

# 自适应核学习的交互式图像分割算法

龙建武, 李继豪

(重庆理工大学计算机科学与工程学院, 重庆 400054)

**摘要:** 针对现有大多数交互式分割方法在原始特征空间易受噪声干扰及非凸结构影响, 致使分割性能受限的问题, 提出一种自适应核学习的交互式图像分割算法。首先, 在 SLIC 超像素分割结果上融合用户标注的空间距离信息和像素邻域拓扑关系, 构建能量函数。其次, 引入核映射机制, 将原始数据嵌入高维特征空间, 增强线性可分性。接着, 基于 RBF 核函数的平滑性与正定性等特性, 设计优化目标函数, 并通过迭代优化策略动态调整核参数  $\sigma$ 。最后, 在 BSDS500 与 MSRC 数据集上, 采用交并比、信息差异、边界漂移误差和兰德指数等标准评估指标进行系统性实验。结果表明, 所提算法在综合评价指标上显著优于对比算法, 验证了其在处理复杂场景时的有效性与普适性。

**关键词:** 交互式图像分割; 超像素分割; 能量函数; 高斯核函数; 参数自适应优化

**中图分类号:** TP391

**文献标志码:** A

**DOI:** 10.11959/j.issn.1000-436x.2025123

## Interactive image segmentation algorithm based on adaptive kernel learning

LONG Jianwu, LI Jihao

College of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China

**Abstract:** To address the issue that most existing interactive image segmentation methods suffer from limited segmentation performance due to their susceptibility to noise interference and non-convex structure impacts in the original feature space, an adaptive kernel learning-based interactive image segmentation algorithm was proposed. Firstly, an energy function was constructed by integrating spatial distance information from user annotations on the results of SLIC superpixel segmentation with the pixel neighborhood topological relationships. Then, a kernel mapping mechanism was introduced to embed raw data into a high-dimensional feature space, enhancing linear separability. Subsequently, leveraging the smoothness and positive definiteness properties of RBF kernel functions, an optimized objective function was designed. Kernel parameter  $\sigma$  was dynamically adjusted through iterative optimization strategies. Finally, systematic experiments were conducted on BSDS500 and MSRC datasets using standard evaluation metrics, including intersection over union, variation of information, boundary display error, and rand index. Experimental results demonstrate that the proposed algorithm overperforms existing approaches across all evaluation metrics, achieving outstanding performance, and it validates the effectiveness and universality of the algorithm in handling complex scenarios.

**Keywords:** interactive image segmentation, superpixel segmentation, energy function, Gaussian kernel function, parameter adaptive optimization

收稿日期: 2025-04-25; 修回日期: 2025-06-23

基金项目: 重庆市教育委员会科学技术研究计划青年项目基金资助项目(No.KJQN202201148); 重庆市教育委员会人文社会科学研究青年项目基金资助项目(No.23SKGH263); 重庆理工大学研究生教育高质量发展行动计划基金资助项目(No.gzlcx20253270)

**Foundation Items:** The Young Scientists Project of Science and Technology Research Program of Chongqing Education Commission of China (No.KJQN202201148), Humanities and Social Sciences Research Program of Chongqing Municipal Education Commission (No.23SKGH263), The Funding Achievements of the Action Plan for High Quality Development of Graduate Education at Chongqing University of Technology (No.gzlcx20253270)

## 0 引言

图像分割是计算机视觉等学科领域的一项基础且重要的任务，其作用在于将图像划分为多个具有相似特性的区域，从而实现目标对象的识别、定位和理解，广泛应用于医学影像、自动驾驶、安防监控等领域，在医学影像诊断、人工智能、交通路况解析以及目标识别等诸多图像处理实际应用场景中扮演着不可或缺的角色。

然而，实际应用中图像分割面临诸多难题。一方面，图像通常存在复杂背景、多样纹理和不规则形状，使分割算法难以精准识别目标区域。另一方面，不同用户对同一图像的关注目标不同，有的用户希望分割出整棵树，有的用户只对树冠感兴趣，这种主观需求差异增加了分割难度。

为解决这些问题，交互式图像分割应运而生。它允许用户干预和控制图像分割过程，通过人机交互协同完成分割任务。用户直接参与处理，为算法提供关键先验信息和约束条件，显著提高了分割精度与可靠性。该方法尤其适用于形状复杂、纹理相似或背景干扰大的目标对象，能有效化解传统自动分割方法的歧义问题。同时，它增强了算法对不同用户需求的适应性，用户可按需灵活调整分割结果，提升了用户体验和工作效率。

不过，交互式图像分割在处理复杂图像时仍面临挑战。在医学影像中，肾肿瘤、皮肤病变、视网膜等图像的目标对象占比极小，却需要精确分割以辅助医生诊断和治疗规划；在遥感图像中，建筑、道路、植被等目标在高分辨率图像中虽小但细节丰富。交互式方法在精准处理这些细微特征上仍存在困难。

针对上述交互式图像分割存在的问题提出了诸多方法。文献[1]提出一种SAM (segment anything model) 分割模型，通过用户提供的点、框等提示快速生成分割结果，支持多种图像类型，如自然图像、医疗图像等。文献[2]提出一种基于简单非层次化 ViT (vision transformer) 的交互式图像分割方法 SimpleClick，通过引入对称的 patch 嵌入层将点击信息编码到骨干网络中，仅对骨干网络进行少量修改，利用预训练的掩码自编码器 (MAE) 实现最佳性能。文献[3]提出一种新的交互式图像分割方法 SegNext，通过引入密集表示和融合视觉提示，实现了低时延、高质量和多样化提示支持的交互式分割。文献[4]提出一种基于像素级深度强化

学习的医学图像精准鲁棒分割方法，通过将图像中每个像素的分割决策建模为马尔可夫决策过程，并引入深度强化学习框架动态优化分割策略，实现了对复杂医学图像的像素级精准分割。文献[5]引入“掩码对齐”辅助任务，对文本词汇进行随机掩码处理并引导算法预测其语义，显式推动模型学习文本与视觉对象的细粒度对应关系，结合核方法与跨模态对齐技术，显著提升交互式图像分割精度。文献[6]提出一种多交互特征学习策略，融合全时多模态基准，经核函数映射与多模态数据融合，显著提升交互式图像分割性能。

基于图论的算法在交互式图像分割领域同样应用广泛。Boykov 等<sup>[7]</sup>提出一种多维图像分割方法，其核心在于借助用户交互对目标与背景像素进行标记，并综合考量边界及区域信息，借助图像分割算法求解全局最优分割方案。该方法在照片处理、视频编辑以及医学图像分割等多个领域均展现出良好的应用效果。文献[8]提出一种基于图论优化的三维体数据表面分割方法，通过将体素间的拓扑关系建模为加权图结构，并引入最小割最大流算法求解全局最优分割曲面，显著提升了复杂解剖结构（如心脏腔室、脑皮层）的分割精度与拓扑一致性，尤其在低对比度区域和分支结构分割任务中表现出色。Liao 等<sup>[9]</sup>提出一种基于多智能体强化学习的迭代细化的交互式 3D 医学图像分割方法，旨在通过多智能体协作和迭代细化，提高分割的准确性和效率，实验结果表明，该方法在多个医学数据集上具有更少的交互次数和更快的收敛速度。文献[10]提出一种高效的交互式 3D 医学图像分割解决方案，旨在通过改进的 GrowCut 算法<sup>[7]</sup>和 Dijkstra 算法<sup>[11]</sup>，结合用户交互，实现快速准确的分割，实验结果表明，该方法在多个数据集上具有较高的分割精度和效率。

