

Restoring underwater images from turbidity and motion blur: A three-step framework

Fatima Iqbal^{1,3}, Behçet Uğur Töreyn^{2,3}

¹ Department of Computer Sciences, Informatics Institute, Istanbul Technical University, Turkey, ² Faculty of Computer and Informatics Engineering, Dept. of AI and Data Engineering, Istanbul Technical University, Turkey, ³ SP4CING - Signal Processing for Computational Intelligence Research Group

Corresponding author: Fatima Iqbal, iqbal23@itu.edu.tr

Underwater image restoration is challenging due to the pervasive and simultaneous presence of multiple degradations, including turbidity, uneven illumination, color distortion, and motion blur. Existing approaches typically address individual degradation types and therefore fail to generalize to complex real-world underwater conditions where these artifacts coexist. To overcome this limitation, we propose a novel three-stage sequential framework for progressive underwater image enhancement. First, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to mitigate uneven illumination and enhance local contrast. The contrast-enhanced image is then processed by an adaptive color correction algorithm that estimates the dominant color channel intensity and compensates for color casts across the red, green, and blue channels. Finally, the refined image is passed to a data-driven diffusion-based restoration model trained using noise parameters derived from a unique turbidity-graded laboratory dataset. This dataset enabled the model to predict and quantify turbidity-related noise and reconstruct an enhanced underwater image. The proposed framework is evaluated on the Underwater Image Enhancement Benchmark (UIEB) and the Turbid Underwater Dataset (TUDS), where quantitative and perceptual results demonstrate consistent improvements at each stage of the pipeline. In particular, we report improved quantitative performance in terms of Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) compared with recent enhancement methods on a paired subset of the UIEB, indicating effective structural and perceptual reconstruction.

Keywords: Diffusion models, image enhancement, underwater image quality assessment, underwater image restoration

1. INTRODUCTION

Water covers approximately 70% of the Earth's surface, yet vast portions remain unexplored. Unlike air, water significantly degrades captured images, causing turbidity, color distortion and uneven illumination. These challenges limit the effectiveness of imaging systems in marine research and biodiversity monitoring. Enhancing underwater image quality is essential not only for human perception but also for autonomous underwater vehicles [1] and remote sensing applications [2] that rely on clear visual data. Previous studies in underwater image enhancement [3] and restoration have made significant progress, yet several challenges remain. Improving underwater image quality can generally be divided into two main categories:

- **Image restoration:** recovers the true scene by correcting degradations like noise, scattering, and color distortion.
- **Image enhancement:** improves visual quality for human perception or machine analysis without relying on physical models.

Image enhancement and restoration in underwater environments are hindered by visual challenges such as severe turbidity and color distortions, as well as the scarcity of large-scale real-world datasets. While recent studies have applied domain adaptation to bridge the gap between synthetic and real underwater data, these methods still struggle

under extreme conditions. In addition, conventional image quality metrics often fail to capture human perception or align with high-level tasks like object detection, emphasizing the need to better understand and address the unique visual challenges present in underwater imagery.

1.1 Visual challenges in underwater imagery

1.1.1 Turbidity

Suspended particles in water scatter and absorb light, resulting in turbidity. This phenomenon produces hazy, blurred, and color-distorted images, significantly reducing visibility and hindering feature extraction. Sources of turbidity include sediments (silt, clay), organic matter, algae, dissolved compounds, and microscopic organisms. High turbidity not only obscures underwater scenes but also impacts marine ecosystems by limiting light penetration, reducing photosynthesis in aquatic plants, and affecting coral reefs and fish populations. From an imaging perspective, turbidity diminishes contrast, sharpness, and color fidelity, complicating accurate data capture and often necessitating artificial lighting or computational restoration methods. Fig. 1(a) illustrates the visual degradation caused by turbidity, highlighting the challenges it poses for underwater image processing.

1.1.2 Low-light conditions

Underwater environments are frequently characterized by poor illumination due to light absorption and scattering by water. As depth increases, red and yellow wavelengths are absorbed more rapidly, leaving predominantly blue-green light. Insufficient illumination results in low-contrast images, color shifts, and increased noise, reducing the clarity of objects and textures. These factors pose significant challenges for both visual inspection and automated image analysis. Fig. 1(b) demonstrates the effects of low-light conditions on underwater imagery.

1.1.3 Motion blur

Motion blur occurs due to camera movement, water currents, or the motion of subjects such as fish and suspended particles. Even minor displacements during exposure can produce smearing, reducing sharpness and obscuring features. Motion blur complicates tasks such as object detection, segmentation, and quantitative measurement of underwater structures or organisms. The impact of motion blur is exemplified in Fig. 1(c).

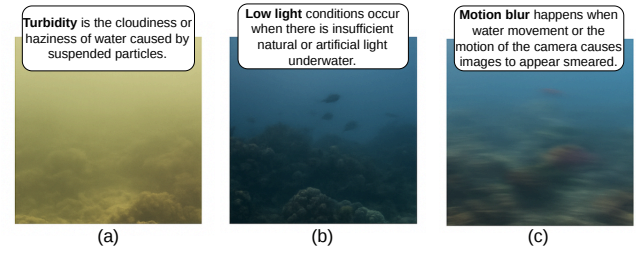


Figure 1 – Representative examples of common underwater image degradations: (a) turbidity causing haziness and reduced visibility, (b) low-light conditions resulting in dark and low-contrast images, and (c) motion blur due to camera or water movement, reducing clarity and sharpness.

2. LITERATURE REVIEW

2.1 Physics-based and model-driven methods

Physics-based underwater image restoration methods rely on modeling the underwater image formation process. The Underwater Image Formation Model (UIFM) [4] describes wavelength-dependent attenuation and backscattering effects. Techniques such as Dark Channel Prior (DCP) [5] and Red Channel Prior (RCP) [6] estimate transmission maps to compensate for color distortion and haze. Depth-dependent methods, including color attenuation models [7] and polarization-based approaches [8], attempt to separate direct and scattered light components.

