# 基于加权的K近邻线性混合显著性目标检测

李 炜 李全龙 刘政怡*

(安徽大学计算机科学与技术学院 合肥 230601)

**摘 要**：显著性目标检测旨在于一个场景中自动检测能够引起人类注意的目标或区域，在自底向上的方法中，基于多核支持向量机(SVM)的集成学习取得了卓越的效果。然而，针对每一张要处理的图像，该方法都要重新训练，每一次训练都非常耗时。因此，该文提出一个基于加权的K近邻线性混合(WKNNLB)显著性目标检测方法：利用现有的方法来产生初始的弱显著图并获得训练样本，引入加权的K近邻(WKNN)模型来预测样本的显著性值，该模型不需要任何训练过程，仅需选择一个最优的K值和计算与测试样本最近的K个训练样本的欧式距离。为了减少选择K值带来的影响，多个加权的K近邻模型通过线性混合的方式融合来产生强的显著图。最后，将多尺度的弱显著图和强显著图融合来进一步提高检测效果。在常用的ASD和复杂的DUT-OMRON数据集上的实验结果表明了该算法在运行时间和性能上的有效性和优越性。当采用较好的弱显著图时，该算法能够取得更好的效果。

**关键词**：显著性目标检测；集成学习；线性混合；加权的K近邻

**中图分类号**：TP391 **文献标识码**：A **文章编号**：1009-5896(2019)10-2442-08

**DOI**: 10.11999/JEIT190093

# Salient Object Detection Using Weighted K-nearest Neighbor Linear Blending

LI Wei LI Quanlong LIU Zhengyi

(*College of Computer Science and Technology*, *Anhui University*, *Hefei* 230601, *China*)

**Abstract**: Salient object detection which aims at automatically detecting what attracts human's attention most in a scene, bootstrap learning based on Support Vector Machine(SVM) has achieved excellent performance in bottom-up methods. However, it is time-consuming for each image to be trained once based on multiple kernel SVM ensemble. So a salient object detection model via Weighted K-Nearest Neighbor Linear Blending (WKNNLB) is proposed. First of all, existing saliency detection methods are employed to generate weak saliency maps and obtain training samples. Then, Weighted K-Nearest Neighbor (WKNN) is introduced to learning salient score of samples. WKNN model needs no pre-training process, only needs selecting K value and computing saliency value by the K-nearest neighbors labels of training sample and the distances between the K-nearest neighbors training samples and the testing sample. In order to reduce the influence of selecting K value, linear blending of multi-WKNNs is applied to generating strong saliency maps. Finally, multi-scale saliency maps of weak and strong model are integrated together to further improve the detection performance. The experimental results on common ASD and complex DUT-OMRON datasets show that the algorithm is effective and superior in running time and performance. It can even perform favorable against the state-of-the-art methods when adopting better weak saliency map.

**Key words**: Salient object detection; Bootstrap learning; Linear blending; Weighted K-Nearest Neighbor(WKNN)

## 1 引言

显著性目标检测旨在模拟人类的视觉注意力机制去关注一个场景中有区别的目标或者区域，在心理学，神经科学和计算机视觉等广泛领域引起了大量的关注。显著性目标检测有广泛的应用如目标识别[1]、图像压缩[2]、场景分类[3]、图像检索[4]和弱监督学习[5]等。尽管显著性目标检测在近些年取得了很大的进步，但在运行时间及效率上仍然是一个挑战性的任务。在近几年来，大量的关于显著性目标

收稿日期：2019-02-01；改回日期：2019-06-03；网络出版：2019-06-12
*通信作者： 刘政怡 liuzywen@ahu.edu.cn

检测的算法被提出。显著性目标检测算法可以简单地划分为基于传统的方法和基于机器学习的方法。目前，基于机器学习的一个分支是深度学习的方法，尤其是全卷积神经网络[6–10]利用图像的高层语义信息取得了惊人的效果。然而，该方法也有其自身固有的缺点：第一，需要大量的算力，强烈依赖于GPU；第二，需要大量的存储空间来保存上亿个模型参数；第三，需要花费大量的时间去调节超参数来训练出一个优秀的模型；第四，需要大量的人工标记的图像去有监督地训练模型。

大量的传统方法[11–13]启发式地使用了一些先验信息，例如局部或者全局对比先验，中心先验，边界先验等，文献[14]有更多的基于传统方法的细节，本文则将重点关注基于机器学习的方法。这类方法是在机器学习的基础上构建的，其中结合了高层次信息和监督方法，以提高显著性图的准确性。文献[15,16]通过集成学习的方式提出了显著性目标检测算法，对每张图像训练一个含有多个支持向量机模型的集合，并通过集成学习的方式融合成一个多核支持向量机模型。文献[17]使用多核支持向量机集成模型来处理RGB-D图像分割问题。文献[18]使用了同样的方式来处理RGB-D协同显著问题。文献[15–18]都是使用基于多核支持向量机集成的方式来构造模型。但是，这些方法处理每张或者每组图片，都需要在训练上消耗较多的时间。

