

# 基于全局和局部信息融合的显著性检测

刘尚旺<sup>a</sup>, 赵欣莹<sup>a</sup>, 杨磊<sup>b</sup>

(河南师范大学 a.计算机与信息工程学院;b.旅游学院,河南 新乡 453007)

**摘要:**为提高低对比度、复杂自然图像显著性检测的准确率和泛化性能,提出一种贝叶斯框架下的全局和局部信息融合的显著性检测模型.首先,构建深度卷积自编码网络,采用对称编解码结构,监督学习图像全局特征,得到全局显著图;然后,根据全局显著图产生前景和背景码本,利用局部约束线性编码算法进行编码,采用稀疏编码描述局部特征,产生局部显著图;最后,提出采用贝叶斯框架,将全局和局部信息融合,生成最终显著图.实验结果表明,所提模型在 ECSSD、DUT-OMRON 和 PASCAL 数据集上 F-measure 值分别为 76.53%、59.45% 和 72.52%, MAE 值分别为 0.143 28、0.137 87 和 0.181 05,且能够有效对低对比度、复杂真实自然图像进行显著性检测.

**关键词:**显著性检测;贝叶斯框架;稀疏编码;深度卷积自编码网络

**中图分类号:**TP391

**文献标志码:**A

随着图像获取设备的普及,图像资源越来越丰富,信息冗余问题随之产生.近年来,降低计算复杂度的图像显著性检测成为图像预处理的一个重要步骤,广受计算机视觉领域的研究关注<sup>[1-10]</sup>.

根据人类视觉系统原理,图像显著性检测方法分为自下而上和自上而下两种.自下而上方法通常基于低层先验知识,属于目标驱动,通过计算视觉特征如颜色、纹理、位置、亮度等得到图像的显著结果.文献[1]综合模式和颜色显著性,通过分析图像块的内部统计信息,以深入了解图像块的特征,采用图像块的主成分分析形式表示图像显著性.文献[2]计算图像的对比度和中心先验作为初始显著图,通过引入基于图的目标分割思想,计算成对超像素块之间的平滑先验,能够均匀地突出显著目标,同时有效抑制背景信息.文献[3]将图像表示为图的形式,利用显著性检测与马尔科夫吸收概率之间的关系,建立吸收马尔科夫链,根据吸收概率矩阵确定显著目标位置,通过前景先验排序得到显著图.文献[4]为建立更加复杂、鲁棒的边界先验,提出一种基于边界连通性的背景检测算法,来表征图像区域相对于图像边界的空间信息.文献[5]对图像进行稠密和稀疏重构,稠密模型表示图像的通用描述且对噪声更敏感,稀疏模型表示图像的独特描述且简洁,通过重构误差分别建立图像显著图,采用贝叶斯模型将二者融合.

自上而下的方法需要结合高层知识来指导检测计算,属于任务驱动.文献[6]为解决自下而上模型中存在的 inaccuracies 和特征表示的限制性,提出一种学习机制,通过融合图像的多种低层特征得到弱显著性检测模型,采用 AdaBoost 增强学习方法迭代训练得到强显著性检测模型.文献[7]建立一种基于低秩矩阵恢复的显著性检测算法,将传统的低层特征和高层指导结合起来,来检测显著目标.文献[8]通过低层和高层对比度计算全局显著图,基于全局显著图生成码本,采用局部约束线性编码(Locality-constrained Linear Coding, LLC)模型生成局部显著图,优化融合得到最终检测结果.文献[9]提出一种基于深度学习架构的堆栈去噪自编码器,以模拟探索潜在模式的背景,通过计算深度自编码的重构残差,将显著目标从背景中有效分离开.

上述方法中,传统显著性检测算法主要依赖于对比度和先验知识,对于低对比度和环境复杂的图像无法

**收稿日期:**2018-12-08;**修回日期:**2019-03-21.

**基金项目:**河南省科技攻关项目(192102210290);河南省高等学校重点科研项目(15A520082);河南师范大学博士科研启动基金(qd12138).

**作者简介:**刘尚旺(1973-),男,河南新乡人,河南师范大学副教授,博士,研究方向为计算机视觉,E-mail:shwl2012@hotmail.com.

**通信作者:**杨磊,E-mail:40119191@qq.com.

取得良好的效果,深度卷积神经网络由于其局部感知和参数共享等特性,在图像处理中取得突出成就.但是,随着越来越多深度网络的提出,发现由于卷积网络的局部感受野结构,检测结果通常会出现目标边界模糊的情况,而且没有明确地将低层和高层显著线索结合起来<sup>[11]</sup>.针对低对比度和环境复杂的图像,传统的基于对比度和先验线索的检测方法不能取得良好的效果,然而深度网络依据其特殊的结构和强大的学习能力,能够获得更精确的显著结果,但也会出现检测失败的情况,通过与传统方法的融合,有利于弥补深度网络的不足,因此本文采用二者结合的思想,使用深度卷积自编码网络提取图像全局特征,在此基础上提取局部特征,采用贝叶斯框架,获取更为可靠的后验概率,产生最终的检测结果.

