

全局信息引导的多尺度显著物体检测模型

陈小伟¹ 张裕^{1*} 林家骏² 张晴¹

¹(上海应用技术大学 上海 201418)

²(华东理工大学 上海 200237)

摘要 目前方法不能有效提取图像的多尺度特征并完全利用全局信息,对此提出一种新的全局信息引导的多尺度显著物体检测模型。设计了多尺度特征提升模块,提升了各个侧输出特征的多尺度表征能力;利用空洞空间卷积池化金字塔模块提取图像的全局特征;将全局特征直接与各个侧边输出的多尺度特征相融合,利用全局特征引导侧边特征聚焦于目标区域中有用的中低层特征;采用由粗至细的方式得到最终的预测结果。该模型在四个广泛使用的数据集上进行了测试,并与8种近3年具有代表性的方法进行比较。实验结果表明,该模型对各种场景的显著物体检测具有较好的鲁棒性,可以均匀高亮显著目标的同时抑制背景噪声。

关键词 显著物体检测 显著性检测 全卷积网络 多尺度特征 显著图

中图分类号 TP391 文献标志码 A DOI:10.3969/j.issn.1000-386x.2022.03.024

GLOBAL CONTEXT GUIDED MULTI-SCALE FEATURE NETWORK FOR SALIENT OBJECT DETECTION

Chen Xiaowei¹ Zhang Yu^{1*} Lin Jiajun² Zhang Qing¹

¹(Shanghai Institute of Technology, Shanghai 201418, China)

²(East China University of Science and Technology, Shanghai 200237, China)

Abstract Current methods cannot represent multi-scale features effectively and utilize global information completely. This paper proposes a novel salient object detection approach using global context guide. We designed a multi-scale feature enhancement module to improve the capability of feature representation of multi-level side-output features. We utilized the atrous spatial pyramid pooling module to obtain the global features. The global features were directly fused with the multi-scale side-output features, which guided the side features to focus on the useful middle and low level features in the target area. The finally estimated results were obtained in a coarse-to-fine manner. We compared our approach with 8 state-of-the-art methods on four public benchmark datasets. Results demonstrate the robustness of the proposed approach on various scenarios, and it can evenly highlight salient targets while suppressing background noise.

Keywords Salient object detection Saliency detection Fully convolutional network Multi-scale features Saliency map

0 引言

显著物体检测的目的是模拟人的视觉感知,从杂乱背景中定位和分割出最引人注意的具有精确轮廓的物体。近年来,显著性物体检测作为一个预处理步骤

被广泛应用于各种计算机视觉任务,包括视频跟踪^[1]、目标识别^[2]和图像编辑^[3]等。

根据算法是否使用深度特征,可将显著性物体检测算法分为两大类:基于手动选择特征的传统方法^[4-5]和使用深度神经网络提取语义特征的方法^[6-7]。传统的显著性检测方法采用颜色、纹理、形状等图像中低层

特征,并利用启发式先验条件(颜色对比度、边界、物体级信息等)进行显著性物体检测计算。虽然基于传统的方法在处理简单场景图像时取得了较好的检测效果,但由于其无法提取图像深层的语义特征,因此在面对复杂图像时,检测效果与人的视觉感知结果存在较大的差异。

近年来,卷积神经网络在计算机视觉任务中显示出了其强大的特征表征和学习能力。受此启发,研究人员将卷积神经网络应用于显著性检测任务。一些基于卷积神经网络的显著性检测算法^[6,8,22]利用了图像的深度特征,取得了比传统方法更好的检测性能。随着全卷积神经网络的兴起,研究人员发现,融入中低层特征的显著性检测模型^[9-10,23-24]相比仅利用深度特征的方法更进一步提高了算法性能,因为中低层特征包含丰富的结构和细节信息,对于勾勒出完整和精确的轮廓信息具有十分重要的作用。

尽管现有算法取得了令人瞩目的成果,显著性物体检测领域仍具有如下问题需要解决:(1) 基于特征金字塔网络(Feature pyramid network, FPN)^[11]结构的显著检测模型,将深度信息逐层传递给浅层,在传递过程中,深度信息必然有损失,不能全部传递给最浅层;(2) 自然场景中包含各种尺度的物体,而某一固定大小的卷积核只能处理固定尺寸的目标物体,因此如何在每一层次的特征中融合多尺度信息值得进一步研究。

本文提出一种简单有效的基于全卷积神经网络的显著性物体检测模型,结合图像的多层次特征,探索多尺度特征表示和融合,并且将全局信息直接与每一层的特征进行融合,指导多层次特征的提取,从而提高模型的检测性能。

