A Novel Underwater Image Enhancement Technique using ResNet

Hema Krishnan, Dept. of CSE, FISAT Cochin, Kerala, India
Anjana A. Lakshmi, Dept. of CSE, FISAT Cochin, Kerala, India
Anamika L.S., Dept. of CSE, FISAT Cochin, Kerala, India
Athira C.H., Dept. of CSE, FISAT Cochin, Kerala, India
Alaikha P. V., Dept. of CSE, FISAT Cochin, Kerala, India
V. M. Manikandan, Dept. of CSE, SRM University-AP, Andhra Pradesh, India

***Abstract*—Underwater image enhancement is an active area of research due to its wide applications in areas like marine research, automated underwater vehicles, etc. In general, the underwater images have low contrast, blurriness, and color cast due to various effects like absorption, scattering, and refraction. Normally, the underwater images are less clear and not suitable for various applications. The underwater images should be enhanced to use for real-life applications and the natural image enhancement techniques may not work well for underwater images. In this paper, we introduce a scheme for enhancing the quality of the underwater images by using a residual neural network (ResNet). The synthetic underwater images for the experimental study are generated using the underwater generative adversarial network (UWGAN). The experimental results show that the new underwater image enhancement techniques perform well, and it can be used for real-life applications where we need good quality underwater images.**

***Index Terms*—Underwater image enhancement, Underwater image colorfulness measure, Underwater ResNet;**

I. Introduction

Image enhancement is the task of improving the perceptual quality of the image by performing a sequence of operations in an input image. The image enhancement techniques will be helpful in the future for further operations such as object identification, object recognition, or image segmentation. The image enhancement techniques are widely used for enhancing the visual quality of natural images and a number of algorithms are available for this. The approaches like histogram equalization and contrast stretching were used long back to distribute the pixels of the image in a better way to give better perceptual quality for the images. Enhancing normal image quality is quite easy as compared to the enhancement of underwater images. Since the underwater images may have entirely different properties, the existing image enhancement techniques cannot directly apply to underwater images. In this research, we addressed the issue related to underwater image enhancement and a machine learning-based approach is introduced in this manuscript.

As we know that many unknown creatures are living in oceans and they will play a crucial role in the life cycle. The capturing of the underwater image is very popular in various applications such as inspection of man-made things staying in the water, for various underwater rescue operations, for monitoring the underwater ecology, etc. But the real issue is getting good quality images suitable for various applications. The various chemical properties of underwater may affect the image quality and result in some greenbluish color [1].

The literature review shows that underwater images generally have poor quality due to the nature of light while going through the water. When the light passing through the water it might be refracted, absorbed, and even scattered since water is a denser medium while comparing with the air. The nature of the light while passing through the water and also the presence of organisms present in the water will lead to getting low-quality images while capturing the underwater images [2].

The underwater images are an essential requirement for tasks like underwater explorations, automated underwater vehicles, scientific research, and many others. For many of the applications that use underwater images expecting high-quality images. Especially the applications such as the 3D reconstruction of underwater objects, tracking of underwater objects, underwater biological research, and so on. But in general, the underwater images obtained have light attenuation, color distortion, blurriness, and low contrast. Due to this, the underwater images may not be used directly for further applications.

The attenuation of light in the water will lead to getting low contrast images with variation from the actual color. The low contrast images with color variation will not have enough perceptual quality for further processing. The absorption light and diffusion while capturing the images also affect the quality of the images. Due to all these challenges, the research in the domain of underwater image enhancement and underwater restoration got much attention in the past few decades.

In this paper, we propose a new underwater image enhancement scheme with better color correction and enhancement of image features for better usage. The blurred images are enhanced using deep learning techniques.

The further sections of the manuscript are organized as follows: the related underwater image enhancement approaches are discussed in section II. The proposed approach is discussed in section III and section IV gives the analysis of experimental results. In section V, we conclude the paper with a few insights to future works.

II. Related Work

Underwater image enhancement (UWI) is one of the areas where lots of research is going on in the past two decades. Sev-

978-0-7381-2447-6/20/$31.00 ©2020 IEEE

eral underwater image enhancement and restoration techniques are already reported in the literature. For a better understanding of this manuscript, a few important works in this area are reviewed in this section.

