

一种轻量化边界感知的极限下采样显著性目标检测网络

慈兆会^{1*}, 聂广远²

¹合肥测准量子科技有限公司, 安徽 合肥

²合肥工业大学计算机与信息学院, 安徽 合肥

收稿日期: 2026年6月22日; 录用日期: 2026年7月20日; 发布日期: 2026年7月29日

摘要

显著性目标检测(SOD)旨在从复杂场景中定位最具视觉吸引力的区域, 是图像分割和目标识别的重要预处理步骤。针对现有极限下采样网络(EDN)存在的深层下采样冗余、多尺度特征融合不足及边界表达能力弱的问题, 本文提出轻量化边界感知极限下采样网络(LB-EDN)。具体包括: 设计轻量极限下采样模块(Light-EDB)以保留高层语义信息并降低计算复杂度; 提出分组尺度相关金字塔卷积(Group-SCPC)实现轻量多尺度特征建模; 构建轻量边界注意力模块(LBA)提升边界清晰度和结构对齐; 并对解码器进行通道压缩与模块轻量化设计。LB-EDN在六个公开数据集上的实验表明, 其总参数量仅1.73 M, 计算量2.04 G FLOPs, 同时在DUTS-TE上达成Max F_β 0.8521、MAE 0.0472、S-measure 0.9513、E-measure 0.9436, 相较原EDN-Lite明显提升。结果验证了方法在结构一致性、边界表达及轻量化部署上的有效性, 适用于资源受限边缘设备。

关键词

显著性目标检测, 轻量化网络, 极限下采样, 边界注意力, 多尺度特征融合

A Lightweight Boundary-Aware Extreme Downsampling Network for Salient Object Detection

Zhaohui Ci^{1*}, Guangyuan Nie²

¹Hefei Cezhun Quantum Technology Co., Ltd., Hefei Anhui

²School of Computer and Information, Hefei University of Technology, Hefei Anhui

Received: June 22, 2026; accepted: July 20, 2026; published: July 29, 2026

*通讯作者。

文章引用: 慈兆会, 聂广远. 一种轻量化边界感知的极限下采样显著性目标检测网络[J]. 计算机科学与应用, 2026, 16(7): 212-225. DOI: 10.12677/csa.2026.167253

Abstract

Salient object detection (SOD) aims to locate the most visually attractive regions in complex scenes and serves as a crucial preprocessing step for image segmentation and object recognition. To address the limitations of existing extremely-downsampled networks (EDN), including redundant deep downsampling, inefficient multi-scale feature fusion, and weak boundary representation, this paper proposes a Lightweight Boundary-aware EDN (LB-EDN). Specifically, we design a Lightweight Extremely-Downsampled Block (Light-EDB) to retain high-level semantic information while reducing computational cost; introduce Group Scale-Correlated Pyramid Convolution (Group-SCPC) for efficient multi-scale feature modeling; develop a Light Boundary Attention (LBA) module to enhance boundary clarity and structural alignment; and perform channel compression and module light-weighting in the decoder. Extensive experiments on six public datasets demonstrate that LB-EDN achieves only 1.73 M parameters and 2.04 G FLOPs, while delivering strong performance. On DUTS-TE, it achieves Max F_β 0.8521, MAE 0.0472, S-measure 0.9513, and E-measure 0.9436, showing clear improvements over the original EDN-Lite. These results validate the effectiveness of LB-EDN in structural consistency, boundary clarity, and lightweight deployment, making it suitable for resource-constrained edge devices.

Keywords

Salient Object Detection, Lightweight Network, Extremely-Downsampled, Boundary Attention, Multi-Scale Feature Fusion

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

显著性目标检测作为计算机视觉领域的基础研究问题之一,其核心目标是模拟人类视觉系统的选择性注意机制,从复杂的自然图像中快速定位最能吸引观察者注意力的前景区域。该技术已被广泛应用于图像分割[1]、目标检测[2]、图像裁剪[3]、视频摘要[4]以及医学影像分析[5]等众多领域。例如,在图像分割任务中,显著性检测结果可以作为先验信息,引导分割模型聚焦于目标区域,从而提高分割精度并降低计算复杂度;在自动驾驶系统中,显著性检测可以帮助车辆快速识别道路上的行人、车辆等关键目标,提升系统的实时性和安全性。

早期的显著性检测方法主要基于手工设计的低层视觉特征,如颜色对比度[6]、纹理分布[7]、边缘信息[8]等。这类方法计算简单,但在复杂背景、光照变化以及目标遮挡等场景下表现有限,难以准确区分显著目标与背景。随着深度学习技术的快速发展,基于卷积神经网络(Convolutional Neural Network, CNN)的显著性检测方法逐渐成为主流。这类方法通过端到端的训练自动学习多层次的特征表示,能够有效捕捉目标的语义信息,显著提升了检测精度。例如, Hou 等人[9]提出的 DSS 网络通过引入短连接和深度监督机制,实现了多尺度特征的有效融合; Qin 等人[10]提出的 U-2-Net 采用嵌套的 U 型结构,在保持轻量化的同时取得了优异的检测性能。

然而,现有基于深度学习的显著性检测方法大多追求检测精度的提升,往往忽略了模型的计算效率。这些方法通常采用复杂的网络架构和大量的参数,需要高性能的 GPU 才能实时运行,难以部署在计算能

力和存储资源有限的边缘设备(如智能手机、嵌入式系统、无人机等)上。为了解决这一问题, 研究者们开始致力于轻量化显著性检测方法的研究。目前的轻量化方法主要分为两类: 一类是对已有的大模型进行压缩, 如通道剪枝[11]、权重量化[12]、知识蒸馏[13]等; 另一类是直接设计高效的网络架构, 如采用轻量级骨干网络(MobileNet [14]、ShuffleNet [15]等)并设计高效的特征融合模块。

极限下采样网络(EDN [16])是近年来提出的一种高效显著性检测架构, 其核心思想是通过将特征图分辨率大幅降低至输入图像的 $1/128$, 在保证高层语义信息表达能力的同时, 显著减少计算量。EDN-Lite 作为 EDN 的轻量化版本, 采用 MobileNetV2 作为骨干网络, 通过极端下采样块(Extremely-Downsampled Block, EDB)和尺度相关金字塔卷积(Scale-Correlated Pyramid Convolution, SCPC)实现了高效的多尺度特征融合, 在保持较低计算复杂度的同时取得了与大模型相当的检测精度。然而, EDN-Lite 仍存在以下三个主要问题:

- 1) 解码器通道冗余: 解码器高层通道数设置过大, 存在大量计算冗余, 限制了模型推理速度的进一步提升;
- 2) 边界信息表达不足: 缺乏显式的边界建模机制, 导致检测结果的边界模糊, 与真值图的结构对齐度较低;
- 3) 多尺度特征建模效率低: 原始 SCPC 模块采用全通道卷积, 参数量和计算量仍然较大, 存在优化空间。

针对上述问题, 本文提出一种轻量化边界感知极限下采样网络 LB-EDN。本文的主要贡献如下:

- 1) 改进轻量极限下采样模块 Light-EDB, 去除原始 EDN 中冗余的 conv7 下采样层, 采用 1×1 卷积进行特征变换, 在保留高层语义结构信息的同时, 将 EDB 模块的参数量减少约 50%;
- 2) 设计分组尺度相关金字塔卷积 Group-SCPC, 将输入特征图沿通道维度分组, 每组采用不同膨胀率的空洞卷积提取多尺度特征, 并通过尺度相关性机制融合相邻组的特征, 在保持多尺度表达能力的同时, 将 SCPC 模块的参数量减少约 40%;
- 3) 构建轻量边界注意力模块 LBA, 采用自适应缩减率策略生成边界注意力图, 以极小的计算代价显式增强边界区域的特征响应, 显著提升了检测结果的边界清晰度和结构对齐度;
- 4) 在六个标准显著性检测数据集上进行了全面的实验, 结果表明本文方法在保持较低计算复杂度的同时, 在结构相关指标上取得了显著提升, 实现了效率与精度的良好平衡。

本文后续内容安排如下: 第 2 节介绍相关工作; 第 3 节详细阐述本文提出的 LB-EDN 网络架构及各个模块的设计细节; 第 4 节介绍实验设置、实验结果及分析; 第 5 节总结全文并展望未来工作。

2. 相关工作

2.1. 基于深度学习的显著性目标检测

基于深度学习的显著性目标检测方法大致可以分为基于全卷积网络(Fully Convolutional Network, FCN)的方法、基于多尺度特征融合的方法和基于注意力机制的方法三类。

基于 FCN 的方法是最早将深度学习应用于显著性检测的方法。Long 等人[17]提出的 FCN 首次实现了端到端的像素级预测, 为显著性检测提供了新的思路。Li 等人[18]将 FCN 应用于显著性检测, 通过融合多层卷积特征, 取得了优于传统方法的性能。然而, 早期的 FCN 方法由于缺乏多尺度特征融合机制, 难以准确检测不同大小的目标。

为了解决多尺度目标检测问题, 研究者们提出了多种多尺度特征融合方法。Zhao 等人[19]提出的 PSPNet 通过金字塔池化模块聚合不同区域的上下文信息, 有效提升了多尺度目标的检测性能。Chen

等人[20]提出的 DeepLab 系列采用空洞卷积和空洞空间金字塔池化(Atrous Spatial Pyramid Pooling, ASPP)模块, 在不降低特征图分辨率的情况下扩大感受野, 取得了优异的分割和检测性能。在显著性检测领域, Hou 等人[9]提出的 DSS 网络通过短连接融合不同层次的特征, 并采用深度监督机制引导网络学习; Liu 等人[21]提出的 PoolNet 通过引入全局指导模块(Global Guidance Module, GGM)和特征聚合模块(Feature Aggregation Module, FAM), 有效融合了全局与局部特征, 提高了显著目标检测性能。

近年来, 注意力机制被广泛应用于显著性检测, 通过自适应地分配特征权重, 引导网络聚焦于显著区域。Liu 等人[22]提出的 PiCANet 采用像素级上下文注意力机制, 能够捕捉每个像素的全局上下文信息, 从而提升显著区域的定位能力; Chen 等人[23]提出的反向注意力机制通过逐步擦除已检测的显著区域, 引导网络聚焦于未被发现的目标部分; Liu 等人[24]提出的 BASNet 引入边界感知损失函数, 显式优化边界区域的预测结果。然而, 这些方法通常计算复杂度较高, 难以在边缘设备上部署。

2.2. 轻量化显著性目标检测

随着移动设备和嵌入式系统的普及, 轻量化显著性检测方法受到了越来越多的关注。目前的轻量化方法主要分为模型压缩和高效架构设计两类。

模型压缩方法通过对已有的大模型进行剪枝、量化或蒸馏, 在尽量保持精度的同时减小模型尺寸。例如, He 等人[11]提出了一种基于通道剪枝的网络压缩方法, 通过移除冗余通道有效减少模型参数量与计算复杂度, 在保持较高精度的同时提升了网络推理效率; Courbariaux 等人[12]提出的二值化神经网络将权重和激活值量化为 1 位, 显著降低了存储和计算开销; Hinton 等人[13]提出的知识蒸馏方法通过将大模型的知识迁移到小模型中, 提升小模型的性能。然而, 模型压缩方法的性能提升有限, 且往往需要复杂的压缩过程。

高效架构设计方法直接针对边缘设备的特点, 设计轻量级的网络架构。这类方法通常采用 MobileNet [14]、ShuffleNet [15]等轻量级骨干网络, 并设计高效的特征融合模块。例如, Qin 等人[10]提出的 U-2-Net 采用嵌套的 U 型结构和残差 U 型块, 在保持 1.1M 参数量的同时取得了优异的检测性能; U-2-Net-Lite 进一步将参数量压缩至 0.2M, 但精度损失较大。Wu 等人[16]提出的 EDN (Extremely-Downsampled Network) 通过极限下采样策略将特征图分辨率降低至输入图像的 1/128, 从而显著降低了模型计算复杂度, 提高了显著目标检测的推理效率。

EDN-Lite 采用 MobileNetV2 作为骨干网络, 参数量仅为 1.8 M, 是目前最具代表性的轻量化显著性检测方法之一。然而, EDN-Lite 仍存在边界模糊、结构对齐度低等问题, 有待进一步改进。

