

DOI:10.19651/j.cnki.emt.2313853

# 基于偏振特征与强度信息融合的工件目标检测\*

杨长虎<sup>1</sup> 袁东<sup>1</sup> 张健<sup>1</sup> 余洋洋<sup>1</sup> 张志良<sup>2</sup>

(1. 山东省水利勘测设计院 济南 250013; 2. 河海大学信息科学与工程学院 常州 213000)

**摘要:** 针对日前工件尺寸测量在工件关键测量点的准确提取阶段存在的问题,因此本文提出一种基于偏振特征与强度信息融合的工件目标检测方法。在工件强度图像的基础上引入偏振特征,建立强度信息与偏振特征差异化式高效交互的双流网络模型,以实现更高效的偏振特征和强度信息融合。为验证算法的效果,本文建立了偏振图像工件目标显著性检测数据集。在此数据集上,本文提出的算法在精确度 Precision、max-F 和相似性值 S-measure 三个指标上和视觉结果均在对比算法中达到了最优结果,充分表明了本文算法出色的工件目标检测性能,具有极佳的工件目标检测效果。

**关键词:** 偏振特征;目标检测;强度信息;双流网络;偏振角

**中图分类号:** TP2 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Workpiece target detection based on fusion of polarization features and intensity information

Yang Changhu<sup>1</sup> Yuan Dong<sup>1</sup> Zhang Jian<sup>1</sup> Yu Yangyang<sup>1</sup> Zhang Zhiliang<sup>2</sup>

(1. Shandong Water Conservancy Survey and Design Institute, Jinan 250013, China;

2. College of Information Science and Engineering, Hohai University, Changzhou 213000, China)

**Abstract:** In response to the current problems of dimensional measurement of workpieces at the stage of accurate extraction of key measurement points of the workpieces, this paper proposes a workpiece target detection method based on the fusion of polarization features and intensity information. The polarization features are introduced on the basis of the workpiece intensity image, and a dual-stream network model with differentiated and efficient interaction between intensity information and polarization features is established to achieve a more efficient fusion of polarization features and intensity information. To validate the algorithm's effectiveness, we have established a dataset for detecting the saliency of workpiece targets in polarization images. On this dataset, the proposed algorithm outperforms comparison algorithms in terms of Precision, max-F, S-measure, and visual results, underscoring its exceptional performance in workpiece target detection and its outstanding results in workpiece target detection.

**Keywords:** polarization imaging; target detection; intensity information; dual stream network; angle of polarization

## 0 引言

生产工件关键测量点的准确提取对提高未来工件尺寸测量精度具有重要意义。传统的显著性目标检测方法(salient object detection, SOD)通常利用手工设计的特征,以自上而下的模式预测显著性图,例如基于对比度、边界背景或基于中心先验等特征。传统的方法主要利用人类直觉感觉或者启发式先验,通过人工提取特征的方式来检测目标,然而人工提取特征是耗时耗力的。

近些年来,随着深度学习在计算机视觉诸多领域展现

了其优势,神经网络在显著性目标检测任务上也取得了很大的进展。Wang等<sup>[1]</sup>利用卷积神经网络融合局部空间特征和全局语义信息,整合局部估计结果和全局搜索结果作为最终的显著性预测结果。Liu等<sup>[2]</sup>设计了一个深度多层次网络来预测粗略的显著性图,然后对其进行分层的逐步细化。由于全卷积网络(fully convolutional network, FCN)在语义分割任务上取得了不错的效果,许多的研究者基于全卷积神经网络来实现效果更好的显著性目标检测。Wei等<sup>[3]</sup>提出了循环全卷积神经网络来迭代式的多次细化预测的显著性图。Hou等<sup>[4]</sup>在较浅的输出层和较深的输出

收稿日期:2023-06-07

\* 基金项目:济南市数字孪生与智慧水利重点实验室开放研究基金(37H2022KY040116)项目资助

层之间增加了多条跳跃连接,以提高预测显著性图的精度。尽管这些研究工作使得 SOD 算法效果更好,但也仍然存在一些问题,主要表现在:1)场景复杂、成像效果差的情况下,目标亮度不均匀、边缘模糊,显著性检测出现目标缺失或语义模糊等问题;2)单一模态信息的局限性,导致目标检测精度受限。

为了进一步提高显著性目标区域检测的完整度和准确度,研究人员提出引入额外辅助图像来提高 SOD 算法的精度。近些年来,偏振凭借其蕴含的丰富的物体表面光波特性和受到越来越多研究人员的关注,许多基于偏振信息的目标检测及分割方法被相继提出。Ren 等<sup>[5]</sup>建立了包含物体反射和杂波反射的偏振模型,根据偏振度的统计特性来实现目标检测任务。Xiang 等<sup>[6]</sup>设计了一种注意力桥接的特征融合模块,实现偏振特征与强度信息注意力级的融合,进而实现物体完整的检测分割。Fan 等<sup>[7]</sup>提出了一种利用偏振信息和彩色图像特性的融合 DPM 模型,利用偏振固有的光波特性提高目标检测算法准确性。

如何应对两种模态特征间的信息偏差问题,更好地实现不同模态的编码特征的高效融合,是引入辅助信息的显著性检测方法的主要的难题。本文在工件强度信息的基础上引入偏振特性,研究偏振特征与强度信息的高效融合方式,以提升工件目标检测的准确度和完整度。本文提出基于偏振特征与强度信息融合的工件目标检测方法,构建了

包含强度流和偏振流的双分支结构神经网络模型,以更加高效的方式实现强度信息与偏振特征的融合。提出的强度特征增强模块利用强度特征指导偏振特征流分辨相同偏振特性的物体,偏振特征增强模块则通过偏振特性丰富强度图像的语义信息和细节,充分利用强度信息和偏振特征在不同网络层次特征的优势,实现偏振和强度特征的高效流动融合和补充增强。最后利用稠密重建解码模块对增强特征进行解析重构,实现准确完整的工件目标区域检测。

