Università degli Studi di Padova

Using JoJoGAN in the thesis project

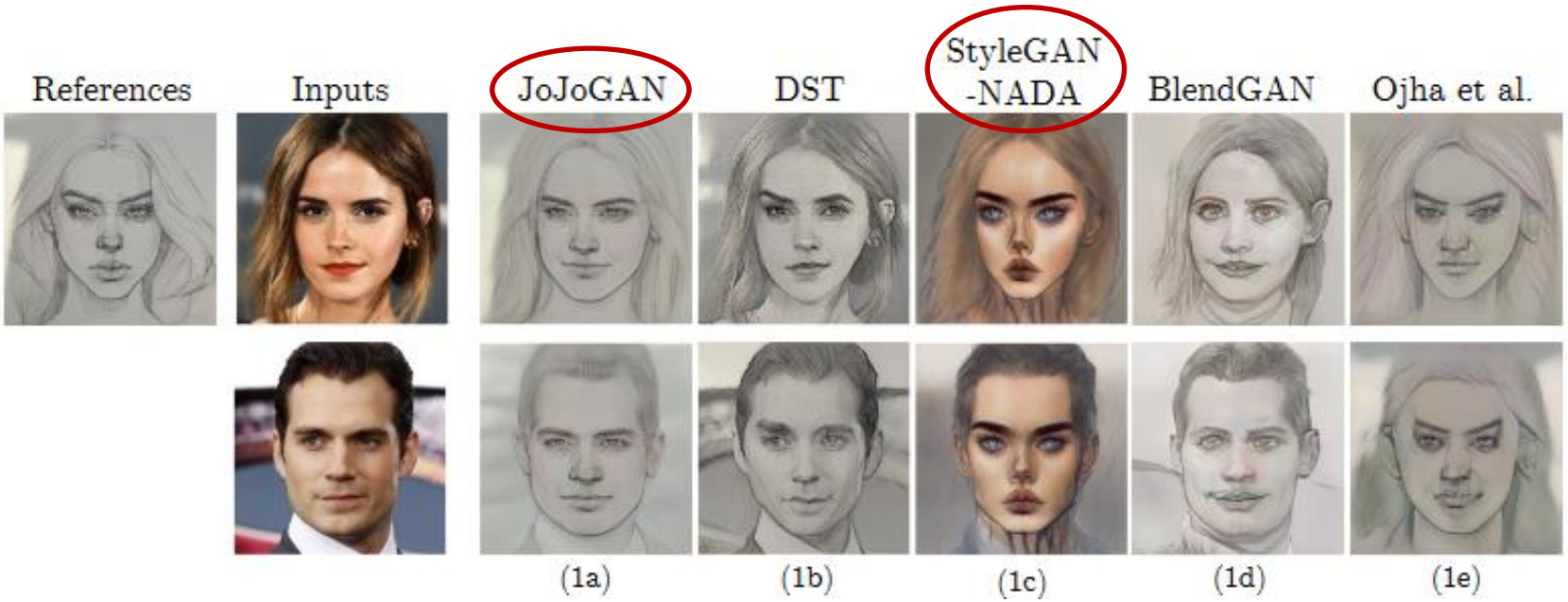
Student: Giorgia Bortoletti
Supervisor: Emanuele Menegatti
Co-supervisor: Gloria Beraldo

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

INTELLIGENT AUTONOMOUS SYSTEMS LAB

JoJoGAN (2022): One shot face stylization

"StyleGAN-NADA[1] fails to capture minute facial details that are important for face stylization."

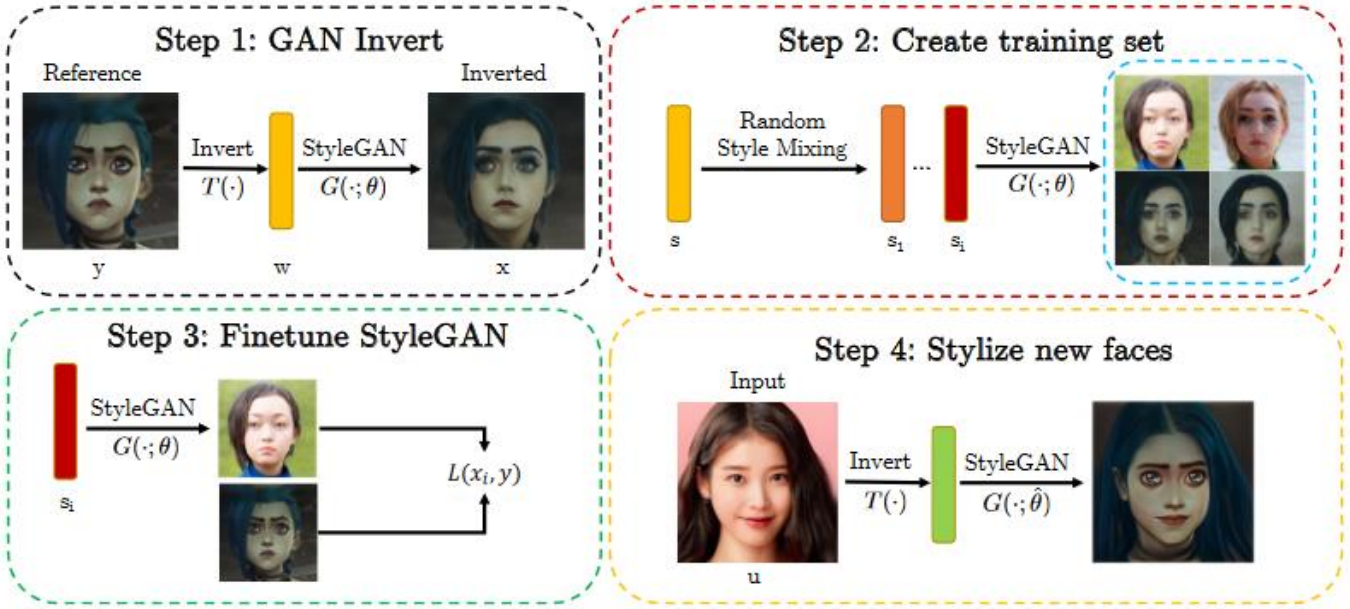


[1] https://stylegan-nada.github.io/

Workflow

JoJoGAN

$$s_i = M \cdot s + (1 - M) \cdot s(FC(z_i))$$



$$\hat{\theta} = \underset{\theta}{\text{argmin}}\, \text{loss}(\theta)$$

$$= \underset{\theta}{\text{argmin}}\, \frac{1}{N} \sum_{i}^{N} \mathcal{L}(G(s_i;\theta), y)$$

Fig. 2. Workflow: JoJoGAN's steps are: **GAN Inversion** to obtain a code s from the style reference; creating a **training set** $\mathcal{S}$ of similar s_i via random style mixing; **finetuning** a StyleGAN to obtain $\hat{\theta}$ so that $G(w_i;\hat{\theta}) \approx y$ using our perceptual loss; and **inference** by computing $G(T(u);\hat{\theta})$ for input u.