2.3. 边界感知显著性目标检测

边界质量是评价显著性检测结果的重要指标。早期的方法通常将边界检测作为辅助任务, 通过引入额外的边界检测分支来提升边界质量。例如, Wang 等人[25]提出的 BDRAR 网络通过递归的边界检测和区域细化模块, 逐步优化检测结果的边界; Zhao 等人[26]提出的 EGNNet 同时预测显著图和边缘图, 通过双向引导机制实现相互增强。然而, 这些方法需要额外的边界标注, 增加了训练的复杂度。

近年来, 注意力机制被广泛应用于边界增强。例如, Liu 等人[24]提出的 BASNet 引入边界感知损失函数, 通过加权交叉熵损失显式优化边界区域的预测; Wu 等人[27]提出的 CAGNet 通过上下文感知的门控机制, 自适应地增强边界区域的特征。这些方法不需要额外的标注, 仅通过网络结构设计即可提升边界质量, 更加轻量高效。本文提出的 LBA 模块采用类似的思路, 通过生成边界注意力图来增强边界特征, 计算开销极小, 适合集成到轻量化网络中。

3. 本文方法

3.1. 整体网络架构

本文提出的 LB-EDN 网络采用经典的编码器 - 解码器架构, 整体结构如图 1 所示。编码器输出 5 层多尺度特征, 经 Light-EDB、Group-SCPC 模块特征增强与 LBA 边界注意力优化后, 生成 5 个不同尺度的显著性预测图; 将各尺度预测图上采样至输入图像尺寸, 通过加权融合得到最终显著性检测结果。编码器采用预训练的 MobileNetV2 [14] 作为骨干网络, 用于提取多尺度的视觉特征。MobileNetV2 采用倒置残差结构和线性瓶颈设计, 在保证特征表达能力的同时具有极低的计算复杂度, 非常适合作为轻量化网络的骨干。

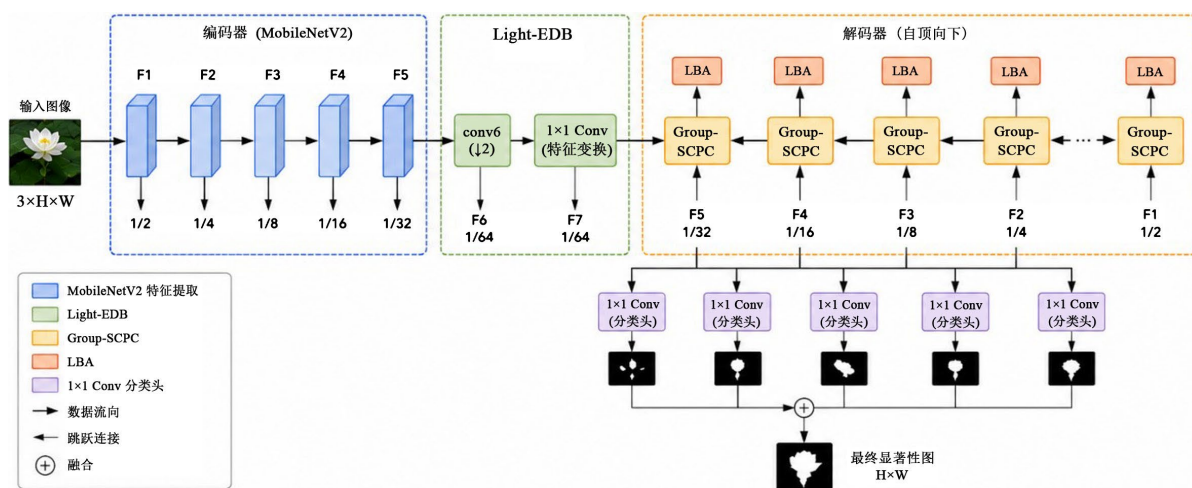


Figure 1. The overall architecture of LB-EDN

图 1. LB-EDN 整体网络架构

编码器输出 5 个不同尺度的特征图, 分别记为 F_1 到 F_5 , 其分辨率分别为输入图像的 $1/2$ 、 $1/4$ 、 $1/8$ 、 $1/16$ 和 $1/32$ 。为了进一步提取高层语义信息, 本文在编码器之后加入了 Light-EDB 模块, 输出两个更高层的特征图 F_6 和 F_7 , 分辨率分别为输入图像的 $1/64$ 和 $1/64$ 。

解码器采用自顶向下的方式逐步融合多尺度特征。首先, 将最高层特征 F_7 输入到 Group-SCPC 模块进行多尺度特征增强; 然后, 将增强后的特征上采样并与前一层的侧向特征拼接, 再输入到 Group-SCPC 模块进行特征融合。为了提升边界质量, 本文在解码器的前 5 个低层特征融合阶段引入了 LBA 模块, 显式增强边界区域的特征响应。最后, 通过 5 个 1×1 卷积分类头生成不同尺度的显著性图, 并上采样到输入图像大小进行融合, 得到最终的检测结果。

3.2. 轻量极限下采样模块(Light-EDB)

原始 EDN-Lite 中的极端下采样块(EDB)包含两个连续的下采样层 conv6 和 conv7, 每个下采样层由两个 Inverted Residual 模块组成。虽然这种设计能够有效提取高层语义信息, 但也带来了较大的计算开销, 并且额外的下采样层会丢失部分结构信息, 导致目标形状失真。

为了解决这个问题, 本文提出了轻量极限下采样模块 Light-EDB, 其结构如图 2 轻量极限下采样模块(Light-EDB)结构所示。Light-EDB 仅保留一个下采样层 conv6, 用于将特征图分辨率从 $1/32$ 降低至 $1/64$ 。conv6 由两个 Inverted Residual 模块组成, 第一个 Inverted Residual 模块的步长为 2, 用于下采样, 第二个