## 2　加权的$k$近邻线性混合模型

如图1所示，本文首先将输入图像分割成多个尺度的超像素，然后将每个尺度的超像素依次通过现有的显著性检测方法产生对应尺度的弱显著图，再通过2.4节介绍的采集样本的方法从多尺度的超像素中采集出训练样本；其次将每个尺度的超像素和训练样本送入本文提出的强显著模型，从而计算出对应尺度的强显著图，然后分别将多尺度的弱、强显著图融合产生弱显著图和强显著图；最后通过将弱显著图和强显著图融合在一起生成最后的显著图。

### 2.1　多尺度超像素分割

本文以超像素为基本单位进行图像的显著性目标检测。超像素是具有相似纹理、颜色、亮度等特征的相邻像素构成的像素块，它利用像素之间特征的相似性将像素分组，用少量的超像素代替大量的像素来表达图像特征，很大程度上降低了图像处理的复杂度。超像素分割的尺度不同，影响着显著目标的检测精度。单一的尺度不能适应不同尺寸的目标，多尺度既可以捕捉到大尺寸的目标也可以捕捉到显著目标的细节。因此，本文采用多尺度的超像

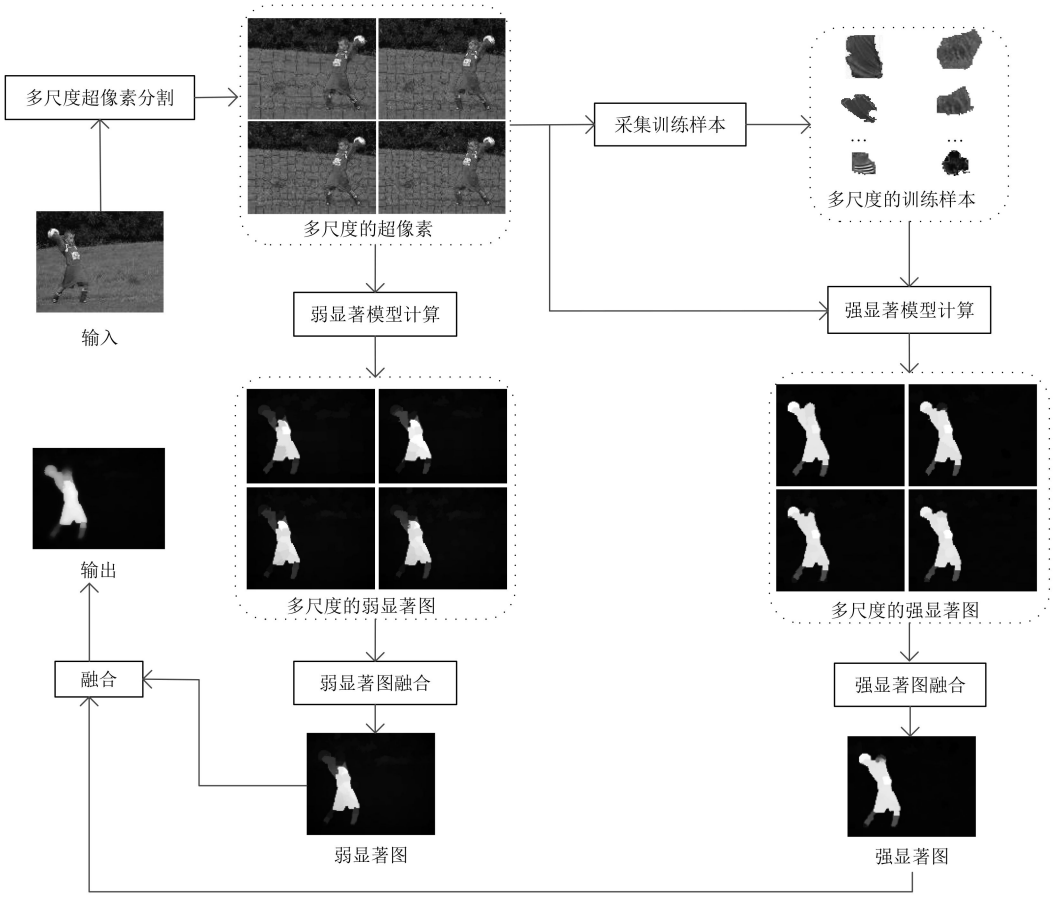



图 1 本文方法的框架图

素分割方法，利用经典的SLIC算法[19]将图像$\boldsymbol{I}$按照$M$个尺度进行分割，形成$M$个尺度下的图像$\{\boldsymbol{I}^j|j=1,2,\cdots,M\}$，其中$\boldsymbol{I}^j=\{\boldsymbol{r}_i^j|i=1,2,\cdots,N_j\}$，$N_j$表示图像$\boldsymbol{I}$按第$j$个尺度分割时的超像素总数。

### 2.2 图像特征

本文使用和BLSVM相同的特征来描述每一个超像素。具体而言，使用RGB, CIELab和LBP这3个特征来表示每一个超像素$\boldsymbol{r}_i^j$，即这3个特征级联成1个特征向量$\boldsymbol{f}_i^j$，它的前3维为超像素的RGB颜色空间特征值，随后的3维表示CIELab颜色空间特征值，最后的59维为LBP特征。前6维的计算公式为

$$\boldsymbol{f}_i^{j(u)}=\frac{1}{|\boldsymbol{r}_i^j|}\sum_{p\in\boldsymbol{r}_i^j}u(p),u\in\{R,G,B,L,a,b\} \tag{1}$$

