

基于眼动点视觉先验与边缘优化的显著性检测

刘翔宇¹ 蹇木伟¹ 鲁祥伟¹ 何为凯² 李晓峰³ 尹义龙³

1 山东财经大学计算机科学与技术学院 济南 250014

2 山东交通学院航空学院 济南 250357

3 山东建筑大学计算机科学与技术学院 济南 250000

(997865855@qq.com)

摘要 图像显著性检测是计算机视觉中的基础研究课题之一。当前基于深度学习的方法虽然能够有效提高显著性检测结果的准确性,但是在显著性目标的物体边缘细节提取方面还不能令人满意。为此,提出了一种基于眼动点预测先验的边缘细化网络用于显著性目标提取。首先,对输入图像进行眼动点预测,将生成的特征图像作为后续显著性检测的视觉先验;其次,利用多注意力机制 VGG16 网络进行显著性目标特征提取;最后,对特征图像进行质量优化处理,进一步提升图像显著图的质量。实验结果表明,在 3 个公开数据集(DUTS, ECSSD, HKU-IS)上,所提方法与其他 6 个主流方法相比,取得了更好的显著性检测效果。

关键词: 显著性检测;注意力机制;眼动点预测;边缘优化

中图分类号 TP391

Saliency Detection Based on Eye Fixation Prediction and Boundary Optimization

LIU Xiang-yu¹, JIAN Mu-wei¹, LU Xiang-wei¹, HE Wei-kai², LI Xiao-feng³ and YIN Yi-long³

1 School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China

2 School of Aeronautics, Shandong Jiaotong University, Jinan 250357, China

3 School of Computer Science and Technology, Shandong Jianzhu University, Jinan 250000, China

Abstract Saliency detection is one of the most fundamental challenges in computer vision. Although the rapid development of deep learning has greatly improved the accuracy of saliency-detection results, the extraction of details of salient object is still unsatisfactory. Therefore, this paper proposes an edge refinement network based on eye-fixation prediction priori for salient object detection. Firstly, eye-fixation extraction is carried out on the original image and the extracted feature image is used as the visual priori of subsequent saliency detection. Secondly, the multi-attention mechanism of VGG16 network is used for feature extraction, and finally the feature image is refined to improve the quality of the saliency image. Experimental results show that, compared with other 6 state-of-the-art methods, the proposed method achieves better results in 3 open-accessed data sets (i. e. DUTS, ECSSD, HKU-IS).

Keywords Saliency detection, Attention mechanism, Eye-fixation prediction, Edge refinement

过去十几年以来,显著性检测在计算机视觉研究领域受到持续关注,显著性检测方法被广泛运用于目标跟踪、自适应图像压缩、图像分割、目标识别和目标语义提取等方面。特别是随着深度学习的快速发展,如何运用深度学习方法进行显著性检测受到了越来越多研究者的重视。虽然相对于传统显著性检测方法,深度学习在显著性检测中的应用使得检测结果有都很大的提升,但是深度模型方法仍然会出现目标检测缺失或者边缘检测结果粗糙的情况。

在当前大部分工作中,眼动点(视点)预测与显著性检测是独立进行研究的。根据对视觉感知机理的理解,受眼动点预测的启发,我们尝试将视点预测的方法应用到显著性检测中,将二者进行有效结合,将视点预测结果作为显著性检测的

视觉先验进行图像预处理。我们的方法舍弃了显著性检测中应用比较广泛的候选框选以及其他的边界预处理传统框架,高效采用眼动点预测的方法对输入图像进行视觉先验的处理。提出模型的优势包括:首先运用视点预测作为视觉先验的方式来粗略预估图像中的显著性区域,在环境复杂图像背景和场景中利用人类视点的视觉感知本能可以更加准确捕捉显著性区域。不同于其他传统方法,该方法更贴合人类视觉的注视习惯,有利于快速、准确地从复杂背景或者前背景模糊的图像中提取出注视区域即显著性区域。

本文提出的模型中显著对象提取网络是基于 VGG-16^[34] 网络进行改进的。VGG-16 网络是当前显著性检测领域最常用的网络结构之一,我们的模型将 conv 1 跟 conv 2 组合来提

基金项目:国家自然科学基金(61976123);山东省泰山青年学者计划;山东省基础研究重点发展计划(ZR2020ZD44)

This work was supported by the National Natural Science Foundation of China(61976123), Taishan Young Scholars Program of Shandong Province and Key Development Program for Basic Research of Shandong Province(ZR2020ZD44).

