Blind quality assessment using channel-based structural, dispersion rate scores, and overall saturation and hue for underwater images.

TOLIE, H.F., REN, J., CAI, J., CHEN, R. and ZHAO, H.

2025

Blind Quality Assessment Using Channel-Based Structural, Dispersion Rate Scores, and Overall Saturation and Hue for Underwater Images

Hamidreza Farhadi Tolie, Jinchang Ren, *Senior Member, IEEE*, Jun Cai, Rongjun Chen, and Huimin Zhao

***Abstract*—In underwater subsea environments light attenuation, water turbidity, and limitations of the optical devices make the captured images suffer from poor contrast and quality, proportional degradation, low visibility, and low color richness. In recent years, various image enhancement techniques have been applied to improve the image quality, resulting in a new challenge, i.e., the quality assessment of the underwater images. In this study, we introduce an innovative and versatile blind quality assessment method for underwater images without using any references. Our approach leverages structural and contour-based metrics, combined with dispersion rate analysis, to quantify image degradation and color richness within an opponent color space. Specifically, we measure the proportional degradation by computing the edge magnitude using the directional Kirsch kernels, strengthened by image contour and saliency maps. To assess the color quality, chrominance dispersion rates and the overall saturation and hue are used to capture color distortions introduced by enhancement methods. The final quality score is obtained via a multiple linear regression model trained on extensive data sets. Experiments on three benchmark data sets have demonstrated the superior accuracy, consistency, and computational efficiency of the proposed method for both raw and enhanced underwater images.**

***Index Terms*—Blind image quality assessment (IQA), dispersion rate (DR)/color richness, image contour (IC), structural features, underwater images.**

I. Introduction

UNDERWATER visual inspection plays a crucial role in marine research and engineering, supporting condition monitoring, asset and natural resource management, and environmental sustainability. However, optical images captured underwater often suffer from degradation due to light scattering, water turbidity, and suspended particles, leading to blurriness, reduced color richness, and poor contrast [1], [2], [3]. These issues significantly impact vision-based analysis including image captioning [4] and monitoring [5], making reliable image quality assessment (IQA) essential. While various enhancement [6], [7], [8] and restoration [9], [10] techniques have been proposed, their effectiveness is often judged using subjective evaluation [11], which is costly and time-consuming. Therefore, developing objective IQA methods tailored for underwater images is crucial for ensuring more consistent, efficient, and automated quality assessment that can guarantee high quality data acquisition. This reliability is particularly important for autonomous vehicle decision-making and serves as a necessary metric for developing and benchmarking the enhancement methods.

Generally, IQA techniques, depending on the presence of the reference/original image, are usually divided into three categories: full-reference (FR), reduced-reference (RR), and blind/no-reference (NR) methods [13]. FR methods usually evaluate the quality of distorted images by measuring the similarity or distance between the features extracted from the reference and distorted images, while blind/NR methods describe the images with various features and rely on a machine learning model to learn the mapping between the descriptors and the subjective scores. In addition, RR methods predict the quality index with limited access to the reference image, based on features available in the reference images. Considering the lack of the reference in an underwater environment, some well-known FR and RR methods become infeasible in this context, these include the structural similarity (SSIM) index [14] and its variants including [15], [16], [17], RR SSIM [18], and RR entropic differencing [19].

Although existing blind IQA methods, such as NR free energy-based robust metric (NFERM) [20], NR quality assessment using statistical structural and luminance features [21], and visual quality evaluation using gradient and chromatic statistics [22] have achieved promising results on natural images, they cannot obtain consistent results with the human visual system (HVS) on underwater images, due to the following reasons: first,

Received 30 May 2024; revised 21 February 2025 and 7 March 2025; accepted 8 March 2025. This work was supported in part by the SeaSense project under Grant SPARK-2294, in part by the Net Zero Technology Centre, U.K., Guangdong Province Key Construction Discipline Scientific Research Ability Promotion Project, under Grant 2022ZDJS015 and Grant 2021ZDJS025, in part by the Special Projects in Key Fields of Ordinary Universities of Guangdong Province under Grant 2021ZDZX1087, and in part by the Guangzhou Science and Technology Plan Project under Grant 2024B03J1361 and Grant 2023B03J1327. *(Corresponding author: Jinchang Ren.)*

Associate Editor: P. Ren.

Hamidreza Farhadi Tolie is with the National Subsea Centre, School of Computing, Engineering and Technology, Robert Gordon University, AB21 0BH Aberdeen, U.K., and also with the Warwick Manufacturing Group, University of Warwick, CV4 7AL Coventry, U.K. (e-mail: h.farhadi-tolie@rgu.ac.uk).

Jinchang Ren is with the National Subsea Centre, School of Computing, Engineering, and Technology, Robert Gordon University, AB21 0BH Aberdeen, U.K., and also with the School of Computer Science, Guangdong Polytechnic Normal University, Guangzhou 610556, China (e-mail: jinchang.ren@ieee.org).

Jun Cai is with the School of Cyber Security, Guangdong Polytechnic Normal University, Guangzhou 610556, China.

Rongjun Chen and Huimin Zhao are with the School of Computer Science, Guangdong Polytechnic Normal University, Guangzhou 610556, China.

Digital Object Identifier 10.1109/JOE.2025.3553888

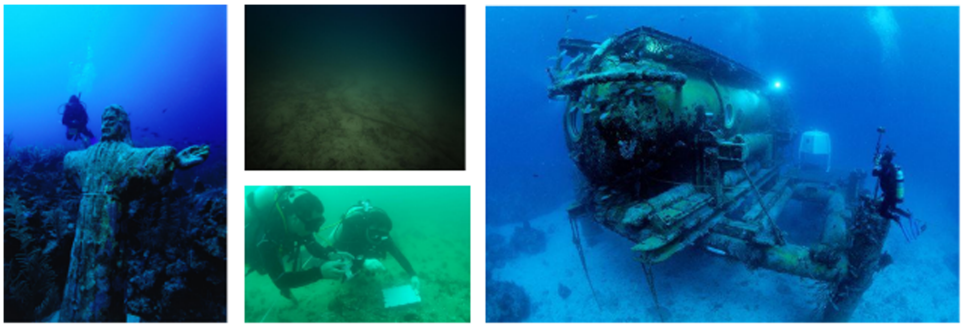


Fig. 1. Examples of underwater images taken from the UWIQA data set [12] for illustrating their low color richness (i.e., blueish and greenish appearance) and low visibility in some parts of the image (i.e., proportional degradation).

natural images are usually degraded during the transmission, storing, or compression stages, where the conditions are less complicated for modeling than imaging in the underwater scenarios. Actually, underwater images are often degraded because of the low light and poor visibility nature of the subsea environments, leading to inconsistent and uneven degradations within the image. Moreover, as shown in Fig. 1, the attenuation of lights is another artifact that leads to distorted color (i.e., bluish or greenish appearance), depending on the depth where the image is taken. This is mainly because when the depth increases the colors attenuate based on their wavelengths, where the red color dissolves much faster than the green and blue ones, respectively [23]. Furthermore, existing blind IQA methods are usually designed to follow the natural scene statistics (NSS) model [24], yet it performs differently in subsea environments [25]. Based on the well-known NSS model [24], a variety of IQA variations have been designed, including the blind/referenceless image spatial quality evaluator (BRISQUE) [26], natural image quality evaluator (NIQE) [27], and the integrated local NIQE (IL-NIQE) [28], however, their performance drops inevitably on underwater images. Even the performance of the deep learning methods, such as the discriminable image pairs inferred quality (dipIQ) index [29], with the RankNet model to rank images, is not as expected on underwater images. Therefore, quality assessment of underwater images becomes a particular challenge, especially for quantitative metrics-based measurements, before carrying out any further application-driven tasks, such as detection, classification, and condition monitoring [30]. In this study, we have proposed a new blind underwater image quality evaluator, which effectively measures the quality of both raw and enhanced images. The major contributions can be highlighted as follows.

1) To assess image color richness and object visibility, we convert images from red, green and blue (RGB) to the CIELAB color space, enabling independent analysis of brightness and chrominance degradations. Leveraging the HVS's sensitivity to color variations, we introduce a chrominance dispersion rate (DR) that captures brightness-to-color richness ratios, enhancing IQA performance.
2) We propose a proportional degradation metric by computing edge magnitudes using directional Kirsch kernels. These edge maps are further refined with contour and saliency information to highlight object boundaries, ensuring a more perceptually relevant assessment of structural integrity.
3) To quantify enhancement-induced distortions, we analyze saturation and hue variations, providing a measure of color purity and distinguishability postenhancement. Finally, we integrate all extracted features using a multiple linear regression model, demonstrating high accuracy, robustness, and efficiency across diverse underwater IQA data sets.