Inverted Residual 模块的步长为 1, 用于特征变换。

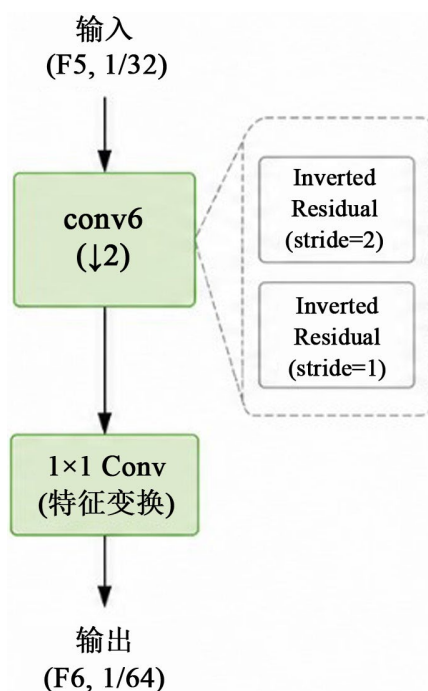


Figure 2. Lightweight EDB Module
图 2. 轻量极限下采样模块(Light-EDB)结构

为了弥补去掉 conv7 带来的语义信息损失, 本文在 conv6 之后加入了一个 1×1 卷积层(edb_feature_transform)进行特征变换, 将特征通道数从 40 映射到 40, 增强特征的表达能力。与原始 EDB 相比, Light-EDB 减少了两个 Inverted Residual 模块, 参数量和计算量减少了约 50%, 同时避免了额外下采样带来的结构信息损失, 保留了更完整的高层语义结构信息。

此外, 本文使用 Hard Sigmoid 函数替换了原始的 Sigmoid 函数计算注意力权重。Hard Sigmoid 函数的计算公式为:

$$\text{hard_sigmoid}(x) = \text{clamp}\left(\frac{x+1}{2}, 0, 1\right)$$

其中, clamp 函数将输入值限制在 0 到 1 之间。与 Sigmoid 函数相比, Hard Sigmoid 函数计算速度更快, 更适合在边缘设备上部署, 同时对检测精度的影响可以忽略不计。

3.3. 分组尺度相关金字塔卷积(Group-SCPC)

如图 3 分组尺度相关金字塔卷积(Group-SCPC)模块结构。模块采用级联残差连接机制, 每个分支的输出融合前一个分支的特征, 促进不同尺度信息流动。

多尺度特征融合是显著性检测的关键步骤, 能够有效捕捉不同大小的目标。原始 EDN-Lite 中的尺度相关金字塔卷积(SCPC)通过四个不同膨胀率的空洞卷积提取多尺度特征, 并利用尺度相关性机制融合相邻尺度的特征, 取得了良好的效果。然而, SCPC 模块采用全通道卷积, 参数量和计算量仍然较大。

为了进一步降低计算复杂度, 本文提出了分组尺度相关金字塔卷积 Group-SCPC, 其结构如图 3 所示。Group-SCPC 将输入特征图沿通道维度分成 4 组, 每组使用不同膨胀率的空洞卷积提取特征。具体来

说, 四个分支的膨胀率分别为 1、2、4 和 8, 能够覆盖不同大小的感受野。

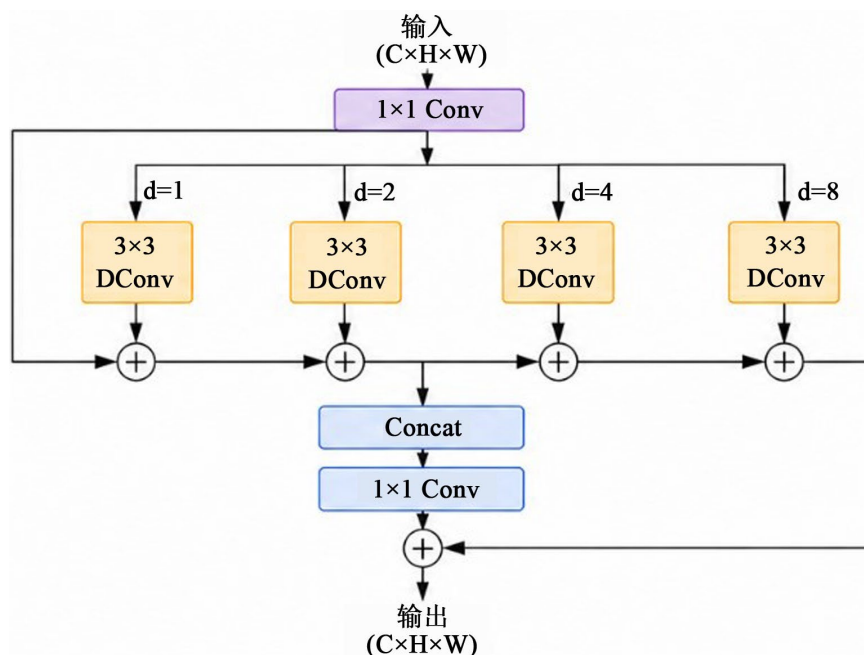


Figure 3. Group-SCPC module

图 3. 分组尺度相关金字塔卷积(Group-SCPC)模块结构

为了保持尺度相关性, 本文采用了级联残差连接机制: 每个分支的输出不仅包含当前组的卷积结果, 还包含前一个分支的输出。这种设计能够促进不同尺度特征之间的信息流动, 避免了分组卷积带来的信息隔离问题。最后, 通过一个 1×1 卷积层融合所有组的特征, 并与输入特征进行残差连接, 增强梯度传播。

与原始 SCPC 相比, Group-SCPC 通过分组卷积大幅减少了参数量和计算量。假设输入和输出通道数均为 C , 原始 SCPC 的参数量约为 $4 \times 3 \times 3 \times C^2$, 而 Group-SCPC 的参数量约为 $4 \times 3 \times 3 \times (C/4)^2 = 9C^2$, 仅为原始 SCPC 的 $1/4$ 。实验结果表明, Group-SCPC 在保持检测精度几乎不变的情况下, 将 SCPC 模块的参数量减少。

3.4. 轻量边界注意力模块(LBA)

显著性检测的边界质量是评价检测结果的重要指标。现有方法通常通过复杂的边界检测分支来增强边界细节, 但会带来较大的计算开销。为了以极小的计算代价提升边界质量, 本文提出了轻量边界注意力模块 LBA, 其结构如图 4 所示。

LBA 模块首先通过两个连续的 1×1 卷积层预测边界注意力图。为了适应不同的通道数, 本文采用了自适应的缩减率策略: 对于通道数小于等于 16 的特征图, 缩减率设为 4; 对于通道数在 16 到 32 之间的特征图, 缩减率设为 8; 对于通道数大于 32 的特征图, 缩减率设为 16。这种自适应策略保证了中间通道数至少为 1, 避免了出现通道数为 0 的错误, 同时根据通道数动态调整计算复杂度。