通信作者:蹇木伟(jianmuweihk@163.com)

取输入图像的低层特征,用 conv 3 层进行进一步的中期优化。为了更好地细化图像的显著性区域的高级上下文语义特征,对 conv3,conv4,conv5 层进行多尺度特征提取,然后通过一个轻量级通道注意力模块对高级语义特征进行最终优化。另外,由于主干网络输出的中间显著性映射只是一个相对粗糙的显著性检测结果,通常会产生一些背景干扰因素。因此,本文设计的基于 VGG 网络的边缘优化是在骨干网络之后完成的。我们的设计方法是在网络优化的同时对输出层图像又进行了更深层次的显著性目标边界优化。具体而言,受到 Unet 网络^[35]的启发,本文的模型在主干网络生成显著性映射图像之后添加一个边缘优化模块去做最后的显著性图处理,实现了输出显著性检测结果的最优化。

本文的主要贡献包括:

(1)将眼动点预测的方法运用到了显著性检测当中,对输入图像进行注意视点的标定作为视觉先验,以此来有效预估显著性区域的相对位置。

(2)基于 VGG-16 网络,提出了一种多尺度-注意力优化网络,“由粗到细”合理利用了注意力机制模块对显著性检测结果进行细化处理。

(3)最后,对输出层显著性图进行更深层次的边缘细化处理,有效提高了显著性物体检测结果的平滑性和一致性。

1 相关工作

本节将系统地回顾与本文相关的眼动点预测方法、注意力机制以及显著性物体检测、边缘细化的方法以及相关文献。

1.1 眼动点预测

眼动点预测,也称视点检测,指的是人在看到一幅图像时眼睛的视觉注意力关注的区域,换句话说,就是人类视觉感知本能上最感兴趣的部分。

视点检测是一个热点研究领域^[1],从最开始的传统方法一直到现在的深度学习方法,关于视点预测的研究一直受到来自不同学科研究者的关注。

眼动点检测一般都是基于注意力模型。早期的注意力模型方法根据视点获得的机制可以分为:矩阵谱分析模型^[10]、信息理论模型^[8]、图形模型^[9]、认知模型^[1-5]、贝叶斯模型^[6]、模式分类模型^[11]、和决策理论模型^[7]等。根据预测精度而言,传统眼动点检测方法有一定的局限性,其预测模型已经很难适应现在的高质量以及高复杂度的大数据量图像分析。随着神经网络的兴起,最近几年很多基于深度学习方法被提出。例如,eDN 模型^[13]是比较早期的一个深度学习模型,它可以自动学习视点预测的深层表示。之后,随着网络层数的加深以及模型复杂度的提高,更多的深度学习视点预测模型被提出。Srinivas 等^[12]提出了一种基于深度卷积神经网络(CNN)的视点预测方法,该方法在一个统一的框架能够预测眼睛注视的显著性对象。Zhou 等^[14]提出了一种基于面部辅助的显著性视频注意力预测模型(AFSnet),用于预测眼动点在电影场景中的关注位置。与传统方法相比,深度学习方法在预测精度上取得了更好的效果。

1.2 显著性检测

同视点预测相似,显著性检测也是视觉研究领域的一个热点课题。从显著性检测被提出,到目前二十多年的时间里,涌现出了大量先进的模型构建方法,特别是随着深度学习在

视觉领域的应用,极大地提高了显著性检测的准确度。同时,一些传统方法与深度学习方法相结合的模型也开辟了一条新的显著性检测新思路。

传统的显著性检测方法一般都是基于低层视觉特征的提取,比如基于图像变换分析的方法。Hou 等^[10]提出了一种基于傅里叶变换的简单、快速的显著性检测方法,称为谱残差。Cui 等^[47]将稀疏表示方法扩展至视频显著性领域,提出了一种快速的运动显著性检测方法,即时间频谱残差。

