ECCV'20 ONLINE
23-28 AUGUST 2020

Deformable Style Transfer

Sunnie S. Y. Kim[1], Nick Kolkin[1], Jason Salavon[2] and Greg Shakhnarovich[1]

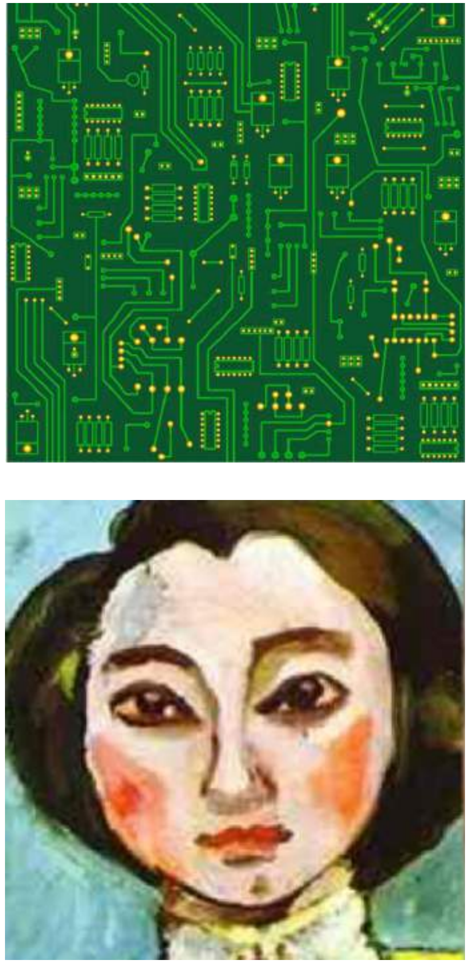



1 TOYOTA TECHNOLOGICAL INSTITUTE AT CHICAGO

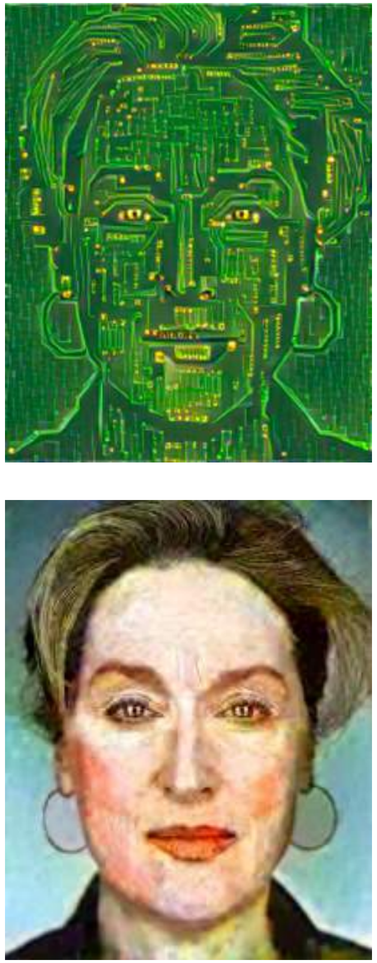
2 THE UNIVERSITY OF CHICAGO

Style Transfer

Content image



Style image



Output image

Early Approaches for Texture Synthesis

Painterly Rendering (SIGGRAPH'98)

Source image | Rough sketch | Intermediate sketch | Final painting

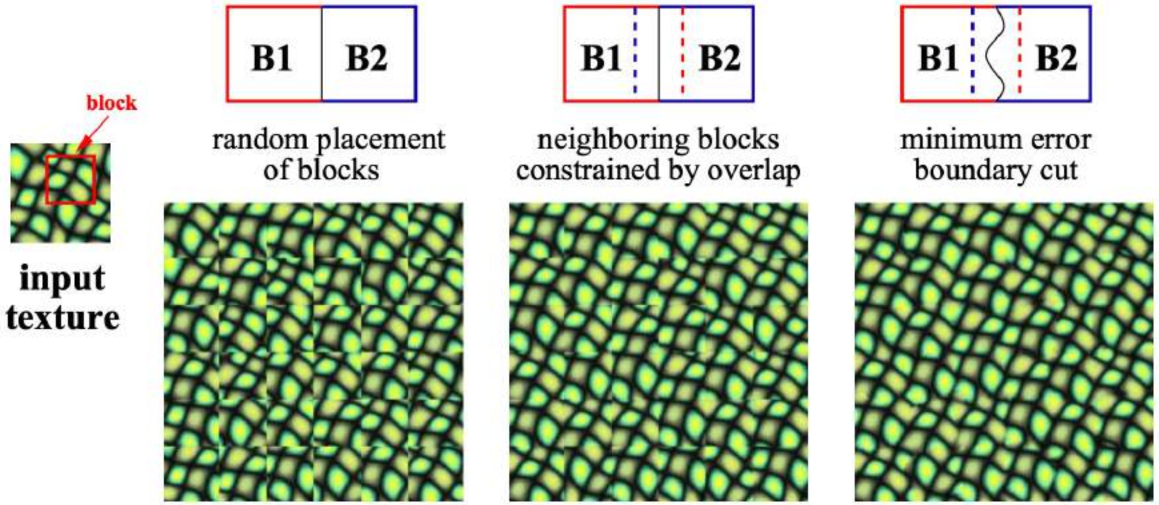
Image Analogies (SIGGRAPH'01)

A | A' | B | B'

"Painterly Rendering with Curved Brush Strokes of Multiple Sizes." Aaron Hertzmann. SIGGRAPH 1998.
"Image Analogies." Aaron Hertzmann, Charles E. Jacobs, Nuria Oliver, Brian Curless and David H. Salesin. SIGGRAPH 2001.

Early Approaches for Texture Synthesis

Image Quilting (SIGGRAPH'01)



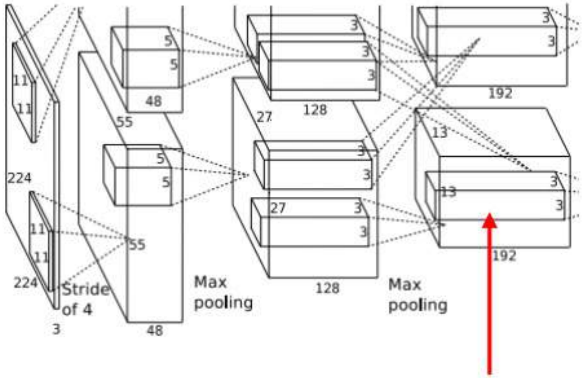
Input images

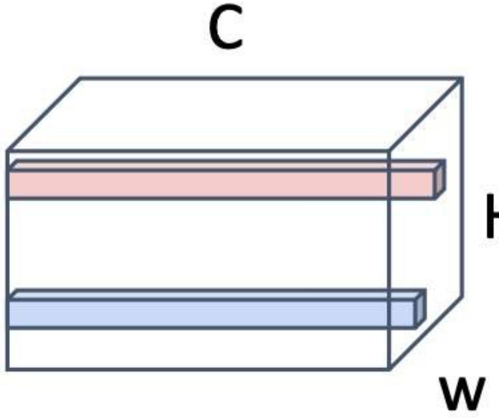
Quilting results

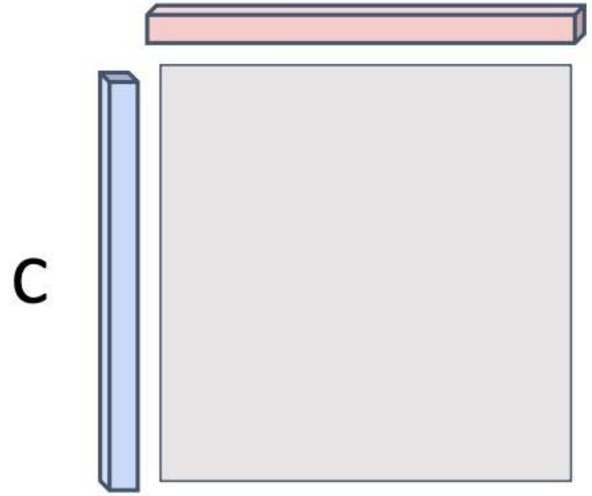
"Image Quilting for Texture Synthesis and Transfer." Alexei A. Efros and Willian T. Freeman. SIGGRAPH 2001.

