ENHANCEMENT OF UNDERWATER IMAGES WITH ARTIFICIAL INTELLIGENCE

Z. Ertan [1], B. Korkut [1], G. Gördük [1], B. Kulavuz [1], T. Bakırman [1 *], B. Bayram [1]

[1] Yildiz Technical University, Civil Engineering Faculty, 34220 Esenler Istanbul, Türkiye - zeynepertannnn@gmail.com, korkutberat78@gmail.com, gunsu.gorduk@gmail.com, bahadir.kulavuz@std.yildiz.edu.tr, bakirman@yildiz.edu.tr, bayram@yildiz.edu.tr

KEY WORDS: Generative Adversarial Networks, Underwater Image, Deep Learning, Image Enhancement, Pix2Pix

ABSTRACT:

Camera systems using optical sensors have made great progress in recent years in obtaining underwater images. Very high-resolution images can be obtained with underwater cameras. However, it becomes difficult to process the image obtained due to many distorting factors such as inhomogeneous underwater lighting, low contrast, blur, sea snow. For this purpose, image enhancement algorithms are used to minimise the problems in the captured images. The physical characteristics of the underwater environment cause degradation effects that are not found in normal images captured in air. Due to the effects such as colour distortion, low contrast and brightness, blurred details in images captured in this type of environment, the usage area of these images is very limited. This project aims to use artificial intelligence methods for colour enhancement of underwater images. In addition, one of the main objectives of the project is to examine the effect of colour enhancement applied to images on other areas of use. Within the scope of the project, An Underwater Image Enhancement Benchmark Dataset and Beyond Dataset and Large Scale Underwater Image Dataset datasets were used. In our study, U-Shape Transformer architecture and Pix2Pix architecture were tested for their usability for image enhancement. According to the results obtained, Pix2Pix architecture achieved the highest accuracy. Accuracy results were obtained as 22.5884 and 0.8764 for PSNR and SSIM, respectively using UIEB dataset. The accuracy values for LSUI dataset are 25.1010 and 0.8440 for PSNR and SSIM, respectively.

1. INTRODUCTION

A large part of our world is covered by oceans and seas (Gleic, 1993). Despite this, underwater imaging studies are not as common and practical as terrestrial studies and only small part of the marine zone has been mapped (Wright & Heyman, 2008). Underwater imaging can suffer from problems that affect image quality, such as low contrast, colour distortion and poor visual appearance (Schettini & Corchs, 2010). Such problems are caused by the scattering and refraction of light as it passes from rarer to denser environments. The scattering of light therefore reduces colour contrast (Anwar & Li, 2020).

Underwater imagery is a good source of information for gaining insight into marine organisms, detecting rich cultural heritage underwater, and studying hydrothermal vents on the seafloor. Currently, underwater images can be obtained with optical cameras and underwater techniques such as laser scanning, distance selective transmission and polarised light. However, raw underwater images cannot be directly exploited due to effects of the water medium. Therefore, there are various pre-processing methods for underwater image enhancement in the literature.

In recent years, many underwater image enhancement methods have been proposed for underwater images, such as underwater polarisation imaging techniques, underwater dark channel prior (UDCP) method, histogram equalisation (HE), adaptive histogram equalisation (AHE), adaptive histogram equalisation with contrast constraint (Roser M., 2014), and various deep learning methods (Schechner Y., 2005; Treitbitz T., 2009). Underwater image enhancement methods can be divided into three categories which are based on physical models, non-physical models and deep learning (Li et al., 2019).

Non-physical model-based methods aim to improve the image by changing pixel values. Although fusion-based methods improve overall contrast and visibility, they do not produce effective results in local regions of the resulting images (Antuci & Antuci, 2012). Rayleigh-based histogram method solves the problems in local regions; however, it generally provides noisy results (Ghani & Isa, 2015). In an underwater image colour correction method based on weakly supervised colour transfer, a cross-domain mapping function is learned between underwater images and aerial images (Li et al. 2018a).

