BLIND UNDERWATER IMAGE RESTORATION USING OPERATIONAL CYCLE-GANS

OZER CAN DEVECIOGLU

Master's Thesis

Graduate School
İzmir University of Economics
İzmir
2022

BLIND UNDERWATER IMAGE RESTORATION USING OPERATIONAL CYCLE-GANS

OZER CAN DEVECIOGLU

A Thesis Submitted to
The Graduate School of Izmir University of Economics
Master Program in Electrical and Electronics Engineering

Izmir
2022

ABSTRACT

BLIND UNDERWATER IMAGE RESTORATION USING OPERATIONAL CYCLE-GANS

Devecioglu, Ozer Can

Master Program in Electrical and Electronics Engineering

Advisor: Prof. Dr. Turker Ince

July, 2022

Monitoring the underwater scenes is important for both continuation of ecological life and exploration of these underwater areas. However, because of the lights' transmission characteristics and water physical attributes, underwater images suffer from artifacts like color distortion, scattering, flickering, poor visibility, and uneven illumination. The set of these abovementioned artifacts with varying severities makes underwater images hard to monitor. Despite numerous studies that have attempted underwater image restoration, they obviously fail to perfectly restore real underwater images corrupted with a random blend of artifacts with their simple restoration models. In this paper, we propose a promising approach for restoring blind underwater images using novel operational cycle-consistent generative adversarial networks (Op-GANs), where the signal quality may be improved regardless of the kind or degree of the artifacts degrading the underwater image. This

is the first study to utilize 2D operational layers with higher learning capacity in powerful cycle-GANs for processing 2D underwater images. Utilizing one of the largest benchmark underwater image datasets from the Large-Scale Underwater Image dataset (LSUI), the proposed technique achieved promising results. The detailed evaluations, both quantitative and qualitative, show that the suggested method outperforms the other competing techniques in the literature.

Keywords: Machine Learning, Convolutional Neural Networks, Operational Neural Networks, Generative Adversarial Networks, Underwater Images, Underwater Image Restoration,

ÖZET

SU ALTI GORUNTULERININ OPERASYONEL DONGU TUTARLI UZETKEN KARSIT AGLAR ILE GOZU KAPALI RESTORASYONU

Devecioglu, Ozer Can

Elektrik ve Elektronik Mühendisliği Yüksek Lisans Programı

Tez Danışmanı: Prof. Dr. Turker Ince

Temmuz, 2022

Sualtı görüntülerinin izlenmesi hem ekolojik yaşamın devamlılığı hem de bu alanların araştırılması açısından önemlidir. Bununla birlikte, ışığın sualtında iletim özellikleri ve suların fiziksel özellikleri gibi sebepler, su altı görüntülerinde renk bozulması, saçılma, titreme, zayıf görüş ve eşit olmayan aydınlatma bozulmalarına yol açmaktadır. Yukarıda bahsedilen ve değişen şiddetteki bu artefaktlar kümesi, su altı görüntülerini izlemeyi ve bilgi çıkartımını anlaşılmaz kılmaktadır. Sualtı görüntü restorasyonu girişiminde bulunan çok sayıda çalışmaya rağmen, basit restorasyon modelleri ile artefakt karışımı ile bozulan gerçek sualtı görüntülerini restore etmekte başarılı olamamışlardır. Bu çalışmada, sualtı görüntülerini rastgele bozan gürültü ve artefaktların türüne ve şiddetine bağlı olmadan resimlerin kalitesinin arttırabilecek, operasyonel döngü ile tutarlı üretken karşıt ağları kullanarak sualtı görüntülerini restore edecek için umut verici yeni bir

yöntem sunuyoruz. Bu çalışma iki boyutlu sualtı resimlerini işlemek amacıyla yüksek öğrenme kapasiteli iki boyutlu operasyonel katmanları döngü ile tutarlı üretken karşıt ağları içinde kullanan ilk çalışmadır. Önerilen yeni yaklaşım büyük ölçekli sualtı görüntü veri setini (LSUI) kullanarak kapsamlı bir şekilde değerlendirilmiştir. Detaylı nicel ve nitel değerlendirmeler, yöntemimizin, literatürde mevcut yöntemlere göre üstünlüğünü göstermektedir.

Anahtar Kelimeler: Makine Öğrenimi, Evrişimsel Sinir Ağları, Operasyonel Sinir Ağları, Üretken Karşıt Ağları, Sualtı Görüntüleri, Sualtı Görüntü Restorasyonu,

Dedicated to my mother and my father…

ACKNOWLEDGEMENTS

I would like to express my endless thanks to Prof. Dr. Turker Ince, who guided me in studies, did not spare his support and efforts, encouraged me and I will always be honored to be his assistant.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

CHAPTER 1: INTRODUCTION

Almost 71% of the earth's surface is covered with water. Because of the reasons like biological/archeological research, wreckage exploration, and damage control of underwater pipelines and cables, there has been always a high demand for exploring these underwater areas. Recently, due to the prevalence of the use of Autonomous underwater vehicles (AUV) and remotely Operating Machines (ROV), the ability of image capturing has increased in areas that are difficult to reach under the sea. However, because of the light's transportation characteristics and the water turbidity or any other interferences, underwater images can be corrupted with severe artifacts such as color deformation, scattering, uneven illumination, flickering and poor visibility.

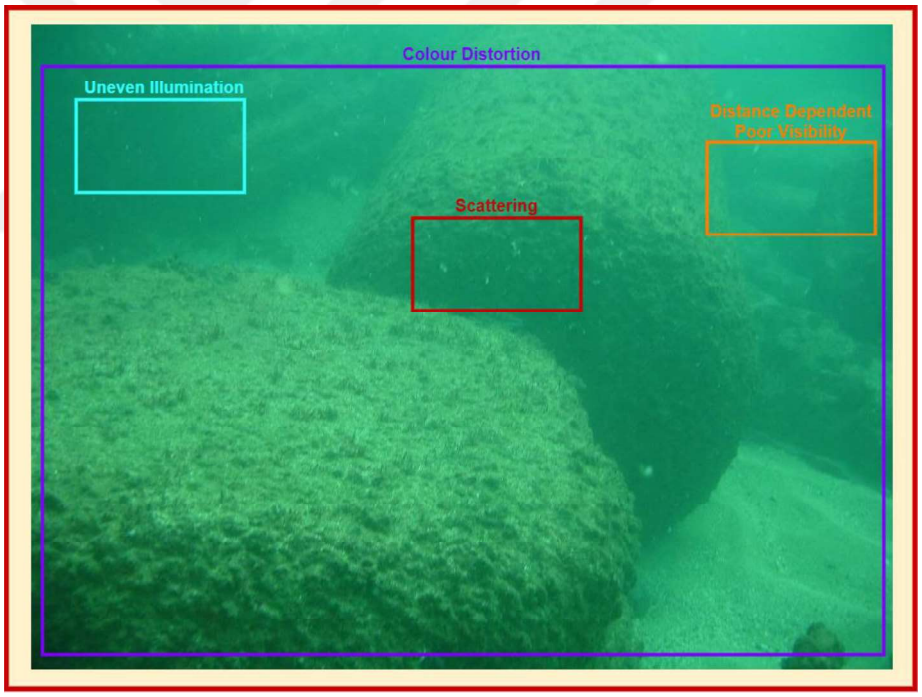

Figure 1. Sample corrupted underwater image

Figure 1 shows a sample underwater image from the LSUI dataset (Peng et al., 2021). The underwater image is distorted by several of the aforementioned distortions, as seen in the figure. As a consequence, most of the

details on the image are diminished. The detailed description of the artifacts of corruption and their sources are presented in the next chapter.

(Liu et al., 2019), proposed an underwater image enhancement method using an underwater ResNet model. They first created synthetic paired training data using CycleGANs to train the model. They used the Underwater ResNet model with 20 convolutional layers to train paired synthetic images. They obtained 43.9 and 0.86 PSNR and SSIM results respectively. (Huimin Lu et al., 2017), proposed an underwater image enhancement convolutional neural network (CNN) model based on underwater scene prior, called UWCNN. First, they synthesized the training data which covers a diverse set of water types and degradation levels using UWCNN. Then, they use 10-Layer CNN to enhance underwater images. (Li et al.,2021) proposed a novel self-similarity-based method for de scattering and super-resolution (SR) of underwater images. They used a self-similarity-based SR algorithm with a convex fusion rule to dispel the noise of underwater images with high turbidity. The proposed algorithm demonstrates a consistent improvement with 21.4 PSNR. (Han et al., 2021), proposed a deep supervised residual network for underwater image enhancement and restoration. To train the proposed system, they used clear in-air images and synthetic underwater degraded images. The proposed method uses residual dense blocks for extracting features to enhance feature utilization. Next, it reduces the semantic disparities between low-level features and high-level features by using the U-Net model. Finally, they use a supervision mechanism to train the system. They obtain 36.2 PSNR.

Most of the works above-mentioned use deep learning models with datasets that are artificially synthesized artifacts by disturbing only the color harmony of clean outside pictures. Eventually, deep learning methods learn to suppress one artifact of the noise. However, in real-life underwater images, there is usually more than one artifact to corrupt the image. Also, most of the real corruption artifacts can not be synthesized by image processing algorithms. Therefore, this method cannot reconstruct the images. As a consequence, In this study, we take this issue as a blind restoration approach to avoid any prior assumptions about the types and severity of artifacts. The proposed method will learn to convert corrupted

underwater images to clean underwater pictures with an unsupervised approach using Operational Cycle-GANs.

Cycle-Consistent Adversarial Networks (Cycle-GANs) (Zhu et al., 2017) are developed for image-to-image translations on the unpaired datasets. Cycle-GANs learn the aspects of domains individually and learn to translate between them while not violating the features not directly related to either domain. To achieve the above-mentioned goal, in this study, we used unpaired clean and corrupted underwater images from the LSUI dataset to train the data Cycle-GANs. The quality of the underwater photos will be enhanced using Cycle-GANs by translating corrupted domain characteristics to clean domain characteristics while preserving details of the images. Self-Organized Operational Neural Networks (Self-ONNs), (Kiranyaz et al., 2021; Malik et al., 2021; Devecioglu et al.,2021; Ince et al.,2021; Kiranyaz et al., 2022). were recently shown to superiority over well-known Convolutional Networks, on many regression and classification tasks with their non-linear generalized neuron models. To reflect this superiority over underwater image reconstruction, the operational layers of the Self-ONNs take the place of the native GANs' convolutional layers. The Self-ONN based generator trained to transform corrupted to clean images can subsequently be utilized for the underwater image restoration after an Operational Cycle-GAN has been trained over the batches.