The rest of this article is organized as follows. In Section II, we review the related works and examine commonly employed attributes for assessing underwater image quality. Section III presents the proposed quality evaluation metric. Section IV provides analyses over components of the proposed metric and compares the obtained experimental results. Finally, Section V concludes this article.

II. Related Work

A. Underwater Image Quality Attributes: An Overview

In subjective quality ratings, the low-level (i.e., physical and detailed) information, such as the brightness, sharpness, and contrast are combined with the high-level (i.e., general and abstract) information including the naturalness, clarity, and objects' 3-D depth measurement to perceive a quality score [31]. Accordingly, objective metric needs to consider these factors for consistent and accurate evaluations. To this end, the aforementioned quality-related features are widely explored with respect to the unique characteristics of the corresponding images (e.g., natural, screen content, underwater, etc.) [32], [33], [34], [35], where different fusion strategies have been proposed to integrate those features. For quality evaluation of underwater images, sharpness, contrast, and colorfulness are the three highly explored features, which represent the discrimination of shapes or objects within an image, the clarity of the image, and the color richness, respectively.

Yang and Sowmya [34] used the linear combination of the asymmetric alpha-trimmed first-order and second-order statistical values [36] in an opponent color space, the enhancement measure estimation (EME) [37] over the grayscale edge map,

and the logAMEE measure [38], to quantify the colorfulness, sharpness, and contrast, respectively. In [39], colorfulness is quantified similarly to the [34] with minor changes on the opponent color space, and the contrast is measured as the sum of the root mean square (rms) contrast value of the edge blocks over the red intensity of all pixels. In [12], the variance of chroma, refined rms contrast value of the edge blocks similar to [39], and EME same as in [37] are considered as colorfulness, contrast, and sharpness measures, respectively.

As discussed earlier, the values of the red color, in underwater especially subsea images, decrease with the increasing depth [35]. Hence, utilizing the red color in measuring the colorfulness, contrast, and other chrominance-related attributes is not working effectively as expected and limits the performance of quality evaluation. Also, the sharpness degradations are among the major reasons for performance drop in the vision-based analysis of underwater images. However, this is not fully investigated in comparison to the chrominance degradations. Therefore, to overcome these challenges we consider the brightness and color richnesses of the image rather than its overall colorfulness to investigate the effect of the edge and contour information in measuring the sharpness.

B. Underwater Image Quality Metrics (UIQM)

Regarding blind underwater image quality assessment (BUIQA), several handcraft feature-based methods have been proposed, such as the UIQM [34], underwater color image quality evaluation (UCIQE) [35], colorfulness, contrast, and fog density (CCF) [39], frequency domain underwater image quality metric (FDUM) [12], and underwater image fidelity (UIF) [25]. Moreover, recently a deep neural network (DNN)-based model, namely, underwater ranker (URanker) [40], is also introduced for BUIQA to mitigate the handcraft feature-based methods challenges.

Inspired by the HVS's properties, which perceives the chrominance and luminance attributes (e.g., edges or object boundaries) in opponent color space, UIQM, proposed by Yang and Sowmya [34], linearly combines the image's colorfulness, sharpness, and contrast in a weighted scheme to assess its corresponding quality index. UCIQE [35], similar to the UIQM [34], linearly combines the chroma, saturation, and contrast to measure the degree of the degradation as blurring, contrast, and nonuniform color cast of the image. Following the studies [34], [35], CCF [39] proposes to take the fog density into account and combines it with other scores (i.e., colorfulness and contrast). More recently, Yang et al. [12] considered both the low- and high-frequency information of the underwater images to measure the proportional degradation. To this end, they combined the colorfulness scores from both the spatial and frequency domains with the sharpness and dark channel prior weighted contrast measure. UIF [25] integrates the naturalness, sharpness, and structure features to compute a feature domain representation before applying a saliency-guided pooling strategy to determine the final quality score. The recently proposed URanker method [40], utilizes a ranking-based model for BUIQA with a conv-attentional image Transformer incorporated with histogram prior to embed the color distribution of the image.

Despite notable efforts, there remains a need for improvement in the existing methods, especially to improve their robustness, reliability, and adaptability across a broader range of underwater imaging conditions and applications, as most existing methods aim to characterize images from a general perspective. For instance, UIQM [34] interprets underwater images using the overall sharpness, colorfulness, and contrast. However, given the challenges of deviated colors in underwater images, relying solely on overall image information may not be optimal. For quality assessment, effective evaluation of both the raw and enhanced images is crucial, where separately determination of the color richness, distribution, and degradation for each color channel is needed. This is because for the raw images, the degradation is channel-dependent as the color attenuation rate for the primary colors, i.e., red, green, and blue, is not uniform. Therefore, our study proposes a novel approach, which integrates the computed structural degradation and the DR of each color channel in the CIELAB color space with the overall saturation and hue of the image. Specifically, the proposed method, inspired by [33] [41], considers mainly the light attenuations of underwater images to introduce a more precise and robust quality metric adapted to the unique characteristics of underwater imagery.

III. Proposed Method

Proportional quality degradation and low color richness are the two main deteriorations that are usually seen in underwater images. Depending on the object–camera distance, water turbidity, and lighting condition, the quality of underwater images degrades proportionally [42], where different regions of the image are degraded with various ratios, resulting in visibility and clarity changes of objects in the image. As a result of proportional degradation, various attributes of the image, such as edges and object boundaries, can be affected with their details blurred. Moreover, due to light attenuation and scattering, underwater images have low color richness, hence they are usually seen in bluish or greenish appearances.

The artifacts mentioned above are commonly observed in raw underwater images. Nevertheless, significant efforts have been made in recent years to mitigate these issues and improve the overall quality of captured images. While enhancement methods, especially in recent years with the emergence of reinforcement learning [43], [44], [45], [46] and foundation models [47], have demonstrated remarkable outcomes, they often face challenges, such as over-enhancement, which can lead to images with excessive saturation or desaturation. Therefore, there is a critical need to introduce a comprehensive approach capable of assessing the quality of both raw and enhanced images.

In this section, to address the abovementioned issues and improve the performance of the existing BUIQA methods, especially for raw underwater images taken in data acquisition for condition monitoring of assets and natural resources, we present a blind/no-reference metric for accurate, efficient, and consistent quality assessment of underwater images. The proposed metric evaluates the quality of underwater images in the CIELAB color

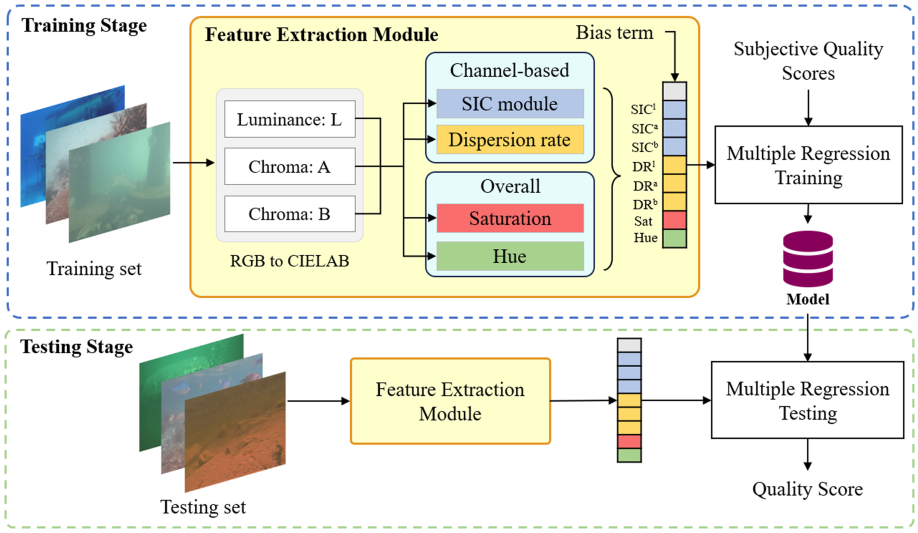


Fig. 2. Framework of the proposed method for BUIQA.

