Computer Science 12 Final Project: StarGAN Training

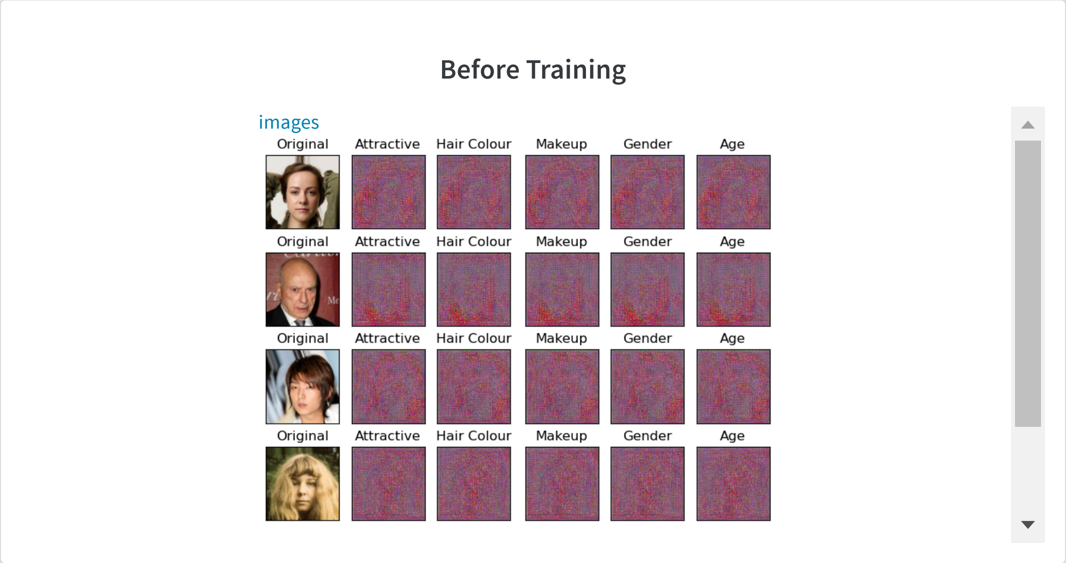
Training an implementation of the StarGAN model from the paper https://arxiv.org/abs/1711.09020 in PyTorch Lightning, starting from pretrained weights.

Stevan Zhuang

Overview

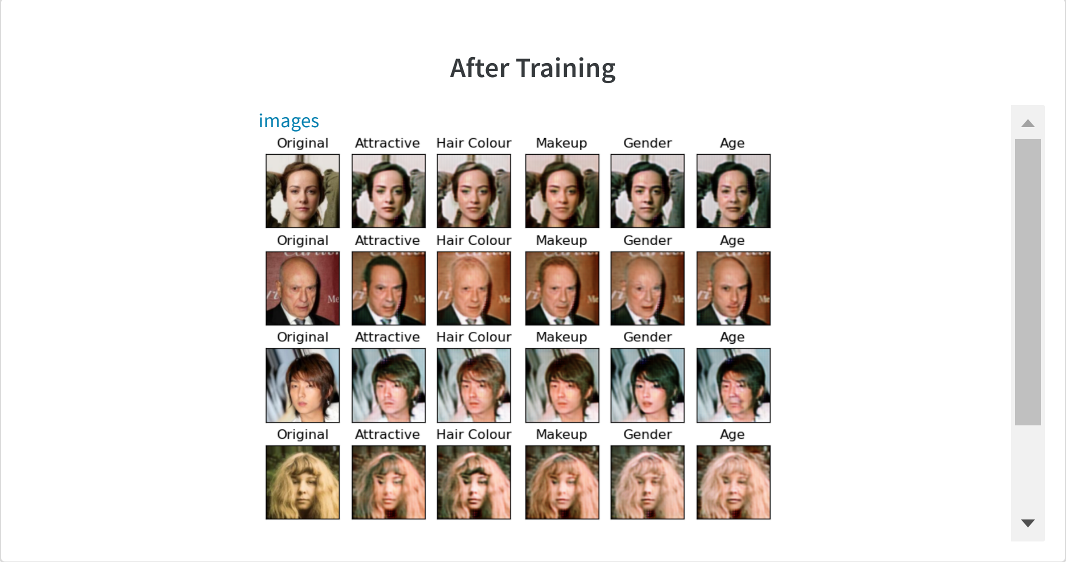
My goal was to implement and train the StarGAN model on a set of labels from the CelebA dataset. To accomplish this, I took three different approaches, all involving training from a set of pretrained weights made available by the publishers of the StarGAN paper. Training the model from scratch was out of the question, as I was running on my home computer's CPU, as well as on a hard deadline. Roughly 24 hours of training was spent on each approach.

