CycleGAN: Unpaired Image-to-Image Translation

Mahshid Alinoori

Introduction

Image-to-Image Translation

Learn the mapping between an input image and an output image using a training set of aligned image pairs.

Unpaired Image-to-Image Translation

Learn the mapping G between two data domains X and Y without having paired examples such that G(X) is indistinguishable from Y.

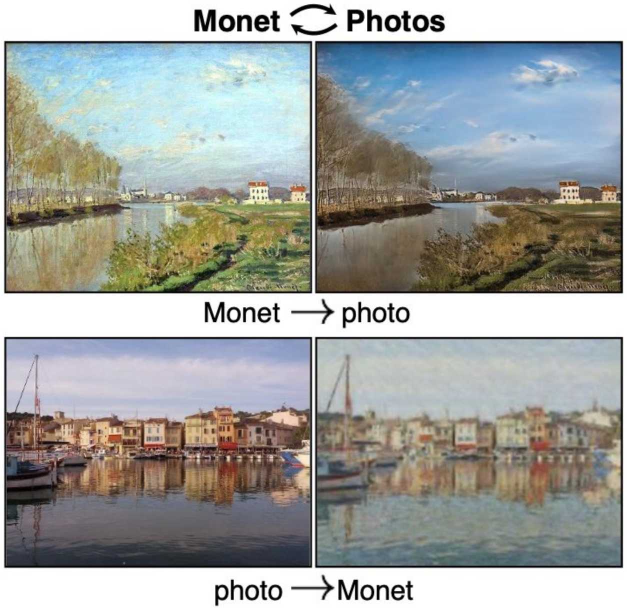
Introduction

GANs

Adopt an adversarial loss to learn the mapping such that the translated images cannot be distinguished from the images in the target domain.

Cycle Consistency

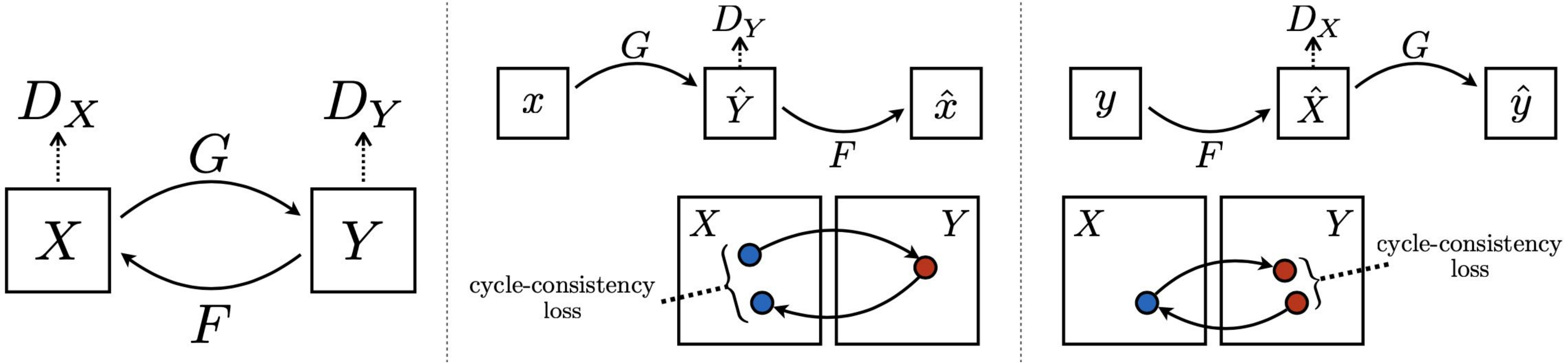
As adversary loss is not enough to find the right mapping, Cycle consistency Loss is adopted to prevent collapse.



CycleGAN

- Captures the characteristics of the input collection and learns to translate them into a target collection without being trained on image pairs
- Implemented by GANs and Cycle consistency loss.

Formulation



- Two mappings
  G : X → Y and F : Y → X
- Dx and Dy are discriminators
- G and F are generators

