Swarm and Evolutionary Computation 98 (2025) 102065

Contents lists available at ScienceDirect

Swarm and Evolutionary Computation

journal homepage: www.elsevier.com/locate/swevo

Weak-prior medical image matting based on microscale-searching evolutionary optimization

Li Kang [a,1], Han Huang [a], Yihui Liang [b,*], Baoxiong Zhuang [c,*]

[a] School of Software Engineering, South China University of Technology, Guangzhou, 510006, Guangdong, China
[b] Zhongshan Institute, University of Electronic Science and Technology of China, Zhongshan 528400, Guangdong, China
[c] Department of General Surgery, Nanfang Hospital, Southern Medical University, 1838 North of Guangzhou Avenue, Guangzhou, 510517, Guangdong, China

ARTICLE INFO

Keywords:
Medical image matting
Microscale-searching
Large-scale combinatorial optimization

ABSTRACT

Image matting is to predict alpha mattes that reflects the opacity of images, recently showing potential in identifying transition regions of lesions in computer aided diagnosis. Image matting can be modeled as a large-scale combinatorial optimization problem that has numerous subproblems. Evolutionary algorithms (EAs) have been applied to predict accurate alpha mattes. The advantage of EAs-based methods is the ability to predict alpha mattes with weak prior like trimaps that provide value of opacity for pixels effortless to annotate compared to recent deep learning-based methods. However, it is challenging for EAs to solve the problem efficiently due to numerous subproblems and the large size of the decision set. Based on the observation that the similarity of subproblems correlates with the similarity of their objective spaces, this paper proposes a method for estimating a microscale subset of the decision set from the solving process of similar subproblems. A framework is designed to reduce the exploration cost of EAs in the large-scale decision set by guiding EAs to search in this estimated microscale subsets. Three medical image matting datasets are used to validate our method's improvement in the efficiency of evolutionary algorithms. Experimental results demonstrate that EAs embedded in the proposed framework obtain the best prediction of alpha mattes on medical images and also in weak scenarios involving natural images. Comparative experimental results on multi-objective performance metrics indicate that our method is capable of finding superior solutions using fewer fitness evaluations. The contribution of our work is to make EAs an efficient approach to solving the medical image matting problem with weak prior.

1. Introduction

Image matting is a fundamental task for computer vision, playing a critical role in film editing [1], preprocessing for remote sensing [2] and video postprocessing [3]. Recently, it has shown potential in computer aided diagnosis by automatically generating masks that preserve ambiguous transition region from target lesions to healthy regions [4]. This advancement provides clinicians with enhanced visual delineation of pathological structures, thereby offering critical decision support in oncological diagnosis. Image matting involves generating an alpha matte α by which an image I can be separated into a foreground image F and a background image B. The color of the ith pixel in I can be calculated by:

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i \tag{1}$$

where α_i, ranging from 0 to 1, represents the foreground opacity of a pixel. Pixels with α values of 0, 1, and those ranging from 0 to 1 correspond to background, foreground, and opaque pixels, respectively. F_i and B_i are the selected foreground and background pixels, respectively. For a three-channel image I, the only known variable in Eq. (1) is I_i, while there are a total of seven unknown variables on the right side of Eq. (1). It means that image matting problem is ill-posed. Prior is needed to solve such ill-posed problem. Trimaps as prior are provided by users to roughly segment all pixels into three

* Corresponding authors.
E-mail addresses: k_conley@foxmail.com (L. Kang), hhan@scut.edu.cn (H. Huang), yihuiliang@outlook.com (Y. Liang), zhuangbaoxiongnfyy@163.com (B. Zhuang).

[1] This work is supported by National Natural Science Foundation of China (62276103), Innovation Team Project of General Colleges and Universities in Guangdong Province (2023KCXTD002),Analysis of Urban Events Based on Low-Altitude Drones (2024BQ010011), the Research and Development Project on Key Technologies for Intelligent Sensing, the 2023 Special Program for Audit Theory Research, Guangdong Provincial Philosophy and Social Science Planning (GD23SJZ09)

https://doi.org/10.1016/j.swevo.2025.102065
Received 15 February 2025; Received in revised form 12 May 2025; Accepted 1 July 2025
Available online 22 July 2025

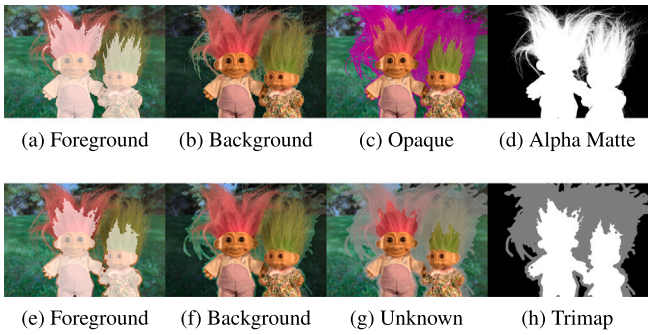


Fig. 1. Comparison of annotations between a trimap (a–d) and an alpha matte (e–h). The locations of white and black pixels are annotated with values of 1 and 0 in both (d) and (h), respectively. The locations of purple pixels are annotated with a real number between 0 and 1 in (d). The locations of gray pixels are annotated with a value of 0.5 in (h).

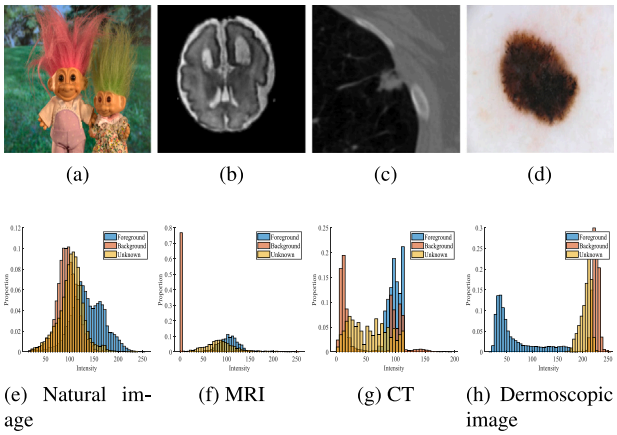


Fig. 2. Comparison of pixel distributions of natural images and medical images. (e)–(h) are the distributions of pixel intensities in (a)–(d), respectively.

types: foreground pixels (shown in white), background pixels (shown in black), and unknown pixels (shown in gray) as shown in the second row of Fig. 1. Both single-objective EAs and multi-objective EAs have been applied to search for the optimal pixel pairs [5–7].

Prior is crucial for solving the image matting problem. Prior in image matting can be classified into weak prior and strong prior according to the number of pixels that need to be annotated and the range of annotation values. The production of these two types of prior are trimaps and detailed alpha mattes, respectively. The differences between them are illustrated in Fig. 1. For detailed annotated alpha mattes, all pixels must be assigned values between 0 and 1, a process that is both time-consuming and labor-intensive. Several image matting datasets have emerged recently, containing a large number of detailed annotated alpha mattes [8–10]. Many efficient deep learning-based matting methods have been developed by training on these datasets. However, the deep models in these methods exhibit high parameter complexity due to the large number of unknown pixels requiring alpha value prediction. This complexity can limit the generalization performance of these methods, making them sensitive to the change of data distributions if the number and variety of detailed annotated alpha mattes are also limited. Especially in medical imaging, it is particularly challenging to produce datasets with detailed annotated alpha mattes on the scale of popular image matting datasets. The production of detailed annotated alpha mattes for medical images requires a high level of expert knowledge. In contrast, traditional image matting methods are not affected by data distribution, as they typically rely on weak priors like trimaps. For the trimaps, all or a portion of the fully transparent pixels and fully opaque pixels need to be labeled as 0 and 1, respectively. The remaining pixels, considered as unknown and to be predicted, are uniformly assigned a value, typically 0.5. In the context of medical imaging, fully transparent pixels represent regions unrelated to the target lesion, while fully opaque pixels correspond to the definitive target lesion.