## 1 本文模型

本文显著性检测模型框架图,如图 1 所示.

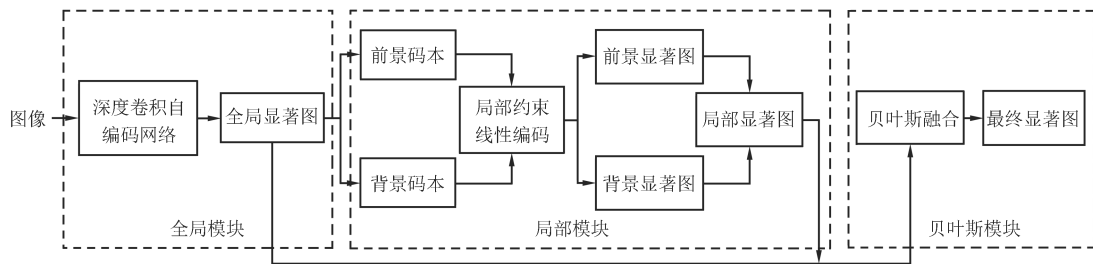


图 1 本文模型整体框架图

Fig.1 The overall architecture of our model

从图 1 可以看出,模型分为 3 部分:全局模块、局部模块和贝叶斯模块.在全局模块中,构建深度卷积自编码网络,提取和重构图像抽象全局特征;在局部模块中,根据全局特征,产生前景和背景码本,采用局部约束线性编码算法产生具有局部细节信息的局部特征;在贝叶斯模块中,将全局和局部特征通过后验概率有效融合,使得显著结果同时包含全局轮廓和局部细节信息.

### 1.1 基于卷积自编码的全局检测

传统显著性检测方法通常依赖于对比度计算图像显著图<sup>[1-2,6-8]</sup>,同时需要依赖有限的先验知识,对于具有低对比度和复杂环境的图像,会具有较大的挑战性<sup>[2-4,6]</sup>.通过采用图 2 所示的深度卷积自编码器提取图像全局特征,以产生全局显著图.卷积自编码模型结构分为两部分:卷积部分通过卷积网络来提取图像显著目标的全局空间信息(如图 2 左半部分所示);反卷积部分通过反卷积网络以产生图像全局显著图(如图 2 右半部分所示).卷积部分和反卷积部分是对称的结构,一个编码层有相应的解码层,均采用 VGG-16 网络结构<sup>[12]</sup>.

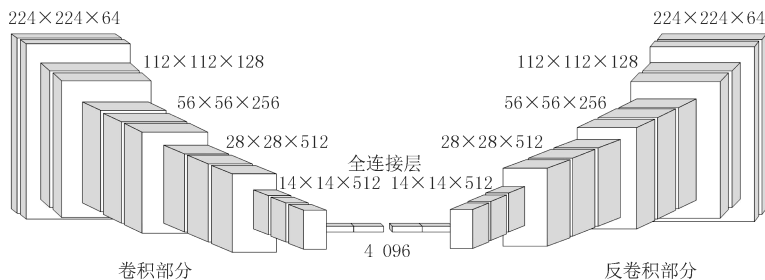


图 2 卷积自编码模型结构图

Fig.2 The architecture of the convolutional autoencoder network

从图 2 可以看出,VGG-16 网络有 5 个卷积组,后面连接 2 个全连接层;每组卷积层之间通过池化层连接,卷积部分具有 5 个池化层.

