



OPEN Underwater image dehazing using a hybrid GAN with bottleneck attention and improved Retinex-based optimization

Amandeep Kaur¹, Shalli Rani¹ & Mohammad Shabaz²✉

Autonomous underwater vehicles (AUVs) are essential for marine exploration, monitoring, and surveillance, especially in hazardous or inaccessible environments for human divers. Underwater imaging systems frequently face considerable difficulties in detecting and tracking objects due to image degradation resulting from light scattering, colour distortion, and haze. Conventional enhancement methods—like histogram equalisation and gamma correction—struggle with non-uniform illumination and frequently do not maintain critical structural details and perceptual quality. To address these limitations, this work proposes a novel framework for underwater image dehazing and enhancement that incorporates three essential components: a generative adversarial network (GAN), a bottleneck attention module (BAM), and an enhanced Retinex-based contrast enhancement technique. The GAN acquires the intricate correspondence between deteriorated and high-quality underwater images, facilitating the restoration of fine textures and the attenuation of noise. The BAM selectively amplifies spatial and channel-specific features, thereby augmenting the network's capacity to preserve natural hues and intricate details. The modified Retinex algorithm adaptively distinguishes between illumination and reflectance components, facilitating context-sensitive contrast enhancement across various lighting conditions. This integrated architecture facilitates collaborative learning among generative modelling, attention-driven feature refinement, and physics-based enhancement. The proposed method undergoes thorough evaluation on the underwater image enhancement benchmark (UIEB) dataset, which consists of 890 authentic underwater images. This study presents exceptional quantitative performance across various evaluation metrics: a UIQM score of 3.71 (indicating image quality), a PSNR of 28.4 dB (reflecting signal fidelity), an SSIM of 0.88 (representing structural similarity), and a perceptual LPIPS score of 0.082. The low LPIPS score underscores the perceptual realism of the enhanced images, correlating effectively with human visual preferences. These results distinctly surpass current classical and learning-based enhancement methods, demonstrating the efficacy and resilience of this approach for practical underwater vision applications.

Keywords Underwater image enhancement, Generative adversarial network (GAN), Bottleneck attention module (BAM), Dehazing, Retinex, Deep learning

In recent years, underwater imagery has become an important area of interest and gained much attention in various public and private organizations and demands clearer and enhanced images for many computer vision applications¹. In terms of maritime defence security, underwater optical imaging technology improves the sensitivity and accuracy of maritime defence early-warning systems by making it easier to find and track suspicious underwater targets^{2,3}. It also plays a crucial role in the identification and recovery of historical wrecks, including sunken ships and aeroplanes, protecting the country's marine interests and cultural legacy. Underwater optical imaging technology provides vital information to promote scientific study and marine ecological conservation by enabling non-invasive observation of the behaviour, population distribution, and environmental changes of marine species⁴.

¹Chitkara University Institute of Engineering & Technology, Chitkara University, Rajpura, Punjab, India. ² Marwadi University Research Center, Department of Computer Engineering, Faculty of Engineering and Technology, Marwadi University, Rajkot 360003, Gujarat, India. ✉email: bhatsab4@gmail.com

The unmanned vehicles operate remotely for underwater exploration, monitoring, and inspection of underwater infrastructure, marine life, mineral resources, and oceanographic conditions. There are remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs), drifting underwater camera system and many more, the trained person on the surface controls these vehicles remotely by capturing live images and videos. However, common underwater cameras cannot guarantee clear images and visibility decreases after 4 to 5 m under the water and usually capture degraded images with haze, blurriness, low color cast, and poor illumination effects^{5,6}. The main factor behind the degradation of image quality is the attenuation of light through absorption and scattering that affect underwater optical imagery. The atmospheric scattering and absorption of light is caused by floating particles and dissolved material in water that absorb light at different wavelengths and varying depths, which leads to loss of color and blurry images. Therefore, recovering underwater image is challenging task for researchers and a lot of work has been done in this area to enhance the visual quality of images using image dehazing and enhancement techniques to overcome these issues, most of the existing dehazing techniques still lack in restoring color and texture information of the image. However, many of these methods still fall short in achieving perceptual naturalness, especially in terms of realistic color rendering and contrast balance. A notable advancement in this direction is the work in⁷, which explored the inherent color disparities in underwater scenes and introduced new perceptual cues that enhance color restoration based on natural color consistency principles. Their method leverages inter-channel color relationships and distribution priors derived from above-water reference data to minimize unnatural color shifts and maintain a consistent appearance across different water types. This approach demonstrated significant improvements in subjective quality and aligns closely with the goal of enhancing underwater images toward visually faithful representations. Over the past decades, researchers have proposed various image enhancement techniques for improving the image quality of degraded image that are collected from underwater environment^{6,8}.

As shown in Fig. 1¹, the underwater image formation model. The camera received light from three resources: the direct transmission is the light received from object to camera, the forward scattering is that light deviates from its original path before reaching to the camera, in the back scattering light come across particles that scatter light before reaching the camera that leads to color distortion, turbidity and hazy images⁹. So, the performance of the dehazing algorithm and enhancement module depends on the attenuated light and water conditions of that particular scene. When light travels underwater, its wavelengths are attenuated to differing degrees; longer wavelengths, like red light, are most significantly attenuated. Underwater photos seem green or blue because shorter wavelengths, including blue and green light, attenuate more slowly. This suggests that the colour spectrum in underwater images are dominated by blue-green tones^{10,11}. Second, light scattering by particles and marine plankton in seawater causes chromatic distortion, diminished contrast, blurred details, and impaired visual acuity due to the dispersion and attenuation of luminous energy. As a result, getting good underwater photos is essential for marine exploration. Over past few years, work is also focused on improving the hardware setups in the water using fluorescence imaging, range-gated imaging, two-view cameras, stereo imaging, three-dimensional cameras etc⁵.

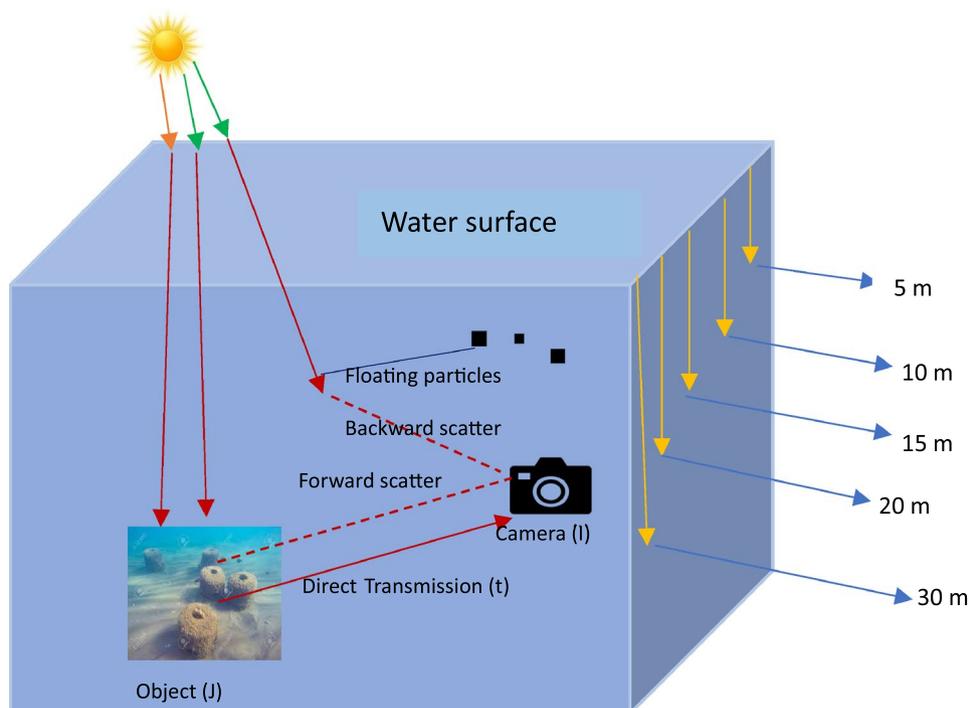


Fig. 1. Underwater image formation model.

In recent years, various underwater image enhancement and dehazing techniques have emerged and gained widespread attention to produce high-quality underwater images. It is one of the active topics of discussion for the improvement of underwater scenes analysis and deals with the issues related to the scattering of light and absorption, that lead to poor visibility, degraded colours, and low contrast issues, as illustrated in Fig. 2¹² where hazy sample images are taken from UIEB dataset.

Various traditional methods are used by researchers for underwater image enhancement and restoration. Image enhancement algorithms modify the value of pixels in an image to generate aesthetically pleasing visuals without using the underwater image formation model. These enhancement-based algorithms can only provide one type of enhancement effect for various underwater images and cannot easily restore high-quality images from degraded ones using the given information¹³. Though many of the modern techniques improve underwater image quality to varying degrees, they frequently nevertheless result in residual haze, blurred features, and unnatural colours. Ancuti et al.¹⁴ presented a fusion framework for edge preservation and noise reduction. Fu et al.¹⁵ proposed an improved Retinex-based method for a single underwater image. Li et al.¹⁶ proposed a method to restore visibility and color appearance.

Image restoration methods build a degradation model for underwater optical imaging and estimate model parameters to recover underwater images. These models adapt the Jaffe–McGlamery image degradation model for underwater images and overcome the degradation caused by light scattering and absorption in water. Some physical model-based image restoration algorithms may reduce underwater image degradation and recover images using a formation model. However, they tend to yield inaccurate results in more dense underwater conditions.

Deep learning enables underwater image enhancement by estimating parameters of physical model or directly generating improved underwater visuals using neural networks. Supervised learning-based underwater image-enhancing systems have improved due to extensive training data. To recover high-quality underwater images, due to the requirement for specialized equipment and skilled divers, collecting large, annotated underwater datasets is expensive and time-consuming, making learning-based approaches challenging. Training and implementing deep learning models are computationally intensive, and researchers have tried many methods to address these difficulties. Flipping, rotation, and color tilting enhance training datasets to give models more instances to learn. These advances might have limitations, particularly in generalizing training models to real-world underwater images and effectively simulating the difficult physics of underwater light scattering. Additional research is essential to integrate underwater physics into deep learning models, create unsupervised and weakly-supervised learning techniques, and enhance deep learning models for real-time processing. In response to the limitations of existing learning-based systems, recent research has explored reinforcement learning (RL) as a promising alternative for adaptive enhancement in underwater environments. The INSPIRATION framework by¹⁷ proposed a human visual perception-driven enhancement paradigm, in which reinforcement learning is guided by perceptual rewards calibrated to human quality judgments. This approach focuses not only on improving standard metrics like PSNR and SSIM but also on aligning the enhancement process with human visual preferences, such as edge clarity, color fidelity, and spatial realism. Unlike conventional deep learning models trained solely on paired data, INSPIRATION adapts dynamically to different underwater conditions, making it well-suited for diverse and unpredictable underwater scenes. So, learning-based image restoration algorithms increase image quality in various underwater environments by using enough training data to restore real-world underwater images.

To achieve more realistic and high-quality image enhancement results, the other category is based on generative adversarial networks (GANs), which mimic the process of creating realistic underwater images by building an adversarial learning mechanism between a generator and a discriminator. Yu et al.¹⁸ introduced an underwater image restoration approach utilizing a conditional generative adversarial network (GAN). For UIE, Guo et al.¹⁹ proposed multi-scale dense GAN to effectively enhance underwater images. The distribution of high-quality images is more accurately represented by this model, which combines adversarial loss and L1 loss. These learning-based methods use enough training data to build an improved image from a degraded image, effectively improving the quality of the underwater image.

Complementing these GAN-based efforts, a recent study²⁰ introduced a large foundation model-powered discriminative underwater image enhancement framework. This method employs pre-trained vision-language



Fig. 2. Hazy Sample images from UIEB dataset.

models, such as CLIP, as semantic priors to guide the enhancement of underwater images by learning contextual cues and object-level information. By incorporating transformer-based architectures and multi-scale attention modules, the framework discriminates between foreground and background regions, allowing selective enhancement of important image areas. This technique demonstrates superior generalization across multiple underwater datasets and addresses the challenge of scene diversity without requiring extensive domain-specific retraining. GAN models are especially well-suited to handling the intricacies of underwater image enhancement because they can quickly build strong generative models through adversarial training. We have chosen a GAN-based architecture as the main technological strategy for improving underwater images because of the unique qualities of underwater imagery as well as the limitations associated with training time and processing resources.

