Transforming aerial images to underwater images

A Project Report

submitted by

POONDLA JYOTHI PRAKASH EE14B045

in partial fulfillment of requirements
for the award of the degree of

BACHELOR OF TECHNOLOGY

DEPARTMENT OF Electrical Engineering
INDIAN INSTITUTE OF TECHNOLOGY MADRAS

MAY 2018

THESIS CERTIFICATE

This is to certify that the thesis titled **Transforming aerial images to underwater images**, submitted by **Poondla Jyothi Prakash, EE14B045**, to the Indian Institute of Technology Madras, for the award of the degree of **Bachelor of Technology**, is a bona fide record of the research work done by him under our supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Prof. Kaushik Mitra
Project Guide
Assistant Professor
Dept. of Electrical Engineering
IIT Madras, 600036

Place: Chennai
Date: 10 May 2018

ACKNOWLEDGEMENTS

I am very grateful to my advisor Prof. Kaushik Mitra for very kindly, patiently accepting me and helping me do this project. I am also grateful to my parents and to my teachers who taught me academic knowledge as well as spiritual knowledge to maintain proper consciousness while doing this project. I dedicate this humble effort of my project work as a responsible duty and as an offering to Supreme Lord Krishna and to my Gurudev and to all my teachers.

ABSTRACT

Underwater image processing is in great demand, while the research is far from enough. The unrestricted natural environment makes it a challenging task. Deep Learning methods are the state-of-the-art machine learning techniques used for the underwater imaging for the age of big data. As they are data-driven techniques, there is a need for a much generalised underwater dataset to carry out experiments. In this report I have presented my work in the direction of creating a simulated underwater dataset to be used for experiments.

TABLE OF CONTENTS

CHAPTER 1
Introduction

Underwater imaging is a prominent field of Computational Photography. It focuses on several challenging problems posed by underwater images such as dehazing, deskewing, decolorization etc., There have been several papers published to address these challenges using deep learning methods. The accuracy and the re-usability of these methods very much depends on the right dataset used for evaluating underwater image restoration methods.

To fulfill this need there have also been papers published to actually produce simulated datasets. These impose several specific initial conditions such as controlling the image degradation, or a particular type of surface waves applied or a particular type of haze applied or the depth of the images considered etc., Therefore the datasets generated through these methods will be very specific to the initial conditions.

There are some deep learning related papers which address this problem to some extent. Some of them are discussed in further sections. They either propose a method to transform an underwater image directly into a decolorized or dehazed version, or they simulate a dataset using some specific test samples and train a network with that dataset so that they can be applied to test samples.

The challenge with obtaining the generalised underwater dataset is that we cannot get the ground truth for the underwater images directly. We need to either simulate them with initial conditions as I have previously mentioned or find ways to obtain the dataset with ground truths, because deep learning methods very well require the ground truths of the images to train the networks.

I have therefore taken up this as my problem statement to create a generalised underwater dataset with ground truths (their aerial versions). I have discussed in chapter 3 about my work in detail as an application of cycleGAN and of some papers which use cycleGAN for their work. In the next chapter I will discuss about the prior works which inspired me to move further.

CHAPTER 2

Prior Work

In this chapter I will briefly discuss the prior works carried out in the direction of creating underwater datasets.

2.1 A paper by OCEANS-2016, Shanghai [1]:

In this paper they try to particularly make an experimental setup to produce an environment to control image degradation and then obtain the dataset for aerial objects which are submerged in water.

The main features of this paper are:

- ➢ They control the property of **turbidity** in regard to image degradation.
- ➢ They apply these conditions on the scenario of 3D objects which represent the seabed characteristics.

(m) Clear(I_0)) (n) I_4 - 20 ml (o) I_{10} - 58 ml (p) I_{16} - 110 ml

Fig 2.1 3D Turbid dataset sample images with different turbidity levels Image courtesy [1]

The problems with this approach are:

- ➢ The dataset formed is only to specific objects submerged in water.
- ➢ And only the property of turbidity is taken into account. There is no generalised dataset.

2.2 WaterGAN:

It a generative adversarial network (GAN) for generating realistic underwater images from in-air image and depth pairings in an unsupervised pipeline used for color correction of monocular underwater images.