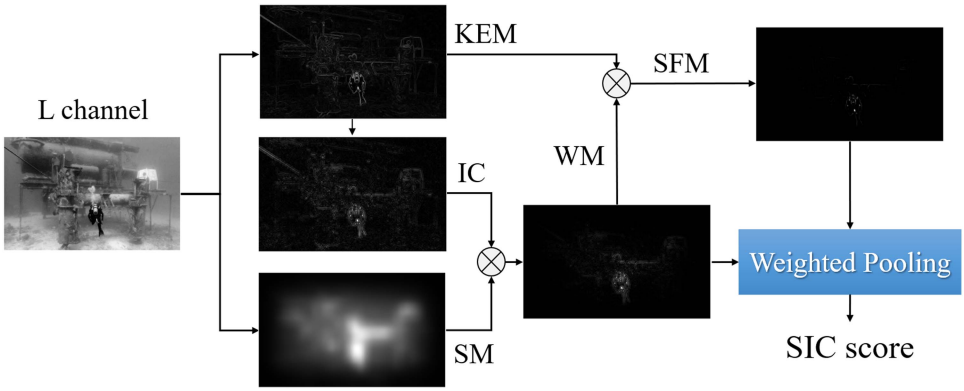


Fig. 3. Illustration of the SIC feature-based scoring module for the L channel of an input image. The KEM is obtained and weighed by the fused WM (i.e., IC and SM/GBVS) to form the SFM. The SIC score is computed by applying a weighted pooling on SFM and WM-based feature maps.

space to take both the luminance and chrominance information of the image into account as well as giving a perceptual uniformity to the image meaning that the differences in the color values correspond more closely to the perceived differences in color by human vision [48], [49]. In general, our method quantifies the proportional degradation using directional Kirsch edge kernels and measures the brightness and color richness rate (i.e., DR) of each channel separately. Moreover, it also computes the overall saturation and hue of the image for accurate quality evaluation of the enhanced images. As illustrated in Fig. 2, after color space conversion, the channel-wise structural and image contour (SIC)-based scores are derived followed by the DRs and the overall saturation and hue from each channel. These are then combined to train a multiple linear regression model to determine the associated quality score of the image.

The SIC feature-based scoring module is shown in Fig. 3. As shown, in this module, we first compute the image's gradient map by applying the Kirsch filters [50] in eight directions. Also, besides the edge features, since image contours (ICs) and salience areas are significantly compatible with HVS [33], [51], we incorporate them to amplify the boundaries of the objects, which are more in the center of HVS's attention [52]. To this end, we formed the image's structural feature map (SFM) by multiplying the image's edge map by its weight map (WM)—computed by fusion of the IC and saliency map (SM). Finally, we have utilized a weighted pooling strategy, in which the WM is used as a WM of SFM to obtain the final SIC feature-based score.

From Fig. 3, it can be seen that after applying the contour and SM on the edge map, the object closer to the camera (i.e., diver) is still visible. This shows the effectiveness of our proposed method in measurings the visibility of the objects, based on the object–camera distance, and addressing the proportional degradation. The following sections explain the directional edge feature extraction, edge strength weighting, pooling strategy, and quality regression stage used to determine the overall quality score.

A. *Directional Edge Features*

Various studies [53], [54] have employed a variety of edge features to describe images due to the significant sensitivity of

the HVS to edges in forming the visual perception of the image. In underwater images, a wide range of objects could be seen in fixed (e.g., pipelines or energy infrastructures) or dynamic positions (e.g., divers, underwater species, or trash). Consequently, the edges occur in different directions. Thus, we proposed to use the nonlinear Kirsch operator to capture the edge strength in eight predetermined directions. The notable advantages of the Kirsch edge detector are fast computing operations and the ability to produce more edge pixels [55], which is necessary for underwater images to capture proportional degradation. In this study, as mentioned earlier, the Kirsch operator [50] is used in eight directions as follows to compute the gradient maps of each channel (i.e., L, A, and B):

$$G_i^j(x,y) = |I_j(x,y) \otimes K_i|, \quad i \in \{1,2,3,\ldots,8\} \tag{1}$$

where K_i is the ith Kirsch operator, $\otimes$ is the convolution operation, and $G_i^j(x,y)$ and $I_j(x,y)$ are the ith gradient map and intensity map of the jth channel ($j \in \{1,2,3\}$ representing the L, A, and B channels) for the pixel at position (x,y), respectively

$$K_1 = \begin{bmatrix} +5 & +5 & +5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}. \tag{2}$$

Equation (2) shows the first Kirsch operator in the North compass direction. The rest seven operators corresponding to the North West, West, South West, South, South East, East, and North East can be obtained by rotating the K_1 by 45°, progressively. In this study, to perceive the visibility of objects, we compute the maximum gradient strength (i.e., edge map) from the image's jth channel, which has been proven to have the maximum correlation with subjective scores [25]. Image's maximum gradient strength, namely, Kirsch edge map (KEM) for the jth channel is computed as follows:

$$\text{KEM}^j(x,y) = \max\{G_1^j(x,y), G_2^j(x,y), \ldots, G_8^j(x,y)\}. \tag{3}$$

B. Edge Strength Weighting and Pooling

As discussed earlier, ICs are also highly considered in HVS's image perception. Generally, ICs are the object's boundaries, which can be computed from the edge map of the image. As underwater images suffer from proportional degradation, degrading the visibility of the objects and edges accordingly, we have used the ICs to weigh the KEM. In this study, similar to [33], [52], we employed a difference of Gaussians (DoG) filter on a larger scale to smooth the edges (i.e., object boundaries/contours) of the image and utilize them to highlight the object boundaries. IC of the jth channel of the image is defined from its KEM as follows:

$$\text{IC}^j_{\sigma_1,\sigma_2}(x,y) = |\text{DoG}_{\sigma_1,\sigma_2}(x,y) \otimes \text{KEM}^j(x,y)| \tag{4}$$

$$\text{DoG}_{\sigma_1,\sigma_2}(x,y) = G(x,y;\sigma_1) - G(x,y;\sigma_2) \tag{5}$$

where $\text{DoG}_{\sigma_1,\sigma_2}(x,y)$ is the difference of two Gaussian functions with close large-scale values of σ_1, σ_2 (i.e., standard deviations), and $G(x,y;\sigma)$ is the Gaussian kernel function as a 2-D normal distribution of $N(0,\sigma)$ given by

$$G(x,y;\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}. \tag{6}$$

After obtaining the IC maps, considering that some portions of the image may be blurred due to low-light conditions, water turbidity, etc., the SM of the image is also utilized to further strengthen the highly important areas in ICs. In [56] and [57], the impact of saliency-guided methods has been demonstrated in real-world IQA methods for improved performance. Our experiments also show the effectiveness of the SM in assessing the quality of the underwater images. As shown in the sample images in Fig. 3, IC is well representing the objects' boundaries, while the SM reflects the regional importance. In other words, in SM, objects closer to the camera are more visible than the distant ones. Here, we have utilized the graph-based visual saliency (GBVS) [58] to compute the SM as it highlights the informative locations of the image by considering human eye movement indices like fixation. The extracted KEM, IC, and SM from the jth channel are fused by multiplication to form the SFM as follows:

$$\text{SFM}^j(x,y) = [\text{KEM}^j(x,y)]^{\alpha^j} \cdot [\text{IC}^j(x,y)]^{\beta^j} \cdot [\text{GBVS}^j(x,y)]^{\gamma^j} \tag{7}$$

where $\alpha = (\alpha^1, \alpha^2, \alpha^3)$, $\beta = (\beta^1, \beta^2, \beta^3)$, and $\gamma = (\gamma^1, \gamma^2, \gamma^3)$ are three vectors with positive parameters that have been employed to adjust the relative importance of each feature map corresponding to each channel. These three parameters are empirically set to $\alpha = (1,1,1)$, $\beta = (0.8, 0.2, 0.8)$, and $\gamma = (1, 0.3, 0.7)$ to yield best performance. Moreover, $\text{GBVS}^j(x,y)$ is the SM of the jth color channel.

