On the Reproduction of Pix2Pix

INTRODUCTION

In 2016, Phillip Isola, et al. published their paper Image-to-Image Translation with Conditional Adversarial Networks. The paper presents a conditional Generative Adversarial Network (cGAN), Pix2Pix, designed for general purpose image-to-image translation. This project aims to reproduce some of the results in the paper.

REPRODUCTION

We attempt to reproduce the results in table 1 of the paper. This table contains the FCN-8s scores for the pix2pix model trained on the Cityscapes dataset. The scores are computed by generating images from the cityscapes masks, feeding these through the FCN segmenter and comparing the output with the original mask. We limited ourselves to reproducing four of the six total rows in the table.

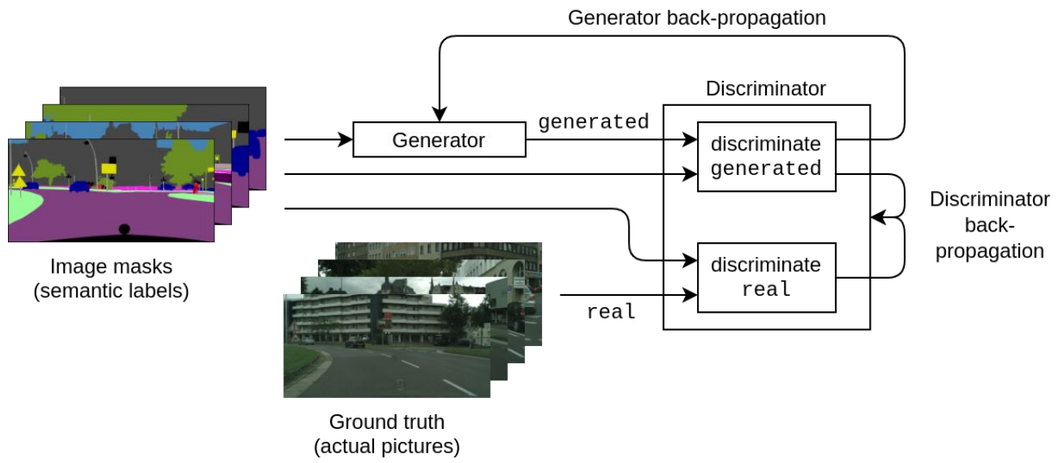
Loss	Per-pixel acc.	Per-class acc.	Class IOU
L1	0.42	0.15	0.11
GAN	0.22	0.05	0.01
cGAN	0.57	0.22	0.16
L1+GAN	0.64	0.20	0.15
L1+cGAN	**0.66**	**0.23**	**0.17**
Ground truth	0.80	0.26	0.21

Table 1: FCN-scores for different losses, evaluated on Cityscapes labels↔photos.

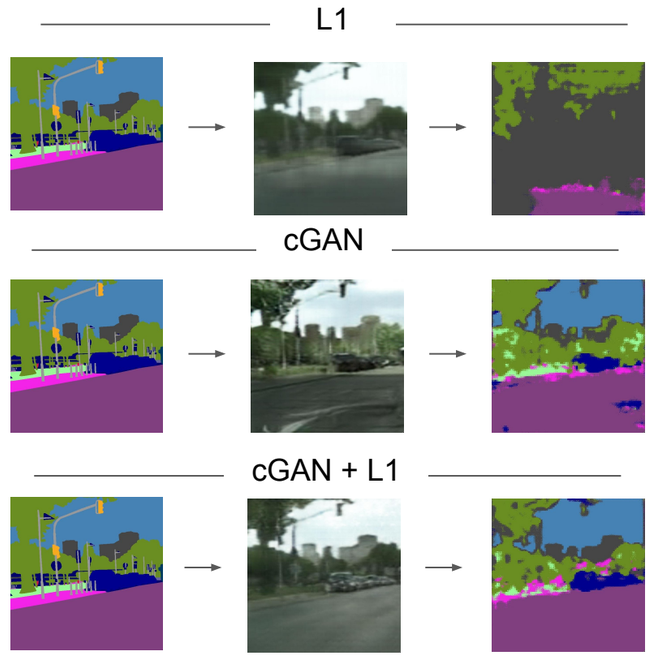
Loss	Per-pixel acc.	Per-class acc.	Class IOU
L1	0.60	0.20	0.15
cGAN	0.66	0.23	0.18
L1+cGAN	**0.55**	**0.20**	**0.14**
Ground truth	0.76	0.25	0.20

Table 2: Our FCN-scores evaluated on Cityscapes dataset

cGAN OVERVIEW



RESULTS



DISCUSSION

Our model produces qualitatively convincing results and table values mostly correspond with the paper although cGAN alone performs better than cGAN + L1. We suspect this stems from excessive training on the L1 loss term as output images become progressively more blurry after many training epochs. Moreover, we find that the discriminator can “win” over the generator, hence zeroing the gradients in directions minimizing the GAN loss and forcing it to only train on L1 loss.

Original Paper

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros, Image-to-image translation with conditional adversarial networks, 2016. https://arxiv.org/abs/1611.07004

Reproducers

Art van Liere,
Pavel Suma,
Sigurd Totland

TU Delft