



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2016
01
VOL.21

ISSN1006-8961
CN11-3758/TB



医学图像处理 P69

中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）



第21卷第1期（总第237期）
2016年1月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 李小文
编辑出版《中国图象图形学报》编辑出版委员会
邮政信箱 北京9718信箱
邮 编 100101
电子信箱 jig@radi.ac.cn
电 话 010-64807995
网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

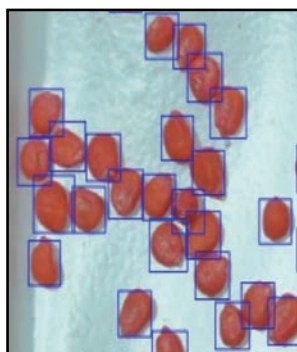
Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
P.O.Box 9718, Beijing, P.R.China
Zip code 100101
E-mail jig@radi.ac.cn
Telephone 010-64807995
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



多尺度特征融合与极限学习机的玉米种子检测
(第0024页)



深度卷积神经网络的显著性检测(第0053页)



双树复小波域的MRI图像去噪(第0104页)

图像处理和编码

空时形状预测与高效编码

朱仲杰, 王玉儿, 蒋刚毅..... 1

第二小方向导数信息熵的兴趣点检测

卢健, 黄杰, 潘峰..... 8

采用面积采样的反走样字符旋转算法

郑凯文, 刘文波..... 17

图像分析和识别

多尺度特征融合与极限学习机的玉米种子检测

柯道, 杜明智..... 24

图像理解和计算机视觉

结合NSCT和压缩感知的红外与可见光图像融合

陈木生..... 39

Plücker直线描述的空间后方交会

盛庆红, 陈姝文, 肖晖, 张斌, 王青, 费利佳..... 45

深度卷积神经网络的显著性检测

李岳云, 许悦雷, 马时平, 史鹤欢..... 53

计算机图形学

以优先点为中心的Delaunay三角网生长算法

尤磊, 唐守正, 宋新宇..... 60

医学图像处理

快速数字影像重建的2维/3维医学图像配准

刘坤, 吕晓琪, 谷宇, 于荷峰, 任国印, 张明..... 69

小波-Lagrange方法进行医学图像层间插值

武士想, 尚鹏, 王立功..... 78

遥感图像处理

多主体框架下结合最大期望值和遗传算法的SAR图像分割

张金静, 李玉, 赵泉华..... 86

第十届图像图形技术与应用学术会议

视差互信息引导下的立体航空影像与LiDAR点云自动配准

吴军, 胡彦君, 饶云, 彭智勇..... 95

双树复小波域的MRI图像去噪

黄学优, 张长江..... 104

快速定位图像尺度和区域的3维跟踪算法

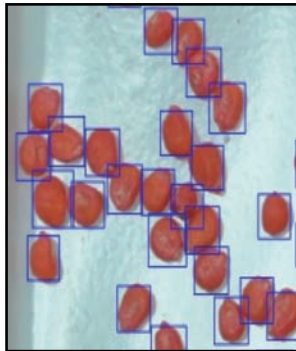
赵奇可, 孙延奎..... 114

局部不变特征点的精度指标

滕日, 周进, 蒋平, 崔雄文..... 122

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Detection of maize seeds based on multi-scale feature fusion and extreme learning machine(P0024)



Saliency detection based on deep convolutional neural network(P0053)



MRI denoising based on dual-tree complex wavelet transform(P0104)

Image Processing and Coding

- Spatio-temporal shape prediction and efficient coding
Zhu Zhongjie, Wang Yuer, Jiang Gangyi 1
- Interest point detection by using information entropy of the second small direction derivative
Lu Jian, Huang Jie, Pan Feng 8
- Anti-aliasing algorithm for character rotation based on area sampling
Zheng Kaiwen, Liu Wenbo 17

Image Analysis and Recognition

- Detection of maize seeds based on multi-scale feature fusion and extreme learning machine
Ke Xiao, Du Mingzhi 24

Image Understanding and Computer Vision

- Image fusion of visual and infrared image based on NSCT and compressed sensing
Chen Musheng 39
- Space resection method based on Plücker line
Sheng Qinghong, Chen Shuwen, Xiao Hui, Zhang Bin, Wang Qing, Fei Lijiao 45
- Saliency detection based on deep convolutional neural network
Li Yueyun, Xu Yuelei, Ma Shiping, Shi Hehuan 53

Computer Graphics

- Growth algorithm centered on priority point for constructing the Delaunay triangulation
You Lei, Tang Shouzheng, Song Xinyu 60

Medical Image Processing

- The 2D/3D medical image registration algorithm based on rapid digital image reconstruction
Liu Kun, Lyu Xiaoqi, Gu Yu, Yu Hefeng, Ren Guoyin, Zhang Ming 69
- Inter-slice interpolation for medical images by using the wavelet-lagrange method
Wu Shixiang, Shang Peng, Wang Ligong 78

Remote Sensing Image Processing

- Combining the expectation maximization and genetic algorithms for SAR image segmentation within a MAS framework
Zhang Jinjing, Li Yu, Zhao Quanhua 86