In the HVS's perspective, the visual gaze point in an image is notably different from its neighbor points [59]. In fact, this difference is a spatial decay from the gaze point caused by the limited visual resolution of the human retina [60]. In other words, HVS does not perceive the image pixels equally [52], thus we employed a weighted pooling strategy to highlight the perceptual importance of all pixels in the image for more effective BUIQA. To this end, the extracted edge features (i.e., KEMs) are weighed and pooled to evaluate the SIC feature-based score of each channel as follows:

$$\text{SIC\_score}^j = \frac{\sum_{(x,y)} \text{SFM}^j(x,y)}{\sum_{(x,y)} \text{WM}^j(x,y)} \tag{8}$$

$$\text{WM}^j(x,y) = [\text{IC}^j(x,y)]^{\beta} \cdot [\text{GBVS}^j(x,y)]^{\gamma} \tag{9}$$

where $\text{WM}^j(x,y)$ is the jth WM computed by multiplying the jth IC and the GBVS SM. As seen in Fig. 4, the obtained WMs differ in channels and they can actually highlight the regional importance based on the spatial location of the main target in the image. The combination of the IC and GBVS SM for the weighted pooling strategy significantly contributes to the overall quality score. Note that, the WM_A computed from the A channel is more pleasant visually as we considered a lower weight, i.e., $\beta^2 = 0.2$, for its IC map. Since the intensity values of both the input image and its computed KEM lie within the range of $[0,1]$, a lower coefficient amplifies these values, leading to brighter regions in the final output. This weighting is applied

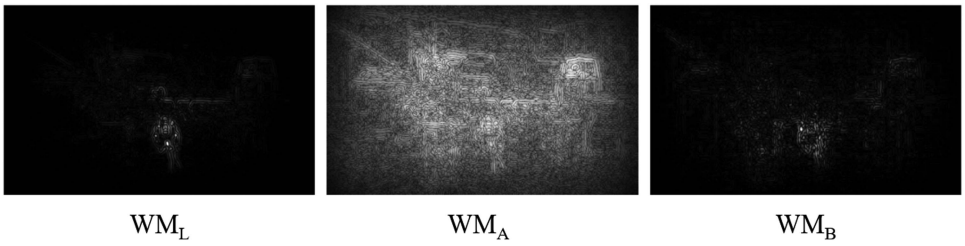


Fig. 4. WM obtained from the L, A, and B channels by multiplying the IC and saliency/GBVS maps.

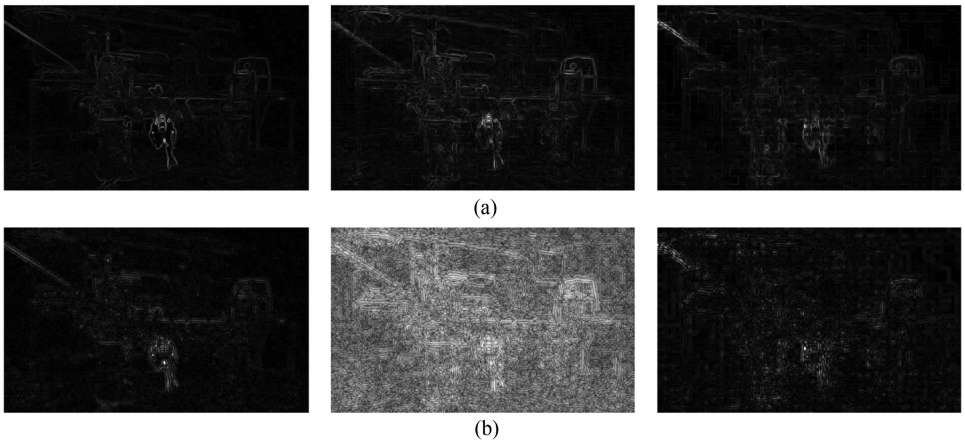


Fig. 5. (a) KEM and (b) IC feature maps of a sample image. From left to right: obtained feature maps corresponding to the L, A, and B channels.

because channel A represents green-to-red color information, and due to their faster dissolution in water, they lose more crucial information compared to other colors, necessitating a higher weighting for accurate representation.

C. Dispersion Rate

As shown in Fig. 5, the extracted edge and contour features from the luminance and chrominance channels can effectively represent the image's content and capture the color information loss. However, the amount of the lost information in luminance and chrominance channels is unequal and thus affects the extracted features. Therefore, to have an accurate method for assessing the quality of the both raw and enhanced underwater images, we proposed to extract the brightness and color richness level of the image, namely DR, alongside the determined quality scores (i.e., SIC_score^j, j $= 1, 2, 3$).

To assess an image's quality and measure its color richness, many researchers measured the image's global colorfulness based on the variance to the mean ratio of the chrominance channels in an opponent color space [61], [62]. In this study, inspired by [33], [62], and [63], we quantified the DR of each channel of the underwater image separately in the CIELAB color space. The DR specifically measures the variability of colors relative to their mean values. For example, when computing the DR for each channel in the CIELAB color space, it assesses the variability or richness of each color attribute, such as luminance, green to red, and blue to yellow. Analyzing the DR of each channel aids in a more comprehensive evaluation of underwater image quality in terms of color distribution. The DR is computed as follows:

$$\mathrm{DR}^j = \log\left(\frac{\sigma_{I^j}^2}{|\mu_{I^j}|^{0.2}}\right) = 2\log\left(\frac{\sigma_{I^j}}{|\mu_{I^j}|^{0.1}}\right) \tag{10}$$

where σ_{I^j} and μ_{I^j} denote the standard deviation and mean values of the jth channel of the image, respectively.

However, compared with the [62] and [63], instead of measuring the image's overall colorfulness, we assessed the DR of each channel. Considering the characteristic of the CIELAB color space, which gives us information about the red and blue color differences, our proposed measure is capable of quantifying the color distribution with regards to underwater images sensitivity (i.e., light attenuation leading to the disappearance of red, green, and blue colors, respectively). Equation (10) measures the DR of intensity values corresponding to each channel with respect to the mean intensity value of that channel. To demonstrate the effectiveness of the proposed DR metric, two sample images with greenish and bluish appearances are compared in Fig. 6. For the greenish image to the left, as the A channel corresponds to the green-to-red range, DR_A is higher than DR_B, as expected. In contrast, for the bluish image to the right, the DR_A channel is lower than DR_B due to the bluish appearance of the image. In addition, since the bluish image contains brighter pixels at

Fig. 6. DR analysis of two sample images with greenish and bluish appearances, showing the relationship between DR_A and DR_B.

the bottom, it results in a higher DR_L compared to the greenish image.

D. Overall Saturation and Hue

In addition to the previously introduced quality scores and DR, it is important to characterize color degradation in images. While the introduced DR effectively represents color variation within an image, it alone is insufficient for assessing the quality of the enhanced images as well as evaluating the efficacy of underwater enhancement methods. These enhancement techniques aim to restore the color information, but they may lead to degradation or over-enhancement of the original image.

Drawing inspiration from [64], we have proposed a comprehensive approach to quantitatively measure the purity and dominance of colors, by computing the overall saturation and hue of the image. While color spaces like the HSV directly provide a saturation map, they are not perceptually uniform. This means that equal steps in hue, saturation, or value do not correspond to equal perceived changes in color [65].

This additional analysis enhances our ability to evaluate the impact of enhancement methods on color fidelity, leading to a more effective quality assessment of underwater images. The image's saturation map is then formulated to measure how balanced or unbalanced the color components are. If one component is significantly lower than the others, the saturation value decreases, indicating less color richness

$$\mathrm{Saturation}_{\mathrm{map}} = 1 - \frac{3 \times \min(L, A, B)}{L + A + B} \tag{11}$$

where the minimum value across the channels can help us to identify the colors with the least presence in a particular pixel. Dividing this value by the sum of the three channels for each pixel, gives us a measure of how dominant or significant a particular channel is in relation to the overall color composition of the pixel.

In this study, the average value of the $\mathrm{Saturation}_{\mathrm{map}}$ is used to denote the overall saturation of the image. Moreover, we compute the hue map of the image as follows and consider its average value as the overall hue

$$\mathrm{Hue}_{\mathrm{map}} = \arctan\left(\frac{\phi}{\zeta}\right) = \arctan\left(\frac{\sqrt{3} \times (L - A)}{L + A - 2B}\right) \tag{12}$$

where $\phi = (L - A)/\sqrt{2}$ and $\zeta = (L + A - 2B)/\sqrt{6}$. Unlike the common approaches to calculate the arctangent of the ratio B/A, we used the ϕ and ζ, by referring to [66], to compute the hue with respect to the green to red color information, which dissolve faster than the blue color in underwater conditions. By subtracting the A color channel from the luminance L, we can assess the information contributed by green and red colors in the content. Similarly, subtracting the B channel from the addition of L and A assists in measuring the impact of blue and yellow colors on the representation of underwater content. Also, $\sqrt{2}$ and $\sqrt{6}$ are empirically set to normalize the obtained chrominance maps.