The rest of the thesis is organized as follows, first, in Chapter 2 source of corruption of underwater images. Then, the methods used to restore the underwater images are explained in Chapter 3. Then, the dataset used, experimental setup, and experimental results are presented in Chapter 4. Finally, in chapter 5, the thesis is concluded.

CHAPTER 2: WHY UNDERWATER IMAGES ARE DISTORTED?

Underwater imaging is a challenging task that requires expensive cameras and lights. Even though these requirements are satisfied, underwater images still suffer from color distortion, uneven illumination, scattering, and flickering. Chemical composition, physical characteristics, and The underwater environment's light transmission characteristics have a significant impact on the quality of the images that are taken there by creating issues that do not exist while photographing above water. Sample underwater images corrupted with mentioned artifacts are given in Figure 2

Even though light can go through a space, it cannot pass through all objects. Light can be transmitted, reflected, or absorbed when it hits an object. The item is made up of molecules, and each molecule has electrons that can absorb energy and jump to higher energy levels. According to its frequency, a light packet has a particular quantity of energy, the higher the frequency, the more energy. The electron will absorb this energy and re-emit it as heat if it corresponds to one of the electron energy levels. Transparent materials, on the other hand, do not absorb the photon's energy. The photon can pass right through since it is not absorbed. Some materials are partly transparent, allowing some photons to pass through while others are absorbed. Because it only passes specific colors of light, the material will appear colored.

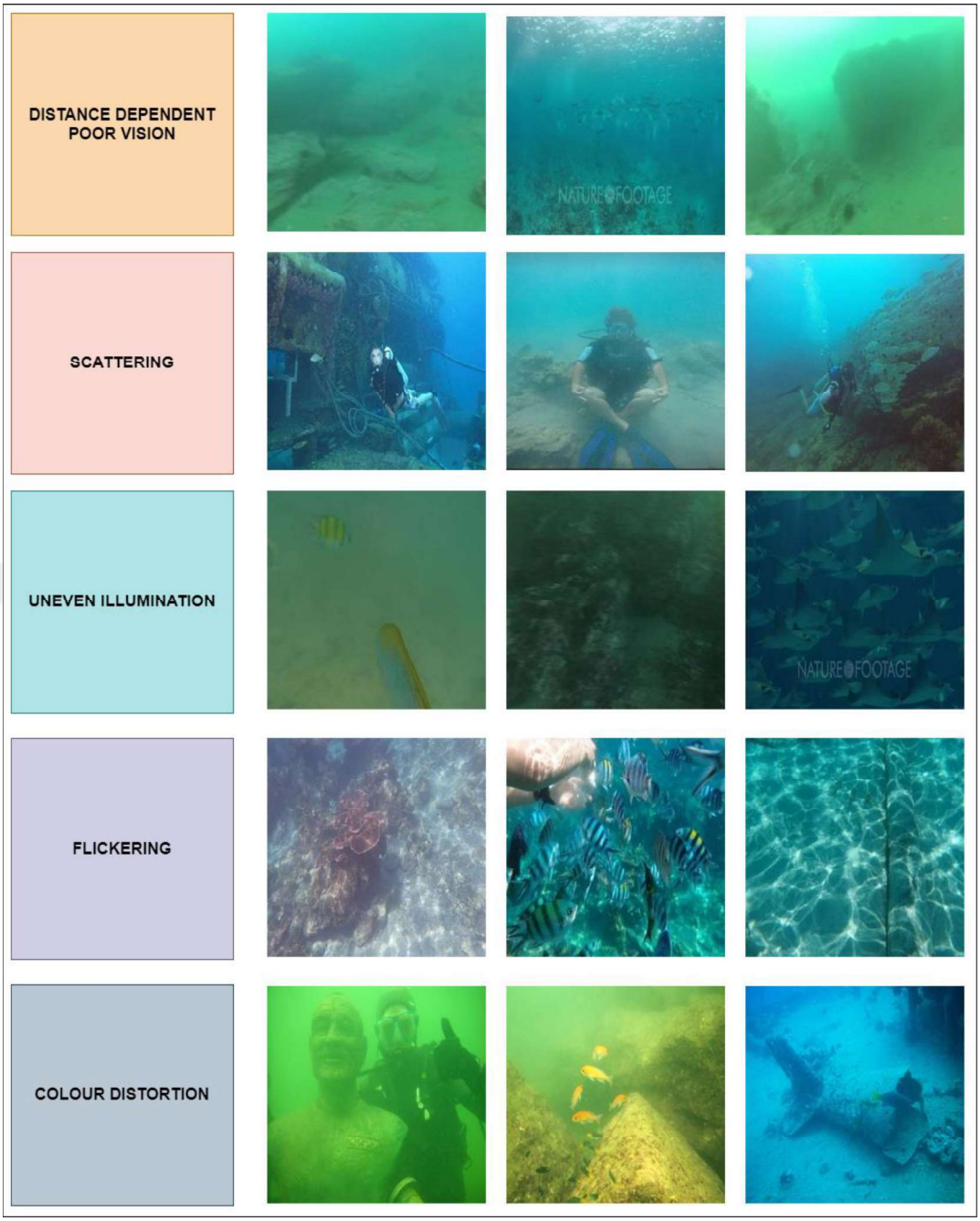

Figure 2. Corrupted Underwater Images with respect to their dominant artifacts

Furthermore, the density of water is approximately 800 times that of the atmosphere above water. As a result, visible light falling on water bounces back in part and passes into the water in part (Water and light–Seafriends, 2022). As it propagates through the water, the amount of energy going into it begins to decrease. Water molecules absorb a portion of the light's energy, which is why subaquatic

images get more unlit as depth increases. The maximum energy of particles floating in water is diminished, and light energy deviates from its route before reaching the camera, resulting in a blur, reduced contrast, and mist (Oakley, 1998). Scene radiance also decreases as the distance between the intended objectives and the camera increases due to absorption (Qingsong et al., 2015).

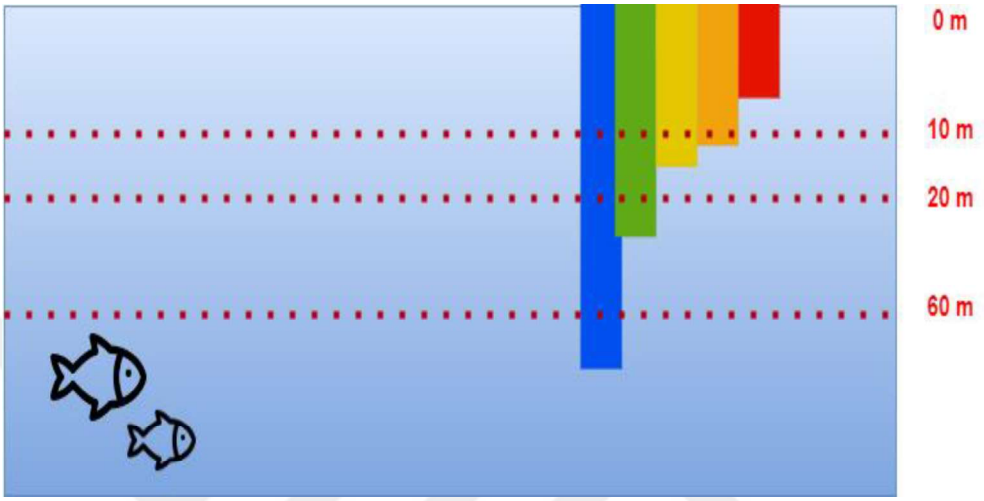

Figure 3. Visualization of underwater color deformation artifact

Figure 3. shows the illustration of how colors are absorbed by light. From the figure, it can be seen that, while going deeper, the red and orange lights of the sun are quickly absorbed by the water. As a consequence, blue/green color is more dominant in photos taken underwater.

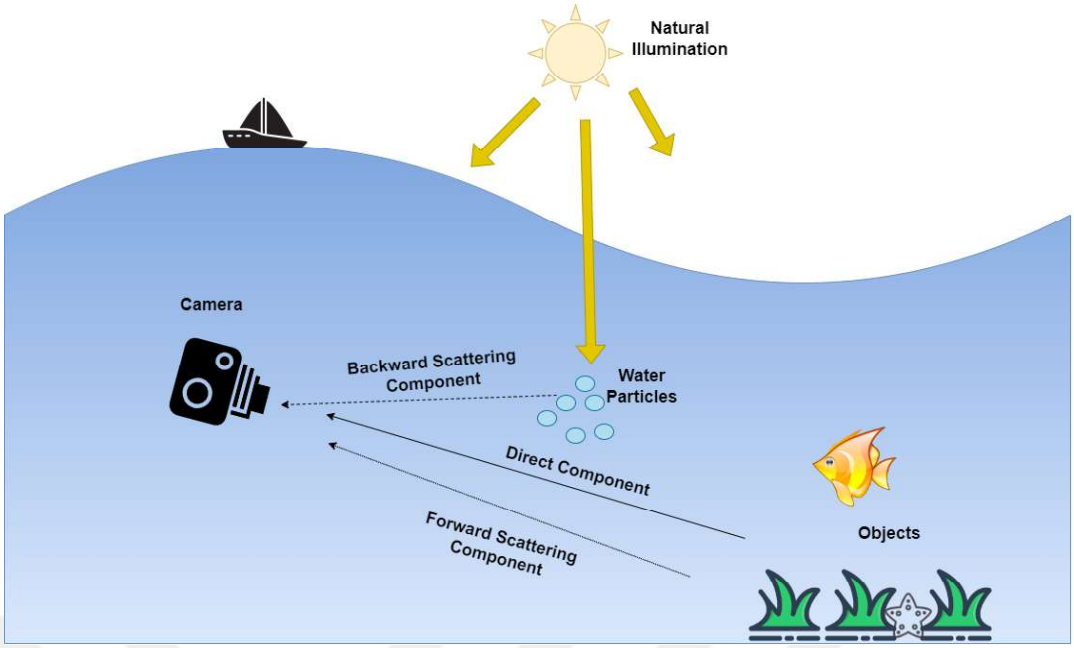

Figure 4. Visualization of underwater scattering artifact

Additionally, water molecules lower a certain percentage of the energy in visible light, which results in the cloudiness of an image. (Sandbhor et al., 2015). A fraction of the visible light energy from the image interacts with the fluid particles as it moves toward the camera's aperture. In consequence of this, these particles absorb and scatter light energy as illustrated in Figure 4. Artificial illumination is frequently used but it is also influenced by scattering and attenuation (Galdran, 2018). Simultaneously, uneven illumination results, producing bright spots in the center of the underwater image. (Li et al., 2018,). The camera collects radiation of 3 different types of light: Energy that is directly reflected from the scene (direct transmission), the energy that encounters minute particles and scatters before it reaches the image-capturing device's aperture (forward scattering), and energy from atmospheric light that is reflected by a water-based particle (background scattering) (Schechner et al., 2005). The underwater image is a linear mixture of the direct-transmission, forward-scattering, and back-scattering components, which is given in Equation 1 (Schettini et al., 2010; McGlamery, 1980; Jaffe, 1990):

$$E_t(l,m) = E_d(l,m) + E_f(l,m) + E_b\ (l,m)$$

Equation 1. The underwater image is a linear mixture of the direct-transmission, forward-scattering, and back-scattering components.

where (l,m) stands for the coordinates of the image's components; $E_t(l,m)$ stands for the total energy directed at the camera's aperture; $E_d(l,m)$ stands for the direct scattering component; $E_b\ (l,m)$ stands for the backscattering component; and $E_f(l,m)$ stands for the forward scattering component.