- Forward
  cycle-consistency loss:
  x → G(x) → F (G(x)) ≈ x

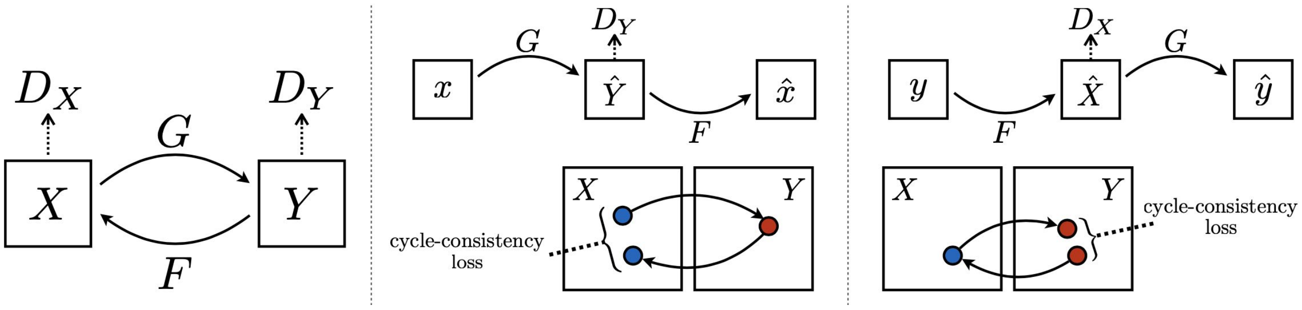
- Backward
  cycle-consistency loss:
  y → F (y) → G(F (y)) ≈ y

Formulation

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

$$\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\|F(G(x)) - x\|_1] + \mathbb{E}_{y\sim p_{\text{data}}(y)}[\|G(F(y)) - y\|_1]$$

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



Implementation

Generator

Two stride-2 convolutions + residual blocks + two fractionally strided convolutions

Discriminator

70x70 PatchGANs with fewer parameters

L_{GAN}

Replacing log-likelihood by least-square loss for stability

Updates

Updating the discriminator using a history of generated images (last 50)

Evaluation

Qualitative Evaluation

25 participants per algorithm using Amazon Mechanical Turk

Quantitative Evaluation

FCN Score used in predicting a label for a generated image and comparing it with the original label

Baselines

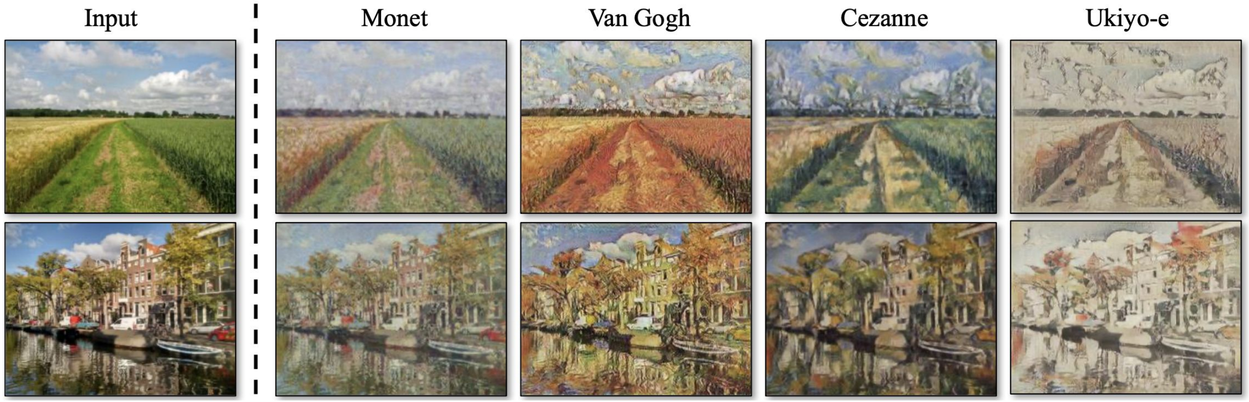
CoGAN
SimGAN
BiGAN/ALI
Pix2pix

Ablations

Both cycle loss and GAN losses are necessary to achieve the best possible results

Application: Collection Style Transfer

Mimic the style of of an entire collection of artworks rather than the style of a single image



Application: Object Transfiguration and Season Transfer

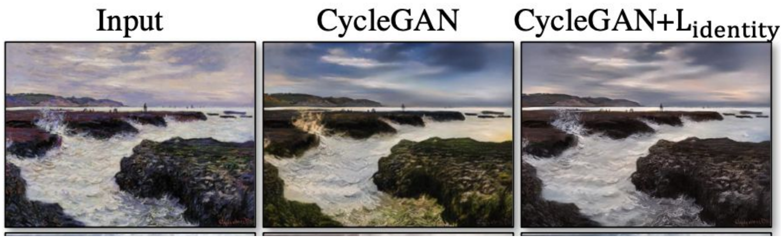
Translate one object into another object of the same category.

