

基于双向学习的弱监督阴影-对象实例检测<sup>\*</sup>王 晓 刘 洪 廖丽琪 邹世奇  
(贵州大学大数据与信息工程学院 贵阳 550025)

**摘 要:** 现有阴影-对象实例检测网络都是基于掩码标签的全监督训练,但掩码标签标注难度和成本较高。仅利用边界框标签进行监督训练可以有效降低数据集的标注难度和成本,但弱监督会导致预测实例掩码精度降低。为了解决这一问题,首次利用弱监督方法进行阴影-对象实例检测,提出了一种基于双向学习结构的弱监督阴影-对象实例检测网络。首先,设计了教师-学生双向学习结构,利用教师网络预测的结果作为学生网络监督训练的的伪掩码标签,通过指数移动平均方法更新教师网络的参数的方法提高弱监督检测的精确度。其次,通过投影损失对预测掩码进行精准定位,并引入了能表征图像色彩先验信息的色彩相似性指标,结合交叉熵损失函数设计了色彩相似性损失函数,提高了整体网络的检测性能。为了验证所提方法的有效性和提高网络的鲁棒性,构建了一个阴影-对象实例检测数据集,并在该数据集与公开数据集 SOBA 上验证了本文网络的预测能力,平均精度值分别达到了 53.3 和 51.5。

**关键词:** 阴影检测;实例分割;弱监督;双向学习;色彩相似性

**中图分类号:** TP394.1;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Weakly supervised shadow-object instance detection with  
bidirectional learning

Wang Xiao Liu Hong Liao Liying Zou Shiqi

(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

**Abstract:** Current shadow-object instance detection methods rely on fully supervised training with mask labels. However, mask labeling is both complex and costly. Utilizing only bounding box labels for training can reduce annotation challenges and costs, but this weak supervision may decrease the accuracy of instance mask predictions. To address this issue, weakly supervised methods were first utilized for shadow-object instance detection. A weakly supervised shadow-object instance detection approach with bidirectional learning network is proposed. Firstly, a teacher-student bidirectional learning structure was designed, utilizing the predicted results of the teacher network as pseudo mask labels for the supervised training of the student network. The accuracy of weakly supervised detection was enhanced by updating the parameters of the teacher network using the exponential moving average method. Secondly, the prediction mask is precisely positioned using projection loss, and a color affinity index is introduced to represent the color prior information of the image. By integrating this with the cross-entropy loss function, a color affinity loss function is designed to enhance the network's overall detection performance. To verify the effectiveness of the proposed method and enhance the network's robustness, a shadow-object instance detection dataset was constructed. The predictive capability of the network was validated using both this dataset and the public dataset SOBA, with average precision values of 53.3 and 51.5, respectively.

**Keywords:** shadow detection;instance segmentation;weak supervision;bidirectional learning;color affinity

## 0 引 言

阴影是一种普遍存在于自然界的光学现象,当光线照射到物体上时,因物体遮挡部分光线传播而形成的较暗区

域被称为阴影。阴影形成的物理特性使得阴影检测在光方向估计、显著性目标检测<sup>[1]</sup>、图像编辑<sup>[2]</sup>等图像处理应用中起到至关重要的作用。此外,图像中存在的阴影也可能会给图像处理与分析带来干扰,例如阴影会干扰目标识别<sup>[3]</sup>、

目标跟踪<sup>[4]</sup>和文本检测等任务。因此,阴影检测具有重要的研究意义,是计算机视觉的重要研究方向之一。

现有的阴影检测方法主要分为基于图像先验知识与特征提取的传统方法和基于深度学习的方法。传统阴影检测利用不同的先验信息,例如色彩、梯度和明度<sup>[5]</sup>等特征检测图像中的阴影。随着深度学习的发展,Khan 等<sup>[6]</sup>首次在阴影检测中引入卷积神经网络。Le 等<sup>[7]</sup>通过引入对抗神经网络进行训练获得高精度的阴影掩码。Hu 等<sup>[8]</sup>利用图像中的空间上下文信息对阴影有不同的作用这一特点检测阴影,一定程度上克服了阴影同时投影在不同表面不能准确识别的缺点。

