Discriminative Region Proposal AdversarialNetworks for High-Quality Image-to-Image Translation

Wenjie Niu

July 19,2018

This writing is also about the paper of *Discriminative Region Proposal Adversarial Networks for High-Quality Image-to-Image Translation*. And this part mainly analyse the method,which I spent 2 days to understand what's the meaning. Therefore, the method is worth considering.

1. Method

Discriminative Region Adversarial Networks(DRPANs) is composed of three components: a generator, a discriminator and a reviser. The overall network architecture and data flow are illustrated in figure. 1.

1.1. Objective

The original GANs suffer from unstability and mode collapse problem [1, 2]. So some recent works [1, 6, 3] improved the training of GAN. To stably train our DRPAN with high-diversity synthesis ability, we modify DRAGAN [5] as the loss of our reviser R, and use the original objective function for training Patch Discriminator.

$$\begin{aligned}\mathcal{L}_D(G, D_P) = \mathbb{E}_y[logD_P(x, y)]+ \\ \mathbb{E}_{x,z}[log(1 - D_P(x, G(x, z)))]\end{aligned} \tag{1}$$

For reviser R, to distinguish between the very similar real and fake-mask real $y_{mask} = M(G(x, z))$ ($M()$ represents the mask operation), they add a regulariza tion to the loss of reviser as the penalty, which is expressed as

$$\begin{aligned}\mathcal{L}_R(G, R) = \mathbb{E}_y[logR(x, y)] + \mathbb{E}_{x,z}[log(1 - R(x, y_{mask}))] \\ +\alpha\mathbb{E}_{x,\sigma}[\|\nabla_x R(x + \sigma) - 1\|]\end{aligned} \tag{2}$$

where α is hyper parameter, σ is random noise on x, and ∇ indicates gradient.

Previous studies have found it beneficial to mix the GAN objective with a more traditional loss, such as L2 and L1 distance [4, 7]. Considering that L1 distance encourages less blurring than L2 [4], they provide extra L1 loss for regularization on the whole input image and the local discriminative region to generator, which is defined as

$$\begin{aligned}\mathcal{L}_{L_1}(G) = \beta\mathbb{E}_(x, y, z)[\|y - G(x, z)\|_1]+ \\ \gamma\mathbb{E}_{d_r,y_r,z}[\|y_r - F_{DRPnet}(G(x, z))\|_1]\end{aligned} \tag{3}$$

where β and γ are hyper parameters, d_r is the discriminative region, and y_r represents the region on the real image corresponding to the discriminative region on the synthesized image. Then the total loss of generator can be expressed as

$$\begin{aligned}\mathcal{L}(G, D_P, R) = -\mathbb{E}_{x,z}[log(1 - D_P(x, G(x, z)))]- \\ \mathbb{E}_{x,z}[log(1 - R(x, y_{mask}))] + \mathcal{L}_{L_1}(G)\end{aligned} \tag{4}$$

Their proposed model totally contains a generator G, a patch discriminator D_P for DRPnet, and a reviser R. G will be optimized by D_p , R and L_1 . And our full objective function is

$$\begin{aligned}\mathcal{L}(G, D_P, R) = (1 - \lambda)\mathcal{L}_D(G, D_P) + \lambda\mathcal{L}_R(G, R) \\ +\mathcal{L}_{L_1}(G)\end{aligned} \tag{5}$$

2. Conlusion

Thanks to the carefully reading of this paper, I learnt a lot about the procedure of GAN training, I understand the method of the paper,which is very strange for me before. The entire procedure is complex and it worth considering again and again.

References

[1] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein GAN. In *ICML*, 2017. 1

[2] S. Arora, R. Ge, Y. Liang, T. Ma, and Y. Zhang. Generalization and equilibrium in generative adversarial nets. *arXiv preprint arXiv:1703.00573*, 2017. 1

[3] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville. Improved training of wasserstein gans. In *NIPS*, 2017. 1

