Underwater Image Enhancement using GAN

Sabnam Jebin C.[1]; Rahamathulla K[2]
Department of Computer Science & Engineering
Government Engineering College Thrissur, India

Abstract:- The process of enhancing the distorted underwater images to clear image is known as Underwater image enhance- ment. Distorted images are the raw underwater images that taken from the deep portion of ocean, river etc by using different cameras. In general underwater images are mainly used in underwater robotics, ocean pasture and environmental monitor- ing, ocean exploration etc. The underwater image enhancement process is done by using underwater image dataset which includes the distorted images (raw underwater images) and the corre- sponding enhanced underwater images. Currently used image enhancement methods cannot provide sufficient satisfaction to the underwater image enhancement. So proposed a new method by using Generative Adversarial Network (GAN), which tries to produce more images from the dataset.

Keywords:- *Generative Adversarial Network(GAN), Under-Water Image Enhancement.*

I. INTRODUCTION

Underwater imaging is an important task for various appli- cations, including marine biology, oceanography, and under- water exploration. However, the quality of underwater images is often degraded due to the effects of water on light, such as scattering, absorption, and attenuation. This degradation can result in poor contrast, color distortion, and blurring, which can make it difficult to extract useful information from the images. In this paper [1] traditional image enhancement methods have limited effectiveness in improving the quality of underwater images. In Figure 1.1 shows that the general overview of the underwater image enhancement. Here we can see that the left image set is the distorted (poor quality) underwater images and the right image set is the corresponding enhanced (good quality) underwater images.

In recent years, deep learning methods, particularly Genera- tive Adversarial Networks (GANs), have shown great promise in image enhancement tasks. GANs are a type of neural network that can learn to generate realistic images by training a generator network to produce images that are similar to a dataset of real images, and a discriminator network to differentiate between real and generated images. GANs have been successfully applied to tasks such as super-resolution, and style transfer.

In this project, we investigate the use of GANs for enhanc- ing underwater images. Our goal is to develop a GAN-based method that can produce high-quality enhanced images from a dataset of degraded underwater images. The proposed method involves training a GAN on a dataset of degraded and high- quality underwater images to learn the statistical properties of the data. In this paper [2] Evaluate the effectiveness of our method by comparing the quality of the generated images with the original degraded images and with other state-of- the-art methods. We will also analyze the performance of our method in terms of its ability to preserve important features and details in the images.

The contributions of this project are two fold. First, we pro- pose a GAN-based method for enhancing underwater images that can produce high-quality images with improved contrast, color, and sharpness. Second, we present an evaluation of the effectiveness of our method in comparison to other state-of- the-art methods. The results of this project have the potential to advance the field of underwater imaging and provide a valuable tool for underwater exploration and research.

Table I: Abbreviations and Acronyms

- GAN : Generative adversarial network
- UW-GAN : Underwater generative adversarial network
- UWC-Net : Underwater coarse-level generative network
- UWF-Net : Underwater fine-level network
- CNN : Convolutional neural network
- PAM : Parallel attention module
- ALM : Adaptive learning module
- EUVP : Enhancing underwater visual perception
- SSIM : Structural similiarity
- PSNR : Peak signal-to-noise ratio

II. LITERATURE SURVEY

A. Underwater Image Enhancement

The image enhancement is done to improve an image suitability for a given task, such as making it more pleasing to the eye. To put it another way, the main goal of image enhance- ment is to modify a given image so that the finished product is better suited than the original image for a given application. In this paper [2] propose a conditional generative adversarial network-based technique for improving real time underwater images. Similarly, [3] present an improved multiscale dense generative adversarial network for the

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUN403

process of underwater image enhancement. Create an objective function that monitors the adversarial training on the dataset and examines the global content, colour, local texture, and style information to assess the perceived image quality. Additionally, propose a sizable dataset with paired and unpaired underwater image collections of (”bad” and ”excellent” quality) collected with seven distinct cameras under a range of visibility conditions during mar- itime trips and experiments using human-robot collaboration [EUVP]. Perform Both qualitative and quantitative analyses show that The suggested model can improve underwater image quality through paired and unpaired training.

The task of predicting underwater single-image depth is challenging because there aren’t any large-scale datasets of underwater depth image data and the difficulties aren’t well-posed. An precise depth estimate is needed for a single underwater image in applications like marine engineering and underwater robots. In this paper [4] from an underwater single image, propose an end to-end underwater generative adversarial network (UW-GAN) for depth estimation . Firstly, by using using the underwater coarse-level generative net- work (UWC-Net) a coarse-level depth map is estimated. The estimated coarse-level depth map and the input picture are concatenated as input into the underwater fine-level network (UWF- Net), which computes a fine-level depth map. Also, suggest a method for creating synthetic underwater images for huge databases. For the performance analysis of the proposed network, both real-world and artificial underwater datasets are used.

