Inverting Facial Embeddings with GANs

**Introduction**

- We seek to recreate an image of a face from only its “identity”, a 128-dimensional vector generated using FaceNet.
- We take the identity of a person’s face as an input and find a latent vector generating a face with a similar identity.
- Given just an identity, our algorithm generates images that fool FaceNet, a state-of-the-art facial recognition system.

**Data**

Examples from FFHQ (training StyleGAN):

We preprocessed data by running face detection to crop and resize images.

**Method**

- We use greedy search algorithm boosted by a large, pre-generated dataset.
- We iteratively narrow down our search by repeatedly adding noise to the previous best latent to find a better one, judged by the FaceNet loss on its image.
- Decrease noise variance over time.

**Results**

Our method presents a significant improvement over previous efforts to recreate faces from embeddings. By basing our face generation on NVIDIA's state-of-the-art StyleGAN, we are able to create images that are both convincingly real and similar in identity to the original face.

- Ours
  - Real
  - Recreated
- Previous
  - Real
  - Recreated

**Future Work**

- Look for better search algorithms that can perform the search for the desired latent vector more efficiently.
- Apply re-sampling and then perform gradient descent to explore other approaches to recovering the desired latent vector.
- Plenty of other potential avenues as this is a relatively unexplored research area.

**Architecture**

Architecture used for traditional models vs architecture used by the StyleGAN model [3].

**References**