Day2Night Image-to-Image Translation

R.Dharia, N.Raichura, A.Sagar, A.Tai

Introduction

Image-to-image translation is a class of vision and graphics problem where the goal is to learn the mapping between an input image and an output image.

It can be applied to a wide range of applications, such as collection style transfer, object transfiguration, season transfer and photo enhancement.

This concept is important in the task of domain adaptation. Transferring characteristics from one image to another is at the core of image-to-image translation.

We address this problem with three different approaches based on the input at hand.

Objectives

- Implement Image-to-Image Translation via mapping images from one domain to the other domain.
- Paired Dataset domain mapping from Day to Night.
- Identify the input instance and apply an appropriate approach.
- Automate Photoshop steps in achieving Day to night translations and create an end-to-end pipeline to achieve the same in a single process.

Approach

Our Methodology includes three different approaches:

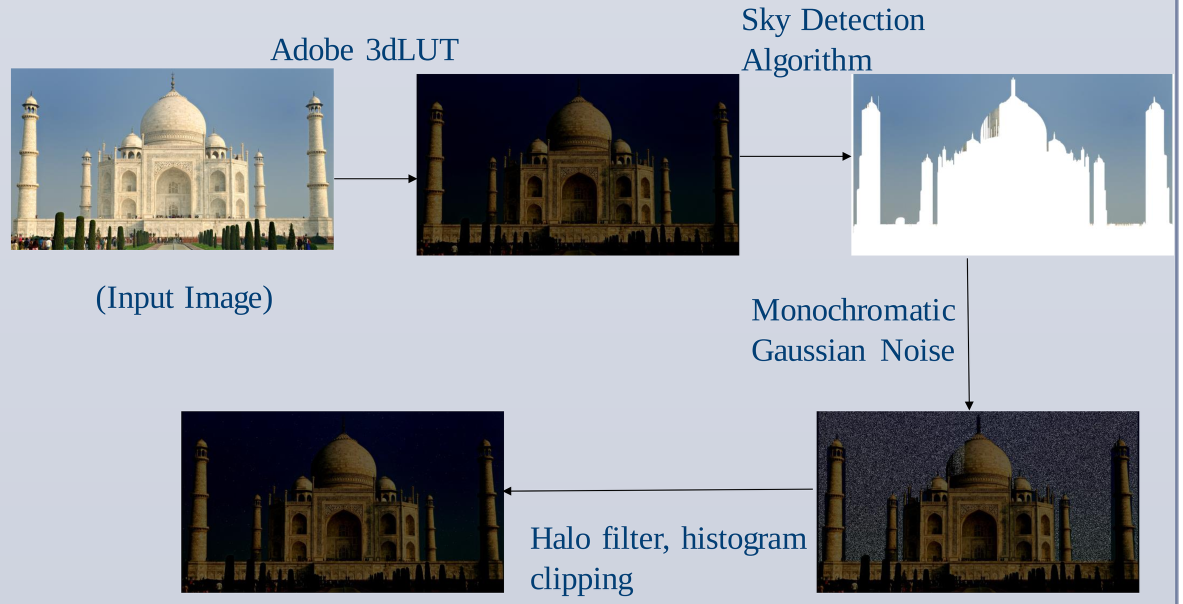
- Naïve Based Approach
- Cycle-GAN
- Conditional Adversarial Network

Naïve Approach

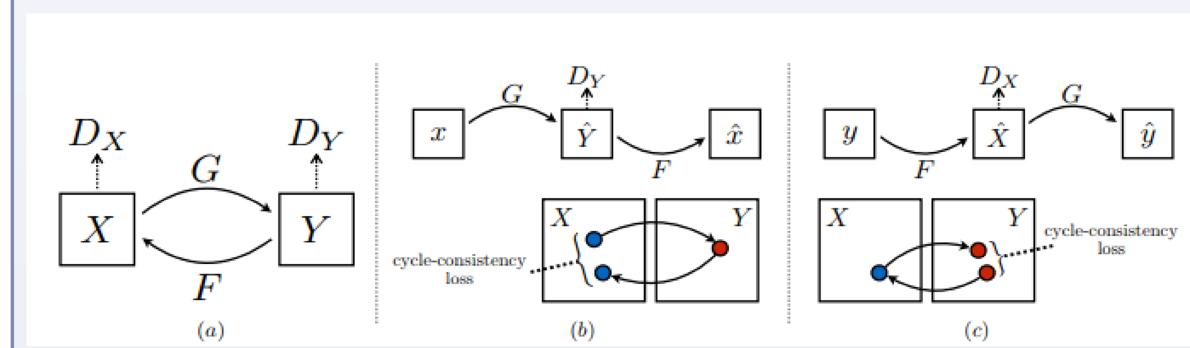
The steps for this approach are as follows:

- Reduce the overall brightness of the image.
- Apply Adobe's Night-from-day 3dLUT for enhancing the blue channel in the image.
- Detect and segment the sky in the image.
- Generate Monochromatic Gaussian noise and then use a neon filter to create some realistic stars in the sky.

Our motivation in this approach is to automate a 15-20 step long process in Adobe Photoshop using Computer Vision techniques. As this approach does not involve a training step, it has certain limitations like no illumination in windows.