A UWI scheme that uses a deep residual framework is discussed in [10]. In this work, the researchers attempted to increase the clarity of the images that had low contrast, color distortion, and blurriness due to natural reasons such as refraction, scattering, and absorption. In this study, the underwater images are generated created synthetically as a training dataset using a cycle-consistent adversarial network (CycleGAN). Further, the very deep super-resolution reconstruction model (VDSR) is used for the enhancement task.

One of the studies shows that the visual quality of underwater images can be improved using many algorithms [3]. A new underwater image quality measure (UIQM) which consists of three components called underwater image colorfulness measure (UICM), underwater image sharpness measure (UISM), and underwater image contrast measure (UIConM) is proposed. The color casting of the underwater images can be improved using UICM. The underwater images will be affected by the severe color-casting problem. Instead the regular statistical values, a asymmetric alpha-trimmed statistical measure is utilized. The UISM is used to evaluate the sharpness of edges. The key assumption is that if the image is sharp the quality of the underwater image is good. Due to the backward scattering issues, the contrast of the underwater images will be less. The UIConM measures the contrast of the underwater images. The contrast of an image is calculated by performing the log AMEE measure on the pixel intensity value of the image. The UIConM measure will be high for an enhanced image and visibility will be good. The UIQM measure helps the researchers to evaluate the perceptual quality of the images. It should be noted that the UIQM measure keeps a better correlation with human visual perception. A few other major works related to underwater image enhancement are discussed in [9]–[11]

III. Methodology

The proposed technique for underwater image enhancement scheme is discussed in this section.

A. Preparation of Training Dataset

The underwater images that we had may not be much realistic and clear. Different techniques can be used for the enhancement of underwater images. We need a lot of underwater images for training. In this phase, we generate a dataset for training. UnderwaterGAN (UWGAN) is used to generate a training dataset in which paired images from enhancing underwater visual perception (EUVP) dataset is used as input. The Pix2Pix is used as the neural network in UWGAN. It has a generator and discriminator, the generator is used to convert the input image to the target image. The structure of the generator contains an encoder that uses a leaky rectified linear unit (LRELU) and a decoder with RELU. There are 8 encoders and decoders. Each layer of the encoder is arranged as convolution layer-batch normalization-LRELU and each decoder layer is arranged as convolution layer-random initializer-RELU-concatenation except for the last decoder layer, which is arranged as convolution layer-initialization.

The discriminator is used to check the similarity between the input image and the target image. The discriminator is 5 layered network, each layer consists of a convolution layer-LReLU. The air images and the corresponding underwater images from the EUVP dataset will be given as input to train UWGAN. Adam Optimizer is used to perform learning in the network. It will compute gradients based on the learning rate. Training is performed based on the equation:

TOTAL STEP = (EPOCHS * num of train / BATCH SIZE)

where EPOCHS specifies the number of times training performed, the number of the training samples used, and BATCH SIZE represent the interval between each updation (number of samples processed). Based on TOTAL STEP generator and discriminator is trained. The WGAN loss function is calculated as follows:

$$\mathrm{L}_{WGAN}(G,D) = E[D(I^C)] - E[D(G(I^D))] + \lambda_{GP} E_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}D(\hat{x})}\|_2 - 1^2] \quad (1)$$

Samples along straight lines between pairs of points coming from the true data distribution and the generator distribution is defined by $P_{\hat{x}}$ and λ_{GP} is a weighing factor. The $L1$ loss is also used to capture low level frequencies in the image.

$$L_{L1} = E[\|I^C - G(I^D)\|_1] \quad (2)$$

After training, testing will be done. It is to check whether the generated image is correct or not. It will identify that generated enhanced image is real or fake. Thus underwater images will be generated from in-air images. Then loss measures of these images will be calculated. The underwater image quality measure (UIQM), structural similarity index (SSIM), and peak signal to noise ratio (PSNR) will be calculated to analyze the quality of the underwater image. By learning this we can enhance the quality of trained images.

1) Peak Signal to Noise Ratio (PSNR): The PSNR calculates the loss method based on noise. This is a reference-based image quality assessment technique in which the degradation of an image can be evaluated by performing the mean square error between the pixels. If the degraded image pixels are exactly matching with the reference image then we will get a PSNR of ∞. The formula to compute the PSNR is given below:

$$RMSE(x,y) = \sqrt{\frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{N}} \quad (3)$$

$$PSNR(x,y) = 20log_{10}[\frac{255}{MSE(x,y)}] \quad (4)$$

Higher PSNR means more noise is removed and the quality of the image increases.

