

增强蒲公英算法优化乳腺癌图像多阈值分割^①

王正红, 王丹, 胡容俊

(北京大学 计算机科学技术学院, 吉林 132013)

通信作者: 王丹, E-mail: wangdanjl_jl@163.com



摘要: 针对显微镜下乳腺癌病理组织图像结构复杂, 细胞边界模糊等情况, 基于传统的阈值分割在乳腺癌图像的分割应用中不能很好地实现把病灶区准确分离开来的问题, 提出一种基于增强蒲公英优化算法 (IDO) 的乳腺癌图像多阈值分割方法. 该方法引入 IDO 计算类间方差的最大值 (Otsu) 作为目标函数寻找最佳阈值, IDO 建立回守策略解决传统蒲公英算法 (DO) 无限制搜索, 超出像素范围的问题; 引入对立式学习 (OBL) 避免算法陷入局部最优. 实验结果表明, 与哈里斯鹰算法 (HHO)、人工猩猩部队优化算法 (GTO)、传统蒲公英优化算法 (DO)、海洋捕食者算法 (MPA) 相比, 在相同阈值个数情况下 IDO 算法适应度值最大、收敛最快, 并且在峰值信噪比 (*PSNR*)、结构相似度 (*FSIM*)、特征相似度 (*SSIM*) 这 3 个性能指标上也比其他对比算法更具有优势.

关键词: 增强蒲公英优化算法; 多阈值分割; 乳腺癌图像; 对立式学习; 回守策略

引用格式: 王正红, 王丹, 胡容俊. 增强蒲公英算法优化乳腺癌图像多阈值分割. 计算机系统应用, 2024, 33(1): 148-156. <http://www.c-s-a.org.cn/1003-3254/9357.html>

Improved Dandelion Algorithm for Optimizing Multi-threshold Segmentation of Breast Cancer Images

WANG Zheng-Hong, WANG Dan, HU Rong-Jun

(College of Computer Science and Technology, Beihua University, Jilin 132013, China)

Abstract: In the context of complex structures and blurred cell boundaries in microscopic breast cancer histopathological images, traditional threshold-based segmentation faces challenges in accurately separating lesion areas of breast cancer images. To address this issue, this study proposes a multi-threshold segmentation method for breast cancer images based on the improved dandelion optimization algorithm (IDO). This method introduces the IDO to calculate the maximum inter-class variance (Otsu) as the objective function for finding the optimal thresholds. The IDO incorporates a defensive strategy to address the issue of unbounded search in the traditional dandelion optimization algorithm (DO) that extends beyond pixel ranges. Additionally, opposition-based learning (OBL) is introduced to prevent the algorithm from getting trapped in local optima. The experimental results indicate that compared with the Harris Hawks optimization (HHO), gorilla troop optimization (GTO), traditional DO, and marine predators algorithm (MPA), the IDO algorithm achieves the highest fitness value and fastest convergence under the same number of threshold levels. Moreover, it outperforms other comparative algorithms in terms of peak signal-to-noise ratio (*PSNR*), structural similarity index (*SSIM*), and feature similarity index (*FSIM*).

Key words: improved dandelion algorithm (IDO); multi-threshold segmentation; breast cancer image; opposition-based learning (OBL); fall-back strategy

① 收稿时间: 2023-06-18; 修改时间: 2023-07-25; 采用时间: 2023-08-17; csa 在线出版时间: 2023-11-28

CNKI 网络首发时间: 2023-11-30

乳腺癌已取代宫颈癌成为威胁妇女身体健康的最常见恶性肿瘤。如果能在乳腺早期检查中识别出异常,将会提高医生治疗成功率并降低患者死亡风险^[1]。浸润性乳腺导管癌(IDC)是乳腺癌中最常见的类型,但因为乳腺结构的复杂和 IDC 的细微性质,导致人工诊断效率低下,甚至出现误诊的情况^[2]。随着医学技术的发展与人工智能的兴起,越来越多的医院利用计算机辅助技术对乳腺癌医学图像进行分析检测,帮助医生提高诊断率。图像分割是计算机辅助技术的关键环节,高质量的医学分割图像可以提高病变区域检测的准确性^[3,4]。

因为阈值分割相对于其他的分割技术更加简单与便捷,而且阈值分割方法中最大化类间方差法(Otsu)因为图像分割效果好,操作简单,所以我们选择阈值分割和 Otsu 用于图像处理^[5,6]。但传统的阈值分割需要通过穷举法来列举排除出一个合适阈值,效率低下,过程复杂,为减小计算量,文献^[7]和文献^[8]提出利用鲸鱼算法来寻找 Otsu 的合适阈值,Wang 等人提出了一种基于改进金枪鱼群优化(ITSO)的混合森林冠层图像分割方法,将对称交叉熵方法引入到森林冠层图像分割中以获取各种植物种群参数^[9],在文献^[10]中,采用狮子优化和猫群优化的混合算法,为有效的多级图像阈值分割提供了最优阈值,Agrawal 等人采用了主色分量和自适应 WOA 对彩色图像进行多级阈值处理^[11],Sharma 等人使用 Kapur、Tsali 和模糊熵的目标函数,提出了一种有效的基于对立的改进萤火虫算法,用于多级图像分割^[12],文献^[13]提出一种基于多目标海洋捕食者算法的多级阈值分割算法,该算法以 9DKapur 熵函数作为目标函数,得到图像分割效果良好,但这些方法要么计算复杂,实用性较为欠缺,要么本身算法的收敛能力不足,无法获得最佳阈值。

所以我们引入增强的蒲公英优化算法(IDO)来探索出最合适的阈值,IDO 在探索阈值的阶段结合布朗运动和莱文(Levy)飞行,以其强大的局部优化能力和对立式学习本身的全局优化能力完美的胜任了探寻最佳阈值的任务。本算法寻到的最佳阈值后得到的图像分割效果通过实验证明满足了我们的需求。

1 相关工作

1.1 最大化类间方差法(Otsu)

