**THE ARCHITECTURE**

- **Input:** An image.
- **Output:** A transformed image.

**The Benefits of pSp**

- **Multi-Modal Results**
  - Using a pretrained StyleGAN and performing the translation between images through the StyleGAN inversion, multi-modal conditional image synthesis, face frontalization, inpainting and super-resolution.

**What Can It Do?**

- **Encoding Results**
  - Inherent support for multi-modal synthesis for ambiguous tasks such as image generation from sketches or super-resolution.
  - Better results for non-local translations, as the generator is governed only by the styles with no direct spatial input.
  - Feature maps are first extracted using a standard feature pyramid over a ResNet backbone.
  - The generated 112 vectors are fed into a pretrained StyleGAN which then generates the output image.

**The Architecture**

- **pSp Encoder:** Takes an image as input and encodes it into a latent space.
- **StyleGAN:** Uses a pretrained StyleGAN to manipulate the encoded image.
- **Deeplat Latent:** Represents the generated latent space for output image.

**What Makes pSp Even More Interesting?**

- It can be applied to more general image-to-image translation tasks by directly encoding the input image into the latent code corresponding to the desired output image. Using this technique one can perform image-to-image translation even when the input image cannot be encoded into the latent space of our pretrained StyleGAN.

**THE LOSSES**

\[
\begin{align*}
    L_{id}(y) &= |y - \alpha \cdot f(y)|_1 \\
    L_{cyc}(y) &= \|f(y) - f(y_0)|_1 \\
    L_{id}(y) &= \|y - \alpha \cdot f(y)|_1 \\
    L_{cyc}(y) &= \|f(y) - f(y_0)|_1 \text{.}
\end{align*}
\]

A curated set of losses allows pSp to learn accurate encodings. \(L_{id}\) is helpful to prevent memorization on top of the pixel-wise \(L_1\) loss. We found \(L_{cyc}\) to be important for preserving identity in facial images. \(L_{id}\) serves as a regularization for training, similarly to the translation trick in StyleGAN.

**References**