Revealing the Mysteries of the Deep - Underwater Image Enhancement

[1] **Anirudh Raghavendra Rao, [2] Gurunagasimha S, [3] H Lakshmish, [4] Galla Srinivas, [5] R C Ravindranath**

[1,2,3,4] Eight Semester, Department of Computer Science and Engineering, Global Academy of Technology, Rajarajeshwarinagar, Bengaluru-560098.

[5] Assistant Professor, Department of Computer Science and Engineering, Global Academy of Technology, Rajarajeshwarinagar, Bengaluru-560098.

Abstract: Images that are captured underwater generally suffer from profound categories related to decreased visibility, low brightness, high contrast, high noise and color distortion. Various image processing techniques are used to improve the quality of images, which are often degraded by environmental influences and constraints associated to optical gears. The repair and improvement of underwater images have drawn more and more scientific attention during the last few decades. The survey of underwater images has a crucial problem in the processing of images, is one of the areas that still needs a thorough and in-depth analysis of associated advancements and successes. By examining the contributions and limits of existing approaches through knowledge of underwater image enhancement and restoration, a method has been proposed to improve underwater picture visibility by pre-processing which includes contrast and brightness improvement, denoising and detail recovery. The proposed strategy offers greater overall performance when it comes to the handling of submerged image improvement. The problem of the brightness of enhanced images is solved. The information on texture in the images is properly recovered. Thereby, demonstrating a more robust approach to the problem.

IndexTerms - **Deep Learning, Underwater Image Enhancement, Feature Fusion, Image Processing, Residual Convolution Neural Network, Attention Mechanism.**

I. INTRODUCTION

Underwater image enhancement is an important process for improving the look and quality of images taken in underwater environments. By adjusting the brightness, contrast, and color of the images, the end result can be a much more attractive and vibrant image. The process can help to reduce the amount of noise and artifacts that are common when shooting in underwater environments. With the right settings, the colors in the images can be made to appear much more vivid, and the contrast and saturation can be increased to make the image even more vibrant and attractive. By making these adjustments, the images can be made much more visually appealing. Additionally, it can provide a more enjoyable viewing experience which can make it look better than ever before and provides much more adaptability to underwater surroundings and environment in.

II. LITERATURE REVIEW

To solve drawbacks such as uneven lighting in deep-sea and night conditions as well as image noise, poor contrast and variation in color due to the suspended particles, *Jiachen Zhang et al.* [1] provides low illumination enhancement strategy on attention mechanism, residual dense blocks, and countermeasure generation to address the difficulties of low illumination pictures. This technique uses GIEM which will direct the future modules in improving the lighting. Then, multiple levels of CRM and CARDM derives characteristics that are merged to gain more detailed information. To enhance the image of the underwater environment, a generative countermeasure network is presented where the textural contrast and luminous parameters of the image are enhanced.

In the interest of improving the image quality of backlit photographs without requiring any human input, such as gamma or other factors, *T. Trongtirakul et al.* [2] uses a full-piecewise, non-linear automated stretching method. The technique aims to disclose hidden information in the dark sections of the picture, retain the features and color of both well-lit and over-brightened regions, and improve local contrast in the dark areas. The results of the simulation have proved the efficiency of the suggested method and its superiority over other cutting-edge ways for backlit picture enhancement in terms of visual quality.

Y. Yang et al. [3] have suggested a medical image enhancement algorithm based on wavelet transform. At first, the low-contrast and poor-quality pictures are analyzed by wavelet transform. The high-frequency sub-images are divided by using Haar transform. In the next step, the noise is decreased through a soft threshold approach. The enhanced picture was then produced

using inverse wavelet transform and inverse Haar transform. Finally, the histogram is lengthened using nonlinear histogram equalization. It is quite good at both increasing the details in a picture and successfully keeping its edge characteristics.

Underwater photos frequently suffer from quality concerns including poor contrast, softened features, unnatural colors, uneven illumination, etc. For many real-world applications, underwater picture restoration and enhancement are crucial, and this poses a considerable barrier to image processing and computer vision. *Miao Yang et al.* [4] provide the effects of common underwater picture enhancement and restoration techniques on blue, green, yellow, offshore, and deep-sea photographs, serving as a guide for choosing the technique that will work best in each situation. The commonly utilized underwater picture quality evaluation's restrictions and precision were also examined. This concludes by looking at the possible challenges and unsolved issues related to underwater image enhancement.

The information is sometimes not well conveyed by the images which are taken underwater. There may be sporadic instances of the images becoming extremely brilliant or very dark, with low or high contrast. The research by *Jeevan K M et al.* [5], proposed a wavelet-domain Gabor and median filtering with a spatial-domain adaptive histogram equalization approach for image enhancement. Brightness and contrast are the two factors along with the time domain which is used to assess the effectiveness of filtered images.

Hao Tang et al. [6] gives a method for integrating the YBbCr color channels and RetinexNet into the brightness channel to decrease distortion and suppress noise in improved pictures then afterward recombining the original color and enhanced brightness in the channel directory. The picture is divided into Y, Cb and Cr channels with Y serving as the input to the noise reduction model and the other channels being mixed throughout the enhancement phase.

Codruta O et al. [7] gives a solution that just needs a single image and doesn't call on specialized equipment or an understanding of the scene's organization or underwater circumstances. It is based on the merging of two images, one of which is a color-compensated version and the other is a white-balanced version of the original deteriorated image. In order to promote the transfer of edges and color contrast to the final image, the two images that will be combined as well as the weight maps linked to them are constructed. Additionally, in order to enhance the transfer of edges and color contrast to the output image and to increase accuracy, a multiscale fusion approach is utilized to eliminate artifacts in the low-frequency components of the reconstructed image produced by abrupt weight map shifts.

III. THEORETICAL FRAMEWORK

The aim of underwater image enhancement is to improve the visual look of low-light and low-contrast pictures. The layering of density approach, gray-scale color conversion method, self-materials based on pixel design, rainbow coding method, metal coding method, and frequency domain-based pseudo color improvement algorithm are the most prominent ways for underwater picture enhancement.

In this publication we describe a unique strategy for underwater picture enhancement by proposing a **Dynamic Hetero-Fusion Neural Network (DH-FNN)** shown in **Figure 6**, utilizing a Residual Convolutional U-Net (RCU-N) Architecture and a De-Noising Neural Network (DnCNN). The suggested technique initially employs a U-net architecture to estimate the residual image between both the original underwater picture and the improved image. To better capture the residual picture, the U-Net design is then enhanced with residual blocks following the estimation of the residual picture. A DeNoising Neural Network is applied on the residual image to increase the quality of the output image.

IV. METHODOLOGY

4.1 DIFFERENT VARIANT OF CONVOLUTIONAL UNITS

There are mainly two different architecture variants as quoted by *Md Zahangir Alom et al.* [8], Forward Convolutional Unit (FCU) is a sort of designed to capture visual local information. It is composed of several convolutional filters, each of which conducts a linear transformation on the input picture. In image processing, this technique extracts the image's local characteristics, such as edges and textures. Each convolution's output is then processed via an activation function like ReLU (or) LeakyReLU before being transferred to the next layer. FCU generates an array of extracted features that reflect the input image's local information. The diagrammatic study of this variant is shown in the first part of **Figure 1**.

A Residual Convolutional Unit (RCU) is a means of gathering an image's global information. It consists of filters applied to an input image rather than being applied directly to the image. The filters are put on the incoming images residual the residual is the discrepancy between the source images and its subsequent convolutional layer. This lets the RCU to receive the image and creates a set of feature maps based on the input photo's global knowledge of the input image. The representational study of this variant can be seen in second part of **Figure 1**.

*Figure 1: Structure of Forward Convolutional Unit (FCU) and Residual Convolutional Unit (RCU)*
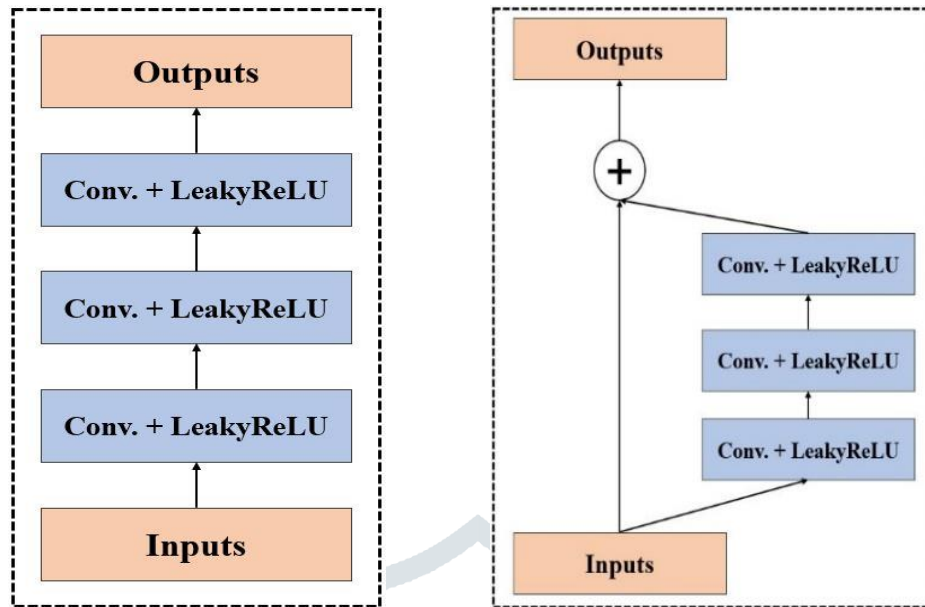
4.2 RESIDUAL CONVOLUTIONAL U-NET (RCU-N) ARCHITECTURE

Residual Convolution U-Net (RCU-N) architecture consists of RCU-N Encoder, RCU-N Decoder in form of RCB; combined with Gating mechanism and Attention blocks along with Skip connections.

Residual Convolution Block (RCB) as shown in **Figure 2** is used to understand the underlying properties of the underwater environment and improve image quality. It provides extra deep learning layers to capture the intricate aspects of the underwater environment. The RCB also aid in the reduction of image noise and outliers, resulting in more accurate and improved images. This function accepts three parameters: filters, number, and tensor. The function begins by generating convolutions with the specified number of filters and a (3x3) kernel size. After the convolutions, it passes through batch normalization [9] layer and LeakyReLU [10] layer. Then the function generates a shortcut link by using (1x1) filters and the same number of filters as the original convolution. After this comes another batch normalizing layer. Ultimately, the function combines the shortcut connection and the original convolutional layers and activates the output using a LeakyReLU.

*Figure 2: Structure of Residual Convolution Block (RCB)*
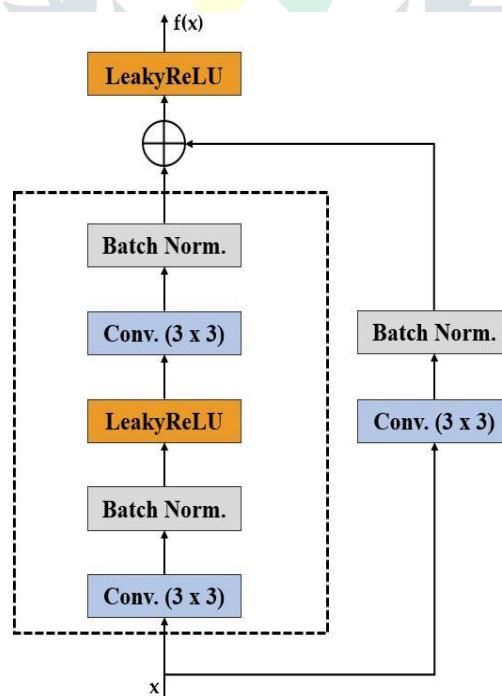
RCU-N Encoder uses RCBs to build a deep convolutional neural network. An input layer is first created of size (256x256x3). A succession of convolutions of RCB is applied with filter sizes ranging from 32 to 1024 and a kernel size of (3x3). After that, a Max Pooling is applied to each block's output to further minimize size. The output layer receives the output of the previous block at the end. This model representation is shown in **Figure 3**.

Figure 3: Residual Convolutional U-Net (RCU-N) Encoder
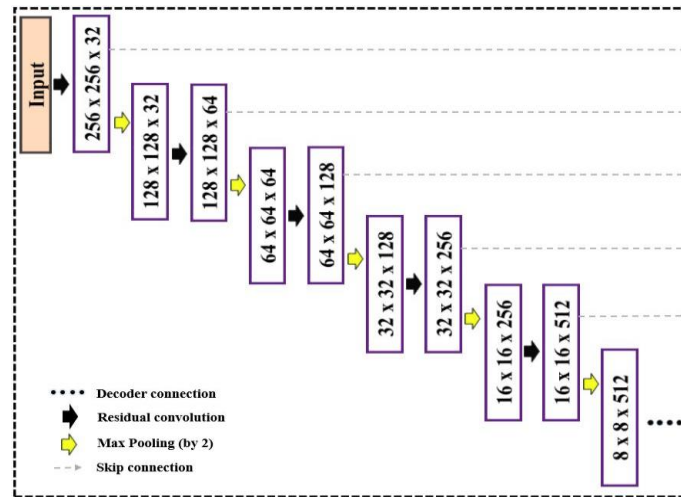
RCU-N Decoder is built with a convolutional neural network with an attention mechanism combined with gating signals and skip connections. The output from each block is then sent through a Gating signal which is used to generate an attention map of the image. The attention map is then concatenated with the output of each up-sampled convolutional block with the help of skip connections and passed through another convolution of kernel size (3x3). This process is repeated until the final output layer is reached, is a convolutional layer with 32 filters. The representation is summarized in **Figure 4**.

Figure 4: Residual Convolutional U-Net (RCU-N) Decoder
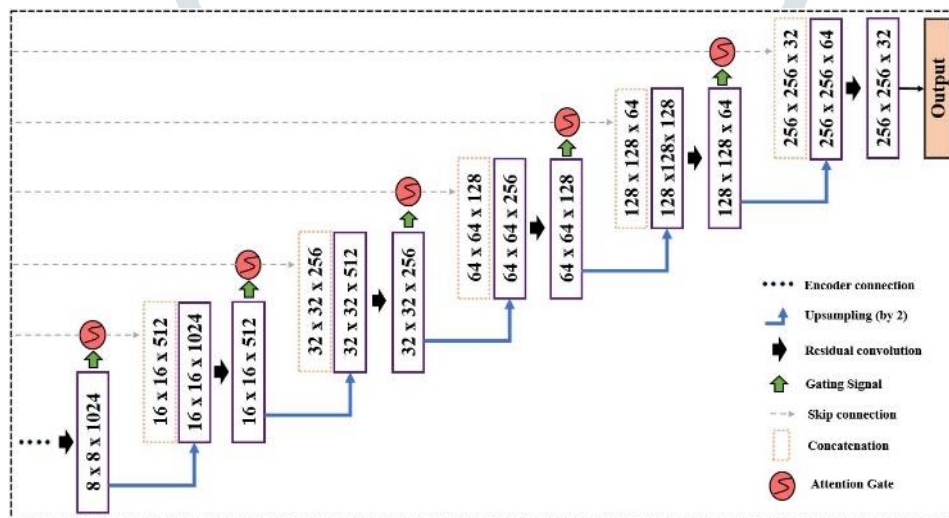
Gating signals (*g*) are utilized as they allow the picture to be divided into several frequency ranges, each of which may be treated individually. This can assist to lessen noise and improve image quality overall. Gating signals may be used to minimize the information that has to be processed and delivered, by minimizing both the cost and power usage of the system. For developing a gating signal, it takes a tensor and applies a convolution with the given number of filters, kernel size, strides, and padding. The output is then batch-normalized before being activated using LeakyReLU. The gating signal is then used to control the flow of information through the network by permitting or disallowing it.

Ozan Oktay et al. [11] explain applying Attention to the areas of the image that contain the most major characteristics and brilliant colors in the form of an Attention gate (AG). As a result, the image is both visually beautiful and analytically valuable. By processing local input layer by layer, convolutional layers gradually extract higher dimensional picture representations (b^l). This eventually divides pixels in a high-dimensional space into groups based on their semantics. The input feature map (b^l) of layer l is produced by progressively using a linear transformation and a non-linear activation function of the Leaky Rectified Linear Unit:

$$\sigma_1 (b^l_{p,q}) = max ((0.1) \times (b^l_{p,q}), b^l_{p,q}) \qquad (1)$$

where p and q are spatial and channel dimensions. The element-wise multiplication of the input feature map (b^l) and the attention coefficient (α) is computed to obtain the modified feature map. A gating signal (g) is used to determine the focus regions based on the activation (σ_1) and contextual data (Eg). To obtain the attention coefficient (α), a convolutional layer is first applied to reduce the tensor size, followed by a transposed convolutional layer to increase the size of the gating signal (g). The two signals are combined and sent through activation of LeakyReLU. A sigmoid activation:

$$\sigma_2 (b_{p,q}) = 1 / (1 + exp (-b_{p,q})) \qquad (2)$$

is then produced using a convolutional layer and trilinear interpolation is used for the resampling of the attention coefficient (α). After multiplying the resampled sigmoid activation by the tensor, the AG is produced using a convolutional layer and a Batch Normalization layer. The Attention mechanism is represented as:

$$z^l_{AM} = \psi^T \left(\sigma_1 \left(E_x^T\, b^l_p + E_g^T\, g_p + j_g \right) \right) + j_\psi \qquad (3)$$

$$\alpha^l_p = \sigma_2 \left(z^l_{AM} \left(b^l_p, g_p ; \Theta_{AM} \right) \right) \qquad\qquad (4)$$

where Θ_{AM} represents a collection of linear transformations Ex, Eg, ψ and bias terms j_ψ, j_g. AG diagram is visualized in **Figure 5** below.

The contrast, sharpness, and clarity of the image is enhanced using the attributes that were extracted. These features are excellent for enhancing underwater images, which are frequently impacted by poor lighting and limited visibility. Feature Evaluation can also be done using Feature Extraction.

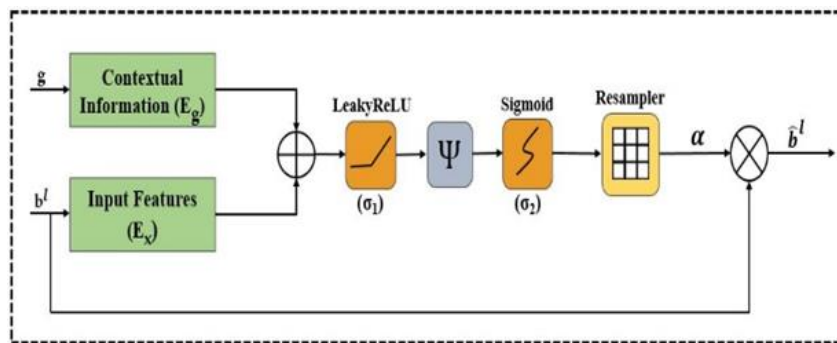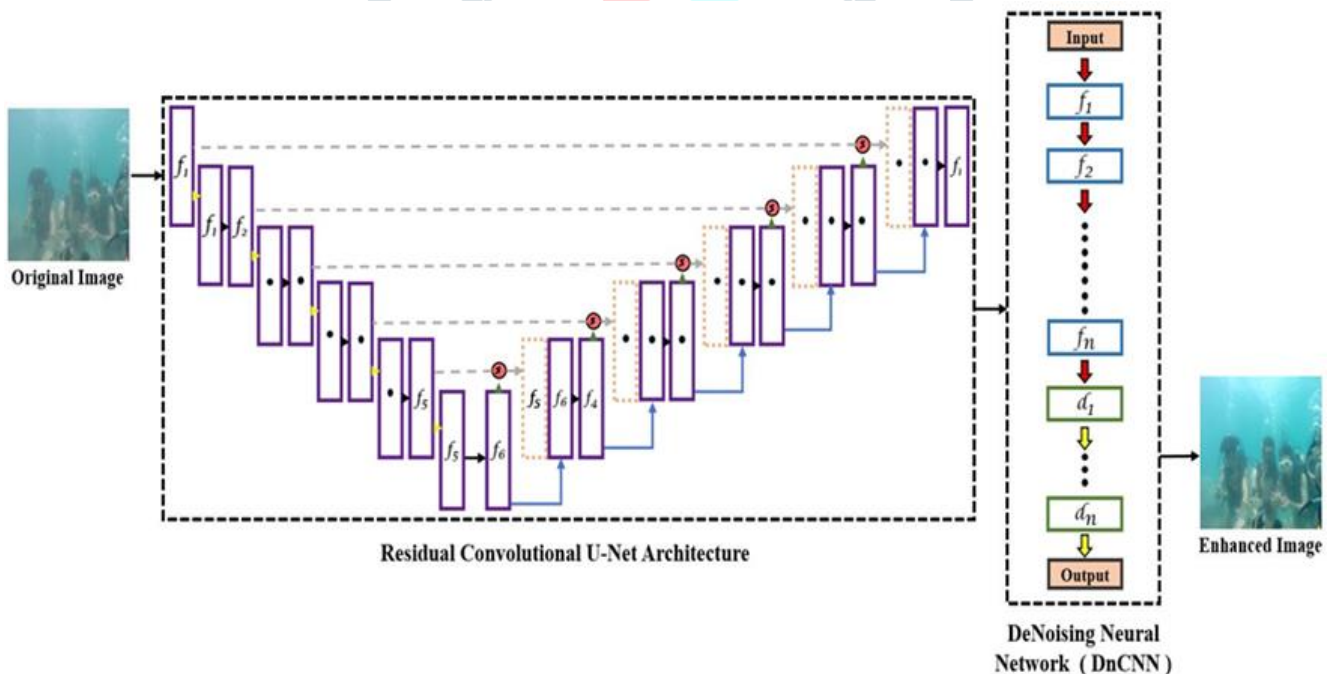*Figure 5: Structure of Attention Gate (AG)*

Figure 6: Dynamic Hetero-Fusion Neural Network (DH-FNN)

4.3 DENOISING NEURAL NETWORK (DnCNN)

DnCNN is a type of CNN which mainly focuses on achieving optimal noise reduction and to enhance the model's adaptability in cacophonous visual settings. It is imperative to undertake additional measures aimed at eradicating any lingering visionary disturbances. DnCNN starts with an input layer, followed by five convolutional layers of filter sizes ranging from 32 to 1024.

The Layer Down Convolution (LDC) function as shown in **Figure 7** starts by creating convolutions of kernel size (3x3). Then it is made to pass through the batch normalization layer and LeakyReLU layer. Use of the convolutional layers is to extract low-level features of the image. It is then followed by a series of dense layers with sigmoid activation layers which are used to

eradicate any noisy parameters still affecting the enhanced image. The output is the enhanced layer of the input image. The DnCNN architecture can be seen in **Figure 8**.
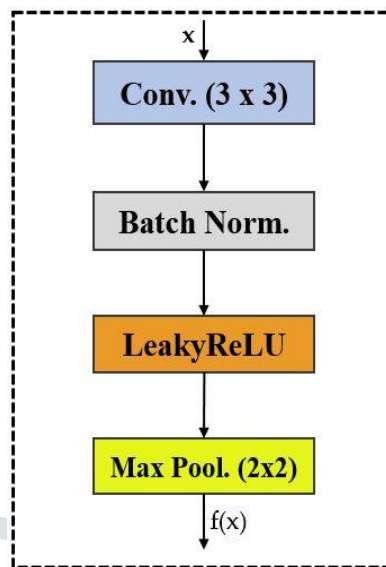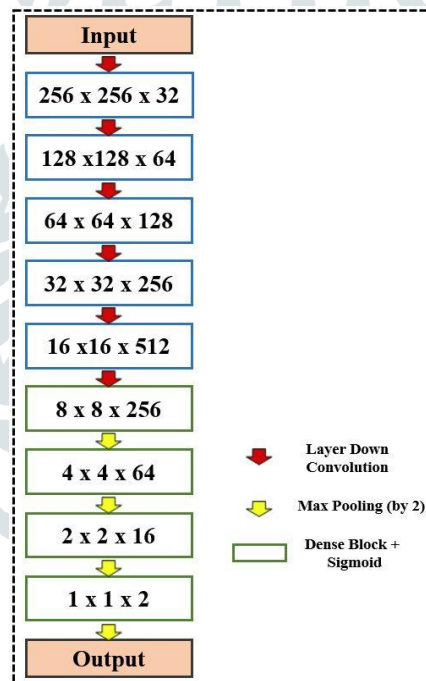
Figure 7: Layer Down Convolution (LDC)

*Figure 8: DeNoising Neural Network (DnCNN)*

V. EXPERIMENT AND RESULT ANALYSIS

The present study employed an experimental method for underwater image enhancement using a network framework implemented on an Intel(R) Core (TM) i9-9900K CPU @ 3.60GHz 3.60 GHz, Intel UHD 630 and NVIDIA GeForce RTX 2070 Graphics and 32 GB memory. The framework was developed with the TensorFlow 2.9.3 framework and Anaconda Python 3.9.16 interpreter on the Windows 10 operating system.

A thorough evaluation of the proposed method was conducted through a comprehensive series of experiments. The Training Loss relative to Validation Loss is shown in the line chart **Figure 9**.

Figure 9: Training and Validation Loss Line Graph
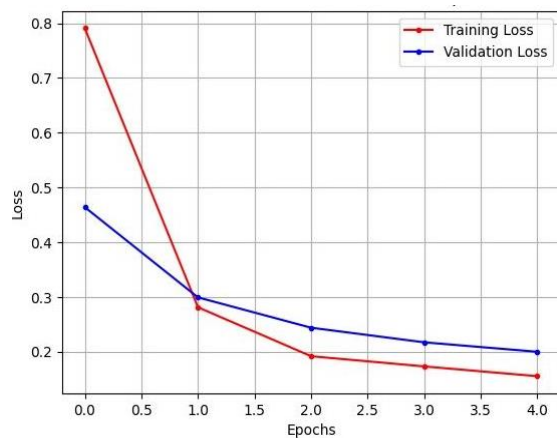
The results of this study indicate the effectiveness of the proposed method in enhancing underwater images which are shown in **Figure 10**.

Figure 10: Image Enhancement based on Subjective Visual Effect Comparison on Test Dataset;
(a) ORIGINAL IMAGE   (b) ENHANCED IMAGE
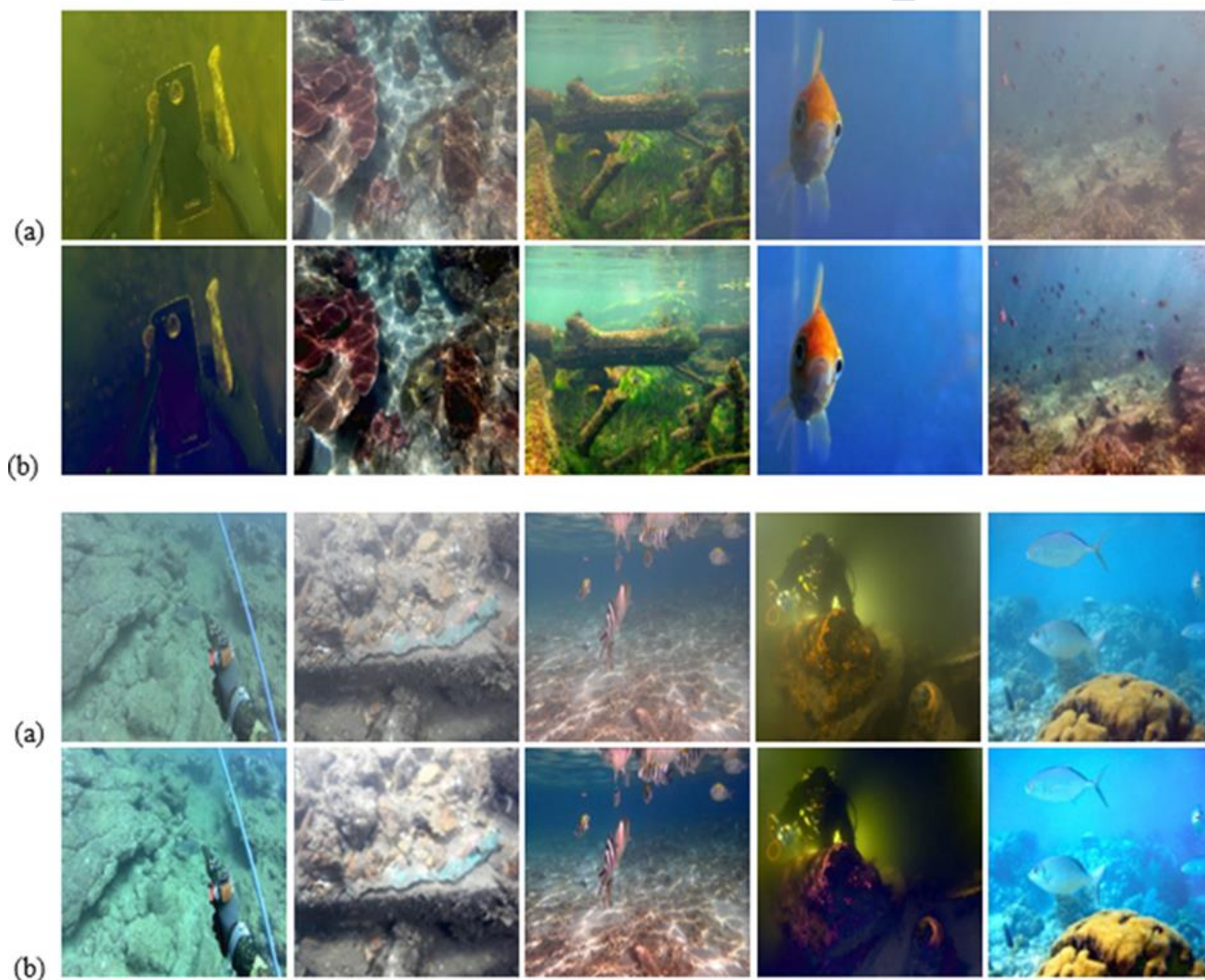
Performance was qualitatively assessed using the Peak Signal to Noise Ratio [12] (PSNR) and Structural Similarity Index [13] [14] (SSIM) indicators on the UIEB [15] test dataset.

Table 1: Comparison of Different Full Reference Image Quality Indexes

Method	PSNR ↑	SSIM ↑
UDCP	16.28	0.70
Fusion	17.90	0.73
MSRNet	20.01	0.8045
MBLLEN	20.25	0.8026
LLCNN	20.63	0.8301
GLADNet	20.90	0.7941
UWCNN	23.62	0.82
FIEGAN	25.66	0.88
DHFFNN	27.58	0.94
ARD-GAN	23.03	0.8475
Ours	**23.71**	**0.8816**

It is evident that the method "Ours" put forth in this paper has demonstrated better performance in the evaluation of Full Reference Image quality [16]. The results obtained from experimentation and comparison with other approaches from *Jiachen Zhang et al.* [1] and *Yuan Tian et al.* [17] can be seen in **Table 1**.

VI. CONCLUSION

An enhancement technique has been provided to address the challenges posed by poor illumination in underwater images as well as the associated problems of contrast loss. Our approach combines Attention mechanism, RCU-N architecture, and DnCNN to effectively extract high and low-level features from underwater images. Specifically, we utilize a U-net with residual convolutional blocks and spatial attention jump to capture high-level features that are robust and accurate, surpassing those achieved by existing methods. The attention mechanism enhances the structure details of underwater images, resulting in more realistic images. Additionally, we incorporate a DeNoising neural network to further eliminate any residual noise, making the model more adaptive to noisy environments. Our model successfully improves the quality of underwater images, making them more visually appealing and providing a powerful tool to enhance the perception of underwater images. Furthermore, our model offers a novel solution to the challenges of feature loss, contrast and luminance distortion in deep convolutional networks for underwater imagery. The findings presented in this work have important implications for improving the quality of underwater imaging and can contribute to the development to enhance the effectiveness and precision of underwater imaging systems.

VII. FUTURE IMPROVEMENTS

Enhancing underwater images is a crucial task in a broad range of applications, including scientific research, environmental monitoring, and exploration. Recent advancements in deep learning techniques have made it possible to use Generative Adversarial Networks (GANs) for underwater image enhancement. By training GANs on still larger underwater image datasets, they can learn to extract features and produce more realistic images, significantly enhancing the quality of underwater images. In addition to luminance and contrast, future enhancements can incorporate other parameters such as color gradients, composition, and depth of field to further improve the quality of underwater images. Recent developments in GANs [18], such as Super-Resolution GANs [19] and StyleGANs [20], have demonstrated the ability to generate higher quality images with more detailed and diverse features. Applying color correction techniques [21] [22] [23] to these generated images can further improve their quality and visual appeal. These advances in deep learning and color correction techniques offer a powerful tool for enhancing underwater images and can have important implications for advancing scientific research, monitoring the environment, and exploring underwater ecosystems. By leveraging these latest advancements, we can significantly improve the accuracy and quality of underwater imagery, leading to a better understanding of the underwater world and its ecosystems.

