iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: V Month of publication: May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82727

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Underwater Image Enhancement Using Deep Learning: A Multi-Stage Processing Approach

Mrs. Nalini H C[1], Meghana Y S[2], Akshatha K R[3], Jeevan H R[4], Likitha D K[5]
[1]*Department of ISE, RIT Hassan*
[2, 3, 4, 5]*Department of Information Science and Engineering, Rajeev Institute of Technology, Hassan — 573201*
Visvesvaraya Technological University, Belagavi, Karnataka, India

Abstract: Capturing images beneath the water surface is fundamentally different from photography in air. Water selectively absorbs different wavelengths of light, scatters photons through suspended particles, and strips images of natural colour, contrast, and sharpness before they ever reach a sensor. The result is a degraded visual record that makes downstream tasks — coral reef surveys, pipeline inspections, autonomous vehicle navigation, and archaeological documentation — significantly harder than they need to be. This paper presents a systematic, multi-stage processing pipeline designed to address these degradation effects without requiring specialised underwater hardware. Starting from a raw degraded image, the system performs histogram-guided colour compensation on the red and blue channels using the green channel as a stable reference, applies Gray World white balancing to neutralise residual colour cast, and then branches into parallel enhancement paths: unsharp-masking for edge and detail recovery, and HSV-domain histogram equalisation for global contrast improvement. The two enhanced streams are subsequently fused through both an averaging strategy and a Principal Component Analysis (PCA) weighted combination. Quality is assessed quantitatively using Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) against reference images. Experimental results demonstrate measurable and visually convincing improvements across a representative set of underwater scenes. The modular design of the pipeline ensures that individual stages can be independently upgraded, making the system readily extensible as algorithmic advances emerge.
Keywords: Underwater image enhancement; colour compensation; Gray World algorithm; histogram equalisation; unsharp masking; image fusion; PCA fusion; PSNR; MSE; deep learning.

I. INTRODUCTION

Underwater photography carries a burden that surface imaging does not: physics works against the camera at every depth. Red wavelengths are absorbed within a few metres of the surface, robbing images of warmth and colour balance. Blue and green photons survive longer but are scattered by suspended particles, wrapping everything in a diffuse haze. By the time an image is formed, it typically presents low contrast, a heavy blue-green cast, and a loss of fine detail that makes it difficult to extract reliable information. These are not merely aesthetic shortcomings — they translate directly into reduced accuracy in any computer vision task that depends on the imagery.

The domains where this matters are broad. Marine biologists monitoring coral health need colour-accurate images to detect early signs of bleaching. Naval engineers inspecting submarine infrastructure depend on high-contrast imagery to spot cracks and corrosion. Operators of autonomous underwater vehicles (AUVs) require sharp, well-lit scenes to navigate safely and avoid collisions. Oceanographers conducting archaeological surveys need to resolve fine texture in sediment and artefact surfaces. In each case, the raw image falling short means the analysis falling short.

Classical image processing approaches — including histogram stretching, Retinex-based methods, and channel-specific equalisation — have offered partial remedies for decades. Their strength lies in computational efficiency and interpretability. Their weakness lies in inflexibility: a method tuned for clear tropical waters tends to fail in turbid estuarine or polar deep-sea environments, because it embeds assumptions about the degradation that simply do not generalise.

Deep learning has changed expectations in this space. Convolutional neural networks (CNNs), generative adversarial networks (GANs), and transformer architectures have all demonstrated the ability to learn complex image-to-image mappings from data, adapting to the statistical properties of whatever degradation they have been trained on. When paired with sufficiently diverse training sets, these models can surpass classical methods by a meaningful margin. Transfer learning further lowers the barrier to deployment, since models pretrained on large general datasets can be fine-tuned on domain-specific underwater corpora with modest computational investment.

The system described in this paper takes a pragmatic stance: it combines principled classical processing steps in a carefully ordered pipeline, evaluated rigorously with quantitative metrics. The contribution is not a single novel algorithm but a coherent workflow — compensation, white balancing, contrast enhancement, sharpening, fusion, and evaluation — that addresses the degradation problem end-to-end and produces results that are both measurable and visually convincing.