E. Quality Regression

For effective and accurate quality evaluation of underwater images, we proposed to linearly combine the SIC scores and DRs obtained from each channel. To this end, we employed multiple linear regression with a bias term to estimate the optimal coefficients for the linear combination of the extracted SIC and dispersion scores, along with overall saturation and hue. Given that our model had only eight features, multiple linear regression was utilized for its simplicity, interpretability, and effectiveness in capturing linear relationships, while avoiding the risk of overfitting that more complex models like support vector regression, though the latter may be explored in the future. Note that the final feature vector for each image is formed by adding the bias term and normalizing the computed scores similar to [67] through the square root operation, thereby amplifying the influence of each individual score. This approach contributes to a more robust and subtle assessment of underwater image quality

$$\mathrm{feat}_i = [b_0, \mathrm{SIC\_scores}, \mathrm{DR}s, \mathrm{Saturation}, \mathrm{Hue}] \tag{13}$$

where b_0 is the bias term set to 1, SIC_scores is $[\sqrt{\mathrm{SIC\_score}^1}, \sqrt{\mathrm{SIC\_score}^2}, \sqrt{\mathrm{SIC\_score}^3}]$, and DRs is set to $[\sqrt{\mathrm{DR}^1}, \sqrt{\mathrm{DR}^2}, \sqrt{\mathrm{DR}^3}]$.

In this study, we have randomly divided each data set to the training and testing subsets for 1000 times. Each time 80% of the data are used to determine the coefficients for the linear regression problem and 20% of the data are used to test the performance of the proposed method.

IV. Experimental Results

The performance of the proposed method is evaluated on three publicly available data sets, namely, underwater image quality assessment (UWIQA) [12], underwater image data set 2021 (UID2021) [68], and the state-of-the-art (SOTA) UWIQA data set [69]. In total four state-of-the-art blind natural image quality assessment (BNIQA), five BUIQA methods, and one DNN method are compared as well as the distortion-based directional edge and gradient similarity maps (DDEGSM) [33] method, ispired us toward directional edges, using three commonly used evaluation metrics, i.e. Pearson linear correlation coefficient (PLCC), Spearman's rand-order correlation coefficient (SRCC), and root-mean-squared error (RMSE). Detailed results and relevant ablation studies are given as follows.

A. Description of the Data Sets

The UWIQA data set [12] has 890 raw underwater images in different resolutions, which are selected from the [70]. Images in the UWIQA data set contain diverse scenes and various object categories such as coral, marine life (e.g., fish, turtles, sharks), divers, infrastructures, etc. The subjective quality scores of these images are determined by 21 observers with different backgrounds, age groups, and genders. The mean opinion scores (MOSs) are then computed by averaging the observers' ratings and mapping to 10 quality levels varying from 0.1 to 1.

In addition to the UWIQA data set which includes only raw underwater images, to evaluate the performance of the proposed method on enhanced underwater images, we conducted experiments also on the recently proposed UID2021 data set [68]. UID2021 has six groups of images i.e., bluish (B), blue–green (BG), greenish (G), hazy (H), low-light (LL), and turbid (T) ones, and there are ten reference images for each group. In UID2021, 15 underwater image enhancement and restoration algorithms have been used to generate the enhanced images for the reference images, leading to 150 enhanced and 10 raw images being contained in each group, or 960 images in total.

Furthermore, we conduct experiments on another recently proposed SOTA [69] data set, which contains 800 raw images and their enhanced versions, where the latter are generated by nine different algorithms, leading to a total of 8000 images with subjective MOSs. The nine enhancement algorithms include attenuation coefficient prior attention block (ACPAB) method [71], dynamic histogram equalization or histogram partition (HP) [72], image blurriness and light absorption (IBLA) [73], fusion-12 [74], fusion-18 [75], two-step approach enhancement [76], water-Net [70], fast underwater image enhancement-generative adversarial network (FUNIE-GAN) [77], and a commercial method (Dive+[1]).

B. Evaluation Metrics and Parameter Setting

As suggested in [78], we utilized three widely used metrics including the PLCC, SRCC, and RMSE to evaluate the objective quality assessment's accuracy, monotonicity, and consistency, respectively [33]. PLCC measures the linear correlation between the objective and ground truth (i.e., MOS) quality scores, while SRCC assesses the rank-order correlation. Also, the amount of deviation between the objective and MOS values is computed by the RMSE. Generally, higher PLCC and SRCC values (close to 1), and smaller RMSE values indicate better performance. These metrics are computed as follows:

$$\text{PLCC} = \frac{\sum_{i=1}^{n}(Z_i - \bar{Z})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{n}(Z_i - \bar{Z})^2 (O_i - \bar{O})^2}} \tag{14}$$

$$\text{SRCC} = 1 - \frac{1}{n(n^2-1)} 6 \sum_{i=1}^{n} d_i^2 \tag{15}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - \bar{O})^2} \tag{16}$$

where n represents the number of input images, O_i and Z_i are the objective and subjective scores for the ith image in the data set, and $\bar{O}$ and $\bar{Z}$ are the mean values of the objective and subjective scores, respectively, and d_i denotes the difference between the ith image's ranks in subjective and objective assessments.

In this study, we followed the same practice in [33] and [78] to first remove the nonlinearity of the objective scores by applying a nonlinear logistic regression with five parameters and then compute the PLCC, SRCC, and RMSE metrics. The mapped predicted score is computed as follows:

$$Q_i = \kappa_1 \left\{ \frac{1}{2} - \frac{1}{1 + \exp\left[\kappa_2 (p_i - \kappa_3)\right]} \right\} + \kappa_4 s_i + \kappa_5 \tag{17}$$

where p_i is the perceived quality score of the ith underwater image computed by the proposed method, Q_i is its corresponding mapped objective score, and κ_{1-5} are the model parameters in the curve fitting process to minimize the sum of squared differences between the objective and ground truth quality scores.

The proposed method has the following parameters to be set, two scales (i.e., σ_1, σ_2) for IC and three coefficients for feature map fusion of each channel (i.e., α, β, and γ). To obtain the best values for each parameter, we have conducted experiments on a subset of the UWIQA data set. Similar to [33] and [52], we have selected 200 random images out of the 890 available and executed the proposed method over them to select the parameters that led to higher SRCC values. Each time we adjust only one selected parameter with all others fixed. To compute the IC map we set the DoG filter scales to 2.1 and 2.0 (i.e., $\sigma_1 = 2.1$ and $\sigma_2 = 2.0$), and the DoG filter size as 11×11, also we have $\alpha = (1, 1, 1)$, $\beta = (0.8, 0.2, 0.8)$, and $\gamma = (1, 0.3, 0.7)$ for feature map fusion in (7).

C. Performance Comparison

In Table I, the performance of the proposed method is compared with four BNIQA methods (including BRISQUE [26], NIQE [27], IL-NIQE [28], and dipIQ [29]) and five state-of-the-art BUIQA methods (including UIQM [34], UCIQE [35], CCF [39], FDUM [12], and URanker [40]) in terms of three commonly-used metrics of PLCC, SRCC, and RMSE over the UWIQA data set. Note that since the UIF and DDEGSM need

[1]https://itunes.apple.com/us/app/dive-video-color-correction/id1251506403?mt=8

TABLE I
COMPARISON OF THE PLCC, SRCC, AND RMSE FROM FOUR BNIQA, FIVE BUIQA, AND OUR PROPOSED METHODS ON THE UWIQA DATA SET

	BNIQA				BUIQA				DNN	
Criteria	BRISQUE	NIQE	IL-NIQE	dipIQ	UIQM	UCIQE	CCF	FDUM	URanker	**Proposed**
PLCC ↑	0.3602	0.3797	0.4573	0.0992	0.2689	0.5752	0.4493	*0.6652*	0.5140	**0.7294**
SRCC ↑	0.3463	0.3603	0.4718	0.1134	0.2964	0.5854	0.4425	*0.6751*	0.5289	**0.7376**
RMSE ↓	0.1421	0.1409	0.1355	0.1516	0.1467	0.1246	0.1361	*0.1138*	0.1307	**0.1040**

The best and second-best result are highlighted in bold and italic, respectively.

TABLE II
COMPARISON OF THE PLCC, SRCC, AND RMSE FROM FOUR BNIQA, SIX BUIQA, DDEGSM, AND OUR PROPOSED METHODS ON SIX SUBSETS OF THE UID2021 DATA SET

		BNIQA				BUIQA					DNN		
Subset	Criteria	BRISQUE	NIQE	IL-NIQE	dipIQ	UIQM	UCIQE	CCF	FDUM	UIF	URanker	DDEGSM	Proposed
B	PLCC ↑	0.2241	0.2200	0.4799	0.1618	0.5781	0.6417	0.2606	**0.715**	0.2570	0.6643	0.2627	*0.6974*
	SRCC ↑	0.2192	0.2250	0.4521	0.2113	0.5430	0.6182	0.4413	**0.7023**	0.2699	0.6218	0.2049	*0.6686*
	RMSE ↓	2.0489	2.0508	1.8444	2.0746	1.6876	1.6124	2.0297	**1.4364**	2.0317	1.5715	2.0285	*1.4725*
BG	PLCC ↑	0.3094	0.1332	0.5775	0.0827	*0.5891*	0.5778	0.3928	0.5839	0.3421	*0.5891*	0.3526	**0.5911**
	SRCC ↑	0.3116	-0.0465	*0.5854*	0.0140	0.5700	0.5437	0.3018	*0.5705*	0.3319	**0.5984**	0.2869	0.5590
	RMSE ↓	2.0955	2.1839	1.7989	2.1960	1.7948	1.7985	2.0265	1.8031	2.0707	*1.7806*	2.0621	**1.7263**
G	PLCC ↑	0.2713	0.0469	0.4817	0.2751	0.6328	0.6902	0.3376	0.6639	0.2936	*0.7454*	0.3746	**0.7736**
	SRCC ↑	0.2772	0.1292	0.4775	0.2945	0.5789	0.6666	0.4159	0.6584	0.2576	*0.7295*	0.3020	**0.7512**
	RMSE ↓	2.1681	2.2502	1.9740	2.1658	1.7324	1.6300	2.1204	1.6732	2.1533	*1.5016*	2.0886	**1.3831**
H	PLCC ↑	0.1970	0.3148	0.2877	0.3165	0.5773	0.6686	0.4215	0.5814	0.4626	*0.6843*	0.4119	**0.7241**
	SRCC ↑	0.2063	0.1900	0.2621	0.3232	0.5378	*0.6239*	0.4946	0.5549	0.4156	0.6221	0.3522	**0.6625**
	RMSE ↓	2.0880	2.0196	2.0378	2.0184	1.7373	1.5823	1.9295	1.7312	1.8863	*1.5515*	1.9389	**1.4270**
LL	PLCC ↑	0.2913	0.4617	0.2073	0.2553	0.5293	**0.6726**	0.3686	0.6276	0.2958	0.4143	0.3894	*0.6618*
	SRCC ↑	0.1870	0.4378	0.1764	0.1900	0.5497	**0.6640**	0.4268	0.6197	0.2443	0.3850	0.3712	*0.6253*
	RMSE ↓	1.8919	1.7542	1.9347	1.9121	1.6779	*1.4634*	1.8384	1.5397	1.8892	1.7999	1.8216	**1.4453**
T	PLCC ↑	0.2576	0.0882	0.6083	0.1657	**0.7164**	0.5834	0.6347	0.6875	0.3168	0.6401	0.3393	*0.7118*
	SRCC ↑	0.1822	0.0492	0.5918	0.1612	**0.7157**	0.5599	0.6259	0.6850	0.2670	0.6386	0.2945	*0.6877*
	RMSE ↓	2.0709	2.1348	1.7011	2.1136	*1.4952*	1.7406	1.6562	1.5564	2.0328	1.6465	2.0160	**1.4728**

The best and second-best result are highlighted in bold and italic, respectively.

both the raw and enhanced images, they could not been evaluated because this data set only contains the raw images. The source codes of these methods are obtained from their author's websites. The best and second-best values in terms of the three metrics are highlighted in bold and italic, respectively. From Table I, it is clear that the proposed method outperforms all compared methods with the highest PLCC and SRCC and the lowest RMSE values. More precisely, our proposed method surpassed the second-best (i.e., FDUM) and the third-best (i.e., UCIQE) methods by 9.65% and 26.80% in terms of PLCC and 9.25% and 26% in terms of SRCC, respectively.

Moreover, Table II reports the performance of all compared methods on the enhanced underwater images on the UID2021 [68] data set. To assess the performance of our method, we divided the entire data set into training and testing subsets, opting for a unified approach instead of conducting separate evaluations for each individual subset. As seen, there is not any dominant method over all six subsets of the UID2021 data set. Our proposed method demonstrates superior performance by achieving the best results in three subsets, namely, blue–green, green, and hazy, while also obtaining the second-best performance in three other subsets, namely, blue, low-light, and turbid, with a competative performance to the best method. Among the compared methods, FDUM yields the best result on the Bluish image subset. Moreover, on the low-light and turbid images, UCIQE and UIQM outperform other methods. It is worth noting that in these two subsets, our approach achieves a performance level that is significantly higher than other compared methods except UIQM and UCIQE, with the lowest RMSE value.

Furthermore, the performance of the compared methods and ours are evaluated on the SOTA data set with the results reported in Tables III–V, where the best and second-best methods are highlighted in bold and italic, respectively. To evaluate the performance of the proposed method on this data set, we have separately evaluated the performance on raw and enhanced subsets by divinding them into training and testing subsets for 1000 times and reporting the median result on the testing subset. Moreover, to mitigate potential bias in the results for

TABLE III
COMPARISON OF THE PLCC, SRCC, AND RMSE FROM FOUR BNIQA, SIX BUIQA, DDEGSM, AND OUR PROPOSED METHODS ON THE RAW, AND ENHANCED IMAGES USING ACPAB, HP, AND IBLA METHODS FROM THE SOTA DATA SET

	Raw			ACPAB			HP			IBLA		
Methods	PLCC ↑	SRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	RMSE ↓
BRISQUE	0.1862	0.1581	6.8184	0.1665	0.1625	14.0042	0.1927	0.1363	13.1389	0.0519	0.0500	8.4378
NIQE	0.1896	0.1734	6.8139	0.1636	0.1607	14.0109	0.2461	0.2828	12.9781	0.0589	0.0198	8.4345
IL-NIQE	0.3847	0.3698	6.4057	0.3398	0.3174	13.3574	0.5215	0.5317	11.4251	0.2657	*0.2684*	8.1456
dipIQ	0.1628	0.1725	6.8472	0.0128	0.0082	14.2012	0.2029	0.1921	13.1115	0.0848	0.0916	8.4188
UIQM	0.3996	0.3694	6.3616	0.2362	0.2200	13.8006	0.2080	0.1854	13.0969	0.1084	-0.0163	8.3994
UCIQE	*0.5545*	*0.5603*	*5.7751*	0.1463	0.1469	14.0495	0.4174	0.367	12.1676	0.1650	0.1488	8.3334
CCF	0.2964	0.2392	6.6279	0.4146	0.3938	12.9239	0.2447	0.2345	12.9827	0.1754	0.1582	8.3182
FDUM	0.5453	0.5122	5.8173	0.3375	0.3319	13.3693	0.2875	0.2981	12.8246	0.0818	0.0688	8.4206
UIF	-	-	-	0.5199	0.4812	12.1320	*0.7071*	0.7375	*9.4676*	0.3475	0.2493	7.9227
URanker	0.6052	0.5768	5.5243	**0.5688**	**0.5699**	**11.6810**	0.6906	**0.7685**	9.6841	*0.3572*	*0.3023*	*7.8918*
DDEGSM	-	-	-	0.4574	0.4476	12.6294	0.3890	0.4321	12.3350	0.3329	0.2366	7.9672
Proposed	**0.6748**	**0.6293**	**5.1032**	*0.5277*	*0.4925*	*12.0195*	**0.7185**	*0.7540*	**9.2748**	**0.4025**	**0.3640**	**7.6956**

The best and second-best result are highlighted in bold and italic, respectively.

