

DOI:10.3969/j.issn.1005-202X.2025.07.007

医学影像物理

基于显著目标检测的人工视觉物体识别图像处理策略

张岩¹, 赵瑛², 曹凤³, 姜广森², 何洋², 王盛², 王楠¹

1. 齐鲁理工学院计算机与信息工程学院, 山东 济南 250200; 2. 内蒙古科技大学数智产业学院, 内蒙古 包头 014010; 3. 齐鲁理工学院智能制造与控制工程学院, 山东 济南 250200

【摘要】目的:为了优化有限分辨率下假体视觉信息的呈现,提出一种基于显著目标检测算法的图像处理策略,旨在检测和增强场景中的显著目标,去除背景信息。**方法:**首先,结合CNN与Transformer提出一种显著目标检测模型提取显著目标,在此基础上利用目标放大,轮廓增强与对比度增强等方法对图像信息进行优化,并在5种分辨率(16×16、24×24、32×32、48×48、64×64)下进行心理物理学实验。**结果:**在仿真假体视觉下该图像处理策略对于提高被试的物体识别能力效果显著。在16×16、24×24、32×32、48×48、64×64分辨率下,本文方法均取得了最高的图像内容识别准确率,分别为34%±6%、56%±9%、72%±9%、89%±4%和96%±2%。**结论:**使用显著目标检测方法和图像处理策略对显著目标进行提取和增强可以帮助假体植入者有效提高物体识别能力。

【关键词】视觉假体; 仿真假体视觉; 显著目标检测; 物体识别; 图像处理策略

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2025)07-0883-09

Image processing strategy for object recognition in artificial vision based on salient object detection

ZHANG Yan¹, ZHAO Ying², CAO Feng³, JIANG Guangmiao², HE Yang², WANG Sheng², WANG Nan¹

1. School of Computer and Information Engineering, Qilu Institute of Technology, Ji'nan 250200, China; 2. School of Digital and Intelligent Industry, Inner Mongolia University of Science and Technology, Baotou 014010, China; 3. School of Intelligent Manufacturing and Control Engineering, Qilu Institute of Technology, Ji'nan 250200, China

Abstract: Objective To propose a image processing strategy based on salient object detection algorithm for optimizing the presentation of prosthetic visual information at a limited resolution level, aiming to detect and enhance the salient objects in the scene and remove the background information. **Methods** A salient object detection model combining CNN and Transformer was used to extract salient objects. On this basis, methods such as object magnification, contour enhancement and contrast enhancement were utilized to optimize the image information. Psychophysical experiments were carried out at 5 resolution levels (16×16, 24×24, 32×32, 48×48 and 64×64). **Results** In the simulated prosthetic vision, this image processing strategy had a remarkable effect on improving the object recognition ability of the subjects. Regardless of the resolutions of 16×16, 24×24, 32×32, 48×48 and 64×64, the proposed strategy achieved the highest recognition accuracies, specifically 34%±6%, 56%±9%, 72%±9%, 89%±4% and 96%±2%. **Conclusion** Using the salient object detection method and image processing strategy to extract and enhance salient objects can help prosthesis implant recipients effectively improve their object recognition ability.

Keywords: visual prosthesis; simulated prosthetic vision; salient object detection; object recognition; image processing strategy

【收稿日期】2025-02-19

【基金项目】国家自然科学基金(81460279);齐鲁理工学院科研项目基金(QIT24NN047)

【作者简介】张岩, 硕士, 研究方向: 计算机视觉, E-mail: zhangyan_5533@163.com

【通信作者】曹凤, 博士, 副教授, 研究方向: 图像处理, E-mail: caofeng5354@163.com

前言

视觉假体能够帮助不可逆病变造成视觉损伤的患者恢复部分视力。该装置将刺激电极植入到患者体内,通过电刺激视觉通路中功能完好的神经组织,进而让盲人能够获取一定的视觉感知,最终实现修复视觉功能的目标。尽管视网膜假体的研究已经取得了较大进展^[1-3],但是由于视网膜假体的植入电极

数目有限,只能呈现给假体植入者低分辨率的视觉信息^[4]。基于此,研究者们相继提出了多种不同的图像处理策略,旨在对假体所展现出的视觉信息予以优化,并通过在仿真假体视觉(Simulated Prosthetic Vision, SPV)下进行心理物理学试验^[5],对这些策略的性能展开评估^[6]。Parikh等^[7]使用基于显著性的算法引导被试将注意力聚焦到感兴趣区域,使得被试在完成搜索任务时的表现有明显改善。Wang等^[8]提出两种基于Itti显著性算法的图像处理策略,借助这些策略能够在简单背景下对前景目标进行提取。实验结果表明,这两种策略让目标识别精度得到显著提升。Li等^[9]先是利用基于生物似然图的视觉显著性算法对应的显著性分割方法,再配合基于GrabCut的自适应迭代优化框架,自动完成前景对象的提取工作。在此基础上,运用分离像素化以及背景像素收缩的方法,进一步对提取出来的前景对象进行增强处理。Guo等^[10]提出了两种基于显著目标检测技术的图像处理策略。心理物理学实验表明,前景缩放与背景杂波去除、前景边缘检测与背景缩减等技术对模拟假体视觉中的目标识别任务有积极的影响。上述结果表明,基于显著性特征提取的图像处理算法能够有效提高假体视觉感知能力^[11-13]。因此,提出了基于显著目标检测的图像优化策略^[14],旨在对显著目标予以增强,同时削减背景里的冗余信息。首先,笔者采用Transformer和CNN并行的双分支结构模型来进行全局关系建模和局部表示学习^[15-17]。其次,笔者设计了一个特征融合模块对两分支的特征进行融合互补^[18-20]。以上设计能够更好地定位显著目标和提取局部细节特征。本研究的目的是优化低分辨率假体视觉下的视觉信息,以提高低分辨率人工视觉中的物体识别性能。因此,在显著目标检测方法的基础上,结合目标放大,对比度增强和轮廓增强等方法提出一种图像优化策略。为了验证所提出策略的有效性,在16×16、24×24、32×32、48×48、64×64共5种分辨率下进行了基于SPV的心理物理学试验。试验结果表明,与原图直接像素化(DP)和显著图直接像素化(PSM)相比,本文方法在识别准确率和识别效率方面均有显著提高。

