

基于多尺度特征提取的图像语义分割

熊志勇 张国丰 王江晴

(中南民族大学 计算机科学学院, 武汉 430074)

摘要 指出了基于深度学习的图像语义分割中,如何充分利用图像上下文信息以达到更好的分割效果,是当前图像语义分割研究的关键问题.为解决这一问题,提出了一种基于多尺度特征提取的图像语义分割方法,通过构建深层卷积神经网络,并利用不同尺度图像作为网络的输入来提取不同尺度图像的特征,最后经过特征融合得到了分割图.在公开数据集 Stanford background dataset 8 类数据集上进行训练和验证,实验结果达到了 84.33% 的准确率.实验表明:通过提取和融合多尺度特征,可以达到更好的图像语义分割效果.

关键词 图像语义分割;多尺度特征;深度学习;卷积网络

中图分类号 TP183 文献标识码 A 文章编号 1672-4321(2017)03-0118-07

Multi-scale Feature Extract For Image Sematic Segmentation

Xiong Zhiyong Zhang Guofeng Wang Jiangqing

(School of Computer Science, South-Central University for Nationalities, Wuhan 430074, China)

Abstract In the semantic segmentation based on deep learning, how to make full use of the image context information to achieve better segmentation results is the key problem of image semantic segmentation. In this paper, a method of image semantic segmentation based on multi-scale feature extraction is proposed. This method extracts the features of different scale images by constructing deep convolution neural network and using different scale images as network input. Finally, a segmentation map is obtained by feature fusion. The experiments on stanford Background Dataset-8 verify the effectiveness of the proposed method, and the accuracy rate is 84.33%. Experimental results show that better image semantic segmentation can be achieved by extracting and merging multiscale features.

Keywords image semantic segmentation; multi-scale feature; deep learning; convolution network

图像语义分割结合了图像分类和图像分割,通过一定的方法将图像中的每个像素分为不同的语义类别,最终得到不同的分割区域.图像语义分割被广泛用于无人驾驶中的街景识别、目标检测,无人机落地点检测,场景理解等许多计算机视觉任务方面的应用.从基于计算机视觉的机器学习方法到目前基于深度学习的方法,图像语义分割算法的研究得到了很大的进展,但是由于工业需求的不断加大,图像语义分割依然是计算机视觉任务中的研究热点之一.

图像语义分割是一项非常具有挑战性的任务,主要体现在:物体层次的不确定性,这是因为图像的明亮程度、模糊程度、图像中物体的大小、方向等诸多因素的影响;物体类别层次的模糊性,同一个物种的不

同类别,也不好分别开.为了减少不确定性和模糊性等外界因素的影响,充分利用图像中的信息(如图像像素值,从中可以提出许多特征,如颜色特征;图像中像素、物体之间的联系等上下文信息),获得更好的特征表示,是一个重要的方法.

早期的图像语义分割是利用手工标注特征,如方向梯度直方图 HOG^[1]和尺度不变特征转换 SIFT^[2].基于机器学习的方法从最简单的像素级别阈值法、基于像素聚类的分割方法到基于图论划分的分割方法^[3].其中基于条件随机场的分割方法^[4-8]表现较为优异.基于机器学习的方法,先将输入图像分为一些独立的区域块,并提取每个区域块的特征,然后根据一定的规则建立图像特征与语义类别之间的概率模

收稿日期 2017-06-06

作者简介 熊志勇(1965-)男,教授,研究方向:信息隐藏、图像处理 E-mail: zhiyx@scuec.edu.cn

基金项目 国家自然科学基金资助项目(60975021)

型,建立起能量函数,并通过手工标注的特征库,迭代计算对能量函数进行优化,得到最优的参数,最终得到图像语义分割模型。这种机器学习的方法过于依赖手工标注的特征库,难以广泛表示图像特征,在实际应用中有很大的局限性。

