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

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. 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 🡪 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 couple it with an inverse mapping F: Y 🡪 X and introduce a

cycle consistency loss to enforce F(G(X)) ≈ X (and vice versa).

1. Introduction

We can reason about the stylistic differences between these two sets, and thereby imagine what a scene might look like if we were to “translate” it from one set into the other.

In this paper, we present a method that can learn to do the same: capturing special characteristics of one image collection and figuring out how these characteristics could be translated into the other image collection, all in the absence of any paired training examples.

We may train mapping G: X 🡪 Y such that the output ^y = G(x), x ∈ X, is indistinguishable from images y ∈ Y by an adversary trained to classify ^y apart from y.

These issues call for adding more structure to our objective. Therefore, we exploit the property that translation should be “cycle consistent”, in the sense that if we translate, e.g., a sentence from English to French, and then translate it back from French to English, we should arrive back at the original sentence.

We apply this structural assumption by training both the mapping G and F simultaneously, and adding a cycle consistency loss that encourages F(G(x)) ≈ x and G(F(y)) ≈y. Combining this loss with adversarial losses on domains X and Y yields our full objective for unpaired image-to-image translation.

2. Related work

**Generative Adversarial Networks (GANs)**

The key to GANs’ success is the idea of an adversarial loss that forces the generated images to be, in principle, indistinguishable from real photos. We adopt an adversarial loss to learn the mapping such that the translated images cannot be distinguished from images in the target domain.

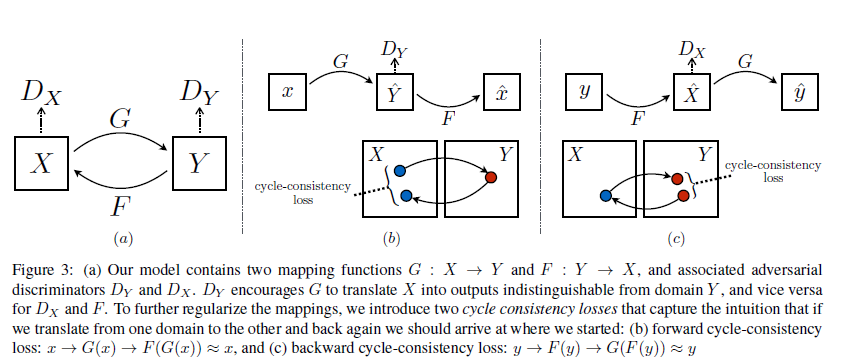
**Image-to-Image Translation**

**Unpaired Image-to-Image Translation**

**Cycle Consistency**

**Neural Style Transfer**

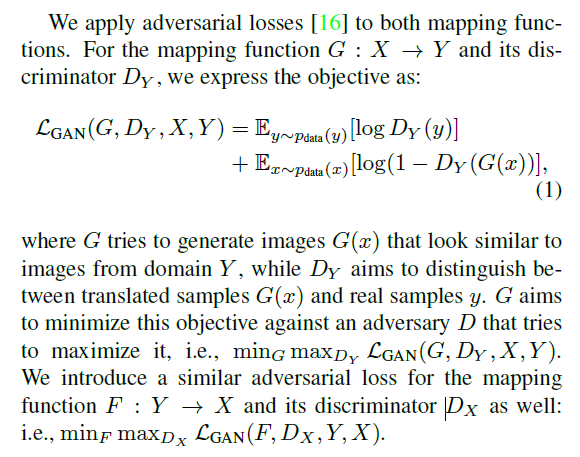
3. Formulation



We introduce two adversarial discriminators DX and DY, where DX aims to distinguish between images {x} and translated images {F(y)}; in the same way, DY aims to discriminate between {y} and {G(x)}.

Our objective contains two types of terms: adversarial losses for matching the distribution of generated images to the data distribution in the target domain; and cycle consistency losses to prevent the learned mappings G and F from contradicting each other.

3.1. Adversarial Loss



3.2. Cycle Consistency Loss

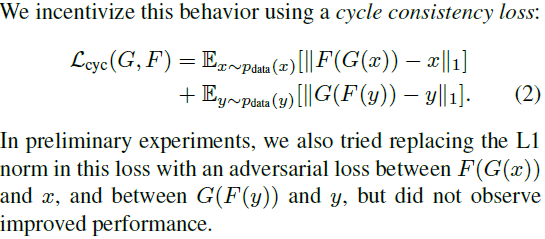
Adversarial losses alone cannot guarantee that the learned function can map an individual input xi to

a desired output yi.

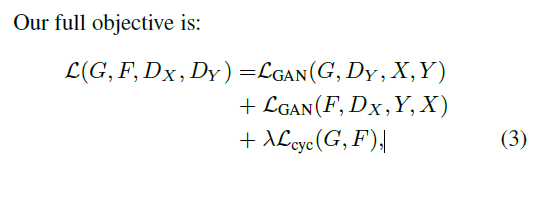
To further reduce the space of possible mapping functions, we argue that the learned mapping functions should be cycle-consistent: for each image x from domain X, the image translation

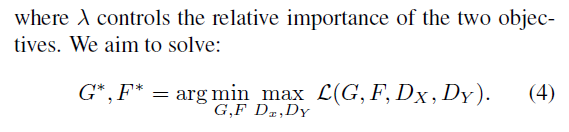
cycle should be able to bring x back to the original image, i.e., x 🡪 G(x) 🡪 F(G(x)) ≈ x. We call this forward cycle consistency. Similarly, for each image y from domain Y , G and F should also satisfy

backward cycle consistency: y 🡪F(y) 🡪 G(F(y)) ≈ y.



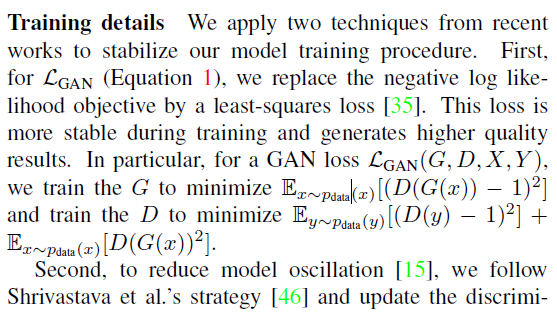
3.3. Full Objective

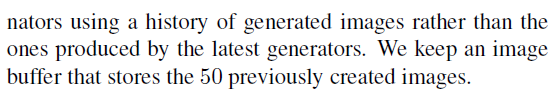




4. Implementation

**Network Architecture**





5. Results

5.1. Evaluation

5.1.1 Evaluation Metrics

**AMT perceptual studies**

**FCN score**

**Semantic segmentation metrics**

5.1.2 Baselines

**CoGAN**

**SimGAN**

**Feature loss + GAN**

**BiGAN/ALI**

**pix2pix**

5.1.3 Comparison against baselines

5.1.4 Analysis of the loss function

5.1.5 Image reconstruction quality

5.1.6 Additional results on paired datasets.

5.2. Applications

**Collection style transfer**

**Object transfiguration.**

**Season transfer**

**Photo generation from paintings**

**Photo enhancement**

6. Limitations and Discussion

Handling more varied and extreme transformations, especially geometric changes, is an important problem for future work.

Some failure cases are caused by the distribution characteristics of the training datasets.

7. Appendix

7.1. Training details

7.2. Network architectures

Generator architectures

Discriminator architectures