

## EGNet: 显著目标检测的边缘引导网络

赵嘉星, 刘姜江, 范登平, 曹洋, 杨巨峰, 程明明\*  
南开大学

<http://mmcheng.net/egnet/>

## Abstract

在目标检测任务中, 全卷积神经网络 (FCN) 展现出了优势。然而, 现有的基于 FCN 的方法仍然存在目标边缘粗糙的问题。为了解决这一问题, 本文着重研究了显著边缘信息与显著目标信息之间的互补性。在此基础上, 提出了一种基于边缘引导网络 (EGNet) 的显著性目标检测方法, 该方法采用三步法实现了两种互补信息在单个网络中的同步建模。在第一步中, 我们采用渐进融合的方法提取显著的目标特征。在第二步中, 我们将局部边缘信息与全局位置信息相结合, 得到显著的边缘特征。最后, 为了充分利用这些互补特征, 我们将相同的显著边缘特征与显著目标特征在不同分辨率下相结合。利用显著边缘特征中丰富的边缘信息和位置信息, 融合特征可以更准确地定位显著目标, 特别是其边缘。实验结果表明, 该方法在不进行任何预处理和后处理的情况下, 相比于最先进的方法在 6 个广泛使用的数据集上取得了更好的效果。代码已开源<http://mmcheng.net/egnet/>。

## 1. 引言

显著目标检测 (SOD) 旨在图像中找到最具视觉特色的目标。近年来, 它受到了广泛的关注, 并被广泛应用于许多与视觉和图像处理相关的领域, 如内容感知图像编辑 [6]、目标识别 [42]、图像合成 [4]、非真实感绘制技术 [41]、弱监督语义分割 [19] 和图像检索 [15]。此外, 还有很多关于视频显著目标检测 [12, 54] 和 RGB-D 显著目标检测 [11, 66] 的研究。

\*程明明 (cmm@nankai.edu.cn) 是通讯作者。本文是 ICCV 2019 论文 [67] 的中译版。

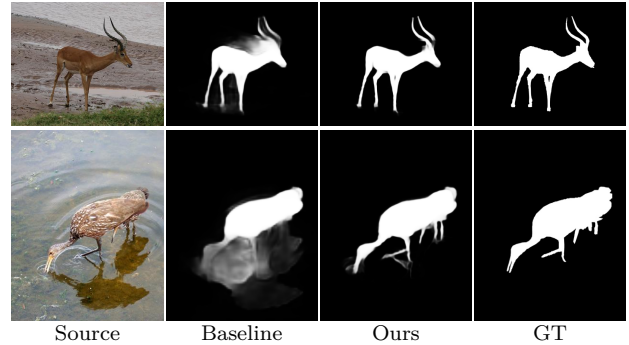


图 1. Visual examples of our method. After we model and fuse the salient edge information, the salient object boundaries become clearer. 本文方法的检测示例。在对显著边缘信息进行建模和融合后, 显著目标的边缘变得更加清晰。

受视觉注意力感知研究的启发 [7, 21, 39], 早期的研究主要基于颜色对比, 其在显著性检测中起着最重要的作用。这些方法主要受益于全局或局部对比线索及其学习到的融合权值。不幸的是, 虽然这些人工选取特征有时可以定位最突出的物体, 但当前景和背景之间的对比度不充分时, 分割方法是不可取和不可靠的, 因而产生了不规则形状的显著图。

近年来, 卷积神经网络 (CNN) [25] 成功地突破了传统人工选取特征的局限, 特别是在全卷积神经网络 (FCN) [34] 出现之后。这些基于 CNN 的方法已经极大地刷新了排行榜上几乎所有广泛使用的基准, 并因其效率和高性能表现, 而逐渐取代传统的显著目标检测方法。在基于 CNN 架构的 SOD 方法中, 大多数以图像块 [64, 65] 为输入, 利用多尺度或多上下文信息来获得最终的显著性图。自完全卷积网络被提出用于像素标记问题以来, 已有几种端到端深度架构 [17, 18, 23, 28, 31, 50, 60, 68] 的显著目标检测出现。输出的显著图的基本单位为图像区域的每个像素。一方面,

结果突出了细节,因为每个像素都有它的显著值。但另一方面,它忽略了 SOD 中重要的结构信息。

随着网络感受野的增大,显著目标的定位越来越精确。但同时也忽略了空间连贯性。最近,为了获得精细的边缘细节,有些基于 SOD U-Net [40] 的方法 [32, 33, 59, 61] 采用了一种双向或递归的方式,利用局部信息来提炼高层特征。然而,显著目标的边缘仍然没有被明确建模。显著边缘信息与显著目标信息之间的互补性尚未被发掘。此外,还有一些使用预处理 (Superpixel) [20] 或后处理 (CRF) [17, 28, 33] 来保留目标边缘的方法。这些方法的主要缺点是推理速度慢。

本文主要研究显著边缘信息与显著目标信息之间的互补性。我们的目标是利用显著的边缘特征来帮助显著的目标特征更准确地定位目标,特别是它们的边缘。综上所述,本文的主要贡献有三点:

- 我们提出了 EGNNet 来明确地建模互补的显著目标信息和网络中的显著边缘信息,以保留显著目标边缘。同时,显著的边缘特征也有助于定位。
- 我们的模型联合优化了这两个互补的任务,使它们相互促进,显著改善了预测的显著图。
- 在 6 个广泛使用的数据集上,我们将提出的方法与 15 个最先进的方法进行了比较。无需任何特殊的训练技巧,我们的方法在三个评估指标下实现了最佳性能。

## 2. 相关工作

近年来,人们提出了一些方法来检测图像中的显著目标。早期的方法是利用人工选取特征以自底向上的模式预测显著性图,如对比度 [5]、边界背景 [57, 69]、中心先验 [24, 44], 以及 [22, 44, 51] 等。更多细节在 [1, 2, 9] 中有介绍。

近年来,卷积神经网络 (CNN) 发挥出了巨大优势,刷新了计算机视觉许多领域的记录。

Li 等人 [27] 将图像区域大小调整为三个不同的尺度来提取多尺度特征,然后对这多个显著性图进行聚合,得到最终的预测图。Wang 等人 [45] 设计了一种神经网络来提取输入小块的局部估计,并将这些特征与全局对比度和几何信息相结合来描述图像小块。然而,检测结果受限于这些方法在图像块上的性能。在 [34] 中, long 等人首先提出了一个网络 (FCN) 来预测每个

像素的语义标签。受 FCN 的启发,人们提出了越来越多的逐像素显著检测方法。Wang 等人 [47] 提出了一种用于显著目标检测的递归 FCN 架构。Hou 等人 HED [55] 的基础上提出了一个短连接 [17, 18] 来集成底层特征和高层特征,从而解决尺度空间问题。在 [62] 中, Zhang 等人引入了一种重新定义的 dropout 和一种有效的混合上采样来学习深度不确定卷积特征,以促进模型鲁棒性和准确性。在 [61] 中, Zhang 等人将多层特征直接聚合为多个分辨率,然后通过双向聚合的方法将这些特征图进行组合。Zhang 等人 [59] 提出了一种双向信息传递模型,该模型集成多级特征,用于显著目标检测。Wang 等人 [53] 利用固定图帮助模型更准确地定位显著目标。在 [35] 中, Luo 等人提出了一种基于 U-Net 的体系结构,其包含 IoU 边缘损失,利用边缘线索检测显著目标。在其他与显著性相关的任务中,出现了一些使用边缘线索的方法。在 [26] 中, li 等人通过生成对象的轮廓来获得显著的实例分割结果。而 [29] 则利用训练好的轮廓检测模型生成显著性掩模,以克服人工标注带来的限制。