当前图像处理过程中，噪声干扰是导致分割效果不佳的常见因素。为应对这一挑战，本文算法借助核函数与自适应调整机制，可高效地应对噪声数据，减弱其对分割结果的负面影响，进而提升分割的精度。本文算法以交互形式融入用户提供的先验信息，能更精确地识别并分割出目标物体；即使在目标与背景特征相近的情形下，也能借助像素点之间的邻域信息和先验信息的距离关系，达到更为清晰的分割效果。此外，大部分算法在处理高维数据分割时效果不太好，自适应核学习算法则不同，它

利用核映射把数据转换至高维空间中实施分割操作,能够更高效地攻克高维数据分割难题,提升分割工作的效率与精准程度。结合上述方法,本文提出一种自适应核学习的交互式图像分割算法。本文主要贡献包括以下几点。

1) 本文融合先验信息的空间距离关系与像素邻域的拓扑结构设计能量函数,并引入核方法<sup>[12]</sup>将原始空间数据映射到高维空间,使原本线性不可分的数据在特征空间内线性可分。

2) 根据RBF(radial basis function)核(或高斯核)函数的平滑和正定等特性构建优化目标函数,使用Adam(adaptive moment estimation)算法<sup>[13]</sup>对RBF中的超参数 $\sigma$ 实现最优化选择,有效解决超参数 $\sigma$ 的选择难题,避免人工选择的主观性,增强模型的适应性和分类性能。

3) 利用SLIC(simple linear iterative clustering)超像素分割算法<sup>[14]</sup>优化计算复杂度,使算法的运行效率得到极大提升,并且能够更好地保持物体的轮廓和边界信息,确保分割结果的精准性。

## 1 相关工作

本文相关工作主要涉及核学习方法、经典核函数的选用,以及SLIC超像素分割算法的应用。

### 1.1 核学习方法

当像素的分布变得稀疏时,分类难度往往会降低。通常,具有高维度(即包含较多光谱带)的样本更容易展现出清晰的类别区分特征。核方法<sup>[15-18]</sup>的核心在于它能够把原始数据通过一种特定的映射 $\phi$ 嵌入一个更高维度的希尔伯特特征空间中,以寻找那些在原始空间中难以发现的有效分类超平面。特征空间映射如图1所示。

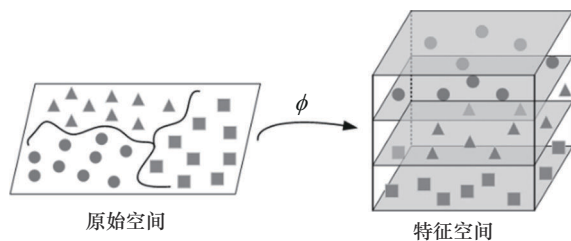


图1 特征空间映射

与之相比的其他同类交互式分割算法各有特点,具体介绍如下。基于图像分割的算法通过构建图结构并利用最小割方法优化分割,擅长处理复杂

的拓扑关系,但在大规模图像处理时计算复杂度较高。基于深度学习的分割算法则依赖于卷积神经网络在线学习图像特征,尽管处理复杂图像时表现出色,但需要大量标注数据进行训练且可解释性较弱。核方法以其灵活的核函数选择和高效的非线性映射机制,为图像分割提供一种独特的解决办法,尤其在处理复杂的非线性数据分布时具有明显优势。

### 1.2 经典核函数

核函数 $\kappa: X \times X \rightarrow \mathcal{R}$ ,其中 $X$ 是向量空间,无论是连续还是有限域的函数,都可以被分解并通过特征映射应用到一个希尔伯特空间 $H$ 中,并且对其2个参数 $x, z$ 都应用这个映射,然后计算 $H$ 中的内积<sup>[19]</sup>为

$$\kappa(x, z) = \langle \phi(x), \phi(z) \rangle \quad (1)$$

其中, $x, z \in X$ 。当且仅当一个对称函数对于空间 $X$ 中的任何有限子集 $\{x_1, x_2, \dots, x_N\}$ 形成的矩阵 $\mathbf{K} =$

$$\left[ \kappa(x_i, x_j) \right]_{1 \leq i, j \leq N}$$

是半正定的,即没有负特征值,

这个函数才是一个有效的核函数。

常见的核函数<sup>[20-24]</sup>如式(2)~式(4)所示。

线性核为

$$\kappa(x, z) = \langle x, z \rangle \quad (2)$$

多项式核为

$$\kappa(x, z) = (\langle x, z \rangle + 1)^t, t \in \mathbf{Z}^+ \quad (3)$$

高斯(RBF)核为

$$\kappa(x, z, \sigma) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right), \sigma > 0 \quad (4)$$

其中, $x, z$ 均为 $R^d$ 中的样本, $\sigma$ 和 $t$ 分别是RBF核与多项式核中的超参数。

特别注意的是,核矩阵的维度是 $N \times N$ ,其中每个元素 $\kappa_{ij}$ 包含了像素 $x_i$ 到 $x_j$ 的距离信息,这一距离通过适当的核函数来计算,且该函数满足核的特性。若采用线性核,则特征映射 $\phi$ 为恒等映射,即线性映射;否则,特征映射可能为非线性。

### 1.3 SLIC超像素分割算法

本文算法是基于像素级的处理方法,因其计算粒度过于精细,会显著增加计算复杂度,尤其在处理高分辨率图像时,实时性难以保障;其次,像素级操作难以有效捕捉图像语义结构,导致目标边缘分割粗糙。

针对上述问题，本文引入 SLIC 超像素分割算法进行优化处理，使算法运行效率得到极大提升，并能够较好地保持物体轮廓和边界信息，生成的超像素能够紧密黏附到图像边界，展现出优秀的边界依从性。此外，SLIC 算法生成的超像素形状规则、大小均匀，如同细胞一般紧凑整齐，SLIC 分割结果如图 2 所示。相较于像素分割，超像素 SLIC 分割能够保持物体轮廓的连续性，降低计算复杂度，提高分割效率和准确性，为后续图像分析提供更简单有效的特征表示。

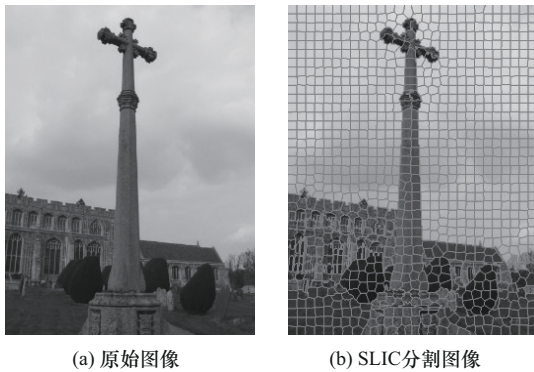


图 2 SLIC 分割结果

对于 SLIC 算法，超像素的数量对最终的分割结果具有不可忽视的影响。当超像素数量处于较低水平时，单个超像素块覆盖的面积较大。该情形下，算法能够较好地维持图像大尺度区域的划分，然而可能会遗漏部分细节和边界信息，故而适用于需要快速提取主要区域的场景。与之相反，若超像素数量较多，那么每个超像素块的面积会相应减小，此时算法能够更精准地捕捉图像中的细节变化和边缘信息，从而更好地适配复杂的纹理以及局部特征，但计算复杂度会大幅上升，算法的运行时间也会随之增加，尤其在处理大规模图像时，这一现象更为显著。因此，超像素数量的确定需要在分割精度与计算效率之间进行权衡，根据具体的应用需求及图像的固有特性来选取最优数量值，进而达到最佳的分割效果与性能平衡。

## 2 算法模型

### 2.1 问题与动机

当前，大部分现有交互式分割方法在原始空间中面临着难以线性分类的问题，像素分布往往呈现非凸结构，使构建有效分类超平面变得极其困难。

同时，RBF 核作为解决此类问题的重要方法，其性能依赖于超参数（核带宽  $\sigma$ ）的合理设置。目前，现有方法大多依赖于人工调整或经验选择来确定超参数，缺乏自适应优化机制。这种主观选择方式不仅效率低下，还可能导致模型泛化能力不足，尤其在处理高维数据时，参数敏感性加剧了分割结果的不稳定。此外，像素级处理方法虽然能够保留丰富的细节信息，但其计算复杂度会随图像分辨率的提高而呈指数增长，难以满足实时性需求；同时，像素级操作难以有效捕捉图像的全局语义结构，导致边界分割粗糙，在处理微小目标或复杂纹理时，分割精度显著下滑。

