

# 基于无监督栈式降噪自编码网络的显著性检测算法

李庆武<sup>1,2</sup>, 马云鹏<sup>1,2</sup>, 周亚琴<sup>1,2</sup>, 邢俊<sup>1,2</sup>

(1. 河海大学物联网工程学院, 江苏常州 213022; 2. 常州市传感网与环境感知重点实验室, 江苏常州 213022)

**摘 要:** 针对现有的显著性检测算法检测目标类型单一、通用性差的问题, 提出一种基于无监督栈式降噪自编码网络的显著性检测算法. 该算法利用无监督栈式降噪自编码网络 (Stacked Denoising Auto Encoder, SDAE) 在多个尺度对原始图像进行稀疏重构, 将原始图像与 SDAE 网络重构图像之间的差作为显著图, 二值化后的显著图作为显著性目标检测结果. 在 SDAE 网络训练过程中, 将原始图像作为原始数据, 网络重构的图像作为观察数据. 为了提升网络训练效率, 首先利用无监督逐层贪婪方法训练同结构的深度信念网络 (Deep Belief Network, DBN), 将训练得到的 DBN 网络参数设为 SDAE 网络的初始参数, 再计算原始数据与观察数据之间的互信息作为网络收敛代价, 利用反向传播进行网络参数微调. 实验表明, 该网络模型可以完成多类型目标的显著性检测, 具有通用性好, 准确度高等优点.

**关键词:** 显著性检测; 无监督网络; 栈式降噪自编码; 深度信念网络; 互信息

**中图分类号:** TP751.1      **文献标识码:** A      **文章编号:** 0372-2112 (2019)04-0871-09

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2019.04.015

## Saliency Detection Based on Unsupervised SDAE Network

LI Qing-wu<sup>1,2</sup>, MA Yun-peng<sup>1,2</sup>, ZHOU Ya-qin<sup>1,2</sup>, XING Jun<sup>1,2</sup>

(1. College of Internet of Things Engineering, Hohai University, Changzhou, Jiangsu 213022, China;

2. Changzhou Key Laboratory of Sensor Networks and Environmental Sensing, Changzhou, Jiangsu 213022, China)

**Abstract:** The traditional saliency detection method is difficult to detect different kinds of saliency target simultaneously. In order to solve this problem, an algorithm based on unsupervised SDAE network is proposed in this paper. The stacked denoising auto-encoder (SDAE) network is used to sparsely reconstruct original image in multiple scales. The difference between the original image and the reconstructed image is used as a saliency map, and the binaryzation of the saliency map is used as salient detection result. In the process of SDAE network training, the original image is used as the original data and the reconstructed images are treated as observed data. In order to improve the efficiency of network training, the deep belief network (DBN) is trained by greedy method in each layer without supervising, and the network parameters are delivered to stacked denoising auto-encoder (SDAE) network as initial parameters. Then, the mutual information between the original data and the observed data is used as loss function, and the network parameters are tuned by backpropagation. The experiments show that the proposed algorithm can accomplish the saliency detection of various targets, which has the advantages of good universality and high accuracy.

**Key words:** saliency detection; unsupervised network; stacked denoising auto-encoder (SDAE); deep belief network (DBN); mutual information

## 1 引言

人类的视觉显著性注意机制是视觉系统中的重要组成部分, 它帮助人类在复杂的场景中实现快速目标

检测. 借鉴人类的视觉显著性注意机制研究显著性检测算法, 可广泛应用于目标检测与识别<sup>[1,2]</sup>、图像检索<sup>[3,4]</sup>、背景抑制<sup>[5]</sup>、计算资源分配等领域.

根据驱动因素的不同, 视觉显著性检测可以分为

自顶向下与自底向上两种检测模式. 自顶向下的模式根据先验信息与任务进行特征选取与融合, 再利用特征检测结果完成视觉显著性检测. 例如基于中心先验的显著性检测、基于背景先验的显著性检测等. Song 等人<sup>[6]</sup>以中心矩形构图为先验信息, 在中心矩形线上构建超级像素, 再根据超像素的流行排序检测图像的显著性区域. Zhu 等人<sup>[7]</sup>将图像块的中心区域作为显著性目标区域, 再对图像的边界连通性与继承性进行分析, 输出图像显著性检测结果. Li 等人<sup>[8]</sup>利用超像素与边缘信息构建背景先验模板, 利用多尺度重建误差检测图像显著性. 该类方法在固定类型的显著性目标检测中可以获得准确的显著性检测结果. 但在显著性目标类型及环境发生改变时, 该类方法的准确度明显降低, 鲁棒性差.

自底向上的模式对图像的底层特征进行分析与处理, 从而完成全局图像中显著性区域的检测. 例如基于颜色信息的显著性检测、基于灰度信息的显著性检测, Ma 等人<sup>[9]</sup>在 LUV 颜色空间计算邻域像素的颜色欧式距离, 根据图像局部颜色对比信息检测图像的显著性. Cheng 等人<sup>[10]</sup>利用图像灰度值分布的直方图计算对比度, 根据像素灰度值之间的差异进行图像显著像素点的检测. 该类方法以图像的基础特征为显著性判断依据, 具有更好的通用性, 但在相似性背景下、复杂环境中的检测准确度较差.

与传统的视觉显著性检测方法相比, 深度神经网络技术在多特征自适应提取、融合方面拥有巨大的优势. 例如, Huang 等人<sup>[11]</sup>提出训练多实例学习网络 (Multiple Instance Learning, MIL) 进行图像视觉显著性检测, 该方法在通用性与准确度方面都取得了很大的进步. 刘等人<sup>[12]</sup>将深度卷积神经网络和手工提取特征

相结合, 提出一种基于全局模型和局部搜索的显著性检测方法. 但利用神经网络进行显著性检测需要大量人工标定的数据样本, 资源消耗大.

为了提升视觉显著性检测的通用性、准确度, 本文提出了一种基于无监督栈式降噪自编码网络的显著性检测算法. 该算法利用无监督栈式降噪自编码网络在多个尺度对图像进行稀疏重构, 将原始图像与重构图像之间的互信息作为收敛代价进行无监督网络训练, 不需要人工标定的数据样本, 将原图像与重构图像之间的差作为显著图, 二值化后的显著图作为显著性目标检测结果. 实验结果表明, 该方法在多种类型的显著性目标检测中具有通用性好、准确度高等优点.