与大多数基于 U-Net 的 SOD 方法 [32, 33, 59, 61] 相比,我们显式地对网络中的边缘信息进行建模,以利用边缘线索。与使用边缘线索的方法 [14, 58, 70] [14, 58, 70] 相比,主要的区别在于我们使用单一的基本网络,并且共同优化显著边缘检测和显著目标检测,使两者相互帮助,由此带来更好的性能。与 NLD [35] 相比,他们实现了一个受 Mumford-Shah 函数启发的损失函数 [38] 来惩罚边缘上的误差。由于显著边缘是通过固定的 sobel 算子从显著目标提取出来的,因此这种惩罚实质上只影响特征图上显著边缘邻域的梯度。这种方法在一定程度上优化了边缘细节,但没有充分利用显著边缘检测与显著目标检测之间的互补性。在我们的方法中,我们设计了两个模块来独立提取这两类特征。然后通过一对一的引导模块对这些互补的特征进行融合。这样,显著边缘信息不仅可以提高边缘的质量,而且可以使定位更加准确。实验部分验证了我们的观点。

## 3. 显著边缘引导网络

总体框架如 Fig. 2 所示。在本节中,我们首先描述 Sec. 3.1 中的动机,然后在 Sec. 3.2 介绍所采用的显著目标特征提取模块以及所提出的 non-local 显著边缘特征提取模块。最后 Sec. 3.3 介绍一对一引导模块。

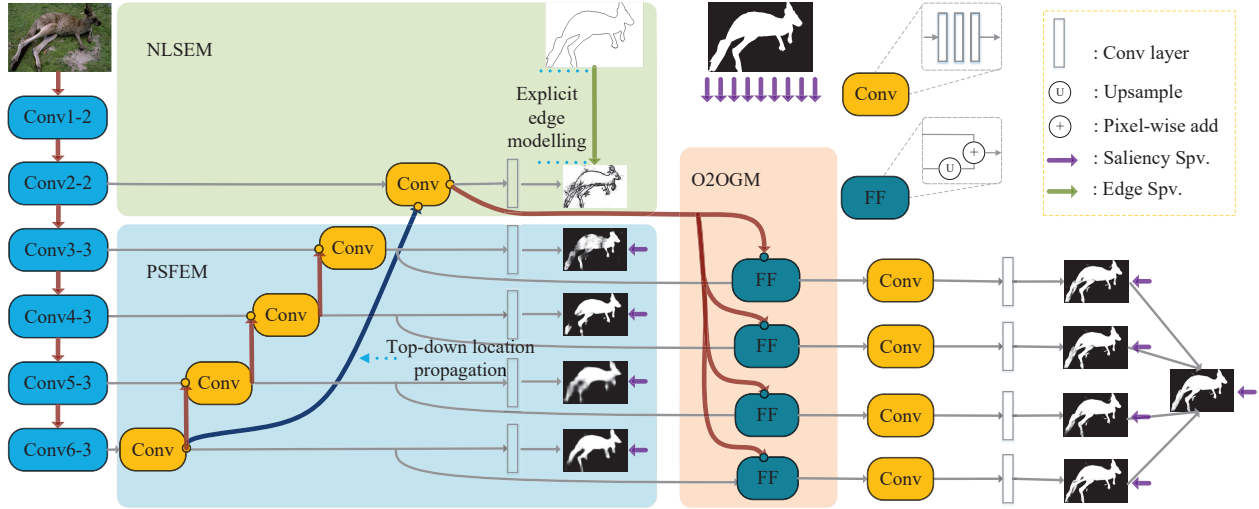


图 2. T 本文方法整体框架。我们用棕色粗线表示尺度之间的信息流。PSFEM：渐进显著目标特征提取模块。O2OGM：一对一引导模块。FF：特征融合。Svp.：监督。

### 3.1. 动机

与基于区域的方法相比，像素级显著目标检测方法显示出了它们的优势。然而，他们忽略了图像的空间相关性，导致显著目标边缘不令人满意。大多数方法 [17, 18, 31, 33, 59, 61] 希望通过融合多尺度信息来解决这一问题。而另一些方法 [17, 28, 33] 则使用 CRF 等后处理来细化显著目标边缘。在 NLDF [35] 中，他们提出了 IOU 损失来影响边缘周围位置的梯度。它们都没有注意到显著边缘检测与显著目标检测之间的互补关系。良好的显著边缘检测结果有助于显著目标检测任务的分割和定位，反之亦然。在此基础上，我们提出了 EGNNet 以端到端方式对单个网络中互补的显著边缘信息和显著目标信息进行建模和融合。

### 3.2. 互补信息建模

我们提出的网络独立于主干网络。这里我们使用其他基于深度学习的方法 [17, 35] 所建议的 VGG 网络来描述所提出的方法。首先，我们删去最后三个全连接层。遵循 DSS [17, 18]，我们将另一条侧支路连接到 VGG 中的最后一个池化层。由此，从主干网络中，我们得到六个侧输出特征 Conv1-2, Conv2-2, Conv3-3, Conv4-3, Conv5-3, Conv6-3。因为 Conv1-2 离输入太近，接受域太小，所以我们丢弃了该条侧支路  $S^{(1)}$ 。于是还剩下 5 条侧支路  $S^{(2)}, S^{(3)}, S^{(4)}, S^{(5)}, S^{(6)}$ 。为简便起见，这五个特性可以用主干网络特性集合  $C$  表示：

$$C = \{C^{(2)}, C^{(3)}, C^{(4)}, C^{(5)}, C^{(6)}\}, \quad (1)$$

其中  $C^{(2)}$  表示 Conv2-2 features，其余类似。Conv2-2 保留了更好的边缘信息 [61]。因此，我们利用  $S^{(2)}$  提取边缘特征，利用其他边路径提取显著目标特征。

#### 3.2.1 渐进显著目标特征提取

如 Fig. 2 的 PSFEM 所示，为了获得更丰富的上下文特征，我们利用了广泛使用的架构 U-Net [40] 来生成多分辨率特性。与原始的 U-Net 不同，为了获得更鲁棒的显著目标特征，我们增加了三个卷积层 (Fig. 2 中的 Conv)，并在每个卷积层后添加 ReLU 层以保证非线性。为了简单说明，我们使用  $T$  (Tab. 1) 来表示这些卷积层和 ReLU 层。此外，对每条侧支进行深度监督。我们采用单个卷积层将特征图转换为单通道预测掩膜，并用  $D$  (Tab. 1) 表示。卷积层的详细信息见 Tab. 1。

#### 3.2.2 Non-local 显著边缘特征提取

在这个模块中，我们的目标是建立显著边缘信息的模型，并提取显著边缘特征。正如上面提到的，Conv2-2 保存了更好的边缘信息。因此我们从 Conv2-2 中提取局部边缘信息。然而，要获得显著的边缘特征，仅靠局部信息是不够的。还需要高级语义信息或位置信息。当信息像 U-Net 体系结构那样从顶层逐渐返回到底层时，高层位置信息就会逐渐被稀释。此外，顶层的感受野最大，位置最准确。因此，我们设计了一个自顶向下的位置传播，将高级位置信息传播到侧支路  $S^{(2)}$ ，以抑制