最大化类间方差法由 Otsu 提出^[14],在一张图片的二维直方图上通过最大化阈值两侧像素数据的类间方

差,使我们分割的各部分区域相差最大^[15-17]。假设一张图片大小为 $M \times N$,其灰度级为 $[0, L]$,计算灰度 i 的频率 P_i :

$$P_i = \frac{n_i}{M \times N} \quad (1)$$

其中, n_i 是图片中灰度值为 i 的像素数量, $i=0, 1, 2, \dots, L$ 。如果阈值个数为 m , $0 < m < L-1$,则分割部分数量为 $m+1$,记为 $[0, T_1, T_2, \dots, T_m, 255]$, $T_{(m+1)}=255$ 。统计每个分割部分灰度级的平均灰度:

$$u_k = \frac{1}{\omega_k} \sum_{j=T_k}^{T_{k+1}} jP_j \quad (2)$$

其中, $k=0, 1, 2, \dots, m$, j 表示的是两个阈值之间的灰度值,参数 ω_k 计算如式(3)所示:

$$\omega_k = \sum_{j=T_k}^{T_{k+1}} P_j \quad (3)$$

图片整体的平均灰度值 μ_T 为:

$$\mu_T = \sum_{i=0}^L iP_i \quad (4)$$

其总体类间方差计算公式为式(5):

$$\sigma_B^2 = \sum_{k=0}^m \omega_k [u_k - \mu_T]^2 \quad (5)$$

1.2 对立式学习(opposition-based learning, OBL)

许多受自然群体影响的启发式算法将函数逼近视为给定问题的答案^[18,19]。在许多情况下,算法起点的初始化位置通常是随机选择的,包括神经网络的权重、群智能优化算法的初始种群和增强剂的主动策略^[20]。当起点接近理想的最佳解时,收敛发生得更快。但是,如果它离理想解的距离很远,比如在最差的情况下,处于和最佳解完全相反的位置,收敛将需要的时间更长,或者获得解可能是无法收敛的不可解^[21]。对立解的数学表现形式如式(6)所示:

$$\bar{x}_0 = a + b - x_0 \quad (6)$$

其中, x_0 是属于 $[a, b]$ 范围内的一个实数, \bar{x}_0 就是它的对立解。将它放入直角坐标系中,对立解就是 $[a, b]$ 范围内的一个实数关于中心直线 $x = \frac{a+b}{2}$ 对称。

在 D 维度决策空间中,某一个体 $x = x_1, x_2, x_3, \dots, x_D$, 它的对立解表现形式 $\bar{x} = [\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_D]$, 群体的对立解为 $\bar{x}_k = a_k + b_k - x_k$, 其中 $k = 1, 2, 3, \dots, D$ ^[22]。

2 蒲公英优化算法(DO)

2.1 算法基本原理

蒲公英优化算法(DO)是一种用于解决连续优化

问题的新型元启发式算法^[23]。元启发式算法简单来说就是模仿自然群体在决策空间内通过个体的探寻和群体的配合不停地迭代搜索可行解,并且优胜劣汰^[24]。在元启发式算法中,探索和开发是值得重点关注的因素。过度关注探索新解会导致算法本身最后收敛困难,但过度关注开发却又会导致算法陷入到局部最优^[25]。所以如何在探索和开发之间取得一个平衡是决定元启发式算法性能优劣的关键指标。

蒲公英优化算法(DO)是受到蒲公英种子依靠风力长途飞行传播的灵感而诞生的算法,蒲公英种子在传播过程中经过3个阶段:第1个阶段是上升阶段,如果是在晴天多风的天气,种子会借助风力向更远的地方传播,如果是在阴天多雨的天气,种子就会漂浮在附近;第2个阶段是下降阶段,种子上升到一定高度时开始逐渐向下降落;第3个阶段是着陆阶段,种子会在天气与风力的影响下落到一个随机位置孕育出下一代蒲公英。而蒲公英算法正是通过这3个阶段更新个体的位置,并且在上升阶段通过服从正态分布的随机数来决定探索和开发,同时通过动态调节种子传播半径来控制蒲公英种子位置的变更。

DO算法框架大致为种群初始化,计算种群适应度,更新种群,挑选全局最优4部分,其包含2个主要参数,分别为种子传播半径 α ,局部搜索系数 K 。种子传播半径 α 和局部搜索系数 K 在迭代过程中随着时间变化而变化, α 用于调节全局搜索步长,局部搜索系数 K 用于调节局部搜索步长,通过服从正态分布的随机数用于调整由探索到开发过程以及避免局部最优。

2.2 初始化阶段

定义一个 N 粒 d 维度决策空间的蒲公英种群矩阵 $seed$,其第 i 粒种子可以表示为 $X_i=[X_i^1, X_i^2, \dots, X_i^d]$, $i=1, 2, \dots, N$ 。对该种群进行初始化,初始化为式(7):

$$X_i = lb + r_i \times (ub - lb) \quad (7)$$

其中, r_i 是服从正态分布的(0, 1)间的随机数字, ub 为决策空间最大值, lb 是决策空间的最小值, N 是种群矩阵的最大行数, d 是种群矩阵的最大列数,后文同字符皆为同义。

2.3 上升阶段

通过一个服从正态分布的随机数 r 决定全局搜索或者局部开发,记为 $r \sim N(0, 1)$ 。

1) 如果 $r < 1.5$,天气晴朗时上升,为全局搜索,种子

此时的位置为式(8):

$$X_{t+1} = X_t + \alpha \times v_x \times v_y \times \ln Y \times (X_s - X_t) \quad (8)$$

其中, X_s 表示初始化的种子位置,如式(9)所示; v_x 和 v_y 表示水平与垂直的风力系数,如式(11)所示; α 表示的是种子传播半径,如式(10)所示; $\ln Y$ 表示服从均值为0,方差为1的对数正态分布。

$$X_s = rand(1, d) \times (ub - lb) + lb \quad (9)$$

$$\alpha = rand() \times \left(\frac{1}{T^2} t^2 - \frac{2}{T} t + 1 \right) \quad (10)$$