CHAPTER 3: METHODOLOGY

In this section, the methods used in this thesis are presented. First, U-Nets (Ronneberger et al.,2015) will be introduced. Then, the main properties of GANs (Goodfellow et al,,2014) are described, Then, the next generation of AI, the ONNs (Kiranyaz et al., 2020) and Self-ONNs (Kiranyaz et al., 2021) are introduced. Finally, the proposed novel underwater image restoration method, Operational Cycle-GANs is explained.

3.1. U-Nets

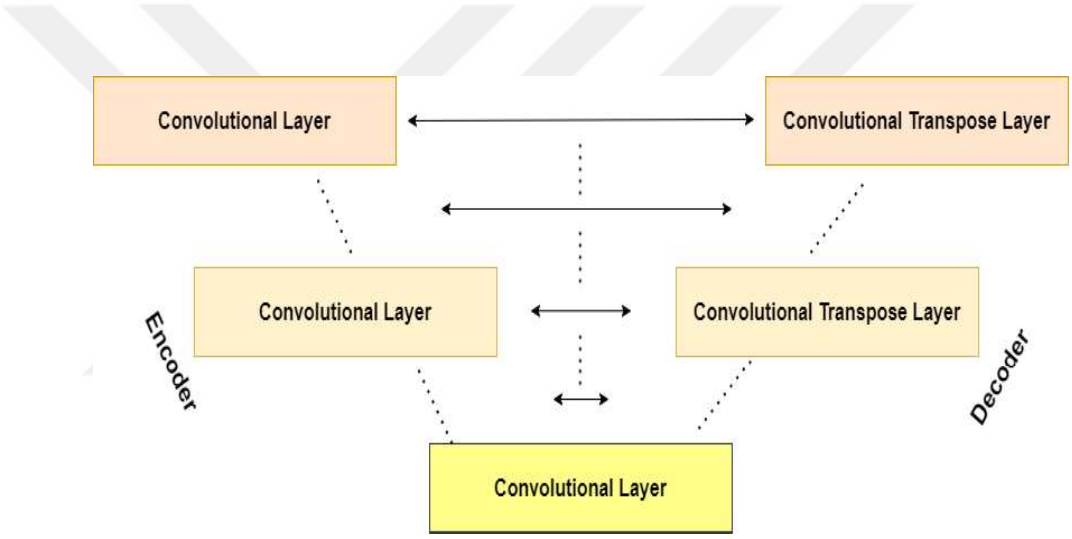

Figure 5. Architecture of U-Net

U-Net (Ronneberger et al., 2015) is an architecture that is derived from convolutional neural networks to process biomedical images. The general framework of U-Nets is given in Figure 5 Unlike traditional convolutional neural networks, U-Net generates an image as output for segmentation purposes rather than one valued label. U-Net network consists of three main parts: encoder, decoder, and skip connections. The encoder network aims to extract features. The decoder network is used to take the features from the encoder and generate a semantic segmentation mask. Skip connections are shortcut connections that help the indirect flow of gradients from the early layers of the encoder to the decoder to generate better

outputs.

3.2. Generative Adversarial Networks

Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) are generative machine learning models that have become popular recently with the abilities to generate fake images, image-to-image translation, and text-to-image translation. GAN contains two machine learning architectures which are generator and discriminator.

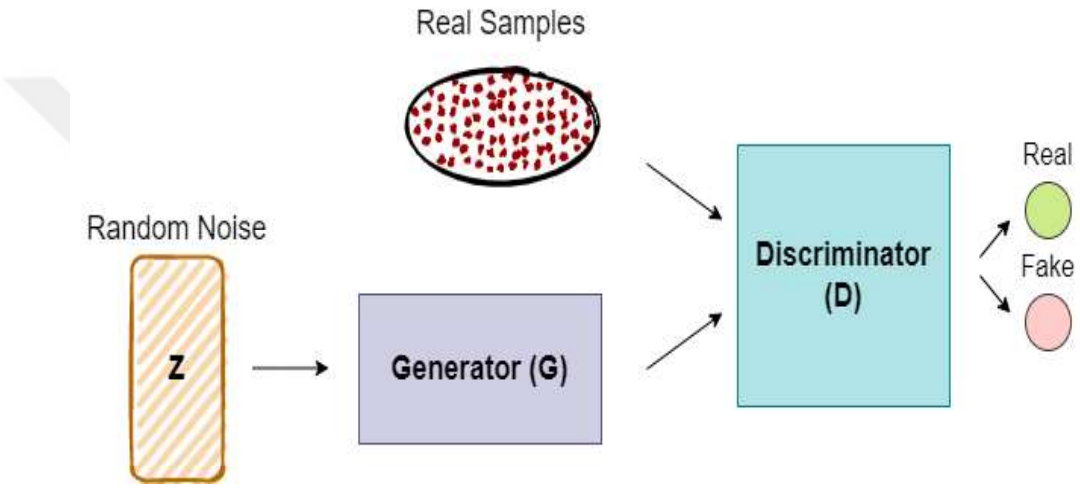

Figure 6. Architecture of conventional GANs

The general framework of GANs is shown in Figure 6. The generator model tries to generate realistic fake samples from random noise. The output of the generator is fed into the discriminator with real samples in the training dataset. Discriminator forces generator to create more realistic samples by discriminating generators fake samples and real samples.

Discriminator plays a two-player min-max game on the objective function V(G, D), where G is the generator, D is the discriminator and z is the random noise:

$$\min_{G} \max_{D} V(D,G) = E_{x(real)} \log\big(D(x)\big) + E_{z(fake)} \log\big(1 - D(G(z))\big)$$

Equation 2. The objective function of GANs

D() gives the probability of the given samples whether real or fake. For the generator's loss function, the aim is to minimize log(1-D(G(z)). For the discriminator, the target is the maximize D(x) and (1-D(G(s))). As a consequence, the optimal output for D will be P(x)=0.5 since we want the discriminator not to discriminate between classes.

3.3. Operational Neural Networks

The two main restrictions imposed by typical CNNs, which employ the conventional "linear" neuron model, are kernel-wise constrained connections and weight sharing. Due to these limitations, the linear weighted sum for MLPs becomes the CNN convolution formula. This is demonstrated in Figure 7 (left) where the three sequential convolutional layers are displayed without the sub-sampling (pooling) layers.

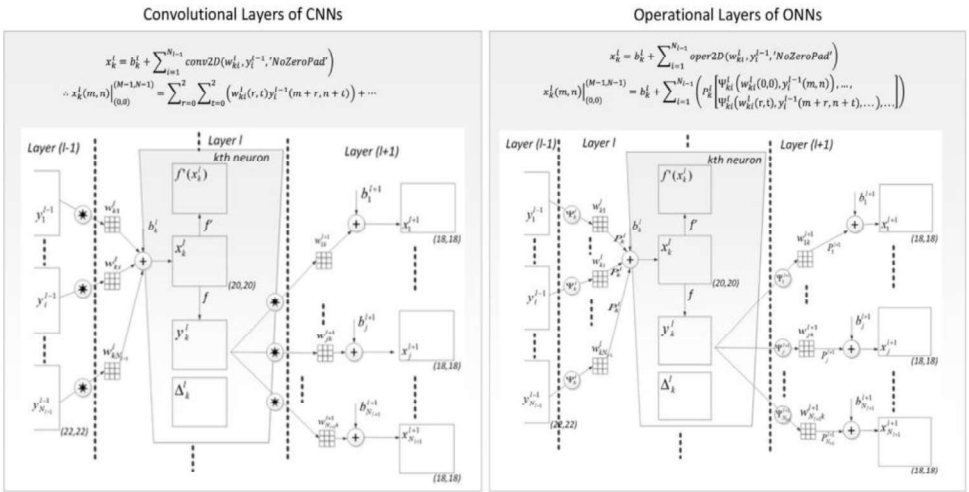

Figure 7. Three following convolutional (left) and operational (right) layers of corresponding the kth neuron of a CNN (left) and an ONN (right).

Using the fundamental GOPs principle, nodal and pool operators extend the exclusive usage of linear convolutions in the convolutional neurons. The operational layers and neurons are made up of these, whereas the two main limitations—weight sharing and constrained (kernel-wise) connectivity—are directly derived from traditional CNNs. This is also shown in Figure 7 (right), where an ONN is represented by an ONN with three operational layers and the kth neuron with three kernels. The input map of the kth neuron at the current layer, shown as x_k^l, is generated by pooling the final output maps, or y_k^l of the neurons at the previous layer that was driven by their respective kernels, w_{ik}^{l+1}, as shown in the example:

$$x_i^{l+1} = b_i^{l+1} + \sum_{k=0}^{N_l} oper2D(w_{ik}^{l+1}, y_k^l, 'NoZeroPad')$$

$$x_i^{l+1}(m,n) = \cdots + P_i^{l+1}\begin{bmatrix} \Psi_i^{l+1}\left(y_k^l(m,n), w_{ik}^{l+1}(0,0)\right), \ldots, \\ \Psi_i^{l+1}\left(y_k^l(m+r,n+t), w_{ik}^{l+1}(\mathrm{r},\mathrm{t}),\right)\ldots \end{bmatrix}$$

Equation 3. Forward Propagation of ONNs

When the pooling operator in Equation 3, is the summation and the nodal operator Ψ_i^{l+1} is linear, the ONN is equal to a CNN.

The training of an ONN via back-propagation (BP) should involve the repetitive execution of the following four steps: 1) Calculating the delta error, L 1, at the output layer, 2) Interacting between two subsequent operational layers, 3) Interacting within an operational neuron, and 4) Calculating the weight (operator kernel) and bias sensitivities to update them at each BP it. When sub-sampling (pooling) operations are used in the neuron, Stage 3 also takes care of them. A detailed information can be found in (Kiranyaz et al., 2020).