Although these approaches provide physically interpretable solutions, their performance depends heavily on accurate estimation of water properties and scene depth, limiting robustness in highly turbid or spatially varying underwater environments.

2.2 CNN-based methods

Convolutional Neural Networks (CNNs) learn nonlinear mappings between degraded and enhanced images without explicitly modeling the physical formation process. Architectures such as VGG-based networks [9], ResNet-based models [10], and encoder-decoder structures have been widely adopted for underwater enhancement.

These models are typically trained using paired datasets and optimized using combinations of pixel-wise loss, perceptual loss, and structural similarity constraints [11]. CNN-based methods improve contrast, color balance, and detail restoration, but may struggle with severe scattering and complex degradation patterns.

2.3 GAN-based methods

Generative Adversarial Networks (GANs) enhance underwater images by introducing adversarial training to

encourage photorealistic outputs [12, 13]. The generator learns to restore degraded images while the discriminator enforces realism, improving texture details and visual plausibility.

GAN-based approaches have demonstrated strong performance in perceptual quality, particularly in challenging underwater scenarios. However, adversarial instability and training complexity may affect consistency and structural preservation.

2.4 Transformer and state-space-based methods

Recent approaches incorporate attention mechanisms and state-space models [14] to capture long-range dependencies and global contextual information. Transformer-based and visual state-space models, such as WaterMamba [15], improve feature representation and global color correction through enhanced contextual modeling.

These models address limitations of local receptive fields in CNN architectures and show promising results in complex underwater scenes.

2.5 Diffusion-based methods

Diffusion models have recently emerged as powerful generative frameworks for image restoration. By modeling a gradual denoising process, diffusion-based methods iteratively refine degraded images, improving structural fidelity and realism. Recent work applies diffusion models specifically to underwater image enhancement, demonstrating improved detail preservation and noise suppression.

Unlike GAN-based approaches, diffusion models offer stable training and explicit noise modeling, making them well-suited for severe turbidity correction and fine-detail restoration.

2.6 Novelty of the proposed method

Diffusion-based underwater image enhancement methods typically employ standard diffusion formulations in which noise is sampled from predefined stochastic schedules during training and inference. These approaches rely on generic noise distributions without explicitly incorporating degradation statistics derived from real underwater environments.

The proposed method differs in two key aspects:

- **Structured three-stage framework:** The method is

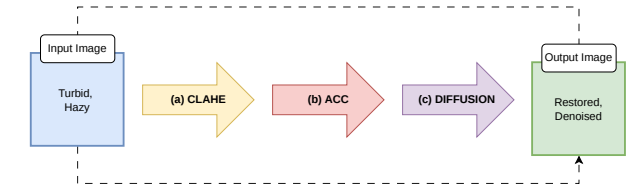


Figure 2 – Processing pipeline illustrating our proposed model to enhance both turbid and hazy images through three distinct steps: (a) Contrasting Limited Adaptive Histogram Equalization (CLAHE), (b) Adaptive Color Correction (ACC) and, finally, while utilizing a (c) diffusion restoration method to obtain an image that has been both restored and denoised.

embedded within a sequential three-stage architecture, where initial preprocessing and intermediate enhancement progressively refine the input prior to diffusion-based reconstruction. The contribution of each stage is summarized in Table 1, and the processing pipeline is illustrated in Fig. 2.

- **Dataset-aware diffusion modeling:** Instead of relying solely on randomly sampled noise schedules, the diffusion stage leverages noise-related parameters extracted from a dataset containing images with graded turbidity levels. This enables the refinement process to reflect realistic underwater degradation patterns.

3. PROPOSED METHODOLOGY

Table 1 – Explanation of enhancement and restoration techniques in the proposed framework

Technique	Main Purpose	Works On	Type
CLAHE	Enhances local contrast and corrects uneven illumination in small regions	Local pixel neighborhoods	Local enhancement (nonlinear histogram-based)
Adaptive Color Balancing (ACB)	Equalizes global brightness and color distribution across RGB channels	Whole image (global adjustment)	Global enhancement (linear and adaptive over time)
Diffusion Model	Restores complex degradations such as turbidity, blur, color distortion, and noise	Entire image (learned model)	Deep generative restoration

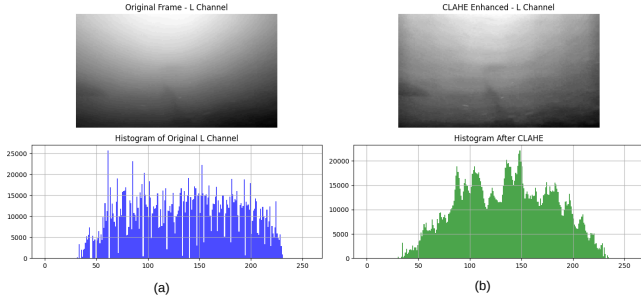


Figure 3 – Comparison of the luminance channel (L-channel) image and its corresponding histogram: (a) before and (b) after CLAHE enhancement.

3.1 Contrast-Limited Adaptive Histogram Equalization (CLAHE)

In the first step of our restoration framework, we implement Contrast-Limited Adaptive Histogram Equalization (CLAHE) [16]. Unlike global histogram equalization, which operates on the entire image and often washes out local details, CLAHE divides the image into non-overlapping blocks or contextual regions. Within each region, the histogram of pixel intensities is computed and equalized independently.