1 资料与方法

1.1 试验对象

24名被试(男12名、女12名),年龄为20~30岁。视力正常或矫正至正常。他们均未曾具备有关SPV方面的经验与知识。在开展SPV试验之前,相关人员向他们详细告知了试验步骤以及需要留意的各类事项。本研究严格遵循《赫尔辛基宣言》的要求。

1.2 试验平台

试验平台包括两台电脑。其中一台电脑(AMD Ryzen Threadripper 1950X 16-Core Processor, 3.40 GHz, 32.0 GB)用来进行模型训练。模型基于Python3.8, CUDA 11.3,在Pytorch1.10.0开源框架下实现。另外一台电脑(AMD Ryzen Threadripper 1950X 16-Core Processor, 3.40 GHz, 32.0 GB)负责图像处理以及显示工作,它能把视觉刺激展示给被试,同时也可供试验人员用来记录数据。被试坐在显示器前50 cm处,图像随机地呈现在显示器中央。试验过程由E-Prime 2.0记录。

1.3 图像获取与预处理

试验用的素材从图像数据集中选取,其中场景和对象常见于日常生活中。由于目前的视网膜植入物一般放置在黄斑上方,其所能提供的视野仅有约20°的视野。在此基础上,对图像进行裁剪,裁剪后的图像尺寸为480×480,占据20°的视角(对角),与Argus II的视野相同^[3]。试验数据库由60张包含不同物体的图像组成,其中包括20张单物体图像和40张多物体图像。考虑到单个物体的识别阈值在(16×16)~(24×24)区间内^[21],通过预试验验证,最终确定在16×16、24×24、32×32、48×48、64×64共5种分辨率下进行试验。虽然试验设计超出了当前植入物的分辨率,但该仿真可能为未来的高密度植入物提供更多的试验参考。

1.4 图像处理策略

视觉假体中的图像处理阶段是将原始图像的分辨率与刺激电极的数量相对应。本文提出一种基于显著目标检测的图像处理策略,旨在优化假体视觉下的视觉信息,进而提高低分辨率条件下物体的识别性能,具体处理流程如图1所示。首先,利用显著目标检测模型获取原始图像的显著图。将显著图作用于原始图像以提取显著目标并放大。然后,由显著图获取目标轮廓,将增强后的轮廓添加到灰度化的显著目标上。经过对比度增强后,通过用高斯点降低分辨率导出像素化图像获得最终的视觉刺激。其中,DP图像的生成方式为在原始灰度图像的基础上直接运用高斯点光幻视模型获得;PSM图像由输入图像的显著图直接运用高斯点光幻视模型获得。

1.4.1 显著性检测 为了准确且简单地提取显著目标,本文提出一种基于深度学习技术的显著目标检测算法。模型架构如图2所示。具体而言,编码器采用Swin-Transformer^[22]与ResNet以双分支并行方式分别建模全局依赖性和提取局部特征,并设计了一个特征融合模块对提取到的特征进行融合互补。解码器由Swin-Transformer模块和双上采样模块构成,用于执行上采样操作以恢复特征图的空间分辨率。