3.4. Self-Operational Neural Networks

As it is explained, ONN aims to break homogeneity in CNNs. However, to use ONNs in optimal conditions, the right transformation should be used. If the right transformation is not in the operator set, this leads to a limit on the diversity.

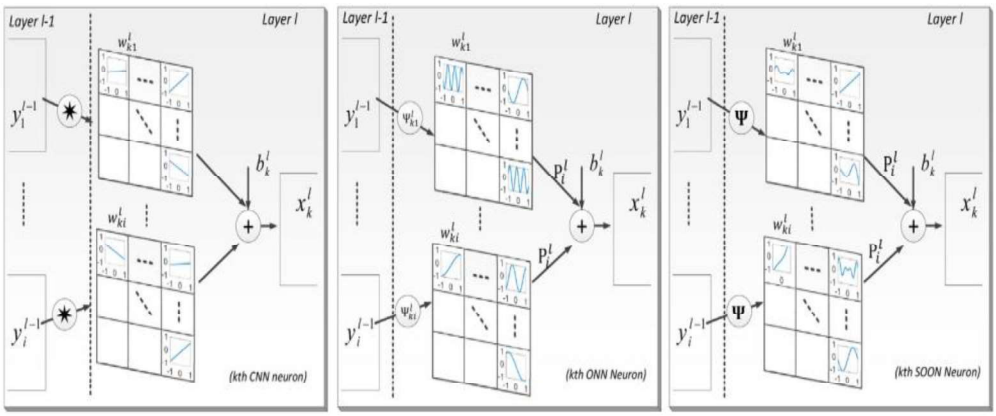

Figure 8. Illustration of the nodal operations using the ith neuron CNN, ONN, and Self-ONN (right).

In Figure 8 (left), the 3x3 kernel of CNN is shown. Weights of the kernels are slopes of the lines in the kernel. It can be seen that all the transformations are linear. In Figure 8 (middle), the operator is set to the sine function. Thus, sine functions with different frequencies can be seen as weights of the network. However, using sine (or any other arbitrary function) in each operation sets limitations. To obtain full freedom for reaching maximum diversity, Self-Operational Neural Networks are proposed. Self-ONNs can generate every nodal operator as any arbitrary function through Backpropagation. In Figure 8 (right), it can be seen that every nodal function fit the different arbitrary function to maximize the performance of the ML algorithm by varying the diversity.

Self-ONNs are defined by a nodal transformation, ψ, which can approximate any arbitrary function. The Taylor approximation of a function, near the origin (a=0)., can be expressed in Equation 4,

$$\psi(\mathrm{w},\mathrm{y}) = \sum_{n=0}^{Q} \frac{\psi^{(n)}(0)}{n!} y^n$$

Equation 4. The Taylor approximation of a function

For each kernel component of every unique inter-neuron connection, the training procedure improves the parameters to create (about) the best-fitting nodal operator. There is a problem right away because this approximation is only accurate up to y = a. The approximation is rougher the further away the points are from a. Self-ONNs are unaffected by this, though, as the nodal operators work with the previous layer's neuron outputs, each of which is constrained by the generating range of the activation operator function. The output, y, operates in the [0, 1] range, for example, if the activation function is sigmoid. If we denote $\frac{\psi^{(n)}(0)}{n!}$ as w_n:

$$\psi(\mathrm{w},\mathrm{y}) = w_0 + w_1 y^1 + w_2 y^2 + \cdots + w_Q y^Q$$

Equation 5. Maclaurin coefficients as weights

from Equation 5, w_0 is the bias and $w_1 - w_Q$ are the Mclauren coefficients which will be calculated through the backpropagation.

3.4.1. Forward Propagation Self-ONNs

The forward propagation formula for Self-ONNs is given in the equation.

$$\Psi\left(y_k^l(m+r,n+t), w_{ik}^{l+1}(r,t)\right) = w_{ik}^{l+1}(r,t,1) y_k^l(m+r,n+t) + w_{ik}^{l+1}(r,t,2)(m+r,n+t)^2 + \cdots + w_{ik}^{l+1}(r,t,Q)(m+r,n+t)^Q$$

Equation 6. The Forward Propagation of Self-ONNs

Where, w_{ik}^{l+1}'s are Mclaurin coefficients that calculate through the backpropagation.

3.4.2. Back Propagation Self-ONNs

The backpropagation of Self-ONNs starts from the output layer. Let's assume that Mean Square Error (MSE) is used as the loss function.

$$E(I) = \sum_p (y_1^L(l_p) - T(l_p))^2$$

Equation 7. MSE loss function

where, l_p is the pixel p of the image I, T is the target output and y_1^L is the predicted output. The derivative of the error with respect to the input feature map x_1^L should be calculated as delta error Δ_1^L .The delta error in the output layer of the input map can be expressed as;

$$\Delta_1^L = \frac{\partial E}{\partial x_1^L} = \frac{\partial E}{\partial y_1^L}\frac{\partial y_1^L}{\partial x_1^L} = \frac{2}{|I|}(y_1^L(I) - T(I))f'(X_1^L(I))$$

Equation 8. The delta error in the output layer of the input map

Which, can be directly calculated from multiplying error $(y_1^L(I) - T(I))$ to the derivative of activation function f. It is needed to find the delta error of every input map of every hidden layer. To Calculate this, the equations are given in Equation 8 with varying the pool operator P and nodal operator Ψ:

$$x_i^{l+1}(m-1,n-1)= \ldots + P_i^{l+1}\left[\Psi\left(y_k^l(m-1,n-1), w_{ik}^{l+1}(0,0)\right), \ldots, \Psi\left(y_k^l(m,n), w_{ik}^{l+1}(1,1)\right)\right] + \cdots$$

$$x_i^{l+1}(m-1,n)= \ldots + P_i^{l+1}\left[\Psi\left(y_k^l(m-1,n), w_{ik}^{l+1}(0,0)\right), \ldots, \Psi\left(y_k^l(m,n), w_{ik}^{l+1}(1,0)\right)\right] + \ldots$$

$$x_i^{l+1}(m,n)= \ldots + P_i^{l+1}\left[\Psi\left(y_k^l(m,n), w_{ik}^{l+1}(0,0)\right), \ldots, \Psi\left(y_k^l(m+r,n+t), w_{ik}^{l+1}(r,t)\right)\right] + \ldots$$

Equation 9. The delta error of every input map of every hidden layer

Where nodal y_k^l operator is Mclauren polynomial given in Equation 5, We can simplfy the equation like Equation 10;

$$\therefore x_i^{l+1}(m-r,n-t)\Big|_{(1,1)}^{(M-1,N-1)} = b_i^{l+1} + \sum_{k=1}^{N_1} p_i^{l+1}\left[\ldots, y_k^l(m,n), w_{ik}^{l+1}(r,t), \ldots\right]$$

$$\therefore x_i^{l+1}(m,n)\Big|_{(1,1)}^{(M-1,N-1)} = b_i^{l+1} + \sum_{k=1}^{N_1} p_i^{l+1}\left[\ldots, y_k^l(m+r,n+t), w_{ik}^{l+1}(r,t), \ldots\right]$$

Equation 10. The delta error of every input map of every hidden layer (simplified)

The derivative of the error can be taken with respect to the previous layer output map y_k^l

$$\therefore \frac{\partial E}{\partial y_k^l}(m,n)\Big|_{(0,0)}^{(M-1,N-1)} = \Delta y_1^l(m,n)\Big|_{(0,0)}^{(M-1,N-1)} =$$

$$\sum_{i=1}^{N_{l+1}}\left(\sum_{r=0}^{Kx-1}\sum_{t=0}^{Ky-1}\frac{\partial E}{\partial x_i^{l+1}(m-r,n-t)} \text{ x } \frac{\partial x_i^{l+1}(m-r,n-t)}{\partial P_i^{l+1}\left[\ldots, \Psi\left(y_k^l(m,n), w_{ik}^{l+1}(r,t)\right), \ldots\right]} \text{ x } \frac{\partial P_i^{l+1}\left[\ldots, \Psi\left(y_k^l(m,n), w_{ik}^{l+1}(r,t)\right), \ldots\right]}{\Psi\left(y_k^l(m,n), w_{ik}^{l+1}(r,t)\right)} \text{ x } \frac{\Psi\left(y_k^l(m,n), w_{ik}^{l+1}(r,t)\right)}{\partial y_k^l(m,n)}\right)$$

Equation 11 The delta error of output pixel

$$\Delta y_1^l(m,n)\Big|_{(0,0)}^{(M-1,N-1)} =$$

$$\sum_{i=0}^{N_{l+1}}\left(\sum_{r=0}^{K_{x-1}}\sum_{t=0}^{K_{y-1}}\Delta_i^{l+1}(m-r,n-t)\,\mathrm{x}\,\Delta_\Psi P_i^{l+1}(m,n,r,t)\mathrm{x}\,\Delta_y\Psi(m,n,r,t)\right)$$

$$Let\ \Delta_y P_i^{l+1}(m,n,r,t) = \Delta_\Psi P_i^{l+1}(m,n,r,t)\mathrm{x}\,\Delta_y\Psi(m,n,r,t),\ then$$

$$\Delta y_k^l = \sum_{i=1}^{N_{l+1}} Conv2Dvar\left\{\Delta y_i^{l+1}, \Delta_y P_i^{l+1}(m,n,r,t)\right\}$$

Equation 12. The delta error of output pixel (simplified)

$$\Delta_y\Psi(m,n,r,t) = w_{ik}^{l+1}(r,t,1) + 2w_{ik}^{l+1}(r,t,2)y_k^l(m,n) + \cdots + Qw_{ik}^{l+1}(r,t,Q)w_{ik}^{l+1}(r,t,2)y_k^l(m,n)^{Q-1}$$

Equation 13 The gradient output of the operator function

Equation 11 and Equation 12 shows the delta error of output pixel and its simplified version respectively. Equation 13 Shows the gradient output of the operator function.

3.5. Operational Cycle GANs

Cycle-GANs determine the aspects of two classes and learn to transform each other while preserving unrelated characteristics of the input image using two generators. As a consequence, if the training set of Cycle-GANs are consist of perfect underwater images as one class and corrupted underwater images as another class, one of the generators learns to detect sources of corruption and learns to transform that source to its possible clean version without damaging main characteristics. This makes Cycle GANs the perfect blind restoration method. Recently Proposed Self-ONNs show superiority over previous state-of-art deep learning models over many ML and CV tasks with less computational complexity. To reflect this superiority over the underwater image restoration problem, convolutional layers of traditional Cycle-GAN are changed to operational layers.

