ISSN No:-2456-2165

GAN based Underwater Image Enhancement

Meenakshi Venkat[1,] Sreelakshmi K V[2], Athul B. Elavunkal[3], Jeshmol P J[4]
Asst.Professor[4], Information Technology
Rajagiri School of Engineering & Technology Kakkanad, Kerala

Abstract:- Underwater image enhancement has drawn a lot of attention because of its significance in marine engineering and aquatic robotics. Many techniques for enhancing underwater photographs have been put forth in recent years. Since light propagation underwater and in the atmosphere are different, a particular set of non-linear visual distortions occur. These distortions are brought on by a variety of factors. Red wavelengths are absorbed in deep water as light travels further, which is why underwater images frequently have a green or blue colour as the dominating hue. Low-contrast, fuzzy, and color-degraded images are the result of such wavelength-dependent attenuation, scattering, and other optical properties of water bodies that cause irregular distortions. The previous CNN-GAN (Generative Adversarial Network) based model for real-time underwater image enhancement is sped up by the upgraded inception model proposed by GAN-Based Underwater Image Enhancement. The suggested model assesses image quality based on its global colour, content, local texture, and style information to construct a perceptual loss function. The dataset being used, called EUVP (Enhancing Underwater Visual Perception), is made up of paired and unpaired collections of underwater images captured by seven distinct cameras under a variety of visibility conditions during maritime explorations and cooperative experiments. The suggested model's accuracy can be increased by learning to enhance and improve underwater image quality from both paired and unpaired training.

***Keywords**:- GAN, EUVP, Image enhancement, Underwater images.*

I. INTRODUCTION

Visually-guided AUVs (Autonomous Underwater Vehicles) and ROVs (Remotely Operated Vehicles) are widely used in important applications such as the monitoring of marine species migration and coral reefs, an inspection of submarine cables and wreckage, underwater scene analysis, seabed mapping, human-robot collaboration, and more. One major operational challenge for these underwater robots is that despite using high-end cameras, visual sensing is often greatly affected by poor visibility, light refraction, absorption, and scattering. These optical artifacts trigger non-linear distortions in the captured images, which severely affect the performance of vision-based tasks such as tracking, detection and classification, segmentation, and visual servoing. Fast and accurate image enhancement techniques can alleviate these problems by restoring the perceptual and statistical qualities of distorted images in real time.

As light propagation differs underwater (than in the atmosphere), a unique set of non-linear image distortions occur which are propelled by a variety of factors. For instance, red wavelengths are absorbed in deep water (where light travels further), giving underwater images a predominating green or blue colour. The irregular non-linear distortions caused by such wavelength-dependent attenuation, scattering, and other optical characteristics of the water bodies lead to poor contrast, frequently blurry, and color-degraded images. For dehazing and colour correction, in particular, physics-based systems can model and accurately estimate some of these factors. However, many robotic applications do not always have access to data like scene depth and optical water-quality measurements. Furthermore, these models frequently require too much computational power for real-time implementations.

A practical alternative is to approximate the underlying solution by learning-based methods, which demonstrated remarkable success in recent years. Several models based on deep Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) provide state-of-the-art performance in learning to enhance perceptual image quality from a large collection of paired or unpaired data. For underwater imagery a number of GAN-based models and CNN-based residual models report inspiring progress for automatic color enhancement, dehazing, and contrast adjustment. However, there is significant room for improvement as learning perceptual enhancement for underwater imagery is a more challenging ill-posed problem (than terrestrial imagery). Additionally, due to the high costs and difficulties associated with acquiring large-scale underwater data, most learning-based models use small-scale and often only synthetically generated images that fail to capture a wide range of natural variability. Moreover, designing robust yet efficient image enhancement models and investigating their applicability for improving real-time underwater visual perception have not been explored in the literature in depth.

Such challenges are addressed by designing a fast underwater image enhancement model and analysing its feasibility for real-time applications. The problem is formulated as an image-to-image translation problem by assuming there exists a non-linear mapping between the distorted (input) and enhanced (output) images. Then, a conditional GAN-based model is designed to learn this mapping by adversarial training on a large-scale dataset named EUVP (Enhancement of Underwater Visual Perception).

