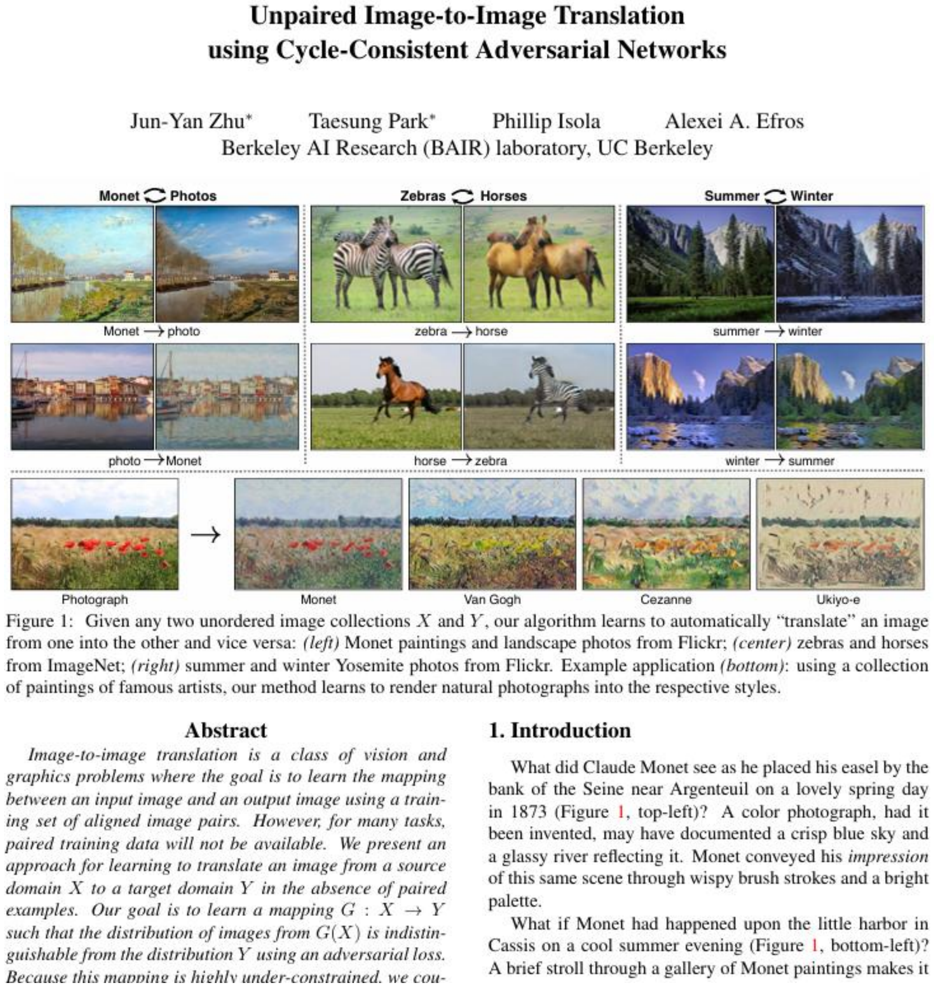
Image-to-Image Translation

CycleGAN

ICCV

2017



Unpaired Image-to-Image Translation
using Cycle-Consistent Adversarial Networks

Jun-Yan Zhu* Taesung Park* Phillip Isola Alexei A. Efros
Berkeley AI Research (BAIR) laboratory, UC Berkeley

Figure 1: Given any two unordered image collections X and Y, our algorithm learns to automatically "translate" an image from one into the other and vice versa: *(left)* Monet paintings and landscape photos from Flickr; *(center)* zebras and horses from ImageNet; *(right)* summer and winter Yosemite photos from Flickr. Example application *(bottom)*: using a collection of paintings of famous artists, our method learns to render natural photographs into the respective styles.

Abstract

Image-to-image translation is a class of vision and graphics problems where the goal is to learn the mapping between an input image and an output image using a training set of aligned image pairs. However, for many tasks, paired training data will not be available. We present an approach for learning to translate an image from a source domain X to a target domain Y in the absence of paired examples. Our goal is to learn a mapping $G : X \rightarrow Y$ such that the distribution of images from $G(X)$ is indistinguishable from the distribution Y using an adversarial loss. Because this mapping is highly under-constrained, we cou-

1. Introduction

What did Claude Monet see as he placed his easel by the bank of the Seine near Argenteuil on a lovely spring day in 1873 (Figure 1, top-left)? A color photograph, had it been invented, may have documented a crisp blue sky and a glassy river reflecting it. Monet conveyed his *impression* of this same scene through wispy brush strokes and a bright palette.