Cycle-Generative Adversarial Network[3]



Cycle-GAN Forward and Backward Mapping Functions[3]

Architecture of Cycle-GAN

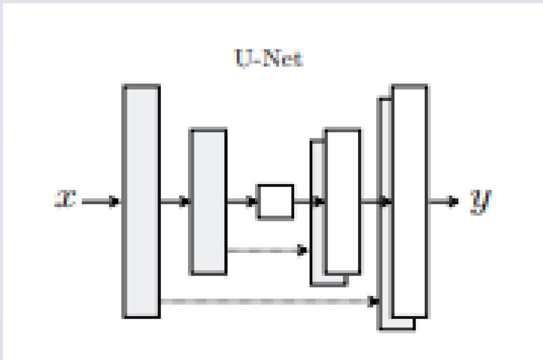
Adapting the principles of GAN[4], the Cycle-GAN architecture consists of two Generators with mapping functions G: X → Y and F: Y → X and associated adversarial Discriminators DY and DX. DY encourages G to translate X into outputs indistinguishable from domain Y, and vice versa for DX and F.

The Loss function is evaluated as follows:

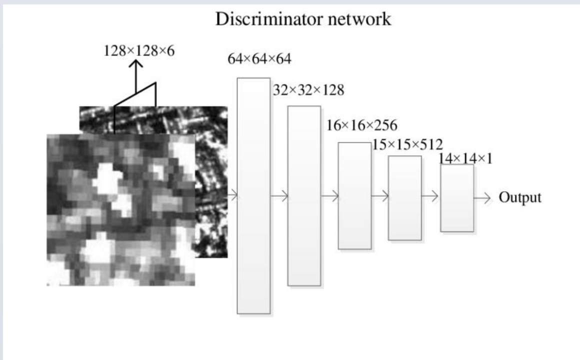
(a) Forward cycle-consistency loss: $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$

(b) Backward cycle-consistency loss: $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$

Conditional Adversarial Network



Generator – U-Net Architecture



Discriminator - PatchGAN

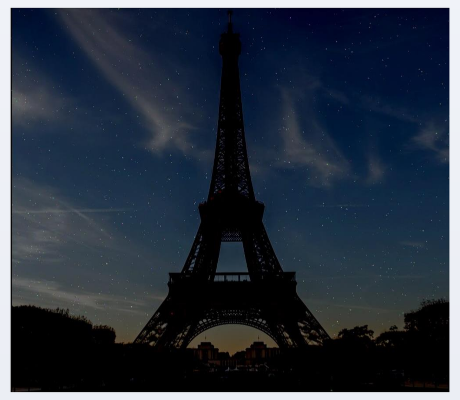
1. U-Net - Encoder Decoder Network with Skip Connections
2. PatchGAN Explanation – Markovian discriminator

- Takes two images - input image and an unknown image.
- Decides if the unknown image was the ground truth or was produced by the generator.
- Works by classifying individual (N x N) patches in the image as "real vs. fake", opposed to classifying the entire image as "real vs. fake".
- Runs convolutionally across the image, averaging all responses to provide the ultimate output.
- N can be much smaller than the full size of the image and still produce good results. This is advantageous because a smaller PatchGAN has fewer parameters, runs faster, and can be applied to arbitrarily large images.

Results

Input Image	Output Image

Naïve Approach

Cycle-Generative Adversarial Network




Conditional Adversarial Network




Conclusion

- Deep Learning based approach produces better outputs for Image to Image Translation on generalized images that span similar context to that of the training dataset.
- Cycle-GAN is slower in terms of training process but generalizes better to unpaired data, while Conditional Adversarial Network performs better with paired image instances.

Future Work

Naïve Based

- Illumination for windows, Street Lamps and any other light source detected in the image.
- Remove shadows cast by sunlight and try to create the ones cast by artificial light.

Cycle-GAN

- Improve training time by ignoring Generator input fed to the discriminator for first few epochs.
- Try different sets of Training Data varying in features, in order to determine efficient subset with fewer image instances.

References

[1] Shen, Yehu, and Qicong Wang. "Sky region detection in a single image for autonomous ground robot navigation." *International Journal of Advanced Robotic Systems* 10.10 (2013): 362.

[2] Neuhausen, Marcel, et al. "Window detection in facade images for risk assessment in tunneling." *Visualization in Engineering* 6.1 (2018): 1.

[3] Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." *Proceedings of the IEEE international conference on computer vision*. 2017.

[4] Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems*. 2014.

[5] Adobe dataset for Nightfromday 3d LUT: <https://github.com/picwellwisher12pk/Presets/blob/master/3DLUTs/NightFromDay.CUBE>

[6] Laffont, Pierre-Yves, et al. "Transient attributes for high-level understanding and editing of outdoor scenes." *ACM Transactions on Graphics (TOG)* 33.4 (2014): 149.

Acknowledgements

1.We would like to thank Prof D. Crandall and Eriya Terada for their valuable guidance.

2.The CycleGAN code was developed by the authors of CycleGAN, Zhu, Jun-Yan and Park, Taesung and Isola, Phillip and Efros, Alexei A. Their code is located at: https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix.