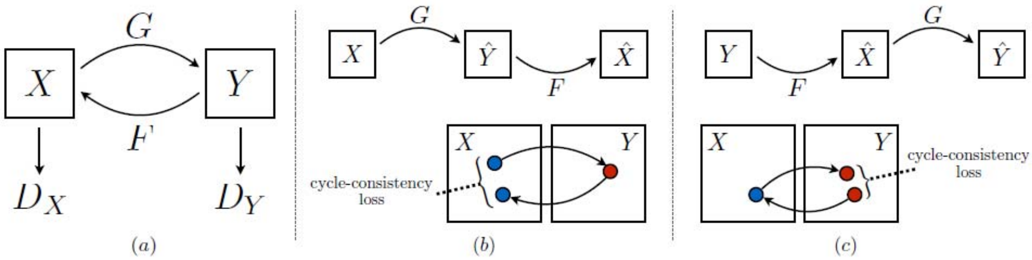
10 月 20 日周报告

1. CycleGAN

- 实现了不同图像领域之间的映射，通过 Unpaired 训练数据找到了一种映射关系
- 模型



- 损失函数

Our full objective is:

$$\begin{aligned}\mathcal{L}(G, F, D_X, D_Y) = & \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) \\ & + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) \\ & + \lambda \mathcal{L}_{\text{cyc}}(G, F), \end{aligned} \tag{3}$$

where λ controls the relative importance of the two objectives. We aim to solve:

$$G^*, F^* = \arg \min_{F,G} \max_{D_x, D_Y} \mathcal{L}(G, F, D_X, D_Y). \tag{4}$$

- 训练细节

为了稳定模型训练步骤，训练时做了两点改动：

（1）. 将

$$\begin{aligned}\mathcal{L}_{\text{LSGAN}}(G, D_Y, X, Y) = & \mathbb{E}_{y \sim p_{\text{data}}(y)}[(D_Y(y) - 1)^2] \\ & + \mathbb{E}_{x \sim p_{\text{data}}(x)}[D_Y(G(x))^2], \end{aligned}$$

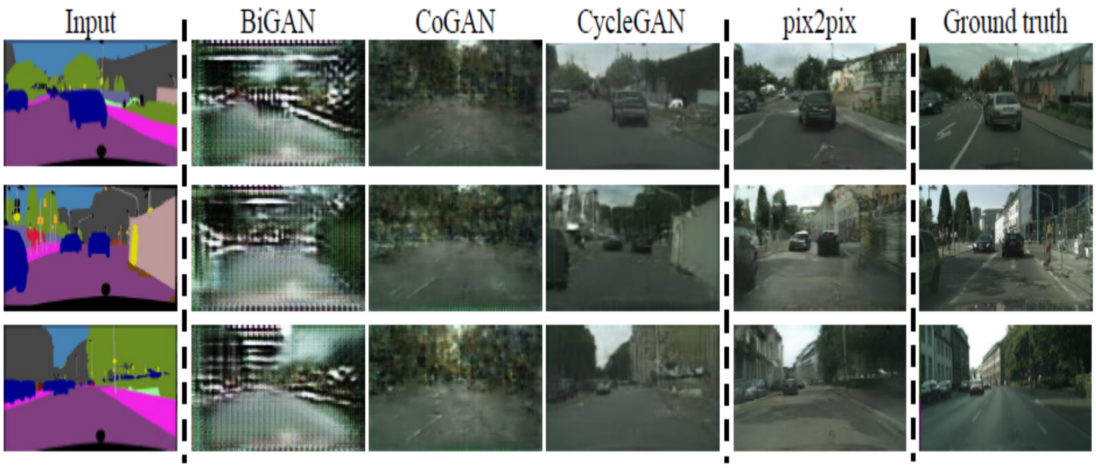
替换为最小二乘损失：

$$\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) = \mathbb{E}_{y\sim p_{\text{data}}(y)}[\log D_Y(y)] + \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log(1 - D_Y(G(x))]$$