卷积层和反卷积层特征表示为  $h^k = f\left(\sum_{p \in P} x^p \otimes \omega^k + b^k\right)$ , 其中,  $h^k$  表示当前层的第  $k$  个特征图的潜在表示;  $f$  是激活函数, 采用 ReLU 激活函数进行非线性计算;  $x^p$  表示前一层的特征图组  $P$  的第  $p$  个特征图;  $\otimes$  表示卷积操作;  $\omega^k$  和  $b^k$  分别表示当前层的第  $k$  个特征图的权重和偏置。

通过监督学习的方式训练深度卷积自编码网络, 训练样本包含原图和对应的真值图. 将每轮训练结果与真值图对比, 来微调网络参数, 使得像素级的交叉熵损失函数值最小化. 交叉熵损失函数的定义为  $L = \frac{1}{HW} \sum_{i=1}^W \sum_{j=1}^H [-G(I)_{i,j} \cdot \ln(D \circ E(I)_{i,j}) - (1 - G(I)_{i,j}) \cdot \ln(1 - D \circ E(I)_{i,j})]$ , 其中,  $E$  表示编码方法;  $D$  表示解码方法;  $G(I)$  表示图像  $I$  的真值图, 其高为  $H$ , 宽为  $W$ ;  $(i, j)$  表示像素的坐标;  $D \circ E(I)$  表示网络的卷积输出。

在网络训练过程中, 采用梯度下降算法和反向传播算法来逐步调整各层的网络参数, 使得损失足够小<sup>[13]</sup>. 权重和偏差的更新, 分别通过  $\omega^k \rightarrow \omega^{k'} = \omega^k - \eta \frac{\partial L}{\partial \omega^k}$  和  $b^k \rightarrow b^{k'} = b^k - \eta \frac{\partial L}{\partial b^k}$  实现, 其中,  $\eta$  表示学习率, 设置为  $10^{-4}$ ;  $\frac{\partial L}{\partial \omega^k}$  和  $\frac{\partial L}{\partial b^k}$  表示当前损失函数对应  $\omega^k$  和  $b^k$  的梯度方向, 梯度的反方向表示函数值下降最快的方向,  $\omega^{k'}$  和  $b^{k'}$  分别表示更新后的权重和偏置。

## 1.2 基于局部稀疏编码的局部检测

采用 LLC<sup>[14]</sup> 模型, 通过最小化重构误差生成局部显著图. 首先, 使用简单线性迭代聚类 (Simple Linear Iterative Clustering, SLIC) 算法<sup>[15]</sup> 将图像分为  $N$  块超像素区域,  $\{r_i\}, i=1, 2, \dots, N$ ; 其次, 在全局显著图基础上计算视觉码本, 通过自适应阈值  $\lambda_1, \lambda_2$  产生前景和背景码本, 其中,  $\lambda_2 < \lambda_1 < 1$ , 根据全局显著图的平均值的  $t$  倍自适应计算得到  $\lambda_1$ , 本文设置  $t = 1.5$ <sup>[8]</sup>, 由于码本对  $\lambda_2$  不敏感, 因此设置  $\lambda_2 = 0.05$ <sup>[8]</sup>; 最后, 超像素区域全局平均显著值大于  $\lambda_1$  则为前景码本, 小于  $\lambda_2$  为背景码本。

LLC 算法对局部描述子  $x_i$  的编码, 如(1)式所示。

$$\begin{aligned} \min_C \sum_{i=1}^N \|x_i - Bc_i\|^2 + \lambda \|d_i \odot c_i\|^2, \\ \text{s.t. } 1^T c_i = 1, \forall i, \end{aligned} \quad (1)$$

其中,  $B$  表示视觉码本; 通过约束条件, 计算得到编码系数向量  $C = \{c_1, c_2, \dots, c_N\}$ ;  $\lambda$  控制惩罚项和正则项之间的平衡;  $\odot$  表示点乘运算;  $1^T c_i = 1$  体现编码的平移不变特性.  $d_i$  表示为每个基向量赋予的局部适配器, 与  $x_i$  的相似性成正比, 由下式计算得出,

$$d_i = \exp\left(\frac{d_s(x_i, B)}{\sigma}\right), d_s(x_i, B) = [d_s(x_i, b_1), \dots, d_s(x_i, b_O)]^T,$$

其中,  $d_s(x_i, b_j)$  表示  $x_i$  与  $b_j$  之间的欧式距离;  $\sigma$  表示  $c_i$  的权重, 用来控制局部适配器的权重衰减速度;  $O$  表示码本的维度。

为此, 可以将每个局部描述子  $x_i$  由尽可能临近的码本中心描述。

通过  $c_i = 1/(C_i + \lambda \times \text{tr}(C_i))$ ,  $\tilde{c}_i = c_i/1^T c_i$ , 对(1)式进行求解. 其中,  $C_i = (B - 1x_i^T)(B - 1x_i^T)^T$  表示数据的协方差矩阵;  $\lambda$  表示正则化参数, 设置为  $\lambda = 0.1$ <sup>[8]</sup>,  $\text{tr}$  表示矩阵的迹。

