

## 基于SOLOv2-RS的人工假体视觉避障研究

鄂宁<sup>1</sup>, 王静<sup>1,2</sup>, 周翔龙<sup>1</sup>, 赵容锋<sup>1</sup>, 何海洋<sup>3</sup>

1. 上海海洋大学信息学院, 上海 201306; 2. 农业农村部渔业信息重点实验室, 上海 201306; 3. 上海商汤智能科技有限公司, 上海 200231

**【摘要】**面向人工假体视觉条件下的避障问题,提出改进的实例分割模型SOLOv2-RS,为植入者在低分辨率人工视觉中更准确地感知导航任务的相关实例对象提供基础。根据视觉注意力机制,采取视野中心距离和目标尺度作为各实例重要性计算准则,以得到的重要性分数作为对需规避障碍物进行分级表达的依据;同时,采用边缘信息提示盲道,并对其进行形态学膨胀处理以避免光幻视有限导致的边缘信息缺失。人工假体视觉仿真结果表明本研究提出的人工假体视觉分级优化处理策略能有效实现盲道和障碍物的优化表达,为植入者更高效地完成室外避障任务提供便利,为人工假体视觉设备图像处理研究提供良好思路。

**【关键词】**视觉假体;避障;SOLOv2-RS;ResNeSt;分级优化处理

**【中图分类号】**R318

**【文献标志码】**A

**【文章编号】**1005-202X(2024)03-0309-07

### Obstacle avoidance in simulated prosthetic vision based on SOLOv2-RS

E Ning<sup>1</sup>, WANG Jing<sup>1,2</sup>, ZHOU Xianglong<sup>1</sup>, ZHAO Rongfeng<sup>1</sup>, HE Haiyang<sup>3</sup>

1. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Fisheries Information, Ministry of Agriculture and Rural Affairs, Shanghai 201306, China; 3. Shanghai SenseTime Intelligent Technology Co., Ltd., Shanghai 200231, China

**Abstract:** Aiming at the obstacle avoidance in simulated prosthetic vision, an improved instance segmentation model SOLOv2-RS is proposed for providing a basis for implant recipients to accurately perceive the relevant instance objects of navigation tasks in low-resolution prosthetic vision. According to the visual attention mechanism, the distance from the center of the visual field and the target scale are adopted as the importance calculation criteria for each instance, and the obtained importance score is used as the basis for the hierarchical representation of the obstacles to be avoided. Meanwhile, edge information is used to cue the tactile paving, and it is morphologically inflated for avoiding the edge information loss caused by the limited phosphene. The prosthetic vision simulation results demonstrate that the hierarchical optimization processing strategy for simulated prosthetic vision can effectively achieve the optimal representation of tactile paving and obstacles, thus facilitating the implant recipients to accomplish outdoor obstacle avoidance tasks more efficiently, and providing ideas for the research on the image processing of visual prosthetic devices.

**Keywords:** visual prosthesis; obstacle avoidance; SOLOv2-RS; ResNeSt; hierarchical optimization processing

### 前言

视觉假体作为一种视觉功能修复手段,对于一些疾病导致的失明患者,有着一定程度恢复视觉的

功能,其中典型的疾病是原发性老年黄斑变性和视网膜色素变性<sup>[1]</sup>。尽管视觉假体设备可以通过植入失明患者体内的微电极阵列刺激视觉通路上完好的部分,从而帮助植入者恢复部分的视觉感知,但是这种电刺激产生的视觉感知受限于植入电极的数目和视野范围,同时由于缺少大量重要的视觉特征信息,植入者完成基本的视觉任务存在一定的难度。为优化视觉感知,在刺激前通过图像处理策略尽可能多提取任务场景中的有效视觉信息,从而改善假体植入者的使用感,这是当前仿真假体视觉研究中的重要方向之一。在视障人士的日常任务中,室外任务