II. LITERATURE SURVEY

A. Color Balance and Fusion for Underwater Image Enhancement

Introduces a cutting-edge method for removing haze from underwater photographs based on a single picture taken with a regular camera [1]. This method relies on the fusion of multiple inputs, but it derives the two inputs to merge by sharpening a white-balanced version of a single native input image and adjusting the contrast. The goal of the white balancing stage is to eliminate the colour cast brought on by underwater light scattering, resulting in a more realistic appearance of the underwater photos. Blending without artefacts is produced via the fusion method multi-scale implementation.

Without explicitly inverting the optical model, the image improvement approach uses a two-step strategy that combines white balancing and picture fusion to enhance underwater photographs. In our method, image fusion is thought to improve the edges and features of the picture and to offset the loss of contrast produced by backscattering. White balancing tries to compensate for the colour cast created by the selective absorption of colours with depth.

B. Underwater Image Processing System for Image Enhancement and Restoration

[2] Proposes a multi-step system, especially for processing an underwater image taken in succession. The system is divided into two parts. Firstly, the process of contrast enhancement and color correction is pre-processing, after which it obtains a haze-free and color-corrected image. In this process, in order to improve clarity, the paper adopted a dark channel prior to getting information on the dark channel and enhancing image contrast. After that, developed a gray world algorithm to make it more eligible for underwater image processing. As the traditional gray world algorithm neglects the special underwater environment where the colors of light principally consist of blue and green, several coefficients are set to adjust for brightening according to the underwater visibility distance of the red, green, and blue light. After getting the enhanced image, the next part is called image restoration.

Firstly, by comparing and subtracting two photos taken in succession, the moving irrelated objects are recognized and removed. Secondly, narrowed the blank according to the direction of its size and applied an improved TV model to fill the blank. When resizing the image after the inpainting step, the BP network is upgraded, and then the trained BP network is applied to realize super-resolution restoration of some details in the image.

C. Underwater Image Enhancement by Dehazing with Minimum Information Loss and Histogram Distribution Prior

This paper shows how to improve a single underwater image methodically so that two copies of the updated output are produced [3]. For presentation, a variation that resembles nature and includes colours that are near to the original are appropriate. For additional investigation, the other version with more contrast and brightness might be employed. This work proposes the first successful underwater image dehazing method based on the least information loss idea and the underwater imaging's optical properties. By restoring the colour and visibility of underwater images, the recommended underwater image dehazing method might enhance them. Then, a simple yet effective histogram distribution prior is recommended to enhance the contrast and brightness of the acquired haze-free underwater images. In this manner, two improved images are created, each of which can be used to a different situation. To evaluate the effectiveness of the suggested method, simulation experiments, qualitative and quantitative comparisons, colour accuracy and application tests are carried out, respectively. According to experimental findings, our haze-free output version has reasonably true colour, a natural appearance, and good visibility. Additionally, this contrast-enhanced output version's improved contrast and brightness are suited for revealing more details and retrieving more useful information.

D. Underwater Image Enhancement With a Deep Residual Framework

Powerful supervised learning methods have achieved remarkable results in many fields of computer vision, but they have rarely been used in underwater image enhancement. Because the blurred underwater images have no corresponding clear images as the ground truth. To address this lack of training data, CycleGAN is employed to generate training data [8]. It can learn the mapping from one data distribution to another data distribution without paired training data. CycleGAN consists of discriminators X and Y and generators F and G. The discriminator X learn the features of the in-air images to judge whether the output result is the in-air image; the discriminator Y learns the features of the underwater images to judge whether the output result is the underwater image. Generator G learns the mapping from the in-air images to the underwater images; generator F learns the mapping from the underwater images to the in-air images to complete the mutual conversion between the in-air images and the underwater images.

