

吴瑞瑾, 王岳继, 罗丽, 等. 基于 Token 复杂度感知分配的高分辨率图像抠图[J]. 智能计算机与应用, 2025, 15(12): 117-122. DOI: 10.20169/j.issn.2095-2163.25101601

基于 Token 复杂度感知分配的高分辨率图像抠图

吴瑞瑾^{1,2}, 王岳继^{1,2}, 罗丽^{1,2}, 王才丽^{1,2}, 梁椅辉^{1,2}

(1 贵州民族大学 数据科学与信息工程学院, 贵阳 550025;

2 贵州民族大学 贵州省模式识别与智能系统重点实验室, 贵阳 550025)

摘要: 近年来, 基于高分辨率图像的抠图方法在性能上取得了显著进展, 为高分辨率图像抠图模型在边缘设备上的高效部署提供了新的可能性。然而, 由于 Token 分配机制在复杂场景下缺乏对区域特征复杂度的有效建模, 现有抠图方法在高分辨率图像上的性能仍存在明显下降。针对这一问题, 本文提出了一种基于 Token 复杂度感知分配的高分辨率图像抠图模型。该模型在原有分配机制的基础上引入了复杂度感知模块, 可自适应地预测每个 Token 的特征复杂性水平, 并与原始分配机制进行融合, 从而实现边缘、半透明等高复杂度区域的精细化处理。通过该设计, 模型能够在保证计算效率的同时, 显著提升抠图结果的精度。实验结果表明, 在高分辨率数据集 UHR-395 上, 本方法在 MSE、SAD、Grad 和 Conn 4 项指标上分别提升了 13.20%、7.08%、0.70% 和 9.30%, 相较于现有主流方法取得了更优的综合性能。

关键词: 高分辨率图像抠图; 特征选择; 复杂度感知; 融合策略; 损失函数

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2025)12-0117-06

High-resolution image matting based on Token complexity aware allocation

WU Ruijin^{1,2}, WANG Yueji^{1,2}, LUO Li^{1,2}, WANG Caili^{1,2}, LIANG Yihui^{1,2}

(1 School of Data Sciences and Information Engineering, Guizhou Minzu University, Guiyang 550025, China;

2 Guizhou Key Laboratory of Pattern Recognition and Intelligent System, Guizhou Minzu University, Guiyang 550025, China)

Abstract: In recent years, significant progress has been made in the performance of high-resolution image matting methods, providing new possibilities for the efficient deployment of high-resolution image matting models on edge devices. However, due to the lack of effective modeling of regional feature complexity in complex scenarios by token allocation mechanisms, the performance of existing matting methods on high-resolution images still shows a significant decline. This article proposes a high-resolution image cutout model based on Token complexity aware allocation to address this issue. This model introduces a complexity perception module on the basis of the original allocation mechanism, which can adaptively predict the feature complexity level of each Token and integrate it with the original allocation mechanism, thereby achieving fine-grained processing of high complexity areas such as edges and semi transparency. Through this design, the model can significantly improve the accuracy of the matting results while ensuring computational efficiency. The experimental results show that on the high-resolution dataset UHR-395, this method improved the MSE, SAD, Grad, and Conn metrics by 13.20%, 7.08%, 0.70%, and 9.30%, respectively, achieving better overall performance compared to existing mainstream methods.

Key words: high-resolution image matting; feature selection; complexity perception; integration strategy; loss function

0 引言

自然图像抠图作为计算机视觉的重要组成部分, 其目的是获取图像中每个像素的透明度, 以此从

复杂的背景图像中精确地分离前景对象。该技术在众多应用中至关重要, 例如视觉内容制作^[1]、图像编辑^[2,3]、影视后期处理^[4]、虚拟数字人^[5]、自动驾驶^[6]。从数学的角度出发, 给定输入图像 I 可以表

基金项目: 国家自然科学基金(62002053); 贵州民族大学基金科研项目(GZMUZK[2023]YB14); 贵州省模式识别与智能系统重点实验室开放课题(GZMUKL[2022]KF01); 贵州民族大学自然科学基金资助项目(GZMUZK[2023]YB04); 广东省基础与应用基础研究基金(2023A1515010066); 贵州省教育厅青年科技人才成长项目(QJJ[2024]273)。