Physical model-based methods treat this problem as an inverse problem, building physical models and then estimating the model parameters. For example, a method based on the different absorption amounts of the three bands underwater removes the effects of light scattering, however, it is not effective for underwater images that do not have a strong difference in attenuation between the three bands (Carlevaris-Bianco et al. 2010). Another method, which integrates an image blurring algorithm with a wavelength-dependent compensation algorithm, is not effective for processing underwater images due to significant colour differences (Chiang & Chen, 2012). Another method, which combines an underwater image blurring algorithm with a contrast enhancement algorithm, simultaneously produces two enhanced underwater images with different levels of detail (Li et al. 2016).

Supplementary information-based methods generally utilize additional information from multiple images or specialized hardware (Sheinin & Schechner, 2015). The use of these methods has been limited, as multiple images or additional hardware are unlikely to be available in most cases.

Deep learning, convolutional neural networks (CNNs), and generative adversarial networks (GANs) have been applied to

* Corresponding author

https://doi.org/10.5194/isprs-archives-XLVIII-4-W9-2024-149-2024

underwater image restoration in recent years. For example, in a study carried out by Wang et al. (2017), an architecture called UIE-Net is proposed that focuses on color correction and haze removal. WaterGAN architecture uses a two-stage network for underwater image restoration by first simulating underwater images from aerial imagery and depth matches, especially to remove color casts (Li et al. 2018b). Water CycleGAN utilizes adversarial architecture and a multi-stage loss function (Li et al. 2018a). In another study, Anwar et al. (2018) developed UWCNN architecture directly reconstructs the clear latent underwater image by leveraging on an automatic end-to-end and data-driven training mechanism. Guo et al. (2019) proposed Dense GAN method that learns the distribution of real images using multiple loss functions as a multi-scale architecture. Islam et al. (2020) presents a conditional GAN based architecture for real-time underwater image enhancement exploiting an objective function that evaluates the perceptual image quality based on its global content, colour, local texture, and style information. Chen et al. (2021) developed HybridDetectionGAN architecture that embeds two perceptual enhancement models, each of which uses a deep enhancement model with a detection preceptor that provides feedback information in the form of gradients. Fu et al. (2022) proposes probabilistic network to learn the enhancement distribution of degraded underwater images which combines conditional variational autoencoder with adaptive instance normalization to construct the enhancement distribution. Wang et al. (2023) focused on domain adaptation for underwater image enhancement and developed a two-phase Underwater Domain Adaptation network (TUDA) that aims to simultaneously minimize the inter-domain and intra-domain gap. In one of the latest studies to date, Zhang et al. (2024) introduces a lightweight network called LiteEnhanceNet for single underwater image enhancement that requires less computational sources for real-time applications and deployment on underwater mobile devices.

As can be seen in the literature review, methods are generally proposed based on CNN and GAN. In our study, the importance of dataset in the enhancement of underwater images was tested by using 2 different datasets with Pix2pix architecture and U-Shape Transformers architecture.

2. MATERIALS AND METHODS

2.1 Deep Learning Architectures

Artificial intelligence is a technology that can create intelligent systems that can imitate human intelligence. Machine learning, one of the most important methods for achieving artificial intelligence, enables machines to learn from past data or experiences without being programmed. Deep learning, the most popular machine learning technique in recent years, offers fast, reliable, and effective solutions using artificial neural networks. Deep learning is one of the fastest developing methods in big data analytics and is considered a breakthrough technology (Zhou et al. 2017).

After the great success of AlexNet (Krizhevsky et al., 2012) on the ImageNet dataset using CNNs, deep learning approaches have become widespread. Many deep learning networks have been proposed in the literature for image enhancement over time.

In this study, we employ two deep learning architectures which are based on GAN and transformers, respectively. In this regard, we exploited Pix2Pix conditional GAN and U-Net based Shape Transformers architectures.

2.1.1 Pix2Pix Conditional GAN

Pix2Pix GAN is a deep convolutional neural network that can transform an image into another image. The most important advantage of this model is that it can transform between images at higher resolution than other GAN models. The Pix2Pix model is a type of conditional GAN in which the generation of the output image requires input and, in this case, depends on a source image.

Many problems in image processing and computer vision require the generation of an output image corresponding to an input image. For example, it is difficult to transform from one image domain to another domain in operations such as colorizing a black and white image, obtaining a semantic label map from an edge map, etc. Pix2Pix is an architecture that takes an important step in solving such problems. Pix2Pix is a special version of conditional GAN architecture. This model includes L1 cost value in addition to the conditional GAN architecture. The L1 metric is a similarity comparison between the generated image and the target image (Isola P, 2017).