其中, T 是总迭代次数, t 是当前迭代次数, $t=1, 2, 3, \dots, T$ 。

$$v_x = r \times \cos \theta, v_y = r \times \sin \theta, r = \frac{1}{e^\theta} \quad (11)$$

其中, θ 是介于 $[-\pi, \pi]$ 之间的随机数。

2) 如果 $r \geq 1.5$,天气下雨时上升,为局部搜索,种子此时的位置如式(12):

$$X_{t+1} = X_t \times k \quad (12)$$

其中,局部搜索系数 k 如式(13)所示:

$$k = 1 - rand() \times q \quad (13)$$

$$q = \frac{1}{T^2 - 2T + 1} t^2 - \frac{2}{T^2 - 2T + 1} t + 1 + \frac{1}{T^2 - 2T + 1} \quad (14)$$

2.4 下降阶段

经过上升一定距离后,此时算法仍然侧重于探索,种子开始根据布朗运动逐渐下降,种子位置如式(15)所示:

$$X_{t+1} = X_t - \alpha \times \beta_t \times (X_{\text{meant}} - \alpha \times \beta_t \times X_t) \quad (15)$$

其中, β_t 是布朗运动,服从正态分布的随机数; X_{meant} 是为了反映蒲公英下降的稳定性,采用了上升阶段后的平均位置信息,表示为式(16):

$$X_{\text{meant}} = \frac{1}{N} \sum_{i=1}^N X_i \quad (16)$$

2.5 着陆阶段

蒲公英此时开始着陆,根据莱文飞行在陆地随机选择着陆地点,算法开始收敛进入开发阶段。种子位置表示为式(17):

$$X_{t+1} = X_{\text{elite}} + levy(\lambda) \times \alpha \times (X_{\text{elite}} - X_t \times \delta) \quad (17)$$

其中, X_{elite} 表示种群的最佳个体; $levy(\lambda)$ 代表莱文飞行,可由式(18)表示, λ 是服从 $levy$ 分布函数的随机数;线性函数 δ 如式(20)所示。

$$levy(\lambda) = s \times \frac{w \times \sigma}{|t|^{\frac{1}{\beta}}} \quad (18)$$

$$\sigma = \begin{cases} \frac{\Gamma(1+\beta) \times \sin(\frac{\pi\beta}{2})}{\Gamma(\frac{1+\beta}{2}) \times \beta \times 2^{(\frac{\beta-1}{2})}} & \beta < 1.5 \\ \frac{\Gamma(1+\beta) \times \sin(\frac{\pi\beta}{2})}{\Gamma(\frac{1+\beta}{2}) \times \beta \times 2^{(\frac{\beta-1}{2})}} & \beta \geq 1.5 \end{cases} \quad (19)$$

$$\delta = \frac{2t}{T} \quad (20)$$

其中, $s=0.01$, $\beta=1.5$, w 和 t 都是 0-1 之间的随机数字, Γ 是伽马函数.

越界判定流程如图 1 所示, 增强的蒲公英优化算法流程图如图 2 所示.