其中，$\boldsymbol{f}_i^{j(u)}$表示超像素$\boldsymbol{r}_i^j$在$R$, $G$, $B$, $L$, $a$和$b$这6个通道上的特征值，$p$为超像素$\boldsymbol{r}_i^j$中的像素点，$u(p)$表示像素点$p$对应的通道$u$上的特征值，$|\boldsymbol{r}_i^j|$表示超像素$\boldsymbol{r}_i^j$中像素点的个数。本文使用LBP直方图值来表示超像素的纹理特征，最后的59维计算公式为

$$\boldsymbol{f}_i^{j(v)}=|U_{p\in\boldsymbol{r}_i^j}(I_{q\in\text{patch}(p)}(g(p),g(q)))|_v,\quad v=1,2,\cdots,59 \tag{2}$$

其中，$\boldsymbol{f}_i^{j(v)}$表示超像素$\boldsymbol{r}_i^j$的LBP直方图为$v$的特征值，patch$(p)$表示以像素点$p$为中心的周围8个像素点集合，$g(p)$表示像素点$p$的灰度值，$I(g(p),g(q))$是指示函数，当$g(p)\le g(q)$时取1，反之取0，所以通过$I_{q\in\text{patch}(p)}(g(p),g(q))$即得到像素点$p$的8位二进制数。$U(\cdot)$是文献[20]提出的等价模式(uniform pattern)方法，将8位二进制数转换成[1,59]之间的一个整数值，$|\cdot|_v$表示值为$v$的个数。在本文中，所有的特征值均归一化到0～1之间。关于特征的详细取值请参考表1。

### 2.3 弱显著模型

本文将输入图像通过弱显著模型，来获得弱显著图。任何显著性检测的方法都可以作为弱显著模型，例如基于对比先验，中心先验和暗通道先验等方法。为了公平地和BLSVM方法对比，本文使用相同的弱显著模型，即超像素的显著值计算为

$$\boldsymbol{S}_{\text{weak}}(\boldsymbol{r}_i^j)=g(\boldsymbol{r}_i^j)\cdot s_d(\boldsymbol{r}_i^j)\cdot\sum_{c\in\{F_1,F_2,F_3\}}\left(\frac{1}{N_B}\sum_{h=1}^{N_B}d_c(\boldsymbol{r}_i^j,\boldsymbol{n}_h^j)\right) \tag{3}$$

**表 1 特征$\boldsymbol{f}_i^j$取值(65维)**

| 特征维度序号 | 特征 | 维度大小 | 取值范围 |
|---|---|---|---|
| 0～2 | 平均RGB值 | 3 | [0,1] |
| 3～5 | 平均CIELab值 | 3 | [0,1] |
| 6～64 | LBP直方图值 | 59 | [0,1] |

其中，$\boldsymbol{r}_i^j$表示第$j$个尺度的第$i$个超像素，$g(\boldsymbol{r}_i^j)$为中心先验权重，$s_d(\boldsymbol{r}_i^j)$表示超像素$\boldsymbol{r}_i^j$内的平均暗通道值，$F_1,F_2,F_3$分别表示RGB, CIELab和LBP特征，$N_B$是位于图像边缘区域的超像素的个数，$\boldsymbol{n}_h^j$表示第$j$个尺度下位于图像边缘区域的第$h$个超像素，$d_c(\boldsymbol{r}_i^j,\boldsymbol{n}_h^j)$表示在$c$所对应的特征上超像素$\boldsymbol{r}_i^j$和$\boldsymbol{n}_h^j$之间的特征欧式距离。式(3)计算出了每个超像素的显著值$\boldsymbol{S}_{\text{weak}}(\boldsymbol{r}_i^j)$，为了得到像素级别的弱显著图，需要将超像素的显著值分配给像素，只要像素点$\boldsymbol{I}_l^j\in\boldsymbol{r}_i^j,l=1,2,\cdots,N$, 则$\boldsymbol{S}_{\text{weak}}(\boldsymbol{I}_l^j)=\boldsymbol{S}_{\text{weak}}(\boldsymbol{r}_i^j)$, 其中$N$表示图像$\boldsymbol{I}$中的像素总数。通过像素显著值的分配，最终得到$j$个尺度下的弱显著图$\boldsymbol{WM}^j=\{\boldsymbol{S}_{\text{weak}}(\boldsymbol{I}_l^j)|l=1,2,\cdots,N\}$。

### 2.4 采集样本

最近，机器学习方法被广泛地应用到显著性目标检测上，使用机器学习方法的关键点是需要使用人工标记的样本去训练出一个模型。本文通过现有的弱显著模型得到弱显著图，并从多尺度的超像素中采集训练样本，学习出一个强显著模型WKNNLB。训练样本集中的正样本来自于显著前景部分，负样本来自于背景部分。本文参考文献[15]，设定固定阈值来分割正负样本。获取正样本的阈值设置为$\text{TH}_p^j=1.5\bar{S}^j$，其中$\bar{S}^j$在这里指的是第$j$个尺度的弱显著图中所有像素的显著值的平均值，即

$$\bar{S}^j=\frac{1}{N}\sum_{l=1}^{N}\boldsymbol{S}_{\text{weak}}(\boldsymbol{I}_l^j) \tag{4}$$

