

Loomis Painter

Reconstructing the Painting Process



Markus Pobitzer¹, Chang Liu¹, Chenyi Zhuang¹, Teng Long¹, Bin Ren^{1,2}, Nicu Sebe¹
¹University of Trento, ²University of Pisa

Can Video Generation Models Create Painting Tutorials?

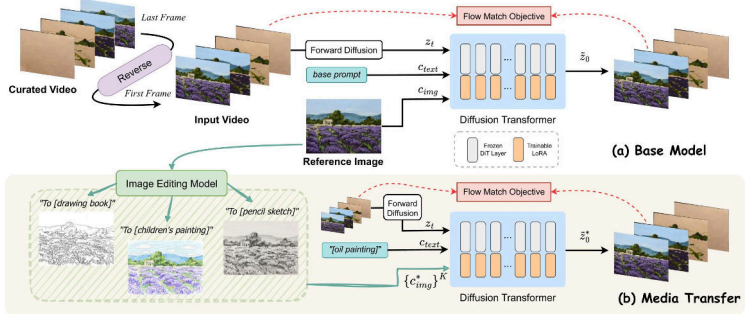
Existing painting tutorials lack interactivity & personalization

We solve this by:

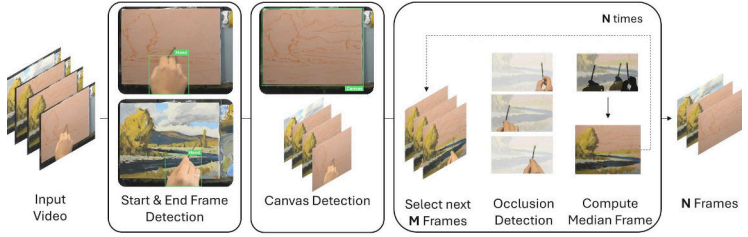
- Collecting a real world dataset: 700+ processed videos
- Fine-tuning a video diffusion model
- Enabling Art Media Transfer

We model the painting task as an image to video task and reverse the painting order. Start from a finished painting to get to the blank canvas.

Pipeline



Dataset Curation



Quantitative Results

Method	Fidelity				Temporal Consistency				Params
	FID ↓	LPIPS ↓	CLIP ↑	DINOv2 ↑	DDC ↓	PDP-LPIPS ↓	PDP-CLIP ↓	PDP-DINOv2 ↓	
Inverse Painting [9]	326.15	0.61	0.66	0.21	48.77	0.31	0.13	0.31	900M
ProcessPainter [44]	282.90	0.53	0.76	0.50	17.79	0.17	0.07	0.14	1.3B
PaintsUndo [48]	236.52	0.55	0.77	0.56	15.05	0.16	0.06	0.12	900M
PaintsUndo [†] [48]	196.66	0.43	0.83	0.66	10.02	0.09	0.04	0.09	900M
Ours (VACE)	183.61	0.42	0.83	0.70	9.18	0.09	0.03	0.08	1.3B
Ours (WAN 2.2)	152.42	0.38	0.85	0.75	7.97	0.08	0.03	0.06	5B
Ours (WAN 2.1)	151.04	0.38	0.86	0.76	8.39	0.08	0.03	0.07	14B

Human Evaluation

Method A	Method B	Win (A) (%)	Tie (%)	Win (B) (%)	Total
Ours	PaintsUndo	95.6	1.3	3.2	158
Ours	Inverse Painting	78.7	9.8	11.6	164
Ours	ProcessPainter	96.2	0.0	3.8	160

Qualitative: Loomis Portrait Drawing



Qualitative: Reconstructing the Painting Process



Qualitative: Art Media Transfer