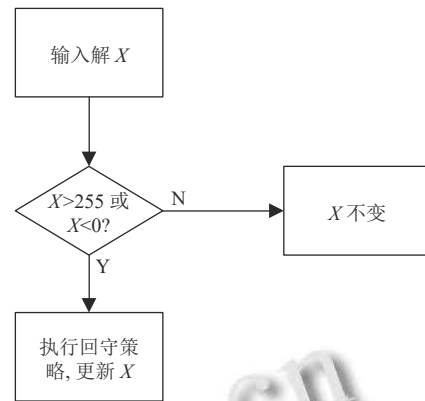


图 1 越界判定流程图

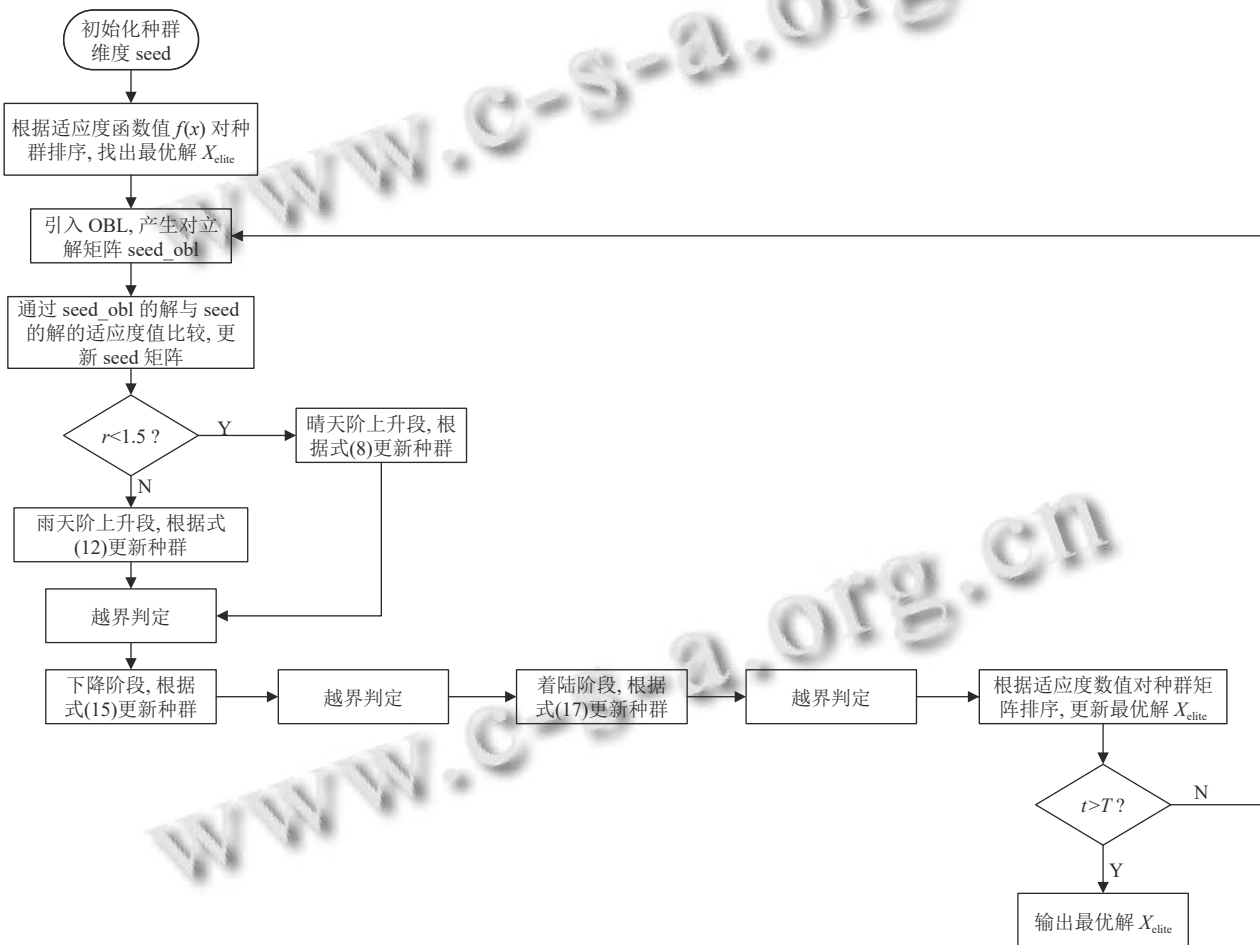


图 2 增强的蒲公英优化算法流程图

3 增强蒲公英优化算法 (IDO)

3.1 IDO 的 OBL 优化

本节对所提出的技术 IDO 进行了全面的解释. 最初只是扩展了 DO 算法的应用, 将它引入乳腺肿瘤图像分割邻域, 以 Otsu 法为适应度函数. IDO 引入了 OBL 规则, 在解 R 迭代更新之前都会与反向解 R 相

比较, 如果 R 的适应度值) 大于 R , R 保持不变, 否则 R 的值更新为 R . 相比原算法在决策空间的孤军奋战, OBL 通过对立解的产生, 相当于在决策空间的两个方向同时探索, 增加了解的多样性, 避免算法陷入局部最优, 提高了算法的收敛性. 它的大致流程如图 3 所示.

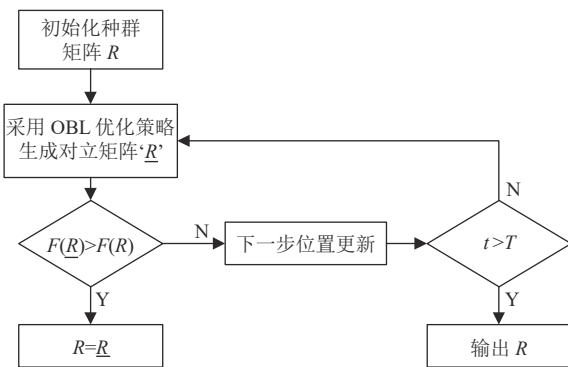


图3 OBL 优化流程图

3.2 回守策略

蒲公英传统算法探索解时超出范围, 会强制等于上下界限 $[0, 255]$, 但在图像分割上会导致图片信息丢失, 模糊不清, 如图4所示. IDO 建立的回守策略却在解 X 超出范围时, 在范围内通过服从正态分布的随机数随机生成一个新解, 然后将新解与 X 的适应度值比较, 优胜劣汰. 通过回守策略, 就可避免阈值取到 0 或 255 的极端情况, 保证图片信息完好.



图4 极端取值 Lenna 测试图

3.3 IDO 算法操作步骤

IDO 算法操作步骤如算法 1.

算法 1. IDO 算法

- 1) 定义一个 N 粒 d 维度决策空间的蒲公英种群矩阵 $seed$, 其第 i 粒种子可以表示为 $X_i = [X_i^1, X_i^2, X_i^3, \dots, X_i^d]$, $i = 1, 2, 3, \dots, N$. 对该种群进行初始化, 初始化为式 (7).
- 2) 引入 OBL 规则, 通过式 (6) 产生 $seed$ 的对立解矩阵 $seed_obs$.
- 3) 通过 OBL 优化规则将对立解矩阵 $seed_obs$ 和种群矩阵 $seed$ 比较, 更新种群矩阵 $seed$.
- 4) 对种群 $seed$ 进行评估, 挑选种群的最佳个体.
- 5) 通过一个服从正态分布的随机数 r 决定全局搜索或者局部开发: 如果 $r < 1.5$, 天气晴朗时上升, 为全局搜索, 种子此时的位置为式 (8); 如果 $r \geq 1.5$, 天气下雨时上升, 为局部搜索, 种子此时的位置为式 (12).
- 6) 经过上升一定距离后, 此时算法仍然侧重于探索, 种子开始根据布朗运动逐渐下降, 种子位置如式 (15) 所示.
- 7) 蒲公英此时开始着陆, 根据莱文飞行在陆地随机选择着陆地点, 算法开始收敛进入开发阶段. 种子位置表示为式 (17).
- 8) 更新最佳个体 X_{elite} .
- 9) 如果迭代没有结束, 返回第 2) 步, 否则输出 X_{elite} .

4 实验分析

为了验证 IDO 算法对乳腺癌图像分割的准确性, 需要将它与经典算法哈里斯鹰算法 (HHO)^[26], 人工猩猩部队优化算法 (GTO)^[27], 传统蒲公英优化算法 (DO), 海洋捕食者算法 (MPA)^[28] 在相同阈值下用于图像分割领域, 通过峰值信噪比 (PSNR), 结构相似度 (FSIM), 特征相似度 (SSIM) 对比, 分析算法的性能优劣. 本实验是在 PyCharm 的 2019 专业版本, 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30 GHz, 2.30 GHz, Windows 10 64 位系统平台上运行, 本文乳腺细胞图像来自文献[29]中的数据集, 数据集中的标本是用苏木精和伊红染色的乳腺组织, 取自 2019–2014 年间在伊朗伊斯法罕医科大学 Poursina Hakim 研究中心诊断出的 124 名患者, 共有 922 张 JPEG 格式的图片, 本文选取 4X 倍数的其中一张. 其原始图像和灰度图如图 5 所示, 数据网址链接 <https://datatubox.com/>.