Formally, let an image be divided into $M \times N$ blocks. For each block, the histogram $H(i)$ is clipped at a maximum value H_{clip} , and the Cumulative Density Function (CDF) is computed as:

$$CDF(i) = \sum_{j=0}^i H_{clipped}(j), \quad (1)$$

where $H_{clipped}(j)$ is the clipped histogram count for intensity level j . Each pixel intensity is then mapped according to its local normalized CDF, and bilinear interpolation is applied between neighboring blocks to prevent abrupt transitions.

By enhancing local contrast while limiting noise amplification, CLAHE brightens dark regions and preserves structural details. This makes it particularly suitable as the first stage in our multi-step framework, providing a strong foundation for subsequent adaptive color correction and diffusion-based restoration. Fig. 3 shows a comparison of the luminance channel and its histogram before and after CLAHE enhancement.

3.2 Adaptive Color Correction (ACC)

The second step in our framework addresses severe color degradation caused by wavelength-dependent absorption and scattering in underwater environments. An Adaptive Color Correction (ACC) technique [17] is ap-

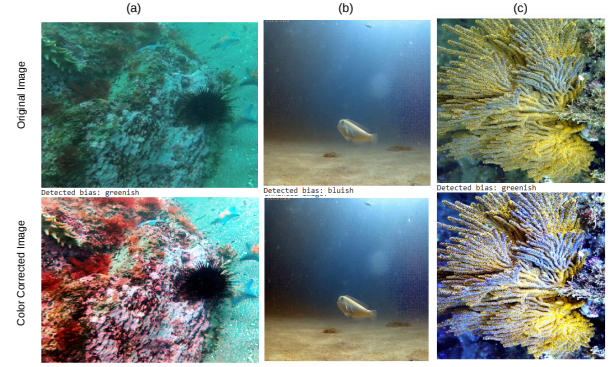


Figure 4 – Three underwater scenes were subjected to adaptive color correction. The technique in (a) restores the absent warm tones and lessens the severe greenish hue. The algorithm in (b) adjusts for the bluish bias brought on by light absorption in deeper water. In (c), it maintains the coral's original texture while balancing the greenish, yellowish tint. In every instance, the ACC method automatically detects the predominant color bias and generates an output that is more visually balanced and natural

plied after contrast enhancement to restore natural color appearance.

The approach first classifies underwater images into four categories, *bluish*, *greenish*, *yellowish*, and *color-balanced*, based on the dominant color in the estimated background region as shown in Fig. 4. The background region is identified using a Quadtree-based search, which recursively partitions the image into subregions and selects the area with the highest brightness and lowest intensity variance as an approximation of the scene's illumination.

For greenish images, histogram normalization balances the gray-level distributions of the three channels. Let $I_c(x)$ denote the normalized pixel value (in $[0, 1]$) of channel $c \in \{r, g, b\}$ at position x . The gray value distribution is normalized according to its cumulative histogram to ensure uniformity among channels.

For bluish images, the red channel suffers the most attenuation. Compensation is applied using the green channel as a reference, following the gray-world assumption. The corrected red channel $I_r(x)$ is computed as:

$$I_r(x) = I_r(x) + \alpha (\bar{I}_g - \bar{I}_r) (1 - I_r(x)) I_g(x), \quad (2)$$

where $\bar{I}_r$ and $\bar{I}_g$ denote the mean pixel values of the red and green channels within the background region, and α is a tunable parameter set to 1. After red-channel compensation, gain adjustment is applied to all three channels to expand their dynamic range:

$$I'_c(x) = K_c (I_c(x) - I_{c,min}), \quad (3)$$

where

$$K_c = \frac{255}{D_c}, \quad D_c = I_{c,max} - I_{c,min}, \quad (4)$$

and $I_{c,min}$, $I_{c,max}$ are the minimum and maximum pixel values of channel c after compensation. Here, K_c is the

gain factor that scales the channel to use the full 0–255 intensity range, ensuring enhanced contrast and color distribution.

For yellowish images, the red channel serves as a reference to compensate the green and blue channels:

$$\begin{aligned} I_{g_c}(x) &= I_g(x) + \alpha (\bar{I}_r - \bar{I}_g) (1 - I_g(x)) I_r(x), \\ I_{b_c}(x) &= I_b(x) + \beta (\bar{I}_r - \bar{I}_b) (1 - I_b(x)) I_r(x), \end{aligned} \quad (5)$$

where α and β are tunable parameters set to 1.

Images where the mean values of all three channels in the background region are approximately equal are considered color-balanced and are passed directly to the next enhancement stage without additional correction.

This adaptive strategy effectively compensates for wavelength-dependent attenuation and restores natural color across diverse underwater scenes. Fig. 5 illustrates the framework and the sequential enhancement process.

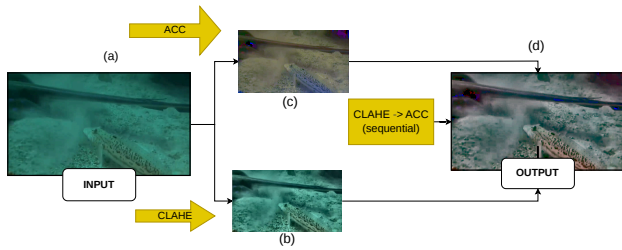


Figure 5 – Illustration of the Adaptive Color Correction (ACC) framework. (a) Input underwater image. (b) CLAHE enhances local contrast. (c) ACC classifies images into greenish, bluish, or yellowish categories and performs color compensation. (d) Sequential CLAHE + ACC enhancement produces the restored image with improved color balance and visual clarity.