Our contribution

The main contribution of this study is the development of a unified underwater image enhancement framework that integrates a generative adversarial network (GAN)¹⁸ with a bottleneck attention module (BAM)²¹ and an improved Retinex-based¹⁵ contrast enhancement technique. This integrated approach effectively addresses common underwater image issues such as low contrast, color distortion, and haze. Specifically:

- i) The proposed hybrid method integrates a Generative Adversarial Network (GAN) with a Bottleneck Attention Module (BAM) and an improved Retinex-based contrast enhancement technique using the UIEB¹² dataset containing 850 underwater paired images.
- ii) The GAN restores fine details and reduces noise by learning complex mappings between degraded and high-quality underwater images. The BAM improves feature extraction by focusing on key spatial and channel information, enhancing detail and natural color preservation. The improved Retinex-based method adaptively enhances contrast and corrects illumination by separating images into reflectance and illumination components.
- iii) This method achieves state-of-the-art performance on several recent benchmarks in terms of both visual quality and quantitative metrics with a UIQM of 3.71, a PSNR of 28.4 dB, and a SSIM of 0.88 and a perceptual LPIPS score of 0.082. These values significantly surpass the performance of baseline approaches.

The outline of the paper is organized as follows. Section 2 presents the related work for underwater dehazing and enhancement techniques, Sect. 3 presents our proposed underwater image dehazing enhancement method. Section 4 presents the experimental results and the discussion part. Section 5 is the conclusion of work. Section 6 is the future work.

Motivation

The key motivation behind this work is to enhance the image quality by enabling cGAN to learn complex mapping representations from low-quality to high-quality images for image dehazing applications and outperforms state-of-the-art models. However, cGANs model alone could lose minor spatial and channel-wise information that are essential for accuracy of the enhancement model. Image enhancement, specifically under challenging conditions such as blurry or poor contrast, remains a critical task in underwater computer vision applications. Traditional enhancement methods often face challenges with balancing global contrast and preserving local details. We proposed a novel integration of a conditional Generative Adversarial Network (cGAN), a Bottleneck Attention Module (BAM), and an improved Retinex-based contrast enhancement technique to address these limitations.

In this work, we integrate the Bottleneck Attention Module (BAM) to refine the feature representations by modelling both spatial and channel domain. BAM enables the network to focus on significant regions and areas where texture restoration is essential and reduce artifacts, particularly in areas susceptible to degradation under low illumination conditions. In addition, we apply a refined Retinex-based contrast enhancement method that more accurately decomposes an image's illumination and reflectance regions. cGAN provides more consistent colors and natural illumination by improving contrast at this decompositional level, which in turn improves the prior. Our suggested approach combines cGANs for end-to-end learning, BAM for attention-guided refinement, and contrast-aware inputs processed using the Retinex method. The goal is to produce improved results that are both robust and aesthetically pleasing and that generalize effectively across various visual conditions.

Problem statement and research gaps

As suspended particles in water regions absorb and scatter light, underwater images frequently experience significant quality degradation, leading to low contrast, color distortion, reduced visibility, and loss of fine details. Although deep learning-based models and conventional image enhancement algorithms have made significant strides, existing enhancement strategies still struggle to successfully restore both fine local features and overall contrast. Many methods in underwater setup find it difficult to strike a balance between color correction, contrast improvement, and visibility consistency. An efficient and adaptable technique is therefore necessary to handle the complicated and uneven deterioration in underwater scenes.

Related work

This section describes underwater image enhancement and dehazing methods that can be elaborated into three groups: (1) Non-Physical based Enhancement Methods, (2) Physical Model-based Methods, (3) Deep learning-based Methods.

Non-physical based enhancement methods

These methods enhance images by manipulating the pixel intensity values directly and grouped into spatial domain, transform domain and fusion-based methods to improve the in-air and underwater images using

an atmospheric scattering model¹. These techniques help to improve the contrast and color brightness of underwater images. focus exclusively on improving visibility and contrast, neglecting the underlying physical causes of foggy pictures. There are various methods for the enhancement of underwater images. Contrast enhancement methods Techniques like histogram equalization (HE), Adaptive histogram equalization (AHE), Contrast limited Adaptive histogram equalization (CLAHE), mean filtering, medium filtering, bilateral filtering, Color correction & white balance, Gamma correction, Contrast stretching, retinex-based methods⁶.

To eliminate haze, Tan²² enhanced the contrast of a localized region by presuming that the ambient light level remains constant, the resultant dehaze image displayed enhanced contrast. Cheng et al.²³ proposed Histogram equalisation technique employed to adaptively improve the contrast of a hazy image. Mohan et al.²⁴ proposed a multiscale fusion by using CLAHE and gamma correction to enhance the contrast of white balanced image to recover the significant features and edges that have faded. Garg et al.²⁵ presented an approach to blend CLAHE and Percentile methodologies to improve the outcome. Zhou et al.²⁶ proposed adaptive feature enhancement by extracting the statistical characteristics of an image. Zhang et al.²⁷ used multi scale retinex to achieve dehazing of turbid water. Zhuang et al.¹⁰ presented an edge-preserving filtering method using retinex theory for underwater image enhancement and effective in artifacts and noise reduction and color preservation. Lin et al.²⁸ presented adaptive color correction with improved Retinex method and further NSST fusion for enhancing underwater images. Vasamsetti et al.²⁹ proposed a Wavelet-based perspective on the Variational enhancement technique for underwater images applied on approximation and detailed coefficients of an image with color correction method, yields an improved color and contrast of underwater images with preserved structural details of features yet not able to remove particle scatter.

Qiao et al.³⁰ proposed an approach that modifies the grey histogram distribution following the Rayleigh distribution by stretching it towards the lower and upper boundaries while diminishing the high value of grey level 255. Lin et al. proposed an enhanced adaptive color correction method using improved retinex theory. Zhou et al. presented an underwater image enhancement method via multi-interval subhistogram perspective equalization to address the issues posed by underwater images. However, improved images frequently experience over-enhancement, loss of information, and increased noise. The technique involves spatial and frequency enhancement through techniques like intensity transformation, homomorphic filter, and transforms, redistributing the histogram to improve the image quality³¹. Traditional underwater image enhancement methods have drawbacks such as color distortion, high contrast, sharp colors, noise amplification, and halo artifacts due to their inability to consider attenuation caused by water particles.

Physical model-based methods

Restoring the image quality with underwater physical models requires an effective degradation model. These models estimate additional parameters using prior models and polarization-based models. These models include dark channel prior (DCP), red channel prior, minimum information prior, etc. He et al.³² developed a DCP that can be used in-air and for underwater images to estimate light and transmittance and produce halo effects and may not recover images with high density. Drews et al.³³ proposed an underwater DCP to estimate transmission that considers blue and green color channels to reduce haze. Chen et al.³⁴ proposed a region-specialized method for underwater image restoration. Li et al.³⁵ proposed a dehazing method using minimum information loss using DCP and histogram distribution prior but producing artifacts and noise. Peng et al.³⁶ proposed a depth estimation method for image blurriness and light absorption for underwater images, which can be used in the image formation model (IFM) to restore and enhance underwater images. Priyadharshini et al.³⁷ proposed a two-stage technique using a DCP based algorithm for enhancing visibility in hazy images, which consists of contrast enhancement and haze removal. Song et al.³⁸ proposed an effective underwater restoration and color correction using new underwater dark channel prior. These methods are undoubtedly solving the dehazing and restore the images using degradation methods, but sometimes fail due to the assumption of inaccurate model parameters in underwater images.

Deep learning-based models

Deep learning methods have recently surpassed traditional methods with remarkable success in dehazing and enhancing underwater images by achieving color balance and contrast, and used by various state-of-the-art methods. Convolutional Neural Networks (CNNs)³⁹, Generative Adversarial Networks (GANs)¹⁸, and Reinforcement learning¹⁷ are popular models for underwater image dehazing and enhancement. These models automatically extract features from training data and, dehaze images through nonlinear mapping. Ling et al. proposed a deep transmission network for complex weather conditions. Mai et al.⁴⁰ implemented neural network to dehaze the images by observing color and depth of training samples and Tang et al.⁴¹ suggested best feature combination to identify the haze relevant features using deep learning. Hussain et al.⁴² proposed a model to approximate fog function using deep learning method.

Cai et al.⁴³ implemented DehazeNet using a convolutional neural network to find transmission map to accurately dehaze the images. Wang et al.³⁹ proposed UIENet for color correction and haze removal using wavelet and CNN. Li et al.⁴⁴ used lightweight CNN (UWCNN) for diverse underwater scenes. Li et al.⁴⁵ used the Ucolor network to improve color and contrast enhancement issues. Fu et al.⁴⁶ proposed SCNet for diverse water conditions but not performed well on all water conditions. Hou et al.⁴⁷ proposed DFFA-Net a lightweight model for producing clear and vivid images.

GANs are most popular dehazing method for underwater applications and address challenges that support to visually enhance underwater images that experience color distortion, low light, and blurry images. GANs generate high-resolution, realistic images for dehazing applications and shown better results compared to traditional methods. A GAN with a self-supervised dehazing network has been developed to improve dehazing efficacy on actual hazy photos, addressing these constraints. This model utilizes generative adversarial learning

to create a link between dehazed and hazy images, enhance the natural appearance of dehazed outcomes. Li et al.⁴⁸ proposed conditional GANs for achieving image dehazing. Fast underwater image enhancement (FUnIE), a supervised learning approach, was used by Islam et al.⁴⁹ to improve underwater images in real time. Nevertheless, there are limits to how well this technique may improve badly damaged photographs. For improved underwater image processing, Yang et al.⁵⁰ introduced a GAN-based architecture with a multi-scale feature discerning mechanism by successfully capturing local semantic information, a discriminator model can provide visual outputs that are more convincing and lifelike.

In order to improve enhancement results in underwater environments, Jiang et al.⁵¹ presented a novel model that addressed turbidity and chromaticity difficulties. Fabbri et al.⁵² presented a GAN to improve underwater visual quality by using CycleGAN⁵³ to generate paired images to enhance the underwater image. Using the underwater image degradation model, In^{54,55} introduced an unsupervised GAN for correction of color cast and underwater image enhancement. Wang et al.⁵⁶ also proposed unsupervised GAN based on improved underwater imaging model and used U-net to dehaze and restore colors of underwater images. Awan et al.⁵⁷ proposed underwater restoration through color correction module and UWNet. In following Table 1 is the summary of selected state-of-art review on underwater images affected by color distortion, low contrast, light scattering and degradation issues.

The traditional enhancement methods face challenges in estimating the parameters accurately while model-based methods produce artifacts and contrast issues. However, deep learning methods have shown better visual quality on reference datasets but face challenges when lack high-quality training sets. The idea in the novel approach is to implement the hybrid model that combines the Generative adversarial networks (GAN) with enhancement techniques to reduce noise and artifacts in GAN-generated outputs. It enhances the contrast and uneven illumination to recover a clear image.

Proposed methodology

Underwater image enhancement is a challenging task because of light scattering, absorption and colour distortion that degrade image quality. Traditional methods like histogram equalisation and white balance often fail to alleviate these issues to a satisfactory extent. In order to overcome these limitations, a novel hybrid model is presented in Fig. 3 incorporating a Generative Adversarial Network (GAN) with a Bottleneck Attention Module (BAM) and an improved Retinex-based contrast enhancement algorithm. It ensures effective dehazing, noise reduction, and lighting adjustment while preserving structural integrity and natural colours.

