ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

RESTORING UNDERWATER IMAGES FROM TURBIDITY AND MOTION BLUR – A THREE-STEP FRAMEWORK

M.Sc. THESIS

Fatima IQBAL

Department of Computer Sciences

Computer Sciences Programme

FEBRUARY 2026

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

RESTORING UNDERWATER IMAGES FROM TURBIDITY AND MOTION BLUR – A THREE-STEP FRAMEWORK

M.Sc. THESIS

Fatima IQBAL
(704231010)

Department of Computer Sciences

Computer Sciences Programme

Thesis Advisor: Prof. Dr. Behçet Uğur TÖREYİN
Co-advisor: Prof. Dr. Serhat KÜÇÜKALİ

FEBRUARY 2026

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BULANIKLIK VE HAREKETTEN ETKİLENMİŞ SU ALTI GÖRÜNTÜLERİNİN İYİLEŞTİRİLMESİ - ÜÇ ADIMLI BİR YAKLAŞIM

YÜKSEK LİSANS TEZİ

Fatima IQBAL
(704231010)

Bilgisayar Bilimleri Anabilim Dalı

Bilgisayar Bilimleri Programı

Tez Danışmanı: Prof. Dr. Behçet Uğur TÖREYİN
Eş Danışman: Prof. Dr. Serhat KÜÇÜKALİ

ŞUBAT 2026

Fatima IQBAL, a M.Sc. student of ITU Graduate School student ID 704231010 successfully defended the thesis entitled "RESTORING UNDERWATER IMAGES FROM TURBIDITY AND MOTION BLUR – A THREE-STEP FRAMEWORK", which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Dr. Behçet Uğur TÖREYİN** Istanbul Technical University	.....
Co-advisor :	**Prof. Dr. Serhat KÜÇÜKALİ** Hacettepe University	.....
Jury Members :	**Assoc. Prof. Dr. Cihan TOPAL** Istanbul Technical University	.....
	Prof. Dr. Ender Mete EKŞIOĞLU Istanbul Technical University	.....
	Asst. Prof. Dr. Nurullah ÇALIK Istanbul Medeniyet University	.....

Date of Submission : **7 January 2026**
Date of Defense : **2 February 2026**

To my parents, with gratitude,

FOREWORD

I would like to express my sincere gratitude to my advisor, Prof. Dr. Behçet Uğur TÖREYİN, and my co-advisor, Prof. Dr. Serhat KÜÇÜKALİ, for their guidance, support, and valuable advice throughout the completion of this thesis. This work was supported by the Scientific and Technological Research Council of Türkiye (TUBITAK) through the 1515 Frontier R&D Laboratories Support Program for BTS Advanced AI Hub: BTS Autonomous Networks and Data Innovation Lab, Project 5239903, and Project 223M055; and partly by the Scientific Research Projects Coordination Department (BAP), Istanbul Technical University, under Projects ITU-BAP MGA-2024-45372 and PMA-2024-45912. I am also thankful to my colleagues and family for their encouragement and support.

FEBRUARY 2026 Fatima IQBAL

TABLE OF CONTENTS

ABBREVIATIONS

UIEBD	**:** Underwater Image Enhancement Benchmark Dataset
TUDS	**:** Turbidity Underwater Dataset
CLAHE	**:** Contrast Limited Adaptive Histogram Equalization
ACC	**:** Adaptive Color Correction
NIQE	**:** Natural Image Quality Evaluator
GANs	**:** Generative Adversarial Networks
CNNs	**:** Convolutional Neural Networks
CDF	**:** Cumulative Distribution Function
MSE	**:** Mean Squared Error
ReLU	**:** Rectified Linear Unit
SSIM	**:** Structural Similarity Index Measure
PSNR	**:** Peak Signal-to-Noise Ratio
FID	**:** Fréchet Inception Distance
KID	**:** Kernel Inception Distance
DCP	**:** Dark Channel Prior
RCP	**:** Red Channel Prior
DDPM	**:** Denoising Diffusion Probabilistic Model
U-Net	**:** U-shaped Convolutional Neural Network Architecture
ACC+CLAHE	**:** Combined Adaptive Color Correction and CLAHE Enhancement
VGG	**:** Visual Geometry Group network used for feature extraction

SYMBOLS

$H(X)$	: Entropy of an image
$I_r(x)$	: Red-channel intensity at pixel x
$I_g(x)$	: Green-channel intensity at pixel x
α	: Compensation parameter for red-channel gain
K_c	: Scaling factor for channel c in dynamic range adjustment
D_c	: Difference between maximum and minimum intensities of channel c
$H_{\text{clipped}}(j)$	: Clipped histogram count at intensity level j
$CDF(i)$	: Cumulative distribution function at intensity i
μ_t, σ_t^2	: Mean and variance of noisy image at timestep t
x_0	: Clean image
x_t	: Noisy image at timestep t
ε_t	: Noise sampled at timestep t in forward diffusion
ΔE	: Color accuracy in CIE-Lab space
f	: Feature vector for NIQE calculation
μ	: Mean of natural image features for NIQE
Σ	: Covariance matrix of natural image features for NIQE
FID	: Fréchet Inception Distance
μ_r, Σ_r	: Mean and covariance of real image features for FID
μ_g, Σ_g	: Mean and covariance of generated/restored image features for FID
$k(\cdot,\cdot)$	: Kernel function in KID calculation
$\mathscr{L}_{\text{noise}}$	: Noise prediction loss (MSE) in diffusion model
$\hat{\varepsilon}_t$	: Predicted noise at timestep t

LIST OF TABLES

Page

LIST OF FIGURES

RESTORING UNDERWATER IMAGES FROM TURBIDITY AND MOTION BLUR – A THREE-STEP FRAMEWORK

SUMMARY

Underwater image restoration is essential for studying and visualizing marine life, aquatic ecosystems, underwater robotics, and optical communication systems. However, underwater imagery processing remains a difficult task because of the combined effects of **turbidity, motion blur, and wavelength-dependent absorption**. Phenomena like forward and backward scattering due to the suspended particles found in water bodies make imaging even more difficult. Degraded aquatic imagery eventually results from this loss of fine structural details, distorted color information, and decreased visibility, making image processing computationally challenging. Despite earlier underwater restoration models offering a suitable mathematical formulation to quantify such degradations, practical restoration of underwater imagery remains a difficult task.

Traditional image enhancement methods that attempt to estimate degradation parameters and restore image quality manually are included in the literature review. However, these are not as reliable, particularly for multi-degradation cases where more than one artifact can be present at the same time. Recent developments in deep learning-based models have led to the widespread use of Generative Adversarial Networks (GANs) for underwater enhancement due to their capacity to produce high-quality synthetic output images. Diffusion models are also becoming increasingly popular for denoising visual data due to their exceptional ability to model real-world noise distributions. Nevertheless, the majority of approaches only deal with one kind of degradation, which restricts their use in actual underwater settings. Another important challenge is the scarcity of paired ground-truth underwater images, which hinders objective evaluation and makes supervised training much more difficult.

Taking into consideration the inadequacies discussed above, the proposed approach in this study consists of enhancement framework based on noise parameters estimated using a human-made turbidity dataset. We make use of the pre-processing image enhancement techniques described in the literature to remove artifacts such as uneven illumination and haziness. In the first stage of our framework, **Contrast-Limited Adaptive Histogram Equalization (CLAHE)** is used to increase local contrast in the image by controlling the extent of noise augmentation. CLAHE basically improves the visibility of structural details and highlights important image features.

In the second stage, an **Adaptive Color Correction (ACC)** module is employed that removes the color hue of the underwater environment by compensating for the absorption of certain wavelengths. In this stage, the images are classified as bluish images, greenish images, yellowish images, and color-balanced images based upon the channel statistics. This intensity adjustment allows the tonal range to be used efficiently while simultaneously enhancing global color fidelity.

In the final stage, a **diffusion-based denoising model** refines the image further. In this case, we have approximated the values of the noise variance by the simulation of a realistic underwater scenario noise data that was prepared using soil and water. The mean and variance parameters extracted from this laboratory data were set to construct the synthesized noisy data images. This created data is then used for model training procedure via the forward diffusion process. Hence, it enables the diffusion model to learn noise patterns that closely resemble the real-life underwater degradations. The reverse diffusion model then finally reconstructs the output images by denoising the degraded images step-by-step.

Conventional evaluation metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) would not be effective in the case of underwater images that lack ground truth. In this work, we evaluated our proposed framework on two datasets: The Underwater Image Enhancement Benchmark Dataset (UIEBD) and the Turbidity Underwater Dataset (TUDS). Both these datasets were chosen as there was variation in their levels of turbidity, distortions in color, and illumination issues.

Here, we have used relevant quantitative metrics, including entropy $H(X)$, color accuracy ΔE, Naturalness Image Quality Evaluator (NIQE), Fréchet Inception Distance (FID), and Kernel Inception Distance (KID) to quantify the merits obtained at every stage of the enhancements. The obtained results demonstrated that there were substantial advancements made by the final step of the diffusion model when it comes to perceptual quality, and detail retention. An examination of the resulting images obtained at every step of the architecture further revealed that there have been progressive enhancements in the local contrast, color, and clarity of the noisy images.

However, in spite of the very good results achieved by the proposed architecture, several impractical but unavoidable drawbacks are pointed out. Firstly, under some conditions, both CLAHE and ACC may correct images too much, which may result in extreme high contrast and unrealistic colors. To avoid such difficulties, it is necessary to manually adjust all parameters pertaining to different images and levels of turbidity. Secondly, it has been observed that a diffusion model performance is greatly dependent on the properties of images being used as the training data and the noise being modeled. This results in a diffusion model with less possibility to generalize well when test images with entirely different properties are considered in the underwater domain.

This novel method is a crucial step ahead compared to other existing methods in the underwater image enhancement literature. This is an innovative technique that is capable of effectively simulating the noise that occurs in the real-life environment. It utilizes the noise parameters that were estimated from the turbidity sequences prepared in the laboratory. Secondly, traditional image processing methods applied in our research are capable of preparing the raw underwater images to be processed through enhancement of structural details as well as color contrast in images. Finally, diffusion model based denoising step follows realistic noise patterns and performs well under varying levels of water turbidity. Therefore, this method provides a useful tool in natural underwater image recovery with several applications in underwater communications research and the preservation of marine life.

BULANIKLIK VE HAREKETTEN ETKİLENMİŞ SU ALTI GÖRÜNTÜLERİNİN İYİLEŞTİRİLMESİ - ÜÇ ADIMLI BİR YAKLAŞIM

ÖZET

Denizaltı görüntüleme, deniz yaşam alanlarının gözlemlenmesi, sualtı robotik sistemlerinin kontrolü, çevresel koşulların izlenmesi, arkeolojik çalışmalar ve hatta biyolojik gözetim gibi birçok alanda büyük öneme sahip temel bir teknolojidir. Bu uygulamalarda doğru ve güvenilir yüksek kaliteli görüntü verilerine olan talep her geçen yıl artmaktadır. Bununla birlikte, su yüzeyinin altında bulunan alışılmadık optik koşullar, iyi görüntüler elde etmeyi son derece zorlaştırmaktadır. Bu nedenle, sualtı görüntü kurtarma üzerine yapılan araştırmalar, akademisyen çevreleri ve endüstri uygulayıcıları arasında en önemli önceliklerden biri haline gelmiştir. Bu makale, sualtı görüntülerinin işlenmesi sırasında karşılaşılan sorunlara ve mevcut literatürde ele alınan yöntemlere genel bir bakış sunmanın yanı sıra, önerilen üç aşamalı gelişmiş bir iyileştirme yaklaşımını da sunmaktadır.

En büyük zorluk, suyun havadan çok daha fazla ışığı emmesi ve saçması nedeniyle su altı görüntülerinin elde edilmesidir. Bu nedenle, askıda kalan katı maddeler, organik madde ve mikroskobik parçacıklar hem doğrudan emilim hem de nispeten yoğun bir sisleme etkisi yaratan geri saçılma ile birlikte, derinlikle birlikte kırmızı kanalda renk zayıflamasına ve görüntülerde yaygın mavi-yeşil ton kaymalarına neden olur. Diğer sorunlar arasında tutarsız aydınlatma, soğuk renk tonları, bulanıklık, hareket bulanıklığı ve düşük aydınlatma seviyeleri yer alır; bunlar su altı görüntülerinin yapısal ve algısal kalitesini birçok açıdan olumsuz etkiler. Geleneksel iyileştirme teknikleri yalnızca bir tür bozulmaya karşı etkilidir; ancak bu bozulmalar gerçek bir su altı ortamında çoğunlukla birlikte bulunur ve tek adımlı yaklaşımlarla tutarlı bir şekilde ele alınamaz. Bu nedenle, bu çalışma birbirini tamamlayan çok aşamalı bir iyileştirme yaklaşımının neden gerekli olduğunu ortaya koymaktadır.

Literatürde su altı görüntü iyileştirme için kullanılan yöntemler genel olarak klasik iyileştirme yöntemleri ve derin öğrenme tabanlı teknikler olarak sınıflandırılabilir. Klasik yöntemler arasında histogram eşitleme, beyaz dengesi teknikleri, Retinex tabanlı yaklaşımlar ve fiziksel model tabanlı sis giderme yöntemleri yer almaktadır. Bu yöntemler kontrastı ve renk dağılımını iyileştirse de, aşırı düzeltme, detay kaybı ve düzensiz renk dağılımlarına neden olmaktadır. Derin öğrenme tabanlı yöntemler ise GAN tabanlı yaklaşımlara, CNN tabanlı modellemeye ve son zamanlarda popüler hale gelen difüzyon tabanlı modellere dayanmaktadır. Bu modeller hem sentetik hem de gerçek su altı veri kümelerinde iyi performans gösterse de, farklı bulanıklık seviyeleri ve aydınlatma koşullarına göre kararlılık sorunları yaşıyor gibi görünmektedir. Bu eksiklikler, daha kontrollü, daha evrensel ve birden fazla bozulmayı aynı anda ele alabilen bir restorasyon yaklaşımını gerekli kılmaktadır.

Bu ihtiyaca yönelik olarak bu çalışma, çoklu bozulmayı aşamalı ve kademeli biçimde düzeltmeyi hedefleyen üç aşamalı bir görüntü iyileştirme çerçevesi önermektedir. Önerilen mimari, klasik iyileştirme yöntemlerinin kararlılığını, öğrenme tabanlı yöntemlerin öğrenilebilirliğini ve modern difüzyon yaklaşımlarının gücünü bir araya getirmektedir. Çerçevenin ilk aşaması olan CLAHE (Contrast Limited Adaptive Histogram Equalization), yerel kontrastı adaptif bir şekilde artırarak ayrıntı ve doku bilgisini güçlendirmektedir. Standart histogram eşitleme tüm görüntüye tek bir global düzeltme uygularken, CLAHE görüntüyü daha küçük alt bölgelere ayırmakta ve her bölgenin kendi dağılım karakterine göre kontrast iyileştirmesi yapmaktadır. Bu sayede gerek derin su ortamları gibi düşük ışıklı sahnelerde gerekse yoğun türbidite içeren görüntülerde ayrıntı kaybına yol açmadan yapısal detayların ortaya çıkarılması mümkün olmaktadır. Bu aşama, sonraki adımlarda uygulanacak renk düzeltimi için sağlam ve dengeli bir temel oluşturmaktadır.

