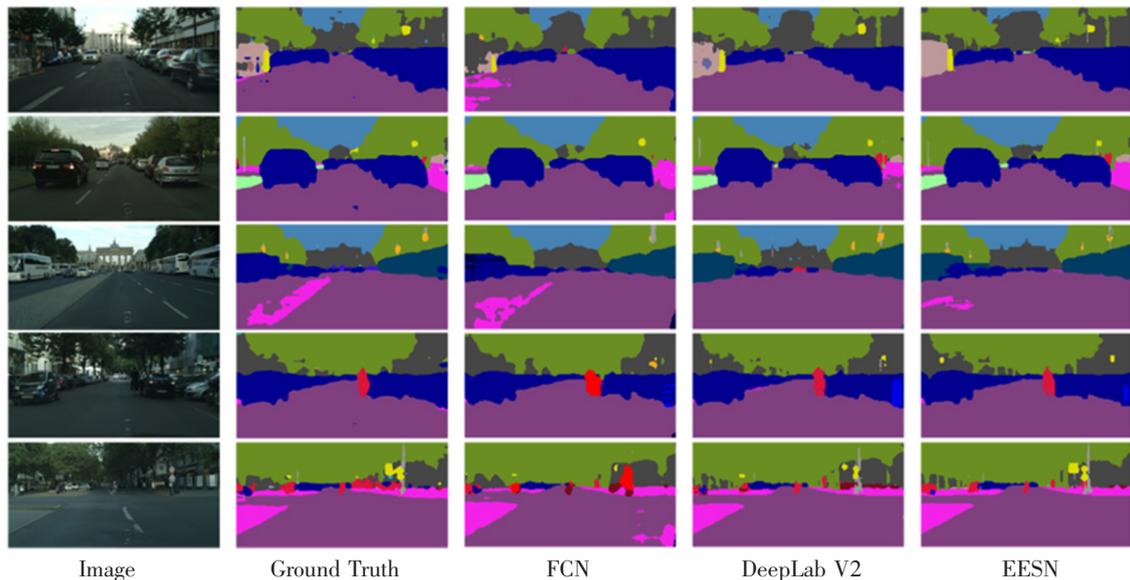


图6 Cityscapes数据集上几种语义分割模型的分割视觉质量图对比

Fig.6 Comparison of visual quality maps of several semantic segmentation models on Cityscapes dataset

最后,为了更直观地观察 EESN 的分割性能,图 5,图 6 分别给出了从 Pascal VOC12 数据集和 Cityscapes 数据集中选取的几幅具有代表性的图像,以及使用 FCN、DeepLab V2 和 EESN 对这些图像的分割效果图。

从图 5 中可以看出,EESN 模型对暗区域中公交车车轮的分割效果优于 FCN 模型和 DeepLab V2 模型的分割效果(第 2 行)。另外,图 5 中第 4 行,在大巴挡风玻璃上由于反射率不同导致的复杂阴面的场景下,EESN 模型对大巴的分割结果的完整性仍然得到了保证,并优于另外两个模型。对于 Cityscape 数据集,从图 6

的整体来看,EESN模型对近地面暗区于的车轮分割效果明显优于FCN和DeepLab V2。另外,图6第2行中,EESN模型能够很好地分割出处于车辆与树间暗区域中的行人;第3行中,EESN模型对站在车辆左暗区域中司机腿部的分割结果也优于其他两个模型。

在本小节中,首先,EESN模型在Pascal VOC12数据集上的分割实验及其结果分析证明了EESN模型的有效性。其次,通过对EESN与几种典型语义分割模型在Pascal VOC12和Cityscapes两个数据集上的分割实验结果分析,证明了对含有低照度边缘图像的分割任务而言,EESN模型比其他几种模型的分割效果更好,并在一定程度上提升了分割精度。可以推断出,所提出的低照度边缘增强算法是合理、有效的,并且EESN模型对提高图像分割精度是有效的,特别是对含有低照度边缘的图像。

4 结束语

为了更加准确地描述语义分割任务中目标边缘特征,提高模型对低照度图像的分割精度,本文提出了EESN模型,其优点如下: EESN以深度残差网络为主干网,可以保证模型对图像特征的学习效率; EESN利用RPN生成高质量的区域建议,可加快网络的运算速度; EESN采用融合算法对候选区域块进行融合,剔除了重复的候选区域,可提高网络的运算速度和分割的精度; EESN采用局部增强算法,有针对性地强化低照度边缘特征,可进一步提高网络对图像特征的描述能力,进而提高网络分割性能。通过比较EESN与几种典型语义分割模型在Pascal VOC12数据集和Cityscapes数据集的分割,证明了EESN能够提升对低照度图像的分割精度。

参考文献:

- [1] 梁新宇,罗晨,权冀川,等.基于深度学习的图像语义分割技术研究进展[J].计算机工程与应用,2020,56(2):18-28.
- [2] LONG J,SHELHAMER E,DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2017,39(4):640-651.
- [3] BADRINARAYANAN V,KENDALL A,CIPOLLA R. SegNet:A Deep Convolutional Encoder-Decoder Architecture for Scene Segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2017,39(12):2481-2495.
- [4] CHEN L C,PAPANDREOU G,KOKKINOS I,et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. Computer Science,2014(4):357-361.
- [5] CHEN L C,PAPANDREOU G,KOKKINOS I,et al. DeepLab:semantic image segmentation with deep convolutional nets,atrous convolution and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2017,40(4):834-848.
- [6] CHEN L C,PAPANDREOU G,SCHROFF F,et al. Rethinking Atrous Convolution for Semantic Image Segmentation[J/OL]. arXiv:1706.05587.
- [7] CHEN L C,ZHU Y,PAPANDREOU G,et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Proceedings of the European conference on computer vision (ECCV),2018:801-818.
- [8] TEICHMANN M T T,CIPOLLA R. Convolutional CRFs for semantic segmentation[J/OL]. arXiv:1805.04777.
- [9] LI H,XIONG P,AN J et al. Pyramid attention network for semantic segmentation[J/OL]. arXiv:1805.1018.
- [10] ZHANG Z,ZHANG X,PENG C et al. Exfuse:enhancing feature fusion for semantic segmentation[J/OL]. arXiv:1804.03821.
- [11] HE K,ZHANG X,REN S et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society,2016.
- [12] REN,SHAOQING,HE,KAIMING,GIRSHICK,ROSS et al. Faster r-CNN:towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2015,39(6):1137-1149.
- [13] MOSTAJABI M,YADOLLAHPOUR P,SHAKHNAROVICH G. Feedforward semantic segmentation with zoom-out features[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE,2015.
- [14] ZITNICK C L,DOLLAR P. Edge boxes:locating object proposals from edges[C]//ECCV. Springer,Cham,2014.
- [15] ARNAB A,JAYASUMANA S,ZHENG S et al. Higher order conditional random fields in deep neural networks[J]. Computer Science,2015,9906:524-540.
- [16] 洪维伟,王元庆.基于神经网络与马尔可夫链的换道次数预测方法[J].华东交通大学学报,2019,36(2):92-98.
- [17] 郑晓芳,邱运霞,傅军栋,等.基于图形识别和GRNN网络的照明设计自动化系统[J].华东交通大学学报,2019,36(5):82-90.

Research on Semantic Segmentation Model of Low-Illumination Edge Enhancement

Luo Hui, Lu Chunyu

(School of Information Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: In order to improve the accuracy of semantic segmentation of low-illumination images, an edge enhanced semantic segmentation model (EESN) based on RPN is proposed. Firstly, higher-order semantic features of images were learned by using a deep residual network, and the region proposals were quickly generated by RPN. Secondly, a fusion algorithm was designed to fuse the region proposals and eliminate the repeated candidate regions. Finally, low-illumination edges were searched by the edge search box, and their features were enhanced by using the local enhancement algorithm with low distortion. EESN was used in semantic segmentation of Pascal VOC12 dataset and Cityscapes dataset, and 81.2% and 67.6% of pixel intersection-over-union (mIoU) on these two datasets were obtained. The experimental results also demonstrate that EESN has good segmentation performance for images with low illumination edge.

Key words: semantic segmentation; RPN; low-illumination; edge detection; local enhancement

(上接第 108 页)

Application of Variable Weighted Multivariate Statistical Process Control Method in Transformer State Evaluation

Wang Xun¹, Wang Yuxin¹, Kang Chen², Wan Hua², Cheng Hongbo¹

(1. School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China;

2. State Grid Jiangxi Electric Power Research Institute, Nanchang 330096, China)

Abstract: According to the characteristics of different correlation and importance between the variables of transformer state evaluation, the weighted Mahalanobis distance is introduced into the traditional multivariate statistical control chart based on Mahalanobis distance, and a transformer evaluation method based on the variable weighted multivariate statistical process control is designed. This method combines the weight of transformer state variable given by association rule method, and uses the wave detection ability of multivariate control chart to transform the multivariate real-time monitoring data corresponding to various faults of transformer into a test statistic, and then evaluates the operation state of transformer. The example analysis of the current circuit over-heat fault shows that compared with the current transformer condition evaluation standard, the method has faster fault detection speed; the retrospective detection of the sample data of known faults shows that the method has higher fault detection rate.

Key words: multivariate statistical process control; association rule; weighted Mahalanobis distance, transformer operation status