Image Colorization using CycleGAN

Zahil Shanis
CSE, University at Buffalo
Person Number: 50291495
zahilsha@buffalo.edu

Yash Narendra Saraf
CSE, University at Buffalo
Person Number: 50290453
ysaraf@buffalo.edu

Sai Varun Alapati
CSE, University at Buffalo
Person Number:50290571
saivarun@buffalo.edu

Abstract

Image-to-image translation is a class of computer vision problem where the goal is to learn the mapping between an input and output image domain using a training set of images from both domains. With the advent of deep learning techniques, several papers have been published which studies image-to-image translation using deep neural networks. Artificially coloring grayscale images using deep learning is an application of image to image translation which has produced several compelling results recently in the literature. CycleGAN[4], a GAN based unsupervised image-to-image translation architecture proposed by Zhu et. al in 2017, have been used for a wide variety of geometric and color transformations. In this project we investigate a promising method for unsupervised image colorization task using CycleGAN[4] architecture using CIFAR-10[1] and Places [2] dataset.

1 Introduction

Image Colorization is an thoroughly studied problem even before the use of deep learning based techniques. Image colorization was initially done using two main methods i.e. interactive colorization methods, and automatic colorization methods[5]. The interactive methods required the user to mark color to different surfaces which are then smoothly propagated across the entire image based on an optimization framework. Many earlier automatic colorization techniques focused on using image processing techniques like Gabor transforms and SIFT keypoint detection [6].

In 2016 Ryan Dahls [7] CNN based architecture for automatically colorizing images used a network that integrated ImageNet-trained layers from VGG16 with an autoencoder like system using residual connections. In terms of results, Dahls system performed extremely well in realistically colorizing foliage, skies, and skin. Here Dahl formulated image colorization as a regression problem where the objective function was to minimise the a sum of Euclidean distances between each pixels blurred color channel values in the target image and predicted image. The regression methods result in inconsistencies if multiple colored copies of same object are present in the dataset as the network would try to average those colors out resulting in a subdued mixture of possible colors.

In 2016, Richard et al. [11] proposed a fully automatic approach by posing it as a classification task and use class-rebalancing at training time to increase the diversity of colors in the result. Deep Colorization [10] of Cheng, Yang, and Sheng (2016) describes a neural network for colorization which input a grayscale image and output a color image in YUV color space. This paper also formulates colorization as a regression problem, with the loss function as a least squares minimization between the predicted and ground-truth color pixel values.

Image-to-image translation [9] has gained considerable attention due to the recent impressive progress based on generative adversarial networks (GANs). Generative Adversarial Networks [12] proposed by Ian Good fellow in 2014 is composed of two smaller networks called the generator and discriminator. The generators task is to produce results that are indistinguishable from real data whereas discriminator tries to classify whether a sample came from the generators model distribution or the original data distribution. Both of these sub networks are trained simultaneously until the generator is able to consistently produce results that the discriminator cannot classify.

Several papers have been published which uses GAN architecture for various image transformation tasks including image-to-image translation. Image-to-Image Translation with Conditional Adversarial Networks(pix2pix) of Isola et al.[12] introduces conditional adversarial networks as a general-purpose solution to image-to-image translation problems like reconstructing objects from edge maps, style transfer, etc.

Even though pix2pix is capable of generating excellent translations, it requires paired input data for training. The two image spaces that we want to learn to translate between needed to be pre-formatted into a single X/Y image that held both tightly-correlated images. This could be infeasible for the cases where we don't have one-to-one matches between the two image profiles (eg: image colorization). CycleGAN [4] proposed by Zhu et. al overcomes this issue.

The key idea behind CycleGANs is that they can build upon the power of the PIX2PIX architecture, but allow you to point the model at two discrete, unpaired collections of images. The way CycleGANs are able to learn such great translations without having explicit X/Y training images involves introducing the idea of a full translation cycle to determine how good the entire translation system is, thus improving both generators at the same time.

CycleGAN framework have been applied for several different image-to-image translation problems, including artists styles and photos, apples and oranges, zebras and horses, winter and summer, and maps and aerial photographs and have achieved excellent results. We plan to use CycleGANs as the base architecture for our image colorization translation problem.

2 Timeline

A tentative timeline for our project is as follows:

Table 1: Project Timeline.

15 March	Literature Review.
22 March	Set up CycleGAN architecture for image colorization..
24 March	Train the model for image colorization using default set of parameters..
30 March	Verify model performance and fix mode collapse issues if there are any..
6 April	Parameter tuning for model architecture..
13 April	Train the architecture on a different dataset and compare the results..
19 April	Performance Analysis: Compare the results of our architecture with that of existing architectures.
25 April	Further improvements and debugging..
3 May	Documentation, Report and Poster preparation..
9 May	Project Presentation.

References

[1] Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton, *CIFAR-10 and CIFAR-100,*. https://www.cs.toronto.edu/ kriz/cifar.html, 2009

[2] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva, *Learning Deep Features for Scene Recognition using Places Database,*. Advances in Neural Information Processing Systems 27 (NIPS), 2014

[3] Alex Krizhevsky, *Learning Multiple Layers of Features from Tiny Images,*. https://www.cs.toronto.edu/ kriz/learning-features-2009-TR.pdf, 2009

[4] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros, *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks,* International Conference on Computer Vision 2017

[5] V. Konushin, and V. Vezhnevets. *Interactive Image Colorization and Recoloring based on Coupled Map Lattices,*. 2006 .

[6] Raj Kumar Gupta, Alex Yong-Sang Chia, Deepu Rajan1 Ee Sin Ng, and Huang Zhiyong, *Image Colorization Using Similar Images,*. MM '12 Proceedings of the 20th ACM international conference on Multimedia Pages 369-378 .

[7] Ryan Dahl, *Automatic Colorization,*. https://tinyclouds.org/colorize/, January 2016.

[8] Medium Blog: Lee Cohn, *Artificial Colorization of Grayscale Satellite Imagery via GANs: Part 1,*. https://medium.com/the-downlinq/artificial-colorization-of-grayscale-satellite-imagery-via-gans-part-1-79c8d137e97b, August 2017

[9] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros, *Image-to-Image Translation with Conditional Adversarial Networks,*. https://medium.com/the-downlinq/artificial-colorization-of-grayscale-satellite-imagery-via-gans-part-1-79c8d137e97b, August 2017

[10] Zezhou Cheng, Qingxiong Yang, Bin Sheng, *Deep Colorization,* International Conference on Computer Vision, 2015

[11] Zhang R., Isola P., Efros A.A, *Colorful Image Colorization,* ECCV 2016. Lecture Notes in Computer Science, vol 9907. Springer, Cham

[12] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, *Generative adversarial nets,* Advances in Neural Information Processing Systems pages 26722680, 2014.