Ensembling with Deep Generative Views
Lucy Chai, Jun-Yan Zhu, Eli Shechtman, Phillip Isola, Richard Zhang
https://chail.github.io/gan-ensembling/

Overview
- Pretrained GAN creates "views" of an original input image
- Ensemble views for classification
- Balance original and GAN views, due to classifier sensitivity to GAN artifacts

Style-mixing
- Swaps a random latent code at specified generator layers
- Perturb coarse or fine layers of StyleGAN2 (middle layers change identity)
- Only using GAN views can decrease accuracy
- Coarse layers change shape, fine layers change coloring
- Style-mixing causes larger visual changes compared to isotropic or PCA perturbations

Perturb optimized latent code in StyleGAN2 W+ space
- Encode and optimize to find image latent code, inversion required be fast and accurate

Properties of GAN Ensembling
- Showing average over 40 CelebAHO binary attributes
- Increasing GAN ensemble size improves classification but saturates
- GAN style-mixing and traditional image cropping perform similarly
- Combining traditional and GAN views offers additional improvement

Creating GAN-generated Image Variants
- $w^* = \arg\min_w L_{\text{img}}(x, G(w)) + \lambda L_{\text{latent}}(w, E(x))$

Adjusting the Ensemble Weight
- Finetuning the classifier on GAN outputs improves ensembling benefit
- Style-mixing perturbation at test time is best

Limitations of Ensembling with GANs
- GAN reconstruction capability: must preserve discriminative attributes
- GAN inversion efficiency: optimization time and reconstruction tradeoff
- Classifier sensitive to GAN artifacts
- Currently limited to simple tasks and small structured datasets
- Similar techniques may yield greater benefits in the future as GANs improve

Visualizing GAN Outputs
- Isotropic Coarse
- Isotropic Fine
- PCA Coarse
- PCA Fine
- Style-mix Coarse
- Style-mix Fine