近年来,深度卷积神经网络(DCNNs)^[9]的发展,计算机视觉任务中的许多问题得到了巨大的突破。由于深度卷积网络可以从大量样本数据中提取出图像的特征,比手工标注特征更好,在图像分类^[10-12]和物体检测^[13-16]等高层计算机视觉任务上获得了巨大的成功。卷积神经网络在图像分类等高层次计算机视觉任务上的出色表现,以图像语义分割为代表的计算机低层次任务的研究也逐渐利用卷积网络实现。Schulz H 等^[17]利用 CNN 产生粗糙的输出图,并施加成对类别定位过滤器,实现图像分割,比传统的计算机视觉方法速度更快效果更好。Farabet C 等^[18,19]为了充分利用图像的上下文信息,提出了多尺度 CNN 的图像语义分割算法,首先利用金字塔池化得到原始图像的三个不同尺度特征,输入到 CNN 中学习对应输出三个尺度特征图,并组合通过优化函数实现图像分割。Dai J 等^[20]提出了通过 CNN 进行目标区域特征提取算法并交叉训练实现图像分割。Long J 等^[21]提出了全卷积网络用于图像语义分割,把传统的图像分类网络直接变为像素分类网络。蒋应锋等^[22]提出了多尺度迭代训练,对每个像素进行分类并通过超像素描绘分割图的轮廓。刘丹等^[23]利用超像素中心产生不同尺度的图像作为输入,提出了多尺度 CNN 模型结构用于图像语义分割。Mostajabi M 等^[24]利用超像素为基本单元,提出缩放特征的图像语义分割模型。Lin G 等^[25]通过多尺度图像作为输入产生不同尺度的特

征图。

应用 DCNN 的图像语义分割模型,可以通过大量的样本自主学习特征,打破了手工标注特征的局限性,可以实现非常好的分割效果。但是基于 DCNN 的图像语义分割要考虑两个关键的因素:1) 由于 DCNN 中池化层的出现最后输出图分辨率较小,构建端到端网络训练模型完成图像分割,有利于提高分割精度;2) 如何考虑图像的上下文信息,物体与物体之间,物体与背景之间的关系,依然是关键。本文主要解决如何充分利用图像的上下文信息,并通过 DCNNs 提取出丰富特征用于图像语义分割。最近多尺度输入训练 CNNs 用于图像语义分割的算法^[19,24]达到了良好的效果。本文提出一种利用多尺度图像作为输入,通过全卷积网络^[21]进行特征提取并完成图像语义分割的模型。

1 本文算法

本方法首先利用深度卷积网络进行特征提取,为了利用图像的多尺度信息,将图像缩放为三个不同的尺度,作为网络的输入;为了最后输出为原始图像大小的分割图,把网络的最后的全连接层改为全卷积层,最后进行上采样获得原始大小的分割图,这种网络通常被称为全卷积网络(FCN)^[21]。在训练阶段,每个尺度的图像通过网络都会产生一个得分图,通过上采样将它们缩放为相同大小,通过多尺度融合层生成输出图,再与分割标签计算损失,再反向传播完成一次训练;测试阶段,每个尺度的图像通过网络生成一个得分图,通过多尺度融合算法生成输出图。算法的基本流程图如图 1 所示。