S	T <sub>1</sub>				T <sub>2</sub>				T <sub>3</sub>				D		
2	3	1	128		3	1	128		3	1	128		3	1	1
3	3	1	256		3	1	256		3	1	256		3	1	1
4	5	2	512		5	2	512		5	2	512		3	1	1
5	5	2	512		5	2	512		5	2	512		3	1	1
6	7	3	512		7	3	512		7	3	512		3	1	1

表 1. 每个侧支输出的详细信息。T 为特征增强模块 (Fig. 2 中的 Conv)，每个 T 包含三个卷积层:  $T_1, T_2, T_3$  及紧随其后的 ReLU 层。我们给出了每个卷积层的卷积核大小、填充和通道数。例如，3, 1, 128 表示一个卷积核大小为 3，填充为 1，通道数为 128 卷积层。D 为转换多通道特征图为单通道激活图的转换层。S 为侧支路径。

非显著边缘。融合特征  $\bar{C}^{(2)}$  可以表示为:

$$\bar{C}^{(2)} = C^{(2)} + \text{Up}(\phi(\text{Trans}(\hat{F}^{(6)}; \theta)); C^{(2)}), \quad (2)$$

其中  $\text{Trans}(*; \theta)$  是带有参数  $\theta$  的卷积层，目的是改变特征的通道数，而  $\phi()$  表示一个 ReLU 激活函数。 $\text{Up}(*; C^{(2)})$  为双线性插值运算，其目标为上采样 \* 到与  $C^{(2)}$  相同的大小。在上式的右边，第二项是来自较高边侧支路径的特征。为了解释清楚，我们使用  $\text{UpT}(\hat{F}^{(i)}; \theta, C^{(i)})$  来表示  $\text{Up}(\phi(\text{Trans}(\hat{F}^{(i)}; \theta)); C^{(i)})$ 。 $\hat{F}^{(6)}$  表示侧支路径  $S^{(6)}$  的增强的特征。增强特性  $\hat{F}^{(6)}$  可以表示为  $f(C^{(6)}; W_T^{(6)})$ ，且  $S^{(3)} S^{(4)} S^{(5)}$  中的增强特征可计算为:  $\text{Up}(\phi(\text{Trans}(\hat{F}^{(i)}; \theta)); C^{(i)})$ 。 $\hat{F}^{(6)}$  denotes the enhanced features in side path  $S^{(6)}$ .

$$\hat{F}^{(i)} = f(C^{(i)} + \text{UpT}(\hat{F}^{(i+1)}; \theta, C^{(i)}); W_T^{(i)}), \quad (3)$$

其中  $W_T^{(i)}$  表示  $T^{(i)}$  中的参数， $f(*; W_T^{(i)})$  表示用参数  $W_T^{(i)}$  进行一系列卷积和非线性运算。

在获得引导特征  $\bar{C}^{(2)}$  之后，其他侧支路径同理，我们添加一系列卷积层来增强引导特征，则  $S^{(2)}$  中的最终显著边缘特征  $F_E$  可计算为  $f(\bar{C}^{(2)}; W_T^{(2)})$ 。配置细节可以在 Tab. 1 中找到。为了显式地对显著边缘特征进行建模，我们增加了一个额外的显著边缘监督来监督显著边缘特征。我们使用交叉熵损失，可定义为:

$$\begin{aligned} \mathcal{L}^{(2)}(F_E; W_D^{(2)}) = & - \sum_{j \in Z_+} \log \Pr(y_j = 1 | F_E; W_D^{(2)}) \\ & - \sum_{j \in Z_-} \log \Pr(y_j = 0 | F_E; W_D^{(2)}), \end{aligned} \quad (4)$$

其中  $Z_+$  和  $Z_-$  分别表示显著边缘像素集和背景像素集。 $W_D$  表示转换层的参数，如 Tab. 1 所示。 $\Pr(y_j = 1 | F_E; W_D^{(2)})$  为预测图，其中每个值代表该像素的显著

边缘置信度。另外，在显著目标检测侧支路径上增加的监督可表示为:

$$\begin{aligned} \mathcal{L}^{(i)}(\hat{F}^{(i)}; W_D^{(i)}) = & - \sum_{j \in Y_+} \log \Pr(y_j = 1 | \hat{F}^{(i)}; W_D^{(i)}) \\ & - \sum_{j \in Y_-} \log \Pr(y_j = 0 | \hat{F}^{(i)}; W_D^{(i)}), \quad i \in [3, 6], \end{aligned} \quad (5)$$

其中  $Y_+$  和  $Y_-$  分别表示显著区域像素集和非显著像素集。因此互补信息模型中的总损失  $\mathbb{L}$  可表示为:

$$\mathbb{L} = \mathcal{L}^{(2)}(F_E; W_D^{(2)}) + \sum_{i=3}^6 \mathcal{L}^{(i)}(\hat{F}^{(i)}; W_D^{(i)}). \quad (6)$$

### 3.3. 一对一引导模块

在获取互补的显著边缘特征和显著目标特征后，我们即可利用显著边缘特征引导显著目标特征更好地进行分割和定位。简单的方法是融合  $F_E$  和  $\hat{F}^{(3)}$ 。充分利用多分辨率显著目标特征将会更好。但是，从下到上逐次融合显著边缘特征和多分辨率显著目标特征的缺点是，在融合显著目标特征的同时，显著边缘特征会被稀释。此外，我们的目标是融合显著目标特征和显著的边缘特征，利用互补的信息来获得更好的预测结果。因此，我们提出了一对一的引导模块。此外，实验部分验证了我们的观点。

具体来说，我们为  $S^{(3)}, S^{(4)}, S^{(5)}, S^{(6)}$  添加了子侧支路径。在每条子侧支路径中，通过将显著边缘特征融合到增强的显著目标特征中，提高了高阶预测的定位精度，更重要的是提高了分割细节。显著边缘引导特征 (s 特征) 可表示为:

$$G^{(i)} = \text{UpT}(\hat{F}^{(i)}; \theta, F_E) + F_E, i \in [3, 6]. \quad (7)$$

然后与 PSFEM 相似，我们在每条子侧支路径上采用一系列卷积层  $T$  进一步增强 s 特征，并采用转换层  $D$  将多通道特征图转换为单通道预测图。为了更清楚地说明这一点，在本模块中，我们将  $T$  和  $D$  分别表示为  $T'$  和  $D'$ 。通过 Eq. (3)，我们得到增强的 s-features  $\hat{G}^{(i)}$ 。这里我们还为这些增强的 s 特征添加了深度监督。对于每条子侧支输出预测图，损失可计算为:

$$\begin{aligned} \mathcal{L}^{(i)'}(\hat{G}^{(i)}; W_{D'}^{(i)}) = & - \sum_{j \in Y_+} \log \Pr(y_j = 1 | \hat{G}^{(i)}; W_{D'}^{(i)}) \\ & - \sum_{j \in Y_-} \log \Pr(y_j = 0 | \hat{G}^{(i)}; W_{D'}^{(i)}), \quad i \in [3, 6]. \end{aligned} \quad (8)$$

