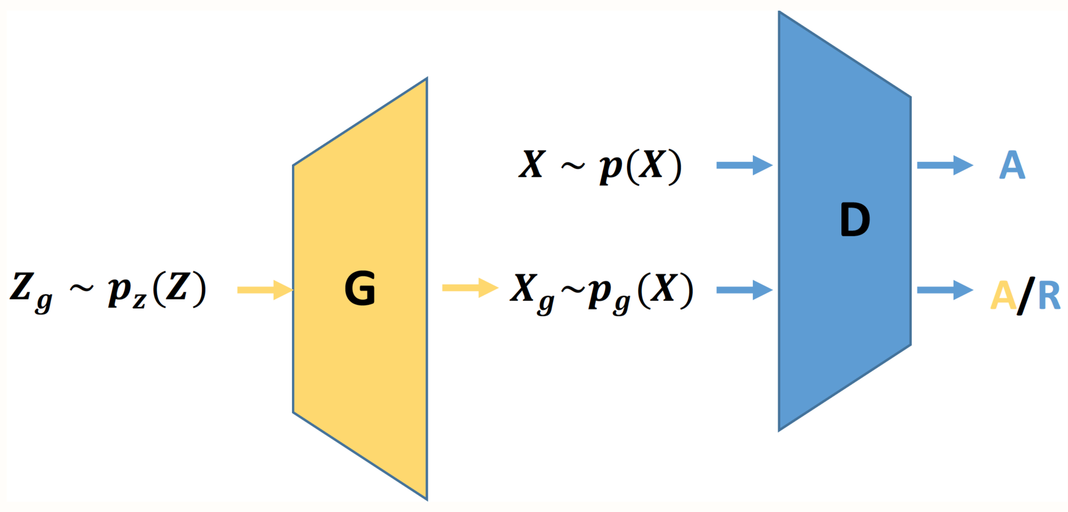
Triple Generative Adversarial Nets

Chongxuan Li, Kun Xu, Jun Zhu and Bo Zhang
Department of Computer Science and Technology, Tsinghua University
{licx14, xu-k16}@mails.tsinghua.edu.cn

Background and Motivation

Generative Adversarial Nets(GANs)

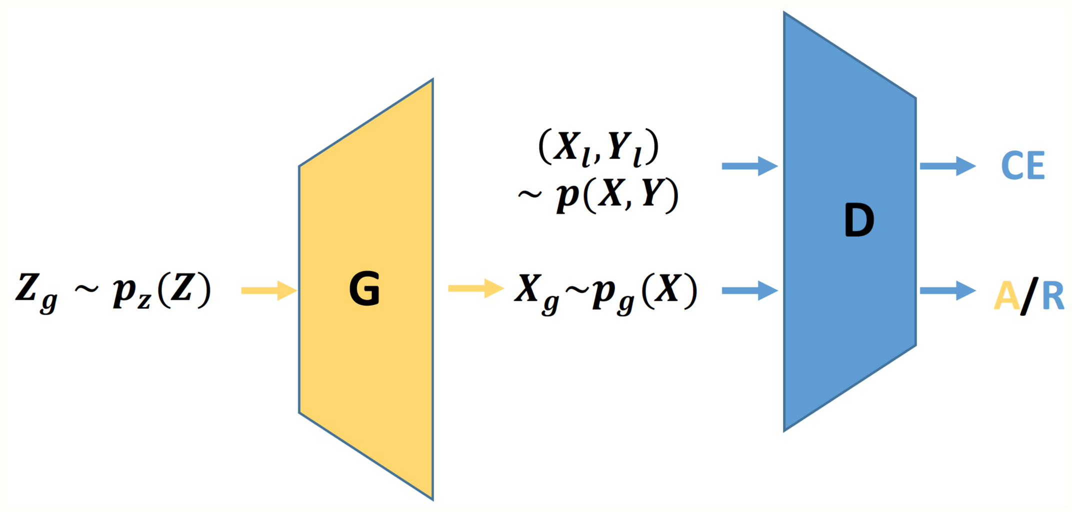
Given samples from $p(x)$,to learn $p_g(x) \approx p(x)$. $D(x)$ outputs a binary unit.