## 2 算法描述

本文提出的一种基于无监督栈式降噪自编码网络的显著性检测算法, 其处理流程如图 1 所示. 与传统的神经网络模型不同, 该算法根据待处理图像的信息无监督训练网络模型, 将原始图像分成若干图像块作为原始数据, 网络重构的图像作为观察数据, 不再需要人工标定的数据样本. 将原始图像作为原始数据进行网络训练, 提升了本文算法的对于当前图像显著性检测的有效性与准确度, 也保证了对于多种类图像的通用性.

在网络训练过程中, 为了提升训练效率, 本文算法将原始数据传递至 DBN 网络, 利用无监督逐层贪婪方法完成 DBN 网络训练. 然后将训练得到的 DBN 网络参数作为 SDAE 网络的初始参数, 再利用原始数据对 SDAE 网络进行训练, 将原始数据与观察数据之间的互信息作为网络收敛代价, 完成 SDAE 网络参数微调. 选取相同结构的 DBN 网络训练 SDAE 网络初始参数, 不仅提升了 SDAE 网络的收敛速度, 还可以避免网络陷入

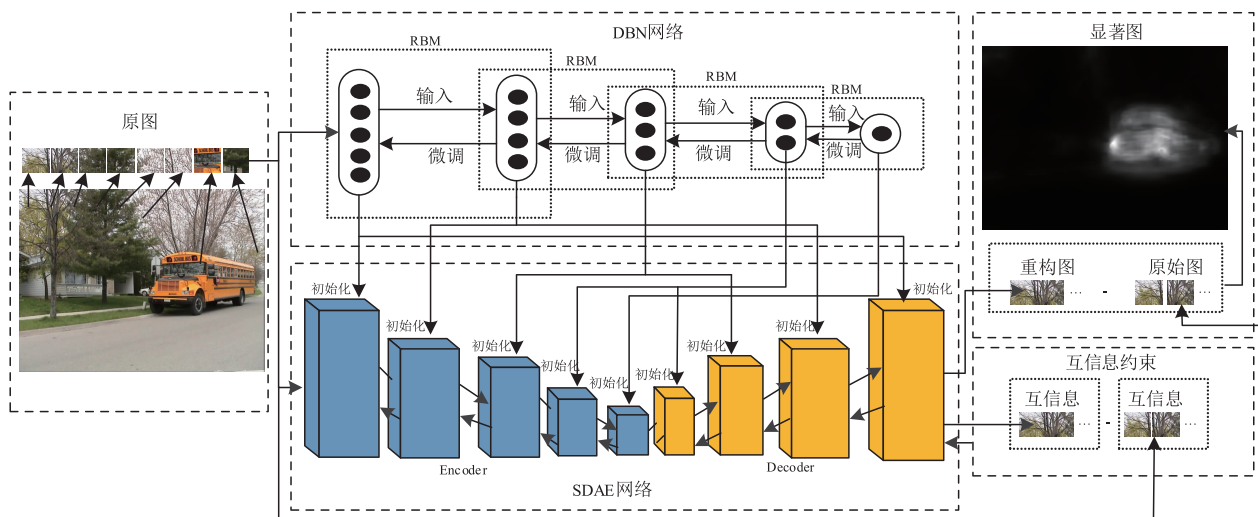


图1 本文方法处理流程图

局部最优. 在 SDAE 网络参数微调过程中, 选取原始数据与观察数据之间的互信息作为网络收敛代价, 当网络收敛代价达到最小值时, 完成 SDAE 网络训练, 保证了显著性区域检测的准确度与完整性.

在显著性检测过程中, 将原图像与 SDAE 网络重构图像的差作为显著图, 二值化后的显著图作为显著性目标检测结果.

## 2.1 DBN 网络参数训练

DBN 网络是一种无监督自编码的过程, 但逐层训练的特点也导致其存在易陷入局部最优解的缺点. 本文建立相同结构的 DBN 网络与 SDAE 网络, 将 DBN 网络训练得到的网络参数作为 SDAE 网络的初始参数, 再对网络参数进行反向微调, 不仅避免了网络陷入局部最优解, 而且提升了网络收敛速度<sup>[13]</sup>. 如图 2 所示, DBN 网络是由多个限制玻尔兹曼机 (Restricted Boltzmann Machines, RBM) 组成, DBN 网络的训练过程可以理解为 RBM 网络逐层的训练过程.

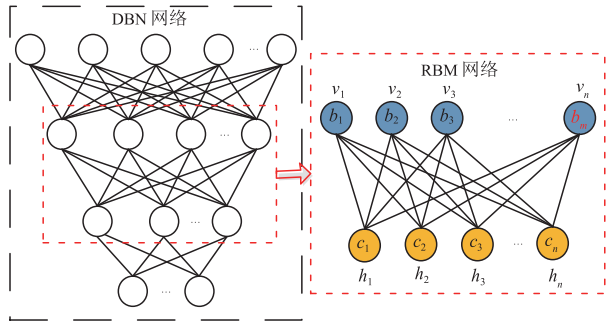


图2 DBN与RBM网络结构

如图 2 所示, RBM 网络共有两层, 第一层为可视层, 作为数据输入层, 另一层为隐藏层, 作为特征提取层<sup>[14]</sup>. 在无监督的训练过程中, 首先随机选取均值为 0、标准差为 0.1 的高斯分布中的数值初始化网络节点参数. 假设训练样本  $X$  的取值为  $X = (v_1, v_2, \dots, v_m)$ , 根据 RBM 网络, 可以得到这个样本的  $m$  维的编码后的样本  $Y = (h_1, h_2, \dots, h_n)$ . 这  $n$  维的编码也可以认为是抽取了  $n$  个特征的样本. 而  $m$  维的编码后的样本是按照以下的规则生成.

(1) 对于给定的  $X = (v_1, v_2, \dots, v_m)$ , 隐藏层的第  $i$  个元素的取值为 1 (编码后的样本的第  $i$  个特征的取值为 1) 的概率为

$$p(h_i = 1 | v) = \sigma(c_j + \sum_{i=1}^m v_i w_{ij}) \quad (1)$$

$$\sigma(x) = 1 / (1 + e^{-x}) \quad (2)$$

其中,  $v$  为可视训练样本集合  $X$ ,  $c_j$  是隐藏节点的偏移量, 设定节点元素取值  $h_i \in \{0, 1\}$ , 可以将  $h_i$  看作为编码后的样本  $Y$ ,  $w_{ij}$  为连接权重.