然后对多尺度细化的预测图进行融合，得到融合图。融合图的损失函数可表示为：

$$\mathcal{L}'_f(\hat{G}; W_{D'}) = \sigma(Y, \sum_{i=3}^6 \beta_i f(\hat{G}^{(i)}; W_{D'}^{(i)})), \quad (9)$$

其中  $\sigma(* *)$  表示预测图与显著真实图之间的交叉熵损失，其形式与 Eq. (5) 相同。因此，这部分的损失和所提出的网络的总损失可以表示为：

$$\mathbb{L}' = \mathcal{L}'_f(\hat{G}; W_{D'}) + \sum_{i=3}^6 \mathcal{L}^{(i)'}(\hat{G}^{(i)}; W_{D'}^{(i)}) \quad (10)$$

$$\mathbb{L}_t = \mathbb{L} + \mathbb{L}'.$$

## 4. 实验

### 4.1. 实现细节

我们遵循 [33, 49, 59, 63] 在 DUTS [46] 数据集上训练我们的模型。为公平对比，我们分别使用 VGG [43] 与 ResNet [16] 作为主干网络。为公平地与其它方法 [17, 28] 进行对比，我们都在 MSRA-B 数据集上训练模型。我们的模型使用 PyTorch 实现。所有新添加的卷积层的权值都用截断的正态分布随机初始化 ( $\sigma = 0.01$ )，偏差初始化为 0。超参数设置如下：学习率 =  $5e-5$ ，权重衰减 = 0.0005，动量 = 0.9，每条侧支输出的损失权重统一为 1。对每一幅图像进行反向传播处理。在训练期间不使用验证数据集。我们训练模型 24 个 epoch，15 个 epoch 后将学习率除以 10。在推理过程中，我们可以得到一个预测的显著边缘图和一组显著图。在我们的方法中，我们直接使用融合的预测图作为最终的显著图。

### 4.2. 数据集与评估指标

我们在六个广泛使用的公共基准数据集上评估了所提出的框架：ECSSD [56]，PASCAL-S [30]，DUT-OMRON [57]，SOD [36, 44]，HKU-IS [27]，DUTS [46]。ECSSD [56] 包含 1000 个有意义的语义图像和各种复杂的场景。PASCAL-S [30] 包含从 PASCAL VOC 分割数据集 [8] 的验证集中选择的 850 幅图像。DUT-OMRON [57] 包含 5168 张高质量但具有挑战性的图像。该数据集中的图像包含一个或多个具有相对复杂背景的显著目标。SOD [36] 包含 300 幅用于图像分割的图像。这是目前最具挑战性的数据集之一。HKU-IS [27] 包含 4447 幅带有高质量标注的图像，其中许多图像有

多个不相关的显著对象。该数据集分为 2500 张训练图像、500 张验证图像和 2000 张测试图像。DUTS [46] 是最大的显著目标检测基准。它包含 10553 张用于训练的图像和 5019 张用于测试的图像。大多数图像具有不同的位置和比例。我们遵循最新的工作 [33, 49, 52]，使用 DUTS 数据集来训练所提出的模型。

我们使用三个广泛使用的标准度量，F 度量，平均绝对误差 (MAE) [2]，和最近提出的基于结构的度量，即 S 度量 [10]，来评估我们的模型和其他最先进的模型。F 度量是平均精度和平均召回率的调和平均值，表达式为：

$$F_\beta = \frac{(1 + \beta^2) Precision \times Recall}{\beta^2 \times Precision + Recall}, \quad (11)$$

我们设置  $\beta^2 = 0.3$  来更多地突出精度而不是召回率，正如 [5] 中所建议的那样。精度表示在预测的显著图中检测到的显著像素的比例。召回率是指在真实图中检测到的显著像素的比例。精度和召回率是在二值图像上计算的。因此，我们应该先对预测图使用阈值二值化。不同的阈值会导致不同的精度和召回率。我们可以绘制不同阈值下的精度-召回率曲线。这里我们使用 [17, 18] 所提供的代码来评估。遵循最新的显著目标检测方法 [17, 18, 32, 59]，我们报告了所有精度-召回率对的最大 F 度量。

MAE 是评估预测图与真实图之间平均差异的一种度量。设  $P$  和  $Y$  表示显著图和归一化为  $[0, 1]$  的真是图。我们通过下式计算 MAE 分数：

$$\varepsilon = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H |P(x, y) - Y(x, y)|, \quad (12)$$

其中  $W$  和  $H$  分别为图像的宽度和高度。

S 度量侧重于评价显著图的结构信息，它比 F 度量更接近于人类视觉系统。因此，我们加入了 S 度量来进行更全面的评价。S 度量由下式计算：

$$S = \gamma S_o + (1 - \gamma) S_r, \quad (13)$$

其中  $S_o$  和  $S_r$  表示区域感知和目标感知的结构相似性， $\gamma$  默认设置为 0.5。更多细节可以在 [10] 中找到。

### 4.3. 消融实验与分析

在本节中，以 DUTS-TR [46] 作为训练集，我们探讨了在相对困难的数据集 SOD [36] 和最近提出的大型数据集 DUTS-TE [46] 上所提出的网络中不同组件的影响。

	ECSSD [56]			PASCAL-S [30]			DUT-O [57]			HKU-IS [27]			SOD [36,37]			DUTS-TE [46]		
	MaxF ↑	MAE ↓	S ↑	MaxF ↑	MAE ↓	S ↑	MaxF ↑	MAE ↓	S ↑	MaxF ↑	MAE ↓	S ↑	MaxF ↑	MAE ↓	S ↑	MaxF ↑	MAE ↓	S ↑
VGG-based																		
DCL* [28]	0.896	0.080	0.863	0.805	0.115	0.791	0.733	0.094	0.743	0.893	0.063	0.859	0.831	0.131	0.748	0.786	0.081	0.785
DSS* [17,18]	0.906	0.064	0.882	0.821	0.101	0.796	0.760	0.074	0.765	0.900	0.050	0.878	0.834	0.125	0.744	0.813	0.065	0.812
MSR [26]	0.903	0.059	0.875	0.839	0.083	0.802	0.790	0.073	0.767	0.907	0.043	0.852	0.841	0.111	0.757	0.824	0.062	0.809
NLDF [35]	0.903	0.065	0.875	0.822	0.098	0.803	0.753	0.079	0.750	0.902	0.048	0.878	0.837	0.123	0.756	0.816	0.065	0.805
RAS [3]	0.915	0.060	0.886	0.830	0.102	0.798	0.784	0.063	0.792	0.910	0.047	0.884	0.844	0.130	0.760	0.800	0.060	0.827
ELD* [13]	0.865	0.082	0.839	0.772	0.122	0.757	0.738	0.093	0.743	0.843	0.072	0.823	0.762	0.154	0.705	0.747	0.092	0.749
DHS [32]	0.905	0.062	0.884	0.825	0.092	0.807	-	-	-	0.892	0.052	0.869	0.823	0.128	0.750	0.815	0.065	0.809
RFCN* [48]	0.898	0.097	0.852	0.827	0.118	0.799	0.747	0.094	0.752	0.895	0.079	0.860	0.805	0.161	0.730	0.786	0.090	0.784
UCF [62]	0.908	0.080	0.884	0.820	0.127	0.806	0.735	0.131	0.748	0.888	0.073	0.874	0.798	0.164	0.762	0.771	0.116	0.777
Amulet [61]	0.911	0.062	0.894	0.826	0.092	0.820	0.737	0.083	0.771	0.889	0.052	0.886	0.799	0.146	0.753	0.773	0.075	0.796
C2S [29]	0.909	0.057	0.891	0.845	0.081	0.839	0.759	0.072	0.783	0.897	0.047	0.886	0.821	0.122	0.763	0.811	0.062	0.822
PAGR [63]	0.924	0.064	0.889	0.847	0.089	0.818	0.771	0.071	0.751	0.919	0.047	0.889	0.841	0.146	0.716	0.854	0.055	0.825
Ours	0.941	0.044	0.913	0.863	0.076	0.848	0.826	0.056	0.813	0.929	0.034	0.910	0.869	0.110	0.788	0.880	0.043	0.866
ResNet-based																		
SRM* [49]	0.916	0.056	0.895	0.838	0.084	0.832	0.769	0.069	0.777	0.906	0.046	0.887	0.840	0.126	0.742	0.826	0.058	0.824
DGRL [52]	0.921	0.043	0.906	0.844	0.075	0.839	0.774	0.062	0.791	0.910	0.036	0.896	0.843	0.103	0.774	0.828	0.049	0.836
PiCANet* [33]	0.932	0.048	0.914	0.864	0.077	0.850	0.820	0.064	0.808	0.920	0.044	0.905	0.861	0.103	0.790	0.863	0.050	0.850
Ours	0.943	0.041	0.918	0.869	0.074	0.852	0.842	0.052	0.818	0.937	0.031	0.918	0.890	0.097	0.807	0.893	0.039	0.875