通过重构误差, 计算出基于码本的局部显著图  $S_l(r_i) = \|x_i - B\tilde{c}_i\|^2$ 。

局部显著图由编码重构误差计算产生. 对于背景, 重构误差等于 0 为理想; 对于前景目标, 则重构效果越差越好. 为此, 针对前景码本进行局部编码, 其重构误差与相应超像素区域显著值是相反的; 针对背景码本进行局部编码, 其重构误差作为相应超像素区域显著值. 局部显著图的计算如(2)式所示。

$$S_l(r_i) = (1 - S_{lf}(r_i)) \times S_{lb}(r_i), \quad (2)$$

其中,  $S_{lf}(r_i)$  表示超像素区域  $r_i$  的前景显著图,  $S_{lb}(r_i)$  表示背景显著图。

## 1.3 基于贝叶斯的全局和局部特征融合

全局显著图  $S_g$  和局部显著图  $S_l$ , 通过贝叶斯模型<sup>[16-17]</sup> 融合成最终显著图. 若  $S_{gf}(z)$  和  $S_{gb}(z)$  分别表

示像素点  $z$  在全局显著图中属于前景和背景;  $S_{lf}(z)$  和  $S_{lb}(z)$  分别表示像素点  $z$  在局部显著图中属于前景和背景.对全局显著图进行归一化操作,将像素点  $z$  对应的显著值  $Y_g(z)$  作为其属于前景的概率  $\rho(S_{gf}(z))$ ,则属于背景的概率  $\rho(S_{gb}(z)) = 1 - \rho(S_{gf}(z))$ .

选取阈值对全局显著图二值分割,产生前景和背景部分,计算对应部分在局部显著图中的像素个数  $N_{lf}^g$  和  $N_{lb}^g$ ,计算出像素点的显著值统计直方图.像素  $z$  在局部显著图中的显著值设为  $Y_l(z)$ ,则像素  $z$  的条件概率可通过(3)式计算得到.

$$\rho(Y_l(z) | S_{gf}(z)) = N_{lf}^g(Y_l(z)) / N_{lf}^g, \rho(Y_l(z) | S_{gb}(z)) = N_{lb}^g(Y_l(z)) / N_{lb}^g, \quad (3)$$

其中,  $N_{lf}^g(Y_l(z))$  和  $N_{lb}^g(Y_l(z))$  分别表示前景和背景部分在显著值统计图中包含  $Y_l(z)$  的位的像素数.

后验概率的计算,如(4)式所示,

$$\rho(S_{gf}(z) | Y_l(z)) = \frac{\rho(S_{gf}(z))\rho(Y_l(z) | S_{gf}(z))}{\rho(S_{gf}(z))\rho(Y_l(z) | S_{gf}(z)) + \rho(S_{gb}(z))\rho(Y_l(z) | S_{gb}(z))}. \quad (4)$$

(4)式将全局显著图作为先验.若将局部显著图作为先验,可得相应的后验概率  $\rho(S_{lf}(z) | Y_g(z))$ .

经过贝叶斯模型融合成最终显著图,如(5)式所示.

$$S = \rho(S_{gf}(z) | Y_l(z)) \times \rho(S_{lf}(z) | Y_g(z)). \quad (5)$$

## 2 实验结果与分析

### 2.1 实验设置

实验平台:CPU, Intel Xeon E5-2690 v4 2.6 Ghz; 内存, 32 G; 显卡, NVIDIA Titan X; 显存, 12 G; 操作系统, Ubuntu17.10; 编程语言, Python 3; 深度学习环境, TensorFlow 1.7.0.

深度卷积自编码网络训练集为 MSRA 10K 数据集<sup>[18]</sup>, 其中包含 10 000 幅图像和对应的真值图.网络在 VGG-16 网络原有权重的基础上完成训练.训练迭代次数设置为 100 000 步.输入图像大小均被调整为  $224 \times 224$  像素.

为验证本文模型的性能,在 3 个广泛使用的标准数据集上进行评价:1) ECSSD 数据集<sup>[19]</sup>: 包含 1 000 幅具有复杂结构的自然图像.2) DUT-OMRON 数据集<sup>[20]</sup>: 包含 5 168 幅具有挑战的图像.3) PASCAL 数据集<sup>[21]</sup>: 包含 850 幅多目标图像.