When freezing all layers of the generator model and discriminator model except for the generator and discriminator's output layer, which were both replaced, the model simply relearned the connections from the old weights. Neither the generator nor discriminator were able to generalize for new labels, and resulted in similar images produced by the original pretrained weights.



Before Training

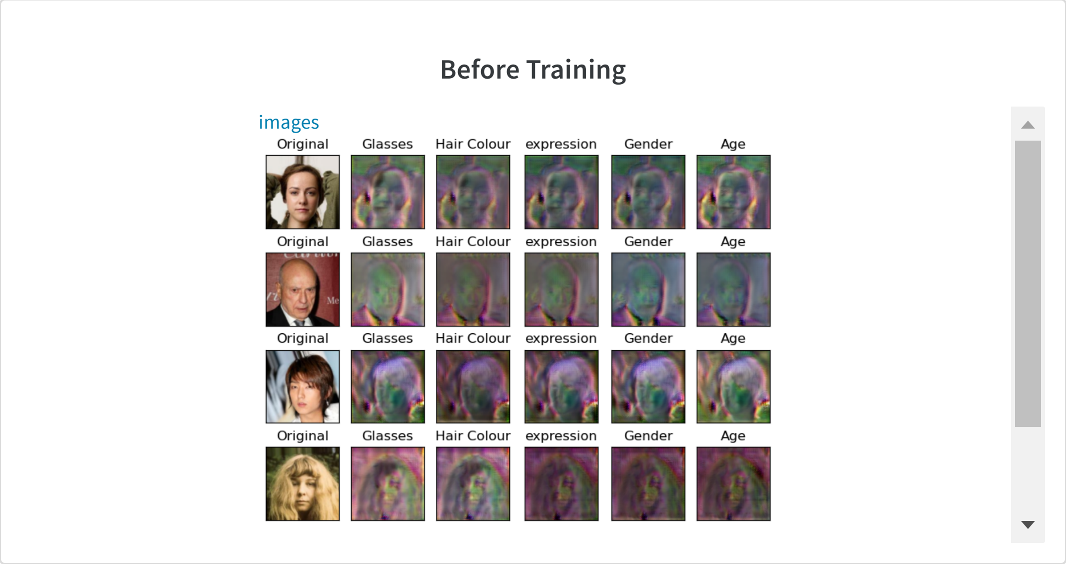
images
Original Attractive Hair Colour Makeup Gender Age
Original Attractive Hair Colour Makeup Gender Age
Original Attractive Hair Colour Makeup Gender Age
Original Attractive Hair Colour Makeup Gender Age