Pix2Pix generative network produces high quality images using the U-Net architecture. The U-Net architecture uses jump links to preserve the details of the image. The encoder section consists of a series of convolution layers that summarize the input image by reducing it to a smaller size. The decoder section consists of a series of convolution layers that generate the output image using the summary from the encoder section. Each layer in the encoder section halves the size of the input image. This allows the encoder section to compress the summary of the input image into a smaller size. Each layer in the decoder section doubles the size of the input image. In this way, the decoder section can generate the output image using the hash. As a result, the Pix2Pix producer network can produce high-quality images using the U-Net architecture. This architecture is used to generate the output image while preserving the details of the input image. A general GAN-based underwater image enhancement architecture is given in Figure 1.

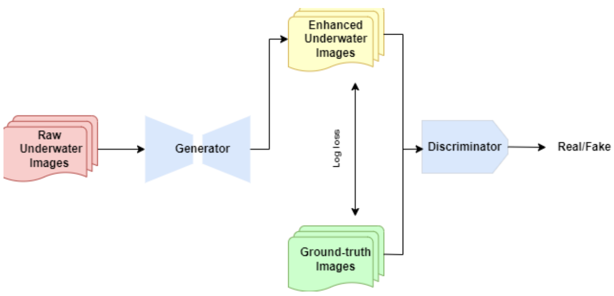


Figure 1. A sample GAN-based underwater image enhancement architecture

2.1.2 U-Net Shaped Transformer

The first example of the use of transformers on images was given by Dosovitskiy et al. (2020). Unlike traditional models, it does not use a repetition mechanism, but instead determines global dependencies between input and output based purely on the attention mechanism.

A transformer model is a neural network that learns context, and hence meaning, by tracking relationships in sequential data, such as the words in this sentence. Transformer models apply a set of still-developing mathematical techniques, called interest or self-interest, to detect subtle connections of even disparate data items in a sequence that influence and depend on each other (Vaswani A., 2023).

In this project, the U-shape Transformer architecture proposed by Peng et al. (2022) is used. The overall architecture of the U-shape Transformer consists of a generator and a discriminator based on CMSFFT (a channel wise multi-scale feature fusion transformer) and SGFMT (spatial-wise global feature modelling transformer). In the coding part of the generator, the original image is subsampled three times respectively, except the input is directly inserted into the network. Then, after 1x1 convolution, the three-scale feature map is input into the corresponding scale convolution block. The outputs of the four-convolution block are the inputs of CMSFFT and SGFMT. After feature remapping in the decoding part of the generator, the output of SGFMT is directly sent to the first convolutional block. Meanwhile, four convolutional blocks with various scales receive four outputs from the CMSFFT: feature map output by its own upper layer, feature map of the corresponding dimension from the decoding part, and feature map generated using the reference image by 1x1 convolution after down sampling to the corresponding dimension. With the defined multi-scale connections, the gradient flow can flow freely at multiple scales between the generator and the discriminator, so that a stable training process can be achieved, enriching the details of the generated images (Peng et al., 2022).

2.2 Underwater Image Enhancement Datasets

Datasets used in deep learning are usually classified as labelled or unlabeled. Labelled datasets are datasets that contain an output value in each instance, such as a class label or numeric value. Labelled datasets are the most used dataset type for training deep learning models. Unlabeled datasets are datasets that do not contain an output value in each instance. Unlabeled datasets can be used to train deep learning models, however, are more difficult compared to labelled datasets.

Datasets used in deep learning are usually large datasets. This is due to the fact that deep learning models need a sufficient amount of data to learn complex patterns. The quality of datasets in deep learning can significantly affect the performance of the models. It is important that the data is accurate and complete. It is also important that the data is representative. That is, the dataset should represent the real-world data to which the model will be applied.

The choice of datasets in deep learning depends on the nature of the problem and the dataset used. It must be ensured that the data is accurate and complete. The size of the dataset should be considered, and its diversity should be taken into account.

In this study, we employ two widely used underwater image enhancement datasets in the literature namely Underwater Image Enhancement Benchmark Dataset and Beyond (UIEB) and Large Scale Underwater Image Veri Seti (LSUI).