Texture Synthesis with Neural Networks: Gram Matrix

This image is in the public domain.







Each layer of CNN gives C x H x W tensor of features; H x W grid of C-dimensional vectors

Outer product of two C-dimensional vectors gives C x C matrix of elementwise products

Average over all HW pairs gives **Gram Matrix** of shape C x C giving unnormalized covariance

Efficient to compute; reshape features from

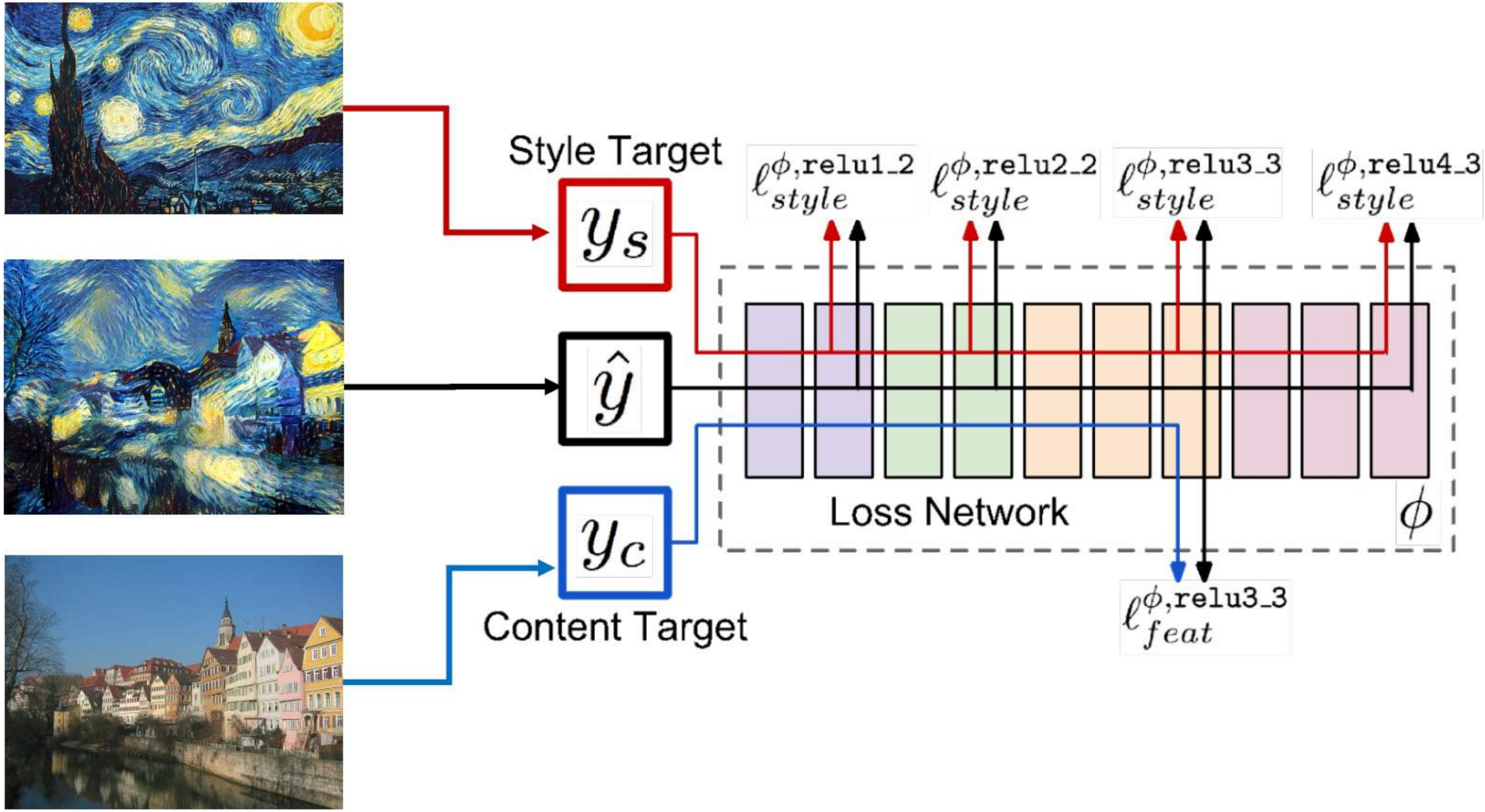
C x H x W to F = C x HW

then compute $G = FF^T$

Style image

Output image

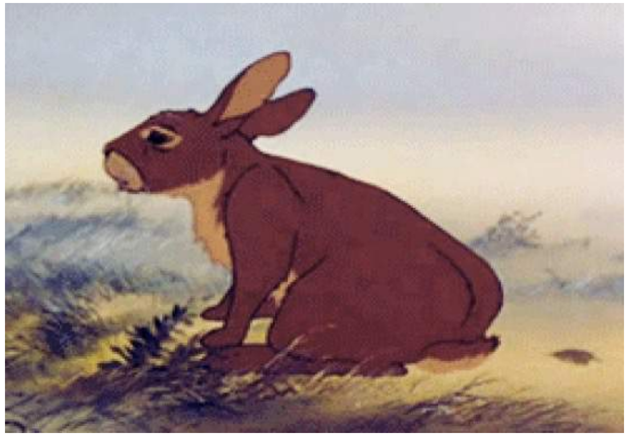
Content image



Gatys, Ecker, and Bethge, "Image style transfer using convolutional neural networks", CVPR 2016

Figure adapted from Johnson, Alahi, and Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution", ECCV 2016.