Among traditional image matting methods, EAs-based methods achieve high accuracy. The goal of EAs-based image matting problem is to solve the following problem [5]:

$$\min\ G(X) = \left(\sum_{k=1}^{c} h_{1k}(x_k), \sum_{k=1}^{c} h_{2k}(x_k), \dots, \sum_{k=1}^{c} h_{Nk}(x_k)\right) \tag{2}$$

$$s.t.\ X = (x_1, x_2, \dots, x_c), c = |\Omega_U| \tag{3}$$

$$x_k = (x_{kF}, x_{kB}), k = 1, 2, \dots, |\Omega_U| \tag{4}$$

$$x_{kF} \in \{1, 2, \dots, |\Omega_F|\} \tag{5}$$

$$x_{kB} \in \{1, 2, \dots, |\Omega_B|\} \tag{6}$$

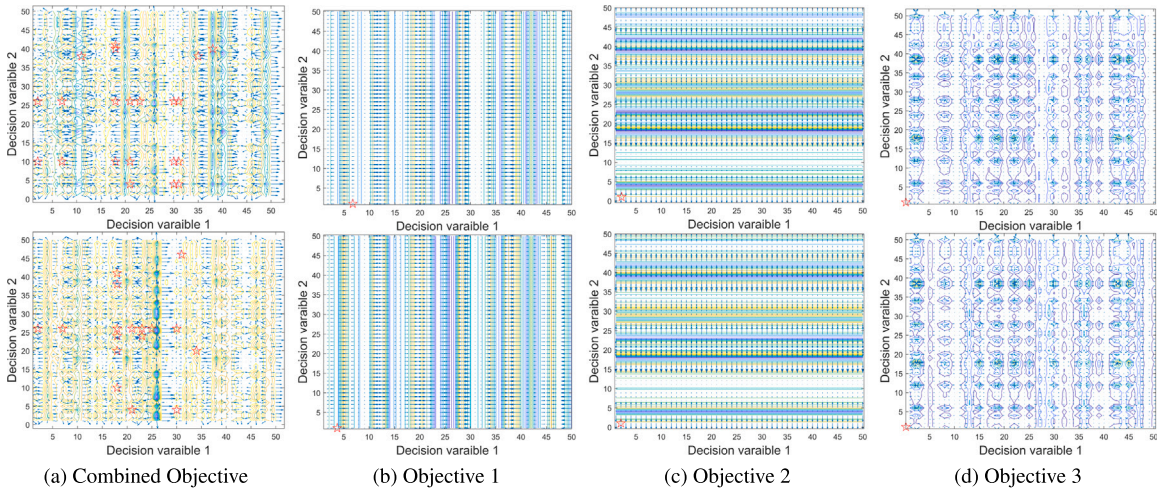
where $h_{1k}, h_{2k}, \dots, k_{Nk}, k = 1, 2, \dots, c$ are N different measurement functions of the kth pixel . Ω_U, Ω_F, and Ω_B denote the sets of unknown pixels, foreground pixels, and background pixels, respectively, as defined in the trimap. $(h_{1k}(x_k), h_{2k}(x_k), \dots, h_{Nk}(x_k))$ is a subproblem of G. The decision set of each subproblem is the combination of foreground pixels and background pixels. The number of subproblems is equal to the number of unknown pixels. The sizes of Ω_U, Ω_F, and Ω_B are summarized in Table 1. As shown in Table 1, the size of the decision set and the number of subproblems are large, which constrains the efficiency of EAs. This constraint is a primary reason why the application of EAs-based methods are restricted compared to other types of methods. Consequently, it is essential to design efficient methods that enhance the effectiveness of EAs-based methods to expand their application scenarios, such as in medical imaging.

Single-objective and multi-objective EAs have been employed to tackle the image matting problem [5,7,16]. In methods based on single-objective EAs, the objective function aggregates various metrics that assess the quality of pixel pairs. This aggregation overlooks conflicts among multiple metrics, which can increase the discrepancy between the optimal pixel pair and the optimal solution of the objective function. Multi-objective optimization models have been proposed to mitigate the impact of these conflicts [6]. MOEA/D was applied to solve the optimization problem [5,17]. To enhance the efficiency of MOEA/D, the number of subproblems is reduced by selecting representatives from groups of subproblems that are clustered based on the spatial correlation of pixels. Spatial correlation is also used in single-objective EAs-based methods to reduce the number of subproblems [7]. However, the reduction of subproblems is limited by the number of groups, as it is essential to maintain similarity among subproblems within each group. The large size of the decision set still constraints the efficiency of EAs. Especially for medical images, as shown in Fig. 2, although the color distribution of pixels is less complex compared to natural images, the color changes in natural images are smoother and more diverse than in medical images. This smoothness and diversity facilitate the application of continuous EAs to solve the image matting problem after the problem has been continuousized, which can improve the efficiency of solving the image matting problem [18]. In contrast, for medical images, the decision set may contain a large number of similar solutions that have large difference from the optimal pixel pairs. This situation poses challenges for EAs in efficiently searching for the optimal solution.

The efficiency of EAs can be enhanced based on the observation that the spaces of objective functions of similar unknown pixels exhibit analogous properties. Two rows in each subfigure of Fig. 3 present contour plots of the objective functions for two similar unknown pixels, x_i and $x_j, i \neq j$. The arrows in Fig. 3 visualize the gradients of the objective functions with respect to the two decision variables. Four objective functions are depicted in the figure. One is the aggregation of three objective functions, and the other three correspond to each

Table 1
The summary of the mean, maximum, and minimum numbers of foreground pixels, background pixels, and unknown pixels in testing data across different datasets. $|\Omega_F|$, $|\Omega_B|$, $|\Omega_U|$ denote the sets of foreground, background, and unknown pixels, respectively.

	Dataset names	Total number of images	Ω_F	Ω_B	Ω_U		
			avg	avg	avg	max	min
Natrual images	AlphaMatting [11]	27	4.74e+05	4.74e+05	4.74e+05	5.75e+05	3.92e+05
	DIM-481 [8]	1000	2.28e+06	2.28e+06	2.28e+06	3.69e+06	5.62e+05
	Distinctions-646 [12]	1000	2.23e+06	2.23e+06	2.23e+06	7.96e+06	2.40e+05
	SIMD [9]	116	2.36e+06	2.36e+06	2.36e+06	3.69e+06	9.21e+05
Medical images	Brain-growth [13]	39	6.97e+02	1.22e+04	2.66e+03	3.70e+03	2.24e+03
	LIDC-IDRI [14]	1609	1.36e+02	1.57e+04	5.06e+02	2.24e+03	1.95e+02
	ISIC [15]	120	5.41e+04	1.18e+05	2.29e+04	1.28e+05	1.85e+02



(a) Combined Objective (b) Objective 1 (c) Objective 2 (d) Objective 3

Fig. 3. Contour plots of the objective spaces for two similar unknown pixels.

individual objective function. The red stars represent non-dominated solutions in the aggregated objective function diagram and denote the best solutions in the single objective function diagrams. Since the large size of the decision set reduces the clarity of visualization, we randomly sampled 50 pixels each from the foreground and background pixel sets to form a subset of the pixel pair set. As shown in the figure, for aggregated objective function and all individual objective functions, the gradient directions and the distribution of objective values are largely consistent across the objective functions of x_1 and x_2. The optimal solutions for these two unknown pixels are also close to each other. This implies that the solutions and heuristic information obtained by EAs while optimizing one unknown pixel are beneficial for optimizing the other.

In this paper, we propose a framework to improve the efficiency of EAs from the perspective of subsets of the decision set. The key in this framework is to estimate a microscale subset of the decision set for a group of similar subproblems. Two key points need to be considered when estimating this subset: (1) The solutions included in this subset should be beneficial for EAs to find the optimal solution. (2) The size of the subset should be microscale. This paper proposes a dynamic fitness evaluation allocation strategy based on convergence and divergence of solutions to estimate a microscale subset online from the solution sets when EAs solve similar subproblems. The information about the explored regions of the decision set is compressed in the microscale subset. This compression process is guided by EAs themselves. The contributions of this paper are summarized as follows:

(1) We propose a framework for enhancing the efficiency of EAs in solving the image matting optimization problem based on a microscale subset of the decision set. Unlike subproblem reduction methods, this approach uses the microscale subset to guide the search process of EAs, thereby reducing the exploration cost to solve numerous subproblems with large-scale decision sets.

(2) We propose a dynamic fitness evaluation allocation method based on the analysis of convergence and diversity, enabling EAs to automatically estimate the microscale subset during the problem-solving process.

(3) Experimental results on medical imaging datasets demonstrate that our method effectively improves the efficiency of EAs. Additionally, results under weak prior scenarios confirm that our approach makes EAs efficient methods for image matting.

The remainder of this paper is organized as follows. Section 2 discusses related work and gaps in existing methods. Sections 3 and 4 introduce our framework. Section 5 demonstrates the experimental validation results. Section 6 concludes the paper.

2. Related work

This section provides a concise review, categorizing existing image matting works into weak prior-based and strong prior-based methods based on the prior required for predicting alpha mattes.

2.1. Strong prior-based methods

Strong prior-based methods primarily rely on deep learning techniques. A large amount of training data with detailed annotated alpha mattes is required to train deep neural network models. The generalization performance of deep learning-based image matting methods is closely tied to the quality and diversity of data in image matting datasets. These methods emerged following the introduction of the Composition-1K dataset [8,19–21]. The diversity of foregrounds in the Composition-1K dataset is insufficient because the dataset includes many consecutive video frames and different cropped patches of the same images. Subsequently, datasets with more foreground diversity, such as Distinctions-646 [12] and SIMD [9], have been proposed. However, synthetic data is broadly produced to increase the number

of training data in most exiting methods due to the high complexity of model parameters. Consequently, the disparity between synthetic and real-world data significantly impacts the generalization performance of these methods [10]. Recently, some researchers have extended the use of image matting for segmenting the transition regions of lesions in medical images [4]. The boundaries of these transition regions are often blurred, ambiguous, and difficult to delineate with precise edges. Researchers have assumed a function mapping between the alpha values computed using the matting formula Eq. (1) and the uncertainty of whether a pixel belongs to lesions. However, these methods find it challenging to achieve consistent results across different imaging modalities constrained by the limited number of detailed annotated alpha mattes.

2.2. Weak prior-based methods

Methods based on weak prior refer to techniques that require prior that is effortless to be obtained, such as a trimap of the target image, for alpha matte prediction. Propagation-based methods, sampling-based methods, and optimization-based methods can be classified under this approach.

Propagation-based methods: Quadratic optimization models are constructed in propagation-based methods under a smoothness assumption. The smoothness assumption is that the optimal pixel pairs of pixels in certain regions change smoothly [22]. If the alpha values of some pixels in these regions are known, these values can be propagated to other pixels with unknown alpha values. The propagation of alpha values is confined to local regions in early methods [22,23]. More accurate prediction of alpha values are obtained by propagating alpha values of pixels in non-local regions that better accommodate the smoothness assumption [24,25]. However, the accuracy of alpha mattes predicted by these methods is limited because violations of the smoothness assumption, such as those with multiple holes, are widespread in natural images.

Sampling-based methods: Pixel features, such as pixel color, spatial, and texture information, are leveraged to find subsets of pixel pair set for unknown pixels in sampling-based methods [26–29]. The best pixel pairs are selected from these subsets using a hand-crafted function. Subsets of pixel pairs are found by selecting foreground and background pixels that have high similarity with unknown pixels [26]. Foreground and background pixels on the boundaries of local regions centered around unknown pixels are also frequently used to form pixel pair subsets [27,29]. For sampling-based methods, the quality of the pixel pairs included in the subset determines the accuracy of the predicted alpha mattes. If pixel pairs in the subsets have large difference from the optimal pairs, the accuracy of predicted alpha mattes will rapidly decline [6].

Optimization-based methods: Optimization-based methods select pixel pairs from the whole pixel pair set using well-designed single-objective functions or multi-objective functions. EAs are used to solve these functions due to their powerful ability of global optimization. Single-objective EAs are employed to search for optimal solutions based on an aggregation of various metric functions of pixel pairs [7,16,18]. However, conflicts among these metric functions complicate the objective function space, increasing the difficulty of finding optimal solutions [6]. Multi-objective models were introduced to mitigate these conflicts [5,6]. Existing methods focus on the improvement of the efficiency of EAs due to the large size of the decision set and numerous subproblems. Cooperative optimization framework is used to share heuristic information among subproblems [7,18]. Some methods leverage spatial correlation of pixels to reduce the number of subproblems [5]. However, The improvement of efficiency for EAs through subproblem reduction is limited, as the number of similar subproblems is significantly smaller compared to the total number of subproblems.

3. Microscale valid decision subset for large-scale combinatorial optimization problem of image matting

This section presents the concept of valid decision subsets, an algorithm to estimate a microscale valid decision subsets based on similar subproblems, and a framework for EAs to solve the image matting problem based on a microscale-searching strategy.

3.1. Assumption of microscale searching

Let $\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_N$ denote a group of similar subproblems and D denote the decision set for each subproblem that is the combination of the indices for foreground and background pixels. A feasible solution $\mathcal{X}$ for a subproblem in the decision set is encoded as a set of vectors (x_F, x_B) where $x_F \in \{1, 2, \ldots, |\Omega_F|\}$ and $x_B \in \{1, 2, \ldots, |\Omega_B|\}$. A subset of all feasible solutions that the algorithm can find within the maximum number of iterations is defined as:

Definition 1 (*Decision Subset*)**.** A decision subset V is the subset of a decision set D that consists of feasible solutions found by algorithms.

For the ith subproblem, a decision subset V_t^i is obtained at the tth iteration. Let S_t^i denote a region which can be theoretically explored by EAs at the tth iteration and V_t^i is a subset of S_t^i. Hausdorff distance $d_H(\cdot,\cdot)$ is used to calculate the distance between two sets. Let $\hat{\mathcal{X}}_t^i$ be the set of non-dominated solutions within V_t^i, we have $d_H(\hat{\mathcal{X}}_t^i, \mathcal{X}^*) \leq d_H(\mathcal{X}, \mathcal{X}^*), \forall \mathcal{X} \subseteq S_t^i$, where $\mathcal{X}$ is a random feasible solution and $\mathcal{X}^*$ is the Pareto set. As the number of iterations increases, there exists a time τ that $\forall t > \tau, P(\mathcal{X}^* \subset S_t^i) > P(\mathcal{X}^* \subset Q), Q = D/S_t^i$, implying that a solution $\hat{\mathcal{X}}_t^i$ closest to the Pareto set exists in the decision subset after the τth iteration. A decision subset can be defined as follows if it contains distinct $\hat{\mathcal{X}}_t^i$ for every tth iteration after the τth iteration:

Definition 2 (*Valid Decision Subset*)**.** Let $\mathcal{V}$ denote a decision subset of D. $\mathcal{V}$ is a valid decision subset if $\mathcal{V}$ satisfies the following conditions

$$d(\mathcal{X}^*, \hat{\mathcal{X}}) < d(\mathcal{X}^*, \overline{\mathcal{X}}) \quad \forall \hat{\mathcal{X}} \subseteq \mathcal{V}, \overline{\mathcal{X}} \nsubseteq \mathcal{V} \tag{7}$$

where $d(\cdot,\cdot)$ is the Hausdorff distance to calculate the difference between two sets of solutions. $\mathcal{X}^*$ is the Pareto set. $\hat{\mathcal{X}}$ is a subset of $\mathcal{V}$. $\overline{\mathcal{X}}$ is not a subset of $\mathcal{V}$.

Eq. (7) indicates that more solutions that have shorter distance to $\mathcal{X}^*$ can be found based on a valid decision subset because the Hausdorff distance suggests the existence of solutions closer to the optimal solution, since it is calculated using the maximum of the minimum distances between elements among two sets. We configure two search strategies for EAs: (1) Search based on v_t^i. (2) Search in the set $\mathcal{V}$. For the above two strategies, EAs needs to search among $|D| - |S_i|$ and $|\mathcal{V}|$ solutions, respectively. Although EAs under the second strategy can only search within a small set, the explored region of EAs intersects with that of EAs under the first strategy theoretically. The exploration cost of EAs is lower than that of the first strategy. The number of fitness evaluations required for searching within $\mathcal{V}$ is proportional to the size of $\mathcal{V}$. The remaining fitness evaluations can be used to explore extra regions based on the solutions found in $\mathcal{V}$.

3.2. Estimating a microscale valid decision subset

Algorithm 1 is proposed to estimate a valid decision subset for a group of similar subproblems. Let $\hat{\mathcal{X}}_k$ and $\hat{\mathcal{X}}_i$ denote the feasible solution sets of the x_k and $x_i, i, k \in \{1, 2, \ldots, N\}$. $\hat{\mathcal{X}}_k$ have a relationship with $\hat{\mathcal{X}}_i$ as follows:

$$\exists \epsilon \geq 0, d(\hat{\mathcal{X}}_k, \hat{\mathcal{X}}_i) \leq \epsilon \quad \forall \hat{\mathcal{X}}_k \subseteq \mathcal{V}_k, \hat{\mathcal{X}}_i \subseteq \mathcal{V}_i \tag{8}$$

where ϵ is a real number that is close to 0. It reflects how x_i is similar to x_k. If ϵ is equal to 0, x_i and x_k can be seen as the same. Let $\mathcal{V}_i$ denote

a valid decision subset of the x_i. $\hat{\mathcal{X}}_i^*$ is the Pareto set of x_i. $\overline{\mathcal{X}}_i$ is any solution set not included in $\mathcal{V}_i$. According to Definition 2, the relation of $\hat{\mathcal{X}}_i$, $\mathcal{X}_i^*$, and $\overline{\mathcal{X}}_i$ satisfies:

$$d(\hat{\mathcal{X}}_i, \mathcal{X}_i^*) < d(\overline{\mathcal{X}}_i, \mathcal{X}_i^*) \quad \forall \hat{\mathcal{X}}_i \subseteq \mathcal{V}_i, \overline{\mathcal{X}}_i \nsubseteq \mathcal{V}_i \tag{9}$$

