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

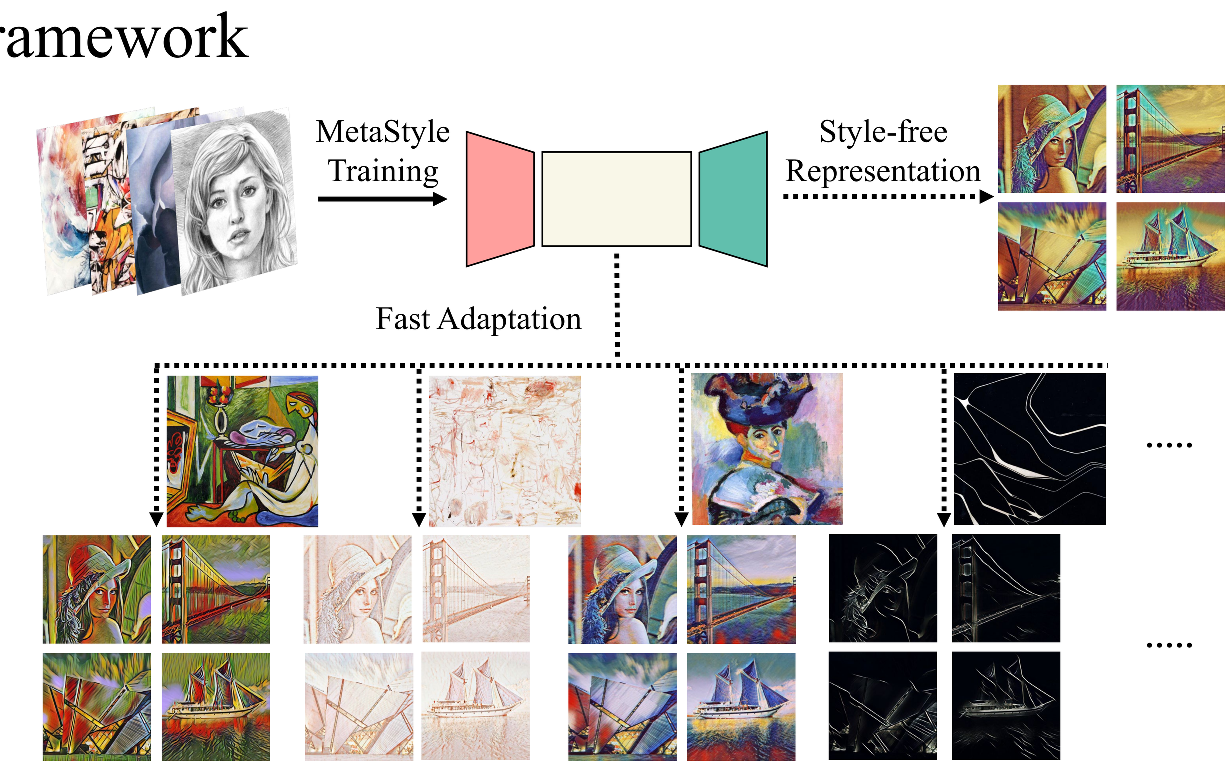
Motivation

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?