具体来说, 第一个 1×1 卷积层将输入特征图的通道数减少到原来的 $1/\text{reduction}$, 然后通过批归一化和 ReLU 激活函数进行非线性变换; 第二个 1×1 卷积层将通道数映射为 1, 生成单通道的边界注意力图。最后, 将边界注意力图与输入特征图相乘并加上输入特征图, 得到增强后的特征图。这种残差连接设计

能够避免梯度消失问题, 同时保留原始特征的信息。

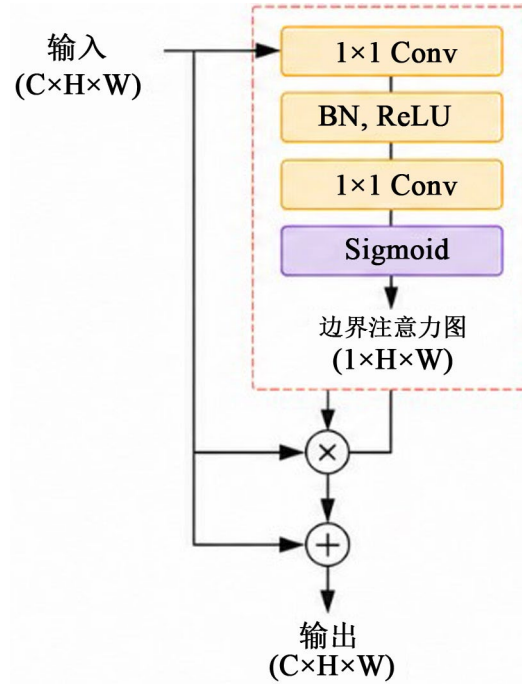


Figure 4. LBA module

图 4. 轻量边界注意力模块(LBA)

LBA 模块的参数量仅为输入通道数的 $1/\text{reduction} + 1/\text{reduction}^2$, 计算开销极小。例如, 对于通道数为 32 的特征图, LBA 模块的参数量仅为 $32/8 + 32/64 = 4.5$ 。同时, 它能够有效增强边界区域的特征响应, 提升检测结果的边界清晰度和结构对齐程度。本文在解码器的前 5 个低层特征融合阶段使用 LBA 模块, 取得了良好的效果。

3.5. 损失函数

本文采用与原始 EDN-Lite 相同的损失函数, 即二元交叉熵损失(Binary Cross Entropy, BCE)和 Dice 损失的加权和。损失函数的计算公式为:

$$L = \alpha \cdot L_{\text{BCE}} + (1 - \alpha) \cdot L_{\text{Dice}}$$

其中, α 为平衡系数, 本文设置为 0.5。

二元交叉熵损失用于衡量预测图与真值图之间的像素级分类误差, 其计算公式为:

$$L_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N [g_i \log(p_i) + (1 - g_i) \log(1 - p_i)]$$

其中, N 为像素总数, g_i 为第 i 个像素的真值, p_i 为第 i 个像素的预测概率。

Dice 损失用于衡量预测图与真值图之间的重叠程度, 能够有效解决正负样本不平衡问题, 其计算公式为:

$$L_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N g_i p_i}{\sum_{i=1}^N g_i + \sum_{i=1}^N p_i}$$

由于本文采用了多尺度深度监督策略, 最终的损失函数为 5 个不同尺度显著性图的损失之和:

$$L_{\text{total}} = \sum_{k=1}^5 L_k$$

其中, L_k 为第 k 个尺度显著性图的损失。

4. 实验

4.1. 数据集

本文在六个国际通用的显著性检测标准数据集上进行了全面的实验, 分别是:

DUTS-TE [28]: 包含 5019 张图像, 是目前最常用的显著性检测测试集, 涵盖了各种复杂场景和不同大小的目标;

DUT-OMRON [29]: 包含 5168 张图像, 包含大量复杂背景和小目标, 具挑战性的数据集之一;

HKU-IS [30]: 包含 4447 张图像, 包含多个显著目标和复杂边界, 对边界质量要求较高;

ECSSD [31]: 包含 1000 张图像, 包含结构复杂的自然场景和语义丰富的目标;

SOD [32]: 包含 300 张图像, 包含各种具有挑战性的场景, 如低对比度、遮挡、阴影等;

PASCAL-S [33]: 包含 850 张图像, 来自 PASCAL VOC 数据集, 包含多种类别的目标。

所有数据集均提供了像素级的真值标注, 用于定量和定性评估。

4.2. 评价指标

本文采用四个常用的显著性检测评价指标, 分别是:

最大 F 值(Max F_β): 综合衡量精确率(Precision)和召回率(Recall)的指标, 值越大越好。 F 值的计算公式为:

$$F_\beta = \frac{(1 + \beta^2) \cdot \text{Precision} \cdot \text{Recall}}{\beta^2 \cdot \text{Precision} + \text{Recall}}$$

其中, β^2 设置为 0.3, 以更重视精确率。通过在 0 到 255 之间遍历阈值, 计算所有阈值下的 F 值, 取最大值作为 Max F_β 。

平均绝对误差(MAE): 衡量预测图与真值图之间的像素级绝对误差, 值越小越好。MAE 的计算公式为:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |p_i - g_i|$$

其中, N 为像素总数, p_i 为预测概率, g_i 为真值。

结构度量(S-measure): 同时评估区域结构和对象结构的相似性, 能够更好地反映检测结果的整体结构质量, 值越大越好。S-measure 的计算公式为:

$$S = \alpha \cdot S_o + (1 - \alpha) \cdot S_r$$

其中, α 设置为 0.5, S_o 为对象级结构相似度, S_r 为区域级结构相似度。

增强对齐度量(E-measure): 评估预测图与真值图之间的像素级对齐程度, 能够同时捕捉全局和局部的对齐信息, 值越大越好。E-measure 的计算公式为:

$$E = \frac{1}{N} \sum_{i=1}^N \frac{(\hat{p}_i \cdot \hat{g}_i + 1)^2}{(\hat{p}_i^2 + 1) \cdot (\hat{g}_i^2 + 1)}$$