What if Monet had happened upon the little harbor in Cassis on a cool summer evening (Figure 1, bottom-left)? A brief stroll through a gallery of Monet paintings makes it

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

Background & Goal

Previous research limitation

- Obtaining paired dataset can be difficult and expensive.
- In some cases, it may even be practically impossible to acquire paired datasets.

Goal

- Achieve image-to-image translation using unpaired dataset.

Fig 1. Examples of paired datasets that are impossible to acquire

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

Background & Goal

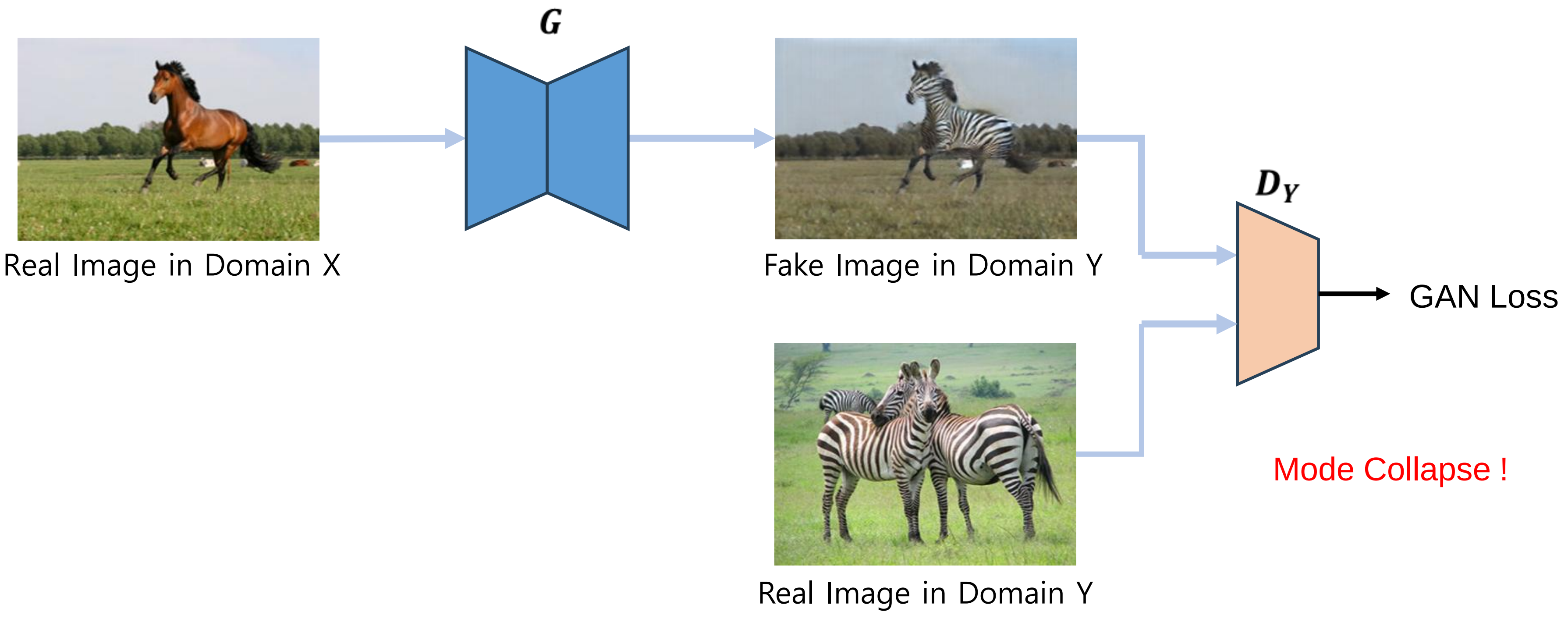


Fig 2. Traditional GAN architecture

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

Cycle Consistency

Hello ⇄ Bonjour

Summer ⇄ Winter

Fig 3. Cycle consistency that requires returning to the original

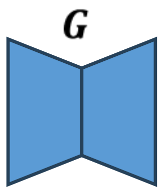
Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

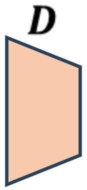
Network Architectures

Pix2Pix

Target Generator (G)



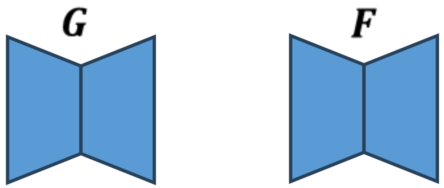
Target Discriminator (D)



CycleGAN

Target Generator (G)

Source Generator (F)



Target Discriminator (D_Y)

Source Discriminator (D_X)

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

Loss Functions

GAN loss

- Mapping function $G: X \rightarrow Y$ and Discriminator D_Y
- Mapping function $F: Y \rightarrow X$ and Discriminator D_X

