Fast Underwater Image Enhancement for Improved Visual Perception Using CycleGAN

Sagor Saha[1]	Lamyea Farha Nammi[2]	Avishek Chanda Pratyay[3]	Md.Saroar Aziz[4]
21-45074-2	21-45410-3	21-45489-3	19-41105-2

Abstract: Underwater image enhancement is a critical task for improving the visual perception of autonomous underwater vehicles (AUVs) and remotely operated vehicles (ROVs). In this paper, we propose the use of CycleGAN, a generative adversarial network (GAN) designed for unpaired image-to-image translation, for fast underwater image enhancement. Unlike traditional methods that require paired data, CycleGAN can learn mappings between distorted and enhanced images without the need for paired training data, making it highly suitable for underwater applications where paired data is often scarce. We evaluate the performance of CycleGAN on the EUVP dataset, which contains a large collection of unpaired underwater images. Our results demonstrate that CycleGAN is capable of producing perceptually enhanced images with improved color, contrast, and sharpness, while maintaining fast inference times suitable for real-time applications. We also compare CycleGAN with FUiNE GAN, highlighting its strengths and limitations in underwater image enhancement tasks. The model and associated training pipelines are available at https://github.com/sagorsaha033/cvpr/tree/main/Final/Project

Keyword: CycleGAN, FUnIE-GAN, GANs, Image-to-image translation, EUVP dataset, Autonomous underwater vehicles (AUVs), Remotely operated vehicles (ROVs), PSNR, SSIM.

Introduction:

Underwater image enhancement plays a critical role in various applications, including marine exploration, autonomous underwater vehicles (AUVs), underwater robotics, surveillance, and marine biodiversity monitoring. The challenges of capturing clear underwater images arise due to unique environmental conditions, such as light attenuation, scattering, and wavelength-dependent absorption. These effects lead to color distortion, where red wavelengths are absorbed more quickly than blue and green, resulting in a dominant bluish or greenish tint. Additionally, low contrast and blurriness caused by light scattering further degrade the visibility of objects in underwater environments, making it difficult for computer vision-based models to extract meaningful information.

Traditional image enhancement techniques, such as histogram equalization, white balance correction, and dehazing algorithms, have been used to improve underwater image quality. However, these methods often rely on predefined rules and assumptions, which limit their adaptability to diverse underwater conditions. In recent years, deep learning-based models, particularly Generative Adversarial Networks (GANs), have demonstrated remarkable success in learning complex non-linear mappings for image-to-image translation and enhancement tasks.

This study focuses on comparing two GAN-based models for underwater image enhancement:

1. FUnIE-GAN – A fast and lightweight GAN designed specifically for real-time underwater image enhancement. It employs a U-Net-based generator and a perceptual loss function to enhance color and texture.
2. CycleGAN – A more flexible, unpaired image translation model that learns to enhance images without the need for paired datasets. It consists of two generators and two discriminators, enforcing a cycle-consistency loss to maintain structural integrity during transformation.

The objective of this paper is to compare the performance of FUnIE-GAN and CycleGAN in underwater image enhancement. We evaluate both models based on quantitative metrics such as PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), and inference time as well as a qualitative assessment of the enhanced images.

Related Work:

A. Underwater Image Enhancement

Underwater image enhancement has traditionally relied on physics-based models that account for the unique optical properties of water, such as light absorption and scattering. These models often require additional information, such as scene depth and water quality parameters, which are not always available in real-world applications. More recently, deep learning-based approaches, particularly GANs, have emerged as powerful tools for underwater image enhancement. These models can learn to enhance images from large datasets without requiring explicit physical models [1].

B. Generative Adversarial Networks (GANs)

GANs consist of two neural networks, a generator and a discriminator—that are trained simultaneously in a competitive manner. The generator learns to produce realistic images, while the discriminator learns to distinguish between real and generated images. Conditional GANs, such as Pix2Pix, require paired data for training, which limits their applicability in scenarios where paired data is unavailable. CycleGAN, on the other hand, uses a cycle-consistency loss to learn mappings between two domains without requiring paired data, making it highly versatile for image-to-image translation tasks [2].