REFERENCES

[1] Jiachen Zhang, Xuejun Zhou, Fangyuan Zhou, "Underwater Image Enhancement Method Based on Dynamic Heterogeneous Feature Fusion Neural Network", vol. 10, 2022.
[2] Thaweesak Trongtirakul, Werapon Chiracharit, and Sos S. Agaian, ''Single Backlit Image Enhancement'', vol. 8, April 2020.
[3] Y. Yang, Z. Su and L. Sun, "Medical image enhancement algorithm based on wavelet transform", vol. 46, 2010.
[4] Miao Yang, Jintong Hu, Chongyi Li, Gustavo Rohde, Yixiang Du1, Ke Hu, "An In-Depth Survey of Underwater Image Enhancement and Restoration," School of Electronic Engineering, Jiangsu Ocean University, Vol 7, August 2019

[5] Jeevan K M, Anne Gowda A B, Padmaja Vijay Kumar, "An Image Enhancement method based on Gabor Filtering in Wavelet Domain and Adaptive Histogram Equalization", Indonesian Journal of Computer Science Vol. 21, No. 1, January 2021, pp. 146-153

[6] Hao Tang, Hongyu Zhu, Huanjie Tao and Chao Xie, "An Improved Algorithm for Low-Light Image Enhancement Based on RetinexNet." Appl. Sci. 2022, 12, 7268.

[7] Codruta O. Ancuti, Cosmin Ancuti, Christophe De Vleeschouwer, and Philippe Bekaert, "Color Balance and Fusion for Underwater Image Enhancement", Vol. 27, no. 1, January 2018.

[8] Md Zahangir Alom, Mahmudul Hasan, Chris Yakopcic, Tarek M. Taha and Vijayan K. Asari, "Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation", University of Dayton, Dayton, OH, 45469, USA.

[9] Johan Bjorck, Carla Gomes, Bart Selman, Kilian Q. Weinberger, "Understanding Batch Normalization", 2018.

[10] Stamatis Mastromichalakis, "ALReLU: A different approach on Leaky ReLU activation function to improve Neural Networks Performance", 29 Apr 2021.

[11] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, Ben Glocker and Daniel Rueckert, "Attention U-Net: Learning Where to Look for the Pancreas", 20 May, 2018.

[12] K. -H. Thung and P. Raveendran, "A survey of image quality measures," 2009 International Conference for Technical Postgraduates (TECHPOS), Kuala Lumpur, Malaysia, 2009.

[13] Jim Nilsson, Tomas Akenine-Möller, "Understanding SSIM", 29 Jun 2020.

[14] Alain Horé, Djemel Ziou, "Image quality metrics: PSNR vs. SSIM", International Conference on Pattern Recognition, 2010.

[15] C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, D. Tao, "An Underwater Image Enhancement Benchmark Dataset and Beyond," IEEE Trans. Image Process., Vol. 29, 2019.

[16] S. Bosse, D. Maniry, K. -R. Müller, T. Wiegand and W. Samek, "Deep Neural Networks for No-Reference and Full-Reference Image Quality Assessment", IEEE Transactions on Image Processing, Vol. 27, No. 1, Jan. 2018.

[17] Yuan Tian, Yuang Xu and Jun Zhou, "Underwater Image Enhancement Method Based on Feature Fusion Neural Network", Vol. 10, 13 October 2022.

[18] Tingting Zhang, Yujie Li, Shinya Takahashi, "Underwater image enhancement using improved generative adversarial network", 29 April 2020.

[19] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi Twitter, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network", 25 May 2017.

[20] Aviv Gabbay, Yedid Hoshen, "Style Generator Inversion for Image Enhancement and Animation", 5 Jun 2019.

[21] Derya Akkaynak, Tali Treibitz, "Sea-thru: A Method for Removing Water from Underwater Images", University of Haifa, 2019.

[22] Weidong Zhang , Xipeng Pan, Xiwang Xie, Lingqiao Li, Zimin Wang, Chu Han, "Color correction and adaptive contrast enhancement for underwater image enhancement", 18 January 2021.

[23] Zheng Lianga, Yafei Wanga, Xueyan Dinga , Zetian Mi, Xianping Fu, "Single underwater image enhancement by attenuation map guided color correction and detail preserved dehazing", 27 March 2020.