Summer2Monsoon: Using CycleGAN for Image-to-Image Translation

Arnav Garg
Department of Computer Science
UCLA
arnavgarg@cs.ucla.edu

Simranjit Singh
Department of Computer Science
UCLA
simranjit@cs.ucla.edu

Tanmay Sardesai
Department of Computer Science
UCLA
tanmays@cs.ucla.edu

Abstract

1   Introduction

In the past couple of years, there has been a lot of research done on cross domain image to image translation, where an image is taken from one domain and then transformed into an image of the target domain. This is particularly important as a large number of Computer Vision and Machine Learning problems can viewed as an image-to-image translation problem. For example, noise reduction can be considered a mapping between a noisy image to a corresponding noise free image. For many tasks, it is either very difficult or impossible to find paired data, and so in-order to find a mapping between different domains of images, an un-supervised setting is required. The goal of unsupervised image to image translation is to learn the mapping of special characteristics of one image collection and finding how these characteristics could be translated into the other image collection, all in the absence of any paired training examples.

In this paper, we perform image-to-image translation using cycle-consistent adversarial network, also known as CycleGAN, to translate a summer scene into a rainy one and visa versa. We also perform single image de-raining. Some of the applications of this include, helping the film industry shoot movies irrespective of the seasonal cycle and helping self-driving car researchers in data collection by transforming summer images to rainy images, which would allow them to train their models on both these environments simultaneously thereby decreasing the time needed for data collection and training.

Some of the other applications of our research include solving the problem of de-raining. There has been various research done on de-raining, as discussed in Section 2 and we believe that our proposed model would be able to solve the problem by training it to transform rainy images to summer.

2   Related Work

Generative Adversarial Networks (GANs) [2] were introduced by Ian Goodfellow in 2014 and have been very successful in image-to-image translations. GANs have one neural network that generates data while the other discriminates between real and fake (generated) data. Over time both networks get better at their tasks and the generator is now able to mimic the distribution of the input

domain. GAN was used for a variety of applications and it has paved a way for series of GAN-family implementations for different applications.

Image-to-Image translation is the core of this project. The work that we will be exploring in this paper is CycleGAN [1]. Given two domains, X and Y, and two generators $G : X \rightarrow Y$ and $F : Y \rightarrow X$ then we try to achieve a cycle-consistency such that $F(G(X)) \approx X$ and $G(F(Y)) \approx Y$. There is also other work done in the domain of image-to-image translation. The authors of CycleGAN previously proposed pix2pix [5] which used Conditional GANs. pix2pix also has multiple applications but one of the constraints is that pix2pix needs paired data for training. This is sometimes impossible or a really difficult task. For example in our case we would need images taken from the same exact location in 2 different seasons for our dataset. Some related work on image-to-image translation is also done by Nvidia using coupled GAN [3]. Concurrent to the work done in CycleGAN, Yi et al [6], published a paper on dual-GAN.

Single Image de-raining is a difficult problem to solve due to its ill-posed nature and unlike video based methods, images do not provide any temporal information. Traditionally it has been approached as considering an image y to be the sum of rain streak r and background image x, i.e. $y = x + r$. This is why most of the research approaches try to decompose the image into a background image and rain streak image. One of the earliest methods is sparse coding based [a] where image is decomposed using learned dictionary atoms that can sparsely represent two components clearly. Gaussian mixture models based priors [c] have also been used in image decomposition frameworks that can capture different orientations and scales of rain streaks. CNNs have also been employed to directly learn non linear mapping between synthesized images and ground truths successfully for image de-raining [d] [e] [f]. Conditional GANs are also used to achieve de-raining [4]. The work in this, and other papers on de-raining, doesn't completely solve the problem as they only remove rain from the image but some aspects of the image still look the same as the sky will still look cloudy and the roads will look wet even after removing the rain. These models only work on monsoon to summer translations but not vice versa.

3 Implementation

4 Results

We performed two different types of studies to help evaluate our results. For our image to image translation, we performed a perceptual study where we created a survey and asked participants to classify if images were real or fake and for the Single Image De-raining, as we had ground truth for these images, we computed the Structural Similarity Index Scores (SSIM) between to de-rained images and the ground-truth to evaluate how well our model performed. We have discussed both of these studies on in the subsection below.

4.1 Perceptual Study

Type	Total	Real	Fake
Sunny	0.591	0.825	0.543
Rainy	0.597	0.716	0.578

Table 1: Survey Results

For evaluating our image to image translation model, we conducted two online surveys, one with only summer images and the other with rainy images, in which we showed multiple fake (our model generated) and real images to the participant. In both the surveys, the participants were shown each image for 2 seconds, and then were asked to identify if the image they saw was fake or real. The results of the surveys are shown in table 1.

@arnav explain the table

4.2 SSIM

@arnav do this

5   Limitation & Discussions

6   Conclusion & Future Work

We have employed the CycleGAN model for the purpose of translating images from summer to monsoon and vice-versa. We were successfully able to train the model to minimize the proposed loss function on our final dataset collected from various sources. We conducted a survey to see if users can identify fake images as real and achieved results comparable to what is discussed in CycleGAN paper. We also employed CycleGAN to de-rain the images as well as add rain to images and evaluated the results based on SSIM scores and compared to some of the previous works.

In future, we want to explore the model used in [3] that involves shared latent space and believe that it would give better results than our current model. In our model the generator loss did not converge but overall loss did and that is something we are keen on analyzing. The loss proposed in CycleGAN is weighted sum of different losses and we believe that these weights can be tuned to achieve even better results.

References

[1] J. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In International Conference on Computer Vision (ICCV), to appear, 2017

[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Ben- gio. Generative adversarial nets. In NIPS, 2014.

[3] M.-Y. Liu, T. Breuel, and J. Kautz. Unsupervised image-to-image translation networks. arXiv preprint arXiv:1703.00848, 2017

[4] H. Zhang, V. A. Sindagi, and V. M. Patel. Image de-raining using a conditional generative adversarial network. arXiv preprint arXiv:1701.05957, 2017.

[5] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In CVPR, 2017

[6] Z. Yi, H. Zhang, T. Gong, Tan, and M. Gong. Dualgan: Unsupervised dual learning for image-to-image translation. In ICCV, 2017

[7] "'Phantom Auto' will tour city". The Milwaukee Sentinel. Google News Archive. 8 December 1926. Retrieved 2013.

[8] "Carnegie Mellon". Navlab: The Carnegie Mellon University Navigation Laboratory. The Robotics Institute. Retrieved 2014.

[9] Hawkins, Andrew J. "Inside the Lab Where Waymo Is Building the Brains for Its Driverless Cars." The Verge, The Verge, www.theverge.com/2018/5/9/17307156/google-waymo-driverless-cars-deep-learning-neural-net-interview, 2018.