The major causes of the quality decline in photographs taken under cloudy settings are 1) various meteorological factors and 2) the attenuation of reflected light. These elements significantly alter the hue and visibility of the acquired photos. To tackle these problems, [5] suggest an end-to-end trainable image de-hazing network [LIGHT-Net]. Haze reduction and colour constancy modules make up the proposed LIGHT- Net. The colour constancy module among them removes the colour tint that the weather supplied to the hazy image. The suggested haze reduction module, which uses an inception- residual block as its building block, aims to lessen the impact of haze as well as to enhance visibility in the foggy image. In contrast to conventional feature concatenation, the haze reduction module proposed in this research uses dense feature sharing to efficiently distribute the features discovered in the network’s first layers.

The underwater image typically has low contrast, colour distortion, and fuzzy features because of the light’s atten- uation and dispersion in the water. By taking into account underwater imaging specifics, an unique two-stage underwater image convolutional neural network (CNN) based on structural decomposition (UWCNN-SD) [6] for underwater picture en- hancement is suggested.

Due to light’s dispersion and absorption as it moves through water, underwater photographs have colour casts and inade- quate illumination. Tasks requiring vision of underwater, such as detection and recognition, may be hampered by these issues. To address these degradation difficulties, [7] offer LANet, an adaptive learning attention network for underwater image enhancement based on supervised learning. A multiscale fu- sion module is first suggested to merge various spatial data. Then, create a brandnew parallel attention module (PAM) that combines pixel and channel attention with a focus on lighted elements and more important colour information. The shallow information can then be retained by an adaptive learning module (ALM), which can then learn crucial feature information.

The effects of absorption and scattering frequently result in the degradation of underwater image quality. When utilised for analysis and display, degraded underwater photos have some restrictions. For instance, low contrast and colour cast in underwater photos reduce accuracy of marine biological recog- nition rate and underwater item detection rate. A contrast en- hancement algorithm an underwater image dehazing algorithm [8] are part of a systematic underwater image enhancement technique that is proposed to get around those restrictions. The vision, colour, and a natural undersea image appearance are restored using an effective underwater image dehazing approach. The lowest information loss principle serves as its foundation. An effective contrast enhancement technique is suggested that enhances the contrast and brightness of underwater images based on a histogram distribution prior type. Two improved output versions can be produced using the suggested procedure. One variant, ideal for display, has fairly authentic colours and a natural appearance. The other version, which has greater brightness and contrast, can be utilised to glean more important data and reveal more specifics.

Underwater image improvement has received a lot of in- terest recently in both underwater vision and image pro- cessing. Enhancement of underwater images is a tough task because, the complex underwater environment and lighting circumstances. Underwater images are typically wavelength-dependent absorption causes degradation and scattering [9], including backward scattering and forward scattering.

B. Underwater Object Tracking

Applications including deep ocean exploration, underwater robot navigation, marine life monitoring [10], and homeland and maritime security all rely heavily on underwater object tracking. These applications call for effective and precise vision-based underwater marine analytics, including methods for image augmentation, picture quality evaluation, and target tracking. Understanding marine image/video analytics faces huge difficulties due to the excessive noise and poor light con- ditions. Because of this, computer vision tasks like detection, recognition, and tracking are substantially more difficult in underwater situations than in open-air environments.

In this paper [5] numerous sophisticated tracking methods have been presented as a result of the availability of big annotated datasets. Depending on how tracking features are gathered and applied, the methodologies used in the majority of these trackers can be divided into different groups. KCF, HCF, DCF, CFNet, STRCF, BACF, and fDSST are examples of trackers that fall within the category of kernelized correlation filters. Modern trackers including MDNet, SiamFC, CCOT, and ECO have primarily used the CNN feature-based tech- nique. The foundation of older trackers like Struck is local and global feature extraction.