Dataset description

This paper uses the Underwater Image Enhancement Benchmark (UIEB) dataset, an openly available dataset of underwater images to be restored and enhanced. The data set consists of 950 high-quality paired images, each consisting of a raw underwater image and its corresponding reference image, 890 of which have the corresponding reference images, in which 60 underwater images which contain degraded challenging data. The raw photos exhibit common underwater degradation effects, including haze, colour distortion (mainly blue-green dominance), low contrast, and non-uniform illumination, simulating real-world underwater environments. These images cover a wide range of underwater environments, such as coral reefs, marine life,

Ref and Year	Method	Datasets	Metrics	Findings	Demerits
li et al. ⁴⁴	UWCNN	RGB-D NYU-v2	PSNR, SSIM, MSE	This method removes the color casts and improve the contrast of the underwater videos without artifacts.	To further improve the performance of model on different dataset.
Li et al. ⁵⁴	WaterNet	UIEB	PSNR, SSIM, MSE, UCIQE, UIQM	This model provides improved image accuracy for both reference and non-reference metrics.	Extend the work using more challenging video datasets.
Li et al. ⁴⁵	Ucolor	UIEB, SQUID	PSNR, MSE, UCIQE, UIQM	This model achieves the color cast and low contrast issues of underwater images with medium transmission-guided multi-color space embedding.	The goal is to work on diverse training data.
Jiang et al. ⁵⁸	Cycle GAN and UNet	EUVP + UIEB + lie et al.	UICM, UISM, UIQM, UICoM	A novel domain adaptation framework based on transfer learning for Underwater image enhancement.	To achieve a faster runtime and reduce the time complexity.
Hou et al. ⁴⁷	DCNN	UIEB	PSNR, SSIM	A novel differential convolutional neural network designed for underwater optical image dehazing for producing clear and vivid optimized images	To achieve a more lightweight model to increase its efficiency and applicability in various underwater imaging scenarios
Priyadharshini et al. ³⁷	White balance (WB), Gamma correction (GC), CLAHE and CNN model	UIEB, EUVP	PSNR, SSIM, MSE, UCIQE, UIQM	The primary focus is to enhance image brightness, sharpening, and reducing over-contrast amplification using the spatial method with CNN based method	More improvements can be done on enhancement of image quality
li et al. ⁵⁹	Adaptive full dynamic range mapping	UIEB, UCCS, U45	UCIQE, UICM, NIQE, PSNR, SSIM	The proposed AFDRM strategy used to maximize image contrast and minimize contrast loss for a color-balanced high-contrast image.	needs more improvements for low light conditions
Wang et al. ⁶⁰	Large foundation model	U45, SUIM, UIEB	UIQM, CCF, AG, BC(e), BC(r), and URanker	The large foundation model uses fusion to improve the color histogram similarity for foreground and background regions to improve color variations and contrast.	More work to be done on empirical fine-tuning through Reinforcement learning.

Table 1. Summary of selected state-of-the-art review papers on underwater enhancement and dehazing.

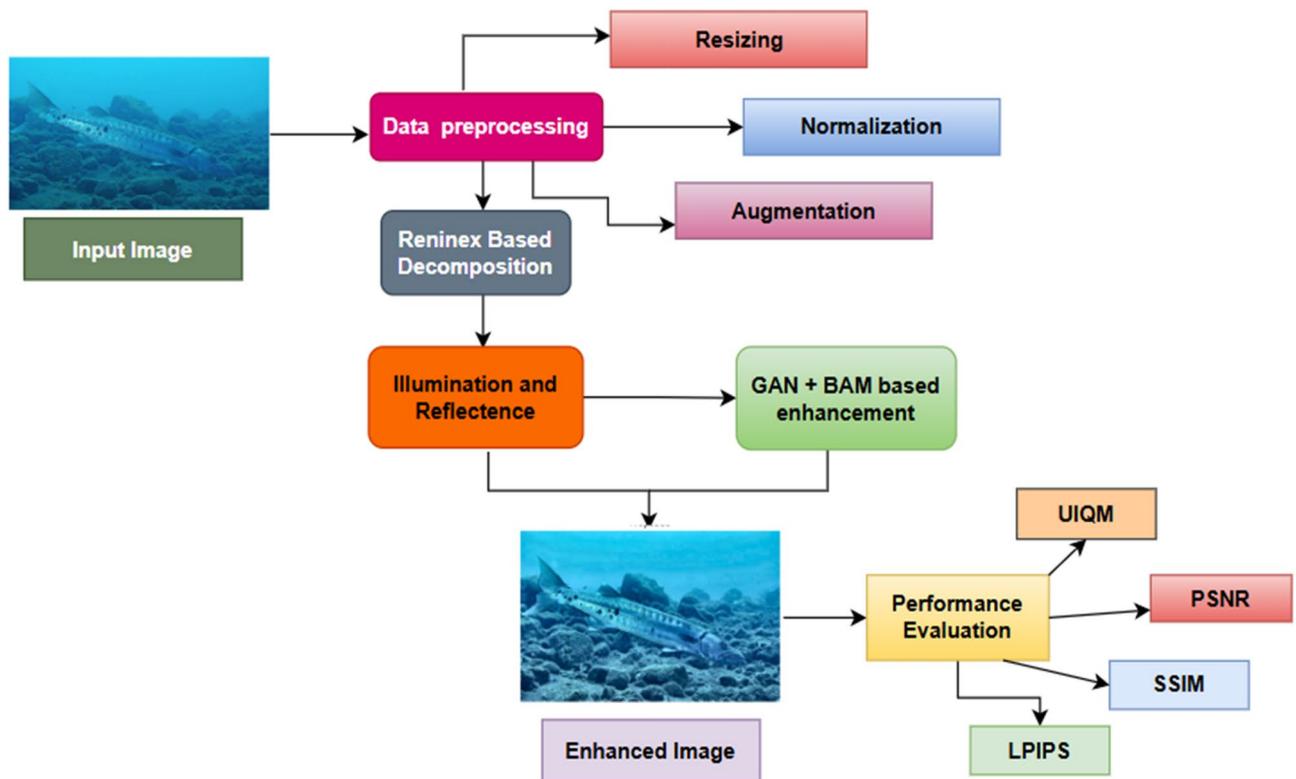


Fig. 3. Flowchart of proposed hybrid model.

and underwater items, thereby ensuring variation in light levels, water visibility, and color distortions. The data set is particularly well-suited for training deep learning networks as it offers an even balance of representative underwater challenges, such as blue-green color casts, poor contrast, and foggy appearance. Additionally, the presence of reference images permits supervised learning as the model may make predictions for actual correlations between corrupted inputs and enhanced outputs. The diversity of the dataset in degradation types and complexity of the scenes sets it as a trustworthy benchmark to evaluate underwater image enhancement methods, ensuring that trained models generalize well to varied underwater environments.

Data preprocessing

In the interest of maximising model performance, the data is put through multiple preprocessing steps, viz., normalisation, augmentation, and resizing. Raw and reference pixel values are first normalised to the range [0, 1] by dividing them by 255 to facilitate stable and rapid convergence during training. Next, data augmentation techniques such as random vertical and horizontal flip, rotation (+/-15°), and small changes in brightness are synthetically applied to augment the dataset and improve generalisation. Images are also resized to a uniform resolution of 256 × 256 pixels, which maximises computation efficiency at the cost of integrity in detail. The preprocessing techniques enhance the model's ability to process varied undersea scenarios and provide uniform input sizes for batch processing at speed.

Model building

This model is explained in Algorithm 1, combines a multi-stage deep learning framework using an improved Retinex-based decomposition, Generative Adversarial Network (GAN), and Bottleneck Attention Module (BAM) as a solution for the intricate process of underwater image restoration as explained in Fig. 4 below. Executing Retinex-based improvement prior to adversarial training helps provide the GAN with more physically consistent and stable input that enables better dehazing as well as detailing preservation during improvement.

Improved Retinex-based enhancement

This framework begins with a improved Retinex decomposition network that achieves non-linear illumination-reflectance separation using a special convolutional sub-network, thus pre-guiding the GAN with physically meaningful features from the very beginning. This initial decomposition allows us to separate illumination effects, caused by underwater lighting inconsistencies, from true reflectance properties, which relate to objects' actual colors and textures. Unlike traditional Retinex approaches that rely on manually defined priors, our approach uses adversarial training to acquire decomposition, with an auxiliary discriminator to ensure that both illumination and reflectance maps are physically valid. The illumination term is adapted to gamma correction to normalize brightness variations, whereas the reflectance term is processed through a color-consistent restoration block, using histogram equalization in LAB color space to reduce color casts.

- 1: Input: Raw underwater image I_{raw} , from UIEB dataset
- 2: Output: Enhanced underwater image I_{enh}
- 3: **Step 1: Data Collection and Description**
- 4: Collect 950 image pairs (I_{raw}, I_{ref}) from UIEB dataset
- 5: Use 890 images with ground truth and 60 challenging samples
- 6: **Step 2: Data Preprocessing**
- 7: for each image pair (I_{raw}, I_{ref}) do
- 8: Normalize pixel values to $[0,1]$
- 9: Resize to 256×256
- 10: Apply augmentation: flips, rotations, brightness variation
- 11: end for
- 12: **Step 3: Generator Design**
- 13: Build Generator G using U-Net with Dense Residual Blocks
- 14: Integrate Bottleneck Attention Module (BAM) into G
- 15: BAM uses channel and spatial attention mechanisms
- 16: **Step 4: Discriminator Design**
- 17: Design multi-scale PatchGAN Discriminator D
- 18: Integrate BAM to refine attention at feature levels
- 19: **Step 5: Retinex-Based Contrast Enhancement**
- 20: Decompose I_{raw} into Illumination (L) and Reflectance (R)
- 21: Enhance L with Gamma correction, R with LAB equalization
- 22: **Step 6: Adaptive Feature Normalization**
- 23: Apply Adaptive Instance Normalization (AdaIN) to handle varying image degradations
- 24: **Step 7: Loss Function Design**
- 25: Compute adversarial loss L_{adv} from D
- 26: Compute Retinex loss $L_{Retinex}$
- 27: Compute BAM attention loss L_{BAM}
- 28: Total loss:

$$\mathcal{L}_{total} = \mathcal{L}_{adv} + \lambda_1 \mathcal{L}_{Retinex} + \lambda_2 \mathcal{L}_{BAM}$$
- 29: **Step 8: Training**
- 30: for each epoch do
- 31: Generate enhanced image $\hat{I}_{enh} = G(I_{raw})$
- 32: Compute L_{total} and backpropagate
- 33: Update weights of G and D
- 34: end for
- 35: **Step 9: Evaluation**
- 36: Compute PSNR, SSIM, LPIPS and UIQM metrics for I_{enh}
- 37: $UIQM = c_1 \cdot UICM + c_2 \cdot UISM + c_3 \cdot UICoM$
- 38: **return** Enhanced underwater image I_{enh}

Algorithm 1. Underwater image enhancement and dehazing using GAN with BAM & Retinex-based enhancement.

The Retinex theory states that illumination and reflectance create images. Source of illumination affects illumination properties. Reflectance depends on object properties. According to Retinex theory, illumination can be calculated by dividing image by reflectance. It's not practical to measure illumination or reflectance. Image illumination cannot be estimated without reflectance. Then, illumination, reflectance, or both assumptions and adjustments are suggested to fix this. Edges are assumed to be the same for both scene and reflectance, and illumination changes gradually. Thus, Retinex-based reflectance estimation is mostly a ratio of the image to its smoothed form, which approximates illumination. The Retinex-based approach to image enhancement mainly recovers from illumination effects on the image. Based on the image formation model of the Retinex approach:

