

π -SUB: A Physics-Informed Synthetic Underwater Benchmark Dataset for Underwater Image Enhancement

Namritha Lasyapriya Maddali¹, Rajini Makam², Suresh Sundaram², and Narasimhan Sundararajan³

Abstract—This paper presents π -SUB, a physics-informed framework for generating synthetic underwater benchmark datasets that bridges the synthetic-to-real gap for Underwater Image Enhancement (UIE). The proposed framework extends the classical underwater image formation model by incorporating depth-dependent downwelling irradiance, biologically resolved absorption, and environmental scattering across all ten Jerlov water types, together with independently controllable residual phenomena. Using this framework, the π -SUB dataset consists of paired synthetic underwater-reference images spanning shallow-to-deep and coastal-to-oceanic environments. Extensive simulation studies have been carried out to evaluate π -SUB along two criteria namely hyper-realism and generalizability. For *hyper-realism*, π -SUB attains a global Fréchet Inception Distance (FID) that is 46% lower than Syrea. For *generalizability*, four state-of-the-art UIE architectures (FUnIE-GAN, Pix2Pix, PUIE-Net, and Phaseformer) are used for comparative evaluation of π -SUB. These models were independently trained on six datasets including one real and five synthetic datasets and tested on six real-world benchmark datasets. Across four UIE architectures and six real benchmark datasets, π -SUB improves UIQM by 4.18% over PHISWID (next best) and 9.46% over Syrea (next best), while reducing NIQE by 48.78% and 23.98%, respectively. These results establish π -SUB as a hyper-realistic and generalizable benchmark for developing the next generation of underwater image enhancement methods. <https://github.com/airli-iisc/pi-SUB>

Index Terms—Physics-Informed, Underwater Image Enhancement, Hyper-realism, Generalization

I. INTRODUCTION

Autonomous underwater vehicles (AUVs) and remotely operated vehicles (ROVs) are widely used for the inspection of subsea communication cables, pipelines, offshore energy infrastructure, and ecological monitoring of coral reefs and marine life [1]. Since sonar provides only limited semantic and textural information, optical imaging remains the primary sensing modality for underwater perception. However, light propagation through water is severely affected by wavelength-dependent absorption and scattering, leading to reduced visibility, color distortion, contrast loss, and veiling light. Underwater image enhancement (UIE) seeks to recover scene radiance by compensating for these degradations and serves as a critical preprocessing step for downstream perception tasks. Supervised

UIE, however, requires paired degraded and clean images of the same scene, where the clean reference represents the scene in the absence of the water medium—a condition that is physically impossible to capture [2]. Consequently, bridging the synthetic-to-real domain gap has become essential for developing generalizable UIE models.

Existing paired datasets therefore substitute approximations. Underwater Image Enhancement Benchmark (UIEB) [3] selects, for each image, the visually preferred output of eleven enhancement algorithms; Enhancing Underwater Visual Perception (EUVP) [2] synthesizes a counterpart for each degraded image with a CycleGAN; and Large Scale Underwater Image (LSUI) [4] extends the same strategy at larger scale. Such approximations are employed because acquiring a distortion-free reference of the same scene without the water medium is physically impossible. The reference images in existing paired datasets exhibit inconsistent colour, contrast, and structural appearance. For example, UIEB and EUVP retain noticeable green colour casts, while LSUI references appear over-bright with a veil-like effect. Consequently, enhancement models tend to learn dataset-specific biases rather than invert the underlying underwater imaging physics. Moreover, the unknown optical conditions of the captured scenes prevent verification against physically valid ground truth, limiting both training and evaluation.

Synthetic data removes this obstacle by constructing the clean reference first and applying a known degradation to it. Two families of methods exist. *Appearance-transfer* methods such as generative translation, physics-guided style transfer, and engine rendering configured from image statistics [5], [6] learn underwater appearance directly from real photographs. They achieve broad visual diversity, but inherit their optical realism from whichever water bodies the reference collection happens to contain and offer no controlled mechanism to extrapolate beyond it. *Physics-based* methods instead derive the degradation from an explicit image formation model, most commonly the Jaffe–McGlamery formulation [7], and are in principle able to span optical conditions that no photograph in a reference set exhibits.

Three gaps recur across them. First, they neglect depth-dependent downwelling irradiance by assuming constant surface illumination. Second, they omit biologically driven optical processes, including chlorophyll absorption, CDOM, and chlorophyll-*a* fluorescence. Third, they do not jointly model the volumetric scattering that arises from suspended particulates and distance-dependent haze. Consequently, no existing

¹Namritha is an Intern at Department of Aerospace Engineering, Indian Institute of Science and Student at Department of Computer Science, PES University, Bangalore, India. namritha.maddali@gmail.com

² Department of Aerospace Engineering, Indian Institute of Science, Bangalore, India. {rajnimakam, vssuresh}@iisc.ac.in

³ Retired Professor, Nanyang Technological University, Singapore. [{ensundara@ntu.edu.sg}](mailto:{ensundara@ntu.edu.sg)

benchmark simultaneously captures the principal physical and biological processes governing underwater image formation, and the synthetic-to-real gap remains still unresolved.

To close these gaps, this paper proposes a **Physics-Informed Synthetic Underwater Benchmark** (π -SUB) framework. Built on an extended Jaffe–McGlamery model, π -SUB combines wavelength-dependent inherent optical properties (IOPs) with separate direct-transmission and backscatter attenuation coefficients, models depth-dependent downwelling irradiance across all ten Jerlov water types, resolves absorption into its biological constituents, and augments the formation model with volumetric haze and suspended particulate scattering. Clean references are drawn from photorealistic Unreal Engine renderings together with curated real underwater images with near-zero medium effects, yielding paired synthetic underwater–reference images.

To assess its effectiveness, we introduce a two-axis evaluation protocol based on (i) hyper-realism, which measures the distributional similarity between synthetic and real underwater imagery, and (ii) generalizability, which quantifies the downstream performance of UIE models trained on the benchmark. Hyper-realism of π -SUB has been validated through analyses against real underwater datasets and compared with synthetic datasets SUID [8], SUIEB [9], PHISWID [10] and Syrea [5] using Fréchet Inception Distance (FID), out-of-distribution rate, and PCA-based manifold analysis. π -SUB achieves a 46% reduction in FID relative to the best existing synthetic benchmark, demonstrating substantially improved distributional alignment with real underwater imagery. To evaluate the generalizability of π -SUB, four representative UIE architectures (FUnIE-GAN [2], Pix2Pix [11], PUIE-Net [12], and Phaseformer [13]) are independently retrained on each competing synthetic dataset and the real paired benchmark UIEB using identical hyperparameters, and subsequently evaluated on six real-world benchmarks. Averaged over all evaluations, π -SUB improves UIQM by 4.2% over PHISWID (next best) and reduces NIQE by 23.9% over Syrea (next best). Furthermore, they preserve 2% more salient keypoints for feature matching, benefiting downstream robotic tasks such as visual localization, mapping, and autonomous underwater navigation.

In summary, the main contributions of this paper are:

- A physics-informed framework, π -SUB, has been developed to generate paired underwater benchmark datasets from either a set of clean real images or Unreal Engine rendered images by modeling the following, viz the water type, the camera depth, visibility, chlorophyll concentration, and finally optical attenuation.
- A modified Jaffe–McGlamery image formation model is developed that incorporates the depth-dependent downwelling irradiance, Jerlov IOPs, separate direct-transmission and backscatter attenuation coefficients, and finally biologically driven optical effects including chlorophyll absorption, CDOM, and fluorescence.
- Extensive performance evaluation of π -SUB has shown that it produces a hyper-realistic and highly generalizable training data, achieving a 46% lower FID than the next best synthetic benchmark dataset Syrea and consistently improving UIE performance across all other state-of-the-art architectures and real-world benchmark datasets.

and evaluated on real-world data and a downstream feature-matching task.

The paper is organized as follows. Section II reviews existing (both) real and synthetic underwater datasets and highlights the uniqueness of π -SUB among them. Section III develops the basic image formation model and the π -SUB image generation framework. Section IV presents the detailed performance evaluation of π -SUB dataset both distributionally and through the generalization of models trained using π -SUB. Further, detailed ablation studies, extended hyper-realism analyses, extended metric analysis and statistical significance tests are provided in the supplementary material. These additional results further confirm the robustness and superior generalizability of the proposed π -SUB framework. For ease of understanding a table with the full expansion of all acronyms is provided. Section V summarizes the conclusions for this study.

II. REVIEW OF EARLIER WORK: EXISTING DATASETS FOR UNDERWATER IMAGE ENHANCEMENT

This section presents a review of both current real-world and synthetic underwater datasets against three properties relevant to supervised training, viz.,: depth-dependent irradiance modeling, Jerlov-grounded inherent optical property (IOP) parameterization, and finally biological optical effects.

A. Current Real-World Underwater Datasets

UIEB [3] provides 890 real underwater images whose references are the visually preferred output of eleven enhancement algorithms, chosen by volunteer pairwise comparison. EUVP [2] trains a CycleGAN on unpaired data to synthesize a counterpart for each degraded image, and LSUI [4] extends this strategy to 4,279 pairs with semantic and transmission maps. In all these three datasets, the reference is an enhanced version of the underwater image rather than a clean capture, so it encodes the biases of the algorithms or raters used to construct it; Hou *et al.* [8] show that model rankings are inconsistent across these benchmarks as a direct consequence. Unpaired collections such as SQUID [14], OceanDark [15], and Real-world Underwater Enhancement (RUIE) [16] provide no supervisory target and are usable only for no-reference evaluation. No real-world dataset therefore supplies a physically valid clean reference for the captured scene.