İkinci aşama olan Adaptive Color Correction (ACC), su altı ortamına özgü renk bozulmalarını hedef almaktadır. Bu yöntemde görüntüler öncelikle *bluish*, *greenish*, *yellowish* veya *balanced* olmak üzere arka plan rengine göre sınıflandırılmaktadır. Bu sınıflandırma, her görüntüde farklı oranlarda baskın olan renk kanalının daha etkin biçimde düzeltilmesine olanak sağlamaktadır. *Bluish* sınıfında yoğunlaşan görüntülerde yeşil kanal referans alınarak kırmızı kanal güçlendirilmektedir. *Yellowish* görüntülerde ise mavi ve yeşil kanallar, kırmızı kanal temel alınarak dengelenmektedir. Renk dağılımı zaten dengeli olan (*balanced*) görüntülerde ise ek bir renk düzeltmesi uygulanmamaktadır. Bu aşamada kullanılan dinamik kazanç ayarlamaları, hem global hem de lokal renk dağılımını daha doğal hale getirerek gerçekçi renk tonlarının elde edilmesini sağlamaktadır. Ayrıca bu aşama, özellikle renk hatalarının nicel olarak ölçüldüğü (ΔE) değerlerinde belirgin bir düşüş sağlamaktadır.

Üçlü çerçevenin son aşaması, restorasyon sürecine difüzyon tabanlı bir modelin entegre edildiği adımdır. Bu aşama, su altı görüntülerinde gürültü, sislenme ve düşük frekanslı bozulmalar gibi türbidite kaynaklı etkilerin giderilmesini amaçlamaktadır. Difüzyon modelleri, ileri difüzyon sürecinde görüntüye kademeli olarak gürültü ekleyerek veri dağılımını öğrenir. Ardından, ters difüzyon aşamasında, öğrenilen olasılıksal yapıdan yararlanılarak görüntü adım adım yeniden oluşturulur.

Bu çalışmada, ters difüzyon süreci, U-Net tabanlı mimari kullanılarak modellenmiştir. Model, istatistiksel olarak gerçek bir sualtı verisi oluşturan bir tür gürültü dağılımını öğrenmek suretiyle, verilere en uygun ortalama ve varyans tahminlerini etkili bir şekilde öğrenmektedir. Böyle bir yaklaşım, restorasyon sürecini gerçekçi ve kararlı kılacaktır. Difüzyon kısmı, yeniden canlandırılan ayrıntılar ve algısal doğal olma açısından en iyi kazancı sağlayan kısımdır. Yapılan renk dağılımlarının doğal rekabete girme seviyeleri, sisin gözle görülür şekilde ışığa doğru yayılmasının azalması ve dokusal bilgilerin daha belirgin oluşu, bu aşamada oluşan en baskın katkılardır.

Önerilen yöntem, Underwater Image Enhancement Benchmark Dataset (UIEBD) ve Turbidity Underwater Dataset (TUDS) gibi yaygın olarak kullanılan iki kapsamlı su altı görüntü iyileştirme veri seti üzerinde değerlendirilmiştir. Bu veri setleri, farklı türbidite seviyeleri, renk bozulmaları ve aydınlatma koşulları açısından geniş bir çeşitlilik sunmaktadır. Nicel değerlendirme amacıyla bilgi içeriğini ölçen entropi ($H(X)$), renk doğruluğunu ifade eden ΔE, algısal kaliteyi değerlendiren NIQE, dağılım

benzerliğini ölçen FID ve istatistiksel farklılıkları analiz etmek için kullanılan KID metrikleri uygulanmıştır.

CLAHE aşaması, entropi değerini artırarak yerel ayrıntıların daha belirgin hale geldiğini göstermiştir. ACC aşaması ise renk doğruluğunu iyileştirerek ΔE değerlerinde kayda değer bir düşüş sağlamıştır. Difüzyon modelinin uygulandığı son aşama, algısal doğallık metriklerinde en yüksek iyileşmeyi ortaya koymuştur; NIQE değeri UIEBD veri setinde 3.1 seviyesinden 2.0 seviyesine gerilemiş, TUDS veri setinde de benzer bir iyileşme gözlemlenmiştir. FID ve KID skorlarındaki azalma, üretilen görüntülerin istatistiksel olarak daha doğal bir dağılıma yaklaştığını doğrulamaktadır. Nitel gözlem analizleri, yoğun türbidite içeren sahnelerde, hareketli nesnelerde ve düşük ışıklı koşullarda dahi önerilen iyileştirme sürecinin tutarlı ve kararlı sonuçlar ürettiğini göstermiştir.

Önerilen çerçevenin birçok nicel ve nitel değerlendirmede öne çıkmasına rağmen, bazı sınırlamaları da bulunuyor. Difüzyon tabanlı aşama, klasik görüntü işleme yöntemlerine göre daha fazla hesaplama gücü gerektiriyor. Bu nedenle, gerçek zamanlı uygulamalar için ek optimizasyonlar yapmak gerekebilir. Ayrıca, aşırı bulanıklık seviyelerine sahip veya renk kanallarında ciddi bozulmalar içeren görüntülerde, modelin veri setine özel ince ayarlamalar yapması muhtemel. Yine de, üç aşamalı çerçevenin genel yapısı, farklı su altı koşullarına uyum sağlayabilen, kararlı ve esnek bir iyileştirme yaklaşımı sunuyor.

Sonuç olarak, bu çalışma, birden fazla su altı bozulmasını tek aşamalı yaklaşımların zorlandığı şekilde kapsamlı bir biçimde ele alan yeni bir çerçeve sunuyor. Önerilen yapı, hem yapısal açıklık hem de renk doğruluğu sağlarken, algısal doğallığı da belirgin bir şekilde artırıyor. Bu çerçevenin, akademik araştırmaların yanı sıra su altı robotik sistemleri, arama-kurtarma operasyonları, biyolojik gözlemleme ve çevresel izleme gibi gerçek dünya uygulamalarında pratik katkılar sunma potansiyeli bulunuyor. Bu bağlamda, çalışma, su altı görüntü işleme literatürüne hem teorik hem de uygulamalı açıdan önemli bir katkı sağlıyor ve gelecekte geliştirilecek yöntemler için sağlam bir temel oluşturuyor.

1. INTRODUCTION

1.1 Background: Importance and Scope of Underwater Image Processing

Approximately 70% of the Earth's surface is covered by water and around 97 percent of this water comprises oceans. A vast portion of this underwater world still remains unexplored. Scientists have long been interested in underwater research due to its crucial role in marine life, biodiversity conservation, environmental monitoring, and industrial inspection. In addition, high-quality underwater images support tasks such as tracking fish migration patterns, assessing coral reef health, and inspecting submerged infrastructure. Despite its wide range of applications, capturing clear underwater images is a challenging task because of the complex interaction between light and water.

There are essentially two ways to approach the underwater image processing problem: either as an image enhancement method or as an image restoration technique.

- **Image restoration** focuses on reconstructing a damaged underwater image by meticulously simulating the deterioration process. Here, an inverse method is used to reverse the imaging-induced artifacts. These restoration techniques rely on various physical parameters, such as attenuation constants and scattering coefficients. However, this estimation is very complicated because of the complexity and volume of real-world datasets.

- **Image enhancement** is the process of improving an underwater image both qualitatively and subjectively. These improvement methods are typically more straightforward and computationally faster than relying on any physical model. Instead of recovering the actual scene, their main goal is to create an aesthetically pleasing output.

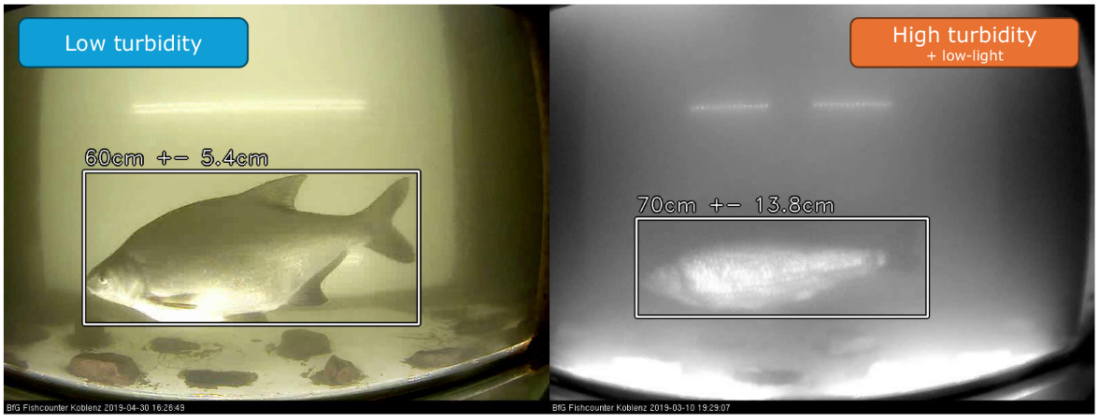

Figure 1.1: Underwater video monitoring is impacted by water turbidity [1]. Fish shape and features are clearly visible in low-turbidity conditions (left), allowing for accurate detection and size estimation. On the other hand, low light and high turbidity (right) greatly impair visibility, resulting in blurry silhouettes, greater measurement uncertainty, and restricted species identification.

1.2 Effect of Turbidity in Fish Observation

Underwater video surveillance is widely used in fish observation and migration studies. However, the site characteristics, specifically the water turbidity, have a large effect on its efficacy. The level of turbidity is found to be the critical factor for decision-making regarding the underwater monitoring.

When water turbidity is greater than 20 Nephelometric Turbidity Units (NTU) or the Secchi depth [4] is less than the bottleneck width, video monitoring is severely limited, leading to poor visibility and inconsistent detection. Although increasing turbidity gradually restricts accurate species identification, more efficient monitoring is possible at turbidity levels below 5 NTU.

In high turbidity conditions, white light illumination may impact fish behavior, and back-lighting techniques reduce visual details, limiting accurate species identification as depicted in Figure 1.1. These drawbacks emphasize turbidity as a significant obstacle and motivate the development of reliable underwater image enhancement methods.

1.3 Underwater Image Formation Model

Underwater image quality is severely reduced by light absorption and scattering, which also causes color distortions. Physical models have been created to describe this phenomenon and allow for more precise restoration. The Jaffe-McGlamery model [5], [6] has been widely adopted because of its physically grounded approach. The model is based on linear superposition and accounts for the optical characteristics of water. As shown in Figure 1.2, the total irradiance recorded by the camera is actually a combination of three main components:

- **Direct component ($I_{\textbf{direct}}(x)$):** Component of light that reaches the camera without scattering after reflection from the target.
- **Forward-scattered component ($I_{\textbf{forward}}(x)$) :** Light reflected from the target that undergoes scattering after colliding with a suspended particle in water and travels towards the camera.
- **Backward-scattered component ($I_{\textbf{backscatter}}(x)$):** Source light scattered by water and suspended particles that reaches the camera without interacting with the target object.

The total irradiance $I(x)$ at a pixel location x can be expressed in the form of a mathematical equation as the sum of each of these components:

$$I(x) = I_{\text{direct}}(x) + I_{\text{forward}}(x) + I_{\text{backscatter}}(x) \tag{1.1}$$

Although the Jaffe-McGlamery model provides an accurate physical basis for underwater image restoration, it also demands precise scene depth and water optical parameters estimation. Hence, it is a computationally intensive approach compared to simpler image enhancement techniques [7], [8]. Figure 1.2 illustrates this phenomenon of optical degradation.

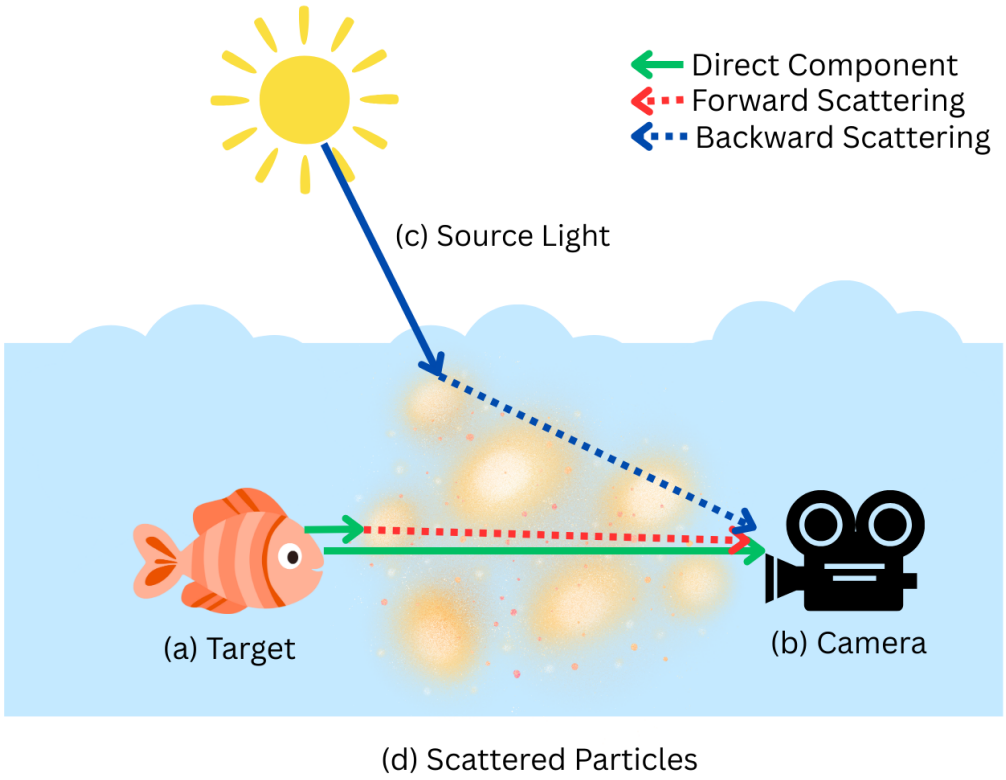

Figure 1.2: An example of how an underwater image is formed. (a) Light is reflected toward the camera by the target item to be photographed. (b) The target's direct light, forward-scattered light, and back-scattered light are all received by the camera. (c) The light from the incident source enters the water and interacts with the medium. (d) Scattering caused by suspended particles in the water results in haze, blur, and decreased visibility.

1.4 Related Work: Traditional Image Enhancement to Deep Learning Models

Scientists have long been interested in the mysterious world of oceans and underwater bodies. Thus, assumptions regarding water type, scene depth, and lighting conditions have been made to enhance and restore this data. Traditional image enhancement based on physics, such as multi-view and polarization-based methods [9], have been implemented to recover scene colors and structure. Other methods like single-image priors, including the dark channel prior [10], have been adapted, but their effectiveness is limited due to inaccurate transmission estimation and fixed attenuation assumptions.