$$I(m, n) = R(m, n) \cdot L(m, n) \quad (1)$$

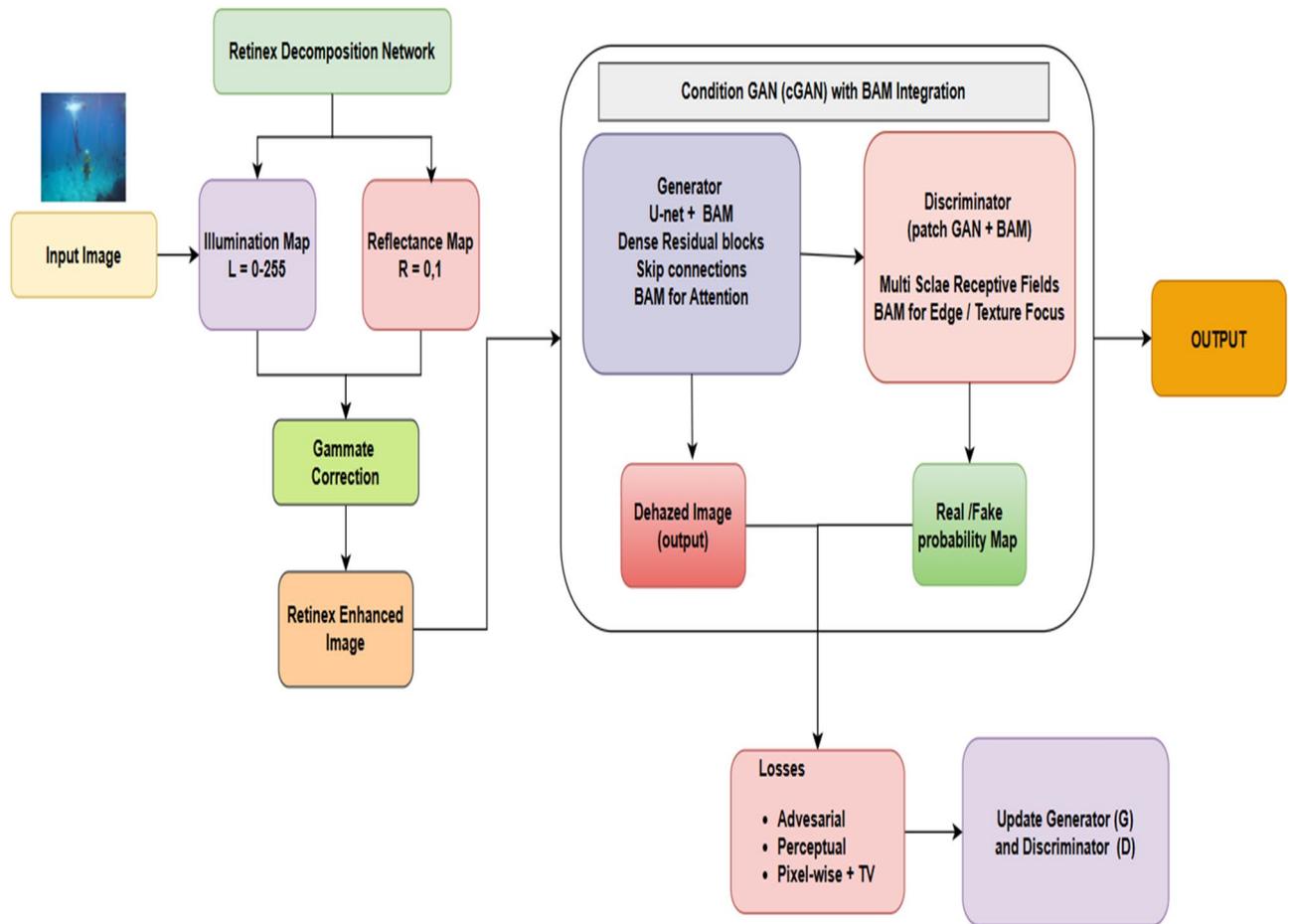


Fig. 4. Proposed hybrid GAN model architecture.

Where $I(m, n)$ is the image, whose value ranges from 0 to 255; $R(m, n)$ is the reflectance of the object, ranging from 0 to 1; and $L(m, n)$ represents the illumination. The range of illumination is within the range of 0 to 255. The logarithmic transformation of the image, as given by Eq. 1, is shown below:

$$\log I = \log R + \log L \quad (2)$$

The obtained Eqs. (1) and (2) from the nature of the logarithmic function. Equation 2 shows that the logarithm of the image is the sum of the logarithm of reflectance and the logarithm of illumination. For seeking to find the reflectance, which is found by rearranging Eq. (2) as the following.

$$\log R = \log I - \log L \quad (3)$$

In Eq. (3), the logarithm of reflectance is used, and by exponentiating Eq. (3), we can obtain the estimation of reflectance.

$$R = \exp(\log I - \log L) \quad (4)$$

To mitigate the ill-posed characteristics of this inverse problem, we utilise an auxiliary discriminator during training that ensures the physical validity and statistical realism of the separated illumination and reflectance components. This adversarial constraint aids the network in acquiring decompositions that align with natural image priors and the principles of scene physics. Recent deep learning-based Retinex methods, including Zero-DCE++⁶¹, URetinex-Net⁶², and EnlightenGAN⁶³, demonstrate that end-to-end networks, when trained with perceptual and adversarial objectives, can attain physically plausible decompositions without the need for explicit handcrafted priors.

In this design, the illumination L is normalised through adaptive gamma correction to mitigate lighting inconsistencies, while the reflectance R is optimised via histogram equalisation in the LAB colour space to improve colour fidelity. In contrast to traditional methods, our model directly learns spatially aware smooth approximations of illumination and edge-preserving reflectance from data. These assumptions—gradual changes in illumination and sudden shifts in reflectance—are inherently integrated into our architecture and loss formulation.

Implementing Retinex decomposition prior to GAN training facilitates superior disentanglement of visual factors, enhances training stability, and yields improved enhancement outcomes, including haze removal, contrast augmentation, and colour correction. Moreover, the physically interpretable intermediate representations (L , R) facilitate stable gradient propagation and expedite convergence.

As per Eq. (4), reflectance is calculated based on the difference between the logarithm of the illumination and the logarithm of the image. Equation 4 demonstrates that reflectance can be estimated by knowing the illumination and image. Different filters proposed can be used to estimate illumination. Filters can also be used for smoothing the image, and then smoothed image may be used as illumination in any of the methods based on Retinex used for image sharpening⁶⁴. Placing this step ahead of the GAN enables clearer differentiation of visual factors, hence better GAN convergence and more accurate haze removal, contrast enhancement, and color correction in later stages. Additionally, Retinex's interpretable middle representations improve training stability and optimize gradient flow.

GAN architecture for dehazing

Following Retinex-based restoration, we use a conditional GAN (cGAN) framework with a variation of the U-Net generator to learn the dehazing mapping. The generator utilizes dense residual blocks and skips connections for maintaining subtle image details and for effective gradient flow. The discriminator employs a multi-scale PatchGAN structure that tests realism on many receptive fields so that it is able to differentiate accurately real high-frequency textures and clean regions. The GAN benefits from feeding Retinex-enhanced images, which improves training stability and dehazing quality by eliminating low-level ambiguities within the image composition. Different from original GAN⁹, cGAN is tasked with generating a specific image. In Eq. 5, conventional cGAN, the generator G synthesises a dehazed image $G(I, z)$, while the discriminator D differentiates between authentic image pairs (I, J) and counterfeit ones $(I, G(I, z))$. Where I is input image, J is ground truth clean image, and z = random noise vector. The objective function is articulated as:

$$\min_G \max_D E_{L,z} \log [1 - D(I, G(I, z))] + E_{I,J} [\log D(I, J)] \quad (5)$$

Conditional Generative Adversarial Networks (cGANs) have achieved significant progress in a variety of image processing tasks such as super-resolution, image inpainting, and style transfer. For example, semantic image inpainting and image-to-image translation techniques show that conditioning the GAN with additional input information achieves more meaningful and visually plausible results. In image super-resolution problems, modifications to the GAN architecture—like adding pixel-wise content loss and perception loss—enable the generation of high-fidelity, realistic outputs. These improvements demonstrate the effectiveness of cGANs in a multitude of visual tasks⁴⁸.

Generator and discriminator The generator's function is to convert a hazy image into a clear image. It preserves the structure and detail information of an input image and also removes the haze. The generator uses skip connections in symmetric layers for decoding process. The generator comprises of an encoding and decoding process, with the encoding process utilizing down-sampling and feature maps, and the decoding process using up-sampling and nonlinear space transfer. The discriminator finds whether an image is real or fake. It uses a neural network and utilizes convolution, batch normalization and LeakyReLU activation function, In the last layer, it applies a sigmoid function for probability score normalization.

Perceptual loss function Let $\{I_i, i = 1, 2, \dots, N\}$ and $\{J_i, i = 1, 2, \dots, N\}$ denote the hazy images and the corresponding clear images.

To recover realistic images, we introduce the perceptual loss based on the pre-trained VGG features to constrain the generator, which is defined as

$$L_P = \frac{1}{N} \sum_{i=1}^N \|F(G(I_i))_i - F_i(J_i)\|_2^2 \quad (6)$$

In Eq. (6), F_i denotes the feature maps of the i -th layer of a pre-trained VGG network. The impact of perceptual loss has been shown to be advantageous in numerous image generation tasks, including super-resolution and image restoration. In contrast to other applications, we note that the implementation of this perceptual loss facilitates the restoration of image details and the elimination of haze. Nonetheless, it may also produce significant artefacts in the reconstructed images, thereby diminishing the visual quality. To resolve this issue and enhance the preservation of structural details, we implement a L_1 regularisation gradient prior on the generator's output in conjunction with a content-based pixel-wise loss.

$$L_T = \frac{1}{N} \sum_{i=1}^N \|G(I_i) - J_i\|_1 + \lambda \|\nabla G(I_i)\|_1 \quad (7)$$

In Eq. (7), $\|\nabla G(I_i)\|_1$ denotes the total variation regularization, $\|G(I_i) - J_i\|_1$ is the content-based pixel-wise loss, and λ is the regularization weight. This loss function is able to remove the artifacts and preserve the details. Ultimately, we integrate the adversarial loss, perceptual loss, L1-regularized gradient prior, and content-based pixel-wise loss to constrain the proposed generative network, which is articulated as

$$\mathcal{L} = \alpha L_A + \beta L_P + \gamma L_T \quad (8)$$

where α , β and γ are the positive weights. The generator G is trained by minimizing Eq. (8). The scalar λ controls the strength of the TV regularization. To determine optimal values for the weights α , β , and γ , we conducted an empirical grid search across the validation set. We evaluated combinations within the following ranges:

- $\alpha \in \{0.1, 0.5, 1.0\}$,
- $\beta \in \{0.5, 1.0, 2.0\}$,
- $\gamma \in \{10, 50, 100\}$.

The configuration $\alpha = 1.0$, $\beta = 1.0$, and $\gamma = 50$ achieved the optimal balance between visual realism and structural integrity, as evidenced by enhancements in PSNR, SSIM, and LPIPS metrics on the validation set. These weights facilitated adversarial training to enhance high-frequency realism, perceptual loss to preserve global structure, and regularised content loss to guarantee pixel-wise coherence and artefact mitigation. After obtaining the intermediate generator G , we update the discriminator D given in Eq. 9 by

$$\max_D \frac{1}{N} \sum_{i=1}^N (\log(1 - D(I_i, \tilde{J}_i)) + \log(D(I_i, \tilde{J}_i))) \quad (9)$$

BAM enhances feature learning

To augment feature representation, we integrate a Bottleneck Attention Module (BAM) at pivotal points in both the generator and discriminator. The BAM functions via parallel channel and spatial attention branches, with the channel attention mechanism dynamically adjusting feature significance through global average pooling and multi-layer perceptron, while the spatial attention branch utilises dilated convolutions to acquire extensive contextual information without compromising resolution. This dual-attention technique enables the network to choose concentrate on prominent areas, including edges, textures, and colour boundaries, while diminishing extraneous background noise.

To augment the representation learning capacity of the network, we integrate a Bottleneck Attention Module (BAM), which combines channel and spatial attention mechanisms. Given an input feature map $F \in R^{C \times H \times W}$, the channel attention mechanism initiates with a squeeze operation via global average pooling to generate a descriptor $z \in R^c$, which encapsulates channel-wise global context. This procedure is articulated in Eq. (10):

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_C(i, j) \quad (10)$$

$F_C(i, j)$ denotes the value at spatial position (i, j) in the c -th channel.

The channel descriptor z is subsequently processed by a two-layer fully connected network utilising a non-linearity (commonly ReLU) and sigmoid activation to produce the channel attention weights $s \in R^c$. The initial feature map is adjusted with these weights through element-wise multiplication:

$$F'_C = s_C \cdot F_C \quad (11)$$

In parallel, the spatial attention module captures spatial dependencies by applying both average pooling and max pooling along the channel axis to produce two 2D maps presented in Eqs. (12)–(14):

$$M_{avg}(i, j) = \frac{1}{C} \sum_{c=1}^C F_C(i, j) \quad (12)$$

$$M_{max}(i, j) = \max_{c \in [1, c]} F_C(i, j) \quad (13)$$

The maps are concatenated and processed through a convolutional layer, succeeded by a sigmoid activation, to produce a spatial attention map $M_s \in R^{H \times W}$, which is employed to scale the feature map:

$$F''(i, j) = M_s(i, j) \cdot F'(i, j) \quad (14)$$

The ultimate result of BAM is the element-wise multiplication of the input feature map with both the channel and spatial attention maps, thereby recalibrating the significance of each feature spatially and across channels. This dual-attention framework allows the network to highlight significant areas, such as object boundaries, while mitigating extraneous background noise, thereby enhancing feature extraction in both the generator and discriminator architectures⁶⁵.