**【收稿日期】**2023-10-26

**【基金项目】**国家自然科学基金(61806123)

**【作者简介】**鄂宁,研究方向:计算机视觉、医学图像处理, E-mail: 2152127@st.shou.edu.cn

**【通信作者】**王静,博士,副教授,研究方向:计算机视觉理论及应用、医学图像处理, E-mail: wangjing@shou.edu.cn

中的核心问题就是导航和避障。对于导航任务, Dagnelie等<sup>[2]</sup>通过虚实场景下的仿真实验探究满足导航任务需要的最小图像分辨率; Vergnecx等<sup>[3]</sup>寻路任务的研究表明在少量微电极产生的低分辨率的图像下,提取到的环境边缘可以满足被试在陌生环境中的基本导航能力。对于避障任务, Parikh等<sup>[4]</sup>评估一种基于显著性检测的线索算法,该算法在移动和搜索等应用场景下为假体植入者提供一定程度的帮助<sup>[5]</sup>,其仿真实验表明该算法具有在陌生环境下检测物体的能力。此外,不管是室内还是室外,场景识别均是视障人士不可缺少的视觉任务, Sanchez-Garcia等<sup>[6]</sup>根据植入者的日常生活需求,提出两种图像处理策略,其中一种策略的仿真实验证明物体掩膜和场景的结构性边缘在有限的分辨率下对室内场景的识别上表现出良好的效果。不同于室内环境,室外环境有着不确定性。Han等<sup>[7]</sup>提出场景简化策略,实验证明实例分割是其研究中的应用场景下的最有效的方法。

上述研究表明,在室外任务中合适的图像处理策略可以优化植入者的视觉感知,提高植入者的室外任务能力。本研究提出一种改进的实例分割的图像处理策略,对障碍物和盲道分别进行优化处理表达<sup>[8]</sup>,在有限的分辨率下表达出更多有效的视觉信息,从而在一定程度上满足视觉假体植入者在陌生室外环境中的避障需求,同时也在导航和场景感知上优化植入者的视觉感知。

## 1 方法

### 1.1 改进实例分割模型 SOLOv2-RS

基于深度学习的实例分割是指在图像中对不同物体进行像素级别分割的过程,其能准确分割出图像中识别到的每一个实例,并标记为不同的类别<sup>[9-10]</sup>。结合目标检测和语义分割的优点,实例分割在医学图像分割<sup>[11-12]</sup>、自动驾驶<sup>[13-14]</sup>、视觉推理等领域上满足研究人员的需求,能更好地解决科学和工程领域的图像处理问题。SOLO<sup>[15-16]</sup>系列模型是目前解决实例分割应用经典的深度学习模型。其中, SOLOv2<sup>[15]</sup>建立一个快速的实例分割框架,是对 SOLOv1<sup>[16]</sup>模型的优化。SOLOv2核心思想是将实例分割重新定义为两个同时存在的类别感知预测(目标检测)和实例感知掩码生成问题(语义分割);其基本结构基于全卷积的实例分割结构,不依赖于边界框,直接优化分割结果,在保证分割精度的同时可以实现目标的快速检测和提取。原始的 SOLOv2 模型采用基于残差的 ResNet<sup>[17]</sup>作为主干网络,以提取输入图像的特征。引入残差可以解决网络加深造成的

梯度爆炸和消失问题,但是由于 ResNet 是针对图像分类而设计的,其缺乏跨通道之间的相互作用和有限的感受野大小,将其作为主干网络时,在特定的下游任务中,如目标检测、实例分割等,模型的效果并不能完全发挥<sup>[15]</sup>。Xie等<sup>[18]</sup>提出的 ResNeXt 是一个简单的、高模块化的图像分类网络体系结构,其网络是通过重复构建块来实现的,该构建块聚合具有相同拓扑结构的一组转换;该策略增加了一个新的维度,即基数(cardinality),代表转换集合的大小。cardinality 作为深度和宽度维度之外的一个重要因素,在保持复杂性的限制条件下,增加 cardinality 可以提高分类精度;并且当增加容量时候,增加 cardinality 相比加深或加宽网络对模型性能更有效。Zhang等<sup>[19]</sup>进一步提出 ResNeSt,一种用于特征提取的拆分注意力网络。ResNeSt 借鉴 ResNeXt,具有普遍改善特征表示,在下游任务上表现出色,同时不需要额外的计算,很容易被用作下游任务的主干网络。