### 2.2 评价标准

采用准确率(Precision)-召回率(Recall)曲线<sup>[22]</sup>验证模型的性能.不同的阈值可能对评价有影响,因此设置自适应阈值范围为  $[0, 255]$ , 将显著图转化为二值图,再与真值图进行比较.每幅图像都有独立的自适应阈值,像素显著值大于该阈值的取 255, 小于则取 0.图像的自适应阈值,如(6)式所示.

$$T = \frac{2}{W \times H} \sum_{x=1}^W \sum_{y=1}^H S(x, y), \quad (6)$$

其中,  $W$  和  $H$  表示图像的宽和高,  $S(x, y)$  表示像素  $(x, y)$  的显著值.

因此,准确率  $P$ 、召回率  $R$  和  $F_\beta$  值可通过(7)式计算得出.

$$P = |M \cap G| / |M|, R = |M \cap G| / |G|, F_\beta = \frac{(1 + \beta^2) \times P \times R}{\beta^2 \times P + R}, \quad (7)$$

其中,  $M$  表示通过自适应阈值得到的二值图,  $G$  表示真值图,参照文献<sup>[22]</sup>设置  $\beta^2 = 0.3$ , 使得准确率权重大于召回率.

上述评价标准对于正确检测显著区域的方法较为公平,对于成功检测背景区域的方法并不公平.因此,计算平均绝对误差(Mean Absolute Error, MAE)<sup>[22]</sup>来进一步评价显著性检测性能,如(8)式所示.

$$MAE = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H |S(x, y) - G(x, y)|, \quad (8)$$

其中,  $S(x, y)$  表示像素  $(x, y)$  的显著值,  $G(x, y)$  表示像素  $(x, y)$  的真值值.

### 2.3 定性分析

一些具有代表性的视觉显著性对比结果,如图 3 所示.

从图 3 可以看出,本文模型生成的显著图更能突显图像的显著区域,并且相比于其他算法,可以较好地去除掉不必要的背景噪声(如图 3(k)第 1~13 幅图像);对于低对比度(如图 3(a)第 1、3~4、6~7、10~12 幅图像)和环境复杂(如图 3(a)第 2、5、8~9 幅图像)图像,依然可以取得较好的检测结果(如图 3(k)第 1~12 幅图像).以图 3 第 3 幅图像为例,多数对比算法将同金字塔相同颜色的地面(如图 3(b-d)、(g-j)第 3 幅图像)和类似颜色的天空(如图 3(d)、(g-j)第 3 幅图像)检测为显著目标,或是金字塔检测不完整(如图 3(e-f)第 3 幅图像),这是由于多数对比算法通常采用对比度计算图像显著图<sup>[1-2,6-8]</sup>,而这幅图像属于低对比度图像,很难通过对比度计算将金字塔与地面、天空良好地分离开.本文采用的深度卷积网络对图像特征提取具有独特的优势,可以有效提取图像的全局化特征,局部约束线性编码可以有效提取到目标的细节信息,并且使用贝叶斯框架将全局和局部特征融合,因此在目标轮廓和细节方面有较好的效果(如图 3(k)第 3 幅图像),本文模型准确提取到金字塔目标,并且有效抑制背景噪声.