After supplying training data for a powerful supervised learning model with CycleGAN, the very-deep super-resolution (VDSR) model was introduced to the underwater image enhancement task. 20 convolution layers make up the VDSR model. A stride of 1 and a zero-padding of 1 pixel are used for each convolution layer, which uses 3*3 size filters. These parameter settings make sure that the input image and output image have the same resolution. Each convolution layer comprises 64 channels, with the first and last layers being the exceptions. The first layer takes three-channel picture data as input, creates 64-channel feature maps, and sends them to the network's core. The reconstruction layer comes last. It takes 64-channel feature maps as input and produces 3-channel residual images. To create the final restored photos, the residual images are combined with the input images.

III. METHODS

A. Generative Adversarial Networks(GAN)

Generative Adversarial Networks (GANs) are a class of neural networks that are used for unsupervised learning [8]. In GANs, there is a generator and a discriminator. The Generator generates fake samples of data and tries to fool the Discriminator. The Discriminator, on the other hand, tries to distinguish between the real and fake samples [4]. Both the networks run in competition with each other in the training phase. The steps are repeated multiple times and as a result, the Generator and Discriminator get better and better in their respective jobs after each repetition.

B. U-Net Generator

U-Net is an encoder-decoder network (e1-e5, d1-d5) with connections between the mirrored layers, i.e., between (e1, d5), (e2, d4), (e3, d2), and (e4, d4). It consists of a contracting path and an expansive path. The contracting path uses the CNN architecture. This path has two 3x3 convolutions, which are followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. The number of feature channels doubles in each downsampling step. The expansive path consists of feature map upsampling and is followed by a 2x2 convolution that reduces the number of feature channels in half. This is followed by skip connections between the mirrored layers and two 3x3 convolutions, each followed by a ReLU [6]. The final layer consists of a 1x1 convolution, to map each feature vector to the desired number of classes. In total the network has 23 convolutional layers. Specifically, the outputs of each encoder are concatenated to the respective mirrored decoders. This idea of skip connections in the generator network is shown to be very effective for image-to-image translation and image quality enhancement problems.

The GAN-Based Underwater Image Enhancement system will be employing a much simpler model with fewer parameters in order to achieve fast inference. The input to the network is set to $256 \times 256 \times 3$ and the encoder (e1-e5) learns only 256 feature maps of size 8×8. The decoder (d1-d5) utilizes these feature maps and inputs from the skip connections to learn to generate a $256 \times 256 \times 3$ (enhanced) images as output. The network is fully-convolutional as no fully-connected layers are used. Additionally, 2D convolutions with 4×4 filters are applied at each layer, which is then followed by a Leaky-ReLU non-linearity and Batch Normalization (BN).

C. Patch GAN Discriminator

A type of discriminator called PatchGAN exclusively penalises structure at the scale of local picture patches in generative adversarial networks [9]. An image's NxN patches are classified as real or fraudulent by the PatchGAN discriminator. The discriminator's final output is produced by convolutionally running it across the image and averaging all replies. With the assumption of independence between pixels separated by more than a patch diameter, such a discriminator effectively models the image as a Markov random field. It could be interpreted as a certain texture or style loss.

The PatchGAN architecture solely discriminates based on patch-level information and assumes the independence of pixels beyond the patch size. To efficiently capture high-frequency elements such local texture and style, this supposition is crucial. Additionally, this arrangement is computationally more efficient because it uses fewer parameters when compared to global image-level discrimination. The image is split into 70x70 patches in the discriminator, and four convolutional layers are utilised to transform a 256 x 256 x 3 input (actual and created image) into a 16 x 16 x 1 output that represents the discriminator's averaged validity answers.

D. System Architecture

The GAN-Based Underwater Image Enhancement system uses conditional GAN to take a distorted underwater image and enhance it to produce a high-quality, clear image in real-time. Figure 1 shows the architecture of the GAN Based Underwater Image Enhancement System.