本研究基于 SOLOv2 和 ResNeSt<sup>[19]</sup>提出一种改进模型 SOLOv2-RS,其继承 SOLOv2 动态分割的同时扩大感受野的大小和跨通道相互作用,提升室外环境场景下面向行动导航任务的实例分割任务表现。SOLOv2-RS 模型结构原理如图 1 所示。根据输入的图片特征图, SOLOv2-RS 将其传入不同的 cardinality,每一个基数拆分为  $r$  个分支,每个分支有着独立的变换函数(Conv 1×1 和 Conv 3×3),并通过拆分注意力模块获取每个基数的表示输出,然后进行连接,最后 ResNeSt 主干网络输出提取的特征至特征金字塔 neck 层,由 SOLOv2 head 对图像特征进行预测,得到最终的掩膜图。

### 1.2 人工视觉分级优化处理 (Hierarchical Optimization Processing, HOP) 策略

在陌生室外环境下的行动导航任务中,盲道的识别和障碍物(遮挡盲道或盲道附近物体)的规避对于视觉假体植入者是十分重要的,需要尽可能将这些任务相关信息传递至假体植入者。利用改进实例分割模型 SOLOv2-RS 提取的室外场景中多个对象,本研究提出 HOP 策略,对盲道和障碍物进行 HOP,用于提示植入者在室外环境下的行动导航任务中盲道识别和障碍物规避。对于障碍物,根据标准重要性分布标准图,计算得出障碍物的整体重要性,并划分不同等级,按照灰度等级评价标准,采用不同的灰度等级处理;对于盲道,直接提取边缘信息后进行形态学膨胀并采用合适的灰度等级处理。处理后的盲道和障碍物均采用相同的光幻视分辨率表达。本研究的图像处理策略示意图见图 2,图像采集帧流 A 经过 SOLOv2-RS 分割提取得到盲道和障碍物掩膜 B。盲

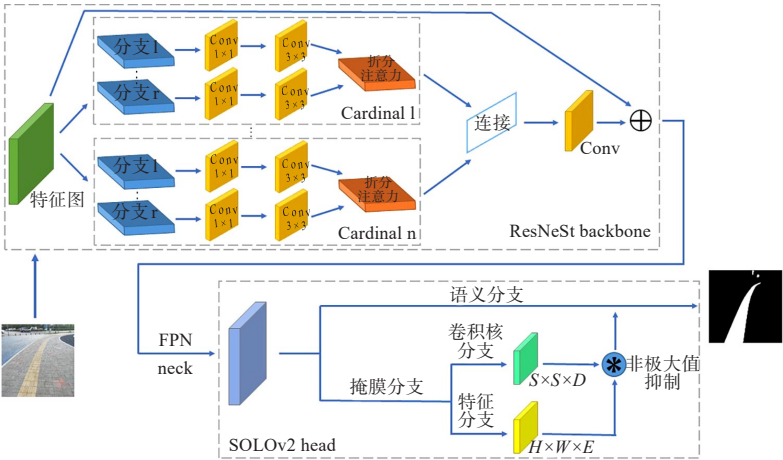


图1 SOLOv2-RS网络结构图

Figure 1 SOLOv2-RS network structure

道掩膜采用Canny算子提取边缘信息,并进行形态学膨胀处理获得盲道边缘图C;障碍物掩膜根据分级表达模型得到显著性图D。盲道边缘图C和显著性图D经过拼接得到拼接图E;拼接图E经过光幻视点仿真,最终作为输出帧流F。