随着深度学习的发展,以及深度学习在视觉显著性检测研究的广泛应用,多种高效的网络模型、方法以及一些显著性检测增强模块被设计出来。从最早的 BP 网络^[22]开始,机器学习就出现了很多显著性检测模型,之后深度学习的快速发展,更多的基于深度学习的显著性检测模型被提出。Wang 等^[18]提出了一种基于全局显著性特征信息的视觉注意力网络,用于从具有全局显著性信息的深层粗糙层到具有局部显著性响应的浅层精细层的视觉显著性检测网络。Marcella 等^[19]提出了一个可以通过结合神经注意机制来预测显著性图的模型。Zhu 等^[20]提出了一种多尺度对抗特征学习(MAFL)模型用于显著性检测。最近,Wei 等^[21]提出了一种基于全卷积网络(FCN)的深度显著性检测方法来解决共显著对象发现与检测问题。

1.3 注意力机制

注意力机制具有很强的特征选择能力,因此非常适用于视觉显著性检测领域。由于注意力机制本身也是符合人类视觉系统的感知机理,所以在计算机视觉领域尤其是在图像以及视频显著性领域都有着非常重要的推动作用。

注意力机制最早在计算机视觉领域被提出。Mnih 等^[15]采用了 RNN 模型,并加入了注意力机制来进行图像的分类,使得注意力机制的研究流行起来。Zhang 等^[16]提出了一种新的注意力引导网络,它有选择地以渐进的方式整合多层次的上下文信息。该网络产生的注意特征可以缓解背景的分散,从而达到更好的性能。Liu 等^[17]提出了一种新的基于像素的上下文注意网络,以学习选择性地关注每个像素的信息上下文位置,合理利用 Attention 机制进行图像的显著性检测,并取得了非常好的效果。

2 本文提出的方法

给定一张输入图像,显著性检测的目标是生成一张像素级的高质量显著性图来突显出人们所关注的区域或物体。在本文提出的模型中,首先通过学习模拟人的眼睛注意力视点的关注度,生成注意力视点图像来预测原图像显著对象的大致位置,以此作为输入图像的视觉先验。然后,提出了基于 VGG-16 网络的优化改进模型,以自下而上进行从低层视觉特征到高层上下文语义特征的提取与增强。最后,针对以往方法对边缘信息提取不足的难题,本文的模型设计了一个基于 Unet 网络的边缘增强模块来进一步增强显著性图的边缘视觉对比。

2.1 眼动点预测

在视觉感知与理解中,眼动点预测有助于显著性目标位置等信息的快速定位。基于此启发,提出的模型首先利用眼球视点预测结果作为显著性检测视觉先验。我们的方法中眼球视觉预测先验部分独立于显著性检测主干网络之外,显著

性目标提取之前单独对输入图像进行相应的视点预测的训练与测试,将得到的视点预测初步结果作为显著性目标提取的视觉先验进行处理。

如图 1 所示,我们的视点预测模块使用的是现在非常流行的 VGG-16^[34,38] 网络进行训练与测试。由于此部分的视点预测属于视觉先验图像,本文提出的框架省略最后的一个池化层来保存更多的空间位置信息。

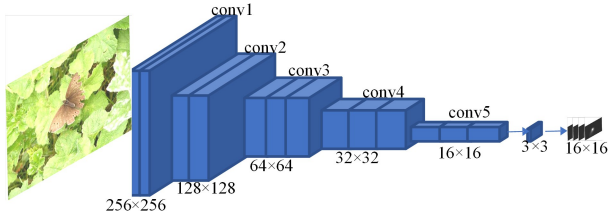


图 1 眼动点预测网络

Fig. 1 Eye-fixation prediction network

图 1 中,我们对一幅 256×256 的图像进行训练时,将其经过 VGG-16 网络的特征提取之后添加一个含有 sigmoid 3×3 的卷积核,将得到的 16×16 的卷积特征概率图进行归一化处理,最后将归一化后的视点图进行映射得到灰度视点图,将得到的视点图对主干网络中的 conv5 - conv3 层进行视觉先验处理。