针对上述问题，本文进行相应改进。首先，通过高斯核函数将原始数据映射至高维空间，利用核技巧解决线性不可分的难题。这种映射不仅增强了类内数据的紧致性与类间数据的可分性，还通过局部敏感性自适应调整分类边界，有效抑制噪声干扰，从而提高分割的准确性和鲁棒性。其次，结合先验标注像素的距离和邻域拓扑关系构建能量函数，该能量函数不仅考虑当前像素点与已知标注像素点的距离关系，而且考虑到了毗邻像素点的约束，从而保证分片光滑性。之后，再结合 RBF 核函数的特性，即相同标签的像素点使核函数值趋于 1，不同标签的像素点使核函数值趋于 0，以此构建目标函数，并利用 Adam 算法最优化选择超参数  $\sigma$ ，从而避免人工选择的主观偏差。最后，采用 SLIC 算法将图像划分为紧凑均匀的超像素单元，显著降低了计算复杂度。超像素不仅保留了目标的轮廓和边界信息，还通过邻域拓扑关系增强了语义特征表示。邻域拓扑信息如图 3 所示，其中，图 3(b)是图 3(a)白色框内一个像素点  $x_i$  的邻域关系。

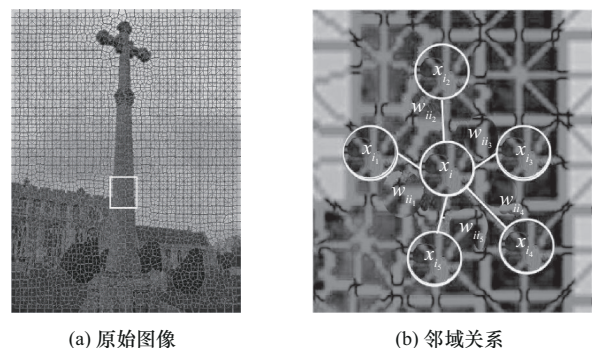


图 3 邻域拓扑信息

像素点 $x_i$ 的近邻像素点集合记为 $\Omega_i = \{x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}, x_{i_5}\}$ , 对应的权重分别为 $\{w_{ii_1}, w_{ii_2}, w_{ii_3}, w_{ii_4}, w_{ii_5}\}$ 。从空间拓扑结构来看, 这些近邻像素点紧密相邻, 其空间位置关系构成了约束的基本架构。基于该约束, 近邻像素点能够维持相对稳定的拓扑关联性, 为后续分割提供更高效的数据抽象, 有助于在保证分割精度的同时提高处理速度。

## 2.2 分割模型

在图像分割过程中, 传统欧几里得距离作为像素相似性度量时, 对噪声干扰和低对比度图像表现出显著的敏感性, 其固有局限性易导致算法收敛于局部最优解而非全局最优解<sup>[25]</sup>。针对这一瓶颈问题, 本文借鉴文献<sup>[26]</sup>的创新思路, 通过引入高斯核诱导的空间距离替代常规欧氏距离。此改进策略不仅有效克服了数据分布非线性带来的挑战, 同时显著增强了算法对图像噪声的鲁棒性, 实现了对复杂图像结构更精准的分割。

记输入图像 $I$ 中的数据集合 $x = \{x_1, x_2, \dots, x_N\}$ ,  $N$ 表示数据总量; 类别数量 $k \in \{1, 2, \dots, C\}$ ,  $C$ 表示类别总量; 类别标签 $y \in \{1, 2, \dots, C\}$ ;  $x = M \cup U$ , 其中 $M$ 表示已知标注像素集,  $U$ 表示未标注像素集;  $M = M_1 \cup M_2 \cup \dots \cup M_C$ , 第 $k$ 类已知标注像素集记为 $M_k = \{x_1^{(k)}, x_2^{(k)}, \dots, x_{|M_k|}^{(k)}\}$ , 属于第 $k$ 类数据的个数 $N_k = |M_k|$ 。

本文根据先验信息距离关系和像素邻域拓扑结构, 构建如下能量函数为

$$E(u) = E_D(u) + \delta E_S(u) \quad (5)$$

$$\text{s.t.} \begin{cases} \sum_{k=1}^C u_{ki} = 1, u_{ki} \in [0, 1] \\ u_{ki} = 1, x_i \in M_k \end{cases}$$

其中,  $E_D(u)$ 为数据项, 反映像素点与先验信息之间的几何关系;  $E_S(u)$ 为约束项, 表示像素点的邻域拓扑关系, 体现了像素点之间的连接性和局部结构信息;  $u_{ki}$ 表示像素点 $i$ 属于第 $k$ 类的概率;  $\delta$ 为平衡参数。数据项和约束项具体定义如下。

1) 数据项为

$$E_D(u) = \sum_{i=1}^N \sum_{k=1}^C u_{ki}^2 d_{ik}^2 \quad (6)$$

其中,  $d_{ik}^2$ 表示在特征空间中像素点 $i$ 到第 $k$ 类中心的距离, 计算式为

$$d_{ik}^2 = \|\phi(x_i) - \phi(v_k)\|_2^2 \quad (7)$$

式(6)可以计算出所有像素点到其对应类别中心的加权平方和, 其目的是调整像素点所属概率来最小化能量函数, 从而找到像素点的最佳分类。

2) 约束项为

$$E_S(u) = \sum_{i=1}^N \sum_{k=1}^C u_{ki}^2 \tilde{d}_{ik}^2 \quad (8)$$

其中,  $\tilde{d}_{ik}^2$ 表示在特征空间中像素点 $i$ 的邻域像素点到第 $k$ 类的距离, 计算式为

$$\tilde{d}_{ik}^2 = \sum_{j \in \Omega_{ij} \neq i} w_{ij} (1 - u_{kj})^2 d_{jk}^2 \quad (9)$$

其中,  $w_{ij}$ 是权重值, 定义为

$$w_{ij} = \frac{1}{\|p_i - p_j\|_2^2 \cdot \|x_i - x_j\|_2^2 + 1} \quad (10)$$

其中,  $p_i$ 和 $p_j$ 分别是像素点 $i$ 和 $j$ 的空间坐标;  $x_i$ 和 $x_j$ 分别是像素点 $i$ 和 $j$ 的颜色信息, 如RGB信息。 $d_{jk}^2$ 是在特征空间中近邻像素点 $j$ 到类别中心 $k$ 的距离, 计算式为

$$d_{jk}^2 = \|\phi(x_j) - \phi(v_k)\|_2^2 \quad (11)$$

约束项式(8)量化描述了距离 $\tilde{d}_{ik}^2$ 对像素点 $i$ 从属第 $k$ 类概率所产生的影响, 以下从4个方面对其进行讨论。

① 当邻域像素点 $j$ 全部属于第 $k$ 类时, 可知 $d_{jk}^2$ 小,  $u_{kj}$ 大即 $(1 - u_{kj})^2$ 小, 导致 $\tilde{d}_{ik}^2$ 小, 那么像素点 $i$ 属于第 $k$ 类的概率就大。

② 当邻域像素点 $j$ 全部不属于第 $k$ 类时, 可知 $d_{jk}^2$ 大,  $u_{kj}$ 小即 $(1 - u_{kj})^2$ 大, 导致 $\tilde{d}_{ik}^2$ 大, 那么像素点 $i$ 属于第 $k$ 类的概率就小。