Under nonparametric assumptions, the global optimum is $p_g(x) = p(x)$.

Semi-Supervsied GANs

Given partially labeled data from $p(x, y)$, to learn $p_c(y|x)$ and $p_g(x)$. $D(x)$ outputs $|\mathcal{Y}|$ or $|\mathcal{Y}| + 1$ dimensional units.



No game-theoretical analysis as in original GANs.

Problems of Two-Player Formulation

- An extreme case, where G is good but C is poor [1]
  - G converges, i.e. $p_g(x) = p(x)$
  - D^*_G rejects $x \sim p_g(x)$ with probability $\frac{1}{2}$
  - C^* should accept it and classify it as certain class confidently
  - Then C and G won't be both optimal because of sharing parameters between D and C
- Cannot control what to generate
  - The classifier predicts accurately while the generator cannot leverage such information

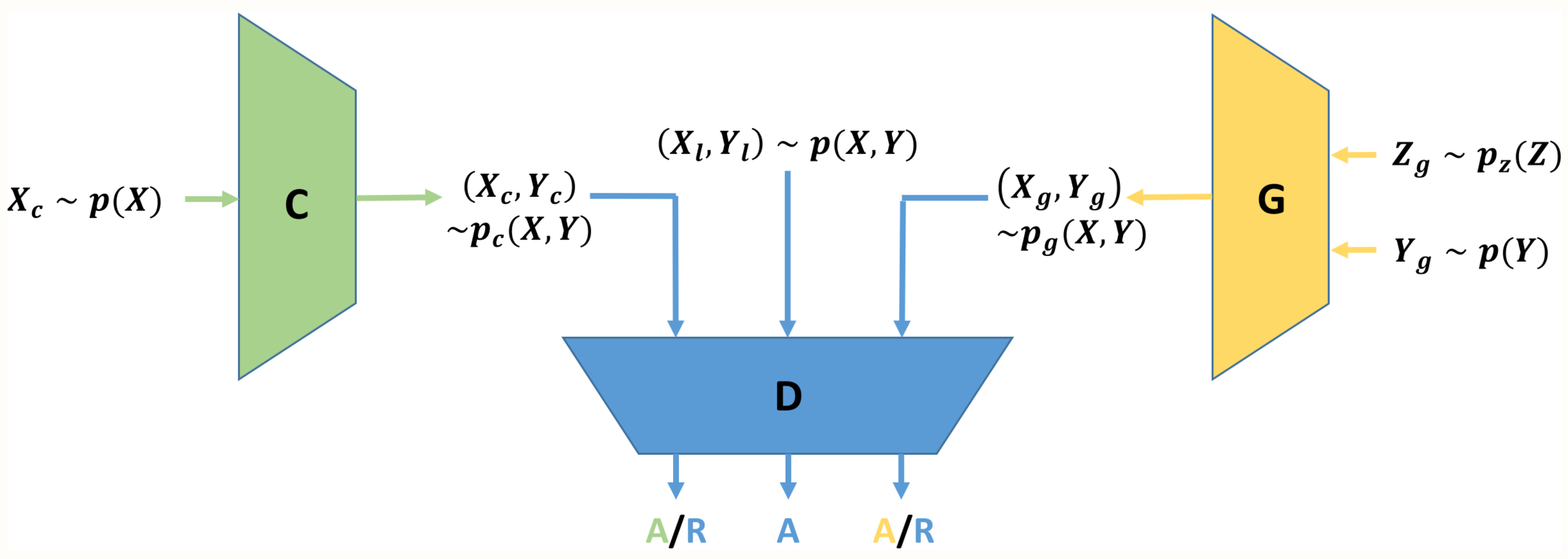
Triple-GAN

Rethink the Problem

- A dual perspective: $p(x, y) = p(y)p(x|y) = p(x)p(y|x)$
- Marginal distributions are known
- Conditional distributions are of interests