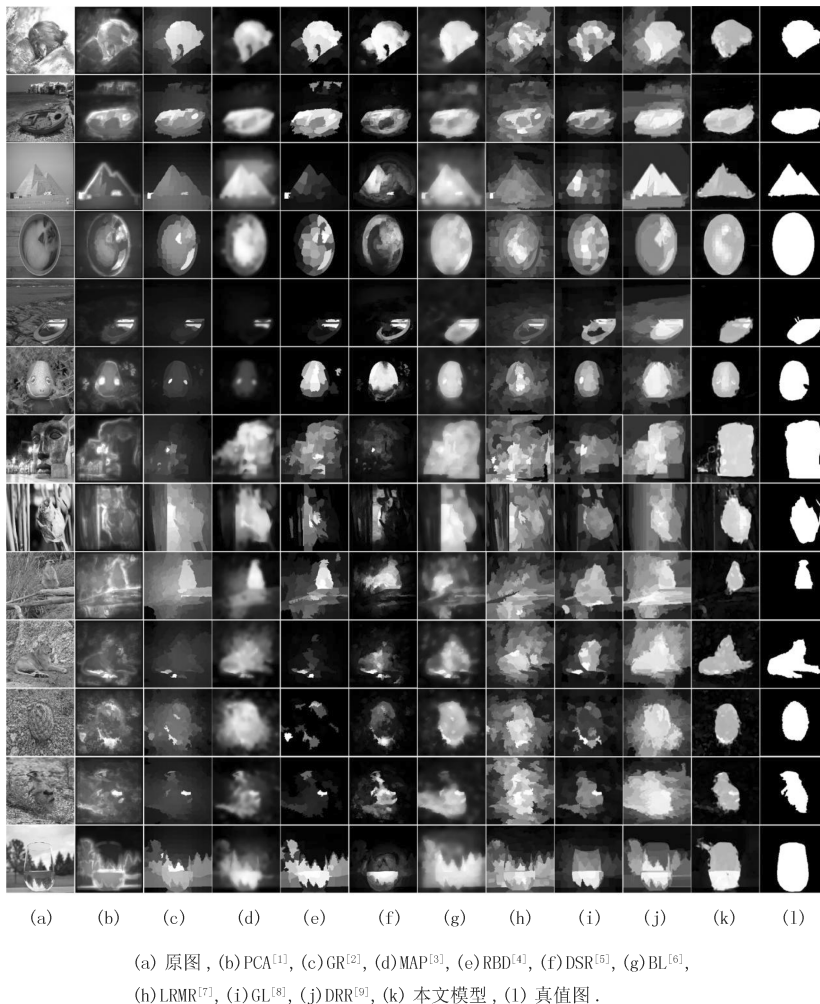


图 3 相关算法的视觉显著图

Fig.3 Visual saliency maps of related methods

## 2.4 定量分析

相关对比算法的准确率-召回率曲线如图 4 所示,以及定量对比结果如表 1 所示.

从图 4 可以看出,本文模型在 3 个标准数据集上的准确率-召回率曲线均在其他对比算法的上方,表明本文模型检测效果优于其他对比算法.

从表 1 可以看出,本文模型的  $F_{\beta}$  值均高于其他对比算法,并且 MAE 值均低于其他对比算法,这说明本文模型既可以更好地提取图像显著目标信息,又可以更好地检测图像背景区域.自编码模型本身对图像具有

强大的描述能力,卷积神经网络利用其参数共享的特性,可以提取到图像的平移、旋转、尺度不变等特征,随着卷积层数的增高,学习到的特征更加全局化,本文采用深度卷积自编码网络,有较好的显著目标整体轮廓提取结果;局部约束线性编码模型利用局部性和稀疏性来编码图像,使得结果更加注重目标的细节信息;贝叶斯模型利用其稳定的后验概率可以更加有效地融合图像的全局和局部信息.

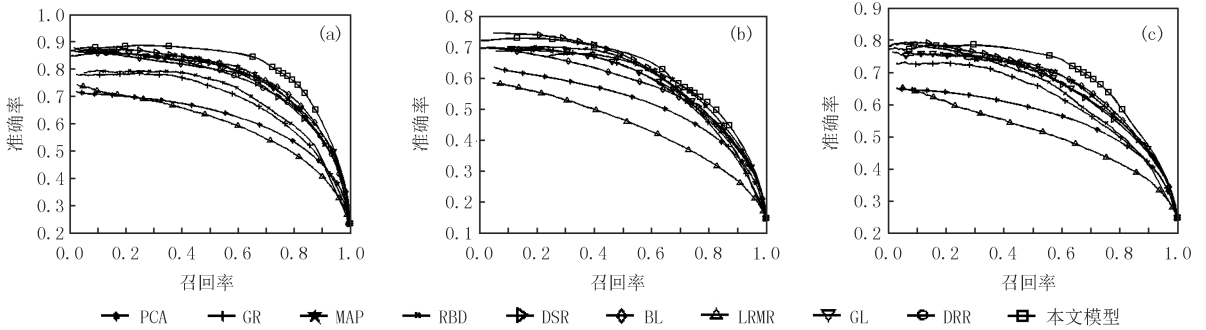


图4 准确率-召回率曲线 (a)ECSSD数据集, (b)DUT-OMRON数据集, (c)PASCAL数据集

Fig.4 Precision-Recall curve (a)ECSSD dataset, (b)DUT-OMRON dataset, (c)PASCAL dataset

表 1 相关算法在 ECSSD, DUT-OMRON, PASCAL 数据集上的定量比较

Tab.1 Quantitative comparison of related methods on ECSSD, DUT-OMRON, PASCAL datasets