③ 当邻域像素点 $j$ 大部分属于第 $k$ 类时, 可知 $d_{jk}^2$ 偏小,  $u_{kj}$ 偏大即 $(1 - u_{kj})^2$ 偏小, 导致 $\tilde{d}_{ik}^2$ 偏小, 那么像素点 $i$ 属于第 $k$ 类的概率就大。

④ 当邻域像素点 $j$ 大部分不属于第 $k$ 类时, 可知:  $d_{jk}^2$ 偏大,  $u_{kj}$ 偏小即 $(1 - u_{kj})^2$ 偏大, 导致 $\tilde{d}_{ik}^2$ 偏大, 那么, 像素点 $i$ 属于第 $k$ 类的概率就小。

通过上述分析, 本文设计的能量函数不仅考虑了像素点与类别之间的相似度, 即通过数据项来最小化数据点到其对应类别的距离, 从而确保模型对数据有良好的拟合; 同时考虑了模型的复杂度和泛

化能力, 对邻域信息通过式(8)进行约束, 避免过拟合。

### 2.3 RBF核函数特征分析

RBF核函数在机器学习中被广泛应用, 本文将围绕RBF核函数开展更深层次的理论分析。

1) 对特征空间中同一像素点的内积进行计算:  
 $\langle \phi(x), \phi(x) \rangle = \kappa(x, x) = 1$ 。

2) 计算特征空间中2个不同像素点 $\phi(x_i)$ 与 $\phi(x_j)$ 的距离为

$$\begin{aligned} \|\phi(x_i) - \phi(x_j)\|_2^2 &= \langle \phi(x_i), \phi(x_i) \rangle - 2\langle \phi(x_i), \phi(x_j) \rangle + \\ &\langle \phi(x_j), \phi(x_j) \rangle = 2 - 2\kappa(x_i, x_j) \end{aligned} \quad (12)$$

其中,  $\kappa(x_i, x_j) \in (0, 1)$ , 可知 $\|\phi(x_i) - \phi(x_j)\|_2^2 \in [0, 2]$ 。

3) 基于特征空间中像素点之间的距离, 可以推导出核函数值的变化规律如下。

① 当2个测试数据点 $x$ 和 $z$ 具有相同标签时, 意味着它们在特征空间中距离较近, 在这种情况下,  $\|\phi(x) - \phi(z)\|_2^2 \rightarrow 0$ , 可得 $\kappa(x, z, \sigma) \rightarrow 1$ 。

② 当2个测试数据点 $x$ 和 $z$ 具有不同标签时, 意味着它们在特征空间中距离较远, 此时,  $\|\phi(x) - \phi(z)\|_2^2 \rightarrow 2$ , 可得 $\kappa(x, z, \sigma) \rightarrow 0$ 。

综上所述, 对于同一标签的像素点, RBF核函数应该输出一个趋于1的值, 从而强调它们之间的相似性; 对于不同标签的像素点, RBF核函数应该输出一个趋于0的值, 从而突出它们之间的差异性。

### 2.4 核参数 $\sigma$ 最优化选择算法

结合RBF核函数以上特性, 本文构建最优化选择 $\sigma$ 的目标函数如下。

1) 同一类样本类内均值 $\kappa_w(\sigma)$ 为

$$\kappa_w(\sigma) = \frac{1}{\sum_{i=1}^C |M_i|^2} \sum_{i=1}^C \sum_{1x \in M_i} \sum_{z \in M_i} (1 - \kappa(x, z, \sigma)) \quad (13)$$

由RBF核函数特性知, 当2个样本为相同标签时, 应该使 $\kappa_w(\sigma)$ 接近于0。

2) 不同类别样本类间均值 $\kappa_b(\sigma)$ 为

$$\kappa_b(\sigma) = \frac{1}{\sum_{i=1}^C \sum_{j=1}^C |M_i| |M_j|} \sum_{i=1}^C \sum_{1x \in M_i} \sum_{z \in M_j} \kappa(x, z, \sigma) \quad (14)$$

同理, 当2个样本为不同标签时, 应使 $\kappa_w(\sigma)$ 接近于0。因此, 本文可以通过解决以下优化问题来获得最优参数 $\sigma$ , 目标函数 $J(\sigma)$ 为

$$\begin{aligned} J(\sigma) &= \kappa_w(\sigma) + \kappa_b(\sigma) = \\ &\frac{1}{\sum_{i=1}^C \sum_{j=1}^C |M_i| |M_j|} \sum_{i=1}^C \sum_{1x \in M_i} \sum_{z \in M_j} \kappa(x, z, \sigma) + \\ &1 - \frac{1}{\sum_{i=1}^C |M_i|^2} \sum_{i=1}^C \sum_{1x \in M_i} \sum_{z \in M_i} \kappa(x, z, \sigma) \end{aligned} \quad (15)$$

最优参数 $\sigma^*$ 为

$$\sigma^* = \arg \min_{\sigma} (J(\sigma)) \quad (16)$$

随后, 本文采用Adam算法对参数 $\sigma$ 进行优化选择, 更新式为

$$\sigma_t = \sigma_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \varepsilon}}, \eta > 0 \quad (17)$$

其中,  $t$ 表示当前时间步,  $\eta$ 表示学习率,  $\sqrt{\hat{v}_t}$ 是修正后对梯度的一个平滑估计, 即

$$\hat{v}_t = \frac{v_t}{(1 - \beta_2^t)} \quad (18)$$

其中,  $v_t$ 是修正的二阶矩估计, 即

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (19)$$

$\hat{m}_t$ 表示当前时间步 $t$ 修正后的一阶矩估计, 即

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (20)$$

其中,  $m_t$ 为未修正的一阶矩估计, 即

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (21)$$

其中,  $\beta_1$ 、 $\beta_2$ 表示矩估计的指数衰减率, 分别取0.9、0.999;  $g_t$ 表示当前时间步 $t$ 的梯度, 即

$$g_t = \nabla_{\sigma} J_t(\sigma_{t-1}) \quad (22)$$

其中,  $\nabla_{\sigma} J_t(\sigma_{t-1})$ 是对目标函数求梯度, 即

$$\begin{aligned} \nabla J(\sigma) &= \frac{1}{\sum_{i=1}^C \sum_{j=1}^C |M_i| |M_j|} \sum_{i=1}^C \sum_{1x \in M_i} \sum_{z \in M_j} \nabla_{\sigma} \kappa(x, z, \sigma) - \\ &\frac{1}{\sum_{i=1}^C |M_i|^2} \sum_{i=1}^C \sum_{1x \in M_i} \sum_{z \in M_i} \nabla_{\sigma} \kappa(x, z, \sigma) \end{aligned} \quad (23)$$

RBF 核函数的梯度  $\nabla_{\sigma} \kappa(x, z, \sigma)$  计算式为

$$\begin{aligned} \nabla_{\sigma} \kappa(x, z, \sigma) &= \frac{\partial}{\partial \sigma} (\kappa(x, z, \sigma)) = \\ \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right) \frac{\partial}{\partial \sigma} \left(-\frac{\|x - z\|^2}{2\sigma^2}\right) &= \\ \exp\left\{-\frac{\|x - z\|^2}{2\sigma^2}\right\} \frac{\|x - z\|^2}{\sigma^3} \end{aligned} \quad (24)$$

本文最终采用 Adam 算法来最优化选择, 这种方法不仅减少了人工试错的时间, 还能更精确地确定最优值, 从而提高模型的分类性能。核参数最优化选择算法如算法 1 所示。