本文工作主要贡献如下:(1) 提出了一种新的全局信息引导的多尺度特征卷积神经网络用于显著物体检测,将全局信息直接与多层次局部特征相结合。该模型能更好利用全局信息,从而提高检测性能。(2) 设计了多尺度卷积模块,利用同一侧输出的不同尺度特征融合,提高网络各层次特征的表达和学习能力。(3) 根据常用的评价指标,在 ECSSD、DUT-OMRON、PASCAL-S 和 DUTS-TE 数据上进行算法性能比较与分析,从而说明本文算法的有效性和鲁棒性。

1 相关工作

显著性检测方法可以分为基于眼动点的显著性预

测和具有精确物体轮廓信息的显著性物体检测,本文主要关注显著性物体检测。

1.1 基于手动选择特征的方法

大部分传统的显著性物体检测方法先将图像进行超像素分割,然后采用手动选择图像的中低层特征进行显著性计算。基于局部的方法^[12]使用每个超像素的对比度或独特性等先验信息来捕获局部显著区域。而基于全局的方法^[13-14]通过使用整个图像的整体信息来计算每个超像素的显著度。由于基于手动选择特征的方法不能有效利用图像蕴含的语义信息,因此无法从复杂图像中精确检测和分割显著性物体。

1.2 基于深度特征的方法

近年来,基于卷积神经网络的显著性物体检测方法^[19]明显提高了检测性能。Wang 等^[6]提出一个深度神经网络,首先计算局部上下文中每个像素的显著性得分,然后用另一个网络在全局视图上重新评估每个对象的显著性得分。Li 等^[8]利用深度神经网络提取图像的多尺度特征,通过融合这些特征计算显著性值。Zhao 等^[15]通过整合全局和局部信息预测显著性图。然而,上述检测方法将图像区域视为基本的计算单元,网络必须运行多次得到整个图像的显著性值。

为了解决该问题,研究人员引入全卷积网络,采用图像到图像的方式进行显著性检测^[7]。利用全卷积网络的各个侧边输出的多层次特征,采用类似 U-Net 结构,进行显著性检测信息的编码和解码。网络的低层侧边输出的特征富含低层特征,但缺乏图像的整体语义信息;而网络的深层侧边输出的特征含有丰富的语义信息,但缺乏图像的结构细节。因此,结合网络的不同侧边输出的多层次特征有助于进一步提高显著性预测的准确性。

Luo 等^[16]通过一个多分辨的 4×5 网格结构融合图像的局部和全局信息,并采用 Mumford-Shah 函数进行边界优化。文献[17]采用反注意力图引导特征选择。文献[18]采用跳层连接方式将深层特征融入各浅层侧输出,从而进行显著性计算。Zhang 等^[10]利用注意力机制,逐层引导各侧边输出整合多层次特征。Zhang 等^[20]提出采用双向信息传递模型整合多层次特征。

虽然这些基于深度学习的方法已经取得了明显的成效,但是仍有很大的进步空间,使其可以在复杂场景中均匀突出整个显著目标并且准确判断边界,同时有效抑制背景噪声。

2 模型设计

2.1 模型整体结构

为了一致高亮显著区域,同时抑制无关背景噪声,本文提出一种新的全局信息引导的多特征网络(GCMF-Net)用于显著物体检测,探索利用全局信息引导多层次特征提取方法及多尺度特征的检测和融合策略。

GCMF-Net 的整体结构如图 1 所示,主要包含多尺度特征提升模块(MFEM)和全局信息引导模块(GCGM)。其中,全局信息引导模块 GCGM 由空洞空间卷积池化金字塔(ASPP)模块^[30]和全局特征融合模块(GFFM)构成。本文使用基于 VGG-16 的全卷积网络结构作为主干网络模型,采用 PFN 结构,以由粗至细的方式更新显著性图。利用主干网络提取图像的多层次特征;利用 MFEM 模块提取不同层次的多尺度特征,并将这些特征进行融合;利用 ASPP 模块获取多尺度的全局语义特征;利用 GFFM 模块融合不同层次的多尺度特征与全局特征,从而准确捕获多层次多尺度的显著区域特征。

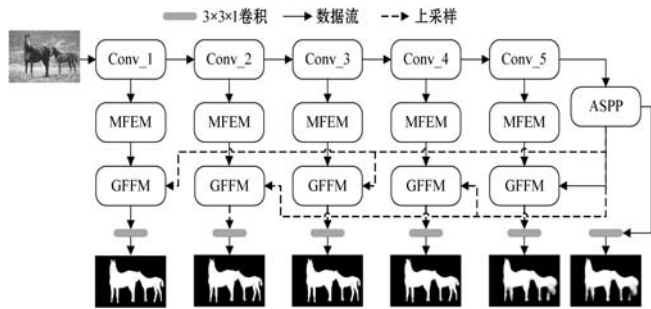


图 1 GCMF-Net 整体结构

2.2 主干网络

GCMF-Net 采用已在 ImageNet 数据集上进行目标分类任务预训练的 VGG16 网络^[21]作为主干网络,提取侧边输出的多层次特征,这些特征的空间分辨率按步幅 2 递减。由于显著性物体检测任务是像素级预测任务,因此移除 VGG16 网络的第 5 个池化层及其后的全连接层。VGG16 网络由 5 个卷积层组构成,将其标记为 $Conv_i (i \in \{1, 2, \dots, 5\})$,其产生的多层次特征图标记为 f_i^a ,其中 f_1^a 具有最大的空间分辨率, f_5^a 具有最小的空间分辨率。