3.3 Diffusion models

The final stage of the proposed framework employs a data-driven diffusion model to restore underwater

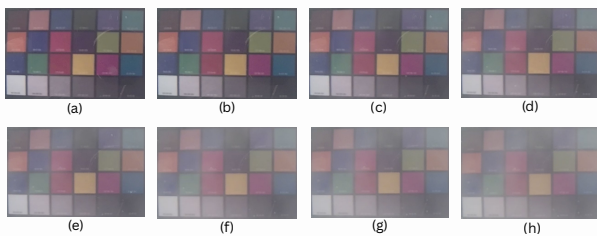


Figure 6 – From (a) to (h), images from the dataset demonstrate progressively increasing turbidity, mimicking the forward diffusion noise addition process. The dataset [18] was generated in a controlled laboratory setup using sand suspended in water to replicate realistic turbid conditions encountered in natural aquatic environments.

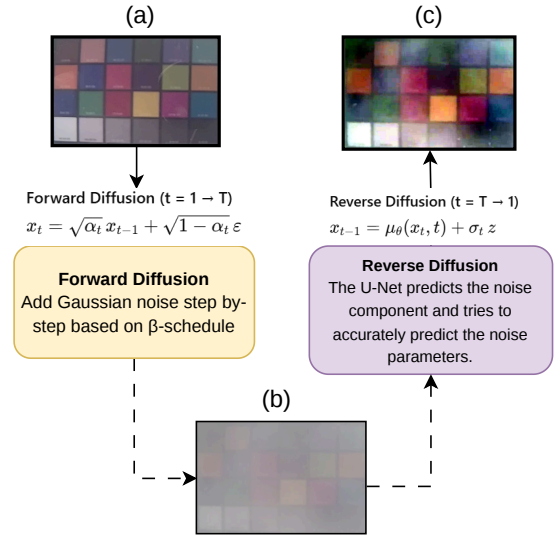


Figure 7 – Flowchart of the proposed diffusion framework. In the forward process ($a \rightarrow b$), Gaussian noise is progressively added according to the β -schedule, $x_t = \sqrt{\alpha_t} x_{t-1} + \sqrt{1 - \alpha_t} \epsilon$. In the reverse process ($b \rightarrow c$), the network estimates the mean (μ_t) and variance (σ_t^2) to iteratively denoise the image using $x_{t-1} = \mu_\theta(x_t, t) + \sigma_t z$.

images using a learned variance schedule [19]. The model consists of two processes: a forward degradation process that models turbidity-induced corruption, and a reverse process that learns to recover the clean image from noisy observations.

3.3.1 Noise parameter estimation from real underwater data

Let the sequence of images captured under progressively increasing turbidity levels (see Fig. 6) be denoted as

$$\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T,$$

where $\mathbf{x}_t \in \mathbb{R}^{H \times W \times C}$ represents the underwater image at turbidity level t , H , W , and C denote the image height, width, and number of channels, respectively, and $\mathbf{x}_0$ corresponds to the clear reference image.

3.3.1.1 Step-wise noise extraction:

The incremental degradation between consecutive turbidity levels is modeled as

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{n}_t, \quad (6)$$

where $\mathbf{n}_t$ denotes the turbidity-induced noise between levels t and $t + 1$.

The empirical noise is computed pixel-wise as

$$\mathbf{n}_t = \mathbf{x}_{t+1} - \mathbf{x}_t, \quad (7)$$

with subtraction applied across all spatial locations and channels.

3.3.1.2 Statistical estimation:

For each diffusion step t , the mean and variance of the extracted noise are estimated over all $N = H \times W \times C$ pixels:

$$\mu_t = \frac{1}{N} \sum_{i=1}^N n_t^{(i)}, \quad (8)$$

$$\sigma_t^2 = \frac{1}{N} \sum_{i=1}^N (n_t^{(i)} - \mu_t)^2, \quad (9)$$

where $n_t^{(i)}$ denotes the i -th pixel value of the flattened noise tensor $\mathbf{n}_t$, μ_t represents the empirical mean, and σ_t^2 represents the empirical variance of the turbidity-induced degradation at timestep t .

3.3.1.3 Beta schedule construction:

The diffusion variance schedule is constructed directly from the estimated noise statistics:

$$\beta_t = \sigma_t^2, \quad t \in \{1, 2, \dots, T\}, \quad (10)$$

where β_t denotes the variance parameter at timestep t .

The corresponding signal retention coefficient is defined as

$$\alpha_t = 1 - \beta_t. \quad (11)$$

This results in a data-driven diffusion process in which the noise magnitude at each timestep reflects physically observed underwater scattering patterns rather than synthetic stochastic perturbations.

3.3.1.4 Noise prediction objective:

During training, the reverse process learns to predict the empirically observed noise:

$$\hat{\mathbf{n}}_t = \epsilon_\theta(\mathbf{x}_t, t), \quad (12)$$

where $\epsilon_\theta(\cdot)$ denotes the neural network parameterized by θ , and $\hat{\mathbf{n}}_t$ is the predicted noise at timestep t .

The model is optimized using the mean squared error objective:

$$\mathcal{L} = \mathbb{E}_{t, \mathbf{x}_t} [\|\hat{\mathbf{n}}_t - \mathbf{n}_t\|_2^2], \quad (13)$$

where $\mathbb{E}_{t, \mathbf{x}_t}$ denotes expectation over sampled timesteps and degraded images, and $\|\cdot\|_2^2$ represents the squared ℓ_2 norm.

3.3.2 Forward diffusion for synthetic dataset generation

Using the estimated noise parameters, forward diffusion generates a synthetic dataset with progressively increasing turbidity. Multiple noisy versions are created from each clean image, exposing the reverse diffusion model to diverse noise patterns. This generated dataset enables effective learning of real-world underwater noise distributions (Fig. 7a–b).