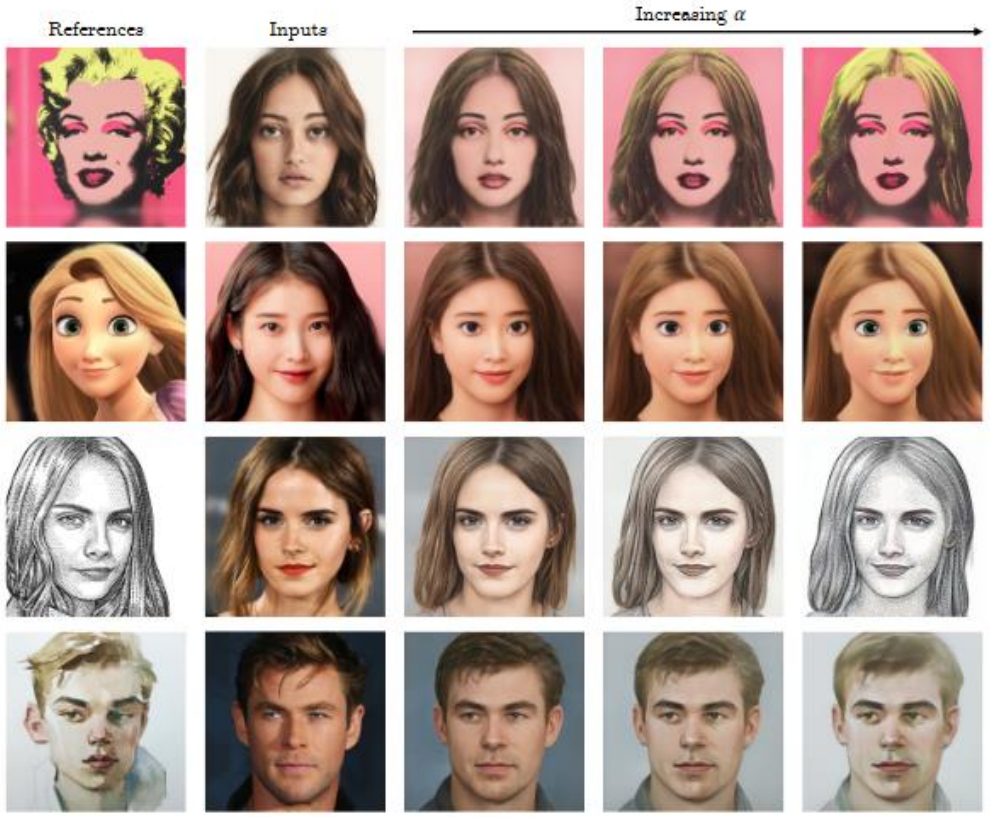


Fig. 5. Feature interpolation allows a user to control style intensity. As α increases, the results take the style of the reference more strongly.