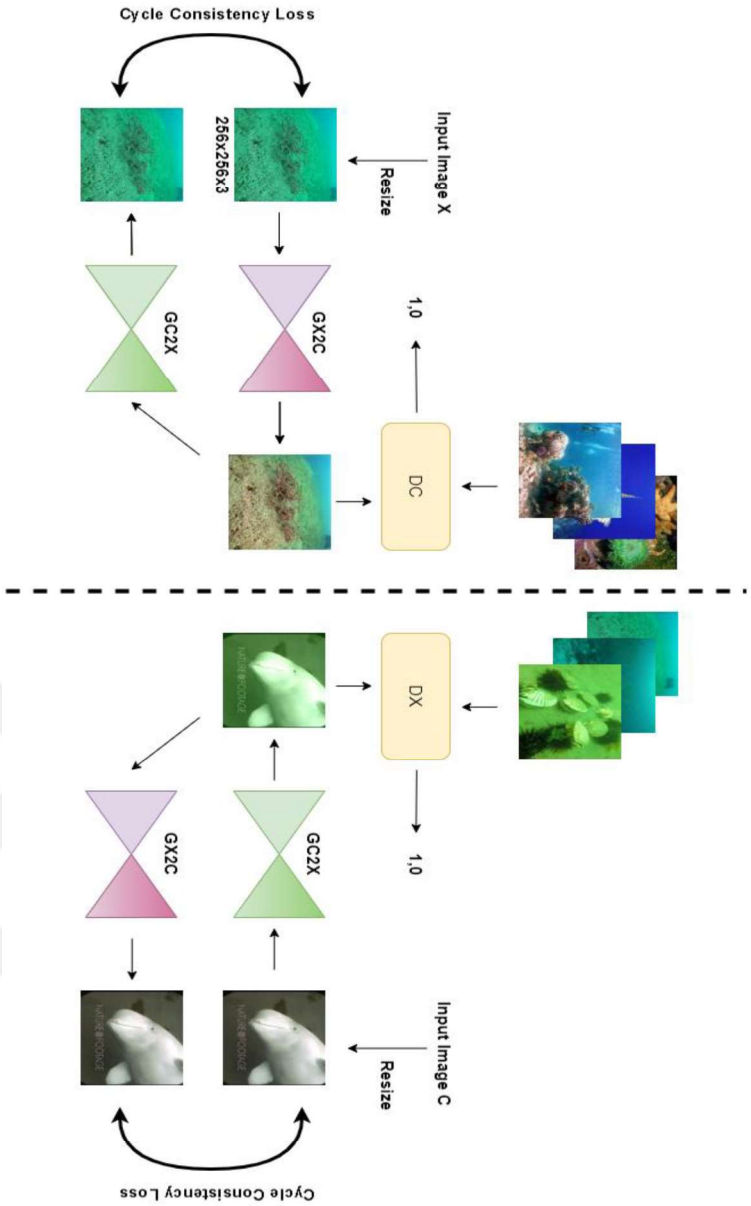

Figure 9. The proposed underwater image restoration framework

Figure 9 shows the basic framework of the proposed underwater image restoration scheme. Each RGB channel of the input image is resized to 256x256 and linearly scaled to a range of [-1. 1]:

$$X_{\mathrm{N}}(i,j) = 2\frac{X(i,j) - X_{min}}{X_{max} - X_{min}} - 1$$

Equation 14. The definition of linear normalization

where $X(i,j)$ and $X_{\mathrm{N}}(i,j)$ are the original and normalized pixel values, The input color channel's maximum and minimum values are designated as X_{max} and X_{min} , respectively. The proposed approach contains two pairs of Self-ONN-based generator and discriminator models. In the first set of networks, The Generator Corrupted-to-Clean (GX2C) algorithm learns how to convert corrupted images into clean ones. Discriminator Clean (DC) helps GX2C to create more realistic clean images by telling apart its synthesized images from real ones. The second set of networks tries to learn the opposite of the first set by creating realistic corrupted images with Generator Clean-to-Corrupted (GC2X) and Discriminator Corrupted (DX). In this case, a generator gets feedback from the other generator to ensure that an image generated by a generator is cycle consistent, meaning that applying consecutively both generators on an image should yield a similar image.

In the training of the Cycle-GANs, there are several important loss functions to minimize. Adversarial losses are calculated from DX and DC to make output samples closer to real ones. The adversarial loss functions are given in Equation 16 and Equation 17,

$$Loss_{cyc}(GX2C, GC2X, X_X, X_C) = \frac{1}{m}\sum_{i=1}^{m}\left[GC2X\left(GX2C\left(X_X(i)\right)\right) - X_X(i)\right] + \frac{1}{m}\sum_{i=1}^{m}\left[GX2C\left(GC2X\left(CS(i)\right)\right) - X_C(i)\right]$$

Equation 15. Cycle-consistency loss function of Cycle-GANs

$$Loss_{adv1}(\mathrm{GX2C}, DC, X) = \frac{1}{m}\sum_{i=1}^{m}(1 - DC(GX2C(X(i))))^2$$

Equation 16. Adversarial loss function with respect to discriminator 1

$$Loss_{adv2}(GC2X, DX, C) = \frac{1}{m}\sum_{i=1}^{m}(1 - DX(GC2X(C(i))))^2$$

Equation 17. Adversarial loss function with respect to discriminator 2

where X and C represent the corresponding corrupted and clean underwater images, respectively. Contrary to traditional GANs, cycle consistency losses are used in Cycle-GANs. This loss leads to preserving characteristic details. The cycle-consistency loss is expressed in Equation 15.

In case the input sample is in the desired output class, GAN should give the same image. To ensure this identity loss is defined Equation 18

$$Loss_{ide}(GX2C, GC2X, X_X, X_C) = \frac{1}{m}\sum_{i=1}^{m}\left[\left(GX2C\left(X_C(i)\right)\right) - X_C(i)\right] + \frac{1}{m}\sum_{i=1}^{m}\left[\left(GC2X\left(X_X(i)\right)\right) - X_X(i)\right]$$

Equation 18. Identity loss function of Cycle-GANs

Any Cycle-GAN training has as its goal minimizing the total loss in Equation 19

$$Loss_{total} = Loss_{adv1} + Loss_{adv2} + \lambda\, Loss_{cyc} + \beta\, Loss_{ide}$$

Equation 19. The total loss function of Cycle-GANs

In the next section, we present the experimental setup and parameters of Cycle-GAN structures.

CHAPTER 4: EXPERIMENTAL RESULTS

The benchmark LSUI dataset will be introduced in this chapter initially. The experimental setup that was used to evaluate the suggested underwater picture restoration strategy will next be described. Then, evolutions in both terms of quantity and quality will be exhibited. The computational complexity of the suggested approach will then be demonstrated.

4.1. Large Scale Underwater Image Dataset

Large Scale Underwater Image dataset (LSUI), the largest real underwater image dataset with high-quality reference images, is utilized to train and test the proposed method. The dataset contains 8018 underwater photos that were both acquired by them and included in other datasets that already existed (Liu et al., 2020; Akkaynak et al., 2019; Li et al., 2017; Fabbri et al., 2018; Li et al., 2020). They decided to eliminate potential bias as much as possible, the reference photos were chosen using two rounds of subjective and objective evaluations. In the first round, the first use of 18 existing optimal UIE methods (Ancuti et al., 2012; Fu et al., 2014; Peng Y.-T. 2017; Drews et al., 2013; Drews et al.,2016; Li et al.,2016 ; Chiang J. Y. et al.,2012; Galdran et al.,2015; Li et al., 2016; Islam et al.,2020; Li et al., 2017; Li et al., 2018; Yang et al., 2011; Fu et al., 2017; Song et al., 2018; Uplavikar et al., 2019; Qi et al., 2022; Ma et al., 2022) to process the collected underwater images successively, and a set with 18*8018 images is generated for the next-step optimal reference dataset selection. After subjective and objective evaluations, the LSUI dataset is formed with 5004 underwater images with their corresponding reference images.

4.2. Experimental Setup

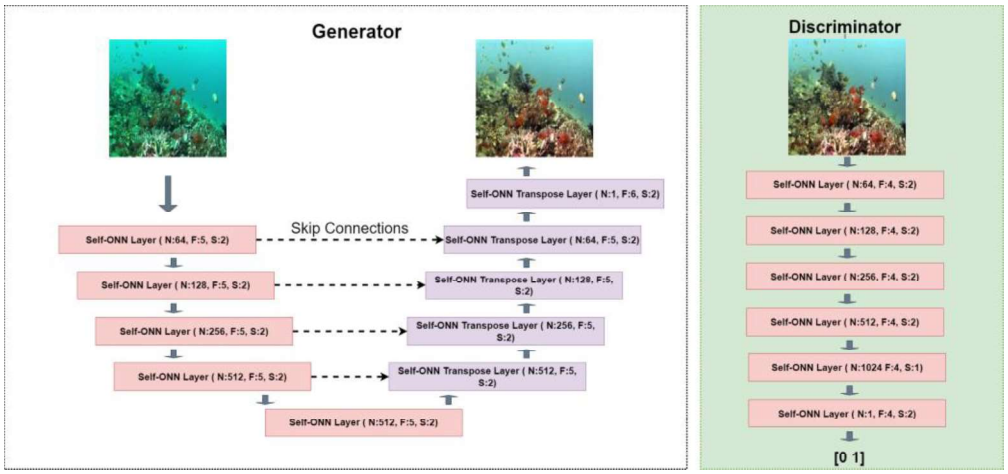

Figure 10. The architectures for the Generator and Discriminator in the proposed technique

A U-Net architecture is utilized, with 5 operational layers and 5 transposed operational layers with residual connections, for both the generators GX2C and GC2X of operational Cycle-GANs. The remaining U-Net connections assist in improving the output image's resolution. With the exception of the last transposed convolutional layer, which has a kernel size of 6, all layers have kernel sizes of 5. The stride is set to 2 for operational and transposed operational layers. Both discriminators have a kernel size of 4 and 6 operating layers. The layers' strides are set at 2, 2, 2, 2, 1, and 2 per layer, accordingly. The generator and discriminator architectures are shown in Figure 10. The loss weights λ and β in (9) are set as 10 and 5. The proposed 1D Self ONN architectures are implemented using the FastONN library (Malik et al., 2020) based on Python and PyTorch. 4500 clean and corrupted underwater images are chosen for the training dataset, while the remaining data segments are used for testing and assessment.