II. RELATED WORK

Research into underwater image enhancement has a decades-long history, and the literature reflects a gradual shift from purely physics-driven models towards data-driven and hybrid approaches.

Nayar and Narasimhan [1][2] established foundational work on imaging in adverse media, demonstrating that scattering and absorption could be modelled chromatically and compensated with white balance correction. Their Gray World framework, which assumes that the average colour of a natural scene should approximate neutral grey, remains a widely used baseline to this day because it is computationally inexpensive and requires no scene-specific tuning.

Tan [3] approached haze removal from a single image using contrast maximisation as the driving principle, observing that haze-free images statistically have higher local contrast than their degraded counterparts. While developed for atmospheric haze, the core insight carries over to the underwater scattering problem. Fattal [4] proposed a complementary single-image dehazing algorithm that estimated the transmission map by analysing independent components of the image colour channels, offering an alternative route to recovering scene radiance from a veiled observation.

He, Sun, and Tang [5] introduced the dark channel prior, one of the most influential ideas in the dehazing literature. The prior observes that, in most local patches of outdoor scenes, at least one colour channel has very low intensity. Deviations from this prior indicate the presence of haze or scattering media, and the prior can be inverted to estimate and subtract the scattering contribution. Although originally applied to atmospheric haze, adaptations for the underwater domain have followed, with modifications to account for the different spectral absorption profile of water.

More recent work has extended these ideas into the deep learning era. Methods based on CNNs, U-Net encoder-decoder architectures, and cycle-consistent GANs have been applied to paired and unpaired underwater image datasets including EUVP, UIEB, and UFO-120, achieving state-of-the-art results on standard benchmarks. Despite this progress, generalisation across dramatically different water types remains an open challenge, and no single method has established dominance across all deployment conditions.

III. PROBLEM FORMULATION

Given a degraded underwater image I(x), acquired under conditions of spectral absorption, forward and backward light scattering, and low ambient illumination, the goal is to compute an enhanced image J(x) that satisfies two simultaneous requirements: it should maximise perceptual quality as measured by objective metrics (PSNR, MSE), and it should support accurate downstream computer vision tasks such as object detection and scene classification.

The specific technical sub-problems that must be addressed are: (a) accurate modelling and inversion of spectrally-dependent light attenuation in the red channel; (b) removal of the blue-green colour cast introduced by differential scattering; (c) recovery of local contrast in regions dominated by haze; (d) restoration of edge sharpness lost to blur and diffraction; and (e) combination of multiple enhancement outputs without introducing artefacts.

The system must generalise across diverse underwater conditions — different depths, water clarity levels, turbidity profiles, and lighting geometries — without requiring manual parameter tuning for each new scene. Computational efficiency is a secondary constraint, but the pipeline should remain feasible on commodity hardware.

IV. PROPOSED SYSTEM ARCHITECTURE

A. Overview

The proposed pipeline is organised as a sequence of modular processing stages, each with a well-defined input-output contract. This architecture simplifies validation, allows independent benchmarking of individual stages, and makes substitution of improved algorithms straightforward.

The five primary stages are: (1) image acquisition and standardisation, (2) colour compensation and white balancing, (3) parallel contrast and sharpness enhancement, (4) image fusion, and (5) quality evaluation.

B. Image Acquisition and Standardisation

The input underwater image is first resized to a standardised resolution to ensure consistent behaviour of all subsequent fixed-architecture processing steps. The image is then decomposed into its Red, Green, and Blue channel representations. Histogram analysis is performed to quantify the distribution of pixel intensities across channels, providing a diagnostic view of colour imbalance and contrast deficiency before any correction is applied.

C. Colour Compensation

Underwater images systematically lose red channel energy as a function of depth, because water absorbs red wavelengths far more aggressively than green or blue. To compensate, the red and blue channels are adjusted using the green channel as a stable reference. Pixel values are first normalised to the unit interval, and the per-channel mean intensities are computed. The correction formula is:

$$R'(i,j) = [R(i,j) + (\mu G - \mu R)(1 - R(i,j)) \cdot G(i,j)] \times maxR$$

An analogous formula is applied to the blue channel, using the same green-channel reference. This approach restores the lost colour energy without relying on any scene-specific depth estimate, making it applicable across a wide range of acquisition conditions.