Beyond Color and Texture Transfer



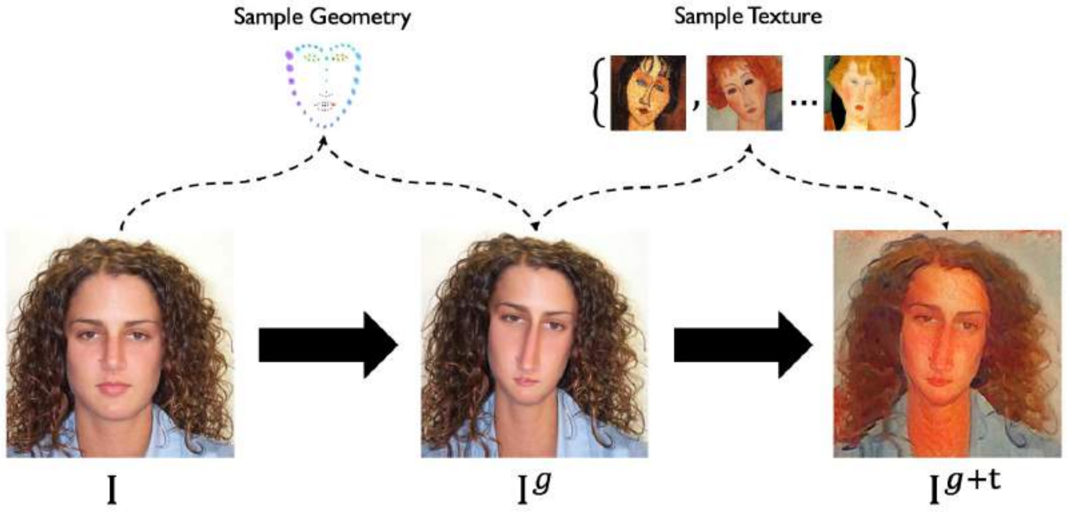
Content

Style

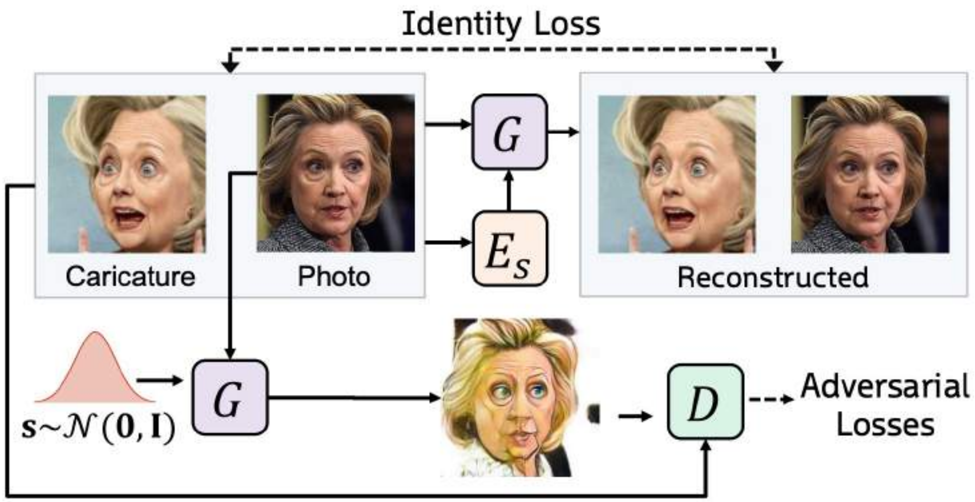
Stylized Output

Prior Work: Limited to Faces

The Face of Art (SIGGRAPH'19)

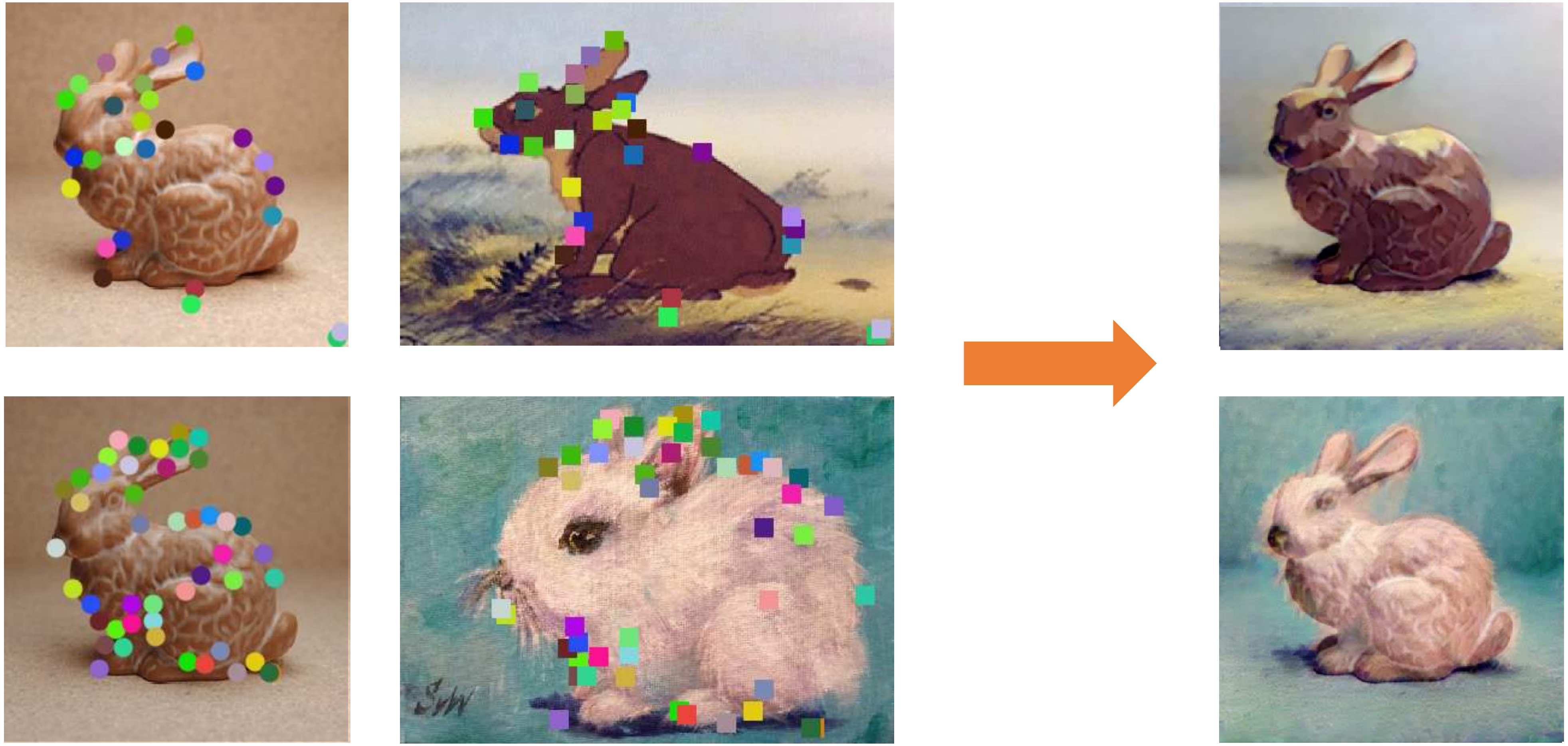


WarpGAN (CVPR'19)