Machine learning, particularly deep learning, addresses these limitations by learning mappings between degraded and clean images directly from data. Convolutional Neural Networks (CNNs) capture spatial patterns effectively while Generative Adversarial Networks (GANs) generate perceptually convincing outputs. Transformer-based models also help improve global context understanding of the captured images. However, the performance of these deep learning models is heavily depended upon the quality and diversity of the training data.

Generative Adversarial Network-based models have faced the following challenges in the literature: (1) lack of convergence during training, and (2) difficulty in simulating more complex types of noise. Diffusion Models represent a more stable way to model complex data distributions. Instead of learning directly to generate samples, they learn how to reverse a gradual process where noise is introduced into the original distribution over time. They are effective because they can handle different types of noise during image restoration, as well as strongly reconstruct high-quality images of marine life.

1.5 Problem Statement

This work focused on solving the following two major problems that persisted in the existing literature:

- **Lack of diverse real-world datasets:** Real underwater datasets that can depict the richness of natural aquatic habitats are severely lacking. The majority of available datasets are either artificially generated or restricted to particular water conditions.

This limits their usefulness for training deep learning models. Supervised learning and evaluation are made more difficult by the lack of associated ground-truth images.

- **Limited generalization to multi-degradations:** Turbidity, color attenuation, poor lighting, and motion blur are just a few of the degradations present in underwater images. Real underwater settings feature complex distortions that vary with depth, turbidity level and camera motion. The majority of models still concentrate on a single type of deterioration and only function well under controlled conditions.

1.6 Novel Contributions and Thesis Outline

The primary contributions of this thesis are summarized as follows:

- **Estimating noise parameters using laboratory-prepared turbidity progression dataset:** A custom dataset prepared using soil and water, consisting of images with increasing turbidity, was included. Using this dataset, turbidity noise parameters were estimated. These parameters then simulate the forward diffusion process through controlled increments of scattering, absorption, and visual noise, which enables fine-grained examination of turbidity progression.
- **A novel three-step restoration framework:** Contrast-Limited Adaptive Histogram Equalization (CLAHE), Adaptive Color Correction (ACC), and a diffusion-based refinement model are all integrated in the suggested enhancement pipeline. This hybrid method gradually restores contrast, color accuracy, and fine texture in murky underwater photos.
- **Comprehensive evaluation metrics:** A no-reference assessment strategy is established using metrics such as Naturalness Image Quality Evaluator (NIQE), entropy, color distortion (ΔE), Fréchet Inception Distance (FID), Kernel Inception Distance (KID), suitable when paired ground-truth images are unavailable.

The remainder of this thesis is divided into the following sections. Section 2 addresses the difficulties encountered when taking photos or videos underwater due to illumination challenges and water properties. Section 3 will introduce a brief summary of related

work from existing published research about approaches undertaken by previous researchers on underwater image enhancement and restoration. The methodology used to develop the three-phased enhancement pipeline, including the formulation of the diffusion models, can be found in Section 4. The specifications of the datasets used in our experiments will be identified in Section 5. Section 6 introduces the evaluation metrics that were employed for assessing no-reference underwater photos and the results obtained. Section 7 summarizes the findings, outlines limitations of this research, and provides recommendations for areas of future study.

2. CHALLENGES IN UNDERWATER IMAGING

2.1 Visual Artifacts

Numerous physical and optical deteriorations are inherent to underwater imaging. These distortions result from mechanical and environmental conditions, as well as the interaction of light with water and suspended particles. In this section, we discuss the main visual difficulties related to underwater image generation as mentioned in the visual artifacts section of Table 2.1.

Table 2.1: Challenges in underwater imaging: Categorization by visual and environmental artifacts

Visual Artifacts	
Absorption of Light	Wavelength-dependent attenuation causes the prominent blue-green tones and warm colors to fade quickly.
Turbidity	Suspended particles scatter light, causing haziness, low contrast, and visibility loss.
Uneven Illumination	Limited natural light and artificial hot spots lead to inconsistent brightness and color.
Motion Blur	Camera or subject movement in dynamic water reduces sharpness and detail.
Environmental and Technical Artifacts	
Lack of Ground Truth Data	Paired clean and degraded images are difficult to obtain, limiting supervised training.
Environmental Variability	Changes in salinity, turbidity, temperature, and biological activity introduce unpredictable distortions.
Technological Limitations	Hardware constraints and limited knowledge of water dynamics contribute to degraded imagery.

2.1.1 Absorption of light

Due to selective absorption, different wavelengths of sunlight penetrate the water column at different rates. Longer wavelengths are absorbed within the first few metres, while

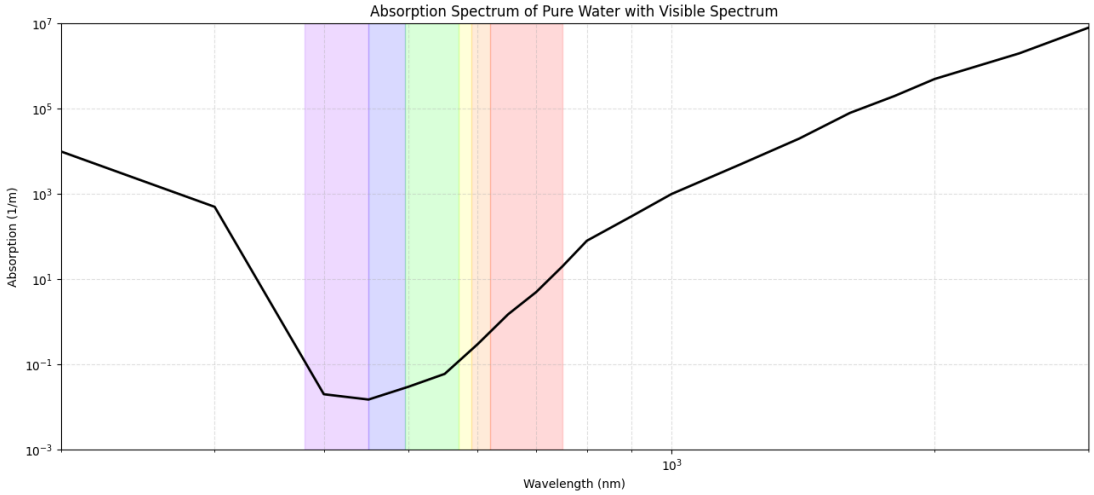

Figure 2.1: The absorption spectrum of water illustrates that blue (450-495 nm) and cyan/green (495-570 nm) wavelengths have greater penetration depth than all other visible light; whereas red light (620-750 nm) is absorbed by less than three meters of water, and infrared radiation (>750 nm) is absorbed within just seconds after entering the water column. Within just a few meters of traversing the water column, longer wavelengths, such as red, orange (590-620 nm), and yellow (570-590 nm), become completely diluted and are therefore not visible in the blue and green tones typical in all underwater photographic images.

some medium-length wavelengths go further and shorter wavelengths generally travel the longest. Longer wavelengths are those associated with red and orange, medium wavelengths include yellow and green, whereas shorter wavelengths are associated with blue and cyan.

The consequence of this wavelength-dependent effect is that when light penetrates through the water column, the resulting image is often heavily blue or green colored. It is often very difficult to accurately identify the colors of aquatic life. A visual representation of how the water absorption spectrum varies with respect to depth and wavelength is shown in Figure 2.1.

2.1.2 Turbidity and suspended particles

Suspended particles such as silt, clay, algae, organic debris, and microorganisms are frequently contained in natural aquatic environments. If this particle concentration is high, it increases scattering and absorption producing turbidity that leads to foggy, milky, or low-contrast imagery. Fine details disappear due to excessive turbidity (lack of clarity), which subsequently causes blurred edges and creates a substantial loss of visibility.

In addition, increased turbidity results in an interference of depth perception; thereby, causing both human observers and automated systems to have a substantially more difficult time interpreting a scene. Turbidity's loss of color and visibility can be seen in Figure 2.2 (a).

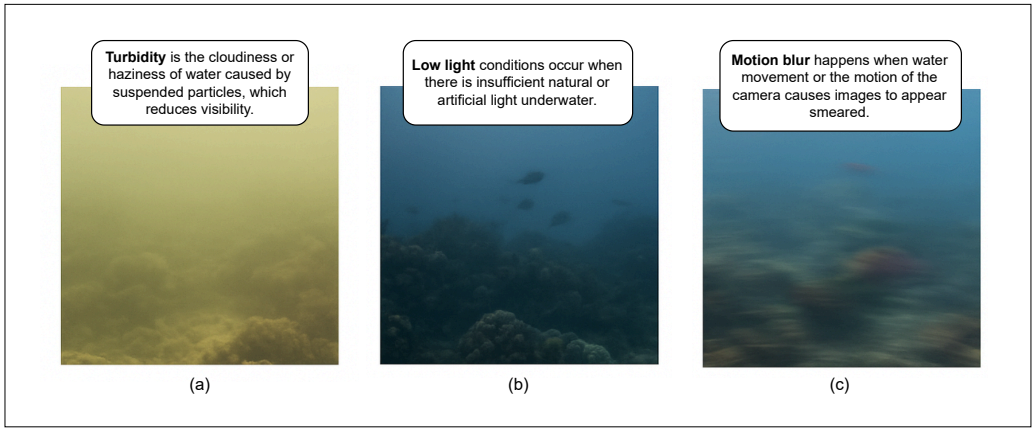

Figure 2.2: Some characteristic types of Underwater Image Degradation: (a) Turbidity, which produces hazy, poorly defined views, and is often combined with diminished visibility; (b) low-light conditions resulting in dark and low-contrast images; (c) Motion Blur, due to either camera or water movement, which produces less clear overall view of the target.

2.1.3 Uneven and limited illumination

Lighting underwater is rarely uniform. Illumination decreases exponentially with depth, and ambient light is often insufficient for accurate imaging. Artificial lighting, such as diver-mounted lamps or vehicle-based Light Emitting Diodes, may introduce strong shadows, flickering, or color inconsistencies. These effects amplify visual distortions and complicate tasks such as segmentation, classification, or geometric reconstruction. Figure 2.2 (b) depicts this uneven illumination and low-light effect.

2.1.4 Motion blur

Underwater images are subject to many factors that cause them to be highly dynamic. The movement of water and the motion of suspended particles can cause an image to exhibit motion blur; therefore, these dynamic interactions create a loss in spatial resolution and important visual details. The existence of motion blur in aquatic photography is problematic due to the loss of detail associated with the aquatic life

captured within the images. Figures 2.2 (c) represents this artifact and its obvious degradation of underwater imagery.

2.2 Environmental and Technical Artifacts

Beyond the inherent challenges, technological limitations and limited understanding of the dynamic properties of water bodies further contribute to distortions in the resulting visual output.

2.2.1 Lack of ground truth data

Capturing paired sets of degraded images and their distortion-free versions from underwater scenes is quite complex because of the unpredictability of the environment. Changes in light, distance between the subject and the camera, and variable water composition make it extremely difficult to record real-world reference images that can serve as true ground truth.

Without paired datasets, many supervised learning approaches are not as effective, and models cannot be accurately assessed. Thus, most image enhancement algorithms are based either on synthetic data or subjective quality assessments, which can introduce additional issues.

2.2.2 Environmental variability

The physical and chemical characteristics of aquatic environments are subject to considerable variability, including temperature, turbidity, salinity, dissolved organic matter, and biological activity. These factors can vary by site (depth), time of the year (season), and are often dynamic, making it very difficult to generalize enhancement models.

3. LITERATURE REVIEW

This chapter reviews existing underwater image restoration and enhancement techniques, ranging from classical physics-based methods to modern deep learning architectures. Methods like supervised, unsupervised, and generative diffusion-based approaches are also discussed in this section and summarized in Table 3.1.

3.1 Traditional Methods

Underwater image degradation is traditionally addressed using reverse modeling of the physical properties of light as it travels through water, which is referred to as the underwater image formation model (UIFM) [21]. UIFM describes how absorption and scattering affect image intensity. Using this model, several image restoration techniques estimate transmission maps to correct the wavelength-dependent attenuation.

Among these, the Dark Channel Prior (DCP) [11] utilizes the fact that under clear conditions (no haze), at least one color channel would be low intensity. Although originally designed for land images, it has been adapted for underwater images. The Red Channel Prior (RCP) [12] relies on the fast attenuation effect of light (red wavelengths) in water to provide depth perception and restore colors to the image. Depth models such as the Color Attenuation Prior [13] utilize both the brightness and color differences to estimate the depth of a scene without explicit sensors. Another method for improving the performance of restoration techniques is using a combination of direct and back-scattered light utilizing polarization techniques [14].

However, while these techniques are very effective, they are also limited by the need for accurate parameter estimation of water characteristics, illumination levels, and scene depth. As such, turbid water, spatially varying illumination, and complex water composition can lead to either over-correction or unreliable color restoration. The application of contrast enhancement techniques, such as Contrast Limited Adaptive

Histogram Equalization (CLAHE) [15], is another useful method for balancing the contrast in restored images.

3.2 Supervised Learning

Supervised deep learning approaches utilize large paired datasets to learn direct mappings from degraded underwater images to restored versions. Convolutional Neural Network (CNN) [16] structures such as VGGNet [17] and ResNet [22] are intended to extract a range of features, i.e. structural and semantic features, from an image to perform the tasks of correcting color casts, sharpening edges, and removing haze-like artifacts from a damaged image.

Most supervised models will require a set of damaged images and their corresponding high-quality ground truth versions to learn from. However, the vast majority of underwater images lack ground truth images due to the difficulty of capturing them under realistic conditions. One alternative to providing this training data is to create a synthetic image dataset, but most synthetic image datasets fail in representing the complexity of a natural underwater environment.

3.3 Unsupervised Learning

Turbid underwater images can be enhanced through unsupervised deep learning techniques without the need for paired ground truth images. These methods make use of iterative refinement loops [18], which employ pseudo labels to progressively enhance the output.

The augmentation is guided by a perceptual loss that takes into account color consistency, contrast, saturation, and clarity; methods like cycle-consistency or content-style disentanglement aid in maintaining scene structure.

3.4 Deep Learning Methods

The latest advances in underwater enhancement now feature increasingly sophisticated generative models. Generative Adversarial Networks (GANs) [19] generate realistic, high-quality output using a "generator-discriminator" pair, which has been trained

together in an adversarial way. The underwater enhancement methodologies developed using GANs learn non-linear mapping functions that can produce the best results for color correction, removal of haze, enhancement of contrast, etc.

Diffusion Models [20] represent a new class of generative frameworks that create meaningful images over time by performing iterative denoising. This approach allows noise to be modeled with very high fidelity and allows for stable training of diffusion models. Recent studies have shown that diffusion models perform very well for common image restoration tasks such as denoising, deblurring, and super-resolution, and are being increasingly used for underwater enhancement, where images have experienced severe degradation.

