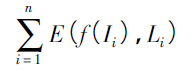
三、显著性物体检测评价标准

对显著性物体检测的算法效果进行评价的相关方法，主要是计算其算法得到的显著图像与标注的真值图像之间的误差，其公式表示为

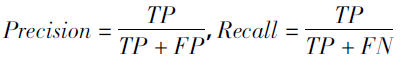


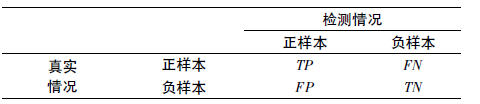
式中，i 为 1，2，…，n 张图片中的任意一张图片; Li 为任意一张图片的真值，即 Li 的取值为0 或1; f( Ii) 为对于任意输入的图片 Ii，通过相关显著性物体检测的算法 f(·) 得到的显著图像的值，这里 f( Ii) ∈［0，1］，0 表示背景，1 表示显著值; E 为显著性物体检测的评价标准。而算法优化的目标是使得非显著物体所在区域的像素值趋向于 0，显著物体所在区域的像素值趋向于 1。

为了进行不同方法之间的性能对比，在建立统一的标准数据集后，需要有统一的评价标准和指标。目前，关于显著性物体检测的评价指标主要有 PＲ 曲线、F 测度值( F-measure) 、平均绝对误差(MAE) 、接收者操作特征曲线(ＲOC) 、 ＲOC 曲线下方的面积(AUC ＲOC) 和 S 测度值( S-Measure) 等，下面分别对各评价指标进行简单的介绍。

**3.1 PＲ 曲线**

PＲ 曲线是反映 Precision 和 Ｒecall 相互关系的曲线。其中，Precision 指的是查准率，Ｒecall 指的是查全率( 也译为召回率) ，一般将 Precision 作为 PＲ 曲线的纵坐标，Ｒecall 作为 PＲ 曲线的横坐标。所谓查准率 ( Precision) 指的是检测出的正样本数与检测的所有正样本数之间的比值，查全率( Ｒecall) 指的是正确检测的正样本数占所有正样本数的比例。给出其定义:





检测结果和真实样本之间的关系

在实际的显著性物体检测过程中，一般用下式进行计算。



式中，S 为检测方法得出的显著值; G 为真实的实际标注的值。在图 5 中用通俗易懂的可视化的形式展示了这几个值的具体对应区域及其相互关系。其中，白色区域( 绿色框显示部分) 为标记的 G 真值为 1 的部分，其他区域( 黄色区域) 为真值 G 的背景 0 部分; 蓝色框的区域为显著性检测方法得到的显著区域，S 与 G 相交的部分为其重叠的区域，即为 TP，蓝色框的区域去除重叠部分为 FP，绿色框的白色区域去除重叠部分为 FN，黑色框的黄色区域去除画线段的部分为 TN。

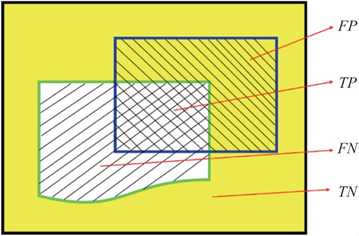
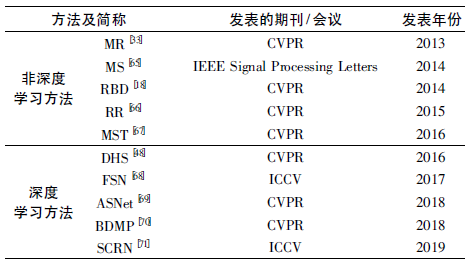


图 5 PＲ 值相关区域示意图

这里，分别选择了另外一些传统显著性计算方法和基于深度学习的计算方法对它们在 ECSSD 数据集上检测效果进行了测试，得到如图 6 所示的 PＲ 曲线图。前 5 种方法主要是传统的显著性计算方法，后 5 种主要是基于深度学习的显著性计算方法。从 PＲ 曲线上来看，一般认为曲线越靠近图的右上方表示该模型的效果越好。从图 6 中可以看出，基于深度学习的显著性计算方法明显优于传统的显著性计算方法，因为基于深度学习方法的 PＲ 曲线的高度都高于传统方法的 PＲ 曲线。