表 2. 对 6 个广泛使用的数据集进行的定量比较，包括最大 F 度量、MAE 和 S 度量。‘-’ 表示在该数据集上训练相应的方法。↑ & ↓ 分别表示越大越好和越小越好。\* 表示使用预处理或后处理的方法。最好的三个结果分别标记在红色、蓝色和绿色中。我们的方法在三个评估指标下，在这六个广泛使用的数据集上达到最先进的水平。

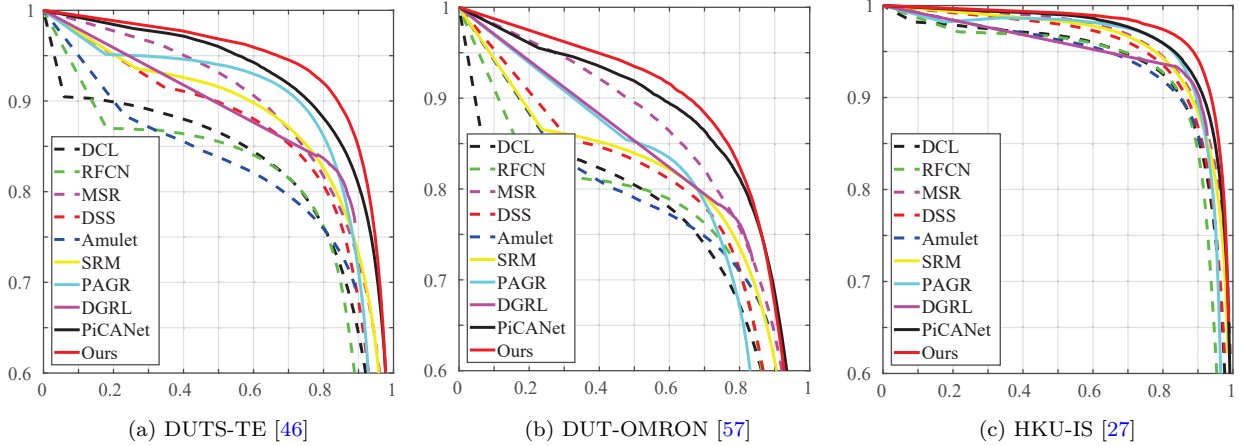


图 3. 在三个流行的显著目标数据集上的精度 (纵轴) 召回率 (横轴) 曲线。可以看出，所提出的方法在目前最先进的技术之上。

#### 4.3.1 互补信息建模

在本小节中，我们探讨了显著边缘信息的作用，这也是我们的基本思想。基线是 U-Net 架构，它以 PSFEM(Fig. 2) 的方式集成了多尺度特征 (从 Conv2-2 到 Conv6-3)。我们去除基线中的侧支路径  $S^{(2)}$ ，然后将最终显著性特征  $\hat{F}^{(3)}$  (Conv3-3 的侧支路径) 与局部 Conv2-2 特征融合，得到显著边缘特征。最后，我们整

合显著边缘特征和显著目标特征  $\hat{F}^{(3)}$ ，得到预测掩模。我们将这种使用边缘的技术记为 edge\_PROG。结果显示在 Tab. 3 的第二行。结果表明，显著边缘信息在显著目标检测任务中是非常有用的。

Model	SOD			DUTS			SOD			DUTS		
	MaxF $\uparrow$	MAE $\downarrow$	S $\uparrow$	MaxF $\uparrow$	MAE $\downarrow$	S $\uparrow$	Recall $\uparrow$	Precision $\uparrow$	MaxF $\uparrow$	Recall $\uparrow$	Precision $\uparrow$	MaxF $\uparrow$
1. B	.851	.116	.780	.855	.060	.851	0.513	0.541	0.527	0.318	0.659	0.429
2. B + edge_PROG	.873	.105	.799	.872	.051	.851	0.637	0.534	0.581	0.446	0.680	0.539
3. B + edge_TDLP	.882	.100	.807	.879	.044	.866						
4. B + edge_NLDF	.857	.112	.794	.866	.053	.860						
5. B + edge_TDLP + MRF_PROG	.882	.106	.796	.880	.046	.869						
6. B + edge_TDLP + MRF_OTO	.890	.097	.807	.893	.039	.875						

表 3. Ablation analyses on SOD [36] and DUTS-TE [46]. Here, B denotes the baseline model. edge\_PROG, edge\_TDLP, edge\_NLDF, MRF\_PROG, MRF\_OTO are introduced in the Sec. 4.3.

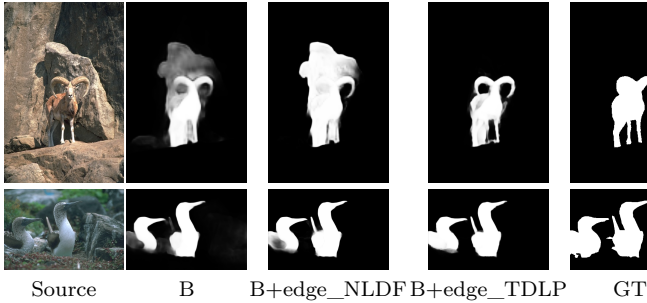


图 4. 添加边缘线索之前和之后的视觉例子。B 表示基线模型。edge\_NLDF 和 edge\_TDLP 表示 NLDF 中使用的边缘惩罚以及本文提出的边缘建模方法。详细内容在 Sec. 4.3 中介绍。