B. Current Synthetic Underwater Datasets

Current synthetic datasets fall into two families. The *physics-based* family derives the degradation from an explicit image formation model. The UWCNN synthetic dataset (named as SUIEB) [9] applies a global per-type coefficient to per-pixel depth from indoor NYU scenes, extending the formulation to all ten Jerlov water types but using a single attenuation coefficient uniformly across each image and omitting any vertical-depth irradiance term. SUID [8] applies thirty heuristically tuned degradation effects to thirty ground-truth images, individually and in combination, yielding 900 images that are neither IOP-derived nor large enough for training. Physics-Inspired Synthesized Underwater Image Dataset (PHISWID) [10] introduces

TABLE I: Physical and structural properties of synthetic underwater image datasets. VD: Vertical depth; IOP: Jerlov IOPs; Bio: Biological effects; Inv.: Invertible; Meta: Metadata.

Dataset	VD	IOP	Bio	Inv.	Range	Meta
SUIEB [9]*	×	~	×	~	~	×
SUID [8]*	×	×	×	~	~	×
Syrea [5]*	×	×	×	×	×	×
PHISWID [10]*	×	×	×	~	✓	~
RSUIGM [17]†	✓	✓	×	✓	✓	×
MUSE [6]†	×	×	×	×	✓	×

* Publicly available synthetic datasets used in the experimental comparison;

† not included in the experimental comparison.

per-pixel scene range from RGB-D imagery together with a physics-based marine snow model, but samples vertical depth, background light B_c , and water type from uniform distributions rather than deriving attenuation coefficients analytically from measured IOPs, and fixes downwelling irradiance at the surface. Realistic Synthetic Underwater Image Generation (RSUIGM) [17] implements the dual-path model with separate direct and backscatter coefficients and incorporates vertical irradiance, but does not model the biological optical effects characteristic of coastal waters.

Common to the family, clean images are typically sourced from terrestrial RGB-D datasets, which limits semantic diversity to land-based scenes rather than marine environments such as coral reefs, shipwrecks, and aquatic ecosystems; attenuation parameters are defined globally, heuristically, or per image rather than from measured Jerlov IOPs; depth-dependent irradiance is absent or only partially modeled; scene range is often conflated with vertical depth; and biological optical effects are ignored. The second family learns underwater appearance directly from data, either generatively (WaterGAN [18], and diffusion-based translation [19]) or through physics-guided style transfer and engine rendering, as in Syrea [5] and MUSE [6]. This family inherits its optical realism from the reference distribution rather than from an explicit image formation model, and therefore provides no controlled mechanism to extrapolate to unseen optical conditions; none of its members employs analytic IOP-grounded parameterization.

A summary comparison of these datasets, evaluated across three primary physical properties—depth-dependent irradiance, Jerlov-grounded IOP parameterization, and biological optical effects—and three auxiliary structural properties—invertibility, range-awareness, and metadata—is presented in Table I. The symbols ✓, ~, and × denote full, partial, and absent support, respectively. PHISWID and RSUIGM model range-dependent degradation but omit biological optical effects; SUIEB and SUID rely on simplified attenuation models and exclude depth-dependent irradiance; Syrea estimates optical parameters from real images rather than measured IOPs; and MUSE employs a graphics engine yet derives scene parameters from image statistics. As the comparison clearly shows, no existing synthetic dataset jointly incorporates depth-dependent illumination, Jerlov-grounded optical properties, biological spectral effects, and controlled physical parameterization. This gap motivated the development of the π -SUB framework described next.

III. DEVELOPMENT OF A NOVEL π -SUB DATASET GENERATION FRAMEWORK

In this section, we describe the development of π -SUB, a physics-informed synthetic underwater benchmark data-generation framework that transforms a clean reference image into an underwater image with fully recorded per-sample ground truth. This provides a new benchmark dataset referred to as π -SUB dataset for realistic assessment in underwater image enhancement studies.

A. The π -SUB Framework

The overall π -SUB generation framework is shown in Figure 1. The framework accepts three user-defined inputs: (i) a reference image source, selected from either a simulated or real image pool; (ii) the physical water configuration, specified by the Jerlov water type, camera depth, and the corresponding inherent and apparent optical properties (IOPs/AOPs); and (iii) an Augmented Realism configuration, which determines whether suspended particulate matter, volumetric haze, biological effects, or their combination are applied to further degrade the images. The framework then proceeds through three sequential stages. *Stage I* estimates a dense scene-range map and converts it to a metric underwater range using the water-type-dependent maximum visibility distance $z_{\max}$. *Stage II* evaluates the proposed modified underwater image formation model to synthesize the deterministic underwater observation I_b . Finally, *Stage III* uses the selected residual optical phenomena to generate the phenomenon-specific and fully augmented subsets that together constitute the π -SUB dataset. The remainder of this section describes each stage in detail.

1) *Reference Image Set*: The leftmost block of Figure 1 represents a clean reference image pool J , which is assembled from complementary simulated and real-image sources. The simulated subset is rendered in Unreal Engine without a water medium, such that the recorded radiance corresponds directly to scene reflectance, free from attenuation, scattering, and backscatter, with the sun remaining the sole source of illumination under a downwelling irradiance normalized to $E_0 = 1$, representing clear, sunny conditions [20]. Every rendered image is accompanied by exact metric range, camera depth, and instance-level semantic annotations. The scenes span representative underwater environments, including coral reefs, ship and aircraft wrecks, rocky seabeds, seagrass, macroalgae, divers, marine animals, and fish schools.

Although Unreal Engine provides physically accurate geometry and controllable ground truth, commercially available asset libraries cannot fully represent the diversity of coral species, marine organisms, and benthic habitats observed in natural underwater ecosystems. To reduce this gap, a real subset is incorporated by choosing enhanced ground-truth references from UIEB and LSUI. This subset accounts for approximately 10% of the reference pool and is curated rather than sampled: candidates are restricted to shallow-water scenes with minimal residual medium effects. They are retained only where the per-channel RGB response is balanced, so that no residual colour cast is carried into the synthesis. Applying the same

underwater image formation model to both subsets ensures that the principal difference between them is scene appearance rather than degradation physics, thereby improving the generalization of downstream underwater image enhancement models.

2) *Stage I — Range Map Estimation*: The underwater image formation model requires a per-pixel horizontal range map z_s . Exact metric range is available only for the simulated subset through Unreal Engine depth buffers. However, to maintain statistical consistency between simulated and real images, a single monocular depth estimator, Depth Anything V2 [21], is applied uniformly to both subsets. The network output is proximity-like, with pixels closer to the camera driven toward 1 and distant pixels toward 0; since the image formation model requires horizontal *range*, this must be inverted before metric scaling. The network predicts a normalized relative depth map $z_n \in [0, 1]$, which is converted into metric underwater range through the water-type-dependent calibration $z_s = z_{\max}(1 - z_n)$, yielding $z_s \in [0, z_{\max}]$. Using the same estimator for both domains prevents a restoration model from exploiting systematic differences in depth statistics. The original Unreal Engine depth buffers are nevertheless retained as reference metadata, enabling quantitative validation of the estimated ranges and providing exact geometry whenever required. Together with the estimated range map, the framework retrieves the assigned physical configuration, namely the camera depth d , Jerlov water type, and the derived optical parameters K_d , β_d , β_b , B_∞ , and $z_{\max}$ from the water-type database and passes on to subsequent block.

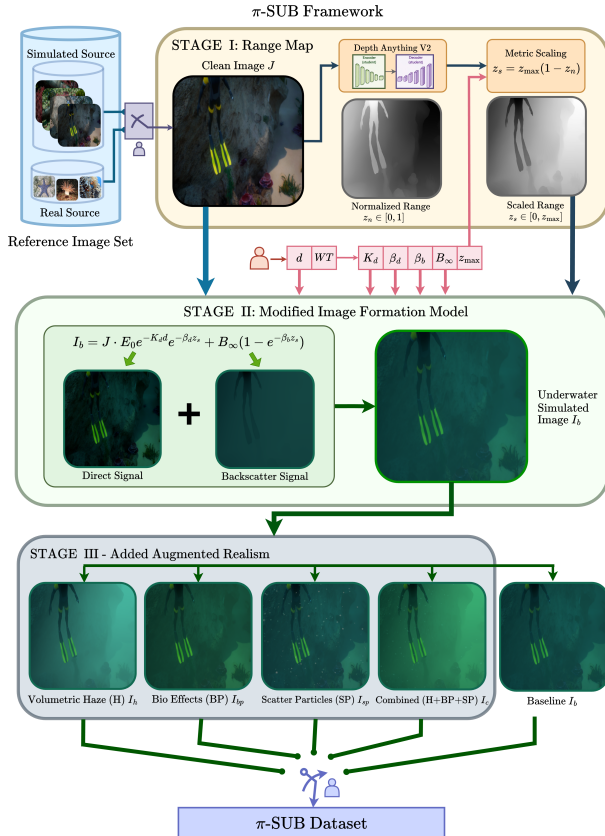


Fig. 1: Overall schematic of the Physics-Informed Synthetic Underwater Benchmark (π -SUB) dataset generation framework.

3) *Stage II — Modified Underwater Image Formation*: Given the clean reference image J , the estimated range map z_s , and the assigned physical configuration, the modified Jaffe–McGlamery image formation model presented in Section III-C synthesizes the underwater image I_b . The model combines wavelength-dependent direct transmission and backscatter using physically derived optical parameters, producing an analytically invertible degradation whose complete parameter set is recorded for every generated sample. This image I_b is referred to as baseline image and fed as an input to the augmented realism stage.

4) *Stage III — Augmented Realism*: The analytical image formation model assumes a homogeneous water column and therefore cannot represent several optical effects commonly observed in natural underwater imagery. Stage III addresses this limitation by augmenting the deterministic underwater image I_b with three independently controlled phenomena: suspended particulate matter, volumetric multiple scattering (Haze), and biologically induced fluorescence. The mathematical description of these augmented realism is given in Section III-D. As illustrated in Figure 1, each phenomenon is applied individually to generate phenomenon-specific subsets (I_h , I_{sp} , I_{bp}), while an additional subset combines all three effects into I_c .

The subsequent sections present the mathematical formulation of Stage I, the physics-based underwater image formation model of Stage II, and the augmented realism model of Stage III.