2.2.1 UIEB Dataset

UIEB dataset has been created by Li et al. (2020) which contains 950 real-world underwater images. These 950 images were divided into two subsets. 890 of these raw underwater images were provided with corresponding high-quality reference images. The remaining 60 images were defined as challenging data since the authors were not able to obtain satisfactory reference images.

In order to create a diverse dataset, the authors gathered images from Google, YouTube, related papers in the literature and self-captured videos considering various scenes and different characteristics of quality degradation including coral and marine life. The reference images are created using a commercial software that are evaluated by 50 volunteers.

The underwater images in the UIEB dataset are 3-band RGB images and have a total of 289 different image sizes. In this study, we exploited the main subset which contains 890 raw underwater images and corresponding enhanced images. Figure 2 shows examples of input and output images in the dataset.

Figure 2. Sample images from UIEB dataset. First row: raw underwater images, second row: corresponding enhanced images

2.2.2 LSUI Dataset

LSUI dataset has been created by Peng et al. (2023) which contains 4279 real-world underwater images. The raw images in the dataset are gathered from self-captured images, websites and public datasets containing rich water scenes, various water types and lighting conditions. The reference images are created using various image enhancement methods and selected by the results of two round subjective and objective evaluations. One of the main highlights of LSUI dataset is that also includes deep-sea and underwater cave scenes.

There are 4279 input and output images in the LSUI dataset. The images in the dataset are 3-band RGB images and there are 19 different image sizes in total. Figure 3 shows examples of input and output images in the dataset.

2.3 Accuracy Metrics

Image quality assessment is the process of objectively measuring the quality of images. In this study, in order to assess the results of the deep learning experiments we utilized two widely used accuracy metrics in the literature namely peak signal-to-noise ratio (PSNR) and structural similarity (SSIM).

2.3.1 Peak Signal-to-Noise Ratio

PSNR is an engineering term for the ratio between the maximum possible power of a signal and the power of interfering noise that affects the accuracy of its representation. Since many signals have a very wide dynamic range, PSNR is usually expressed as a logarithmic quantity using the decibel scale.

Figure 3. Sample images from LSUI dataset. First row: raw underwater images, second row: corresponding enhanced images

PSNR is a metric commonly used to measure reconstruction quality for images and video subject to loss due to compression. PSNR is used in various applications such as image compression, image enhancement and image quality assessment. Image compression is the process of converting images to take up less space. PSNR can be used to evaluate the performance of image compression algorithms. For example, an image compression algorithm will achieve a high PSNR value if it produces an image similar to the original image.

Image enhancement is the process of improving the quality of degraded images. PSNR can be also used to evaluate the performance of image enhancement algorithms. For example, if an image enhancement algorithm can remove noise from a noisy image, it will achieve a high PSNR value (Hore, 2010). PSNR can be used to measure the quality of an image by comparing it with images obtained using various image processing algorithms. It has a high reliability and repeatability. The equation that is used to calculate PSNR is given in Equation (1):

$$PSNR(f,g) = 10\log_{10}(255^2 / MSE(f,g)), \quad (1)$$

where f = reference image
g = enhanced image
255 = maximum grey value of a pixel
MSE = mean squared error

Mean squared error (MSE) can be obtained with Equation (2):

$$MSE(f,g) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(f-g)^2, \quad (2)$$

where M = the number of rows in the images
N = the number of columns in the images
i = the index of the row
j = the index of the column

2.3.2 Structural Similarity

SSIM is a method for estimating the perceived quality of digital television and cinematic pictures, as well as other types of digital pictures and videos. SSIM is used to measure the similarity between two images. The SSIM index is a complete reference metric; in other words, the measurement or estimation of image quality is based on an uncompressed or distortion-free first image as a reference. The difference from other techniques such as MSE or PSNR is that these approaches estimate absolute errors, measure the details and contrast of the image.

