

# Discriminative Region Proposal Adversarial Networks for High-Quality Image-to-Image Translation

Wenjie Niu

July 19, 2018

After several days of reading the paper of Discriminative Region Proposal Adversarial Networks for High-Quality Image-to-Image Translation, I have access to the framework, structure and basic professional knowledge such as Convolutional Neural Network(CNN) and Generative Adversarial Network(GAN). However, there are some difficulties for me like the deprivation of formula in the understanding of GAN.

## Abstract

*In this paper, authors present Discriminative Region Proposal Adversarial Network(DRPAN) for high-quality image-to-image translation. In order to reduce the artifacts and blur problems during the process of translation, they explore the patch discriminator. It produces the score map and find the most artificial part by a sliding window. Then put the discriminative region mask the corresponding real image to obtain a fake-mask real image. A reviser based on GANs is used to distinguish the real to fake-mask real to help improving the translation quality.*

## 1. Introduction

The procedure of image-to-image translation task is decomposed into three iterated steps. First is to generate a fake image with some local artifacts by generator. Second is to propose the most fake region from the generated image using DRPnet as shown in figure. 1. Third is to mask real image with the discriminative region, so that DRPAN(as shown in figure. 3) can be optimized on the most fake part.

There are three main contributions in this paper:

- discriminators based on patch for producing discriminative region;
- reviser based on GANs to provide constructive revisions for generator
- building a DRPAN model

In detail, semantic segmentation graph is put in the generator, the fake image is the output. Then the discriminator is to find and extract the discriminative region in the fake image with sliding window. The corresponding real image mask with the discriminative region producing a fake-mask real sample. Last, the sample is put in the reviser to judge the image whether fake-mask real or real image. The whole architecture is shown in figure. 3.

## 2. Conclusion

Thanks to the carefully reading of this paper, I learnt a lot about the procedure of GAN training, the working mechanism and concrete understanding of this paper.

## References

- [1] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The cityscapes dataset for semantic urban scene understanding. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016. 2
- [2] R. Tyleek and R. ra. Spatial pattern templates for recognition of objects with regular structure. In *Global Conference on Psychology Researches*, 2013. 3
- [3] C. Wang, H. Zheng, Z. Yu, Z. Zheng, Z. Gu, and B. Zheng. Discriminative region proposal adversarial networks for high-quality image-to-image translation. *arXiv preprint arXiv:1711.09554*, 2017. 2, 3

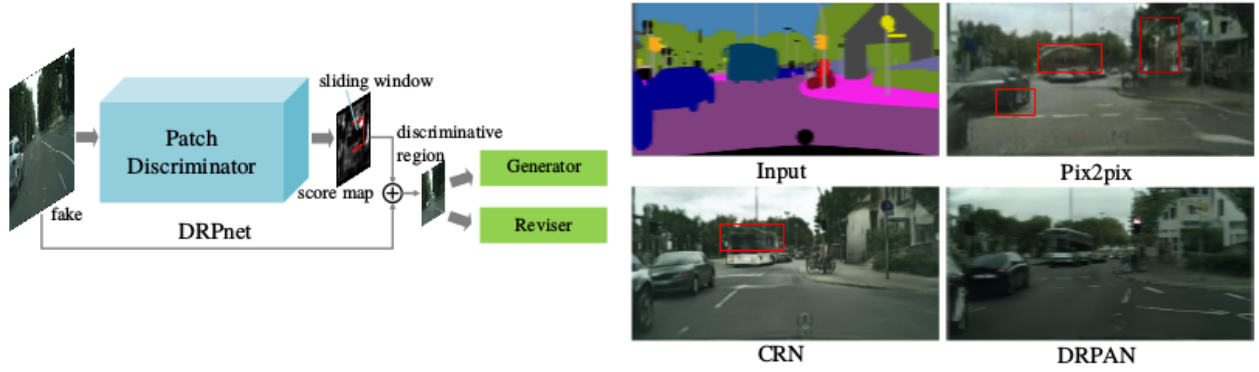


Figure 1. **Left:** Their Discriminative Region Proposal network (DRPnet). **Right:** Synthesized samples compared with previous works on Cityscapes validation dataset [1]. The regions within red window show obvious artifacts or deformation. Their method can synthesize images with clear structure and vivid details. [3]

