Generative Adversarial Networks

Ian Goodfellow, Staff Research Scientist, Google Brain

ICCV Tutorial on GANs

Venice, 2017-10-22
Generative Modeling

• Density estimation

• Sample generation

Training examples  Model samples
Maximum Likelihood

$$\theta^* = \arg \max_\theta \mathbb{E}_{x \sim p_{\text{data}}} \log p_{\text{model}}(x \mid \theta)$$
Adversarial Nets Framework

$D(x)$ tries to be near 1

Differentiable function $D$

$x$ sampled from data

$D$ tries to make $D(G(z))$ near 0,
$G$ tries to make $D(G(z))$ near 1

$D$ tries to make $D(G(z))$ near 0,
$G$ tries to make $D(G(z))$ near 1

$x$ sampled from model

Differentiable function $G$

Input noise $z$

(Goodfellow et al., 2014)
What can you do with GANs?

- Simulated environments and training data
- Missing data
  - Semi-supervised learning
- Multiple correct answers
- Realistic generation tasks
  - Simulation by prediction
- Solve inference problems
- Learn useful embeddings
Apple’s first research paper tries to solve a problem facing every company working on AI
GANs for simulated training data

(Shrivastava et al., 2016)
GANs for domain adaptation

(Bousmalis et al., 2016)
What can you do with GANs?

- Simulated environments and training data
- Missing data
  - Semi-supervised learning
- Multiple correct answers
- Realistic generation tasks
- Simulation by prediction
- Solve inference problems
- Learn useful embeddings
Generative modeling reveals a face

(Yeh et al., 2016)
What can you do with GANs?

• Simulated environments and training data
• Missing data
  • Semi-supervised learning
• Multiple correct answers
• Realistic generation tasks
• Simulation by prediction
• Solve inference problems
• Learn useful embeddings
Supervised Discriminator

(Odena 2016, Salimans et al 2016)
What can you do with GANs?

• Simulated environments and training data
• Missing data
  • Semi-supervised learning
• Multiple correct answers
• Realistic generation tasks
• Simulation by prediction
• Solve inference problems
• Learn useful embeddings
Next Video Frame Prediction

Ground Truth

What happens next?

(Lotter et al 2016)

(Goodfellow 2017)
CHAPTER 15. REPRESENTATION LEARNING

Figure 15.6: Predictive generative networks provide an example of the importance of learning which features are salient. In this example, the predictive generative network has been trained to predict the appearance of a 3-D model of a human head at a specific viewing angle.

(Left) Ground truth. This is the correct image, that the network should emit.

(Center) Image produced by a predictive generative network trained with mean squared error alone. Because the ears do not cause an extreme difference in brightness compared to the neighboring skin, they were not sufficiently salient for the model to learn to represent them.

(Right) Image produced by a model trained with a combination of mean squared error and adversarial loss. Using this learned cost function, the ears are salient because they follow a predictable pattern. Learning which underlying causes are important and relevant enough to model is an important active area of research. Figures graciously provided by Lotter et al. (2015).

A benefit of learning the underlying causal factors, as pointed out by Schölkopf et al. (2012), is that if the true generative process has $x$ as an effect and $y$ as a cause, then modeling $p(x|y)$ is robust to changes in $p(y)$. If the cause-effect relationship was reversed, this would not be true, since by Bayes’ rule, $p(x|y)$ would be sensitive to changes in $p(y)$. Very often, when we consider changes in distribution due to different domains, temporal non-stationarity, or changes in the nature of the task, the causal mechanisms remain invariant (the laws of the universe are constant) while the marginal distribution over the underlying causes can change. Hence, better generalization and robustness to all kinds of changes can be expected.

Next Video Frame Prediction

Ground Truth

MSE

Adversarial

(Lotter et al 2016)
Next Video Frame(s)

Prediction

Mean Squared Error
Mean Absolute Error
Adversarial

(Mathieu et al. 2015)
What can you do with GANs?

• Simulated environments and training data
• Missing data
  • Semi-supervised learning
• Multiple correct answers
  • Realistic generation tasks
• Simulation by prediction
• Solve inference problems
• Learn useful embeddings
Which of these are real photos?

(work by vue.ai covered by Quartz)
What can you do with GANs?

- Simulated environments and training data
- Missing data
  - Semi-supervised learning
- Multiple correct answers
- Realistic generation tasks
- Simulation by prediction
- Solve inference problems
- Learn useful embeddings
Vector Space Arithmetic

Figure 15.9: A generative model has learned a distributed representation that disentangles the concept of gender from the concept of wearing glasses. If we begin with the representation of the concept of a man with glasses, then subtract the vector representing the concept of a man without glasses, and finally add the vector representing the concept of a woman without glasses, we obtain the vector representing the concept of a woman with glasses. The generative model correctly decodes all of these representation vectors to images that may be recognized as belonging to the correct class. Images reproduced with permission from Radford et al. (2015).

A common is that one could imagine learning about each of them without having to see all the configurations of all the others. Radford et al. (2015) demonstrated that a generative model can learn a representation of images of faces, with separate directions in representation space capturing different underlying factors of variation. Figure 15.9 demonstrates that one direction in representation space corresponds to whether the person is male or female, while another corresponds to whether the person is wearing glasses. These features were discovered automatically, not fixed a priori. There is no need to have labels for the hidden unit classifiers: gradient descent on an objective function of interest naturally learns semantically interesting features, so long as the task requires such features. We can learn about the distinction between male and female, or about the presence or absence of glasses, without having to characterize all of the configurations of the other features by examples covering all of these combinations of values. This form of statistical separability is what allows one to generalize to new configurations of a person’s features that have never been seen during training.

(Radford et al, 2015)
How long until GANs can do this?

Training examples

Model samples
AC-GANs

monarch butterfly  goldfinch  daisy

(Odena et al., 2016)
Track updates at the GAN Zoo

https://github.com/hindupuravinash/the-gan-zoo
Questions?