若超像素的显著值大于$\text{TH}_p^j$，则分配为正样本，标签为+1。获取负样本的阈值设置为$\text{TH}_n^j=0.05$，若超像素的显著值小于$\text{TH}_n^j$，则分配为负样本，标签为–1。考虑到显著性目标通常出现在图像的中心区域，而并非边界，所以沿着图像边界的超像素也被作为负样本，标签为–1。

训练样本的个数会直接影响模型的性能，本文在多尺度的弱显著图上来给对应的多尺度的超像素分配显著值，通过与阈值的比较来收集训练样本。所有的训练样本被表示为$\boldsymbol{D}_{\text{T}}=\{\boldsymbol{x}_i,y_i\}_{i=1}^H$, $\boldsymbol{x}_i$表示选作训练样本的超像素，$y_i$表示其标签(-1或+1)，$H$为多尺度的正负训练样本的总数。测试样本是图像在$M$个尺度下的超像素全体$\{\boldsymbol{r}_i^j|j=1,2,\cdots,M,i=1,2,\cdots,N_j\}$，其总数为$\sum_{j=1}^{M}N_j$。

### 2.5 强显著模型

本文使用选取最优$k$值的WKNN模型来估计超像素的显著值。但是，单个模型会受$k$值的选取而影响模型性能，所以为了增加模型的鲁棒性，使用

$n$个WKNN模型通过有权重的线性混合方式来构造一个强显著模型，如图2所示。然后，利用这个强显著模型来估计显著图。

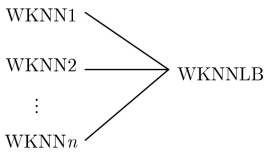



图 2 强显著模型示意图

#### 2.5.1 加权的$k$近邻模型(WKNN)

K近邻模型KNN(K-Nearest Neighbor)是一个懒惰学习模型，即没有训练过程。在KNN模型中，测试样本的标签是由离该测试样本最近的$k$个训练样本的标签确定的。KNN模型仅仅考虑了最近的训练样本的标签值，而忽略了这最近的$k$个训练样本距测试样本之间的特征距离大小。因此，本文使用的WKNN模型是加权的KNN模型，WKNN模型不仅考虑距离测试样本最近的$k$个训练样本的标签，而且还考虑了测试样本和训练样本之间的特征距离。如图3所示，测试样本(中心圆点)的标签不仅取决于其周围最近邻的$k$个训练样本的标签，还取决于他们之间的特征距离。因此，多尺度的图像$\boldsymbol{I}$中的测试样本超像素$\boldsymbol{r}_i^j$的显著值是被定义为

$$\boldsymbol{S}_{\text{strong}}(\boldsymbol{r}_i^j)=\sum_{v=1}^{k} w_{iv}y_v \tag{5}$$

其中，$k$表示距离测试样本的最近训练样本的个数，$y_v$表示K近邻的第$v$个训练样本的标签，$w_{iv}$是测试样本$\boldsymbol{r}_i^j$与第$v$个训练样本之间的权重，定义为

$$w_{iv}=-\ln\left(\frac{d(\boldsymbol{f}_i^j,\boldsymbol{f}_v)}{\sum_{v=1}^{k} d(\boldsymbol{f}_i^j,\boldsymbol{f}_v)}\right) \tag{6}$$

其中，$d(\boldsymbol{f}_i^j,\boldsymbol{f}_v)$表示为在65维的空间内，测试样本$\boldsymbol{r}_i^j$和最近邻的$k$个训练样本之间的特征欧式距离，即数学表达式为

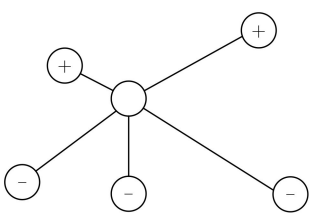


图 3 加权$k$近邻模型示意图

$$d(\boldsymbol{f}_i^j,\boldsymbol{f}_v)=\sqrt{\sum_{n=1}^{65}\left(\boldsymbol{f}_i^{j(n)}-\boldsymbol{f}_v^{(n)}\right)^2} \tag{7}$$