**算法 1** 核参数最优化选择算法

**输入** 已知标签像素  $M$ , 类别数  $C$

**输出** 最优化超参数  $\sigma$

- ① 初始化: 优化参数  $\sigma_0 = 0$ ; 一阶矩估计  $m_0 = 0$ ; 二阶矩估计  $v_0 = 0$ ;  $\beta_1 = 0.9$ ;  $\beta_2 = 0.999$ ; 防止除以零的微小常数  $\varepsilon = 10^{-6}$ ; 时间步  $t = 0$ ; 收敛阈值  $\gamma$
- ② while  $|\sigma_{t+1} - \sigma_t| > \gamma$  do
- ③  $t = t + 1$
- ④ 根据式(22)更新梯度
- ⑤ 根据式(21)更新一阶矩估计
- ⑥ 根据式(19)更新二阶矩估计
- ⑦ 根据式(20)修正一阶矩估计
- ⑧ 根据式(18)修正二阶矩估计
- ⑨ 根据式(17)获取  $\sigma_t$
- ⑩ end while
- ⑪ return  $\sigma_t$

**2.5 模型求解**

求解本模型中像素点  $i$  属于第  $k$  类的最大概率采用拉格朗日乘数法得到, 主要过程如下。

1) 构造拉格朗日函数计算式为

$$\min_{(u,v)} \left\{ E(u) + \sum_{i=1}^N \lambda_i \sum_{k=1}^C (u_{ki} - 1) \right\} \quad (25)$$

其中,  $\lambda_i$  是拉格朗日乘子。由于类别中心  $v$  会随着分割结果  $u$  变化而改变, 因此, 本文采用交替优化方案对其逐步求解。

2) 固定  $v$  求解  $u$ , 原问题等价于

$$\min_u \left\{ E(u) + \sum_{i=1}^N \lambda_i \sum_{k=1}^C (u_{ki} - 1) \right\} \quad (26)$$

对  $u_{ki}$  求偏导, 并令导数为 0, 即

$$\begin{aligned} 0 &= 2u_{ki}(1 - \kappa(x_i, v_k)) + \\ 2\delta u_{ki} \sum_{\substack{j \in \Omega_i \\ j \neq i}} w_{ij} (1 - u_{kj})^2 (1 - \kappa(x_j, v_k)) &+ \lambda_i \end{aligned} \quad (27)$$

结合约束条件  $\sum_{k=1}^C u_{ki} = 1$ , 解得  $u_{ki}$  为

$$u_{ki} = \left[ \frac{(1 - \kappa(x_i, v_k)) + \delta \sum_{\substack{j \in \Omega_i \\ j \neq i}} w_{ij} (1 - u_{kj})^2 (1 - \kappa(x_j, v_k))}{\sum_{l=1}^C (1 - \kappa(x_i, v_l)) + \delta \sum_{\substack{j \in \Omega_i \\ j \neq i}} w_{ij} (1 - u_{lj})^2 (1 - \kappa(x_j, v_l))} \right]^{-1} \quad (28)$$

3) 固定  $u$  求解  $v$ , 原问题等价于

$$\min_v \{E(u)\} \quad (29)$$

对  $v_k$  求偏导, 并令导数为 0, 即

$$\begin{aligned} 0 &= \sum_{i=1}^N u_{ki}^2 \kappa(x_i, v_k) (x_i - v_k) + \\ \delta \sum_{i=1}^N u_{ki}^2 \sum_{\substack{j \in \Omega_i \\ j \neq i}} w_{ij} (1 - u_{kj})^2 \kappa(x_j, v_k) (x_j - v_k) \end{aligned} \quad (30)$$

解得  $v_k$  为

$$v_k = \frac{\sum_{i=1}^N u_{ki}^2 [\kappa(x_i, v_k) x_i + \delta \sum_{\substack{j \in \Omega_i \\ j \neq i}} w_{ij} (1 - u_{kj})^2 \kappa(x_j, v_k) x_j]}{\sum_{i=1}^N u_{ki}^2 [\kappa(x_i, v_k) + \delta \sum_{\substack{j \in \Omega_i \\ j \neq i}} w_{ij} (1 - u_{kj})^2 \kappa(x_j, v_k)]} \quad (31)$$

4) 最后, 获得像素点  $i$  的从属标签为

$$y = \arg \max_{k \in \{1, 2, \dots, C\}} \{u_{ki}\} \quad (32)$$

综上, 本文使用拉格朗日乘数法来求解最大概率。首先, 定义拉格朗日函数, 其结合了能量函数  $E(u)$  并引入拉格朗日乘数  $\lambda_i$  来处理权重为 1 的约束; 其次, 分别对  $u$  和  $v$  求偏导, 并令导数等于 0; 最后, 得到每个像素点  $i$  属于第  $k$  类的最大概率。此过程对所有像素点和所有类别进行迭代计算, 直到收敛得到最优解。

本文在算法 2 中总结了最小化能量函数 (式(5)) 的求解过程, 并结合算法 1 得到最终的图像分割结果。

**算法 2** 自适应核学习算法

**输入** 图像  $I$ , 用户交互信息

**输出** 分割结果

初始化: 收敛阈值  $\eta$ ; 迭代次数  $t = 0$ ; 平衡参

数  $\delta = 1$ ；通过 SLIC 超像素算法分割图像得  $I_S$

- ①根据算法 1 计算最优核参数  $\sigma^*$
- ②根据交互信息获得初始概率  $u^{(0)}$ ，其中，  
 $u_{ki} = 1, x_i \in M_k$
- ③根据式(31)计算中心  $v^{(0)}$
- ④while  $\|u^{(t)} - u^{(t-1)}\|_2^2 < \eta$  do
- ⑤  $t = t + 1$
- ⑥ 根据式(28)更新概率  $u^{(t)}$
- ⑦ 根据式(31)更新中心  $v^{(t)}$
- ⑧end while
- ⑨根据式(32)获取从属标签  $y$
- ⑩return  $y$

### 3 算法实验

为检验本文算法的分割性能，本节将其与 GRO (grow cut) [27]、RW (random walk) [28]、NRW (normalized random walk) [29]、NLRW (normalized lazy random walk) [30]、LC (laplacian coordinate) [31]、MRF (multi-layer random forest) [32]和 GSSR (global seed information and sparse regular) [33]这 7 种交互式图像分割算法进行比较。这些算法均在 MATLAB2020b 编程环境中实现，且均已开源，并且所有测试均使用上述算法的默认参数进行，本文算法的超参数主要是式(5)中的平衡常数  $\sigma$  和超像素数量，分别取 1.00 和 1 000。

#### 3.1 评价指标与数据集

为全面评估算法的分割性能，本文选取 RI (rand index)、VoI (variation of information)、BDE (boundary display error) 和 IoU (intersection over union) 作为评价指标，为算法提供全面且客观的评价。在实验验证中，本文选择了 2 个具有代表性的开源数据集：BSD 数据集和 MSRC 数据集。BSD 数据集包含 500 张高质量的自然场景图片，MSRC 数据集包含 50 张细节丰富的自然图片。这 2 个数据集均提供了精确的真实分割标注，并且涵盖了多种复杂的场景。

#### 3.2 BSD 数据集实验

为评估不同算法的性能，将文献[27-33]所涉及的分割算法与本文算法应用于 BSD 数据集。各算法在 BSD 数据集上的实验结果如表 1 所示，表中加粗数字表示该指标的最优值。部分分割结果如图 4 所示。图 4 中，第一行为原始图像，第二行为文

献[27]算法分割结果，第三行为文献[28]算法分割结果，第四行为文献[29]算法分割结果，第五行为文献[30]算法分割结果，第六行为文献[31]算法分割结果，第七行为文献[32]算法分割结果，第八行为文献[33]算法分割结果，第九行为本文算法分割结果。