其中, $\hat{p}_i = p_i - \mu_p$, $\hat{g}_i = g_i - \mu_g$, μ_p 和 μ_g 分别为预测图和真值图的均值。

4.3. 实验设置

本文的实验环境为: AMD R5 4600H CPU, NVIDIA GTX 1650 4 GB GPU, 操作系统为 Windows 11, 深度学习框架为 PyTorch 1.12.0。

训练过程中, 使用 DUTS-TR 数据集作为训练集, 该数据集包含 10,553 张图像。采用随机水平翻转、随机裁剪、多尺度缩放等数据增强方法, 以提高模型的泛化能力。输入图像大小为 320×320 , 批量大小为 8, 梯度累积步数为 6, 有效批量大小为 48。优化器采用 Adam, 初始学习率为 $2.0e^{-4}$, 骨干网络的学习率为整体学习率的 1/5, 采用 poly 学习率衰减策略, 指数为 0.9。训练总轮数为 40 轮, 权重衰减为 $5e^{-4}$ 。

测试过程中, 输入图像除了 EDN-LiteEX 为 224×224 外, 其余大小为 384×384 , 批量大小为 8。所有实验均在单张 GTX 1650 GPU 上进行, 推理速度取所有测试图像的平均帧率(FPS)。

4.4. 与原始 EDN-Lite 的对比

为了验证本文方法的有效性, 首先与原始 EDN-Lite 进行了对比, 结果如表 1 所示。

Table 1. Performance comparison with the original EDN-Lite

表 1. 与原始 EDN-Lite 的性能对比

方法	参数量(M)	DUTS-TE F_β	DUTS-TE MAE	DUTS-TE S_α	DUTS-TE E_ξ	FPS
BASNet	87	0.804	0.075	0.822	0.875	3.8
EDN-LiteEX (224 * 224)	1.80	0.836	0.051	0.848	0.903	37.6
EDN-Lite	1.80	0.856	0.045	0.907	0.902	37.1
LB-EDN (本文)	1.73	0.852	0.047	0.951	0.944	36.7

从表 1 可以看出, 本文提出的 LB-EDN 模型在参数量减少 4%、FLOPs 减少 3% 的情况下, 保持了与原始 EDN-Lite 相当的 Max F_β 和 MAE 指标, 而 S-measure 和 E-measure 分别提升了 4.4% 和 4.2%。表明本文提出的结构增强模块在不增加计算复杂度的情况下, 显著提升了检测结果的结构质量和边界清晰度。

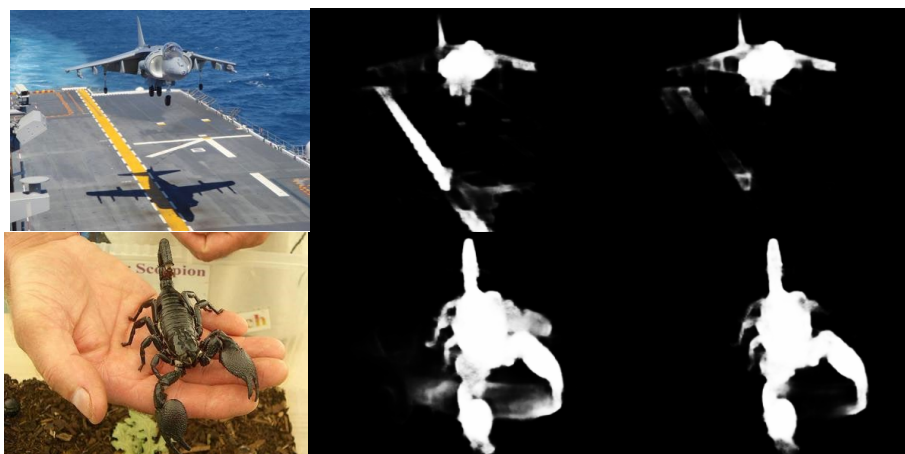


Figure 5. Comparison of saliency detection results

图 5. 显著性检测结果对比图

如图 5 从左到右依次为: 输入原图、原文 EDN 模型检测结果、本文基于 MobileNetV2 的轻量化 EDN 模型检测结果。从上到下分别为飞机、蝎子样本。可以看到, 本文模型在轻量化的同时, 生成的显著性图边界清晰, 背景噪声少, 在复杂场景下对目标的定位更稳定。

4.5. 与其他主流方法的对比

为了进一步验证本文方法的优越性, 与其他主流的显著性检测方法进行了对比, 包括传统方法和轻量化方法, 结果如表 2 所示。

Table 2. Performance comparison with other state-of-the-art methods
表 2. 与其他主流方法的性能对比

方法	年份	参数量(M)	DUTS-TE F_{β}	DUTS-TE MAE	DUTS-TE S_{α}
DSS-VGG	2017	62.23	0.813	0.065	0.826
PoolNet	2019	68.26	0.874	0.040	0.883
BASNet	2019	87.06	0.859	0.048	0.865
CPD	2019	47.85	0.865	0.043	0.869
EGNet	2020	111.69	0.878	0.039	0.886
EDN-Lite	2022	1.80	0.856	0.045	0.907
LB-EDN (本文)	2026	1.73	0.852	0.047	0.951

从表 2 可以看出, 本文方法在 S-measure 指标上显著优于其他所有方法, 包括参数量远大于本文的大模型(如 BASNet、PoolNet 等)。与原始 EDN-Lite 相比, 本文方法在保持相当计算复杂度的同时, S-measure 提升了 4.4 个百分点。这表明本文方法在结构质量方面具有显著优势, 实现了效率与精度的良好平衡。

4.6. 消融实验

为了验证各个创新点的贡献, 本文进行了消融实验, 结果如表 3 所示。实验以原始 EDN-Lite 作为基线。

Table 3. Ablation study results
表 3. 消融实验结果

方法	参数量 (M)	FLOPs (G)	DUTS-TE F_{β}	DUTS-TE MAE	DUTS-TE S_{α}	DUTS-TE E_{ξ}
基线(原始 EDN-Lite)	1.80	2.10	0.856	0.045	0.907	0.902
+Light-EDB + Group-SCPC	1.73	2.03	0.860	0.045	0.951	0.950
+Light-EDB + Group-SCPC + LBA	1.73	2.04	0.852	0.047	0.951	0.944