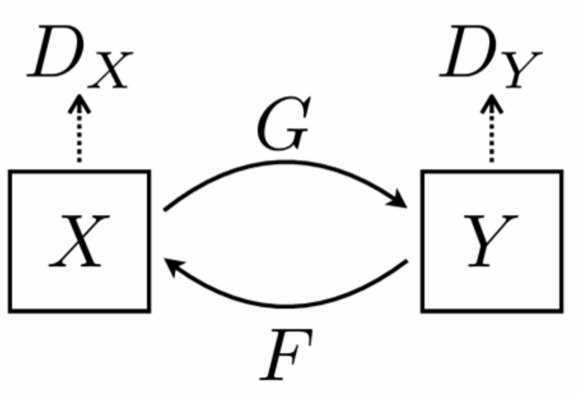


Fig 4. CycleGAN Architecture

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

Eq 1. Adversarial loss for G

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

Eq 2. Adversarial loss for F

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

Image-to-Image Translation

CycleGAN

Loss Functions

Cycle consistency loss

- $x \rightarrow G(x) \rightarrow F\big(G(x)\big) \approx x$
- $y \rightarrow F(y) \rightarrow G\big(F(y)\big) \approx y$

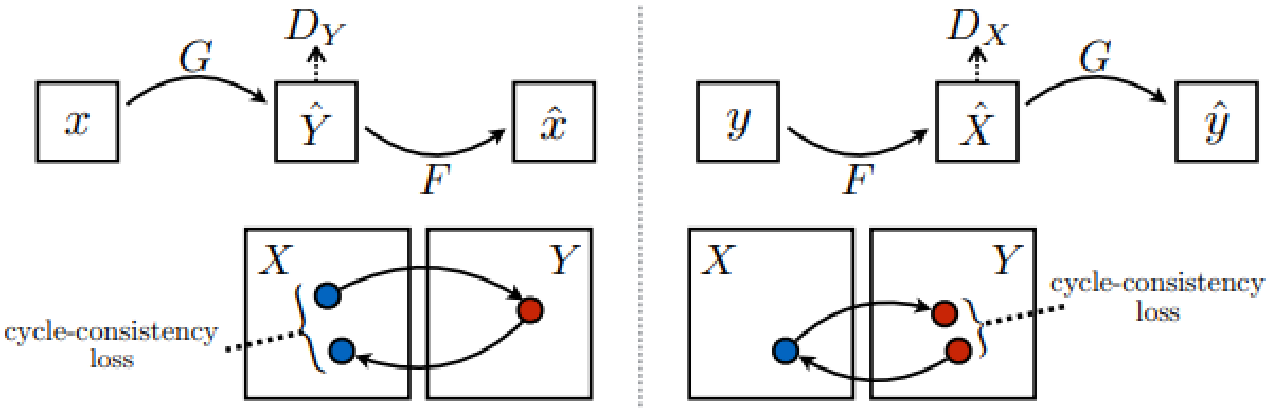