In this paper, we used a diffusion model to estimate the noise characteristics of underwater environments, replicating the actual noise characteristics of underwater photographs. The process was a probabilistic framework with forward and reverse diffusion steps, as shown in Figure 3.1. After a series of forward and reverse diffusion steps, our proposed method was able to remove turbidity and generate a visually coherent reconstruction of the original underwater photograph.

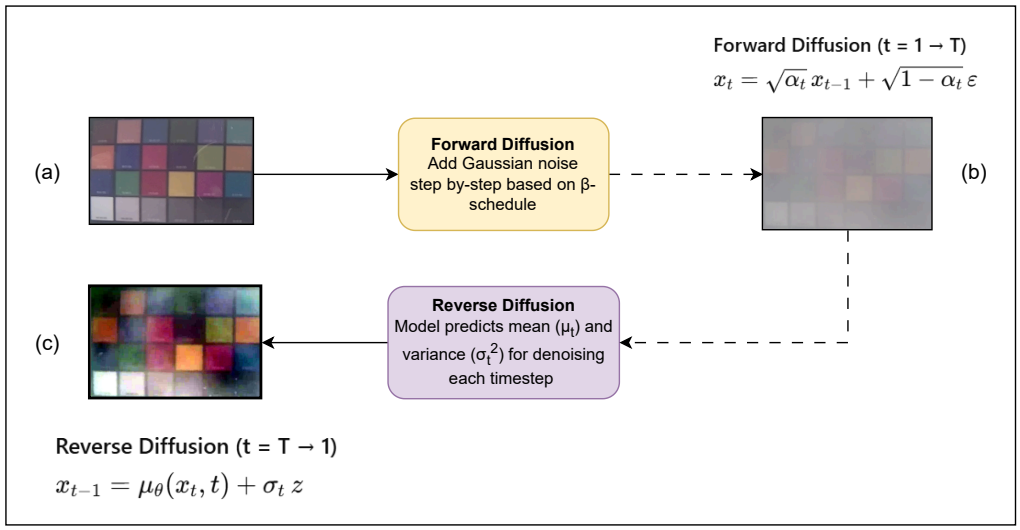

Figure 3.1: Flowchart of the forward and reverse diffusion processes used in the proposed model. (a) The clean input image. (b) The noisy image obtained after the forward diffusion stage, where Gaussian noise is progressively added at each timestep using $x_t = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, \varepsilon$. (c) The restored output generated by the reverse diffusion stage, in which the model predicts the variance (σ_t^2) at each timestep to iteratively denoise the image.

Table 3.1: Overview of underwater image restoration methods classified as traditional, supervised, unsupervised, and deep learning approaches

Method	Techniques	Advantages	Limitations
Traditional	• Dark Channel Prior (DCP) [11] • Red Channel Prior (RCP) [12] • Color Attenuation Prior [13] • Polarization-based [14] • Contrast Limited Adaptive Histogram Equalization (CLAHE) [15]	• No training needed • Interpretable • Works for mild degradations	• Requires accurate water/depth info • Struggles in turbid/complex conditions • Artifacts possible
Supervised	• CNNs (VGG [16], ResNet [17]) • Pixel-wise, perceptual, adversarial losses	• High-quality restoration • Strong downstream performance	• Needs paired datasets • Poor generalization in diverse waters
Unsupervised	• Cycle-consistency [18] • Content-style disentanglement • Perceptual losses	• No paired data required • Preserves structure • Robust to distortions	• Can hallucinate textures • Sensitive to domain mismatch
Deep Learning	• GANs [19] • Diffusion models [20]	• Realistic results • Effective color/haze correction • Handles severe degradations	• Data-hungry • Computationally heavy • Real-time deployment challenging

4. PROPOSED METHODOLOGY

Several steps are needed to improve and restore underwater photos using the innovative methodology presented in this work. In this work, the artifacts that cause underwater image degradation are addressed using a three-step sequential architecture as shown in Figure 4.1. Image pre-processing techniques are first used to eliminate haziness, color distortion, and uneven lighting. In order to gradually denoise and restore the data from turbidity, the images are subsequently run through a deep learning framework based on diffusion models.

Contrast Limited Adaptive Histogram Equalization (CLAHE) is used in the first stage to improve local contrast and rectify uneven lighting. Adaptive color correction (ACC) is used in the second step to equalize the image's overall brightness and color distribution. Lastly, a diffusion-based restoration model is used to decrease residual degradations such as noise, turbidity, and motion blur while recovering finer structural details. The contribution of each stage is summarized in Table 4.1.

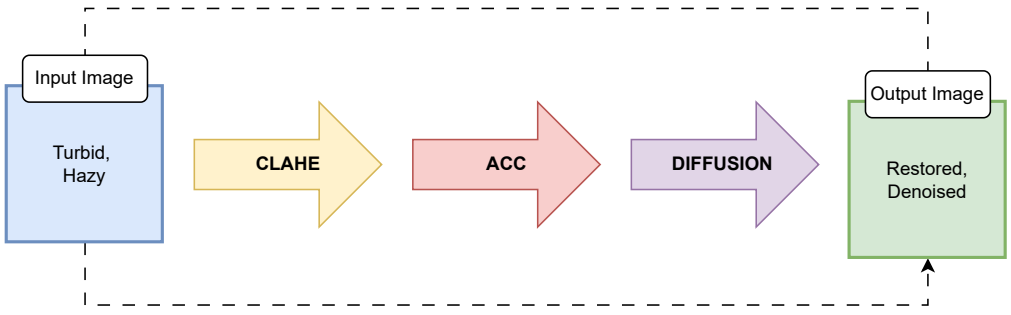

Figure 4.1: Processing pipeline illustrating our proposed model to enhance both turbid and hazy images through three distinct steps, Contrasting Limited Adaptive Histogram Equalization (CLAHE), Adaptive Color Correction (ACC) and, finally, while utilizing a diffusion restoration method to obtain an image that has been both restored and denoised.

Table 4.1: Explanation of enhancement and restoration techniques in the proposed framework

Technique	Main Purpose	Works On	Type
Contrast Limited Adaptive Histogram Equalization (CLAHE)	Enhancement of the **local contrast** and correction pf uneven illumination in smaller regions	Local pixel neighborhoods	Local enhancement (nonlinear histogram-based)
Adaptive Color Correction (ACC)	Balances and corrects the **global brightness and color distribution** across Red, Green and Blue channels	Entire image (global adjustment)	Global enhancement (linear and adaptive over time)
Diffusion Model	Restoration of **complex degradations** like turbidity, motion blur, and noise	Entire image (learned model)	Deep generative restoration

4.1 Contrast-Limited Adaptive Histogram Equalization (CLAHE)

Contrast-Limited Adaptive Histogram Equalization (CLAHE) is a technique used to enhance the visual quality of an image. The technique first divides an image into uniformly-sized tiles, and then applies local equalisation to improve the look of each tile separately. The steps used for carrying out CLAHE are outlined below:

1. **Divide the image into blocks:** The input image is split into smaller, uniformly-sized blocks (sometimes referred to as "tiles") so that the local contrast of each tile can be improved independently.

2. **Computing the histogram for each block:** For each of the tiles, a pixel intensity histogram is created that shows how many pixels in that tile have a certain brightness level. The histograms form the basis of the contrast adjustments.

3. **Clip the histograms:** A predefined threshold is imposed on all histogram bins to prevent over-amplification of contrast. Any histogram bin that exceeds this threshold will be clipped, and the clipped intensity levels will be uniformly redistributed

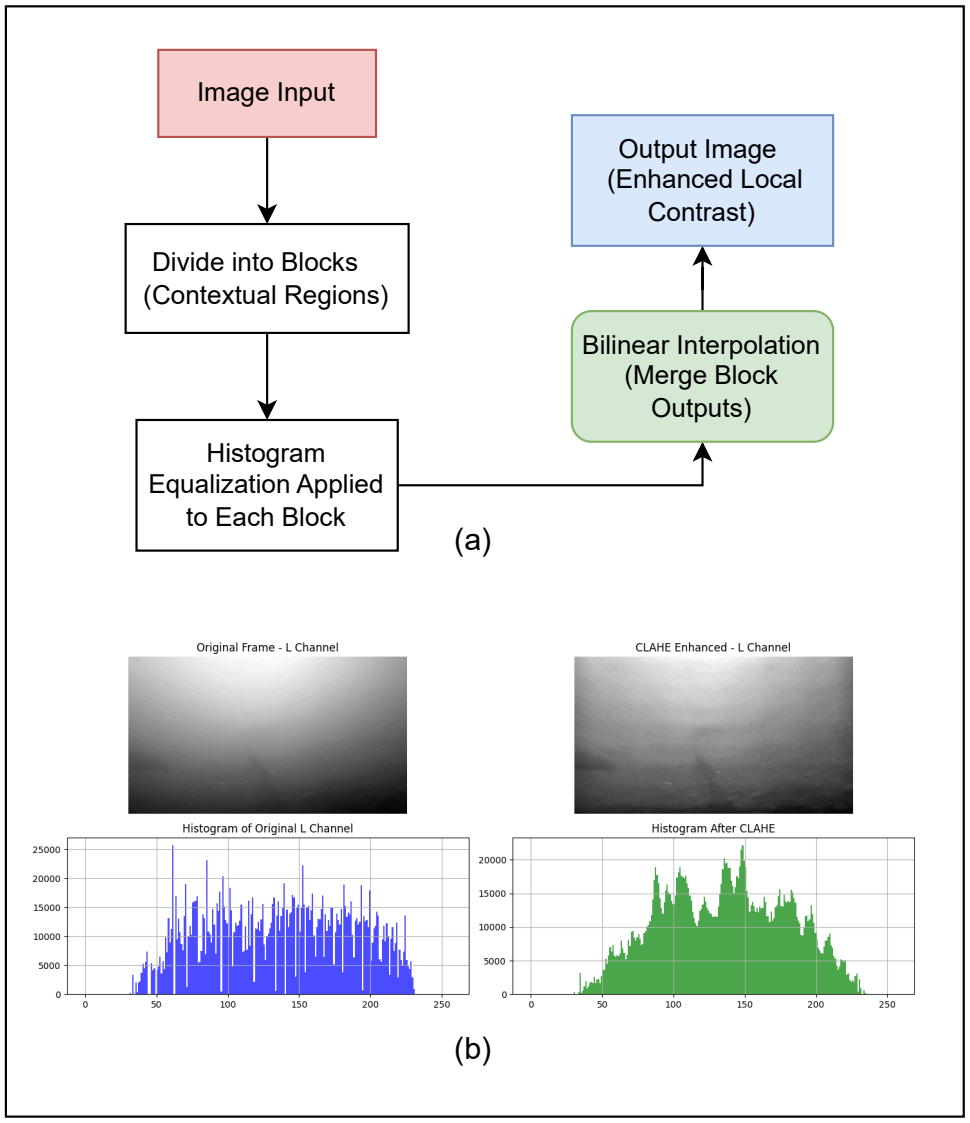

Figure 4.2: (a) A flowchart showing how Contrast Limited Adaptive Histogram Equalization (CLAHE) works. Each contextual region (block) in the input image is subjected to histogram equalization. The output image with better local contrast is then produced by combining the locally enhanced regions using bilinear interpolation. (b) Luminance channel (L-channel) image and related histogram comparison prior to and following CLAHE enhancement

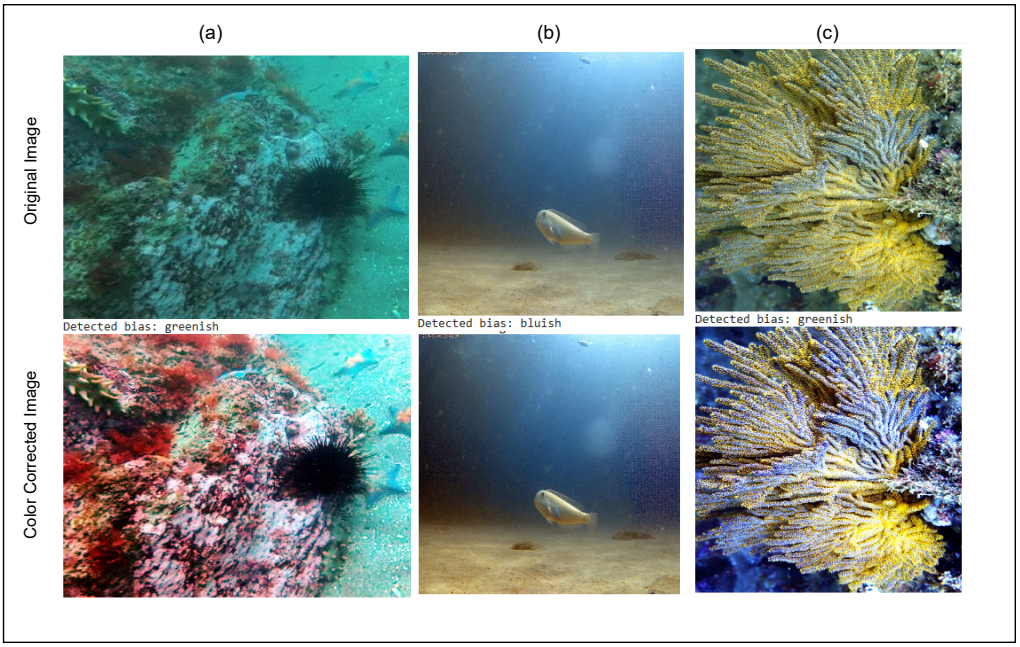

Figure 4.3: Three underwater scenes were subjected to adaptive color correction. The technique in (a) restores the absent warm tones and lessens the severe greenish hue. The algorithm in (b) adjusts for the bluish bias brought on by light absorption in deeper water. In (c), it maintains the coral's original texture while balancing the greenish, yellowish tint. In every instance, the Adaptive Color Correction method automatically detects the predominant color bias and generates an output that is more visually balanced and natural.

among all the bins. This clipping process reduces the amount of noise present in homogeneous sections of the image.

4. **Perform local histogram equalization on each block:** After the clipping process is complete, equalization of the histograms is carried out on each of the blocks separately. This equalization process redistributes the pixel intensities and enhances the local contrast, and reveals more details within each of the blocks.

5. **Apply bilinear interpolation between adjacent blocks:** Since the tiles are processed separately from one another, there can be abrupt transitions at the boundaries. Therefore, bilinear interpolation is applied to these transition areas to smooth the contrast adjustments and produce a seamless blending across all the adjacent tiles.

6. **Merge the processed blocks:** The processed blocks are then merged to reconstruct the enhanced image. The output image obtained as a result exhibits improved local contrast and detail visibility.

In short, this step first divides the image into non-overlapping blocks or contextual regions. Within each region, the histogram of pixel intensities is computed and then equalized as depicted in Figure 4.2.

4.1.1 Mathematical formulation

Consider an image that is divided into $M \times N$ blocks. For each block of this image, the histogram is given by $H(i)$, clipped at a maximum value H_{clip}. Its cumulative distribution function (CDF) is computed as:

$$\text{CDF}(i) = \sum_{j=0}^{i} H_{\text{clipped}}(j), \tag{4.1}$$

where $H_{\text{clipped}}(j)$ represents the clipped histogram count for intensity level j. The intensity of each pixel is mapped based on a normalized cumulative density function (CDF) of pixel values reconstructed from local areas in the image. Bilinear interpolation is applied between neighboring blocks to avoid abrupt intensity transitions.