(2) 设  $b_i$  是可视节点的偏移量, 那么反向重构可视

单元中第  $i$  个元素的取值为 1 的概率为

$$p(v'_i = 1 | h) = \sigma(b_i + \sum_{j=1}^n h_j w_{ij}) \quad (3)$$

(3) 根据以下规则对 RBM 网络中的连接权重  $w_{ij}$ , 隐藏节点的偏移量  $c_j$ , 可视节点的偏移量  $b_i$  进行更新:

$$\begin{cases} \Delta w_{ij} = v_i(p(h_i = 1 | v_i)) - v'_i(p(h'_i = 1 | v'_i)) \\ \Delta c_j = p(h_j = 1 | v_j) - p(h'_j = 1 | v'_j) \\ \Delta b_i = v_i - v'_i \end{cases} \quad (4)$$

在训练 DBN 网络与 SDAE 网络的过程中, 将从原图像中随机提取的 10000 个图像块作为输入数据, 图像块的尺寸为  $7 \times 7 \times 3$  (单位: 像素), 分成 100 批次训练网络模型. 当 DBN 中每一层 RBM 网络逐步完成训练后, 将训练好的 DBN 网络参数对应地传递至相同结构的 SDAE 网络作为网络初始参数.

## 2.2 互信息约束的 SDAE 网络模型

SDAE 网络是通过多个降噪自编码器堆叠形成的深度学习网络. 由于 SDAE 的初始参数是由 DBN 网络训练得到, 再选取原图像与重构图像间的互信息作为网络收敛代价进行网络参数微调, 解决了传统的 SDAE 网络收敛难度的大的问题<sup>[15]</sup>.

如图 3 所示, 本文使用的 SDAE 网络模型具有 4 个编码层、4 个解码层、1 个中间层, 在网络编码器训练过程中, 将网络中上一层的输出作为下一层的输入. 设 SDAE 网络在编码部分具有  $L$  层, 则第  $k$  层的激活函数为

$$\begin{cases} y^{(k+1)} = f_e(W^{(k+1)}y^{(k)} + p^{(k+1)}) \\ f_e(W\tilde{x} + p) = \max(0, W\tilde{x} + p) \end{cases} \quad k = 0, 1, \dots, L-1 \quad (5)$$

$y^{(0)}$  为输入的原始图像块数据,  $y^{(L)}$  为编码器末层的输出, 并且  $y^{(L)}$  可以作为原始图像块数据的高维特征数据,  $W$  为连接系数,  $\tilde{x}$  为输入数据,  $p$  为编码层隐藏节点偏移量. 在解码器训练过程中, 将网络中上一层的输出作为下一层的输入, 则第  $k$  层的激活函数如式 (6) 所示, 其中  $z^{(0)}$  为解码器第一层输入的数据, 即  $y^{(L)}$ , 解码器末层输出的数据  $z^{(L)}$  是重构的原始图像数据,  $q$  为解码层隐藏节点偏移量.

$$\begin{cases} z^{(k+1)} = f_d(W^{(L-k)T}z^{(k)} + q^{(k+1)}) \\ f_d(a) = \begin{cases} \ln(1 + e^a), & \tilde{x} \in [0, 1] \\ a, & \tilde{x} \notin [0, 1] \end{cases} \end{cases} \quad k = 0, 1, \dots, L-1 \quad (6)$$

在 SDAE 网络训练过程中, 将原始输入图像与网络重构输出图像的互信息作为为网络收敛代价进行网络参数微调. 不同于相关系数, 互信息并不局限于实值随机变量, 它不仅可以表征联合分布  $p(X, Y)$ , 还反映了分解的边缘分布的乘积  $p(X)p(Y)$  之间的相似程度. 原始图像  $X$  和网络重构图像  $Y$  的互信息可以定义为

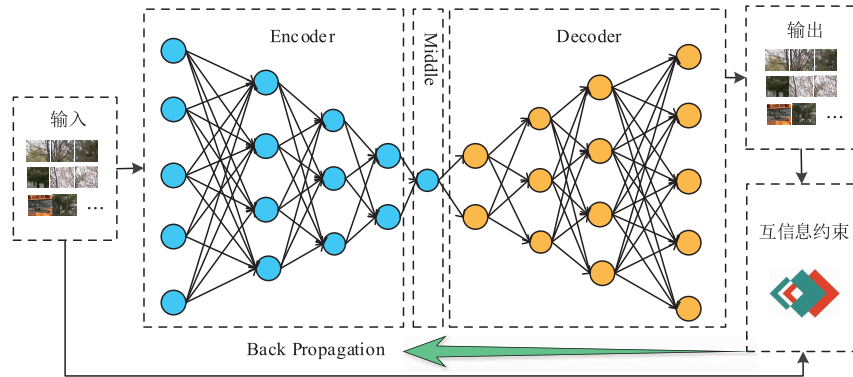


图3 互信息约束的SDAE网络模型

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \ln \left( \frac{p(x,y)}{p(x)p(y)} \right) \quad (7)$$

其中  $p(x,y)$  是原始图像  $X$  和网络重构图像  $Y$  的联合概率分布函数, 而  $p(x)$  和  $p(y)$  分别是原始图像  $X$  和网络重构图像  $Y$  的边缘概率分布函数. SDAE 网络训练与显著性检测的处理流程如下:

**步骤 1** 将 DBN 网络参数传递至 SDAE 网络作为网络初始化参数;

**步骤 2** 从原图随机提取若干个图像块, 图像块的尺寸为  $7 \times 7 \times 3$  (单位: 像素), 将提取的所有图像块作为原始图像数据分批输入至 SDAE 网络模型;

**步骤 3** 训练 SDAE 网络, 调整编码层与解码层的连接系数与节点偏移量;

**步骤 4** 计算原始图像与网络重构之间的互信息, 并将其作为网络收敛代价, 利用小批量梯度下降法的反馈信息微调 SDAE 网络参数;

**步骤 5** 重复以上步骤, 当网络收敛代价最小时, 完成网络训练, 并对原图像进行 SDAE 网络稀疏重构;

**步骤 6** 将原始图像与 SDAE 网络重构图像之间的差作为显著图, 再运用自适应阈值分割方法 Otsu 对显著图进行二值化处理, 得到显著性目标检测结果.

### 3 实验结果与分析

本文构建的网络模型结构及参数如表 1 所示, 其中 Hid1 ~ Hid4 为编码层, Hid6 ~ Hid9 为解码层, Hid5 为中间层. 为了定量分析本文提出算法在显著性检测方面的通用性与准确度, 我们选取了显著性检测公共数据集 MIT data set<sup>[16]</sup> 和 Toronto data set<sup>[17]</sup> 中 1123 幅图像对本文算法进行测试. 算法运行环境为 Intel Core I5, 3.00GHz CPU, 8GB 内存, 编译环境为 MATLAB R2014.