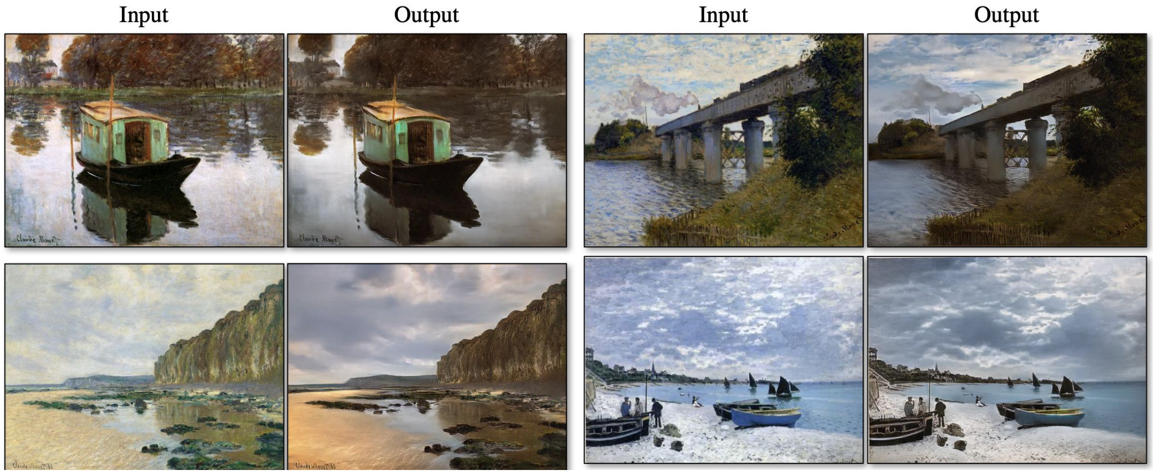
zebra → horse

winter Yosemite → summer Yosemite

Application: Photo generation from paintings

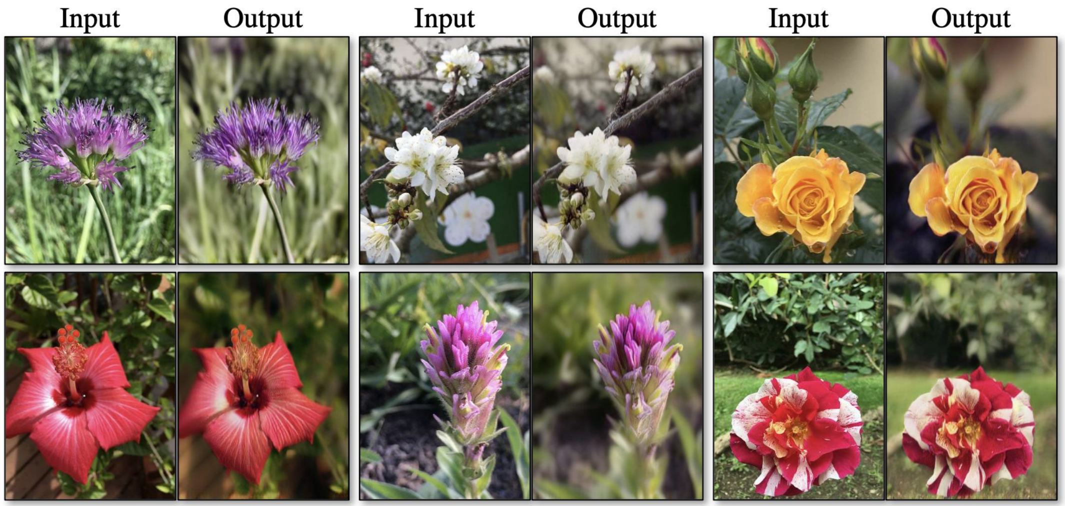
Painting to photo translation using an additive loss $L_{identity}$ to preserve the coloring





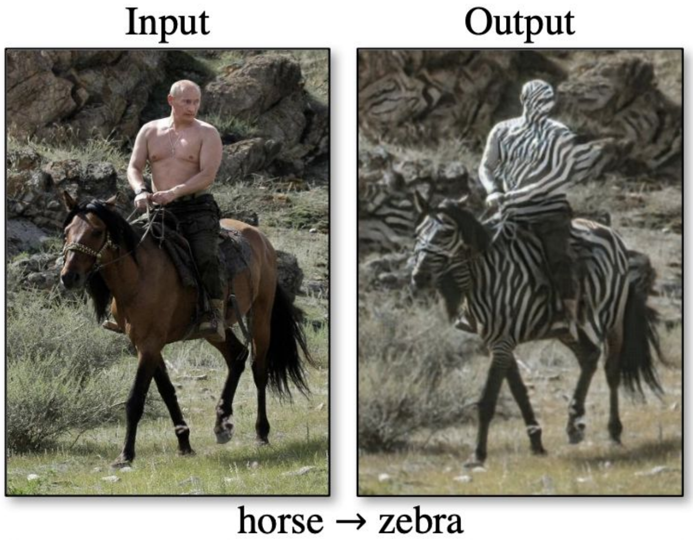
Application: Photo enhancement

Decreasing the depth of field in photos taken by smartphones



Conclusion

- CycleGAN is the first GAN-based unpaired image-to-image translator
- Models the mapping in a cycle with GAN Loss and cycle-consistency loss
- Limitations in geometric changes and failures due to distribution characteristics of the training set



Thanks!