Initially, the EUVP dataset is collected and given to the system for fine-tuning the generator and discriminator in the pre-trained conditional GAN model. After fine-tuning the model, the user provides a distorted underwater image as input to the generator. The generator, which uses U-Net architecture, generates an image based on the input image. The output generated by the generator is then given as input to the PatchGAN discriminator for classification as real/fake. The PatchGAN discriminator classifies every patch as real or fake and takes an average of all the results to obtain the discriminator output. If the discriminator classifies the image as fake, then the output is discarded and a generator loss is sent back to the generator, which then updates the weights by backpropagation and then generates another image If the discriminator classifies the image as real, then it qualifies as an enhanced image. Once the enhanced image is obtained, using the object detection module, the objects in the enhanced image are detected and identified. The enhanced image and the objects identified are displayed to the user.

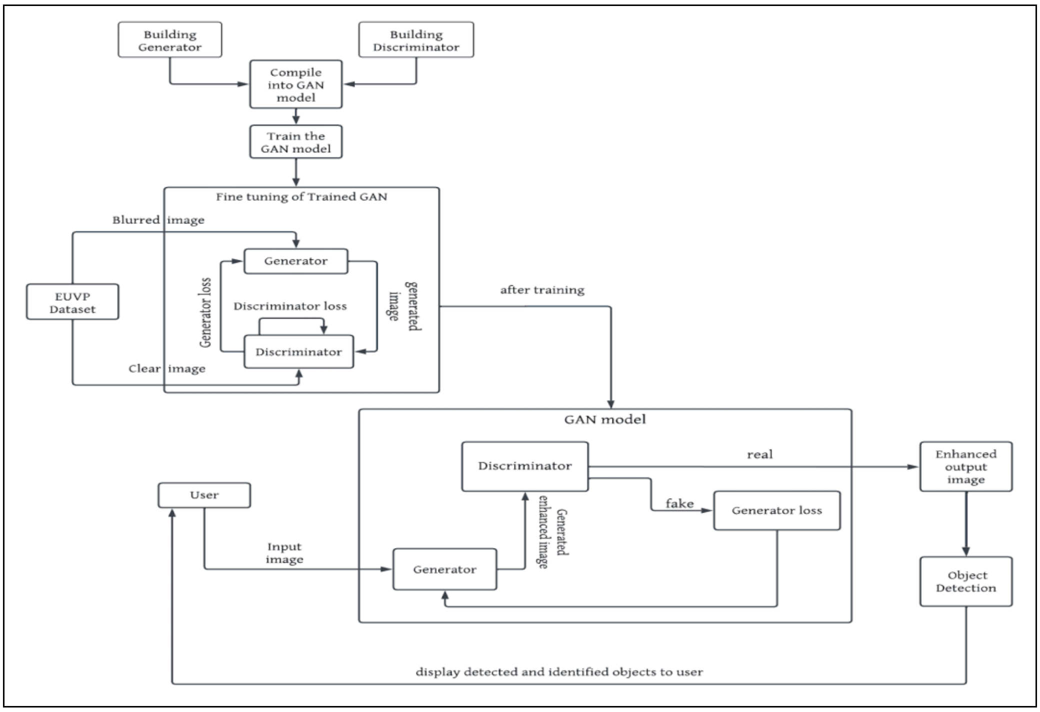

Fig. 1: Architecture of the system

IV. DATASET

EUVP dataset is a large collection of paired and unpaired underwater images of both good and poor perceptual quality. The dataset is collected from Kaggle and is imported to IDE. The dataset was created using multiple cameras, including Go Pros, Aqua AUV's uEye cameras, low-light USB cameras, and Trident ROV's HD camera, during oceanic explorations and human-robot cooperative experiments in different locations under various visibility conditions. The unpaired data in the EUVP dataset is prepared by visually inspecting the images by six human participants. The participants examine several image properties, such as color, contrast, and sharpness, and determine whether the foreground/objects in the scene are identifiable, i.e., visually interpretable [5]. This process separates the images into good and poor-quality categories. Overall, the EUVP dataset contains over 12K paired and 8K unpaired instances, covering a wide range of natural variability in scenes, waterbody types, lighting conditions, etc. The dataset's focus is to facilitate perceptual image enhancement to boost robotic scene understanding [7], rather than to model the underwater optical degradation process for image restoration, which requires scene depth and waterbody properties.

