S^2–Flow: Joint Semantic and Style Editing of Facial Images

Krishnakant Singh[1]
Simone Schaub-Meyer[1,2]
Stefan Roth[1,2]

[1] TECHNISCHE UNIVERSITÄT DARMSTADT
[2] hessian.AI

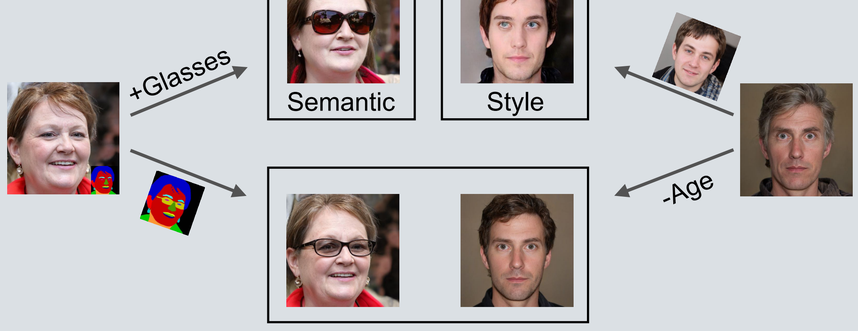
BMVC 2022

Motivation

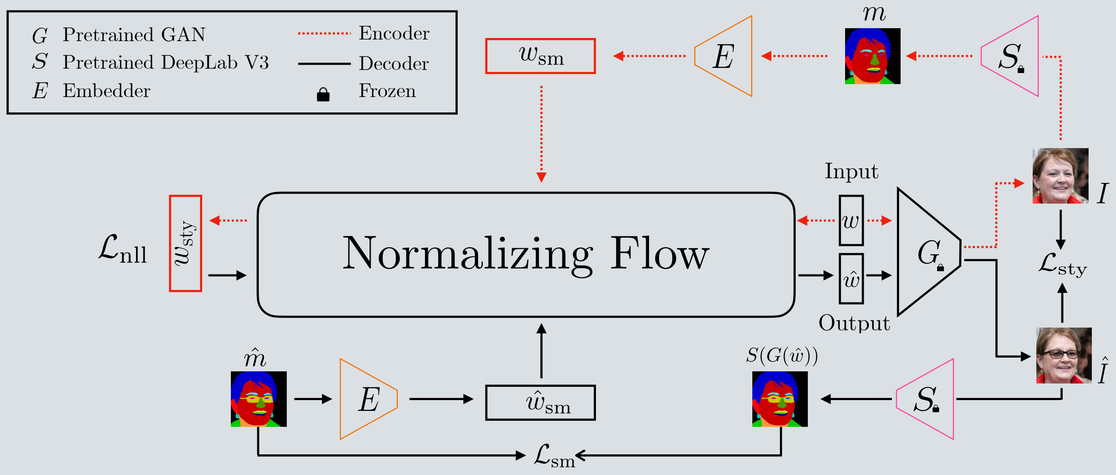
For controlled semantic and style edits, we need to disentangle the semantic and style spaces. Once disentangled, we can apply semantic and style edits in their respective domains for more controlled edits.

	Conditional			Unconditional	Ours
Method	Text	Semantic	Attribute		
Semantic	◆	✓	◆	◆	✓
Style	◆	◆	✓	◆	✓

◆: Limited



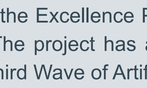
Framework of S^2–Flow



- **Encoder:** The encoder model takes the latent code w and the corresponding semantic mask m, and outputs a style code w_{sty} .
- **Decoder:** The decoder model combines the style code w_{sty} and the edited semantic code $\hat{w}_{sm}$ to output an edited latent code $\hat{w}$.

Use **transformation consistency** for disentangling the latent space of GAN into semantic and style:

- **Semantic consistency loss**: Edits to the semantic mask should be reflected in the generated image