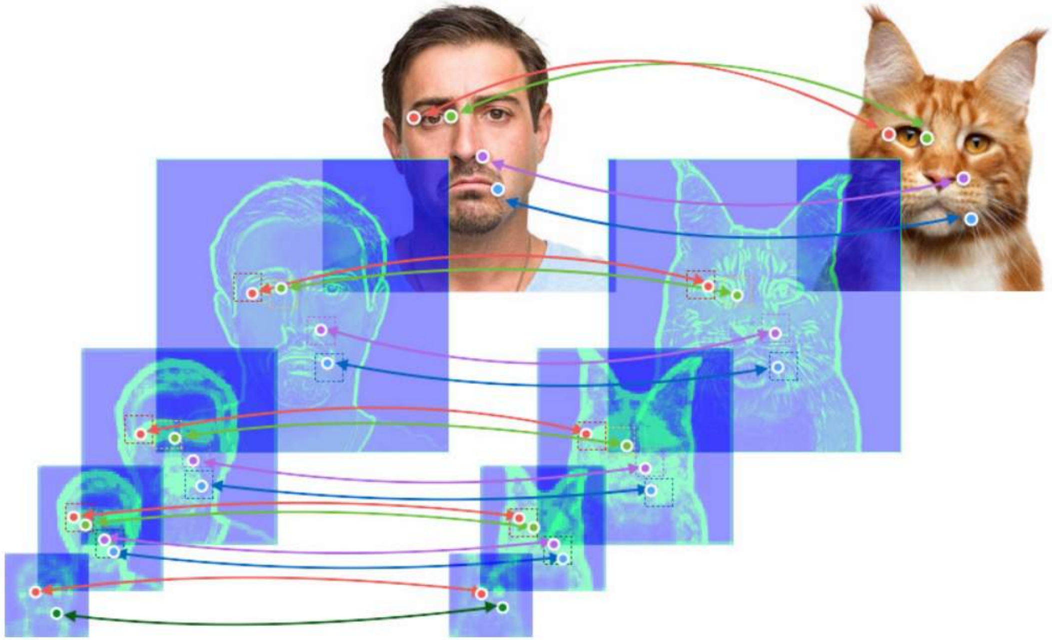
"The Face of Art: Landmark Detection and Geometric Style in Portraits." Jordan Yaniv, Yael Newman and Ariel Shamir. SIGGRAPH 2019.
"WarpGAN: Automatic Caricature Generation." Yichun Shi, Debayan Deb and Anil K. Jain. CVPR 2019.

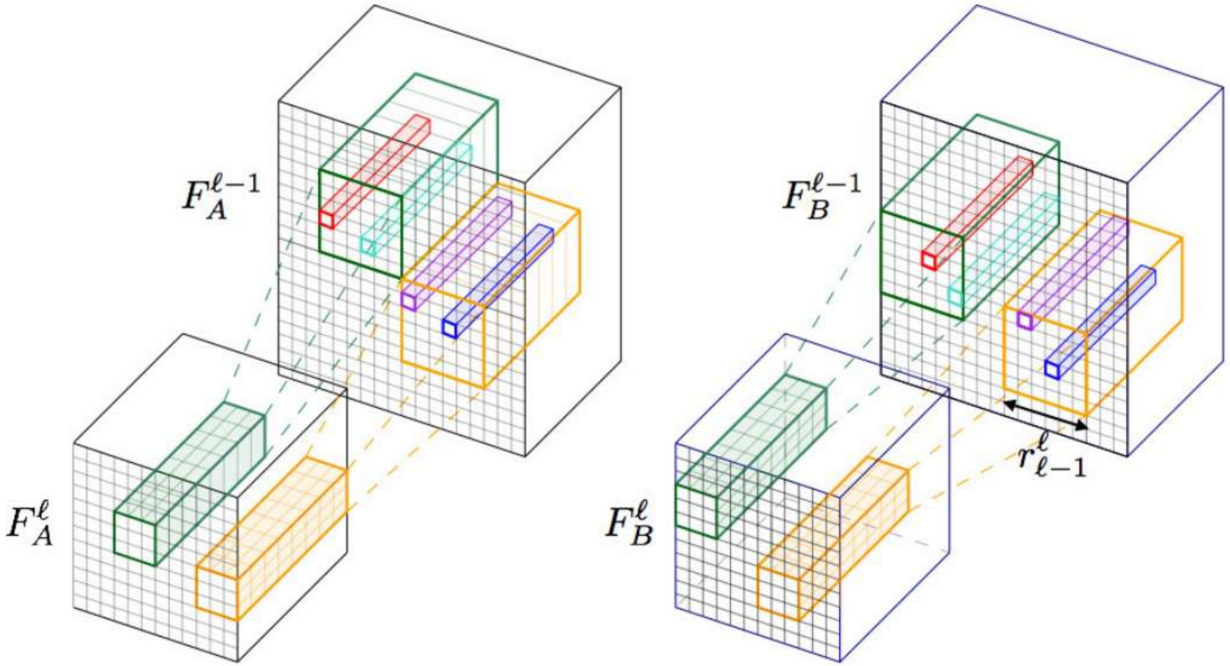
Deformable Style Transfer (DST)



Content Style Stylized Output

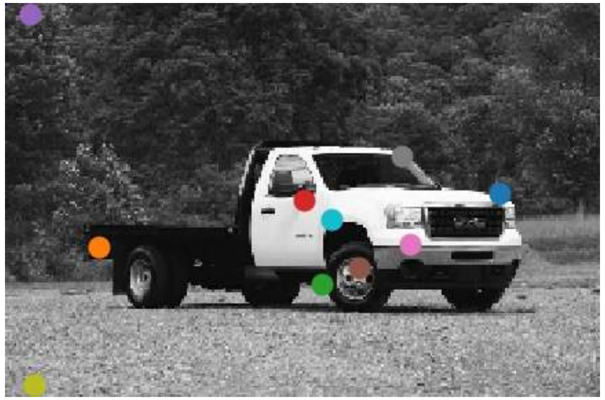
Geometry Transfer via Correspondences





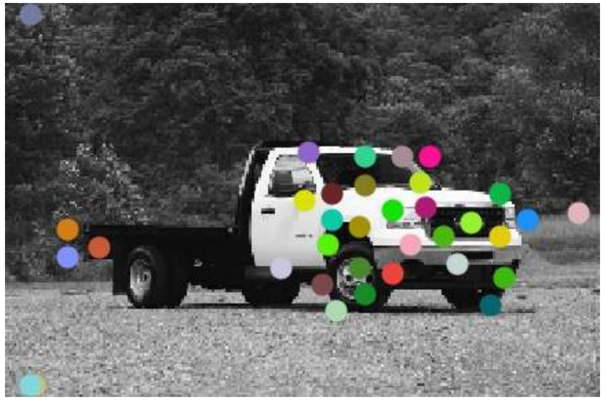
"Neural Best-Buddies: Sparse Cross-Domain Correspondence."
Kfir Aberman, Jing Liao, Mingyi Shi, Dani Lischinski, Baoquan Chen and Daniel Cohen-Or. SIGGRAPH 2018.

Geometry Transfer via Correspondences



Content image with original NBB points

Style image with original NBB points



Content image with modified points

Style image with modified points

“Neural Best-Buddies: Sparse Cross-Domain Correspondence.”
Kfir Aberman, Jing Liao, Mingyi Shi, Dani Lischinski, Baoquan Chen and Daniel Cohen-Or. SIGGRAPH 2018.

Geometry Transfer via Correspondences

⚫: Source points
⬛: Target points

Content image with
aligned points

Image warped by moving
source points to target points

Differentiable Image Warping

- $P = \{p_1, \ldots, p_k\}, \theta = \{\theta_1, \ldots, \theta_k\}, P + \theta = \{p_1 + \theta_1, \ldots, p_k + \theta_k\}$
- **Thin-plate spline interpolation** produces a dense flow field from the coordinates of an image I to a warped image $W(I, \theta)$.
- This is a closed-form procedure which finds parameters w, v, b that minimize $\sum_{i=1}^{k} \| f_\theta(p_i + \theta_i) - p_i \|^2$ subject to a curvature constraint.
- With these parameters, we have the inverse mapping function

$$f_\theta(q) = \sum_{i=1}^{k} w_i \phi(\| q - p_i - \theta_i \|) + v^T q + b$$

where q denotes the location of a pixel in the warped image and ϕ is a kernel function.

