

Underwater Image Restoration Through a Prior Guided Hybrid Sense Approach and Extensive Benchmark Analysis

Summarized on: May 21, 2024

1. Introduction / Abstract

The Core Problem: Underwater photographs are essential for exploring and understanding our oceans, but they almost always suffer from poor quality. Because water absorbs and scatters light unevenly—absorbing red light first, then green, then blue—images often have a strong blue or green color cast and appear blurry and washed out. This makes it difficult to see details, which hinders scientific research like marine biology and archaeology.

Proposed Solution and Main Finding: To solve this, researchers developed a new Artificial Intelligence (AI) framework called **GuidedHybSensUIR**. This “hybrid” system cleverly combines two different AI approaches: one that excels at restoring fine, local details and another that is great at understanding the “big picture” to fix the overall color balance. The system is guided by a novel **Color Balance Prior** (a pre-calculated guide that tells the AI what a natural color balance should look like). The researchers also created a comprehensive benchmark by combining several real-world datasets to fairly test their model. Their main finding is that their method significantly outperforms 37 other state-of-the-art techniques, producing clearer, more color-accurate underwater images.

2. Methodology

The researchers designed their AI model using a U-shaped architecture, a popular structure for image restoration tasks. The process can be broken down into three main stages:

1. **Encoding (Analyzing the Image):** The original, distorted underwater image is fed into the network. As it moves down the “U,” it passes through several **Detail Restorer** modules. These modules use **CNNs (Convolutional Neural Networks)**, a type of AI that is excellent at identifying local patterns like edges, textures, and shapes) to focus on restoring small-scale, blurry details in the image. At each step, the image is downsized to focus on progressively larger features.
2. **Bottleneck (Global Understanding):** At the bottom of the “U,” the most compressed representation of the image is processed by the **Feature Contextualizer** module. This powerful module uses **Transformers** (a type of AI originally designed for language that excels at understanding long-range relationships between different parts of the data). Instead of analyzing patches of the image, this Transformer looks at the relationships

between the color channels themselves. Guided by the **Color Balance Prior**, it works to correct the global color cast across the entire image.

3. **Decoding (Reconstructing the Image):** The model then works its way back up the “U” to reconstruct the final, high-quality image. At each step, **Scale Harmonizer** modules fuse the global color information from the bottleneck with the fine-grained detail information saved from the corresponding encoding step. This ensures that the final image is both sharp and has realistic, balanced colors.

3. Theory / Mathematics

The researchers used two key theoretical concepts to guide their model’s learning process: a Color Balance Prior and a composite loss function.

Color Balance Prior The idea behind the prior comes from the “Gray World Assumption,” which states that for a typical image taken in normal light, the average of all red, green, and blue values across the image should be a neutral gray. Underwater images violate this, as they are heavily skewed towards blue or green.

To guide the AI, the researchers created a simple but effective prior by averaging the color channels for each pixel. The formula is:

$$\text{Prior}_i(x, y) = \frac{R(x, y) + G(x, y) + B(x, y)}{3}$$

What this means: For every pixel at position (x, y) , this formula calculates the average of its Red (R), Green (G), and Blue (B) values. This average value is then used for all three channels, creating a perfectly color-balanced (effectively grayscale) reference. This prior doesn’t represent the final desired image, but it acts as a strong directional hint for the AI, teaching it how to neutralize the unnatural color cast.

Loss Function A loss function is a formula that measures how “wrong” the AI’s prediction is compared to the ideal target. The researchers combined three different loss functions to evaluate the restored image from multiple perspectives.

$$L = w_1 \cdot L_f + w_2 \cdot L_s + w_3 \cdot L_p$$

What this means: The total error (L) is a weighted sum of three components: * L_f (**Fidelity Loss**): This measures the direct, pixel-by-pixel difference between the AI’s restored image and the ground-truth clear image. It ensures the basic colors and brightness are accurate. * L_s (**Structural Loss**): This measures the similarity in structure, contrast, and luminance. It’s better at capturing how humans perceive similarity than just comparing raw pixel values. * L_p

(Perceptual Loss): This uses another pre-trained neural network to judge whether the restored image *looks* realistic to a human. It focuses on preserving textures and complex features, leading to more visually pleasing results.

4. All Figures and Tables Explained

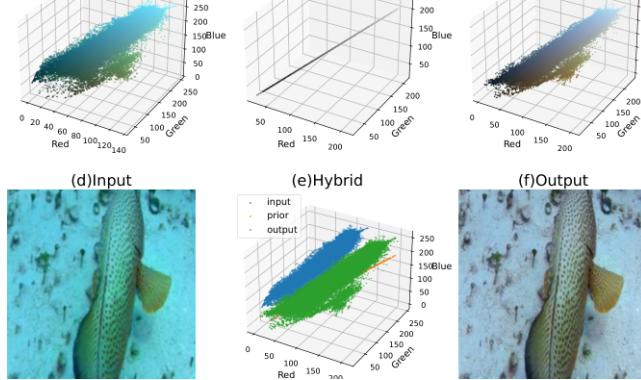


Figure 1: This figure

visualizes the color distribution of images as 3D scatter plots. (a) shows the input image’s colors are skewed towards green/blue. (b) shows the Color Balance Prior as a perfect diagonal line (balanced colors). (c) shows the output image’s colors are much more balanced and centered. (e) overlays all three, showing how the prior guides the input towards the restored output. (d) and (f) show the actual before-and-after images.



Figure 2: This is the main diagram of the GuidedHybSensUIR architecture. It shows the U-shaped path: the input image goes through the encoder (with Detail Restorers), into the bottleneck (Feature Contextualizer guided by the prior), and back up through the decoder (with Scale Harmonizers) to produce the final restored image. The gray arrows represent “skip connections” that carry fine-detail information from the encoder directly to the decoder.



Figure 3: This shows the architecture of the Nonlinear Activation-Free Block (NAFB), a key component of the Detail Restorer. It uses simple gating and channel attention mechanisms to efficiently restore intricate details without adding heavy computational cost.



Figure 4: This illustrates the ContextBlock, a part of the Detail Restorer. It is designed to capture important contextual information across the image, helping the model make more informed decisions about detail restoration.



Figure 5: This diagram details the Feature Contextualizer module. It shows how the image features and prior features are processed by four Multi-Attention Quaternion (MAQ) blocks. Each MAQ block contains three types of Transformers (ACT, KFT, SAT) that work in parallel to analyze color relationships.



Figure 6: This conceptually compares the three different attention mechanisms inside the MAQ blocks. It shows how they select their Query (Q), Key (K), and Value (V) to focus on different things: adjusting the image color (ACT), refining the prior (KFT), or analyzing the image’s internal relationships (SAT).



Figure 7: This shows the detailed architecture of the Adjust Color Transformer (ACT). Its job is to adjust the image’s color channels by paying attention to the guidance provided by the Color Balance Prior.

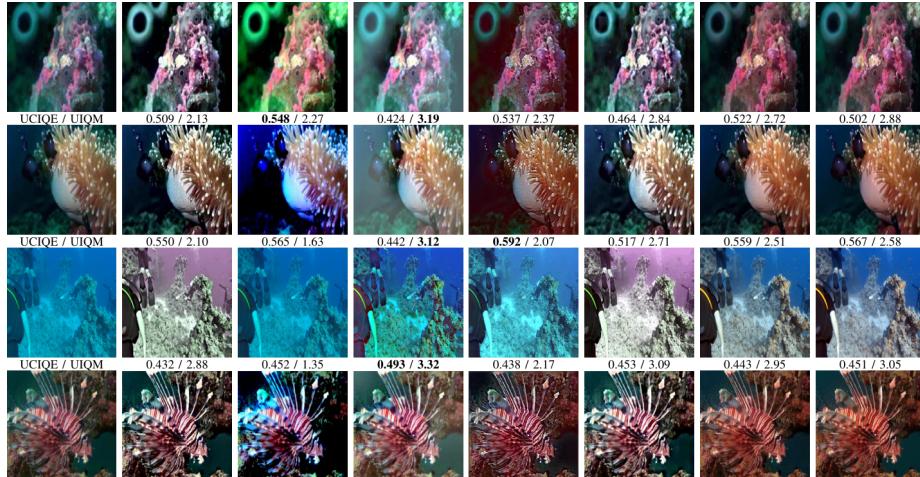


Figure 8: This shows the architecture of the Scale Harmonizer module. Its role

in the decoder is to intelligently blend the coarse, global color information from the lower layers with the fine-grained detail information from the encoder’s skip connections.

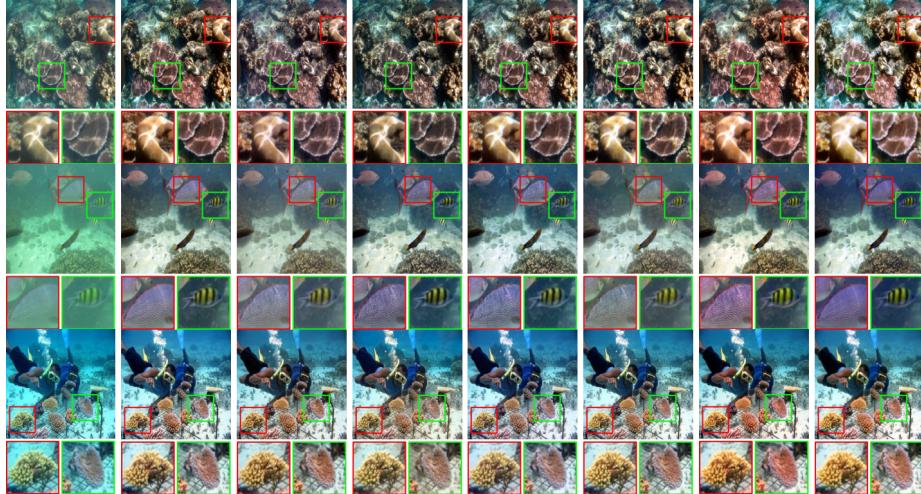


Figure 9: This provides a visual comparison with traditional (non-AI) restoration methods. The results show that traditional methods often fail, leaving color casts or creating overexposed areas. In contrast, the proposed method (“Ours”) produces a much clearer and more color-accurate image, very close to the “Reference” (ground truth).

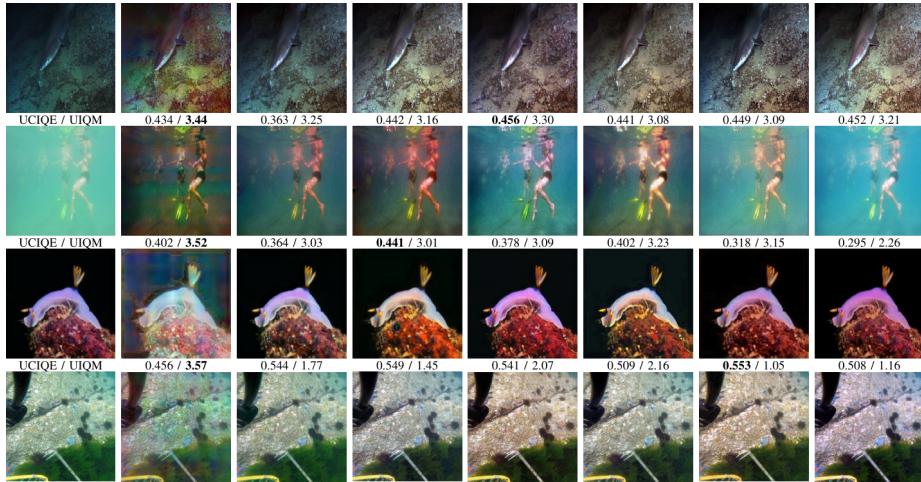


Figure 10: This is a visual comparison with other state-of-the-art deep learning methods. The magnified insets highlight that the proposed method produces images with better color saturation, contrast, and richer details (e.g., in the red and green boxes) than its competitors.



Figure 11: This shows a visual comparison on “unpaired” test images (where no perfect reference image exists). This tests how well the model generalizes to new, unseen data. The images show that the proposed method produces visually superior results compared to other top-performing AI models, even in challenging scenarios.

Table 1: This table presents a quantitative comparison on test sets where reference images are available. It uses metrics like PSNR and SSIM (higher is better) and LPIPS (lower is better). The results show that the proposed method (“Ours”) achieves the best or second-best scores across all datasets and metrics, proving its superior performance.

Table 2: This table shows results using non-reference metrics (UCIQE and UIQM), which attempt to judge quality without a ground truth image. The proposed method scores competitively, demonstrating strong performance in attributes like colorfulness and contrast.

Table 3: This table shows a quantitative comparison on unpaired test sets using the UCIQE metric. It demonstrates that the proposed method performs robustly and generalizes well to diverse and challenging underwater environments.

Table 4: This is an ablation study, which tests the importance of each component of the model by removing it and measuring the drop in performance. The results show that every single module—the Detail Restorer (DR), Feature Contextualizer (FC), Scale Harmonizer (SH), and the Color Balance Prior—contributes positively to the final result. The full model (bottom row) achieves the highest scores, confirming the effectiveness of the proposed design.

5. Conclusion

The researchers successfully developed a novel framework, GuidedHybSensUIR, that effectively restores degraded underwater images. By combining the local detail-finding strengths of **CNNs** with the global context-understanding power of **Transformers**, and guiding the entire process with a simple yet powerful **Color Balance Prior**, their model corrects severe color casts and restores blurry details. As demonstrated in comprehensive experiments against 37 other methods on a newly established benchmark (Table 1, Figures 9-11), their approach sets a new state-of-the-art in the field.

Why It Matters: This research provides a powerful new tool for scientists, marine biologists, and ocean explorers, allowing them to see the underwater

world with unprecedented clarity, which can accelerate discovery and aid in the conservation of marine ecosystems.