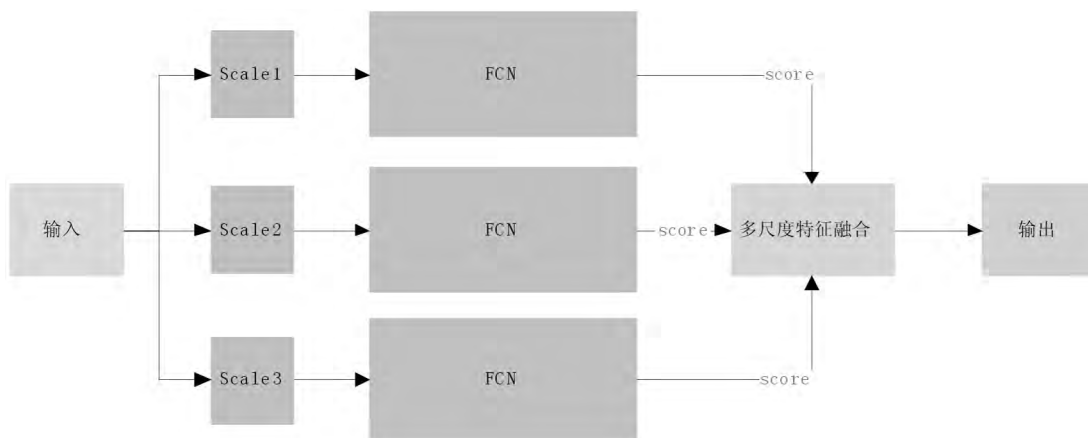


图 1 算法流程图

Fig. 1 The flow chart of the algorithm

1.1 全卷积网络

全卷积网络(FCN)实际上是把普通卷积网络的最后一层的全连接层变为卷积层,因为图像语义分割是输出的二维图像,全连接层会把空间信息隐藏,输出展开为一维向量,因此通过卷积操作来保留空间信息.文献[21]通过修改16层的分类网络VGG-16^[11],将最后的全连接层改为卷积层,并修改输出为21通道的二维图像,可以用分类网络训练好的模型进行初始化参数,并在此基础上训练优化,可以达到非常好的分割效果.

深度神经网络利用信息的冗余,可以通过池化(pooling)来减少计算量,经过池化后图像的分辨率会下降,如文献[21]中最后的输出是原始图像的1/32倍,这是需要使用上采样(Upsampling)或者反卷积(Deconvolution)来增加图像的分辨率,保证输出图与原始图像大小相同.公式(1)、(2)说明了卷积操作和反卷积计算的过程.假设输入图像大小为 $w_0 \times h_0$,经过卷积计算后的图像大小为 $w_1 \times h_1$,填充值大小padding,卷积核大小kernelsize,滑动窗大小stride. w 和 h 的计算相同,因此以 w 计算为例.

$$w_1 = (w_0 + 2 \times padding - kernelsize) / stride + 1, \quad (1)$$

$$w_0 = (w_1 - 1) \times stride + kernelsize - 2 \times padding. \quad (2)$$

反卷积计算往往是利用双线性插值法,这样就会丢失一些图像的空间信息,导致分割图像非常粗糙,FCN的作者提出了一种跳跃网络的结构,通过融合不同层的信息来弥补信息的丢失,如FCN-8s,通过融合pool4和pool3的输出和最后一层的输出产生最后的分割图.

1.2 网络模型结构

本文的网络是全卷积网络的扩展,用共享网络的结构替换跳跃网络结构,通过不同尺度图像输入产生不同尺度特征图来保证输出图的效果.网络的结构如图2所示,首先将输入图像缩放为3个不同的大小,放入卷积网络中,3个卷积网络在训练时共享网络结构如图2,浅色方框围着的区域,分别产生3个得分图,通过多尺度融合层fuse使之融合成一张得分图,这样利用不同尺度的图像,可以更好地感知图像中的丰富空间信息.与传统的FCN的不同之处在于多尺度图像的输入以及多尺度特征图的提取和融合.