2) Structural Similarity Index (SSIM): The SSIM is used to measure image quality degradation due to losses in data transmission or data compression. It will compare the images based on luminance, contrast, and structure. It calculates absolute error.

$$SSIM(x,y) = \frac{2\mu_x\mu_y + c1}{\mu_x^2 + \mu_y^2 + c1} \times \frac{2\sigma_{xy} + c2}{\sigma_x^2 + \sigma_y^2 + c2} \quad (5)$$

The SSIM value should be a small positive value for better quality images.

3) Underwater Image Colorfulness Measure (UICM): The UICM is used to enhance the color casting of the images which is captured from underwater. Based on the depth of the water while taking an underwater image, the attenuation may change and we will get images with different color casting [4].

In UICM we suppose to use two color components related with the chrominance RG and YB are used in the UICM [3].

$$RG = R - G \quad (6)$$

$$YB = \frac{R+G}{2} - B \quad (7)$$

Let us assume the image is having a size $M \times N$, the total number of pixels $K = M \times N$ and all pixels of the image are sorted. Let $T_{\alpha_L} = \alpha_L K$ and $T_{\alpha_R} = \alpha_R K$ The asymmetric alpha-trimmed mean [5] is defined as

$$\mu_{\alpha,RG} = \frac{1}{K - T_{\alpha_L} - T_{\alpha_R}} \sum_{i=T_{\alpha_L}+1}^{K-T_{\alpha_R}} Intensity_{RG,i} \quad (8)$$

Where theμ represents first-order statistic mean value, and σ^2 represents second order statistics variance.

$$\sigma^2_{\alpha,RG} = \frac{1}{N}\sum_{p=1}^{N}(Intensity_{RG,p} - \mu_{\alpha,RG})^2 \quad (9)$$

The $UICM =$ is defined as follows:

$$-0.0268\sqrt{\mu^2_{\alpha,RG} + \mu^2_{\alpha,YB}} + 0.1586\sqrt{\sigma^2_{\alpha,RG} + \sigma^2_{\alpha,YB}} \quad (10)$$

4) Underwater Image Sharpness Measure (UISM): The UISM is used to improve the sharpness of the underwater image. Enhancement measure estimation (EME) technique is used to evaluate the sharpness of the edges [6] and descatterring algorithm is used to remove the scattering effect of the water while capturing the images [7].

The Sobel edge detector is one of the well-known edge detection algorithms, and firstly, we need to apply it to each plane from RGB color components. The grayscale edge map can be computed by multiplying the resultant edge map obtained from the Sobel edge detector with the original image [3].

$$UISM = \sum_{c=1}^{3}\lambda_c EME(grayscale_c) \quad (11)$$

$$EME = \frac{1}{k_1k_2} \times \sum_{l=1}^{k_1}\sum_{k=1}^{k_2}\log(\frac{I_{max,k,l}}{I_{min,k,l}}) \quad (12)$$

The image will divided into k_1k_2 blocks, the fraction $(I_{max,k,l})/(I_{min,k,l})$ is the relative contrast ratio in each block, and the EME computed in each RGB color component values and are combined linearly with the coefficient λ_c [8].

where
$\lambda_R = 0.299$,
$\lambda_G = 0.587$, and
$\lambda_B = 0.144$.

5) Underwater Image Contrast Measure (UIConM): The UIConM is used to improve the contrast of the underwater image. The backward scattering will cause contrast degradation [4].

logAMEE measure on the intensity image is used to measure contrast [9].

$$UIConM = logAMEE(Intensity) \quad (13)$$

The logAMEE measure is defined as:

logAMEE= $\frac{1}{k_1k_2} \otimes \sum_{l=1}^{k_1}\sum_{k=1}^{k_2}\log(\frac{I_{max,k,l}\ominus I_{min,k,l}}{I_{max,k,l}\oplus I_{min,k,l}}) \times \log(\frac{I_{max,k,l}\ominus I_{min,k,l}}{I_{max,k,l}\oplus I_{min,k,l}})$

The image is partitioned into k_1k_2 blocks and $\otimes$, $\oplus$ and $\ominus$ indicates the PLIP operations, and it provides consistent nonlinear representations with human visual perceptions [9]. Log and PLIP operations give more importance to the areas with low luminance [6]. An image that is enhanced will have greater UIConM value and visibility will be greatly improved.

