Image Super-Resolution for Facial Recognition

Corey Shih
ctshih@stanford.edu

Abstract and Motivation

Facial recognition systems are becoming increasingly common; however, low-resolution or blurry images can severely detract from the performance of such technology. Using deep learning to produce accurate estimates of high-resolution images from low-resolution ones would thus greatly enhance the performance of facial recognition systems. I implemented a conditional generative adversarial network (cGAN) for the task of super-resolution on images of human faces. The resulting images a bit clearer than images obtained from bicubic interpolation, but still leave much room for improvement.

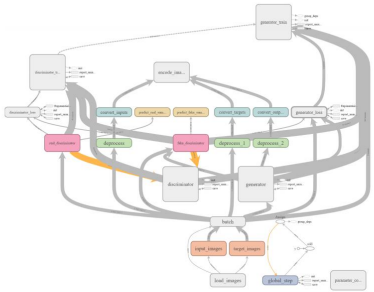
Data and Features

Data was taken from the Labeled Faces in the Wild (LFW) database, which contains 13233 color images of 5749 unique people. Images from the database are 250 × 250 in resolution, and were downscaled to 25 × 25 and 50 × 50 using nearest-neighbor interpolation in TensorFlow. The low resolution images were used as inputs to the model, and the original resolution images from the database were used as the ground-truth images. Data was split into 10586 training and 2647 test images.

250 × 250 50 × 50 25 × 25

Models

I implemented a cGAN using the pix2pix architecture by Isola et al. The model trains a generator to generate 250 × 250 outputs from the 25 × 25 or 50 × 50 inputs, and a discriminator to classify how probable the input/output and input/target pairs look. Both models were trained for 10 epochs.



Results

Results (cont.)

input output target

Discussion and Future Work

Despite an increase in resolution, generated images were still relatively blurry and occasionally featured distortion or artifacts. This could in part be due to the L1 loss term in the generator loss, which for the task of super-resolution pushes generated images to be blurrier. Adjusting the weight of this term in the loss function could potentially improve results. Additional data preprocessing such as deep funneling might help, as the model appears to perform more poorly on angled or tilted faces.

GANs can be notoriously hard to train, and hyperparameter tuning proved to be difficult. Tuning certain parameters led to the discriminator becoming too powerful, and leaving the generator without useful gradients to train on. Adding Gaussian noise to the data could help with this issue.

References

1) Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. University of Massachusetts, Amherst, Technical Report 07-49, October, 2007.
2) Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. Image-to-Image Translation with Conditional Adversarial Networks. arXiv 2016.