The equation that is used to calculate SSIM is given in Equation (3):

$$SSIM(f,g) = l(f,g)c(f,g)s(f,g), \quad (3)$$

where f = reference image
g = enhanced image
l = luminance
c = contrast
s = structure

In Equation (3), luminance, contrast and structure is calculated by Equation (4), Equation (5) and Equation (6), respectively:

$$l(f,g) = \frac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1}, \quad (4)$$

$$c(f,g) = \frac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2}, \quad (5)$$

$$s(f,g) = \frac{\sigma_{fg} + C_3}{\sigma_f\sigma_g + C_3}, \quad (6)$$

where μ_f = reference image pixel sample mean
μ_g = enhanced image pixel sample mean
σ_f^2 = reference image variance
σ_g^2 = enhanced image variance
σ_{fg} = covariance of reference and enhanced image
$C_1 = (0.01 \times 255)^2$
$C_2 = (0.03 \times 255)^2$
$C_3 = C_2 / 2$

3. RESULTS & DISCUSSION

The images in both datasets have varying sizes as mentioned in the previous section. Therefore, all images in the LSUI and UIEB datasets were resized to 512x512 in the training of Pix2Pix architecture. The parameters used in the training phase of Pix2Pix are shown in Table 1. These hyperparameters were determined empirically following number of experiments.

	UIEB	LSUI
Number of Train Images	800	3849
Number of Test Images	90	430
Image Size	512 x 512	512 x 512
Number of Epochs	100	100
Learning rate	0.0002	0.0002
Optimizer	Adam	Adam
Activation function	Sigmoid, tanh, relu	Sigmoid, tanh, relu

Table 1. Hyperparameters for Pix2Pix architecture

In the training on the U-Net Shaped Transformer architecture, the images in the LSUI and UIEB datasets were resized to 256x256. The parameters used in the training phase are shown in Table 2. These hyperparameters were determined empirically following number of experiments.

	UIEB	LSUI
Number of Train Images	800	3849
Number of Test Images	90	430
Image Size	256 x 256	256 x 256
Number of Epochs	360	360
Learning rate	0.0002	0.0002
Optimizer	Adam	Adam
Activation function	Sigmoid, tanh, relu	Sigmoid, tanh, relu

Table 2. Hyperparameters for U-Net Shaped Transformers architecture

The accuracy results calculated from the test dataset are given in Table 3 for both architectures and datasets. Based on our experiments conducted within the scope of this study, the accuracy results show that both architectures provide better results with LSUI dataset. This is most likely caused by the number of images in the dataset. However, SSIM metric obtained with UIEB dataset using Pix2Pix architecture is higher than the one with LSUI dataset. It can be also said that Pix2Pix architecture provided better results than U-Net Shaped Transformer in all accuracy metrics for both dataset using the hyperparameter setting in this study.

		Pix2Pix	Transformer
UIEB	**PSNR/SSIM**	22.5884/0.8764	18.2722/0.7142
LSUI	**PSNR/SSIM**	25.1010/0.8440	21.0529/0.7870

Table 3. Accuracy results for both architectures

The authors of UIEB dataset (Li et al. 2020) proposed Water-Net architecture for image enhancement which is a gated fusion network that fuses the inputs with the predicted confidence maps to achieve the enhanced underwater image. In their study, Water-Net achieved 19.1130 db PSNR and 0.7971 SSIM accuracy. Compared to our results, Pix2Pix outperformed proposed Water-Net architecture with 3.4754 db increase in PSNR and 7.94% increase in SSIM metric. However, we can see that U-Net Shaped Transformer was not able to catch up with Water-Net in terms of both accuracy metrics.

The authors of LSUI dataset (Peng et al. 2023) proposed the U-Net Shaped Transformer architecture which is also used within this study. In their study, U-Net Shaped Transformer achieved 24.16 db PSNR and 0.93 SSIM accuracy using 400 images in the LSUI dataset which they refer to Test-L400. In our study, it looks like we were not able reproduce the results with the U-Net Shaped Transformer as the original authors. However, the authors do not indicate which 400 images are chosen for the test dataset. Therefore, it was not possible to perform a matching comparison since we used 430 images in our assessments. On the other hand, we outperformed their accuracy results using Pix2Pix architecture with 0.941 db increase in PSNR. However, their SSIM metric is still higher than Pix2Pix results. They also run their model on UIEB test dataset which they refer to Test-U90 and achieved 22.91 db PSNR and 0.91 SSIM accuracy which are slightly higher than our best results with Pix2Pix.