#### 2.5.2 最优$k$值选取

$k$值的选取对于WKNN模型是至关重要的，并会影响最终的模型效果。当$k$=1时，K近邻模型就变成了最近邻模型，它不考虑周围样本对其提供的有用信息，仅仅考虑距离它最近的一个样本，所以本文不考虑$k$=1时的情况。为了获取最优的$k$值，本文从训练样本集$\boldsymbol{D}_{\text{T}}$中随机地选取30%的样本作为验证样本$\boldsymbol{D}_{\text{TK}}$，剩下的70%样本作为选取最优$k$值的训练样本。最优的$k$值$k^*$具体定义为

$$k^*=\arg\max_{k=2,3,\cdots,m} P_{\boldsymbol{D}_{\text{TK}}}(k) \tag{8}$$

其中，$P_{\boldsymbol{D}_{\text{TK}}}(k)$表示使用$k$个最近邻样本模型在验证样本集$\boldsymbol{D}_{\text{TK}}$上的准确率，在计算$k^*$时，$k$是从2遍历到$m$，那么这$m$–1个WKNN模型在验证样本集$\boldsymbol{D}_{\text{TK}}$上的准确率有高有低，$k^*$取的就是使$P_{\boldsymbol{D}_{\text{TK}}}(k)$到达最高准确率的最优$k$值，其中，准确率$P_{\boldsymbol{D}_{\text{TK}}}(k)$具体定义为

$$P_{\boldsymbol{D}_{\text{TK}}}(k)=\frac{\sum_{i=1}^{|\boldsymbol{D}_{\text{TK}}|} I(y_i\boldsymbol{S}_{\text{strong}}(\boldsymbol{x}_i))}{|\boldsymbol{D}_{\text{TK}}|} \tag{9}$$

其中，$|\boldsymbol{D}_{\text{TK}}|$表示为验证样本集中的样本个数，$I(y_i\boldsymbol{S}_{\text{strong}}(\boldsymbol{x}_i))$是指示函数，当$y_i\boldsymbol{S}_{\text{strong}}(\boldsymbol{x}_i)>0$时取值为1，反之取0。$m$的取值是根据实验得出的，参考图4，随着$m$的增大，评价指标$F$-度量值在逐渐地增大。综合考虑计算成本和最佳$k$值的选择，取$m$=15。

#### 2.5.3 有权重的线性混合

WKNN中选取不同个数的近邻对结果有较大的影响，因此为了减小这种影响，选择$n$个WKNN模型通过有权重的线性混合方式，形成强显著模型。这$n$个模型是选择在验证样本集$\boldsymbol{D}_{\text{TK}}$上的准确率较高的前$n$个，具体如下：将$k$从2遍历到$m$，计算他们的$P_{\boldsymbol{D}_{\text{TK}}}(k)$值，并从高到低进行排序，选择

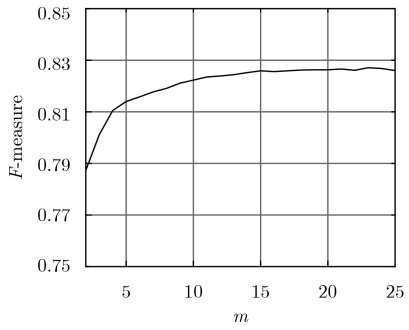



图 4 $m$取不同值在ASD数据集上的$F$-measure曲线

前$n$个WKNN模型，他们所对应的$k$值用符号$k^q$来表示，$q=1,2,\cdots,n$。$n$的取值是根据实验得出的，如图5，评价指标$F$-度量值随着$n$值增加而逐渐增加，但当$n$>5时，$F$值接近稳定，因此$n$值设置为5。所以，测试样本$\boldsymbol{r}_i^j$的显著值计算式(5)更新为

$$\boldsymbol{S}_{\text{strong}}\left(\boldsymbol{r}_i^j\right)=\sum_{q=1}^{n}\left(P_{\boldsymbol{D}_{\text{TK}}}(k^q)\cdot\sum_{v=1}^{k^q}w_{iv}y_v\right) \tag{10}$$

其中，$P_{\boldsymbol{D}_{\text{TK}}}(k^q)$表示为取$k$值为的$k^q$第$q$个WKNN模型在验证样本集$\boldsymbol{D}_{\text{TK}}$上的准确率，以此作为线性混合的权重。据此，多尺度的超像素$\boldsymbol{r}_i^j$经过式(10)的计算，能够得到以超像素为单位的多尺度下的显著图，再经过显著值的分配，只要像素点$\boldsymbol{I}_l^j\in\boldsymbol{r}_i^j$，$l=1,2,\cdots,N$，则$\boldsymbol{S}_{\text{strong}}(\boldsymbol{I}_l^j)=\boldsymbol{S}_{\text{strong}}(\boldsymbol{r}_i^j)$。通过像素显著值的分配，最终得到$j$个尺度下的强显著图$\mathbf{SM}^j=\{\boldsymbol{S}_{\text{strong}}(\boldsymbol{I}_l^j)|l=1,2,\cdots,N\}$。

注意，学习过程是严格限制在每一张输入图像，对于不同的输入图像，前$n$个$k$值和对应的系数是不同的，并会根据式(9)的定义自动调节。本文是在ASD数据集上进行实验来确定最优的$m$和$n$值的。