6) Underwater Image Quality Measure (UIQM): As mentioned earlier, the UIQM included three attribute measures such as UICM, UISM, and UIConM [3]. The UIQM is used to assess the underwater image quality in the same way as human observers. The UIQM always keeps a correlation with human perception and it also provides better color quality [3].

The overall underwater image quality measure is:

$$UIQM = c_1 \times UICM + c_2 \times UISM + c_3 \times UIConM \quad (14)$$

Three parameters c_1,c_2 and c_3 are application dependent.The values are taken as c_1=0.0282, c_2=0.2953 and c_3=3.5753.For better quality image UIQM have greater value.

Sample in-air and generated image is shown in Fig. 1.

B. Underwater Image Enhancement Using Underwater ResNet

The UWIE or image reconstruction using Underwater ResNet (UResNet) is discussed in this section. The UResNet is constructed by ResBlocks, in which the input of one convolution layer is connected to the output of the preceding convolution layer. The UResNet model consists of three parts:

Fig. 1. Samples of in-air and generated underwater images

a head, body, and tail. The head section consists of one convolution layer. The body part stacks 16 ResBlocks organized in such a way to reduce the training time. The order of layers is as follows: [ConvBN-ReLU-Conv-BN]. The tail consists of one convolution layer. This network uses a 3×3 convolution with zero paddings. Due to this, the UResNet can receive inputs with some arbitrary size. In UResNet, edge difference loss (EDL) is also be considered, and the training in asynchronous mode can be used. The addition of BN layers in the network will help in further restoring the details in the image at the time of contrast enhancement. The UResNet considered the loss function as a combination of both MSE loss and EDL [10].

1) Edge Difference Loss (EDL): In most of the image translation models, mean square edge (MSE) Loss is calculated to find the loss. But we are using edge difference loss (EDL) also as the Underwater images suffer significant detail loss of edge information. Models of EDL are penalized and the features of the generated images are promoted to a higher level. The edge information will get sharpened by this. For computing EDL, a Laplacian template represented as the lap is used as a convolutional kernel for performing the convolutional operation.

The Laplacian template $lap = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

The EDL is defined below:

$$EDL = \mathbb{E}\left[\|I^c \otimes lap - I^g \otimes lap\|_2\right] \tag{15}$$

MSE Loss is mathematically defined in the equation below:

$$MSELoss = \mathbb{E}\left[\|I^c - I^g\|_2\right] \tag{16}$$

The loss function is computed from the MSE loss and the EDL in the following way:

$$Loss = MSE\ Loss + k * EDL \tag{17}$$

The proportion of the two parts of the loss is adjusted by k. The value of k will be different for each dataset. This is used to keep the size of the two sections of the loss function at the same level [10].

Fig. 2. In-air and corresponding underwater image samples for training

2) Asynchronous Training Mode: Since the Laplacian operator is sensitive to noise it is hard to find the value of k in equation (17). When the k value is not appropriate, it will lead to the low quality of the output image. So asynchronous training scheme is used. In asynchronous training mode, training is done twice. The EDL is used in the first training round in which the ability of EDL to provide edge information will make use of and helps to restore edge information and details in the network. The influence of EDL in the network is constrained by the second training, which focuses on the pixel level difference between label images and the output that will limit the network for this.

In the first round, the gradient will be computed using EDL, and the updation of weights in the network is performed by backpropagation. In the second round, MSE loss is used to for calculating the gradients and propagated back for the modification of the weights in the network. Training is performed twice for each batch and the updation of weights in the network is performed twice for each batch. The amplification behavior of the Laplacian operator on noise will be suppressed. The Laplacian operator is very sensitive to the edge information and it is also susceptible to noise [10].

IV. Experimental Result

The experimental results are discussed in this section.

A. Dataset Generation

The enhancing underwater visual perception (EUVP) dataset which contains separate sets of paired and unpaired image samples is used for training the UWGAN. We have used the paired images for training the generator and discriminator. Fig. 2 shows samples for training. For testing, we have used a set of 23 test images. Fig. 3 shows the test image samples. The output of testing (shown in Fig. 1) is given as a label image for the UResNet

Fig. 3. Sample test images

B. Comparison to CycleGAN

The generated images are evaluated using the UIQM measures. The below table shows the comparison of CycleGAN and UWGAN based on variuos quality measures. Values in the table shows better quality for images generated using UWGAN compared to that of CycleGAN

TABLE I
EVALUATION OF VARIOUS SCHEMES

Method	CycleGAN	UWGAN
UICM	4.31	5.44
UISM	8.34	7.07
UIConM	0.22	0.25
UIQM	4.36	4.55

C. Enhancement and Evaluation

Input image is enhanced using UResNet. The output is shown in Fig. 4.