DST Objective

$$
\begin{aligned}
L(X,\theta,I_c,I_s,P,P') = {} & \alpha\, L_{content}\,(I_c,X) \\
& + L_{style}\,(I_s,X) + L_{style}\big(I_s,W(X,\theta)\big) \\
& + \beta\, L_{warp}(P,P',\theta) + \gamma\, R_{TV}(\theta)
\end{aligned}
$$

DST can use any one-shot, optimization-based style transfer method with a *content loss* and a *style loss*!

“Image Style Transfer Using Convolutional Neural Networks.” Leon A. Gatys, Alexander S. Ecker and Matthias Bethge. CVPR 2016.
“Style Transfer by Relaxed Optimal Transport and Self-Similarity.” Nicholas Kolkin, Jason Salavon and Gregory Shakhnarovich. CVPR 2019.

Base Style Transfer Methods

	Gatys et al.	STROTSS
$\boldsymbol{L_{content}\ (I_c, X)}$	Difference between $f_{deep}(I_c)$ and $f_{deep}(X)$	Difference between $D\big(f(I_c)\big)$ and $D\big(f(X)\big)$
$\boldsymbol{L_{style}\ (I_s, X)}$	Difference between $Gram\big(f(I_s)\big)$ and $Gram\big(f(X)\big)$	1. $\mathrm{REMD}\big(\mathrm{f}(\mathrm{I_s}), \mathrm{f}(\mathrm{X})\big)$ 2. Difference between $\mathrm{m}\big(\mathrm{f}(\mathrm{I_s})\big)$ and $\mathrm{m}\big(\mathrm{f}(\mathrm{X})\big)$ 3. Difference between $\mathrm{c}(\mathrm{I_s})$ and $\mathrm{c}(X)$

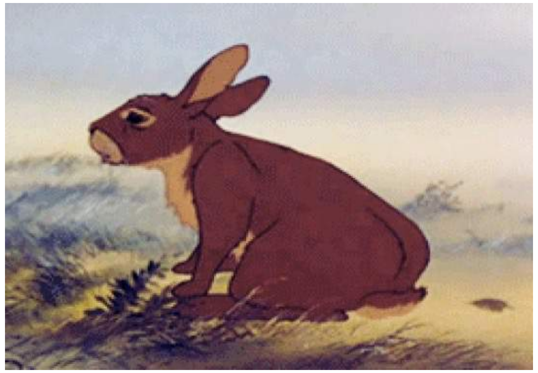
"Image Style Transfer Using Convolutional Neural Networks." Leon A. Gatys, Alexander S. Ecker and Matthias Bethge. CVPR 2016.
"Style Transfer by Relaxed Optimal Transport and Self-Similarity." Nicholas Kolkin, Jason Salavon and Gregory Shakhnarovich. CVPR 2019.

DST Objective

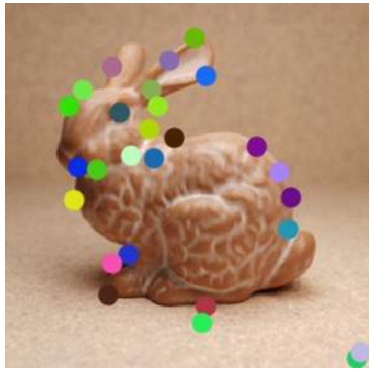
$$
\begin{aligned}
L(X, \theta, I_c, I_s, P, P') = {} & \alpha\, L_{content}\left(I_c, X\right) \\
& + L_{style}\left(I_s, X\right) + L_{style}\left(I_s, W(X, \theta)\right) \\
& + \beta\, L_{warp}\left(P, P', \theta\right) + \gamma\, R_{TV}(\theta)
\end{aligned}
$$

Input

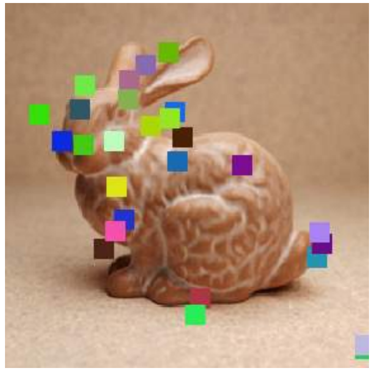
Content
I_c



Style
I_s

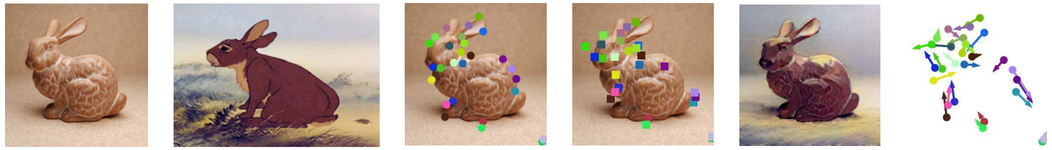


Source points
P



Target points
P'

DST Objective

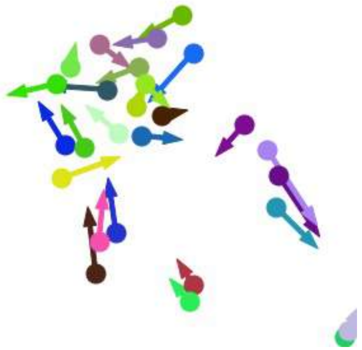


I_c I_s P P' X θ

$$
\begin{aligned}
L(X, \theta, I_c, I_s, P, P') = {} & \alpha\, L_{content}\,(I_c, X) \\
& + L_{style}\,(I_s, X) + L_{style}\big(I_s, W(X, \theta)\big) \\
& + \beta\, L_{warp}(P, P', \theta) + \gamma\, R_{TV}(\theta)
\end{aligned}
$$

Parameters

Stylization
X

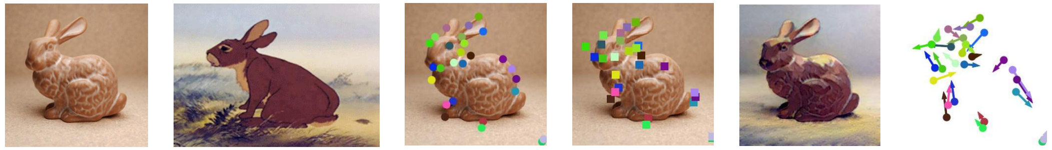


Deformation
θ

Output

Warped stylized image
$W(X, \theta)$

DST Objective

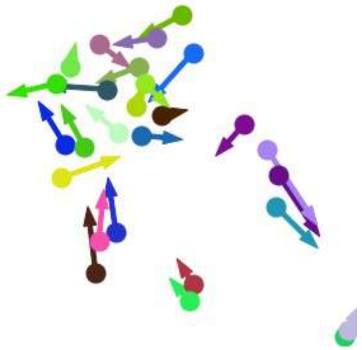


I_c I_s P P' X θ

$$
\begin{aligned}
L(X, \theta, I_c, I_s, P, P') = \; & \alpha \, L_{content}\,(I_c, X) \\
& + L_{style}\,(I_s, X) + L_{style}\big(I_s, W(X, \theta)\big) \\
& + \beta \, L_{warp}(P, P', \theta) + \gamma \, R_{TV}(\theta)
\end{aligned}
$$