CLAHE enhances the contrast in darker areas of the photograph while also preventing noise from becoming amplified. Therefore, it is a useful pre-processing step that perfectly leads to an adaptive color correction step and diffused-based restoration of a noisy image. Figure 4.2 illustrates this process and a comparison of the luminance channel of an underwater image with its histogram before and after CLAHE enhancement.

4.2 Adaptive Color Correction (ACC)

In stage two of our proposed framework, we work through the challenge of underwater color distortion created by the wavelength-dependent absorption and scattering of light. Blue and green, as they have short wavelengths, penetrate better in a water column than longer wavelengths (red and yellow). Therefore, a significant amount of underwater imagery exhibits an imbalance of color among wavelength ranges. To mitigate these effects, we integrate an adaptive color correction strategy [23] after contrast enhancement to restore a more natural and visually consistent color appearance.

The adaptive color correction approach starts by categorizing the input image into four groups, *bluish*, *greenish*, *yellowish*, and *color-balanced*. The categorization of each image is according to the dominant color characteristics of the background. We identify the background of the input image through the use of a quadtree-based recursive subdivision process [24]. In this process, we segment the imagery into blocks, determining the block with the lowest local variance and the maximum overall luminance. This block provides a reasonable approximation of the illumination and can therefore be used to calculate the attenuation of each channel.

4.2.1 Greenish images

For images with a dominant greenish hue, there is an uneven distribution of the grayscale levels in the RGB channels. This lack of balance can be corrected by normalizing the histogram of each channel. Let $I_c(x)$ denote the normalized pixel value (in $[0,1]$) of channel $c \in \{r,g,b\}$ at pixel position x. To establish consistency across channels, excessive green dominance is reduced and global color balance is

improved by equalizing each channel based on its cumulative histogram.

$$I_{r_c}(x) = I_r(x) + \alpha\left(\overline{I_g} - \overline{I_r}\right)(1 - I_r(x))\, I_g(x), \tag{4.2}$$

where $I_{r_c}(x)$ denotes the corrected red channel, and $\overline{I_r}$ and $\overline{I_g}$ represent the mean intensities of the red and green channels used to compensate for excessive green dominance.

$$I_{b_c}(x) = I_b(x) + \beta\left(\overline{I_g} - \overline{I_b}\right)(1 - I_b(x))\, I_g(x), \tag{4.3}$$

where $I_{b_c}(x)$ denotes the corrected blue channel, and $\overline{I_b}$ and $\overline{I_g}$ are the mean intensities of the blue and green channels guiding the degree of enhancement.

$$I_{g_c}(x) = I_g(x), \tag{4.4}$$

where $I_{g_c}(x)$ indicates the green channel, kept unchanged since $\overline{I_g}$ (its mean intensity) is already dominant in greenish underwater images. Here, α and β are positive scalar parameters that control the strength of the red- and blue-channel corrections, respectively. These scalars determine how strongly each channel is adjusted toward the mean intensity of the dominant green channel to compensate for the greenish color imbalance.

4.2.2 Bluish images

The red channel is usually severely attenuated in bluish images. In order to make up for this, the ACC approach uses a compensation model to improve the red channel intensity by referencing the green channel. The corrected red channel $I_{r_c}(x)$ is computed as:

$$I_{r_c}(x) = I_r(x) + \alpha\left(\overline{I_g} - \overline{I_r}\right)(1 - I_r(x))I_g(x), \tag{4.5}$$

where $\overline{I_r}$ and $\overline{I_g}$ are the mean red and green intensities within the background region, and α is a compensation parameter (set to 1). This formulation increases red-channel gain in proportion to its attenuation while preventing over-amplification in highly saturated regions.

After compensation, a dynamic range adjustment is applied to all channels to stretch their pixel intensities across the full displayable range. This is defined as:

$$I'_c(x) = K_c\left(I_c(x) - I_{c,\min}\right), \tag{4.6}$$

where

$$K_c = \frac{255}{D_c}, \quad D_c = I_{c,\max} - I_{c,\min}, \tag{4.7}$$

and $I_{c,\min}$, $I_{c,\max}$ are the minimum and maximum intensities of channel c after compensation. This step ensures improved contrast, channel balance, and tonal distribution.

4.2.3 Yellowish images

When it comes to yellowish images, they suffer from the opposite issue. Here, the green and blue channels are typically weakened relative to the red channel. In this instance, both the green and blue channels are compensated for, with the red channel acting as the reference:

$$I_{b_c}(x) = I_b(x) + \beta\left(\overline{I_{rg}} - \overline{I_b}\right)(1 - I_b(x))\,\frac{I_r(x) + I_g(x)}{2}, \tag{4.8}$$

$$I_{r_c}(x) = I_r(x) + \alpha\left(\overline{I_b} - \overline{I_r}\right)(1 - I_r(x))\,I_b(x), \tag{4.9}$$

$$I_{g_c}(x) = I_g(x) + \gamma\left(\overline{I_b} - \overline{I_g}\right)(1 - I_g(x))\,I_b(x), \tag{4.10}$$

where x denotes a spatial pixel location, and all image intensities are assumed to be normalized to the range $[0,1]$. $I_r(x)$, $I_g(x)$, and $I_b(x)$ represent the original red, green, and blue channel intensities at pixel location x, respectively. $I_{r_c}(x)$, $I_{g_c}(x)$, and $I_{b_c}(x)$ denote the corresponding color-corrected red, green, and blue channel intensities. The other quantities like $\overline{I_r}$, $\overline{I_g}$, and $\overline{I_b}$ denote the mean intensity values of the red, green, and blue channels, respectively, computed over the entire image. The reference yellow-channel intensity is defined as

$$\overline{I_{rg}} = \frac{\overline{I_r} + \overline{I_g}}{2}. \tag{4.11}$$

The parameters α, β, and γ control the strength of the red, blue, and green channel compensation, respectively, and are set to 1 in this work. By strengthening channels

that are excessively muted in yellowish settings, this adjustment enhances the overall spectral balance.

4.2.4 Color-balanced images

An image is considered color-balanced if the background region's three-channel mean intensities are comparable. In these situations, the image moves straight on to the next enhancement stage, and the ACC module does not apply any additional correction.

The results from the adaptive color correction are shown in Figure 4.3. Figure 4.4 provides a visual overview of the ACC process and its integration with the CLAHE-based contrast enhancement module.

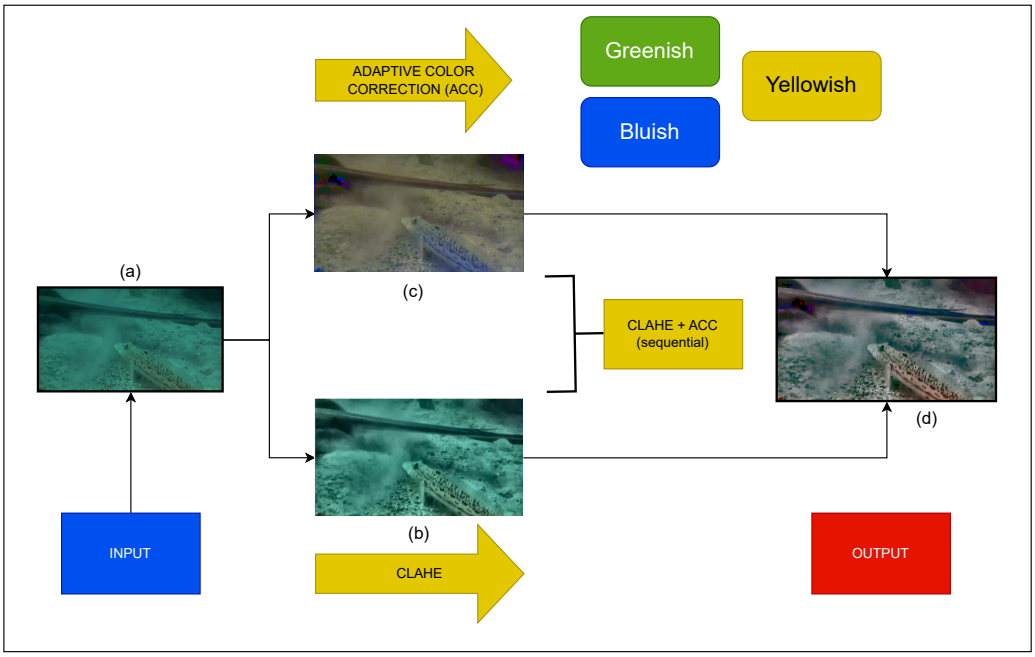

Figure 4.4: Adaptive Color Correction (ACC) process: (A) Original image taken underwater. (B) CLAHE has increased Local Contrast (local contrast enhancement) in a local region. (C) The ACC Process determines whether an underwater image is classified as 'greenish', 'bluish', or 'yellowish' then appropriate compensation adjustments are used to correct the problems that exist with the Underwater image. (D) By applying combined CLAHE + ACC, a visually restored image is achieved with improved naturalness and clarity of the colors.

4.3 Diffusion Models

The last phase of the proposed structure utilizes a diffusion model to eliminate murkiness and retrieve small structural features from underwater photos.

Two processes are used by diffusion models, the *forward diffusion process*, where an image is gradually converted from a clean image to a turbid version with added noise based on a planned rate of deterioration, and the *reverse diffusion process*, a model that learns how to reverse damage from the forward diffusion process by estimating what type of noise statistics will be the most precise. Figure 3.1 depicts both the forward process and the model's operation in reconstructing accurate images from noisy input.

4.3.1 Probabilistic vs data-driven diffusion models

Probabilistic diffusion models start by randomly sampling the noise variance parameters and the model estimation is made during the reverse diffusion process. Meanwhile, data-driven diffusion model, in this case, starts from the turbidity-graded dataset and extracts the noise variance parameter by one-on-one correspondence.

4.3.2 Noise parameter estimation from real underwater data

The standard denoising diffusion probabilistic model (DDPM) formulation assumes a fixed Gaussian noise schedule. However, underwater imaging presents noise that is highly variable and dependent on turbidity level. To get rid of this issue, the statistical noise parameters, mean (μ_t) and variance (σ_t^2), are estimated directly from real underwater images to ensure realistic modeling.

A controlled tank-based experiment described in [25] is used to replicate variable turbidity conditions. This data is created by using different concentrations of suspended particles matter, like sand and turbid water. Images are then captured with multiple levels of turbidity. This enables a direct correspondence between degraded images and the underlying noise distribution. Figure 4.5 shows these progressively increasing noisy images.

Let x_0 denote the clean underwater image and x_t the corresponding degraded image at diffusion timestep t. The forward diffusion process progressively introduces degradation by adding Gaussian noise to the clean image, which can be expressed as

$$x_t = \sqrt{\alpha_t}\, x_0 + \sqrt{1-\alpha_t}\, \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, I), \tag{4.12}$$

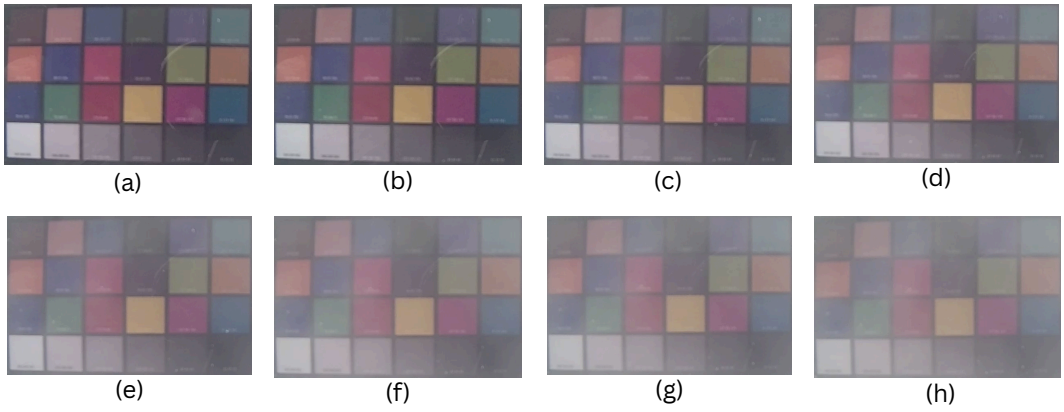

Figure 4.5: From (a) to (h), images from the dataset have increasing turbidity. This linear progression of noise in images replicates the forward diffusion noise addition process. Using this dataset, mean and variance values for the creation of synthetic images are estimated.

where x_t represents the degraded image at step t, and ε_t denotes zero-mean, unit-variance Gaussian noise sampled from a normal distribution. The coefficient $\alpha_t \in (0,1)$ controls the fraction of the original signal preserved at each diffusion timestep.

Conventional denoising diffusion probabilistic models rely on an assumed fixed noise schedule of the model; however, this assumption becomes untenable here because the degradation characteristics of images strongly depend on the amount of turbidity in the water column. To circumvent these assumptions, we shift our focus of attention from learning noise schedules to actually modeling the noise statistics directly from real-world underwater photographs. Specifically, the proposed model estimates the mean μ_t and variance σ_t^2 of the noise distribution at each diffusion step t, enabling adaptive and physically consistent modeling of turbidity-induced degradation. Given a dataset of real underwater images, these parameters are optimized using the following loss function:

$$\mathscr{L}_{\text{noise}} = \mathbb{E}_{\mathbf{x}_0,\boldsymbol{\varepsilon}_t,t}\left[\|\boldsymbol{\varepsilon}_t - \hat{\boldsymbol{\varepsilon}}_t(\mathbf{x}_t,t;\theta)\|_2^2\right], \tag{4.13}$$

where:

- $\hat{\boldsymbol{\varepsilon}}_t(\mathbf{x}_t,t;\theta)$ is the noise predicted by the U-Net model with parameters θ,
- $\|\cdot\|_2^2$ denotes the mean squared error (MSE) between the predicted and actual noise.

This loss guides the model to correctly deduce the underlying noise properties that exist at every stage of the diffusion process, as illustrated in Figure 4.6. To produce synthetic noisy images and enhance subsequent restoration performance, the model must be able to replicate underwater degradation patterns.