#### 4.3.2 自顶向下的位置传播

在本小节中，我们将探讨自顶向下的位置传播的作用。与前面 Sec. 4.3.1 中提到的 edge\_PROG 相比，我们利用自顶向下的位置传播从顶层提取更精确的位置信息，而不是从侧支路径  $S^{(3)}$ 。我们称这种使边缘的策略为 edge\_TDLP。通过比较 Tab. 3 的第二行和第三行，可以证明自顶向下的位置传播的有效性。此外，对比 Tab. 3 的第一行和第三行，我们可以发现，通过对这两种互补信息在网络内的显式建模，在数据集上的性能有了很大的提高 (F 度量提升 3.1%, 2.4%)，而没有额外的时间和空间成本。

#### 4.3.3 使用边缘线索的机制

为了证明它优于 NLDF [35]，在 NLDF 中，IoU 损失被添加到网络的末端以惩罚边缘的错误。我们将相同的 IOU 损失添加到基线。该技术称为 edge\_NLDF。性能显示在 Tab. 3 的第 4 行。与基线模型相比，改进是有限的。这也证明了所提出的利用边缘信息的方法

表 4. NLDF 和我们方法生成的显著边缘对比。

更加有效。可视化结果显示在 Fig. 4 中。与没有边缘约束的基线模型相比，在加入 NLDF [35] 中使用的边缘惩罚后，边缘信息只能帮助细化边缘。特别是，这种惩罚既不能去除显著性预测掩模中多余的部分，也不能弥补缺失的部分。互补信息建模方法考虑了显著边缘信息与显著目标信息之间的互补性，在分割和定位上都有更好的表现。

此外，为了进一步证明显著边缘检测和显著目标检测是相互帮助和互补的。我们将 NLDF 生成的显著边缘与我们生成的显著边缘进行了对比。预训练模型和代码均由其作者提供。如 Tab. 4 所示，我们的方法生成的显著边缘要好得多，特别是在召回率和 F 度量指标下。结果表明，我们方法的边缘精度更高。

#### 4.3.4 互补特征融合

在获得显著边缘特征和多分辨率显著目标特征之后。我们的目标是融合这些互补的特征。本文对三种融合方法进行了比较。第一种方法是默认方法，它集成了 U-Net 框架顶部的显著边缘特征 ( $F_E$ ) 和显著目标特征  $\hat{F}^{(3)}$ 。第二种方法是逐步融合多分辨率特征  $\hat{F}^{(3)}$ ,  $\hat{F}^{(4)}$ ,  $\hat{F}^{(5)}$ ,  $\hat{F}^{(6)}$ ，称为 MRF\_PROG。第三种方法是提出的一对一引导，记为 MRF\_OTO。这里 MRF 表示多分辨率融合。结果分别显示在 Tab. 3 的第三、第五、第六行。可以看出，我们提出的一对一引导方法是最适合我们整个框架的。

#### 4.4. 与最先进方法的比较

在本节中，我们将我们提出的 EGNNet 与之前的 15 种最先进的方法进行比较，包括 DCL [28], DSS [17,18], NLDF [35], MSR [26], ELD [13], DHS [32], RFCN [48], UCF [62], Amulet [61], PAGR [63], PiCANet [33], SRM [49], DGRL [52], RAS [3] and C2S [29]。注意，上述方法的所有显著图都是通过运行源代码或其作者预先计算生成的。评估代码在 [10,17,18] 中提供。

**F 度量, MAE, 以及 S 度量。**我们从 F 度量、MAE 和 S 度量三个方面对我们提出的方法与其他显著目标



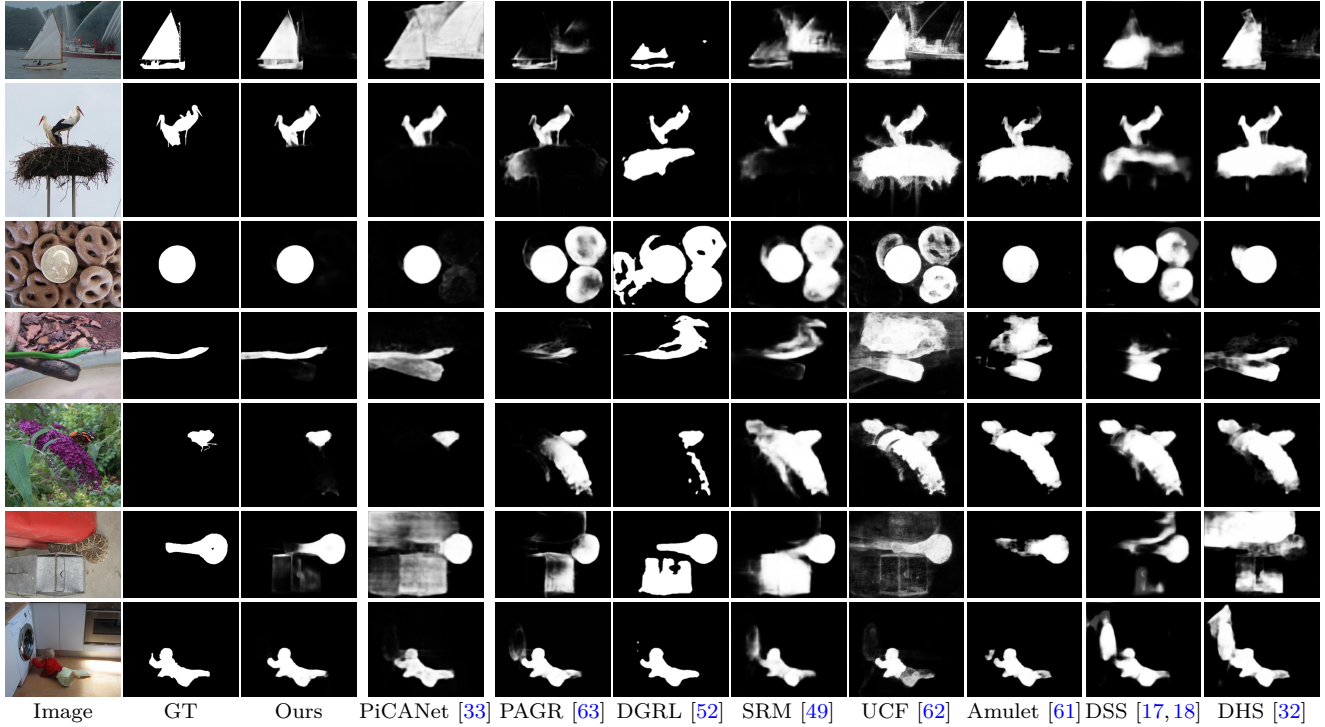


图 5. 与最先进方法的定量比较。

检测方法进行评估和比较, 如 Tab. 2 所示。我们可以看到, 不同的方法可能使用不同的主干网络。为了比较公平, 我们分别在 VGG [43] 和 ResNet [16] 上训练我们的模型。可以看到, 在所有比较的数据集的所有评估指标下, 我们的模型优于最先进的方法, 特别是在相对具有挑战性的数据集 SOD [36, 44] 上 (2.9% 与 1.7% 的 F 度量与 S 度量提升), 以及在最大型数据集 DUTS [46] 上 (3.0% 与 2.5%)。具体来说, 与当前的最佳方法相比, 6 个数据集上的平均 F 度量改进了 1.9 注意, 这是没有任何预处理和后处理的情况下实现的。

**精度-召回曲线。**除了在 Tab. 2 中显示的数值比较外, 我们还绘制了三个数据集上所有比较方法的精度召回曲线 Fig. 3。可以看出, 实红线表示所提方法在大多数阈值上优于其他方法。由于互补的显著边缘信息的帮助, 结果产生了清晰的边缘信息和精确的定位, 从而得到了更好的 PR 曲线。