Fig 5. Idea of cycle consistency

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

Eq 3. Cycle consistency loss

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

Loss Functions

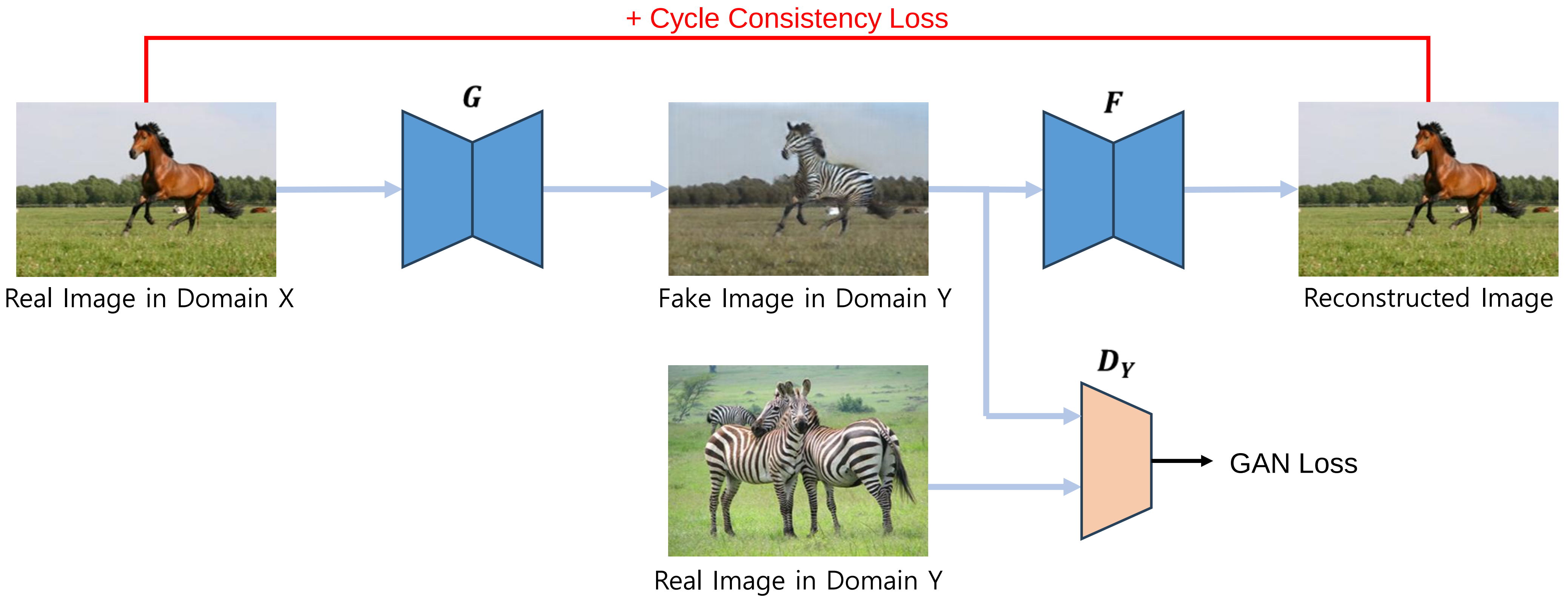


Fig 6. GAN + Cycle consistency loss architecture

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

Loss Functions

Identity Loss

- $Y = G(Y)\ (G: X \rightarrow Y)$
- $X = F(X)\ (F: Y \rightarrow X)$

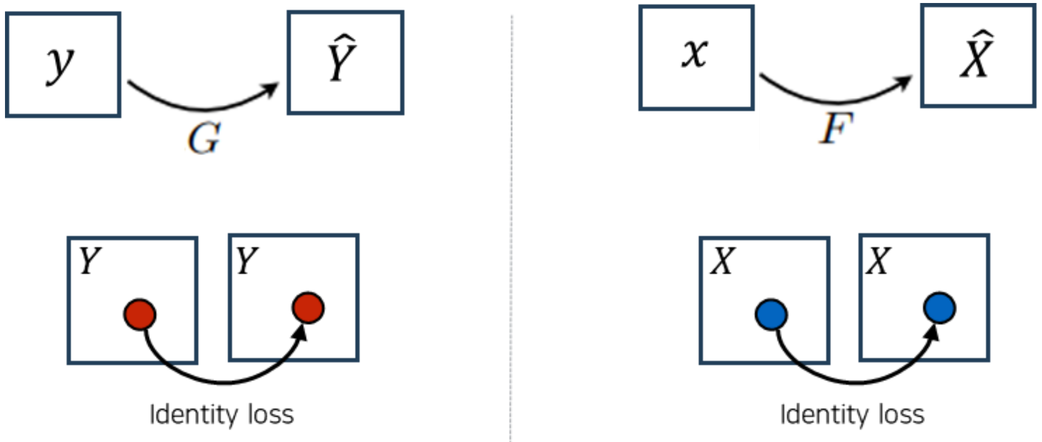


Fig 7. Idea of Identity

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