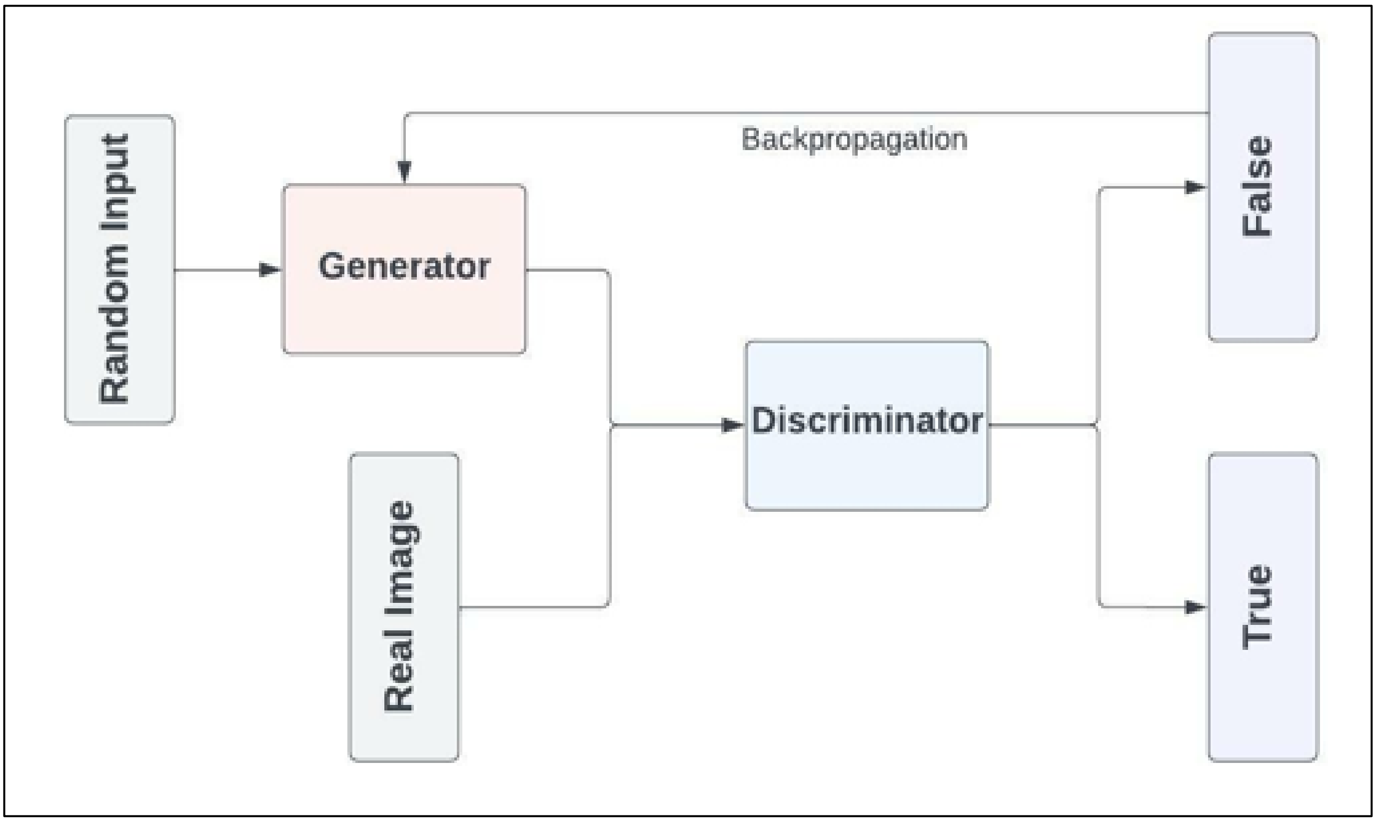

Fig 1: Block Diagram of GAN

III. PROPOSED METHOD

A. Generative Adversarial Networks

The field of machine learning has taken positive interest in GANs [11] due to both their theoretical appeal and their capacity the target probability distribution should be learned. To learn a nonlinear mapping between the distorted image and the non distorted image, suggest a GAN. The suggested network uses an end-to-end, data-driven training mechanism to produce improved output.

Here provide an overview of GANs and their components, including generator and discriminator networks. And also discuss the advantages of GANs over other deep learning methods, such as their ability to generate realistic and diverse samples. Finally, should review some popular GAN-based applications, such as image synthesis, image-to-image trans- lation, and style transfer. The generative adversarial network's general design is depicted in Figure 1. It is made up of a generator and a discriminator.

Generative Adversarial Networks (GANs) are a type of deep learning model that can learn to generate new data samples that are similar to a given dataset. GANs consist of two networks such as a generator network and a discriminator network. The generator network takes a random noise vector as input and generates a new sample that is intended to be similar to the real data. The discriminator network takes the generated sample and the real data as input and tries to distinguish between them. The two networks are trained simultaneously, with the goal of the generator network learning to generate samples that are indistinguishable from the real data, while the discriminator network learns to correctly classify real and generated samples. Figure 1 shows that the general working of GAN.

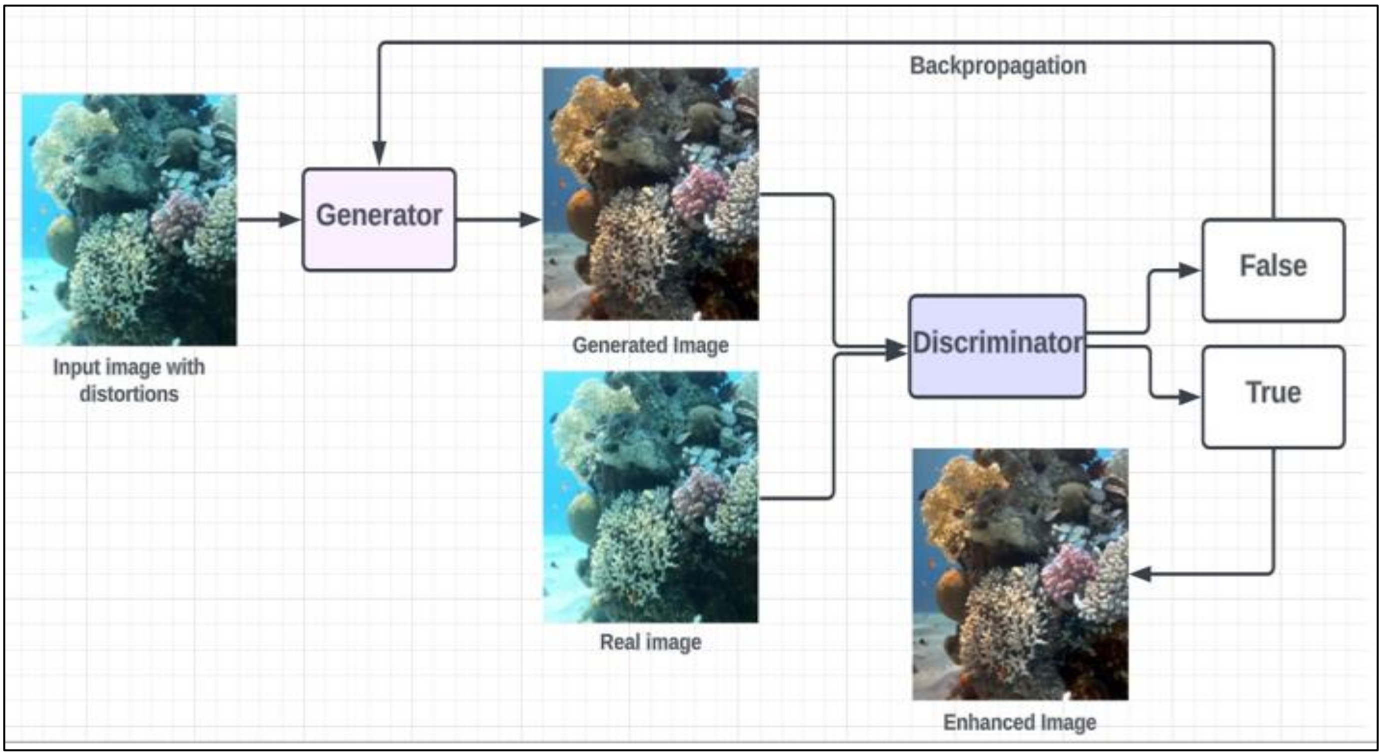

Fig 2: General Overview of the Proposed Framework

However, GANs also have their limitations, such as the need for large amounts of high-quality training data and the diffi- culty of training stable and robust models. Nevertheless, GANs are a powerful and promising deep learning technique that has shown impressive results in a wide range of applications.

B. Framework Overview

The new underwater images (duplicate) are generated us- ing a Generative Adversarial Network (GAN). The proposed framework includes a generator and discriminator, as seen in Figure 2. This architecture shows that the input taken by the generator from the dataset and generate it's duplicates of images. After that the discriminator takes the output of generator as it's input for the discrimination of whether the image is real or fake. The output of discriminator indicates that true or false (1 or 0) corresponding to real or fake respectively.

C. Generator Network

There are many modules for feature extraction that have been created. In the paper [12] includes the popular architec- ture of inception looks for the structure of optimal local sparse inside a structure of network. Even so, at the conclusion of the block, these various scale characteristics concatenate in a straightforward manner, contributing in part to the underuti- lization of feature maps. Following the U-Net principles [13], our generator network has been modified from UGAN. The encoder and decoder networks make up this system.

Figure 3 shows the generator network that consist convolu- tional layer, batch normalization layer, Leaky ReLU, decon- volutional layer and tanh function. This convolutional layer is used for the feature extractionby using kernal or filter. Batch normalization is used for to standardize each input and output of each layer. Leaky ReLU is an activation function. Tanh function is an activation function that produce [0, 1] result.

➢ *Here is a General Overview of How a Generator in a GAN Works:*

- **Input:** The generator receives random noise vectors as input, which are typically drawn from a simple probabil- ity distribution like a Gaussian distribution. These noise vectors are low-dimensional representations that capture the variability of the data.

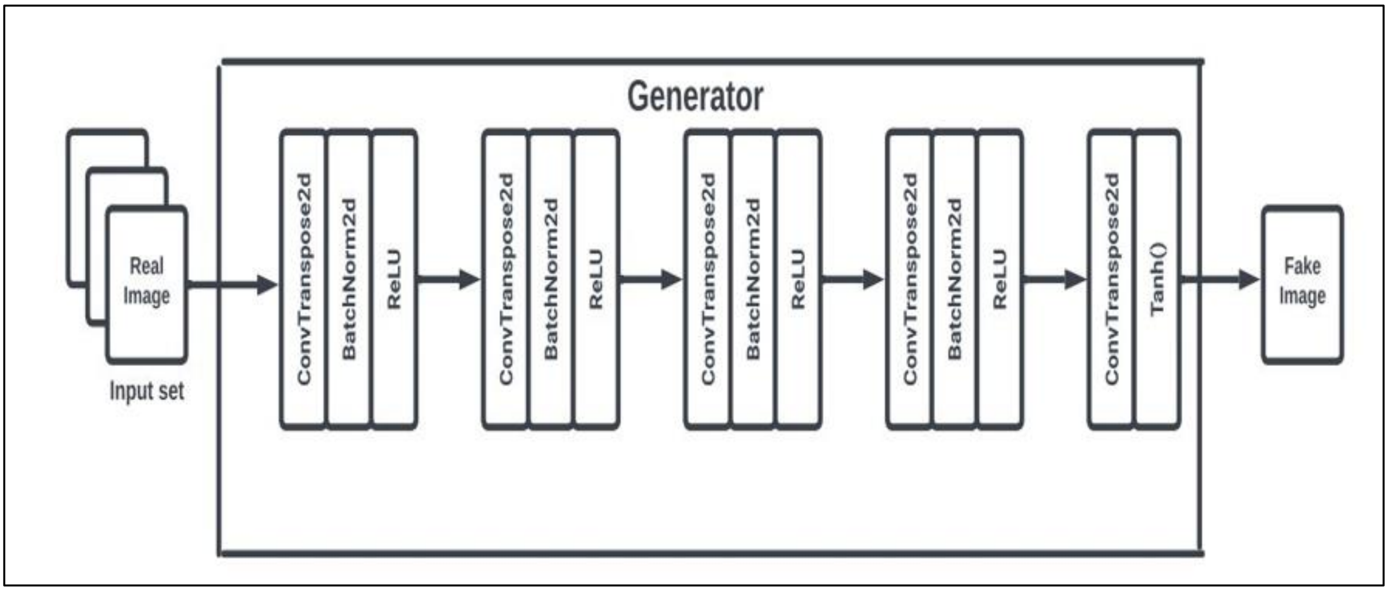

Fig. 3: Architectures of Generator.

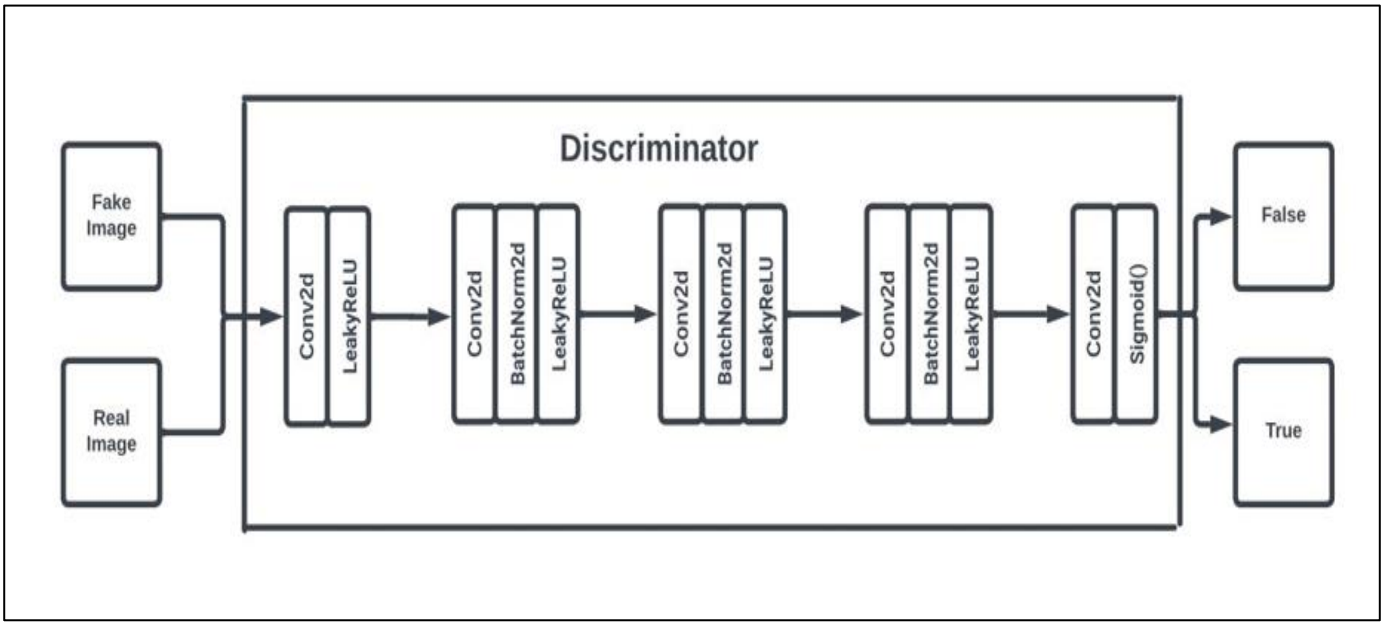

Fig 4: Architectures of Discriminator

- **Transformation:** The generator takes the input noise vectors and applies a series of mathematical transformations through its neural network layers. These layers gradually increase the complexity and dimensionality of the input to produce a higher-dimensional output.
- **Output:** The final layer of the generator typically employs an activation function, such as a sigmoid or a hyperbolic tangent, to squash the output values into the desired range. The generated output is meant to resemble the real data samples from the training set.
- **Training:** During the training phase of the GAN, the generator's objective is to generate synthetic samples that can deceive the discriminator, another component of the GAN architecture (discussed further below). The generator's weights are updated based on the feedback it receives from the discriminator.

D. Discriminator Network

Figure 4 shows the network of discriminator. The proposed discriminator network includes four layers. It consist of convolutional layer, batch normalization, Leaky ReLU and sigmoid function. Convolutional layer is used for the feature extraction by using the kernal or filter map. And the batch normalization is used for to standardize each input and output of each layer. The Leaky ReLU is an activation function. The sigmoid activation function is the last layer of discriminator which predict the output as true or false.