前述研究主要集中于独立阴影检测,然而遮挡光线的对象与阴影紧密相关,通过对阴影及其关联对象的联合检测能为推断光源位置、照明条件及场景几何等信息提供重要线索。为此,Wang 等<sup>[9]</sup>首次提出一种基于光方向引导的阴影-对象实例检测(light guided instance shadow-object association, LISA)网络,该网络基于 Mask R-CNN<sup>[10]</sup>设计双通道结构以检测对象与阴影及其匹配关系,把检测任务分别设置为独立阴影、独立对象以及阴影及其关联对象 3 个检测类别,并创建了首个用于阴影-对象实例检测的 SOBA(shadow object associate)数据集,但独立阴影、独立对象以及阴影及其关联对象三者之间的检测并未有内在联系,且阴影与对象实例的外观有极大差异,导致二者关联关系时常忽略。Wang 等<sup>[11]</sup>基于单阶段全卷积网络 CondInst<sup>[12]</sup>设计了单阶段实例阴影检测(single-stage instance shadow detection, SSIS)网络,这个网络只包括为阴影实例、对象实例和阴影-对象关联生成掩码的全卷积操作,并利用偏移向量设计了阴影-对象关系学习模块,学习阴影与对象之间的空间关系,提高了网络检测阴影的能力。在此基础上,Wang 等<sup>[13]</sup>通过添加边缘检测模块、可变形卷积掩码头、图像增强策略等设计了 SSISv2,显著地提高了网络在复杂场景中检测阴影的能力,是现有阴影-对象实例检测中性能最好的网络。

尽管学者们设计了优秀的网络模型用于阴影及其关联对象的检测,但由于标注掩码费时费力、成本极高、手工标注的掩码存在误差,导致阴影-对象实例检测网络很难在大型数据集上进行监督训练,使得网络可能存在泛化能力不强,对特殊场景的图片检测性能弱的问题。而弱监督利用边界框标签进行训练,大大地降低了标注成本,为阴影-对象实例检测模型在大型数据集上进行监督训练提供了可能。综上所述,本文借鉴 BoxInst<sup>[14]</sup>颜色相似性和 BoxTeacher<sup>[15]</sup>教师-学生双向学习结构的设计理念,首次利用弱监督方法进行阴影-对象实例检测,提出了基于双向学习结构的弱监督阴影-对象实例检测网络。设计教师-学生双向学习结构,教师网络预测实例掩码并筛选高质量预测掩码作为监督学生网络训练的伪掩码标签,学生网络通过指数移动平均(exponential moving average, EMA)方法

更新教师网络的参数。针对阴影-对象实例检测,本文在 SSISv2 的基础上引入投影损失替换原网络的掩码损失,利用边界框标签对预测掩码进行精准定位,实现弱监督阴影-对象实例检测,并引入色彩相似损失函数,提高模型的检测性能。最后,构建了阴影-对象实例检测(shadow object instance detection, SOID)数据集,在 SOBA 数据集的基础上丰富了数据集的多样性,提高了网络的检测性能与鲁棒性。

## 1 模型设计

### 1.1 网络结构

本文提出了一种基于双向学习结构的弱监督阴影-对象实例检测网络,其总体结构如图 1 所示。教师网络和学生网络共享同一个由特征提取模块、阴影-对象关系学习模块构成的阴影-对象实例检测网络。输入图像  $I$  经 ResNeXt101-BiFPN<sup>[16]</sup>构成的特征提取模块处理获得多尺度特征图  $P_i$ ,通过引用 SSIS 的双向学习模块实现对阴影及其关联对象实例掩码的预测。同时,借鉴了 BoxInst 的设计理念,在不改变网络结构的前提下引入投影损失和色彩相似性损失代替原网络的掩码损失实现弱监督训练。

### 1.2 教师-学生双向学习结构

如图 1 所示,教师网络和学生网络共用同一个阴影-对象实例检测网络,给定输入图像预测实例掩码。教师网络  $f_t$  利用边界框标签进行监督训练输出预测结果,包括阴影-对象的实例掩码、分类分数和预测掩码置信度,并对预测掩码进行评分筛选,保留高质量预测结果作为学生网络  $f_s$  的伪掩码标签进行监督训练。通过与学生网络预测的掩码计算掩码损失,通过 EMA 更新教师网络的参数,即:

$$f_t = \alpha \cdot f_t + (1 - \alpha) \cdot f_s \quad (1)$$

其中,  $\alpha$  为经验参数,设置为 0.999。然而教师网络预测的伪掩码可能存在预测结果不准确、精度低等问题,不能为学生网络训练的提供有效监督。受 SOLO<sup>[17]</sup>的启发,引入感知掩码置信度  $S_i$  用以评估伪掩码的质量,将  $m_{i,x,y}^b \in \{0,1\}$  和  $m_{i,x,y} \in (0,1)$  分别表示为第  $i$  个伪掩码中像素点  $(x,y)$  的二值掩码及其 Sigmoid 概率值,  $C_i$  为第  $i$  个伪掩码的预测掩码置信度,  $S_i$  定义如下:

$$S_i = \sqrt{C_i \cdot \frac{\sum_{x,y} 1(m_{i,x,y} > \tau_m) \cdot m_{i,x,y} \cdot m_{i,x,y}^b}{\sum_{x,y} 1(m_{i,x,y} > \tau_m) \cdot m_{i,x,y}}} \quad (2)$$