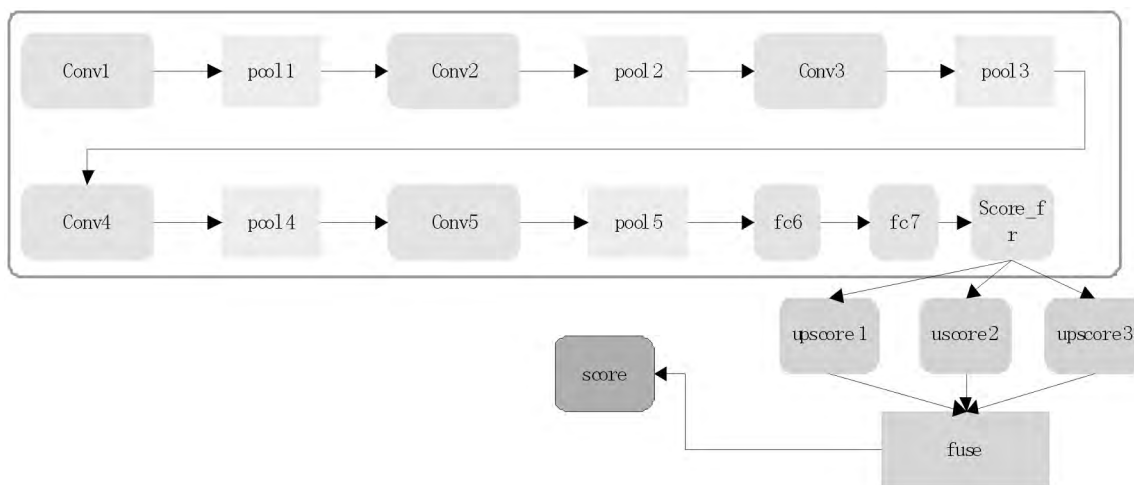


图2 本文所用网络结构

Fig. 2 Network structure of this paper

对于每个尺度的图像,将其放入图2的网络中训练学习,其中网络的组成为:5个卷积层(Conv1、Conv2、Conv3、Conv4、Conv5),每个卷积层后面用一层最大池化层(pool1、pool2、pool3、pool4、pool5)来减少计算量,卷积网络中的全连接层(fc6、fc7)变为卷积层,然后卷积输出8个特征图(数据集中有8类物体),即Score_fr层,此时的特征图大小是原始图像大小的1/32倍,需要通过公式(2)的反卷积运算将

特征图像变为原始图像大小(upscore1、upscore2、upscore3).每一层卷积层后都用ReLU(Rectified-Linear Units)函数激活,主要特点是简单效果好、网络收敛速度快.不同尺度对应了不同的输出特征图,经过融合函数fuse可以得到最终的分割图.

给定输入尺度为 s 的图像 I_s ,所要训练网络的参数为 $\theta_s = (w^l, b^l)$,其中 $l = 1, 2, \dots, n$ 为网络层次. w 表示网络的权重, b 表示偏置,假设经过 l 层卷积、

池化操作后的输出为 O_s^l , 则有公式 (3):

$$O_s^l = \text{pool}(\text{ReLU}(w^l \times O_s^{l-1} + b^l)). \quad (3)$$

从输入图像 I_s 开始, 依次计算网络中每个神经元的输出值, 直到计算出最后的得分图为止, 这个过程被称为前向传播 (Forward Propagation). 最后一层的输出 O_s^L , 需要 *softmax* 激活函数来进行归一化处理, 如公式 (4):

$$O_s^L = \text{softmax}(w^L \times O_s^L + b^L). \quad (4)$$

在训练的过程中, 还需要从输出结果向输入图像的方向计算网络中每个神经元的实际输出和正确输出之间的误差并利用随机梯度下降算法 (SGD) 更新网络中的参数, 这一过程被称为反向传播 (Backward Propagation).

通常有了网络计算的输出图, 和正确的标签分割图, 可以计算出每个像素被分为每一类的概率. 设像素为 a_i 其中 $i = 1, 2, \dots, N$ 表示像素的索引, p_{ij} 表示像素 a_i 被分到标签 c_j 的概率, 其中 $j = 1, 2, \dots, C$ 表示物体的类别. 则损失函数可以表示为公式 (5):

$$e = - \sum_{i=1}^N \sum_{j=1}^C 1\{y_i = j\} \ln p_{ij}, \quad (5)$$

其中 y_i 为像素 a_i 经过网络输出的标签, 整个训练的过程就是要优化这个损失函数.