2.3 多尺度特征提升模块

由于主干网络的卷积层组采用固定大小的卷积核得到侧边输出特征,因此各侧边输出特征均对固定尺寸的显著物体具有较好的响应值。然而,自然场景中

包含各种尺度的物体,多尺度特征检测和融合模块的研究可以提高模型处理多尺度目标的能力,从而提升整个网络的显著物体检测性能。

因此,本文设计了多尺度特征提升模块 MFEM 进行多尺度特征的提取和融合,进一步提高侧边输出特征的表达力。MFEM 的总体结构如图 2 所示。对于空间分辨率为 $H_i \times W_i$ 的侧边输出特征 f_i^a ,MFEM 采用 3 个全局平均池化层对其进行下采样,采样率为 r 取 2、4、8 以得到不同尺度的特征;将侧边输出特征和其根据 3 个不同采样率得到的多尺度特征,分别通过 3×3 大小的卷积核将维度降至 32 维;为了充分利用同一侧输出的多尺度特征,通过上采样和下采样操作融合相邻尺度的特征;将所有融合后的特征进行特征连接,并通过一个 1×1 的卷积层得到融合后的空间分辨率为 $H_i \times W_i$ 的多尺度特征 f_i^b 。

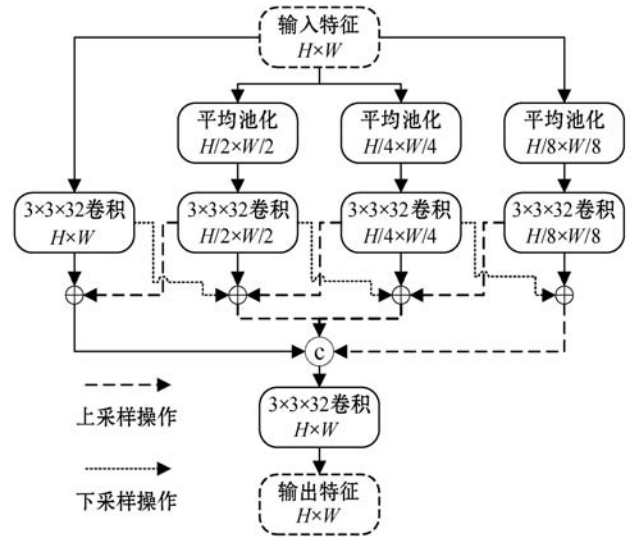


图 2 MFEM 结构

2.4 全局信息引导模块

本文利用空洞空间卷积池化金字塔 ASPP 模块捕获多尺度的全局语义信息,从而生成包含显著物体位置的全局特征。实验将 ASPP 的三个空洞卷积层膨胀率分别设置为 4、6 和 8。

卷积神经网络产生的高层特征包含图像丰富的语义信息,而卷积神经网络低层侧边输出特征包含更多的颜色、纹理和形状等中低层图像特征,因此融合各个侧边输出的不同层次特征,能提高整个网络的特征表达和学习能力。

现有模型一般采用 FPN 结构,将高层语义特征逐渐与低层特征相融合,采用由粗到细的方式逐渐更新显著性图,提高检测质量。但是这种逐层更新的方式在一定程度上削弱了高层语义信息对低层特征提取和

更新的引导作用。因此,本文提出将全局语义信息直接与各个侧边输出的多尺度融合特征相结合,利用高层语义信息直接引导侧边输出的中低层特征集中于有效目标区域的特征表征与学习。

在本文提出的全局特征融合模块中,将侧边输出、深层的显著特征图、ASPP模块生成的全局语义特征图进行维度方向的连接,再使用32维的 3×3 卷积核进行特征融合,可表示为:

$$s_i = \text{Conv}_{(3 \times 3, 32)}(\text{Conat}(f_i^b, f^g, s_{i+1})) \quad (1)$$

式中: s_i 表示第*i*侧的显著特征图; f_i^b 表示第*i*侧MFEM模块输出的多尺度特征; f^g 表示ASPP模块输出的全局特征; $\text{Conat}(\cdot)$ 表示特征维度上的连接操作; $\text{Conv}_{(3 \times 3, 32)}$ 表示32维的 3×3 卷积操作。