➢ *Here is a General Overview of How a Discriminator in a GAN Works:*

- **Input:** The discriminator takes as input either real data samples or generated data samples from the generator.

- **Network Layers:** The input data is processed through a series of layers in the discriminator network. These layers can include convolutional or fully connected layers, depending on the architecture chosen.
- **Feature Extraction:** The discriminator's layers gradually extract meaningful features from the input data, which allow it to differentiate between real and fake samples. These features capture the distinguishing characteristics of the data, such as texture, shape, or patterns.
- **Activation Function:** Typically, the layers in the discriminator are followed by an activation function, such as the sigmoid or softmax function. This activation function squashes the output values to a specific range (e.g., [0, 1] for sigmoid) and provides a probability score that represents the discriminator's confidence in classifying the input as real or fake.
- **Binary Classification:** The output of the discriminator is interpreted as a binary classification decision: either the input data is classified as real (assigned a label of 1) or fake (assigned a label of 0). The discriminator's objective is to accurately distinguish between real and fake samples.
- **Training:** During the training phase of the GAN, the discriminator is trained independently from the generator. It is presented with labeled real and fake samples and is optimized to correctly classify them. The weights of the discriminator are updated based on a loss function, such as binary cross-entropy, which measures the difference between the predicted labels and the true labels.

IV. DATASET

The EUVP (Enhancing Underwater Visual Perception) dataset [4] contains separate sets of paired and unpaired image samples of poor and good perceptual quality to facilitate supervised training of underwater image enhancement models. The EUVP dataset, developed by the Interactive Robotics and Vision Lab at the University of Minnesota, is designed to facilitate research in enhancing underwater visual perception. It aims to improve the visibility and quality of underwater images and videos through the development of advanced computer vision algorithms and image processing techniques.

V. COMPARISON STUDY WITH OTHER METHODS

Proposed framework is compared to three other methods for the underwater image enhancement with GAN. The compar- ison methods are [2] FUnIE-GAN, [2] FUnIE-GAN-UP, [14] UGAN.

We can evaluate our results using two different approaches such as qualitative comparison and quantitative comparison. Figure 6 shows that the qualitative and quantitative evalua- tion of an underwater image. Here SSIM and PSNR as the quantitative metrices used for the evaluation.

A. Qualitative Analysis

In the qualitative comparison, involves assessing the visual quality and realism of the generated samples. It focuses on subjective judgments made by human evaluators rather than relying on quantitative metrics. Figure 5 shows the qualitative comparison of our proposed method with other deep learning methods.

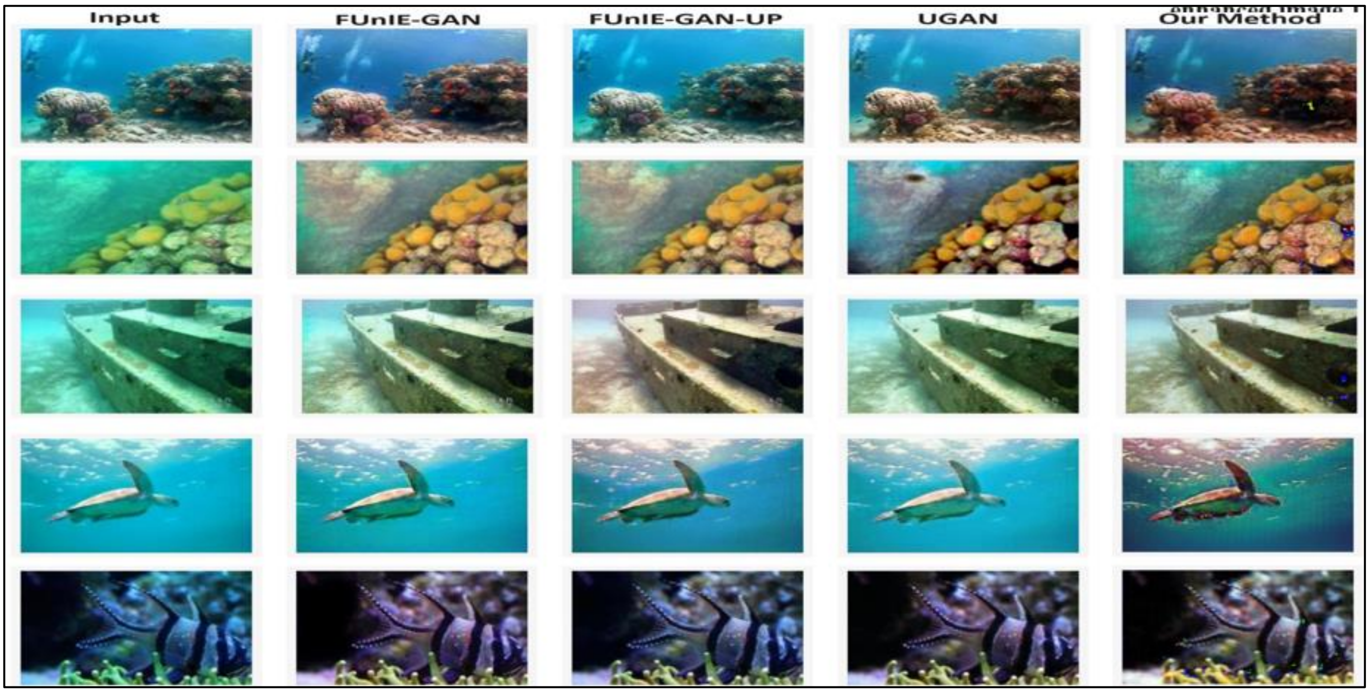

Fig 5: Qualitative Comparison with other Method

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUN403

Table 2: Quantitative Comparison with Euvp Dataset

Model	PSNR	SSIM
FUnIE-GAN	21.92	0.88
FUnIE-GAN-UP	21.36	0.81
UGAN	19.59	0.66
Ours	23.92	0.92

B. Quantitative Analysis

In the quantitative comparison, involves using objective metrics to measure various characteristics of the generated samples. These metrics provide numerical scores or measure-ments that can be used to compare different models or track progress during training. In this paper [2], two commonly used metrics in the quantitative evaluation are SSIM and PSNR. Table II shows the both PSNR and SSIM values obtained from different methods. Based on this analysis, it is obvious that our method does not demonstrates good performance compared to the other previously used methods, as evidenced by the PSNR and SSIM value.

- *SSIM:* SSIM is a metric that quantifies the similarity between two images, considering their structural information. It measures the perceptual similarity by evaluating three components: luminance, contrast, and structure. SSIM [15] produces a score between 0 and 1, where 1 indicates perfect similarity.

Fig 6: SSIM and PSNR Value of Proposed Method

$$SSIM(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (1)$$

Where, $\mu_x(\mu_y)$ denotes the mean, and $\sigma^2(\sigma^2)$ is the variance of $x(y)$; whereas σ_{xy} denotes the cross-correlation between x and y. Additionally, $c_1 = (255 \times 0.01)^2$ and $c_2 = (255 \times 0.03)^2$ are constants that ensure numeric stability.

SSIM takes into account both low-level pixel-wise compar- isons and higher-level structural comparisons. It considers how well the local structures, textures, and edges of the generated image match those of the ground truth (real) image. Higher SSIM scores indicate higher visual similarity.

- *PSNR:* PSNR [16] is a metric commonly used to mea- sure the quality of reconstructed or generated images by comparing them to a reference image. It measures the ratio of the peak power of the signal [17] (the maximum possible value) to the noise power (the difference between the generated and reference images).

$$PSNR(x, y) = 10 \log_{10} \frac{255^2}{\mathrm{MSE}(x, y)} \quad (2)$$

The PSNR approximates the reconstruction quality of a gen- erated image x compared to its ground truth y based on their Mean Squared Error (MSE). PSNR is expressed in decibels (dB) and provides a quantitative assessment of image fidelity. Higher PSNR values indicate lower perceptual differences between the generated and reference images, implying higher image quality.

VI. CONCLUSION

In this paper focused on exploring the application of Gen- erative Adversarial Networks (GANs) for underwater image enhancement. The objective was to improve the visual quality and clarity of underwater images, which often suffer from poor visibility, color distortion, and low contrast. Through the implementation and experimentation with GAN-based ap- proaches, significant advancements have been made in under- water image enhancement.