D. White Balancing (Gray World Algorithm)

Following channel compensation, residual colour cast is addressed using the Gray World assumption: the mean value of each colour channel in a natural, well-illuminated scene should be approximately equal. The algorithm computes the global mean intensity across all three channels, determines an overall grey mean, and scales each channel by the ratio of the grey mean to the channel-specific mean:

$$C'(i,j) = C(i,j) \times \mu grey / \mu C$$

where C denotes each of the three colour channels. This equalisation produces images in which no single channel dominates, effectively neutralising the blue-green tint that characterises most underwater photography.

E. Parallel Enhancement: Contrast and Sharpness

After colour correction, the pipeline branches into two independent enhancement paths that run in parallel. The first path addresses global contrast. The white-balanced image is converted from RGB colour space to HSV, isolating the Value (V) channel which encodes luminance independently of hue and saturation. Histogram equalisation is applied to the Value channel, redistributing the pixel intensity distribution to span the full dynamic range. The image is then converted back to RGB. This approach improves brightness and tonal separation without introducing colour artefacts.

The second parallel path addresses local sharpness. Unsharp masking is applied: a Gaussian-blurred copy of the white-balanced image is generated, and the difference between the original and the blurred version is computed at the pixel level. Adding a scaled version of this difference back to the original amplifies edge gradients and recovers fine detail. The result is an image with cleaner boundaries around objects such as fish fins, coral branches, and sediment textures.

F. Image Fusion

The outputs of the two parallel enhancement paths are combined through fusion. Two fusion strategies are implemented. Average fusion computes a simple per-pixel mean of the contrast-enhanced and sharpness-enhanced images, providing a balanced blend with equal weight given to both enhancements. PCA-based fusion treats each enhanced image as a data source and uses Principal Component Analysis to determine optimal combination weights based on the variance structure of the two images. Greater weight is assigned to the image (or combination) that carries more informative variance. This tends to produce a fused result that is both sharper and better exposed than either constituent image alone.

G. Quality Evaluation

The final enhanced image is compared against a reference image using two objective metrics. Mean Squared Error (MSE) is computed as the average squared per-pixel difference between the reference and enhanced images. Peak Signal-to-Noise Ratio (PSNR) is derived from MSE by the standard formula:

$$PSNR = 10 \times \log_{10} (255^2 / MSE)$$

A lower MSE and a higher PSNR indicate closer correspondence between the enhanced and reference images. These metrics provide an objective, reproducible basis for comparing the present pipeline against alternative approaches.

V. IMPLEMENTATION DETAILS

A. Software Environment

The entire pipeline is implemented in Python 3.8 using the PIL/Pillow library for image input/output and colour space operations, NumPy for array-level computation, and Matplotlib for visualisation. The Streamlit framework is used to provide an interactive graphical interface, allowing users to upload images, adjust sharpening parameters, and inspect results side by side. No proprietary software or specialised hardware accelerators are required for inference.

B. Dataset

Three benchmark datasets are used for training, validation, and evaluation. The EUVP (Enhancing Underwater Visual Perception) dataset provides over 12,000 paired images of degraded and reference-quality underwater scenes spanning diverse environments. The UFO-120 dataset contributes 1,500 image triplets with semantic annotations. The Sea-Thru dataset provides calibrated raw images with associated depth maps, enabling evaluation of depth-conditioned approaches. Hardware specifications for full training runs are an Intel Core i7 processor, an NVIDIA GeForce GTX 1650 or RTX 3050 GPU, 16 GB RAM, and a 512 GB SSD.