3 实验

3.1 基准数据集

为评价算法性能,在四个公开的基准数据集上进行了一系列定性和定量评价实验。数据集包括ECSSD、DUT-OMRON、PASCAL-S和DUTS-TE,各数据集的图像数如表1所示。

表1 基准数据集信息

数据集	ECSSD	DUT-OMRON	PASCAL-S	DUTS-TE
图像数	1 000	5 168	800	5 199

3.2 评价指标

本文采用常用的准确率-召回率(PR)曲线、F测度(F_β)、平均F测度($\text{avg } F_\beta$)、加权F测度(wF_β)和平均绝对误差(MAE)作为算法性能评价的量化指标。

根据从0到255的阈值,将显著图二值化,并与真值图进行比较来计算准确率和召回率。PR曲线显示了在某一数据集上不同阈值下的显著图的平均准确率和召回率。

F_β 用于综合考虑准确率和召回率:

$$F_\beta = \frac{(1 + \beta^2) \times p \times r}{\beta^2 \times p + r} \quad (2)$$

式中: p 和 r 分别代表准确率和召回率; β 是衡量准确率和召回率的平衡参数,与文献[18]参数选取一致,本文实验将 β^2 设为0.3。

wF_β 使用加权准确率 p^w 和加权召回率 r^w 进行计算:

$$F_\beta^w = \frac{(1 + \beta^2) \times p^w \times r^w}{\beta^2 \times p^w + r^w} \quad (3)$$

MAE用来评价显著图和真值图之间的平均像素误差:

$$MAE = \frac{1}{h \times w} \sum_{i=1}^h \sum_{j=1}^w |S_{ij} - G_{ij}| \quad (4)$$

式中: S 表示最终显著图; G 表示真值图; h 和 w 分别表示图像的高度和宽度, i 和 j 表示像素点的位置。

3.3 实施细节

所提模型的训练和测试是在具有Intel i7-7700k CPU(4.2 GHz)、32 GB RAM和一块英伟达GTX TITAN GPU的台式计算机上,使用Python实现。

所提出的网络基于公开的Pytorch框架。本文模型使用DUTS-TR作为训练集。DUTS-TR包含10 553幅训练图像。为了提高模型的鲁棒性,本文通过随机水平翻转进行训练集增强。所有的训练图像的分辨率均 320×320 ,每次只加载一个图像。学习率设置为 $5e-5$,使用Adam作为优化算子,权重衰减为 $5e-4$,一共训练25期。

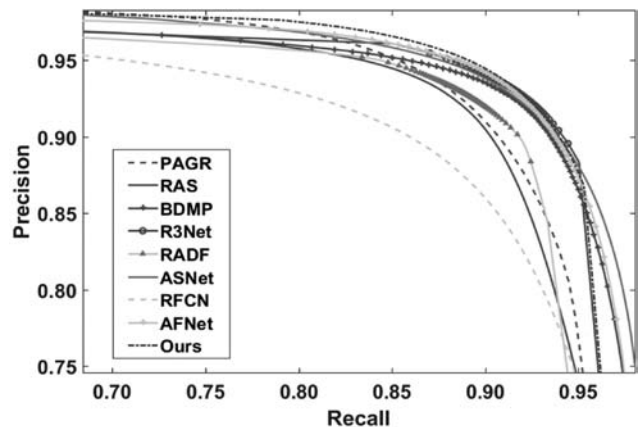
为了进一步提高边缘像素的检测精度,使用CRF^[29]对本文模型输出的预测结果进行显著图优化。

3.4 算法性能对比

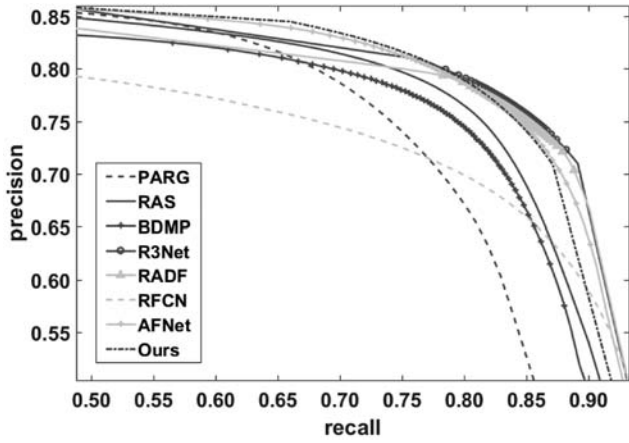
将本文方法与8种近三年发表的具有代表性的方法进行了比较,包括PAGR^[10]、RAS^[17]、BDMP^[20]、R3Net^[28]、RADF^[25]、ASNet^[26]、RFCN^[7]和AFNet^[27]。为了比较的公平性,所有对比算法的显著图均使用作者提供的实现方法获取或作者公开的显著预测图。