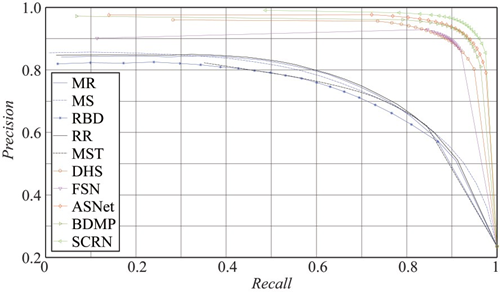


图 6 10 种方法在 ECSSD 数据集中的 PＲ 曲线

**3.2 F-measure**

因 PＲ 曲线的值不能很好地评估模型的有效性，提出了 F 测度值( F-measure) ，简称 F值。F 值是根据查准率( Precision) 和查全率( Ｒecall) 加权和的平均值求得的结果，当一个显著性物体检测模型的 F 值较高时，表示该模型越有效。F 值可以计算得出:



式中 β2 的值，一般经验性地取 0. 3。

同样地，计算了上述 10 种方法的 F 值，一般认为F 值越大，该方法的检测效率越好。10 种方法在 EC- SSD 数据集中的 F 值如图 7 所示。从图 7 中可以看出，基于深度学习的方法比传统的方法高出 0. 2，表明了基于深度学习的方法要优于传统的显著性计算方法。

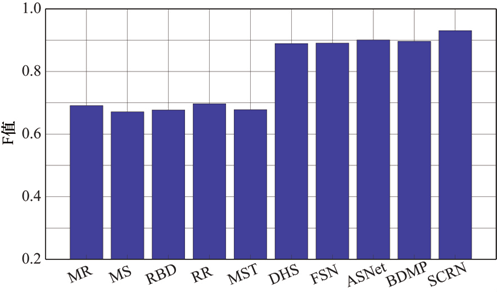


图 7 10 种方法在 ECSSD 数据集中的 F 值

另外一种关于 F 值的曲线图将统计的阈值和对应的 F 值进行关联，如图 8 所示。

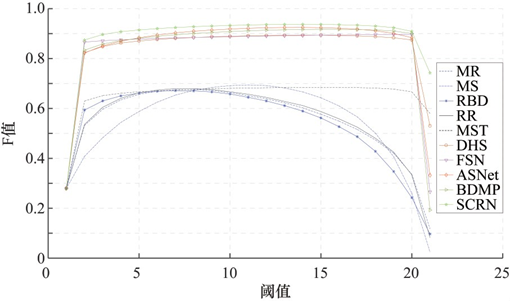


图 8 10 种方法在 ECSSD 数据集中的 F 值和对应的阈值曲线图

**3.3 MAE**

MAE ( Mean Absolute Error) 即为平均绝对误差，反映的是模型的检测值与实际真值的误差情况，可表示为式



式中，w 为检测图像的宽度; h 为检测图像的高度; S 为检测模型计算出的显著值; G 为标注图像的真值。将 S 和 G 二值化后，通过计算每个像素点在显著值与真值之间的误差的平均值，反映出图像中各个像素点通过模型检测的正确率。因此，平均绝对误差值越低，表示模型检测的精确程度越高。

**3.4 ＲOC 和 AUC**

ＲOC( Ｒeceiver Operating Characteristic curve) 译为接收者操作特征曲线，而 AUC( Area Under the Curve of ＲOC) 为 ＲOC 曲线下的面积。通过表 3 可知，TP 表示检测为正样本，实际也是正样本，而 FP 表示检测为正样本，但实际为负样本。基于此，得出两个比值，即真正率 TPＲ ( True Positive Ｒate) 和假正率 FPＲ ( False Positive Ｒate) ，用式表示



式中，TPＲ 表示正确被检测出的正样本和实际的正样本的比值，FPＲ 表示被检测为正样本的负样本占实际的负样本的比值。在 ＲOC 曲线中，一般将 FPＲ 设为横坐标，TPＲ 设为纵坐标。因得到的显著值 S 的灰度值范围为［0，255］，二值化后的区间为［0， 1］，通过设置阈值对像素进行标记，若 S 中的某个像素的值大于阈值则标记为 1，否则为 0，最终能计算出 TPＲ 和 FPＲ 的值。

ＲOC 曲线及其 AUC 如图 9 所示。从图 9 中可以看出，如果 ＲOC 曲线越靠近左上方，那么检测结果效果越好。同时，通过积分可以计算得到 AUC 的值，其值越大，则表示显著目标检测算法效果越好。