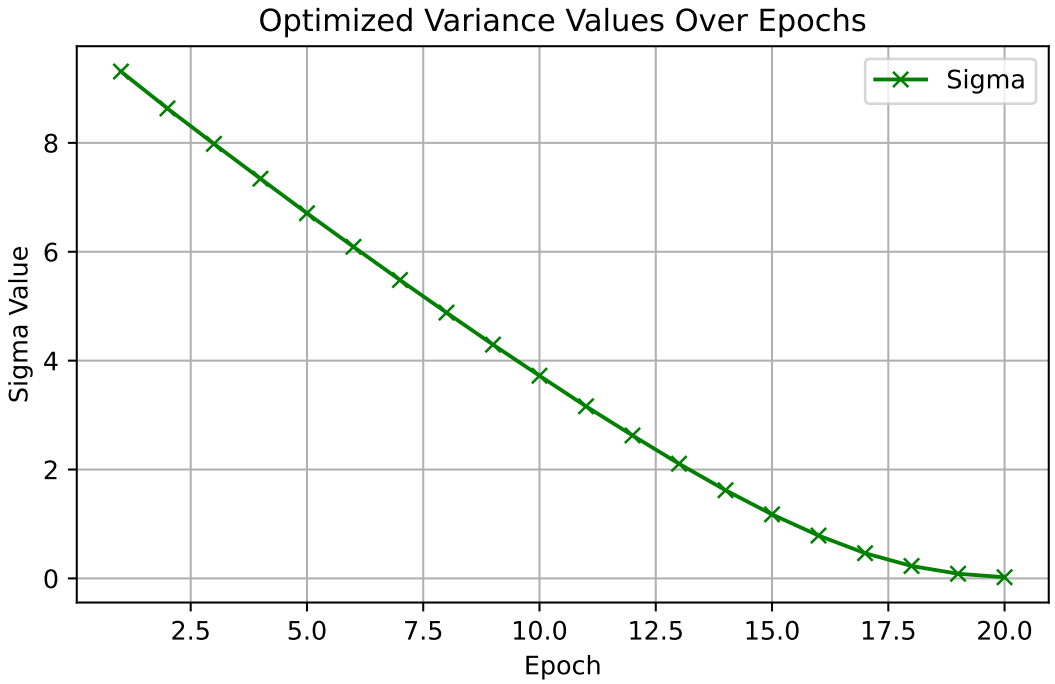

Figure 4.6: Variance estimation for the turbidity noise with respect to the training epochs. Initially, the sigma value is kept as 10 with each epoch; the training loss reduces, and finally, at 20 epochs, we get an estimated variance value.

4.3.3 Forward diffusion for synthetic dataset generation

Each clean underwater image is subjected to forward diffusion using the empirically derived noise parameters. Figure 4.7 depicts the synthetic noisy images that are produced across a number of diffusion steps. This results in a dataset where the turbidity progressively rises with the timestep index. The model may see a wide range of noise intensities and spatial patterns since each clean image produces multiple noisy variations (Figure 3.1).

By training the reverse diffusion model using realistic underwater noise that closely resembles natural turbidity, this synthetic dataset enhances the model's capacity to generalize beyond the initial training set.

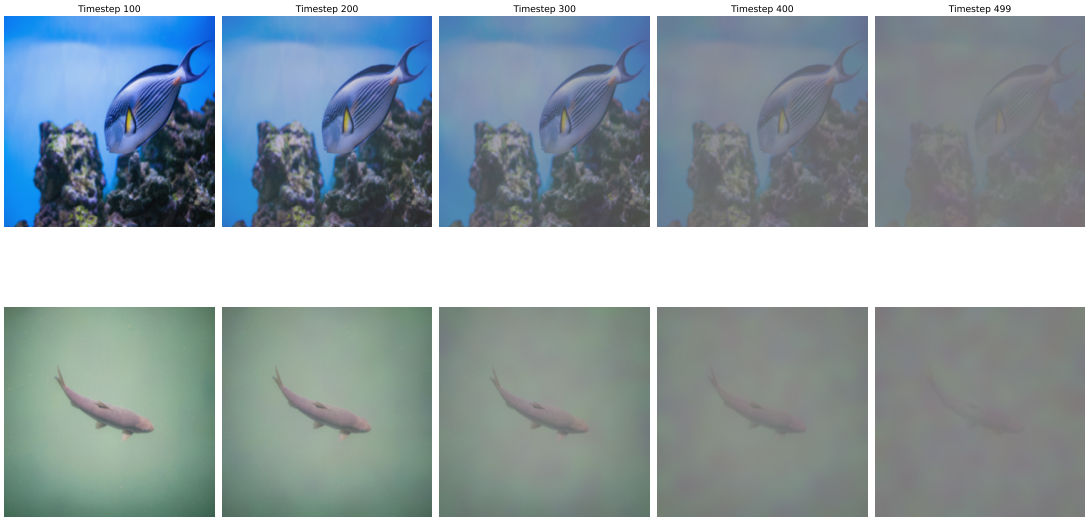

Figure 4.7: The diffusion-based U-Net model is trained using progressive turbidity simulation. Synthetic underwater photos produced during diffusion time-steps (100, 200, 300, 400, 499) are displayed in both rows, demonstrating progressive haze, color attenuation, and scattering. This sequence shows how the model improves noise estimation and restoration performance by learning the transition from clear to severely degraded circumstances.

4.3.4 Reverse diffusion for image denoising

The reverse diffusion process attempts to invert the forward process by reconstructing $\mathbf{x}_0$ from $\mathbf{x}_t$. At each timestep, the model predicts the mean $\mu_\theta(\mathbf{x}_t, t)$ and variance $\sigma_\theta^2(t)$ of the conditional distribution $p_\theta(\mathbf{x}_{t-1} \mid \mathbf{x}_t)$, enabling iterative noise removal:

$$\mathbf{x}_{t-1} = \mu_\theta(\mathbf{x}_t, t) + \sigma_\theta(t)\,\mathbf{z}, \qquad \mathbf{z} \sim \mathcal{N}(0, \mathbf{I}). \tag{4.14}$$

where

- $\mathbf{x}_t$ is the noisy image at timestep t in the reverse diffusion process.
- $\mathbf{x}_{t-1}$ is the denoised (or less noisy) image predicted for timestep $t-1$.
- $\mu_\theta(\mathbf{x}_t, t)$ is the predicted mean of the reverse diffusion distribution, produced by the neural network with parameters θ.
- $\sigma_\theta(t)$ is the predicted standard deviation (noise level) for timestep t, also learned by the model.
- $\mathbf{z}$ is a random noise sample drawn from a standard normal distribution.

- $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I})$ indicates that $\mathbf{z}$ follows a Gaussian distribution with zero mean and identity covariance.
- θ represents the learnable parameters of the neural network (e.g., U-Net weights).

The reverse model is implemented using a U-Net architecture, which is well-suited for capturing multi-scale features. Skip connections ensure that fine-grained details lost at deeper layers are preserved during reconstruction (Figure 3.1: b–c).

4.3.5 Training procedure

The full diffusion model is trained end-to-end using supervised noise prediction. The training procedure consists of the following steps:

1. **Initialization of Noise Schedule:** A β_t schedule is defined to determine α_t values, controlling the noise level introduced at each diffusion step.
2. **Forward diffusion:** Gaussian noise is progressively added to each input image according to the estimated noise parameters, producing noisy samples $(\mathbf{x}_t)$.
3. **Noise prediction:** The U-Net predicts the noise component injected at step t, serving as an estimate of the perturbation.
4. **Loss optimization:** The mean squared error (MSE) between predicted and true noise is minimized. The Adam optimizer is used for stable convergence.
5. **Reverse diffusion:** Once trained, the model performs iterative denoising from timestep T to 0, reconstructing a clean image from its noisiest representation.

In the U-Net structure, there are four encoder and decoder stages, and there is a connection between the encoder and decoder stages that is called skip connections. In the encoder and decoder stages, there are five 3×3 convolutional filters and five ReLU activation functions are used; transposed convolutions are used to create the upsampling layers. The highest number of filters used in the bottleneck layer is 1024, and therefore, this layer contains the semantic information of the image to a high level. At the end of the reverse process of denoised image creation, the final output is produced from the denoised image 1×1 filter to the RGB color space.

5. DATA SET SPECIFICATIONS

When evaluating the performance of the proposed underwater image restoration framework, it is essential to select datasets that provide both real-world variability and controlled turbidity levels.

To do this, we used two complementary benchmark datasets that served as good example data to test our model. The first dataset is called the Underwater Image Enhancement Benchmark Dataset (UIEBD), which consists of a large collection of images taken in the natural environment with an extensive system to choose reference images. This dataset provides variability found in real-life conditions.

The other dataset is the Turbidity Underwater Dataset (TUDS), which was created in a controlled laboratory setting and allows us to study how degradation occurs with increasing turbidity in water. Therefore, this dataset helps us in examining how well our underwater image restoration framework can handle the different levels of degradation.

This section will have additional technical information about both datasets, how they were constructed, characteristics, and the ways both datasets will be useful for research and development of new underwater image restoration techniques. A summary of the properties of both datasets is described in Table 5.1.

5.1 Underwater Image Enhancement Benchmark Dataset (UIEBD)

The development of the Underwater Image Enhancement Benchmark Dataset occurred in three major steps: acquisition of varied types of underwater imagery, application of multiple enhancement methodologies, and the selection of reference images based upon feedback from observers.

5.1.1 Data collection and diversity

A primary objective of UIEBD [2] is to meet three requirements: (1) Collect a variety of underwater images taken in different locations, with varying levels of degradation;

(2) Acquire enough images to create a representative sample population for comparison purposes; (3) Provide the opportunity to compare results against high-quality enhanced reference images.

The pictures used to achieve these objectives were gathered from various sources, including Google Images, YouTube, academic datasets relevant to this topic, and original underwater photographs. A comprehensive set of pictures was retrieved and refined based on deterioration and scene type. The final collection includes around 950 underwater pictures that were taken under a variety of lighting conditions. It contains several categories of underwater scenes, including coral reef structures, marine animals (such as turtles, fish, and sharks), in-water objects, and differently configured seascapes.

5.1.2 Reference image generation

A distinguishing characteristic of UIEBD is the availability of reference images for 890 of the 950 samples, which are of high quality. The reference images were created using twelve different techniques for improving images, which included nine underwater-specific techniques as well as two dehazing techniques: Dark Channel Prior (DCP) and Multi-Scale Convolutional Neural Network (MSCNN).

A total of twelve enhanced versions of each raw image were produced by applying all techniques separately. To determine the best-enhanced image, a pair-wise human assessment study was conducted on 50 individuals, with 25 having an understanding of image processing and 25 without. The individuals compared each pair of enhanced images in a series, selecting the one that they thought was the highest quality image each time until only one image remained.

This final reference image was then validated through individual ratings by those 50 volunteers; if the image was rated unsatisfactory (more than half of the participants rated it that way), it was moved to a set of images that were considered difficult to process (60 images total).

The resulting dataset, therefore, consists of:

- 890 raw underwater images with corresponding high-quality human-validated reference images,

- 60 challenging underwater images without any valid references consisting of extreme degradation.

UIEBD is the first extensive, real-life underwater dataset with dependable reference photos that can be used for training CNN, complete-reference assessment (PSNR, SSIM), and investigating perceived quality.

5.2 Turbidity Underwater Dataset (TUDS)

This turbidity dataset was designed by introducing a controlled laboratory-scale environment that simulates progressive turbidity levels. Unlike UIEBD, which focuses on real-world natural variability, TUDS emphasizes systematic turbidity manipulation to support model testing and turbidity-dependent restoration.

5.2.1 Laboratory environment and turbidity control

A specially designed experimental system was built to model underwater conditions, ensuring a one-for-one physical representation. The water tank was built and the turbidity level was controlled using materials like clay and suspended particles. Through gradually increasing turbidity levels, the researchers were able to create a series of conditions known as a turbidity gradient from slightly hazy to very opaque scene conditions as shown in Figure 4.5.

This data provides us with the ability to conduct comparative analysis of both direct and indirect factors, such as attenuation, light scattering, and density of particulates. In short, this setup enabled objective analysis by studying the impacts of turbidity on image visibility and degradation.

5.2.2 Real-world video capture and frame extraction

In order to provide ecological relevance for our study, we also obtained more video data of the underwater part of the vertical-slot fish passage system at Çatlarlolk HEPP on the Tekir Stream in Kahramanmaraş–Ceyhan Basin. This footage was obtained using two underwater cameras affixed to the system that were set to different exposure settings.

The underwater conditions captured by these cameras included floating debris, fish swimming around, changing light levels, and sediment-laden waters

The final dataset consisted of randomly selected 400 video frames from a total of 53 five-minute recordings taken from the two cameras. These frames represent some of the most difficult visibility scenarios, including:

- extreme turbidity and high particulate density,
- low-light and non-uniform illumination,
- motion blur from water flow and fish movement,
- color distortion due to suspended sediments.

Because TUDS lacks ground-truth reference images, it is ideal for evaluating model performance and generalization under harsh underwater conditions by using no-reference metrics.

5.3 Evaluation Rationale

For the assessment of our suggested underwater enhancement framework, UIEBD and TUDS play complementary roles.

- **UIEBD** allows us to deliver a controlled comparison against reference images. It supports the qualitative analysis of color fidelity and detailed recovery of typical underwater scenes.
- **TUDS** consists of extreme degradation conditions, leading to the evaluation of the model performance using no-reference perceptual metrics.

With these datasets combined, we can provide a complete assessment across the full spectrum of underwater visibility. Images range from natural environments with available references to those captured in laboratory-controlled, as well as extremely turbid fish tank scenarios.

Table 5.1: Comparison of UIEBD and TUDS datasets used for evaluating the proposed underwater enhancement framework.

Characteristic	UIEBD	TUDS
Type of Data	Real-world underwater photographs	Controlled-tank and real-world video frames
Total Images	$\sim$950 images (890 with reference, 60 without)	400 frames from 53 videos
Ground Truth Availability	Human-validated high-quality reference images for 890 samples	No reference images; highly degraded turbidity scenarios
Degradation Type	Color casts (green, yellow, blue), moderate haze, natural attenuation	Severe turbidity, visibility loss, low-light noise, motion blur
Capture Conditions	Natural underwater scenes, diverse lighting and content	Lab-controlled turbidity tank + fish-pass camera recordings
Primary Use Cases	Color correction, enhancement, full-reference comparison	Robustness testing, turbidity-aware evaluation, no-reference metrics
Access and Licensing	Academic, non-commercial use	Released for research with citation requirement

6. RESULTS AND EVALUATION

The two distinct datasets with differing degrees of turbidity and deterioration were used to test and execute the suggested methodology. Figures 6.1, 6.2 represent the sequential results for each respective dataset.

6.1 Underwater Image Enhancement Benchmark Dataset (UIEBD)

Chapter 5 has comprehensively discussed that the UIEBD [2] includes underwater images exhibiting a wide range of color distortions, such as greenish, yellowish, and bluish haze. Figure 6.1 shows the outputs that are generated by the proposed enhancement framework.

Example images Figure 6.1 (a) and (d) that were primarily green showed that the combination of Contrast Limited Adaptive Histogram Equalization (CLAHE) and Adaptive Color Correction (ACC) can reduce the bias created by green illumination in those images. However, for more turbid examples like (b) and (c), the use of CLAHE allowed for a significant increase in the local contrast while also improving the visibility of structural detail. At the final stage of the pipeline, an integrated version of CLAHE, ACC and a diffusion-based restoration method consistently produced the best and most balanced perceptual quality results. The combined effect of the three restoration techniques led to the most visually enhanced results for all images within the dataset.