3.4.1 定量比较

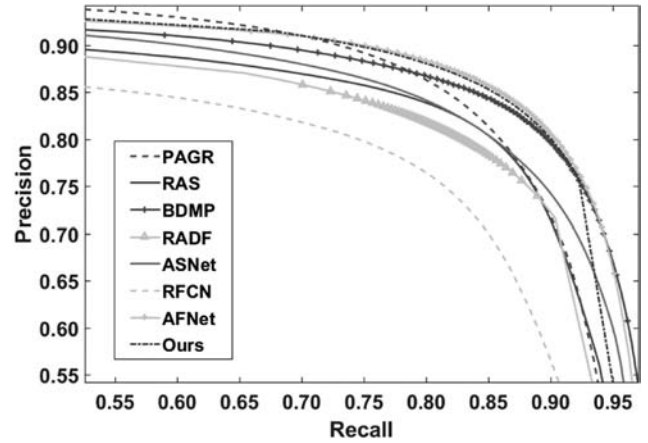
本文提出的GCMF-Net模型与具有代表性的8种主流方法的PR曲线图如图3所示。由图3可知,GCMF-Net在ECSSD、DUTS-TE和DUT-OMRON这3个基准数据集上,较现有的代表性算法具有竞争力,只在PASCAL-S数据集上略逊于ASNet和AFNet。



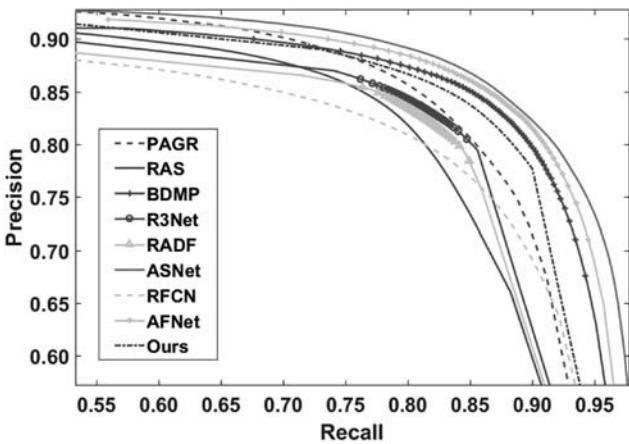
(a) ECSSD数据集



(b) DUT-OMRON 数据集



(d) DUTS-TE 数据集



(c) PASCAL-S 数据集

图3 不同方法在不同测试集上的PR曲线比较

此外,GCMF-Net模型与8种主流方法在四个公开基准数据集上就 F_β 、 $avg F_\beta$ 、 wF_β 和 MAE 指标进行了定量比较,结果如表2和表3所示,“/”表示原文作者没有提供该数据集上的显著性图。

由表2和表3可知:(1) GCMF-Net在PASCAL-S数据集上的 $avg F_\beta$ 指标略逊于AFNet,而MAE指标位居第四;(2) GCMF-Net在ECSSD、DUT-OMRON和DUTS-TE数据上,均具有最佳表现。

综合本文方法在四个基准数据集上的PR曲线、 F_β 、 $avg F_\beta$ 、 wF_β 和MAE评价结果,GCMF-Net较近年的主流方法具有一定的优越性。

表2 不同方法在ECSSD和DUT-OMRON基准数据集上的性能比较

算法	ECSSD 数据集				DUT-OMRON 数据集			
	F_β	$avg F_\beta$	wF_β	MAE	F_β	$avg F_\beta$	wF_β	MAE
PAGR	0.894 2	0.904 7	0.850 7	0.060 9	0.711 4	0.750 7	0.643 6	0.070 9
RAS	0.889 4	0.902 4	0.864 8	0.056 4	0.712 9	0.763 9	0.688 9	0.061 7
BDMP	0.868 4	0.901 8	0.863 2	0.044 5	0.691 7	0.747 2	0.668 5	0.063 6
R3Net	0.914 0	0.926 0	0.902 8	0.040 2	0.747 7	0.780 7	0.713 9	0.062 5
RADF	0.904 9	0.917 5	0.889 3	0.048 7	0.741 3	0.773 6	0.707 6	0.061 5
ASNet	0.875 2	0.900 9	0.859 4	0.046 8	/	/	/	/
RFCN	0.871 5	0.881 1	0.842 0	0.066 7	0.678 6	0.718 1	0.632 8	0.077 0
AFNet	0.907 6	0.917 3	0.885 1	0.041 8	0.738 6	0.778 9	0.709 9	0.057 4
GCMF-Net	0.915 7	0.930 3	0.905 8	0.039 3	0.755 6	0.790 7	0.727 7	0.054 7