3.3.3 Reverse diffusion for image denoising

The reverse diffusion model reconstructs the clean image $\mathbf{x}_0$ from a noisy image $\mathbf{x}_t$. At each timestep, the model predicts the denoised mean $\mu_\theta(\mathbf{x}_t, t)$ and variance $\sigma_\theta^2(t)$, iteratively removing noise according to:

$$\mathbf{x}_{t-1} = \mu_\theta(\mathbf{x}_t, t) + \sigma_\theta(t) \mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(0, \mathbf{I}). \quad (14)$$

This process recovers structural details while maintaining natural textures. A U-Net architecture with skip connections is used for the reverse model to preserve the fine details of the input image (Fig. 7b–c).

4. IMPLEMENTATION DETAILS

The proposed diffusion model is trained using the noise prediction objective described in Section 3.3. To improve computational efficiency, the precomputed coefficients $\bar{\alpha}_t$ and β_t are used during both sampling and optimization.

During training, input images are resized to 256×256 and normalized to the range $[-1, 1]$. Random horizontal flipping is applied for data augmentation.

The model is optimized using the Adam optimizer with an initial learning rate of 1×10^{-4} , batch size of 16, and 20 training epochs. The experiments are conducted in a

GPU-accelerated computing environment.

The forward diffusion process follows the data-driven variance schedule β_t derived from laboratory turbidity statistics. At each timestep, spatially correlated turbidity noise is generated by sampling from a uniform distribution in $[-1, 1]$ and applying Gaussian smoothing with standard deviation $\sigma = 10$. The generated noise is subsequently normalized to match the empirically estimated variance β_t , ensuring consistency with the observed degradation distribution.

5. RESULTS

The proposed framework was implemented and tested on two different datasets exhibiting varying levels of turbidity and degradation. In the next section, we will discuss the quantitative and qualitative results obtained and the evaluation metrics used. Figures 8 and 9 represent the sequential results for each respective dataset.

5.1 Underwater Image Enhancement Benchmark Dataset (UIEBD)

The Underwater Image Enhancement Benchmark Dataset (UIEBD) [20] contains underwater images with varying degrees of color degradation, including greenish, yellowish, and bluish haze. Fig. 8 illustrates the progressive outputs of our enhancement pipeline. For images with a dominant greenish hue, such as (a) and (d), the combined use of Contrast Limited Adaptive Histogram Equalization (CLAHE) and Adaptive Color Correction (ACC) effectively mitigated the green illumination bias. In more turbid cases, represented by (b) and (c), CLAHE substantially improved local contrast and revealed hidden scene details. The final stage, integrating CLAHE, ACC, and the diffusion-based restoration model, consistently produced the most visually-balanced and perceptually-enhanced results across all images.

5.2 TUDS dataset

To assess performance under highly turbid conditions, the proposed method was evaluated on the TUDS dataset [18], which contains real-time underwater imagery with moving fish, introducing challenges such as motion blur and low-light conditions. For images like Fig. 9(d), the pipeline effectively corrected color casts, while the diffusion-based restoration model further enhanced visual quality by suppressing noise and restoring fine details. These improvements demonstrate the method's robustness in complex underwater environments and its potential for practical applications in underwater imaging and analysis.

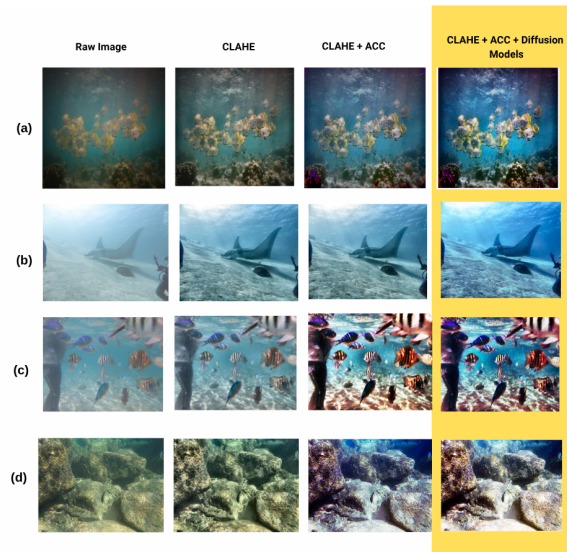


Figure 8 – Comparison of underwater image enhancement results on UIEBD [20]. From left to right: Raw image, CLAHE enhancement, CLAHE + Adaptive Color Correction (ACC), and the proposed sequential CLAHE + ACC + Diffusion model framework. Rows (a)–(d) show different underwater scenes and the improvement in contrast, color balance, and visual clarity achieved by the proposed method.

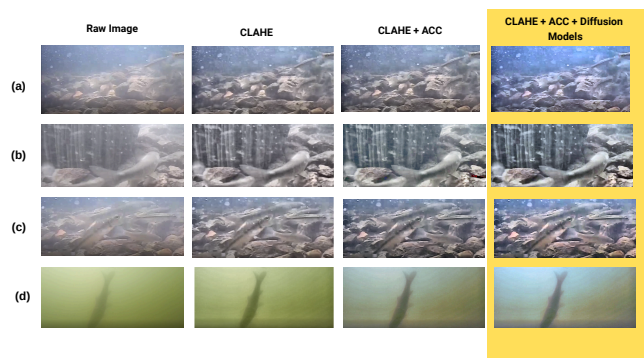


Figure 9 – Underwater image enhancement results on TUDS [18]. Columns represent the stages: Raw image, CLAHE, CLAHE + Adaptive Color Correction (ACC), and the full pipeline with CLAHE + ACC + Diffusion-based restoration model. Rows (a)–(d) depict images with varying turbidity and color casts, demonstrating improved color contrast and detail recovery at each enhancement stage. The final stage produces the most visually-balanced and perceptually-enhanced results.

6. EVALUATION METRICS

For underwater image restoration, objective evaluation remains challenging due to the absence of corresponding ground truth references in most real-world datasets. However, the paired subset of the UIEB dataset provides reference images, enabling full-reference evaluation. Therefore, PSNR and SSIM are computed on the UIEBD paired subset, while no-reference and perceptual metrics are used for unpaired datasets such as TUDS.