其中,  $1(\cdot)$  为指示函数,当  $m_{i,x,y} > \tau_m$  时为 1,反之为 0。 $\tau_m$  为评估掩码阈值,设置为 0.5。不同于 Boxteacher 对单个目标进行感知掩码质量评价,需要同时检测阴影及其关联对象,故对两个掩码进行了综合掩码质量评价。

### 1.3 阴影-对象实例检测网络

#### 1) 特征提取模块

特征提取是指从原始图像中提取出具有代表性的特

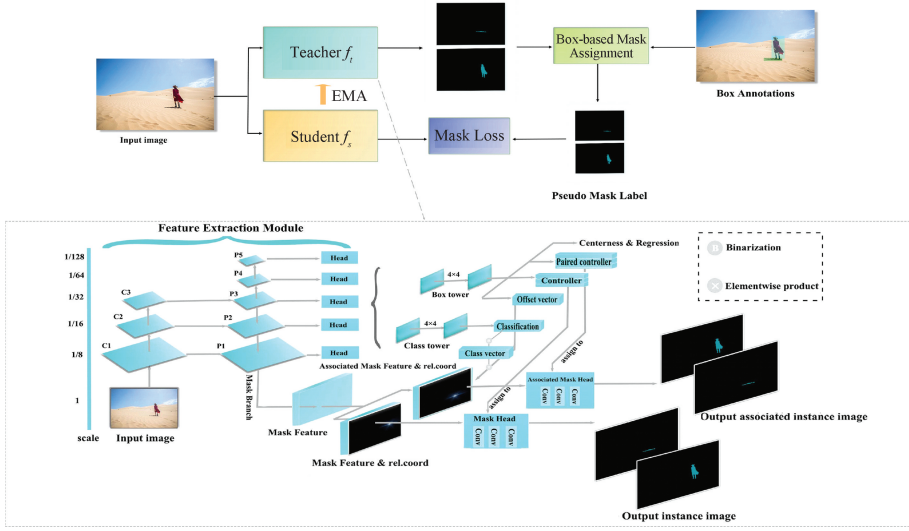


图 1 基于弱监督的阴影-对象实例检测网络架构

Fig. 1 Shadow-object instance detection network architecture with weakly supervision

征,用于分析、建模和预测等任务。本模块参考了 SSIS 的特征提取分支,将 ResNeXt101-BiFPN 结构作为特征提取模块,如图 1 所示。不同于常规金字塔结构,该模块对特征层  $P_3$  进行下采样获得尺度更小的特征图层  $P_4$  和  $P_5$ ,用以更全面地描述图像全局信息。最终,输入图像  $I$  通过特征提取模块获得预测头,分别预测分类分数、偏移向量、掩码预测器和关联掩码预测器、回归度与中心度等参数。

## 2) 阴影-对象关系学习模块

为精准描述阴影及其关联对象之间的位置关系,SSIS 使用偏移向量预测二者的空间位置关系,并提出了基于偏移向量的阴影-对象关系学习模块。如图 2 所示,通过关系学习模块,网络为每对阴影和关联对象预测两组掩码,其中仅采用主分支(图 2(a)和(b)的左分支)预测的掩码为输出掩码,关联分支(图 2(a)和(b)的右分支)用以预测阴影及其关联对象的匹配关系。假设网络预测第  $n$  个实例的位置为  $L_n$ ,利用偏移向量  $O_n$  与分类分数  $C_n$  计算关联位置  $A_n$ :

$$A_n = L_n + O_n \times C_n \quad (3)$$

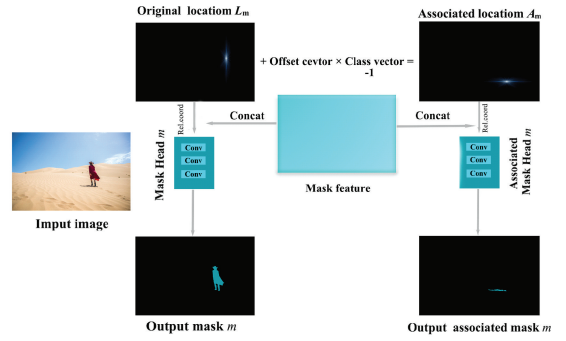
其中,  $C_n \in \{-1, 1\}$ , 当  $C_n = -1$  表示通过关联对象预测阴影,  $C_n = 1$  时通过阴影预测关联对象。

## 1.4 投影损失与色彩相似性损失

为了实现弱监督阴影-对象实例检测这一任务,受 BoxInst 的启发,本文利用由投影损失和色彩相似性损失组成的掩码损失取代 SSISv2 的逐像素监督的掩码损失。