Column of IGIA' 2015

- Automatic registration of aerial stereo imagery to LiDAR data through iterative parallax mutual information computation
Wu Jun, Hu Yanjun, Rao Yun, Peng Zhiyong 95
- MRI denoising based on dual-tree complex wavelet transform
Huang Xueyou, Zhang Changjiang 104
- Three dimensional tracking with fast locating of image scale and area
Zhao Qike, Sun Yankui 114
- Extraction precision of local invariant feature points
Teng Ri, Zhou Jin, Jiang Ping, Cui Xiongwen 122

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2016)01-0053-07

论文引用格式: Li Y Y, Xu Y L, Ma S P, Shi H H. Saliency detection based on deep convolutional neural network[J]. Journal of Image and Graphics, 2016, 21(1): 0053-0059. [李岳云, 许悦雷, 马时平, 史鹤欢. 深度卷积神经网络的显著性检测[J]. 中国图象图形学报, 2016, 21(1): 0053-0059.][DOI:10.11834/jig.20160107]

深度卷积神经网络的显著性检测

李岳云, 许悦雷, 马时平, 史鹤欢

空军工程大学航空航天工程学院, 西安 710038

摘要: **目的** 显著性检测问题是近年来的研究热点之一, 针对许多传统方法都存在着特征学习不足和鲁棒检测效果不好等问题, 提出一种新的基于深度卷积神经网络的显著性检测模型。 **方法** 首先, 利用超像素的方法聚类相似特征的像素点, 仿人脑视皮层细胞提取目标边缘, 得到区域和边缘特征。然后, 通过深度卷积神经网络学习图像的区域与边缘特征, 获取相应的目标检测显著度置信图。最后, 将深度卷积神经网络输出的置信度融入到条件随机场, 求取能量最小化, 实现显著性与非显著性判别, 完成显著性检测任务。 **结果** 在两个常用的视觉检测数据库上进行实验, 本文算法的检测精度与当前最好的方法相比, 在 MSAR 数据库上检测精度相对提升大约 1.5%, 在 Berkeley 数据库上提升效果更加明显, 达到了 5%。此外, 无论是自然场景还是人工建筑场景、大目标与小目标, 检测的效果都是最好的。 **结论** 本文融合多特征的深度学习方法与单一浅层人工特征的方法相比更有优势, 它避免了手工标定特征所带来的不确定性, 具有更好的鲁棒性与普适性, 从主观视觉愉悦度和客观检测准确度两方面说明了算法的有效性。

关键词: 显著性检测; 超像素; 卷积神经网络; 条件随机场

Saliency detection based on deep convolutional neural network

Li Yueyun, Xu Yuelei, Ma Shiping, Shi Hehuan

Aeronautics and Astronautics Engineering College, Air Force Engineering University, Xi'an 710038, China

Abstract: **Objective** Saliency detection has become a highly active research field in recent years. Considering that many traditional methods suffer from insufficient feature learning and bad robust detection, this study proposes a novel saliency detection model based on deep convolutional neural networks. **Method** First, a pixel with similar characteristics is clustered by using superpixels and the target edge is extracted by imitating the human visual cortex cell to obtain the region and edge features. Thereafter, image regions and edge features are identified by convolutional neural networks to obtain the corresponding target-detection decision confidence images. Finally, we introduce the output of the deep-convolution neural network confidence coefficient into the conditional random field to calculate energy minimization. The discrimination of saliency and non-saliency is realized to complete the saliency detection task. **Result** Compared with the state-of-the-art method, the detection accuracy of our algorithm increases by approximately 1.5% in the MSAR database and 5% in the Berkeley database. Furthermore, our detection algorithm produces the best results whether in natural/artificial construction scenarios or large/small objects. **Conclusion** Our detection algorithm can avoid the uncertainty brought by manual features and has high

收稿日期: 2015-07-16; 修回日期: 2015-09-07

基金项目: 国家自然科学基金项目 (61372167, 61379104)

第一作者简介: 李岳云 (1991 年—), 男, 空军工程大学信息与通信工程专业硕士研究生, 主要研究方向为图像处理与目标识别。

E-mail: 171515576@qq.com

Supported by: National Natural Science Foundation of China (61372167, 61379104)

robustness and universality. Experimental results show the superiority of our proposed algorithm to the method using shallow artificial features. Both subjective visual pleasure and objective detection accuracy attest the effectiveness of the proposed algorithm.

