
CS 391D - Data Mining: A Mathematical Perspective

Project Proposal

Qi Huang (EID:qh2466) Zhiwen Qiang (EID: zq686)

March 7, 2020

1 PROBLEM STATEMENT

CycleGAN proposed by Zhu et.al.[5] is a novel generative adversarial network approach in learning image translation. By using two generator/discriminator pairs and a cycle consistency loss, *CycleGAN* can learn mappings of both direction between two unpaired image domains. We are fascinated by *CycleGAN*'s result, and would like to further investigate potential improvements or new applications of *CycleGAN*. i.e. increase *CycleGAN*'s stability during training, output images with higher resolution and applied the model to a different field.

2 GOALS

In this project, we will firstly try to read and fully understand the paper[5], implement the *CycleGAN* and select a small dataset to train a sample generator to test its effectiveness. Then we will explore potential improvements on *CycleGAN*, or new applications of it. For one, the original *CycleGAN* model only learns mappings between image distribution with the same dimension. It's not clear whether *CycleGAN* can (stably) learn image translation between image collections of different resolution. As a first step, we would like to take LAPGAN[2] and DCGAN [3] to explore transferring image style from low resolution image to high resolution, and vice versa. For another, we would like to explore potential application of *CycleGAN* in other image processing tasks. As far as we know, most state-of-the-art method in image deblurring requires pairwise sharp and blurred images. Yet there is no doubt that it is nearly impossible to get the real blurred image and its real sharp version (instead of the synthetic one) at the same time. If we consider treating blur and sharpness as image "styles", successful image deblurring may be achieved with an unpaired image dataset based on *CycleGAN* or some modified version of it we created.

We summarized our goals we try to achieve in this project below :

- Learn more about *CycleGAN*, potentially through ablation study and implementation.
- Explore image translation between domain of different dimensions, i.e. translate from lower resolution images to higher resolutions.

- Applied the model to a certain field (possibly image deblurring), and analyze the result.

3 DATASET

Some related datasets are as follows:

- facades: 400 images from the CMP Facades dataset [4]
- apple2orange: 996 apple images and 1020 orange images downloaded from ImageNet[1]
- horse2zebra: 939 horse images and 1177 zebra images downloaded from ImageNet
- GOPRO dataset: sharp and blurred images from multiple street views

4 TIMELINE

- By March 23: literature review, model improvement prototyping, dataset search
- By Apr 20: model design, experiment
- By May 4: experiment analysis, paper drafting

REFERENCES

- [1] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.
- [2] E. L. Denton, S. Chintala, R. Fergus, et al. Deep generative image models using a laplacian pyramid of adversarial networks. In *Advances in neural information processing systems*, pages 1486–1494, 2015.
- [3] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.
- [4] R. Š. Radim Tyleček. Spatial pattern templates for recognition of objects with regular structure. In *Proc. GCPR*, Saarbrücken, Germany, 2013.
- [5] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.