Context-Preserving Dermoscopic Editing: Mask-Guided Lesion-Aware Diffusion for Attribute Modification

Tao Sun[1], Yun Jiang[1], Yarong Jin[1], Huanting Guo[3], Zequn Zhang[2,4]

[1]NWNU, [2]USTC, [3]BIT, [4]Xiong'an Anying Technology

WACV 2026 TUCSON, AZ 3/6 – 3/10

Introduction

Background

- Dermoscopic diagnosis relies on fine-grained lesion cues as well as global structural information, including the lesion boundary and peri-lesional context.
- Key pathological attributes are often long-tailed and under-sampled, which limits model learning and interpretability and motivates controllable editing and counterfactual generation for data augmentation and diagnostic analysis.

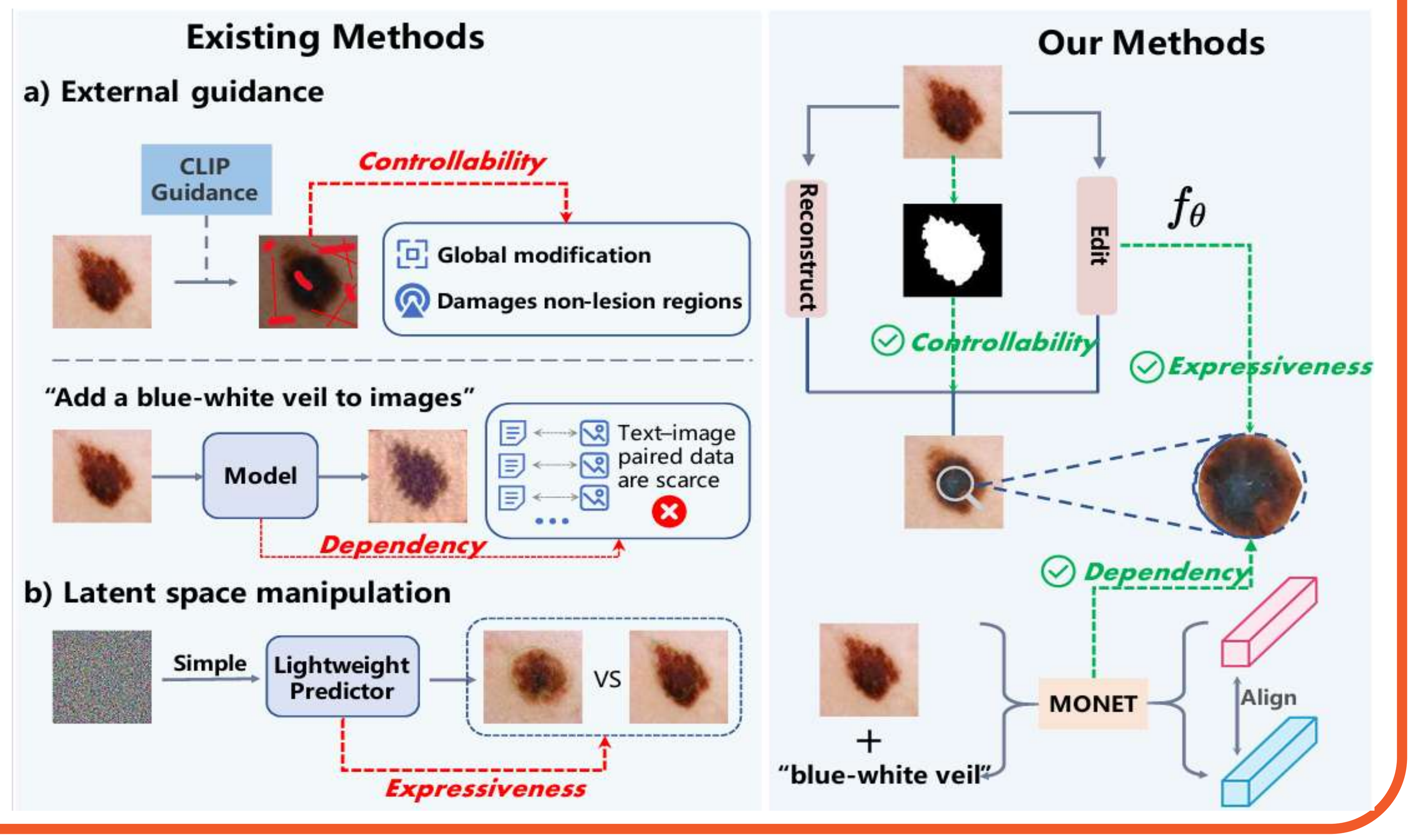
Motivation

- Existing diffusion editing is either globally disruptive or semantically weak, motivating lesion-aware editing that preserves background context.
- Controllable, clinically meaningful synthesis can alleviate long-tailed imbalance and improve diagnostic robustness on rare signs.
- High-quality counterfactuals reveal model-relevant pathological cues, boosting interpretability and clinician trust.

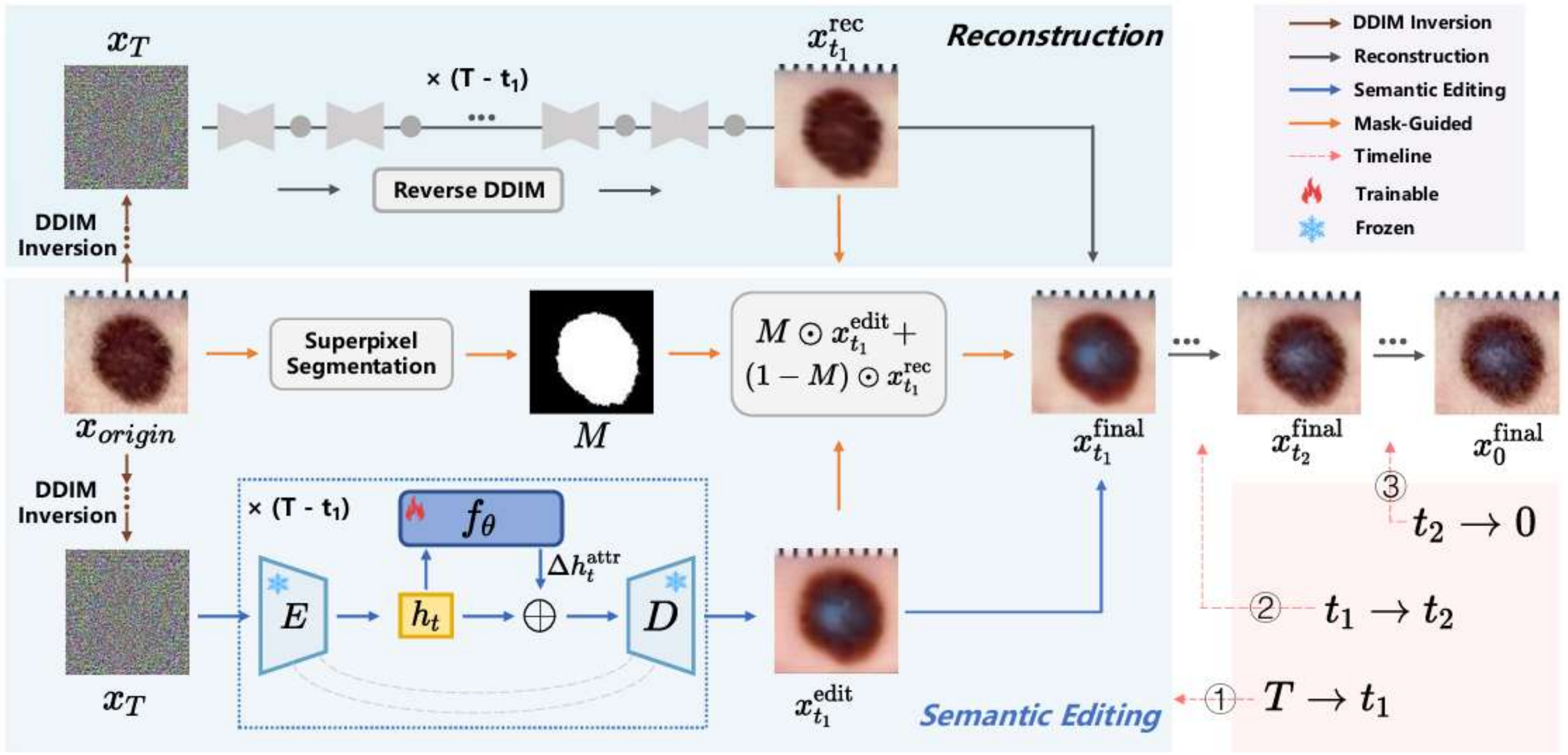
Contribution

- **CPDE**: Dual-branch diffusion for lesion editing with background preservation.
- **SCT + lesion-aware loss**: Expressive residuals and mask-guided semantic control.
- **Evaluation**: Better fidelity/semantics on ISIC 2017/2018, with lesion preservation and diagnostic consistency.

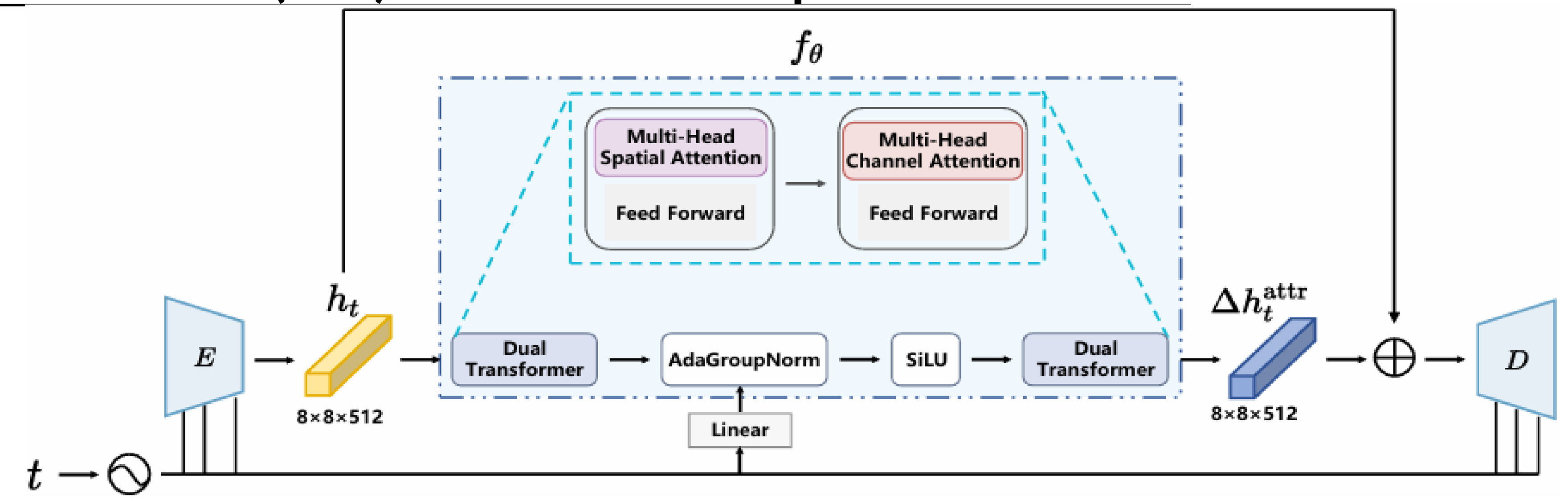
Existing Methods vs Our Method:



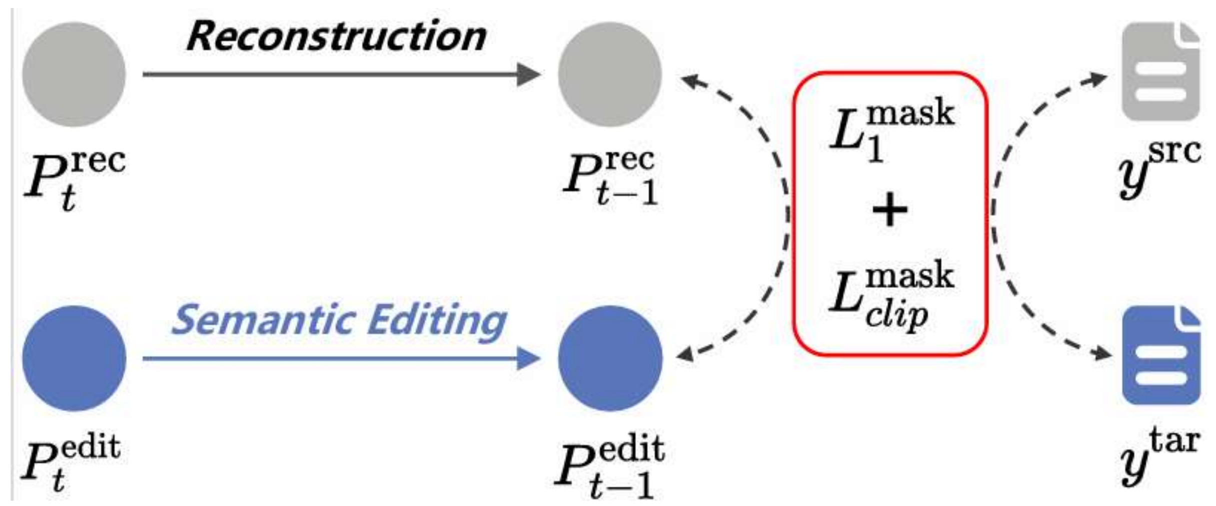
CPDE Framework



(a) fθ network (SCT) semantic residual prediction module



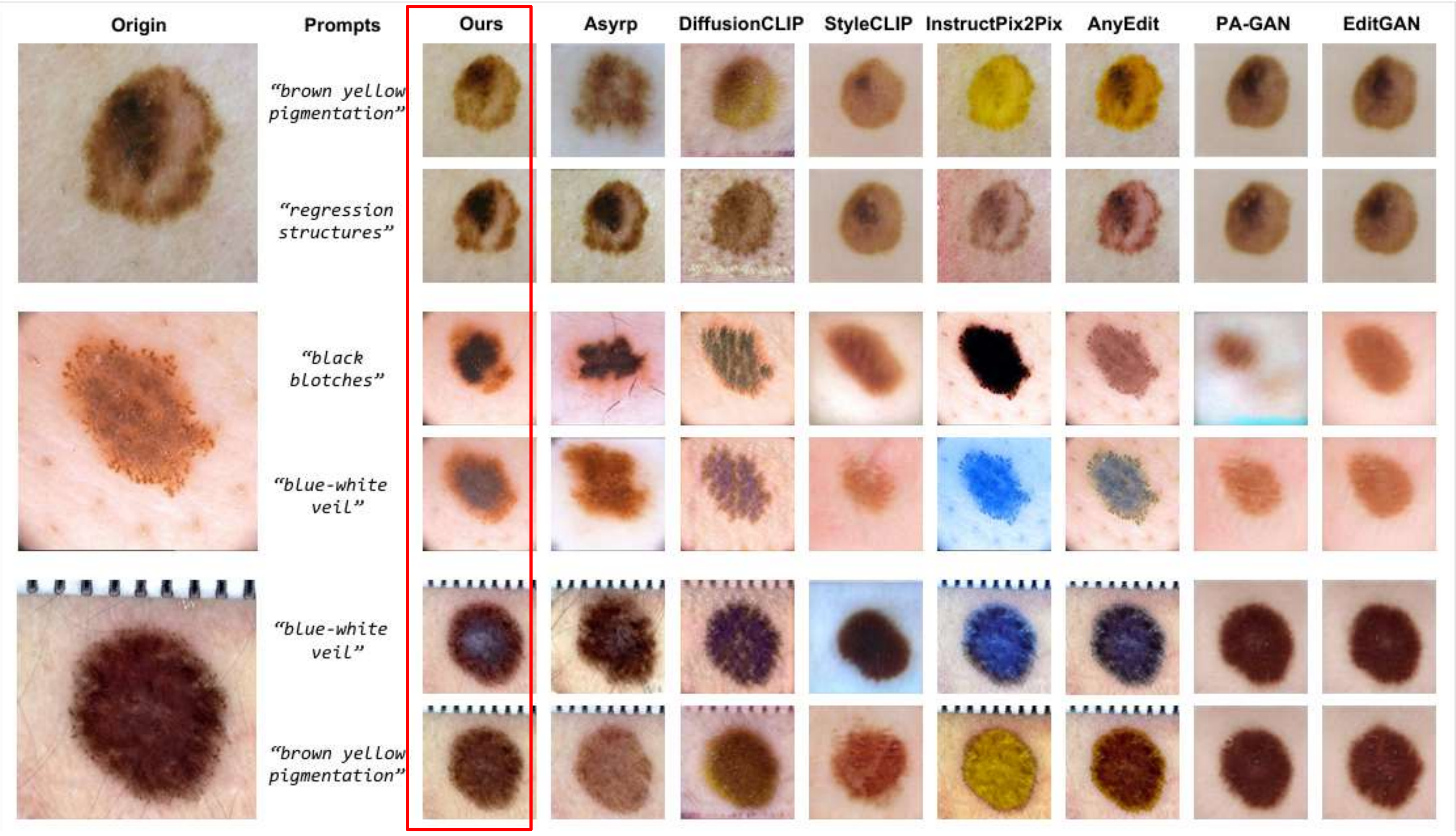
(b) Mask-guided dual-branch training loss design diagram



Training Objective: $$L^{(t)} = \lambda_{L1} L_1^{mask} + \lambda_{clip} L_{clip}^{mask}$$

Experiments

Qualitative Comparison



Quantitative Comparison

Method	FID ↓	S_{dir} ↑	NS-LPIPS ↓
Asyrp	0.482	0.452	0.151
DiffusionCLIP	1.761	0.413	0.247
StyleCLIP	2.132	0.167	0.329
InstructPix2Pix	1.069	0.385	0.147
AnyEdit	0.731	0.432	0.106
PA-GAN	2.329	0.117	0.521
EditGAN	2.377	0.082	0.483
Ours (CPDE)	**0.274**	**0.486**	**0.012**

Ablation study on the SCT and Masking components of CPDE

Method	FID ↓	S_{dir} ↑	NS-LPIPS ↓
Baseline (Asyrp)	0.482	0.452	0.151
w/ Masking	0.372	0.461	0.022
w/ SCT	0.391	0.477	0.163
Ours (CPDE)	**0.274**	**0.486**	**0.012**

- Segmentation consistency visualization: green contours represent the ground-truth lesion boundaries in the original image, and red contours correspond to the segmentation results on the edited image.