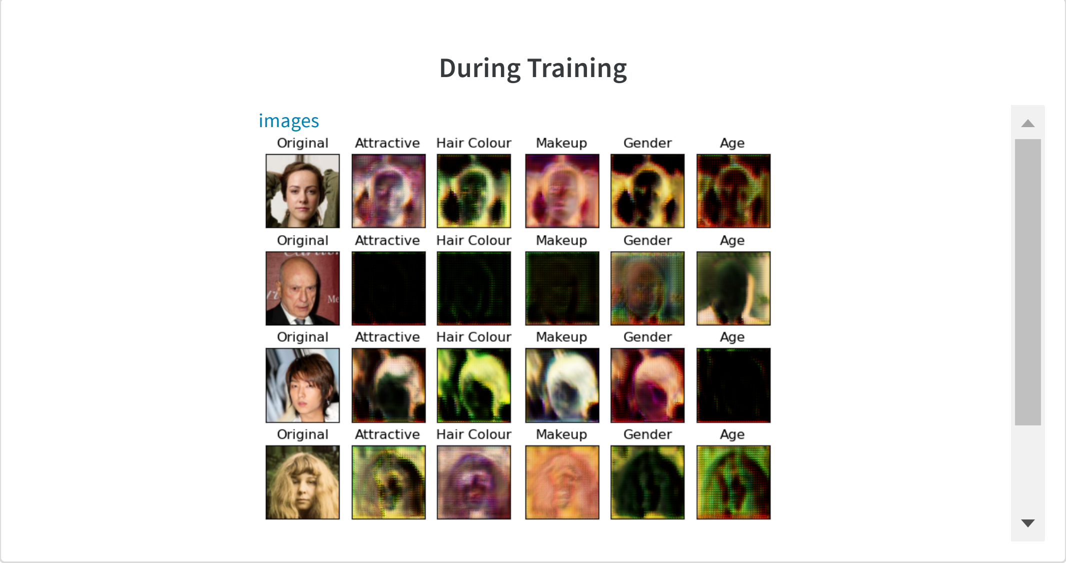


After Training

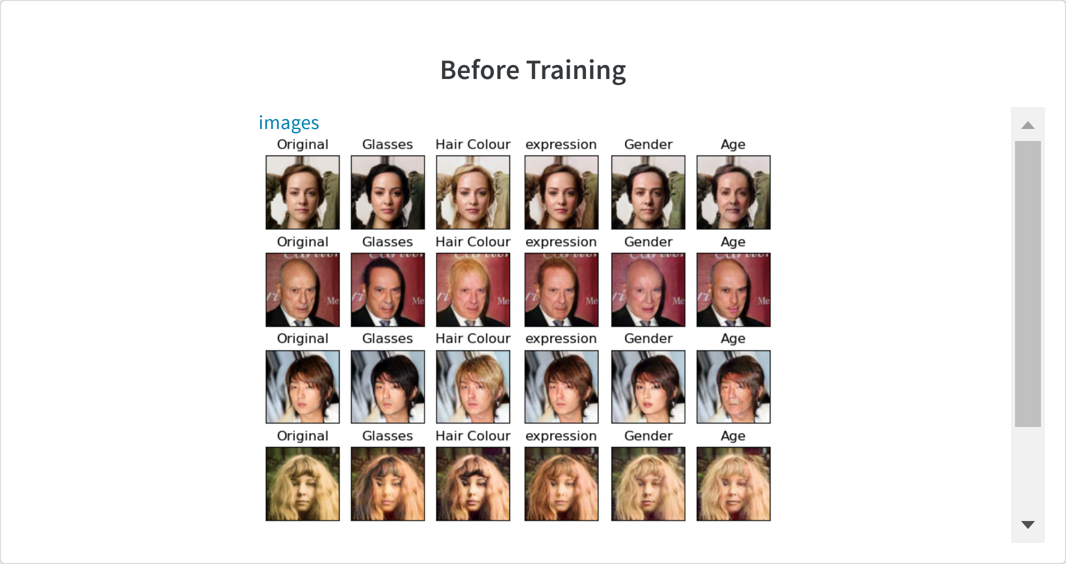
images
Original Attractive Hair Colour Makeup Gender Age
Original Attractive Hair Colour Makeup Gender Age
Original Attractive Hair Colour Makeup Gender Age
Original Attractive Hair Colour Makeup Gender Age

When freezing all layers of the generator model and discriminator model except for the generator's input layer and discriminator's output layer, which were both replaced, the model failed to converge. The generator became stuck on producing images with the colours of features inverted.


Before Training

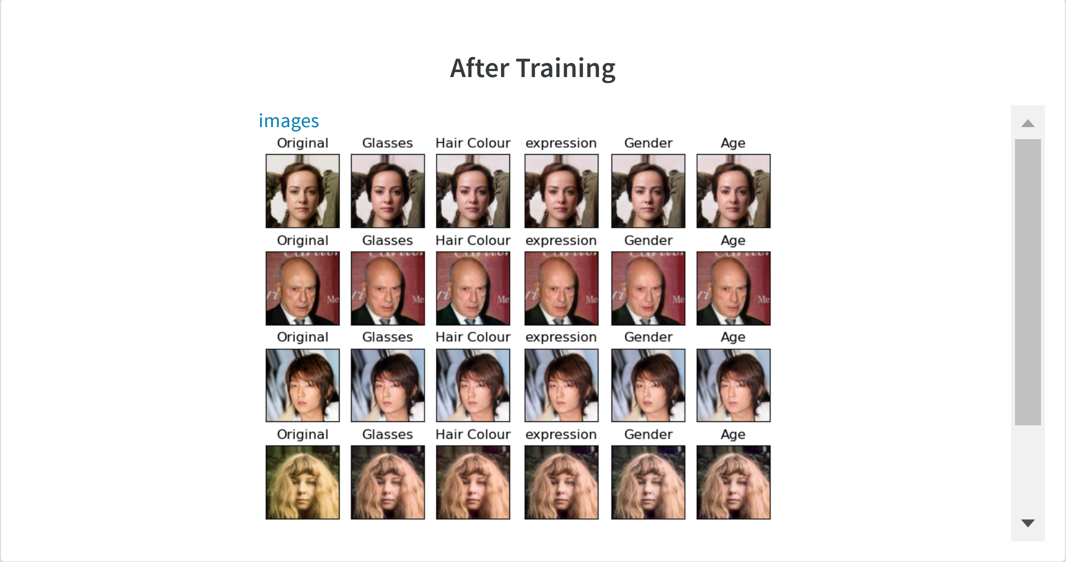

During Training

When fine tuning the pretrained weights, the model became stuck on only recreating the input images which little change, regardless of of the labels. Given much more training, the model likely would have converged to accurately translate the images with attributes changed.