TABLE IV
COMPARISON OF THE PLCC, SRCC, AND RMSE FROM FOUR BNIQA, SIX BUIQA, DDEGSM, AND OUR PROPOSED METHODS ON THE ENHANCED IMAGES USING FUSION-12, FUSION-18, TWO-STEP, AND WATER-NET METHODS FROM THE SOTA DATA SET

	Fusion-12			Fusion-18			Two-Step			Water-Net		
Methods	PLCC ↑	SRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	RMSE ↓
BRISQUE	0.2107	0.2095	9.7681	0.2315	0.1878	8.1210	0.0538	0.0646	10.9140	0.1392	0.1139	6.2757
NIQE	0.3278	0.3077	9.4404	0.1313	0.1034	8.2756	0.1388	0.0957	10.8241	0.0575	0.0731	6.3269
IL-NIQE	0.5422	0.5263	8.3964	0.3594	0.2979	7.7900	0.3539	0.3332	10.2224	0.2294	0.1966	6.1684
dipIQ	0.2510	0.2565	9.6726	0.1668	0.1500	8.2309	0.1077	0.1080	10.8662	0.1286	0.1260	6.2848
UIQM	0.2178	0.2601	9.7525	0.3144	0.3018	7.9245	0.1614	0.1562	10.7866	0.2388	0.2278	6.1541
UCIQE	0.2023	0.2220	9.7858	0.1114	0.0968	8.2958	0.3460	0.2457	10.2548	0.1743	0.1680	6.2404
CCF	0.2060	0.1877	9.7781	0.4120	*0.3973*	7.6065	0.3270	0.2699	10.3289	0.1413	0.2013	6.2738
FDUM	0.1860	0.2176	9.8181	0.2408	0.2019	8.1023	0.3053	0.2158	10.4079	0.1539	0.1674	6.2619
UIF	0.5714	*0.5318*	8.2004	0.3810	0.3193	7.7183	0.4833	0.4317	9.5686	0.3149	0.2900	6.0151
URanker	*0.5783*	0.5150	*8.1519*	**0.4984**	**0.4436**	**7.2370**	*0.4854*	*0.4768*	*9.5557*	*0.4394*	*0.4399*	*5.6929*
DDEGSM	0.4544	0.2719	8.9013	0.1822	0.1353	8.2080	0.2205	0.1876	10.6609	0.1313	0.1028	6.2826
Proposed	**0.7373**	**0.6519**	**6.7293**	*0.4749*	0.3808	*7.2994*	**0.6497**	**0.6225**	**8.2583**	**0.4817**	**0.4400**	**5.5072**

The best and second-best result are highlighted in bold and italic, respectively.

the enhanced subset, we grouped all enhanced images from various subsets. Subsequently, we conducted training and testing on this merged set. Note that the UIF and DDEGSM methods results for the Raw subset are not reported, as these two methods require both raw and enhanced images for assessment, and they can only evaluate the quality of enhanced images. Based on performance comparisons on the SOTA data set, our proposed method demonstrates superior results on both raw and enhanced images across six of the nine enhanced groups within the SOTA data set: HP, IBLA, fusion-12, two-step, water-Net, and Dive+. In addition, it achieves the second-best results on the remaining subsets and the overall best result, as reported in the ALL column of Table V. The overall result is assessed by dividing the entire data set into training and testing subsets. The obtained results provide evidence of the proposed method's excellence in assessing the quality of both raw and enhanced images.

From Tables III through V the performance of our method slightly drops over the enhanced images in ACPAB, Fusion-18, and FUNIE-GAN subsets, while it has still close results to the best method. This is mainly due to the over-enhancement, blockiness, and inconsistent color correction shown in Fig. 7, but it still has competitive results over these subsets compared with other BUIQA methods. There is no superior method over

TABLE V
COMPARISON OF THE PLCC, SRCC, AND RMSE FROM FOUR BNIQA, SIX BUIQA, DDEGSM, AND OUR PROPOSED METHODS ON THE ENHANCED IMAGES USING FUNIE-GAN, DIVE+ METHODS AND ALL IMAGES FROM THE SOTA DATA SET

	FUNIE-GAN			Dive+			ALL		
Methods	PLCC ↑	SRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	RMSE ↓
BRISQUE	0.2331	0.1574	9.2503	0.1876	0.2042	8.8757	0.1172	0.1149	12.2039
NIQE	0.2044	0.1787	9.3116	0.1595	0.1664	8.9204	0.0902	0.0651	12.2385
IL-NIQE	*0.4257*	0.4050	*8.6077*	0.2704	0.2619	8.6996	0.3553	*0.3639*	11.4868
dipIQ	0.2153	0.1775	9.2893	0.1127	0.0998	8.9785	0.0794	0.0903	12.2498
UIQM	0.3549	0.3587	8.8931	0.2489	0.2508	8.7518	0.1630	0.1016	12.1243
UCIQE	0.4223	*0.4395*	8.6228	0.2639	0.2424	8.7158	*0.4490*	0.3414	*10.9800*
CCF	0.4787	*0.4716*	8.3516	0.3306	0.3353	8.5279	0.2937	0.3050	11.7467
FDUM	0.3629	0.3722	8.8640	0.2086	0.1455	8.8373	0.2536	0.1804	11.8869
UIF	0.3478	0.3524	8.9187	0.2983	0.2422	8.6247	0.2777	0.2595	12.1966
URanker	**0.6336**	**0.6089**	**7.3592**	*0.4664*	*0.4775*	*7.9932*	0.4587	0.4415	10.9034
DDEGSM	0.0999	0.1100	9.4649	0.1284	0.1352	8.9613	0.3334	0.2415	11.9696
Proposed	*0.4984*	0.4197	*8.1925*	**0.5481**	**0.5332**	**7.5080**	**0.5396**	**0.5225**	**7.5680**

The best and second-best result are highlighted in bold and italic, respectively.

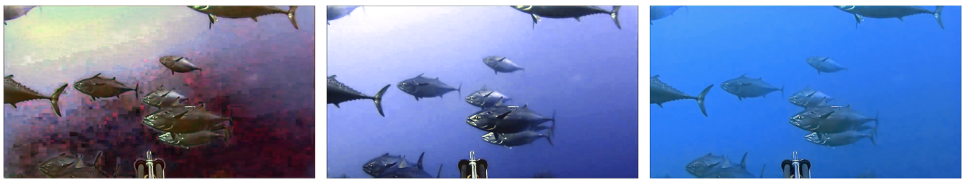


Fig. 7. Sample raw image taken from the SOTA data set [69] and from left to right its corresponding enhanced images produced by the ACPAB, Fusion-18, and FUNIE-GAN methods, respectively. The enhanced images suffer from over-enhancement, blockiness, and inconsistent color correction.

all others on this data set. However, our method achieves the top performance on raw images and six out of nine subsets of the enhanced images, which is the highest among the compared methods. From our observations, we have identified a limitation in our method coming from cases where the enhanced images exhibit a similar color distribution across the all images. In these instances, the features we currently extract struggle to differentiate between the images in a same group, resulting in lower IQA results.

Also, Table V, reports the overall performance of each method evaluated on all images within the SOTA data set, i.e., column ALL. Based on the results, our proposed method excels in performance across all images within the SOTA data set, establishing itself as the top-performing approach. This validates the conclusion that our method exhibits superior generalization ability compared to the other methods under consideration, even the URanker DNN-based method. Notably, the URaner method, while performing well in the ACPAB, fusion-18, and FUNIE-GAN subsets and achieving competitive results in Raw, HP, IBLA, fusion-12, fusion-18, two-step, and water-Net, falls short of delivering consistently high performance across all data set images. Considering these findings and the fact that each enhancement method gives specific characteristics to the resulting images, our proposed method emerges as a reliable and robust metric for evaluating the quality of both raw and enhanced underwater images.

D. Statistical Significance Test

To compare the statistical significance of our proposed method with both the BNIQA and BUIQA models we have utilized the commonly-used F-test, similar to [79] on the UWIQA data set. For comparison, we compute the ratio between their corresponding residual variances and denote it as F. If the F is greater than a confidence level of 95% then we consider a significant difference between the performance of the two compared methods. The corresponding result is shown in Table VI, where an element with value "1" and green color means that the method in the row is significantly better than the method in the column. Meanwhile, an element with value "0" and red color means that the method in the row is not significantly better. As seen from the obtained results, the FDUM and our proposed methods perform the best. The incorporated statistical significance test has clearly shown that the FDUM and the proposed method have the best performance compared with both the BNIQA and BUIQA methods, which is in line with their overall performance comparison made in the previous section.