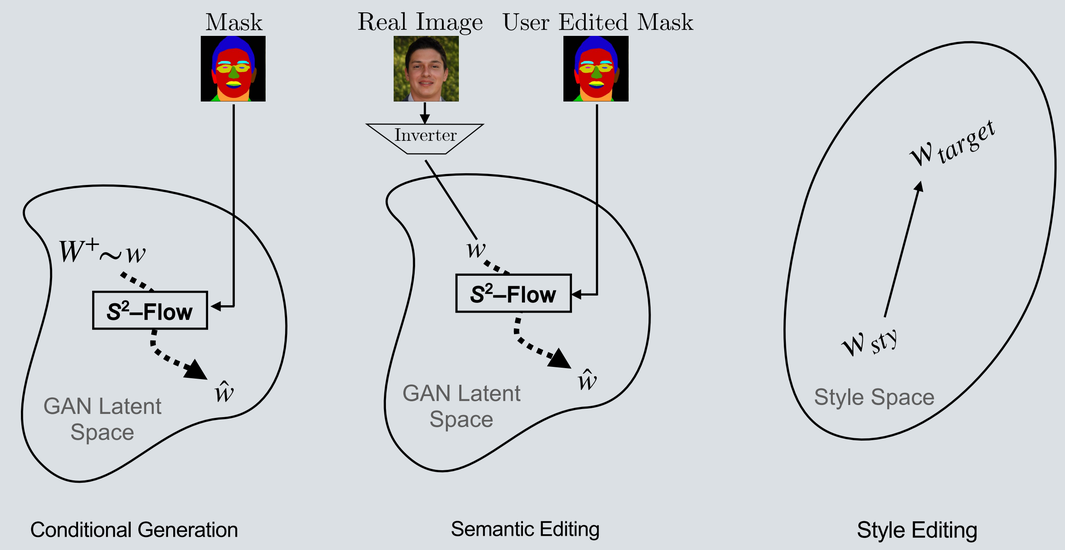
$$\mathcal{L}_{\text{sm}}\left(\text{  }\right)$$

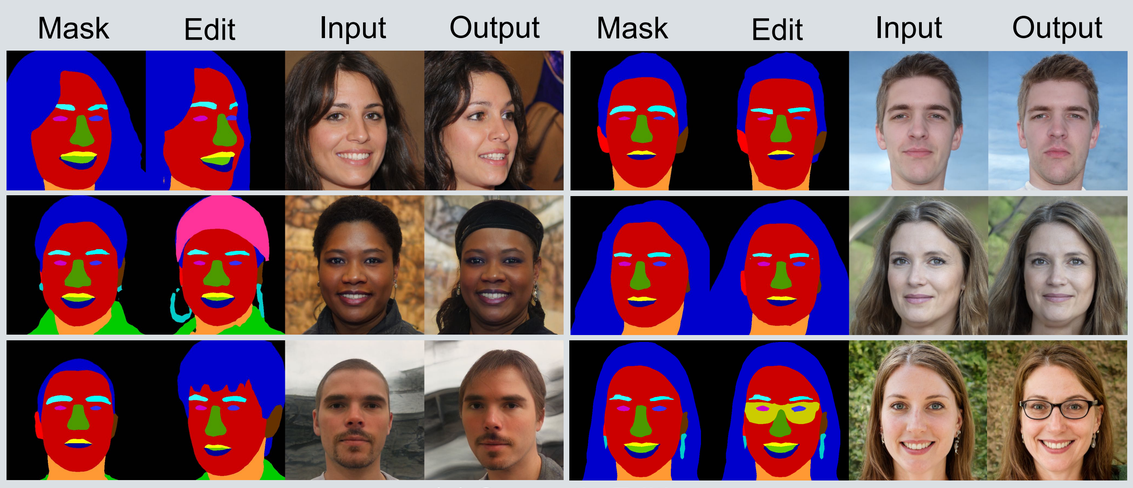
- **Style consistency loss**: Edits made in the semantic domain should not affect the style of the generated image

$$\mathcal{L}_{\text{sty}}\left(\text{}\right)$$

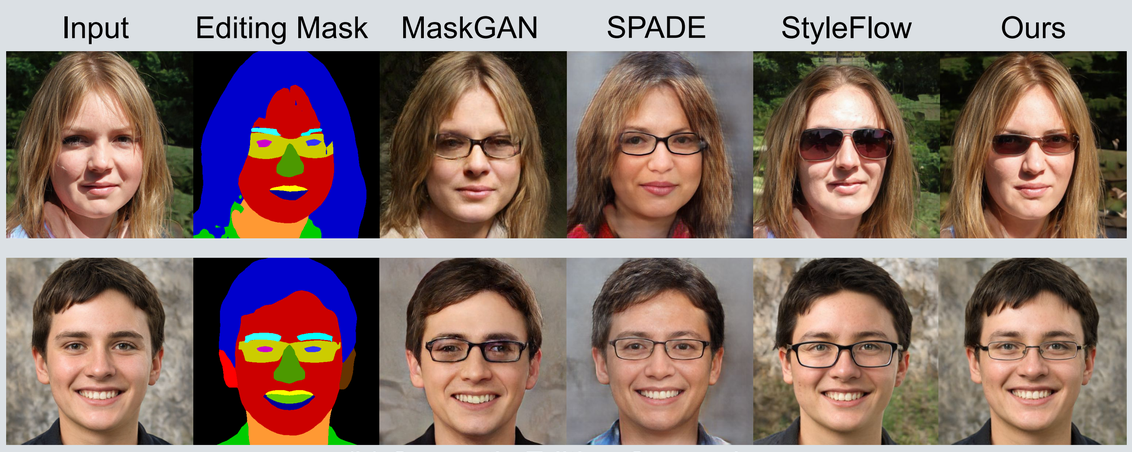
Consequence: Disentanglement of semantics and style enables us to perform semantic and style edits in their respective spaces, leading to more controlled and disentangled edits.