Eq 4. Identity loss

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

Image-to-Image Translation

CycleGAN

Loss Functions

$G: X \rightarrow Y$
$F: Y \rightarrow X$

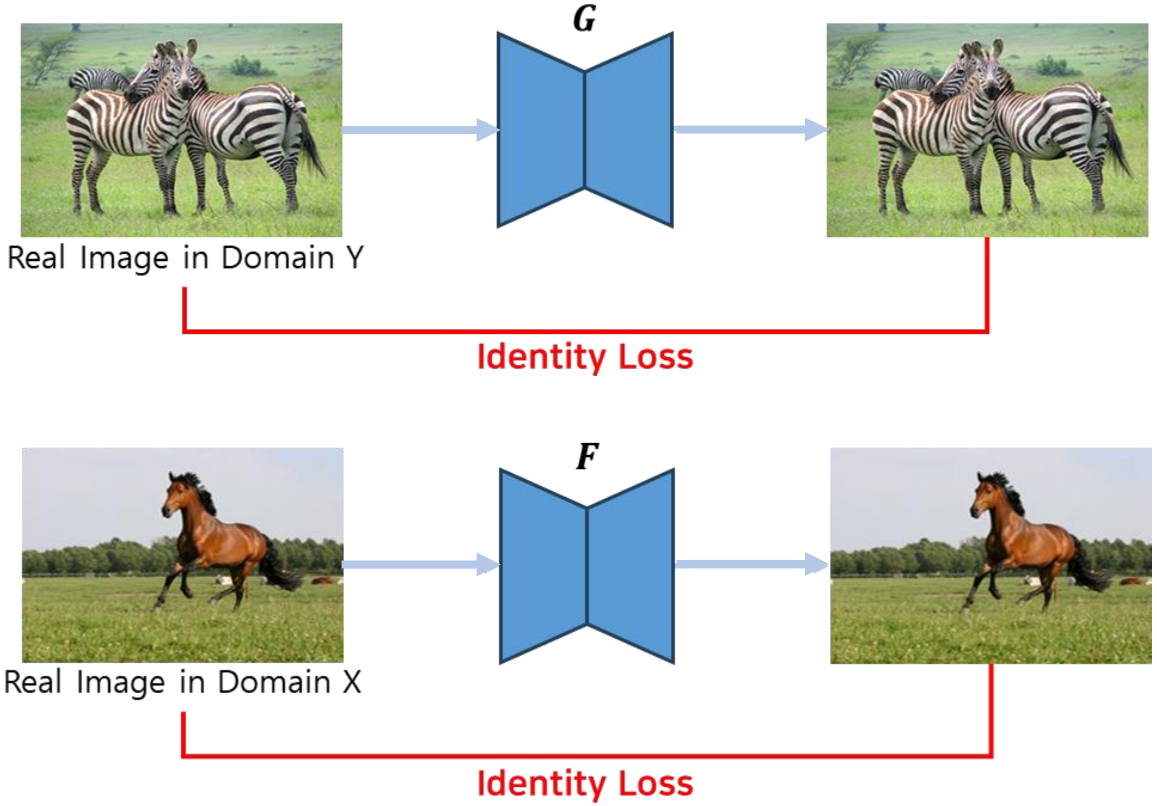


Fig 8. Identity loss Architecture

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

Experiment Results

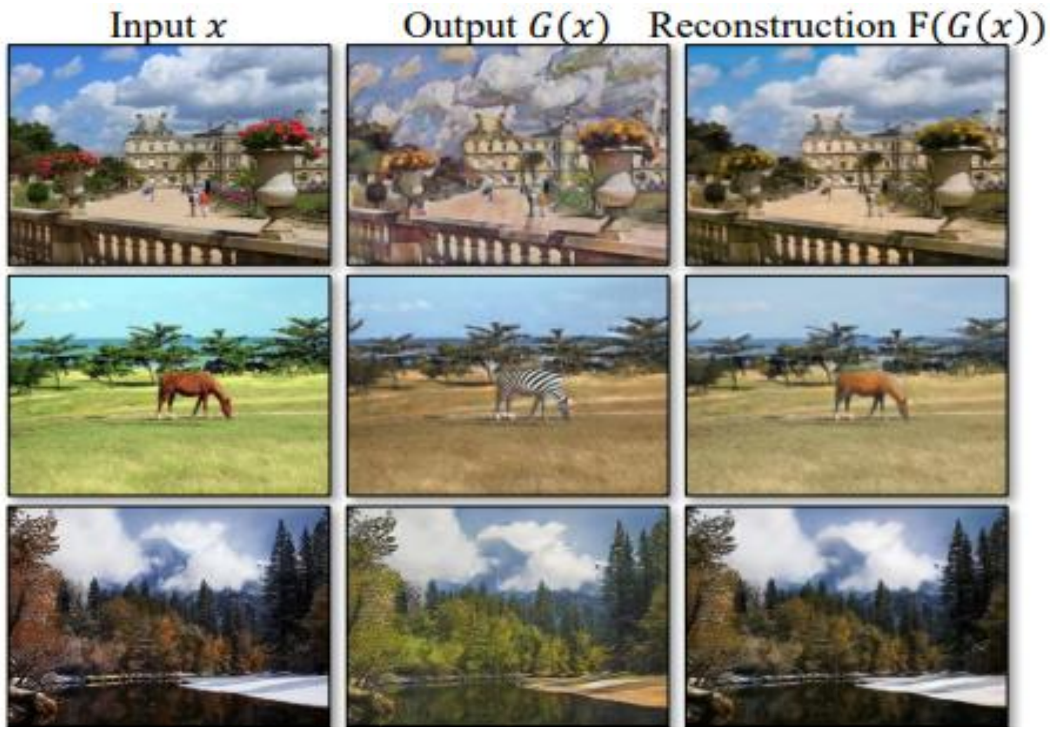


Fig 9. Cycled result images

Cycle consistency loss enforces that an image translated to the target domain and then back to the source domain remai ns unchanged.