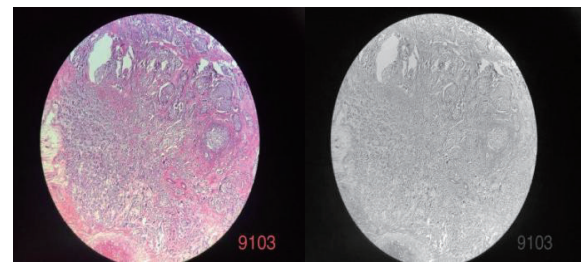


图5 乳腺细胞图

4.1 参数设置

所有的算法我们将种群大小 N 设为 20, 最大迭代次数 T 设为 350, 选取阈值个数为 $k=2, 3, 4$, 我们得到的最终实验数据是在每一个算法单独执行 100 次后取得的平均数. 我们对实验中算法的参数设置如表 1 所示. IDO 和 DO 算法以及 HHO 算法中使用莱文飞行部分的参数设置为 1.5, MPA 算法涡流效应参数 FADs 设为 0.2, 步长 $p1$ 设为 1.5, 参数 C 设为 0.05 设为 1.5.

4.2 数据分析

表 2 展示了针对实验中使用的乳腺癌测试图像, 使用不同阈值 $[k=2, 3, 4]$ 的 IDO 算法、DO 算法、HHO 算法、MPA 算法、GTO 算法分割后的图像以及它们的最佳阈值取值大小. 表 3 表示了使用不同阈值 $[k=2, 3, 4]$ 时 IDO 算法、DO 算法、HHO 算法、MPA 算法、GTO 算法的适应度值曲线和适应度值.

从表 3 的各算法适应度图中可以看出, IDO 算法, 在 $k=2, 3, 4$ 任意阈值时, 它的收敛速度都快于其他算

法,并且它获得的适应度值要比其他算法高.在阈值个数 $k=2$ 时,IDO 算法在 50 次迭代以内便开始完成收敛,GTO 算法在迭代次数达到 250 次时开始收敛,HHO 算法适应度曲线波动起伏大并且在迭代次数达到 150 次后才开始收敛,DO 算法和 MPA 算法在迭代次数达到 250 次后才开始收敛,而且相较于其他算法,IDO 算法和 MPA 算法所获的适应度值最高.在阈值个数 $k=3$ 时,IDO 算法仍然在 50 次迭代以内便开始完成收敛,但收敛时的时间开始增加,GTO 算法在迭代次数达到 150 次时开始收敛,HHO 算法适应度曲线波动起伏大并且也在迭代次数达到 150 次后才开始收敛,MPA 算法在迭代次数达到 300 次后开始收敛,DO 算法在迭代次数达到 250 次后开始收敛,而且相较于其他算法,IDO 算法和 MPA 算法所获的适应度值最高. $k=4$ 时,IDO 算法在 50 次迭代以内便开始完成收敛,GTO 算法在迭代次数达到 150 次时开始收敛,HHO 算法适应度曲线波动起伏大并且在迭代次数达到 200 次后才开始收敛,MPA 算法在迭代次数达到 300 次后开始收敛,DO 算法在迭代次数达到 250 次后开始收敛且适应度值最低,而且相较于其他算法,IDO 算法所获的适应度值最高.这些充分证明了 IDO 算法的收敛能力比其他算法更强,搜索最佳阈值更快,实时性更好.

表1 算法参数表

算法	参数	值
GTO	T	350
	N	20
IDO	T	350
	N	20
	β	1.5
DO	T	350
	N	20
	β	1.5
HHO	T	350
	N	20
	$Beta$	1.5
MPA	T	350
	N	20
	FADs	0.2
	$p1$	0.5
	C	0.05
	α	1.5

4.3 性能指标

我们很难通过肉眼对分割后的图像质量进行主观评价,所以需要通过峰值信噪比 (PSNR),结构相似指数 (SSIM),特征相似指数 (FSIM) 这 3 个指标进行衡

量^[30].

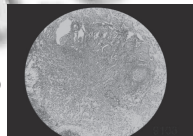
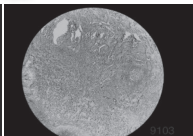
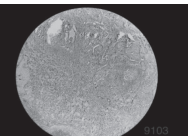
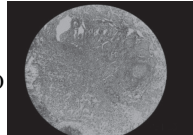
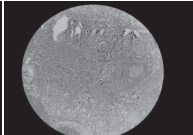
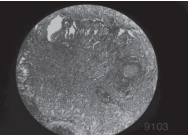
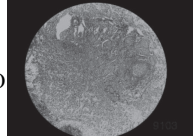
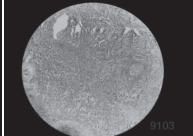
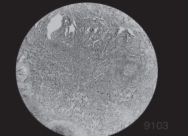
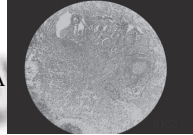
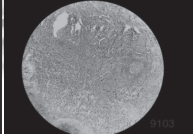


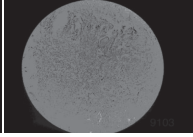
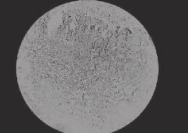
PSNR 计算公式如式 (21) 所示:

$$PSNR = 20 \times \log_{10} \left(\frac{255}{RMSE} \right) \quad (21)$$

$$RMSE = \sqrt{\frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (I_{ij} - Seg_{ij})^2}{M \times N}} \quad (22)$$

其中, $RMSE$ 是每个像素的均方根误差, $M \times N$ 表示图像的大小, I_{ij} 表示原始图像的像素灰度值, Seg_{ij} 表示分割后的图像中的像素灰度值^[31].