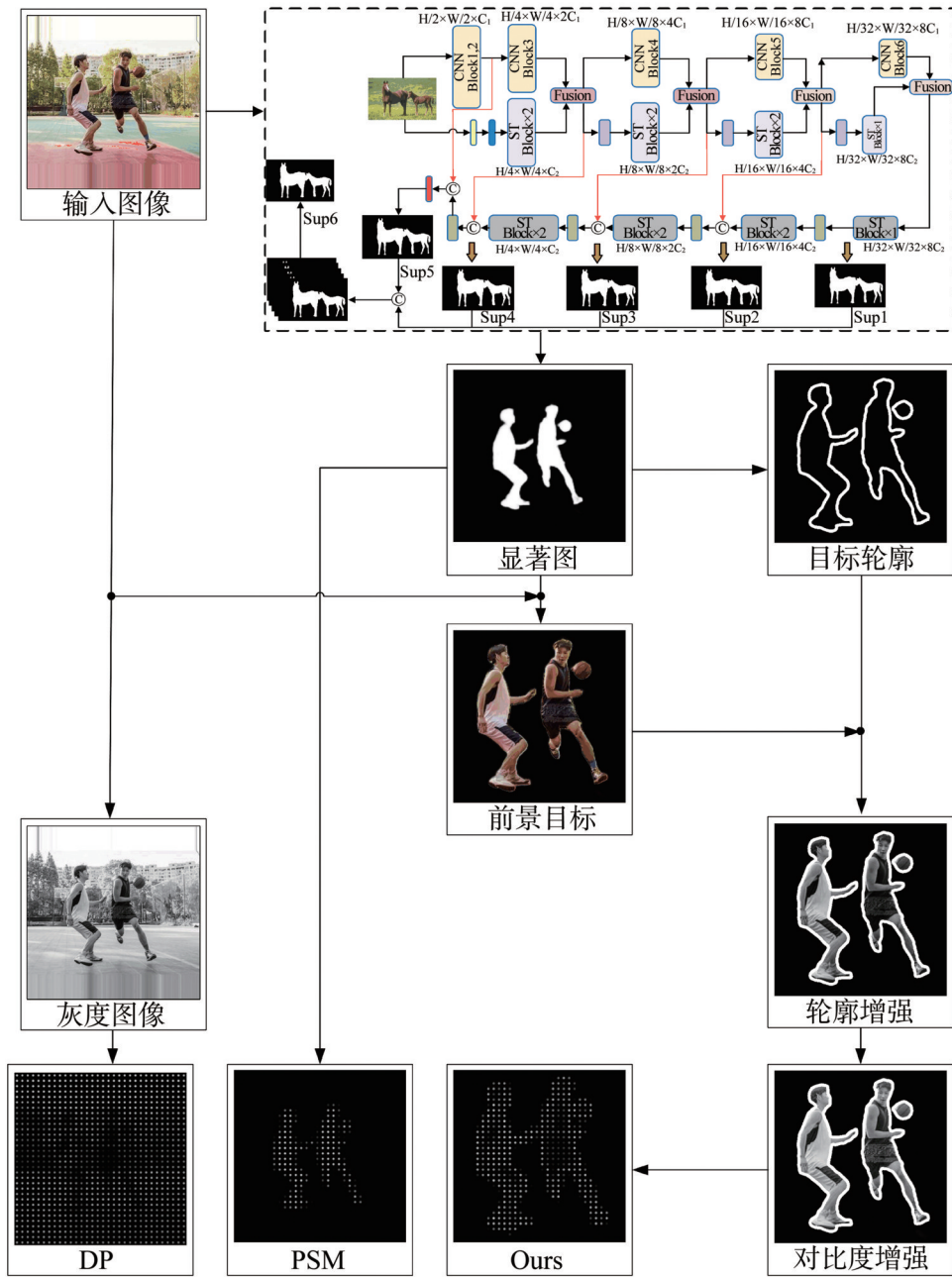


图1 3种不同图像处理策略的过程展示

Figure 1 Process demonstration of 3 different image processing strategies

受 HED (Holistically-nested Edge Detection) 模型^[23]启发,为了减少过拟合,在每个解码器阶段的最后一层使用真值进行监督。该模型具有若干优点:首先,模型融合了卷积神经网络的局部建模能力和 Swin-Transformer 的长程建模能力。其次,通过多尺度 U-Net 结构进一步增强了模型的局部和长程建模能力。第三,特征融合模块可以有效促进卷积模块和 Transformer 模块之间的信息融合。第四,双上采样模块可以有效防止棋盘伪影。

1.4.2 图像分割 得益于显著目标检测模型的良好性能,模型输出的显著图具有清晰的边界。笔者将原图与显著图进行简单的逐像素相与操作来进行目标

分割。分割图像中的目标区域将保留原始图像中对应像素位置上的颜色信息,而背景区域则变为黑色。分割结果如图3所示。

1.4.3 轮廓增强 本文轮廓检测算法是基于 Suzuki 等^[24]提出的一种轮廓跟踪算法。该算法基于连接式组件标记原理,通过对图像进行扫描,并利用像素之间的连接关系来识别连通区域,从而得到轮廓。在进行低像素化处理时,笔者将 15×15 区域内灰度的平均值作为该位置光幻视点的灰度值,这导致了物体轮廓会受到背景像素的稀释。为此笔者增加了轮廓的宽度来降低背景像素的干扰。