### 1 网络结构

基于偏振特征与强度信息融合的工件目标检测网络的整体结构如图 1 所示。总体上,该网络的结构为经典的编-解码模型,主要包括两个编码器和一个解码器,分别为强度图像编码器,偏振图像编码器和共享重建解码器。其中,编码器中的强度特征增强模块(intensity feature enhancement module, IFE)、偏振特征增强模块(polarization feature enhancement module, PFE)和解码器中的稠密解码重建模块(dense decoding reconstruction module, DDR)为本模型中主要的 3 个模块。偏振特征与强度信息的融合与传统的融合方式不同,本文的特征融合是采用特征交互的方式,分别使用 IFE 和 PFE 两个增强模块,在强度和偏振特征流流动阶段,通过特征流交互的方式实现强度特征与偏振特征的相互融合与增强。

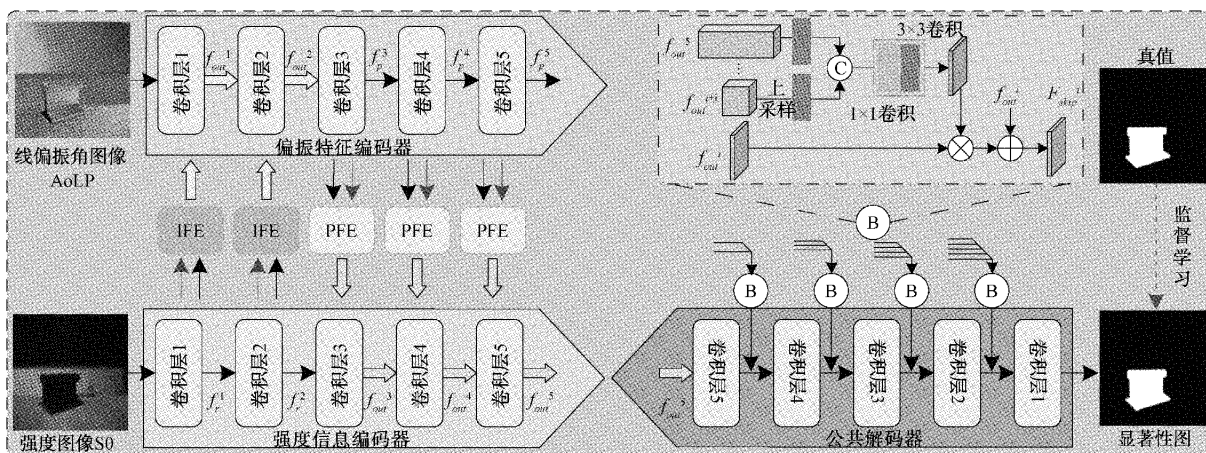


图 1 基于偏振特征与强度信息融合的工件目标检测网络结构图

两个分支的编码器均采用移除全连接层的预训练 VGG16 网络,用于提取强度图像和偏振特征图像的多层次深度特征表示以及用于未来偏振特征与强度的特征的整合交互。为了便于描述,使用  $f_s^i$  和  $f_p^i$  表示提取的强度特征和偏振特征,其中  $r$  和  $p$  分别表示强度图像分支和偏振特征分支,  $i \in \{1, 2, \dots, 5\}$  为对应特征层的索引。对于编码器中的低级特征(VGG16 网络中的前两层),利用 IFE 模块将细节纹理特征从强度特征空间传输至偏振特征空间,用于协助提高偏振特征中相似特征的可区分度。对于高级的编码器特征,则通过 PFE 模块的高级偏振类别特

征,协助强度分支捕获更清晰,更加细粒度的语义信息,抑制背景目标的干扰,提高显著性物体的预测准确度。此外,对于解码器中的上采样基本模块,DDR 模块的基于传统的跳跃连接进行了改进,利用更高层次的跳跃连接特征作为引导信息,实现更有效的编码器信息传输。最终解码器最后的卷积层对输入图像的显著性目标区域进行估计,重建对应的显著性结果。

### 2 强度特征增强模块

相比于强度图像,偏振特征可以更直观的描述显著性

目标的形状和位置。因此,对于包含更详细信息的低级编码器特征(如边界和形状),偏振特征可以提供比强度特征更直接和有指导意义的特征表示,有利于细节特征的学习。但是相同入射角的不同目标可能会具有相同的偏振特性,而且偏振特征图像包含较多的干扰噪声,导致在检测工件目标区域时困难。但是在对应的强度图像中,通常

可以通过颜色差异等视觉特征区分不同物体。因此,本节提出了强度特征增强模块(IFE),通过在网络浅层引入强度图像分支的细节特征指导,利用低层的强度特征来加强和补充偏振特征,实现在显著性预测过程中使用更多的信息来处理较难判断的感兴趣区域,进而提升预测精度。IFE 模块的详细结构如图 2(a)所示。

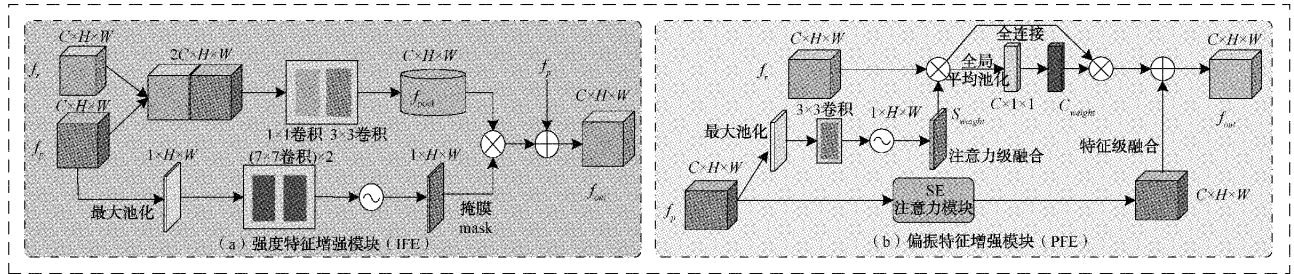


图 2 模块结构图

如图 2(a)所示,IFE 模块首先采用两个级联的卷积层来融合沿通道方向拼接的强度及偏振特征图,其中首个卷积层的卷积核是  $1 \times 1$  的,用于缩减特征图的通道数,聚合重要特征,降低计算负担;另一个卷积层的卷积核则为  $3 \times 3$ ,实现更全面的非线性特征融合,从而生成融合后的特征池  $f_{pool}$ :

$$f_{pool}^i = Conv_3(Conv_1([f_s^i, f_p^i])) \quad (1)$$

其中,  $i \in \{1, 2\}$  表示底层编码器特征层的索引,  $[\cdot, \cdot]$  表示特征图的通道级拼接操作,  $Conv_k$  表示卷积核为  $k \times k$  的卷积层。相比于直接将强度特征传递到偏振分支,在生成特征池  $f_{pool}$  的过程中,可以增强强度和偏振两种特征之间的共同细节特征,削弱对于预测无用的特征。