This study further employs perceptual and statistical quality measures, including Entropy ($H(X)$), the Naturalness Image Quality Evaluator (NIQE), Fréchet Inception Distance (FID), and Kernel Inception Distance (KID), to quantitatively assess restoration performance.

6.1 Metric definitions

(a) Peak Signal-to-Noise Ratio (PSNR): PSNR measures pixel-level fidelity between restored and reference images and is defined as:

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right), \quad (15)$$

where MAX is the maximum possible pixel value and MSE is the mean squared error between restored and reference images. Higher PSNR indicates better reconstruction quality.

(b) Structural Similarity Index Measure (SSIM): SSIM evaluates structural similarity between restored and reference images by comparing luminance, contrast, and structure:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (16)$$

where μ , σ , and σ_{xy} denote mean, variance, and covariance terms. Higher SSIM values indicate greater structural similarity.

(c) Entropy ($H(X)$): Entropy [21] quantifies the amount of information retained in an image and is defined as:

$$H(X) = - \sum_i P(i) \log_2 P(i), \quad (17)$$

where $P(i)$ denotes the normalized histogram of pixel intensities. Higher entropy values indicate richer image details and enhanced sharpness obtained through progressive restoration.

(d) Naturalness Image Quality Evaluator (NIQE): NIQE [22] assesses perceptual image quality by quantifying the deviation of natural scene statistics from those of the test image:

$$NIQE = \sqrt{(f - \mu)^T \Sigma^{-1} (f - \mu)}, \quad (18)$$

where f is the feature vector, μ is the mean, and Σ is the covariance matrix of natural image features. Lower NIQE values indicate a closer resemblance to natural scene statistics.

(e) Fréchet Inception Distance (FID): FID measures the similarity between feature distributions of restored and reference natural images:

$$FID = \|\mu_r - \mu_g\|_2^2 + \text{Tr}(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2}), \quad (19)$$

where (μ_r, Σ_r) and (μ_g, Σ_g) represent the mean and covariance of features extracted from real and generated image sets, respectively. Lower FID scores indicate a closer alignment with natural image distributions.

(f) Kernel Inception Distance (KID): KID measures distributional similarity using polynomial kernel statistics:

$$KID = \mathbb{E}[k(f_r, f_r')] + \mathbb{E}[k(f_g, f_g')] - 2\mathbb{E}[k(f_r, f_g)], \quad (20)$$

where $k(\cdot, \cdot)$ denotes the kernel function, and expectations are computed over feature pairs from real (f_r) and generated (f_g) images. Lower values correspond to improved perceptual quality.

6.2 Comparison with existing methods

We compare the proposed method against representative recent approaches, including both diffusion-based and non-diffusion frameworks. It should be noted that prior work such as Osmosis [23] reports quantitative evaluation on synthetic underwater images generated from the NYU-v2 RGBD dataset, whereas our evaluation is conducted on the UIEB paired subset. Therefore, direct numerical comparison should be interpreted with caution due to differences in dataset characteristics, degradation modeling, and evaluation protocols.

On the UIEB paired subset, Water-Net [20] and the Fusion-based method [24] achieve PSNR values of 19.11 dB and 18.75 dB, with SSIM scores of 0.79 and 0.81, respectively. WaterMamba [15] attained 24.72 dB PSNR and 0.93 SSIM, while Diffusion-UIE [25] achieves 25.44 dB PSNR and 0.94 SSIM. The partial fine-tuning variant of Diffusion-UIE further improves the PSNR to 26.88 dB, although the SSIM decreases to 0.78. The proposed hybrid framework (CLAHE + ACC + Diffusion) achieves 26.12 dB PSNR and 0.91 SSIM, demonstrating a favorable balance between reconstruction fidelity and structural preservation. Quantitative results evaluated using PSNR and SSIM are summarized in Table 2.

It can be noted that although PSNR and SSIM are widely-adopted metrics, they do not fully reflect perceptual quality in underwater imaging scenarios. The complex effects

of scattering, absorption, and color shift are not entirely captured by pixel-wise similarity measures. Therefore, the numerical improvements should be interpreted in conjunction with qualitative visual analysis to provide a more comprehensive assessment of restoration performance.

Table 2 – Quantitative comparison with representative underwater image enhancement methods on the UIEBD paired subset. Bold values denote the highest PSNR and SSIM obtained across these methods.

Methods	PSNR ↑	SSIM ↑
Water-Net [20]	19.11	0.79
Fusion-based [24]	18.75	0.81
WaterMamba [15]	24.72	0.93
Diffusion-UIE [25]	25.44	0.94
Diffusion-UIE (Fine-tuning) [25]	26.88	0.78
Proposed (CLAHE+ACC+Diffusion)	26.12	0.91

6.3 No-reference evaluation

Table 3 – Performance metrics for underwater image restoration on UIEBD and TUDS datasets across sequential enhancement stages

Dataset	CLAHE	CLAHE+ACC	Diffusion
Entropy (H(X))			
UIEBD	6.9	7.5	7.8
TUDS	6.5	6.9	7.1
NIQE			
UIEBD	3.1	2.6	2.0
TUDS	3.5	3.0	2.8
FID			
UIEBD	18.2	12.3	9.4
TUDS	22.4	16.1	14.7
KID ($\times 10^{-3}$)			
UIEBD	10.8	8.2	6.1
TUDS	12.5	9.6	8.3

The evaluation results in Table 3 present the performance of the proposed three-stage restoration framework across the UIEBD and TUDS datasets. For Entropy ($H(X)$), which measures the amount of visual information, the diffusion-based restoration yields the highest scores (7.8 for UIEBD and 7.1 for TUDS), indicating superior texture and detail preservation relative to CLAHE or CLAHE+ACC. .