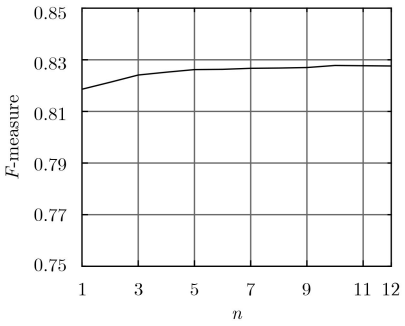



图 5 $n$取不同值在ASD数据集上的$F$-measure曲线

### 2.6 融合

本文将$M$个尺度的弱显著图$\{\mathbf{WM}^j|j=1,2,\cdots,M\}$和强显著图$\{\mathbf{SM}^j|j=1,2,\cdots,M\}$分别通过$1/M\sum_{j=1}^{M}\mathbf{WM}^j$和$\overline{\mathbf{SM}}=1/M\sum_{j=1}^{M}\mathbf{SM}^j$来得到融合后的弱显著图和强显著图。本文使用4种不同粒度的超像素，即$M=4$，超像素个数分别设置为100, 150, 200, 250。弱显著图和强显著图具有互补的特性。弱显著图由于采用了基于对比度的测量方法，更倾向于检测出细微的细节，捕捉到局部的结构信息，相反，强显著图通过关注于图像中的全局目标形状来捕捉到全局的结构信息。弱显著图和强显著图的融合定义为

$$\mathbf{FM}=\eta\overline{\mathbf{SM}}+(1-\eta)\overline{\mathbf{WM}} \tag{11}$$

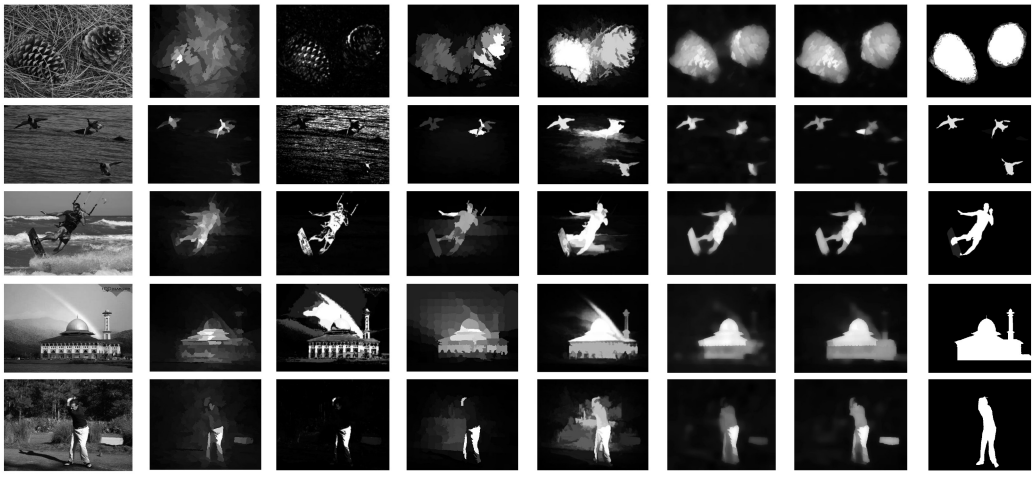
其中，$\eta$是平衡因子，$\mathbf{FM}$是最终的显著图。在本文中，考虑强显著图的贡献要大于弱显著图，所以$\eta$设置为0.7。

## 3 实验结果

本文主要做了两大实验。第1个实验：将本文所提WKNNLB模型作用于现存的5种优秀方法上，即采用现存的5种方法作为弱显著模型来提供弱显著图，再通过本文所提模型产生最后的显著图。第2个实验：本文所提WKNNLB模型重点和BLSVM模型相比，为了保证公平地对比，均使用文献[15]提出相同的方式来得到弱显著图。本文比较了各种经典的方法，正如图6所示，本文所提方法产生的显著图更接近真值图。本文实验是使用MATLAB编写，电脑的配置参数为Intel i5-7500 CPU(3.4 GHz)和16 GB RAM。

### 3.1 数据集及评价指标

在本文中使用ASD和DUT-OMRON这两个公



(a) 原图 (b) LRMR (c) GC (d) MR (e) MBD (f) BLSVM (g) 本文算法 (h) GT

图 6 各种方法产生的显著图的视觉对比

开的数据集及$P$-$R$(Precision-Recall)曲线和$F$-度量值这两个评价指标。在某些情况下，$P$-$R$曲线不能全面评估显著性图的质量。因此，本文使用综合考虑$P$和$R$值的$F$-度量来进行评估，$F$-度量值计算为

$$F_\beta = \frac{(1+\beta^2)\cdot P\cdot R}{\beta^2\cdot P + R} \tag{12}$$

参考文献[21], $\beta^2$取值为0.3，即更多地关注于$P$值。

### 3.2 经典显著性检测方法的提高

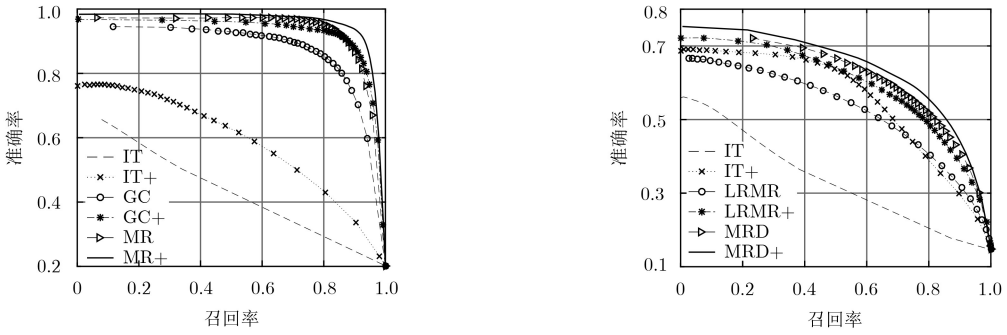
强显著模型WKNNLB能够很好地提升现存的显著性检测目标模型的性能。本文使用5种传统的显著性检测方法作为本文的弱显著性模型，这5种经典方法简写分别为IT[22], LRMR[23], GC[12], MR[11]和MBD[13]。正如图7和表2所示，本文提出的强显著模型是能够显著地提升这5种经典方法的性能。