为了更高效地提供偏振编码器支路所需要的关键信息,需要进一步从偏振角度对强度特征进行筛选提取,利用最大池化、卷积和 Sigmoid 函数处理强度特征并生成空间注意力权重图。在两个串行的卷积层中,使用了更大的  $7 \times 7$  卷积核来感知大接受域中的重要细节区域。最后将空间注意力掩膜与特征池  $f_{pool}$  相乘,抑制无关的强度特征,从而获得补充的强度图像特征信息。IFE 模块的计算过程可以用式(2)表示。

$$f_{out}^i = \sigma(Conv_7(Conv_7(\maxpool(f_p^i)))) \odot f_{pool}^i + f_p^i \quad (2)$$

其中,  $\maxpool(\cdot)$  和  $\sigma(\cdot)$  分别表示沿通道维度的最大池化操作和 Sigmoid 激活函数,  $\odot$  代表元素级的逐个相乘操作,生成的增强特征  $f_{out}^i (i \in \{1, 2\})$  将会在偏振分支中被用于下一层的输入。为了更好的利用偏振分支中细节特征,通过跳跃连接,将其传输到解码器的前两层中,更好的利用偏振特征凸显目标形状等特性。

### 3 偏振特征增强模块

虽然编码器的深层通常可以学习更多的高级语义信息(如类别等),对于强度图像,由于包含丰富的颜色外观和内容,其语义信息也比偏振特征更加全面。但偏振特征

图的结构和数据特征相对简单,学习到的高级语义特征能够更好地实现显著物体的定位,抑制背景目标干扰。针对某些环境下目标对比度低的问题,强度图像往往会在阴暗处丢失纹理细节,而偏振可以自适应增强图像中的对比度,可以有效的保留目标信息。因此,偏振特征增强模块(PFE)被提出用于利用偏振特征丰富强度图像特征的语义信息。但是简单的特征融合策略(如相加或拼接)可能会存在信息丢失、固定权重、计算成本高昂、过拟合风险等问题,在某些情况下可能会限制模型的性能和灵活性。考虑到简单的特征融合策略不能高效地融合特征,PFE 模块使用通道注意力级和特征级两种整合交互模式,能够更加全面地进行强度图像和偏振图像的特征融合。PFE 模块的详细结构如图 2(b)所示。

如图 2(b)所示,首先根据偏振特征学习注意力向量  $S_w \in R^{1 \times H \times W}$ ,以空间注意力的方式指导强度特征感兴趣区域,其中  $H$  和  $W$  分别代表特征图的高度和宽度。使用空间注意力的方式,不仅有助于加强已经提取到的突出特征区域,还能够矫正强度图像特征中被忽略或被错误关注的特征。该过程用公式表述如式(3)所示。

$$\begin{cases} S_w = \sigma(Conv_3(\maxpool(f_p^i))) \\ f_{rs}^i = S_w \odot f_s^i \end{cases} \quad (3)$$

其中,  $f_p^i$  代表偏振分支中的编码器高层特征,  $f_s^i$  代表强度图像分支提取的高层强度特征,  $f_{rs}^i$  表示由偏振特征  $f_p^i$  的空间注意力增强后的强度图像编码器特征,  $i \in \{3, 4, 5\}$  为编码器层索引。此外,为了减少高层特征中的冗余通道,在使用注意力级融合策略后,网络还使用了通道注意力模块来建模不同通道之间的相关性,学习更有用的差异性特征,进行特征级融合。通道注意力模块主要包括一个全局平均池化层,两个全连接层(fully connected layer, FC)和 Sigmoid 激活函数,用于学习通道注意力权重向量  $C_w \in R^{C \times 1 \times 1}$ ,其中  $C$  即代表特征图的通道数。

PFE 中的注意力级的指导操作使用公式表示如式(4)所示,其中  $D_{add}^i$  即为注意力级的强度增强特征。

$$\begin{cases} C_w = \sigma(FC(GAP(f_{rs}^i))) \\ D_{add}^i = C_w \odot f_{rs}^i \end{cases} \quad (4)$$

对于 PFE 模块中在特征层面上的引导,使用像素级的加法操作直接融合两种特征,可以加强显著对象的内部响应,获得更好的内部一致性。其中特征层面增强特征  $D_{add}^i$  是使用级联注意力和空间注意力机制增强偏振特征后生成的。因此,最终到达强度图像分支的下一层网络的特征可以使用公式表述为:

$$f_{out}^i = D_{add}^i + D_{add}^i \quad (5)$$

最后强度图像分支中的增强特征  $f_{out}^i (i \in \{3,4,5\})$  将会被引入解码器阶段,用于实现显著性目标检测结果的

解码重建。

## 4 实验结果与分析

### 4.1 数据集与评价指标

#### 1) 偏振显著性检测数据集

偏振图像被广泛应用在语义分割<sup>[8-9]</sup>、图像复原<sup>[10-12]</sup>和目标检测<sup>[13-14]</sup>等计算机视觉任务中,但是相关的数据集<sup>[15]</sup>却十分缺乏,尤其是显著性目标检测,暂时还没有公开的数据集。本文建立了偏振图像工件目标显著性检测数据集。数据集图像使用 LUCID Vision 实验室研发的 TRI050S-QC 分焦平面彩色偏振相机采集,包含多种不同材料、形状和大小的工件目标。如图 3 所示为工件显著性目标检测数据集部分样本示例图。