#### 3.1 显著图预测实验

在显著图预测对比实验中, 将本文算法得到的显著性分布图直接与其他方法得到的显著性分布图进行对比, 测试的数据集为本文选取的显著性检测公共数

据集 MIT data set 和 Toronto data set 中 1123 幅图像.

表 1 DBN 与 SDAE 网络模型结构及初始参数

参数	DBN 网络	SDAE 网络
可视节点数量	147	147
隐藏节点数量	Hid1 = 256, Hid2 = 128, Hid3 = 64, Hid4 = 32, Hid5 = 8	Hid1 = 256, Hid2 = 128, Hid3 = 64, Hid4 = 32, Hid5 = 8, Hid6 = 32, Hid7 = 64, Hid8 = 128, Hid9 = 256
学习率	Hid1 ~ Hid4 : 0.1 Hid5 : 0.001	Hid1 ~ Hid4, Hid6 ~ Hid9 : 0.01 Hid5 : 0.0001
正则化系数	0.0002	0.0001
初始动量	0.5	无
5 次迭代后动量	0.9	无
迭代次数	100	100

#### 3.1.1 显著图评价指标

在显著图预测实验中, 选取 Area Under the Curve (AUC) 与 Normalized Scanpath Saliency (NSS) 作为本文算法性能的评价指标.

##### (1) AUC

AUC 为 Receiver Operating Characteristic (ROC) 曲线下的区域面积<sup>[18]</sup>. 其中, ROC 是评价分类器性能的一种二维曲线, 其横坐标是 False Positive Rate (FPR), 纵坐标是 True Positive Rate (TPR). 通常情况下, AUC 的值介于 0.5 到 1.0 之间, AUC 的数值越大, 表示分类器的性能越好. 在定量分析过程, 将当前图像的显著性真值图中显著区域的像素点作为 Positive Samples, 随机选取真值图中同样数量的非显著性区域的像素点作为 Negative Samples, 输入检测结果获取该方法的 FPR 与 TPR 值, 并绘制 ROC 曲线. 经过多次循环试验, 取 AUC 的平均值作为当前算法的定量评价数据.

##### (2) NSS

NSS<sup>[19]</sup> 可以定量估计显著性检测结果与人眼注意力集中点之间的响应关系, 其计算方法如下

$$NSS = \frac{1}{N} \sum_{i=1}^N [(S(x_f^i) - \mu_s) / \sigma_s] \quad (8)$$

其中,  $N$  为人眼注意力分布图中的所有非零点, 也就是数据库提供的真值图中所有显著值非零点,  $S(x_j^i)$  为真值图中第  $i$  个注视点上显著性的值,  $\mu_s$  为显著性检测结果中显著值的平均值,  $\sigma_s$  为显著性检测结果中显著值的均方差. NSS 的值与显著性检测准确度正相关, NSS 的值越大代表显著性检测结果越符合人眼注意力的分布.

### 3.1.2 实验对比

将本文算法与现有 9 种显著图预测算法进行对比, 它们包括 Graph-Based Visual Saliency (GBVS)<sup>[20]</sup>, Attention based on Information Maximization (AIM)<sup>[21]</sup>, Spatially Weighted Dissimilarity (SWD)<sup>[22]</sup>, The Saliency Model from Itti. (Itti)<sup>[23]</sup>, Itti-Koch (IttiK) Saliency Model<sup>[24]</sup>, Random Rectangular Saliency (RCS)<sup>[25]</sup>, Visual Saliency Estimation using Region Covariances (CovSal)<sup>[26]</sup>, Learning Discriminative Subspaces (LDS)<sup>[27]</sup>, Unsupervised Hierarchical Models (UHM)<sup>[28]</sup>.

在显著图预测的实验中, 本文将评价指标的平均值作为当前测试算法的最终 AUC 值与 NSS 值. 图 4 为部分对比算法与本文算法的显著图预测结果, 表 2 为所有对比算法与本文算法显著图预测的定量分析结果. 由图 4 和表 2 可以看出, 与其他算法相比, 本文算法取得了更高的 AUC 值与 NSS 值, 具有准确度高的优点, 并且对包含多种类型显著性目标的图像拥有更加稳定的预测效果, 具有更好的通用性.

表 2 显著图预测定量分析

算法	GBVS <sup>[20]</sup>	AIM <sup>[21]</sup>	SWD <sup>[22]</sup>	Itti <sup>[25]</sup>	IttiK <sup>[24]</sup>
AUC	0.731	0.673	0.716	0.654	0.720
NSS	1.233	0.966	1.242	0.716	1.103
算法	RCS <sup>[25]</sup>	CovSal <sup>[26]</sup>	LDS <sup>[27]</sup>	UHM <sup>[28]</sup>	本文算法
AUC	0.720	0.673	0.666	0.719	<b>0.746</b>
NSS	1.103	1.181	1.126	1.143	<b>1.316</b>

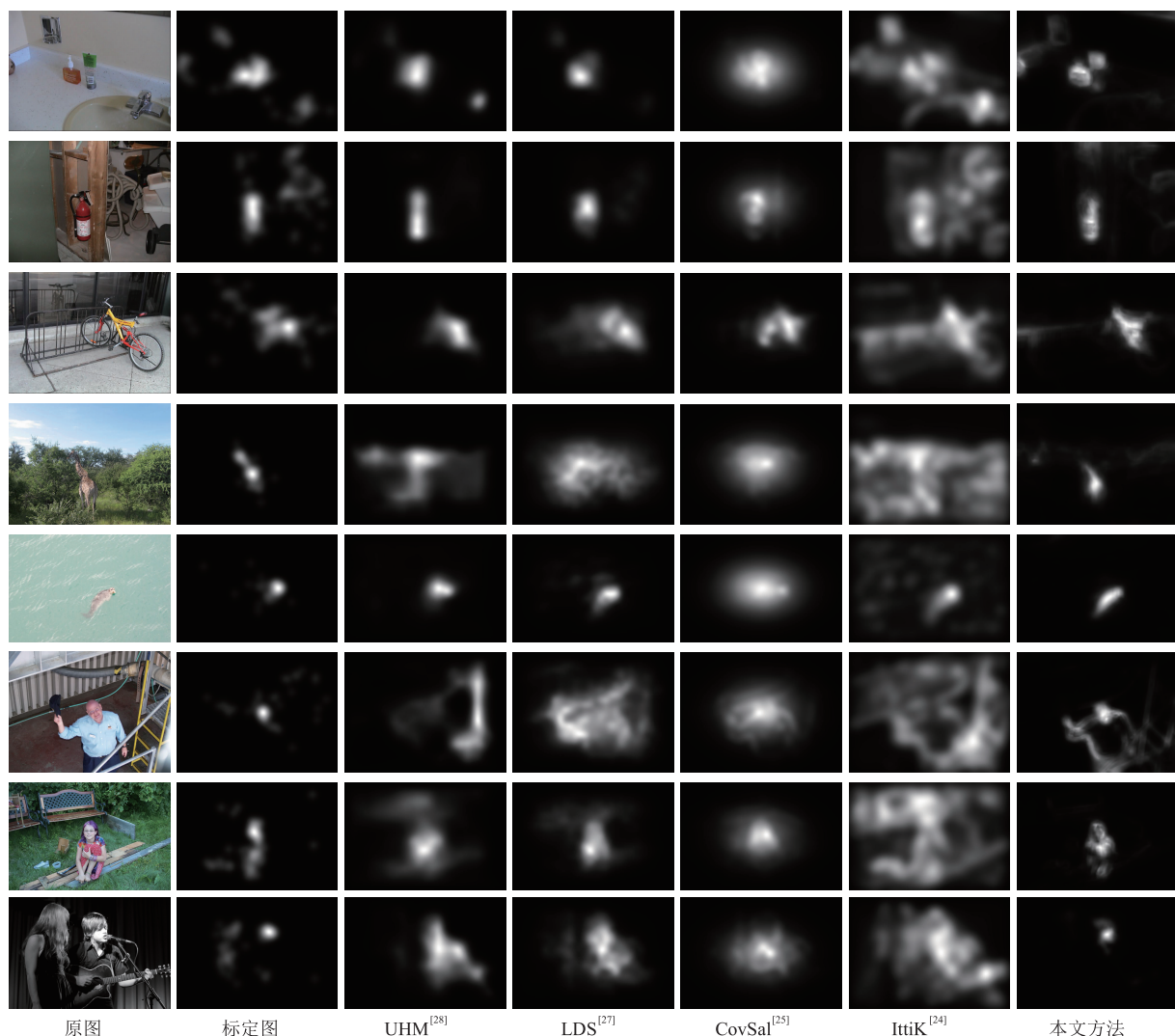


图4 显著图预测结果对比

### 3.2 显著性目标检测实验

在显著性目标检测实验中,选取自适应阈值分割法 Otsu 进行二值化,将二值化结果与现有算法显著性目标检测结果行对比.测试的数据集为显著性检测公共数据集 MIT data set 和 Toronto data set,显著性目标的标定图 GT(Ground-Truth)由显著性预测图中人眼注意力分布结果获得,即将注意力集中最多的完整目标分割出来作为显著性目标的标定图 GT.

#### 3.2.1 显著性目标检测评价指标

在显著性目标检测实验中,选取 F-measurement ( $F$  值)与平均绝对误差值 Mean absolute error (MAE) 评价算法性能.

##### (1) $F$ 值

如果单独使用查准率 (Precision,  $P$ ) 或者查全率 (Recall,  $R$ ) 进行评价,它们之间以负相关的关系互相影响,不能准确地完成图像分割结果的定量评价.  $F$  值为  $P$ - $R$  (Precision-Recall) 曲线中图像分割查准率  $P$  与查全率  $R$  的综合表达,可以定量地对图像分割结果进行评价, $F$  值越大,显著性目标检测的准确度越高.  $F$  值计算方法如式(9)所示, $\beta$  为权重参数,设  $\beta = 0.3$ ,  $GT(x, y)$  为人工标注图,  $BS(x, y)$  为分割二值图.

$$\begin{cases} F = \frac{(1 + \beta^2) \cdot P \cdot R}{\beta^2 \cdot P + R} \\ P = \frac{\sum_{(x,y)} GT(x,y) BS(x,y)}{\sum_{(x,y)} BS(x,y)} \\ R = \frac{\sum_{(x,y)} GT(x,y) BS(x,y)}{\sum_{(x,y)} GT(x,y)} \end{cases} \quad (9)$$

##### (2) MAE

MAE 为平均绝对误差值,其可以避免图像二值化过程中阈值对于结果评价的影响. MAE 通过计算分割结果中每个像素点的值  $S(x, y)$  与人工标注图  $GT(x, y)$  之间绝对差的均值对分割结果进行定量分析, MAE 的值越小代表准确度越高,其计算方法为

$$MAE = \frac{\sum_{x=1}^M \sum_{y=1}^N |S(x, y) - GT(x, y)|}{MN} \quad (10)$$

#### 3.2.2 实验对比

将本文算法与现有的 11 种显著性目标检测算法进行对比,它们包括 Luminance Contrast (LC)<sup>[29]</sup>, A Saliency Detection Method from Achanta (AC)<sup>[30]</sup>, Frequency-Tuned Salient Region Detection (FT)<sup>[31]</sup>, Context-Aware Saliency Detection (CA)<sup>[32]</sup>, Segmenting Salient Objects from Images and Videos (SEG)<sup>[33]</sup>, Histogram Contrast (HC)<sup>[16]</sup>, Principal Component Analysis (PCA)<sup>[34]</sup>, Dense and Sparse Re-

construction (DSR)<sup>[8]</sup>, Graph-Regularized (GR)<sup>[35]</sup>, Saliency Optimization from Robust Background Detection (RBD)<sup>[7]</sup>, Label Propagation Saliency (LPS)<sup>[36]</sup>, Multiple Instance Learning and Proposals for Saliency Detection (MILPS)<sup>[11]</sup>.

在显著性目标检测实验中,随机选取了 MIT data set 和 Toronto data set 数据库中 600 幅图像进行算法测试,最后求取评价指标的平均值作为当前测试算法的最终  $F$  值与 MAE 值. 图 5 为部分对比算法与本文算法的显著性目标检测结果,图 6 为所有对比算法与本文算法显著性目标检测结果的定量评价,图 6(a)、图 6(b)分别为算法的  $F$  值与  $F$  值误差柄,图 6(c)、图 6(d)分别为算法的 MAE 值与 MAE 值误差柄. 由图 5 和图 6 可以看出,与现有的算法相比,本文提出的算法取得了最高的  $F$  值与最低的 MEA 值,说明在本文算法得到的显著图中,运用自适应阈值分割方法得到的显著性目标检测结果具有准确度高的优点. 在误差柄图中,被测算法的指标数值以高概率落在误差柄线段包含的区间内,线段的长度反映了算法检测结果的不确定度,线段长度越短,表明算法的不确定度越低. 可以看出本文算法的  $F$  值误差柄与 MEA 误差柄的线段长度较短,按照从小到大排列分别处于第 4 和第 1 的位置,说明本文算法在多种环境下对多种显著性目标检测具有稳定性强、鲁棒性好的优点.

#### 3.3 算法时间消耗分析

本文分别通过处理不同尺寸图像、不同显著性区域面积的图像来进行算法时间消耗分析. 不同尺寸图像指的是内容相同、尺寸不同的多幅图像;不同显著性区域面积的图像指的是相同尺寸、显著性区域面积不同的多幅图像. 将本文算法与现有 9 种显著性检测算法进行对比,它们包括 GBVS<sup>[20]</sup>, AIM<sup>[21]</sup>, SWD<sup>[22]</sup>, Itti<sup>[23]</sup>, IttiK<sup>[24]</sup>, RCS<sup>[25]</sup>, CovSal<sup>[26]</sup>, LDS<sup>[27]</sup>, UHM<sup>[28]</sup>, 实验图像的原始分辨率都为  $681 \times 511$  (单位:像素). 时间消耗结果如图 7 所示,与现有的算法相比,本文算法对不同尺寸图像、不同显著性区域面积图像都有稳定的处理效率,时间消耗较短,具有较低的时间复杂度.

## 4 总结

本文提出了一种基于无监督栈式降噪自编码网络的显著性检测算法,该算法构建了一种根据图像信息对参数进行学习及优化的网络模型. 为了提升学习效率,该网络利用 DBN 网络进行预训练,获取网络初始参数. 在网络训练过程中,根据原始图像与重构图像之间的互信息对网络参数进行微调. 最后,将原始图像与 SDAE 网络重构图像的差作为显著图,将二值化后的显著图作为显著性目标检测结果. 与现有的显著性检测算法相比,该方法对不同种类的目标具有更好的通用性与准确度.