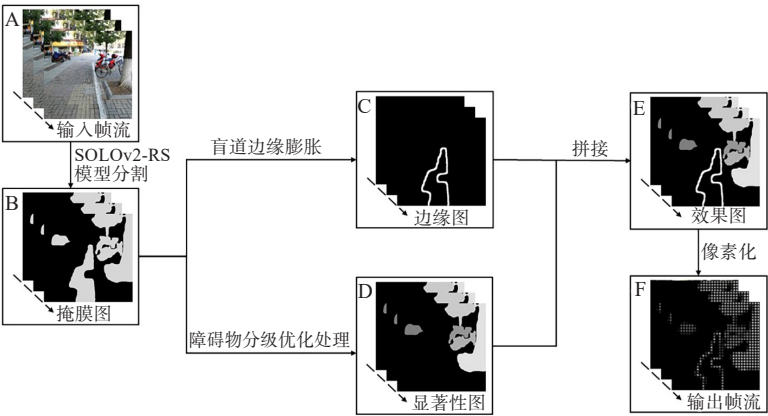


图2 基于SOLOv2-RS的人工假体视觉避障研究图像处理策略示意图

Figure 2 Schematic diagram of image processing strategy for SOLOv2-RS based obstacle avoidance in simulated prosthetic vision

**1.2.1 障碍物分级表达模型** SOLOv2-RS实例分割模型提取室外环境中障碍物后,根据重要性分布图,计算得到不同障碍物的重要性分数,并根据重要性分数,将其划分不同重要性等级,然后采用不同灰度等级处理表达具有不同重要性等级的障碍物。具体处理流程见图3,未经处理的障碍物掩膜图A分割提取得到多张单掩膜分割图B,根据重要性标准分布图C,计算各掩膜的重要性分数,并以此排序,对不同等级的掩膜分配相应的灰度等级,形成不同灰度等级各掩膜图D,经拼接后得到场景中障碍物分级优化表达结果E。

**1.2.2 障碍物重要性计算** 有研究表明视野中心的物体更有吸引力,这意味着靠近视野中心位置的物体

比远离视野中心位置的物体更有可能是重要的<sup>[20]</sup>。基于此,在避障任务中,靠近植入者视野中心的障碍物是重要的,更容易影响植入者。由于视觉注意力分布近似二维高斯分布,同时考虑光幻视点亮度应由中心向四周均匀递减,因此采用式(1)得到与输入图像大小一致的标准重要性分布(Importance Distribution, IMD):

$$IMD_{i,j} = \frac{1}{m} e^{-\frac{(i-c_x)^2 + (j-c_y)^2}{2\sigma^2}}, (i,j) \in I_{imd} \quad (1)$$

其中, $I_{imd}$ 是和输入规格相同大小的标准重要性分布图; $(c_x, c_y)$ 是分布图中心位置坐标; $\sigma^2 = \frac{w^2 + h^2}{4}$ ,  $w$ 和 $h$ 分别为重要性分布标准图的宽和高; $m$ 是归一化



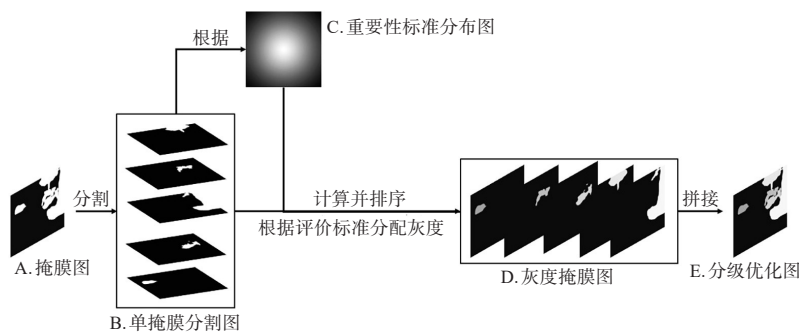


图3 障碍物分级表达处理流程图

Figure 3 Flowchart of obstacle hierarchical representation

系数,用来进行归一化。

根据得到的与输入规格相同大小的标准重要性分布图,通过式(2)计算障碍物的整体重要性:

$$IMO = \sum \text{mask} \times I_{\text{imd}} \quad (2)$$

其中,mask是障碍物经过实例分割模型提取出来的二进制掩膜

**1.2.3 障碍物分级表达** 按照障碍物整体重要性进行等级划分,不同重要性等级下的障碍物采用对应的灰度等级处理,具体分配方式见式(3)和式(4):

$$\text{Level} = \text{index} \{ \text{sortedIMOs} [\text{mask}_i] \}, i \in [1, n] \quad (3)$$

$$\text{GL}_{\text{mask}_i} = \text{ECIML} [\text{Level}] \quad (4)$$

其中,Level<sub>i</sub>是mask<sub>i</sub>的重要性等级分数;sortedIMOs是分割出来的所有障碍物经过从高到低排序后的整

体重要性序列;mask<sub>i</sub>是第i个具体障碍物;index{sortedIMOs[mask<sub>i</sub>]}表示第i个障碍物经过排序后的索引,即第i个障碍物的重要性等级;ECIML是灰度等级评价标准;GL是对应的灰度等级。

**1.2.4 盲道边缘增强处理** 盲道的边缘信息是影响盲道识别效果的重要信息。本研究采用5×5的Canny边缘算子对分割后的盲道掩膜提取边缘信息。由于提取得到的盲道边缘过细并且存在非连续情况,导致仿真结果中的光幻视点亮度动态范围较低。为有效表达盲道边缘信息,对提取到的盲道边缘进行形态学膨胀增强,然后对其分配合适的灰度等级。具体处理流程见图4,未经处理的掩膜图A筛选得到盲道掩膜B,然后进行边缘提取和形态学膨胀,得到盲道边缘图C。

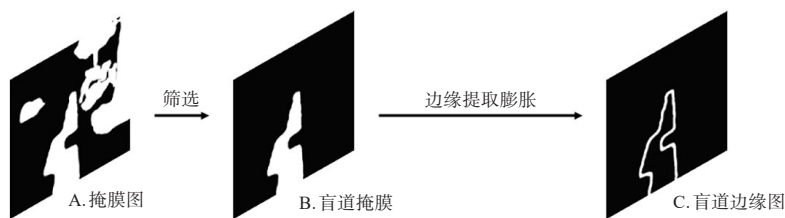


图4 盲道边缘提取膨胀处理流程图

Figure 4 Flowchart of tactile paving edge extraction and dilation

## 2 实验设计

### 2.1 数据集

数据集选取于模拟人眼视角下的室外场景拍摄图片(图5)。原始数据共1034张图像,其中实例目标共17类,包括盲道以及场景中盲道附近的目标,如汽车、行人、树木、垃圾桶等。为仿真视网膜假体Argus II的4×4电极阵列所能诱发理论视野,保持图像宽高比不变,将原始图像高度缩放到640像素。为

提升模型的鲁棒性和泛化能力,通过镜像、旋转等操作对原始数据进行增强,最终实验数据集扩充至4136张图像。自建数据集每类物体标签和数量如表1所示,并按照8:2的比例划分为训练集和测试集。

### 2.2 实验环境

实验硬件配置为2块Intel Xeon Silver 4210R Processor CPU处理器,采用GeForce GTX 3060 12 GB GPU的服务器,操作系统是64位Ubuntu22.04。



图5 数据集样例  
Figure 5 Samples in the dataset

2.3 评价指标

选用平均精度均值(mAP)、处理时间(FPS)、计算量(Flops)、参数量(Params)对模型性能进行评价。其中,AP<sub>0.5</sub>和AP<sub>0.75</sub>分别是IoU取0.5和0.75的AP值;mAP是IoU以步长为0.05,从0.5到0.95的AP平均值;APs、APm、APl分别是模型分割小目标物体(区域范围小于32×32像素)、中等目标物体(区域范围在32×32像素和96×96像素之间)和大目标物体(区域范围大于96×96像素)的AP值。