Performance metrics

Underwater image quality measure (UIQM)

The underwater image quality measure (UIQM) is comprehensive metric used to evaluate the quality of underwater images by assessing their colorfulness, sharpness, and contrast presented in Eq. (15).

$$UIQM = c_1 \times UICM + c_2 \times UISM + c_3 \times UICoM \quad (15)$$

where:

- *UICM* (underwater image colorfulness measure): Evaluates color distortion.
- *UISM* (underwater image sharpness measure): Assesses image sharpness.
- *UICoM* (underwater image contrast measure): Measures contrast enhancement.
- c_1, c_2, c_3 are weights that balance the contribution of each component.

Peak signal-to-noise ratio (PSNR)

The Peak Signal-to-Noise Ratio (PSNR) is a commonly employed metric for evaluating the quality of a reconstructed or compressed image in relation to its original counterpart. It is characterised by the Mean Squared Error (MSE) and the maximum attainable pixel value (I_{max}) shown in Eq. (16).

$$PSNR = 10 \cdot \log_{10} \left(\frac{I_{max}^2}{MSE} \right) \quad (16)$$

Structural similarity index (SSIM)

The Structural Similarity Index (SSIM) is a perceptual metric that quantifies image quality by comparing structural information rather than just pixel differences. In Eq. (17), it considers three factors: luminance, contrast, and structure.

$$SSIM(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (17)$$

Where:

- x and y are the original and distorted images.
- μ_x and μ_y = Mean intensities of x and y .
- σ_x^2, σ_y^2 = Variances of x and y (measure of contrast).

$x \rightarrow y$

- σ_{xy} = Covariance between x and y (measures structural similarity).
- C_1 and C_2 = Small constants to stabilize the division.

Learnt perceptual image patch similarity (LPIPS)

This metric effectively measures the perceptual quality of underwater image enhancement. In contrast to conventional full-reference metrics like PSNR and SSIM, which focus on pixel-wise or structural similarity, LPIPS aims to emulate human visual perception by analysing deep feature activations derived from a pre-trained neural network⁶⁶.

The calculation of LPIPS entails multiple stages. The original image x and the generated image $F(G(x))$ are processed through a predetermined, pre-trained feature extraction network, such as AlexNet, SqueezeNet, or VGG. These networks produce multi-layer feature representations that are subsequently normalised across the channel dimension. The normalised feature maps from each layer are represented as x_i^* and $F(G(x))_i^*$.

Subsequently, each pair of feature maps is scaled element-wise by a set of learnt weights ω_i , and the Frobenius norm is calculated over their differences. The average norm is computed across the spatial dimensions of the feature maps, and the contributions from all layers are aggregated to derive the final LPIPS score. This procedure is officially delineated in Eq. (18):

$$LPIPS(x, F(G(x))) = \sum_i \frac{1}{H_i W_i} \omega_i \left\| (x_i^* - F(G(x))_i^*)_F \right\| \quad (18)$$

In this context, H_i and W_i signify the height and width of the i -th feature map layer, respectively, while ω_i denotes the learnt weight tensor, and \odot indicates element-wise multiplication. The Frobenius norm $\| \cdot \|_F$ for a matrix $A \in \mathbb{R}^{(m \times n)}$ is defined in Eq. (19):

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \quad (19)$$

Incorporating LPIPS into the evaluation framework allows for the assessment of the model's performance based on both low-level pixel accuracy and high-level perceptual fidelity. This is particularly vital in underwater

environments, where preserving colour fidelity, structural integrity, and semantic information is imperative for practical uses like underwater navigation and object recognition.

Results and discussion

Quantitative results

To thoroughly assess the efficacy of the proposed underwater image enhancement model (GAN + BAM + Retinex), we performed both quantitative and perceptual comparisons with a variety of classical and deep learning-based techniques. The assessment encompasses various objective quality metrics, including UIQM, PSNR, SSIM, and LPIPS, ensuring a comprehensive evaluation of visual quality, structural similarity, and perceptual fidelity as illustrated in (Table 2). Supplementary benchmarks on computational complexity (MACs, Params) and ablation studies further corroborate the efficacy and modular contribution of the proposed methodology.

In the initial comparative assessment, we evaluated our proposed GAN + BAM + Retinex model against various classical and GAN-based underwater image enhancement methods utilising three fundamental metrics: UIQM, PSNR, and SSIM. The proposed model attained a UIQM score of 3.7, significantly surpassing all rival methods, including UMGAN (3.49) and GuidedHybSensUIR (3.22), thereby indicating superior perceptual quality. Our model achieved a PSNR of 28.4 dB, exceeding all baselines; MGDC-Net was the nearest competitor at 27.81 dB, yet it considerably lagged in SSIM performance. Our method achieved a competitive SSIM of 0.88, near to the performance of GuidedHybSensUIR (0.923), WaterNet (0.91) and UMGAN (0.899), while exhibiting superior balance across all three metrics. A particularly significant finding is evident in the instance of MGDC-Net, which exhibits an unexpectedly low SSIM value of 0.134 despite its elevated PSNR. This unusual behaviour likely arises from the significant dependence on contrast enhancement methods such as CLAHE, gamma correction, and DCP, which may have enhanced overall brightness and pixel-level accuracy (increasing PSNR) while compromising structural integrity and texture coherence (reducing SSIM). This significant discrepancy indicates that the model either generates unnatural local distortions or inadequately maintains intricate spatial details. Moreover, potential complications in metric computation—such as incongruent reference images, inadequate normalisation, or varying resolutions—could further intensify the reported SSIM shortfall. These results establish the proposed method as perceptually superior and structurally consistent, effectively balancing contrast, texture, and fidelity.

This work extended the comparison by integrating more recent and perceptually relevant metrics—specifically LPIPS—while assessing a wider array of contemporary enhancement models, including GuidedHybSensUIR, UDCP, and UWNet. This proposed model exhibited exceptional performance with an LPIPS score of 0.082, which is superior to UMGAN (0.0856), Sea-Pix GAN (0.091), UGAN (0.0935). LPIPS, grounded in deep perceptual similarity and more closely aligned with human visual assessments, substantiates the assertion that our model not only achieves pixel accuracy but also maintains visual realism. The proposed method achieves the highest PSNR (28.4 dB) among all evaluated models, demonstrating robust fidelity to the original content, while preserving perceptual similarity with a minimal LPIPS value. The dual validation employing both traditional (PSNR, SSIM) and perceptual (LPIPS) metrics confirms the robustness of our enhancement pipeline. The findings indicate that although traditional metrics such as PSNR may be exaggerated in specific models due to pixel-level modifications, the integration of LPIPS offers a more precise representation of qualitative enhancements and structural coherence in improved underwater scenes.

In Fig. 5, the UIQM plot evaluates the perceptual quality of underwater images based on conventional no-reference quality assessment criteria. The proposed model achieves the highest UIQM of 3.7, signifying exceptional improvement in underwater visibility and realism. UMGAN produce comparable results of 3.49, GuidedHybSensUIR achieves a UIQM of 3.22, closely followed by CLAHE + Zero Shot + White balancing at 3.2. UGAN and WaterGAN exhibits a UIQM of 2.91 and 1.963, UDCP and UW-Net exhibit the lowest UIQM values of 0.8547 and 1.58, respectively. Nonetheless, MGDC-Net, despite exhibiting a high PSNR, registers the lower UIQM score of 1.75, suggesting that its pixel-level enhancements may have undermined perceptual quality. This plot distinctly demonstrates the proposed method's capacity to generate more aesthetically pleasing underwater images compared to traditional and deep learning alternatives.

Method	UIQM	PSNR (dB)	SSIM	LPIPS
WaterNet ⁶⁷	2.83	24.60	0.91	–
MGDC-Net ⁶⁸	1.75	27.81	0.134	0.112
UGAN (Underwater GAN) ⁵²	2.91	22.15	0.81	0.0935
UMGAN (Underwater multi-scene GAN) ⁶⁹	3.49	23.55	0.899	0.0856
Sea-pix GAN ⁷⁰	2.63	23.36	0.76	0.091
GuidedHybSensUIR ⁷¹	3.22	23.63	0.923	0.1
UDCP ⁷²	0.8547	17.35	0.842	0.1215
WaterGAN ⁵⁴	1.963	20.25	0.836	0.0994
CLAHE + Zero Shot + White balancing ⁷³	3.2	17.95	0.807	0.189
UW-Net ⁵⁷	1.58	15.54	0.587	0.399
Proposed work	3.7	28.4	0.88	0.082

Table 2. Quantitative comparison of underwater image enhancement methods based on UIQM, perceptual metric (LPIPS) and fidelity metrics (PSNR, SSIM).

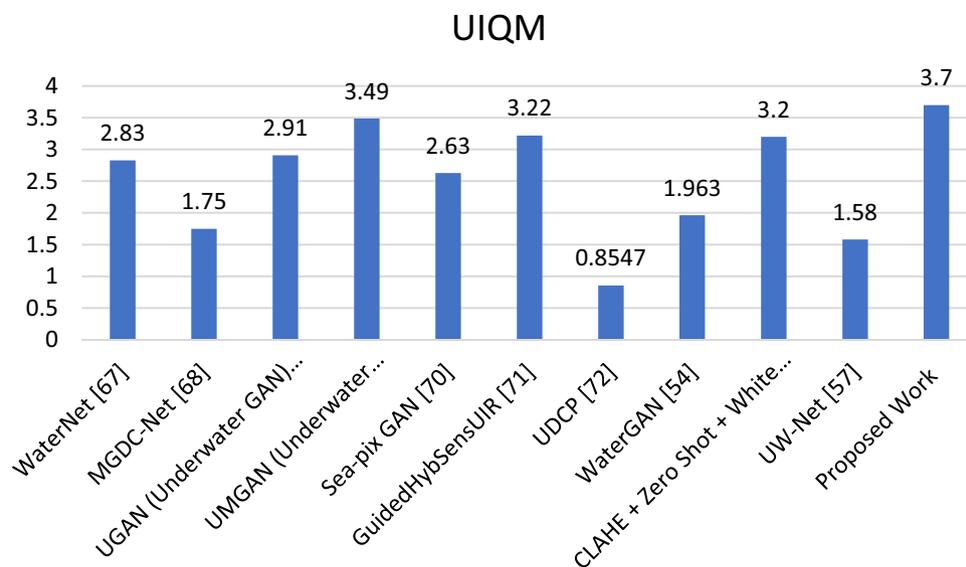


Fig. 5. UIQM value comparison for each model.

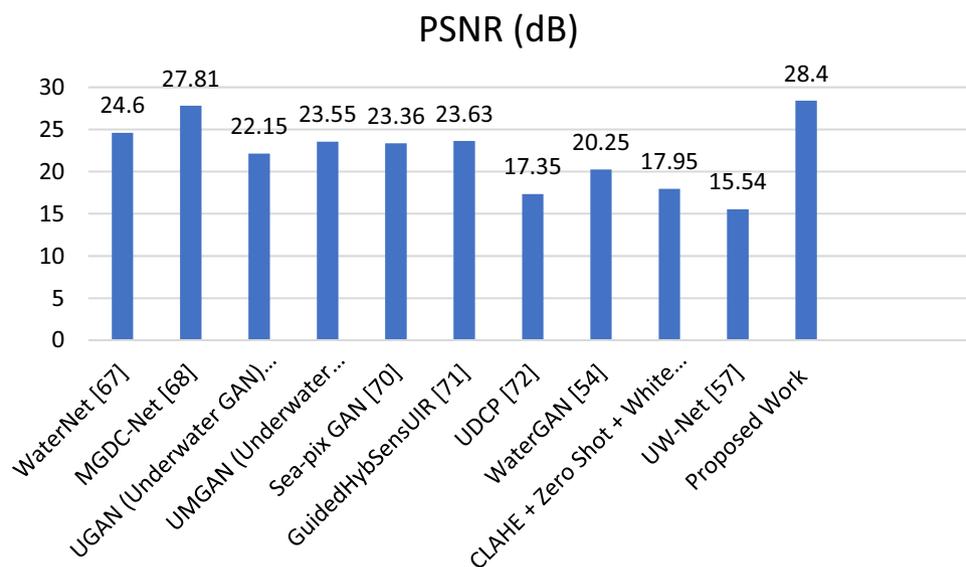


Fig. 6. PSNR values for all models.