Before Training

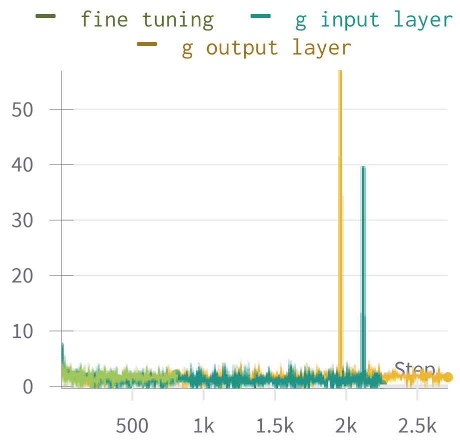
images
Original Glasses Hair Colour expression Gender Age
Original Glasses Hair Colour expression Gender Age
Original Glasses Hair Colour expression Gender Age
Original Glasses Hair Colour expression Gender Age

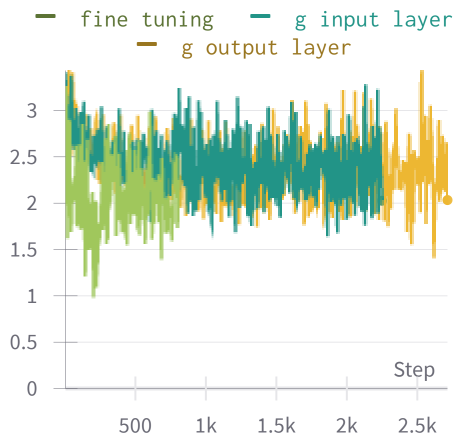


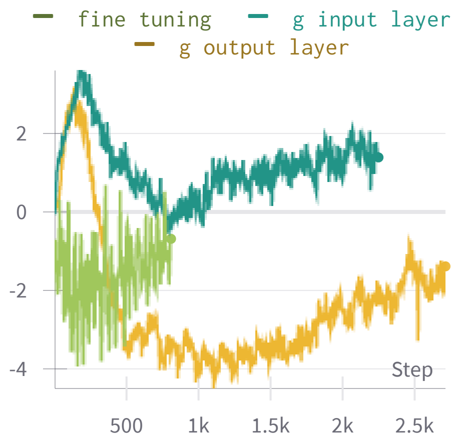
After Training

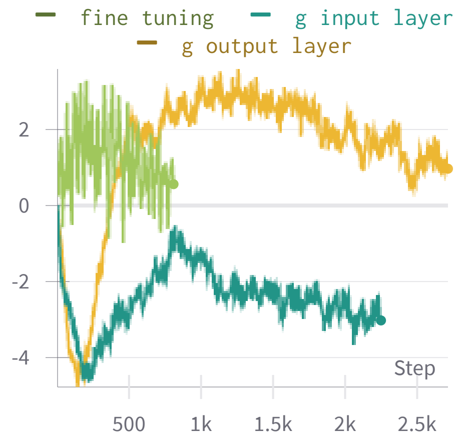
images
Original Glasses Hair Colour expression Gender Age
Original Glasses Hair Colour expression Gender Age
Original Glasses Hair Colour expression Gender Age
Original Glasses Hair Colour expression Gender Age