4.3. Experimental Results

Peak Signal-to-noise Ratio (PSNR) is used to quantitatively evaluate results. PSNR which is the signal's maximum power to corrupting noise power ratio is the most commonly used restoration metric. PSNR can be described as;

$$PSNR = 10 * log_{10} \frac{(MAX_{R(i,j)})^2}{\frac{1}{m*n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[R(i,j) - I(i,j)]^2}$$

Equation 20. The definition of PSNR

Where R(i,j) and I(I,j) are the corresponding m*n sized restored and ground truth images.

Table 1. Restoration Performance of Operational Cycle-GANs and Benchmark Underwater Restoration Algorithms.

Methods	PSNR	Parameters	Computational Time (s)
UIBLA (Peng et al., 2017)	13.5	x	42.1 s
UDCP (Drews et al., 2013)	11.8	x	30.8 s
Fusion (Ancuti et al., 2012)	17.4	x	6.5 s
Retinex based (Fu et al., 2014)	13.8	x	1.0 s
RGHS (Chiang J. Y.)	14.2	x	8.9 s
WaterNet (Li et al., 2020)	17.7	24.8 M	0.6 s
FUnIE (Islam et al., 2020)	19.3	7.0 M	0.09 s
UGAN (Fabbri et al., 2018)	19.7	57.1 M	0.05 s
UIE-DAL (Ma Z. 2019)	17.4	18.8 M	0.07 s
Ucolor (Li et al., 2021)	22.9	157.4 M	2.75 s
U-Shaped Transformer (Peng et al., 2021)	24.1	65.6 M	0.07 s
Original (CNN) Cycle-GAN	18.9	5.1 M	0.01 s
Operational Cycle-GAN (q=5)	22.3	12.2 M	0.02 s

Table 1 demonstrates the PSNR performances over the restored images by the Operational Cycle-GAN, original Cycle-GAN, and benchmark underwater restoration models. Operational Cycle-GAN has more parameters than their equivalent CNN Cycle-GAN because of the nonlinear parameters. However, this nonlinearity increases the learning capacity of Operational Cycle-GAN by increasing the PSNR by about 4 dB. When we compare Operational Cycle-GAN to benchmark underwater restoration models, the proposed model outperforms UIBLA, UDCP, Fusion, Retinex based, RGHS, WaterNet, FUnIE, UGAN, UIE-DAL but, both Ucolor and U-Shaped Transformer achieved better than our proposed model with corresponding 0.6 dB and 1.8 dB improvements. However, Our model significantly decreases the number of parameters. This makes out model more real-timely implementable.

Figure 11 and Figure 12 show the sample underwater images and their corresponding Operational GAN restored outputs. The first and most significant observation is that the restored underwater images' quality is superior to the original underwater images: The details are sharpened, the color of images is corrected, the water turbidity effect is suppressed, and contrast is enhanced. The proposed algorithm maintains the original image without any artificial alterations or degradation, whether it isn’t corrupted with artifacts.

When we take a closer look at Figure 11, in a b and c, the algorithm corrected the colors of the images. In b, it smooths the flickering effects without completely suppressing them. In c Operational Cycle-GAN enhances the contrast of the image. The details on the wrecked boat sharpened and become more visible.

When we examine Figure 12 more closely, in the first row, the input image is extremely blurred. The proposed method sharpens the image. In the d and e, both input images are green-scaled. The Operational Cycle-GAN corrected the colors of images.

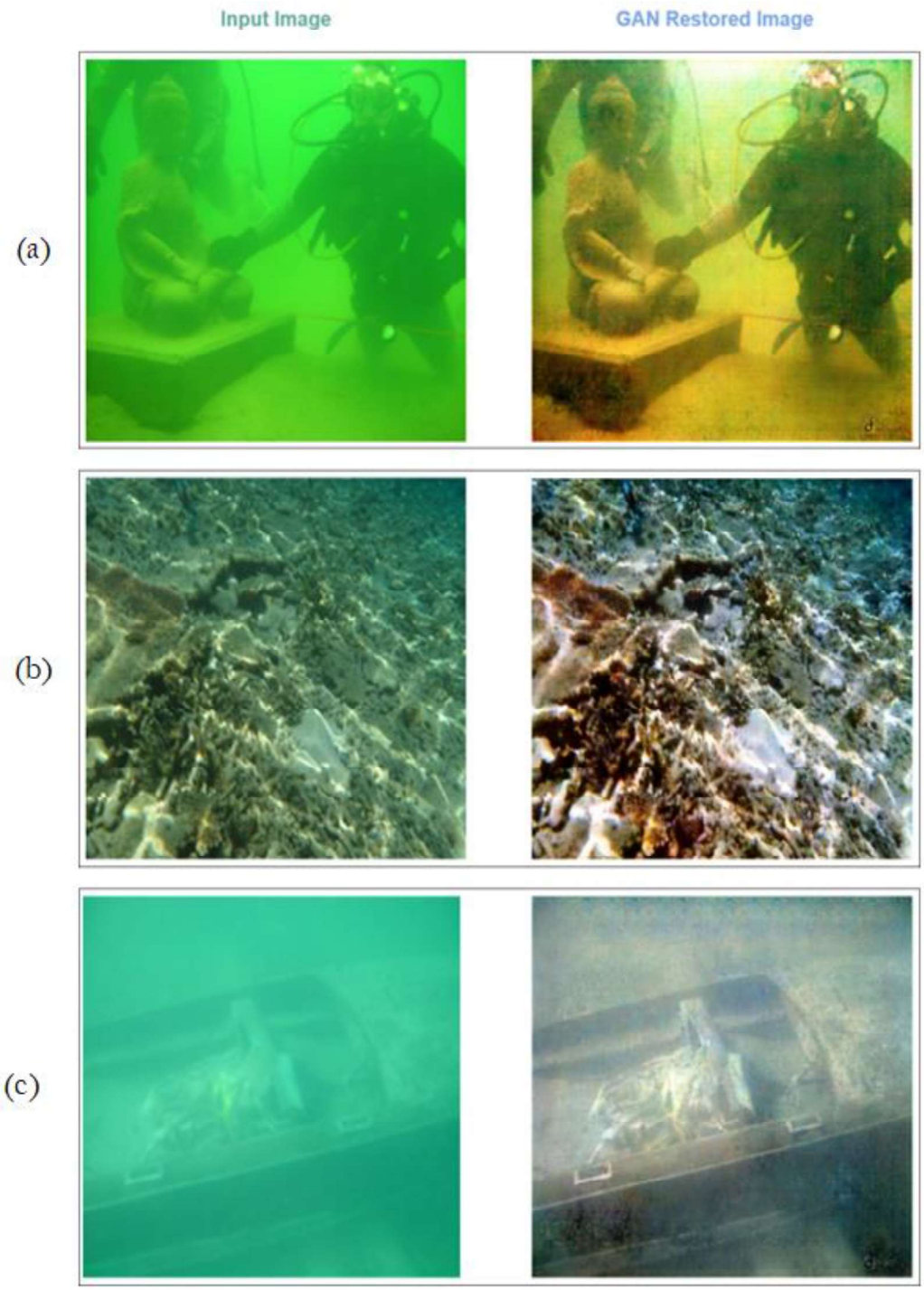

Figure 11. Sample underwater images and its corresponding Operational Cycle-GAN output Images.

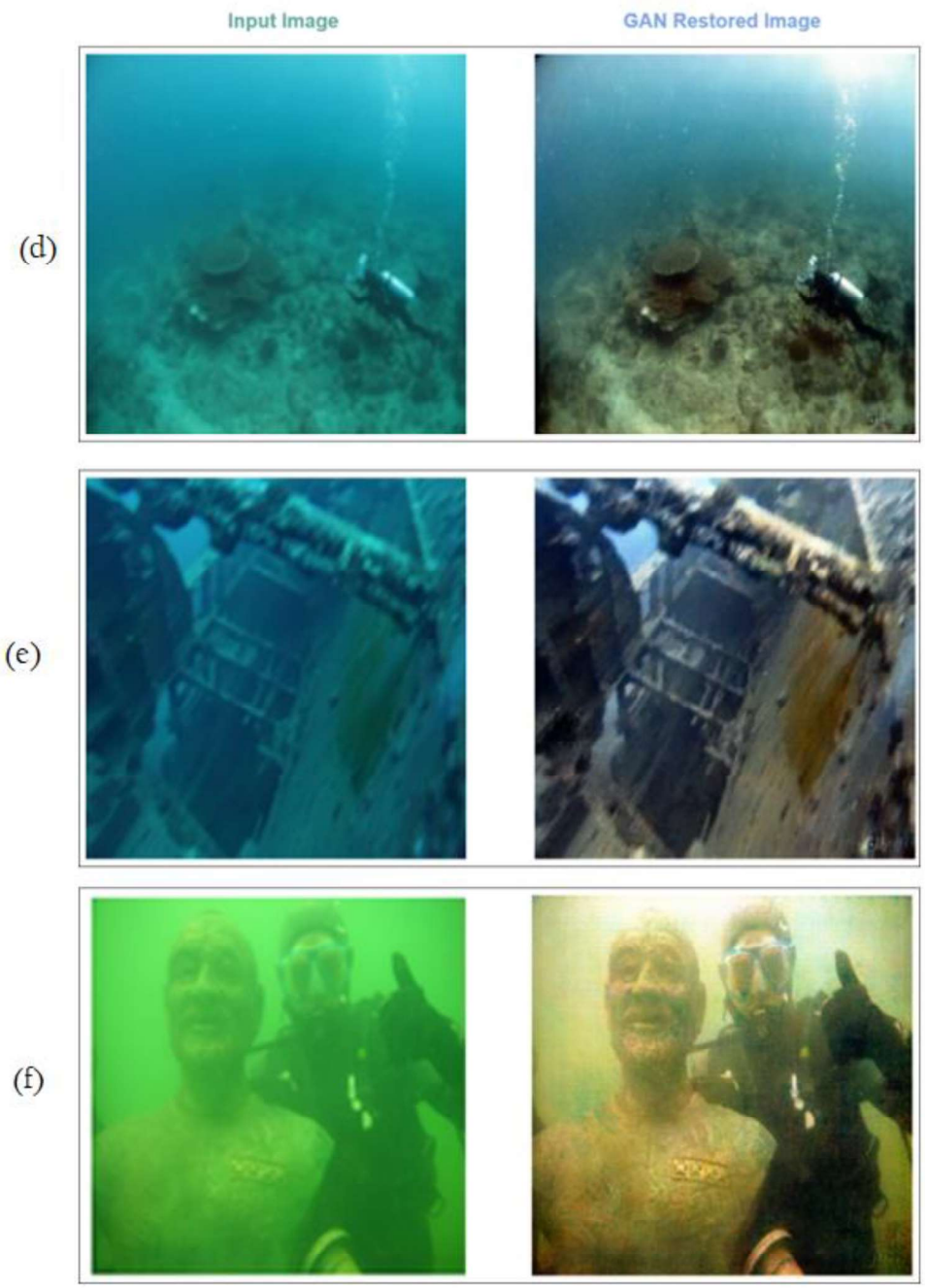

Figure 12. Sample underwater images and their corresponding Operational Cycle-GAN output Images.