Key words: saliency detection; superpixel; convolutional neural network; conditional random field

0 引言

显著性检测是计算机视觉领域中非常具有代表性的问题,它的目的是定位出那些最吸引人视觉注意的像素或区域。随着信息科技的发展与快速交通工具的推广,人类所接收的信息量呈指数级增长,如何筛选出人类感兴趣的目标和区域具有重要的意义。显著性区域与人视觉感知关系极为紧密,并具有一定的主观性,开展显著性检测的研究非常有利于图像处理基本任务的完成。例如,能够大致定位出兴趣目标,减小图像分割与检测的搜索范围和计算量;精确的显著性检测使目标跟踪与识别更加容易。

当前,关于显著性检测的研究成果颇为丰硕,主要分为在自顶向下与自底向上视觉机制两种方法。自底向上的方法就是从区域纹理、颜色、对比度、HOG、SIFT 等大量中低层特征出发,通过逐层特征抽取,实现显著与非显著性的区分。Itti 等人^[1]模拟人脑视皮层神经细胞最早提出显著性检测问题,通过融合颜色、亮度和方向特征,并对 3 种特征进行中心一周边机制算子求差分,从而能够检测出对比度明显的显著性区域。Yang 等人^[2]构建了联合条件随机场(CRF)与区分字典学习的自顶向下显著性模型,建立一个复杂的能量函数,通过最小化能量函数能够实现基于任务和目的的显著性检测,但它只能大致检测出目标位置,精度还远远不够。Zhan 等人^[3]使用区域相似算法计算区域之间方向通道与距离通道的相似性,实现区域的聚类,但该算法往往区域边界追踪不好,导致较大的偏差。Sun 等人^[4]提出一种边界先验图和凸壳软分割的方法,计算区域显著性值,并结合贝叶斯判别模型能够较好地定位出显著区域。Jiang 等人^[5]提出基于吸收马尔可夫链的显著性检测模型,从背景与前景的角度来认识显著性与非显著性区域,但它处理背景复杂的图像效果一般。深度学习方法近年也开始广泛用于检测研究中,Chen^[6]将深度卷积神经网络用于卫星图

像的目标检测,Wen^[7]则结合深度玻尔兹曼机特征学习来判别显著性区域,但该方法需要经过大量学习训练,复杂度比较高。

从区域和边缘特征出发,首先,通过超像素的方法聚类相似像素的区域,并模拟人视觉初级皮层端点细胞提取目标边缘,同时将超像素区域图和边缘图进行窗化处理,便可以归一化深度卷积神经网络的输入。其次,用两个独立的深度卷积神经网络分别学习区域和边缘特征,得到检测区域的显著性和非显著性置信度。最后,将基于特征学习的后验预测概率反馈到 CRF 得到区域的显著性值,从而对显著性与非显著性区域进行判别,完成显著性区域的检测。

1 卷积神经网络(CNN)

CNN 是一种区分型深度神经网络结构,它通过描述数据的后验概率从而提高模式分类的能力。CNN 由特征提取层和特征映射层所组成,特征提取层又称为卷积层,选取一系列不同的滤波器与图像进行卷积,得到的结果称为特征图谱。特征映射层也称为下采样层,它属于计算层,同样是多层结构,并且按神经网络的方式连接,一个映射层为一个平面,通过约束神经元共享网络权值,使得这种网络结构对图像的平移、缩放、旋转等形变具有高度不变性,是最早真正实现训练多层学习结构的神经网络模型。最后,多级卷积层和下采样层级联到一个或多个全连层,全连层的输出就是最终输出。图 1 是卷积神经网络一般结构图。

卷积神经网络含有多个卷积层与下采样层,这两者的操作有一定的区别,前者重在特征的提取上,而后者旨在特征的计算。卷积层节点输出表示为

$$a_n^l = f(\sum_{m \in \mathbb{M}} (a_m^{l-1} * k_{m,n}^l) + b_n^l) \quad (1)$$

式中, a_n^l 和 a_n^{l-1} 是当前层的特征图和上一层特征图, $k_{m,n}^l$ 表示从上一层第 m 个特征图到当前层第 n 个特征图的卷积核, $f(x) = 1/[1 + \exp(-x)]$ 为神经

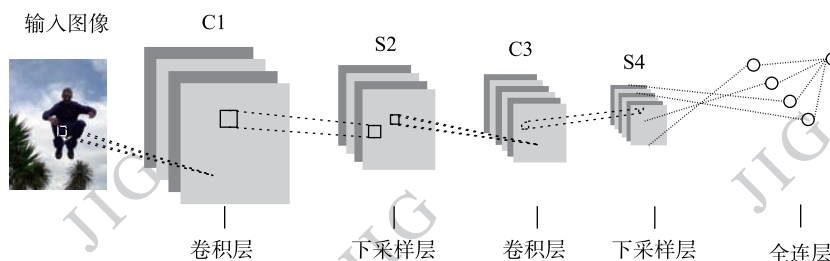


图1 卷积神经网络结构图

Fig. 1 Convolution neural network structure diagram

元激活函数, b_n^l 代表神经元偏置。它是卷积核与上一层特征图谱卷积和的响应, 采用不同的卷积核则可以提取出不同的特征。

下采样层节点输出可以表示为

$$a_n^l = f\left(k_n^l \times \frac{1}{s^2} \sum_{s \times s} a_n^{l-1} + b_n^l\right) \quad (2)$$

式中, $s \times s$ 是下采样模板尺度, k_n^l 为模板的权值。一般下采样的方法有最大值下采样、均值下采样和随机下采样等。因此, 卷积神经网络 CNN 全连接输出层的最终输出结果可以表示为

$$a_n^{\text{out}} = f\left(\sum_{\forall m} (a_m^{\text{out}-1} * k_{m,n}^{\text{out}}) + b_n^{\text{out}}\right) \quad (3)$$