表2 分割图像对比图

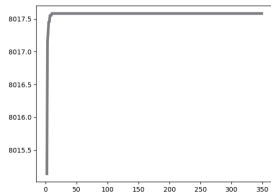
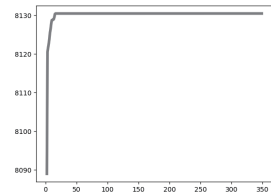
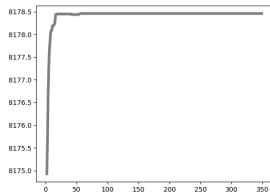
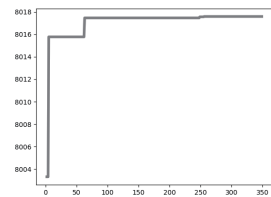
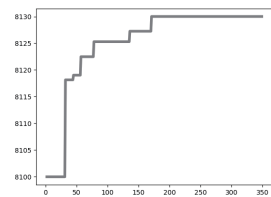
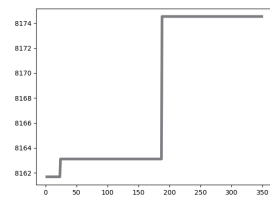
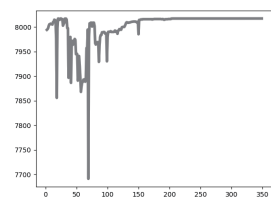
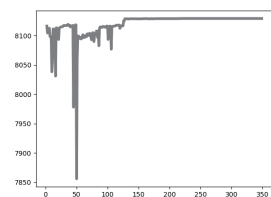
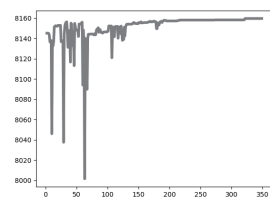
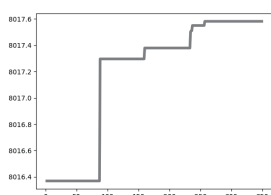
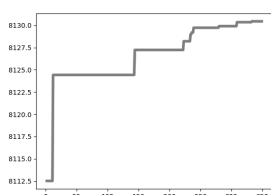
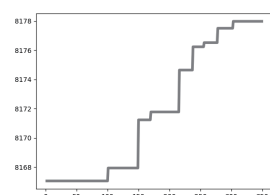

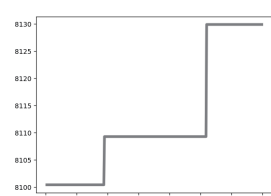
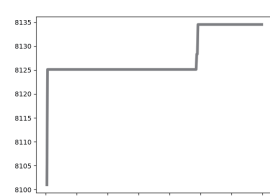
算法	2阈值	3阈值	4阈值
IDO	 86, 186	 74, 159, 202	 67, 143, 181, 213
GTO	 91, 186	 123, 139, 201	 17, 71, 203, 255
HHO	 78, 186	 67, 159, 202	 51, 52, 159, 203
MPA	 86, 186	 75, 16202	 7143, 18211
DO	 127, 199	 105, 127, 137	 79, 81, 91, 109

SSIM 计算公式如式 (23) 所示:

$$SSIM = \frac{(2\mu_I \mu_{seg} + c_1)(2\sigma_{I,seg} + c_2)}{(\mu_I^2 + \mu_{seg}^2 + c_1)(\sigma_I^2 + \sigma_{seg}^2 + c_2)} \quad (23)$$

其中, μ_I 和 μ_{seg} 是原始图像和分割图像的像素平均值, σ_I 和 σ_{seg} 是原始图像和分割图像的标准差, $\sigma_{I,seg}$ 是原始图像和分割图像的协方差, c_1 和 c_2 是用来保证稳定性的常数^[32].

表3 适应度对比图

算法	2阈值	3阈值	4阈值
IDO	 适应度值: 8017.5810	 适应度值: 8130.4463	 适应度值: 8178.4607
GTO	 适应度值: 8017.4548	 适应度值: 8128.5015	 适应度值: 8174.9578
HHO	 适应度值: 8017.4589	 适应度值: 8129.2776	 适应度值: 8177.5435
MPA	 适应度值: 8017.5810	 适应度值: 8130.4463	 适应度值: 8178.3196
DO	 适应度值: 8012.4936	 适应度值: 8066.1557	 适应度值: 8155.0160

FSIM 计算公式如式 (24) 所示:

$$FSIM = \frac{\sum_{I \in \Omega} S_L(X) PC_m(X)}{\sum_{I \in \Omega} PC_m(X)} \quad (24)$$

$$S_L(X) = S_{PC}(X) S_G(X) \quad (25)$$

$$S_{PC}(X) = \frac{2PC_1(X)PC_2(X) + T_1}{PC_1^2(X)PC_2^2(X) + T_1} \quad (26)$$

$$S_G(X) = \frac{2G_1(X)G_2(X) + T_2}{G_1^2(X)G_2^2(X) + T_2} \quad (27)$$

$$G = \sqrt{G_x^2 + G_y^2} \quad (28)$$

$$PC(X) = \frac{E(X)}{\varepsilon + \sum_m A_n(X)} \quad (29)$$

其中, Ω 代表原始图像的所有像素域. $S_L(X)$ 表示相似度得分. T_1 和 T_2 是常数, G_1 和 G_2 分别为原始图像和分割图像的梯度, $E(X)$ 是位置 X 和尺度 n 中的向量大小. 此外, ε 是一个小数, $A_n(X)$ 是尺度 n 中的局部大小, 采用文献[33]的参数设置.

表 4 给出了使用不同阈值 [$k=2, 3, 4$] 的 IDO 算法、DO 算法、HHO 算法、MPA 算法、GTO 算法分割图像后性能指标 $PSNR$ 、 $SSIM$ 、 $FSIM$ 的平均结果.