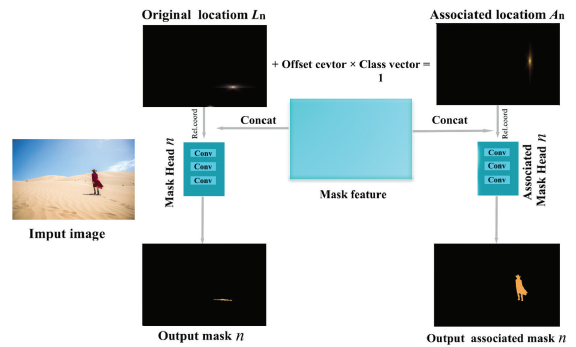
### 1) 投影损失

如图 3 上部所示,通过将掩码标签  $b \in \{0, 1\}^{H \times W}$  向水平与垂直两个方向分别投影获得一个矩形掩码,计算该矩形掩码的边框与边界框标签之间的 dice 损失。投影损失可以确保预测掩码的位置能与边界框标签相匹配,对实例进行精准定位。为此,对投影损失作如下定义:



(a) 对象→阴影

(a) Object→shadow



(b) 阴影→对象

(b) Shadow→object

图 2 阴影-对象关系学习模块

Fig. 2 Shadow-object relationship learning module

$$Proj_x(b) = l_x \quad (4)$$

$$Proj_y(b) = l_y \quad (5)$$

$$L_{proj} = L(Proj_x(\tilde{m}), Proj_x(b)) +$$

$$L(Proj_y(\tilde{m}), Proj_y(b)) = L(\tilde{l}_x, l_x) + L(\tilde{l}_y, l_y) \quad (6)$$



其中,  $Proj_x: \mathbb{R}^{H \times W} \rightarrow \mathbb{R}^W$  和  $Proj_y: \mathbb{R}^{H \times W} \rightarrow \mathbb{R}^H$  分别表示将掩码投影到 X 轴与 Y 轴。 $\mathbf{l}_x \in \{0,1\}^W$  表示在 X 轴的一维分割掩码,  $\mathbf{l}_y$  同理。对预测掩码  $\tilde{\mathbf{m}}$  应用式(1)、(2)相同的投影获得  $\tilde{\mathbf{l}}_x$  和  $\tilde{\mathbf{l}}_y$ 。 $L_{proj}$  为 dice 损失,用于计算边界框标签与预测掩码投影之间的差异。

## 2) 色彩相似性损失

然而想要获得准确的预测结果,仅靠掩码损失是不够的,在绝大多数实例分割模型中,如 Mask R-CNN 和 CondInst 都是以逐像素的方式进行监督训练。但在弱监督任务中并没有掩码标签,所以逐像素的监督方法并不可用。为此,本文借鉴了 BoxInst 利用成对像素进行监督训练的方法,实现弱监督阴影-对象实例检测。首先,基于预测框内的像素集合构建无向图  $G=(V,E)$ ,  $V$  代表框内像素点的集合,  $E$  代表相邻像素点构成的边集,无向图示例如图 3 下部分所示。将  $y_e \in \{0,1\}$  定义为边的标签,  $y_e=1$  表示构成边的相邻像素点具有相同标签,而  $y_e=0$  则表明它们的标签不同。设相邻像素  $(i,j)$  和  $(l,k)$  构成边  $e_{(i,j) \rightarrow (l,k)}$ , 网络预测  $\tilde{m}_{i,j} \in \{0,1\}$  表示像素点为目标概率,进而可以描述  $y_e=1$  的概率和网络预测由相邻像素构成的边是否为实例边缘的概率分布的损失函数  $L_{pairwise}$  为:

$$P(y_e=1) = \tilde{m}_{i,j} \cdot \tilde{m}_{l,k} + (1-\tilde{m}_{i,j}) \cdot (1-\tilde{m}_{l,k}) \quad (7)$$

$$L_{pairwise} = -\frac{1}{N} \sum_{e \in E_{in}} y_e \log P(y_e=1) + (1-y_e) \log P(y_e=0) \quad (8)$$

其中,  $E_{in}$  是包含框内至少一个像素的边的集合,使用  $E_{in}$  代替  $E$  可以防止框外大量像素参与监督训练,减轻训练压力。



图 3 投影损失与可能预测的边缘

Fig. 3 Projection loss and possible predicted edges

同时,色彩特征在实例分割的研究中具有重要的意义,通过大量观察结果可以发现如果相邻像素具有相似颜色,则相邻像素大概率具有相同标签。为此,定义色彩相似性指标  $S_e$ :

$$S_e = S(\mathbf{h}_{i,j}, \mathbf{h}_{l,k}) = \exp\left(-\frac{h_{i,j} - h_{l,k}}{\theta}\right) \quad (9)$$

其中,  $\mathbf{h}_{i,j}$  和  $\mathbf{h}_{l,k}$  分别是与该边连接的相邻像素点  $(i,j)$  和  $(l,k)$  的色彩向量。 $\theta$  是超参数,设置为 2。通过比较相邻像素的色彩相似度就可以获得值得信赖的边缘像素点,