Editing with S^2–Flow

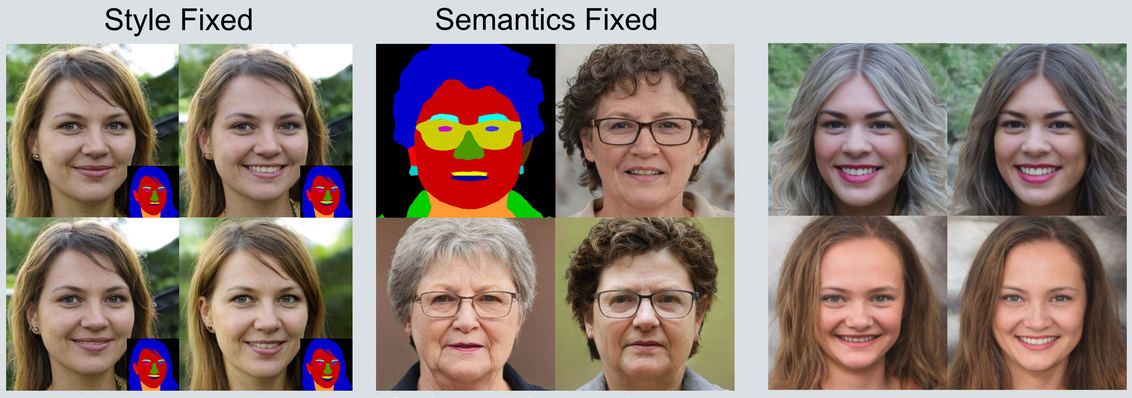




(a) Semantic Editing

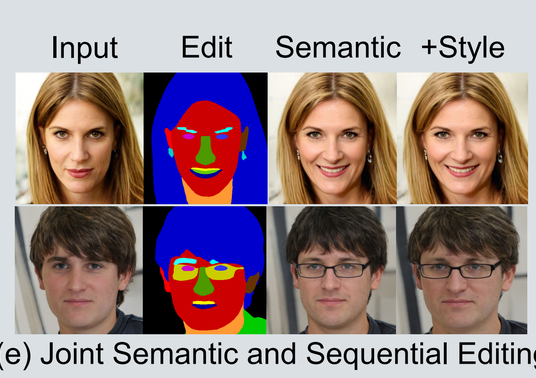


(b) Semantic Editing Comparison

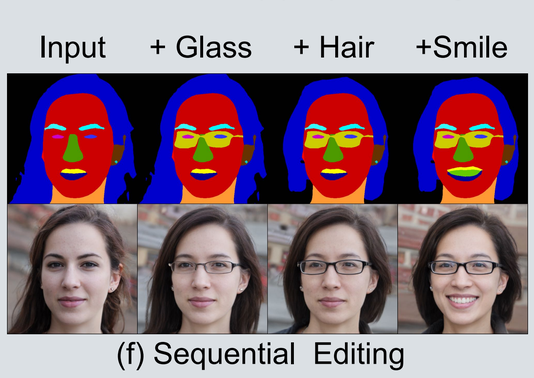


(c) Disentanglement of Style and Semantics

(d) Style Editing



(e) Joint Semantic and Sequential Editing



(f) Sequential Editing

References & Disclosure of Funding

[1] Abdal et al. (2021). "StyleFlow: Attribute-conditioned exploration of StyleGAN-generated images using conditional continuous normalizing flows." In: *ACM ToG*

[2] Lee et al. (2020). "MaskGAN: Towards diverse and interactive facial image manipulation." In: *CVPR*

[3] Park et al. (2019). "Semantic image synthesis with spatially-adaptive normalization." In: CVPR

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 866008). This research was supported by "The Adaptive Mind", funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art. The project has also been supported in part by the State of Hesse through the cluster project "The Third Wave of Artificial Intelligence (3AI)".

Take Home

- Transformation consistencies, namely semantics and style, help in disentangling the latent space.
- Disentangling the latent space of a StyleGAN allows for more controlled semantic and style edits.
- Our model is one of the first to enable controlled joint semantic and style editing within the same framework.

Code