Using GANs for new designs of chairs

Matias Castillo - matiasct@stanford.edu

Motivation and Objective

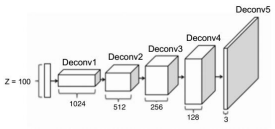
- Most successful image generation results using GANs have been achieved using huge datasets (over 100,000 images) and generating images with very similar high level features.
- If we could use GANs for generating new images of specific objects, we could use them to create new designs.

Project objective:

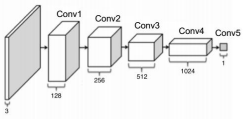
- Try to generate images of individual complex objects using a small dataset and generative adversarial techniques.
- Analyze how transferable are the features learned from large datasets into the generation of specific objects.

GAN implementation: DC-GAN

Generator:



Discriminator:



- Batch Normalization in most layers in both G and D.
- Convolutions with stride instead of max pooling layers
- ReLU for all hidden layers in G and Tanh for output.
- Leaky ReLU for all hidden layers in D and Sigmoid for output.

- **Binary Cross-Entropy Loss**, with non-saturating cost function for the generator:

$$J^{(D)} = -\frac{1}{m_d}\sum_{i=1}^{m_d} y_d^{(i)}.\log(D(x^{(i)})) - \frac{1}{m_g}\sum_{i=1}^{m_g}(1-y_g^{(i)}).\log(1-D(G(z^{(i)})))$$

$$J^{(G)} = -\frac{1}{m_g}\sum_{i=1}^{m_g}\log(D(G(z^{(i)})))$$

Model validation in large datasets

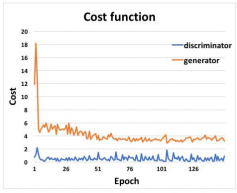
- The model implementation was first trained with 60,000 images from the MNIST dataset and also with the 1 million images from Imagenet.

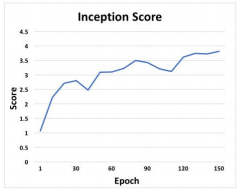
Chairs dataset collection & augmentation

- 2,300 images of chairs downloaded from Google Images, with different styles, sizes, colors and orientations.
- All images were downsized to 64x64 and used for training.
- Data augmentation: horizontal flip randomly while training.

Results

- Graph of cost function over training shows how both D and G are improving at similar rates.
- Inception score (30,000 images, Inception_v3) shows an improvement in generating more "realistic" images with the training.





Results (cont.)

- We can see how the model learned the main features of chairs.
- Apparently, the model did not collapse: diversity of colors, styles and orientations is maintained.

Improvements

Transfer learning from Imagenet:

- Retraining all parameters: discriminator learned too fast, generator was not able to improve (even restricting D with a threshold).
- Retraining last layers: discriminator also learned too fast.
- Retraining last layers of generator only: D and G were able to learn at similar rates extracting the main features of chairs, but the result was worse than without pretraining.

Bigger dataset, more data augmentation and longer training:

- Added 1,000 images to the training.
- Applied horizontal flip augmentation.

Future Work

- Try other improvement techniques like feature matching and virtual BN.
- Try bigger dataset and higher resolution for further improvement.
- Try transfer learning with more similar datasets (LSUN) and after longer pretraining.
- Try to create specific designs using conditional GAN.