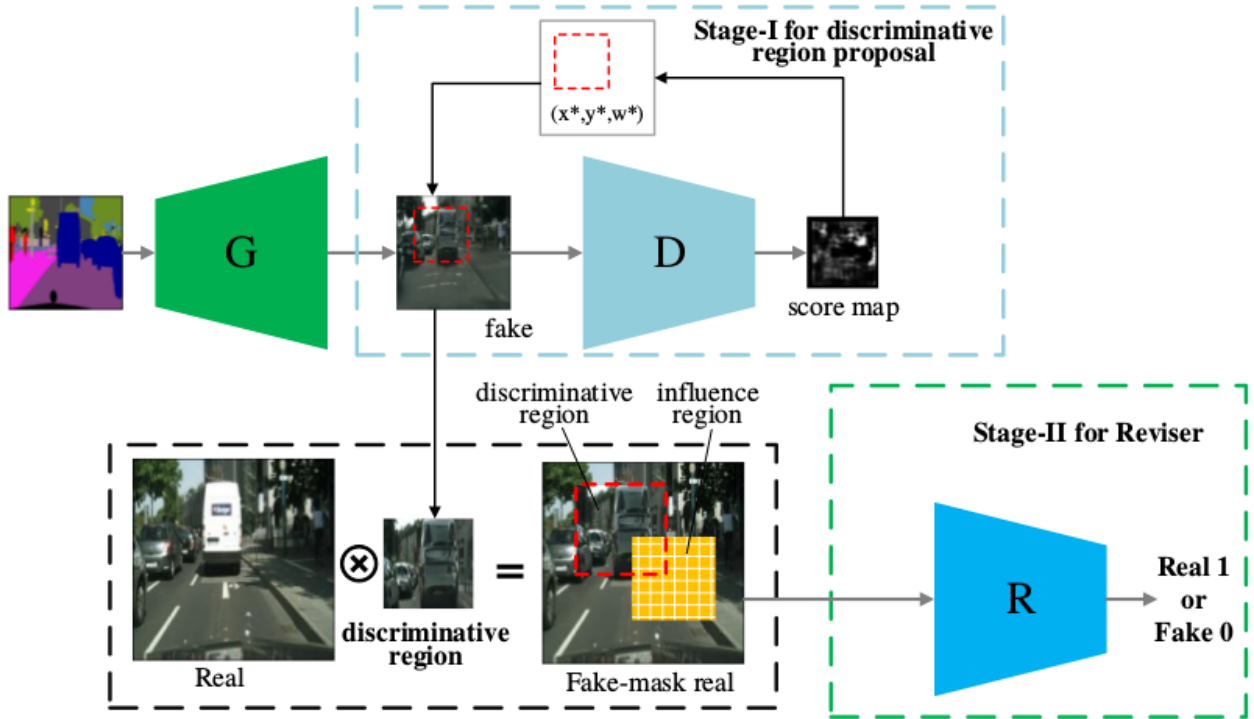


Figure 2. The overall network architecture and data flow of our proposed Discriminative Region Proposal Adversarial Network (DRPAN), which is composed of three components: a generator, a discriminator, and a reviser, and is a unified model for image-to-image translation tasks. [3]

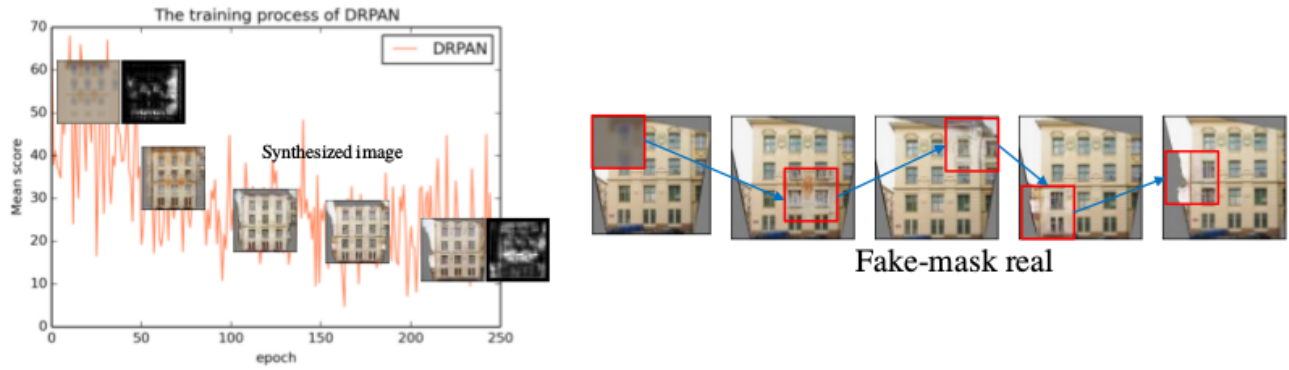


Figure 3. The training process of DRPAN on facades dataset [2]. Left: The plotting curve shows mean value of score map on synthesized samples. Right: Step by step synthesis image on different discriminative regions. [3]