CycleGAN Paper

Dr. Alireza Aghamohammadi

Introduction to CycleGAN

- In this lecture, we will discuss the paper titled "Image-to-Image Translation Using Cycle-Consistent Adversarial Networks."
- For many tasks, paired training data is not available.
- CycleGAN captures the unique characteristics of one image collection and translates these characteristics into another image collection, all without any paired training examples.

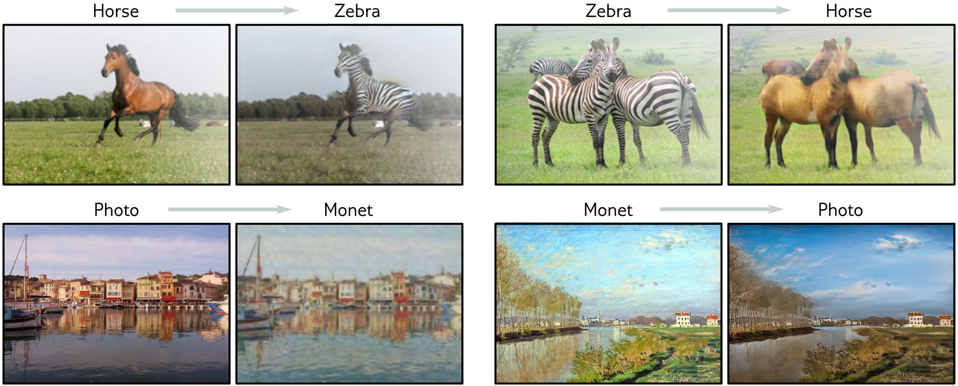

Image credit: Understanding Deep Learning

CycleGAN Building Blocks

Consider unpaired training data consisting of:

- A source set $\{x_i\}_{i=1}^{N}$ ($x_i \in X$)
- A target set $\{y_j\}_{j=1}^{M}$ ($y_j \in Y$)

with no information provided as to which x_i matches y_j.
We denote the data distribution as:

- $x \sim P_{data}(x)$
- $y \sim P_{data}(y)$

The model consists of:

- Two generators: G_A and G_B
- Two discriminators: D_A and D_B

- G_A generates images for the target set.
- G_B generates images for the source set.

Cycle Consistency: If we translate a sentence from English to French, and then translate it back from French to English, we should arrive back at the original sentence.

$$x_i \rightarrow G_A \rightarrow y_j \rightarrow G_B \rightarrow x_i$$

Training CycleGAN

The full objective for CycleGAN is:

$$\min_{G_A, G_B} \max_{D_A, D_B} \mathcal{L}_{GAN}(G_A, D_A, X, Y) + \mathcal{L}_{GAN}(G_B, D_B, Y, X) + \lambda \cdot \mathcal{L}_{CYC}(G_A, G_B)$$

Where $\mathcal{L}_{GAN}$ is the adversarial loss:

$$\mathcal{L}_{GAN}(G_A, D_A, X, Y) = \mathbb{E}_{y \sim P_{data}(y)} [\log D_A(y)] + \mathbb{E}_{x \sim P_{data}(x)} [\log(1 - D_A(G_A(x)))]$$
$$\mathcal{L}_{GAN}(G_B, D_B, Y, X) = \mathbb{E}_{x \sim P_{data}(x)} [\log D_B(x)] + \mathbb{E}_{y \sim P_{data}(y)} [\log(1 - D_B(G_B(y)))]$$

We introduce two cycle consistency losses to ensure that translating from one domain to the other and back again returns us to the original domain:

$$\mathcal{L}_{CYC}(G_A, G_B) = \mathbb{E}_{x \sim P_{data}(x)} [\|G_B(G_A(x)) - x\|_1] + \mathbb{E}_{y \sim P_{data}(y)} [\|G_A(G_B(y)) - y\|_1]$$

Sometimes, it can be beneficial to add an identity mapping loss to the objective function:

$$\mathcal{L}_{id}(G_A, G_B) = \mathbb{E}_{x \sim P_{data}(x)} [\|G_B(x) - x\|_1] + \mathbb{E}_{y \sim P_{data}(y)} [\|G_A(y) - y\|_1]$$