1.3 多尺度特征提取和融合

长期的实践证明, 多尺度特征在计算机视觉任务上非常有用, 在物体识别和图像分类、图像分割方面应用广泛. 如何提取和利用图像的多尺度信息在图像语义分割领域依然是研究的热点问题. 本文利用 Lin 等提出的缩放输入图像来感知图像的多尺度方法^[25]. 将图像缩放为 1, 0.6, 1.2 三个不同尺度, 作为网络的输入, 分为三个尺度可以更好地感知过大或过小的物体. 由网络分别产生 3 个不同尺度的得分图, 经过融合得到最终的输出图, 公式 (3) ~ (5) 给出了单个网络的计算和误差函数, 利用图 2 的网络结构提取不同尺度的特征图, 对于多个尺度特征图的融合, 利用层 fuse 来计算, 具体过程如下所述.

假设输入图像 I 缩放为 S 个输入图像, 用 $f_s(I_s, \theta_s)$ 表示图像 I_s 在第 s 个网络上的输出得分图, 其中 $s = \{1, 2, \dots, S\}$. 这 S 个得分图感知了不同尺度信息, 这时利用公式 (6) 的尺度融合公式将它们融合得到最终的输出结果图 $G(I, C)$, 其中 C 表示语义类别标签.

$$G(I, C) = \sum_{s=1}^S (w_s \times f_s(I_s, \theta_s)), \quad (6)$$

其中权重 w_s 表示第 s 个尺度图像产生的得分图的重

要性, 本文中设置为 1/3, 表示不同尺度的得分图之间等价, 融合只需要线性叠加产生最终的输出图.

2 实验

实验采用深度学习框架 Caffe 进行模型的训练和测试, 数据集来自 Stanford Background Dataset 8 类数据集, 此数据集由已知数据集如 PASCAL VOC、MSRC 中挑选出了 715 张户外场景图像. 图像包含原始图像和图像语义标签文件, 语义分为: 天空 (sky)、树 (tree)、道路 (road)、草地 (grass)、水 (water)、建筑物 (building)、山 (mountain) 和前景目标 (foreground) 共 8 类, 分别对应语义标签 0 ~ 7, -1 表示未知目标类别.

2.1 实验设计

数据集: 首先, 准备实验数据集, 采用 Stanford background dataset 8 类数据集, 随机挑选出 600 张图像用于训练, 115 张图像用于验证和测试.

训练网络和测试网络的搭建: 利用 Caffe 深度学习框架提供的接口, 并根据 FCN^[21] 提供的 VGG-16 源码构建自己的 CNN 网络, 输入层改为 3 个不同尺度的图像作为输入, 输出层添加尺度融合层. 网络的基本结构如图 2 所示.

训练: 实验用 VGG-16 分类网络的训练好的模型进行初始化网络参数, 基础学习率设置为 $1e - 10$, 冲量设置为 0.99, 权重衰减设置为 0.0005, 迭代一次变化一次学习率, 用反向传播和随机梯度下降 (SGD) 来计算误差并更新参数.

评价指标: 为了评价图像语义分割算法的分割结果的精度, 本文采用平均像素精度 (MPA)、平均类别精度 (MCA) 和平均 IOU (mIOU) 3 个指标进行衡量和对比. 假设总共有 C 类语义标签, n_{ij} 表示真实语义类别为 i 但是被预测为 j 的像素个数, $i, j = 0, 1, \dots, C - 1$, t_i 表示真实语义类别为 i 的像素个数, 则有以下计算公式:

$$MPA = \frac{\sum_i n_{ii}}{\sum_i t_i}, \quad (7)$$

$$MCA = \frac{1}{C} \sum_i \frac{n_{ii}}{t_i}, \quad (8)$$

$$mIOU = \frac{1}{C} \sum_i \frac{n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}}. \quad (9)$$

代码实现: 利用 Caffe 深度学习框架提供的接口, 搭建了训练和验证用的 DCNN 和配置运行文件, 可以

在 <https://github.com/abumaster/SemanticSegmentation> 下载使用.