Bugs

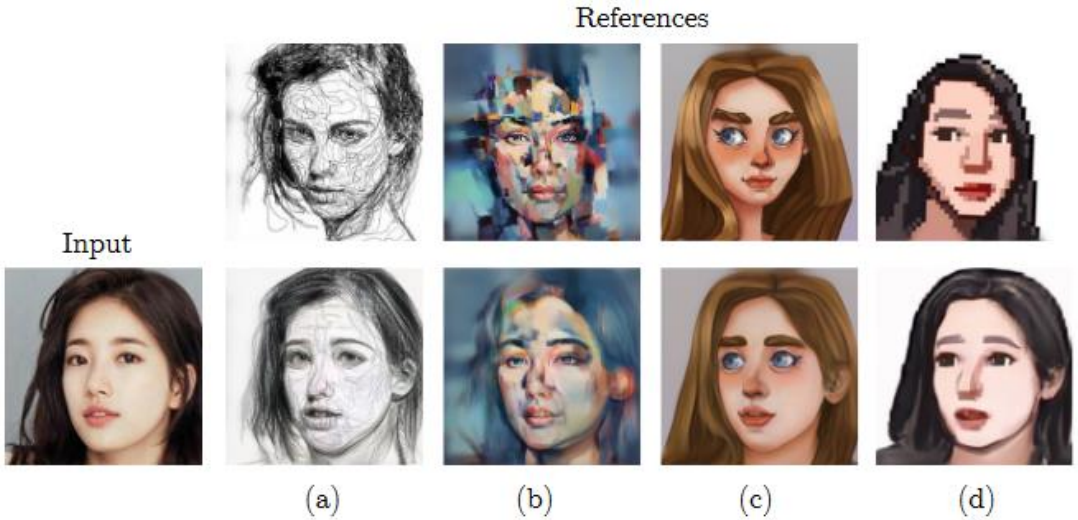
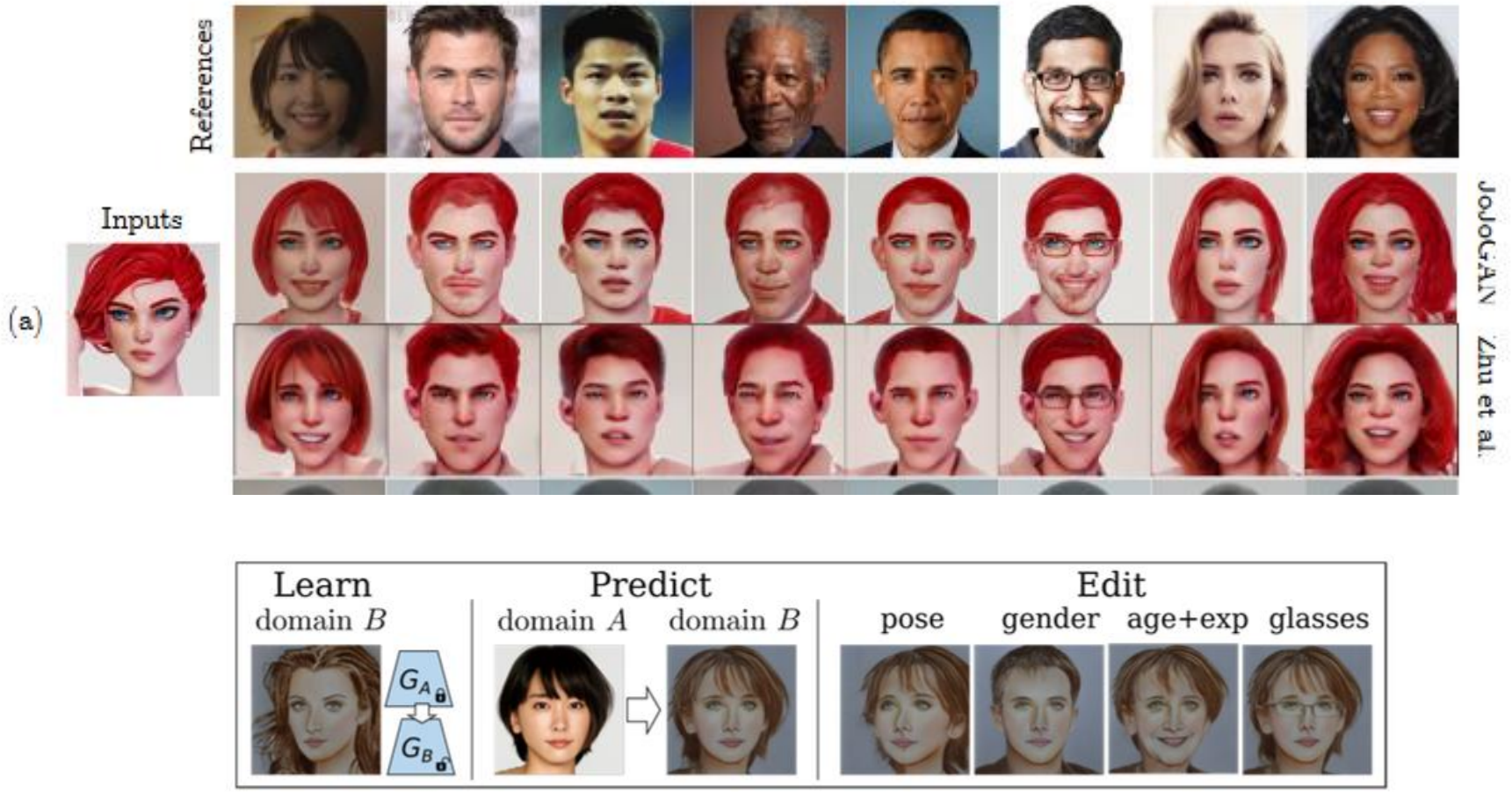


Fig. 13. Some style references are hard for JoJoGAN, likely a result of complicated structures in the style reference that are unfamiliar to StyleGAN. Note: loops in (a) mapped to strokes in the output; structure of brush strokes in (b) being broken up in output; gaze direction in (c) controlled by style reference rather than by input; high frequency pixel grids in (d) map to smooth strokes.

«Mind the Gap»



Zhu, Peihao, Rameen Abdal, John Femiani, and Peter Wonka. 'Mind the Gap: Domain Gap Control for Single Shot Domain Adaptation for Generative Adversarial Networks', 2022. https://openreview.net/forum?id=vqGi8Kp0wM.

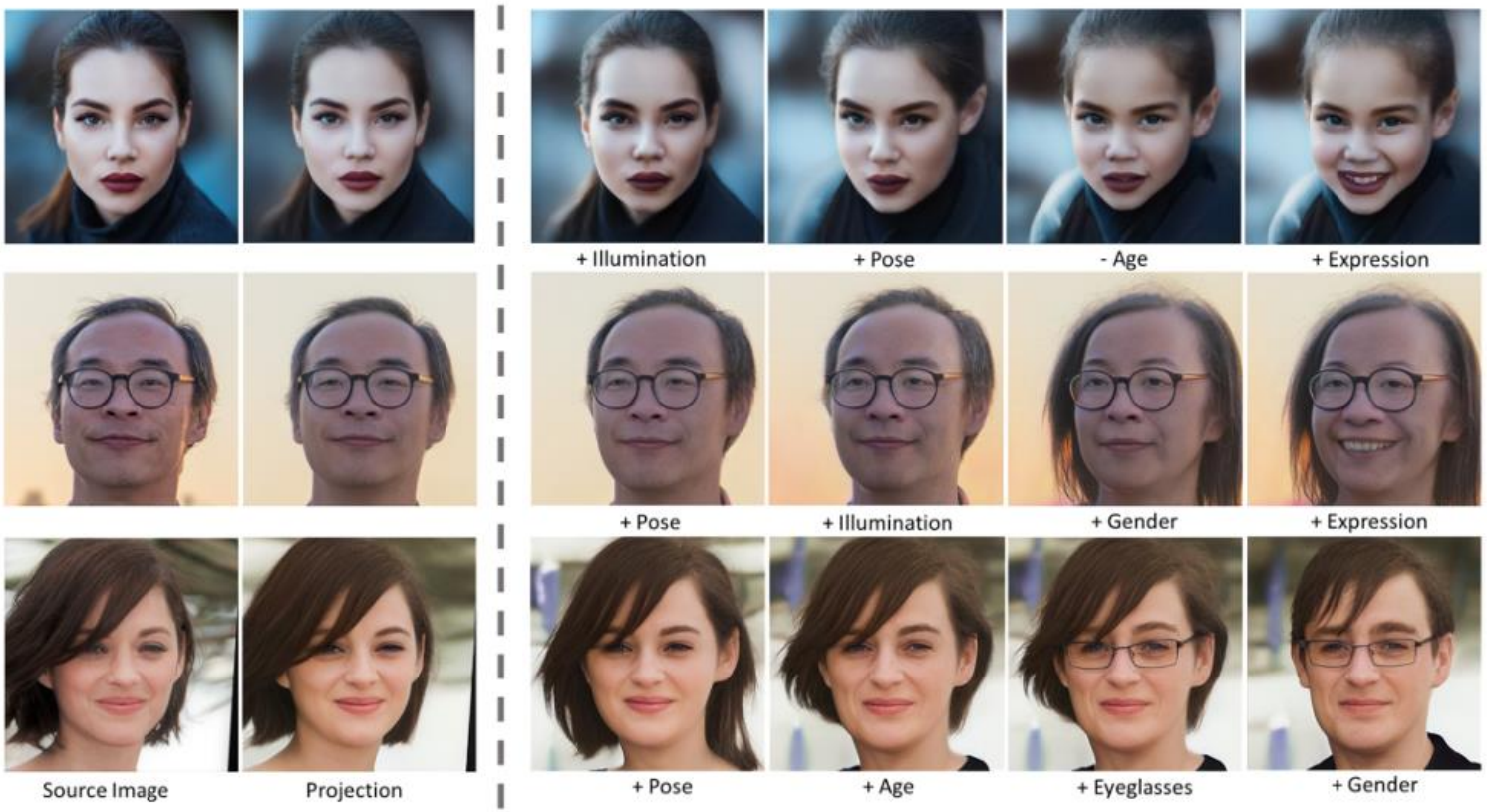


Fig. 1. We present *StyleFlow* to enable attribute-conditioned semantic edits on projected real images and StyleGAN generated images. For each of these examples, the user sequentially changes (camera) pose, illumination, expression, eyeglasses, gender, and age of a real image. Please judge, where applicable, the extent of identity preservation of the respective person under the applied edits. In this figure, all the source images are real images.

Abdal, Rameen, Peihao Zhu, Niloy J. Mitra, and Peter Wonka. 'StyleFlow: Attribute-Conditioned Exploration of StyleGAN-Generated Images Using Conditional Continuous Normalizing Flows'. *ACM Transactions on Graphics* 40, no. 3 (Maggio 2021): 21:1-21:21. https://doi.org/10.1145/3447648.