B. Classical Jaffe–McGlamery Underwater Image Formation Model (IMF)

The physics of underwater light propagation is governed by the Radiative Transfer Equation (RTE), which models light absorption and scattering within a participating medium [22]. Under the single-scattering assumption, the RTE reduces to the Jaffe–McGlamery image formation model [23], where the recorded radiance is decomposed into two physically meaningful components: direct transmission from the scene and backscatter generated by the water column.

$$I_b = \underbrace{J e^{-\beta_d z}}_D + \underbrace{B_\infty(1 - e^{-\beta_b z})}_B, \quad (1)$$

where J is the clean scene radiance, z the horizontal camera-to-scene range, and β_d , β_b the attenuation coefficients of the direct D and backscatter B paths [20]. The two components behave differently with range: direct transmission decays exponentially, whereas backscatter accumulates along the imaging path and asymptotically approaches the veiling light B_∞ , so distant regions converge toward a uniform radiance field rather than fading to darkness [24]. They are also governed by different attenuation mechanisms. The direct signal is dominated by absorption, since underwater scattering is strongly forward directed and many scattered photons remain within the camera’s acceptance angle, while backscatter is produced by photons redirected toward the sensor and therefore experiences the full beam attenuation rate. Using a single coefficient $\beta = a + b$ for both paths overestimates extinction of the

direct signal, so following Akkaynak and Treibitz [20] separate coefficients are adopted,

$$\beta_d(\lambda) = a(\lambda) + \delta b(\lambda), \quad \beta_b(\lambda) = a(\lambda) + b(\lambda) = c(\lambda), \quad (2)$$

where $\delta = 0.05$ accounts for the limited contribution of forward scattering to direct-path extinction, while the backscatter component follows the full beam attenuation coefficient $c(\lambda)$.

The equation (1) in this section models radiative transport only along the horizontal imaging path. However, the available illumination reaching the scene depends on the downwelling irradiance, which is attenuated over the vertical water column before interacting with scene surfaces. Both the direct transmission and background veiling light vary with camera depth. Section III-C extends the classical formulation by incorporating depth-dependent apparent optical properties and biological effects, leading to the modified underwater image formation model used throughout π -SUB.

C. Modified Underwater Image Formation (MUIF)

In this section, physics and mathematical formulation underlying Stage II of the π -SUB framework is presented. Building on the image formation model in Section III-B, the direct attenuation (β_d), backscatter (β_b and B_∞), and ambient irradiance are parameterized from the inherent optical properties (IOPs), (a, b), of the water column. The IOPs are intrinsic properties of the medium, whereas the apparent optical properties (AOPs), namely K_d , β_d , β_b , and B_∞ , depend on both the medium and the illumination and viewing geometry. This section derives the AOPs from measured Jerlov IOPs and determines the water-type-specific visibility limits required by the modified underwater image formation (MUIF) model.

The MUIF is parameterized using the ten canonical Jerlov water types, comprising five oceanic (I, IA, IB, II, III) and five coastal (1C, 3C, 5C, 7C, 9C) classes [25]. These span the full range of natural optical conditions, from clear oceanic waters to highly turbid coastal environments. The wavelength-dependent absorption $a(\lambda)$ and scattering $b(\lambda)$ are sampled at the representative RGB wavelengths $\lambda = \{650, 550, 450\}$ nm following [20], [23]. The IOPs are obtained from the experimentally derived measurements of Solonenko and Mobley [26]. Since measurements are available for six Jerlov types (IB–5C), the remaining types (I, IA, 7C, and 9C) are obtained by monotonic interpolation along the Jerlov clarity scale. Oceanic water types primarily attenuate the red channel, whereas coastal types attenuate all three channels more strongly. Consequently, increasing depth or range progressively reduces scene brightness and contrast.

Although the classical underwater image formation model is parameterized by the inherent optical properties (IOPs) of the water column, it assumes a vertically homogeneous medium and therefore does not distinguish the camera's vertical depth from the horizontal scene range. Consequently, the attenuation of downwelling irradiance E_d with camera depth is neglected, causing the global scene illumination to remain unchanged irrespective of the imaging depth. The proposed MUIF addresses this limitation by explicitly modeling

depth-dependent downwelling irradiance as a separate physical process.

1) *Vertical Irradiance E_d* : The direct transmission component depends not only on the horizontal propagation distance but also on the downwelling irradiance available at the scene. The vertical camera depth must be explicitly modeled, since the available illumination decreases with depth even when the scene range remains unchanged.

The downwelling irradiance E_d sets the illumination level available at the scene and decays exponentially with depth d at a rate set by the diffuse attenuation coefficient K_d . Following Morel and Loisel [27] and Williamson [28], K_d is computed from the IOPs at an optical depth $\zeta = 1$:

$$K_d(\lambda) = \frac{a(\lambda)}{\mu_d} \sqrt{1 + \frac{b(\lambda)}{a(\lambda)} \cdot G(\eta, \mu_d)} \quad (3)$$

where μ_d is the mean cosine of the downwelling light just below the water surface ($\mu_d = 0.89$ for oceanic and 0.85 for coastal waters), and G is the geometry factor:

$$G(\eta, \mu_d) = \mu_d(1.3433\eta - 0.01414) + (0.1304 + 0.272\eta) \quad (4)$$

Here $\eta = b_w(\lambda)/b(\lambda)$ is the fraction of total scattering attributable to pure-water (molecular) scattering [28], with $b_w(\lambda) = 5.83 \times 10^{-3} (400/\lambda)^{4.322}$. The parameter η captures the transition from particle-dominated scattering ($\eta \rightarrow 0$, turbid coastal water) to molecular scattering ($\eta \rightarrow 1$, clear ocean), and G varies accordingly. The downwelling irradiance at depth d then follows as

$$E_d(\lambda) = E_0 e^{-K_d(\lambda)d} \quad (5)$$

with E_0 the surface irradiance. Because $K_d < a + b$, light decays more slowly along the vertical than along the horizontal path; depth and range are therefore *not* interchangeable, and each requires its own coefficient.

2) *Veiling Light B_∞* : The Jerlov IOP tables report the *total* scattering $b(\lambda)$, whereas B_∞ is physically driven by the *backscattering* coefficient $b_b(\lambda) \leq b(\lambda)$ [29], since only photons scattered into the rear hemisphere reach the sensor. It is reconstructed from its molecular and particulate contributions:

$$b_b(\lambda) = \frac{1}{2} b_w(\lambda) + \tilde{b}_{bp} (b(\lambda) - b_w(\lambda)), \quad \tilde{b}_{bp} \approx 0.018 \quad (6)$$

where one half of the pure-water scattering b_w is back-directed owing to the near-symmetric molecular phase function, and $\tilde{b}_{bp} \approx 0.018$ is the backscatter fraction of the Petzold average particle phase function, obtained by integrating the measured volume scattering function over the backward hemisphere ($\theta \in [90^\circ, 180^\circ]$) as tabulated by Mobley [29]. In the single-scattering regime the veiling light is the backscattered fraction of the depth-attenuated ambient irradiance, saturating at the beam-attenuation rate $c = a + b$:

$$B_\infty(\lambda) = \frac{b_b(\lambda) \cdot E_d(\lambda)}{a(\lambda) + b(\lambda)} \quad (7)$$

Note that B_∞ inherits the depth dependence of E_d : the veil dims as the camera descends, exactly as the scene does.

3) *Biological Constituents of the Medium*: The preceding formulation models absorption and scattering through bulk optical coefficients. Although sufficient for image formation,

these coefficients do not explicitly identify the biological constituents responsible for wavelength-dependent attenuation. In coastal and shallow water the regime represented in the EUVP [2], LSUI [4] and RUIE [16] datasets where the absorption is dominated by an active phytoplankton population that is largely absent from clear oceanic water. Resolving these constituents enables biologically meaningful parameterization while preserving consistency with measured optical properties. First, the chlorophyll concentration is the physical variable that separates the oceanic types from the coastal ones; exposing it converts water type from a categorical label into a continuous, controllable degree of freedom. Second, the same chlorophyll that absorbs light also *re-emits* it, and that emission cannot be represented by any attenuation coefficient.

Three processes contribute. First, selective absorption by chlorophyll-*a* and its accessory pigments. Second, co-absorption by Chromophoric Dissolved Organic Matter (CDOM), a photo-degradation byproduct of phytoplankton metabolism whose concentration is correlated with that of chlorophyll. Third, inelastic chlorophyll fluorescence, in which absorbed blue and red photons are re-emitted near 685 nm. The first two are *elastic* losses and belong in the medium; the third is an *additive emission* called Fluorescence.

Following the experimentally-derived constituent decomposition of Solonenko and Mobley [26], the measured Jerlov absorption is reconstructed as the sum of pure water, chlorophyll and CDOM:

$$a(\lambda) = a_w(\lambda) + a_{chl}(\lambda) + a_{cdom}(\lambda) \quad (8)$$

where a_w is the pure-water absorption [30]. The chlorophyll term follows the Bricaud *et al.* [22] power law:

$$a_{chl}(\lambda) = A_\lambda C_{chl}^{E_\lambda} \quad (9)$$

where A_λ and E_λ are wavelength-dependent pigment coefficients and C_{chl} is the chlorophyll concentration (mg/m³). CDOM is coupled to chlorophyll rather than treated as an independent constituent, since it is predominantly a photo-oxidation byproduct of phytoplankton material and co-varies with C_{chl} [22]:

$$a_{cdom}(\lambda) = a_{chl}(440) M e^{-\alpha(\lambda-440)} \quad (10)$$

where M (unitless) and α (nm⁻¹) are per-type CDOM fitting parameters. The per-type values of C_{chl} and (M, α) are taken from the fits of [26] and details are given in the supplementary material.