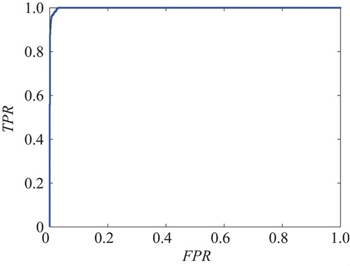


图 9 ＲOC 曲线及其 AUC

**3.5 S-measure**

以上所有的评测方法主要针对像素级别的误差评估，但对显著性物体的结构性特性没有进行评测，Fan等针对以往的评测标准忽略了结构相似性，提出新的评测指标 S-measure。该评测指标主要由区域结构相似性度量和物体结构相似性度量组成，从而更好地评价前景物体，如显著性物体。具体公式如下:



式中，S 为结构性评测 S 度量值( S-measure) ; a 为权重参数，其取值范围为［0，1 ］，一般地，a 的值设置为0. 5。So 为结构相似性度量值:



式中，λ 为前景区域与整个图像区域的比值。作者通过实验得出前景和背景具有强烈的对比特性以及内部近似均匀分布特性，因此，OFG 和 OBG 分别为前景和背景相似性度量值，将两者通过 λ 参数加权求和，作为最终的物体结构性度量值 So。Sr 为区域结构相似性度量值:



即在计算区域相似性度量值时，将真值和检测的结果切分成 k × k 块，然后用结构性评价指标 SSIM［74］计算每一块的 ssim( k) ，再加上每块占前景的比例权重 w ，最后将每块的乘积相加得到区域相似性度量值 Sr。通过在相关数据集上进行实验，表明了 S-measure 评测指标具有很强的鲁棒性和稳定性，因其计算方式简单、计算速度快，将被广泛运用于显著性物体检测当中。

四、 显著性物体检测的应用

显著性物体检测的相关方法的不断出现、数据集的不断更新和评价标准的更加有效，使得显著性物体检测的应用范围越来越广。正如在本文开始所述，显著性物体检测既可以应用于各种任务的预处理，也可以用于完成相关任务的检测。本节介绍了近年来显著物体检测在多个领域的应用。显著性物体检测可以应用在计算机视觉、计算机图形学、多媒体等众多领域的许多任务中。下面分别介绍其代表性的工作。

**4.1 计算机视觉**

计算机视觉中有各种各样的视觉任务，这里主要列举了显著性物体检测在目标检测、目标识别、目标分割和目标跟踪等方面的应用。

在目标检测任务当中，主要任务是把目标的轮廓从背景中检测出来。针对水下复杂环境，如何从大量的水下场景的视频和图像中检测物体的轮廓，是一项具有挑战的任务。

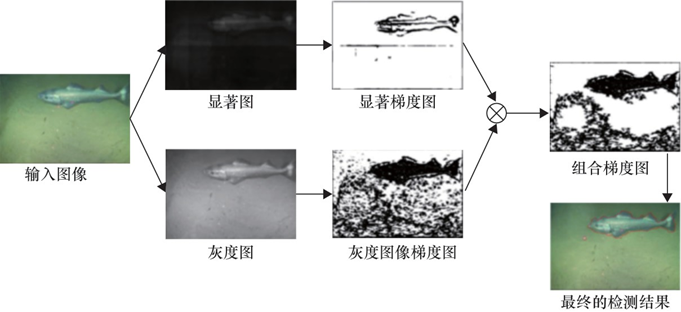


图 10 基于显著物体检测的水下目标轮廓检测流程图

在目标识别的任务当中，需要检测到物体，并识别出物体所属于的类别。针对细粒度图像识别问题，判别类间物体类别，即在识别出物体所属鸟类( 大类) 的同时判断该物体具体属于哪种鸟类( 小类) 。