The main features of this work are:

- ➢ It uses monocular underwater image samples for training the rgb-d dataset.
- ➢ Generates synthetic underwater images
- ➢ Can be applied to train a network for testing on specific test samples.

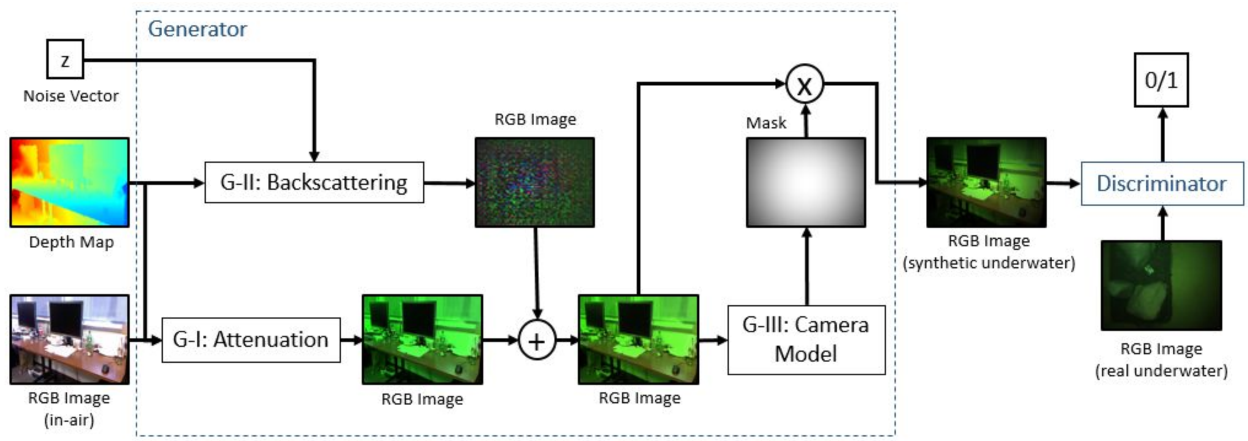


Fig 2.2 WaterGAN Architecture Image courtesy [2]

Till now I have presented the work done by OCEANS-2016 and WaterGAN. In the next chapter we will see how I have taken the work further.