Furthermore, both FID and KID metrics demonstrate significant improvements through the enhancement sequence. The FID values decrease progressively from 18.2 to 9.4 for UIEBD and from 22.4 to 14.7 for TUDS, while the KID metric shows consistent reductions, indicating that the diffusion model effectively aligns the restored image features with those of real-world natural scenes.

Overall, the results validate that each enhancement stage

contributes positively to visual quality. The diffusion-based restoration consistently outperforms preceding stages, achieving the best balance between perceptual realism, color fidelity, and statistical similarity to natural image distributions.

7. CONCLUSION

The proposed three-stage underwater image restoration framework, consisting of Contrast Limited Adaptive Histogram Equalization (CLAHE), Adaptive Color Correction (ACC), and diffusion-based enhancement, showed clear improvements in both perceptual and statistical image quality. The step-by-step structure made it possible to compare the results after each stage, helping to understand how each process contributed to the final image quality. Experiments on the UIEBD and TUDS datasets confirmed the effectiveness of this approach. Overall, combining traditional contrast and color correction methods with a diffusion-based model proved to be a simple, reliable, and interpretable solution for underwater image restoration.

A. APPENDIX EXAMPLE

This is an example of the appendix sections. You can also change the contents of the following sections using the settings on the template.

ACKNOWLEDGEMENT

This work was supported by the Scientific and Technological Research Council of Türkiye (TUBITAK) with 1515 Frontier R&D Laboratories Support Program for BTS Advanced AI Hub: BTS Autonomous Networks and Data Innovation Lab. Project 5239903, in part by the Turk Telekom 6G Research and Development Laboratory under Project 5249902, and Project 223M055; partly by the Scientific Research Projects Coordination Department (BAP), Istanbul Technical University, under ITU-BAP MGA-2024-45372.

REFERENCES

- [1] Y. Yildirim et al. "Image analysis based fish tail beat frequency estimation for fishway efficiency". In: *2018 26th European Signal Processing Conference (EUSIPCO)*. Rome, Italy, 2018, pp. 621–625. doi: [10.23919/EUSIPCO.2018.8553131](https://doi.org/10.23919/EUSIPCO.2018.8553131).
- [2] S. Kucukali, B. Verap, A. Alp, D. Turan, T. Mutlu, C. Kaya, Y. Yildirim, B. U. Toreyin, and D. Ozelci. "Flow structure and fish passage performance of a brush-type fish way: a field study in the İyidere River, Turkey". In: *Marine and Freshwater Research* 70.11 (2019), pp. 1619–1632.
- [3] Mingjie Wang, Keke Zhang, Hongan Wei, Weiling Chen, and Tiesong Zhao. "Underwater image quality optimization: Re-