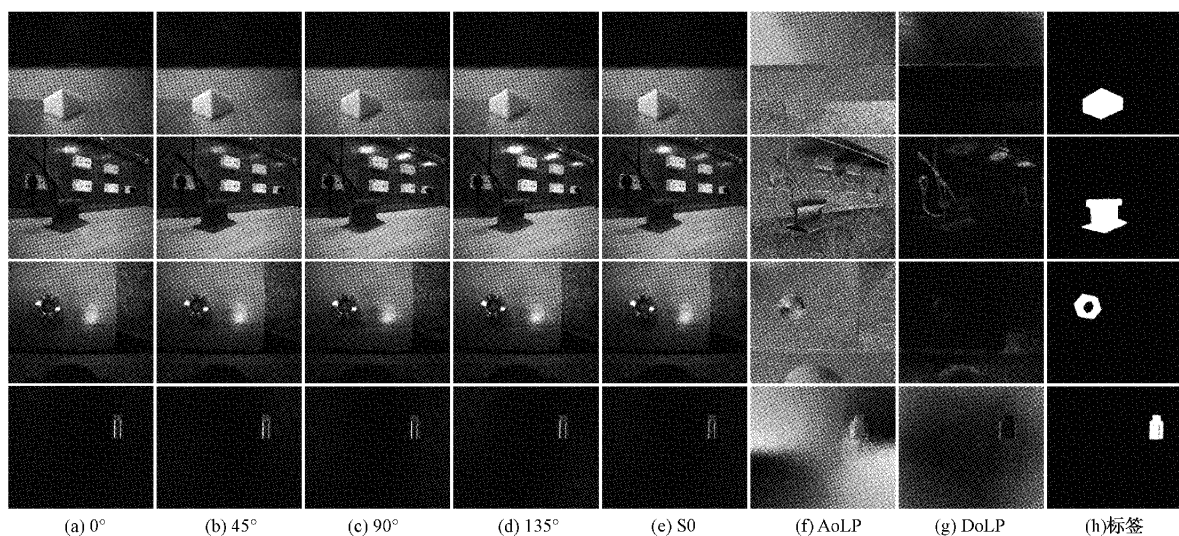


图 3 偏振图像工件目标检测数据集

如图 3 所示,偏振显著性检测数据集中每组图像共 8 张图像,由 3 个部分组成,第 1 部分为偏振角度分别为  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$  和  $135^\circ$  时采集的偏振图像,由偏振探测仪直接捕获,如图 3(a)~(d)所示,均为 3 通道 RGB 彩色图像。此外自然光经过反射和折射后,偏振光中线偏振光分量为主要成分,其中的圆偏振光通常可以忽略不计<sup>[16]</sup>,因此可以利用前文的斯托克斯参数计算公式估计描述该线偏振光偏振态的斯托克斯参数。

如图 3 所示,第 2 部分包括根据 4 个角度的偏振图像计算获得的偏振特性斯托克斯矢量:强度图像(S0)、线偏振角(AoLP)图像和线偏振度(DoLP)图像,如图 3(e)~(g)所示,其中 S0 为 3 通道彩色图像,AoLP 和 DoLP 为单通道灰度图像。最后一部分为强度图像对应的显著性标签真值,为二值图像,取值范围为 0 或 1,表示该像素点是否为显著性目标区域。显著性目标真值标签图像是使用 labelme 工具标注所得,为了得到较准确的显著性目标真值标签,本文采取了多次标注求平均的方法,对每张图像前后多次标注,然后对多张标注结果取公共显著性区域

获取最终标注结果。如图 4 所示,其为显著性标签标注工具 labelme 的软件界面,最终建立偏振图像工件目标显著性检测数据集共包含 3 360 张图像,420 组数据。

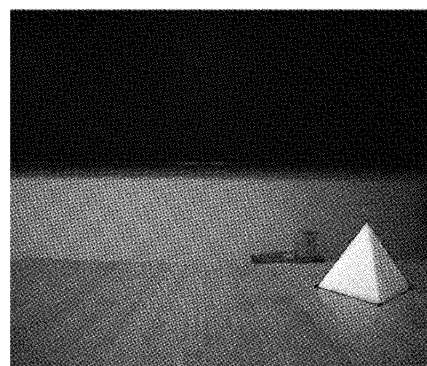


图 4 标注工具界面

#### 2) 评价指标

为了客观的评估算法的有效性,本文选择了平均绝对误差<sup>[17]</sup>(mean absolute error, MAE)、精确率(Precision)、

召回率 (Recall)、F-measure<sup>[18]</sup> 和结构相似性指标<sup>[19]</sup> (S-measure) 多个主流的评价指标, 从不同的角度来全面验证算法的显著性目标检测效果, 其中 Precision、Recall、F-measure 和 S-measure 均为越大越好, 而 MAE 越小, 表明算法效果越好。

通过计算显著性预测结果与对应真值标签间的平均像素绝对差可以获得 MAE<sup>[17]</sup> 指标, 直接反映了显著性结果与标签的相似程度, 显著性结果越接近真值, MAE 越小, 则对应算法效果越好。其计算公式如式(6)所示。

$$MAE = \frac{1}{M \times N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |\hat{S}(m, n) - S(m, n)| \quad (6)$$

其中,  $M$  和  $N$  分别代表图像的长和宽,  $\hat{S}(m, n)$  和  $S(m, n)$  分别表示在像素  $(m, n)$  处的显著性预测值和标签值。

Precision 与 Recall 可以通过二值化显著性预测结果, 再与实际真值比较计算得到, 又被称为查准率与查全率, 其计算公式如式(7)所示。

$$Precision = \frac{n_{TP}}{n_{TP} + n_{FP}}, Recall = \frac{n_{TP}}{n_{TP} + n_{FN}} \quad (7)$$

其中,  $n_{TP}$  表示被模型预测为真的正样本(真阳性)的数量;  $n_{FP}$  表示被模型预测为假的正样本(假阳性)的数量;  $n_{FN}$  代表被模型预测为假的负样本(假阴性)的数量。综合考虑精确率和召回率, 通过调和平均的计算方式, 可得到 F-measure 指标的结果, 计算公式如下:

$$F = \frac{(1 + \beta^2) \times Precision \times Recall}{\beta^2 \times Precision + Recall} \quad (8)$$

式(8)中,  $\beta^2$  为调和系数, 用于平衡精确率和召回率所占的比重, 通常为 0.3。可以看到, 相比于精确率和召回率, F-measure 能更全面的评价算法的效果。由于不同阈值下, 可以计算多个 F-measure 结果, 本节取最大 F 值 max F-measure(max-F) 作为评价指标。