[4] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *CVPR*, 2016. 1

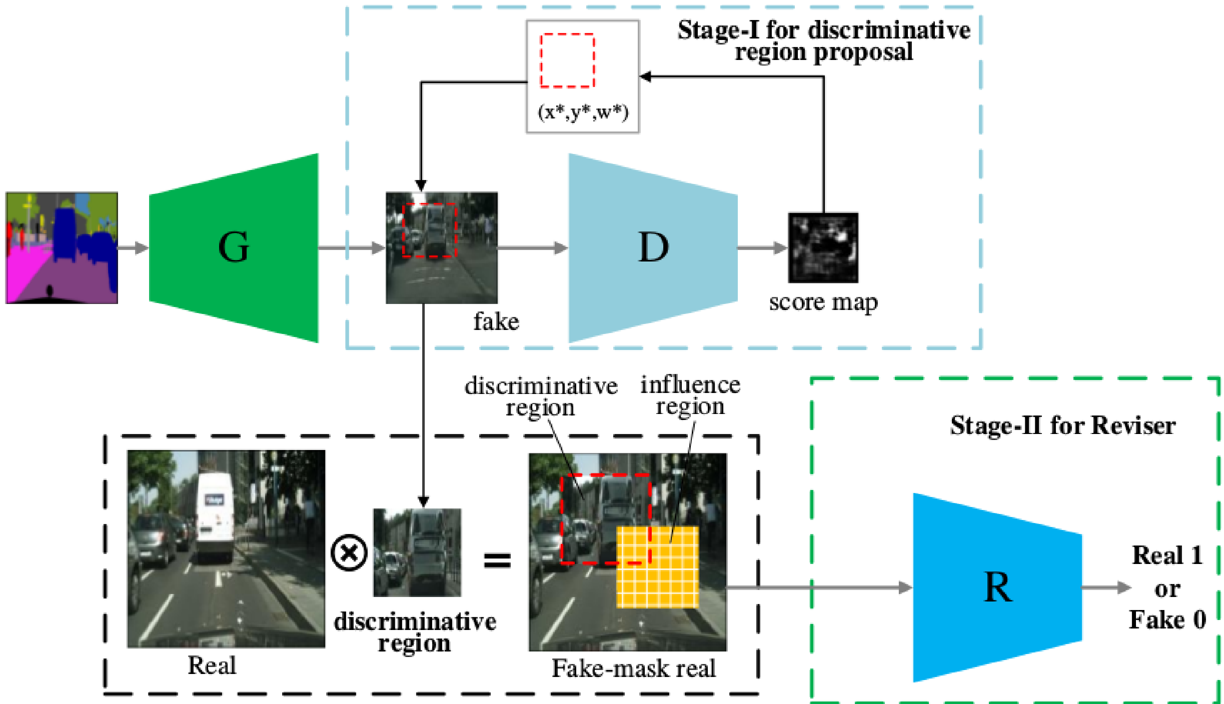

Figure 1. The overall network architecture and data flow of our proposed Discriminative Region Proposal Adversarial Network (DRPAN), which is composed of three components: a generator, a discriminator, and a reviser, and is a unified model for image-to-image translation tasks [9].

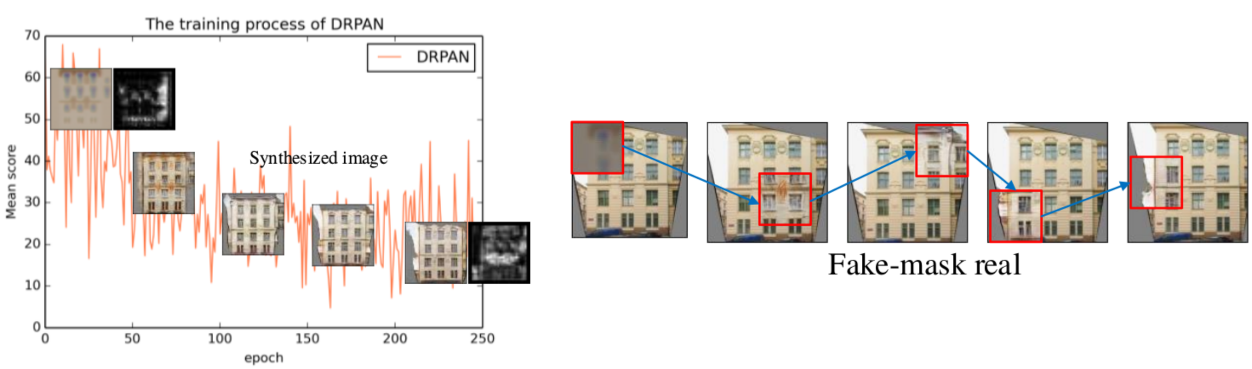

Figure 2. The training process of DRPAN on facades dataset [8]. Left: The plotting curve shows mean value of score map on synthesized samples. Right: Step by step synthesis on different discriminative regions [9].

[5] N. Kodali, J. Abernethy, J. Hays, and Z. Kira. How to train your dragan. *arXiv preprint arXiv:1705.07215*, 2017. 1

[6] G. J. Qi. Loss-sensitive generative adversarial networks on lipschitz densities. *arXiv preprint arXiv:1701.06264*, 2017. 1

[7] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb. Learning from simulated and unsupervised images through adversarial training. In *CVPR*, 2017. 1

[8] R. Tyleek and R. ra. Spatial pattern templates for recognition of objects with regular structure. In *GCPR*, 2013. 2

[9] 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