MetaStyle

- Framework



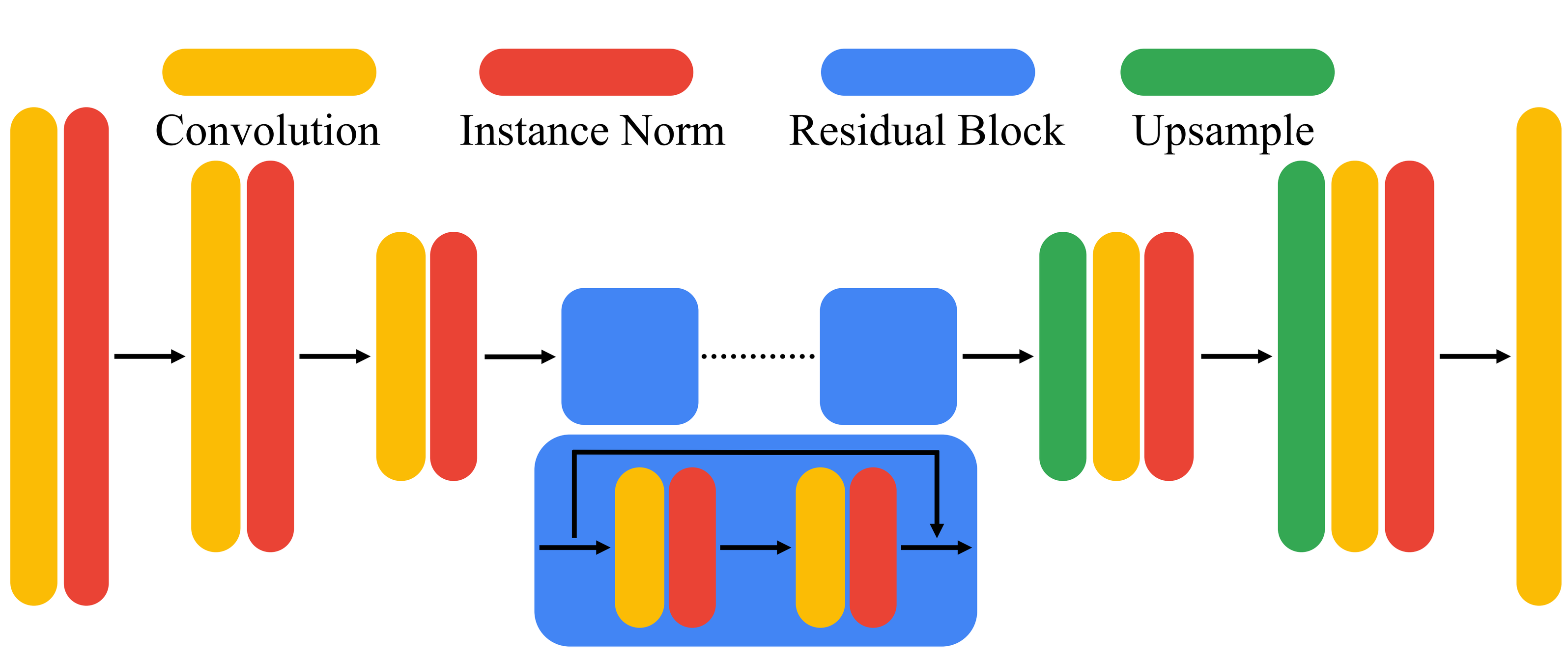
- Training

$$\begin{aligned} &\underset{\theta}{\text{minimize}} && \mathbb{E}_{c,s}[\ell(I_c, I_s, M(I_c; w_{s,T}))] \\ &\text{subject to} && w_{s,0} = \theta \\ & && w_{s,t} = w_{s,t-1} - \delta \nabla \mathbb{E}_c[\ell(I_c, I_s, M(I_c; w_{s,t-1}))] \end{aligned}$$

- Adaptation

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

Network



Algorithm

Algorithm 1: MetaStyle

Input : content training dataset $\mathcal{D}_{tr}$, content validation dataset $\mathcal{D}_{val}$, style dataset $\mathcal{D}_{style}$, inner learning rate δ, outer learning rate η, number of inner updates T

Output: trained parameters θ

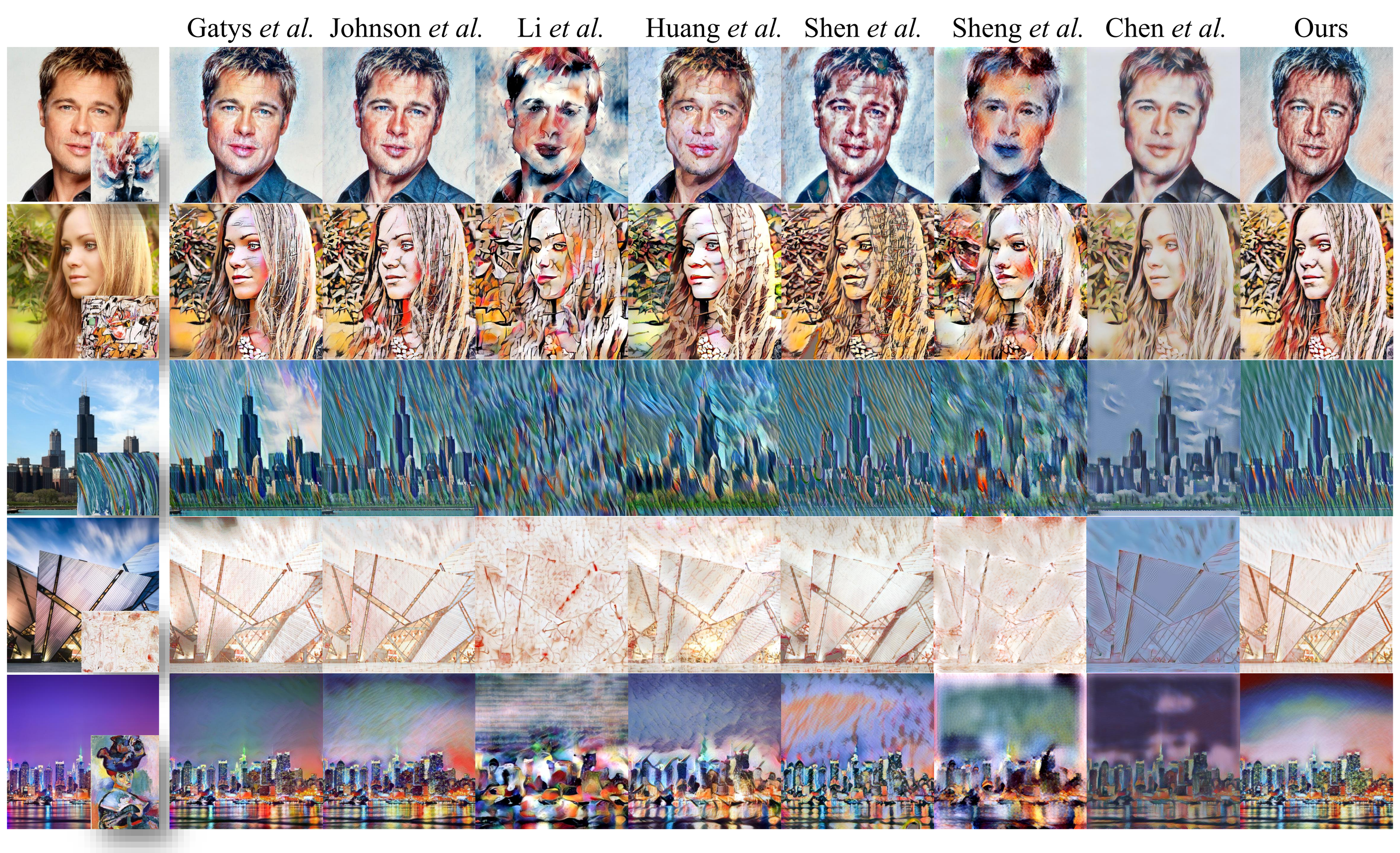
randomly initialize θ
while *not done* **do**
  initialize outer loss $E \leftarrow 0$
  sample a batch of styles from $\mathcal{D}_{style}$
  for *each style I_s* **do**
    $w_s \leftarrow \theta$
    for $i \leftarrow 1$ **to** T **do**
      sample a batch $\mathcal{B}_{tr}$ from $\mathcal{D}_{tr}$
      compute inner loss L_θ using I_s and $\mathcal{B}_{tr}$
      $w_s \leftarrow w_s - \delta \nabla L_\theta$
    end
    sample a batch $\mathcal{B}_{val}$ from $\mathcal{D}_{val}$
    increment E by loss from I_s and $\mathcal{B}_{val}$
  end
  $\theta \leftarrow \theta - \eta \nabla E$
end