Three-Player Game

- Given partially labeled data, learn $p_c(y|x)$ and $p_g(x|y)$



where $D(x)$ outputs a binary unit to justify (x, y) pairs. A minimax game:

$$\min_{C,G}\max_{D} E_p[\log D(x,y)] + \alpha E_{p_c}[\log(1 - D(x,y))] + (1-\alpha)E_{p_g}[\log(1 - D(x,y))].$$

Game-Theoretical Analysis and Regularizations

- The optimum is $p(x, y) = (1 - \alpha)p_g(x, y) + \alpha p_c(x, y)$.
- G and C will benefit each other.
- With cross entropy losses on labeled data and generated data, the unique optimum is $p(x, y) = p_g(x, y) = p_c(x, y)$.

Experiments

Classification

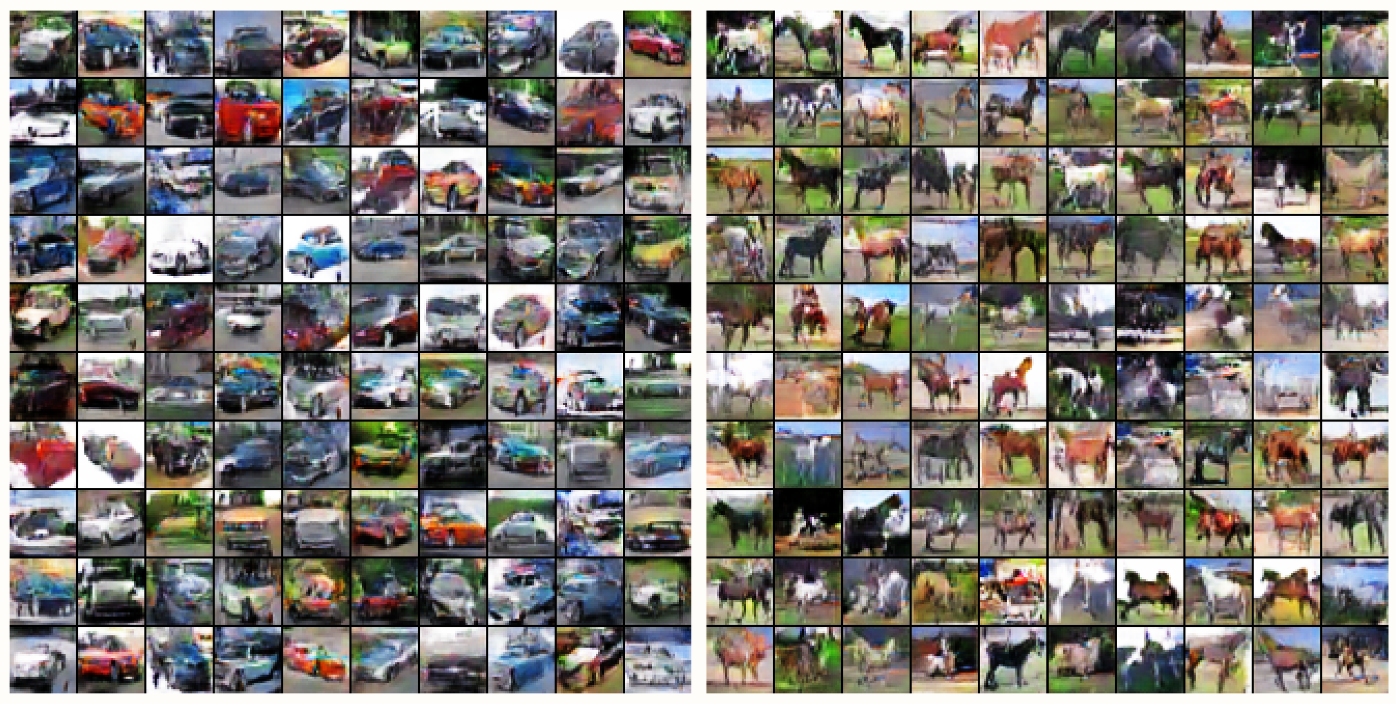
State-of-the art classification results on MNIST, SVHN and CIFAR10.