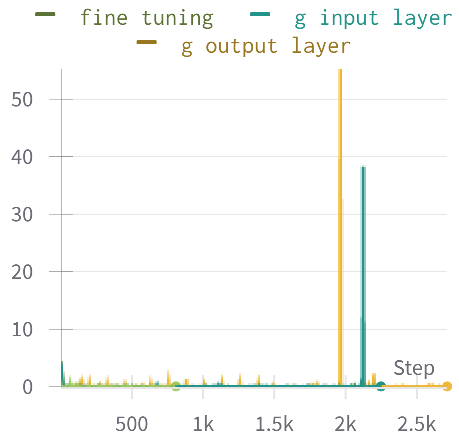
Discriminator Losses


discriminator loss

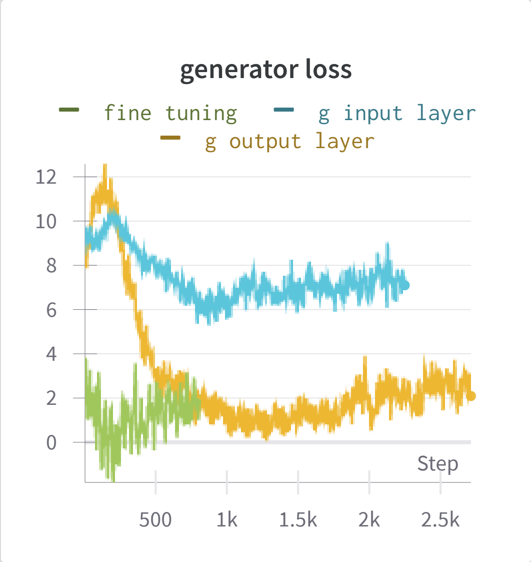

discriminator classification loss

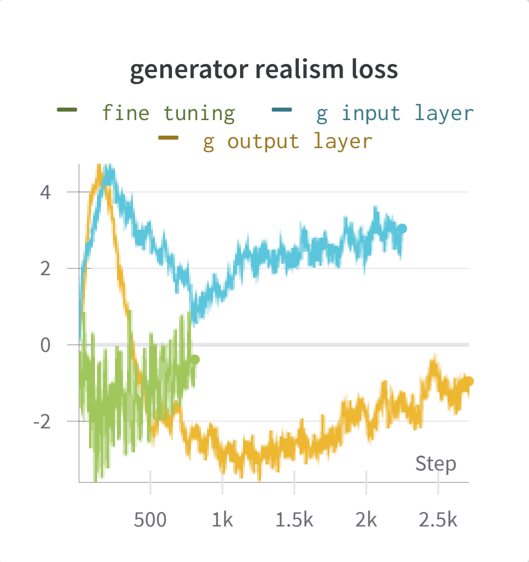

discriminator real images loss

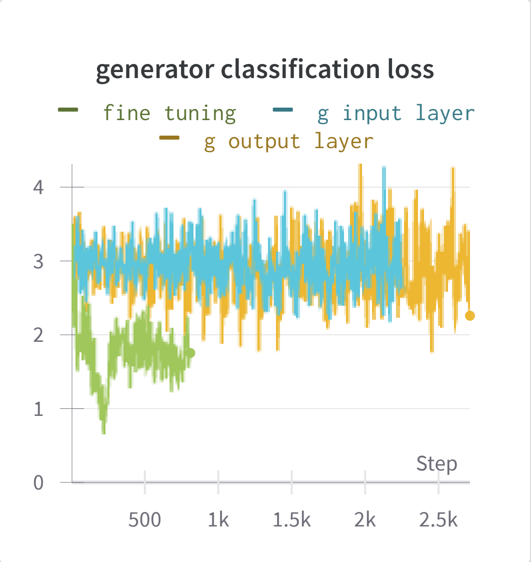

discriminator fake images loss

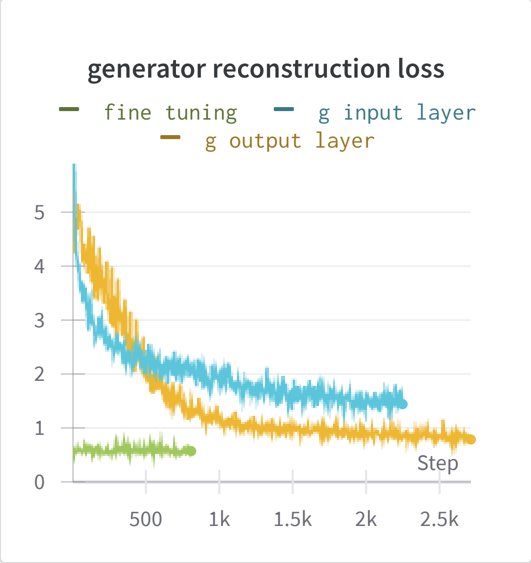

discriminator gradient loss

Generator Losses


generator loss


generator realism loss


generator classification loss


generator reconstruction loss

Created with ♡ on Weights & Biases.

https://wandb.ai/stevan-zhuang/Image%20Domain%20Transfer%20GAN/reports/Computer-Science-12-Final-Project-StarGAN-Training--Vmlldzo1NTQ2MzY