The scattering coefficient admits an analogous decomposition into a molecular term and two particulate modes:

$$b(\lambda) = b_w(\lambda) + \underbrace{B_s b_s(\lambda) + B_l b_l(\lambda)}_{b_p(\lambda)} \quad (11)$$

where b_w is the pure-water (molecular) scattering term introduced in Section III-C1, and b_p splits into small- and large-particle modes $b_s(\lambda) = 1.1513(400/\lambda)^{1.7}$ and $b_l(\lambda) = 0.3411(400/\lambda)^{0.3}$, with per-type concentrations (B_s, B_l) from the same Solonenko–Mobley fit [26]. Absorption is thus biologically resolved while scattering enters through bulk particulate modes, but both are drawn from a single measured constituent basis, so a and b remain mutually consistent. Because C_{chl}

now appears explicitly, it can be varied independently of the nominal water type, allowing a fixed scene geometry to be rendered across a productivity gradient.

Having resolved the absorption and scattering coefficients into their physical constituents, the complete set of inherent optical properties is now available. These properties not only parameterize the image formation model but also determine the maximum distance over which scene contrast can be preserved. The corresponding visibility limits are therefore derived next and used to bound the scene range during dataset synthesis.

4) *Visibility Limits*: The backscatter term saturates at a range beyond which no scene contrast survives, and this defines the physical extent of the scene rather than an arbitrary rendering choice. Koschmieder’s Law [31] gives the extinction of visible contrast along the line of sight as $C_z = C_0 e^{-\beta z}$. The beam contrast analysis of Lee *et al.* [32] reports a detection threshold of $\varepsilon \approx 0.05$ for the human eye and $\varepsilon \approx 0.01$ for camera systems [22]. Since π -SUB models images captured by a camera, the camera threshold $C_z/C_0 = \varepsilon \approx 0.01$ is used throughout. The maximum scene range is therefore

$$z_{\max} = \frac{-\ln \varepsilon}{\beta_{\min}}, \quad \beta_{\min} = \min_{\lambda} (a(\lambda) + b(\lambda)) \quad (12)$$

where $\beta_{\min}$ is the minimum beam attenuation across the three channels, since visibility is governed by the channel that penetrates furthest. The corresponding vertical limit $d_{\max}$ follows from K_d . Horizontal visibility $z_{\max}$ exceeds 200 m in clear oceanic water (Type I) and collapses below two metres in turbid Type 9C coastal water, while $d_{\max}$ remains consistently larger because $K_d < a + b$. Both bounds are strongly type-dependent, and π -SUB treats them as such rather than fixing a single scene extent across all water types; the complete per-type $(z_{\max}, d_{\max})$ values are tabulated in the supplementary material.

Combining the separated attenuation coefficients, depth-dependent irradiance, biologically resolved absorption, and veiling light with the classical Jaffe–McGlamery formulation yields the image formation model adopted by π -SUB:

$$I_b(\lambda) = \underbrace{J E_d(\lambda) e^{-\beta_d(\lambda) z_s}}_{\text{direct}} + \underbrace{B_\infty(\lambda) (1 - e^{-\beta_b(\lambda) z_s})}_{\text{backscatter}}, \quad (13)$$

where I_b is referred as Baseline image, $E_d = E_0 e^{-K_d d}$, $\beta_d = a + \delta b$, $\beta_b = a + b$, $B_\infty = b_b E_d / (a + b)$, and $a = a_w + a_{chl} + a_{cdom}$. Unlike the conventional Jaffe–McGlamery model, the proposed formulation explicitly separates direct and backscatter attenuation, incorporates depth-dependent illumination through E_d , and resolves absorption into measured biological constituents while remaining deterministic and analytically invertible.

For each clean reference image, π -SUB evaluates (13) in linear radiometric space using the assigned Jerlov water type, vertical camera depth d , surface irradiance E_0 , and per-pixel scene range z_s . Every image is synthesized under all ten Jerlov water types and multiple camera depths, with z_s obtained by calibrating the normalized Depth Anything V2 prediction z_n to the water-type-specific visibility limit,

$$z_s = z_{\max} (1 - z_n), \quad (14)$$

where the same calibration is applied to both simulated and real image pools. Unlike previous formulations that conflate camera depth and scene range, π -SUB models vertical depth through the global illumination term E_d and horizontal range through the spatially varying attenuation and backscatter terms.

Fig. 2 visualizes the resulting joint degradation surface produced by evaluating Eq. (13) across all ten Jerlov types, horizontal range, and vertical depth. Oceanic types (I–III) retain a blue-green cast since only the red channel is stripped, while coastal types (1C–9C) darken across all channels. Along either axis, intensity falls monotonically as the corresponding attenuation term decays, confirming that range and depth are governed by distinct, non-interchangeable coefficients.

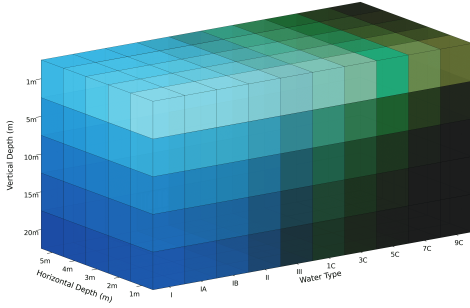


Fig. 2: Joint degradation over Jerlov type, range, and depth: oceanic types (I–III) retain a blue–green cast, coastal types (1C–9C) darken uniformly.

The modified image formation model accurately reproduces the dominant deterministic effects of underwater light propagation. However, several residual phenomena observed in real underwater imagery arise from complex volumetric interactions and stochastic environmental variability that are not captured by analytical radiative transport models. The following section addresses these effects through the proposed Augmented Realism stage.

D. Augmented Realism

Having established the deterministic image formation model in Stage II, this section describes Stage III of the π -SUB framework, which augments I_b with residual optical phenomena that are not captured by the analytical model. The proposed Augmented Realism overlays three independently controlled residual phenomena on the deterministic underwater image I_b : suspended particulate matter, volumetric haze, and biological effects. Each phenomenon is generated as a separate subset, enabling controlled evaluation of individual optical degradations as well as their combined effects.

1) *Suspended Particulate (SP) Matter*: Suspended particles are composited as a procedural layer on top of I_b following recent underwater rendering practice [33], [34]. Each particle is rendered as an oriented anisotropic Gaussian whose opacity α_i is modulated by an occlusion term derived from the normalized range map, so that particles lying behind scene geometry are suppressed. The particulate layer is composited as

$$I_{sp} = I_b + \lambda \sum_{i=1}^N \alpha_i \mathbf{c}_i, \quad (15)$$

where N is the particle count, $\mathbf{c}_i$ the particle colour, and λ the overall intensity. The full parameterization of α_i , including

the size, elongation, and occlusion terms, is given in the supplementary material.

2) *Volumetric Haze*: The single-scattering term in (13) underestimates the diffuse veil produced by multiple scattering. π -SUB approximates this effect with a continuous range-dependent haze [23], [29]:

$$I_h = I_b(1 - \rho) + \mathbf{c}_{haze}\rho, \quad \rho = (1 - z_n)^{1/s_h}, \quad (16)$$

where z_n is the normalized scene range, $\mathbf{c}_{haze}$ is the spectrally tinted haze colour, and s_h controls the severity.

3) *Biological Effects (BE)*: Chlorophyll and CDOM absorption are already encoded in the base model through a in Section III-C3. The remaining biological phenomenon is inelastic fluorescence, which introduces additive radiance and therefore cannot be represented by attenuation alone [35], [36]. The fluorescence yield is not constant with depth. Near the surface, high photosynthetically active radiation (PAR) drives photo-inhibition, which dissipates absorbed energy as heat and suppresses the quantum yield. As the scalar irradiance attenuates with depth, photo-inhibition relaxes, the yield rises to a characteristic subsurface maximum, and the fluorescence signal peaks before photon starvation at depth extinguishes it [37]. This non-monotonic depth profile distinguishes fluorescence from every other modeled phenomenon, all of which decrease monotonically with depth. The resulting RGB fluorescence source term is

$$\mathbf{F}_{em} = \phi_C \cdot a_\phi \cdot \mathbf{f}_{em}, \quad \mathbf{f}_{em} = [0.97, 0.03, 0.00]^\top \quad (17)$$

where ϕ_C is the depth-resolved quantum yield obtained from the photo-inhibition model of Zhai *et al.* [36], a_ϕ is the phytoplankton absorption evaluated at the two chlorophyll excitation peaks following the Bricaud *et al.* parameterization [22], and $\mathbf{f}_{em}$ maps the 685nm emission band onto the RGB channels. The complete four-stage derivation, all photo-physical constants, and the green-normalized chromatic refinement applied to productive coastal waters [29], [38] are given in the supplementary material. Fluorescence and chromatic refinement become increasingly pronounced with chlorophyll concentration, producing negligible changes for clear oceanic waters (Types I–IB) and stronger coastal coloration for productive Jerlov types (1C–9C). The proposed fluorescence model is physically motivated and consistent with established radiative transfer principles; it has not yet been independently validated against *in-situ* measurements of chlorophyll fluorescence and absorption, and is identified as future work.

Together with the suspended particulate matter and volumetric haze models presented previously, the biological effects complete the proposed Augmented Realism stage. The independently controlled combinations of these residual phenomena define the final organization of the π -SUB benchmark.

The π -SUB comprises a deterministic MUIF subset (I_b) together with phenomenon-specific variants: MUIF+Haze (I_h), MUIF+BE (I_{bp}), and MUIF+SP (I_{sp}), as well as a fully augmented subset, MUIF+Haze+BE+SP (I_c), which combines all three residual phenomena. This organization enables controlled evaluation of individual optical phenomena across Jerlov water types while also providing a hyper-realistic composite

benchmark that more closely reflects natural underwater imaging conditions. Each sample is accompanied by complete metadata, including the Jerlov water type, camera depth d , chlorophyll concentration C_{chl} , visibility limits ($z_{\max}$, $d_{\max}$), surface irradiance E_0 , and channel-wise optical parameters, ensuring full reproducibility of the synthesis process. By combining physically grounded modified image formation with independently controllable residual effects, π -SUB spans a broad distribution of underwater optical conditions, reducing the simulation-to-reality gap and improving the generalizability of underwater image enhancement models trained on the benchmark.