Quantitative Results

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

Qualitative Results

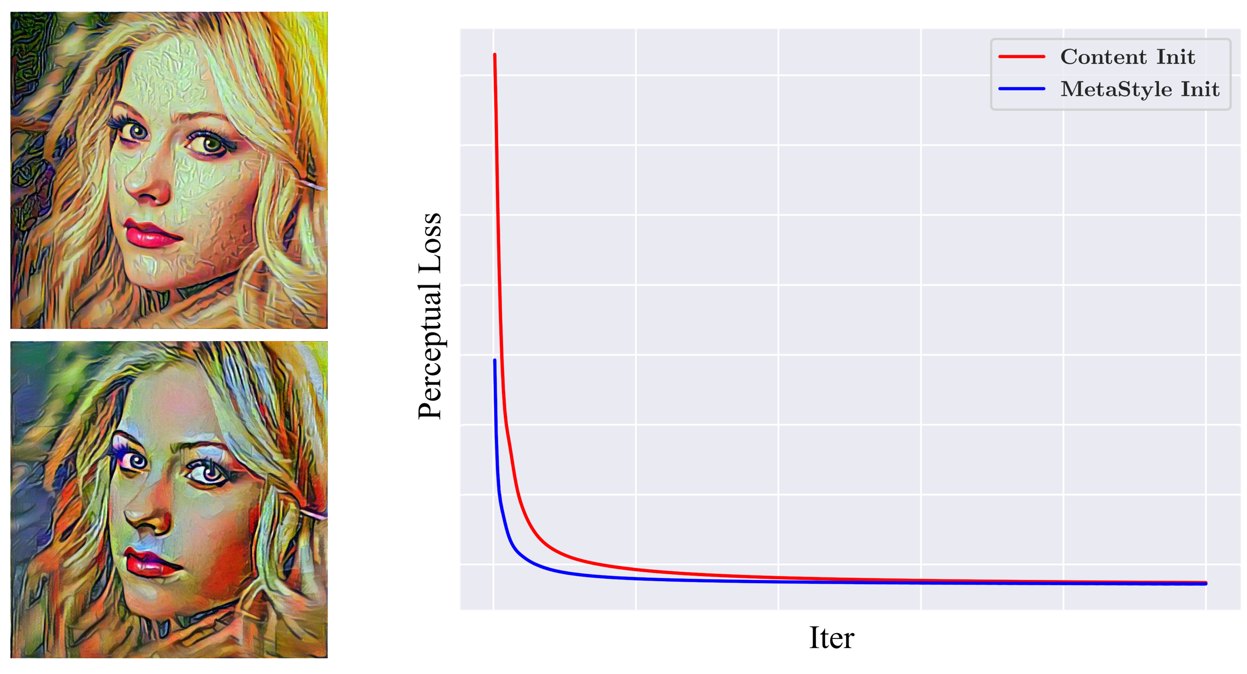
- Comparison with Existing Methods



- Style Interpolation

- Video Style Transfer

- Comparison to Gatys *et al.* with MetaStyle Init



- Comparison to Johnson *et al.* with MetaStyle