6.2 TUDS Dataset

To determine the efficacy of the developed method in severe turbidity environments, the TUDS dataset [3] was utilized. The dataset comprised in-situ captured aquatic imagery with moving fish and possessed inherent difficulties, including motion blur and low-light environment. For images with medium turbidity, such as Figure 6.2(d), the combined ACC+CLAHE step efficiently corrected color casts, and the subsequent diffusion-based restoration further improved visual quality.

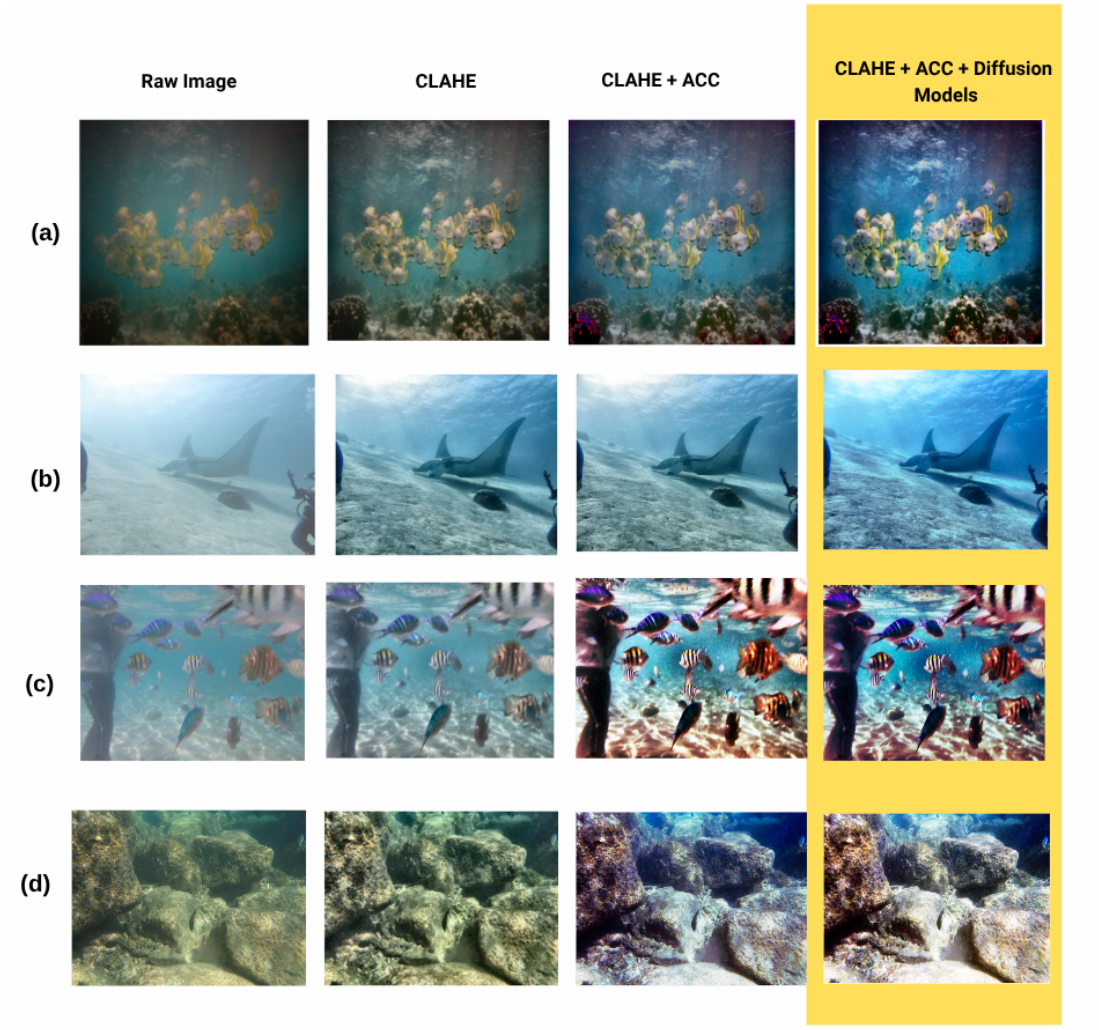

Figure 6.1: Comparative images of enhanced under-water images on UIEBD [2]. Figure shows the processing stages: Raw Image, A CLAHE enhancement, A CLAHE enhancement + an Adapted Colour Correction, and A gradual workflow of CLAHE + Adapted Colour Correction + Diffusion Model, from left to right on the images. The images (a)–(d), show four separate underwater scenes with their respective improvements in the visual clarity, color balance, and contrast compared to the raw images achieved by our method.

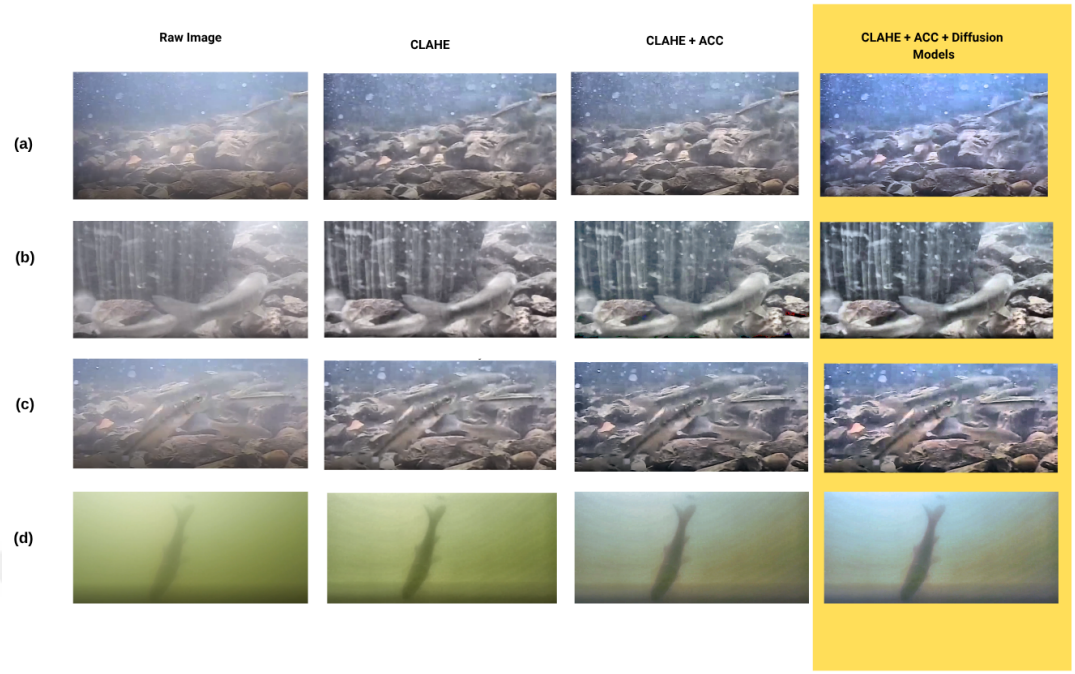

Figure 6.2: Results of underwater image enhancement using TUDS [3]. The columns indicate the different processing stages for the selected image, from left to right: Raw image, Contrast Limited Adaptive Histogram Equalization (CLAHE), CLAHE with adaptive color correction (ACC), and the full enhancement pipeline consisting of CLAHE, ACC, and diffusion model. The rows (a-d) show images with different levels of turbidity and color casts, with their four processing stages. The results show improved color contrast and recovery of detail at each stage. The final stage produces the best result both visually and perceptually in terms of color balance and enhancement.

The diffusion model, which uses estimated turbidity coefficients to help denoise the image, was used to restore images with extreme turbidity, such as Figure 6.2 (a) and (b). By denoising the images according to the anticipated turbidity parameters, the diffusion model demonstrated a notable restoration success.

6.3 Evaluation

The lack of corresponding ground truth photos is the primary reason why objective analysis remains a difficult undertaking in the evaluation of underwater images. In these situations, it is not possible to immediately apply traditional full-reference measurements like the Structural Similarity Index Measure (SSIM) and the Peak Signal-to-Noise Ratio (PSNR).

Therefore, this study employs perceptual and statistical quality measures, including Entropy ($H(X)$), Color Accuracy (ΔE), the Naturalness Image Quality Evaluator (NIQE), Fréchet Inception Distance (FID), and Kernel Inception Distance (KID), to quantitatively assess restoration performance. For quantitative evaluation, all image-wise metrics, including NIQE, entropy, and the color error metric ΔE, are computed independently for each enhanced image. The final performance values for each dataset are then obtained by averaging the corresponding metric scores over all images, which ensured a fair and unbiased dataset assessment.

Table 6.1: Performance metrics for underwater image restoration on UIEBD and TUDS

Dataset	Raw	CLAHE	CLAHE+ACC	Diffusion
Entropy ($H(X)$)				
UIEBD	5.6	6.9	7.5	7.8
TUDS	5.3	6.5	6.9	7.1
Color Accuracy (ΔE)				
UIEBD	8.2	5.2	4.1	3.7
TUDS	8.7	5.8	4.7	4.4
NIQE				
UIEBD	5.2	3.1	2.6	2.0
TUDS	5.6	3.5	3.0	2.8
FID				
UIEBD	35.8	18.2	12.3	9.4
TUDS	38.9	22.4	16.1	14.7
KID ($\times 10^{-3}$)				
UIEBD	21.3	10.8	8.2	6.1
TUDS	23.7	12.5	9.6	8.3

6.3.1 Metric definitions

(a) Entropy ($H(X)$): Entropy [26] quantifies the amount of information retained in an image and is defined as:

$$H(X) = -\sum_{i} P(i) \log_2 P(i), \tag{6.1}$$

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

(b) Color Accuracy (ΔE): Color fidelity is evaluated in the CIE-Lab color space using:

$$\Delta E = \frac{1}{n}\sum \sqrt{(L^*_{pred} - L^*_{ref})^2 + (a^*_{pred} - a^*_{ref})^2 + (b^*_{pred} - b^*_{ref})^2}, \tag{6.2}$$

where (L^*, a^*, b^*) denote the lightness and chromaticity components. Lower ΔE values represent more accurate color reproduction.

(c) NIQE (Naturalness Image Quality Evaluator): NIQE [27] assesses perceptual image quality by quantifying the deviation of natural scene statistics from those of the test image:

$$NIQE = \sqrt{(f-\mu)^T \Sigma^{-1} (f-\mu)}, \tag{6.3}$$

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.

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

$$FID = ||\mu_r - \mu_g||_2^2 + Tr(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2}), \tag{6.4}$$

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.

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

$$KID = \mathbb{E}\big[k(f_r, f'_r)\big] + \mathbb{E}\big[k(f_g, f'_g)\big] - 2\mathbb{E}\big[k(f_r, f_g)\big], \tag{6.5}$$

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.

6.3.2 Discussion on evaluation metrics

The quantitative results presented in Table 6.1 and Figure 6.3 provide insights into the performance of different enhancement stages across the UIEBD and TUDS datasets.

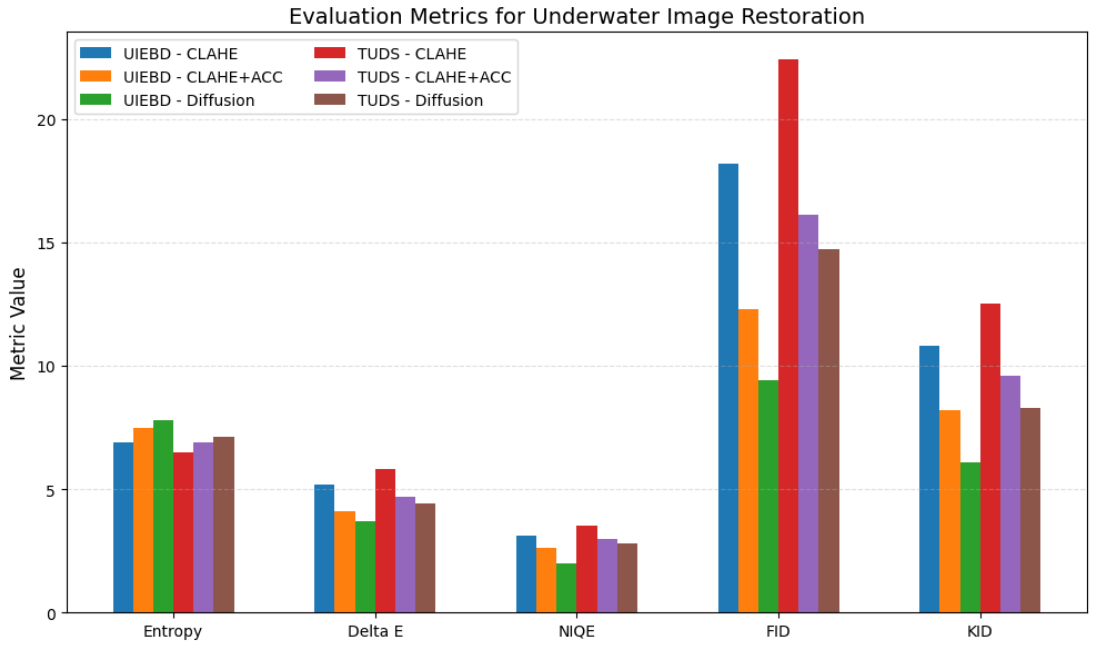

Figure 6.3: Quantitative comparison of enhancement stages on the UIEBD and TUDS datasets using entropy, color accuracy, and perceptual quality metrics.

6.3.2.1 Entropy ($H(X)$)

An image's information richness is measured by its entropy. Because details are obscured by haze and turbidity, raw images have the lowest entropy. Entropy in UIEBD is raised from 6.9 to 7.5 by CLAHE, and then it is further improved to 7.8 by the CLAHE + ACC stage. Entropy in TUDS increases from 6.5 (raw) to 6.9 with CLAHE and then to 7.1 with ACC. In both datasets, the diffusion-based model achieves the highest entropy, demonstrating its capacity to suppress noise while restoring subtle textures and structural details.

6.3.2.2 Color accuracy (ΔE)

Lower ΔE values indicate closer alignment with natural colors. For UIEBD, CLAHE reduces ΔE from 5.2 to 4.1, and CLAHE + ACC lowers it further to 3.7. The diffusion model slightly improves it beyond this stage. In TUDS, which is more challenging due to turbidity and low light, ΔE decreases from 5.8 (raw) to 4.7 (CLAHE + ACC) and 4.4 after diffusion-based restoration. These results highlight ACC's effectiveness

in correcting dominant color casts, while the diffusion model fine-tunes residual discrepancies.

6.3.2.3 NIQE

NIQE assesses perceptual naturalness; lower values indicate images closer to natural scene statistics. For UIEBD, NIQE decreases from 3.1 (CLAHE) to 2.6 (CLAHE + ACC) and reaches 2.0 after diffusion restoration. TUDS images follow a similar trend: $3.5 \rightarrow 3.0 \rightarrow 2.8$. The diffusion model consistently produces the lowest NIQE, demonstrating enhanced perceptual quality and naturalness.