- searches, challenges, and future trends". In: *Image and Vision Computing* 146 (2024), p. 104995. ISSN: 0262-8856. doi: [10.1016/j.mavis.2024.104995](https://doi.org/10.1016/j.mavis.2024.104995). URL: <https://www.sciencedirect.com/science/article/pii/S0262885624000994>.
- [4] Derya Akkaynak and Tali Treibitz. "A revised underwater image formation model". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 6723–6732.
- [5] R. Sathya, M. Bharathi, and G. Dhivya. "Underwater image enhancement by dark channel prior". In: *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*. Coimbatore, India, 2015, pp. 1119–1123. doi: [10.1109/ECS.2015.7124757](https://doi.org/10.1109/ECS.2015.7124757).
- [6] Zhengping Sun et al. "Underwater image processing method based on red channel prior and Retinex algorithm". In: *Optical Engineering* 60.9 (2021), p. 093102. doi: [10.1117/1.OE.60.9.093102](https://doi.org/10.1117/1.OE.60.9.093102). URL: <https://doi.org/10.1117/1.OE.60.9.093102>.
- [7] Weidong Zhang, Yudong Wang, and Chongyi Li. "Underwater Image Enhancement by Attenuated Color Channel Correction and Detail Preserved Contrast Enhancement". In: *IEEE Journal of Oceanic Engineering* 47.3 (2022), pp. 718–735. doi: [10.1109/JOE.2022.3140563](https://doi.org/10.1109/JOE.2022.3140563).
- [8] Linghao Shen and Yongqiang Zhao. "Underwater image enhancement based on polarization imaging". In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 43 (2020), pp. 579–585.
- [9] Xu Liu, Hanhao Zhu, Weihua Song, Jiahui Wang, Lengeng Yan, and Kelin Wang. "Research on improved VGG-16 model based on transfer learning for acoustic image recognition of underwater search and rescue targets". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2024).
- [10] Hema Krishnan, Anjana A Lakshmi, LS Anamika, CH Athira, PV Alaiikha, and VM Manikandan. "A novel underwater image enhancement technique using ResNet". In: *2020 IEEE 4th Conference on Information & Communication Technology (CICT)*. IEEE, 2020, pp. 1–5.
- [11] YL Lai, TF Ang, UA Bhatti, CS Ku, Q Han, and LY Por. "Color correction methods for underwater image enhancement: A systematic literature review". In: *PLoS ONE* 20.3 (2025), e0317306. doi: [10.1371/journal.pone.0317306](https://doi.org/10.1371/journal.pone.0317306). URL: <https://doi.org/10.1371/journal.pone.0317306>.
- [12] Kancharagunta Kishan Babu, Ashreen Tabassum, Bommakanti Navaneeth, Tenneti Jahnava, and Yenka Akshaya. "Underwater image enhancement using generative adversarial networks: a survey". In: *International Journal of Computers and Applications* 47.4 (2025), pp. 356–372.
- [13] Qiong Liu, Qi Zhang, Wei Liu, Wenbai Chen, Xinwang Liu, and Xiangke Wang. "WSDS-GAN: A weak-strong dual supervised learning method for underwater image enhancement". In: *Pattern Recognition* 143 (2023), p. 109774. ISSN: 0031-3203. doi: <https://doi.org/10.1016/j.patcog.2023.109774>. URL: <https://www.sciencedirect.com/science/article/pii/S0031320323004727>.
- [14] Yi Tang, Hiroshi Kawasaki, and Takafumi Iwaguchi. "Underwater Image Enhancement by Transformer-based Diffusion Model with Non-uniform Sampling for Skip Strategy". In: *Proceedings of the 31st ACM International Conference on Multimedia*. MM '23. Ottawa ON, Canada: Association for Computing Machinery, 2023, pp. 5419–5427. ISBN: 9798400701085. doi: [10.1145/3581783.3612378](https://doi.org/10.1145/3581783.3612378). URL: <https://doi.org/10.1145/3581783.3612378>.
- [15] Meisheng Guan, Haiyong Xu, Gangyi Jiang, Mei Yu, Yeyao Chen, Ting Luo, and Yang Song. "WaterMamba: Visual State Space Model for Underwater Image Enhancement". In: *arXiv preprint arXiv:2405.08419* (2024). URL: <https://api.semanticscholar.org/CorpusID:269761991>.
- [16] Lintao Zheng, Hengliang Shi, and Shibao Sun. "Underwater image enhancement algorithm based on CLAHE and USM". In: *2016 IEEE International Conference on Information and Automation (ICIA)*. 2016, pp. 585–590. doi: [10.1109/ICInfA.2016.7831889](https://doi.org/10.1109/ICInfA.2016.7831889).
- [17] Weihong Zhang, Xiaobo Li, Shuping Xu, Xujin Li, Yiguang Yang, Degang Xu, Tiegeng Liu, and Haofeng Hu. "Underwater Image Restoration via Adaptive Color Correction and Contrast Enhancement Fusion". In: *Remote Sensing* 15.19 (2023). ISSN: 2072-4292. doi: [10.3390/rs15194699](https://doi.org/10.3390/rs15194699). URL: <https://www.mdpi.com/2072-4292/15/19/4699>.
- [18] Tansel Akgül, N. Calik, and B. U. Töreyn. "Deep Learning-Based Fish Detection in Turbid Underwater Images". In: *Proceedings of the SIU Conference*. Gaziantep, Turkey, 2020.
- [19] Fatima Iqbal and Behçet Uğur Töreyn. "Underwater Turbid Image Restoration Using Diffusion Models". In: *2025 33rd Signal Processing and Communications Applications Conference (SIU)*. 2025, pp. 1–4. doi: [10.1109/SIU66497.2025.11112226](https://doi.org/10.1109/SIU66497.2025.11112226).
- [20] Chongyi Li, Chunle Guo, Wenqi Ren, Runmin Cong, Junhui Hou, Sam Kwong, and Dacheng Tao. "An Underwater Image Enhancement Benchmark Dataset and Beyond". In: *IEEE Transactions on Image Processing* 29 (2020), pp. 4376–4389. doi: [10.1109/TIP.2019.2955241](https://doi.org/10.1109/TIP.2019.2955241).
- [21] Pouria Khanzadi, Babak Majidi, and Ehsan Akhtarkavan. "A novel metric for digital image quality assessment using entropy-based image complexity". In: *2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI)*. 2017, pp. 0440–0445. doi: [10.1109/KBEI.2017.8325017](https://doi.org/10.1109/KBEI.2017.8325017).
- [22] Rudy Gunawan, Yvonne Tran, Jinchuan Zheng, Hung Nguyen, and Rifai Chai. "Implementing Natural Image Quality Evaluator for Performance Indicator on Noise Artefacts Recovery in CT Scan". In: *2023 45th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*. 2023, pp. 1–4. doi: [10.1109/EMBC40787.2023.10340822](https://doi.org/10.1109/EMBC40787.2023.10340822).
- [23] Opher Bar Nathan, Deborah Levy, Tali Treibitz, and Dan Rosenbaum. "Osmosis: RGBD Diffusion Prior for Underwater Image Restoration". In: *Computer Vision – ECCV 2024*. Vol. 15120. Lecture Notes in Computer Science. Cham, Switzerland: Springer Science and Business Media Deutschland GmbH, 2024, pp. 302–319. doi: [10.1007/978-3-031-73033-7_17](https://doi.org/10.1007/978-3-031-73033-7_17). URL: https://doi.org/10.1007/978-3-031-73033-7_17.
- [24] Cosmin Ancuți, Codruta Orniana Ancuti, Tom Haber, and Philippe Bekaert. "Enhancing underwater images and videos by fusion". In: *2012 IEEE Conference on Computer Vision and Pattern Recognition* (2012), pp. 81–88. URL: <https://api.semanticscholar.org/CorpusID:17006174>.
- [25] Shuaixin Liu, Kunqian Li, Yilin Ding, and Qi Qi. "Underwater Image Enhancement by Diffusion Model with Customized CLIP-Classifier". In: *Pattern Recognition* 171.Part B (2026), p. 112232. ISSN: 0031-3203. doi: [10.1016/j.patcog.2025.112232](https://doi.org/10.1016/j.patcog.2025.112232). URL: <https://doi.org/10.1016/j.patcog.2025.112232>.

AUTHORS

FATIMA IQBAL received her Master's degree in Computer Sciences from the Informatics Institute, Istanbul Technical University, Turkey.

BEHÇET UĞUR TÖREYN is a Professor in the Department of Artificial Intelligence and Data Engineering, Faculty of Computer and Informatics Engineering, Istanbul Technical University, Turkey