S-measure 通过结合显著性预测结果和真值之间的区域感知相似性和目标感知相似性来确定预测结果的准确度, 计算公式如式(9)。

$$S_\alpha = \alpha \times S_{object} + (1 - \alpha) \times S_{region} \quad (9)$$

其中,  $S_{object}$  和  $S_{region}$  分别为目标感知相似度和区域相似感知度,  $\alpha$  为二者的平衡系数, 通常取 0.5。

#### 4.2 实验参数设置

本文网络模型利用 Python 编程语言和 PyTorch 深度

学习框架搭建, 使用 24 G NVIDIA RTX 3090 显卡加速模型训练、测试及相关实验实施。所有的训练和测试图像的尺寸均被调整为 3 通道  $608 \times 512$  空间分辨率的大小, 对应的偏振信息图像 ( $DoLP$  和  $AoLP$ ) 则通过通道复制, 由单通道图像转变为 3 通道图像。训练开始时模型初始参数使用在 ImageNet 预训练完成的 VGG16 的权重参数, 其他的参数则根据 PyTorch 的默认设置初始化。为了避免模型过拟合, 随机翻转和旋转的方法被用于提高训练样本的多样性, 提升模型的泛化性能。训练过程中, 本文使用显著性目标检测常用的二进制交叉熵损失计算模型代价损失, 并使用 ADAM 优化算法来优化模型参数收敛至最优解。训练 BatchSize 为 4, 初始学习率为  $1 \times 10^{-4}$ , 采用可变学习率的方法, 每迭代 50 次学习率衰减 1/5, 模型在数据集上训练总迭代次数为 200 次。

#### 4.3 实验结果分析

为了验证本文提出的工件目标检测算法的有效性, 本文与前沿的多种基于深度学习的显著性目标检测算法进行了定性和定量的比较, 对比方法包括 BASNet<sup>[20]</sup>、U2Net<sup>[21]</sup>、CPD<sup>[22]</sup>、RAS<sup>[23]</sup>、CTDNet<sup>[24]</sup>、EDN<sup>[25]</sup>、F3Net<sup>[26]</sup>、PFSNet<sup>[27]</sup> 及 PiCANet<sup>[28]</sup> 算法。为了保证评估指标的公平性, 对比算法均在默认设置下使用已发布的代码测试, 获得显著性图。在求得所有算法的显著性预测图后, 所有指标都在同一度量指标计算代码下评估检测精度。

##### 1) 定量评估指标分析

如表 1 所示, 其展示了不同对比算法在 MAE、Precision、Recall、max-F 和 S-measure 五种客观评价指标上的结果, 其中最优结果加粗标注。可以看到, 本文算法在精确度 Precision、max-F 和相似性值 S-measure 三个指标上均达到了最优结果, 分别高于次优算法 RAS<sup>[23]</sup> 为 0.006 2、0.002 及 0.004 6, 主要在于本文引入的偏振信息能够很好地提取显著性目标的空间位置信息, 提出的偏振特征增强模块则可以有效地增强边缘细节特征, 削弱干扰无用信息, 最大化利用偏振特征和强度信息的优势, 提高显著性目标检测的准确度。在平均绝对误差 MAE 和召回率 Recall 指标上, 本文算法稍微低于 RAS<sup>[23]</sup> 和 U2Net<sup>[21]</sup> 算法, 但与最优结果差距很小, 也达到了令人满意的效果。因此定量实验结果充分表明了本文算法出色的工件目标检测性能。

表 1 不同对比算法的定量指标结果

方法	BASNet	CPD	CTDNet	EDN	F3Net	PFSNet	PiCANet	U2Net	RAS	本文算法
MAE(↓)	0.038 7	0.017 6	0.026 8	0.018 6	0.048 5	0.018 0	0.049 0	0.018 1	<b>0.010 9</b>	0.011 6
Precision(↑)	0.846 8	0.888 9	0.897 1	0.897 9	0.724 6	0.930 6	0.841 0	0.903 1	0.935 2	<b>0.941 4</b>
Recall(↑)	0.949 4	0.949 4	0.946 0	0.940 5	0.931 2	0.944 8	0.760 1	<b>0.953 9</b>	0.939 0	0.952 0
max-F(↑)	0.868 5	0.902 1	0.907 9	0.907 4	0.769 7	0.938 3	0.820 8	0.911 9	0.938 4	<b>0.940 4</b>
S-measure(↑)	0.910 9	0.932 3	0.939 5	0.934 0	0.856 7	0.953 9	0.852 3	0.938 4	0.956 1	<b>0.960 7</b>

