Image-to-Image Translation with Conditional Adversarial Networks

Abstract

We investigate conditional adversarial networks as a general-purpose solution to image-to-image translation problems.

1. Introduction

CNNs learn to minimize a loss function –an objective that scores the quality of results – and although

the learning process is automatic, a lot of manual effort still goes into designing effective losses.

GANs learn a loss that tries to classify if the output image is real or fake, while simultaneously training a generative model to minimize this loss.

2. Related work

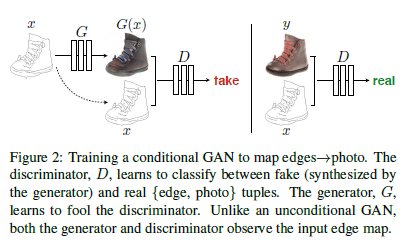
**Structured losses for image modeling**

Conditional GANs instead learn a structured loss. Structured losses penalize the joint configuration of the output. In theory, penalize any possible structure that differs between output and target.

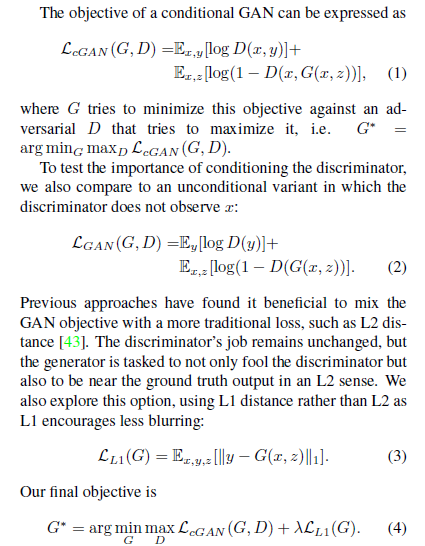
**Conditional GANs**

3. Method

GANs are generative models that learn a mapping from random noise vector z to output image y, G : z → y . In contrast, conditional GANs learn a mapping from observed image x and random noise vector z, to y, G : {x; z} → y. The generator G is trained to produce outputs that cannot be distinguished from “real” images by an adversarially trained discriminator, D, which is trained to do as well as possible at detecting the generator’s “fakes”. This training procedure is diagrammed in Figure 2.



3.1. Objective



3.2. Network architectures

3.3. Optimization and inference

4. Experiments

4.1. Evaluation metrics

To more holistically evaluate the visual quality of our results, we employ two tactics. First, we run “real vs. fake” perceptual studies on Amazon Mechanical Turk (AMT). Second, we measure whether or not our synthesized cityscapes are realistic enough that off-the-shelf recognition system can recognize the objects in them. This metric is similar to the “inception score” from, the object detection evaluation, and the “semantic interpretability” measures.

4.2. Analysis of the objective function

4.3. Analysis of the generator architecture

4.4. FromPixelGANs to PatchGANs to ImageGANs

4.4. Perceptual validation

4.6. Semantic segmentation

4.7. Community driven Research

5. Conclusion