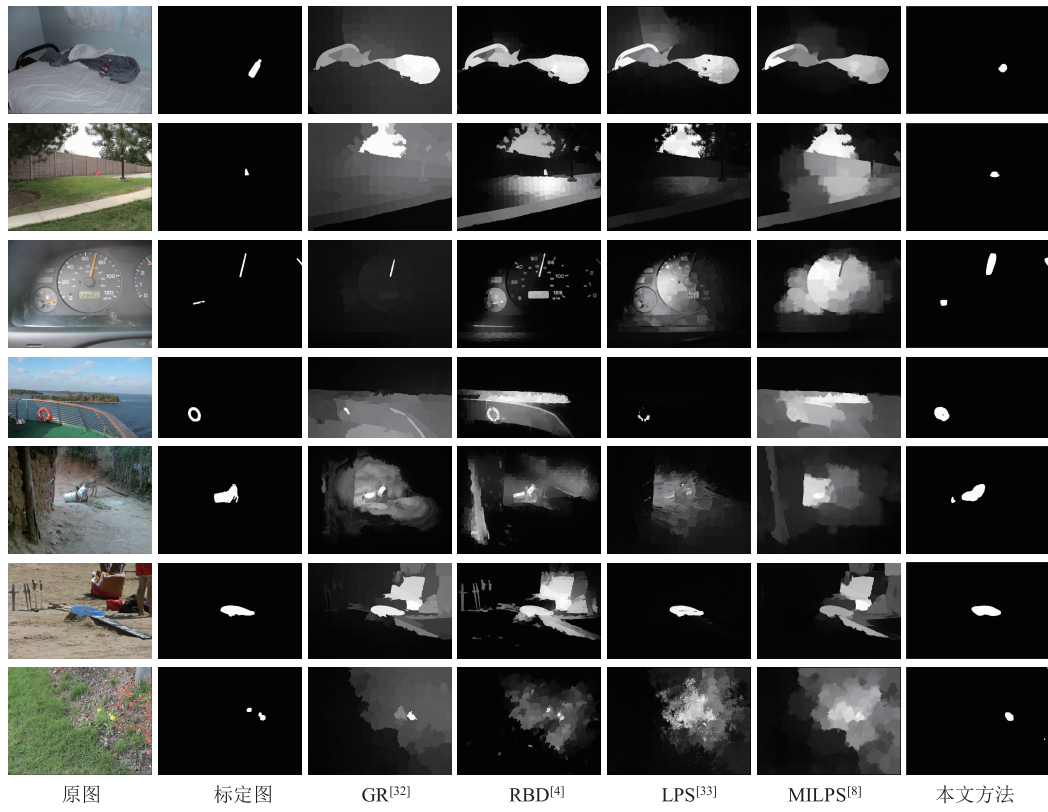


图5 显著性目标检测结果对比

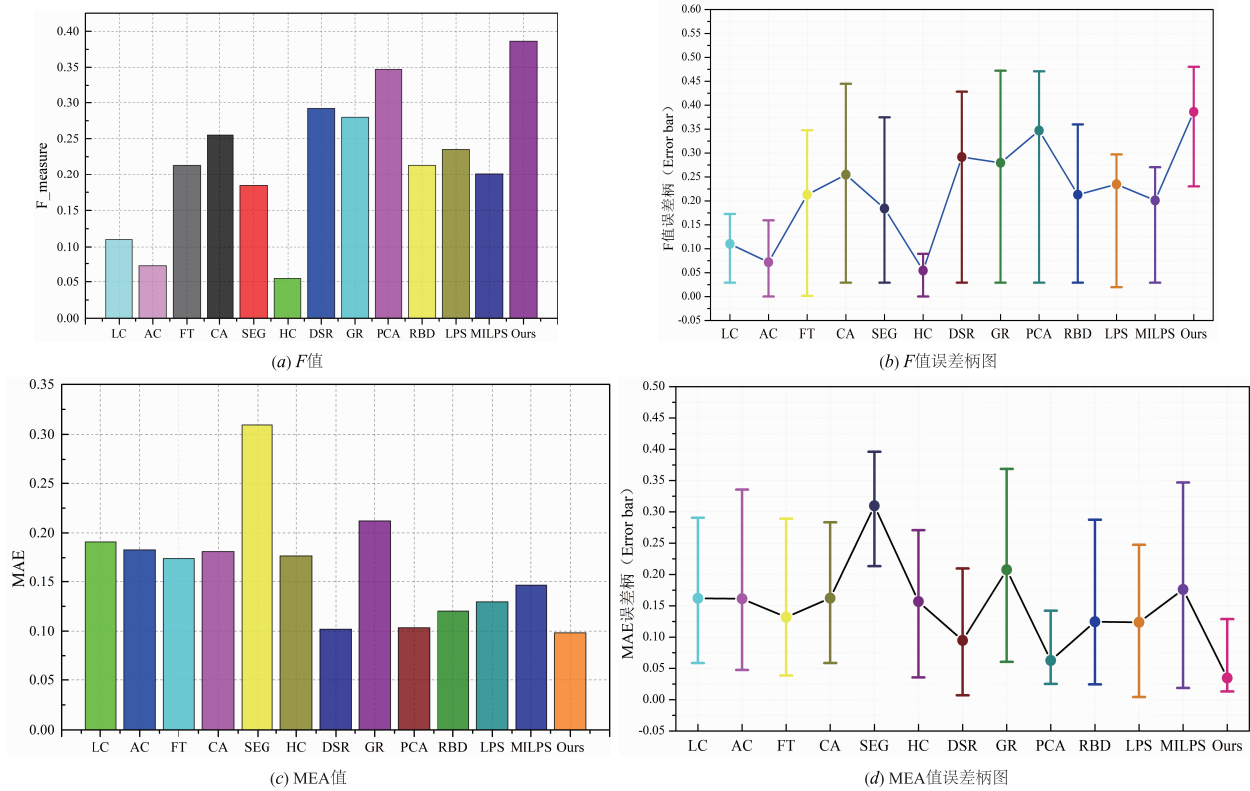


图6 显著性目标检测结果定量评价

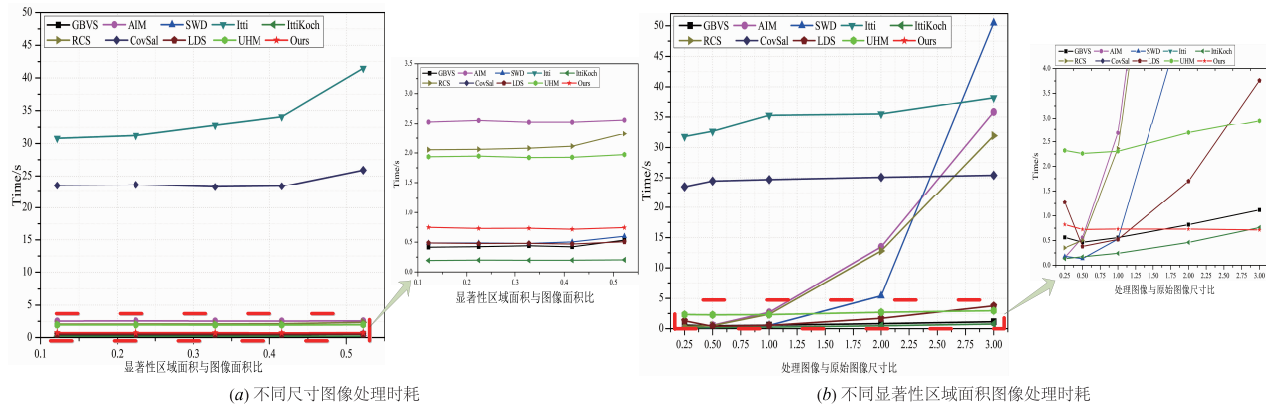


图7 算法时间消耗图

利用深度学习进行显著性检测是该领域未来重要的研究方向之一。本文提出的算法将当前处理图像作为网络训练的原始数据,不再需要人工标定的数据样本,根据图像信息进行深度学习,进而完成参数优化,实现显著图预测与显著性目标检测,具有重要的研究价值与较好的应用前景。