Figure 4 shows prediction outputs from our best resulting trainings in terms of PSNR which is Pix2Pix architecture with both datasets. The top row in the figure shows that the network trained on UIEB dataset can provide sufficient results even in environments with low visibility. On the other hand, the prediction seems to lack brightness compared to reference image. The second row shows a better prediction example in terms of both quality and color characteristics. The assessment of Pix2Pix architecture trained on LSUI dataset shows that even though the prediction results are quite satisfactory in terms of color and brightness, the resolution of the predictions seem to have slightly decreased (third and last row on Figure 4).

4. CONCLUSION

The application of deep learning techniques in underwater image enhancement has exhibited remarkable advancements, transcending the limitations posed by challenging underwater conditions. Through the deployment of CNNs, GANs, Transformers, and other innovative architectures, this field has witnessed substantial progress in improving image quality, contrast, and overall visibility in underwater environments.
The adaptability of these models to learn from limited and often noisy underwater data has been a focus point, allowing for the extraction of meaningful features and the restoration of degraded images with unprecedented accuracy. However, while these methodologies showcase promise, there remains a scope for further refinement and optimization to address the complexity of underwater scenes comprehensively.

The robustness of these models across diverse underwater scenarios, including varying water types, depths, and lighting conditions, remains a significant challenge. Achieving consistent performance in real-time applications, where rapid decision-making and processing are crucial, presents another avenue for exploration. Moreover, ensuring the generalizability of these deep learning approaches beyond the training data distributions is imperative for their practical deployment in real-world underwater settings.

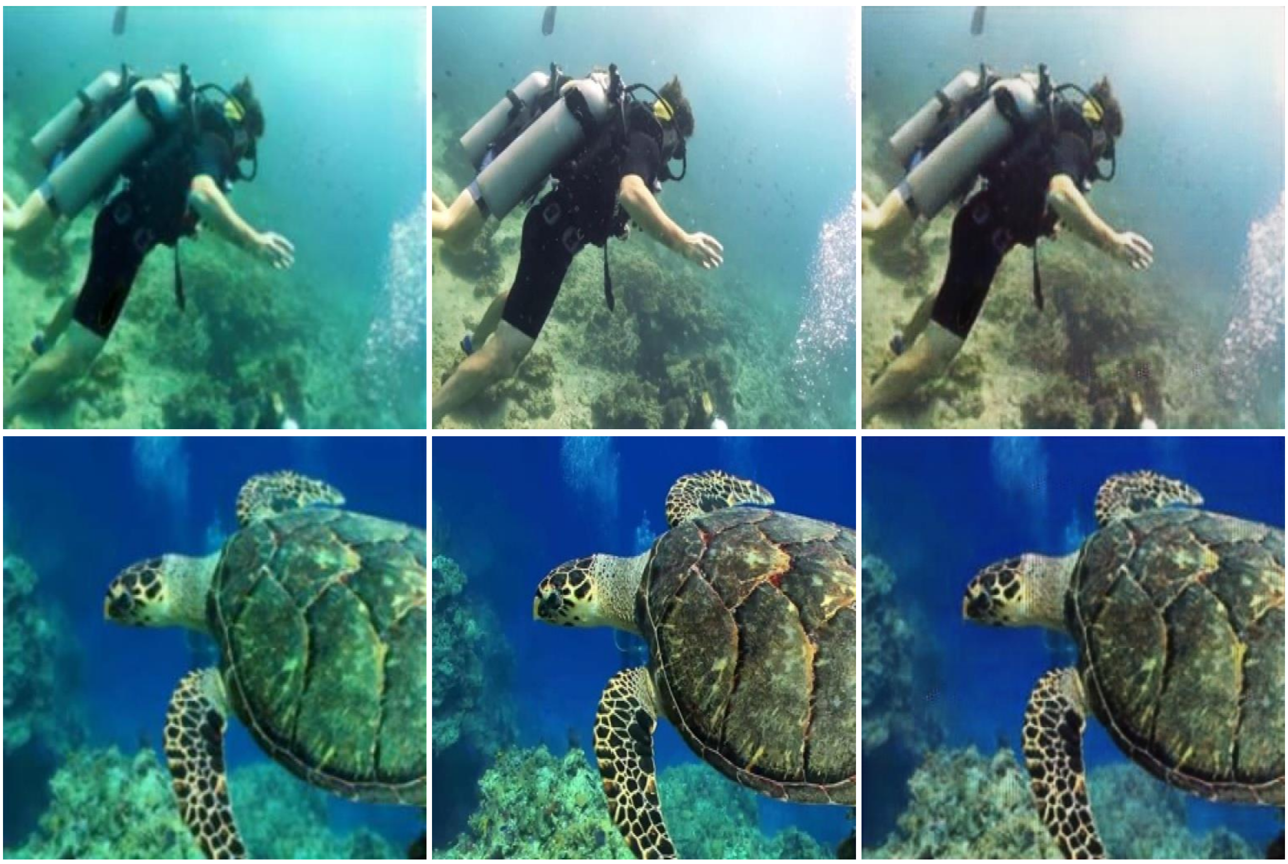
In this study, we employed GAN and Transformer based deep learning architectures using two widely used open access datasets for underwater image enhancement. Our results show that GAN-based methods have a great potential in this field in terms of accuracy and visual results. However, the number of images in the underwater image enhancement datasets are still quite limited which also limits the performance of the networks.