The released benchmark is generated from $N = 117$ clean reference images drawn from the simulated and real pools of Section III-A. Each reference is synthesized under all ten Jerlov water types for the I_b , I_{sp} , and I_{bp} subsets, under ten water types at three severity settings for I_h and I_c , and under thirty randomized effect–depth–water-type combinations for the held-out test split, giving $117 \times 120 = 14,040$ paired synthetic underwater–reference images in total. The per-subset breakdown and the metadata schema are tabulated in the supplementary material.

IV. PERFORMANCE EVALUATION OF π -SUB

In the preceding section, the mathematical foundations of π -SUB, introducing a physics-informed synthesis framework that combines measured Jerlov optical properties, depth-dependent irradiance, and biologically driven spectral modulation to generate paired underwater images was presented. In this section, validation of the π -SUB against its two major design objectives: (i) reproducing the distribution of real underwater imagery (*hyper-realism*), and (ii) to show how good the generalization of underwater image enhancement UIE models trained using the proposed dataset is achieved (*generalizability*).

Hyper-realism is assessed by comparing the feature-space distribution of π -SUB against SUID [8], SUIEB [9], Syrea [5], and PHISWID [10], using the combined distribution of five real benchmarks (UIEB [3], RUIE [16], SQUID [14], FishTrac [39], OceanDark [15]) as reference, quantified through Fréchet Inception Distance (FID) [40], out-of-distribution (OOD) rate, and principal component analysis. Generalizability is assessed by independently training four UIE architectures — Pix2Pix [11], FUnIE-GAN [2], Phaseformer [13], and PUIE-Net [12] — on each candidate training dataset and evaluating them on six real-world benchmarks, which isolates the influence of the training data from that of the architecture.

A. Hyper-realism Validation

Each image is first mapped to a physically interpretable perceptual feature space, on which four complementary analyses are performed: perceptual clustering, global and cluster-wise FID, the out-of-distribution rate, and a PCA projection of the combined distributions.

1) **Perceptual Feature Clustering:** Each image is first represented by a physically interpretable perceptual feature vector comprising the mean and standard deviation of brightness (L^*) and chroma, together with the mean dark-channel value,

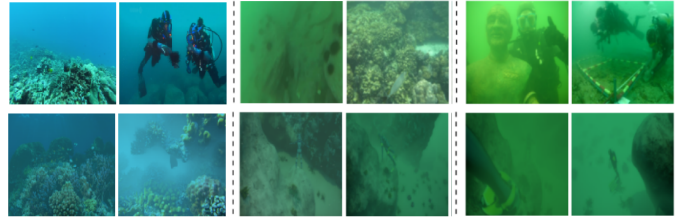


Fig. 3: Representative images from three perceptual clusters obtained using $K=7$ K-means on the combined real feature space (UIEB, RUIE, SQUID, FishTrac, and OceanDark). (a) Real underwater images. (b) Corresponding π -SUB images.

the mean HSV saturation, and the chromatic coordinates (a^* , b^*). These features are directly related to the optical effects of absorption, scattering, and illumination, unlike generic deep features that lack physical interpretability. A K -means model ($K = 7$) is fitted to the standardized feature vectors extracted from the combined real datasets (UIEB, RUIE, SQUID, FishTrac, and OceanDark), partitioning the real data into seven perceptually distinct clusters. π -SUB is the only synthetic dataset represented in all seven clusters, indicating broad coverage of the perceptual characteristics observed in real underwater imagery. Figure 3 shows a sample cluster containing a real underwater image and the corresponding π -SUB-generated images. The generated images are visually close to the real images. In contrast, the other synthetic datasets exhibit incomplete coverage: SUIEB is not represented in cluster C4, PHISWID is not represented in cluster C6, and the remaining synthetic datasets show similar gaps. Representative images from each cluster are provided in the supplementary material, confirming that π -SUB consistently matches the color, contrast, and haze characteristics of the corresponding real images, whereas competing datasets either have sparse representation or exhibit noticeable visual deviations in several clusters.

2) **Global and cluster-wise FID:** FID measures distributional similarity by passing each set through an Inception-v3 network and computing the Fréchet distance between the resulting feature distributions; lower is better. It is computed both globally and per perceptual cluster, so that the comparison reflects performance within each physically grounded optical regime rather than only in aggregate.

The cluster-wise and global FID scores of each synthetic dataset against the combined real benchmark pool are reported in Table II. π -SUB achieves the lowest global FID at 95. This compares to 176 for Syrea, 192 for SUIEB, 227 for PHISWID, and 230 for SUID. This result holds across every cluster individually. SUIEB does not populate cluster C4, PHISWID does not populate cluster C6, and the remaining datasets, including PHISWID, score roughly 50 to 200 FID points above π -SUB in every cluster where they are populated. The cluster-wise breakdown shows that π -SUB's lower distance to the real distribution is consistent across the full range of perceptual conditions present in real underwater imagery, and not just on average.

3) **Out-of-distribution (OOD) rate:** FID does not indicate what fraction of synthetic images lie entirely outside the real manifold. An image is flagged as out-of-distribution (OOD) if its distance to the nearest real-data cluster centroid, in the

TABLE II: Cluster-wise and global FID against real distributions (UIEB + RUIE + SQUID + OceanDark + FishTrac). Lower is better. Best per column in cyan.

Dataset	C1	C2	C3	C4	C5	C6	C7	Global
SUID	278	348	258	207	274	355	258	230
SUIEB	254	248	316	—	290	394	215	192
PHISWID	291	271	278	256	392	—	272	227
Syrea	249	235	232	203	234	262	225	176
π -SUB	192	196	126	105	183	232	142	95

same perceptual feature space, exceeds the 95th percentile of distances observed among real images.

Only 2.00% of the 14,040 π -SUB images are flagged as OOD, against 4.02% for PHISWID (9,426 images), 12.23% for Syrea (21,826), 19.26% for SUIEB (11,592), and 27.33% for SUID (900). This reduction reflects the benefit of grounding optical parameters in measured Jerlov IOPs and modelling depth-dependent irradiance and biological spectral effects rather than approximating them heuristically. The π -SUB images that are flagged arise predominantly from deep coastal water with strong biological absorption and no artificial illumination: physically valid conditions that are structurally absent from real benchmark collections, since field deployments rarely operate under them. The OOD fraction therefore reflects coverage beyond the reach of current real benchmarks rather than a mismatch with underwater optics.

4) *Principal Component Analysis (PCA)*: Projecting the perceptual feature space onto its first two principal components, which explain 40.2% and 22.2% of the total variance, provides a visual comparison of the dataset distributions. The convex hull of π -SUB closely overlaps that of the real data, with nearby centroids, whereas Syrea occupies a partially overlapping but shifted region. SUIEB and PHISWID cluster together in an offset region along PC1, reflecting their characteristic brightness and chromatic differences, while SUID occupies only a small region of the feature space because of its limited size. The corresponding PCA projection is provided in the supplementary material.

Further checks are reported in the supplementary material and support the same conclusion. Cluster-wise visual inspection confirms that π -SUB images assigned to each real cluster carry the correct water type and depth label, a per-image nearest-neighbour analysis shows that 90% of real images find a π -SUB counterpart at a smaller feature distance than for any competing dataset, and cross-validation against four SQUID survey sites recovers mean matched depths that fall within each site’s reported range.

Together these analyses establish the hyper-realism of π -SUB: the lowest FID to the real distribution both globally (95 against 176 for the next-best dataset) and in every perceptual cluster, an OOD rate of 2.00%, the greatest overlap with real data in PCA space, and the only synthetic dataset covering all seven perceptual regimes. The ranking is not an artifact of the clustering configuration: repeating the analysis with eight clusters over a partially overlapping benchmark pool leaves π -SUB lowest overall and in every populated cluster, as reported in the supplementary material. Distributional similarity alone does not guarantee learning effectiveness, so the next subsection evaluates whether it translates into better generalization.

B. Generalization Performance Evaluation

Generalization is measured through the enhancement quality of models trained on each candidate dataset. Four architectures spanning distinct inductive biases are used: Pix2Pix [11], a conditional adversarial baseline; FUNIE-GAN [2], a generative adversarial network; PUIE-Net [12], a probabilistic CNN; and Phaseformer [13], a frequency-domain vision transformer. Each is trained independently on the real paired dataset UIEB [3] and on the synthetic datasets SUID [8], SUIEB [9], PHISWID [10], Syrea [5], and π -SUB, giving twenty-four trained models. All models are retrained from scratch under identical hyperparameters, input resolution, and schedule, with the number of training pairs equalized across datasets; the complete protocol, including the subsampling used to equalize dataset scale, is given in the supplementary material.

These twenty-four models are evaluated on six real benchmark underwater datasets, each contributing a distinct visual and optical character to the evaluation suite. UIEB [3] is the largest and most heterogeneous of these, comprising a broad collection of shallow-water scenes that include divers, fish, coral, rocks, and underwater cables and structures, captured under a wide range of illumination and turbidity conditions. U45 is a designated subset of UIEB, isolating forty-five of its most severely degraded images, characterized by low contrast, non-uniform lighting, and pronounced color casts, and is used to probe enhancement performance under the hardest conditions the parent dataset offers. RUIE [16] spans a wide variety of blue- and green-tinted underwater scenarios collected for underwater robotic perception tasks, and its imagery was later used as the source material from which LSUI [4] generated its enhanced model-based references. OceanDark [15] consists of low-light scenes captured at deep-sea sites under artificial illumination, which impart a characteristic green-dominant cast. SQUID [14] and FishTrac [39], by contrast, are both drawn from open, relatively clear oceanic waters and exhibit a predominantly blue color cast, with SQUID additionally providing stereo image pairs of coral reef scenes and FishTrac capturing fish-tracking footage in similarly blue-dominant conditions.

1) *Evaluation Metrics*: Because clean references are unavailable for real underwater imagery, enhancement is scored with no-reference metrics. The Underwater Image Quality Measure (UIQM) [41] is a weighted sum of colorfulness, sharpness, and contrast terms,

$$\text{UIQM} = c_1 \times \text{UICM} + c_2 \times \text{UISM} + c_3 \times \text{UIConM}, \quad (18)$$

where UICM, UISM, and UIConM quantify colour quality, edge sharpness, and local contrast, with empirically determined weights $c_1 = 0.0282$, $c_2 = 0.2953$, $c_3 = 3.5753$ [41]; higher is better. The Natural Image Quality Evaluator (NIQE) [42] measures the deviation of patch statistics from a model of natural scene statistics learned from undistorted images, requiring neither reference images nor human opinion scores; lower values indicate outputs that are statistically closer to natural imagery. A secondary colour-based metric, UCIQE [43], is reported in the supplementary material.