If we replace $\hat{\mathcal{X}}_i$ by $\hat{\mathcal{X}}_k$, Eq. (9) can be rewritten as follows by substituting Eq. (8) into Eq. (9)

$$\forall \hat{\mathcal{X}}_k \subseteq \mathcal{V}_k, \hat{\mathcal{X}}_i \subseteq \mathcal{V}_i, \overline{\mathcal{X}}_k \nsubseteq \mathcal{V}_k \tag{10}$$

$$\exists \epsilon \geq 0, d(\hat{\mathcal{X}}_k, \mathcal{X}_i^*) \leq d(\overline{\mathcal{X}}_k, \mathcal{X}_i^*) + d(\overline{\mathcal{X}}_k, \hat{\mathcal{X}}_i) + \epsilon \tag{11}$$

Eq. (11) indicates that the solutions in $\mathcal{V}_i$ can be estimated by $\hat{\mathcal{X}}_k$. According to Eq. (11), the upper bound of estimation accuracy is determined by ϵ and $d(\overline{\mathcal{X}}_k, \mathcal{X}_i^*) + d(\overline{\mathcal{X}}_k, \hat{\mathcal{X}}_i)$. Once k is selected, the minimal ϵ is fixed. An accurate estimation of $\mathcal{V}_i$ needs to find $\mathcal{V}_k$ that minimizes $min(\{d(\overline{\mathcal{X}}_k, \mathcal{X}_i^*) + d(\overline{\mathcal{X}}_k, \hat{\mathcal{X}}_i) | \hat{\mathcal{X}}_i \subseteq \mathcal{V}_i, \overline{\mathcal{X}}_k \nsubseteq \mathcal{V}_k\})$. It can be achieved by maximizing the size of the intersection of $\overline{\mathcal{V}}_k$ and $\mathcal{V}_i$ where $\overline{\mathcal{V}}_k$ is the complement set of $\mathcal{V}_k$. According to Fig. 3, similar local regions exist in the objective spaces of similar subproblems. A large intersection is more likely to contain solutions in these local regions. $d(\overline{\mathcal{X}}_k, \hat{\mathcal{X}}_i)$ diminishes as solutions within these local regions are included in $\overline{\mathcal{V}}_k$. $d(\overline{\mathcal{X}}_k, \mathcal{X}_i^*)$ decreases along with $d(\overline{\mathcal{X}}_k, \hat{\mathcal{X}}_i)$. The maximization of the size of $\overline{\mathcal{V}}_k$ is equal to minimizing the size of $\mathcal{V}_k$. It means that the estimation of a valid decision subset for a group of similar subproblems can be achieved by finding a microscale subset from the decision subset.

We estimate a valid decision subset by gradually constructing a subset during the optimization of subproblems in **Algorithm** 1. The non-dominated solutions are selected from the population of each generation to limit the increase in the size of the estimated subset. The size of the estimated valid decision subset is also constrained by controlling the allocation of fitness evaluations. Two indicators are introduced to control this allocation.

The first indicator utilizes the non-dominated solutions from both the current and history populations to measure the convergence trend of EAs. The indicator is defined as follows

$$I_1 = |Q_h \cap Q_c| / |Q_h| \tag{12}$$

where Q_h and Q_c are the sets of non-dominated solutions from the current and history populations, respectively. A high value of I_1 indicates that non-dominated solutions have been barely updated in recent iterations. This suggests that EAs struggle to benefit from the currently explored region. The allocation of fitness evaluations should decrease as the value of I_1 increases. The second indicator is used to assess the quality of the regions currently being explored by EAs, based on the repetitive evaluation of solutions across different subproblems. The indicator is defined as follows

$$I_2 = \frac{1}{|X_R|} \sum_{i=1}^{|X_R|} \exp(-1 * c_i) \tag{13}$$

$$c_i = \sigma(\hat{x}_i) \quad \hat{x}_i \in X_R, i = 1, 2, \ldots, |X_R| \tag{14}$$

where X_R is the set of solutions repetitively evaluated for different subproblems. σ is used to measure the quality of the regions that contain these solutions. We use crowding distance as σ [30]. A high value of I_2 suggests EAs explore similar regions for these subproblems, as the objective function values are close. If this high value occurs early in the optimization process, it may result from random initialization. EAs can explore different regions in subsequent iterations. If this high value continues to appear during optimization, EAs might explore these similar regions in future iterations, as EAs tends to converge as the number of iterations increases. The aggregation of these two indicators is adaptively adjusted for the ith subproblem according to the number of iterations as follows

$$I_i = \eta_i \cdot I_1 + (1 - \eta_i) \cdot I_2 \tag{15}$$

$$\eta_i = \gamma^{c_i/\tau} \tag{16}$$

where γ is a decay parameter close to 1. c_i is the number of iterations consumed by EAs to optimize the ith subproblem. The number of fitness evaluations to be allocated is positively proportional to I_i. Eq. (15) indicates that in the early stages of optimization, the allocation of fitness evaluations is primarily governed by the convergence trend. If EAs converge early, fitness evaluations will be preserved for EAs to continue search under the guidance of the valid decision subset to find better solutions. As iterations increase, the allocation of fitness evaluations is mainly controlled by the quality of the explored regions. If the explored region is low-quality, fitness evaluations will also be preserved. I_i is updated after certain iterations. The algorithm to estimate a valid decision subset is shown in Algorithm 1.

Algorithm 1 Estimation of a Valid Decision Subset

Input: γ,τ,t_1,e,population size n_p,number of subproblems n_s
Output: estimated valid decision subset $\mathcal{V}$,
preserved fitness evaluations $\{v_1, v_2, ..., v_{n_s}\}$

```
Initialize a population P_i for each subproblem, i = 1, 2, ..., n_s
Initialize H_i as an empty set, i = 1, 2, ..., n_s;
Initialize r_i as 1 for i = 1, 2, ..., n_s
Initialize v_i as 0 for i = 1, 2, ..., n_s
Initialize V as empty set
Allocate ⌊e/n_s⌋ fitness evaluations for each subproblem
for i = 1 to ⌊e/(n_s · n_p · t_1)⌋ do
  for j = 1 to n_s do
    s ← t_1 · n_p · r_j
    v_j ← v_j + t_1 · n_p · (1 − r_j)
    if s < n_p then
      continue
    end if
    Update P_j by EAs with s fitness evaluations
    Calculate I_1 using Eq. (12) with P_j and H_j
    Calculate I_2 using Eq. (13) with P_j and H_1, H_2, ..., H_{n_s}
    Calculate r_j using Eq. (15) with I_1 and I_2
    H_j ← H_j ∪ P_j
    Obtain the non-dominated solutions from P_j and add them into V
  end for
end for
return V, {v_1, v_2, ..., v_{n_s}}
```

Let all subproblems be clustered into M groups, with each group containing $n_i, i = 1, 2, \ldots, M$ subproblems. The size of the valid decision subset estimated by **Algorithm** 1 is less than the number of fitness evaluations used for n_i subproblems:

$$|\mathcal{V}| \leq n_i \cdot e. \tag{17}$$

Let D denote the decision set. The size of the D is $|D| = |\Omega_F| \cdot |\Omega_B|$. Let η denote the ratio of sizes of the valid decision subset to that of the decision set. η can be written as:

$$\eta = \frac{|\mathcal{V}|}{|D|} \leq \frac{e}{|\Omega_F| \cdot |\Omega_B|} \tag{18}$$

In practice, the number of fitness evaluations is far less than $|\Omega_F| \cdot |\Omega_B|$. The value of η will become much smaller than 1 as the number of unknown pixels increases. Therefore, the size of the estimated valid decision subset $\mathcal{V}$ is microscale.

4. Framework of microscale-searching evolutionary optimization for image matting

This section presents a framework called microscale-searching evolutionary optimization (MSEO) for EAs to solve the image matting problem based on microscale-searching as shown in Fig. 5. Microscale-searching refers to the process of searching for the best pixel pairs under the guidance of the microscale valid decision subset estimated by **Algorithm** 1. Fig. 4 illustrates the differences between the microscale-searching evolutionary optimization image matting method and existing EAs-based matting methods. Assume that the EA is used to optimize three similar subproblems: x_1, x_2, and x_3. Fig. 4 depicts solutions found

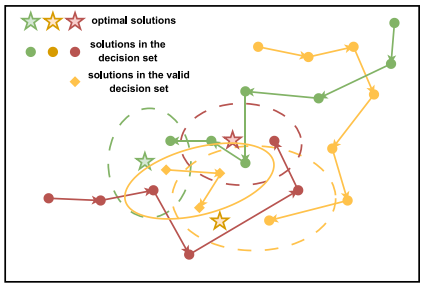


Fig. 4. The black rectangle represents the decision set. Red, green, and yellow solid circles represent feasible solutions for subproblems x_1, x_2, and x_3, respectively. Yellow diamonds also represent feasible solutions for x_3. Solid lines indicate the EA's search process. Three different colored dashed ellipses denote the optimal valid decision subsets for different subproblems. A yellow solid ellipse represents the valid decision subset estimated from the feasible solutions of the other two subproblems.

during the optimization of x_3 as yellow circles and yellow diamonds. The diamond specifically represent solutions within the valid decision subset. This subset is estimated based on the solutions of x_1 and x_2 (shown as green and brown circles, respectively), leveraging subproblem similarity. We assume the EA can find better solutions using fewer evaluations since the valid decision subset is close to the optimal solution for x_3.

According to Definition 2, solutions within a valid decision subset are supposed to be closer to the Pareto set than other solutions. The allocation of fitness evaluations should be prioritized to search in the valid decision subset. Fitness evaluations should also be allocated to search outside the valid decision subset considering the estimation error between the estimated valid decision subset and the optimal valid decision subset. EAs can search outside the valid decision subset more efficiently under the guidance of the valid decision subset, as repetitive evaluations can be reduced. In the framework, EAs search alternatively inside and outside the valid decision subset. The best solution found in the valid decision subset serves as the initial solution for EAs to search outside the valid decision subset. Non-dominated solutions found outside the valid decision subsets are added into the valid decision subset.

On the other hand, the similarity among subproblems within the same group is inconsistent. For subproblems with lower similarity, the deviation of estimated valid decision subset from the optimal valid decision subset tends to be larger compared to those with higher similarity. The allocation of fitness evaluations should be prioritized to search outside the valid decision subset. The similarity of subproblems is calculated using the Euclidean distance of computer vision feature vectors of pixels. As suggested in [25], the feature vector f_i of the x_i is defined as $f_i = (r_i, g_i, b_i, x_i, y_i)^T$, where $(r_i, g_i, b_i)^T$ and $(x_i, y_i)^T$ are the RGB color vector and coordinate vector of the ith unknown pixel. Subproblems with smaller distance values exhibit higher similarity. In the framework, a roulette wheel method determines whether fitness evaluations are allocated to search within or outside the valid decision subset based on a probability vector W. W is obtained by the feature vectors as follows:

$$\overline{f} = \frac{1}{n_j}\sum_{i=1}^{n_j} f_i, w_i = \frac{\|f_i - \overline{f}\|_2}{\sum_{k=1}^{n_j}\|f_k - \overline{f}\|_2} \tag{19}$$

$$W = (w_1, w_2, \ldots, w_n) \tag{20}$$

$\overline{f}$ is the centroid vector of feature vectors. Subproblems that are far from the centroid vector are more likely to be selected to search outside the valid decision subsets.

The algorithm to find the best solution with the valid decision subset is shown in **Algorithm** 2.

In the algorithm, EAs search alternatively in and out of the valid decision subset according to a_i (Lines 12–18) every t_2 iterations. EAs uses the population updated in the subset as the initial solution if EAs search out of the valid decision subset (Line 13). EAs are not guided by

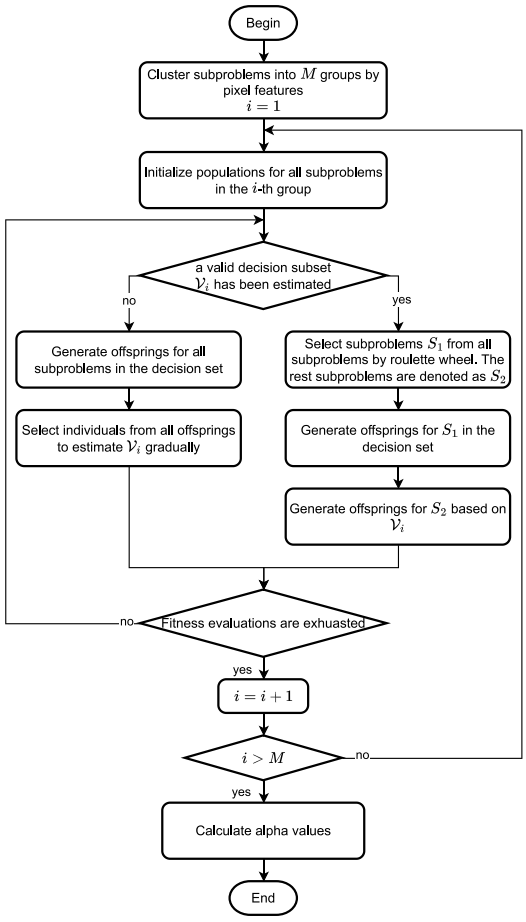


Fig. 5. Framework of microscale-searching evolutionary optimization.

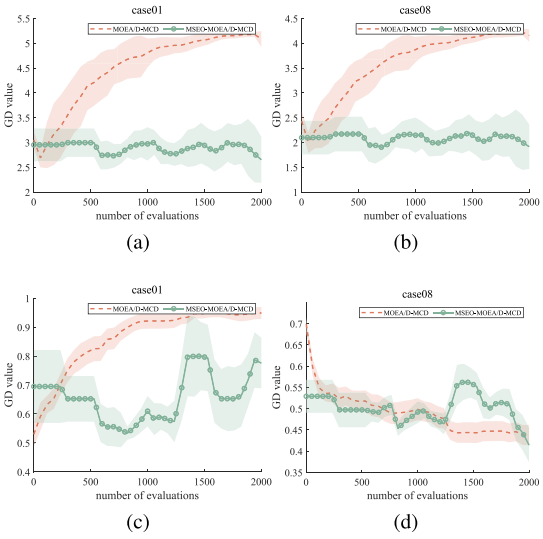


Fig. 6. Convergence comparison of MOEA/D-MCD and MS-MOEA/D. (a) and (b) present the changes of GD metric for unknown pixels with the highest similarity. (c) and (d) present the changes of GD metric for unknown pixels with the lowest similarity. The curves in the figure represent the average GD changes over 30 experiments, while the bandwidth indicates the variance of the GD changes.

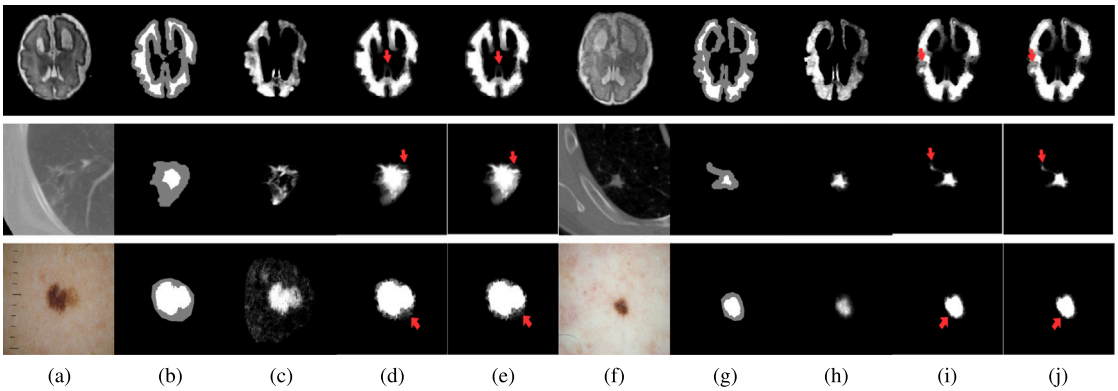


Fig. 7. Visual comparison on samples from Brain-growth, LIDC, and ISIC datasets. (a)–(c) and (f)–(h) are original image, trimap, ground truth alpha matte, respectively. (d)–(e) and (i)–(j) are predicted alpha mattes of MOEA/D-MCD, MS-MOEA/D, respectively. The red arrows indicate regions where the differences in the predicted alpha mattes from these methods are obvious.

the valid decision subset if optimizing subproblems with low similarity (Line 10). Finally, the best pixel pairs are selected from the non-dominated solutions of the three populations based on an aggregate evaluation function (Line 22).

5. Experiments

This section will present several experiments to address the following questions:

Q1: What is the effect of the microscale-searching strategy on the performance of EAs in solving the medical image matting problem?

Q2: How does the EAs based on the microscale-searching strategy perform in weak prior scenarios?

E1: The microscale-searching strategy-based EAs are compared with the original versions of EAs on three medical datasets from the perspectives of matting metrics, visualization results, and convergence trends (for answering Q1).

E2: The microscale-searching strategy-based EAs are compared with other image matting methods under different weak prior scenarios (for answering Q2).

5.1. Experimental setup

In this study, three medical image matting datasets, Brain-growth of the QUBIQ dataset [13], LIDC-IDRI [14], and ISIC [15], and a natural image matting dataset AlphaMatting [11] are used to evaluate the performance of different methods. Brain-growth consists of low-intensity contrast T2-W MR images for the newborn brain's white matter tissue myelination process. LIDC-IDRI dataset includes thoracic CT scans for lung nodules diagnosis. ISIC is a large-scale dermoscopic images dataset consisting of 10,000 images. 39 images of Brain-growth, 1609 images of LIDC-IDRI, and 120 images of ISIC are selected and labeled alpha mattes [4]. 27 natural images are available for evaluation in AlphaMatting.