2.2 上下文高级特征提取

高层视觉信息在显著性检测中起着非常重要的作用。现有的深度学习网络模型一般是通过堆叠较多的卷积层和池化层来学习显著性对象的语义特征。但是,随着各领域图像数据量的不断增加,显著性对象在位置、形状、颜色以及语义特征上变化很大,从而导致图像背景与前景的颜色等因素的混叠给显著性检测工作带来了很大的挑战。过去的处理方法一般是直接使用自下而上的深度卷积与池化,但是这种方法通常无法有效处理上述各种特殊情况。受到 VGG-16 网络模型的启发,我们设计了一种新的显著性特征提取网络模型来进行上下文高级特征提取。

目前,VGG-16^[34,36] 网络是显著性检测应用最为广泛的深度学习骨干网络模型之一,我们的模型以 VGG-16 网络为基础进行了改进。对于大小为 256×256 的输入图像,通过 VGG-16 网络我们可以提取 5 个层次的特征图像,分别记为 $\{f_i; i=1, \dots, 5\}$ 。本文运用 VGG-16 中的 conv3, conv4 和 conv5 作为特征提取中基本的高层视觉特征。在图 2 中,我们展示了所用的高层特征提取框架。通过 conv3, conv4, conv5 层的信息进行多尺度高级语义特征提取,将视觉先验干预后的图像进行逐级上采样,并通过轻量级通道注意力模块^[24,36] 作为高响应通道分配更大的权重,将显著性图经过卷积以及上采样之后,生成与低级语义特征图相同尺度的高层特征图:

$$S = F(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

其中, H 和 W 分别表示图像的高度与宽度,将图像矩阵压缩成 1×1 ,每个通道的信息都用该通道的全局平均池化来表示:

$$E = F(z, W) = \sigma(f_2(W_2 \delta(f_2(W_1 z)))) \quad (2)$$

其中, W 表示通道注意力模块的模型参数; f_c 表示全连接层,通道注意力模块使用两个全连接层来增加非线性,尽可能减少参数量和计算量; z 表示每一个通道上的特征向量; σ 表示

sigmoid 操作,将变量进行归一化处理; δ 值实现 ReLU 功能,解决梯度消失的问题,同时也能提高收敛速度。

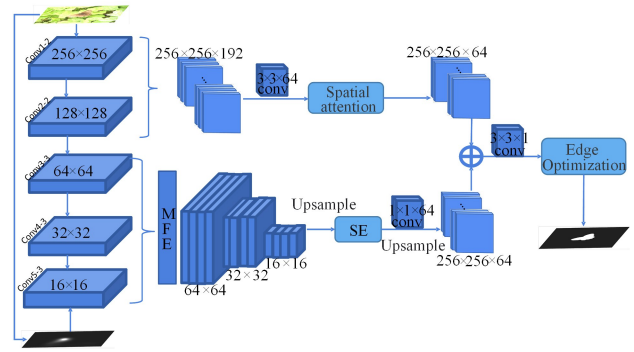


图 2 显著性特征提取网络结构图

Fig. 2 Saliency feature extraction network structure diagram

2.3 低层特征提取

图像通常包含了复杂的背景以及相对显著的前景物体对比信息。在显著性检测过程中,前景与背景差异较大的图像,可以较为容易地检测到显著性目标。因此,图像的低层视觉特征有助于实现显著性检测。在显著性模型设计时,需要差异化地去考虑图像中显著性目标所在的空间位置,采用空间注意力的方法可以更好地过滤掉图像的背景。利用低层特征更好地关注显著性前景,有助于得到更为准确的显著性检测结果。

我们的框架中借鉴了空间注意力模块^[23] 对低级特征进行提取。类似地,我们的低层特征需要将较小的特征上采样到最大的特征,将它们组合起来经过一个 $3 \times 3 \times 64$ 的卷积核进行处理,通过空间注意力模块实现低层语义特征的提取。