1.4.4 图像信息优化 在通过显著目标检测模型得到

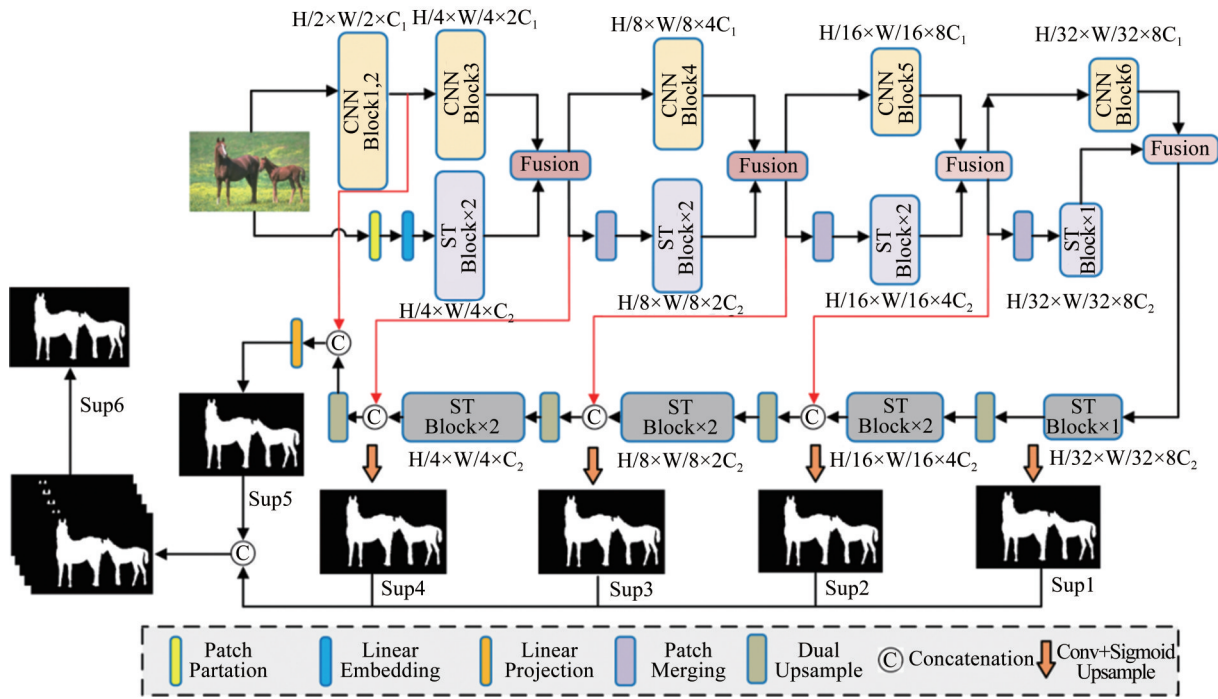


图2 显著目标检测模型架构

Figure 2 Architecture of salient object detection model

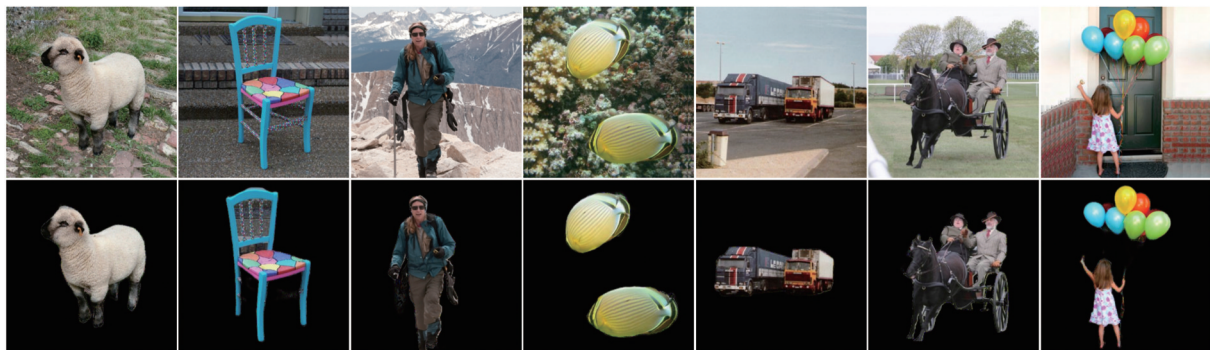


图3 部分试验图像的目标分割结果示例

Figure 3 Examples of object segmentation results of some experimental images

显著图的基础上,笔者结合目标放大、轮廓增强与对比度增强提出了一种图像处理策略。为了充分利用低分辨率下的像素资源,突出目标轮廓,尤其对于图像中的小目标,单纯的低像素化处理目标仅由少数光幻视点组成,难以达到识别要求。因此笔者首先对目标进行放大处理,然后将提取到的轮廓增强后叠加到灰度化的目标图像上,最后对图像进行对比度增强后生成试验所用的像素化图像。同时笔者将原始图像的灰度图和显著图直接降采样为像素化图像进行对比试验。

1.4.5 光幻视模拟 通过预试验以及早期的研究可知,SPV下单个物体的分辨率识别阈值为 16×16 到 24×24 。因此,笔者将图像降采样到 16×16 、 24×24 、 32×32 、 48×48 、 64×64 5种分辨率。针对图像里每一

个 15×15 像素的区域,笔者采取了平均灰度值的方式,进而得到了直径为7个像素的光幻视点。为了更加接近假体植入者描述的光幻视点形态,笔者使用灰度分布呈高斯曲线分布的高斯光幻视点。图4所示是 32×32 分辨率时,DP、PSM和本文方法这三种策略的处理结果。

1.5 心理物理学试验设计与程序

为了对所提出的图像处理策略在现实场景识别方面的有效性予以验证,笔者设计了SPV情境下的心理物理学试验。在开展正式试验前,会先安排预试验,且预试验用到的试验素材和正式试验的素材不存在重复情况。图像由电脑屏幕进行呈现,在被试者看到图像后,需尽可能地描述屏幕上呈现的图像内容。该试验主要划分成两个任务板块:任务1主