We select MOEA/D-MCD [5] to be embedded into the framework for our experiments, which we designate MS-MOEA/D. MOEA/D-MCD has been verified as a method based on multi-objective evolutionary algorithms capable of accurately predicting alpha mattes. However, some solutions that contribute negatively to the whole optimization are repeatedly evaluated when similar subproblems are optimized. We verify the effectiveness of the microscale-searching strategy in our experiments by comparing the performance of MS-MOEA/D against MOEA/D-MCD. The parameter settings for the proposed method are as follows: The parameters of MOEA/D-MCD are set according to the recommended settings. In Algorithm 1, the decay factor γ is set to 0.99, with τ and t_1 set to 10 and 5, respectively. In Algorithm 2, t_2 is set to 5.

Algorithm 2 Scheme of Allocation of FEs among Different Subproblems

Input: $\mathcal{V}$,n_s,n_p,t_2,D,W,$\{v_1, v_2, ..., v_{n_s}\}$;
Output: the best pixel pairs $p_1^*, p_2^*, ..., p_{n_s}^*$;
1: Initialize $a_i = 0, i = 1, 2, ..., n_s$
2: Initialize populations $P_i^{in}, P_i^{out}, \hat{P}_i^{out}$ for each subproblem, $i = 1, 2, ..., n$
3: **while** $v_i \geq n_p, i = 1, 2, ..., n$ **do**
4: **for** $i = 1$ to n_s **do**
5: **if** $v_i < n_p$ **then**
6: continue
7: **end if**
8: $idx \leftarrow$ the index of a decision vector selected by roulette-wheel selection based on W
9: **if** $i == idx$ **then**
10: Update $\hat{P}_i^{out}$ by searching in D
11: **else**
12: **if** $a_i == 1$ **then**
13: Use P_i^{in} as the initial population to update P_i^{out} by searching in D and Update $\mathcal{V}$
14: $a_i \leftarrow 0$
15: **else**
16: Update P_i^{in} by searching in $\mathcal{V}$
17: $a_i \leftarrow 1$
18: **end if**
19: **end if**
20: $v_i \leftarrow v_i - t_2 \cdot n_p$
21: **if** $v_i < n_p$ **then**
22: Obtain the best pixel pair p_i^* from the non-dominated solutions of $P_i^{in}, P_i^{out}, \hat{P}_i^{out}$
23: **end if**
24: **end for**
25: **end while**
26: **return** $p_1^*, p_2^*, ..., p_{n_s}^*$

Three measurement functions are used as objectives suggested by [5]. For the ith unknown pixel, selected pth foreground pixel, and q background pixel, $h_{1i}(p, q)$ is color chromatic criteria, $h_{2i}(p)$ is spatial closeness criteria for foreground pixels, and $h_{3i}(q)$ is spatial closeness criteria for background pixels. The formulations of three criteria are presented as:

$$h_{1i}(p, q) = \|(C_U)_i - \hat{\alpha}(C_F)_p - (1 - \hat{\alpha})(C_B)_q\|^2 \tag{21}$$

$$h_{2i}(p) = \|(S_U)_i - (S_F)_p\|^2 \tag{22}$$

$$h_{3i}(q) = \|(S_U)_i - (S_F)_p\|^2. \tag{23}$$

$(C_x)_i$ and $(S_x)_i$ are the ith color vector and the ith spatial coordinate vector of the pixel in $x, x \in \{F, B, U\}$, where F, B, U represent the sets of foreground pixels, background pixels, and unknown pixels, respectively. According to [5], the combination of measurement functions H

Table 2
The results of the mean and standard deviation for all samples in three datasets. Numbers in boldface indicate the best results.

(a) Brain-growth		
	MOEA/D-MCD	MS-MOEA/D
MSE	0.0744 ± 0.0237	**0.0740 ± 0.0237**
SAD	0.5634 ± 0.1016	**0.5619 ± 0.1005**
CON	0.8834 ± 0.2851	**0.8758 ± 0.2836**
GRAD	0.5684 ± 0.1050	**0.5667 ± 0.1035**
(b) LIDC		
	MOEA/D-MCD	MS-MOEA/D
MSE	0.0262 ± 0.0250	**0.0262 ± 0.0250**
SAD	0.0525 ± 0.0397	**0.0524 ± 0.0396**
CON	0.0788 ± 0.0767	**0.0786 ± 0.0765**
GRAD	0.0507 ± 0.0395	**0.0507 ± 0.0394**
(c) ISIC		
	MOEA/D-MCD	MS-MOEA/D
MSE	0.1784 ± 0.0706	**0.1765 ± 0.0710**
SAD	7.5443 ± 6.5885	**7.4518 ± 6.5220**
CON	4.8961 ± 2.4095	**4.8656 ± 2.4076**
GRAD	7.7882 ± 6.7793	**7.6810 ± 6.6958**

Table 3
Number of samples on the Brain-growth dataset for which the proposed method is significantly better than, equal to, or significantly worse than MOEA/D-MCD.

MSE	SAD	CON	GRAD
29/0/10	27/0/12	30/0/9	27/0/12

is defined as:

$$H(h_{1i}, h_{2i}, h_{3i}) = \frac{h_{1i} \cdot 0.5(h_{2i} + h_{3i})}{1 - (1 - h_{1i}) \cdot (1 - 0.5 \cdot (h_{2i} + h_{3i}))} \tag{24}$$

H is used to select the best pixel pairs from the non-dominated solutions of the final population.

5.2. Effect of microscale-searching strategy

In this subsection, we validate the effect of the microscale-searching strategy by comparing MOEA/D-MCD [5] and MOEA/D-MCD based on microscale-searching strategy (MS-MOEA/D). Four metrics, MSE, SAD, CON, and CON, are used to measure the quality of predicted alpha mattes. MSE and SAD measure the numerical difference between the predicted alpha mattes and the ground truth. CON and GRAD assess the visual difference between the predicted alpha mattes and the ground truth. Small values indicate high-quality predicted alpha mattes for these four metrics.

Table 2 summarize the mean and standard deviation of the matting metrics for all methods across all datasets. The best results are indicated in boldface. The result indicates that the microscale-searching strategy enables MOEA/D-MCD to predict alpha mattes with higher accuracy. The improvement of MOEA/D-MCD is not significant on the LIDC dataset. This is because images in the LIDC contain the fewest unknown pixels among the three datasets, which reduces the number of similar subproblems. Consequently, the diversity of solutions in the valid decision subset decreases, making it difficult to effectively enhance the efficiency of MOEA/D-MCD. We conducted the Wilcoxon rank-sum test on the results from 30 independent experiments on the Brain-growth dataset to demonstrate the statistical superiority of the proposed method. Table 3 presents the number of samples, out of a total of 39, for which the proposed method is significantly better than, equal to, or significantly worse than MOEA/D-MCD across four matting metrics. The results in Table 3 indicate that the proposed method is significantly better than MOEA/D-MCD on a majority of samples for these four matting metrics.

Fig. 7 visualizes several predicted alpha mattes. It shows that the microscale-searching strategy can improve the accuracy of predicted alpha values of regions where the foreground and background are very similar.

We demonstrate the enhancement in the EA's optimization efficiency due to the microscale-searching strategy by tracking the change in Generational Distance (GD) during optimization. GD provides a more direct reflection of solution quality compared to matting metrics. The enhancement is related to the valid decision subsets according to the proposed method which are estimated based on subproblem similarity. We observed the GD changes using two groups selected from two samples case01 and case08 in the Brain-growth dataset.

The curves in Fig. 6 represent the average GD changes over 30 independent experiments, while the bandwidth indicates the variance of the GD changes. A small value of GD indicates that the solutions found by the algorithm are closer to the reference set. The Pareto fronts of subproblems are used as the reference sets. The results in Fig. 6 show that without the aid of the microscale-searching strategy, GD cannot diverge to a small value when MOEA/D-MCD optimizing the two groups of subproblems in case01 and case08 with the highest similarity. Under the guidance of the microscale-searching strategy, GD can converge to a small value smoothly. This indicates that the microscale-searching strategy enables MOEA/D-MCD to find high-quality solutions with fewer fitness evaluations. During the optimization of the group of subproblems with the lowest similarity in case08, although the effect of the estimated valid decision subset is limited by the low similarity of subproblems, the small size of valid decision subset ensures that there are still sufficient fitness evaluations for MOEA/D-MCD to search beyond the valid decision subset to converge. Overall, the microscale-searching strategy can facilitate EAs find high-quality solutions, resulting in more accurate alpha mattes.