数据集	指标	PCA	GR	MAP	RBD	DSR	BL	LRMR	GL	DRR	本文模型
ECSSD	$F_{\beta}$	0.574 1	0.627 8	0.734 3	0.687 3	0.724 4	0.717 9	0.529 1	0.732 3	0.685 7	0.765 3
	MAE	0.282 37	0.277 12	0.186 20	0.179 84	0.171 18	0.220 56	0.315 45	0.200 81	0.280 73	0.143 28
DUT-OMRON	$F_{\beta}$	0.495 2	0.571 0	0.556 9	0.564 8	0.563 2	0.545 4	0.449 4	0.569 4	—	0.594 5
	MAE	0.250 73	0.256 73	0.180 28	0.144 11	0.138 76	0.242 26	0.313 88	0.183 23	—	0.137 87
PASCAL	$F_{\beta}$	0.566 0	0.587 7	0.676 8	0.665 6	0.678 6	0.670 6	0.483 0	0.662 6	—	0.725 2
	MAE	0.289 05	0.293 14	0.225 81	0.210 18	0.204 92	0.250 99	0.337 00	0.232 48	—	0.181 05

### 2.5 实际应用

为验证本文模型的泛化能力,将本文模型应用于车辆检测和火焰检测中.车辆数据拍摄环境:采用 M5 Note 设备,3.5 mm 的焦距,F/2.2 的光圈,于河南省新乡市牧野区,中午 11:30 到 12:30,天气晴;火焰数据从公开数据(<http://signal.ee.bilkent.edu.tr/VisiFire/>)和网络搜集得到,其检测效果,如图 5 所示.

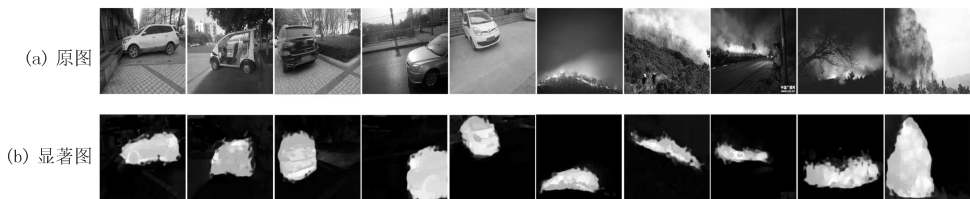


图 5 本文模型在车辆检测和火焰检测中的应用举例

Fig.5 Application examples of our model in vehicle detection and flame detection

从图 5 可以看出,本文模型对于不同型号的车辆以及不同形态的火焰均能较好地检测出来(如图 5(b)第 1~10 幅图像).本文模型的泛化能力主要来源于深度卷积自编码网络,深度网络由于其记忆数据的能力,具有良好的泛化性能,对除训练集以外的未知数据,有较好的检测效果.

### 3 结 论

为有效提取低对比度、复杂自然图像的显著信息,提出一种基于深度卷积自编码和局部约束线性编码的显著性检测模型.通过训练深度卷积自编码网络以产生全局显著图,从而使得模型能够更好地分离出显著目标;使用自适应阈值分割算法针对全局显著图产生图像的前景和背景码本,采用局部约束线性编码算法产生基于前景和背景码本的显著图,融合得到局部显著结果;将全局和局部显著图通过贝叶斯框架融合,使得显著结果同时包含全局和局部信息.下一步研究工作需对深度卷积神经网络本身的结构进行改进,加入诸如边界先验等先验知识,以使显著目标边缘更加清晰.