Parameters

Stylization
X

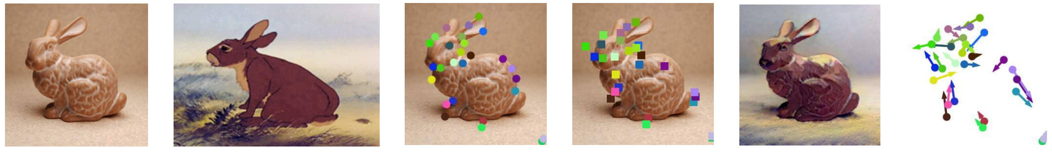


Deformation
θ

Output

Warped stylized image
$W(X, \theta)$

DST Objective



I_c I_s P P' X θ

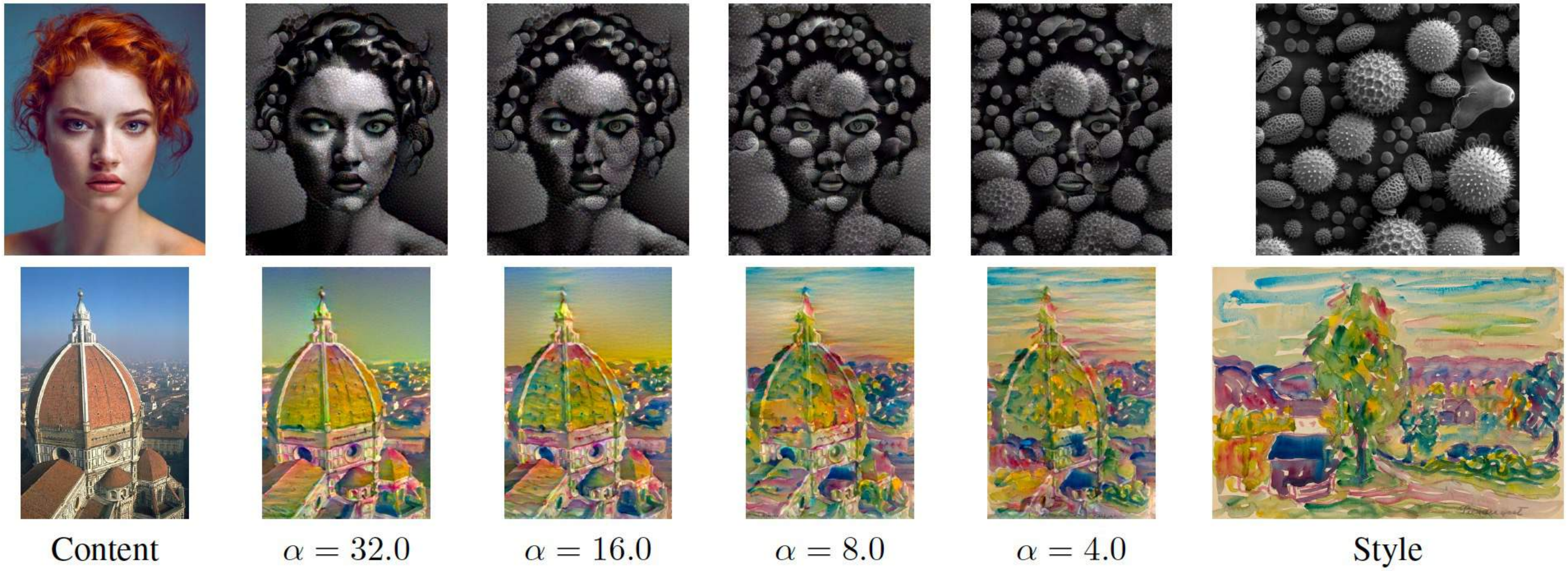
$$
\begin{aligned}
L(X,\theta,I_c,I_s,P,P') = {} & \alpha\, L_{content}\,(I_c,X) \\
& + L_{style}\,(I_s,X) + L_{style}\big(I_s,W(X,\theta)\big) \\
& + \beta\, L_{warp}(P,P',\theta) + \gamma\, R_{TV}(\theta)
\end{aligned}
$$

- $L_{content}\,(I_c,X)$: Content loss of the base style transfer method
- $L_{style}\,(I_s,X)$: Style loss of the base style transfer method
- $L_{style}\,(I_s,W(X,\theta))$: Style loss applied to the warped stylized image
- $L_{warp}(P,P',\theta)$: Mean l_2 distance between optimized and target points
- $R_{TV}(\theta)$: Total variation norm of the 2D warp field

Effect of Varying α

$$L(X,\theta,I_c,I_s,P,P') = \alpha\, L_{content}(I_c,X) + L_{style}(I_s,X) + L_{style}(I_s,W(X,\theta)) + \beta\, L_{warp}(P,P',\theta) + \gamma\, R_{TV}(\theta)$$

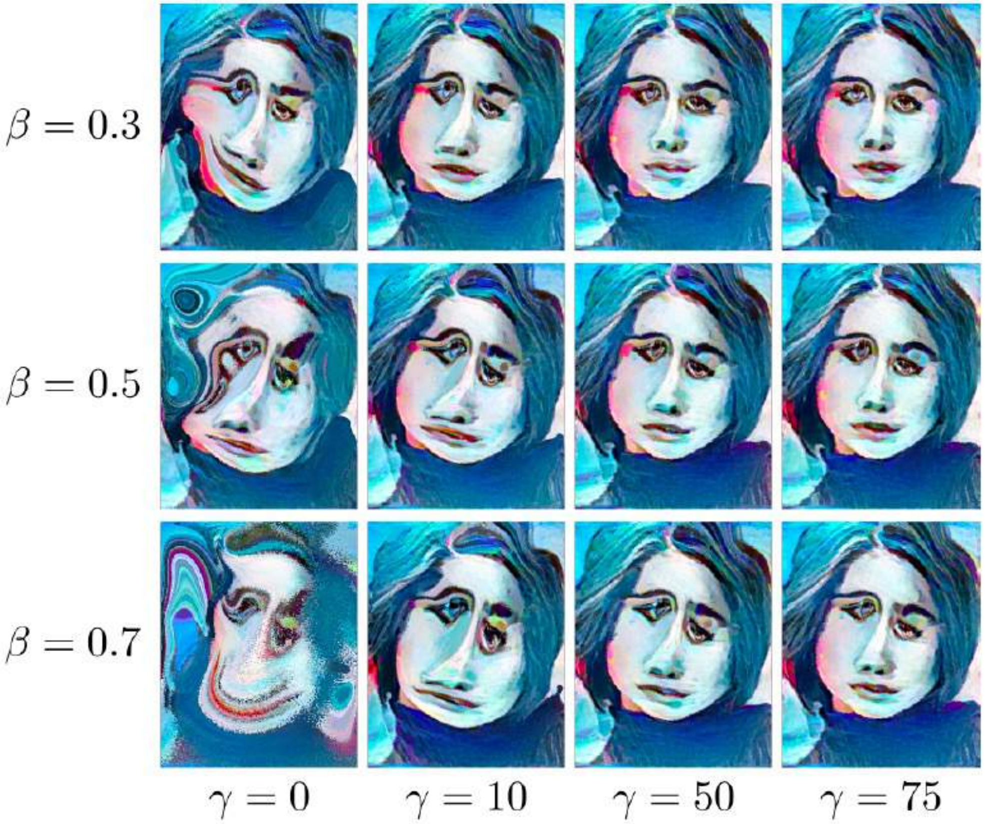
[Figure 6 from the STROTSS paper]



“Style Transfer by Relaxed Optimal Transport and Self-Similarity.” Nicholas Kolkin, Jason Salavon and Gregory Shakhnarovich. CVPR 2019.