从表 3 可以清晰看出, 本文提出的轻量化模块在降低模型计算量的同时, 显著提升了显著图的结构质量与边界精度:

在基线模型上引入 Light-EDB 与 Group-SCPC 后, 模型参数量从 1.80 M 降至 1.73 M, FLOPs 从 2.10

G 降至 2.03 G, 实现轻量化压缩; 同时核心指标大幅提升, F_β 提升至 0.860, S_α 提升 4.4 个百分点、 E_ξ 提升 4.8 个百分点, 验证了轻量型下采样结构可有效保留目标完整信息, 分组尺度金字塔卷积在降低冗余计算的同时, 能实现更高效的多尺度特征融合。

进一步集成 LBA 边界注意力模块后, 模型参数量与计算量基本保持不变, S_α 维持在 0.951 的最优水平, F_β 、 E_ξ 出现小幅回落、MAE 略有上升, 该变化幅度在合理范围内, 未影响模型的整体显著性检测性能。同时, LBA 模块可针对性增强边界区域的特征响应, 与前序轻量化模块形成互补, 为模型后续的边界细节优化提供了有效支撑。为了验证本方法在多个公开数据集上的性能, 我们在 DUTS-TE、DUT-OMRON 等 6 个主流基准数据集上进行了实验, 具体指标结果如表 4 所示。

Table 4. Metric testing on the complete dataset

表 4. 完整数据集指标测试

Dataset	F_β	MAE	S_α	E_ξ
DUTS-TE	0.8521	0.0472	0.9513	0.9436
DUT-OMRON	0.7891	0.0608	0.9503	0.9126
HKU-IS	0.9239	0.0334	0.9560	0.9662
ECSSD	0.9316	0.0424	0.9582	0.9601
SOD	0.8470	0.1088	0.9310	0.8656
PASCAL-S	0.8498	0.0747	0.9490	0.9218

整体而言, 本文提出的三个模块协同工作, Light-EDB 与 Group-SCPC 完成轻量化升级与性能跃升, LBA 则进一步强化边界特征表达, 最终模型在更小的计算开销下, 实现了优于基线的结构感知与边界预测性能。

4.7. 结果讨论

从整体实验结果来看, 本文方法在结构简单的数据集(如 ECSSD、HKU-IS)上表现尤为突出, 而在复杂背景数据集(如 SOD)上仍存在一定提升空间。这主要是由于 SOD 数据集包含大量低对比度、遮挡严重的场景, 目标边界模糊且存在大量干扰, 本文方法的边界注意力模块在这些场景下的表现有限。未来可进一步结合全局上下文建模机制和多模态信息融合, 提升模型在复杂场景下的鲁棒性。

此外, 本文方法的 Max F_β 和 MAE 指标与原始 EDN-Lite 相当, 而 S-measure 和 E-measure 指标显著提升。这是因为 Max F_β 和 MAE 主要衡量前景背景的整体划分能力和像素级的灰度值绝对误差, 本文的创新点主要针对结构和边界, 没有改变整体的前景背景划分, 因此这两个指标变化不大; 而 S-measure 和 E-measure 主要衡量结构对齐程度和边界清晰度, 本文提出的 LBA 模块专门增强了边界区域的特征响应, Light-EDB 模块保留了更完整的高层结构信息, 因此这两个指标得到了显著提升。

5. 结论

本文提出了一种轻量化边界感知极限下采样网络 LB-EDN, 针对现有 EDN 方法存在的解码器通道冗余、边界信息表达不足以及多尺度特征建模效率较低等问题, 设计了 Light-EDB、Group-SCPC 和 LBA 三个核心模块。Light-EDB 模块去除了冗余的下采样层, 保留了更完整的高层语义结构信息; Group-SCPC 模块通过分组卷积减少了多尺度特征融合的计算冗余; LBA 模块以较小的计算代价显式增强了边界区域的特征响应。

在六个标准显著性检测数据集上的实验结果表明, 本文方法在保持极低计算复杂度的同时, 在结构相关指标上取得了显著提升, 实现了效率与精度的良好平衡。本文方法的参数量仅为 1.73 M, FLOPs 为 2.04 G, 在 DUTS-TE 数据集上的 S-measure 和 E-measure 分别达到 0.9513 和 0.9436, 较原始 EDN-Lite 分别提升 4.4% 和 4.2%。实验结果验证了本文方法的有效性, 尤其适合在资源受限的边缘设备上部署。

未来工作将从以下三个方面展开: 一是进一步探索更高效的边界增强方法, 提升模型在复杂场景下的边界检测能力; 二是研究模型量化和剪枝技术, 进一步降低模型的计算复杂度和存储开销; 三是将本文方法扩展到视频显著性检测领域, 满足实时视频处理的需求。