因此将  $L_{pairwise}$  替换为色彩相似性损失函数:

$$L_{color} = -\frac{1}{N} \sum_{e \in E_{in}} 1_{(S_e > \tau_{color})} \log(P(y_e=1)) \quad (10)$$

其中,  $1_{(S_e \geq \tau_{color})}$  为指示函数,当  $S_e \geq \tau_{color}$  为 1,反之则为 0;  $\tau_{color}$  为阈值设置为 0.3。

## 2 实验结果及分析

### 2.1 数据集及实验设置

SOID 数据集:为了增加数据集的多样性,本文构建了 SOID 数据集,通过从 SBU<sup>[18]</sup>、ISTD<sup>[19]</sup>、Microsoft COCO<sup>[20]</sup> 以及利用人行道、公园、建筑等关键词在谷歌图像和百度图像等多种图像检索引擎中获得互联网图像。共选用 1 360 张图像,并使用 Adobe Photoshop 和 DaVinci Resolve 软件描绘阴影掩码以及关联对象掩码,总计 4 871 对阴影-对象实例。如图 4 所示,展示了 SOID 的部分图像,并给出了 SOID 数据集中图像实例对数的统计关系,如图 5 所示,从右部的直方图可以观察到该数据集不同图像含有的阴影-对象实例对占比情况,平均每张图片有 2.6 对阴影-对象实例,超过 7.7% 的图像含有 6 对以上的阴影-对象实例。



图 4 SOID 数据集部分示例图

Fig. 4 Sample diagram of SOID dataset

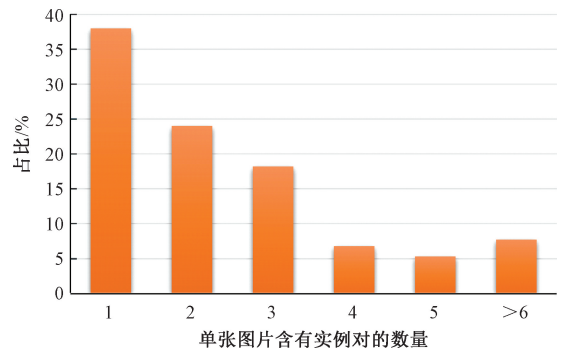


图 5 SOID 数据集统计特性

Fig. 5 SOID dataset statistical properties

SOBA 数据集:是 LISA、SSIS、SSISv2 共同使用的阴影-对象实例检测数据集,包含 1 000 张测试图像,3 623 对阴影-对象实例。为验证本文方法的有效性,在两个数据集上进行了实验。



实验设备选取的 CPU 为 Inter i7-12700F 处理器,内存为 64 G,GPU 选用 RTX3080Ti。参考 SSISv2 的训练方法,设置基本学习率为 0.001,采用 warm-up<sup>[21]</sup>策略,前 10 000 次迭代中学习率从 0.000 1 提升至 0.001,经过 40 000 次迭代后学习率降低至 0.000 1,并在 45 000 次迭代后停止学习。Batch-size 设置为 2。

2.2 评价指标

为了更好地评估模型性能,本文选用了平均精度 (average precision, AP) 和阴影-对象平均精度 (shadow object average precision, SOAP) 作为评价指标,SOAP 是 LISA、SSIS、SSISv2 中提出的用以评估阴影-对象匹配关系的指标。AP 通过计算每一个目标的召回率 (Recall) 和精确率 (Precision) 获得的 PR 曲线的面积评估模型的检测性能,如下所示:

$$R = \frac{TP}{TP + NP} \times 100\%$$
 (11)

$$R = \frac{TP}{TP + NP} \times 100\%$$
 (12)

$$AP = \int_0^1 P(R) dR$$
 (13)

其中, $P$  为精确率, $R$  为召回率, $TP$  是识别为正样本的数量, $FP$  是为负样本的数量, $FN$  是未检测的正样本数量。SOAP 在与计算 AP 的基础上,引入了新的正样本选择策略:须同时满足预测阴影实例与真实阴影实例、预测关联对象实例与真实关联对象实例、预测阴影-对象实例与真实阴影-对象实例的 IOU 都不小于阈值  $\tau$  才考虑该样本为正样本,阈值  $\tau$  的选择策略为 $[0.5 : 0.05 : 0.95]$ ,计算不同阈值后求平均值获得 SOAP。

2.3 消融实验

为了验证色彩相似性损失函数和教师-学生双向学习模块的有效性,本文在 SOBA 数据集上进行了验证实验。baseline 的基础网络是单阶段目标检测器 CondInst,但为了实现弱监督与阴影-对象实例检测两项任务,分别添加了投影损失与阴影-对象关系学习模块,而“ $+L_{color}$ ”表示添加色彩相似性损失函数;“ $+teacher-student$ ”表示添加教师-学生双向学习模块。实验结果如表 1 所示。