2.2 实验结果

随着训练迭代次数的增加,图像语义分割的精度也在不断提高,变化趋势如图3所示.在迭代训练12000次后精度变化平缓,趋于收敛.如表1所示为迭代训练12000次后的每个语义类的正确率(对角线为正确分类的语义类,其他为混淆率),可以很好地把不同的语义类区别开,部分分割结果图如图4所示.

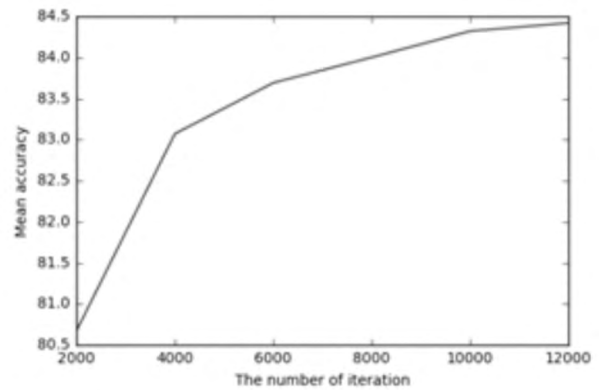


图3 训练精度随着迭代次数的变化

Fig. 3 The training accuracy changing with the number of iteration

表1 每类语义正确率矩阵

Tab. 1 per-class accuracy

Accuracy	Sky	Tree	Road	Grass	Water	Building	Mountain	Foreground
Sky	91.85	3.92	0.00	0.56	2.25	2.03	1.09	1.85
Tree	3.86	76.12	0.75	3.57	0.64	7.05	5.37	3.14
Road	0.00	0.37	94.30	3.52	1.81	0.68	1.20	4.25
Grass	0.25	0.99	0.58	88.39	0.83	0.21	4.27	3.14
Water	0.18	0.00	1.79	0.31	85.38	0.00	3.43	1.21
Building	3.17	8.53	0.93	0.81	0.49	82.49	0.41	6.55
Mountain	1.48	1.27	0.66	1.86	4.64	0.52	27.06	1.00
Foreground	2.48	5.26	5.49	6.04	4.97	6.91	1.58	74.87

为了对比实验结果,本文方法与文献[22]和文献[23]提出的方法进行了对比,结果如表2所示,它们都是利用CNN进行特征提取并完成图像语义分割和理解任务.文献[22]利用图像块作为输入,进行特征提取并得出语义类别,然后利用超像素修正错误的语义类别;文献[23]则是利用超像素块为基本处理单元,首先把图像分为超像素块,然后输入不同尺度的超像素斑块进行训练,对超像素块进行语义分类以完成图像的语义分割.

2.3 分析和讨论

对以上实验结果和对比实验结果进行分析和讨论:从表1中图像中像素大致能分类为正确语义标签,在山(mountain)的语义类正确率偏低,是因为在Stanford background dataset数据集中,有山的训练图像偏少,因而准确率偏低.

与文献[22]和文献[23]的实验结果对比中,本文方法在各项衡量指标中明显优于其他两种,这主要有两方面的原因,一个是因为CNN的深度,本文方法利用了更深层次的CNN可以提取和发现更多的图像特征;另一个原因是本文方法提取了不同尺度图像的特征,并将之融合训练,使得在语义类别边界可以达到很好的分割效果.

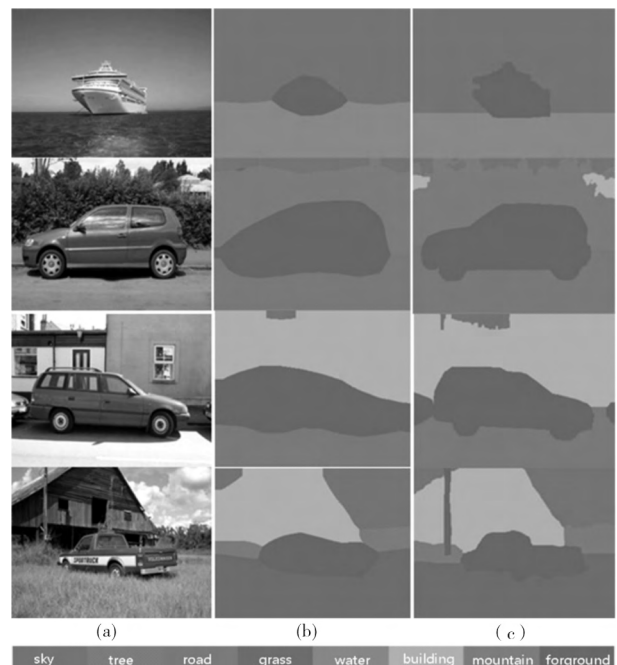


图4 多尺度特征提取的图像语义分割结果图

Fig. 4 Multi-scale feature extract for image semantic segmentation results

表 2 3 种实验方法对比

Tab. 2 Comparison of the three experimental methods

方法	MPA/%	MCA/%	mIOU/%
文献[22]	77.40	69.50	-
文献[23]	84.52	76.60	37.15
本文方法	84.33	77.55	66.98

3 结语

本文提出了一种利用不同尺度图像作为输入的 CNN 深度学习模型,通过多次的迭代训练,可以提取出图像的多尺度的特征信息,并使之融合以完成图像语义分割的任务. 在 Stanford background dataset 公开数据集上的实验结果表明,不同尺度的图像特征对图像语义分割的准确性有着至关重要的影响,结合了多尺度特征提取,可以正确对像素进行分类并感知物体的轮廓信息. 但是,本文方法在描绘分割图像轮廓边界方面仍有不足,对于小的物体还存在难以识别的情形. 为了更好地描绘边缘和融合不同尺度特征,未来的工作会考虑加入条件随机场和边缘检测模型,考虑将尺度融合改为可学习的权重,从而提高图像语义分割的准确性.

参 考 文 献

- [1] Kadota R, Sugano H, Hiromoto M, et al. Hardware architecture for HOG feature extraction [C]//IEEE. Intelligent Information Hiding and Multimedia Signal Processing. Piscataway: IEEE, 2009: 1330-1333.
- [2] Zhou H, Yuan Y, Shi C. Object tracking using SIFT features and mean shift[J]. Computer vision and image understanding, 2009, 113(3): 345-352.
- [3] Rother C, Kolmogorov V, Blake A. Grabcut: Interactive foreground extraction using iterated graph cuts [C]//ACM. transactions on graphics (TOG). New York: ACM, 2004, 23(3): 309-314.
- [4] Krahenbühl P, Koltun V. Efficient inference in fully connected crfs with gaussianedgepotentials [C]//Advances in neural information processing systems. Massachusetts: MIT Press, 2011: 109-117.
- [5] He X, Zemel R S, Carreira-Perpinán M A. Multiscale conditional random fields for image labeling[C]//IEEE. Proceedings of the 2004 IEEE computer society conference on Computer vision and pattern recognition. Piscataway: IEEE, 2004: 695-703.
- [6] He X, Zemel R S, Ray D. Learning and Incorporating Top-Down Cues in Image Segmentation [J]. Lecture notes in computer science, 2006: 338-351.
- [7] Gould S, Fulton R, Koller D. Decomposing a scene into geometric and semantically consistent regions [C]//IEEE. 12th International Conference on Computer Vision. Piscataway: IEEE, 2009: 1-8.
- [8] Zheng S, Cheng M M, Warrell J, et al. Dense semantic image segmentation with objects and attributes [C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 3214-3221.
- [9] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [10] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [C]//Advances in neural information processing systems. Massachusetts: MIT Press 2012: 1097-1105.
- [11] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J/OL]. (2015-04-10). <https://arxiv.org/abs/1409.1556>.
- [12] Papandreou G, Kokkinos I, Savalle P A. Modeling local and global deformations in deep learning: Epitomic convolution, multiple instance learning, and sliding window detection [C]//IEEE. Proceedings of the Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 390-399.
- [13] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//IEEE. Proceedings of the conference on computer vision and pattern recognition. Piscataway: IEEE 2014: 580-587.
- [14] Erhan D, Szegedy C, Toshev A, et al. Scalable object detection using deep neural networks [C]//IEEE. Proceedings of the Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 2147-2154.
- [15] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks [C]//Advances in neural information processing systems. Massachusetts: MIT Press, 2015: 91-99.
- [16] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//IEEE. Proceedings of the Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.
- [17] Schulz H, Behnke S. Learning object-class segmentation with convolutional neural networks [C]. Bruges: i6doc.com Press, 2012: 151-156.
- [18] Farabet C, Couprie C, Najman L, et al. Scene parsing