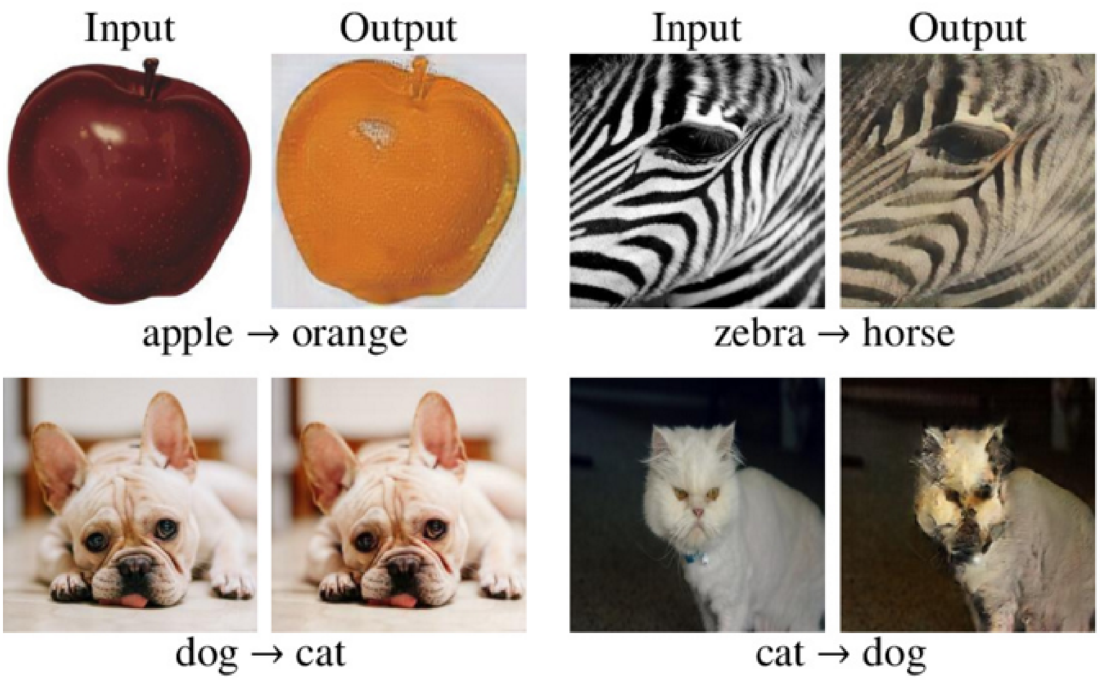
（2）. 判别器利用生成的历史图像数据，而不是仅仅使用当前最新生成的图片数据

- 实验结果

整体实验效果还是可以的，但不是最优



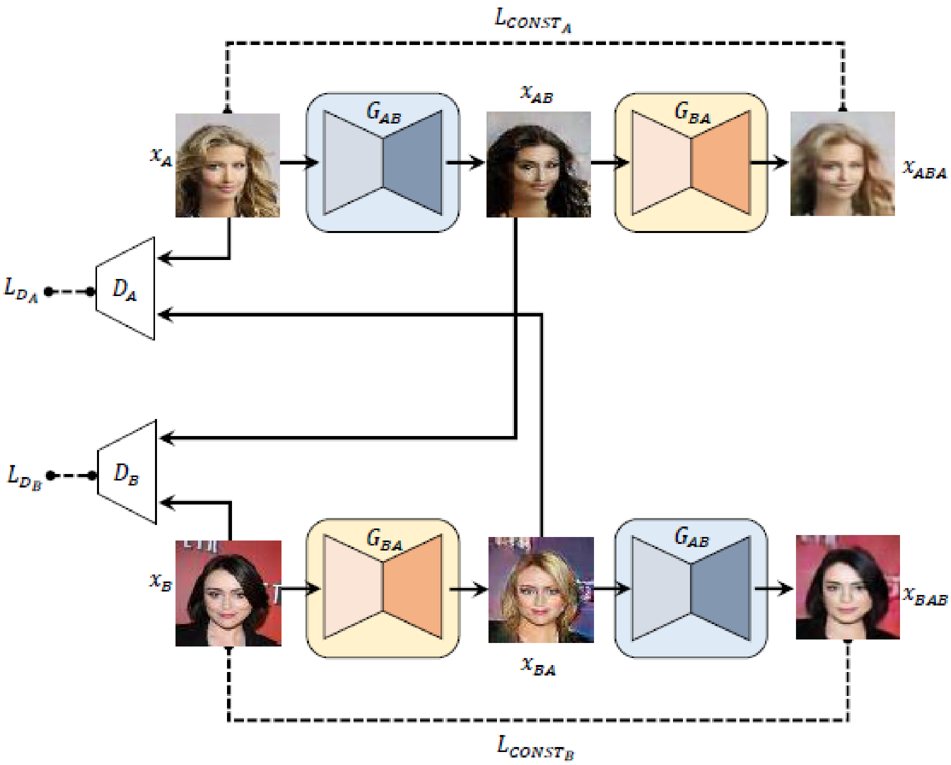
一些失败的例子：



- 该模型在形状转换还存在一定的改进空间

2. DiscoGAN

- 与 CycleGAN 思想完全类似，实现了不同图像领域之间的映射，通过 Unpaired 训练数据找到了一种映射关系
- 模型



- 损失函数

$$
\begin{aligned}
x_{AB} &= \mathbf{G}_{AB}(x_A) && (1)\\
x_{ABA} &= \mathbf{G}_{BA}(x_{AB}) = \mathbf{G}_{BA} \circ \mathbf{G}_{AB}(x_A) && (2)\\
L_{CONST_A} &= d(\mathbf{G}_{BA} \circ \mathbf{G}_{AB}(x_A), x_A) && (3)\\
L_{GAN_B} &= -\mathbb{E}_{x_A \sim P_A}[\log \mathbf{D}_B(\mathbf{G}_{AB}(x_A))] && (4)
\end{aligned}
$$

$$
L_{G_{AB}} = L_{GAN_B} + L_{CONST_A} \quad (5)
$$

$$
\begin{aligned}
L_{D_B} = & -\mathbb{E}_{x_B \sim P_B}[\log \mathbf{D}_B(x_B)] \\
& -\mathbb{E}_{x_A \sim P_A}[\log(1 - \mathbf{D}_B(\mathbf{G}_{AB}(x_A)))] \quad (6)
\end{aligned}
$$

$$
\begin{aligned}
L_G &= L_{G_{AB}} + L_{G_{BA}} \quad (7)\\
&= L_{GAN_B} + L_{CONST_A} + L_{GAN_A} + L_{CONST_B}
\end{aligned}
$$

$$L_D = L_{D_A} + L_{D_B} \tag{8}$$

- 实验结果（人脸，椅子，汽车）

(a) Chair to Car (b) Car to Face

- 提出了跟 cycleGAN 十分类似的模型，不使用成对数据训练，获得不同领域之间图像的映射；未来工作：将 DiscoGAN 应用在不同的模式领域中（文本，图像等）