Generation

Good C results in good G

(a) Improved-GAN (FM) (b) Triple-GAN

Conditional Generation



(c) Automobile (d) Horse

Disentangle Classes and Styles

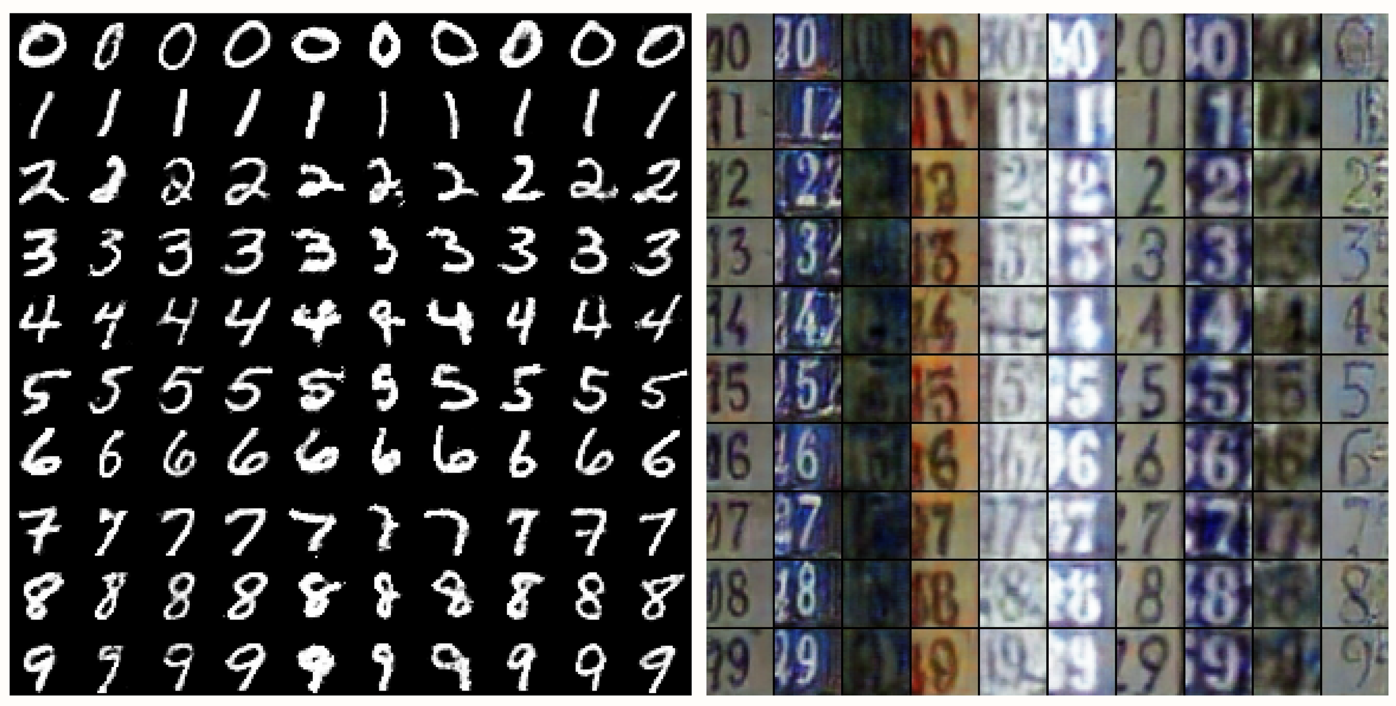


Figure 1: Same y for each row. Same z for each column.

Latent Space Interpolation

References [1] Salimans T, Goodfellow I, Zaremba W, et al. Improved techniques for training gans[C]//Advances in Neural Information Processing Systems. 2016: 2234-2242.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, US