2)定性视觉效果分析

图5~7展示了本文所提算法与其他对比算法在工件

目标检测任务上的视觉效果对比,包括多种不同的基于深度学习

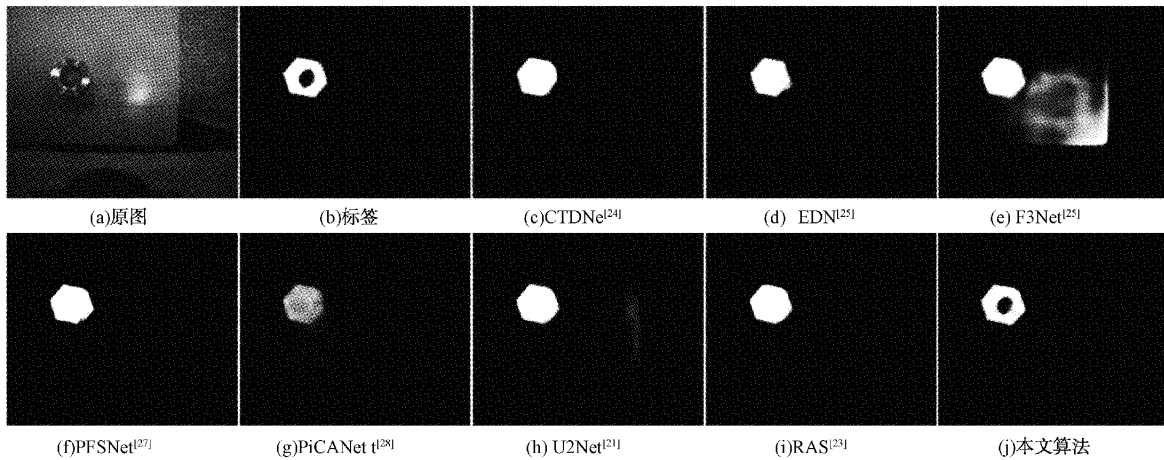


图5 工件目标检测结果图 I

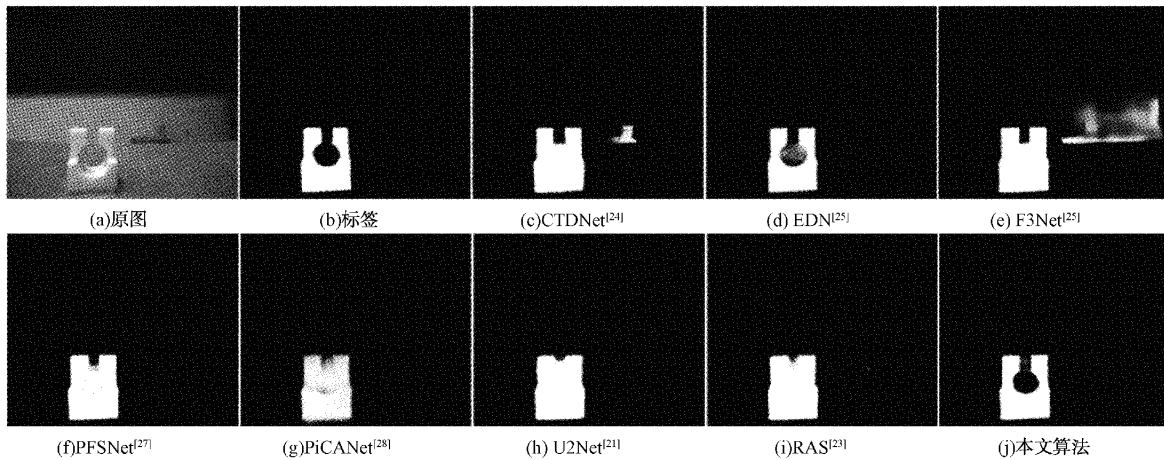


图6 工件目标检测结果图 II

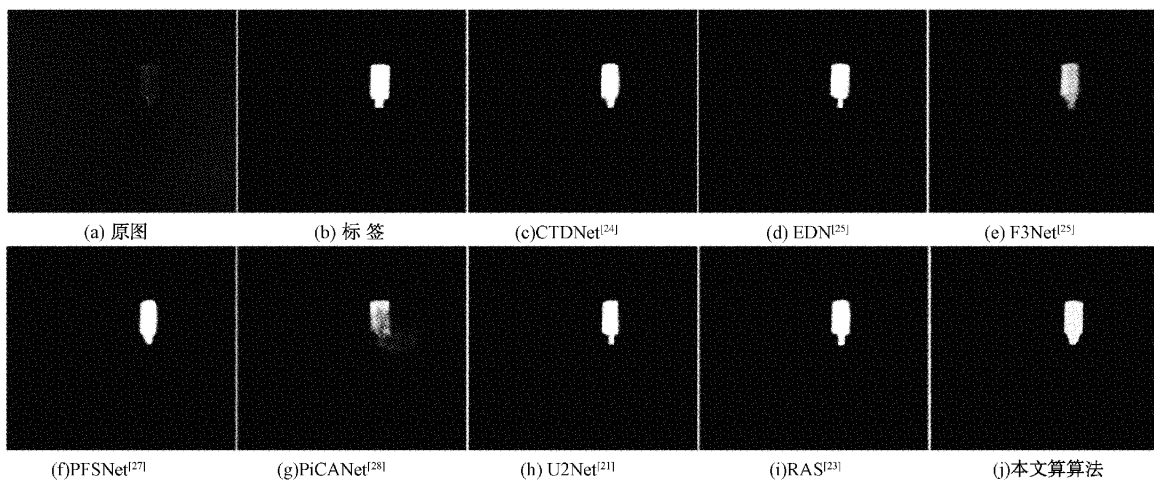


图7 工件目标检测结果图 III

如图5~7所示,可以看到本文提出的工件目标检测方法具有更强的显著性目标信息捕捉能力,F3Net<sup>[26]</sup>、

PiCANet<sup>[28]</sup>以及 U2Net<sup>[21]</sup>不能完全准确地检测到显著性目标,结果中存在错误的显著性区域。CTDNet<sup>[24]</sup>、

EDN<sup>[25]</sup>、PFSNet<sup>[27]</sup>以及RAS<sup>[23]</sup>方法虽然能够准确的检测显著性目标,但是它们对于显著性目标的边界轮廓等细节部分的处理能力较差,与标签存在较大差异。而本文的方法不仅能准确的预测显著性目标,而且对显著性目标的边缘检测也相当不错,具有最佳的工件目标检测效果,这是因为本文算法使用了偏振特征与强度信息的交互增强,由于工件目标与背景不同的材质特征,导致二者的偏振特征也不同,进而可以利用偏振特征加强了模型对边界细节特征的提取能力,提高了工件目标检测的准确度和边界完整度。本文提出的工件目标检测方法由于引入了工件的偏振特征,为算法提供了更丰富的物理先验知识和目标信息,使得本文算法在面对不同形状的工件目标时,也可以实现较准确地检测显著性目标。综上所述,定性和定量实验均证明了本文提出的基于偏振特征与强度信息融合的工件目标显著性检测算法对于普通目标、形状结构较复杂的工件目标,均能实现满意的工件目标显著性检测效果,在本文的应用场景中具有很好的效果。

#### 4.4 偏振信息消融实验分析

场景中显著性目标的偏振信息包括线偏振度(DoLP)和线偏振角(AoLP)两种,其中DoLP的值范围为 $[0,1]$ ,代表线性偏振光的偏振成分所占的比重,光波中偏振分量越多,则对应DoLP值越大。与DoLP不同,AoLP表示的是图像中某点的偏振角度,其变化范围为 $0^\circ \sim 180^\circ$ ,更重要的是,AoLP可以反映显著性目标的轮廓信息,因为相同材质的物体通常具有相近的AoLP值。因此,AoLP图像可以看作是显著性目标的掩膜图像,对于完整准确地检测目标具有重要作用。