**视觉比较。**在 Fig. 5 中, 我们展示了一些可视化结果。可以看出, 我们方法在显著目标分割和定位方面有更好的效果。值得一提的是, 得益于显著边缘特征, 我们的结果不仅可以突出显著区域, 而且还可以产生连贯的边缘。例如, 对于第一个例子, 由于复杂场景的影响, 其他方法无法准确定位和分割显著目标。然而, 得

益于互补的显著边缘特征, 我们的方法表现得更好。对于第二个例子, 其中显著目标相对较小, 我们的结果仍然非常接近真实标注。

## 5. 总结

在本文中, 我们的目标是保留显著目标边缘。不同于其他综合多尺度特征或利用后处理的方法, 我们关注显著边缘信息与显著目标信息之间的互补性。基于这一思想, 我们提出了 EGNet 来建模网络内部的这些互补特征。首先, 基于 U-Net 提取多分辨率显著目标特征; 然后, 我们提出了一个 non-local 显著边缘特征提取模块, 该模块将局部边缘信息与全局位置信息相结合, 得到显著边缘特征。最后, 我们采用一对一的引导模块来融合这些互补特征。在显著边缘特征的帮助下, 显著目标的边缘和定位得到了改善。我们的模型在没有任何预处理或后处理的情况下, 在 6 个广泛使用的数据集上超过了目前最先进的方法。我们还提供了 EGNet 的有效性分析。

**致谢** 本研究受到 NSFC (61572264), 国家”万人计划”青年拔尖人才支持计划, 天津市自然科学基金 (17JCJQJC43700, 18ZXZNGX00110) 的支持。



## 参考文献

- [1] Ali Borji, Ming-Ming Cheng, Qibin Hou, Huaizu Jiang, and Jia Li. Salient object detection: A survey. *CVM*, 5(2):117–150, 2019.
- [2] Ali Borji, Ming-Ming Cheng, Huaizu Jiang, and Jia Li. Salient object detection: A benchmark. *IEEE TIP*, 24(12):5706–5722, 2015.
- [3] Shuhan Chen, Xiuli Tan, Ben Wang, and Xuelong Hu. Reverse attention for salient object detection. In *ECCV*, pages 234–250, 2018.
- [4] Tao Chen, Ming-Ming Cheng, Ping Tan, Ariel Shamir, and Shi-Min Hu. Sketch2photo: Internet image montage. *ACM TOG*, 28(5):124:1–10, 2009.
- [5] Ming Cheng, Niloy J Mitra, Xumin Huang, Philip HS Torr, and Song Hu. Global contrast based salient region detection. *IEEE TPAMI*, 37(3):569–582, 2015.
- [6] Ming-Ming Cheng, Fang-Lue Zhang, Niloy J Mitra, Xiaolei Huang, and Shi-Min Hu. Repfinder: finding approximately repeated scene elements for image editing. *ACM TOG*, 29(4):83, 2010.
- [7] Wolfgang Einhäuser and Peter König. Does luminance-contrast contribute to a saliency map for overt visual attention? *European Journal of Neuroscience*, 17(5):1089–1097, 2003.
- [8] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *IJCV*, 88(2):303–338, 2010.
- [9] Deng-Ping Fan, Ming-Ming Cheng, Jiang-Jiang Liu, Shang-Hua Gao, Qibin Hou, and Ali Borji. Salient objects in clutter: Bringing salient object detection to the foreground. In *ECCV*, pages 186–202. Springer, 2018.
- [10] Deng-Ping Fan, Ming-Ming Cheng, Yun Liu, Tao Li, and Ali Borji. Structure-measure: A new way to evaluate foreground maps. In *ICCV*, pages 4548–4557, 2017.
- [11] Deng-Ping Fan, Zheng Lin, Jia-Xing Zhao, Yun Liu, Zhao Zhang, Qibin Hou, Menglong Zhu, and Ming-Ming Cheng. Rethinking rgb-d salient object detection: Models, datasets, and large-scale benchmarks. *arXiv preprint arXiv:1907.06781*, 2019.
- [12] Deng-Ping Fan, Wenguan Wang, Ming-Ming Cheng, and Jianbing Shen. Shifting more attention to video salient object detection. In *CVPR*, pages 8554–8564, 2019.
- [13] Lee Gayoung, Tai Yu-Wing, and Kim Junmo. Deep saliency with encoded low level distance map and high level features. In *CVPR*, 2016.
- [14] Wenlong Guan, Tiantian Wang, Jinqing Qi, Lihe Zhang, and Huchuan Lu. Edge-aware convolution neural network based salient object detection. *IEEE SPL*, 26(1):114–118, 2018.
- [15] Junfeng He, Jinyuan Feng, Xianglong Liu, Tao Cheng, Tai-Hsu Lin, Hyunjin Chung, and Shih-Fu Chang. Mobile product search with bag of hash bits and boundary reranking. In *CVPR*, pages 3005–3012, 2012.
- [16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *ICCV*, pages 770–778, 2016.
- [17] Qibin Hou, Ming-Ming Cheng, Xiaowei Hu, Ali Borji, Zhuowen Tu, and Philip Torr. Deeply supervised salient object detection with short connections. In *CVPR*, pages 3203–3212, 2017.
- [18] Qibin Hou, Ming-Ming Cheng, Xiaowei Hu, Ali Borji, Zhuowen Tu, and Philip Torr. Deeply supervised salient object detection with short connections. *IEEE TPAMI*, 41(4):815–828, 2019.
- [19] Qibin Hou, Peng-Tao Jiang, Yunchao Wei, and Ming-Ming Cheng. Self-erasing network for integral object attention. In *NIPS*, 2018.
- [20] Ping Hu, Bing Shuai, Jun Liu, and Gang Wang. Deep level sets for salient object detection. In *CVPR*, pages 2300–2309, 2017.
- [21] Laurent Itti and Christof Koch. Computational modeling of visual attention. *Nature reviews neuroscience*, 2(3):194–203, 2001.
- [22] Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE TPAMI*, 20(11):1254–1259, 1998.
- [23] Sen Jia and Neil DB Bruce. Richer and deeper supervision network for salient object detection. *arXiv preprint arXiv:1901.02425*, 2019.
- [24] Dominik A Klein and Simone Frintrop. Center-surround divergence of feature statistics for salient object detection. In *ICCV*, pages 2214–2219. IEEE, 2011.
- [25] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to