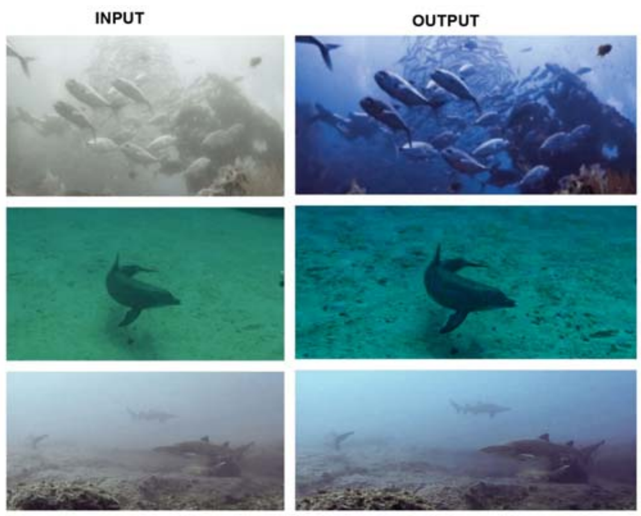

Fig. 4. Input images and corresponding enhanced underwater image

V. CONCLUSION

This paper proposed a method for enhancing the perceptual quality of the underwater image. Firstly, synthetic underwater images are produced using UwGAN. Using UResNet incorporating various loss functions underwater image enhancement is done. It includes edge difference loss (EDL) and mean square edge (MSE) loss. An asynchronous training mode is used to improve the functioning of multi-term loss function. The restoration of the underwater image has a key role in underwater vision operations. It is also used to find the marine pollution level, identify coral reefs destruction, and also for deep-sea explorations. The proposed method provides clear underwater images and this can be used for marine explorations, automated vehicles, and so on. Our system ensures more realistic and trustworthy underwater images. The proposed method tries to enhance the quality of the image and improve the visual effect of underwater images. It is very helpful in the field of vision-based tasks, such as tracking and segmentation.

REFERENCES

[1] Y. Wang, W. Song, G. Fortino, L. Qi, W. Zhang, A. Liotta, "An Experimental-based Review of Image Enhancement and Image Restoration Methods for Underwater Imaging", IEEE Access, pp. 1-19, 2017.

[2] Yakun Gao, Haibin Li, and Shuhuan Wen, "Restoration and Enhancement of Underwater Images Based on Bright Channel Prior", Mathematical Problems in Engineering, 2016. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.

[3] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired under- water image quality measures," IEEE J. Ocean. Eng., vol. 41, no. 3, pp. 541–551, Jul. 2015.

[4] S. Raimondo and C. Silvia, "Underwater image processing: State of the art of restoration and image enhancement methods," EURASIP J. Adv. Signal Process., vol. 2010, 2010, DOI: 10.1155/2010/746052.

[5] J. Bednar and T. L. Watt, "Alpha-trimmed means and their relationship to median filters," IEEE Trans. Acoust. Speech Signal Process., vol. ASSP-32, no. 1, pp. 145–153, Feb. 1984.

[6] K. Panetta, A. Samani, and S. Agaian, "Choosing the optimal spatial domain measure of enhancement for mammogram images," J. Biomed. Imag., 2014, Article ID 937849.

[7] T. Treibitz and Y. Y. Schechner, "Active polarization descattering," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 3, pp. 385–399, Mar. 2009.

[8] W. K. Pratt, Ed., Digital Image Processing. New York, NY, USA: Wiley, 1991, ch. 2.

[9] K. Panetta, S. Agaian, Y. Zhou, and E. J. Wharton, "Parameterized logarithmic framework for image enhancement," IEEE Trans. Syst. Man Cybern. B, Cybern., vol. 41, no. 2, pp. 460–473, Apr. 2011.

[10] Peng Liu, Guoyu Wang , Hao Qi, Chufeng Zhang, Haiyong Zheng , "Underwater Image Enhancement With a Deep Residual Framework", IEEE Access vol.7 pp. 94614 - 94629, July 2019. Available: https://arxiv.org/abs/1801.04011

[11] Jie Li, Katherine A. Skinner,Ryan M. Eustice and Matthew Johnson-Roberson , "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, Jan. 2018.