Text-to-Image-to-Text Translation using Cycle Consistent Adversarial Networks

Jeremy Ma, Satya Krishna Gorti

Abstract

1 Introduction

Text-to-Image synthesis is a challenging problem that has a lot of room for improvement considering the current state-of-the-art results. Synthesized images from existing methods give a rough sketch of the described image but fail to capture the true essence of what the text describes. The recent success of Generative Adversarial Networks [1] indicate that they are a good candidate for the choice of architecture to approach this problem.

However the very nature of this problem is such that a piece of text can map to multiple valid images. The lack of such a direct one-to-one mapping means that traditional conditional GANs [3] cannot be used directly. We draw our inspiration from the recent works of image-to-image translation [2][4] where a cycle consistent GANs have been trained and achieved very impressive results.

We believe that using a cycle GAN for text-to-image-to-text translation will generate better results than existing approaches and give more photo-realistic images. The added advantage of framing the problem in a cycle consistent manner would also mean that the architecture can not only be a text-to-image synthesizing network but also an image captioning network.

Therefore we have two generators G and F. We train a a mapping $G : T_{emb} \mapsto Y$ and inverse mapping $F : Y \mapsto T_{emb}$ in a cycle consistent manner, where T_{emb} is an embedding for the text that describes an image. The generators G and F have their corresponding discriminator D_g and D_f.

2 Related work

3 Project plan

4 Nice-to-halves

References

[1] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[2] Ming-Yu Liu, Thomas Breuel, and Jan Kautz. Unsupervised image-to-image translation networks. In *Advances in Neural Information Processing Systems*, pages 700–708, 2017.

[3] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.

[4] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv preprint arXiv:1703.10593*, 2017.