The application of GANs in inverse problems and Image to Image translation

Application part

2017-7-14

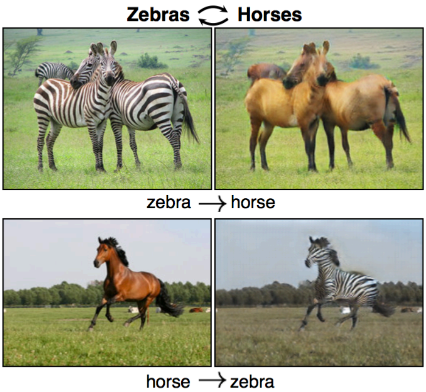
Table

Image to Image Translation

Traget

Condition: Some sampels $\{x_i\}_{i=1}^{N}$, $\{y_j\}_{j=1}^{M}$ from two domains X and Y, in which $x_i \sim P_x$, $Y_j \sim P_y$

Target : find $G : X \to Y$, such that $\forall x \sim P_x$, $G(x) \sim P_y$, while the content of $G(x)$ keeps in consistent with the content of x。

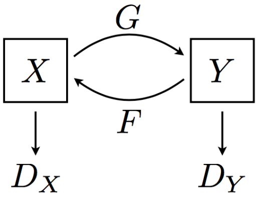


Problems:

GANs are good at transforming the points from P_X to P_y. If contents are wanted to be consistent, vanilla GANs need to be modified。

Vanilla GANs can transform the style but can't keep the content

Architecture[1]



[1] Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

Aiming at:

$$F(G(x)) = x,\ \forall x \in X \quad G(F(y)) = y,\ \forall y \in Y \tag{1}$$

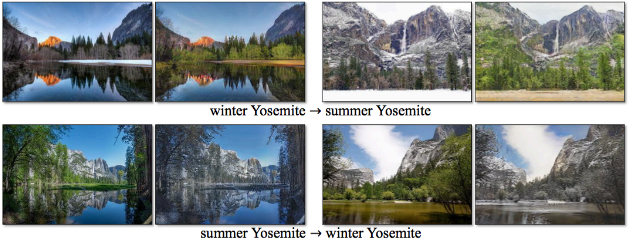
The loss function for Generator:

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_x, Y, X) + \lambda L_{cyc}(G, F) \tag{2}$$

in which:

$$\begin{aligned} L_{GAN}(G, D_Y, X, Y) &= \mathrm{E}_{y \sim P_Y}[\log D_Y(y)] + \mathrm{E}_{x \sim P_x}[\log(1 - D_Y(G(x))] \\ L_{cyc}(G, F) &= \mathrm{E}_{x \sim P_x}[||F(G(x)) - X||_1] + \mathrm{E}_{y \sim P_y}[||G(F(y)) - y||_1] \end{aligned} \tag{3}$$

$$G^*, F^* = \arg \min_{F,G} \max_{D_x, D_y} L(G, F, D_x, D_y) \tag{4}$$


winter Yosemite → summer Yosemite
summer Yosemite → winter Yosemite

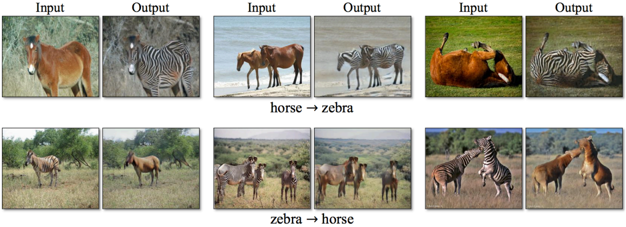

Input
Output
Input
Output
Input
Output
horse → zebra
zebra → horse

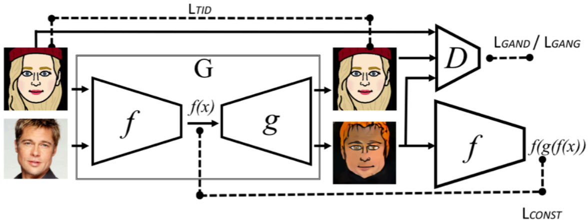
Image to Image Translation

Traget

Given: Some points $\{x_i\}_{i=1}^{N}$, $\{y_j\}_{j=1}^{M}$ from two domains X and Y, in which $x_i \sim P_x$, $y_j \sim P_y$, what't more, another function f, $X \cup Y \in dom(f)$ is given。

Find a transformation $G: X \to Y$, such that $\forall x \sim P_x$, $G(x) \sim P_y$, what's more, $f(x) = f(G(x))$。

Architecture[2]



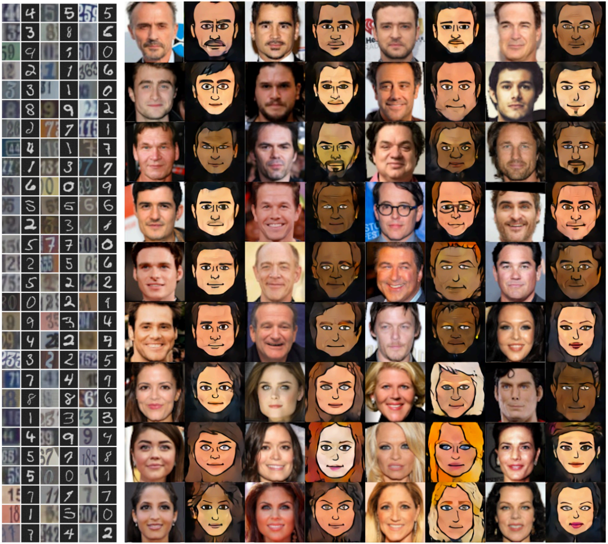
[2] Unsupervised Cross-Domain Image Generation