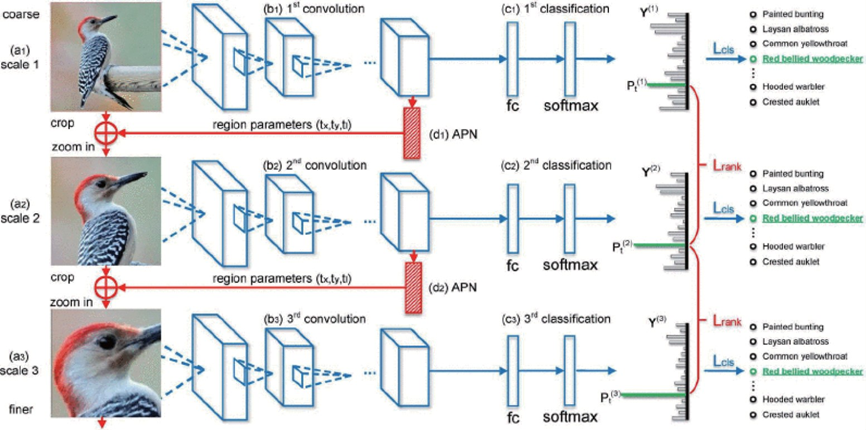


图 11 递归注意卷积神经网络的细颗粒度图像识别流程图

在目标分割任务当中，主要是将目标从背景中分离出来，并且找到目标物体的轮廓，在视频中的目标分割使用较多。

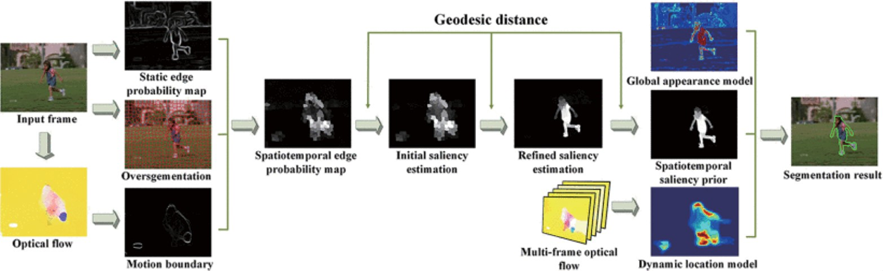


图 12 基于显著性检测的视频目标分割流程图

目标跟踪任务中，需要先检测出目标，在目标后续的运动帧当中，根据目标的位置变化和大小变化等，实现目标的锁定。

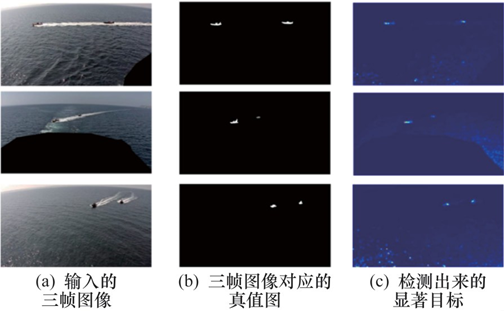


图 13 基于显著性检测的海上目标跟踪示意图

**4.2** 计算机图形学

在计算机图形学领域，显著性物体检测也有广泛的应用。显著性物体检测算法一般先找到图片中的最显著的目标，通过其他方法再次优化该目标，然后实现背景和目标的分离、目标的大小调整、颜色渲染和缩放等操作。

**4.3** 多媒体

在多媒体领域，主要涉及图像和视频的压缩和图像的检索。如在文献［83］中，作者将显著性检测方法应用在图像压缩方面，从而以较少的计算量来产生高质量的压缩图像，实现高压缩率。

五、总结与展望

主要介绍了显著性物体检测的方法、数据集、评价标准及其在多个领域的应用。总的说来，显著性物体检测仍然是一个非常具有挑战性的工作，具有十分重要的研究价值。目前，显著性物体检测的首选方法是基于深度学习的方法，具有较高的检测精度，可适应于复杂场景下的物体检测。未来，显著性物体检测将更加关注网络模型的大小、检测的精度以及实时性方面，作为视觉任务的前序的预处理，为各种实际的应用任务服务。

另外，本文初始部分提到过，人类的选择性视觉注意机制包含两个方面，一个是选择场景中比较醒目的区域，既涉及自下而上的视觉处理过程，也涉及自顶向下的视觉处理过程; 另一个是选择场景中感兴趣的区域，主要涉及自顶向下的视觉处理过程。比较醒目的区域与感兴趣的区域在视觉场景中并不总是一致的。例如，驾驶员在驾驶车辆行驶在道路上时，需要经常关注交通标志牌和信号灯，以便于按照交规采取正确的行动，然而在复杂的交通场景中，交通标志牌和信号灯常常不处于醒目位置。到目前为止，在视觉注意力相关的研究中，人们关注的基本都是比较醒目的区域，对于感兴趣区域的视觉选择注意方面，未有有效的研究。关于感兴趣区域的视觉选择注意模型和计算方法的研究具有极大的挑战性，并且在这一方面的突破将有助于视觉信息处理技术、类脑计算等领域的发展，值得关注研究。