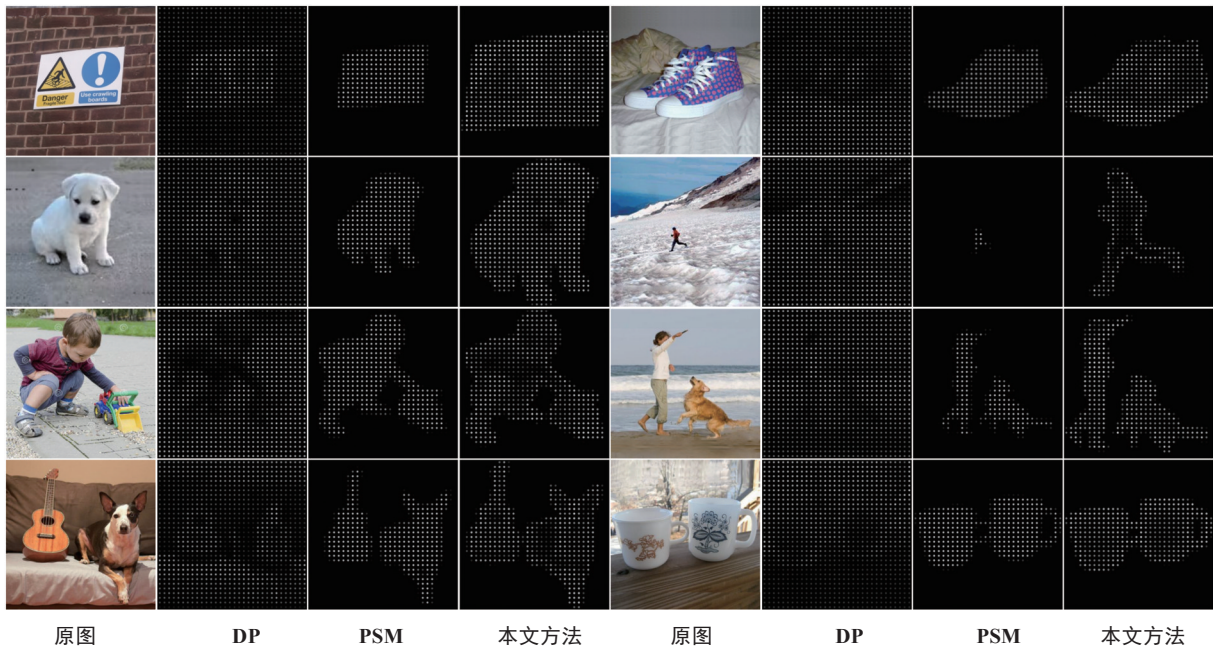


图4 32×32分辨率下DP、PSM和本文方法3种处理策略的结果示例

Figure 4 Examples of results of 3 processing strategies (DP, PSM and the proposed method) at a resolution of 32×32

要针对图像中物体数量的识别。被试者若能精准报出图像里物体的数量,可得1分,反之则计0分。任务2是对图像场景内容的识别判定,假如被试者完全不清楚图片中物体情况且无法描述,记为0分;若被试者可描述图片中的物体,却不能明确具体为何物,则记为1分;若被试者正确识别出物体,记为2分。正式试验所采用的图片数量为60张,为尽可能降低学习效应的影响,把24名被试平均划分成3个小组,依照不同的先后次序展示3种处理策略的图像。试验顺序采用拉丁方样式,详细顺序见表1。

表1 针对不同分组的图像处理策略排列顺序
Table 1 Arrangement order of image processing strategies for different groups

被试组	顺序一	顺序二	顺序三
第一组	DP	PSM	本文方法
第二组	PSM	本文方法	DP
第三组	本文方法	DP	PSM

1.6 数据分析

基于被试完成试验的识别准确率(Recognition Accuracy, RA)、识别时间(Reaction Time, RT)和识别效率(Recognition Efficiency, RE)3项指标,对DP、PSM和本文方法3种图像处理策略展开评估。不同任务中,每张图片的RA由式(1)得出:

$$RA = (\text{score}/S) \times 100\% \quad (1)$$

其中,RA代表任务完成的准确率,score是被试在完成的任务后所获得的分数,S表示任务的分值,任务1中

S=1,任务2中S=2。

RT指的是被试对每幅图像进行识别时所耗费时间的平均值,即从图像呈现到被试描述完毕的平均时间。根据RA和RT计算各组的RE:

$$RE = RA/RT \quad (2)$$

结果以均值±标准差表示。同一图像处理策略下不同分辨率之间性能指标比较采用SPSS中的单因素方差分析,同一分辨率下不同图像处理策略性能指标两两比较采用LSD(Least Significant Difference)多重效应分析,通过P值来衡量差异的显著性,P<0.05代表结果存在显著性差异。

2 结果

2.1 识别时间结果

表2为不同分辨率下各处理策略的识别时间。结果表明,同一图像处理策略下,5个不同的分辨率下RT均存在显著性差异(P<0.001)。同一分辨率下,3种图像处理策略之间进行两两比较时,如图5a所示,当分辨率为16×16或32×32时,PSM和本文方法的RT没有显著性差异;当分辨率为48×48或64×64时,DP和PSM方法的RT也没有显著性差异;除此之外,同一分辨率下,DP、PSM和本文方法的RT两两比较均存在显著性差异。

2.2 物体数量识别准确率结果

如表3所示,在物体数量识别的任务中,同一图像处理策略下,5个不同的分辨率下RA均存在显著性差异(P<0.001)。同一分辨率下,3种图像处理策

表2 不同分辨率下,3种图像处理策略所对应的识别时间结果($\bar{x}\pm s, s$)
Table 2 Recognition time corresponding to 3 image processing strategies at different resolution levels (Mean \pm SD, s)

方法	分辨率					P值
	16×16	24×24	32×32	48×48	64×64	
DP	11.24±1.17	9.60±0.99	8.00±1.01	6.36±1.13	4.63±0.72	<0.001
PSM	9.48±1.03	8.43±1.06	6.77±1.07	5.73±1.25	4.30±0.70	<0.001
本文方法	8.79±1.42	7.60±1.25	6.27±1.15	5.08±0.96	3.85±0.62	<0.001