3. Sketch + Color

- 通过图像素描+粗糙色彩条，生成对应的图像，使用成对数据训练网络
- 模型

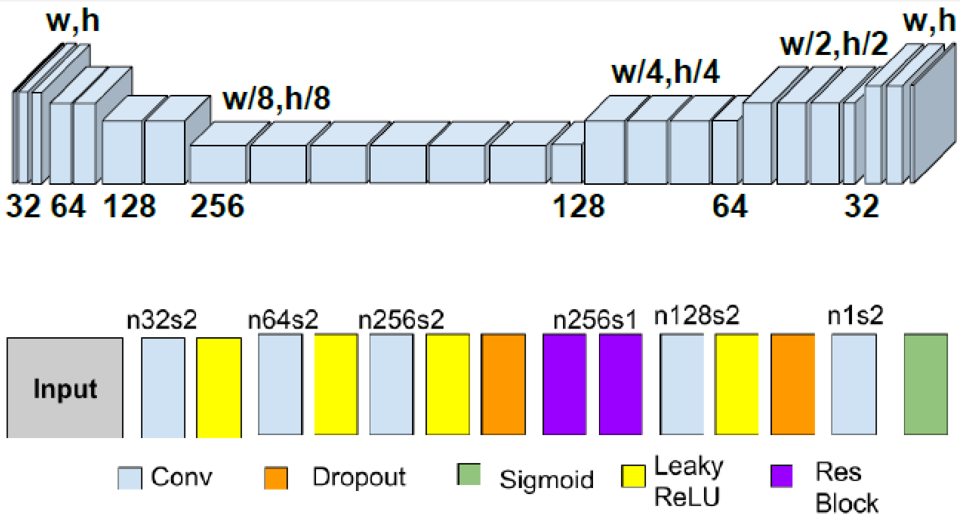


Figure 2. Network Architecture. For the generator (top), we follow the encoder-decoder design, and use three downsampling steps, seven residual blocks at the bottleneck resolution and three upsampling steps. Residual blocks use stride 1. Downsampling uses convolutions with stride 2. Upsampling uses bilinear upsampling followed by residual blocks. We use fully convolutional network for the discriminator (bottom). See scribbler.eye.gatech.edu for code and architecture details.

- 损失函数：像素损失，特征损失，对抗损失，整体变异损失（使输出图像更平滑）

Our final objective function becomes:

$$\mathbf{L} = \mathbf{w_p L_p} + \mathbf{w_f L_f} + \mathbf{w_{adv} L_{adv}} + \mathbf{w_{tv} L_{tv}}$$

- 数据准备与实验效果 作者使用了三种数据：卧室，汽车，人脸数据

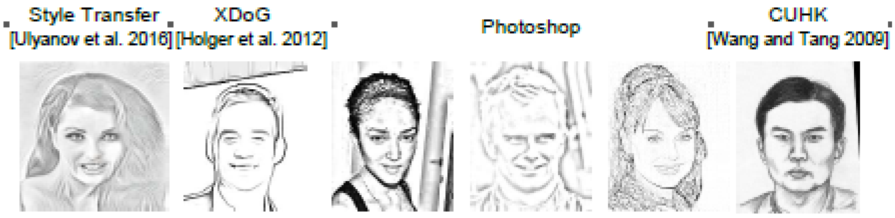
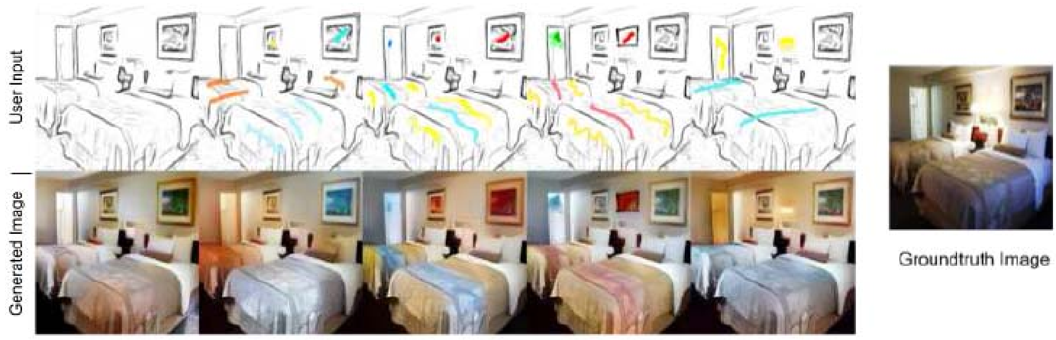


Figure 4. We generate synthetic sketches from photos using five different algorithms. We also include and augment a small set of hand-drawn sketch-photo pairs to help generalize the network to handle real hand-drawn sketch inputs.



References

- Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks
- DiscoGAN
- DiscoGAN information and code
- Sketch and Color