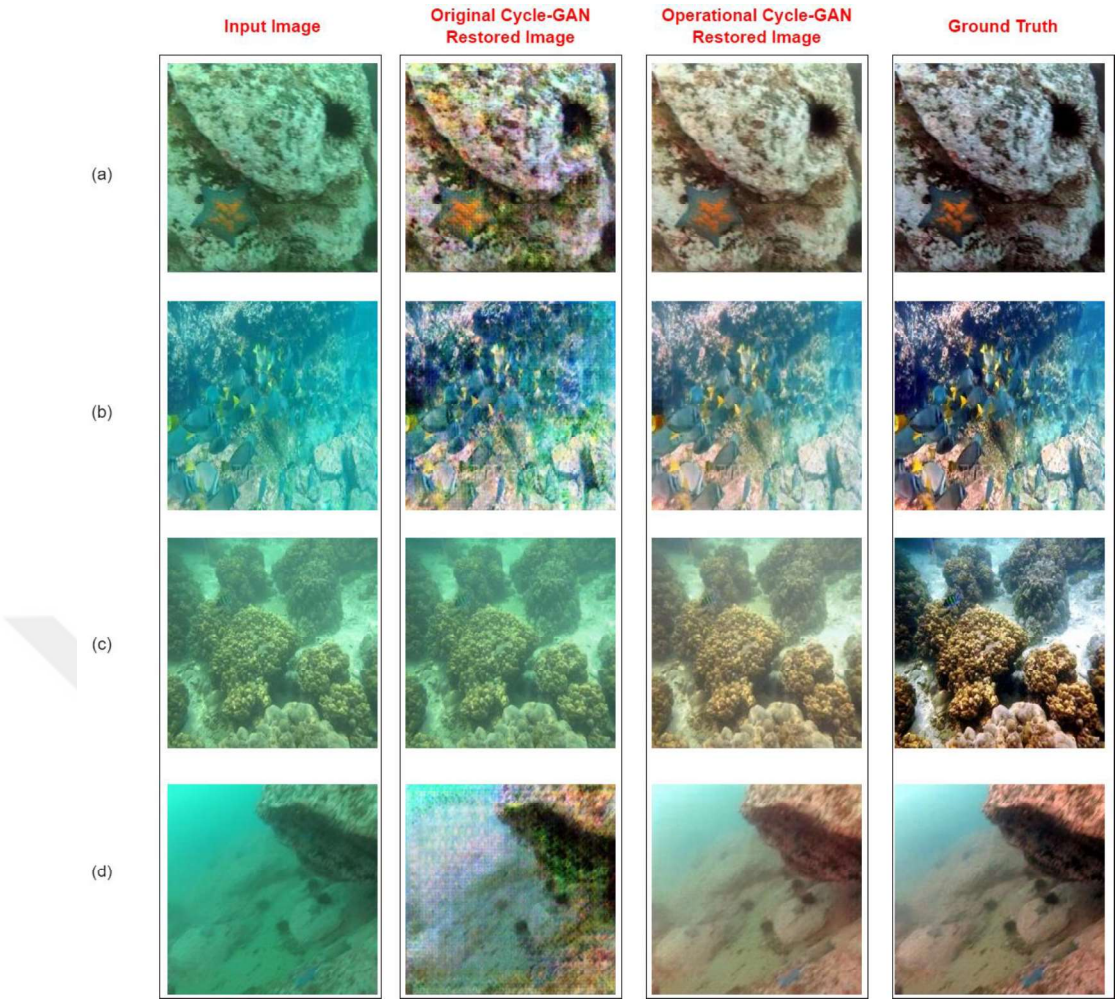

Figure 13. Sample underwater images, its corresponding GAN output images and its ground truths

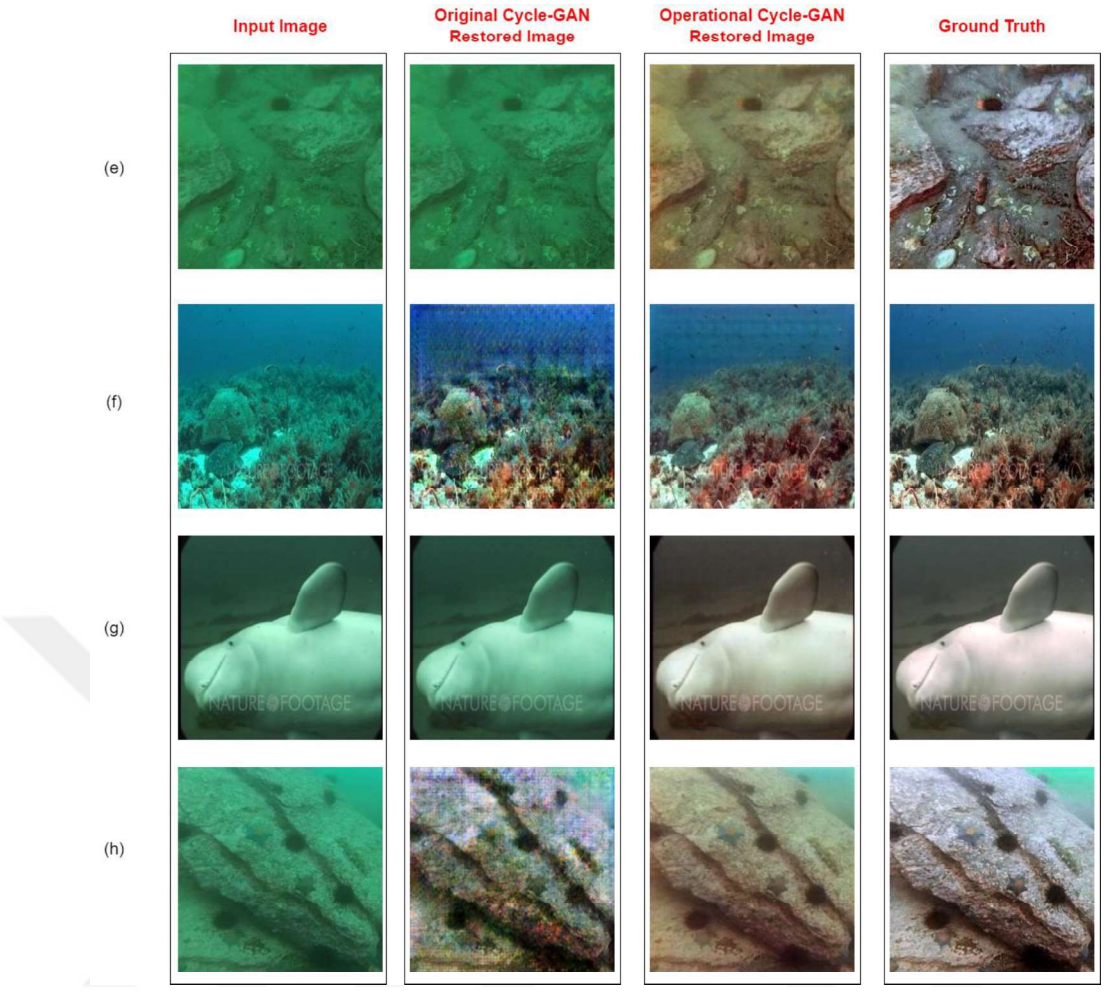

Figure 14. Sample underwater images, their corresponding GAN output images, and their ground truths

In Figure 13 and Figure 14, Original underwater images are given with their corresponding original Cycle-GAN restored, Operational Cycle-GAN restored and their ground truths are given. A closer look at both figures reveals that Operational Cycle-GANs are better restoration performance in terms of color-correcting, blur removal, and contrast enhancement than original Cycle-GANs. Besides the lower restoration performance original cycle gan adds synthetic noise in Figure 13 b, d and Figure 14 f, a. In addition to the artificial noise adding, the original Cycle-GAN decides that Figure 13 c and Figure 14 e-g are not corrupted with artifacts and give exact output as input. As a consequence, when both inputs take into consideration, Operational Cycle-GANs have superiority over restoring the underwater images.

4.4. Computational Complexity

For GX2C network setup, the network size, total number of parameters (PARs), and inference time (to recover an underwater image) are calculated and provided in Table 1 for the computational complexity. (Malik et al.,) contains the comprehensive formulas of the PARs computations for Self-ONNs. A 2.2 GHz Intel Core i7 computer with 16 GB of RAM and an NVIDIA GeForce RTX 3080 graphics card was used for all of the trials. Python and the PyTorch and FastONN libraries are used to implement operational Cycle-GANs. GPU cores were used to process the classifier's training and testing phases. The operational Cycle-GAN is substantially quicker in terms of the inference time and PARs. Less parameter size of our model makes our model more real time implementable than the models which are slightly get mode PSNR value than our model.

CHAPTER 5: CONCLUSIONS

Even though expensive equipment is used, underwater images are severely corrupted because of both lights' transmission properties and water's physical characteristics. These artifacts lose the details of the picture and render these pictures useless. To avoid this problem, prior works usually find the solution to enhance these corrupted images using deep learning models with using synthetic noise added paired datasets as a supervised solution. However, even though these methods can be useful to enhance a certain type of artifact which trained, they are useless to remove a blend of artifacts. As a consequence, these typical regression-based methods are not practical, which is why this work addressed the problem as a blind approach, without making any assumptions about the types or severity of the artifacts. We provide a novel Operational Cycle-GAN strategy to repair the underwater images in order to solve this issue, independent of the kind or severity of the artifacts. The clean and corrupted batches of images are utilized to train the operational cycle-GANs. The generator, GX2C, develops the ability to convert corrupted underwater images into their uncorrupted counterparts.

The wide collection of real underwater images used in the quantitative and qualitative evaluations shows that the corrupted underwater images can be restored to the appropriate vision quality. During the qualitative evaluations, the proposed work achieves superiority over the many benchmark underwater image restoration methods in terms of PSNR and the number of parameters used in the model. The proposed method outperforms the models using fewer parameters to double of parameters by 3-11 dB and get about 1-2 dB less PSNR value than methods using more double parameters than our model. This is not surprising that Self-ONNs consistently outperform (deep) CNN models by using compact models in a wide range of challenging ML and CV tasks (Kiranyaz et al., 2021; Malik et al., 2021; Devecioglu et al., 2021; Kiranyaz et al., 2022; Malik et al., 2021; Ince et al., 2021).