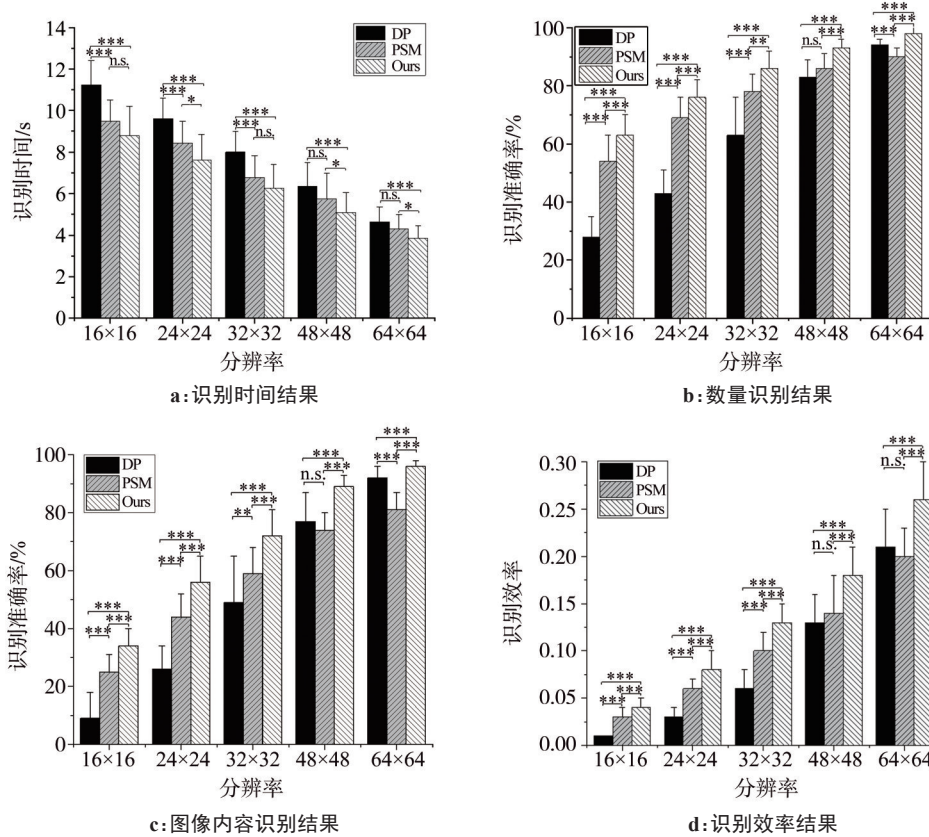


图5 图像识别结果

Figure 5 Image recognition results

*表示 $P < 0.05$; **表示 $P < 0.01$; ***表示 $P < 0.001$

表3 不同分辨率下,3种图像处理策略所对应的物体数量识别准确率结果($\bar{x}\pm s, \%$)
Table 3 Object quantity recognition accuracies corresponding to 3 image processing strategies at different resolution levels (Mean \pm SD, %)

方法	分辨率					P值
	16×16	24×24	32×32	48×48	64×64	
DP	28±7	43±8	63±13	83±6	94±2	<0.001
PSM	54±9	69±7	78±6	86±5	90±3	<0.001
本文方法	63±7	76±6	86±6	93±3	98±2	<0.001

略之间进行两两比较时,如图5b所示,当分辨率为48×48时,DP和PSM方法的RA没有显著性差异($P > 0.05$),除此之外,同一分辨率下,DP、PSM和本文方法的RA两两比较均存在显著性差异。

2.3 图像内容描述准确率结果

如表4所示,在描述图像内容的任务中,同一图像处理策略下,5个不同的分辨率下RA均存在显著

性差异($P < 0.001$)。同一分辨率下,3种图像处理策略之间进行两两比较时,如图5c所示,当分辨率为16×16、24×24、32×32时,DP方法的RA均小于PSM方法;当分辨率为48×48、64×64时,DP的RA大于PSM的RA。当分辨率为48×48时,DP和PSM的RA没有显著性差异,除此之外,同一分辨率下,DP、PSM和本文方法的RA两两比较均存在显著性差异。

表4 不同分辨率下,3种图像处理策略所对应的图像内容描述准确率结果($\bar{x}\pm s$,%)Table 4 Image content description accuracies corresponding to 3 image processing strategies at different resolutions levels (Mean \pm SD, %)

方法	分辨率					P值
	16×16	24×24	32×32	48×48	64×64	
DP	9±9	26±8	49±16	77±10	92±4	<0.001
PSM	25±6	44±8	59±9	74±6	81±6	<0.001
本文方法	34±6	56±9	72±9	89±4	96±2	<0.001

2.4 图像识别效率结果

图像识别效率由完成上述两个任务的整体识别准确率和识别时间得出。如表5所示,同一图像处理策略下,5个不同的分辨率下RE均存在显著性差异($P<0.001$)。同一分辨率下,3种图像处理策略之间进行两两比较时,如图5d所示,当分辨率

低于64×64时,DP的RE均小于PSM的RE,当分辨率为64×64时,DP的RE大于PSM的RE。当分辨率为48×48时,DP和PSM的RE没有显著性差异,除此之外,同一分辨率下,DP、PSM和本文方法的RE两两比较均存在显著性差异。

表5 不同分辨率下,3种图像处理策略所对应的图像识别效率结果($\bar{x}\pm s$)Table 5 Image recognition efficiencies corresponding to 3 image processing strategies at different resolution levels (Mean \pm SD)

方法	分辨率					P值
	16×16	24×24	32×32	48×48	64×64	
DP	0.01±0	0.03±0.01	0.06±0.02	0.13±0.03	0.21±0.04	<0.001
PSM	0.03±0.01	0.06±0.01	0.10±0.02	0.14±0.04	0.20±0.03	<0.001
本文方法	0.04±0.01	0.08±0.02	0.13±0.02	0.18±0.03	0.26±0.04	<0.001

3 讨论

引入某种图像处理技术被认为是优化视觉假体中视觉信息的有效方法^[25]。在本文中,笔者证明了基于深度学习的显著目标检测技术对日常场景中主要对象的分割有显著的效果。并且提出了基于显著目标检测的图像处理策略以提取并凸显现实场景关键信息,辅助被试完成物体感知与识别任务。心理物理学试验证明,该方法可以显著提高被试的物体定位和识别能力。