表 1 各算法在 BSD 数据集上的实验结果

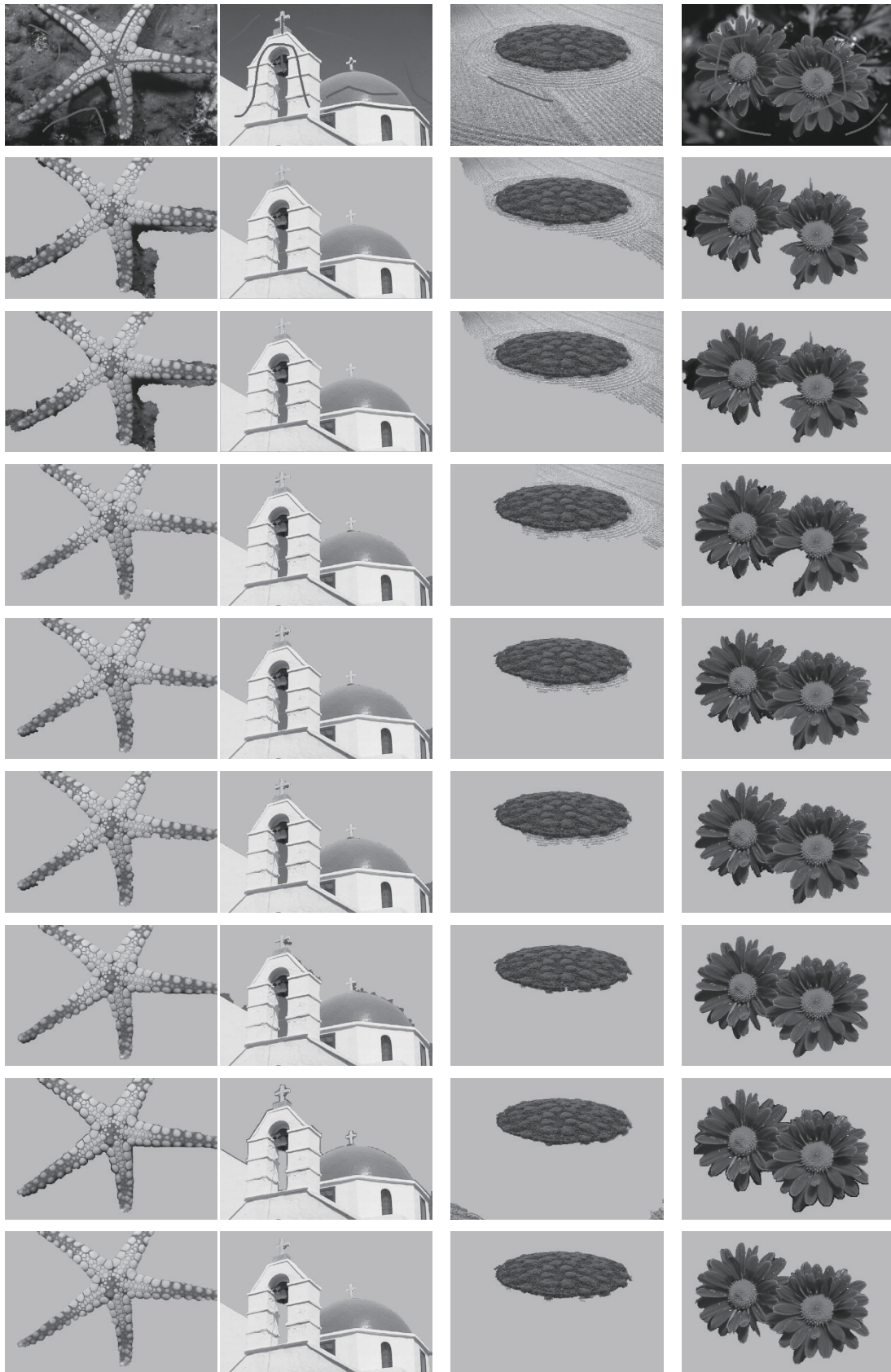
算法	mIoU(↑)	mVoI(↓)	mBDE(↓)	mRI(↑)
GRO	0.803 1	0.468 7	13.110 5	0.889 1
LC	0.837 5	0.314 3	6.359 1	0.930 5
RW	0.849 0	0.366 4	8.313 2	0.915 7
NRW	0.882 6	0.272 0	4.649 9	0.943 9
NLRW	0.883 1	0.270 3	4.621 6	0.944 2
MRF	0.859 3	0.288 3	6.328 6	0.942 9
GSSR	0.877 3	0.271 2	5.276 4	0.949 4
本文算法	<b>0.920 1</b>	<b>0.182 6</b>	<b>2.892 0</b>	<b>0.965 4</b>

从表 1 所呈现的数据可知，本文算法在平均交并比 (mIoU) 和平均兰德指数 (mRI) 上分别提升了 0.064 和 0.035，在平均边界位移误差 (mBDE) 和平均信息变化量 (mVoI) 上分别降低了 4.059 和 0.139，并且在各个评价指标维度上的表现均全面超越了其他分割方法。这一现象有力地表明，在 BSD 数据集这一特定场景下，无论是在图像的复杂背景还是在目标形态多样的情况下，本文算法都能保持较为一致的分割效果，为后续的图像分析和处理提供了更为可靠的基础。

通过观察图 4 中展示的分割可视化结果可知，本文通过核技巧这一关键特性，在处理具有丰富纹理的图像上显著提升了分割准确性，有效减少了误分情况，例如，在第 3 张草坪图的分割结果上，大部分对比算法对草坪的边缘分割细节不足，而本文算法可以准确地将草坪目标从地毯背景中分割出来，展现了本文算法在处理复杂数据集时优势更为明显。

#### 3.3 MSRC 数据集实验

各算法在 MSRC 数据集上的实验结果如表 2 所示，表中加粗数字表示该指标的最优值。部分分割结果如图 5 所示，图 5 中对比算法按行排序同图 4。通过表 2 中各项性能指标数据可以发现，本