V. RESULTS AND DISCUSSION

Underwater images suffer from a variety of degradations, such as color distortion, blurring, and low contrast. These problems make it difficult to obtain useful information from underwater images, and thus, image enhancement techniques are required. Generative Adversarial Networks (GANs) have been successfully applied to image enhancement tasks. In this project, we implemented a GAN-based model, for enhancing underwater images.

The results of our evaluation indicate that our model performs well on both qualitative and quantitative metrics [11]. In terms of qualitative evaluation, we found that the model was able to recover true color and sharpness in the enhanced images and rectify the greenish hue in underwater images. Moreover, our GAN model shows much better performance in the qualitative evaluation and the proposed method's average accuracy is 83%.

In quantitative evaluation, we compared the GAN-enhanced images with their respective ground truths using three standard metrics, PSNR, UIQM, and SSIM. The results as shown in Table 1 indicate that it performs best on both metrics.

Table 1. Performance Evaluation Results

Evaluation Metrics	**Mean**	**Standard Deviation**
PSNR	46.2208	3.0687
SSIM	0.8021	0.0608
Input image UIQM	2.5754	0.5486
Enhanced image UIQM	3.9857	0.3014

A. PSNR (Peak Signal-to-Noise Ratio)

A higher PSNR value indicates that generated images are closer to ground truth images in terms of visual quality. The obtained PSNR mean/value of 46.22 with a standard deviation of 3.06 in Table 1 means that the generated images have an average PSNR of 46.22 dB (decibels) compared to their corresponding ground truth images, with a variation of 3.06 dB around the mean. The mean PSNR value of 46.22 indicates that generated images are of reasonably good quality

B. SSIM (Structural Similarity Index Measure)

It is a performance evaluation metric used to assess the similarity between two images. SSIM value ranges from -1 to 1, where a value of 1 indicates perfect similarity between the two images, while a value of -1 indicates complete dissimilarity. A value of 0 indicates that the two images are uncorrelated. The obtained SSIM value of 0.802 and SSIM standard deviation of 0.06 in Table 4.1 indicates that the images being compared have a high degree of structural similarity, with a small variation across the samples being evaluated. SSIM value of 0.802 is relatively high, suggesting that the images are similar, and a low standard deviation of 0.06 indicates that the similarity between the images is consistent across the dataset.

C. UIQM (Universal Image Quality Measure)

It is a performance evaluation metric that quantifies the quality of an image. The higher the UIQM value, the better the image quality. Input image UIQM calculates the UIQM value for the input distorted underwater images. The mean UIQM value of the input images is 2.57 with a standard deviation of 0.55. This indicates that input images have lower quality than the enhanced output images, which is expected since the purpose of enhancing the images is to improve their quality. Enhanced image UIQM calculates the UIQM value for the enhanced output images. The mean UIQM value of enhanced output images are 3.99 with a standard deviation of 0.40. This indicates that the enhanced images have higher quality than the input images.

VI. CONCLUSION

The marine sector has been more interested in exploring the deep sea and ocean in recent years. Diverse, uncommon attractions, including plants and other species, can be found in the underwater environment. Deep-sea divers, marine scientists, researchers, and others can benefit greatly from underwater images. However, due to light attenuation and dispersion, it is practically hard to capture crisp, high-quality photos underwater, making the photographs unusable as they become distorted.

In order to improve underwater images in real-time, a GAN-Based Underwater Image Enhancement System is presented. For the purpose of enhancing underwater images in real time, the system suggests an effective conditional GAN-based model. The suggested model assesses image quality based on its global colour, content, local texture, and style information to construct a perceptual loss function. For supervised training, a sizable dataset called EUVP Dataset that includes both paired and unpaired collections of underwater photos is used.

