

MetaStyle: Three-Way Trade-Off Among Speed, Flexibility and Quality in Neural Style Transfer

Chi Zhang, Yixin Zhu, Song-Chun Zhu

International Center for AI and Robot Autonomy

{chizhang, yzhu, sczhu}@cara.ai



MetaStyle

- Training

minimize θ

$$\mathbb{E}_{c,s}[\ell(I_c, I_s, M(I_c; w_{s,T}))]$$

subject to

$$w_{s,0} = \theta$$

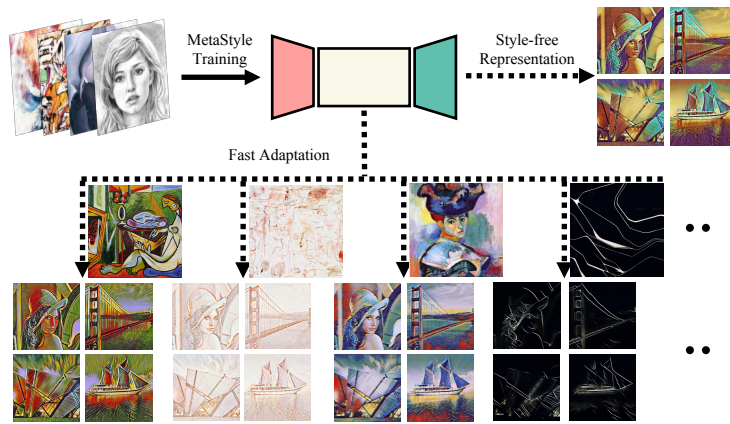
$$w_{s,t} = w_{s,t-1} - \delta \nabla \mathbb{E}_c[\ell(I_c, I_s, M(I_c; w_{s,t-1}))]$$