For future work, it is possible to further lower the depth and complexity of the operational Cycle-GANs while improving the restoration performance using

the next-generation AI method Super Neurons (Kiranyaz et al., 2021). Using Super Neurons instead of the generative neurons in the Self-ONN layers will be the topic of our future research.

REFERENCES

Akkaynak d. and Treibitz T. (2019) *Sea-thru: A method for removing water from underwater images*, CVPR, pp. 1682–1691.

Ancuti C., Ancuti C. O., Haber T., and Bekaert P. (2012) *Enhancing underwater images and videos by fusion*, CVPR, pp. 81–88.

Chiang J. Y. and Chen Y.-C. (2012) *Underwater image enhancement by wavelength compensation and dehazing*, IEEE T. Image Process., vol. 21, no. 4, pp. 1756–1769.

Devecioglu O. C., Malik J., Ince T., Kiranyaz S., Atalay E. and Gabbouj M. (2021) *Real-Time Glaucoma Detection From Digital Fundus Images Using Self-ONNs*, IEEE Access, vol. 9, pp. 140031-140041.

Drews Jr P., do Nascimento E., Moraes F., Botelho S., and Campos M. (2013) *Transmission estimation in underwater single images*, ICCV workshops, pp. 825–830.

Drews P. L., Nascimento E. R., Botelho S. S., and Montenegro Campos M. F. (2016) *Underwater depth estimation and image restoration based on single images*, IEEE Comput. Graph. Appl., vol. 36, no. 2, pp. 24– 35.

Fabbri C., Islam M. J., and Sattar J. (2018) *Enhancing underwater imagery using generative adversarial networks*, ICRA, pp. 7159–7165.

Fu X., Zhuang P., Huang Y., Liao Y., Zhang X.-P., and Ding X. (2014) *A retinex-based enhancing approach for single underwater image*, ICIP, pp. 4572–4576.

Fu X., Fan Z., Ling M., Huang Y., and Ding X. (2017) *Two-step approach for single underwater image enhancement*, ISPACS, 2017, pp. 789–794.

Galdran, A. (2018) *Image dehazing by artificial multiple-exposure image fusion*, Signal Process, vol. 149 pp. 135–147.

Galdran A., Pardo D., Picon A., and Alvarez-Gila A. (2015) *Automatic red channel underwater image restoration*, JVCIR, vol. 26, pp. 132–145.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014) *Generative adversarial nets*. In Advances in neural information processing systems. pp. 2672–2680.

Han, Y., Lihua H., Zhonghua H., Shouqi C., Yun Z., and Jing W. (2021) *Deep Supervised Residual Dense Network for Underwater Image Enhancement Sensors* 21, no. 9 pp. 3289.

He K., Sun J., and Tang X. (2009) *Single image haze removal using dark channel prior*, CVPR, pp. 1956–1963.

Huang D., Wang Y., Song W., Sequeira J., and Mavromatis S. (2018) *Shallowwater image enhancement using relative global histogram stretching based on adaptive parameter acquisition*, MMM. Springer, pp. 453–465.

Huimin Lu, Yujie Li, Shota Nakashima, Hyongseop Kim, And Seiichi Serikawa (2017) *Underwater image super-resolution by descattering and fusion*. IEEE Access, vol. 5, pp.670-679.

Ince T., Malik J., Devecioglu, O. C. Kiranyaz S.,Avci O., Eren L. and Gabbouj M. (2021) *Early Bearing Fault Diagnosis of Rotating Machinery by 1D Self-Organized Operational Neural Networks*, IEEE Access, Vol. 9, pp. 139260–139270.

Islam M. J., Xia Y., and Sattar J. (2020) *Fast underwater image enhancement for improved visual perception*, IEEE Robot. Autom. Lett., vol. 5, no. 2, pp. 3227–3234.

Jaffe, J.S. (1990) *Computer modelling and the design of optimal underwater imaging systems*. IEEE J. Ocean. Eng. 15(2), pp. 101–111.

Kiranyaz S., Malik J., Abdallah H. B., Ince T., Iosifidis A., Gabbouj M. (2021) *Self-organized Operational Neural Networks with Generative Neurons*, Neural Networks (Elsevier), pp. 140:294-308.

Kiranyaz S., Devecioglu O. C., Ince T., Malik J., Chowdhury M.E.H., Hamid T., Mazhar T., Khandakar A., Tahir A., Rahman T., and Gabbouj M., (2022) *Blind ECG Restoration by Operational Cycle-GANs*, in IEEE Transactions on Biomedical Engineering.

Kiranyaz S., Malik J., Yamac M., Guldogan E., Ince T., Gabbouj M. (2021) *Super Neurons*, arXiv preprint.

Li C., Anwar s., and Porikli F. (2019) *Underwater scene prior inspired deep underwater image and video enhancement*, Pattern Recognition, vol. 98, pp.1-11.

Li C., Guo J., and Guo C (2018) *Emerging from water: Underwater image color correction based on weakly supervised color transfer*, IEEE Signal. Process. Lett., vol. 25, no. 3, pp. 323–327.

Li C., Guo C., Ren W., Cong R., J. Hou, S. Kwong, and D. Tao (2020)*An underwater image enhancement benchmark dataset and beyond*, IEEE T. Image Process., vol. 29, pp. 4376–4389.

Li C., Guo J., Chen S., Tang Y., Pang Y., and Wang J. (2016) *Underwater image restoration based on minimum information loss principle and optical properties of underwater imaging*, in ICIP, pp. 1993– 1997.

Li C.-Y.,. Guo J.C,. Cong R.-M,. Pang Y.-W, and Wang B. (2016) *Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior*, IEEE T. Image Process., vol. 25, no. 12, pp. 5664–5677.

Li J., Skinner K. A., Eustice R. M., and Johnson-Roberson M. (2017) *Watergan: Unsupervised generative network to enable real-time color correction of monocular underwater images*, IEEE Robot. Autom. Lett., vol. 3, no. 1, pp. 387–394.

Li, Y., Lu H., Li K.C., and Kim H. (2018) *Non-uniform descattering and de-blurring of underwater images*. Mobile Networks Appl. 23(2), 352–362.

Liu P., Wang G., Qi H., Zhang C., Zheng H., Yu Z. (2019) *Underwater Image Enhancement with a Deep Residual Framework*. IEEE Access, vol. 7 pp. 94614-94629.

Liu R., Fan X., Zhu M., M. Hou, and Luo Z. (2020) *Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light*, IEEE Trans. Circuits. Syst. Video Technol., vol. 30, pp. 4861– 4875.

Ma Z. and Oh C. (2022) *A wavelet-based dual-stream network for underwater image enhancement*, arXiv preprint.

Malik J., Devecioglu O.C., Kiranyaz S., Ince T. and Gabbouj M. (2021) *Real-Time Patient-Specific ECG Classification by 1D Self-Operational Neural Networks*, IEEE Transactions on Biomedical Engineering, vol. 69, no. 5, pp. 1788-1801.

Malik J., Kiranyaz S. , and Gabbouj M. (2020) *FastONN—Python-based opensource GPU implementation for operational neural networks*, arXiv preprint.

Malik J., Kiranyaz S., Gabbouj M. (2021) *Self-Organized Operational Neural Networks for Severe Image Restoration Problems*, Neural Networks (Elsevier), vol. 135, pp. 201-211.

McGlamery, B.L. (1980) *A computer model for underwater camera systems*, Duntley, S.Q. (ed.) Proceedings of SPIE, vol. 0208, pp. 221–232.

Oakley, J.P. (1998) Satherley, B.L.: *Improving image quality in poor visibility conditions using a physical model for contrast degradation*, IEEE Trans. Image Process. 7(2), pp. 167–179 .

Peng L., Chunli Z. and Liheng B. 2021 *U-shape Transformer for Underwater Image Enhancement*, arXiv preprint.

Peng Y.-T. and Cosman P. C. 2017 *Underwater image restoration based on image blurriness and light absorption*, IEEE T. Image Process., vol. 26, no. 4, pp. 1579–1594.

Qi Q., Li K., Zheng H., Gao X., Hou G., and Sun K. (2022) *Sguie-net: Semantic attention guided underwater image enhancement with multiscale perception*, arXiv preprint.

Qingsong, Z. and Mai J. (2015) *A fast single image haze removal algorithm using color attenuation prior*. IEEE Trans. Image Process. vol 24, pp. 3522–3533.

Ronneberger, O., Fischer, P., Brox, T. (2015*). U-Net: Convolutional Networks for Biomedical Image Segmentation*. LNCS, vol. 9351, pp. 234-241.

Sahu P., Gupta N., and Sharma N. 2014 *A survey on underwater image enhancement techniques*, IJCA, vol. 87, no. 13.

Sandbhor, B., Kharat, G.U. (2015) *A review on underwater image enhancement techniques*. Int. J. Adv. Res. Comput. Sci. Software Eng. 5, vol. 5, pp. 676–680.

Schechner, Y.Y., Karpel, N. (2004) *Clear underwater vision*. in In: Proceedings of the 2004 IEEE Computer Society Conference on Comput Vision and Pattern Recognition (CVPR), Washington DC, pp. 536–543.

Schettini, R., Corchs, S. (2010) *Underwater image processing: State of the art of restoration and image enhancement methods*. EURASIP J. Adv. Signal Process., pp. 1–14.

Song W., Wang Y., Huang D., and Tjondronegoro D. (2018) *A rapid scene depth estimation model based on underwater light attenuation prior for underwater image restoration*, in PCM. Springer, pp. 678–688.

Uplavikar P.M., Wu Z., and Wang Z. (2019) *All-in-one underwater image enhancement using domain-adversarial learning*, CVPR Workshops, pp. 1–8.
Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., and Polosukhin I. (2017) *Attention is all you need*, NIPS, 2017, pp. 5998–6008.

Water and light–Seafriends Available at: www.seafriends.org.nz/phgraph/water.html. (Accessed: 20 June 2022).

Yang H.Y., Chen P.Y., Huang C.C., Zhuang Y.Z., and Shiau Y.H. (2011) *Low complexity underwater image enhancement based on dark channel prior*, IBICA, pp. 17–20.

Zhu J.-Y., Park T., Isola P., Efros A. (2017) *Unpaired image-to-image translation using cycle-consistent adversarial networks*, Proceedings of the IEEE international conference on computer vision , pp. 2223–2232.