2.4 边缘强化

在以往的研究中,为了提高精度,显著性检测往往都是更关注图像先验的预处理,后期优化也有助于提升显著性检测的准确性。以前大多显著性检测方法常常忽视显著性物体的边缘优化,导致生成的显著性图中的物体边缘模糊或者不完整。针对以往方法的不足,本文的模型在预处理的基础上还增加了显著性物体的边缘优化模块。

对于显著性检测中的边缘优化模块,之前的研究通常将其设计为残差模块,对预测的初步显著性映射进行进一步优化。一般地,初始的显著性图像通过学习显著性图与标准显著性图之间的残差来得到最终的优化图像,具体实现如式(3)所示:

$$U_{\text{refined}} = U_{\text{initial}} + U_{\text{residual}} \quad (3)$$

如图 2 所示,为了能够更好地优化显著性检测结果中的边界模糊或者噪声等不利影响,我们受到解码编码模型的启发,在 VGG 网络之后添加了一个基于区域上下文高级语义特征的残差边界优化模块来增强显著性物体的边缘提取^[35,37]。该模块主要包括了一组编码器、一个优化链接器和一组解码器。每一组的编码器与解码器都有 4 个阶段,每一个阶段都只有一个卷积层。优化连接器中的过滤器设置大小以及数量与第四阶段编码器的过滤器相同。编码器的下采样采用最大池化的方式,解码器采用了双线性插值。从边缘强化模块中输出的图像即为优化处理的最终显著性检测结果。

2.5 损失函数

在机器学习和数学优化中,损失函数代表了分类问题中预测不准确所付出的代价。在眼动点检测中,生成的检测图

作为视觉先验线索使用。我们运用常用的最小化相对熵作为损失函数进行网络的训练,损失函数如(4)所示:

$$L(P, Q) = \frac{1}{16 \times 16} \sum_x^{(16 \times 16)} P(x) \log \frac{P(x)}{Q(x)} \quad (4)$$

其中, P 表示真值(Ground Truth), Q 表示预测图像。

在显著性目标检测中,我们通常使用最终显著性图与真值之间的交叉熵损失。损失函数定义为:

$$L(G, S) = - \sum_{i=1}^N (G_i \log(S_i) + (1 - G_i) \log(1 - S_i)) \quad (5)$$

其中, G 表示真值(Ground Truth), S 表示预测图像。

3 实验

3.1 实验数据集

我们的模型训练以及测试共用到了当前 3 个广泛使用的开源图像数据集: DUTS^[25], ECS^[26], 和 HKU-IS^[27]。其中, DUTS 数据集是一个大型的显著性检测图像数据集, 该数据集中包含了 10553 张训练图像以及 5019 张测试图像; ECS^[26] 则提供了 1000 张语义特征丰富、结构比较复杂、背景与前景差别较小的图像供显著性检测与评估; HKU-IS^[27] 数据集包含 4447 张原图及对应的 Ground Truth 图, 即每张图像都提供了像素级别标注。

3.2 实验环境

基于 pytorch 1.5.0 框架进行实验, 并且使用了 NVIDIA GeForce RTX 2070 SUPER 的 GPU 进行加速。实验的环境是 Windows 10 操作系统, Python 版本是 3.7。

3.3 评价标准

在显著性检测结果的评价中, 采用常规的 P-R 曲线与 F_β -measure 评估本文方法与 6 种主流方法在 3 个数据集上的对比结果。

P-R 曲线是评价显著性概率图的一种标准方法。准确率 P 表示得到的显著区域中实际目标物体所占的比例, 召回率 R 表示检测结果中的目标物体占真值图(Ground Truth)中目

标物体的比例。通过将二值化显著性图与真实 Ground Truth 图进行比较, 计算显著性图的查全率和查准率。计算公式:

$$P = \frac{TP}{TP + FP} \quad (6)$$

$$R = \frac{TP}{TP + FN} \quad (7)$$

其中, TP 表示前景区域中具有正确显著性的目标像素; FP 表示背景区域中具有正确显著性的目标像素; FN 表示该区域属于显著性物体但是却被检测为背景区域的像素个数。