In Fig. 6, the PSNR plot, which indicates image fidelity regarding noise reduction and pixel accuracy, shows that the proposed method attains a peak PSNR of 28.4 dB, markedly surpassing all alternative methods. GuidedHybSensUIR achieves 23.63 dB, whereas UGAN, Sea-Pix GAN, and WaterGAN exhibit moderate outcomes of 22.15 dB, 23.35 dB, and 20.25 dB, respectively. CLAHE combined with Zero Shot and white balancing achieves 17.95 dB, slightly surpassing UDCP's 17.35 dB. UW-Net exhibits the poorest performance, achieving a PSNR of merely 15.54 dB; however, they remain inferior to the proposed method. This plot verifies that the proposed model achieves robust pixel-level precision and clarity under various image conditions.

In Fig. 7, the SSIM plot demonstrates each method's capacity to maintain the structural integrity and textural nuances of underwater scenes. GuidedHybSensUIR attains a maximum SSIM of 0.923, indicating nearly flawless structural preservation. The proposed model achieves a robust SSIM of 0.88, underscoring their capacity to preserve spatial coherence. WaterNet achieves a marginally superior score of 0.91, yet exhibits subpar performance in PSNR. Notably, UMGAN (0.899) exhibits robust structural retention. In sharp contrast, MGDC-Net exhibits a notably low SSIM of 0.134, despite its elevated PSNR, indicating significant structural distortions—presumably resulting from excessive processing during its median filtering and contrast stretching procedures. Other classical methods such as UDCP and WaterGAN exhibit diminished structural similarity, with scores of 0.842 and 0.836, respectively. CLAHE combined with Zero Shot and white balancing attains an SSIM

SSIM

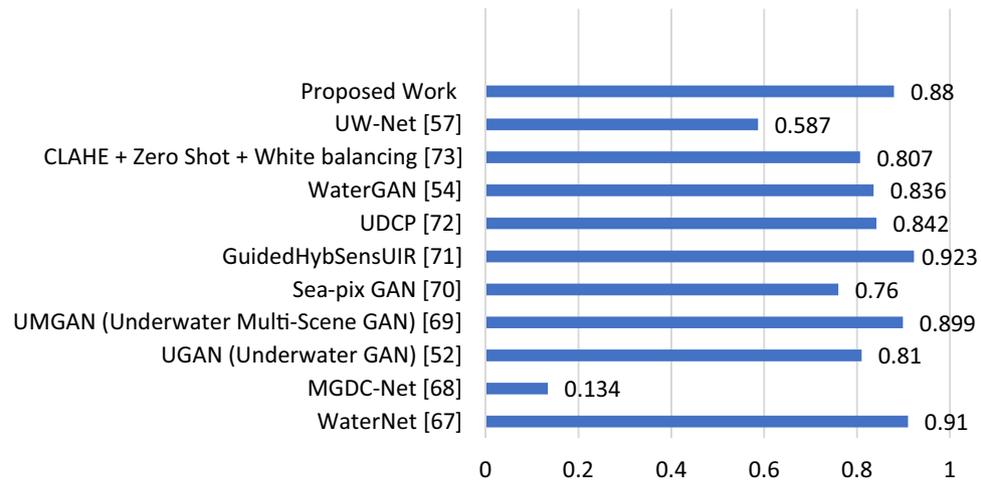


Fig. 7. SSIM values for all models.

LPIPS

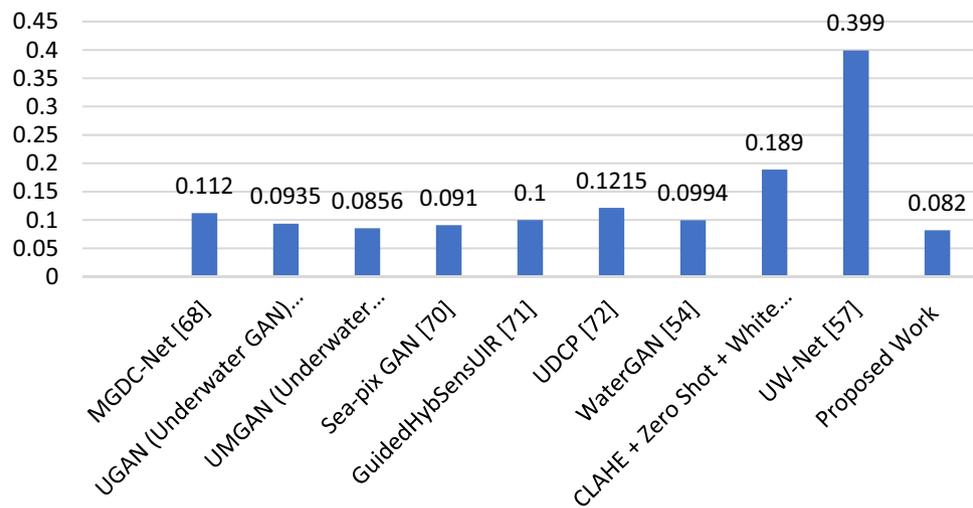


Fig. 8. LPIPS values for all models.

of 0.807, whereas UW-Net considerably trails at 0.587. This plot demonstrates that the proposed model achieves an advantageous equilibrium between improvement and the preservation of natural structure.

In Fig. 8, the LPIPS plot, which assesses perceptual realism through human visual characteristics, substantiates the superiority of the proposed method, achieving the lowest score of 0.082, signifying outputs that are perceptually nearest to ground truth images. GuidedHybSensUIR (0.1) and MGDC-Net (0.112) demonstrate comparable efficacy. Conversely, techniques such as, UDPC (0.1215), CLAHE + ZeroShot (0.189), and UWNet (0.399) exhibit elevated LPIPS values, indicating a decline in perceptual quality despite satisfactory SSIM and PSNR metrics. The LPIPS metric effectively complements UIQM, demonstrating that the proposed method excels in conventional metrics while also conforming to human perceptual standards, thereby guaranteeing both technical and subjective visual satisfaction.

Computational complexity

To emphasise the computational efficiency of the proposed method, this work compare its Multiply-Accumulate Operations (MACs) and parameter count (Params) with various state-of-the-art underwater image enhancement techniques. To evaluate this, Table 3 compares the Multiply-Accumulate Operations (MACs) and parameter counts (Params) of the proposed model with a variety of existing underwater image enhancement methods. The proposed work demonstrates significantly low computational complexity, featuring merely 36.9 G MACs and

Methods	MACs (G) ↓	Params (M) ↓
WaterNet ⁶⁷	45.2	2.6
MGDC-Net ⁶⁸	78.5	5.1
UGAN (Underwater GAN) ⁵²	62.7	3.8
UMGAN (underwater multi-scene GAN) ⁶⁹	70.3	4.4
Sea-pix GAN ⁷⁰	58.1	3.5
GuidedHybSensUIR ⁷¹	61.4	13.2
UDCP ⁷²	5.7	0.03
WaterGAN ⁵⁴	63.5	12.1
CLAHE + Zero Shot + White balancing ⁷³	2.3	0.01
UW-Net ⁵⁷	24.6	5.4
Proposed work	36.9	2.2

Table 3. Computational complexity using multiply-accumulate operations (MACs) and parameter counts (params).

2.2 million parameters, Other techniques such as CLAHE, Zero Shot, White Balancing, UDCP and UW-Net are remarkably lightweight, requiring 2.3G, 5.7G and 24.6MMACs respectively, and possess minimal parameters; however, they substantially compromise enhancement quality. Deep learning models like GuidedHybSensUIR (61.4G MACs, 13.2 M parameters), MGDC-Net (78.5G MACs, 5.1 M), UMGAN (70.3G MACs, 4.4 M) and SEA-pix-GAN (58.1G MACs, 3.5 M) exhibit significant computational demands, potentially limiting their utility in real-time or edge applications.

In comparison to conventional GAN-based models like UGAN (62.7 G, 3.8 M), and WaterGAN (63.5 G, 12.1 M), the proposed model attains a more favourable balance, delivering competitive or superior enhancement outcomes with reduced computational requirements. The equilibrium renders the proposed method especially attractive for practical applications, such as underwater robotics, portable marine devices, or remote sensing situations where runtime efficiency and model size are paramount constraints.

Quantitative results

The visual comparison depicts the performance of various underwater image enhancement models for a variety of underwater scenes, such as coral reefs, sea creatures, divers, and underwater objects. The image shows, from left to right, the original degraded inputs, the results of the proposed model, the results from other models and finishes with the ground truth images as references. In Fig. 9, the suggested model significantly outperforms sample image from dataset and skillfully recovering natural color balance, enhancing contrast, and preserving critical textures and structural details.

Unlike other models that often produce oversaturated green or blue colors, the suggested model provides visually pleasing results that are highly similar to the ground truth. Figure 9 presents a comparative visual analysis of underwater image enhancement methodologies applied to a sample image from the UIEB dataset from left to right Hazy underwater image, Ground Truth image, WaterNet⁶⁷, MGDC-Net⁶⁸, UGAN (Underwater GAN)⁵², UMGAN (Underwater Multi-Scene GAN)⁶⁹, Sea-pix GAN⁷⁰, GuidedHybSensUIR⁷¹, UDCP⁷², WaterGAN⁵⁴, CLAHE + Zero Shot + White balancing⁷³, UW-Net⁵⁷ and the Proposed Work. Of all the methods, the proposed model (GAN + BAM + Retinex) yields the most aesthetically pleasing results, exhibiting superior colour balance, restored structural details, and enhanced contrast, closely mirroring the ground truth. Other models demonstrate prevalent problems such as colour casting, blurriness, or texture loss, thereby underscoring the efficacy of the proposed method in difficult underwater environments.

Ablation study

To assess the individual and collective contributions of each element in the proposed underwater image enhancement framework, we conducted a comprehensive ablation study across various iterations. Here, Table 4 presents the ablated study of each key variant using UIEB dataset. The most basic version utilising solely the Retinex algorithm functions as a traditional benchmark. It attains moderate perceptual enhancement (UIQM: 2.91) through illumination correction and local contrast enhancement, yet it inadequately preserves structure and details, as indicated by its comparatively low SSIM (0.70) and elevated LPIPS (0.205). This suggests that although Retinex is beneficial for overall appearance modification, it lacks semantic or structural direction.

Conversely, employing a GAN-only architecture enhances fidelity metrics (PSNR: 25.1, SSIM: 0.82), illustrating its capacity to establish a significant mapping from hazy to clear images. This variant continues to generate artefacts and inconsistent textures due to the lack of explicit contrast and feature refinement mechanisms.

The integration of GAN with Retinex results in enhanced PSNR and SSIM scores (26.7 and 0.84, respectively), signifying improved structural fidelity and perceptual clarity. This indicates that Retinex assists in rectifying minor degradations that enhance GAN's generative ability. The introduction of BAM (Bilateral Attention Module) with Retinex, absent of GAN, enhances edge consistency and localised contrast (UIQM: 3.1), yet demonstrates limited efficacy in maintaining deeper structural integrity (SSIM: 0.76), thereby revealing constraints in learning global feature distributions. The GAN + BAM variant enhances contextual refinement and structural alignment (SSIM: 0.85, LPIPS: 0.118), underscoring the significance of attention in efficient feature selection