5.3. Performance on weak-prior scenarios

In this section, we demonstrate that microscale-searching strategy can make MOEA/D-MCD an efficient method in weak prior scenarios by comparing different image matting methods and the microscale-searching-based MOEA/D-MCD. The strong prior-based methods include DIM [8], FBA [31], DiffMatte [32] and MedicalMatting [4]. DIM, FBA, and DiffMatte are pre-trained on the Composition-1K dataset [8]. The MedicalMatting models pre-trained on the Brain-growth, LIDC, and ISIC datasets are denoted as MM†, MM‡, and MM$'$, respectively. The weak prior methods include: (1) Propagation-based methods: Closed-Form [22], KNN [25], Information-Flow [33]. (2) Sampling-based methods: Bayesian Matting [34], PDMS [6]. We construct a weak prior dataset in the following manner: randomly selecting 27 images from one medical dataset and combining them with the AlphaMatting dataset to form a new dataset. The prediction of alpha mattes for this dataset represents a weak prior scenario for MedicalMatting, DIM, and FBA, as the training data for MedicalMatting does not include natural images, and the training data for DIM and FBA does not include medical images. When predicting the mask of the target image, only a trimap is provided.

The sub-tables within Table 4 represent the results of data from the Brain-growth, LIDC, and ISIC datasets combined with the natural image dataset, respectively. The results of Table 4 show that the proposed method achieve the best result on MSE over all methods. MSE reflects the difference between the predicted alpha matte and the ground truth alpha matte. Compared to SAD, MSE is more sensitive to variations in these differences. This result indicates that most weak-prior methods can predict the alpha matte more accurately than other methods. Most weak prior methods achieve results comparable to, or even better than, those of deep learning-based methods on MSE. One reason is that the deep learning-based methods were not trained on datasets that have data with the same distribution as test datasets. Additionally, the efficiency of weak prior methods is less affected by variations in data distribution compared to deep learning-based methods. PDMS is the best-performing sampling-based method. The strong performance of

Table 4
The results of the mean and standard deviation for all samples in three datasets. Numbers in boldface indicate the best results.

(a) Brain-growth				
Methods	MSE	SAD	CON	GRAD
DIM	0.0530 ± 0.0012	2.2397 ± 0.0073	2.4806 ± 0.0136	2.1518 ± 0.0076
FBA	0.0725 ± 0.0031	1.7979 ± 0.0120	1.7552 ± 0.0253	1.7624 ± 0.0122
MM‡	0.3372 ± 0.0025	20.1287 ± 0.0140	17.9438 ± 0.0307	20.3772 ± 0.0143
MM′	0.3746 ± 0.0036	22.0245 ± 0.0151	20.8340 ± 0.0356	22.0070 ± 0.0153
DiffMatte	0.1695 ± 0.0827	**0.8596** ± 0.2528	**0.8746** ± 0.2588	**1.4369** ± 0.5392
Closed-Form	0.0500 ± 0.0013	2.9354 ± 0.0058	5.0067 ± 0.0134	2.8837 ± 0.0062
KNN	0.0564 ± 0.0007	3.6230 ± **0.0050**	4.1823 ± **0.0098**	3.5249 ± **0.0054**
Information-Flow	0.0464 ± 0.0013	2.7059 ± 0.0054	2.5809 ± 0.0103	2.6365 ± 0.0057
Bayesian	0.1548 ± 0.0016	14.8008 ± 0.0065	17.5435 ± 0.0114	14.9725 ± 0.0067
PDMS	0.0443 ± 0.0016	2.7769 ± 0.0057	2.6824 ± 0.0131	2.6692 ± 0.0060
MS-MOEA/D	**0.0426 ± 0.0015**	2.5310 ± 0.0052	2.4297 ± 0.0119	2.4392 ± 0.0055
(b) LIDC				
Methods	MSE	SAD	CON	GRAD
DIM	0.0339 ± 0.0042	1.9666 ± 0.0049	1.9250 ± 0.0164	1.8745 ± 0.0048
FBA	0.0232 ± 0.0072	1.4407 ± 0.0051	1.2064 ± 0.0151	1.3984 ± 0.0052
MM†	0.3300 ± 0.0083	19.8434 ± 0.0080	20.6576 ± 0.0199	19.7624 ± 0.0081
MM′	0.3862 ± 0.0053	21.5872 ± 0.0123	19.0615 ± 0.0267	21.5728 ± 0.0125
DiffMatte	0.1480 ± 0.1282	**0.1194** ± 0.0661	**0.1197** ± 0.0683	**0.2568** ± 0.2373
Closed-Form	0.0257 ± **0.0015**	2.6831 ± 0.0057	4.5389 ± 0.0085	2.6265 ± 0.0058
KNN	0.0422 ± 0.0030	3.3692 ± 0.0053	3.6968 ± 0.0106	3.2625 ± 0.0053
Information-Flow	0.0252 ± 0.0006	2.4565 ± **0.0021**	2.1586 ± **0.0042**	2.3821 ± **0.0021**
Bayesian	0.1649 ± 0.0043	14.5393 ± 0.0071	17.1334 ± 0.0113	14.7074 ± 0.0071
PDMS	0.0193 ± 0.0022	2.5187 ± 0.0049	2.2635 ± 0.0079	2.4071 ± 0.0048
MS-MOEA/D	**0.0184** ± 0.0028	2.2751 ± 0.0040	2.0356 ± 0.0095	2.1797 ± 0.0041
(c) ISIC				
Methods	MSE	SAD	CON	GRAD
DIM	0.2558 ± 0.0141	21.0238 ± 0.3167	21.2210 ± 0.2194	20.6760 ± 0.3327
FBA	0.3574 ± 0.0176	24.2613 ± 0.8494	16.8167 ± 0.1655	24.6705 ± 0.8887
MM†	0.2576 ± 0.0117	21.0875 ± **0.3235**	21.2290 ± 0.2150	20.6894 ± **0.3093**
MM‡	0.3644 ± 0.0196	24.6034 ± 1.0938	16.8411 ± 0.2019	25.0629 ± 1.1702
DiffMatte	0.1451 ± 0.1019	6.6806 ± 6.9632	**3.0475** ± 1.8488	6.7881 ± 7.1212
Closed-Form	0.1350 ± 0.0039	6.9599 ± 0.7856	7.0953 ± 0.2204	7.0338 ± 0.8047
KNN	0.0955 ± 0.0044	6.6171 ± 0.4830	7.1436 ± 0.3343	6.5674 ± 0.5003
Information-Flow	0.1289 ± 0.0089	6.4981 ± 0.9441	5.1209 ± 0.3253	6.5239 ± 0.9852
Bayesian	0.2557 ± 0.0101	20.8585 ± 1.3120	19.2447 ± **0.1555**	21.2693 ± 1.3133
PDMS	0.1042 ± 0.0079	6.4836 ± 0.6320	4.9668 ± 0.3318	6.5051 ± 0.6405
MS-MOEA/D	**0.0939 ± 0.0034**	**5.9425** ± 0.5536	4.3701 ± 0.1764	**5.9452** ± 0.5438

Table 5
Mean and standard deviation of matting metrics for predicted alpha mattes from all samples of Brain-growth dataset under different settings of t_1 and t_2.

t_1/t_2	MSE	SAD	CON	GRAD
5/5	7.350E−02(2.348E−02)	5.598E−01(9.783E−02)	8.548E−01(2.742E−01)	5.646E−01(1.009E−01)
5/10	7.356E−02(2.341E−02)	5.602E−01(9.742E−02)	8.544E−01(2.727E−01)	5.652E−01(1.000E−01)
5/15	7.363E−02(2.305E−02)	5.606E−01(9.645E−02)	8.548E−01(2.727E−01)	5.654E−01(9.901E−02)
10/5	7.361E−02(2.355E−02)	5.607E−01(9.935E−02)	8.613E−01(2.774E−01)	5.657E−01(1.022E−01)
10/10	7.360E−02(2.354E−02)	5.603E−01(9.764E−02)	8.607E−01(2.753E−01)	5.652E−01(1.006E−01)
10/15	7.357E−02(2.340E−02)	5.606E−01(9.825E−02)	8.616E−01(2.737E−01)	5.655E−01(1.011E−01)
15/5	7.371E−02(2.348E−02)	5.609E−01(9.934E−02)	8.676E−01(2.794E−01)	5.658E−01(1.025E−01)
15/10	7.328E−02(2.318E−02)	5.594E−01(9.884E−02)	8.641E−01(2.747E−01)	5.643E−01(1.017E−01)
15/15	7.356E−02(2.329E−02)	5.606E−01(9.924E−02)	8.670E−01(2.782E−01)	5.654E−01(1.023E−01)

PDMS indicates that designing efficient methods to find higher quality non-dominated solution sets can improve the accuracy of predicted alpha mattes, as PDMS uses the non-dominated solution set as the sampling subset of pixel pairs. MS-MOEA/D achieved the best results in all weak prior scenarios. This indicates that the microscale-searching strategy can make EA an efficient matting method in weak prior scenarios.