### 3.3 WKNNLB和BLSVM之间的对比

本文提出的模型和BLSVM[15]算法是相似的。为了公平对比，相同的弱显著模型被采用。本文在$P$-$R$曲线，$F$-度量值和运行时间上进行了对比。为了进一步验证本文方法的有效性，本文在另外常用的SED2和PASCAL-S数据集上进行了$F$-度量值和运行时间上的对比。正如图8，在ASD数据集上$P$-$R$曲线几乎重合，但在DUT-OMRON数据集上，WKNNLB完全超过BLSVM。正如表3所示，在4个数据集上的实验结果也表明了本文提出的算法仅用一半的时间便能达到或者稍优于BLSVM算法。在WKNNLB和BLSVM之间的时间复杂度分析：处理时间包括训练时间和测试时间，对于BLSVM算法训练时间为$O(n^2)$，测试时间为$O(1)$；而对于WKNNLB算法，训练时间为0,测试时间为$O(n)$。本文为了进一步提高运行效率使用kd树算法来搜索训练样本。

## 4　结束语

本文提出一个基于加权的K近邻线性混合模型的显著性目标检测方法，测试样本的显著值是由它的邻居及与邻居之间的欧式距离来决定的。该模型的时间复杂度是依赖于训练样本的个数。实验结果

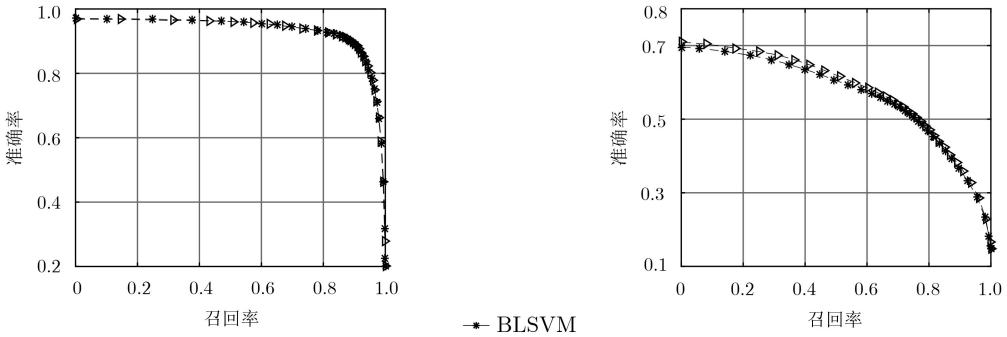



(a) ASD数据集上的$P$-$R$曲线　　(b) DUT-OMRON数据集上的$P$-$R$曲线

图 7 各方法及其提高在ASD和DUT-OMRON数据集上的$P$-$R$曲线

**表 2　5种经典方法及其提高在$F$-度量值的对比**

| 算法 | IT | IT+ | LRMR | LRMR+ | GC | GC+ | MR | MR+ | MBD | MBD+ |
|---|---|---|---|---|---|---|---|---|---|---|
| ASD | 0.381 | **0.542** | 0.727 | **0.829** | 0.811 | **0.848** | 0.869 | **0.876** | 0.855 | **0.867** |
| DUT-OMRON | 0.343 | **0.541** | 0.498 | **0.545** | 0.520 | **0.554** | 0.574 | **0.576** | 0.850 | **0.854** |




(a) ASD数据集上的$P$-$R$曲线　　(b) DUT-OMRON数据集上的$P$-$R$曲线

图 8 WKNNLB和BLSVM在ASD和DUT-OMRON数据集上的$P$-$R$曲线

表 3 WKNNLB和BLSVM在4个数据集上$F$-度量和运行时间(s)对比

| | ASD | | SED2 | | PASCAL-S | | DUT-OMRON | |
|---|---|---|---|---|---|---|---|---|
| | $F$-measure | Time | $F$-measure | Time | $F$-measure | Time | $F$-measure | Time |
| WKNNLB | **0.820** | **4058** | **0.758** | **332** | **0.655** | **5000** | **0.534** | **30864** |
| BLSVM | 0.810 | 8093 | 0.740 | 720 | 0.651 | 11184 | 0.524 | 65120 |

表明，该算法是有效的，它可以显著提高现有其他显著性检测模型的性能。它也达到了和基于多核支持向量机集成学习方法相同的性能，且仅用了其一半的处理时间，另外当采用较好的弱显著图时，该算法能够取得更优异的效果。下一步工作，将要尝试将加权的$k$近邻线性混合模型引入到协同显著性检测中，并期望在准确率及效率方面上具有优异的性能。

## 参 考 文 献


[1] BORJI A and ITTI L. State-of-the-art in visual attention modeling[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(1): 185–207. doi: 10.1109/TPAMI.2012.89.