Effect of Varying β and γ

$$
\begin{aligned}
L(X,\theta,I_c,I_s,P,P') = {} & \alpha\, L_{content}(I_c,X) \\
& + L_{style}(I_s,X) + L_{style}\big(I_s,W(X,\theta)\big) \\
& + \beta\, L_{warp}(P,P',\theta) + \gamma\, R_{TV}(\theta)
\end{aligned}
$$



Results

Content

Style

STROTSS

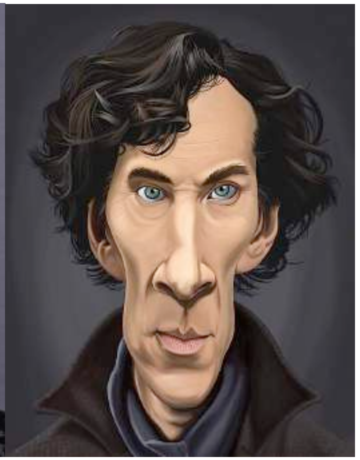
STROTSS +
Naïve warp

DST (Ours)

Content +
DST warp

Results

Content



Style

STROTSS

STROTSS +
Naïve warp

DST (Ours)

Content +
DST warp

Results

Content Style DST output

Multiple Contents to One Style

One Content to Multiple Styles

Content Style DST output

Comparison with FoA and WarpGAN

Content Style DST FoA Content Style DST WarpGAN

"The Face of Art: Landmark Detection and Geometric Style in Portraits." Jordan Yaniv, Yael Newman and Ariel Shamir. SIGGRAPH 2019.
"WarpGAN: Automatic Caricature Generation." Yichun Shi, Debayan Deb and Anil K. Jain. CVPR 2019.

DST with Non-NBB Keypoints

FoA facial landmarks

Content | Style | DST output | Naïve warp

Manually selected points

Content | Style | DST output | Naïve warp

Human Evaluation

Measure 1: Content preservation

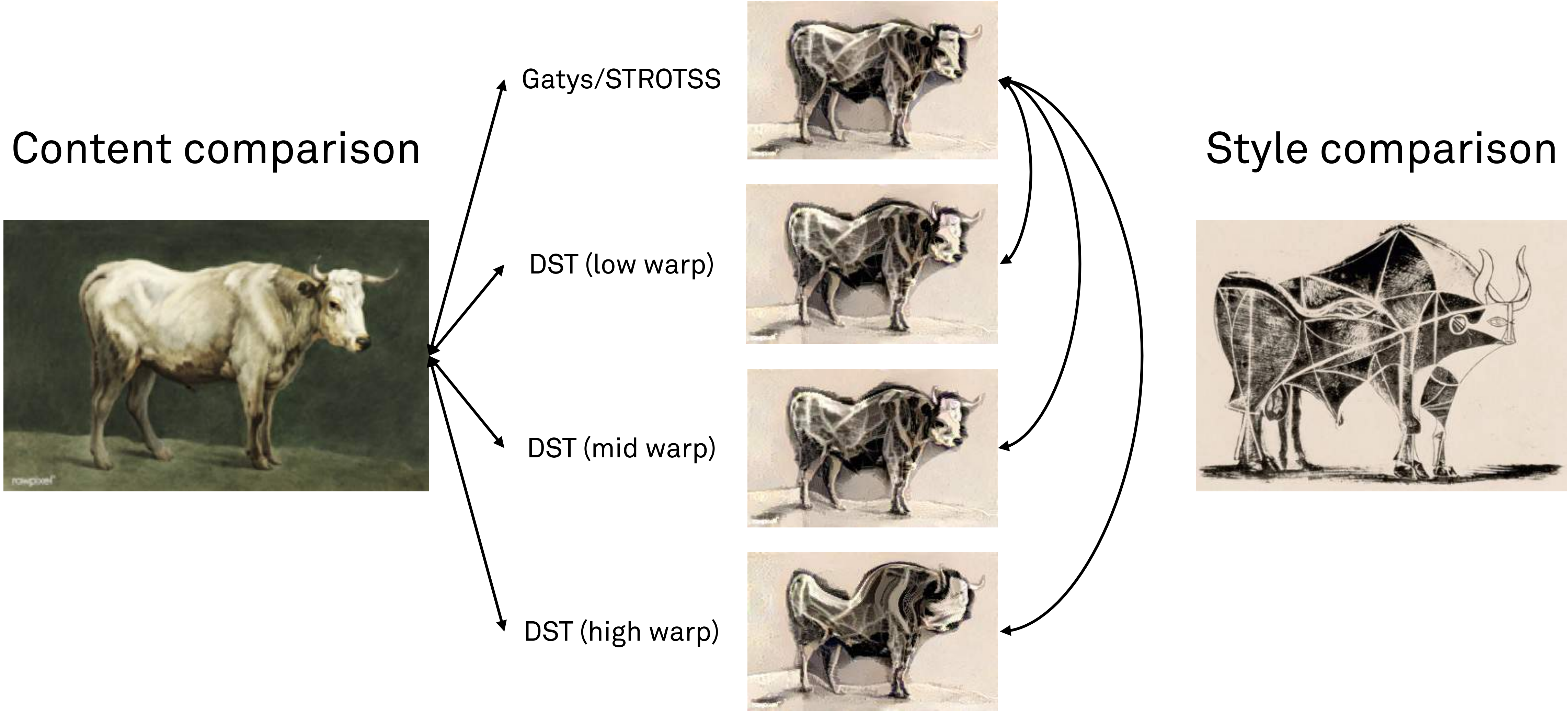
How would you respond to the following statement?

"Image A represents the same scene as Image B"