5.4. Parameter investigation

This subsection presents the results for the proposed method with varying values of t_1 and t_2 to further investigate its underlying mechanism. Table 5 summarizes the mean and standard deviation of matting metrics for the predicted alpha mattes generated by the proposed method on all samples of the Brain-growth dataset. The parameter t_2 is proportional to the total number of iterations allocated for the EA to search for solutions outside the valid decision subset. Table 5 indicates that when t_1 equals 10 and 15, the allocation of an excessive number of fitness evaluations leads to a decrease in the accuracy of the predicted alpha mattes. This also suggests that the valid decision subset contains better solutions compared to other regions of the decision set. When $t_1 = 5$, the algorithm obtains better alpha mattes if more fitness evaluations are allocated outside the valid decision subset. This occurs because t_1 is related to the diversity of solutions within the valid decision subset. A small value of t_1 reduces the diversity of

solutions within this subset. An increase in t_1 enriches the diversity of solutions within the valid decision subset. However, this also raises the exploration cost for the valid decision subset. Consequently, the quality of the solutions can be unstable, as evident from the results when $t_1 = 15$.

6. Conclusion

In this paper, we propose a framework for MOEAs to solve the medical image matting problem based on a microscale-searching strategy. The exploration cost of MOEAs in large-scale decision set is reduced by guiding EAs to search in the valid decision subset estimated in the framework. Experimental results demonstrate that EAs with the microscale-searching strategy are able to find competitive solutions with fewer FEs compared to the results obtained solely by EAs. The alpha mattes predicted by the EAs with the microscale-searching strategy are more accurate than those of EAs-based methods. In the future, we will try to design more efficient EAs-based approaches based on the concept of microscale valid decision subsets for other computer vision tasks.

CRediT authorship contribution statement

Li Kang: Writing – review & editing, Validation, Formal analysis, Writing – original draft, Methodology, Conceptualization, Visualization, Investigation. **Han Huang:** Supervision, Funding acquisition, Conceptualization. **Yihui Liang:** Validation, Supervision. **Baoxiong Zhuang:** Data curation, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Minglun Gong, Yiming Qian, Li Cheng, Integrated foreground segmentation and boundary matting for live videos, IEEE Trans. Image Process. 24 (4) (2015) 1356–1370.

[2] Wenyuan Li, Zhengxia Zou, Zhenwei Shi, Deep matting for cloud detection in remote sensing images, IEEE Trans. Geosci. Remote. Sens. 58 (12) (2020) 8490–8502.

[3] Bingke Zhu, Yingying Chen, Jinqiao Wang, Si Liu, Bo Zhang, Ming Tang, Fast deep matting for portrait animation on mobile phone, in: Proc. ACM Int. Conf. Multimedia, 2017, pp. 297–305.

[4] Lin Wang, Xiufen Ye, Lie Ju, Wanji He, Donghao Zhang, Xin Wang, Yelin Huang, Wei Feng, Kaimin Song, Zongyuan Ge, Medical matting: Medical image segmentation with uncertainty from the matting perspective, Comput. Biol. Med. 158 (2023) 106714.

[5] Yihui Liang, Han Huang, Zhaoquan Cai, Zhifeng Hao, Multiobjective evolutionary optimization based on fuzzy multicriteria evaluation and decomposition for image matting, IEEE Trans. Fuzzy Syst. 27 (5) (2019) 1100–1111.

[6] Han Huang, Yihui Liang, Xiaowei Yang, Zhifeng Hao, Pixel-level discrete multi-objective sampling for image matting, IEEE Trans. Image Process. 28 (8) (2019) 3739–3751.

[7] Fu-Jian Feng, Han Huang, Qiu-Xia Wu, X Ling, YH Liang, ZQ Cai, An alpha matting algorithm based on collaborative swarm optimization for high-resolution images, Sci. Sin. Informationis 50 (3) (2020) 424–437.

[8] Ning Xu, Brian L. Price, Scott Cohen, Thomas S. Huang, Deep image matting, in: Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2017, pp. 311–320.

[9] Yanan Sun, Chi-Keung Tang, Yu-Wing Tai, Semantic image matting, in: Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2021, pp. 11120–11129.

[10] Jizhizi Li, Jing Zhang, Stephen J. Maybank, Dacheng Tao, Bridging composite and real: Towards end-to-end deep image matting, Int. J. Comput. Vis. 130 (2) (2022) 246–266.

[11] Christoph Rhemann, Carsten Rother, Jue Wang, Margrit Gelautz, Pushmeet Kohli, Pamela Rott, A perceptually motivated online benchmark for image matting, in: Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2009, pp. 1826–1833.

[12] Yu Qiao, Yuhao Liu, Xin Yang, Dongsheng Zhou, Mingliang Xu, Qiang Zhang, Xiaopeng Wei, Attention-guided hierarchical structure aggregation for image matting, in: Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2020, pp. 13673–13682.

[13] Hongwei Bran Li, Fernando Navarro, et al. Ivan Ezhov, QUBIQ: uncertainty quantification for biomedical image segmentation challenge, 2024, CoRR, arXiv: 2405.18435.

[14] Samuel G. Armato, Geoffrey McLennan, et al., The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans, Med. Phys. 38 (2) (2011) 915–931.

[15] Noel C.F. Codella, David A. Gutman, M. Emre Celebi, et al., Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC), in: ISBI, IEEE, 2018, pp. 168–172.

[16] Yihui Liang, Han Huang, Zhaoquan Cai, Liang Lv, Particle swarm optimization with convergence speed controller for sampling-based image matting, in: Int. Conf. Intelligent Computing, 2018, pp. 656–668.

[17] Qingfu Zhang, Hui Li, MOEA/D: a multiobjective evolutionary algorithm based on decomposition, IEEE Trans. Evol. Comput. 11 (6) (2007) 712–731.

[18] Zhaoquan Cai, Liang Lv, Han Huang, Hui Hu, Yihui Liang, Improving sampling-based image matting with cooperative coevolution differential evolution algorithm, Soft Comput. 21 (15) (2017) 4417–4430.

[19] Jingwei Tang, Yagiz Aksoy, Cengiz Öztireli, Markus H. Gross, Tunç Ozan Aydin, Learning-based sampling for natural image matting, in: CVPR, Computer Vision Foundation / IEEE, 2019, pp. 3055–3063.

[20] Yaoyi Li, Hongtao Lu, Natural image matting via guided contextual attention, in: Proc. AAAI Conf Artif. Intell., 2020, pp. 11450–11457.

[21] Hao Lu, Yutong Dai, Chunhua Shen, Songcen Xu, Index networks, IEEE Trans. Pattern Anal. Mach. Intell. 44 (1) (2022) 242–255.

[22] Anat Levin, Dani Lischinski, Yair Weiss, A closed-form solution to natural image matting, IEEE Trans. Pattern Anal. Mach. Intell. 30 (2) (2008) 228–242.

[23] Jian Sun, Jiaya Jia, Chi-Keung Tang, Heung-Yeung Shum, Poisson matting, ACM Trans. Graph. 23 (3) (2004) 315–321.

[24] Philip Greggory Lee, Ying Wu, Nonlocal matting, in: Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2011, pp. 2193–2200.

[25] Qifeng Chen, Dingzeyu Li, Chi-Keung Tang, KNN matting, IEEE Trans. Pattern Anal. Mach. Intell. 35 (9) (2013) 2175–2188.

[26] Ehsan Shahrian Varnousfaderani, Deepu Rajan, Weighted color and texture sample selection for image matting, IEEE Trans. Image Process. 22 (11) (2013) 4260–4270.

[27] Levent Karacan, Aykut Erdem, Erkut Erdem, Alpha matting with KL-divergence-based sparse sampling, IEEE Trans. Image Process. 26 (9) (2017) 4523–4536.

[28] Jubin Johnson, Ehsan Shahrian Varnousfaderani, Hisham Cholakkal, Deepu Rajan, Sparse coding for alpha matting, IEEE Trans. Image Process. 25 (7) (2016) 3032–3043.

[29] Kaiming He, Christoph Rhemann, Carsten Rother, Xiaoou Tang, Jian Sun, A global sampling method for alpha matting, in: Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2011, pp. 2049–2056.

[30] Kalyanmoy Deb, Samir Agrawal, Amrit Pratap, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2) (2002) 182–197.

[31] Marco Forte, François Pitié, F, b, alpha matting, 2020, CoRR, arXiv:2003.07711.

[32] Yihan Hu, Yiheng Lin, Wei Wang, Yao Zhao, Yunchao Wei, Humphrey Shi, Diffusion for natural image matting, in: ECCV (57), in: Lecture Notes in Computer Science, vol. 15115, Springer, 2024, pp. 181–199.

[33] Yagiz Aksoy, Tunç Ozan Aydin, Marc Pollefeys, Designing effective inter-pixel information flow for natural image matting, in: CVPR, IEEE Computer Society, 2017, pp. 228–236.

[34] Yung-Yu Chuang, Brian Curless, David Salesin, Richard Szeliski, A Bayesian approach to digital matting, in: CVPR (2), IEEE Computer Society, 2001, pp. 264–271.