S²–Flow: Joint Semantic and Style Editing of Facial Images

Krishnakant Singh¹
Simone Schaub-Meyer¹,²
Stefan Roth¹,²

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.

<table>
<thead>
<tr>
<th>Method</th>
<th>Conditional</th>
<th>Unconditional</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic</td>
<td>✔</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Style</td>
<td>✔</td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>

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.

Framework of S²–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 \( \tilde{w}^{sm} \) to output an edited latent code \( \tilde{w} \).

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
  \[ L_{sm} (\tilde{m}, m) \]
- **Style consistency loss**: Edits made in the semantic domain should not affect the style of the generated image
  \[ L_{sty} (\tilde{w}, w^{sty}) \]

References & Disclosure of Funding


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. 899208). 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.