3.1 显著目标提取效果的影响

显著图的生成是图像处理策略中的关键步骤,它直接影响到目标提取的精度以及提取后的进一步增强处理。3种处理策略的结果分析表明,显著图的提取效果对识别准确率有显著影响。对于常见的物体类别,如人、车、动物等,利用传统算法进行分割时会导致目标的缺失或包含一些不必要的背景信息。这会导致物体的轮廓发生改变,进而影响识别准确率。而笔者提出的算法所提取的对象具有清晰边界,通常能够实现较高的准确性。因此,在同样基于显著目标检测的图像处理策略中应用本文提出的检测模型会带来更稳健的识别性能。

3.2 图像处理策略的影响

本文提出的图像处理策略在16×16、24×24、

32×32、48×48和64×64 5种分辨率下的识别性能显著优于DP与PSM方法。在16×16分辨率下,图像难以提供有用的边缘信息,导致背景对被试的判断产生了极大的干扰。物体的轮廓信息成为被试对图像中物体数量和种类进行判断的重要依据。与DP相比,PSM和本文方法基于显著目标检测将背景去除,使得物体具有清晰的轮廓,所以PSM和本文方法的识别准确率显著提升。本文方法相对于PSM的显著提升,主要原因是本文方法应用了目标放大的策略,使得物体的轮廓更加突出。同时,显著目标提取能够帮助被试快速定位目标位置,大大减少被试的识别时间。但是对于存在遮挡的物体,物体轮廓发生了改变,导致被试很难做出正确的判断。随着分辨率提高到24×24和32×32,对比度较高的图像中,物体轮廓逐渐清晰,DP的识别准确率得到显著提升。但是对于对比度较低的图像,物体轮廓并不明显。并且边缘信息所提供的纹理和细节特征不能弥补轮廓模糊带来的损失。所以相对于DP,PSM和本文方法的识别性能有着显著提升。可以看出当前分辨率下物体轮廓仍然是影响被试进行判断的重要因素。本文方法相对于PSM的显著提升同样是因为本文方法应用了目标放大的策略。当分辨率进一步增加到48×48,DP策略下物体的轮廓更加清晰。同时,边缘信息为被试提供了更多的判断依据。对于存在相互

遮挡的物体,边缘信息能够帮助被试对不同物体进行有效区分。所以PSM相较于DP已经没有了显著提升,甚至在图像内容描述的任务中,DP已经超过了PSM。通过对比度增强,本文方法相对于DP具有更加清晰的边缘与轮廓,所以本文方法仍然有着显著提升。在分辨率为 64×64 时,PSM方法已经不具备优势,清晰的边缘信息使DP显著超越了PSM,但是低对比度图像的存在与对比度增强的应用,本文方法在物体轮廓与细节特征上要优于DP,所以本文方法与DP的性能指标仍旧存在显著性差异。

所以,在较低分辨率时,物体轮廓是图像识别的主要影响因素。围绕轮廓增强进行图像处理是提高物体识别准确率的有效方法。在较高分辨率时,突出物体的轮廓与边缘信息是提高识别准确率的有效手段。所以提出能够对边缘进行精细分割的分割算法与相应的图像增强策略是有限分辨率下提高物体识别准确率的重要研究内容之一。

值得注意的是,笔者采用直接去除背景的方式来突出物体轮廓,而没有采用弱化背景信息或其他保留背景的处理方法。一方面是因为当前假体只能提供较低分辨率的视觉感受,背景信息的作用并不突出;另一方面是考虑到去除背景减少了电极刺激的频率,高频刺激会导致电流通过电极的峰值更高、更频繁,这可能会引起电极和周围组织之间的耦合效应,增加电极腐蚀或损伤的风险^[4]。

3.3 图像分辨率的影响

试验显示图像分辨率对物体识别准确率有影响。高分辨率对于视觉识别任务有积极的作用。不同分辨率下图像处理的有效性不同。在分辨率较低时,被试很难从图像中获取有用信息,同时背景信息会干扰被试的判断。突出物体的轮廓信息的图像处理方法能够显著提高物体的识别性能。但是随着分辨率持续提高到 48×48 和 64×64 ,高分辨率使被试能够获取更多的边缘信息,结合轮廓与边缘信息能够达到更好的识别性能。

3.4 轮廓增强的作用

轮廓可以提供物体的轮廓边界信息,帮助植入者理解物体的几何特征和形状。轮廓增强可以减少周围像素对轮廓灰度值的稀释作用。据被试反应,经过轮廓增强的图像,物体的轮廓信息更加突出。清晰的轮廓能够帮助被试更快更准确地做出判断,尤其是具有特殊轮廓的物体。因此,图像轮廓增强在SPV方面起着重要的作用,对于提高SPV的感知和理解能力具有显著的影响。

3.5 学习效应的影响

在本文试验中,共有60幅物体图像。被试仅被

告知图像中出现物体的种类,比如车、动物、人等。他们既没有看到实际物体,也没有看到原始图像。虽然我们采用拉丁方对被试进行分组试验,并根据艾宾浩斯遗忘曲线安排实验间隔来减少学习效应。但是对于PSM处理策略,其仅能提供物体的轮廓信息,而不具备物体内部的细节特征。在分辨率较高时,PSM提供的轮廓信息应该接近饱和。但是不同分辨率之间PSM的识别结果仍然具有显著性差异,这可能是因为在多种分辨率下的重复试验,使得被试对图像产生了一定的学习效应。在将来的研究中会采用更多更合理的图像来进行试验,减少学习效应产生的影响。