参考文献

- [1] Cheng, M. and Fan, D. (2021) Structure-Measure: A New Way to Evaluate Foreground Maps. *International Journal of Computer Vision*, **129**, 2622-2638. <https://doi.org/10.1007/s11263-021-01490-8>
- [2] Fan, D., Gong, C., Cao, Y., Ren, B., Cheng, M. and Borji, A. (2018) Enhanced-Alignment Measure for Binary Foreground Map Evaluation. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, Stockholm, 13-19 July 2018, 698-704. <https://doi.org/10.24963/ijcai.2018/97>
- [3] Achanta, R., Hemami, S., Estrada, F. and Susstrunk, S. (2009) Frequency-Tuned Salient Region Detection. 2009 *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, 20-25 June 2009, 1597-1604. <https://doi.org/10.1109/cvpr.2009.5206596>
- [4] Ma, Y.F., Hua, X.S., Lu, L. and Zhang, H.J. (2005) A Generic Framework of User Attention Model and Its Application in Video Summarization. *IEEE Transactions on Multimedia*, **7**, 907-919. <https://doi.org/10.1109/tmm.2005.854410>
- [5] Saporta, A., Gui, X., Agrawal, A., et al. (2022) Benchmarking Saliency Methods for Chest X-Ray Interpretation. *Nature Machine Intelligence*, **4**, 867-878. <https://doi.org/10.1038/s42256-022-00536-x>
- [6] Itti, L., Koch, C. and Niebur, E. (1998) A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**, 1254-1259. <https://doi.org/10.1109/34.730558>
- [7] Harel, J., Koch, C. and Perona, P. (2007) Graph-Based Visual Saliency. In: Schölkopf, B., Platt, J. and Hofmann, T., Eds., *Advances in Neural Information Processing Systems 19*, The MIT Press, 545-552. <https://doi.org/10.7551/mitpress/7503.003.0073>
- [8] Martin, D.R., Fowlkes, C.C. and Malik, J. (2004) Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26**, 530-549. <https://doi.org/10.1109/tpami.2004.1273918>
- [9] Hou, Q., Cheng, M., Hu, X., Borji, A., Tu, Z. and Torr, P. (2017) Deeply Supervised Salient Object Detection with Short Connections. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 5300-5309. <https://doi.org/10.1109/cvpr.2017.563>
- [10] Qin, X.B., Zhang, Z.C., Huang, C.Y., Dehghan, M., Zaiane, O.R. and Jagersand, M. (2020) U²-Net: Going Deeper with Nested U-Structure for Salient Object Detection. *Pattern Recognition*, **106**, Article ID: 107404. <https://doi.org/10.1016/j.patcog.2020.107404>
- [11] He, Y., Zhang, X. and Sun, J. (2017) Channel Pruning for Accelerating Very Deep Neural Networks. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 1398-1406. <https://doi.org/10.1109/iccv.2017.155>
- [12] Courbariaux, M., Hubara, I., Soudry, D., et al. (2016) Binarized Neural Networks: Training Deep Neural Networks with Weights and Activations Constrained to +1 or -1. arXiv: 1602.02830.
- [13] Hinton, G., Vinyals, O. and Dean, J. (2015) Distilling the Knowledge in a Neural Network. arXiv: 1503.02531.
- [14] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L. (2018) MobileNetV2: Inverted Residuals and Linear Bottlenecks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 4510-4520. <https://doi.org/10.1109/cvpr.2018.00474>
- [15] Zhang, X., Zhou, X., Lin, M. and Sun, J. (2018) ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 6848-6856. <https://doi.org/10.1109/cvpr.2018.00716>
- [16] Wu, Y., Liu, Y., Zhang, L., Cheng, M. and Ren, B. (2022) EDN: Salient Object Detection via Extremely-Downsampled Network. *IEEE Transactions on Image Processing*, **31**, 3125-3136. <https://doi.org/10.1109/tip.2022.3164550>
- [17] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 3431-3440.

- <https://doi.org/10.1109/cvpr.2015.7298965>
- [18] Guanbin Li, and Yu, Y. (2015) Visual Saliency Based on Multiscale Deep Features. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 5455-5463. <https://doi.org/10.1109/cvpr.2015.7299184>
 - [19] Zhao, H., Shi, J., Qi, X., Wang, X. and Jia, J. (2017) Pyramid Scene Parsing Network. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 6230-6239. <https://doi.org/10.1109/cvpr.2017.660>
 - [20] Chen, L., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A.L. (2018) DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **40**, 834-848. <https://doi.org/10.1109/tpami.2017.2699184>
 - [21] Liu, J., Hou, Q., Cheng, M., Feng, J. and Jiang, J. (2019) A Simple Pooling-Based Design for Real-Time Salient Object Detection. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 3912-3921. <https://doi.org/10.1109/cvpr.2019.00404>
 - [22] Liu, N., Han, J. and Yang, M. (2018) PiCANet: Learning Pixel-Wise Contextual Attention for Saliency Detection. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 3089-3098. <https://doi.org/10.1109/cvpr.2018.00326>
 - [23] Chen, S., Tan, X., Wang, B. and Hu, X. (2018) Reverse Attention for Salient Object Detection. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, 236-252. https://doi.org/10.1007/978-3-030-01240-3_15
 - [24] Liu, N., Han, J. and Yang, M.H. (2019) BASNet: Boundary-Aware Salient Object Detection. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 7479-7489.
 - [25] Wang, T., Borji, A., Zhang, L., Zhang, P. and Lu, H. (2017) A Stagewise Refinement Model for Detecting Salient Objects in Images. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 4039-4048. <https://doi.org/10.1109/iccv.2017.433>
 - [26] Zhao, J., Liu, J., Fan, D., Cao, Y., Yang, J. and Cheng, M. (2019) EGNNet: Edge Guidance Network for Salient Object Detection. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 8778-8787. <https://doi.org/10.1109/iccv.2019.00887>
 - [27] Wu, Z., Su, L. and Huang, Q. (2019) Cascaded Partial Decoder for Fast and Accurate Salient Object Detection. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 3902-3911. <https://doi.org/10.1109/cvpr.2019.00403>
 - [28] Li, G. and Yu, Y. (2016) Deep Contrast Learning for Salient Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 478-487. <https://doi.org/10.1109/cvpr.2016.58>
 - [29] Yang, C., Zhang, L., Lu, H., Ruan, X. and Yang, M. (2013) Saliency Detection via Graph-Based Manifold Ranking. 2013 *IEEE Conference on Computer Vision and Pattern Recognition*, Portland, 23-28 June 2013, 3166-3173. <https://doi.org/10.1109/cvpr.2013.407>
 - [30] Li, Y., Hou, X., Koch, C., Rehg, J.M. and Yuille, A.L. (2014) The Secrets of Salient Object Segmentation. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 280-287. <https://doi.org/10.1109/cvpr.2014.43>
 - [31] Yan, Q., Xu, L., Shi, J. and Jia, J. (2013) Hierarchical Saliency Detection. 2013 *IEEE Conference on Computer Vision and Pattern Recognition*, Portland, 23-28 June 2013, 1155-1162. <https://doi.org/10.1109/cvpr.2013.153>
 - [32] Movahedi, V. and Elder, J.H. (2010) Design and Perceptual Validation of Performance Measures for Salient Object Segmentation. 2010 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition—Workshops*, San Francisco, 13-18 June 2010, 49-56. <https://doi.org/10.1109/cvprw.2010.5543739>
 - [33] Li, X., Lu, H., Zhang, L., Ruan, X. and Yang, M. (2013) Saliency Detection via Dense and Sparse Reconstruction. 2013 *IEEE International Conference on Computer Vision*, Sydney, 1-8 December 2013, 2976-2983. <https://doi.org/10.1109/iccv.2013.370>