- with multiscale feature learning, purity trees, and optimal covers [J/OL]. (2012-07-13). <https://arxiv.org/abs/1202.2160>
- [19] Farabet C, Couprie C, Najman L, et al. Learning hierarchical features for scene labeling [J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 35(8): 1915-1929.
- [20] Dai J, He K, Sun J. Convolutional feature masking for joint object and stuff segmentation [C]//IEEE. Proceedings of the Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3992-4000.
- [21] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C]//IEEE. Proceedings of the Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3431-3440.
- [22] 蒋应锋, 张桦, 薛彦兵, 等. 一种新的多尺度深度学习图像语义理解方法研究[J]. 光电子·激光, 2016, (02): 224-230.
- [23] 刘丹, 刘学军, 王美珍. 一种多尺度 CNN 的图像语义分割算法[J]. 遥感信息, 2017, (01): 57-64.
- [24] Mostajabi M, Yadollahpour P, Shakhnarovich G. Feed-forward semantic segmentation with zoom-out features [C]//IEEE. Proceedings of the Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3376-3385.
- [25] Lin G, Shen C, van den Hengel A, et al. Efficient piecewise training of deep structured models for semantic segmentation [C]//IEEE. Proceedings of the Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 3194-3203.

(上接第 105 页)

参 考 文 献

- [1] 吴迪, 朱青松. 图像去雾的最新研究进展[J]. 自动化学报, 2015, 41(2): 221-239.
- [2] Tan R T. Visibility in bad weather from a single image [C]//IEEE. 2008 Computer Vision and Pattern Recognition. Alaska: IEEE, 2008: 1-8.
- [3] Nishino K, Kratz L, Lombardi S. Bayesian Defogging [J]. International Journal of Computer Vision, 2012, 98(3): 263-278.
- [4] Fattal R. Single image dehazing [J]. ACM Transactions on Graphics, 2008, 27(3): 1-9.
- [5] He K, Sun J, Tang X. Single image haze removal using dark channel prior [C]//IEEE. Computer Vision and Pattern Recognition. Miami: IEEE, 2011: 1956-1963.
- [6] Tarel J P, Hautiere N, Caraffa L, et al. Vision Enhancement in Homogeneous and Heterogeneous Fog [J]. Intelligent Transportation Systems Magazine IEEE, 2012, 4(2): 6-20.
- [7] McCartney E J, Jr F F H. Optics of the Atmosphere: Scattering by Molecules and Particles [J]. Journal of Modern Optics, 1977, 14(7): 521-521.
- [8] 李锋, 阚建霞. 基于 Sobel 算子的图像快速二维最大熵阈值分割算法[J]. 计算机科学, 2015, 42(6A): 209-210 + 220.
- [9] Chu H L, Li Y X, Zhou Z M, et al. Optimized fast dehazing method based on dark channel prior [J]. Acta Electronica Sinica, 2013, 41(4): 791-797.
- [10] 刘书琴, 毋立芳, 宫玉, 等. 图像质量评价综述 [J]. 中国科技论文, 2011, 06(7): 501-506.
- [11] 李大鹏, 禹晶, 肖创柏. 图像去雾的无参考客观质量评测方法 [J]. 中国图象图形学报, 2011(9): 1753-1757.