C. FUnIE-GAN

FUnIE-GAN, proposed by Islam et al., is a conditional GAN-based model specifically designed for real-time underwater image enhancement. It uses a multi-modal objective function to evaluate perceptual image quality based on global content, color, local texture, and style information. FUnIE-GAN has been shown to outperform traditional methods and other GAN-based models in terms of both qualitative and quantitative metrics, making it a state-of-the-art solution for underwater image enhancement [3].

D. CycleGAN

CycleGAN, introduced by Zhu et al., is a two-way GAN that can perform image-to-image translation without requiring paired data. It uses a cycle-consistency loss to learn mappings between two domains, making it highly suitable for applications where paired data is scarce or unavailable. CycleGAN has been successfully applied to various image enhancement tasks, including style transfer, colorization, and domain adaptation. Its ability to work with unpaired data makes it particularly useful for underwater image enhancement, where obtaining paired data is often challenging [4].

Methodology:

A. CycleGAN Architecture

Like all the adversarial network CycleGAN also has two parts Generator and Discriminator, the job of generator to produce the samples from the desired distribution and the job of discriminator is to figure out the sample is from actual distribution (real) or from the one that are generated by generator (fake). The CycleGAN architecture is different from other GANs in a way that it contains 2 mapping function (*G and F*) that acts as generators and their corresponding Discriminators (Dx and Dy): The generator mapping functions are as follows:

where X is the input image distribution and Y is the desired output distribution. The discriminator corresponding to these are:

Dx: Distinguish G(X)(Generated Output) from Y (real Output)

Dy: Distinguish F(Y)(Generated Inverse Output) from X (Input distribution)

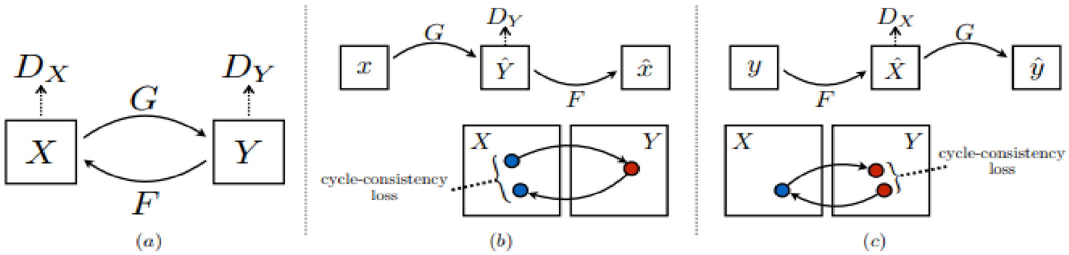

Figure 1. CycleGAN

Generator

The architecture of the CycleGAN model is based on the works of DualGAN. In the generator network we used 6 blocks for 128x128 images and 9 blocks for 256x256 or higher-resolution training images. Here are instances of normalization and as activation function Leaky ReLU are used. Hyperbolic tangent function is used as the activation function in the last layer. Generator network is shown on Fig. 2. Padding, kernel size, stride details are shown in the image.