表3 不同方法在PASCAL-S和DUTS-TE基准数据集上的性能比较

算法	PASCAL-S				DUTS-TE			
	F_β	$avg F_\beta$	wF_β	MAE	F_β	$avg F_\beta$	wF_β	MAE
PAGR	0.805 5	0.828 3	0.750 1	0.089 4	0.788 3	0.824 5	0.745 3	0.054 9
RAS	0.784 3	0.816 0	0.748 2	0.101 3	0.755 1	0.806 3	0.735 7	0.059 0
BDMP	0.763 5	0.827 2	0.761 7	0.073 6	0.750 6	0.813 5	0.748 2	0.048 4
R3Net	0.801 3	0.829 2	0.764 3	0.092 0	/	/	/	/

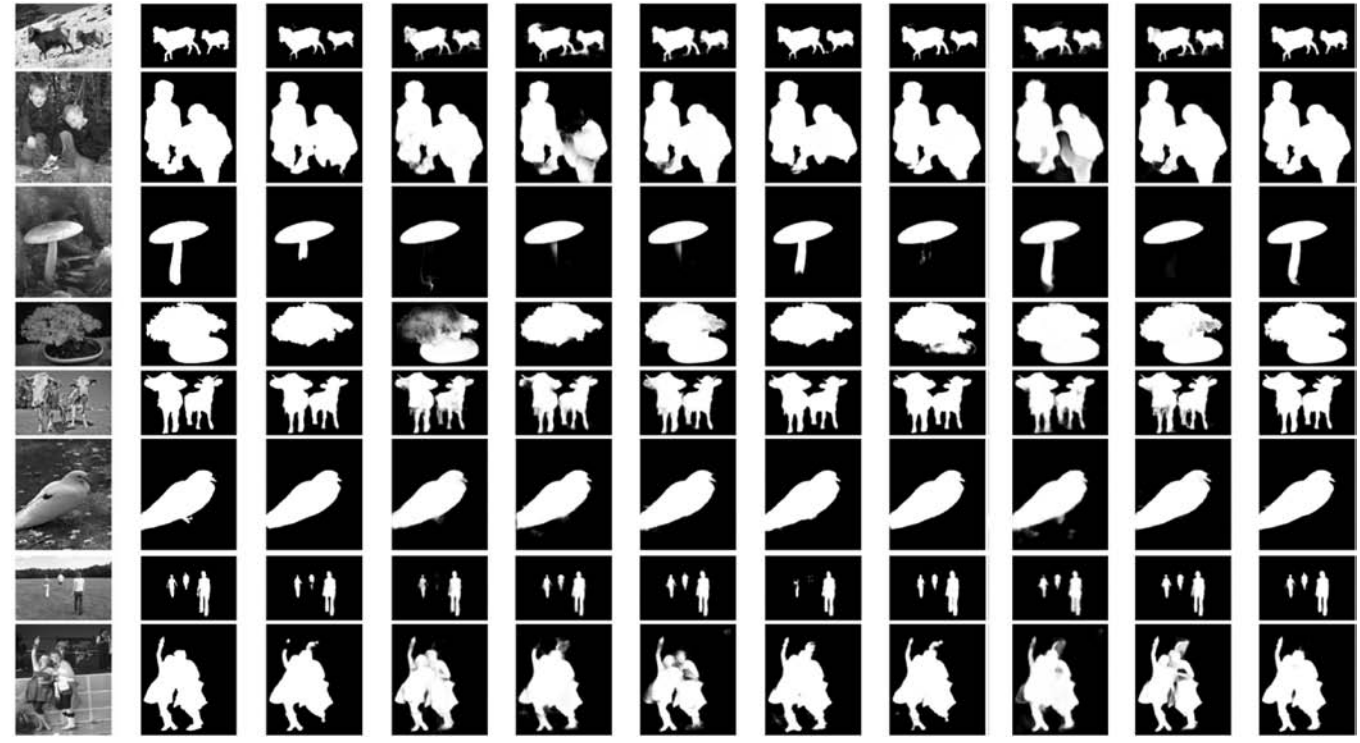
续表 3

算法	PASCAL-S				DUTS-TE			
	F_β	avg F_β	wF_β	MAE	F_β	avg F_β	wF_β	MAE
RADF	0.795 3	0.821 3	0.757 5	0.096 6	0.772 3	0.805 4	0.742 7	0.061 4
ASNet	0.793 3	0.838 1	0.774 4	0.069 9	0.732 6	0.794 9	0.794 9	0.060 1
RFCN	0.775 6	0.799 1	0.730 3	0.103 6	0.713 9	0.735 2	0.567 9	0.089 5
AFNet	0.820 9	0.842 8	0.784 6	0.070 0	0.797 3	0.839 1	0.779 5	0.045 4
GCMF-Net	0.820 9	0.841 6	0.788 5	0.074 9	0.821 7	0.851 5	0.804 2	0.043 6