- document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [26] Guanbin Li, Yuan Xie, Liang Lin, and Yizhou Yu. Instance-level salient object segmentation. In *CVPR*, 2017.
- [27] Guanbin Li and Yizhou Yu. Visual saliency based on multiscale deep features. In *CVPR*, pages 5455–5463, 2015.
- [28] Guanbin Li and Yizhou Yu. Deep contrast learning for salient object detection. In *CVPR*, 2016.
- [29] Xin Li, Fan Yang, Hong Cheng, Wei Liu, and Dinggang Shen. Contour knowledge transfer for salient object detection. In *ECCV*, pages 355–370, 2018.
- [30] Yin Li, Xiaodi Hou, Christof Koch, James M Rehg, and Alan L Yuille. The secrets of salient object segmentation. In *CVPR*, pages 280–287, 2014.
- [31] Zun Li, Congyan Lang, Yunpeng Chen, Junhao Liew, and Jiashi Feng. Deep reasoning with multi-scale context for salient object detection. *arXiv preprint arXiv:1901.08362*, 2019.
- [32] Nian Liu and Junwei Han. Dhsnet: Deep hierarchical saliency network for salient object detection. In *CVPR*, pages 678–686, 2016.
- [33] Nian Liu, Junwei Han, and Ming-Hsuan Yang. Picanet: Learning pixel-wise contextual attention for saliency detection. In *CVPR*, pages 3089–3098, 2018.
- [34] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *CVPR*, pages 3431–3440, 2015.
- [35] Zhiming Luo, Akshaya Kumar Mishra, Andrew Achkar, Justin A Eichel, Shaozi Li, and Pierre-Marc Jodoin. Non-local deep features for salient object detection. In *CVPR*, 2017.
- [36] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *ICCV*, volume 2, pages 416–423, 2001.
- [37] Vida Movahedi and James H Elder. Design and perceptual validation of performance measures for salient object segmentation. In *IEEE CVPRW*, pages 49–56. IEEE, 2010.
- [38] David Mumford and Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *CPAM*, 42(5):577–685, 1989.
- [39] Derrick Parkhurst, Kinton Law, and Ernst Niebur. Modeling the role of salience in the allocation of overt visual attention. *Vision research*, 42(1):107–123, 2002.
- [40] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [41] Paul L Rosin and Yu-Kun Lai. Artistic minimal rendering with lines and blocks. *Graphical Models*, 75(4):208–229, 2013.
- [42] Ueli Rutishauser, Dirk Walther, Christof Koch, and Pietro Perona. Is bottom-up attention useful for object recognition? In *CVPR*, 2004.
- [43] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *ICLR*, 2015.
- [44] Jingdong Wang, Huaizu Jiang, Zejian Yuan, Ming-Ming Cheng, Xiaowei Hu, and Nanning Zheng. Salient object detection: A discriminative regional feature integration approach. *IJCV*, 123(2):251–268, 2017.
- [45] Lijun Wang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang. Deep networks for saliency detection via local estimation and global search. In *ICCV*, pages 3183–3192, 2015.
- [46] Lijun Wang, Huchuan Lu, Yifan Wang, Mengyang Feng, Dong Wang, Baocai Yin, and Xiang Ruan. Learning to detect salient objects with image-level supervision. In *CVPR*, pages 136–145, 2017.
- [47] Linzhao Wang, Lijun Wang, Huchuan Lu, Pingping Zhang, and Xiang Ruan. Saliency detection with recurrent fully convolutional networks. In *ECCV*, pages 825–841. Springer, 2016.
- [48] Linzhao Wang, Lijun Wang, Huchuan Lu, Pingping Zhang, and Xiang Ruan. Saliency detection with recurrent fully convolutional networks. In *ECCV*, 2016.
- [49] Tiantian Wang, Ali Borji, Lihe Zhang, Pingping Zhang, and Huchuan Lu. A stagewise refinement model for detecting salient objects in images. In *ICCV*, pages 4019–4028, 2017.
- [50] Tiantian Wang, Yongri Piao, Li Xiao, Lihe Zhang, and Huchuan Lu. Deep learning for light field saliency detection. In *ICCV*, 2019.

- [51] Tiantian Wang, Lihe Zhang, Huchuan Lu, Chong Sun, and Jinqing Qi. Kernelized subspace ranking for saliency detection. In ECCV, pages 450–466, 2016.
- [52] Tiantian Wang, Lihe Zhang, Shuo Wang, Huchuan Lu, Gang Yang, Xiang Ruan, and Ali Borji. Detect globally, refine locally: A novel approach to saliency detection. In CVPR, pages 3127–3135, 2018.
- [53] Wenguan Wang, Jianbing Shen, Xingping Dong, and Ali Borji. Salient object detection driven by fixation prediction. In ICCV, pages 1711–1720, 2018.
- [54] Ziqin Wang, Jun Xu, Li Liu, Fan Zhu, and Ling Shao. Ranet: Ranking attention network for fast video object segmentation. In ICCV, Oct 2019.
- [55] Saining Xie and Zhuowen Tu. Holistically-nested edge detection. In ICCV, pages 1395–1403, 2015.
- [56] Qiong Yan, Li Xu, Jianping Shi, and Jiaya Jia. Hierarchical saliency detection. In CVPR, pages 1155–1162, 2013.
- [57] Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang. Saliency detection via graph-based manifold ranking. In CVPR, pages 3166–3173, 2013.
- [58] Jing Zhang, Yuchao Dai, Fatih Porikli, and Mingyi He. Deep edge-aware saliency detection. arXiv preprint arXiv:1708.04366, 2017.
- [59] Lu Zhang, Ju Dai, Huchuan Lu, You He, and Gang Wang. A bi-directional message passing model for salient object detection. In ICCV, pages 1741–1750, 2018.
- [60] Pingping Zhang, Wei Liu, Huchuan Lu, and Chunhua Shen. Salient object detection with lossless feature reflection and weighted structural loss. IEEE TIP, 2019.
- [61] Pingping Zhang, Dong Wang, Huchuan Lu, Hongyu Wang, and Xiang Ruan. Amulet: Aggregating multi-level convolutional features for salient object detection. In ICCV, pages 202–211, 2017.
- [62] Pingping Zhang, Dong Wang, Huchuan Lu, Hongyu Wang, and Baocai Yin. Learning uncertain convolutional features for accurate saliency detection. In ICCV, pages 212–221. IEEE, 2017.
- [63] Xiaoning Zhang, Tiantian Wang, Jinqing Qi, Huchuan Lu, and Gang Wang. Progressive attention guided recurrent network for salient object detection. In CVPR, pages 714–722, 2018.
- [64] Jiaxing Zhao, Ren Bo, Qibin Hou, Ming-Ming Cheng, and Paul Rosin. Flic: Fast linear iterative clustering with active search. CVM, 4(4):333–348, Dec 2018.
- [65] JiaXing Zhao, Bo Ren, Qibin Hou, and Ming-Ming Cheng. Flic: Fast linear iterative clustering with active search. In AAAI, 2018.
- [66] Jia-Xing Zhao, Yang Cao, Deng-Ping Fan, Ming-Ming Cheng, Xuan-Yi Li, and Le Zhang. Contrast prior and fluid pyramid integration for rgbd salient object detection. In CVPR, 2019.
- [67] Jia-Xing Zhao, Jiangjiang Liu, Deng-Ping Fan, Yang Cao, Jufeng Yang, and Ming-Ming Cheng. Egned: edge guidance network for salient object detection. In ICCV, Oct 2019.
- [68] Kai Zhao, Shanghua Gao, Wenguan Wang, and Ming-Ming Cheng. Optimizing the f-measure for threshold-free salient object detection. In ICCV, Oct 2019.
- [69] Wangjiang Zhu, Shuang Liang, Yichen Wei, and Jian Sun. Saliency optimization from robust background detection. In CVPR, pages 2814–2821, 2014.
- [70] Yunzhi Zhuge, Gang Yang, Pingping Zhang, and Huchuan Lu. Boundary-guided feature aggregation network for salient object detection. IEEE SPL, 25(12):1800–1804, 2018.