如图3所示,偏振图像显著性检测数据集中,DoLP图像中大部分区域具有较小的值,只有部分极少的区域的DoLP值才较大,导致可视化的DoLP图像整体亮度偏低,也直接导致DoLP只能提供有限的偏振特性。相反,对于AoLP,可以发现同一显著性物体的区域基本显示出相近的AoLP值,即同一目标具有基本的偏振信息连续性。AoLP图像提供了更好的空间信息表示,它在同一物体或材料上保持一致的分布,可以为显著性检测任务提供有用的空间先验信息。

如图8所示,展示了训练集上DoLP和AoLP值的统计分布图,可以看到,AoLP的分布与DoLP有着显著差异。训练集中大部分像素点的DoLP值都很小,范围在0到0.4之间,而DoLP值大于0.4的像素点的比例很低,说明DoLP提供的信息有限,只能提供部分高偏振度的相关特征。与DoLP的分布不同,AoLP的分布则是均匀的,也就是说,AoLP图像中几乎所有的像素都能提供对于显著性检测有意义的特征。由于DoLP和AoLP图像的分布不同,可以预测它们提供的偏振信息侧重于不同的偏振特性。

基于上述观测,本小节进行了消融实验来研究引入偏

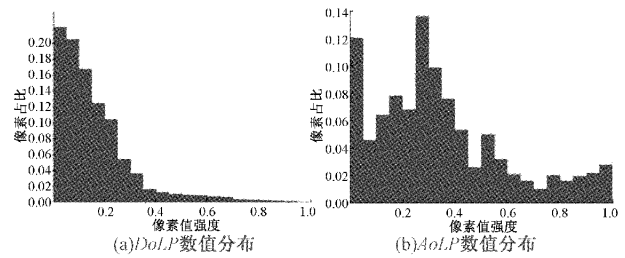


图8 训练集DoLP和AoLP图像数值统计分布图

振特性与强度信息融合方法的有效性,以及AoLP在为显著性检测提供互补信息上是否优于DoLP。如图1所示,本文提出的基于偏振特征与强度信息融合的工件目标显著性检测方法是双支路的网络结构,因此,基础模型(BaseLine)的构建是通过移除偏振信息支路,仅保留强度图像支路得到的,基础模型的训练也仅使用强度图像S0完成。然后,搭建了4个对比模型来进行比较验证。为了探索更有利于检测的偏振特征,分别选择AoLP和DoLP作为输入偏振信息,建立Model-AoLP和Model-DoLP,其中Model-AoLP即为本文算法模型。另外考虑到AoLP和DoLP都能提供偏振信息,还将AoLP和DoLP图像沿通道进行拼接,使用混合的偏振表示图像进行训练,得到网络Model-A|D。最后,Model-3Path的建立则是通过扩展本文模型为3支路得到,3条支路分别输入强度图像、AoLP图像和DoLP图像。模型其他部分均保持不变,训练测试等超参数设置也保持一致。

偏振信息的定量消融实验结果如表2所示。可以看到,Model-AoLP是最佳的偏振信息引入策略,尽管其在MAE和Recall上略低于Model-A|D,但是在其他的定量指标上都优于其他模型,整体上达到了最优的效果。对比BaseLine模型和Model-AoLP、Model-DoLP和Model-A|D,可以看到,偏振信息的引入确实有效地提高了显著性目标检测效果。对比Model-AoLP和Model-DoLP,前者有着比后者更好的显著性检测效果。通过对AoLP和DoLP数值分布的分析,AoLP比DoLP提供了更好的空间表示,如物体轮廓信息,而DoLP只能提供高偏振区域的有意义的信息,因此从该角度上说,AoLP图像为偏振图像显著性检测提供了更丰富的先验和互补信息。

此外,如表2所示,Model-A|D达到了与Model-AoLP相同的效果,但是AoLP和DoLP某些像素处偏振信息不同,对显著性图的预测造成了一定的干扰,影响了部分客观指标。Model-3Path是效果最差的模型,是因为该模型的复杂结构影响了模型利用最具信息量的特征,而且3条支路的结构削弱了强度图像特征的代表能力,而强度图像对于在颜色、形状等信息感知上具有重要的作用。最后,从上述定量分析中可以得出结论,在显著性目标检测中引入AoLP偏振特性图像,能够最佳地提高本文算法的工件目标显著性检测的效果。如图9所示,其展示

表2 偏振信息消融实验定量结果

模型	MAE (↓)	Precision (↑)	Recall (↑)	max-F (↑)	S-measure (↑)
基础模型	0.037 4	0.851 4	0.875 5	0.879 6	0.834 0
Model-DoLP	0.031 2	0.902 1	0.915 7	0.890 8	0.910 6
Model-AoLP	0.011 2	<b>0.941 4</b>	0.952 0	<b>0.940 4</b>	<b>0.960 7</b>
Model-A   D	<b>0.010 7</b>	0.929 5	<b>0.944 5</b>	0.936 4	0.949 2
Model-3Path	0.038 5	0.846 5	0.867 1	0.871 0	0.826 9

了不同消融模型的视觉显著性结果图。可以看到,主观视觉效果图与定量指标相符,引入偏振信息确实可以提高工件目标检测的准确度,而且本文提出的算法能够更精确地检测目标的轮廓边界细节信息,更完整准确地实现工件目标的检测提取。

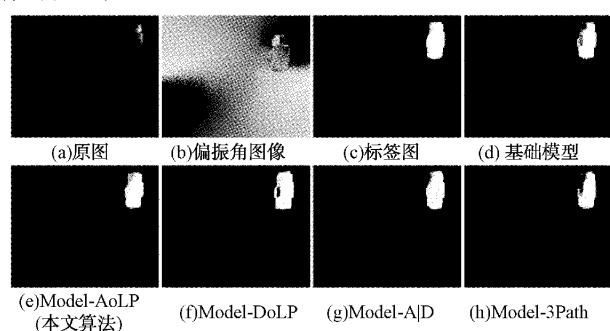


图9 不同消融模型的工件目标检测结果图

## 5 结 论

本文提出了基于偏振特征与强度信息融合的工件目标检测算法,搭建了双分支结构的编-解码网络模型,综合利用强度信息和偏振特征的高、低层特征优势,实现更准确的工件目标检测。提出的偏振特征增强模块和强度图像特征模块能够根据不同网络层次特征的特点,显式地建模特征提取阶段两种特征的依赖关系,利用偏振特征补充强度特征细节,丰富特征图的语义信息和局部细节,充分利用偏振特征和强度图像的信息,高效地实现了强度图像与偏振特征的交互融合,提高显著性目标检测结果的准确度和完整度。丰富的定量与定性结果验证了所提算法可以提高工件目标检测的准确度和完整度,对提高工件关键尺寸测量精度具有重要意义。