3.4.2 定性比较

图 4 显示了 GCMF-Net 模型与不同算法在四个公开数据集上生成显著性图的视觉比较。由图 4 可知,本文提出的 GCMF-Net 方法在各种复杂场景(例如多

目标、复杂背景、大目标、小目标、目标接触边界、复杂目标等)中表现良好,一致高亮了显著区域,并有效抑制了背景噪声。



(a) 输入 (b) 真值 (c) PAGR (d) RAS (e) BAMP (f) R3Net (g) RADF (h) ASNet (i) RFCN (j) AFNet (k) 本文
图 4 不同方法生成显著图的视觉比较

3.5 本文模型分析

3.5.1 不同侧输出的预测结果分析

本文对 GCMF-Net 模型的各个侧边得到的显著性图,在 ECSSD 基准数据集上的进行检测性能比较,结果如表 4 所示。可以看出,最浅层的侧边得到的显著性图性能最佳,这表明了所提模型采用由粗至细的逐层更新策略是有效的。

表 4 不同侧输出的性能比较

侧输出	F_β	avg F_β	wF_β	MAE
side6	0.8649	0.8823	0.8332	0.0560
side5	0.866 4	0.883 7	0.835 3	0.055 5

续表 4

侧输出	F_β	avg F_β	wF_β	MAE
side4	0.877 9	0.894 9	0.850 9	0.051 1
side3	0.882 4	0.899 2	0.856 9	0.049 5
side2	0.883 4	0.900 1	0.858 2	0.049 2
side1	0.883 5	0.900 1	0.858 3	0.049 2

3.5.2 不同模块性能分析

为了验证本文方法的可靠性和有效性,从本文模型中分别移除不同的模块和结构,并且测试这些新的网络模型在 ECSSD 和 PASCAL-S 数据集上的检测性能。

采用 F_β 、avg F_β 、 wF_β 和 MAE 对未采用 CRF 进行

显著图优化的结果进行评价,如表 5 所示,本文方法中的各个模块能有效提高检测性能。其中: $Ours_{wo_MFEM}$ 表示移除多尺度特征提取的特征优化模块, $Ours_{wo_GCCM}$ 表示移除整个全局信息引导模块, $Ours_{wo_GFFM}$ 表示保

留全局信息引导模块中的全局特征提取部分,移除全局特征与其他侧输出的局部特征进行融合部分, $Ours_{wo_ASPP}$ 表示移除全局信息引导模块中的全局特征提取部分。

表 5 不同模块的性能比较

算法	ECSSD				PASCAL-S			
	F_{β}	avg F_{β}	wF_{β}	MAE	F_{β}	avg F_{β}	wF_{β}	MAE
$Ours_{wo_MFEM}$	0.878 9	0.895 4	0.852 7	0.050 4	0.789 3	0.817 7	0.749 2	0.081 9
$Ours_{wo_GCCM}$	0.882 8	0.899 5	0.857 8	0.048 9	0.793 9	0.817 8	0.751 6	0.082 3
$Ours_{wo_ASPP}$	0.876 3	0.891 5	0.846 7	0.055 0	0.784 4	0.811 1	0.740 2	0.086 8
$Ours_{wo_GFFM}$	0.875 8	0.891 6	0.846 7	0.058 7	0.786 0	0.811 0	0.743 8	0.087 9
本文方法	0.883 5	0.900 1	0.858 3	0.049 2	0.791 8	0.819 2	0.751 6	0.082 2

4 结 语

本文提出了一种利用多尺度和多层次特征进行显著性目标检测的方法。针对主干网络的每个侧输出进行多尺度特征的提取和融合;利用从最深层侧输出提取的全局语义特征引导侧输出的多尺度特征聚焦于主要区域的特征表达和学习;采用类似于 FPN 的网络结构,逐渐更新预测显著性图;采用深监督方式进行网络训练,从而得到预测显著性图。在测试阶段,为了进一步获得具有精确轮廓和均匀一致内部区域的显著物体检测结果,使用基于全连接的 CRF 进行显著性图更新。在四个公开的基准数据集上的实验结果表明,本文方法较 8 种近年发表的主流算法具有优越性。未来研究将考虑引入显著性物体的轮廓信息以获得更清晰的目标边界。