2 基于深度卷积神经网络的显著性检测模型

显著性区域, 主观上理解是人眼视觉集中注意的区域, 它与人眼视觉系统(HVS)紧密相关。客观上而言则是针对图像的某种特征, 存在一个该特征最明显的子区。所以, 显著性检测问题的关键也就集中在特征学习和提取上, 鉴于深度学习在这一方面具有的强大功能, 本文将深度卷积神经网络用于显著性检测问题, 提出了基于深度卷积神经网络的显著性检测模型, 其具体模型如图2所示。

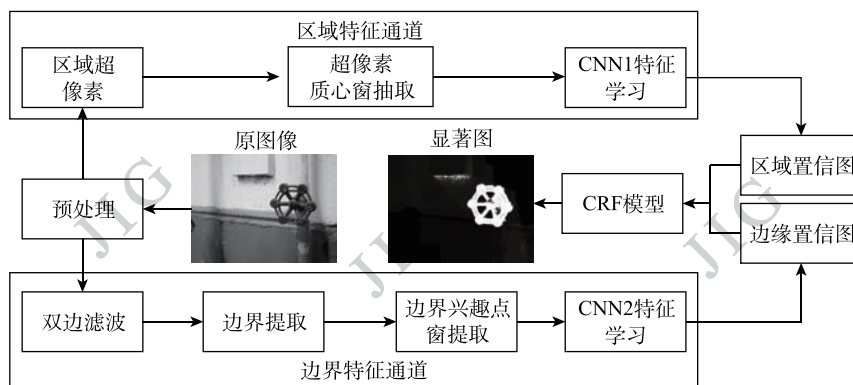


图2 基于深度卷积神经网络的显著性检测模型

Fig. 2 Saliency detection model based on deep convolutional neural network

给定一幅图像, 检测的目的就是确定显著性区域的位置, 对于局部小块 $x_i \in \mathbf{R}^p$, 设置 $L = \{y_i | y_i \in \{0, 1\}, 1 \text{ 为显著}, 0 \text{ 不显著}\}$ 的标记图像, y_i 是估计标签, 那么可以构建基于 CRF 的显著性检测模型

$$p(y_i | x_i; \mathbf{w}) = \frac{1}{Z(\mathbf{w})} \exp[-E(y_i, x_i; \mathbf{w})] \quad (4)$$

式中, 参数 $\mathbf{w}$ 表示 CRF 权值向量, $Z(\mathbf{w})$ 是归一化分布函数, $E(\cdot)$ 是能量函数。因此可以衡量局部

块的显著性值为

$$s_i = p(y_i = 1 | x_i; \mathbf{w}) \quad (5)$$

当有 m 个局部块显著时, 可得该图像的显著区域图 $S = \{s_1, s_2, s_3, \dots, s_m\}$ 。CRF 是一个综合观测数据和邻域标记的判别分类框架, 其能量函数能够分解为数据能量项和平滑能量项两部分叠加的形式, 即

$$E(y_i, x; \mathbf{w}) = \sum_{i \in v} \varphi_i(y_i', x; \mathbf{w}_i) +$$

$$\sum_{(i,j) \in \mathcal{S}} \varphi_{ij}(\mathbf{y}_i, \mathbf{y}_j; \mathbf{w}_{ij}) \quad (6)$$

式中, $\mathbf{y}'_i$ 表示对第 i 局部小块的标签估计, $\mathbf{y}_i, \mathbf{y}_j$ 指两相邻的局部图像块, 前一部分对应为数据能量项, 表示像素属于显著区域与非显著区域的置信度, 后一项对应为平滑能量项, 表示显著区域与非显著区域的局部相似性。对数据能量项和平滑能量项可以进一步分解为式(7), 数据能量项与其所在的区域(region)和区域边界(boundary)是紧密相连的, 它是由这两部分共同决定的。

$$\begin{aligned} \varphi_i(\mathbf{y}'_i, \mathbf{x}; \mathbf{w}_i) &= \overbrace{\varphi_i^r(\mathbf{y}'_i, \mathbf{x}; \mathbf{w}_i^r)}^{\text{region}} + \overbrace{\varphi_i^b(\mathbf{y}'_i, \mathbf{x}; \mathbf{w}_i^b)}^{\text{boundary}} \\ \varphi_{ij}(\mathbf{y}_i, \mathbf{y}_j; \mathbf{w}_{ij}) &= w_{ij} \phi_{p1}(\mathbf{y}_i, \mathbf{y}_j) \phi_{p2}(\mathbf{x}) \end{aligned} \quad (7)$$

$$\text{式中, } \phi_{p1}(\mathbf{y}_i, \mathbf{y}_j) = \begin{cases} 0 & \mathbf{y}_i = \mathbf{y}_j \\ \alpha & \mathbf{y}_i \neq \mathbf{y}_j \end{cases}, \quad \phi_{p2}(\mathbf{x}) =$$