### 参 考 文 献

- [1] MARGOLIN R, TAL A, ZELNIK-MANOR L. What makes a patch distinct? [C]//2013 IEEE Conference on Computer Vision and Pattern Recognition, Washington D C: IEEE Computer Society, 2013: 1139-1146.
- [2] YANG C, ZHANG L H, LU H C. Graph-regularized saliency detection with convex-hull-based center prior[J]. IEEE Signal Processing Letters, 2013, 20(7): 637-640.
- [3] SUN J G, LU H C, LIU X P. Saliency region detection based on Markov absorption probabilities[J]. IEEE Transactions on Image Processing, 2015, 24(5): 1639-1649.
- [4] ZHU W J, LIANG S, WEI Y C, et al. Saliency optimization from robust background detection[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, Washington D C: IEEE Computer Society, 2014: 2814-2821.
- [5] LI X H, LU H C, ZHANG L H, et al. Saliency detection via dense and sparse reconstruction[C]//2013 IEEE International Conference on Computer Vision, Washington D C: IEEE Computer Society, 2013: 2976-2983.
- [6] TONG N, LU H C, RUAN X, et al. Salient object detection via bootstrap learning[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE press, 2015: 1884-1892.
- [7] WU Y, SHEN X H. A unified approach to salient object detection via low rank matrix recovery[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE press, 2012: 853-860.
- [8] TONG N, LU H C, ZHANG Y, et al. Salient object detection via global and local cues[J]. Pattern Recognition, 2015, 48(10): 3258-3267.
- [9] HAN J W, ZHANG D W, HU X T, et al. Background prior-based salient object detection via deep reconstruction residual[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25(8): 1309-1321.
- [10] 刘尚旺, 陈平, 王统昊. 实时语义图像分割模型研究[J]. 河南师范大学学报(自然科学版), 2016, 44(2): 139-147.  
LIU S W, CHEN P, WANG T H. Research on real-time semantic image segmentation model[J]. Journal of Henan Normal University (Natural Science Edition), 2016, 44(2): 139-147.
- [11] ZHANG J, DAI Y C, PORIKLI F. Deep salient object detection by integrating multi-level cues[C]//IEEE Winter Conference on Applications of Computer Vision, Piscataway: IEEE press, 2017: 1-10.
- [12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//International Conference on Learning Representations, Piscataway: IEEE press, 2015: 1-14.
- [13] TURCHENKO V, LUCZAK A. Creation of a deep convolutional auto-encoder in caffe[C]//IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, Piscataway: IEEE press, 2017: 651-659.
- [14] WANG J J, YANG J C, YU K, et al. Locality-constrained linear coding for image classification[C]//2010 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE press, 2010: 3360-3367.
- [15] ACHANTA R, SHAJI A, SMITH K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274-2282.
- [16] ZHU D D, LUO Y, DAI L, et al. Salient object detection via a local and global method based on deep residual network[J]. Journal of Visual Communication and Image Representation, 2018, 54: 1-9.
- [17] 杨咪, 屈文岗, 钱会. 基于熵权的贝叶斯模型及其在水质评价中的应用[J]. 灌溉排水学报, 2018, 37(1): 85-90.  
YANG M, QU W G, QIAN H. Bayesian model based on entropy weight and its application in water quality assessment[J]. Journal of Irrigation and Drainage, 2018, 37(1): 85-90.
- [18] CHENG M M, MITRA N J, HUANG X, et al. Global contrast based salient region detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 569-582.
- [19] SHI J P, YAN Q, XU L, et al. Hierarchical image saliency detection on extended cssd[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 38(4): 717-729.
- [20] YANG C, ZHANG L H, LU H C, et al. Saliency detection via graph-based manifold ranking[C]//2013 IEEE Conference on Computer

Vision and Pattern Recognition. Washington D C: IEEE Computer Society, 2013: 3166-3173.

- [21] LI Y, HOU X D, KOCH C, et al. The secrets of salient object segmentation[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. Washington D C: IEEE Computer Society, 2014: 280-287.
- [22] BORJI A, CHENG M M, JIANG H, et al. Salient object detection: a benchmark[J]. IEEE Transactions on Image Processing, 2015, 24(12): 5706-5722.

## Saliency detection via fused global and local information

Liu Shangwang<sup>a</sup>, Zhao Xinying<sup>a</sup>, Yang Lei<sup>b</sup>

(a. College of Computer and Information Engineering; b. College of Tourism, Henan Normal University, Xinxiang 453007, China)

**Abstract:** Aiming at improving the accuracy and generalization ability of natural images with low contrast and complex background, a saliency detection model based on Bayesian framework is proposed via fusing global and local information. Firstly, a deep convolution autoencoder network was constructed, and symmetrical encoder and decoder structure was adopted to supervise learning the global features, and global saliency map was generated; Secondly, foreground and background codebooks were obtained from the global saliency map, and the locality-constrained linear coding method was used for encoding, and the sparse coding was employed to describe the local features, so local saliency map was generated; Finally, the Bayesian framework was adopted to integrate the global with local information, and the final saliency map was thus obtained. The experimental results show that the F-measure of our model on ECSSD, DUT-OMRON and PASCAL datasets is 76.53%, 59.45% and 72.52% respectively, and the MAE is 0.143 28, 0.137 87 and 0.181 05 respectively. Furthermore, our method can effectively detect the salient object in real natural images with the low contrast and complex background.

**Keywords:** saliency detection; Bayesian framework; sparse coding; deep convolutional autoencoder network

[责任编辑 陈留院 赵晓华]