#### 参考文献

- [1] REN Z, GAO S, CHIA L T, et al. Region-based saliency detection and its application in object recognition[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2014, 24(5): 769 – 779.
- [2] 丁鹏, 张叶, 贾平, 等. 基于视觉显著性的海面舰船检测技术[J]. 电子学报, 2018, 46(1): 127 – 134.  
DING Peng, ZHANG Ye, JIA Ping, et al. Ship detection on sea surface based on visual saliency[J]. Acta Electronica Sinica, 2018, 46(1): 127 – 134. (in Chinese)
- [3] YANG X, QIAN X, XUE Y. Scalable mobile image retrieval by exploring contextual saliency[J]. IEEE Transactions on Image Processing, 2015, 24(6): 1709 – 1721.
- [4] 冯松鹤, 郎丛妍, 须德. 一种融合图学习与区域显著性分析的图像检索算法[J]. 电子学报, 2011, 39(10): 2288 – 2294.  
FENG Song-he, LANG Cong-yan, XU De. Combining graph learning and region saliency analysis for content-based image retrieval[J]. Acta Electronica Sinica, 2011, 39(10): 2288 – 2294. (in Chinese)
- [5] 姜维, 卢朝阳, 李静, 等. 基于视觉显著性与文字置信图的场景文字的背景抑制方法[J]. 电子学报, 2015, 43(9): 62 – 68.  
JIANG Wei, LU Zhao-yang, LI Jing, et al. Visual saliency and text confidence map based background suppression for scene text[J]. Acta Electronica Sinica, 2015, 43(9): 62 – 68. (in Chinese)
- [6] 宋腾飞, 刘政怡. 中心矩形构图先验的显著目标检测[J]. 中国图象图形学报, 2017, 22(3): 315 – 326.
- [7] SONG T-F, LIU Z-Y. Saliency detection based on center rectangle composition prior [J]. Journal of Image and Graphics, 2017, 22(3): 315 – 326. (in Chinese)
- [7] ZHU W, LIANG S, WEI Y, et al. Saliency optimization from robust background detection[A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C]. US: IEEE CVPR, 2014. 2814 – 282.
- [8] LI X, LU H, ZHANG L, et al. Saliency detection via dense and sparse reconstruction[A]. IEEE International Conference on Computer Vision[C]. US: IEEE ICCV, 2013. 2976 – 2983.
- [9] MA Y F, ZHANG H J. Contrast-based image attention analysis by using fuzzy growing[A]. Eleventh ACM International Conference on Multimedia[C]. US: ACM, 2003. 374 – 381.
- [10] CHENG M M, ZHANG G X, MITRA N J, et al. Global contrast based salient region detection[A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C]. US: IEEE CVPR, 2011. 409 – 416.
- [11] HUANG F, JINQING Q, LU H, et al. Salient object detection via multiple instance learning[J]. IEEE Transactions on Image Processing, 2017, 26(4): 1911 – 1922.
- [12] 刘峰, 沈同圣, 娄树理, 等. 全局模型和局部优化的深度网络显著性检测[J]. 光学学报, 2017, 37(12): 1215005.  
LIU F, SHEN T-S, LOU S-L, et al. Deep saliency detection based on global model and local optimization[J]. Acta Optica Sinica, 2017, 37(12): 1215005. (in Chinese)
- [13] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527.
- [14] ROUX N L, BENGIO Y. Representational power of restricted boltzmann machines and deep belief networks[J]. Neural Computation, 2008, 20(6): 1631 – 1649.
- [15] SAINATH T N, KINGSBURY B, RAMABHADRAN B.

- Auto-encoder bottleneck features using deep belief networks [A]. IEEE International Conference on Acoustics, Speech and Signal Processing [C]. US: IEEE, 2012. 4153 – 4156.
- [16] JUDD T, EHINGER K, DURAND F, et al. Learning to predict where humans look [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. US: IEEE CVPR, 2009. 2106 – 2113.
- [17] BRUCE N D B, TSOTSOS J K. Saliency based on information maximization [A]. International Conference on Neural Information Processing Systems [C]. US: MIT Press 2005. 155 – 162.
- [18] BORJI A, ITTI L. State-of-the-art in visual attention modeling [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(1): 185.
- [19] RICHE N, MANCAS M, DUVINAGE M, et al. RARE2012: A multi-scale rarity-based saliency detection with its comparative statistical analysis [J]. Signal Processing Image Communication, 2013, 28(6): 642 – 658.
- [20] HAREL J, KOCH C, PERONA P. Graph-based visual saliency [J]. Advances in Neural Information Processing Systems, 2007, 19: 545 – 552.
- [21] BRUCE N, TSOTSOS J. Attention based on information maximization [J]. Journal of Vision, 2007, 7(9): 950 – 950.
- [22] DUAN L, WU C, MIAO J, et al. Visual saliency detection by spatially weighted dissimilarity [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. US: IEEE CVPR, 2011. 473 – 480.
- [23] BORJI A, ITTI L. Exploiting local and global patch rarities for saliency detection [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. US: IEEE CVPR, 2012. 478 – 485.
- [24] HOU X, HAREL J, KOCH C. Image signature: highlighting sparse salient regions [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(1): 194.
- [25] VIKRAM T N, TSCHEREPANOW M, WREDE B. A saliency map based on sampling an image into random rectangular regions of interest [J]. Pattern Recognition, 2012, 45(9): 3114 – 3124.
- [26] EEDEM E, ERDEM A. Visual saliency estimation by nonlinearly integrating features using region covariances [J]. Journal of Vision, 2013, 13(4): 11.
- [27] FANG S, LI J, TIAN Y, et al. Learning discriminative subspaces on random contrasts for image saliency analysis [J]. IEEE Transactions on Neural Networks & Learning Systems, 2017, 28(5): 1095.
- [28] TAVAKOLI H R, LAAKSONEN J. Bottom-up fixation prediction using unsupervised hierarchical models [A]. Computer Vision ACCV 2016 Workshops [C]. Berlin Heidelberg: Springer, 2017. 287 – 302.
- [29] ZHAI Y, SHAH M. Visual attention detection in video sequences using spatiotemporal cues [A]. ACM International Conference on Multimedia [C]. US: ACM, 2006. 815 – 824.
- [30] ACHANTA R, ESTRADA F, WILS P, et al. Salient region detection and segmentation [A]. Proc ICVS [C]. 2008, 5008: 66 – 75.
- [31] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency-tuned salient region detection [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. US: IEEE CVPR, 2009. 1597 – 1604.
- [32] GOFERMAN S, ZELNIK-MANOR L, TAL A. Context-aware saliency detection [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. US: IEEE CVPR, 2010. 2376 – 2383.
- [33] RAHTU E, KANNALA J, SALO M, et al. Segmenting salient objects from images and videos [A]. Computer Vision, ECCV 2010 [C]. Berlin Heidelberg: Springer, 2010. 366 – 379.
- [34] Ran M, Tal A, ZELNIK-MANOR L. What makes a patch distinct? [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. US: IEEE CVPR, 2013. 1139 – 1146.
- [35] YANG C, ZHANG L, LU H. Graph-regularized saliency detection with convex-hull-based center prior [J]. IEEE Signal Processing Letters, 2013, 20(7): 637 – 640.
- [36] LI H, LU H, LIN Z, et al. Inner and inter label propagation; salient object detection in the wild [J]. IEEE Transactions on Image Processing, 2015, 24(10): 3176 – 3186.

#### 作者简介



**李庆武** 男, 1964年12月出生, 河南新乡人. 分别于郑州大学获得学士学位、西安电子科技大学获得硕士学位、河海大学获得博士学位. 现为河海大学物联网工程学院副院长、教授、博士生导师. 主要研究方向为智能视觉感知、信息获取与智能系统.  
E-mail: li\_qingwu@163.com



**马云鹏** 男, 1993年8月出生, 山东鄄城人. 2015年于河海大学获得学士学位, 现为河海大学博士研究生. 主要研究方向为数字图像处理、智能视觉感知.  
E-mail: yunpengma\_hhu@163.com