$\exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / \beta_x \langle \|\mathbf{x}_i - \mathbf{x}_j\|^2 \rangle)$, 参数 α 和 β_x 是常量, $\langle \cdot \rangle$ 为图像的平均对比度。 $\phi_{p2}(\mathbf{x})$ 称为空间跃变函数, 它可以用来衡量局部块之间的像素亮度差异。CRF 中最重要的部分便是数据能量项, 通过两个单独的 CNN 神经网络学习而得到区域置信度, 每一节点的区域与边界能量按照两个单独的 CNN 神经网络输出状态来定义, 其形式为

$$\begin{aligned} \phi_i^r(\mathbf{y}'_i, \mathbf{x}; \mathbf{w}_i^r) &= -\mathbf{w}_i^r \log p_{\text{CNN1}}(\mathbf{y}'_i | \mathbf{x}_r) \\ \phi_i^b(\mathbf{y}'_i, \mathbf{x}; \mathbf{w}_i^b) &= -\mathbf{w}_i^b \log p_{\text{CNN2}}(\mathbf{y}'_i | \mathbf{x}_b) \end{aligned} \quad (8)$$

显而易见, 边界和区域具有的特征是不同的, 因此, 必须单独分开训练这两个 CNN。 $p_{\text{CNN1}}(\mathbf{y}'_i | \mathbf{x}_r)$ 与 $p_{\text{CNN2}}(\mathbf{y}'_i | \mathbf{x}_b)$ 是两个单独的 CNN 神经网络中得到的区域与边缘置信度。但是, CNN 的每一层运算都是针对相同大小的窗来实现的, 所以在此之前, 需要对边缘与区域特征都进行等窗处理。针对传统的平均分块一幅图像来获得图像局部块的方法, 它没有考虑到块与块之间的互相关信息, 而只是简单的划分图像, 存在着许多不足。Achanta 等人^[8]发现对图像进行超像素处理, 不仅能够很好的保持目标边界而且可以将具有相似特征的像素进行聚类, 一定程度上消除了图像的冗余信息, 而又保证了块之间的互相关信息。Superpixel 处理的方法也有很多, 这里采用简单线性迭代聚类算法^[8](SLIC)来获得超像素图, 该方法相对简单, 能够大大减小计算量, 对边界的保持效果非常不错。对于 Superpixel 获得的每一个图像区域块, 找到该块的质心, 以质心为中心抽取大小为 $\tau \times \tau$ 的窗口, 这样得到相同大小的 CNN1

运算窗。上述过程用数学式表示为

$$\Gamma^r(\mathbf{z}_i, \mathbf{z}_j) = W^{(\mathbf{z}_i, \mathbf{z}_j)} [F_{\text{SLIC}}(\mathbf{x}_i), \tau] \quad (9)$$

式中, $\Gamma^r(\mathbf{z}_i, \mathbf{z}_j)$ 表示对质心 $(\mathbf{z}_i, \mathbf{z}_j)$ 取窗结果, $W^{(\mathbf{z}_i, \mathbf{z}_j)}$ 是加窗操作, $F_{\text{SLIC}}(\mathbf{x}_i)$ 为 SLIC 算法的 Superpixel 处理, $\mathbf{x}_i$ 是区域块。下图 3 所示是显著性检测的中间过程图, 方框表示窗化处理。对于边界部分, 首先对图像进行双边滤波, 它对高频信息有很好的保护作用, 因而能够起到保持边缘的作用, 然后通过拟人脑视皮层端点细胞技术^[9]提取图像边缘和边缘兴趣点, 在兴趣点基础上执行如上所述的区域取窗操作, 该过程同样可以数学描述为

$$\Gamma^b(\mathbf{z}_i, \mathbf{z}_j) = W^{(\mathbf{z}_i, \mathbf{z}_j)} [F_{\text{edg}}(\mathbf{x}_i), \tau] \quad (10)$$



(a) 原图 (b) 超像素图 (c) 边缘图 (d) 显著图

图 3 显著性检测过程图

Fig. 3 Saliency detection process diagram((a) original image; (b) superpixel map; (c) boundary map; (d) saliency map)

取窗后的两类特征块此时便可以作为 CNN 的输入, 经过其多级分层处理, 便能够得到模型

$$\begin{aligned} p_{\text{CNN1}}(\mathbf{y}'_i | \Gamma^r(\mathbf{z}_i, \mathbf{z}_j)) &= c(\theta(\Gamma^r(\mathbf{z}_i, \mathbf{z}_j))) \\ p_{\text{CNN2}}(\mathbf{y}'_i | \Gamma^b(\mathbf{z}_i, \mathbf{z}_j)) &= c(\theta(\Gamma^b(\mathbf{z}_i, \mathbf{z}_j))) \end{aligned} \quad (11)$$

式中, c 表示 CNN 运算输出, θ 是图形预处理操作, 输出层的响应值即是后验概率的分布, 即

$$p_{\text{CNN}}(\mathbf{y}_i | \Gamma(\mathbf{z}_i, \mathbf{z}_j)) = \mathbf{a}_n^{\text{out}} = [a_{\text{sal}}^{\text{out}}, a_{\text{no-sal}}^{\text{out}}] \quad (12)$$

式中, $\mathbf{a}_n^{\text{out}}$ 表示第 n 个输出, $[a_{\text{sal}}^{\text{out}}, a_{\text{no-sal}}^{\text{out}}]$ 指区域显著与否的后验概率大小。