Future research endeavors should focus on not only enhancing visual perception but also considering the integration of real-time processing, robustness to varying conditions, and adaptability across diverse underwater settings to ensure practical applicability and efficacy of these techniques. Collaborative efforts between computer vision experts, marine biologists, and underwater imaging specialists could pave the way for more comprehensive and domain-specific advancements, leading to the development of tailored algorithms capable of addressing the nuanced challenges inherent in underwater image enhancement. By combining domain knowledge with cutting-edge deep learning methodologies, the potential for unlocking clearer, more detailed representations of the underwater world remains a compelling frontier in scientific exploration and technological innovation. In this regard, we aim to investigate the effects of enhanced images to underwater image segmentation or detection task to understand deep learning-based image enhancement on other tasks.

ACKNOWLEDGEMENTS

This study has been supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under 2209-B project number: 1139B412200670.

(a)



(b)

Figure 4. Pix2Pix prediction outputs from the test dataset. (a) UIEB dataset, (b) LSUI dataset. Left Column: Raw input image, middle column: ground-truth image, right column: prediction

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-4-W9-2024-149-2024

REFERENCES

Ancuti, C., Ancuti, O., 2012. Enhancing underwater images and videos by fusion, Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR), 2012, 81-88.

Anwar, S., Li, C., 2020. Diving deeper into underwater image enhancement: A survey. Signal Processing: Image Communication 89, 115978.

Anwar, S., Li, C., Porikli, F., 2018. Deep Underwater Image Enhancement. arXiv:1807.03528.

Carlevaris-Bianco, N., Mohan, A., Eustice, R., 2010. Initial results in underwater single image dehazing. Proc. of IEEE OCEANS, 1-8. DOI: 10.1109/OCEANS.2010.5664428.

Chen, L., Jiang, Z., Tong, L., Liu, Z., Zhao, A., Zhang, Q., Dong, J., Zhou, H., 2021. "Perceptual Underwater Image Enhancement with Deep Learning and Physical Priors," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 8, pp. 3078-3092.

Chiang, J., Chen, Y., 2012. Underwater image enhancement by wavelength compensation and dehazing. IEEE Transactions on Image Processing, 21(4), 1756-1769.
DOI: 10.1109/TIP.2011.2179666.

Dosovitskiy, A., 2020. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv: 2010.11929.

Fu, Z., Wang, W., Huang, Y., Ding, X., Ma, KK., 2022. “Uncertainty Inspired Underwater Image Enhancement,” In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds) Computer Vision – ECCV 2022. Lecture Notes in Computer Science, vol 13678. Springer, Cham. https://doi.org/10.1007/978-3-031-19797-0_27.

Ghani, A., Isa, M., 2015. Enhancement of low quality underwater image through integrated global and local contrast correction. Applied Soft Computing 37 (2015): 332-344. https://doi.org/10.1016/j.asoc.2015.08.033.

Gleick, P., 1993. Water in crisis. Vol. 100. New York: Oxford University Press.

Guo, Y., Li, H., Zhuang, P., 2019. Underwater image enhancement using a multiscale dense generative adversarial network,” IEEE J. Oceanic. Eng., 1-9. DOI: 10.1109/JOE.2019.2911447.