从表 4 观测数据得知, 在阈值个数 $k=2, 3$ 时 IDO 算法, MPA 算法的 $SSIM$ 值同占第 1 位; 在阈值个数

$k=4$ 时, IDO 算法的 $SSIM$ 值独占第 1 位, 超过了 MPA 算法. 在阈值个数 $k=2, 3$ 时 IDO 算法, MPA 算法的 $FSIM$ 值占第 1 位; 在阈值个数 $k=4$ 时, IDO 算法的 $FSIM$ 值占第 2 位, 稍逊 MPA 算法一筹. 在阈值个数 $k=2$ 时, IDO 算法, MPA 算法的 $PSNR$ 值同占第 3 位, GTO 算法占第 1 位, HHO 算法占第 2 位; 在阈值个数 $k=3$ 时, IDO 算法的 $PSNR$ 值占第 4 位, MPA 算法的 $PSNR$ 值占第 3 位, DO 算法占第 1 位, HHO 算法占第 2 位; 在阈值个数 $k=4$ 时, IDO 算法的 $PSNR$ 值超越了其他算法占第 1 位. 从数据的纵向看, IDO 在 3 个指标中总体比其他算法更具优势, 并且随着阈值个数的增加, IDO 的优势在不断地增加.

表 4 性能指标对比数据

指标	k	IDO	DO	GTO	HHO	MPA
$SSIM$	2	0.9168	0.8969	0.8815	0.9100	0.9168
	3	0.9440	0.7633	0.8313	0.9325	0.9440
	4	0.9547	0.8021	0.9072	0.9281	0.9524
$FSIM$	2	0.5811	0.4106	0.3880	0.5789	0.5811
	3	0.6726	0.3278	0.4476	0.6313	0.6726
	4	0.7058	0.3391	0.4868	0.6241	0.7124
$PSNR$	2	52.5939	52.5509	52.6113	52.5995	52.5939
	3	52.6338	52.6518	52.6248	52.6508	52.6349
	4	52.8101	52.6126	52.6604	52.6768	52.7798

5 结论与展望

通过实验数据可知, 本文提出的基于增强蒲公英优化算法 (IDO) 的乳腺癌图像多阈值分割方法通过传统蒲公英优化算法结合布朗运动和莱文飞行的三阶段搜索过程快速寻找最佳阈值, 提升了图像分割的效率, 并且建立的回守策略解决了传统蒲公英算法 (DO) 在图像分割领域无限制搜索, 超出像素范围的问题; 引入的对立式学习 (OBL) 避免算法陷入局部最优. IDO 的性能超越了对比的 4 种算法: GTO 算法、HHO 算法、DO 算法、MPA 算法, 我们提出对传统蒲公英优化算法的改进是合理的, 对其他启发式算法的改进具有借鉴意义, 并且 IDO 算法在图像分割的优良性能, 也让它具有较好的推广价值.

虽然 IDO 算法在数据上显示有很大的优势, 但其仍然有不足之处, 首先本算法使用优化策略调高了算法的复杂度, 其次本算法还不能处理其他高难度的分割任务, 并且本算法只测试了部分低阈值个数的情况, 没有测试高阈值情况, 没有找到最佳阈值个数, 这些将是我们今后的重点研究方向.

参考文献

1 Kumar A, Singh SK, Saxena S, *et al.* Deep feature learning

for histopathological image classification of canine mammary tumors and human breast cancer. *Information Sciences*, 2020, 508: 405–421. [doi: 10.1016/j.ins.2019.08.072]

2 Yang X, Wang R, Zhao D, *et al.* Multi-level threshold segmentation framework for breast cancer images using enhanced differential evolution. *Biomedical Signal Processing and Control*, 2023, 80: 104373. [doi: 10.1016/j.bspc.2022.104373]

3 李雪梅, 曹琼, 曹慧敏, 等. 基于最大熵阈值分割法的颅脑 CT 图像血肿自动诊断系统研究. *中国医学装备*, 2022, 19(8): 1–5. [doi: 10.3969/J.ISSN.1672-8270.2022.08.001]

4 常君杰, 李东兴, 钟欣, 等. 改进乌鸦算法的二维 Tsallis 熵多阈值图像分割算法. *山东理工大学学报 (自然科学版)*, 2021, 35(5): 21–26. [doi: 10.13367/j.cnki.sdgc.2021.05.004]

5 Lan K, Zhou JQ, Jiang XL, *et al.* Group theoretic particle swarm optimization for multi-level threshold lung cancer image segmentation. *Quantitative Imaging in Medicine and Surgery*, 2023, 13(3): 1312–1322. [doi: 10.21037/qims-22-295]

6 卢建宏, 刘海鹏, 王蒙. 改进海鸥算法的多阈值图像分割算法. *光电子·激光*, 2022, 33(9): 932–939. [doi: 10.16136/j.joel.2022.09.0049]