综上所述, 显著性检测便是在一组随机变量 $y \in \mathbf{R}^N$ 中寻找最可能显著的标签, 即求取最大后验概率(MAP)估计。在推导过程中很容易发现, 由于分布函数 $\mathbf{Z}(\mathbf{w})$ 并不依赖于 y , 所以 MAP 问题转化为求取能量最小化的问题, 即

$$\begin{aligned} y^* &= \arg \max_{y \in L^*} p(\mathbf{y} | \mathbf{x}; \mathbf{w}) = \\ &\arg \min_{y \in L^*} E(\mathbf{y}, \mathbf{x}; \mathbf{w}) \end{aligned} \quad (13)$$

3 实验结果及分析

3.1 实验条件及评价指标

为了评估提出算法显著性检测的有效性, 在

ASD 和 Berkeley 图像数据库上进行实验。ASD 数据库是 MSAR 库的子库,它含有 1 000 幅图像,包括人、动植物和人工建筑等多种目标,是图像识别和显著性检测的常用数据库。Berkeley 图形库包括了多种不同尺寸和不同位置的前景目标,它作为图像分割的一个标准数据库,也被广泛用于检测任务中,选择该图形库做显著性检测比 MSAR 更具有挑战性。实验中 CNN 结构与 ImageNet 类似^[10],包括了 5 级卷积层与 2 级全连层。

显著性检测评价标准包括主观评价与客观评价两种,主观评价主要是针对人眼主观感受,也就是视觉感受的愉悦性;客观评价检测指标包括查准率 P 、查全率 R 和 F-measure。查准率是指正确检测的显著区域与检测总区域的比值,它反映了检测的准确性,衡量检测系统和检测方法拒绝误检测的能力,其补数是误检率;查全率是正确检测的显著区域与总显著区域的比值,它反映了检测的全面性,衡量检测系统和检测方法正确检测的能力,其补数为漏检率。 P 和 R 的值越大检测效果越好,但这两者之间存在一定矛盾,例如,当检测区域很大时,可以保证查全率比较高,但查准率不够。因此,需要采用 F-measure 值来综合两种指标评价标准,使两者之间保持均衡,其定义式为

$$F_{\beta} = \frac{(1 + \beta^2) P \times R}{\beta^2 P + R} \quad (14)$$

式中, β 为常数,根据文献^[11]取 $\beta^2 = 0.3$ 。一般为了更加直观有效地反映显著性检测结果的好坏,根据查全率 R 与查准率 P 绘制 PR 响应曲线,并计算曲线下面积 AUC (area under curve)。

3.2 实验分析

将本文算法与 MR^[11]、LMC^[12]、BSF^[4]、GR^[13]、BM^[14]、AMC^[5]、DSR^[15] 当前几种比较流行的显著性检测方法进行比较。图 4 是本文算法与对比算法在 Berkeley 和 ASD 数据库中实验所得的部分显著性检测结果图,前 3 行是 Berkeley 数据库上的实验结果,最后一行是 ASD 数据库上的实验结果。从图 4 中可以看出,本文显著性区域检测与手工检测结果 (GT) 重合度是最好的,同时满足人眼视觉的愉悦性,与视觉注意点吻合。此外,无论在显著性区域相对较小 (蜗牛) 或较大 (建筑物),还是显著性区域与背景具有很大的相似性时 (麋鹿),本文算法的效果都很不错。从图 4 (b)

(c)(e)(f)(g) 中观察,效果图中都有一些块状片段,这严重影响视觉的愉悦度,产生该现象的原因正是对超像素区域显著与非显著的判断不正确造成的。

为了定量说明本文算法的性能,根据前文提到的客观评价标准,可以绘制 7 种对比算法与本文算法的 PR 曲线与 P 、 R 和 F-measure 柱状图,如图 5 所示。并分别计算各算法的检测精度与 PR 曲线下所围成的面积 AUC,结果列于表 1 中。从两个实验数据库的 PR 曲线与柱状图中不难发现:其一,本文算法 PR 曲线更靠近右上角,也就是说本文算法在某一情况下能得到相对最佳的 P 、 R 值;其二,随着横坐标的增大, P 的值均为递减的趋势,即使在 R 比较大时也能够保证查准率 P 比较大;其三,从柱状图看,虽然 P 的值不总是最大的,但 R 的数值在两个数据库上都是最大的,而且综合指标 F-measure 都比其他算法好。