参 考 文 献

- [1] Hong S, You T, Kwak S, et al. Online tracking by learning discriminative saliency map with convolutional neural network[C]//Proceedings of the 32nd International Conference on International Conference on Machine Learning. JMLR. org,2015:597-606.
- [2] Wang H Y, Yan Y X, Hua J, et al. Pedestrian recognition in multi-camera networks using multi-level important salient feature and multicategory incremental learning[J]. Pattern Recognition,2017,67:340-352.
- [3] Zhang Q, Yuan G Z, Xiao C X, et al. High-quality exposure correction of underexposed photos[C]//Proceedings of the 26th ACM International Conference on Multimedia. ACM,2018:582-590.
- [4] Qin Y, Lu H C, Xu Y Q, et al. Saliency detection via cellular automata[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015: 110-119.
- [5] Wang J D, Jiang H Z, Yuan Z J, et al. Salient object detection: A discriminative regional feature integration approach [J]. International Journal of Computer Vision, 2017, 123: 251-268.
- [6] Wang L J, Lu H C, Ruan X, et al. Deep networks for saliency detection via local estimation and global search[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015: 3183-3192.
- [7] Wang L Z, Wang L J, Lu H C, et al. Salient object detection with recurrent fully convolutional networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(7): 1734-1746.
- [8] Li G B, Yu Y Z, Visual saliency based on multiscale deep features[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015: 5455-5463.
- [9] Liu N, Han J W. DHSNet: Deep hierarchical saliency network for salient object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 678-686.
- [10] Zhang X N, Wang T T, Qi J Q, et al. Progressively attention guided recurrent network for salient object detection [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 714-722.
- [11] Lin T Y, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 936-944.

- [12] Itti L, Koch C, 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.
- [13] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection[C]//2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2011: 409 – 416.
- [14] 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. IEEE, 2013: 3166 – 3173.
- [15] Zhao R, Ouyang W L, Li H S, et al. Saliency detection by multi-context deep learning[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015: 1265 – 1274.
- [16] Luo Z M, 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.
- [17] Chen S S, Tan X L, Wang B, et al. Reverse attention for salient object detection[C]//Proceedings of 2018 European Conference on Computer Vision. Springer, 2018: 236 – 252.
- [18] Hou Q B, Cheng M M, Hu X W, et al. Deeply supervised salient object detection with short connections [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(4): 815 – 828.
- [19] Li G B, Yu Y Z. Deep contrast learning for salient object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 478 – 487.
- [20] Zhang L, Dai J, Lu H C, et al. A bi-directional message passing model for salient object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 1741 – 1750.
- [21] Simonyan K, Zisserman A. Very deep convolutional networks for largescale image recognition[C]//Proceedings of International Conference on Learning Representations, 2015.
- [22] Lee G, Tai Y W, Junmo K. Deep saliency with encoded low level distance map and high level features[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 660 – 668.
- [23] Zhang P P, Wang D, Lu H C, et al. Amulet: Aggregating multi-level convolutional features for salient object detection [C]//2017 IEEE International Conference on Computer Vision (ICCV). IEEE, 2017: 202 – 211.
- [24] Li X, Yang F, Cheng H, et al. Contour knowledge transfer for salient object detection [C]//Proceedings of European Conference on Computer Vision (ECCV). Springer, 2018: 370 – 385.
- [25] Hu X W, Zhu L, Qin J, et al. Recurrently aggregating deep features for salient object detection[C]//Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence. AAAI Press, 2018: 6943 – 6950.
- [26] Wang W G, Shen J B, Dong X P, et al. Salient object detection driven by fixation prediction[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 1711 – 1720.
- [27] Feng M Y, Lu H C, Ding E. Attentive feedback network for boundary-aware salient object detection [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019: 1623 – 1632.
- [28] Deng Z J, Hu X W, Zhu L, et al. R³Net: Recurrent residual refinement network for saliency detection[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. AAAI Press, 2018: 684 – 690.
- [29] Krahenbuhl P, Koltun V. Efficient inference in fully connected CRFs with gaussian edge potentials[C]//25th Annual Conference on Neural Information Processing Systems, 2011: 109 – 117.
- [30] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834 – 848.
-
- (上接第 131 页)
- [10] Jin F L, Liu Q F, Liu H, et al. A low complexity signal detection scheme based on improved newton iteration for massive MIMO systems [J]. *IEEE Communications Letters*, 2019, 23(4): 748 – 751.
- [11] 赵书峰, 申滨, 杨芙蓉. 大规模 MIMO 系统低复杂度混合迭代信号检测 [J]. *电信科学*, 2017, 33(7): 39 – 46.
- [12] Xu Y H, Wang J. Signal detection algorithm based on SOR algorithm in Massive MIMO system [C]//2018 14th IEEE International Conference on Signal Processing (ICSP). IEEE, 2018: 662 – 667.
- [13] 范莉花. 面向 5G 的大规模 MIMO 无线传输技术研究 [J]. *通讯世界*, 2019(7): 97 – 98.
- [14] Xue Y, Zhang C, Zhang S Q, et al. Steepest descent method based soft-output detection for massive MIMO uplink [C]//2016 IEEE International Workshop on Signal Processing Systems (SiPS). IEEE, 2016: 273 – 278.