C. Algorithmic Challenges

Several practical difficulties arose during implementation. Pixel value overflow occurred when scaling adjusted channel values beyond the 8-bit range; this was resolved by clamping values to [0, 255] after each arithmetic operation. Pixel-level loop operations in Python are inherently slow for high-resolution images; these were replaced with vectorised NumPy operations wherever possible. Maintaining natural colour balance after multi-stage processing required careful ordering of operations — white balancing after compensation, not before — to avoid compounding corrections that partially cancel each other. Dataset size limitations were mitigated through careful selection of representative test images from across the available benchmarks.

VI. RESULTS AND DISCUSSION

A. Visual Assessment

The pipeline was applied to a representative selection of seven underwater images drawn from the EUVP and UFO-120 datasets. All input images exhibit the characteristic underwater degradation pattern: pronounced blue-green tint, reduced red channel energy, low contrast in mid-tone regions, and soft edges around foreground objects such as fish and coral structures.

After colour compensation and white balancing, the images display a markedly more neutral tone. The blue-green cast is substantially reduced and, in most cases, nearly eliminated. Red channel energy is visibly restored: warm tones appear in coral and rock surfaces that were previously washed out. This stage alone represents the most significant perceptual improvement in the pipeline.

The contrast enhancement stage further separates tonal regions, making mid-tone objects stand out more clearly against background water. The sharpening stage adds crispness to fish fins, coral branch tips, and rock surface textures. The PCA-fused output consistently presents better detail preservation than the average-fused output, at the cost of slightly stronger contrast transitions in some scenes. The average-fused output tends to appear more natural and conservative.

B. Quantitative Results

Table I presents MSE and PSNR values for the average and PCA fusion outputs across the seven test images, compared against corresponding reference images.

TABLE I
Performance Metrics: MSE and PSNR for Average and PCA Fusion Outputs

Image	Avg Fusion MSE	Avg Fusion PSNR (dB)	PCA Fusion MSE	PCA Fusion PSNR (dB)
Image 1	42.7	31.8	38.2	32.3
Image 2	55.3	30.7	49.6	31.2
Image 3	38.1	32.3	34.0	32.8
Image 4	61.4	30.2	57.1	30.6

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Image 5	47.6	31.4	43.3	31.8
Image 6	33.9	32.8	29.5	33.4
Image 7	50.2	31.1	45.8	31.5
Mean	47.0	31.5	42.5	31.9

PCA-based fusion consistently outperforms average fusion on both metrics, reducing MSE by an average of 4.5 points and improving PSNR by approximately 0.4 dB. While these margins may appear modest, they are consistent across all seven test images, indicating that the PCA weighting strategy reliably extracts more informative features from the two enhancement streams than simple averaging. The absolute PSNR values in the 30–33 dB range are competitive with published results for non-deep-learning enhancement methods on similar benchmarks.

C. Discussion

The results confirm that the ordered, multi-stage pipeline produces consistent improvements across a range of underwater scenes. Several observations merit discussion.

First, the ordering of processing stages matters considerably. Applying white balancing before colour compensation produces inferior results because the Gray World scaling modifies channel means before the red and blue channels have been boosted, causing the subsequent compensation to overshoot. The ordering implemented here — compensation first, then white balancing — avoids this artefact. Second, the choice of colour space for contrast enhancement matters. Applying histogram equalisation directly in RGB space produces colour distortion because equalising each channel independently disrupts the relationships between channels. Isolating the Value channel in HSV space and equalising only luminance preserves hue and saturation faithfully.

Third, PCA fusion's advantage over simple averaging reflects the statistical structure of the enhancement problem. The sharpened and contrast-enhanced images carry different kinds of information: the former is richer in high-frequency edge content, the latter in low-frequency tonal structure. PCA identifies the combination that maximises information content across this two-dimensional space, rather than treating both streams as equally informative.

Limitations of the current system include sensitivity to extremely turbid water, where the assumption that the green channel is a reliable reference for compensation breaks down. Additionally, the Gray World assumption fails in scenes that are genuinely dominated by a single colour, such as coral gardens where red hues cover most of the frame. Future work will address these failure modes through adaptive, scene-aware selection of correction parameters.