由表 1 中的统计结果可知,AUC 在两个实验数据库中都是最大的,这说明在相同情况下 P 和 R 的数值都比较大;客观上的检测精度计算结果在各类算法中比较也是最高的。对比算法中效果最好的要数 AMC,本文算法检测精度在 MSAR 数据库上相对提升大约 1.5%,在 Berkeley 数据库上提升效果更加明显,达到了 5%。因此,无论主观评价标准还是客观评价标准都一致表明本文算法较 7 种比较算法检测效果都好。DSR^[15]、LMC^[12]、BM^[14] 是基于贝叶斯的概率推论模型,其先验概率的获取比较粗糙,因而对比算法视觉愉悦度差,这决定了检测的实际效果,本文则通过 CNN 强大的抽象特征学习能力,将网络响应作为条件随机场先验,准确度比较高。GR^[13]、AMC^[5] 和 MR^[11] 算法都只用到了单一的简单特征,本文综合了显著性检测最重要的区域与边缘特征,以双通道的途径获取区域与边缘置信图,通过 CRF 融合这两种置信图,得到更为精确的显著性区域。此外,双边滤波后基于人眼视觉的边缘提取,对显著边缘处理效果很好,同时有一定抑制非显著边缘的能力,这与本文的任务相一致。转化检测问题为回归问题,是深度学习处理检测问题的一大技巧和优势,通过这种转化,充分利用深度学习描述数据后验概率从而提高检测准确率的能力。

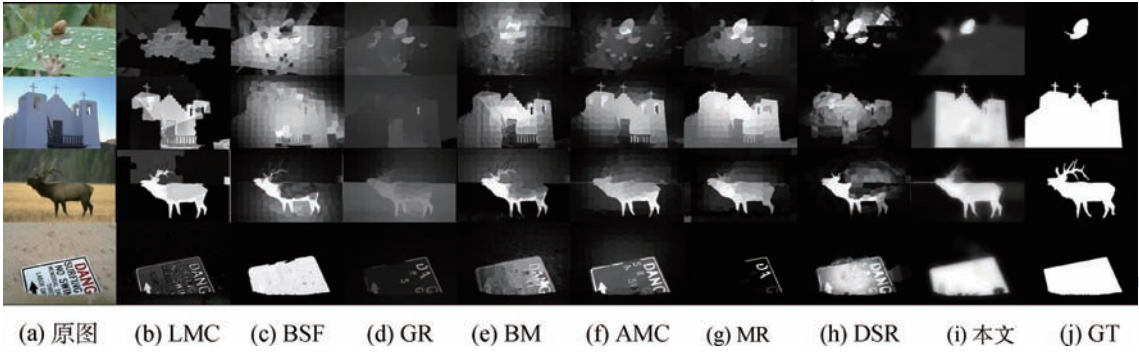


图4 本文算法与7种当前流行算法对比结果

Fig. 4 Comparison of our results with seven state-of-the-art methods((a) original images; (b) LMC; (c) BSF; (d) GR; (e) BM; (f) AMC; (g) MR; (h) DSR; (i) our; (j) GT)

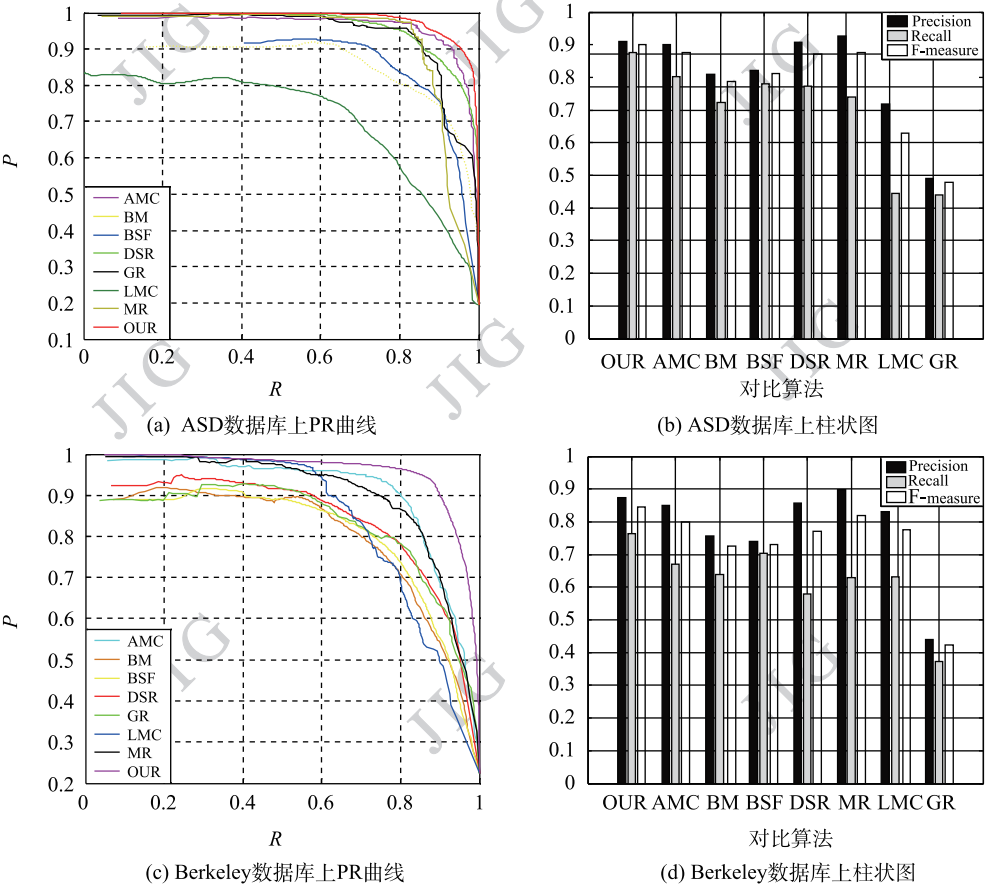


图5 ASD 和 Berkeley 数据库实验结果

Fig. 5 Experimental results on the ASD and Berkeley dataset((a) precision-recall curves on the ASD dataset; (b) bar chart on ASD dataset; (c) precision-recall curves on the Berkeley dataset; (d) bar chart on ASD dataset)