3.6 局限性

本文所提出的算法使用了Transformer进行建模,尽管其具有良好的性能,但较大的计算量使得模型无法对图像进行实时处理。今后的研究可以对模型进行压缩优化或选择轻量级模型结构,以获得更好的结果。现今视觉假体带来的视觉体验,会被扭曲、缺失以及形状不规整等多种因素左右,尽管在仿真试验中已尽可能地去贴近假体植入者的实际感受,但与真实患者所具有的视觉感知相比,仍存在真实性不足的问题。然而,因本研究侧重基于显著性图像优化策略提升物体识别性能,故未考量上述影响因素。本文提出的处理策略选择去除背景信息与目标放大来突出物体轮廓,这导致了物体空间位置的丢失,可能会使患者迷失方向。

4 结论

显著目标提取在视觉假体中起着关键作用,它使得视觉假体能够更加智能地感知和理解周围环境,为各种应用场景提供更加精准和高效的视觉信息处理。因此,本文提出了一种有效的显著目标检测模型,并以此模型为基础提出一种图像处理策略,旨在优化SPV下视觉信息的展示方式。试验结果显示,所提出的显著目标检测模型在提取目标物体方面成效显著。经SPV下的心理物理学试验验证,本研究的图像增强策略对低分辨率假体视觉下物体识别性能的提升作用明显。随着视觉假体硬件,特别是电极植入模块的进一步发展,希望本文提出的算法和图像处理策略能够在提高假体性能方面发挥重要的作用。

【参考文献】

- [1] Hallum LE, Dakin SC. Retinal implantation of electronic vision prostheses to treat retinitis pigmentosa: a systematic review[J]. *Transl Vis Sci Technol*, 2021, 10(10): 8.
- [2] Zrenner E, Bartz-Schmidt KU, Besch D, et al. The subretinal implant ALPHA: implantation and functional results[M]//Gabel

- VP. Artificial Vision: A Clinical Guide. Cham: Springer International Publishing, 2017: 65-83.
- [3] Zhou DD, Dorn JD, Greenberg RJ. The Argus® II retinal prosthesis system: an overview[C]//2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW). Piscataway, NJ, USA: IEEE, 2013: 1-6.
- [4] Wu KY, Mina M, Sahyoun JY, et al. Retinal prostheses: engineering and clinical perspectives for vision restoration[J]. Sensors (Basel), 2023, 23(13): 5782.
- [5] 赵瑛. 视觉假体中最小信息需求研究[D]. 上海: 上海交通大学, 2010.
Zhao Y. Research on minimum information requirements in visual prostheses[D]. Shanghai: Shanghai Jiao Tong University, 2010.
- [6] Wang J, Zhao RF, Li PT, et al. Clinical progress and optimization of information processing in artificial visual prostheses[J]. Sensors (Basel), 2022, 22(17): 6544.
- [7] 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.
- [8] Wang J, Li H, Fu WZ, et al. Image processing strategies based on a visual saliency model for object recognition under simulated prosthetic vision[J]. Artif Organs, 2016, 40(1): 94-100.
- [9] Li H, Su XF, Wang J, et al. Image processing strategies based on saliency segmentation for object recognition under simulated prosthetic vision[J]. Artif Intell Med, 2018, 84: 64-78.
- [10] Guo F, Yang Y, Xiao Y, et al. Recognition of moving object in high dynamic scene for visual prosthesis[J]. IEICE Trans Inf Syst, 2019, E102.D(7): 1321-1331.
- [11] Wang J, Zhu HY, Liu JY, et al. The application of computer vision to visual prosthesis[J]. Artif Organs, 2021, 45(10): 1141-1154.
- [12] 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: ACM, 2021: 45-54.
- [13] 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.
- [14] Wang WG, Lai QX, Fu HZ, et al. Salient object detection in the deep learning era: an in-depth survey[J]. IEEE Trans Pattern Anal Mach Intell, 2022, 44(6): 3239-3259.
- [15] Wang Y, Wang RL, Fan X, et al. Pixels, regions, and objects: multiple enhancement for salient object detection[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2023: 10031-10040.
- [16] Wu ZY, Li S, Chen CLZ, et al. Salient object detection via dynamic scale routing[J]. IEEE Trans Image Process, 2022, 31: 6649-6663.
- [17] Zhu JY, Qin XB, Elsaddik A. DC-Net: divide-and-conquer for salient object detection[J]. Pattern Recognit, 2025, 157: 110903.
- [18] Pang YW, Zhao XQ, Zhang LH, et al. Multi-scale interactive network for salient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 9410-9419.
- [19] Wei J, Wang SH, Huang QF. F³Net: fusion, feedback and focus for salient object detection[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence and the Thirty-Second Conference on Innovative Applications of Artificial Intelligence and the Tenth Symposium on Educational Advances in Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2020: 12321-12328.
- [20] Qin XB, Zhang ZC, Huang CY, et al. U²-Net: going deeper with nested U-structure for salient object detection[J]. Pattern Recognit, 2020, 106: 107404.
- [21] Zhao Y, Lu YY, Tian YK, et al. Image processing based recognition of images with a limited number of pixels using simulated prosthetic vision[J]. Inf Sci, 2010, 180(16): 2915-2924.
- [22] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at scale[EB/OL]. (2021-06-03). <https://arxiv.org/abs/2010.11929>.
- [23] Xie SN, Tu ZW. Holistically-nested edge detection[J]. Int J Comput Vis, 2017, 125(1): 3-18.
- [24] Suzuki S, be K. Topological structural analysis of digitized binary images by border following[J]. Comput Vis Graph Image Process, 1985, 30(1): 32-46.
- [25] Wang CQ, Fang CM, Zou Y, et al. Artificial intelligence techniques for retinal prostheses: a comprehensive review and future direction [J]. J Neural Eng, 2023, 20(1): 011003.

(编辑:薛泽玲)