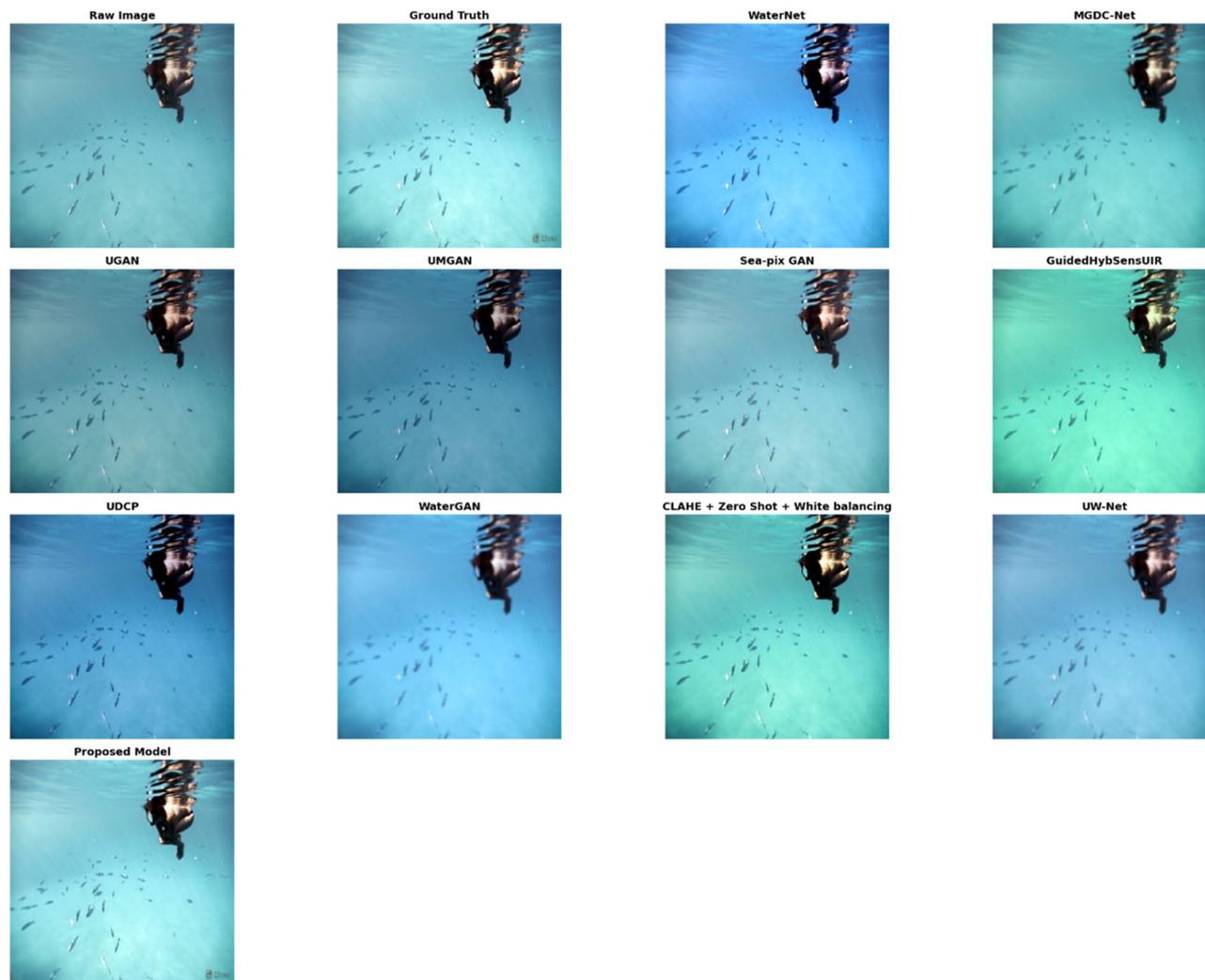


Fig. 9. Comparative analysis of underwater image enhancement techniques on UIEB dataset from left to right (a) Hazy underwater image, (b) ground truth image, (c) WaterNet⁶⁷ (d) MGDC-Net⁶⁸, (e) UGAN (Underwater GAN)⁵², (f) UMGAN (underwater multi-scene GAN)⁶⁹, (g) Sea-pix GAN⁷⁰, (h) GuidedHybSensUIR⁷¹, (i) UDCP⁷², (j) WaterGAN³⁴, (k) CLAHE + Zero Shot + White balancing⁷³ (l) UW-Net⁵⁷, proposed work.

Module	PSNR (dB)	SSIM	UIQM	LPIPS
GAN only	21.9	0.74	2.42	0.228
GAN + Retinex	23.1	0.77	2.65	0.198
GAN + BAM	24.3	0.79	2.81	0.172
Proposed complete model	28.4	0.88	3.7	0.082
CNN + BAM + Retinex	22.6	0.76	2.55	0.212
Retinex only	19.8	0.7	2.91	0.205

Table 4. Ablation study on each key variant on UIEB dataset.

during generation. The complete configuration (GAN + BAM + Retinex) surpasses all other combinations in every assessed metric (PSNR: 28.4, SSIM: 0.88, UIQM: 3.7, LPIPS: 0.082). This illustrates the integration of generative adversarial networks (GAN), attention-based feature augmentation (BAM), and physics-informed colour adjustment (Retinex). Each module uniquely enhances various facets of image restoration—Retinex for colour and illumination, BAM for spatial coherence, and GAN for generative authenticity. This confirms the necessity of the integrated design and emphasises its superiority for effective underwater image enhancement.

Discussion

The experimental assessment reveals that the suggested underwater image dehazing model—incorporating a Generative Adversarial Network (GAN), Bottleneck Attention Module (BAM), and Retinex-based enhancement—attains exceptional performance in both traditional and perceptual quality metrics. Regarding UIQM, our approach attains a maximum value of 3.7, surpassing traditional methods such as WaterNet (2.83) and UGAN (2.91), signifying a notable enhancement in the perception of underwater image quality. This is validated by the peak signal-to-noise ratio (PSNR) of 28.4 dB and a robust structural similarity index measure (SSIM) of 0.88, which together confirm the model's efficacy in noise reduction while maintaining structural integrity. The LPIPS comparison further substantiates the perceptual robustness of our model. Our approach achieved the minimal LPIPS score (0.082), indicating a significant perceptual resemblance to the ground truth images. This significantly surpasses advanced models like UWCNN (0.203), Shallow-UWNet (0.206), and Semi-UIR (0.12), suggesting that the incorporation of BAM and Retinex aids in maintaining essential perceptual attributes.

The ablation study underscores the importance of each module within the pipeline. When evaluated separately or in partial combinations, each module contributed incrementally to the overall performance. For example, the incorporation of BAM into a basic GAN significantly enhanced PSNR and diminished LPIPS, whereas the full integration of GAN, BAM, and Retinex produced optimal outcomes. The Retinex module's inclusion, while advantageous for contrast correction, did not yield competitive performance in isolation, underscoring the necessity of synergy among modules for effective dehazing. The proposed model achieves a balance between performance and efficiency regarding computational complexity. Boasting 36.9G MACs and 2.2 M parameters, it demonstrates superior computational efficiency compared to various GAN-based alternatives such as MGDC-Net (78.5G, 5.1 M) and UMGAN (70.3G, 4.4 M), thereby rendering it suitable for real-time implementation on edge devices. This compactness, attained without sacrificing quality, is especially advantageous for underwater robotics and marine monitoring applications. An intriguing observation is the discrepancy in MGDC-Net, which exhibits a high PSNR (27.81 dB) yet an exceedingly low SSIM (0.134). This indicates possible implementation challenges or metric misalignment, likely resulting from excessive contrast enhancement that undermines spatial coherence. It emphasises the significance of employing both traditional and perceptual metrics to guarantee equitable assessment.

Conclusion

This study introduces a novel and efficient underwater image dehazing framework using a Generative Adversarial Network (GAN), a Bottleneck Attention Module (BAM), and Retinex-based enhancement. The proposed architecture solves underwater problems like colour distortion, low contrast, and poor visibility. After extensive quantitative evaluation using traditional metrics (UIQM, PSNR, SSIM) and perceptual quality assessment (LPIPS), the model outperforms classical and state-of-the-art deep learning-based approaches. It has the highest UIQM (3.7), PSNR (28.4 dB), competitive SSIM (0.88), and lowest LPIPS (0.082), proving its visual appeal and structural coherence.

The ablation study emphasises the importance of each architecture component individually and collectively. Retinex and BAM alone are useful, but their combination with GAN improves colour fidelity and finer textures. Low computational footprint (36.9G MACs and 2.2 M parameters) makes the model suitable for resource-constrained environments like underwater autonomous vehicles, robotic systems, and edge-based marine monitoring. The evaluation highlighted the model's accuracy, perceptual realism, and efficiency. Our architecture is optimised for all-around performance, unlike other enhancement methods that sacrifice perceptual quality for high PSNR or computational costs for marginal gains. Competing methods, such as MGDC-Net's high PSNR but abnormally low SSIM, demonstrate the need for a variety of evaluation metrics, including perceptual ones, to assess quality.

Future work

Although the proposed model demonstrates robust performance, numerous avenues for future investigation persist. Initially, broadening the assessment to encompass supplementary underwater datasets with varied turbidity and illumination conditions can facilitate the validation of model generalisability. Secondly, the incorporation of adaptive learning methodologies, such as reinforcement learning or neural architecture search, could enhance performance in dynamic underwater environments. Integrating multi-scale or transformer-based attention mechanisms may improve global contextual comprehension. Furthermore, perceptual quality may be assessed utilising learning-based no-reference metrics such as the Underwater Ranker, which adapts dynamically to scene characteristics, thereby addressing the constraints of fixed-weight metrics like UIQM. Ultimately, real-time deployment and validation of underwater autonomous vehicles (UAVs/ROVs) would facilitate practical implementation in marine exploration, surveillance, and ecological research.

Data availability

Data has been taken from publicly available source: <https://paperswithcode.com/dataset/uieb>.

Received: 23 April 2025; Accepted: 14 July 2025

Published online: 18 July 2025

References

1. Saoud, L. S. et al. Seeing through the haze: A comprehensive review of underwater image enhancement techniques. *IEEE Access* **12**, 145206–145233. <https://doi.org/10.1109/ACCESS.2024.3465550> (2024).