3 实验结果

3.1 SOLOv2-RS模型与其他模型比较

模型改进前后分割性能对比见表2和表3。相较于主干网络为ResNet和ResNeXt的SOLOv2模型,

表1 数据集各类标签数量 Table1Quantityofdifferenttypeoflabels	
标签	数量
行人	2 420
树	1 828
汽车	1 352
自行车带人	48
路障	2 016
配电箱	708
消防栓	328
公交站牌	160
自行车	116
电线杆	280
垃圾箱	1 072
盲道	4 860
电动车带人	336
电动车	1 528
指示牌	740
井盖	2 964
摩托车	72
总计	20 828

SOLOv2-RS模型(主干网络为ResNeSt)的mAP指标分别提高5.2%和3.3%,但是Flops和Params分别增加6.610 G、1.926 M和5.852 G、2.454 M,而3个主干网络的FPS基本没有差距。在Flops和Params方面上,表现最好的分别是ResNet和ResNeXt。虽然ResNeSt提高了mAP指标,并在一定程度上增加了Flops和Params,但是这对模型整体的性能基本没有影响,可见SOLOv2-RS模型在性能基本不改变的情况下,提高模型分割精度,特别是在小目标物体分割方面,相较于ResNet和ResNeXt,ResNeSt在APs指标分别提高6.4%和3.9%。

表2 不同主干网络的平均精度(%)  
Table 2 Comparison of average precision among different backbone networks (%)

主干网络	AP <sub>0.5</sub>	AP <sub>0.75</sub>	APs	APm	APl	mAP
ResNet	76.0	57.9	13.3	58.5	80.4	50.9
ResNeXt	78.4	59.2	15.8	60.5	81.8	52.8
ResNeSt	81.7	62.7	19.7	64.8	83.9	56.1

3.2 人工视觉仿真处理

通过障碍物的分层表达和盲道边缘增强策略将

室外避障任务中的盲道和障碍物分配相应的灰度等级,然后采用局部平均法<sup>[21]</sup>,均降采样到32×32光幻

表3 不同主干网络的处理时间、计算量及参数量

Table 3 Comparison of FPS, Flops and Params among different backbone networks

主干网络	FPS/ $\text{ms} \cdot \text{batch}^{-1}$	Flops/G	Params/M
ResNet	82.0	54.105	31.200
ResNeXt	81.3	54.863	30.672
ResNeSt	82.6	60.715	33.126

视点数。本研究对比了以下4种处理方法:DP(Direct Pixelization)为原始输入灰度化后降采样为高斯点;IS(Instance Segmentation)是采用SOLOv2-RS实例分割模型提取障碍物和盲道掩膜,然后分配相同的灰度等级并进行低像素化处理;MEP(Mask Edge Pixelization)参考 Sanchez-Garcia 等<sup>[6]</sup>,采用Canny边缘算子对SOLOv2-RS实例分割模型分割得到的盲道及障碍物的掩膜提取边缘,然后对边缘信息进行5×5的膨胀增强处理,并进行人工视觉仿真处理<sup>[22]</sup>;HOP是采用SOLOv2-RS实例分割模型提取障碍物和盲道的掩膜并进行HOP,然后降采样为高

斯点。

图6为分别在4种图像处理策略下的部分图片处理过程及仿真结果示意。仿真结果表明,在DP处理下,虽然一定量的光幻视点可以感知部分场景信息,但是在32×32的光幻视点下视觉感知效果并不理想,无法辨识室外复杂环境中的盲道和多类障碍物。IS方法能准确分割提取盲道和障碍物,增强视觉感知;但是当盲道上存在障碍物时(如图6原图列第3、4、5张图片),由于盲道和障碍物的灰度等级和分辨率相同,无法有效感知盲道上的状况。MEP方法在一定程度上解决了IS方法中存在的问题,但是在室外环境较为复杂的情况下,HOP方法处理后的效果明显优于MEP方法。当路况较为复杂时(如图6原图列第6张图片),MEP方法的处理方式是将各实例以边缘表达,导致边缘信息相互交叠,盲道与各障碍物难以区分;HOP方法则采用HOP策略,可易于区分盲道和障碍物,且前方障碍物也更容易地被感知到。因此在较为复杂的环境中,基于SOLOv2-RS的HOP方法优于其他人工假体视觉导航任务应用中的图像处理方法。

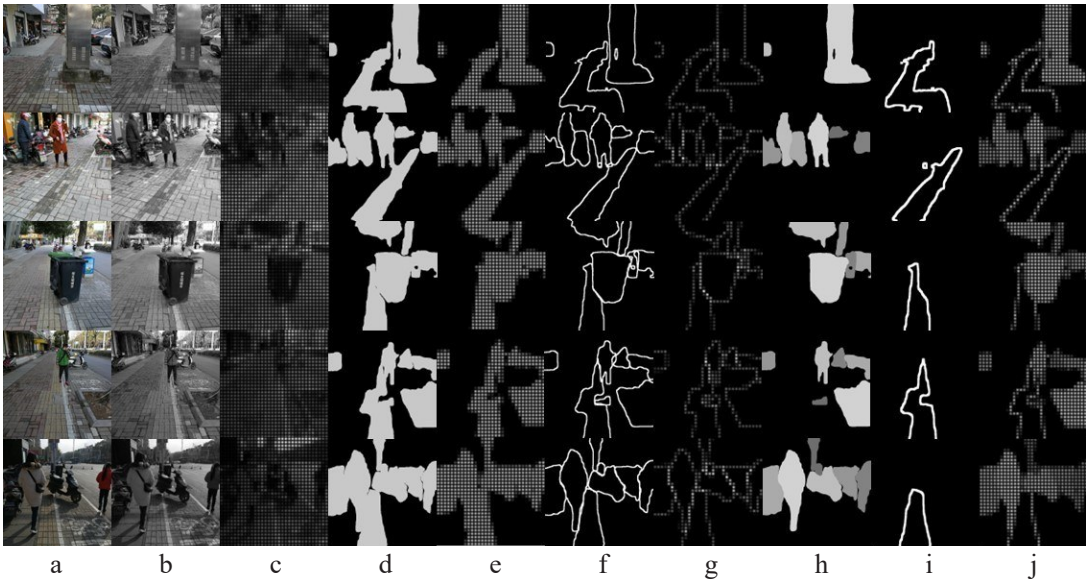


图6 不同图像处理策略下光幻视点仿真结果图

Figure 6 Results of phosphenes simulation using different image processing strategies

a:原图;b:原图的灰度图;c:DP;d:mask;e:IS处理结果;f:盲道和障碍物边缘提取结果;g:MEP处理结果;h:障碍物分级表达 mask;i:盲道边缘增强结果;j:HOP处理结果

4 结 论

本研究面向室外环境避障任务,基于SOLOv2-RS模型实例分割模型,提出一种图像处理策略HOP。相较于原始SOLOv2模型,SOLOv2-RS在分割结果

和效果上表现更加优秀;本研究提出的HOP策略可以提高在较为复杂的环境中的视觉感知能力,特别在盲道上或附近存在障碍物情况下,效果更显著。本研究为低分辨率人工假体视觉下的复杂室外场景



任务提供一种可行的信息优化处理策略,同时为其他自主系统在较为复杂的环境下避障任务提供一种图像信息处理思路。

## 【参考文献】

- [1] 闫妍,柴新禹,陈焱,等.视觉假体的研究进展[J].生理学报,2016,68(5): 628-636.  
Yan Y, Chai XY, Chen Y, et al. Research progress of visual prosthesis [J]. *Acta Physiologica Sinica*, 2016, 68(5): 628-636.
- [2] Dagnelie G, Keane P, Narla V, et al. Real and virtual mobility performance in simulated prosthetic vision[J]. *J Neural Eng*, 2007, 4(1): S92-S101.
- [3] Vergnieux V, Macé MJ, Jouffrais C. Simplification of visual rendering in simulated prosthetic vision facilitates navigation[J]. *Artif Organs*, 2017, 41(9): 852-861.
- [4] Parikh N, Itti L, Weiland J. Saliency-based image processing for retinal prostheses[J]. *J Neural Eng*, 2010, 7(1): 016006.
- [5] Parikh N, Itti L, Humayun M, et al. Performance of visually guided tasks using simulated prosthetic vision and saliency-based cues[J]. *J Neural Eng*, 2013, 10(2): 026017.
- [6] Sanchez-Garcia M, Martinez-Cantin R, Guerrero JJ. Indoor scenes understanding for visual prosthesis with fully convolutional networks [C]//*Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. Setúbal, Portugal: SciTePress, 2019: 218-225.
- [7] Han N, Srivastava S, Xu AW, et al. Deep learning-based scene simplification for bionic vision[C]//*Proceedings of the Augmented Humans International Conference 2021*. New York, NY, USA: Association for Computing Machinery, 2021: 45-54.
- [8] 王静,刘建云,韩彦岭,等.基于实例分割和显著性计算的人工视觉多目标优化处理[J].中国医学物理学杂志,2023,40(3): 320-327.  
Wang J, Liu JY, Han YL, et al. Multi-object optimization approach for artificial vision based on instance segmentation and saliency detection [J]. *Chinese Journal of Medical Physics*, 2023, 40(3): 320-327.
- [9] He KM, Gkioxari G, Dollár P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2017: 2980-2988.
- [10] Cai ZW, Vasconcelos N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 6154-6162.
- [11] Liu QD, Chen C, Qin J, et al. FedDG: federated domain generalization on medical image segmentation via episodic learning in continuous frequency space[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 1013-1023.
- [12] Hsu J, Chiu W, Yeung S. DARCNN: domain adaptive region-based convolutional neural network for unsupervised instance segmentation in biomedical images[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 1003-1012.
- [13] Pei JL, Cheng TY, Fan DP, et al. OSFormer: one-stage camouflaged instance segmentation with transformers[C]//*Computer Vision-ECCV 2022*. Cham: Springer Nature Switzerland, 2022: 19-37.
- [14] Li WT, Liu WY, Zhu JK, et al. Box-supervised instance segmentation with level set evolution[C]//*Computer Vision-ECCV 2022*. Cham: Springer Nature Switzerland, 2022: 1-18.
- [15] Wang XL, Zhang RF, Kong T, et al. SOLOv2: dynamic and fast instance segmentation [C]//*Advances in Neural Information Processing Systems*. Red Hook, NY, USA: Curran Associates Inc., 2020: 17721-17732.
- [16] Wang XL, Kong T, Shen CH, et al. SOLO: segmenting objects by locations [C]//*Computer Vision-ECCV 2020*. Cham: Springer International Publishing, 2020: 649-665.
- [17] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [18] Xie SN, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 5987-5995.
- [19] Zhang H, Wu CR, Zhang ZY, et al. ResNeSt: split-attention networks [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway, NJ, USA: IEEE, 2022: 2735-2745.
- [20] Jiang HC, Li H, Liang JL, et al. A hierarchical image processing strategy for artificial retinal prostheses [C]//2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE). Piscataway, NJ, USA: IEEE, 2020: 359-362.
- [21] Li H, Han TT, Wang J, et al. A real-time image optimization strategy based on global saliency detection for artificial retinal prostheses[J]. *Inf Sci*, 2017, 415-416: 1-18.
- [22] 朱海怡.基于深度学习实例分割的人工假体视觉行动能力障碍避免研究[D].上海:上海海洋大学,2022.  
Zhu HY. Study of mobility and obstacle avoidance in artificial prosthetic vision based on deep learning instance segmentation[D]. Shanghai: Shanghai Ocean University, 2022.

(编辑:谭斯允)