- It helps the network preserve the original content while le arning the transformation.
- By requiring the image to return to its original form, it pre vents mode collapse.

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

Image-to-Image Translation

CycleGAN

Experiment Results

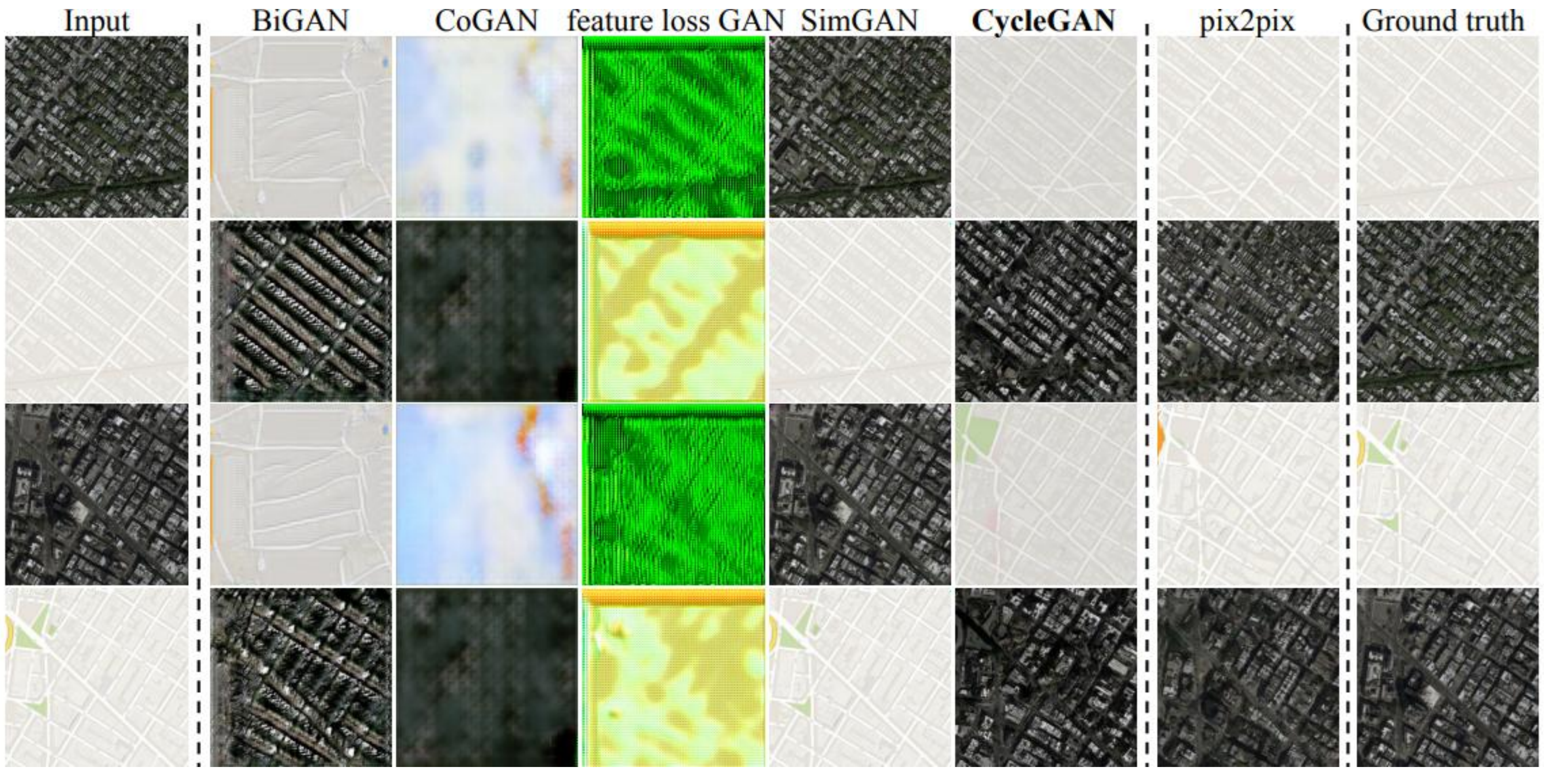


Fig 10. Comparison with conventional methods

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

Image-to-Image Translation

CycleGAN

Experiment Results

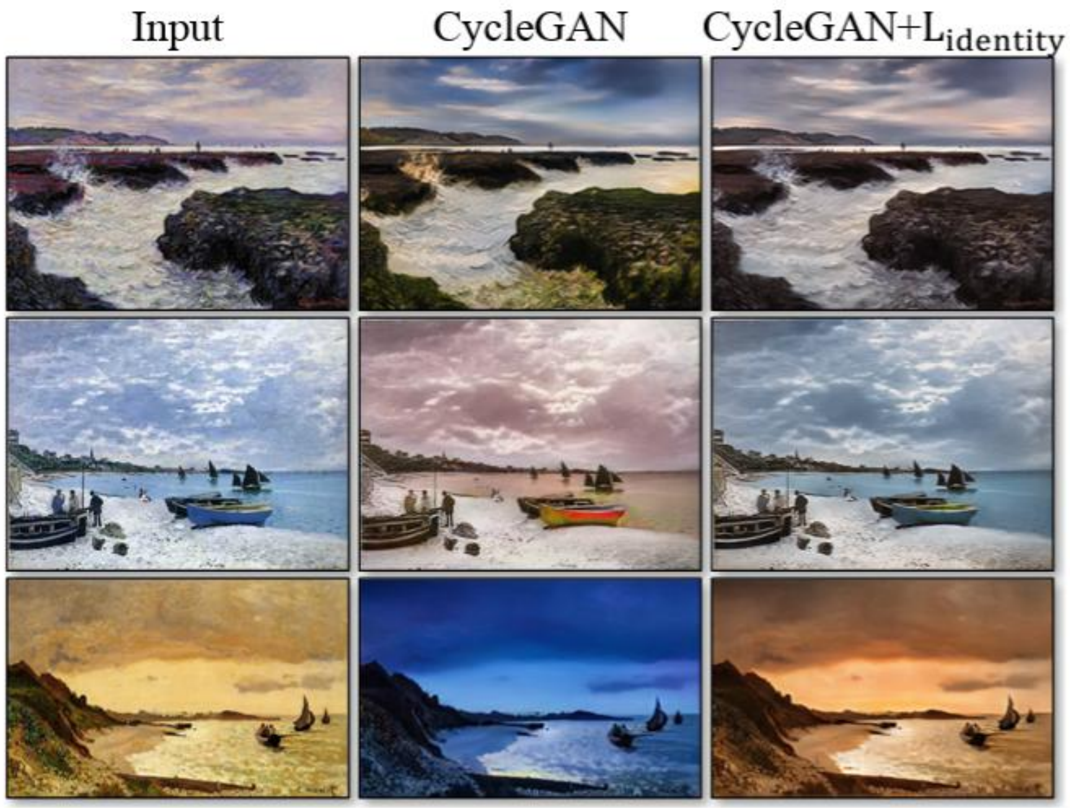


Fig 11. The effect of the identity loss

Identity loss enforces that data already belonging to the targ et domain remains unchanged.

- It helps the network better understand the fundamental c haracteristics of domain differences.
- It guides the model to suppress unnecessary transformat ions.

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

Limitations

Geometric change

- Failure in handling more diverse and extreme transformations, including geometric changes

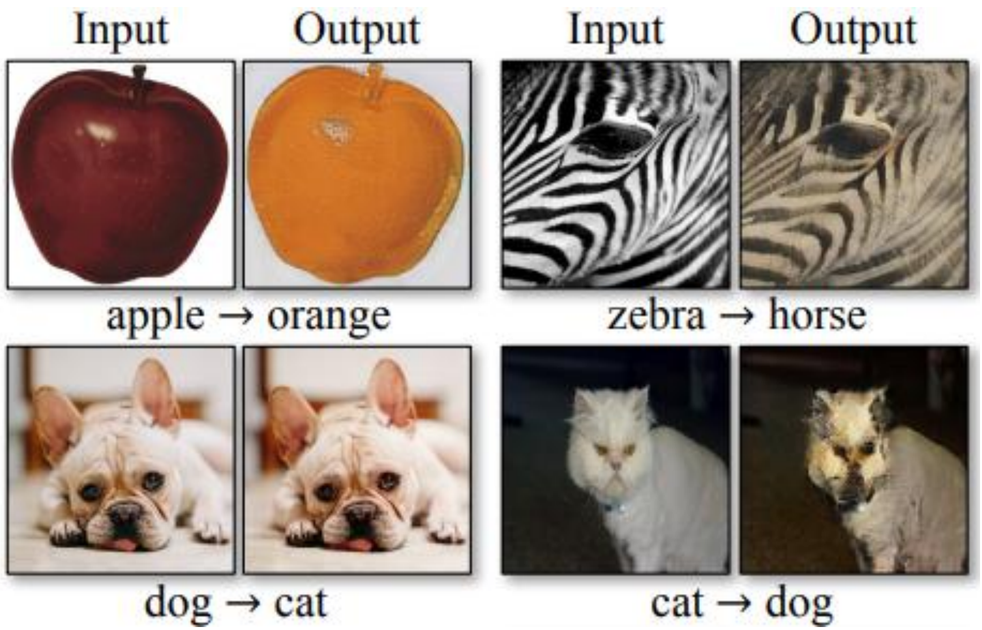


Fig 12. Cases of failure in geometric transformation

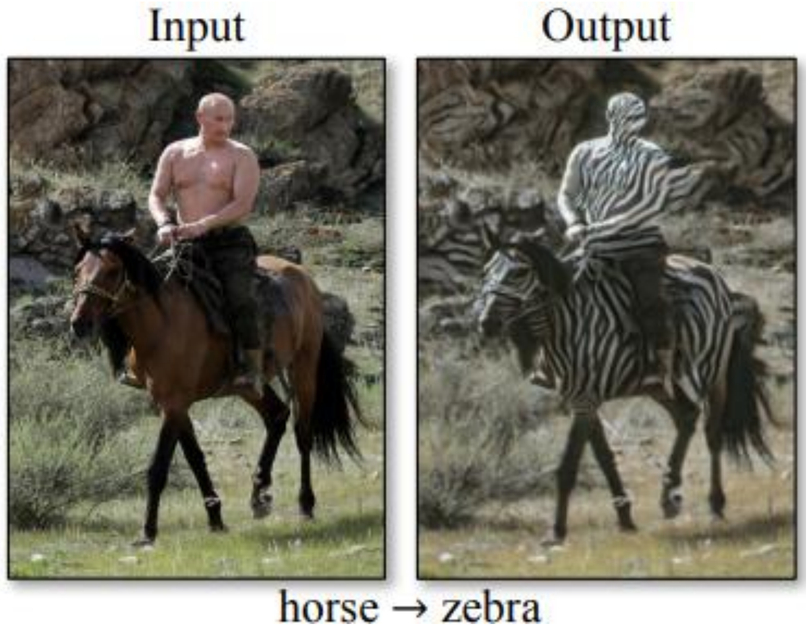
Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

CycleGAN

Limitations

Dataset distribution

- Sensitive to dataset distribution



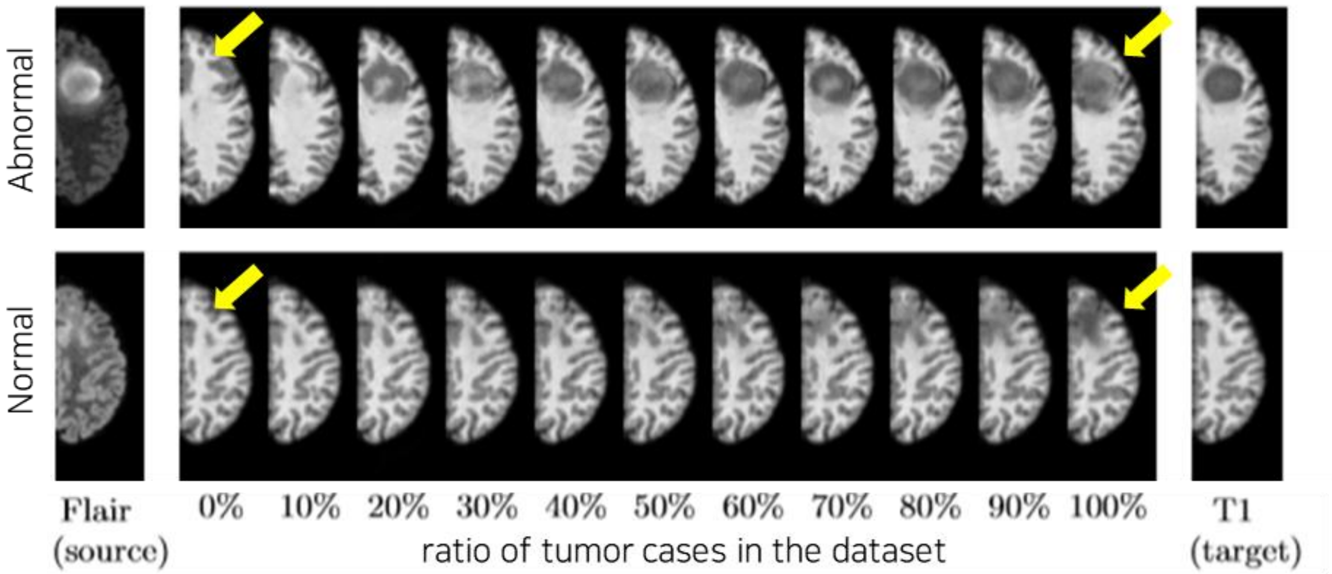


Fig 13. Failure in transforming datasets not used during training

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).
Cohen, J. P., Luck, M., & Honari, S. (2018). Distribution matching losses can hallucinate features in medical image translation. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I* (pp. 529-536). Springer International Publishing.