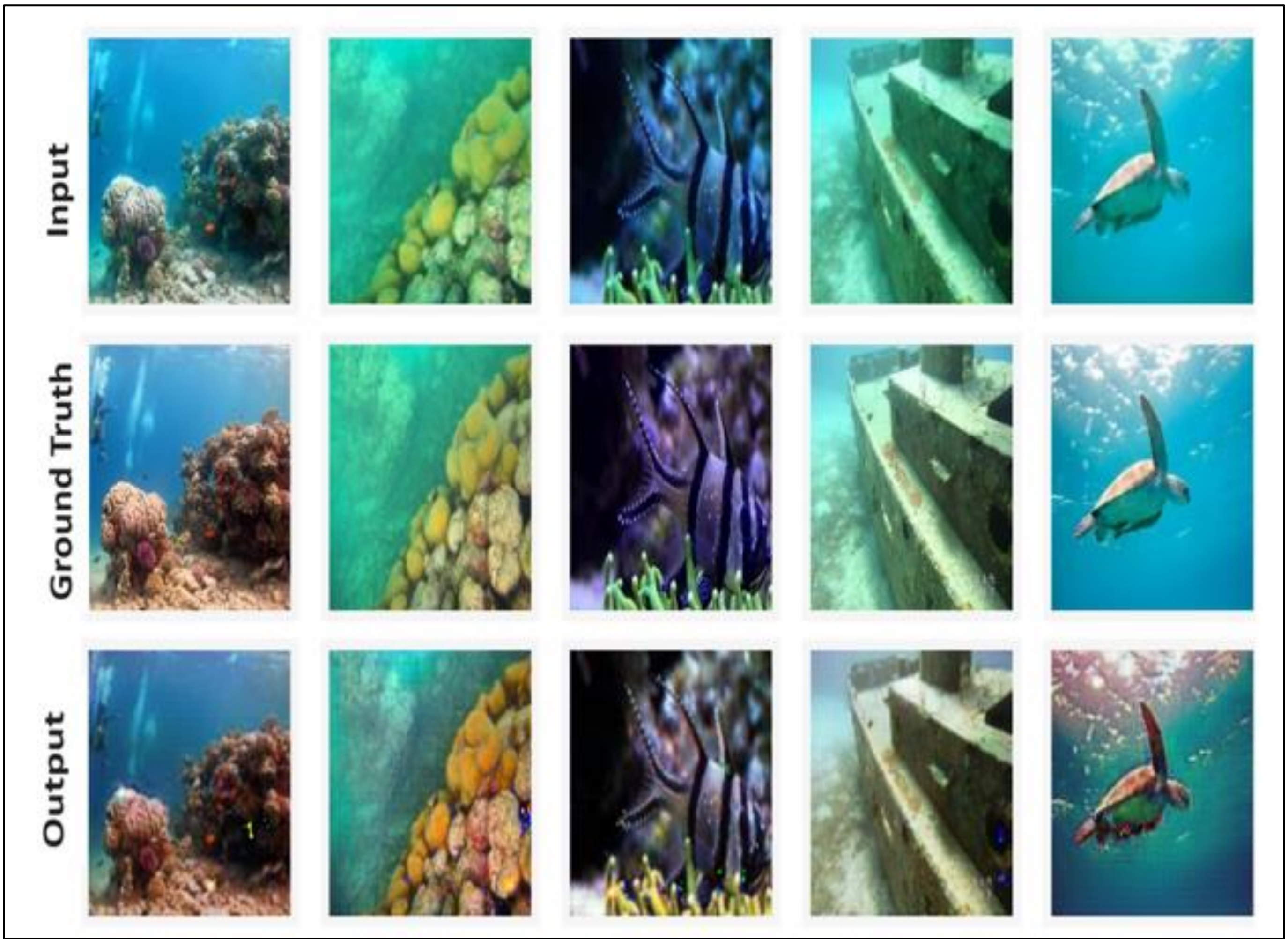

Fig 7: Output of our Proposed Method

The project began by conducting a comprehensive literature review to understand the existing techniques and challenges in underwater image enhancement. Various GAN architectures and image enhancement algorithms were studied to identify the most suitable approach for the task. The selected GAN model was trained using a large dataset of underwater images, aiming to learn the underlying structure and characteristics of the underwater environment.

The evaluation of the proposed GAN-based underwater image enhancement approach involved both qualitative and quantitative measures. Qualitative evaluation involved visual inspection and subjective assessment of the enhanced images by experts and users, while quantitative evaluation utilized established metrics such as Structural Similarity Index (SSIM), Peak Signal-to-Noise Ratio (PSNR). The results indicated a significant improvement in image quality and confirmed the superiority of the proposed GAN-based method compared to traditional image enhancement techniques.

REFERENCES

[1]. K. Panetta, L. Kezebou, V. Oludare, and S. Agaian, "Comprehensive underwater object tracking benchmark dataset and underwater image enhancement with gan," *IEEE Journal of Oceanic Engineering*, vol. 47, no. 1, pp. 59–75, 2021.

[2]. M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3227–3234, 2020.

[3]. Y. Guo, H. Li, and P. Zhuang, "Underwater image enhancement using a multiscale dense generative adversarial network," *IEEE Journal of Oceanic Engineering*, vol. 45, no. 3, pp. 862–870, 2019.

[4]. P. Hambarde, S. Murala, and A. Dhall, "Uw-gan: Single-image depth estimation and image enhancement for underwater images," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–12, 2021.

[5]. Dudhane, P. W. Patil, and S. Murala, "An end-to-end network for image de-hazing and beyond," *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2020.

[6]. S. Wu, T. Luo, G. Jiang, M. Yu, H. Xu, Z. Zhu, and Y. Song, "A two-stage underwater enhancement network based on structure decom- position and characteristics of underwater imaging," *IEEE Journal of Oceanic Engineering*, vol. 46, no. 4, pp. 1213–1227, 2021.
[7]. S. Liu, H. Fan, S. Lin, Q. Wang, N. Ding, and Y. Tang, "Adaptive learning attention network for underwater image enhancement," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5326–5333, 2022.
[8]. 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," *IEEE Transactions on Image Processing*, vol. 25, no. 12, pp. 5664–5677, 2016.
[9]. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, and D. Tao, "An underwater image enhancement benchmark dataset and beyond," *IEEE Transactions on Image Processing*, vol. 29, pp. 4376–4389, 2019.
[10]. M. Han, Z. Lyu, T. Qiu, and M. Xu, "A review on intelligence dehazing and color restoration for underwater images," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 5, pp. 1820–1832, 2018.
[11]. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.
[12]. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.
[13]. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234– 241, Springer, 2015.
[14]. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," in *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 7159–7165, IEEE, 2018.
[15]. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
[16]. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool, "Dslr-quality photos on mobile devices with deep convolutional net- works," in *Proceedings of the IEEE international conference on com- puter vision*, pp. 3277–3285, 2017.
[17]. 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 *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6306–6314, 2018.