表 1 消融实验  
Table 1 Ablation experiment

基线	$+L_{color}$	$+teacher-student$	Associate AP <sub>segm</sub>	Associate AP <sub>bbox</sub>	Instance AP <sub>segm</sub>	Instance AP <sub>bbox</sub>	SOAP <sub>segm</sub>	SOAP <sub>bbox</sub>
✓			48.3	56.8	40.6	39.3	29.1	26.1
✓	✓		50.4	58.3	42.3	40.5	29.5	26.6
✓		✓	49.8	57.6	41.5	40.0	29.3	26.4
✓	✓	✓	<b>51.5</b>	<b>58.7</b>	<b>42.9</b>	<b>40.8</b>	<b>29.8</b>	<b>26.9</b>

由实验结果可得,仅添加  $L_{color}$  和仅添加教师-学生双向学习模块使得评价阴影-对象实例匹配关系的指标 Associate AP<sub>segm</sub> 相较于基础网络提升了 2.1 和 1.5,评价预测掩码精度的指标 Instance AP<sub>segm</sub> 提高了 1.7 和 0.9,SOAP<sub>segm</sub> 提升至 29.5 与 29.3。使用本文方法后,Associate AP<sub>segm</sub>、Instance AP<sub>segm</sub>、SOAP<sub>segm</sub> 对比基础网络分别提升了 3.2、2.3、0.7。综上所述,该消融实验证明了引入色彩相似性损失函数和教师-学生双向学习模块有助于提升模型的检测性能。

2.4 对比实验

为进一步验证本文方法的有效性,与当前主流的阴影-对象实例检测方法 LISA、SSIS、和 SSISv2 进行了对比,以上 3 种网络均为全监督阴影-对象实例检测网络。除 LISA 基于 Mask R-CNN 结构外,SSIS、SSISv2 及本文方法的特征提取模块均为 ResNeXt101-BiFPN。

表 2 展示了 5 种模型在 SOBA 数据集上的对比实验结果。本文提出的方法在所有评估指标上均高于 LISA 和通过在 SSISv2 中添加投影损失和色彩相似性损失实现弱

表 2 主流网络在 SOBA 数据集上的性能对比  
Table 2 Mainstream networks performance comparison on SOBA datasets

网络模型	Associate AP <sub>segm</sub>	Associate AP <sub>bbox</sub>	Instance AP <sub>segm</sub>	Instance AP <sub>bbox</sub>	SOAP <sub>segm</sub>	SOAP <sub>bbox</sub>
LISA <sup>F</sup>	40.9	48.4	39.2	37.6	23.5	21.9
SSIS <sup>F</sup>	52.2	59.6	43.4	41.3	30.2	27.1
SSISv2 <sup>F</sup>	<b>59.2</b>	<b>63.1</b>	<b>50.2</b>	<b>44.4</b>	<b>35.3</b>	<b>29.0</b>
SSISv2+ $\mathcal{L}^W$	50.4	58.3	42.3	40.5	29.5	26.6
Ours <sup>W</sup>	<b>51.5</b>	<b>58.7</b>	<b>42.9</b>	<b>40.8</b>	<b>29.8</b>	<b>26.9</b>

注:LISA、SSIS 和 SSISv2 右上角的 F 代表全监督实例分割,W 代表弱监督实例分割,SSISv2+ $\mathcal{L}$ 代表通过在 SSISv2 中添加投影损失和色彩相似性损失实现弱监督。

监督的网络, Associate AP<sub>segm</sub> 达到了 51.5(约为拥有最佳性能的全监督阴影-对象实例检测网络 SSISv2 的 87%),与 SSIS 性能接近,结果表明了本文方法能够在仅有边界框标签的弱监督下完成阴影-对象实例检测这一任务,并与全监督模型的精度相差不大。另一方面,因为 SSISv2 的性能明显优于 SSIS 和 LISA,故本文在 SOID 数据集上对比了 SSISv2 与本文方法,实验结果如图 6 所示,可以直观地观察到通过在 SOID 上训练有效地提高了弱监督与全监督的检测性能。且该实验结果表明,通过 SOID 数据集训练,有效地提高了模型的检测性能,并在 SOBA 的基础上进一步丰富了阴影-对象实例检测数据集的多样性。

为了能更直观地评价 LISA,SSIS,SSISv2 以及本文设计的网络,给出了以上 4 种方法在 SOID 数据集上预测的部分主观视觉对比结果,如图 7 所示。从图 7 中可以观察到,本文方法与 SSIS,SSISv2 都能较为出色的完成检测任