当图像显著性检测算法在 P-R 曲线下进行数据对比时, 两种算法的曲线若发生了交叉, 则不管是查准率 Precision, 还是查全率 Recall, 它们对最终显著性结果的评估均有一定误差, 因此需要进一步进行综合评估比较。通过 F_β -measure 对精度和召回率进行综合评估, 计算公式如下:

$$F_\beta = \frac{(1 + \beta^2) Precision \times Recall}{\beta^2 Precision + Recall} \quad (8)$$

其中, 关于 β^2 的取值, 我们通过借鉴以往的评估经验将实验评估的最终数值设置为 0.3, 即增加了查准率的权重, 因为相对查全率而言, 查准率对显著性检测任务更加重要。

3.4 实验结果

本小节将本文的模型与 6 种代表性的显著性检测模型进行了比较, 即 Amulet^[28], DCL^[32], DHS^[33], MDF^[29], NLDF^[30] 和 UCF^[31]。分别在上文提到的 3 个数据集上进行系统测试。为了公平起见, 我们使用了文中推荐的参数和作者所提供的显著性检测模型。

3.4.1 视觉对比分析

图 3 提供了本文的模型与上述其他模型的显著性检测结果可视化比较, 直观地反映出了本文方法相较于其他方法的优越性。从图 3 中可以清晰地看出, 对于第三组以及第五组图像, 本文的方法在鸟类腿部的检测方面更加准确; 另外 3 组图像上, 本文方法在边缘的检测处理方面更加细致。

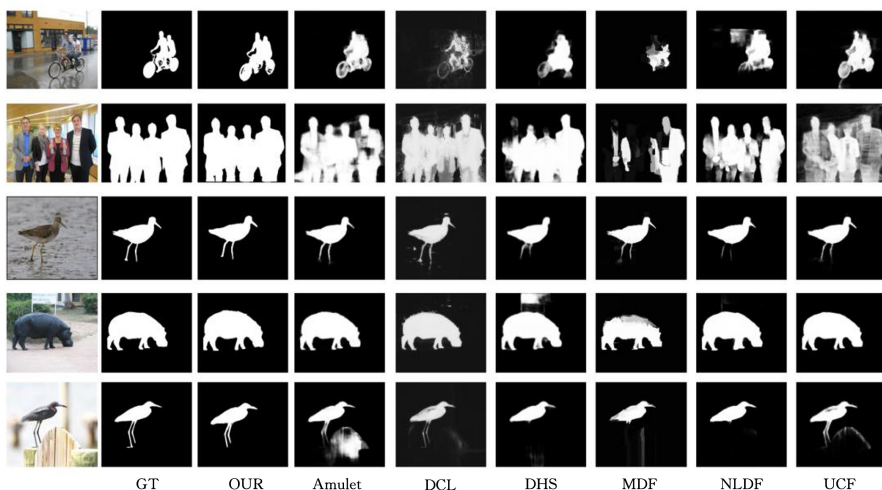


图 3 显著图的视觉效果对比

Fig. 3 Visual comparison of saliency maps

在不同数据集中具有挑战性的场景中, 我们方法的结果相较于其他主流方法的结果能更加接近 Ground Truth。具体来说:

我们所提出的模型不仅能够有效地突出图像的显著区域或显著性物体, 而且因为加入了人眼视点的处理, 能够很好地

抑制背景区域的影像以及噪声干扰, 使得显著性结果在捕捉空间位置等方面更加精确。在边缘细化模块的加持下, 我们的模型在边沿细节处理方面比其他方法做得更加清晰, 提出的模型有效结合了低层特征与高层语义特征等方面的优点。

3.4.2 定量比较

为了客观地比较不同方法的性能,图4和图5给出了本

文方法与其他6种代表性显著性检测方法在3个开源数据集上的P-R曲线与 F_{β} -measure的定量评估结果。

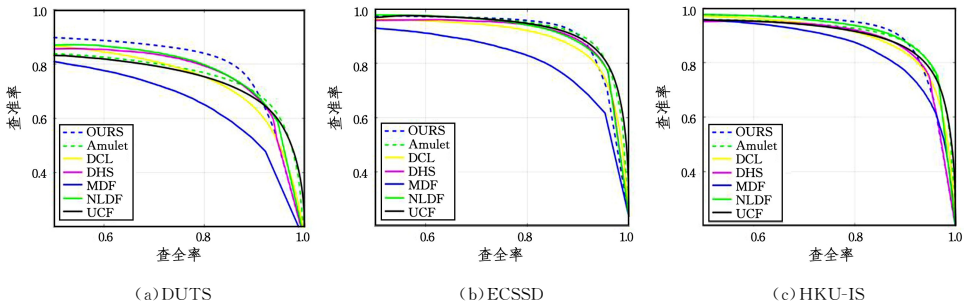
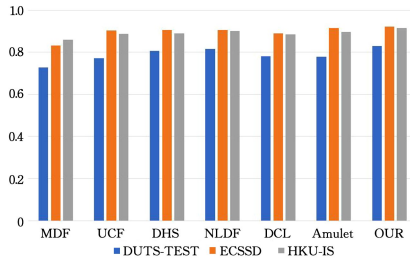


图4 P-R 曲线图

Fig. 4 P-R graph

图5 F_{β} -measure 值对比图Fig. 5 F_{β} -measure value

由图中可以看出,本文方法的准确率与召回率都有一定的提升,而且在3个数据集上 F_{β} -measure都是最高的,我们的方法在3个数据集上的显著性检测效果都优于其他6种方法。在3个数据集上的P-R曲线与 F_{β} -measure统计结果,可以反映出我们模型相比于其他几种先进模型的优越性。

结束语 在计算机视觉领域,显著性检测仍是一个挑战性研究课题。针对以往的多数传统方法容易忽略显著性图像边缘细节信息的缺陷,本文提出了一种基于眼动点预测的视觉先验与边缘优化相融合的显著性检测方法。利用眼动点检测后的视觉线索作为显著性检测的视觉先验进行引导,再利用多注意力机制的显著性检测网络进行低层与高层特征提取,最后添加边缘细化模块进行边缘优化处理提高显著性检测的精度。在多个公开可用的图像数据集上的实验结果表明,我们的方法相对于其他一些主流方法取得了更好的视觉显著性检测结果。

参 考 文 献

[1] LAURENT I, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2002, 20(11): 1254-1259.

[2] LE MEUR O, LE CALLET P, BARBA D, et al. A coherent computational approach to model bottom-up visual attention [J]. *IEEE TPAMI*, 2006, 28(5): 802-817.

[3] MATHE S, SMINCHISESCU C. Dynamic eye movement datasets and learnt saliency models for visual action recognition [C]//ECCV. 2012.

[4] MATHE S, SMINCHISESCU C. Action from still image dataset and inverse optimal control to learn task specific visual scan-

paths[C]//NIPS. 2013.

[5] MATHE S, SMINCHISESCU C. Actions in the eye: Dynamic gaze datasets and learnt saliency models for visual recognition [J]. *IEEE TPAMI*, 2015, 37(7): 1408-1424.

[6] ZHANG L, TONG M H, MARKS T K, et al. SUN: A bayesian framework for saliency using natural statistics[J]. *Journal of Vision*, 2008, 8(7): 32-32.

[7] GAO D, VASCONCELOS N. Discriminant saliency for visual recognition from cluttered scenes[C]//NIPS. 2005.

[8] BRUCE N, TSOTSOS J. Saliency based on information maximization[C]//NIPS. 2006.

[9] HAREL J, KOCH C, PERONA P. Graph-based visual saliency [C]//NIPS. 2007.

[10] HOU X, ZHANG L. Saliency detection: A spectral residual approach[C]//CVPR. 2007.

[11] JUDD T, EHINGER K, DURAND F, et al. Learning to predict where humans look[C]//ICCV. 2009.

[12] KRUTHIVENTI S, GUDISA V, DHOLAKIYA J, et al. Saliency Unified: A Deep Architecture for simultaneous Eye Fixation Prediction and Salient Object Segmentation[C]//Computer Vision & Pattern Recognition. IEEE, 2016.

[13] VIG E, DORR M, COX D. Large-scale optimization of hierarchical features for saliency prediction in natural images [C] // CVPR. 2014.

[14] ZHOU Z, ZHANG L. AFSnet: Fixation Prediction in Movie Scenes with Auxiliary Facial Saliency[C]// International Conference on Brain Inspired Cognitive Systems. Cham: Springer, 2018.

[15] VOLODYMYR M, HEES N, GRAVES A. Recurrent models of visual attention[C]// Advances in Neural Information Processing Systems. 2014.

[16] ZHANG X. Progressive Attention Guided Recurrent Network for Salient Object Detection[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018.

[17] LIU N, HAN J, YANG M H. PiCANet: Learning Pixel-Wise Contextual Attention for Saliency Detection[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018.

[18] WANG W, SHEN J. Deep Visual Attention Prediction[J]. *IEEE Transactions on Image Processing*, 2017(99): 1-1.

[19] CORNIA M. Predicting Human Eye Fixations via an LSTM-

- based Saliency Attentive Model[J]. IEEE Transactions on Image Processing, 2016(27):5142-5154.
- [20] DANDAN Z, MAFL; Multi-scale Adversarial Feature Learning for Saliency Detection [C] // 2018 International Conference. 2018.
- [21] LINA W. Deep Group-Wise Fully Convolutional Network for Co-Saliency Detection With Graph Propagation[J]. IEEE Transactions on Image Processing, 2019(99).
- [22] HECHT-NIELSEN R. Theory of the backpropagation neural network[J]. Neural Networks, 1988, 1 (Supplement-1): 445-448.
- [23] ZAKARIA L, KANNALA J. Context Aware Query Image Representation for Particular Object Retrieval [C] // Scandinavian Conference on Image Analysis. Cham: Springer, 2017: 88-99.
- [24] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7132-7141.
- [25] WANG L, LU H, WANG Y, et al. Learning to detect salient objects with image-level supervision [C] // Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR). 2017: 136-145.
- [26] YAN Q, XU L, SHI J, et al. Hierarchical saliency detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013: 1155-1162.
- [27] LI G, YU Y. Visual saliency based on multiscale deep features [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 5455-5463.
- [28] ZHANG P, WANG D, LU H, et al. Amulet: Aggregating multi-level convolutional features for salient object detection [C] // IEEE International Conference on Computer Vision (ICCV 2017). Venice, Italy, 2017: 202-211.
- [29] LI G, YU Y. Visual saliency detection based on multiscale deep cnn features[J]. IEEE Transactions on Image Processing, 2016, 25(11): 5012-5024.
- [30] LUO Z, MISHRA A, ACHKAR A, et al. Non-local deep features for salient object detection [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 6593-6601.
- [31] ZHANG P, WANG D, LU H, et al. Learning uncertain convolutional features for accurate saliency detection [C] // IEEE International Conference on Computer Vision (ICCV 2017). Venice, Italy, 2017: 212-221.
- [32] LI G, YU Y. Deep contrast learning for salient object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 478-487.
- [33] LIU N, HAN J. Dhsnet: Deep hierarchical saliency network for salient object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 678-686.
- [34] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. arXiv: 1409.1556, 2014.
- [35] RONNEBERGER O, FISCHER P, BROX T. U-net Convolutional networks for biomedical image segmentation [C] // International Conference on Medical Image Computing and Computer-assisted Intervention. Springer, 2015: 234-241.
- [36] ZHAO T, WU X. Pyramid feature attention network for saliency detection [C] // IEEE Conf. Comput. Vis. Pattern Recog. . 2019: 3085-3094.
- [37] QIN X. BASNet: Boundary-Aware Salient Object Detection [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019.
- [38] WANG W. Salient Object Detection Driven by Fixation Prediction [C] // IEEE CVPR. IEEE, 2018.



LIU Xiang-yu, master. His main research interests include image processing and visual saliency detection.



JIAN Mu-wei, professor, doctoral supervisor, CCF computer vision committee, CCF multimedia committee, CCF machine learning and pattern recognition communications committee, etc. His main research interests include image processing, pattern recognition, multimedia computing, and so on.