VII. TESTING AND VALIDATION

System testing was conducted at three levels: unit, functional, and integration. Unit testing verified each processing module in isolation, confirming correct histogram generation, accurate channel separation, proper colour scaling in compensation and white balancing, effective edge amplification in sharpening, and correct metric computation in the evaluation stage. Common issues identified at this level included division-by-zero in degenerate images with constant channels and pixel value overflow in the compensation arithmetic — both resolved before integration.

Functional testing examined end-to-end behaviour for a representative set of input conditions: images with strong blue tints, images with low ambient contrast, images with motion blur, and images at different resolutions. All passed qualitative acceptance criteria (visible improvement in colour and contrast) and quantitative thresholds (PSNR improvement over the unenhanced input).

Integration testing traced the data flow across all stages, verifying format compatibility (PIL Images and NumPy arrays) and dimensional consistency from input to output. A formal test case table covering twelve test scenarios — including valid image processing, colour correction verification, sharpness validation, fusion method comparison, performance metric correctness, multi-image batch processing, variable resolution handling, and invalid input management — was constructed and executed. All twelve cases passed.

VIII. FUTURE SCOPE

Several directions for extending this work are identified. Real-time processing for embedded deployment aboard AUVs and ROVs is the most immediate practical goal; achieving this requires model compression through knowledge distillation, quantisation, and neural architecture search to produce lightweight inference pipelines that fit within the memory and power budgets of underwater hardware.

Expanded and more geographically diverse benchmark datasets would enable more confident claims about generalisation. Current benchmarks, including EUVP and UIEB, are valuable but contextually limited; a dataset spanning polar, tropical, estuarine, and deep-sea environments with comprehensive metadata would substantially raise the bar for evaluation.

Integration with autonomous navigation systems is a natural downstream application. Enhanced imagery serving as perceptual input to AUV planning and control loops would provide a direct demonstration of end-to-end improvement in mission performance, not just image quality scores.

Finally, recent architectural innovations — vision transformers, diffusion models, and neuromorphic computing platforms — present new opportunities that have not yet been systematically explored in the underwater domain. Their ability to model long-range spatial dependencies and generate perceptually realistic outputs makes them well-suited to the enhancement problem and merits thorough investigation.

IX. CONCLUSION

This paper presented a modular, multi-stage pipeline for underwater image enhancement, addressing the interrelated degradation effects of wavelength-selective absorption, scattering-induced haze, and loss of contrast and sharpness. The pipeline proceeds from histogram-guided RGB channel compensation through Gray World white balancing, parallel HSV contrast enhancement and unsharp masking, and PCA-weighted image fusion, culminating in quantitative evaluation using MSE and PSNR.

Experiments on seven representative underwater images demonstrate that the pipeline consistently reduces MSE and improves PSNR relative to the unenhanced input, with PCA fusion outperforming simple averaging across all test images. The improvements are visually compelling: colour casts are neutralised, natural tones are restored, local contrast is heightened, and fine structural details become legible where previously they were obscured. The modular architecture makes the system readily extensible. Individual stages can be replaced with more sophisticated alternatives — a learned colour correction module, a GAN-based enhancement component, or a transformer-based fusion mechanism — without disrupting the rest of the pipeline. The system therefore serves both as a practical tool for current deployment and as a foundation for future research in this domain.

REFERENCES

[1] S. K. Nayar and S. G. Narasimhan, "Vision in bad weather," in Proc. 7th IEEE Int. Conf. on Computer Vision, vol. 2, pp. 820–827, IEEE, 1999.

[2] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 598–605, IEEE, 2000.

[3] R. T. Tan, "Visibility in bad weather from a single image," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1–8, IEEE, 2008.

[4] R. Fattal, "Single image dehazing," ACM Trans. Graph., vol. 27, no. 3, pp. 1–9, 2008.

[5] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 12, pp. 2341–2353, 2010.

[6] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," IEEE Robot. Autom. Lett., vol. 5, no. 2, pp. 3227–3234, 2020.

[7] C. Li et al., "An underwater image enhancement benchmark dataset and beyond," IEEE Trans. Image Process., vol. 29, pp. 4376–4389, 2020.