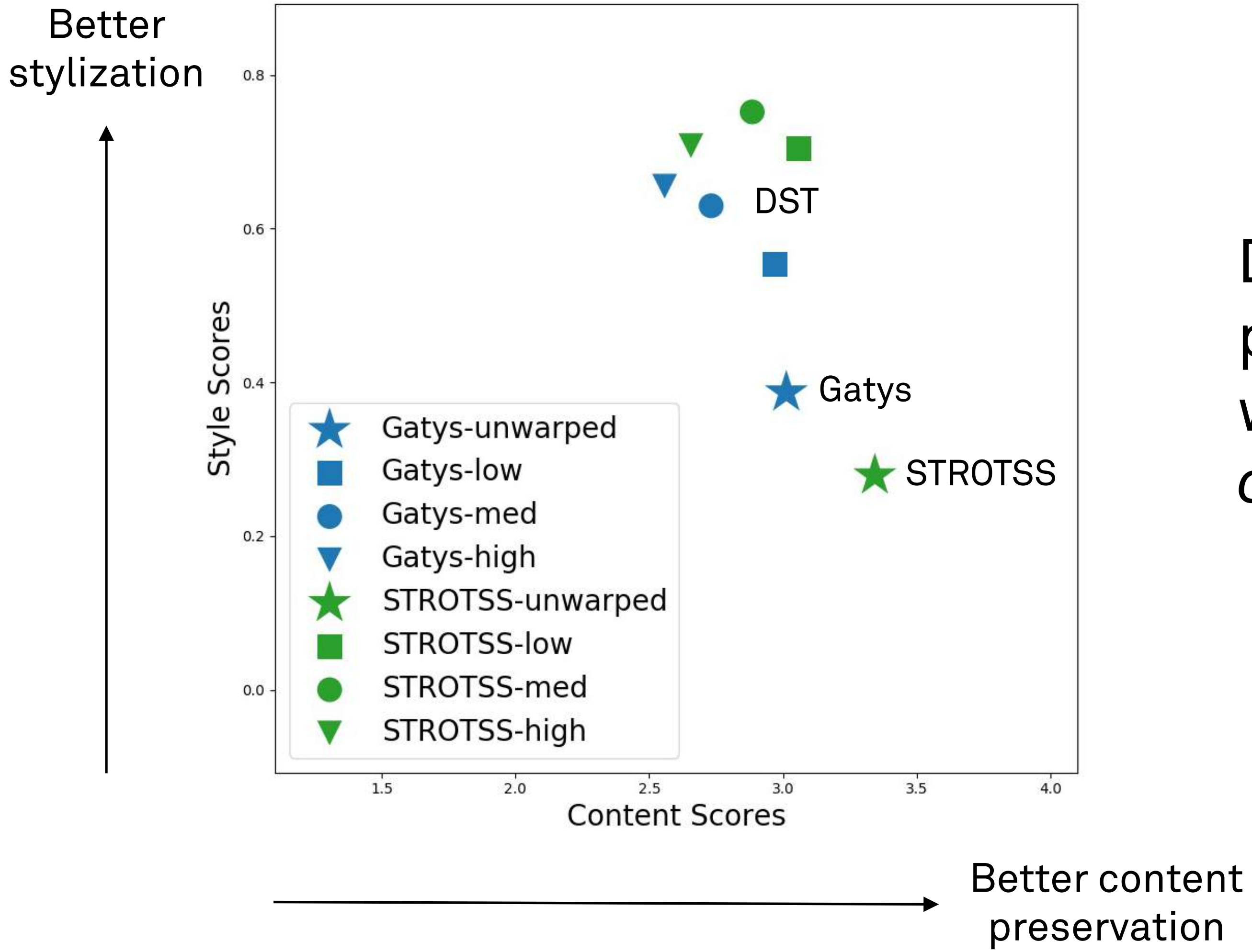
Measure 2: Stylization

Which of Image A or Image B better matches the style of the reference?

Human Evaluation

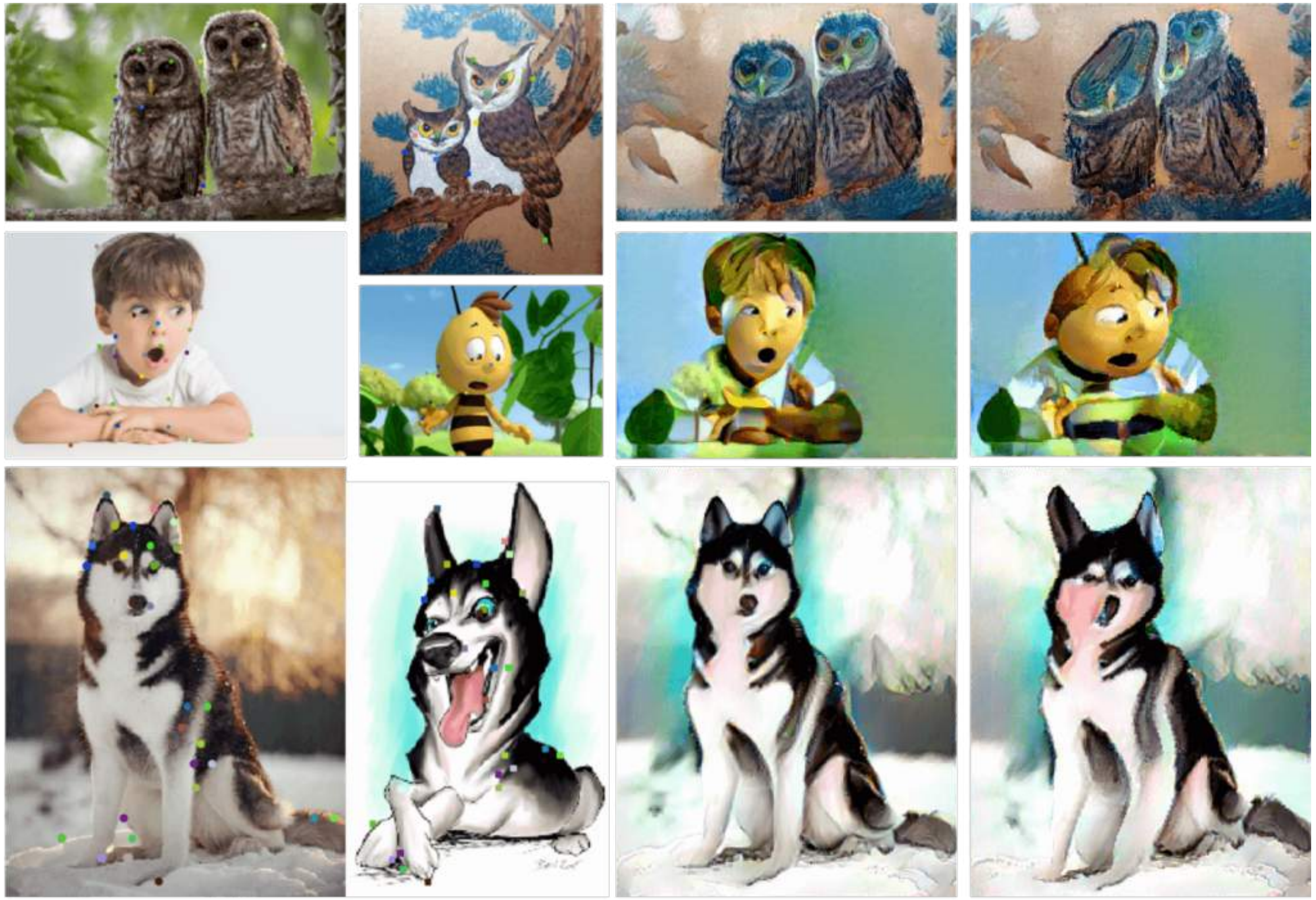


Human Evaluation



DST provides a higher perceived degree of *stylization* without a significant sacrifice in *content preservation*!

Limitations



Content Style STROTSS DST

Deformable Style Transfer (DST)

- Demonstrates for the first time geometry-aware style transfer in a one-shot setting
- Transfers geometric style via automatic deformation of images integrated into an optimization-based method
- Works on non-face images with the assumption that they have some approximate alignment
- Allows explicit user guidance and control of stylization tradeoffs

Project page: https://sunniesuhyoung.github.io/DST-page

Paper: https://arxiv.org/abs/2003.11038

Code: https://github.com/sunniesuhyoung/DST

Demo: https://bit.ly/DST-demo