- Adaptation

minimize w

$$\mathbb{E}_c[\ell(I_c, I_s, M(I_c; w))]$$

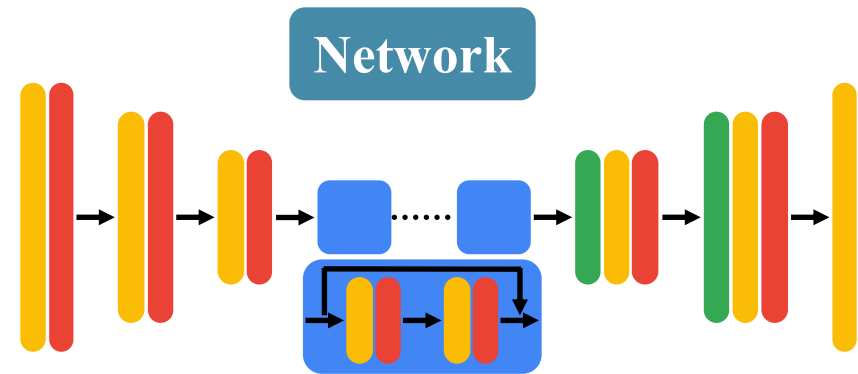
- Framework



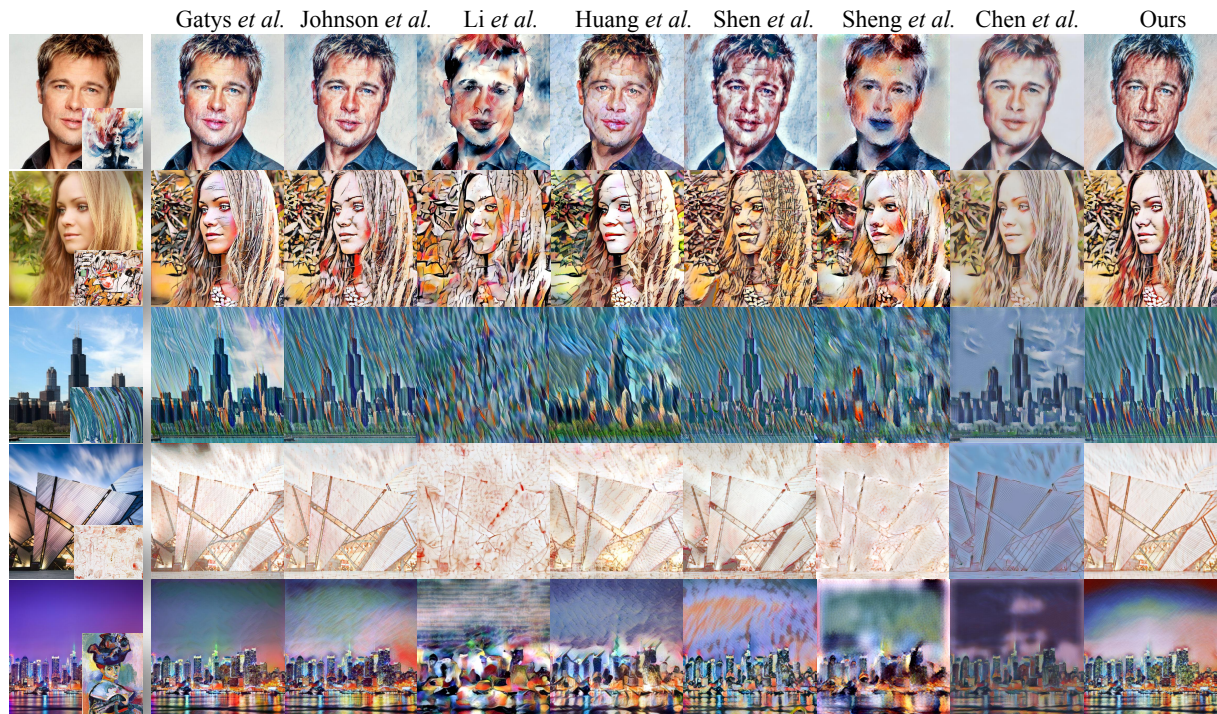
Motivation

Method	Speed	Flexibility	Quality	Drawback
Optimization-based	Slow	Any	High	Run for each content-style pair
Fast approximation	Fast	Single	High	Train long for each new style
Feature matching	Fast	Any/Several	Compromised	Limited set of styles, low quality

Can we find a style transfer algorithm that could quickly adapt to any style, while the adapted model maintains high efficiency and good image quality?



Comparison with Prior Arts



Style Interpolation



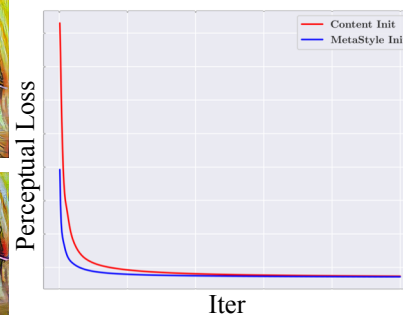
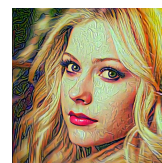
Video Style Transfer



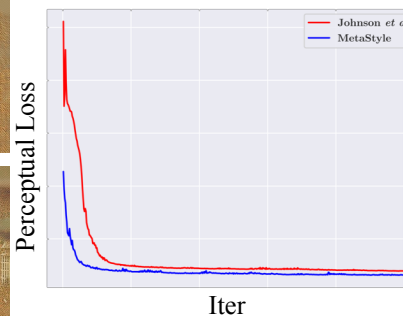
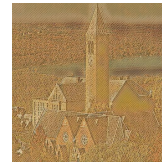
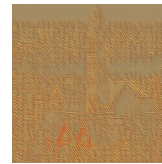
Quantitative Results

Method	Param	256 (s)	512 (s)	# Styles
Gatys <i>et al.</i>	N/A	7.7428	27.0517	∞
Johnson <i>et al.</i>	1.68M	0.0044	0.0146	1
Li <i>et al.</i>	34.23M	0.6887	1.2335	∞
Huang <i>et al.</i>	7.01M	0.0165	0.0320	∞
Shen <i>et al.</i>	219.32M	0.0045	0.0147	∞
Sheng <i>et al.</i>	147.22M	0.5089	0.6088	∞
Chen <i>et al.</i>	1.48M	0.2679	1.0890	∞
Ours	1.68M	0.0047	0.0145	∞^*

Investigate the Representation



Compare with Gatys *et al.* using contents preprocessed by MetaStyle



Compare with Johnson *et al.* using network initialized by MetaStyle