2) *Qualitative Analysis*: Figure 4 compares all four architectures on UIEB, OceanDark, and SQUID when trained on the real paired dataset UIEB, on the four competing synthetic

datasets, and on π -SUB. Complete qualitative comparisons on all six benchmarks are provided in the supplementary material, and the observations below hold across them.

Each training source imparts a recognizable character to the enhanced outputs. Models trained on UIEB perform strongly on UIEB and U45, benefiting from the overlap between training and evaluation distributions, but transfer poorly to RUIE, OceanDark, SQUID, and FishTrac, consistent with close adaptation to the colour and haze character of UIEB’s human-selected references. SUID-trained models show the strongest deviations, pushing bright regions toward saturation and mid-tones toward reddish-orange. SUIEB-trained models are inconsistent, sometimes leaving the raw cast intact and sometimes introducing warm casts of their own. PHISWID-trained models systematically under-correct, remaining washed-out and low in contrast, and on the yellow-dominant U45 scene PUIE-Net trained on PHISWID produces severe structural artifacts, indicating that its degradation model does not span this portion of the real colour-cast distribution. Syrea-trained models achieve balanced global correction but flatten chromatic content toward grey and appear over-smoothed. Models trained on π -SUB restore natural, saturated colour across all six benchmarks while preserving fine structure and object boundaries; the one caveat is a mildly warm tint on the blue-dominant SQUID and FishTrac scenes when paired with FUnIE-GAN, which the other three architectures do not exhibit.

a) **Quantitative Analysis:** Table III reports UIQM and NIQE for all twenty-four trained models on UIEB, OceanDark, and SQUID. The supplementary material has the full table across all six benchmarks.

TABLE III: UIQM ($\uparrow$) and NIQE ($\downarrow$) of enhancement models trained on the six candidate training datasets and evaluated on UIEB, OceanDark (OD), and SQUID. Within each architecture group, the best and second-best score per benchmark and metric are shaded **best** and **second**.

Method	UIEB		OD		SQUID	
	UIQM $\uparrow$	NIQE $\downarrow$	UIQM $\uparrow$	NIQE $\downarrow$	UIQM $\uparrow$	NIQE $\downarrow$
Pix2Pix@UIEB	3.2702	7.4912	3.0191	8.5609	2.7705	8.9478
Pix2Pix@SUID	3.1284	5.3240	2.8849	7.9288	2.8128	8.9778
Pix2Pix@SUIEB	3.2587	5.7889	3.3383	6.2387	3.1036	6.1689
Pix2Pix@PHISWID	3.3577	6.6060	3.2570	6.6190	3.0979	8.5969
Pix2Pix@Syrea	3.2587	4.2789	3.2441	5.9893	3.2033	7.4289
Pix2Pix@π-SUB	3.4129	2.6666	3.3495	3.4896	3.2777	3.1016
FUnIE-GAN@UIEB	3.2630	6.8676	2.7247	7.8561	2.4233	7.8850
FUnIE-GAN@SUID	3.2028	4.5057	2.6783	4.8080	3.0792	5.4466
FUnIE-GAN@SUIEB	3.2075	5.9874	2.7160	5.0544	2.7825	5.0768
FUnIE-GAN@PHISWID	3.2169	6.2559	2.9518	6.5030	3.0582	7.6149
FUnIE-GAN@Syrea	3.2345	3.8422	2.5171	5.6667	3.0516	4.7502
FUnIE-GAN@π-SUB	3.2393	2.4807	2.8874	3.9624	3.1080	3.4308
Phaseformer@UIEB	3.0298	6.4296	2.4974	7.8588	1.9655	7.3066
Phaseformer@SUID	2.9828	6.9559	2.6325	5.3154	2.3539	7.1548
Phaseformer@SUIEB	2.6954	5.4515	2.3843	4.4672	1.7822	5.0577
Phaseformer@PHISWID	3.1306	8.4086	2.6996	6.1256	2.9865	7.1385
Phaseformer@Syrea	2.8950	5.2779	2.2358	4.4180	2.3305	3.0397
Phaseformer@π-SUB	3.2329	2.8519	2.7367	3.9225	2.7771	3.0414
PUIE-Net@UIEB	3.0821	6.3810	2.6740	6.9020	2.5961	6.8937
PUIE-Net@SUID	3.1358	5.1542	2.6414	7.4799	2.7530	7.8411
PUIE-Net@SUIEB	3.2987	5.7597	2.8244	5.1229	2.7265	7.8757
PUIE-Net@PHISWID	3.0735	6.0496	2.8551	6.8193	2.7862	6.5016
PUIE-Net@Syrea	3.0412	3.2931	2.5540	5.1417	2.4545	3.4849
PUIE-Net@π-SUB	3.2269	2.6860	2.8579	2.9919	2.8298	3.1233

Models trained on π -SUB rank first or second in UIQM on UIEB, OceanDark, and SQUID in all twentyfour architecture–benchmark combinations shown, the best win rate of any candidate training set on these three benchmarks. For Pix2Pix, π -SUB is best on all three, with the largest margins on OceanDark (16.1% over SUID) and SQUID (16.5% over SUID). FUnIE-GAN is the setting in which baselines remain most competitive, with π -SUB best on SQUID and second on UIEB and OceanDark. The corresponding margins on U45, RUIE, and FishTrac are reported in the supplementary material and follow the same pattern. UIEB-trained models perform competitively with π -SUB on the in-distribution UIEB benchmark, reflecting their ability to reproduce the appearance characteristics of the training data. However, their performance degrades substantially on out-of-distribution datasets such as SQUID, which exhibits different water types, illumination conditions, and optical characteristics. For example, the π -SUB-trained FUnIE-GAN achieves a 28.3% higher UIQM on SQUID, indicating that the enhancement learned from UIEB is largely dataset-specific rather than physically generalizable. In contrast, the greater optical diversity and physically grounded degradation modeled by π -SUB enable the learned enhancement to transfer more effectively across unseen underwater environments. Finally, averaged across four UIE architectures and six real benchmark datasets, models trained on π -SUB improve UIQM by 9.46% over two strong competing synthetic datasets Syrea and 4.18% over PHISWID, while reducing NIQE by 23.98% and 48.78%, respectively.

Table IV compares the π -SUB-trained models against a classical prior-based method (ULAP) and state-of-the-art learned enhancement models. A π -SUB-trained model achieves the highest UIQM on every benchmark, with margins of 4.4% on UIEB, 5.9% on U45, 7.7% on RUIE, 10.9% on OceanDark, 12.5% on FishTrac, and 18.3% on SQUID over the strongest competing method in each case. The NIQE margins are larger still, from 21.8% on U45 to 64.1% on UIEB, where FUnIE-GAN@ π -SUB reduces the best baseline of 6.9152 to 2.4807. ULAP and the transformer-based Spectroformer remain competitive on individual benchmarks, but the π -SUB-trained models dominate both metrics overall. Because these gains persist across four architectures with distinct inductive biases, they are attributable to the training distribution rather than to any single model design.

Further, statistical analysis using the Friedman rank test with Nemenyi post-hoc comparison across all four architectures and six benchmarks confirms that π -SUB yields significantly better mean ranks on both UIQM and NIQE. On the challenging U45 subset, Phaseformer trained on π -SUB further exhibits the lowest variance (UIQM 10.92 %, NIQE 18.89 %) compared with Syrea (14.60% and 26.44%) and PHISWID (12.31% and 35.34%), indicating greater robustness. Full details are given in the supplementary material.

3) **Ablation Study:** To verify that each residual phenomenon in the Augmented Realism stage contributes to generalization, five training variants of π -SUB are constructed: *Base* (the deterministic MUIF subset I_b only), *Base+Bio* (I_b with biological effects), *Base+Haze* (I_b with volumetric haze), *Base+SP* (I_b with suspended particulate matter), and *Full* (all

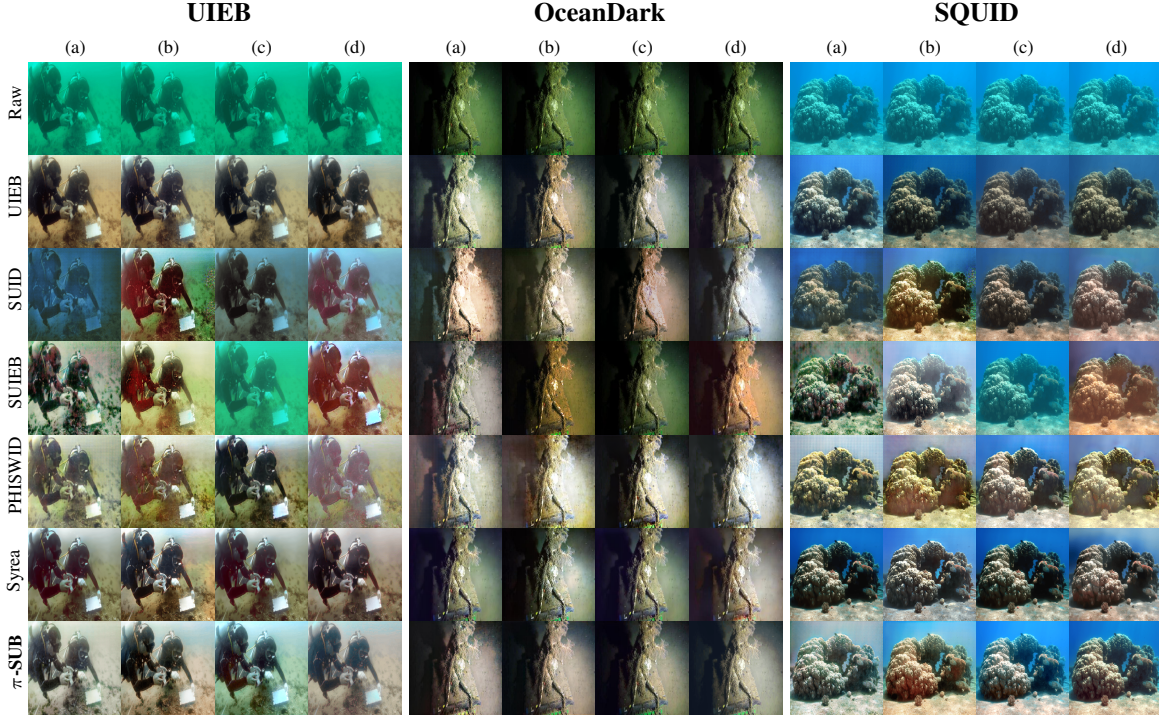


Fig. 4: Representative qualitative enhancement results on three real benchmarks: UIEB (mixed shallow-water scenes), OceanDark (low-light deep-sea scenes), and SQUID (clear oceanic scenes). Columns (a)–(d) correspond to Pix2Pix, FUnIE-GAN, Phaseformer, and PUIE-Net, respectively.