作者简介: 吴瑞瑾(2000—), 男, 硕士研究生, 主要研究方向: 图像抠图; 王岳继(2000—), 男, 硕士研究生, 主要研究方向: 工业检测; 罗丽(2002—), 女, 硕士研究生, 主要研究方向: 图像抠图; 王才丽(1999—), 女, 硕士研究生, 主要研究方向: 图像抠图。

通信作者: 梁椅辉(1990—), 男, 博士, 副教授, 硕士生导师, 主要研究方向: 图像抠图, 图像处理及分析。Email: yihuiliang@outlook.com。

收稿日期: 2025-10-16

示为图像中的前景色和背景色凸线性组合,即:

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i, \alpha_i \in [0, 1] \quad (1)$$

其中 i 表示输入图像中的像素索引, α_i 、 F_i 和 B_i 分别指像素处的不透明度、前景颜色值和背景颜色值。

现有的抠图技术主要应用于低分辨率图像上,具体可分为基于传统的抠图算法以及基于深度学习的抠图算法等方法来分离前景和背景。

基于传统的抠图算法具体可划分为基于采样的方法以及基于传播的抠图方法。Chuang 等^[7]提出了一种基于贝叶斯推断的图像抠图方法 (A Bayesian Approach to Digital Matting), 将抠图建模为统计推断任务, 通过从已知前景与背景中对像素对进行采样, 在此基础上利用贝叶斯推断求解最优的透明度遮罩, 从而获得较为精确的前景分离结果。Feng 等^[8]将聚类采样与稀疏编码引入到图像抠图中。通过在传统采样的基础上引入聚类机制, 针对未知区域的像素, 从不同簇中选择具有代表性的样本对, 通过稀疏编码的方式重建未知像素颜色。从而预测透明度遮罩值。He 等^[9]进一步提出了全局采样抠图算法, 通过全局优化策略求解透明度遮罩。Levin 等^[10]提出了经典的 Closed-Form Matting 方法, 该方法假设图像在局部窗口内的前景与背景颜色可视为线性可分, 将透明度遮罩的求解建模为二次优化问题。Chen 等^[11]提出了 KNN Matting 算法, 该方法将颜色与空间的特征进行结合, 通过构建 K 邻近图传播透明度遮罩的信息。利用全局的样本邻域关系实现透明度遮罩值的传播。Zheng 和 Kambhammettu^[12]提出的 Learning-Based Matting 将抠图问题转化为一个监督学习任务。利用局部特征训练回归模型预测透明度遮罩的值, 并结合传播机制在未知区域中平滑过渡。

随着卷积神经网络 (CNN) 以及 Transformer 的

出现, 基于深度学习的抠图方法凭借其出色的特征提取能力以及全局建模的特点, 为自然图像抠图带来了显著的突破。Xu 等^[13]首次提出了端到端的深度抠图框架——Deep Image Matting (DIM)。该方法为基于三分图的抠图方法, 以经典的编码器解码器架构为基础, 输入原始图像与对应的三分图, 通过多尺度卷积特征提取前景、背景及未知区域的上下文信息。利用跳跃链接融合编码器解码器的信息, 以此恢复细节结构, 从而预测透明度遮罩。Ke 等^[14]提出了一种不依赖于三分图的抠图方法, 该方法主要由轻量化网络结构组合而成, 将抠图任务分解为 3 个子任务: 语义分支、细节分支和融合分支。语义分支主要提取全局语义信息, 细节分支捕获边缘结构, 而融合分支在多尺度特征空间中实现透明度遮罩的精细预测。Yao 等^[15]提出了基于预训练 Vision Transformer (ViT) 的图像抠图方法。凭借 ViT 强大的全局建模能力, 将 ViT 的预训练模型在抠图上进行微调, 通过混合注意力机制, 实现对局部到全局的建模。引入细节捕获模块对细节信息进行补充, 以此得到更加鲁棒的透明度遮罩。

高分辨率图像为抠图任务提供了更加丰富的细节信息以及视觉效果, Ye 等^[16]提出的 HRIMF-AMR^[16]将任务分解为低分辨率抠图与高分辨率 alpha matte 细化两个阶段, 为现有模型无法对高分辨率图像进行处理提供了思路。Lin 等提出了 MEMatte^[17]通过将 Token 的重要性分别处理, 以此减低显存占用和计算复杂度。