CHAPTER 3

Transforming aerial images to underwater images

3.1 WaterGAN:

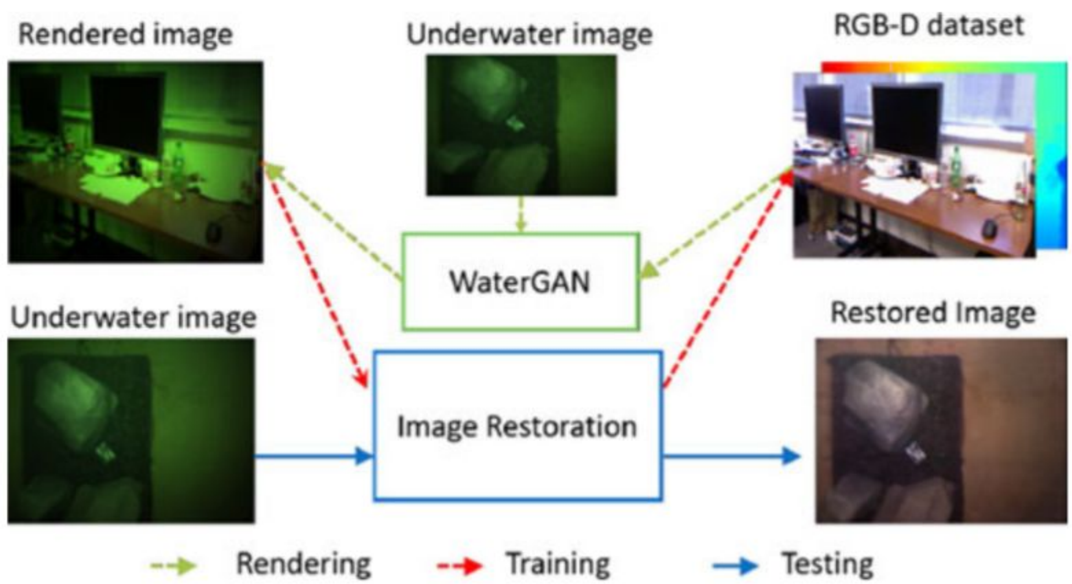
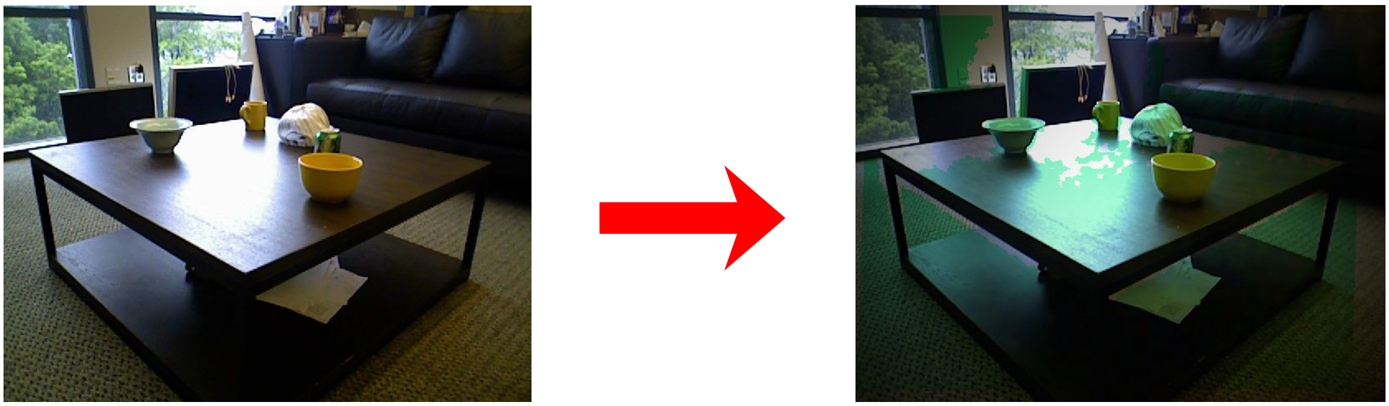


Fig 3.1 WaterGAN Overview   Image courtesy [2]

In WaterGAN, they used the sample underwater images from single video taken at a place. And then when the synthetic images are created they trained a network to decolourise specific test samples of underwater images.

We thought of using a generalised dataset of underwater images with variegatedness in haze, color, turbidity and location etc., in place of the sample underwater images from a single video. But when we performed the experiment with the WaterGAN code and with the same given datasets and configuration, the result is shown below.



Input aerial image Synthetic underwater image

Fig 3.2 WaterGAN experiment result sample

We could not produce even the synthetic underwater images as proposed in the paper. There is no clear colorization and most of the green color is concentrated at the center only. So we moved further for some other ways to generate datasets.

3.2 CycleGAN:

CycleGAN is an adversarial network which is used for image style transformations.