## 参考文献

- [1] WANG L, LU H, RUAN X, et al. Deep networks for saliency detection via local estimation and global search[C]. IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2015: 3183-3192.
- [2] LIU N, HAN J. Dhsnet: Deep hierarchical saliency network for salient object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2016: 678-686.
- [3] WEI J, ZHONG B. Saliency detection using fully convolutional network[C]. 2018 Chinese Automation Congress(CAC), IEEE, 2018: 3902-3907.
- [4] HOU Q, CHENG M M, HU X, et al. Deeply supervised salient object detection with short connections[C]. IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017: 3203-3212.
- [5] REN B, SHI L, WANG G. Polarimetric target detection using statistic of the degree of polarization[J]. Progress In Electromagnetics Research M, 2016, 46: 143-152.
- [6] XIANG K, YANG K, WANG K. Polarization-driven semantic segmentation via efficient attention-bridged fusion[J]. Optics Express, 2020, 29(4): 4802-4820.
- [7] FAN W, AINOUS S, MERIAUDEAU F, et al. Polarization-based car detection [C]. IEEE International Conference on Image Processing(ICIP), 2018: 3069-3073.
- [8] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [9] HUANG Z, WANG X, HUANG L, et al. Ccnet: Criss-cross attention for semantic segmentation[C]. IEEE/CVF International Conference on Computer Vision(ICCV), 2019: 603-612.
- [10] ZHANG F, CHEN Y, LI Z, et al. Acfnnet: Attentional class feature network for semantic segmentation[C]. IEEE/CVF International Conference on Computer Vision(ICCV), 2019: 6798-6807.
- [11] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 3146-3154.
- [12] ZHAO H, JIA J, KOLTUN V. Exploring self-attention for image recognition [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2020: 10076-10085.
- [13] PENG G, ZHANG J, LI H, et al. Attentional pyramid pooling of salient visual residuals for place recognition[C]. IEEE/CVF International Conference



- on Computer Vision(ICCV), 2021: 885-894.
- [14] DAI R, DAS S, MINCIULLO L, et al. Pdan: Pyramid dilated attention network for action detection[C]. IEEE/CVF Winter Conference on Applications of Computer Vision(WACV), 2021: 2970-2979.
- [15] XIANG K, YANG K, WANG K. Polarization-driven semantic segmentation via efficient attention-bridged fusion[J]. Optics Express, 2020, 29(4): 4802-4820.
- [16] KALRA A, TAAMAZYAN V, RAO S K, et al. Deep polarization cues for transparent object segmentation [ C ]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 8602-8611.
- [17] SHEN L, ZHAO Y. Underwater image enhancement based on polarization imaging [J]. The International Archives of Photogrammetry, Remote Sensing & Spatial Information Sciences, 2020, 43: 579-585. doi: 10.5194/isprs-archives-XLIII-B1-2020-579-2020.
- [18] REN Q, XIANG Y, WANG G, et al. The underwater polarization dehazing imaging with a lightweight convolutional neural network[J]. Optik, 2022, 251: 168381-168381. doi: 10.1016/j.ijleo.2021.168381.
- [19] ZHU Y, ZENG T, LIU K, et al. Full scene underwater imaging with polarization and an untrained network [ J ]. Optics Express, 2021, 29 ( 25 ): 41865-41881.
- [20] FAN W, AINOUS S, MERIAUDEAU F, et al. Polarization-based car detection [ C ]. IEEE International Conference on Image Processing(ICIP), 2018: 3069-3073.
- [21] BLIN R, AINOUS S, CANU S, et al. Road scenes analysis in adverse weather conditions by polarization-encoded images and adapted deep learning[C]. IEEE Intelligent Transportation Systems Conference (ITSC), IEEE, 2019: 27-32.
- [22] BLIN R, AINOUS S, CANU S, et al. A new multimodal rgb and polarimetric image dataset for road scenes analysis [ C ]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 216-217.
- [23] LEI C, HUANG X, ZHANG M, et al. Polarized reflection removal with perfect alignment in the wild[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2020: 1750-1758.
- [24] CONG R, LEI J, FU H, et al. Review of visual saliency detection with comprehensive information[J]. IEEE Transactions on Circuits Systems for Video Technology, 2018, 29(10): 2941-2959.
- [25] XU X, CHEN J, ZHANG H, et al. Sa-dpnet: Structure-aware dual pyramid network for salient object detection[J]. Pattern Recognition, 2022, 127: 108624-108624.
- [26] FAN D P, CHENG M M, LIU Y, et al. Structure-measure: A new way to evaluate foreground maps[C]. IEEE/CVF International Conference on Computer Vision(ICCV), 2017: 4548-4557.
- [27] QIN X, ZHANG Z, HUANG C, et al. Basnet: Boundary-aware salient object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2019: 7479-7489.
- [28] QIN X, ZHANG Z, HUANG C, et al. U2-net: Going deeper with nested u-structure for salient object detection [ J ]. Pattern Recognition, 2020, 106: 107404-107404.

### 作者简介

杨长虎,高级工程师,主要研究方向为信息感知与图像处理。

E-mail:798592226@qq.com

袁东,高级工程师,主要研究方向为水利信息感知与智能化处理。

E-mail:1031489560@qq.com

张健,工程师,主要研究方向为水利信息采集与传输。

E-mail:593279188@qq.com

余洋洋,工程师,主要研究方向为深度学习及其在水利图像中的应用。

E-mail:2521347377@qq.com

张志良,硕士研究生,主要研究方向为图像获取与处理等。

E-mail:zhangzl@hhu.edu.cn