The proposed network's accuracy is found to be higher than that of the existing approaches, demonstrating the resilience of the suggested strategy. When compared to raw input photographs without any processing, the results demonstrate that the recovered images from the suggested method produce good and effective outcomes. As a result, the suggested technique can be seen as a crucial step in recovering underwater photos before using them in underwater applications, with an effective result.

Future research on underwater robotics is possible. Underwater robotics is typically expensive since it requires the use of expensive cameras. In-depth investigation and the creation of a more reliable algorithm can be used to make robots camera-independent. Robots can use this technique to make underwater exploration simpler and more affordable. The improvement of underwater photographs makes a crucial contribution to underwater robots and maritime engineering [12].

REFERENCES

[1.] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and P. Bekaert, "Color Balance and Fusion for Underwater Image Enhancement," in *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 379-393, Jan. 2019, doi: 10.1109/TIP.2017.2759252.

[2.] C. Cai, Y. Zhang and T. Liu, "Underwater Image Processing System for Image Enhancement and Restoration," 2019 *IEEE 11th International Conference on Communication Software and Networks (ICCSN)*, 2019, pp. 381-387, doi: 10.1109/ICCSN.2019.8905310.

[3.] C. -Y. Li, J. -C. Guo, R. -M. Cong, Y. -W. Pang and B. Wang, "Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior," in *IEEE Transactions on Image Processing*, vol. 25, no. 12, pp. 5664-5677, Dec. 2021, doi: 10.1109/TIP.2016.2612882.

[4.] Y.-S. Chen, Y.-C. Wang, M.-H. Kao, and Y.-Y. Chuang, "Deep Photo Enhancer: Unpaired Learning for Image Enhancement from Photographs with GANs", in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6306– 6314. IEEE, 2018.

[5.] Z. Cheng, Q. Yang, and B. Sheng, "Deep Colorization",in *IEEE International Conference on Computer Vision (ICCV)*, pages 415–423, 2015.

[6.] Y. Cho, J. Jeong, and A. Kim, "Model-assisted Multiband Fusion for Single Image Enhancement and Applications to Robot Vision", *IEEE Robotics and Automation Letters (RAL)*, 3(4):2822–2829, 2018.

[7.] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. FeiFei, "ImageNet: A Large-scale Hierarchical Image Database", in *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 248–255. IEEE, 2009

[8.] P. Liu, G. Wang, H. Qi, C. Zhang, H. Zheng, and Z. Yu, "Underwater Image Enhancement With a Deep Residual Framework," *in IEEE Access*, vol. 7, pp. 94614-94629, 2019, doi: 10.1109/ACCESS.2019.2928976.

[9.] Saleem, S. Paheding, N. Rawashdeh, A. Awad and N. Kaur, "A Non Reference Evaluation of Underwater Image Enhancement Methods Using a New Underwater Image Dataset," in IEEE Access, vol. 11, pp. 10412-10428, 2023, doi: 10.1109/ACCESS.2023.3240648.

[10.] Y. Wang, H. Fan, S. Liu, J. Liu, S. Li and Y. Tang, "BA-GAN: Block Attention GAN model for Underwater Image Enhancement," 2021 IEEE International Conference on Unmanned Systems

(ICUS), Beijing, China, 2021, pp. 809-813, doi: 10.1109/ICUS52573.2021.9641455.

[11.] Vijay. A. R, Rukmani. D. S, Preetham. S, Pratheep. K,, Punuru. Bhanu. P. Reddy and Ram. Aravind. U, "QUALITATIVE ANALYSIS OF UNDERWATER IMAGE ENHANCEMENT," 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2021, pp. 1225-1230, doi: 10.1109/ICOEI51242.2021.9453082.

[12.] J. Zhang, L. Zhu, L. Xu and Q. Xie, "Research on the Correlation between Image Enhancement and Underwater Object Detection," 2020 Chinese Automation Congress (CAC), Shanghai, China, 2020, pp. 5928-5933, doi: 10.1109/CAC51589.2020.9326936.