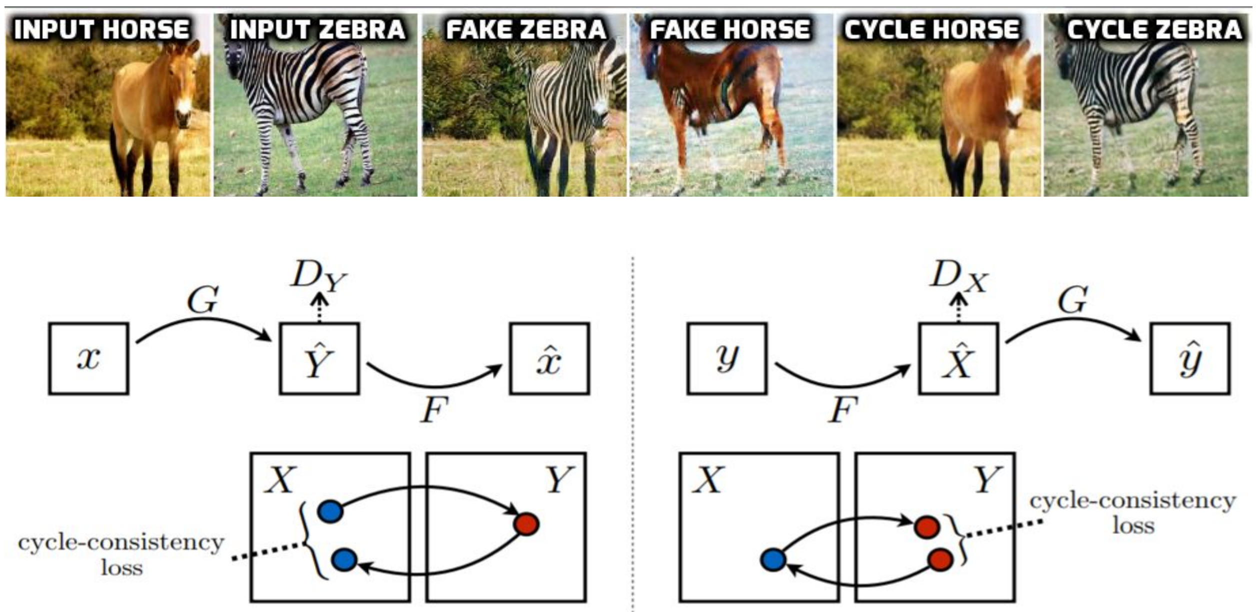


Fig 3.3 CycleGAN framework and a sample style transform (horse → zebra)

Image courtesy [3]

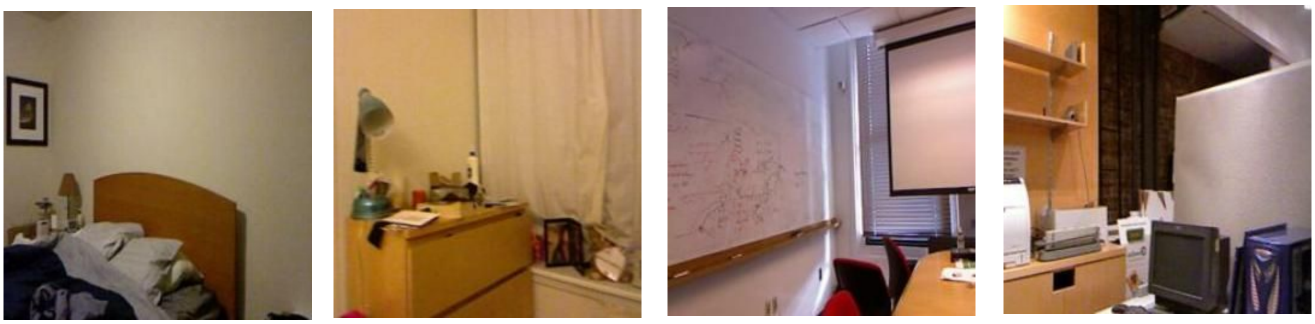
This takes in two datasets of different style images (ex. Horses and Zebras) and convert them from one to another and then convert back cyclically to generate the original images again.

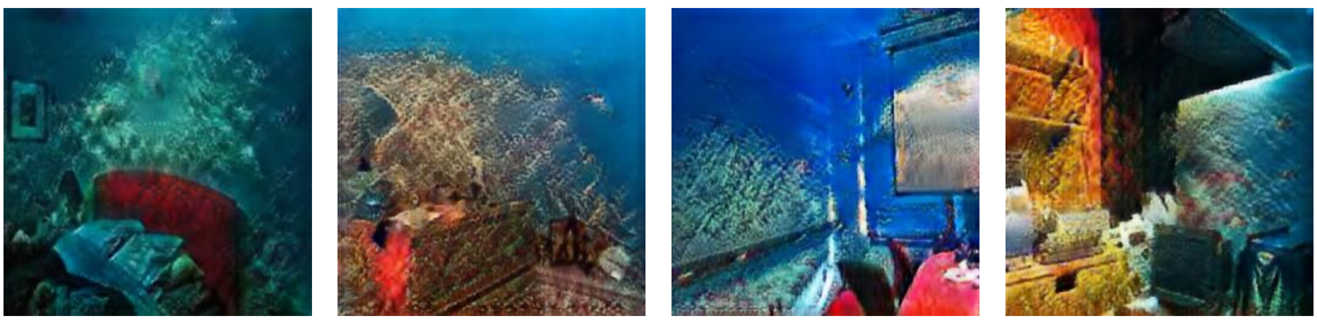
Inspired by "Emerging from Water" [4], which used cycleGAN with SSIM loss to decolorize the underwater images, we used this approach by giving one set of aerial images from any rgb-d dataset and another set of random underwater images to generate a dataset. We collected around 1700 such images from net and trained the network without SSIM loss. The results obtained are as follows:

3.3 Results:

Without SSIM loss factor:

Aerial Images





Underwater Images

But the results are not preserving the structure of original ground truth aerial images. So we improved the images with including SSIM loss factor.Below is the formula for SSIM loss and total loss factors included.

SSIM Loss:

$$\text{SSIM}(p) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

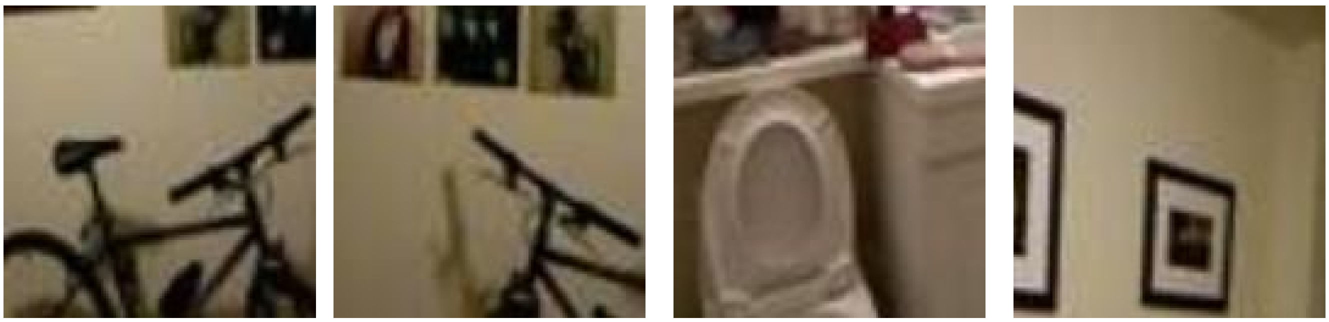
$$L_{\text{SSIM}}(x, G(x)) = 1 - \frac{1}{N} \Sigma_{p=1}^{N} (\text{SSIM}(p))$$

Total Loss:

$$L_{\text{loss}} = \lambda_1 L_{\text{GAN}}(G, D_Y, X, Y) + \lambda_2 L_{\text{cyc}}(G, F) + \lambda_3 L_{\text{SSIM}}(x, G(x)).$$

With SSIM loss factor:

Aerial images



Underwater images

CHAPTER 4

Conclusion and Future Work

We tried to use WaterGAN, but the results of the code for the given configurations itself are not proper. So we being inspired by “Emerging from Water” used CycleGAN with aerial images and randomly collected underwater images. The results obtained without SSIM loss are not so desirable. So we introduced SSIM loss with varying weightage to produce the depicted results.

Although we have successfully produced some desirable output samples, we do not have a dataset as a whole ready yet. There is a lot of work to be carried out in this direction for future scope.

References

[1] Duarte, A., Codevilla, F., Gaya, J.D.O. and Botelho, S.S., 2016, April. A dataset to evaluate underwater image restoration methods. In *OCEANS 2016-Shanghai* (pp. 1-6). IEEE.

[2] Li, J., Skinner, K.A., Eustice, R.M. and Johnson-Roberson, M., 2018. WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robotics and Automation Letters*, *3*(1), pp.387-394.

[3] Zhu, J.Y., Park, T., Isola, P. and Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv preprint arXiv:1703.10593*.

[4] Li, C., Guo, J. and Guo, C., 2018. Emerging from Water: Underwater Image Color Correction Based on Weakly Supervised Color Transfer. *IEEE Signal Processing Letters*.