6.3.2.4 FID and KID

FID and KID quantify the similarity of restored images to reference distributions. In UIEBD, FID drops from 18.2 (CLAHE) $\rightarrow$ 12.3 (CLAHE + ACC) $\rightarrow$ 9.4 (diffusion), while KID decreases from $10.8 \times 10^{-3} \rightarrow 8.2 \times 10^{-3} \rightarrow 6.1 \times 10^{-3}$. TUDS shows higher values due to stronger turbidity and motion blur: FID decreases $22.4 \rightarrow 16.1 \rightarrow 14.7$, and KID $12.5 \times 10^{-3} \rightarrow 9.6 \times 10^{-3} \rightarrow 8.3 \times 10^{-3}$. These trends confirm that the diffusion-based model produces the most perceptually accurate and statistically natural images in both datasets.

Overall, CLAHE enhances local contrast and details, ACC substantially improves color fidelity, and the diffusion-based model provides comprehensive restoration, improving both structural and perceptual quality. The higher metric values observed in TUDS reflect its more challenging conditions, including higher turbidity, motion blur, and low-light scenarios.

6.4 Processing Time Analysis

Computational Efficiency (i.e., the overall speed at which data can be processed) is another key consideration in the implementation of underwater image enhancement systems, along with restoration quality. In this section, we analyze how long it takes to process images at each phase (Stage) of the proposed framework, as well as how the

resolution (size) of the image used for enhancement affects the processing time for each phase.

6.4.1 Experimental Setup

Each component of the enhancement pipeline was measured separately, including CLAHE, ACC, and the diffusion-based restoration model.

The timing results do not account for the time spent loading, saving, or visualizing images because they were compiled based only on the behavioral inference performance of each operation.

The reported processing time represents the average per-image runtime when images are processed at their native dataset resolutions. The computational cost usually scales based on the resolution of the images. The values represented here are considered representative of the UIEBD and TUDS datasets, but not guaranteed as absolute runtimes for processing at all resolution levels.

6.4.2 Per-Stage Processing Time

Table 6.2 summarizes the average processing time per image for each enhancement stage on the UIEBD and TUDS datasets.

Table 6.2: Average processing time per image at native dataset resolution

Processing Stage	**UIEBD (s)**	**TUDS (s)**
CLAHE	0.07	0.09
CLAHE + ACC	0.12	0.15
Diffusion Model	1.88	2.10
Total Pipeline	**1.97**	**2.29**

The least expensive stage of the CLAHE process is the one that uses localized histogram functions and very simple redistribution of contrast. When the ACC is added, the processing cost increases somewhat due to the requirement for more transformations from the original color space and additional normalization (per channel), but still produces a cost-effective result.

The majority of the time required to complete the restoration process is attributed to the diffusion-based model. The time increase in processing that occurs with this model

results from the large number of convolutional operations that are required at each of the many steps that occur when denoising and turbidity removal of an image. However, for a region or image with moderate or severe turbidity (e.g., the TUDS dataset), increased complexity will result in an increase in processing time as well.

6.4.3 Resolution Dependency and Discussion

As the resolution of an image increases within the proposed framework, so does the computational cost. With regard to both CLAHE and ACC, the time it takes to process these two types of algorithms is approximately linear in relation to the number of pixels contained within each image, as both algorithms use local image neighborhoods to operate and then apply pixel-based transformations.

The diffusion-based restoration model primarily relies on convolutional feature extraction and iterative denoising, which means that the computational demands increase significantly as the resolution of an image increases. More operations are required for each diffusion step, and therefore, the amount of time needed to create a final inference will be longer with an increase in the resolution of the image. Once again, the performance characteristics of the deep generative models used for image restoration are the same in that there is always a trade-off between the quality of the results produced by a restoration algorithm and the efficiency with which that restoration is performed.

For video-based applications, the time needed to process a video is also directly correlated to the number of frames within the video, as the proposed framework processes each frame independently. However, in order to increase efficiency, a batch processing or parallel inference strategy could be applied to allow for faster processing time. Therefore, the proposed framework is ideally suited for both offline and semi-offline applications that deal with underwater imagery.

7. CONCLUSION

In summary, the main goal of this study was to restore underwater images that had been distorted by low light, turbidity, color loss, and motion blur. Enhancing visual clarity, improving detail recovery, and correcting color distortions present in actual underwater scenarios were the objectives of this work. This study used a combination of advanced deep learning techniques and conventional enhancement procedures to accomplish these goals.

7.1 Important Aspects

Here, we have summarized the most important aspects of our work that helped us in obtaining visually enhanced output images:

- **CLAHE (Contrast Limited Adaptive Histogram Equalization):** is a useful pre-processing technique for enhancing local contrast and improving the appearance of fine details that may have been hidden by turbidity or haze. By enhancing the structural composition and texture visibility of underwater photos, it results in a notable improvement.
- **ACC (Adaptive Color Correction):** demonstrated a strong ability to correct color casts, producing more color-fidelity and aesthetically pleasing images. Photos with a lot of green or blue tones are particularly well-suited for ACC.
- **Diffusion-based Model:** From the results obtained, it can be seen that our pretrained diffusion model based on simulating real-life turbidity showed superior restoration performance. By effectively denoising and reconstructing underwater images while maintaining both structural integrity and perceptual naturalness.
- **Usage of Iterative Noisy Dataset to Estimate Noise Parameters:** The use of a noisy dataset created in the lab with turbidity that increases iteratively with each image is another innovative contribution of our work. We were able to precisely

estimate the noise parameters that helped determine the architecture of the diffusion model by training our model on these kinds of images.

7.2 Contribution Towards Existing Literature:

Looking at the literature, previous research conducted in the field of underwater image enhancement primarily focused on eliminating one of the three major artifacts among turbidity, motion blur, and color distortion. This work is a step forward because it addresses each of these three issues one after the other while concentrating on them all at once.

Secondly, our work bridges the gap between contemporary deep learning architectures and traditional pre-processing techniques found in previous works. In order to create a reliable, high-performing architecture, this contribution builds upon earlier research.

Lastly, evaluation metrics like entropy, color accuracy, and the Naturalness Quality Indicator (NIQE) are presented in our work. In order to objectively assess output images without ground truth, these metrics are especially useful for quantifying the results.

7.3 Limitations

Although promising visual and objective results were obtained on the test images from datasets (TUDS and UIEBD), there were still some limitations that were observed:

- The performance of the diffusion model is directly correlated with the degree of similarity between the test images and the training dataset. Conditions like extreme turbidity, extremely low light levels, or invisible underwater environments reduce the effectiveness of restoration.
- ACC may require manual parameter tuning in some cases for optimal performance. This limits their adaptability in fully automated systems where no manual adjustments are preferred.
- To make this system more diverse, the training dataset can be further expanded to include region-specific variations, depth measurements, and different water conditions.

- The processing time for large underwater videos is high, when maintaining the original resolution, is important.

7.4 Directions for Future Work

Taking into account the above challenges, there are areas that can be explored further through research that can extend the work that has been done above:

- Development of adaptive diffusion models that can be trained on a a wider range of underwater conditions with regards to turbidity, lighting conditions, scene complexities
- We can combine multi-modal inputs with rich information related to water depth, location, and multi-camera viewpoints.
- Real-time capability can be further extended towards live underwater video support in this framework.
- To facilitate object detection or tracking systems further, we can add a computer vision system with scene understanding that can restore and do semantic analysis of the underwater scenes simultaneously.

Our research proves that the combination of traditional image enhancement methods with the most advanced diffusion models can greatly enhance the quality of an underwater image. Hence, this study is of great value and it provides a comprehensive solution to the complex problem of underwater imaging.

REFERENCES

[1] **Ivanov, A.**, **Kolosova, A.**, **Mandre, R.**, **Udu, H.C. and Tuhtan, J.A.** (2015). Turbid Water and Human Errors: Practical Insights into Real-World ML for Fish Monitoring, *Proceedings of the SUSHP 2015 Conference*, Karlsruhe, Germany.

[2] **Li, C.**, **Guo, C.**, **Ren, W.**, **Cong, R.**, **Hou, J.**, **Kwong, S. and Tao, D.** (2020). An Underwater Image Enhancement Benchmark Dataset and Beyond, *IEEE Transactions on Image Processing*, *29*, 4376–4389.

[3] **Akgül, T.**, **Calık, N. and Töreyin, B.U.** (2020). Deep Learning-Based Fish Detection in Turbid Underwater Images, *Proceedings of the SIU Conference*, Gaziantep, Turkey.

[4] **Larson, G.L.**, **Hoffman, R.L.**, **Hargreaves, B.R.** *et al.* (2007). Predicting Secchi disk depth from average beam attenuation in a deep, ultra-clear lake, *Hydrobiologia*, *574*, 141–148.

[5] **Vasamsetti, S.**, **Mittal, N.**, **Neelapu, B.C. and Sardana, H.K.** (2017). Wavelet based perspective on variational enhancement technique for underwater imagery, *Ocean Engineering*, *141*, 88–100, `hiips://wwwsciencedirectcom/science/article/pii/S0029801817303074`.

[6] **Schettini, R. and Corchs, S.** (2010). Underwater Image Processing: State of the Art of Restoration and Image Enhancement Methods, *EURASIP Journal on Advances in Signal Processing*, *2010*, 746052, `hiips://doi.org/10.1155/2010/746052`.

[7] **Chen, Z.**, **Wang, H.**, **Shen, J.**, **Li, X. and Xu, L.** (2014). Region-specialized underwater image restoration in inhomogeneous optical environments, *Optik*, *125*(9), 2090–2098, `hiips://wwwsciencedirectcom/science/article/pii/S0030402613013843`.

[8] **Berman, D.**, **Levy, D.**, **Avidan, S. and Treibitz, T.** (2021). Underwater Single Image Color Restoration Using Haze-Lines and a New Quantitative Dataset, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(8), 2822–2837.

[9] **Schechner, Y.Y. and Karpel, N.** (2005). Recovery of Underwater Visibility and Structure by Polarization Analysis, *IEEE Journal of Oceanic Engineering*, *30*(3), 570–587.

[10] **He, K.**, **Sun, J. and Tang, X.** (2009). Single Image Haze Removal Using Dark Channel Prior, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1956–1963.

[11] **Sathya, R.**, **Bharathi, M. and Dhivyasri, G.** (2015). Underwater image enhancement by dark channel prior, *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*, Coimbatore, India, pp.1119–1123.

[12] **Sun, Z.** *et al.* (2021). Underwater image processing method based on red channel prior and Retinex algorithm, *Optical Engineering*, *60*(9), 093102, `hiips://doi.org/10.1117/1.OE.60.9.093102`.

[13] **Zhang, W.**, **Wang, Y. and Li, C.** (2022). Underwater Image Enhancement by Attenuated Color Channel Correction and Detail Preserved Contrast Enhancement, *IEEE Journal of Oceanic Engineering*, *47*(3), 718–735.

[14] **Shen, L. and Zhao, Y.** (2020). Underwater image enhancement based on polarization imaging, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *43*, 579–585.

[15] **Garg, D.**, **Garg, N.K. and Kumar, M.** (2018). Underwater image enhancement using blending of CLAHE and percentile methodologies, *Multimedia Tools and Applications*, *77*, 26545–26561, `hiips://doi.org/10.1007/s11042-018-5878-8`.

[16] **Wang, Y.**, **Zhang, J.**, **Cao, Y. and Wang, Z.** (2017). A deep CNN method for underwater image enhancement, *2017 IEEE international conference on image processing (ICIP)*, IEEE, pp.1382–1386.

[17] **Naik, A.**, **Swarnakar, A. and Mittal, K.** (2021). Shallow-UWNet: Compressed Model for Underwater Image Enhancement (Student Abstract), *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp.15853–15854.

[18] **Zhou, W.H.**, **Zhu, D.M.**, **Shi, M.**, **Li, Z.X.**, **Duan, M.**, **Wang, Z.Q.**, **Zhao, G.L. and Zheng, C.D.** (2022). Deep images enhancement for turbid underwater images based on unsupervised learning, *Computers and Electronics in Agriculture*, *202*, 107372.

[19] **Liang, D.**, **Chu, J.**, **Cui, Y.**, **Zhai, Z. and Wang, D.** (2024). NPT-UL: An Underwater Image Enhancement Framework Based on Nonphysical Transformation and Unsupervised Learning, *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 1–19.

[20] **Yang, L.**, **Zhang, Z.**, **Song, Y.**, **Hong, S.**, **Xu, R.**, **Zhao, Y.**, **Zhang, W.**, **Cui, B. and Yang, M.H.** (2023). Diffusion models: A comprehensive survey of methods and applications, *ACM computing surveys*, *56*(4), 1–39.

[21] **Akkaynak, D. and Treibitz, T.** (2018). A revised underwater image formation model, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.6723–6732.

[22] **Krishnan, H.**, **Lakshmi, A.A.**, **Anamika, L.**, **Athira, C.**, **Alaikha, P. and Manikandan, V.** (2020). A novel underwater image enhancement technique using ResNet, *2020 IEEE 4th Conference on Information & Communication Technology (CICT)*, IEEE, pp.1–5.

[23] **Zhang, W.**, **Li, X.**, **Xu, S.**, **Li, X.**, **Yang, Y.**, **Xu, D.**, **Liu, T. and Hu, H.** (2023). Underwater Image Restoration via Adaptive Color Correction and Contrast Enhancement Fusion, *Remote Sensing*, *15*(19), `hiips://www.mdpi.com/2072-4292/15/19/4699`.

[24] **Kim, Y.J. and Patel, J.** (2009). Performance Comparison of the {\rm R}^{\ast}-Tree and the Quadtree for kNN and Distance Join Queries, *IEEE Transactions on Knowledge and Data Engineering*, *22*(7), 1014–1027.

[25] **Iqbal, F. and Töreyin, B.U.** (2025). Underwater Turbid Image Restoration Using Diffusion Models, *2025 33rd Signal Processing and Communications Applications Conference (SIU)*, pp.1–4.

[26] **Khanzadi, P.**, **Majidi, B. and Akhtarkavan, E.** (2017). A novel metric for digital image quality assessment using entropy-based image complexity, *2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, pp.0440–0445.

[27] **Gunawan, R.**, **Tran, Y.**, **Zheng, J.**, **Nguyen, H. and Chai, R.** (2023). Implementing Natural Image Quality Evaluator for Performance Indicator on Noise Artefacts Recovery in CT Scan, *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp.1–4.

CURRICULUM VITAE

Fatima IQBAL

EDUCATION:

- **B.Sc.:** 2023, National University of Sciences and Technology, School of Electrical Engineering and Computer Science, Electrical Engineering
- **M.Sc.:** 2026, Istanbul Technical University, Informatics Institute, Computer Science

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **F. Iqbal** and B. U. Töreyin, "Underwater Turbid Image Restoration Using Diffusion Models," 2025 33rd Signal Processing and Communications Applications Conference (SIU), Sile, Istanbul, Turkiye, 2025, pp. 1-4, doi: 10.1109/SIU66497.2025.11112226.