2. Wibisono, A., Piran, M. J., Song, H. K. & Lee, B. M. A survey on unmanned underwater vehicles: challenges, enabling technologies, and future research directions. *Sensors* **23** (17), 1–29. <https://doi.org/10.3390/s23177321> (2023).
3. Elmezain, M. et al. Advancing underwater vision: A survey of deep learning models for underwater object recognition and tracking. *IEEE Access*. **13**, 17830–17867. <https://doi.org/10.1109/ACCESS.2025.3534098> (2025).
4. Sugunapriya & Markkandan, S. Studies on underwater image processing using artificial intelligence technologies. *IEEE Access*. **13**, 3929–3969. <https://doi.org/10.1109/ACCESS.2024.3524593> (2024).
5. Raihan, J., Abas, A. P. E. & De Silva, L. C. Review of underwater image restoration algorithms. *IET Image Proc.* **13** (10), 1587–1596 (2019).
6. Perez, J., Sanz, P. J., Bryson, M. & Williams, S. B. A benchmarking study on single image dehazing techniques for underwater autonomous vehicles. *OCEANS 2017 - Aberd.* <https://doi.org/10.1109/OCEANSE.2017.8084771> (2017).
7. Wang, H., Sun, S. & Ren, P. Underwater color disparities: cues for enhancing underwater images toward natural color consistencies. *IEEE Trans. Circuits Syst. Video Technol.* **34** (2), 738–753. <https://doi.org/10.1109/TCSVT.2023.3289566> (2024).
8. Kim, W. Y. et al. A modified single image dehazing method for autonomous driving vision system. **83** <https://doi.org/10.1007/s11042-023-16547-8> (Springer, 2024).
9. Fayaz, S. et al. Intelligent underwater object detection and image restoration for autonomous underwater vehicles. *IEEE Trans. Veh. Technol.* **73** (2), 1726–1735. <https://doi.org/10.1109/TVT.2023.3318629> (2024).
10. Zhuang, P. & Ding, X. Underwater image enhancement using an edge-preserving filtering retinex algorithm. *Multimedia Tools Appl.* **79** (25), 17257–17277. <https://doi.org/10.1007/s11042-019-08404-4> (2020).
11. Jiang, Q., Gu, Y., Li, C., Cong, R. & Shao, F. Underwater image enhancement quality evaluation: Benchmark dataset and objective metric. *IEEE Trans. Circuits Syst. Video Technol.* **32** (9), 5959–5974 <https://doi.org/10.1109/TCSVT.2022.3164918>. (2022).
12. Li, C., Guo, C., Ren, W. & Cong, R. An underwater image enhancement benchmark dataset and beyond. *IEEE Trans. Image Process.* **29**, 4376–4389 (2019).
13. Alenezi, F., Armghan, A. & Santosh, K. C. Underwater image dehazing using global color features. *Eng. Appl. Artif. Intell.* **116**, 105489. <https://doi.org/10.1016/j.engappai.2022.105489> (2022).
14. Ancuti, C., Ancuti, C. O., Haber, T. & Bekaert, P. Enhancing underwater images and videos by fusion. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* <https://doi.org/10.1109/CVPR.2012.6247661> (2012).
15. Fu, X. et al. A retinex-based enhancing approach for single underwater image. In *IEEE International Conference on Image Processing (ICIP)* 4572–4576 (IEEE, 2014).
16. Li, C. et al. Single underwater image enhancement based on color cast removal and visibility restoration. *J. Electron. Imaging*. **25** (3), 33012 (2016).
17. Wang, H., Sun, S., Chang, L. & Li, H. INSPIRATION: A reinforcement learning-based human visual perception-driven image enhancement paradigm for underwater scenes. *Eng. Appl. Artif. Intell.* **133**, 108411 (2024).
18. Yu, X., Qu, Y. & Hong, M. Underwater-GAN: Underwater image restoration via conditional generative adversarial network. In *International Conference on Pattern Recognition* 66–75 (Springer, 2018).
19. Guo, Y., Li, H. & Zhuang, P. Underwater image enhancement using a multiscale dense generative adversarial network. *IEEE J. Oceanic Eng.* **45** (3), 862–870 (2019).
20. Wang, H., Köser, K. & Ren, P. Large foundation model empowered discriminative underwater image enhancement. *IEEE Trans. Geosci. Remote Sens.* **63**, 1–17. <https://doi.org/10.1109/TGRS.2025.3525962> (2025).
21. Park, J., Woo, S., Lee, J. Y. & Kweon, I. S. Bam: Bottleneck attention module. *arXiv* (2018).
22. Tan, R. T. Visibility in bad weather from a single image. In *2008 IEEE Conference on Computer Vision and Pattern Recognition* 1–8 (IEEE, 2008).
23. Cheng, H. D. & Shi, X. J. A simple and effective histogram equalization approach to image enhancement. *Digit. Signal Proc.* **14** (2), 158–170 (2004).
24. Mohan, S. & Simon, P. Underwater image enhancement based on histogram manipulation and multiscale fusion. *Procedia Comput. Sci.* **171**, 941–950. <https://doi.org/10.1016/j.procs.2020.04.102> (2020).
25. Garg, D., Garg, N. K. & Kumar, M. Underwater image enhancement using blending of CLAHE and percentile methodologies. *Multimedia Tools Appl.* **77**, 26545–26561 (2018).
26. Zhou, J., Pang, L., Zhang, D. & Zhang, W. Underwater image enhancement method via multi-interval subhistogram perspective equalization. *IEEE J. Oceanic Eng.* **48** (2), 474–488 (2023).
27. Zhang, S., Wang, T., Dong, J. & Yu, H. Underwater image enhancement via extended multi-scale retinex. *Neurocomputing* **245**, 1–9 (2017).
28. Lin, S., Li, Z., Zheng, F., Zhao, Q. & Li, S. Underwater image enhancement based on adaptive color correction and improved retinex algorithm. *IEEE Access*. **11**, 27620–27630 (2023).
29. Vasamsetti, S., Mittal, N., Neelapu, B. C. & Sardana, H. K. Wavelet based perspective on variational enhancement technique for underwater imagery. *Ocean Eng.* **141**, 88–100. <https://doi.org/10.1016/j.oceaneng.2017.06.012> (2017).
30. Qiao, X., Bao, J., Zhang, H., Zeng, L. & Li, D. Underwater image quality enhancement of sea cucumbers based on improved histogram equalization and wavelet transform. *Inform. Process. Agric.* **4** (3), 206–213. <https://doi.org/10.1016/j.inpa.2017.06.001> (2017).
31. Patel, Z. et al. Framework for underwater image enhancement. *Proc. Comput. Sci.* **171**, 491–497 <https://doi.org/10.1016/j.procs.2020.04.052> (2019).
32. He, K., Sun, J. & Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **33** (12), 2341–2353. <https://doi.org/10.1109/TPAMI.2010.168> (2011).
33. Drews, P. L. J., Nascimento, E. R., Botelho, S. S. C. & Campos, M. F. M. Underwater depth Estimation and image restoration based on single images. *IEEE Comput. Graph. Appl.* **36** (2), 24–35 (2016).
34. Chen, Z., Wang, H., Shen, J., Li, X. & Xu, L. Region-specialized underwater image restoration in inhomogeneous optical environments. *Optik* **125** (9), 2090–2098 (2014).
35. Li, C. Y., Guo, J. C., Cong, R. M., Pang, Y. W. & Wang, B. Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior. *IEEE Trans. Image Process.* **25** (12), 5664–5677 (2016).
36. Peng, Y. T. & Cosman, P. C. Underwater image restoration based on image blurriness and light absorption. *IEEE Trans. Image Process.* **26** (4), 1579–1594 (2017).
37. Ahila Priyadarshini, R. & Ramajeyam, K. A combined approach of color correction and homomorphic filtering for enhancing underwater images. In *International Conference on Computational Intelligence in Pattern Recognition* pp. 475–487 (Springer, 2022).
38. Song, W., Wang, Y., Huang, D., Liotta, A. & Perra, C. Enhancement of underwater images with statistical model of background light and optimization of transmission map. *IEEE Trans. Broadcast.* **66** (1), 153–169 (2020).
39. Wang, Y., Zhang, J., Cao, Y. & Wang, Z. A deep CNN method for underwater image enhancement. In *IEEE International Conference on Image Processing (ICIP)* 1382–1386. (IEEE, 2017).
40. Mai, J., Zhu, Q., Wu, D., Xie, Y. & Wang, L. Back propagation neural network dehazing. In *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)* 1433–1438 (IEEE, 2014).
41. Tang, K., Yang, J. & Wang, J. Investigating haze-relevant features in a learning framework for image dehazing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 2995–3000 (2014).

42. Hussain, F. & Jeong, J. Visibility enhancement of scene images degraded by foggy weather conditions with deep neural networks. *J. Sens.* 3894832 (2016).
43. Cai, B., Xu, X., Jia, K., Qing, C. & Tao, D. DehazeNet: an end-to-end system for single image haze removal. *IEEE Trans. Image Process.* 25 (11), 5187–5198. <https://doi.org/10.1109/TIP.2016.2598681> (2016).
44. Li, C., Anwar, S. & Porikli, F. Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recogn.* 98, 107038. <https://doi.org/10.1016/j.patcog.2019.107038> (2020).
45. Li, C. et al. Underwater image enhancement via medium Transmission-Guided Multi-Color space embedding. *IEEE Trans. Image Process.* 30, 4985–5000. <https://doi.org/10.1109/TIP.2021.3076367> (2021).
46. Fu, Z., Lin, X., Wang, W., Huang, Y. & Ding, X. Underwater image enhancement via learning water type desensitized representations. In *ICASSP 2022–2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 2764–2768 (IEEE, 2022).
47. Hou, X. et al. DFFA-Net: A differential convolutional neural network for underwater optical image dehazing. *Electron. (Switzerland)*. 12 (18), 1–15. <https://doi.org/10.3390/electronics12183876> (2023).
48. Li, R., Pan, J., Li, Z. & Tang, J. Single image dehazing via conditional generative adversarial network. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* <https://doi.org/10.1109/CVPR.2018.00856> (2018).
49. Islam, M. J., Xia, Y. & Sattar, J. Fast underwater image enhancement for improved visual perception. *IEEE Rob. Autom. Lett.* 5 (2), 3227–3234. <https://doi.org/10.1109/LRA.2020.2974710> (2020).
50. Yang, M. et al. Underwater image enhancement based on conditional generative adversarial network. *Sig. Process. Image Commun.* 81, 115723. <https://doi.org/10.1016/j.image.2019.115723> (2020).
51. Jiang, Z., Li, Z., Yang, S., Fan, X. & Liu, R. Target oriented perceptual adversarial fusion network for underwater image enhancement. *IEEE Trans. Circuits Syst. Video Technol.* 32 (10), 6584–6598. <https://doi.org/10.1109/TCSVT.2022.3174817> (2022).
52. Fabbri, C., Islam, M. J. & Sattar, J. Enhancing underwater imagery using generative adversarial networks. *Proceedings - IEEE International Conference on Robotics and Automation* 7159–7165 <https://doi.org/10.1109/ICRA.2018.8460552> (2018).
53. Zhu, J. Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision* 2223–2232 (2017).
54. Li, J., Skinner, K. A., Eustice, R. M. & Johnson-Roberson, M. WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Rob. Autom. Lett.* 3 (1), 387–394 (2017).
55. Ye, X., Li, Z., Sun, B., Wang, Z. & Xu, R. Deep joint depth Estimation and color correction from monocular underwater images based on unsupervised adaptation networks. *IEEE Trans. Circuits Syst. Video Technol.* 30 (11), 3995–4008 (2019).
56. Wang, N., Zhou, Y., Han, F., Zhu, H. & Yao, J. UWGAN: Underwater GAN for real-world underwater color restoration and dehazing. *arXiv* (2019).
57. Awan, H. S. A. & Mahmood, M. T. Underwater image restoration through color correction and UW-Net. *Electronics* 13 (1), 199 (2024).
58. Jiang, Q. et al. Two-step domain adaptation for underwater image enhancement. *Pattern Recogn.* 122, 108324. <https://doi.org/10.1016/j.patcog.2021.108324> (2022).
59. Li, F., Li, X., Peng, Y., Li, B. & Zhai, Y. Maximum information transfer and minimum loss dehazing for underwater image restoration. *IEEE J. Oceanic Eng.* 49 (2), 622–636. <https://doi.org/10.1109/JOE.2023.3334478> (2024).
60. Wang, H., Köser, K. & Ren, P. Large foundation model empowered discriminative underwater image enhancement. *IEEE Trans. Geosci. Remote Sens.* (2025).
61. Li, C., Guo, C. & Loy, C. C. Learning to enhance Low-Light image via Zero-Reference deep curve Estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* 44 (8), 4225–4238. <https://doi.org/10.1109/TPAMI.2021.3063604> (2022).
62. Wu, W. et al. URetinex-Net : Retinex-based deep unfolding network for low-light image enhancement. 5901–5910.
63. Jiang, Y., Gong, X., Liu, D., Cheng, Y. & Fang, C. EnlightenGAN: deep light enhancement without paired supervision. *IEEE Trans. Image Process.* 30, 2340–2349. <https://doi.org/10.1109/TIP.2021.3051462> (2021).
64. Parihar, A. S. & Singh, K. A study on Retinex based method for image enhancement. *Proceedings of the 2nd International Conference on Inventive Systems and Control* pp. 619–624 <https://doi.org/10.1109/ICISC.2018.8398874> (ICISC, 2018).
65. Bakht, A. B., Gia, Z., Din, M. & Akram, W. MuLA-GAN: Multi-level attention GAN for enhanced underwater visibility ecological informatics. 81, 102631 (2024).
66. Hepburn, A., Laparra, V., Malo, J., McConville, R. & Santos-Rodriguez, R. Perceptnet: A human visual system inspired neural network for estimating perceptual distance. In *IEEE International Conference on Image Processing (ICIP)* pp. 121–125 (IEEE, 2020).
67. Li, C., Guo, C., Ren, W. & Cong, R. An underwater image enhancement benchmark dataset and beyond. *IEEE Trans. Image Process.* 29, 4376–4389. <https://doi.org/10.1109/TIP.2019.2955241> (2020).
68. Yan, J., Hu, H., Wang, Y., Nawaz, M. W. & Rehman, N. U. Underwater image enhancement via multiscale disentanglement strategy. *Sci. Rep.* 15 (1), 6076. <https://doi.org/10.1038/s41598-025-89109-7> (2025).
69. Zou, G. et al. UMGAN: multi-scale graph attention network for grid parameter identification. *Electr. Eng.* pp. 1–14 (2024).
70. Chaurasia, D. & Chhikara, P. Sea-Pix-GAN: underwater image enhancement using adversarial neural network. *J. Vis. Commun. Image Represent.* 98, 1–10. <https://doi.org/10.1016/j.jvcir.2023.104021> (2024).
71. Guo, X., Chen, X., Wang, S. & Pun, C. Underwater image restoration through a prior guided hybrid sense approach and extensive benchmark analysis. 14 (8), 1–17 (2021).
72. Drews, P., Nascimento, E., Moraes, F., Botelho, S. & Campos, M. Transmission estimation in underwater single images. In *Proceedings of the IEEE International Conference on Computer Vision Workshops* 825–830 (2013).
73. Murugan, T. K., Sharma, S., Ganguly, A., Banerjee, A. & Kejrival, K. An enhanced Multi-stage approach for dehazing underwater images. *IEEE Access* 12, 156803–156822. <https://doi.org/10.1109/ACCESS.2024.3486456> (2024).

Author contributions

Amandeep Kaur: Writing—Original draft, Visualization, Methodology, Validation. Shalli Rani: Conceptualization, Data Curation, Formal analysis, Resources. Ayush Dogra: Data Curation, Formal analysis, Resources, Supervision. Mohammad Shabaz: Validation, Investigation, Data Curation, Writing—editing and reviewing, Supervision, Project administration.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025