综上所述, 尽管现有的高分辨率图像抠图方法已提出了多种可行的改进思路, 但在复杂场景下仍存在一定的局限性。由于现有的 Token 分配机制缺乏对区域特征复杂度的有效建模, 模型在边缘或透明区域的抠图精度仍会出现明显下降, 如图 1 所示。

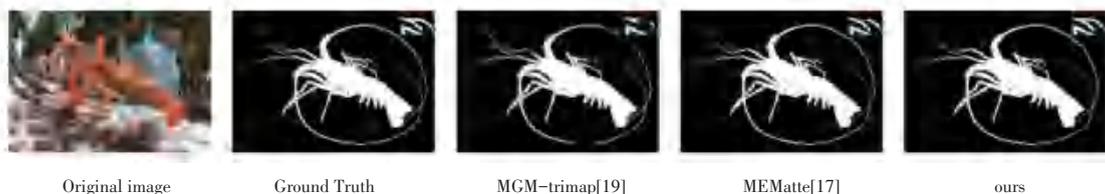


图 1 高分辨率图像抠图结果视觉比较

Fig. 1 Visual comparison of high-resolution image matting results

因此, 如何在高分辨率图像抠图任务中实现更加优化的 Token 分配策略, 仍需深入研究。针对上述问题, 本文提出了一种基于 Token 复杂度感知分配的高分辨率图像抠图模型。该模型在 MEMatte^[17]的基础上引入了复杂性感知模块, 以增

强模型对高复杂度区域的自适应处理能力。传统 MEMatte 的 Token 分配机制无法有效判断当前 Token 的预测可靠性, 导致在高分辨率图像中边缘或透明区域^[18]的抠图精度下降, 而复杂性感知模块能够自适应地评估各 Token 的预测稳定性与潜在复

杂性,从而识别出易出错或预测不稳定的 Token,并将其送入 ViT 模块进行重点细化处理。

1 Token 复杂性感知分配抠图模型

本文方法以 MEMatte^[17] 为基线模型,在原有分配机制的基础上引入复杂性感知策略 (Complexity-Aware Router—CAR),用以预测特征置信度,并通过

置信度动态调整分配策略,最后将原有分配机制的输出与 CAR 进行融合,以此增加对 Token 原有的限制,实现对复杂区域的自适应精修。此外,本文还引入复杂性感知的加权的损失,通过 CAR 输出的复杂性图 (Complexity Map),让模型在训练阶段能更加地关注预测复杂的区域,如图 2 所示。

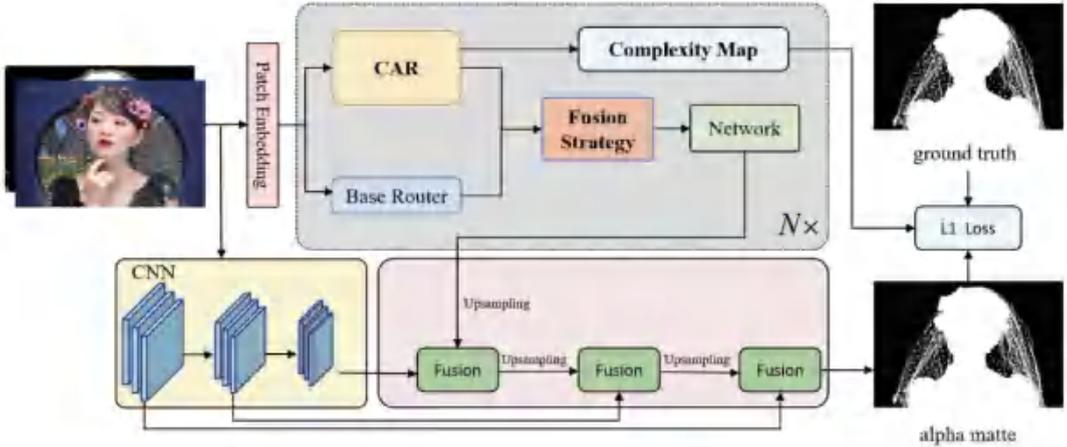


图 2 模型结构图

Fig. 2 Diagram of model structure

1.1 复杂性感知模块

为增强 MEMatte 的动态分配机制,引入复杂性感知分配策略 (Complexity-Aware Router—CAR),其结构如图 3 所示。

缘细节提取不够完整等情况。因此,CAR 能根据每个区域的复杂性、该区域是否预测不自信、边缘、细节、半透明区域是否预测准确做出合理的估计,为 Token 分配决策提供额外的指导信息。

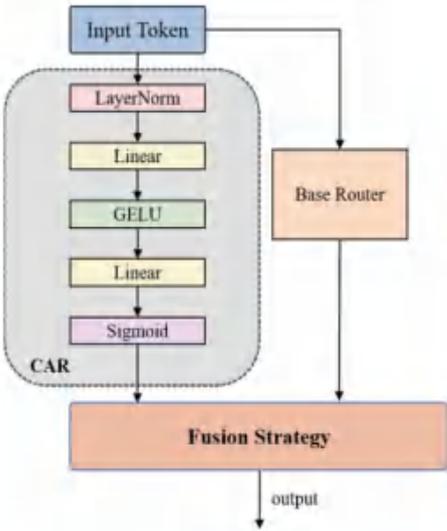


图 3 CAR 结构图

Fig. 3 CAR Structure Diagram

Complexity-Aware Router—CAR 由一组轻量 MLP 组成,主要包括 LayerNorm、Linear、GELU、Linear、Sigmoid,评估输出每个 Token 的复杂性值 $u \in [0,1]$,当 $u = 0$ 时,表示当前区域的预测较为自信,当 $u = 1$ 时,表示当前区域可能存在预测模糊,边

1.2 融合策略

本文的融合策略 (Fusion Strategy) 将两个不同的输出进行加权融合,帮助分配策略更加精准地选择每个 Token 是否需要进入 ViT 分支。具体地,在原有的 MEMatte 网络中,分配机制仅根据特征相似性进行二值或连续化判断,本文将其定义为 $policy_{raw}$ 。融合策略可以表示为:

$$policy = policy_{raw} \times (1 + u - 0.5) \quad (2)$$

当模型对某一区域预测复杂性较高时 ($u \rightarrow 1$), $policy$ 的激活强度随之增强,从而提升该区域进入 ViT 分支的概率;反之,对于置信度较低区域 ($u \rightarrow 0$), $policy$ 值被抑制,避免无效计算。该模型实现了按需精修的自适应推理过程,显著提升了动态分配决策的合理性与效率,增加了模型的稳定。

1.3 复杂性加权损失函数

为了在训练阶段进一步强化复杂性区域的学习能力,更加关注预测复杂的区域,本文在原始未知区域损失的基础上引入了复杂性权重项。具体地,给出 $u \in [0,1]$ 为 CAR 预测的复杂性,损失函数可以

定义为:

$$L_{\text{unknown}} = \frac{1}{|\Omega|} \sum_{(i,j) \in \Omega} (1 + u_{i,j}) |\hat{\alpha}_{i,j} - \alpha_{i,j}| \quad (3)$$

其中, Ω 表示未知区域, 通过损失约束, 模型在复杂性区域, 比如边缘或者半透明区域会自动根据梯度反馈, 从而促使 alpha 细化。结合之前提到的损失, 总损失函数为:

$$L_{\text{total}} = L_{\text{unknown}} + L_{\text{known}} + L_{\text{lap}} + L_{\text{gp}} \quad (4)$$

2 实验结果与分析

本文在 Composition-1k 数据集^[13]进行训练, 该数据集包含 43 100 个图像, 每个图像都由 431 个不同的前景以及对应数量的背景组合而成, 背景在 Microsoft COCO^[19]通过随机选择得到的。本文在 UHR-395 高分辨率数据集上进行测试, 主要包含 1 000 张高分辨率图像, 其平均分辨率达到 $4\ 872 \times 6\ 017$ 。本文实验中的训练在 NVIDIA A100 GPU (40 GB memory) 进行训练, 编程语言为 Python, 其版本为 3.8.8, 并以 torch-2.0.0 作为深度学习框架。推理过程因测试图像分辨率过大, 在 NVIDIA A800 GPU (80 GB memory) 以及配置 A800 的服务器的 CPU 内存 (RAM) 为 503.51 GB 进行推理。本文的评估主要采用 MSE (均方误差)、SAD (绝对误差)、Grad (梯度误差) 和 Conn (连通性误差)^[20]对模型的性能进行评估。

为了验证所提出各个模块对模型性能的贡献, 本文在基准模型上进行了系统的消融实验。在 MEMatte 的基线模型上逐步添加了本文的模块, 以此验证本文方法所提出的模型的有效性。实验结果见表 1。具体包括以下:

(1) Baseline 为本文的基线模型 MEMatte;

(2) CAR 为本文复杂性感知分配策略, 因为 CAR 和融合策略是一起工作的, 所以将其合并到一起做消融实验;

(3) Loss 为复杂性加权损失函数, 用于监督。

表 1 消融实验结果

Table 1 Ablation experiment results

Methods	MSE	SAD	Grad	Conn
Baseline	11.66	696.27	225.44	478.58
Baseline + CAR	11.21	672.79	226.94	452.43
Baseline + CAR + Loss	10.12	646.93	223.86	434.03

由表 1 可以看出, 仅采用原始的 Token 分配策略时, 模型在高分辨率图像上的抠图精度相对较低, 尤其在边缘区域和细节过渡处容易出现伪影或断层

现象。当引入本文提出的 CAR 融合策略后, 模型能够在物体边缘和细节区域获得更加精确和平滑的预测结果。进一步地, 在融合阶段加入了复杂性的损失监督, 使模型在训练过程中能够更好地平衡全局一致性与局部精细化, 显著提升了整体性能。综合来看, 完整模型在各项评价指标上均达到了最优效果, 充分验证了所提出策略的有效性与鲁棒性。

本文的方法与 DIM^[13], GCA-Matting^[21], A2U^[22], MGM-trimap^[23], TIMI-Net^[24], Matteformer^[25], ELGT-Matting^[26], HRAMR-Matting^[16], ViTMatte^[15], MEMatte^[17] 进行比较, 结果见表 2。其中加粗的为效果最好的指标。

表 2 UHR-395 测试集定量结果

Table 2 Quantitative results on UHR-395 test set

Methods	MSE	SAD	Grad	Conn
DIM*	161.41	2 970.18	2 202.36	1 456.57
GCA-Matting†	-	-	-	-
A2U*	58.26	1 626.73	791.35	954.14
MGM-trimap*	25.38	997.39	432.81	632.41
TIMI-Net†	-	-	-	-
Matteformer*	46.39	1 453.87	537.54	765.77
ELGT-Matting*	39.18	1 195.57	705.43	733.77
HRAMR-Matting	14.98	692.30	384.47	483.10
ViTMatte†	-	-	-	-
MEMatte-S	11.66	696.27	225.44	478.58
Ours-S	10.12	646.93	223.86	434.03
MEMatte-B	13.17	692.34	207.81	474.64
Ours-B	12.59	694.64	209.88	481.99

注: * 表示在 A800 上无法以 GPU 进行推理, 会出现显存溢出, 因此在配置 A800 的服务器上以 CPU 进行推理, † 表示无法在 A800 上以 GPU 或者服务器上以 CPU 进行推理, 都会出现显存溢出, GCA-Matting^[21], TIMI-Net^[24], ViTMatte^[15] 的实验耗尽了服务器上 CPU 的 503.51 GB 内存。

由表 2 可知, 在 UHR-395 高分辨率数据集上, 本文的方法与其他的方法相比, 在 4 个评估指标上, 本文的模型都达到了最优的结果, 表明本文所提出的方法针对高分辨率图像抠图存在一定的优势。验证了本文的复杂性感知模块针对高分辨率图像的细节边缘处理等方面表现突出。相比起经典的抠图算法 DIM 模型, 本文的 MSE、SAD、Grad、Conn 4 个指标都有显著提升, 分别提升 93.73%、78.21%、89.83%、70.20%, 与 2025 年最近提出的两篇高分辨率图像抠图 HRAMR-Matting 和基线模型 MEMatte 相比, 本文的方法在 4 个指标上均取得了显著的优势。在

MEMatte 上, 本文的性能分别提升了 13.20%、7.08%、0.70%、9.30%。

为了更加直观的比较本文的方法和其他方法的抠图结果, 本文做了可视化分析, 如图 4 所示。在图 4 中, 本文分别与 (b) DIM、(c) MGM-trimap、(d) HRAMR-Matting、(e) MEMatte 做了可视化比较, 其中 (g) 为真实的 alpha 遮罩。从图 4 的可视化中可以看出, 本文的方法在细节毛发以及透明区域相比起其他的方法更接近于真实的 alpha 遮罩。例如,

在透明物体上, DIM、MGM-trimap、HRAMR-Matting 都与真实的 alpha 差距过大, 出现大量伪影, 与基线模型 MEMatte 相比, 本文的方法可视化结果很接近, 但是在细微的部分, 本文的效果相比起基线是有优势的。

综上所述, 本文所提出的复杂性感知模块对细节的处理以及边缘的优化明显优于现有的方法, 进一步验证多提出的模型在边缘或透明区域具有一定的优势。

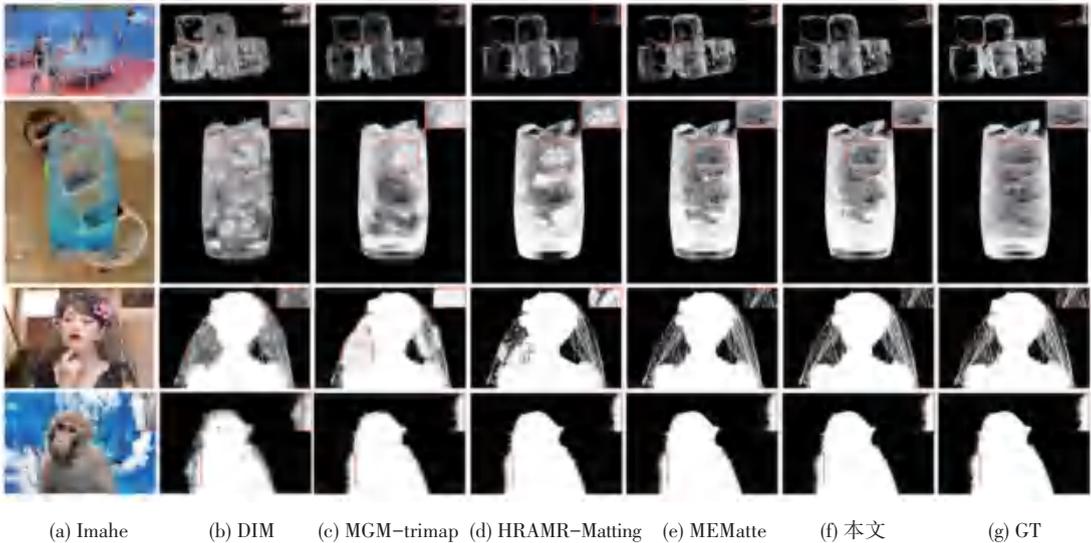


图 4 UHR-395 高分辨率数据集可视化结果

Fig. 4 Visualization results of UHR-395 high resolution dataset

3 结束语

本文针对现有模型在高分辨率图像抠图时因 Token 分配机制缺乏对区域特征复杂度的有效建模, 导致预测性能下降的问题, 提出了 Token 复杂性感知的高分辨率图像抠图模型。模型主要包括 3 个部分, 即复杂性感知模型、融合策略、复杂性加权损失。首先利用复杂性感知模型对 Token 进行预测得到复杂性值, 之后利用原有分配机制的输出与本文的模块进行加权融合, 以此解决原有分配机制仅依赖局部特征分布, 对模型预测置信度缺乏感知, 容易在边缘与半透明结构处出现错误分配, 造成“过修”或“欠修”的现象。最后在训练阶段加入复杂性加权损失, 让模型加强对复杂性区域的学习预测。为高分辨率图像抠图提供了新的解决思路, 未来将考虑如何在低计算资源的前提下, 取得高精度的预测结果。

参考文献

[1] SHEN X, TAO X, GAO H, et al. Deep automatic portrait matting [C]// Proceedings of European Conference on Computer Vision.