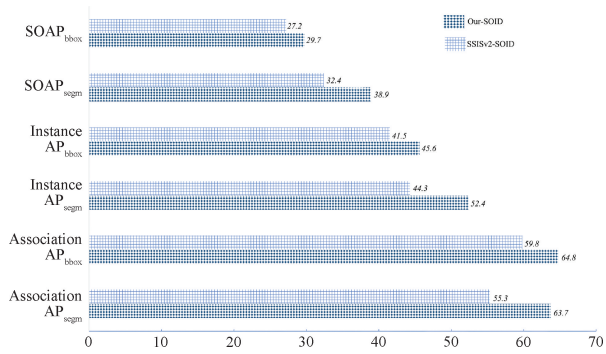


图 6 主流网络在 SOID 数据集上的性能对比

Fig. 6 Mainstream networks performance comparison on SOID

务,但是 LISA 在检测过程会出现漏检或误检情况。由此可见,本文所提方法能够生成较为精确的实例掩码完成阴影-对象实例检测任务。

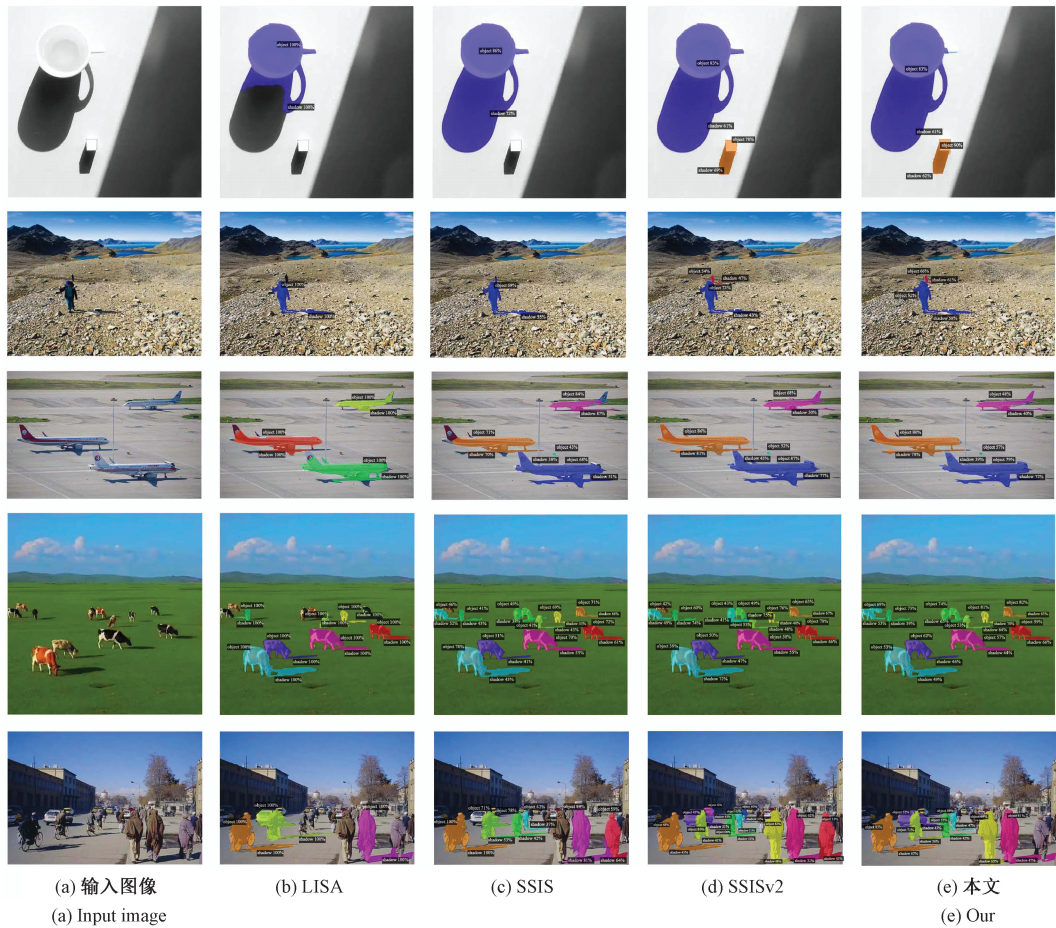


图 7 部分预测结果视觉对比

Fig. 7 Visual contrast experiment

3 结 论

本文提出了一个基于双向学习结构的弱监督阴影-对象实例检测网络,通过引入投影损失与色彩相似性损失捕

捉相邻像素间的色彩关联完成弱监督阴影-对象实例检测这一任务,并利用教师-学生双向学习结构提高了网络的检测性能,且因为教师网络与学生网络的结构相同,实现了兼顾网络复杂度的同时实现较好的分割结果。实验结果

表明,本文提出的方法在两个数据集上相关指标均有所提升,在 SOBA 中有关实例掩码的指标 Associate AP<sub>segm</sub> 和 SOAP<sub>segm</sub> 达到了 51.5 和 29.8,在 SOID 数据集上达到了 53.3 和 32.4,且通过主观视觉对比可以观察本文方法具有较好掩码检测能力,证明了本文方法的有效性。在未来的工作中,将进一步考虑构建大型阴影-对象数据集,利用大型数据集提高弱监督阴影-对象实例检测网络网络的泛化能力。同时,从本文方法中也可以看出色彩特征可以有效提高网络的检测性能,接下来将尝试引入合适的特征强化网络的学习能力。