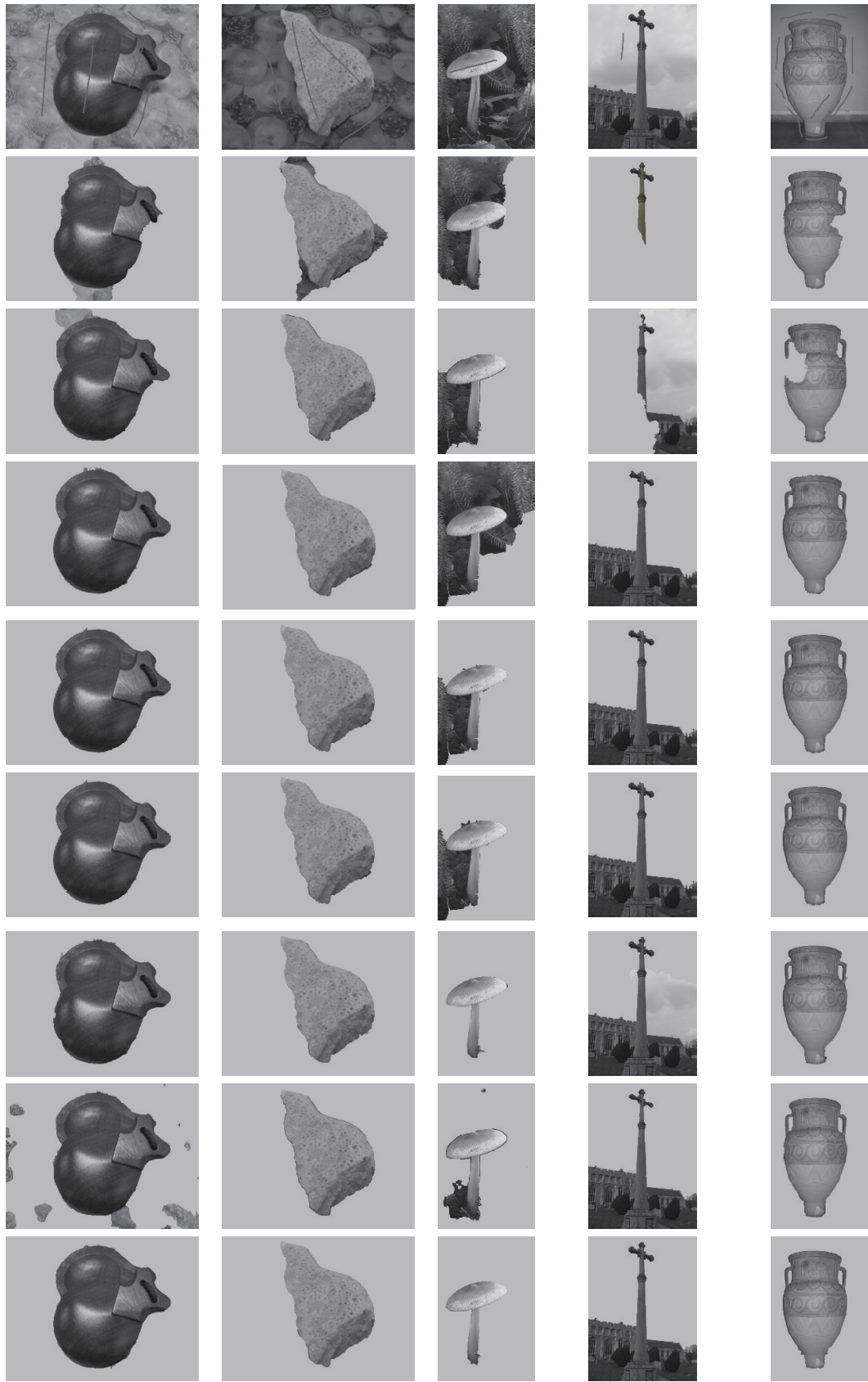
(a) 12003分割结果

(b) 118035分割结果

(c) 86016分割结果

(d) 124084分割结果

图4 BSD数据集上的部分分割结果



(a)music分割结果

(b)stone2分割结果

(c)208001分割结果

(d)cross分割结果

(e)227092分割结果

图5 MSRC数据集上的部分分割结果

文算法在 mIoU 指标和 mRI 指标上分别提升了 0.113 和 0.066, 在 mBDE 和 mVoI 指标上分别降低了 6.138 和 0.230。对图 5 可视化结果进行分析发现, 本文算法在处理具有复杂背景的图片上均展现出显著优势, 例如, 在第 4 张建筑图中, 由于待分割的建筑目标较为复杂, 其他几种对比算法在该图上难以实现精确分割, 而本文算法能够准确地将该建筑从背景中分割出来, 体现出在处理复杂图像中更具备出色的竞争力。

表 2 各算法在 MSRC 数据集上的实验结果

算法	mIoU(↑)	mVoI(↓)	mBDE(↓)	mRI(↑)
GRO	0.676 7	0.711 2	18.493 1	0.797 4
LC	0.789 8	0.514 2	8.190 3	0.892 1
RW	0.843 0	0.390 5	8.726 7	0.904 1
NRW	0.914 7	0.240 2	2.603 5	0.956 4
NLRW	0.912 8	0.245 6	2.720 6	0.955 4
MRF	0.932 1	0.244 8	3.966 9	0.953 5
GSSR	0.909 8	0.291 9	6.798 0	0.953 6
本文算法	<b>0.967 0</b>	<b>0.147 1</b>	<b>1.219 2</b>	<b>0.981 6</b>

### 3.4 多分类实验

本文算法在多分类任务的应用场景中依然能展现出良好的适用性, 应用效果较为理想。

图 6 呈现了多类分割结果, 共包含 6 张图片, 以三行两列的形式排列。具体而言, 第一行左侧是已标注的原始图像, 算法将其分为月亮、树木以及天空 3 类的模式, 这种分割方式较为简洁, 对于那些仅需对少量对象或区域进行分辨的场景而言, 具有较高的契合度, 能够在满足基本需求的同时, 降低分割的复杂度与计算成本。第二行左侧是山峦起伏景象, 颜色有层次感, 近处深灰, 远处浅灰; 右侧分割结果以不同灰度代表不同地形或物体类别, 使山脉结构和组成更清晰。第三行左侧是包含山脉、湖泊和森林的原始图像, 右侧分割结果用不同灰度清晰划分区域, 明确各元素分布。本文算法在各类分割任务中均能有效、精确地分割不同类别目标, 展现出强适应性。



(a) 原始图像 (b) 分割可视化结果

图 6 多类分割结果

## 4 消融研究及讨论

### 4.1 平衡参数 $\delta$

式(5)中  $\delta$  选择不同值对本文分割算法是有影响的, 超参数  $\delta$  对分割性能的影响如表 3 所示, 表中加粗数字表示该指标的最优值。实验选取  $\delta$  分别为 0.96、0.98、1.00、1.02、1.04。由表 3 可知, 当  $\delta = 1.00$  时, 4 项指标均达到最优, 模型性能最佳; 随着  $\delta$  从 1.00 增加到 1.04, 所有指标均逐渐变差,  $\delta$  过大导致模型过平滑或欠拟合, 影响分割精度和边界检测; 当  $\delta < 1.00$  时, 性能显著劣于  $\delta = 1.00$ , 但优于  $\delta > 1.00$  的情况。

表 3 超参数  $\delta$  对分割性能的影响

$\delta$	mIoU(↑)	mVoI(↓)	mBDE(↓)	mRI(↑)
0.96	0.823 0	0.380 4	7.101 8	0.931 5
0.98	0.822 4	0.381 5	7.119 4	0.931 1
1.00	<b>0.920 1</b>	<b>0.182 6</b>	<b>2.892 0</b>	<b>0.965 4</b>
1.02	0.821 9	0.381 8	7.104 8	0.930 8
1.04	0.821 8	0.382 1	7.106 9	0.930 5

基于前文对超参数进行的实验与分析过程, 本文首先明确寻找能使模型在性能评价指标上达到最佳平衡的超参数取值。通过对表 3 中实验数据的分析和反复验证发现, 当  $\delta = 1.00$  时, 模型在评价指标上均展现出了卓越性能。这一发现不仅为算法参数的合理确定提供了可靠的实验依据, 也为后续算法

的进一步优化和应用拓展提供了重要的参考方向。

## 4.2 超像素数量

针对 SLIC 超像素分割算法中超像素个数选择问题, 本文进行消融研究, 以了解不同超像素个数设置对本文算法性能的影响。超像素个数对分割性能的影响如表 4 所示, 表中加粗数字表示该指标的最优值。

表 4 超像素个数对分割性能的影响

$N$	mIoU(↑)	mVoI(↓)	mBDE(↓)	mRI(↑)
400	0.866 0	0.327 9	4.566 9	0.946 9
600	0.830 2	0.364 9	5.987 4	0.934 8
800	0.837 6	0.350 1	6.037 2	0.939 9
1 000	<b>0.920 1</b>	<b>0.182 6</b>	<b>2.892 0</b>	<b>0.965 4</b>
1 200	0.807 1	0.412 8	7.435 3	0.922 4

本文设定了超像素个数分别为 400、600、800、1 000 和 1 200, 并计算 mIoU、mVoI、mBDE 和 mRI 这 4 项指标数据。通过分析表 4, 可以观察到如下趋势。

1) mBDE 指标: 随着超像素个数的增加, mBDE 指标呈先降后升的趋势。当超像素个数为 1 000 时, mBDE 达到最低点, 此时的分割效果在边界离散误差方面表现最佳。然而, 当超像素个数进一步增加到 1 200 时, mBDE 指标显著上升, 提示分割效果有所下降。

2) mIoU 指标: mIoU 指标随着超像素个数的增加而逐渐上升, 在超像素个数为 1 000 时达到峰值, 之后略有下降。这一趋势表明, 当超像素个数为 1 000 时, 分割算法在交并比方面的表现最为优异。