Baseline:

$$
\begin{aligned}
L_{GAN} &= \max_{D} \mathbb{E}_{x\sim P_X} \log[1 - D(G(x))] + \mathbb{E}_{x\sim P_y} \log[D(x)] \\
L_{const} &= \mathbb{E}_{x\sim P_x} d(f(x), f(G(x)))
\end{aligned} \tag{5}
$$

Improvement:

$$
\begin{aligned}
G = g \circ f \quad & dom(G) = X \cup Y \\
L_{TID} &= \sum_{y\sim P_y} d(y, G(y)) \\
L_{const} &= \sum_{x\sim P_x} d(f(x), f(g(f(x)))) \\
L_D = -\mathrm{E}_{x\sim P_x} \log D_1(g(f(x))) & - \mathrm{E}_{y\sim P_y} \log D_2(g(f(x))) - \mathrm{E}_{y\sim P_y} \log D_3(y) \\
L_{GANG} = -\mathrm{E}_{x\sim P_x} \log D_3(g(f(x))) & - \mathrm{E}_{y\sim P_y} \log D_3(g(f(x))) \\
L_G &= L_{GANG} + \alpha L_{const} + \beta L_{TID}
\end{aligned} \tag{6}
$$



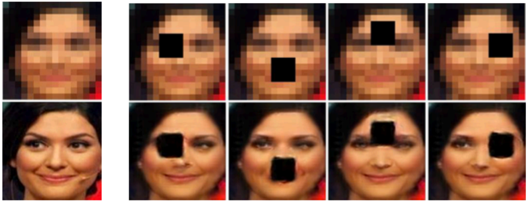
Human Face Super Resolution[3]

The distribution for super resolution images P_H, the distribution for low resolution images P_L, the pair-wised training data $\{l_i, h_i\}$

$$\min_G \max_D F(G, D) = \mathbb{E}_{h_i \sim P_H}[\log D(h_i)] + \mathrm{E}_{l_i \sim P_L}[\log(1 - D(G(l_i)))] + \lambda \mathbb{E}_{\{h_i, l_i\} \sim P_{HL}}[||G(l_i) - h_i||_F^2$$

Contribution:

- reconstructing 64 pixels from one pixel
- high efficiency for reconstrunction
- can deal with misalignment, occlusion



(a)LR,HR (b) UR-DGN

[3] Ultra-Resolving Face Images by Discriminative Generative Networks

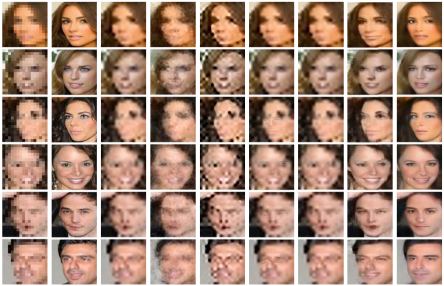


图 1: unaligned faces

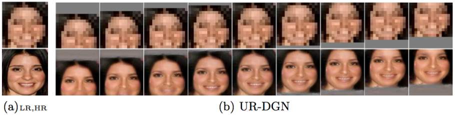


图 2: misalignment

Compressed sensing[4]

Advantages：No assumption on sparsity, only assume that the reconstructed images lies on the manifold that is generated by G

Definition

S-REC：Given a set S, matrix A is said to satisfies S-REC(S,γ, δ) condition, if $\forall x_1, x_2 \in S$:

$$||A(x_1) - A(x_2)|| \geq \gamma ||x_1 - x_2|| - \delta \tag{7}$$

Properties

if generator G is the composition of a linear function and a pointwise non-linearty, Then a random matrix A satisfies the S-REC(G, $1 - \alpha$, 0) with great probability.

[4] Compressed Sensing using Generative Models

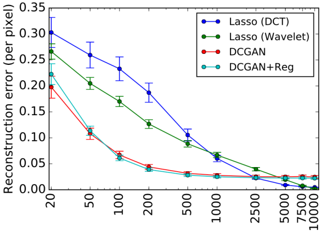
Conclusion

Let $A \in \mathbb{R}^{m\times n}$ satisfies $S-REC(S,\gamma,\delta)$ with probability $1-p$, and for every x it also satisfies $||Ax|| \leq 2||x||$ with probability $1-p$. For any $x^* \in \mathbb{R}^n$, $y = Ax^* + \eta$, suppose

$$||y - A\hat{x}|| \leq \min_{x\in S} ||y - Ax|| + \epsilon \tag{8}$$

then:

$$||\hat{x} - x^*|| \leq (\frac{4}{\gamma} + 1) \min_{x\in S} ||x^* - x|| + \frac{1}{\gamma}(2||\eta|| + \epsilon + \delta) \tag{9}$$



(b) Results on celebA