7 El Aziz MA, Ewees AA, Hassanien AE, *et al.* Multi-

- objective whale optimization algorithm for multilevel thresholding segmentation. *Advances in Soft Computing and Machine Learning in Image Processing*. Cham: Springer.
- 8 郭松林, 巴艳坤. 改进的鲸鱼优化算法与图像阈值分割. *黑龙江科技大学学报*, 2022, 32(5): 672–678.
- 9 Wang JY, Zhu LK, Wu BW, *et al.* Forestry canopy image segmentation based on improved tuna swarm optimization. *Forests*, 2022, 13(11): 1746. [doi: [10.3390/f13111746](https://doi.org/10.3390/f13111746)]
- 10 Vijh S, Saraswat M, Kumar S. Automatic multilevel image thresholding segmentation using hybrid bio-inspired algorithm and artificial neural network for histopathology images. *Multimedia Tools and Applications*, 2023, 82(4): 4979–5010. [doi: [10.1007/s11042-022-12168-9](https://doi.org/10.1007/s11042-022-12168-9)]
- 11 Agrawal S, Panda R, Choudhury P, *et al.* Dominant color component and adaptive whale optimization algorithm for multilevel thresholding of color images. *Knowledge-based Systems*, 2022, 240: 108172. [doi: [10.1016/j.knosys.2022.108172](https://doi.org/10.1016/j.knosys.2022.108172)]
- 12 Sharma A, Chaturvedi R, Bhargava A. A novel opposition based improved firefly algorithm for multilevel image segmentation. *Multimedia Tools and Applications*, 2022, 81(11): 15521–15544. [doi: [10.1007/s11042-022-12303-6](https://doi.org/10.1007/s11042-022-12303-6)]
- 13 Xing ZK, He YG. Many-objective multilevel thresholding image segmentation for infrared images of power equipment with boost marine predators algorithm. *Applied Soft Computing*, 2021, 113: 107905. [doi: [10.1016/j.asoc.2021.107905](https://doi.org/10.1016/j.asoc.2021.107905)]
- 14 Otsu N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, 9(1): 62–66.
- 15 余扬. 基于二维 Otsu 的舰船目标图像分割方法. *舰船电子工程*, 2022, 42(1): 36–39.
- 16 张永刚. 基于小波变换的医学影像图像阈值分割实现设计. *贵州大学学报(自然科学版)*, 2021, 38(2): 37–39, 43. [doi: [10.15958/j.cnki.gdxbzrb.2021.02.07](https://doi.org/10.15958/j.cnki.gdxbzrb.2021.02.07)]
- 17 李晓峰, 焦洪双, 李东. 基于量子蚁群算法的医疗图像阈值分割算法. *沈阳大学学报(自然科学版)*, 2020, 32(6): 490–495, 522. [doi: [10.16103/j.cnki.21-1583/n.2020.06.008](https://doi.org/10.16103/j.cnki.21-1583/n.2020.06.008)]
- 18 Upadhyay P, Chhabra JK. Kapur's entropy based optimal multilevel image segmentation using crow search algorithm. *Applied Soft Computing*, 2020, 97: 105522. [doi: [10.1016/j.asoc.2019.105522](https://doi.org/10.1016/j.asoc.2019.105522)]
- 19 Yu XB, Xu WY, Li CL. Opposition-based learning grey wolf optimizer for global optimization. *Knowledge-based Systems*, 2021, 226: 107139. [doi: [10.1016/j.knosys.2021.107139](https://doi.org/10.1016/j.knosys.2021.107139)]
- 20 Houssein EH, Abdelkareem DA, Emam MM, *et al.* An efficient image segmentation method for skin cancer imaging using improved golden jackal optimization algorithm. *Computers in Biology and Medicine*, 2022, 149: 106075. [doi: [10.1016/j.compbiomed.2022.106075](https://doi.org/10.1016/j.compbiomed.2022.106075)]
- 21 Xu W, Zhang RF, Chen L. An improved crow search algorithm based on oppositional forgetting learning. *Applied Intelligence*, 2022, 52(7): 7905–7921. [doi: [10.1007/s10489-021-02701-y](https://doi.org/10.1007/s10489-021-02701-y)]
- 22 Oliva D, Esquivel-Torres S, Hinojosa S, *et al.* Opposition-based moth swarm algorithm. *Expert Systems with Applications*, 2021, 184: 115481. [doi: [10.1016/j.eswa.2021.115481](https://doi.org/10.1016/j.eswa.2021.115481)]
- 23 Zhao SJ, Zhang TR, Ma SL, *et al.* Dandelion optimizer: A nature-inspired metaheuristic algorithm for engineering applications. *Engineering Applications of Artificial Intelligence*, 2022, 114: 105075. [doi: [10.1016/j.engappai.2022.105075](https://doi.org/10.1016/j.engappai.2022.105075)]
- 24 Elaziz MA, Ewees AA, Oliva D. Hyper-heuristic method for multilevel thresholding image segmentation. *Expert Systems with Applications*, 2020, 146: 113201. [doi: [10.1016/j.eswa.2020.113201](https://doi.org/10.1016/j.eswa.2020.113201)]
- 25 常君杰. 启发式算法在图像分割中的应用研究[硕士学位论文]. 淄博: 山东理工大学, 2021. [doi: [10.27276/d.cnki.gsdgc.2021.000127](https://doi.org/10.27276/d.cnki.gsdgc.2021.000127)]
- 26 Heidari AA, Mirjalili S, Faris H, *et al.* Harris hawks optimization: Algorithm and applications. *Future Generation Computer Systems*, 2019, 97: 849–872. [doi: [10.1016/j.future.2019.02.028](https://doi.org/10.1016/j.future.2019.02.028)]
- 27 Abdollahzadeh B, Gharehchopogh FS, Mirjalili S. Artificial gorilla troops optimizer: A new nature-inspired metaheuristic algorithm for global optimization problems. *International Journal of Intelligent Systems*, 2021, 36(10): 5887–5958. [doi: [10.1002/int.22535](https://doi.org/10.1002/int.22535)]
- 28 Faramarzi A, Heidarinejad M, Mirjalili S, Gandomi AH. Marine predators algorithm: A nature-inspired metaheuristic. *Expert Systems with Applications*, 2020, 152: 113377. [doi: [10.1016/j.eswa.2020.113377](https://doi.org/10.1016/j.eswa.2020.113377)]
- 29 Bolhasani H, Amjadi E, Tabatabaeian M, *et al.* A histopathological image dataset for grading breast invasive ductal carcinomas. *Informatics in Medicine Unlocked*, 2020, 19: 100341. [doi: [10.1016/j.imu.2020.100341](https://doi.org/10.1016/j.imu.2020.100341)]
- 30 吕鑫, 慕晓冬, 张钧. 基于改进麻雀搜索算法的多阈值图像分割. *系统工程与电子技术*, 2021, 43(2): 318–327.
- 31 刘俊梅, 马永刚. 基于新型混合智能算法的图像分割方法. *河南科学*, 2022, 40(5): 709–713.
- 32 王仕儒. 基于分数阶布谷鸟优化的 Otsu 图像分割算法研究[硕士学位论文]. 银川: 宁夏大学, 2022. [doi: [10.27257/d.cnki.gnxhc.2022.000623](https://doi.org/10.27257/d.cnki.gnxhc.2022.000623)]
- 33 Tang KZ, Xiao X, Wu J, *et al.* An improved multilevel thresholding approach based modified bacterial foraging optimization. *Applied Intelligence*, 2017, 46(1): 214–226. [doi: [10.1007/s10489-016-0832-9](https://doi.org/10.1007/s10489-016-0832-9)]

(校对责编: 牛欣悦)