## 参考文献

- [1] 李可新,何丽,刘哲凝,等. 基于跨模态特征融合的 RGB-D 显著性目标检测[J]. 国外电子测量技术, 2024,43(6):59-67.  
LI K X, HE L, LIU ZH N, et al. RGB-D salient object detection based on cross-modal feature fusion[J]. Foreign Electronic Measurement Technology, 2024, 43 (6): 59-67.
- [2] ZHANG X C, BARRON J T, TSAI Y T, et al. Portrait shadow manipulation[J]. ACM Transactions on Graphics(TOG), 2020, 39(4): 78: 1-14.
- [3] 陈元妹,王凤随,王路遥. 细化特征引导对抗性解纠缠学习的无监督行人重识别[J]. 电子测量与仪器学报, 2024,38(5):130-138.  
CHEN Y M, WANG F S, WANG L Y. Unsupervised person re-identification of adversarial disentangling learning guided by refined features[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(5):130-138.
- [4] 霍旭,盖绍彦,洪濡,等. 基于多阶段关联的多目标跟踪算法[J]. 仪器仪表学报, 2023,44(11):205-214.  
HUO X, GAI SH Y, HONG R, et al. Multi-object tracking algorithm based on multi-stage association[J]. Chinese Journal of Scientific Instrument, 2023, 44(11):205-214.
- [5] 杨雳,刘胜全,贾李睿智,等. 基于融合 LDA 与双层 CNN 的文本分类研究[J]. 电子测量技术, 2023, 46(7):1-6.  
YANG L, LIU SH Q, JIA L R ZH, et al. Research on text classification based on fusion of LDA and two-layer CNN[J]. Electronic Measurement Technology, 2023,46(7): 1-6.
- [6] KHAN S H, BENNAMOUN M, SOHEL F, et al. Automatic feature learning for robust shadow detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1931-1938.
- [7] LE H, VICENTE T F Y, NGUYEN V, et al. A+D Net: Training a shadow detector with adversarial shadow attenuation [C]. European Conference on Computer Vision(ECCV), 2018: 662-678.
- [8] HU X W, ZHU L, FU CH W, et al. Direction-aware spatial context features for shadow detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7454-7462.
- [9] WANG T Y, HU X W, WANG Q, et al. Instance shadow detection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1880-1889.
- [10] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]. IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [11] WANG T Y, HU X W, FU CH W, et al. Single-stage instance shadow detection with bidirectional relation learning [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:1-11.
- [12] TIAN ZH, SHEN CH H, CHEN H. Conditional convolutions for instance segmentation[C]. Computer Vision-ECCV 2020: 16th European Conference, Springer International Publishing, 2020: 282-298.
- [13] WANG T Y, HU X W, HENG P A, et al. Instance shadow detection with a single-stage detector [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(3): 3259-3273.
- [14] TIAN ZH, SHEN CH H, WANG X L, et al. Boxinst: High-performance instance segmentation with box annotations[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 5443-5452.
- [15] CHENG T H, WANG X G, CHEN SH Y, et al. Boxteacher: Exploring high-quality pseudo labels for weakly supervised instance segmentation[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 3145-3154.
- [16] TAN M X, PANG R M, LE Q V. Efficientdet: Scalable and efficient object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.
- [17] WANG X L, KONG T, SHEN CH H, et al. Solo: Segmenting objects by locations [C]. Computer Vision-ECCV 2020: 16th European Conference, Springer International Publishing, 2020: 649-665.
- [18] HOU L, VICENTE T F Y, HOAI M, et al. Large scale shadow annotation and detection using lazy annotation and stacked CNNs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(4): 1337-1351.



[19] WANG J F, LI X, YANG J. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 1788-1797.

[20] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context [C]. Computer Vision-ECCV 2014: 13th European Conference, Springer International Publishing, 2014: 740-755.

[21] GOYAL P, DOLLÁR P, GIRSHICK R, et al. Accurate, large minibatch SGD: Training imagenet in

1 hour[J]. ArXiv preprint arXiv:1706.02677, 2017.

作者简介

王 晓, 硕士研究生, 主要研究方向为图像处理与识别。  
E-mail: 1299554370@qq.com

刘 洪(通信作者), 博士, 副教授, 主要研究方向为信号处理、图像处理。  
E-mail: laniliu@sina.com

廖丽瑛, 硕士研究生, 主要研究方向为图像处理与识别。  
E-mail: 1571656196@qq.com

邹世奇, 硕士, 主要研究方向为图像处理与识别。  
E-mail: 1579964335@qq.com