[2] ITTI L. Automatic foveation for video compression using a neurobiological model of visual attention[J]. *IEEE Transactions on Image Processing*, 2004, 13(10): 1304–1318. doi: 10.1109/TIP.2004.834657.

[3] ZHANG Fan, DU Bo, and ZHANG Liangpei. Saliency-guided unsupervised feature learning for scene classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(4): 2175–2184. doi: 10.1109/TGRS.2014.2357078.

[4] LU Xiaoqiang, ZHENG Xiangtao, and LI Xuelong. Latent semantic minimal hashing for image retrieval[J]. *IEEE Transactions on Image Processing*, 2017, 26(1): 355–368. doi: 10.1109/TIP.2016.2627801.

[5] WEI Yunchao, XIAO Huaxin, SHI Honghui, *et al*. Revisiting dilated convolution: A simple approach for weakly-and semi-supervised semantic segmentation[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 7268–7277. doi: 10.1109/CVPR.2018.00759.

[6] ZHANG Xiaoning, WANG Tiantian, QI Jinqing, *et al*. Progressive attention guided recurrent network for salient object detection[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 714–722. doi: 10.1109/CVPR.2018.00081.

[7] CHEN Shuhan, TAN Xiuli, WANG Ben, *et al*. Reverse attention for salient object detection[C]. The 15th European Conference on Computer Vision, Munich, Germany, 2018: 236–252. doi: 10.1007/978-3-030-01240-3_15.

[8] ZHANG Lu, DAI Ju, LU Huchuan, *et al*. A bi-directional message passing model for salient object detection[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 1741–750. doi: 10.1109/CVPR.2018.00187.

[9] WANG Tiantian, ZHANG Lihe, WANG Shuo, *et al*. Detect globally, refine locally: A novel approach to saliency detection[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 3127–3135. doi: 10.1109/CVPR.2018.00330.

[10] HOU Qibin, CHENG Mingming, HU Xiaowei, *et al*. Deeply supervised salient object detection with short connections[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(4): 815–828. doi: 10.1109/TPAMI.2018.2815688.

[11] YANG Chuan, ZHANG Lihe, LU Huchuan, *et al*. Saliency detection via graph-based manifold ranking[C]. 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, USA, 2013: 3166–3173. doi: 10.1109/CVPR.2013.407.

[12] CHENG Mingming, WARRELL J, LIN Wenyan, *et al*. Efficient salient region detection with soft image abstraction[C]. 2013 IEEE International Conference on Computer Vision, Sydney, Australia, 2013: 1529–1536. doi: 10.1109/ICCV.2013.193.

[13] ZHANG Jianming, SCLAROFF S, LIN Zhe, *et al*. Minimum barrier salient object detection at 80 FPS[C]. 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 1404–1412. doi: 10.1109/ICCV.2015.165.

[14] BORJI A, CHENG Mingming, JIANG Huaizu, *et al*. Salient object detection: A benchmark[J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 5706–5722. doi: 10.1109/TIP.2015.2487833.

[15] TONG Na, LU Huchuan, RUAN Xiang, *et al*. Salient object detection via bootstrap learning[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 1884–1892. doi: 10.1109/CVPR.2015.7298798.

[16] LU Huchuan, ZHANG Xiaoning, Qi Jinqing, *et al*. Co-bootstrapping saliency[J]. *IEEE Transactions on Image Processing*, 2017, 26(1): 414–425. doi: 10.1109/TIP.2016.2627804.

[17] SONG Hangke, LIU Zhi, DU Huan, *et al*. Depth-aware salient object detection and segmentation via multiscale

discriminative saliency fusion and bootstrap learning[J]. *IEEE Transactions on Image Processing*, 2017, 26(9): 4204–4216. doi: 10.1109/TIP.2017.2711277.

[18] WU Lishan, LIU Zhi, SONG Hangke, *et al*. RGBD co-saliency detection via multiple kernel boosting and fusion[J]. *Multimedia Tools and Applications*, 2018, 77(16): 21185–21199. doi: 10.1007/s11042-017-5576-y.

[19] 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. doi: 10.1109/TPAMI.2012.120.

[20] OJALA T, PIETIKAINEN M, and MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971–987. doi: 10.1109/tpami.2002.1017623.

[21] ACHANTA R, HEMAMI S, ESTRADA F, *et al*. Frequency-tuned salient region detection[C]. 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, USA, 2009: 1597–1604. doi: 10.1109/CVPR.2009.5206596.

[22] ITTI L, KOCH C, and NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11): 1254–1259. doi: 10.1109/34.730558.

[23] SHEN Xiaohui and WU Ying. A unified approach to salient object detection via low rank matrix recovery[C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012: 853–860. doi: 10.1109/CVPR.2012.6247758.

李　炜：女，1969年生，教授，研究方向为计算机视觉.

李全龙：男，1995年生，硕士生，研究方向为图像显著性检测.

刘政怡：女，1978年生，副教授，研究方向为计算机视觉.