Hore, A., Ziou, D., 2010. Image quality metrics: PSNR vs. SSIM. In 2010 20th international conference on pattern recognition (pp. 2366-2369). IEEE. DOI:10.1109/ICPR.2010.579.

Islam, J., Xia, Y., Sattar, J., 2020. "Fast Underwater Image Enhancement for Improved Visual Perception," in IEEE Robotics and Automation Letters, vol. 5, no. 2, pp. 3227-3234, April 2020. DOI: 10.1109/LRA.2020.2974710.

Isola, P., Zhu, Y., Zhou, T., Efros, A., 2017. Image-to-Image Translation with Conditional Adversarial Networks. Proceedings- 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 5967–5976. https://doi.org/10.48550/arXiv.1611.07004.

Krizhevsky, A., Sutskever, I., Hinton, E., 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25, 1097-1105. https://doi.org/10.1145/3065386.

Li, C., Ji, C., Cong, R., Pang, Y., Wang, B., 2016. Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior, IEEE Transactions on Image Processing, 25(12), 5664-5677. DOI: 10.1109/TIP.2016.2612882.

Li, C., Guo, J., Guo, C., 2018a. Emerging from water: Underwater image color correction based on weakly supervised color transfer, IEEE Signal Process. Lett., 25(3), 323-327. DOI: 10.1109/LSP.2018.2792050.

Li, C., Guo, C., Ren, W., R, Cong, H, Ho, S, Kwong, Tao, D., 2020. “An underwater image enhancement benchmark dataset and beyond,” IEEE T. Image Process., vol. 29, pp. 4376–4389. DOI: 10.1109/TIP.2019.2955241.

Li, J., Sinner, K., Eustice, R., 2018b. WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images. IEEE Robot. Autom. Lett., 3(1), 387-394. DOI: 10.1109/LRA.2017.2730363.

Peng, L., Zhu, C., Bian, L., 2023. "U-Shape Transformer for Underwater Image Enhancement," in IEEE Transactions on Image Processing, vol. 32, pp. 3066-3079. DOI: 10.1109/TIP.2023.3276332.

Roser, M., Dunbabin, M., Geiger, A., 2014. Simultaneous underwater visibility assessment, enhancement and improved stereo, IEEE International Conference on Robotics and Automation, Hong Kong, China.

Schechner, Y., Karpel, N., 2005. Recovery of underwater visibility and structure by polarization analysis. IEEE J. Ocean. 30(3):570–587. DOI: 10.1109/JOE.2005.850871.

Schettini, R., Corchs, S., 2010. Underwater image processing: state of the art of restoration and image enhancement methods. EURASIP journal on advances in signal processing, 1-14. DOI:10.1155/2010/746052.

Sheinin, M., Schechner, Y., 2015. The next best underwater view, Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR) , 3764-3773.

Treibitz, T., Schechner, Y., 2009. The next best underwater view, Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., Polosukhin I.,2023. Attention Is All You Need. https://doi.org/10.48550/arXiv.1706.03762.

Wang, Y., Zhang, J., Cao, Y., Wang, Z.,2017. A deep CNN method for underwater image enhancement. IEEE International Conference on Image Processing (ICIP), Beijing, China. DOI: 10.1109/ICIP.2017.8296508.

Wang, Z., Shen, L., Xu, M., Yu, M., Wang, K., Lin, Y., 2023. "Domain Adaptation for Underwater Image Enhancement," in IEEE Transactions on Image Processing, vol. 32, pp. 1442-1457, DOI: 10.1109/TIP.2023.3244647.

Wright, D., Heyman, D., 2008. Introduction to the special issue: Marine and coastal GIS for geomorphology, habitat mapping, and marine reserves. Marine Geodesy, 31(4), 223–230. https://doi.org/10.1080/01490410802466306.

Zhang, S., Zhao, S., An, D., Li, D., Zhao, R., 2024. LiteEnhanceNet: A lightweight network for real-time single underwater image enhancement. Expert Systems with Applications, 240, 122546.

Zhou, L., Pan, S., Wang, J., Vasilakos, A.V. 2017. "Machine learning on big data: Opportunities and challenges," Neurocomputing, 237, 350 – 361.