GANs

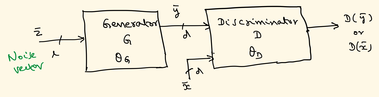
Saturday, 2 May 2020 09:48

Generative Adversarial Networks (GANs)

GAN: Generative model with adversarial learning

Purpose of GAN:

To synthesize examples that are similar to the real examples



Input to GAN: $\bar{z} \sim N(\bar{0}, \sigma^2 I)$

Synthesized example: $\bar{g} \sim p(\bar{g})$

Real example: $\bar{x} \sim p(\bar{x})$

Objective of GAN: $p(\bar{g}) \approx p(\bar{x})$

Desired output for D:

$D(\bar{x}) = 1$

$D(\bar{g}) = 0$

Training a GAN:

Optimization problem

$$\underset{\theta_D}{\text{Maximize}} \quad \underset{\bar{x} \sim p(\bar{x})}{E}\left[\log D(\bar{x})\right] + \underset{\bar{g} \sim p(\bar{g})}{E}\left[\log(1 - D(\bar{g}))\right]$$

$$\underset{\theta_D}{\text{Maximize}} \quad \underbrace{\underset{\bar{x} \sim p(\bar{x})}{E}\left[\log D(\bar{x})\right]}_{\text{independent of } G} + \underset{\bar{z} \sim N(\bar{0}, \sigma^2 I)}{E}\left[\log(1 - D(G(\bar{z})))\right]$$

$$\underset{\theta_G}{\text{Minimize}} \quad \underset{\bar{g} \sim p(\bar{g})}{E}\left[\log(1 - D(\bar{g}))\right]$$

$$\underset{\theta_G}{\text{Minimize}} \quad \underset{\bar{z} \sim N(\bar{0}, \sigma^2 I)}{E}\left[\log(1 - D(G(\bar{z})))\right]$$

Minmax optimization problem

$$\underset{\theta_G}{\text{Minimize}} \quad \underset{\theta_D}{\text{Maximize}} \quad \underset{\bar{x} \sim p(\bar{x})}{E}\left[\log D(\bar{x})\right] + \underset{\bar{z} \sim N(\bar{0}, \sigma^2 I)}{E}\left[\log(1 - D(G(\bar{z})))\right]$$

Discriminator training using gradient-ascent:

$$\theta_D^{new} = \theta_D^{old} + \eta \, \bar{\nabla}_{\theta_D}$$

Generative training using gradient-descent:

$$\theta_G^{new} = \theta_G^{old} - \eta \, \bar{\nabla}_{\theta_G}$$

Deep Convolutional GAN for Image data:

Discriminator: CNN with one output node
- Two-class classifier

Generator: CNN with transposed convolution
or
CNN with fractional-strided convolution

Transposed Convolution: (Fractionally-Strided Convolution)

Convolutional layer using 2-d kernel

$$\underbrace{W_1 \times H_1}_{\bar{x}} \xrightarrow[S=1,\ P=0]{K:\ F \times F} \underbrace{W_2 \times H_2}_{\bar{g}}$$

$W_2 = W_1 - F + 1$

$H_2 = H_1 - F + 1$

$W_1 = H_1 = 4$

$F = 3$

$W_2 = H_2 = 2$

Kernel, K: $\begin{bmatrix} w_{00} & w_{01} & w_{02} \\ w_{10} & w_{11} & w_{12} \\ w_{20} & w_{21} & w_{22} \end{bmatrix}$

Convolution matrix: $\underset{4 \times 16}{C}$

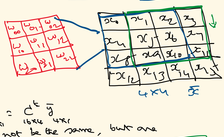
$$C = \begin{bmatrix} w_{00} & w_{01} & w_{02} & 0 & w_{10} & w_{11} & w_{12} & 0 & w_{20} & w_{21} & w_{22} & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{00} & w_{01} & w_{02} & 0 & w_{10} & w_{11} & w_{12} & 0 & w_{20} & w_{21} & w_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{00} & w_{01} & w_{02} & 0 & w_{10} & w_{11} & w_{12} & 0 & w_{20} & w_{21} & w_{22} & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{00} & w_{01} & w_{02} & 0 & w_{10} & w_{11} & w_{12} & 0 & w_{20} & w_{21} & w_{22} \end{bmatrix}$$

Flattened input for 4×4: $\bar{x}$: 16×1

$$\bar{x} = [\underbrace{x_0, x_1, x_2}, x_4, \underbrace{x_5 = x_6}, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}]$$

Flattened output for 2×2: $\bar{g} = 4 \times 1$

$$\bar{g} = C\,\bar{x}$$



$g_0 = \bar{C}_0^t \bar{x}$

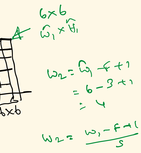
$g_1 = \bar{C}_1 \bar{x}$

$g_2 = \bar{C}_2 \bar{x}$

$g_3 = \bar{C}_3 \bar{x}$

Transposed convolution: $\underset{W \times 1}{\hat{\bar{x}}} = \underset{16 \times 4}{C^t} \underset{4 \times 1}{\bar{g}}$

$\hat{\bar{x}}$ and $\bar{x}$ will not be the same, but are of same size



It can be extended for 3-d kernel:

$$\underbrace{W_1 \times H_1 \times D}_{\bar{x}} \xrightarrow[K:\ F \times F \times D]{} \underbrace{W_2 \times H_2}_{\bar{g}}$$

Alternate method:

Standard convolution

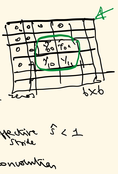
$$\underset{4 \times 4}{X} \xrightarrow[3 \times 3]{K:\ F \times F} \underset{2 \times 2}{Y}$$

$W_2 < W_1$

$H_2 < H_1$

Fractional-strided convolution

$$\underset{2 \times 2,\ W_1 \times H_1}{\hat{Y}} \xrightarrow[3 \times 3]{K} \underset{4 \times 4,\ W_2 \times H_2}{\hat{X}}$$



Pad $\hat{Y}$ with 2 rows and 2 columns of zeros

$$\underset{6 \times 6,\ \hat{W}_1 \times \hat{H}_1}{\bar{Y}} \xrightarrow[3 \times 3]{K} \underset{4 \times 4,\ \hat{W}_2 \times \hat{H}_2}{\bar{X}}$$

Effective stride $\hat{S} < 1$

$W_2 = W_1 - F + 1$

$= 6 - 3 + 1$

$= 4$

$$W_2 = \frac{W_1 - F + 1}{S}$$

$$W_2 = \frac{\hat{W}_1 - F + 1}{\hat{S}}$$

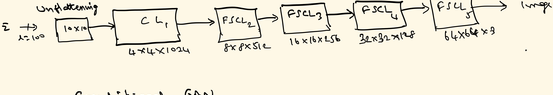
$$\hat{S} = \frac{\hat{W}_1 - F + 1}{W_2}$$

< 1

3-d kernel fractional-strided convolution

$$\underset{6 \times 6 \times 3}{\hat{W}_1 \times \hat{H}_1 \times D} \xrightarrow[3 \times 3 \times D,\ 3 \times 3 \times 3]{K} \underset{4 \times 4}{W_2 \times H_2}$$

Generator in DC GAN



Conditional GAN

Cyclic GAN for Image-to-Image Translation

Text Adaptive GAN for Image Manipulation

GANformer: Combination of principles of GAN and transformer for Image generation

Text GAN for text generation.