表1 检测精度与 AUC 统计结果

Table 1 Detection accuracy and AUC statistical results

数据库	评价指标	LMC	BSF	GR	BM	AMC	MR	DSR	本文
MSAR	精度/%	79.29	89.91	83.68	88.85	92.47	92.86	92.24	93.79
	AUC	0.707 8	0.792 7	0.890 1	0.714 4	0.876 2	0.879 0	0.787 0	0.890 1
Berkeley	精度/%	86.02	82.95	77.81	82.10	86.33	88.03	86.49	91.63
	AUC	0.755 6	0.791 3	0.813 4	0.740 4	0.856 9	0.854 3	0.782 1	0.911 3

4 结 论

从数据特性学习中提出一个深度卷积神经网络结构的单幅图像显著性检测模型,该模型并不依赖于目标形状、环境状况和场景类型的变化,具有较好的鲁棒性。第一、将深度学习的方法用于显著性检测,能够鲁棒有效地学习提取显著性特征,模拟了视觉自底向上的注意机制。第二、超像素方法的使用,将相似特征的像素聚类而不是以单一像素为单元,既可以减少检测的错误,还可以减少运算量。第三、仿人脑视皮层细胞方法的边缘提取,能够获得对比明显的边缘,同时抑制那些虚弱边缘,这与显著性检测需求刚好契合。第四、运用条件随机场的两种置信图融合,其效果好于贝叶斯原则的融合方法。进一步需研究的方向是多显著目标的检测,现有大多数算法包括本文算法,都无法很好地处理显著性多目标的检测。

参考文献 (References)

- [1] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE TPAMI, 1998, 20(11):1254-1259.
- [2] Yang J, Yang M H. Top-down visual saliency via joint CRF and dictionary learning [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI: IEEE Press, 2012, 2296-2303.
- [3] Zhan J, Hu B. Salient object contour detection based on boundary similar region [C]//Proceedings of the 4th International Conference on Digital Home. Guangzhou: IEEE Press, 2012, 335-339.
- [4] Sun J, Lu H C, Li S F. Saliency detection based on integration of boundary and soft-segmentation [C]//Proceedings of IEEE International Conference on Image Processing. Orlando, FL: IEEE Press, 2012, 1085-1088.
- [5] Jiang B W, Zhang L H, Lu H C, et al. Saliency detection via absorbing markov chain [C]//Proceedings of IEEE Conference on International Conference on Computer Vision. Sydney, NSW: IEEE Press, 2013, 1665-1672.
- [6] Chen X Y, Xiang S M, Liu C L, et al. Vehicle detection in satellite images by parallel deep convolutional neural networks [C]//Proceedings of the 2nd IAPR Asian Conference on Pattern Recognition. Japan, Naha: IEEE Press, 2013, 181-185.
- [7] Wen S F, Han J W, Zhang D W, et al. Saliency detection based on feature learning using deep boltzmann machines [C]//Proceedings of IEEE International Conference on Multimedia and Expo. Chengdu: IEEE Press, 2014, 1-6.
- [8] Achanta R, Shaji A, Smith K, et al. Slic superpixels compared to state-of-the-art superpixel Methods [J], TPAMI, 2012, 34(11):2274-2282.
- [9] Csapo, Adam B, Andras R, et al. Visual cortex inspired vertex and corner detection [C]//IEEE International Conference on Mechatronics. Hungary, Budapest: IEEE Press, 2006, 551-556.
- [10] Jia D, Wei D, Socher, et al. ImageNet: a large-scale hierarchical image database [C]//Proceedings of Conference on Computer Vision and Pattern Recognition. Miami, FL: IEEE Press, 2009, 248-255.
- [11] Yang C, Zhang L H, Lu H C, et al. Saliency detection via graph-Based manifold Ranking [C]//Proceedings of Conference on Computer Vision and Pattern Recognition. Portland, OR: IEEE Press, 2013, 3166-3173.
- [12] Xie Y L, Lu H C, Yang M H. Bayesian saliency via low and mid level cues [J]. Image Processing, 2013, 22(5):1689-1398.
- [13] Yang C, Zhang L H, Lu H C. Graph-regularized saliency detection with convex-hull-based center prior [J]. Signal Processing Letters, 2013, 20(7): 637-640.
- [14] Xie Y L, Lu H C. Visual saliency detection based on bayesian, model [C]//Proceedings of IEEE International Conference on Image Processing. Brussels: IEEE Press, 2011, 645-648.
- [15] Li X H, Lu H C, Zhang L H, et al. Saliency detection via dense and sparse reconstruction [C]//Proceedings of International Conference on Computer Vision. Sydney, NSW: IEEE Press, 2013, 2976-2983.