TABLE IV: Underwater image enhancement performance of state-of-the-art methods and models trained on π -SUB, in terms of UIQM ($\uparrow$) and NIQE ($\downarrow$). The best and second-best score per benchmark and metric are shaded **best** and **second**.

Method	UIEB		U45		RUIE		FishTrac		OceanDark		SQUID	
	UIQM $\uparrow$	NIQE $\downarrow$	UIQM $\uparrow$	NIQE $\downarrow$	UIQM $\uparrow$	NIQE $\downarrow$	UIQM $\uparrow$	NIQE $\downarrow$	UIQM $\uparrow$	NIQE $\downarrow$	UIQM $\uparrow$	NIQE $\downarrow$
ULAP [44]	1.3230	7.3794	2.2044	6.3523	2.2723	5.5231	1.3429	6.5523	1.8888	4.8517	0.8095	5.2161
P2CNet [45]	2.5195	7.8159	3.1237	6.2819	2.9362	7.4776	2.3442	11.3184	2.7917	8.6685	2.4927	8.0566
Phaseformer [13]	2.5371	9.0984	3.0136	10.0297	2.1309	7.9462	1.3561	6.9397	2.1360	6.5882	2.2675	9.5063
FUnIE-GAN [2]	3.0493	8.1176	2.7083	9.0159	3.0465	9.3881	2.8346	8.1613	2.9723	7.6801	2.1556	6.8013
PUIE-Net [12]	3.1128	6.9152	2.6778	7.9918	3.0292	7.7949	2.4554	6.4421	2.9079	8.1561	2.1935	7.7716
Pix2Pix [11]	3.2702	7.4912	3.2145	8.9644	2.9561	7.7668	3.0552	8.8543	3.0191	8.5609	2.7705	8.9478
UShape [4]	2.7857	7.4041	3.1912	7.2647	2.9351	7.4888	1.8180	10.7056	2.9122	7.8756	2.3135	8.9676
Spectroformer [46]	2.4380	7.0620	2.9508	7.4322	2.8367	6.2236	2.9669	4.9601	2.5432	5.1699	2.3157	4.4353
UDNet [47]	2.3913	7.5062	3.1420	7.1917	2.9282	6.6007	2.7387	6.6421	2.0922	4.9538	1.9418	7.3046
WaterNet [3]	2.5067	7.4418	3.2734	8.7466	3.0550	6.2991	2.1123	5.2207	2.3290	4.8337	2.0060	6.4154
Pix2Pix@MUSE [6]	-	-	3.0279	-	3.0726	-	-	-	2.7836	-	-	-
Pix2Pix@ π -SUB	3.4129	2.6666	3.2898	4.6579	3.3081	3.2704	3.4383	3.3387	3.3495	3.4896	3.2777	3.1016
FUnIE-GAN@ π -SUB	3.2393	2.4807	3.3986	5.3609	3.0059	3.3553	3.4147	3.7558	2.8874	3.9624	3.1080	3.4308
Phaseformer@ π -SUB	3.2329	2.8519	3.4655	5.2308	2.9786	3.3232	3.2278	4.0125	2.7367	3.9225	2.7771	3.0414
PUIE-Net@ π -SUB	3.2269	2.6860	3.4183	6.0644	3.1082	4.1418	3.3990	3.8792	2.8579	2.9919	2.8298	3.1233

three residual phenomena combined). FUnIE-GAN is retrained from scratch on each variant under identical hyperparameters and evaluated on the six real benchmarks. Figure 5 illustrates the contribution of each augmentation component to underwater image enhancement. The complete π -SUB configuration, combining all augmentation categories, consistently produces the best enhancement performance. This trend is confirmed quantitatively in Table V, where the UIQM and NIQE scores, averaged across four UIE architectures and six real-world underwater benchmark datasets, consistently favor the complete π -SUB configuration.

Removing any single residual phenomenon from the training distribution degrades performance relative to the full configuration. Biological effects alone improve warm-wavelength

TABLE V: Ablation study of π -SUB across six real-world benchmarks.

Variant	UIQM $\uparrow$	NIQE $\downarrow$
Base	2.8713	4.7283
Base+Bio	3.1138	3.7961
Base+Haze	3.0941	3.9254
Base+SP	2.9851	4.3382
Full	3.1829	3.4129

recovery in green-dominant scenes such as those in OceanDark, since biological absorption primarily attenuates the red and yellow bands, but they do not reconstruct the low-contrast haze artifacts of turbid open-water scenes in FishTrac and SQUID. Haze alone improves contrast but over-compensates in detail reconstruction. Only the *Full* configuration achieves

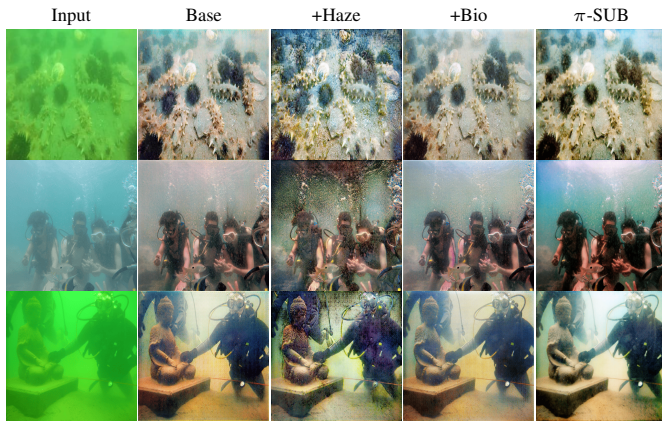


Fig. 5: Qualitative ablation of FUNIE-GAN under π -SUB training datasets.

top performance consistently across benchmarks, confirming that the residual phenomena are complementary rather than redundant components of a physically complete training distribution.

4) **Keypoint Feature Matching:** The practical value of enhancement is ultimately reflected in downstream tasks such as visual odometry, SLAM, and 3D reconstruction, which depend on reliable feature correspondences across frames. Matching is evaluated with SIFT [48], whose scale-space detector suits the low contrast, blur, and residual haze of underwater imagery; initial correspondences from Lowe’s ratio test are refined by RANSAC-based fundamental matrix estimation [49] to retain only geometrically consistent matches.

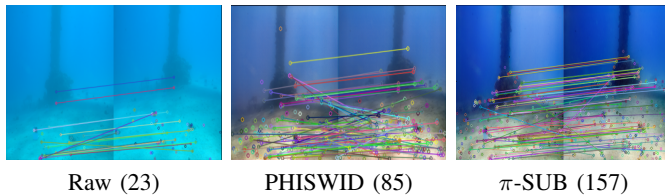


Fig. 6: SIFT keypoint matching between consecutive SQUID Katjaa dataset

On consecutive SQUID Katjaa frames as shown in Figure 6, the raw pair yields only 23 verified matches. PhaseFormer trained on PHISWID [10] raises this to 85, while training on π -SUB yields 157: a $6.8\times$ improvement over the raw frames and $1.8\times$ over PHISWID. The substantially higher count of verified correspondences shows that π -SUB better preserves the local image structure required for feature-dependent downstream applications.

In summary, the results validate the π -SUB framework from two complementary perspectives. Distributionally, the π -SUB dataset aligns more closely with real underwater imagery than existing synthetic datasets, achieving the lowest global and cluster-wise FID, an OOD rate of 2.00%, and complete coverage of the seven perceptual regimes observed in real benchmarks. Functionally, models trained on π -SUB consistently generalize better across six real benchmark datasets, achieving superior UIQM and NIQE while preserving natural color and structural detail. In the downstream feature-matching task, π -SUB nearly doubles the number of geometrically consistent keypoint matches compared with PHISWID. These results demonstrate that the combination of measured Jerlov optical properties, depth-dependent irradiance, and biological spectral

modulation produces a training distribution that improves both enhancement quality and downstream underwater perception.

V. CONCLUSION

This paper presents a novel physics-informed synthetic underwater benchmark framework named π -SUB. Using the proposed framework, the π -SUB dataset comprising of paired synthetic underwater–reference images spanning shallow-to-deep and coastal-to-oceanic environments. The framework makes three principal contributions: (i) a physically consistent degradation pipeline that models depth-dependent downwelling irradiance, biological optical processes including chlorophyll-*a* absorption, CDOM, and fluorescence, together with forward and backward environmental scattering across all ten Jerlov water types; (ii) a comprehensive paired benchmark accompanied by complete physical metadata; and (iii) a two-axis validation methodology based on *hyper-realism* and *generalizability*.

Hyper-realism was demonstrated through distributional analysis against the pooled distribution of five real-world benchmarks. π -SUB achieved a global FID of 95, a 46% reduction relative to the best existing synthetic benchmark (176, Syrea; the remaining comparators being SUID, SUIEB, and PHISWID), an out-of-distribution rate of only 2.00%, and was the only synthetic benchmark covering all seven perceptual regimes observed in real underwater imagery.