3) mVoI 指标: mVoI 指标在整个实验过程中变化不大, 呈现出相对平稳的趋势。这表明本文的分割算法在处理不同数量的超像素时, 体积重叠度能保持稳定。

4) mRI 指标: 随着超像素个数的增加, mRI 指标呈先升后降的趋势, 当超像素个数为 1 000 时, 分割算法在区域相似性方面的性能最佳。

综合前文进行的实验发现, 在不同超像素数量设置下, 各算法的性能表现存在明显差异。通过对实验数据的统计与对比, 可以得出结论: 当超像素数量设置为 1 000 时, 各算法指标均达到了最佳状态。无论是从分割结果的准确性, 还是从

算法运行的稳定性来看, 该参数设置都展现出了明显的优势。这一发现为算法参数的合理选取提供了有力的实验依据, 在后续的算法应用中能够更加准确地选择合适的参数, 提高算法的性能和应用效果。

## 5 结束语

目前, 大多数现有交互式分割算法在原始空间中虽能较好地处理线性可分数据, 但对具有非线性结构分布的数据仍有一定的挑战性, 限制了其在复杂数据场景中的应用。针对该问题, 本文基于空间距离先验信息与像素邻域的拓扑关联, 构建能量函数模型。通过引入核映射技术, 将原始数据映射至高维特征空间, 实现数据线性可分, 并构造目标优化函数最优选择 RBF 核函数中的超参数  $\sigma$ , 有效解决了超参数选择的难题。尽管算法时间复杂度是一大阻碍, 但通过引入 SLIC 超像素分割技术, 显著降低了时间复杂度, 大幅提升了算法的运行效率。未来, 笔者将继续致力于提高算法的效率和精度, 积极拓展该算法在更多实际场景中的应用, 充分展现本文算法的长期发展潜力, 为图像交互式分割技术的发展做出更大贡献。

## 参考文献:

- [1] KIRILLOV A, MINTUN E, RAVI N, et al. Segment anything[C]//Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2023: 3992-4003.
- [2] LIU Q, XU Z L, BERTASIUS G, et al. SimpleClick: interactive image segmentation with simple vision transformers[C]//Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2023: 22233-22243.
- [3] LIU Q, CHO J, BANSAL M, et al. Rethinking interactive image segmentation with low latency, high quality, and diverse prompts[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2024: 3773-3782.
- [4] LIU Y, YUAN D, XU Z, et al. Pixel level deep reinforcement learning for accurate and robust medical image segmentation[J]. Scientific Reports, 2025, 15(1): 8213.
- [5] CHNG Y X, ZHENG H, HAN Y Z, et al. Mask grounding for referring image segmentation[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2024: 26563-26573.
- [6] LIU J Y, LIU Z, WU G Y, et al. Multi-interactive feature learning and a full-time multi-modality benchmark for image fusion and segmentation[C]//Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2023: 8081-8090.
- [7] BOYKOV Y Y, JOLLY M P. Interactive graph cuts for optimal bound-

- ary & region segmentation of objects in N-D images[C]//Proceedings of the Eighth IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2001: 105-112.
- [8] LI K, WU X D, CHEN D Z, et al. Optimal surface segmentation in volumetric images—a graph-theoretic approach[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(1): 119-134.
- [9] LIAO X, LI W H, XU Q S, et al. Iteratively-refined interactive 3D medical image segmentation with multi-agent reinforcement learning[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 9391-9399.
- [10] GAO Y, CHEN X H, YANG Q Z, et al. An effective and open source interactive 3D medical image segmentation solution[J]. Scientific Reports, 2024, 14(1): 1-19.
- [11] DIJKSTRA E W. A note on two problems in connexion with graphs[J]. Numerische Mathematik, 1959, 1(1): 269-271.
- [12] MÜLLER K R, MIKA S, TSUDA K, et al. Handbook of neural network signal processings[M]. Boca Raton: CRC Press, 2018.
- [13] KINGMA D P, BA J, HAMDAD M M. Adam: a method for stochastic optimization[J]. arXiv Preprint, arXiv: 1412.6980, 2014.
- [14] ACHANTA R, SHAJI A, SMITH K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274-2282.
- [15] SHAWE-TAYLOR J, CRISTIANINI N. Kernel methods for pattern analysis[M]. Cambridge: Cambridge University Press, 2004.
- [16] HOFMANN T, SCHÖLKOPF B, SMOLA A J. Kernel methods in machine learning[J]. The Annals of Statistics, 2008, 36(3): 1171-1220.
- [17] LAMPERT C H. Kernel methods in computer vision[J]. Foundations and Trends® in Computer Graphics and Vision, 2008, 4(3): 193-285.
- [18] HERBRICH R. Learning kernel classifiers: theory and algorithms[M]. Cambridge: MIT Press, 2001.
- [19] GLICK J R, GUJARATI T P, CÓRCOLES A D, et al. Covariant quantum kernels for data with group structure[J]. Nature Physics, 2024, 20(3): 479-483.
- [20] REDDY S, FARUQ M S U, LAKSHMI K R. Image segmentation by using linear spectral clustering[J]. Journal of Telecommunications System & Management, 2016, 5(3): 1-5.
- [21] TANG X T, ZHANG X F, GAO H J, et al. FCM-based image segmentation with kernel functions[C]//Proceedings of the 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC). Piscataway: IEEE Press, 2017: 916-919.
- [22] WU C M, ZHANG X. A novel kernelized total bregman divergence-driven possibilistic fuzzy clustering with multiple information constraints for image segmentation[J]. IEEE Transactions on Fuzzy Systems, 2022, 30(6): 1624-1639.
- [23] WU C M, WANG Z R. Quadratic surface center-based possibilistic fuzzy clustering with kernel metric and local information for image segmentation[J]. Multimedia Tools and Applications, 2024, 83(15): 44147-44191.
- [24] PEI Y G, ZHU D T. Local convergence of a trust-region algorithm with line search filter technique for nonlinear constrained optimization[J]. Applied Mathematics and Computation, 2016, 273: 797-808.
- [25] CHEN S C, ZHANG D Q. Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure[J]. IEEE Transactions on Systems, Man, and Cybernetics -Part B: Cybernetics, 2004, 34(4): 1907-1916.
- [26] LI C H, LIN C T, KUO B C, et al. An automatic method for selecting the parameter of the RBF kernel function to support vector machines[C]//Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium. Piscataway: IEEE Press, 2017: 1-4.
- [27] VEZHNEVETS V., KONOUCHE V. GrowCut: interactive multi-label N-D image segmentation by cellular automata[C]//Proceedings of Graphicon 2005. Piscataway: IEEE Press, 2005: 150-156.
- [28] GRADY L. Random walks for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(11): 1768-1783.
- [29] BAMPIS C G, MARAGOS P. Unifying the random walker algorithm and the SIR model for graph clustering and image segmentation[C]//Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP). Piscataway: IEEE Press, 2015: 2265-2269.
- [30] BAMPIS C G, MARAGOS P, BOVIK A C. Graph-driven diffusion and random walk schemes for image segmentation[J]. IEEE Transactions on Image Processing, 2017, 26(1): 35-50.
- [31] CASACA W, GOIS J P, BATAGELO H C, et al. Laplacian coordinates: theory and methods for seeded image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(8): 2665-2681.
- [32] SHAN Y L, MA Y, LIAO Y, et al. Interactive image segmentation based on multi-layer random forest classifiers[J]. Multimedia Tools and Applications, 2023, 82(15): 22469-22495.
- [33] LONG J W, LIU Y Q, ZHANG K X, et al. Interactive image segmentation combining global seeding and sparse local reconstruction[J]. Pattern Analysis and Applications, 2025, 28(2): 55.

#### [作者简介]



龙建武 (1984-), 男, 湖北恩施人, 博士, 重庆理工大学副教授、硕士生导师, 主要研究方向为图像处理与机器学习。



李继豪 (2001-), 男, 江西赣州人, 重庆理工大学硕士生, 主要研究方向为图像处理、机器学习、神经网络等。