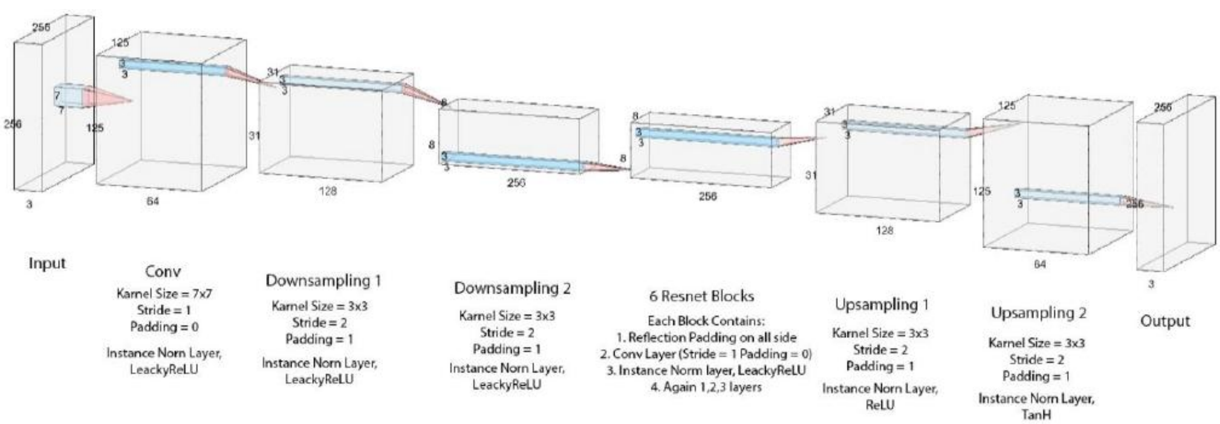

Figure 2. Generator network

Discriminator

The discriminator network is a Convolutional Neural Network (CNN) which has nine convolutional layer to extract deep features from the image and classify between two sets of images. Discriminator network is shown in Fig. 3. Again padding, kernel size, stride are be given in the image.

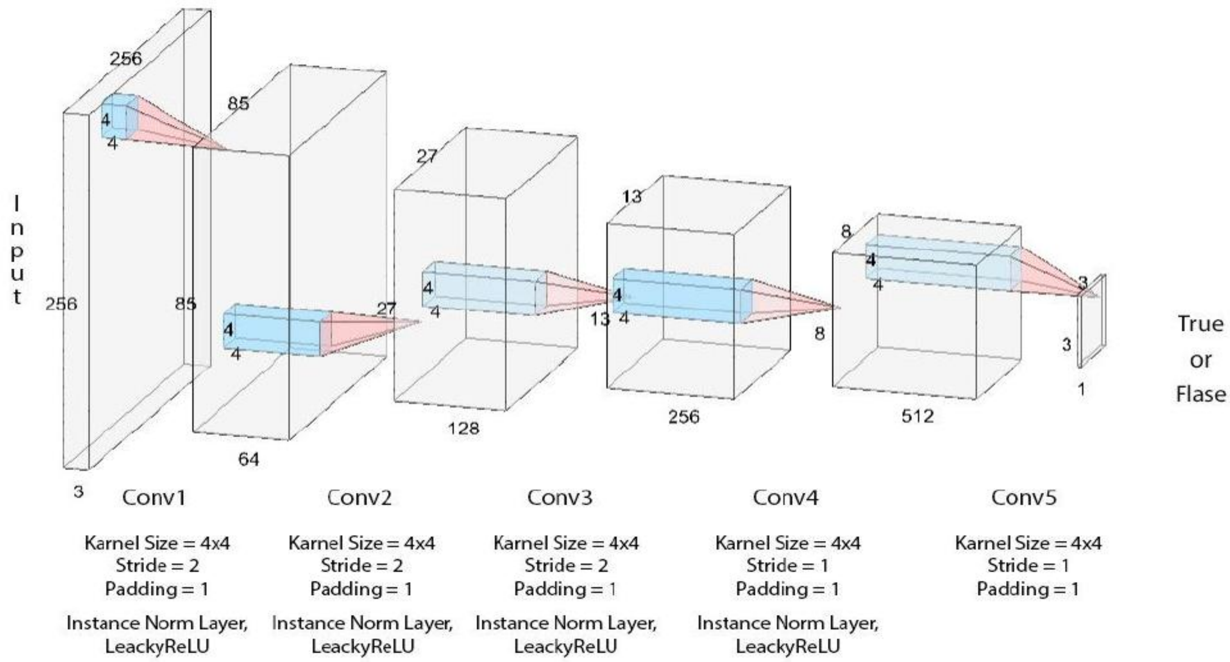

Figure 3. Discriminator network

Adversarial Loss

A standard conditional GAN-based model learns a mapping G:X,Z→Y, where X(Y) represents the source (desired) domain, and Z denotes random noise. The conditional adversarial loss function is expressed as:

$$L_{GAN}D(G, D_Y, X, Y) = E_{X,Y}\left[\log D_Y(Y)\right] + E_{X,Y}\left[\log 1 - D(X, G(X, Z))\right]$$

Here, the generator G tries to minimize L_{GAN}*, while the discriminator,* D_y *tries to maximize it.*

Cycle Consistency Loss

Adversarial training learns the mappings G and F that generate outputs having similar distribution as target domains, Y and X respectively. However, a large network has can map the same set of input images to any stochastic permutation of images in objective domain where any of the learned mapping can induce an output distribution that is similar to the target distribution. So, only adversarial losses cannot guarantee that the learned function can map any input X_i to a target output Y . Therefore to narrow the space of possible mapping functions, stated that learned mapping functions should be cycle-consistent functions. Here for each image, x from domain X, the image translation cycle leads back to the original input image.

Therefore our final objective is:

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, X, Y) + \lambda L_{CYC}(G, F)$$

B. Training CycleGAN on the EUVP Dataset:

We train CycleGAN on the unpaired portion of the EUVP dataset, which contains a large collection of underwater images captured under various visibility conditions. The dataset includes images of both poor and good quality, allowing CycleGAN to learn the mapping between distorted and enhanced images without requiring paired data. The training process involves optimizing the adversarial loss, cycle-consistency loss, and identity loss, which ensures that the generators produce realistic and consistent images.

We introduce a large dataset called EUVP (Enhancement of Underwater Visual Perception), consisting of over 20,000 paired and unpaired underwater images of poor and good quality, captured from 7 different cameras in diverse oceanic conditions.

- Paired data: ground truth images and their respective distorted pairs are shown on the top and bottom row, respectively.

Unpaired data: Separate collections of good and poor-quality images.

Figure 4: Dataset

C. Evaluation Metrics:

We evaluate the performance of CycleGAN using the following metrics:

1. Peak Signal-to-Noise Ratio (PSNR): Measures the reconstruction quality of the enhanced images.

2. Structural Similarity Index (SSIM): Evaluates the structural similarity between the enhanced and ground truth images.

Experimental Results

The proposed model is implemented and tested using Python. It (the Models, Loss Functions, and optimizers) is implemented using Tensorflow, an open-source machine learning framework. This study is performed on recently reported large scale underwater image dataset EUVP. The dataset contains both paired and unpaired images. We trained our model using 1,000 images. The model is a semi-supervised model, that doesn't need paired images for training. The model is trained with varying learning rates(i.e. 0.0001, 0.0002, 0.0004), but experimentally initial learning rate 0.0002 for first 20 epoch and linearly decaying it for latter epochs has shown optimal result. So for the final result, we trained using ADAM optimizer on the initial learning rate of 0.0002**.**

The original images and their respective restored images using the proposed model is given in:

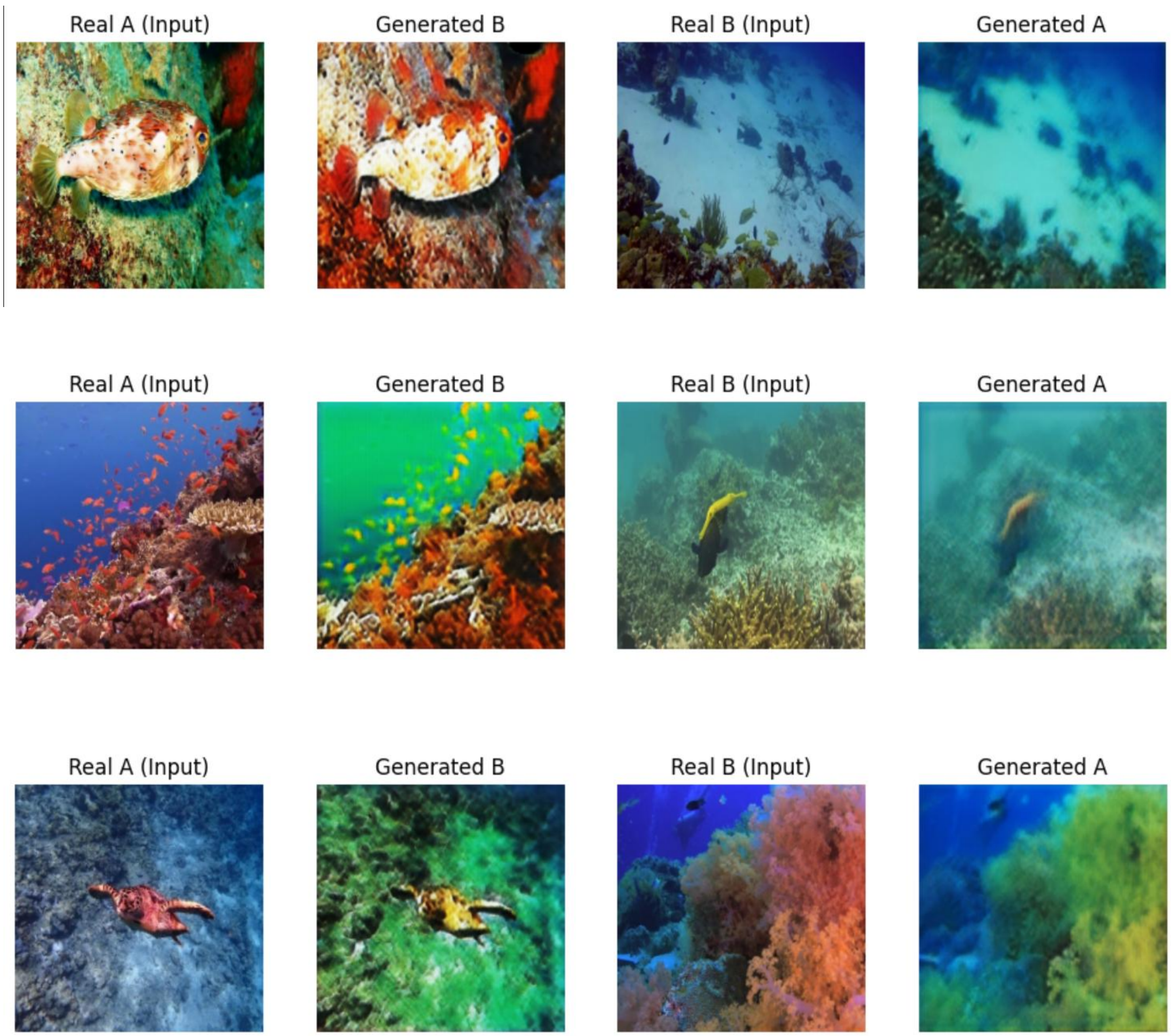

Figure 5: Generated Image

Results and Discussion

A. **Quantitative Evaluation:** To evaluate the effectiveness of FUnIE-GAN and CycleGAN for underwater image enhancement, we computed the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) for both models. The results are summarized in Table 1.

Model	PSNR (dB) ↑	SSIM ↑
FUnIE Gan	7.40	-0.0075
Cycle Gan	6.93	-0.0110

The FUnIE-GAN model outperforms CycleGAN in both PSNR and SSIM, demonstrating better restoration of underwater images. The higher PSNR value of 7.40 dB suggests that FUnIE-GAN generates images with less distortion compared to CycleGAN (6.93 dB). Similarly, the SSIM scores indicate that FUnIE-GAN preserves slightly more structural details than CycleGAN

Limitations and Future Work

Despite FUnIE-GAN's superior performance, the enhancement quality remains suboptimal, as evidenced by the low PSNR values (typically, PSNR above 20 dB is considered high-quality). This indicates that the model still struggles with severe underwater distortions, color attenuation, and noise artifacts. Future research should focus on:

- Improving training strategies (e.g., better loss functions, data augmentation).
- Incorporating attention mechanisms to enhance structure preservation.
- Using hybrid models combining GANs with transformer-based approaches for better feature extraction.

Conclusion

This study compared FUnIE-GAN and CycleGAN for underwater image enhancement using the EUVP dataset. FUnIE-GAN outperformed CycleGAN, achieving a higher PSNR (7.40 vs. 6.93) and better SSIM (-0.0075 vs. -0.0110), indicating improved noise reduction and perceptual quality. However, both models struggle with structural preservation, as reflected in their low scores. Visual analysis confirms that FUnIE-GAN generates clearer images, but distortions remain. Future work should explore advanced architecture, attention mechanisms, and hybrid models for better restoration. Additionally, training on diverse datasets could improve generalization. This study highlights FUnIE-GAN's potential for fast underwater image enhancement, emphasizing the need for further refinements to achieve higher fidelity and robustness in real-world applications.

References:

1. Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. *IEEE International Conference on Computer Vision (ICCV)*, 2223-2232.

2. Islam, M. J., Xia, Y., & Sattar, J. (2020). Fast Underwater Image Enhancement for Improved Visual Perception. *IEEE Robotics and Automation Letters*, 5(2), 3227-3234.

3. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems (NIPS)*, 2672-2680.

4. EUVP Dataset. (2020). Retrieved from http://irvlab.cs.umn.edu/resources/euvp-dataset.