Generalizability was demonstrated using twenty-four enhancement models spanning generative, convolutional, and transformer-based architectures. Since Syrea and PHISWID are the strongest competing synthetic benchmarks in terms of realism and enhancement performance, they are used as the primary baselines. Averaged across four UIE architectures and six real benchmark datasets, models trained on π -SUB improve UIQM by 9.46% over Syrea and 4.18% over PHISWID, while reducing NIQE by 23.98% and 48.78%, respectively. In a downstream feature-matching task, π -SUB increases the number of geometrically consistent SIFT correspondences from 23 on the raw images to 157, representing a $1.8\times$ improvement over PHISWID.

These results demonstrate that physically complete modeling of underwater optical variability as presented in π -SUB produces synthetic data that is both hyper-realistic and generalizable for underwater image enhancement. Future work will extend the framework toward artificially illuminated deep-sea environments, highly turbid coastal water, and caustics that remain underrepresented in existing real-world benchmarks.

REFERENCES

- [1] X. Cong, Y. Zhao, J. Gui, J. Hou, and D. Tao, “A comprehensive survey on underwater image enhancement based on deep learning,” *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2026.
- [2] M. J. Islam, Y. Xia, and J. Sattar, “Fast underwater image enhancement for improved visual perception,” *IEEE robotics and automation letters*, vol. 5, no. 2, pp. 3227–3234, 2020.
- [3] C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, and D. Tao, “An underwater image enhancement benchmark dataset and beyond,” *IEEE transactions on image processing*, vol. 29, pp. 4376–4389, 2019.
- [4] L. Peng, C. Zhu, and L. Bian, “U-shape transformer for underwater image enhancement,” *IEEE Transactions on Image Processing*, vol. 32, pp. 3066–3079, 2023.

- [5] J. Wen, J. Cui, Z. Zhao, R. Yan, Z. Gao, L. Dou, and B. M. Chen, "SyreaNet: A physically guided underwater image enhancement framework integrating synthetic and real images," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5177–5183, 2023.
- [6] S. Li, T. Liu, Q. Jiang, Y. Li, J. Guo, L. Jiao, Y. Guo, and Z. Ni, "Realistic simulation of underwater scene for image enhancement," *IEEE Transactions on Geoscience and Remote Sensing*, 2025.
- [7] B. McGlamery, "A computer model for underwater camera systems," in *Ocean optics VI*, vol. 208, pp. 221–231, SPIE, 1980.
- [8] G. Hou, X. Zhao, Z. Pan, H. Yang, L. Tan, and J. Li, "Benchmarking underwater image enhancement and restoration, and beyond," *IEEE Access*, vol. 8, pp. 122078–122091, 2020.
- [9] C. Li, S. Anwar, and F. Porikli, "Underwater scene prior inspired deep underwater image and video enhancement," *Pattern recognition*, vol. 98, p. 107038, 2020.
- [10] R. Kaneko, T. Ueda, H. Higashi, and Y. Tanaka, "Phiswid: physics-inspired underwater image data set synthesized from rgb-d images," *APSIPA Transactions on Signal and Information Processing*, vol. 15, no. 1, pp. 1–25, 2026.
- [11] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5967–5976, 2017.
- [12] Z. Fu, W. Wang, Y. Huang, X. Ding, and K.-K. Ma, "Uncertainty inspired underwater image enhancement," in *European Conference on Computer Vision (ECCV)*, pp. 465–482, 2022.
- [13] M. R. Khan, A. Negi, A. Kulkarni, S. S. Phutke, S. K. Vipparthi, and S. Murala, "Phaseformer: Phase-based attention mechanism for underwater image restoration and beyond," in *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 9618–9629, 2025.
- [14] D. Berman, D. Levy, S. Avidan, and T. Treibitz, "Underwater single image color restoration using haze-lines and a new quantitative dataset," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 8, pp. 2822–2837, 2021.
- [15] T. P. Marques and A. B. Albu, "L²UWE: A framework for the efficient enhancement of low-light underwater images using local contrast and multi-scale fusion," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops - NTIRE*, pp. 483–492, IEEE/CVF, 2020. Paper associated with the OceanDark dataset, which contains low-light underwater images.
- [16] R. Liu, X. Fan, M. Zhu, M. Hou, and Z. Luo, "Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 12, pp. 4861–4875, 2020.
- [17] C. Desai, S. Benur, U. Patil, and U. Mudanagudi, "Rsuigm: Realistic synthetic underwater image generation with image formation model," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 21, no. 1, pp. 1–22, 2024.
- [18] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, "Watergan: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robotics and Automation letters*, vol. 3, no. 1, pp. 387–394, 2017.
- [19] F. Zhang, S. You, Y. Li, and Y. Fu, "Atlantis: Enabling underwater depth estimation with stable diffusion," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11852–11861, 2024.
- [20] D. Akkaynak and T. Treibitz, "A revised underwater image formation model," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6723–6732, 2018.
- [21] L. Yang, B. Kang, Z. Huang, Z. Zhao, X. Xu, J. Feng, and H. Zhao, "Depth anything v2," *Advances in Neural Information Processing Systems*, vol. 37, pp. 21875–21911, 2024.
- [22] A. Bricaud, A. Morel, M. Babin, K. Allali, and H. Claustre, "Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models," *Journal of Geophysical Research: Oceans*, vol. 103, no. C13, pp. 31033–31044, 1998.
- [23] J. S. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *IEEE Journal of Oceanic Engineering*, vol. 15, no. 2, pp. 101–111, 1990.
- [24] D. Akkaynak and T. Treibitz, "Sea-thru: A method for removing water from underwater images," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1682–1691, 2019.
- [25] N. G. Jerlov, *Marine optics*, vol. 14. Elsevier, 1976.
- [26] M. G. Solonenko and C. D. Mobley, "Inherent optical properties of jerlov water types," *Applied optics*, vol. 54, no. 17, pp. 5392–5401, 2015.
- [27] A. Morel and H. Loisel, "Apparent optical properties of oceanic water: dependence on the molecular scattering contribution," *Applied Optics*, vol. 37, no. 21, pp. 4765–4776, 1998.
- [28] C. A. Williamson and R. C. Hollins, "Depth profiles of jerlov water types," *Limnology and Oceanography Letters*, vol. 8, no. 5, pp. 781–788, 2023.
- [29] C. D. Mobley, *Light and Water: Radiative Transfer in Natural Waters*. San Diego, CA: Academic Press, 1994.
- [30] H. Buiteveld, J. H. M. Hakvoort, and M. Donze, "The optical properties of pure water," in *Ocean Optics XII*, vol. 2258, pp. 174–183, SPIE, 1994.
- [31] H. Koschmieder, *Theorie der horizontalen Sichtweite*. Keim & Nemnich, 1925.
- [32] Z. Lee and S. Shang, "Visibility: how applicable is the century-old koschmieder model?," *Journal of the Atmospheric Sciences*, vol. 73, no. 11, pp. 4573–4581, 2016.
- [33] R. Kaneko, Y. Sato, T. Ueda, H. Higashi, and Y. Tanaka, "Marine snow removal benchmarking dataset," in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 771–778, 2023.
- [34] O. Bagoren, S. Isaacson, S. Sundar, Y.-C. Sun, A. Sheppard, H. Ma, A. Shariff, R. Vasudevan, and K. A. Skinner, "Surfslam: Sim-to-real underwater stereo reconstruction for real-time slam," *arXiv preprint arXiv:2601.10814*, 2026.
- [35] H. R. Gordon, "Diffuse reflectance of the ocean: the theory of its augmentation by chlorophyll a fluorescence at 685 nm," *Applied Optics*, vol. 18, no. 8, pp. 1161–1166, 1979.
- [36] P.-W. Zhai, E. Boss, B. Franz, P. J. Werdell, and Y. Hu, "Radiative transfer modeling of phytoplankton fluorescence quenching processes," *Remote sensing*, vol. 10, no. 8, p. 1309, 2018.
- [37] S. Maritorena, A. Morel, and B. Gentili, "Determination of the fluorescence quantum yield by oceanic phytoplankton in their natural habitat," *Applied Optics*, vol. 39, no. 36, pp. 6725–6737, 2000.
- [38] J. T. Kirk, *Light and photosynthesis in aquatic ecosystems*. Cambridge university press, 1994.
- [39] M. Dawkins, J. Prior, B. Lewis, R. Faillettaz, T. Banez, M. Salvi, A. Rollo, J. Simon, M. Campbell, M. Lucero, *et al.*, "Fishtrack23: An ensemble underwater dataset for multi-object tracking," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 7167–7176, 2024.
- [40] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," *Advances in neural information processing systems*, vol. 30, 2017.
- [41] K. Panetta, A. Samani, and S. Agaian, "A robust no-reference, no-parameter, transform domain image quality metric for evaluating the quality of color images," *IEEE Access*, vol. 6, pp. 10979–10985, 2018.
- [42] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, pp. 209–212, Mar. 2013.
- [43] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 6062–6071, 2015.
- [44] W. Song, Y. Wang, D. Huang, and D. Tjondronegoro, "A rapid scene depth estimation model based on underwater light attenuation prior for underwater image restoration," in *Pacific rim conference on multimedia*, pp. 678–688, Springer, 2018. doi: 10.1007/978-3-030-00776-8_62.
- [45] Y. Rao, W. Liu, K. Li, H. Fan, S. Wang, and J. Dong, "Deep color compensation for generalized underwater image enhancement," *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.
- [46] R. Khan, P. Mishra, N. Mehta, S. S. Phutke, S. K. Vipparthi, S. Nandi, and S. Murala, "Spectroformer: Multi-domain query cascaded transformer network for underwater image enhancement," in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 1454–1463, 2024.
- [47] A. Saleh, M. Sheaves, D. Jerry, and M. R. Azghadi, "Adaptive deep learning framework for robust unsupervised underwater image enhancement," *Expert Systems with Applications*, vol. 268, p. 126314, 2025.
- [48] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [49] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.