Cham; Springer, 2016: 92-107.
 [2] XU Y, YAO X, LIU B, et al. Text-guided portrait image matting [J]. IEEE Transactions on Artificial Intelligence, 2024, 5(8): 4149-4162. DOI: 10.1109/TAI.2024.3363120
 [3] LIANG Y, FU Q, ZOU K, et al. Enhancing transparent object matting using predicted definite foreground and background [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2025, 35(1): 62-74. DOI: 10.1109/TCSVT.2024.3452512
 [4] WANG J, COHEN M F. Image and video matting: A survey [J]. Foundations and Trends in Computer Graphics and Vision, 2007, 3(2): 97-175.
 [5] CAI H, XUE F, XU L, et al. Trans-matting: Enhancing transparent objects matting with transformers [C]// Proceedings of European Conference on Computer Vision. Cham; Springer, 2022: 253-269.
 [6] LIANG Y, HUANG H, CAI Z, et al. Deep infrared pedestrian classification based on automatic image matting [J]. Applied Soft Computing, 2019, 77: 484-496.
 [7] CHUANG Y Y, CURLESS B, SALESIN D H, et al. A Bayesian approach to digital matting [C]// Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2001, 2: 264-271.
 [8] FENG X, LIANG X, ZHANG Z. A cluster sampling method for image matting via sparsecoding [C]// Proceedings of European Conference on Computer Vision. Cham; Springer, 2016: 204-219.

- [9] HE K, RHEMANN C, ROTHER C, et al. A global sampling method for alpha matting [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2011: 2049-2056.
- [10] LEVIN A, LISCHINSKI D, WEISS Y. A closed-form solution to natural image matting [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 228-242.
- [11] CHEN Q, LI D, TANG C K. KNN matting [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(9): 2175-2188.
- [12] ZHENG Y, KAMBHAMMETTU C. Learning-based digital matting [C]// Proceedings of IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2009: 889-896.
- [13] XU N, PRICE B, COHEN S, et al. Deep image matting [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 2970-2979.
- [14] KE Z, SUN J, LI K, et al. MODNet: Real-time trimap-free portrait matting via objective decomposition [C]//Proceedings of the AAAI Conference on Artificial Intelligence. AAAI, 2022: 1140-1147.
- [15] YAO J, WANG X, YANG S, et al. ViTMatte: Boosting image matting with pre-trained plain vision transformers [J]. Information Fusion, 2024, 103: 102091.
- [16] YE X, LIANG Y, TAN M, et al. High-resolution natural image matting by refining low-resolution alpha mattes [J]. IEEE Transactions on Image Processing, 2025, 34: 3323-3335.
- [17] LIN Y, HU Y, ZHANG C, et al. Memory-efficient matting with adaptive token routing [C]//Proceedings of the AAAI Conference on Artificial Intelligence. AAAI, 2025: 5298-5306.
- [18] 张岚岚. 面向自然图像透明区域的特征感知抠图 [J]. 智能计算机与应用, 2024, 14(12): 139-144. DOI: 10.20169/j.issn.2095-2163.24071102.
- [19] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context [C]//Proceedings of European Conference on Computer Vision. Cham: Springer, 2014: 740-755.
- [20] PORTER T, DUFF T. Compositing digital images [C]// Proceedings of the 11th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM, 1984: 253-259.
- [21] LI Y, LU H. Natural image matting via guided contextual attention [C]//Proceedings of the AAAI Conference on Artificial Intelligence. AAAI, 2020: 11450-11457.
- [22] DAI Y, LU H, SHEN C. Learning affinity-aware upsampling for deep image matting [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 6841-6850.
- [23] YU Q, ZHANG J, ZHANG H, et al. Mask guided matting via progressive refinement network [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 1154-1163.
- [24] LIU Y, XIE J, SHI X, et al. Tripartite information mining and integration for image matting [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2021: 7555-7564.
- [25] PARK G T, SON S J, YOO J Y, et al. MatteFormer: Transformer-based image matting via prior-tokens [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2022: 11696-11706.
- [26] HU L, KONG Y, LI J, et al. Effective local-global transformer for natural image matting [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(8): 3888-3898.