E. Ablation Study

In this section, the effectiveness of the major components of the proposed method is analyzed. We investigated the contribution of the CIELAB color space, the impact of the SIC, DR, and overall saturation and hue scores to the overall performance, and

TABLE VI
STATISTICAL SIGNIFICANCE COMPARISON OF THE PROPOSED, BNIQA, AND BUIQA MODELS WITH THE F-TEST ON THE UWIQA DATA SET

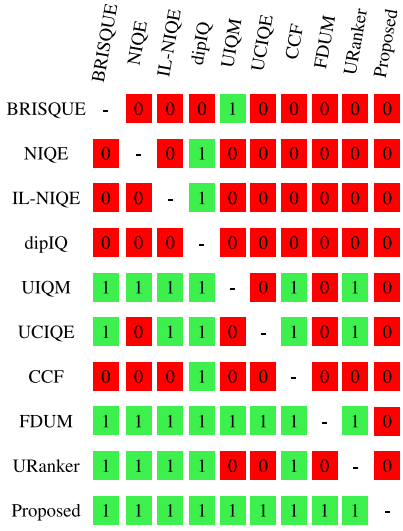


TABLE VII
COMPARISON OF THE PLCC, SRCC, AND RMSE OF OUR PROPOSED METHOD OBTAINED FROM DIFFERENT COLOR SPACES

Color space	PLCC ↑	SRCC ↑	RMSE ↓
RGB	0.6948	0.7014	0.1088
YUV	0.7064	0.7131	0.1074
LIN	0.6804	0.6904	0.1112
HSV	0.7066	0.7126	0.1072
YC_bC_r	0.7094	0.7157	0.1066
LAB	**0.7294**	**0.7376**	**0.1040**

The best result is highlighted in bold.

the structural components used to compute the SIC scores (i.e., IC and GBVS).

1) Discussion of the Color Spaces: To validate the effectiveness and consistency of the CIELAB color space in UWIQA, we have conducted experiments by evaluating the performance of the proposed method in various color spaces including RGB, CIELAB (LAB), YUV, linearized RGB (LIN), hue saturation value (HSV), and YC $_b$ C$_r$. As shown in Table VII, the performance of the proposed method in the YC $_b$ C$_r$ color space is better than those in the RGB, YUV, LIN, and HSV color spaces due to characteristic of the underwater images (i.e., nonuniform color cast). YC $_b$ C$_r$ color space is well separating the color information, i.e., the luminance and the blue and red color differences, of the image. However, underwater images are usually seen with more blue and green colors due to light attenuation. Thus, since the CIELAB is separating the green to red and blue to yellow colors, it helps the IQA model to measure the amount of green and red color information loss and blue color cast and consequently leads to better results for underwater image quality evaluation, i.e. the best in Table VII.

2) Discussion of the SIC, DR, Saturation, and Hue Scores: In this section, we analyzed the impact of the SIC and DR scores extracted from each channel as well as the overall saturation and hue of the image on the overall performance of the method on the UWIQA data set. Table VIII reports the obtained PLCC, SRCC, and RMSE scores for seven variants of the proposed method using the SIC and DR scores. As discussed above in the Section III-B, the proposed method obtains SIC scores in three channels (i.e., L, A, and B). Therefore, in Table VIII we have listed the individual performance of each SIC score and their combinations. As seen, the SIC score obtained from the L channel is the highest and among the chrominance channels (i.e., A and B), A leads to the highest correlation scores. Also, when the obtained scores are combined together, the combination of the $L + B$ SIC scores leads to the best results, while the $L + A$ and $A + B$ scores stand in the second and third place, respectively.

Seamlessly, we conducted the same experiments for the DR scores to see their impact on the overall performance. The results show that combining the DR scores from the three channels results in a better performance compared with using them individually. Note that for the DR scores, channel B has the highest correlation with subjective scores. Given that blue is the last color to dissolve, and underwater images predominantly feature this color more than any other, it can be inferred that DR scores exhibit a strong correlation with both underwater images and human visual perception. Overall, combining the DR scores with the SIC scores, improves the results from 0.7155 to 0.7210 and from 0.7227 to 0.7242 in terms of the PLCC and SRCC metrics.

Furthermore, to validate the effectiveness of the image's overall saturation and hue on the proposed method's performance, we did ablation study in Table IX. In this table, various combination channel-based SIC and DR scores with the saturation and hue score are used to evaluate the quality of the images. Based on the obtained results, although the overall scores has a very low performance individually, their combination with the channel-based scores, increases the performance of the method.

In addition to the previous analysis, it is essential to validate the effectiveness of the IC and saliency component of the SIC scores. Thus, the results from different combinations of the IC and GBVS SM are compared in Table X. From the results, it is obvious that the proposed method has significantly better performance when the IC is utilized for weighted pooling rather than the SM. However, combining the IC and GBVS together has improved the performance of the proposed method.

3) Discussion of the Baseline Methods: The proposed approach draws partial inspiration from [25], [33], [41], [62], and [63], with a focus on leveraging the DR and directional edges. This choice is motivated by the role of directional edges in facilitating the HVS's identification of image content. In addition, the DR, influenced by considerations of colorfulness, provides a metric for assessing the vividness of luminance and color within an image.

Given the common problem of color richness degradation in underwater images, it becomes crucial to evaluate both color information loss in original images and color refinement in enhanced images. Consequently, the DR is employed to measure these aspects, particularly relevant in the context of underwater images facing challenges, such as low-light conditions and color loss.

TABLE VIII
COMPARISON OF THE PLCC, SRCC, AND RMSE FROM OUR PROPOSED METHOD AND ITS SIX VARIANTS ON THE UWIQA DATA SET

	Criteria	L	A	B	$L+A$	$L+B$	$A+B$	$(L+A+B)$
SIC scores	PLCC ↑	0.6850	0.5189	0.5914	0.6974	0.7115	0.5959	**0.7155**
	SRCC ↑	0.6963	0.5107	0.5825	0.7050	0.7207	0.5833	**0.7227**
	RMSE ↓	0.1107	0.1293	0.1227	0.1099	0.1071	0.1222	**0.1061**
DR scores	PLCC ↑	0.4923	0.4038	0.5313	0.5179	0.5852	0.5394	**0.5875**
	SRCC ↑	0.4865	0.3863	0.5106	0.5021	0.5776	0.5208	**0.5795**
	RMSE ↓	0.1327	0.1396	0.1295	0.1289	0.1230	0.1271	**0.1226**

The best result is highlighted in bold.

TABLE IX
COMPARISON OF THE PLCC, SRCC, AND RMSE FROM OUR PROPOSED METHOD AND ITS COMBINED VARIANTS OF CHANNEL-BASED SIC AND DR SCORES WITH THE OVERALL SATURATION, AND HUE SCORES

Criteria	$Saturation$	Hue	$Saturation + Hue$	$SIC + DR + Saturation$	$SIC + DR + Hue$	$Proposed$
PLCC ↑	0.2588	0.3498	0.3173	0.7202	0.7237	**0.7294**
SRCC ↑	0.2304	0.3396	0.3051	0.7252	0.7316	**0.7376**
RMSE ↓	0.1468	0.1426	0.1459	0.1054	0.1047	**0.1040**

The best result is highlighted in bold.

TABLE X
COMPARISON OF THE PLCC, SRCC, AND RMSE FROM THE IC AND SALIENCY (I.E., GBVS) COMPONENTS OF OUR PROPOSED METHOD ON THE UWIQA DATA SET

Criteria	IC	Saliency	**Proposed** (IC + Saliency)
PLCC ↑	0.6975	0.6757	**0.7294**
SRCC ↑	0.7088	0.6858	**0.7376**
RMSE ↓	0.1086	0.1118	**0.1040**

The best result is highlighted in bold.

TABLE XI
COMPUTATIONAL RUN-TIME COMPARISON OF THE PROPOSED AND COMPARED METHODS ON THE UWIQA DATA SET

BNIQA Methods	BRISQUE	NIQE	IL-NIQE	dipIQ	
Time (s)	0.0719	0.0723	2.7905	1.2403	
BUIQA Methods	UIQM	UCIQE	CCF	FDUM	Proposed
Time (s)	0.8204	0.1051	1.3522	2.4083	0.5049

To assess the effectiveness of the proposed method in comparison to baseline approaches, we conducted evaluations using the available source code of the DDEGSM [33] and UIF [25] methods, and the results are presented in Tables II–V. The findings show a significant improvement in performance on all three data sets, which contain both the raw and enhanced images, demonstrating the efficacy of the proposed method in addressing the challenges associated with low-light conditions and color degradation in underwater images.

F. Computational Run-Time

In addition to the accuracy, consistency, and monotonicity of the IQA methods, they are also expected to have a reasonable computational complexity. Here, we have reported the average run-time of the proposed and compared methods in Table XI on the UWIQA data set. For a fair comparison, we have carried out experiments on a computer with Intel Core i9-10885H CPU @ 2.40 GHz and 32-GB memory using MATLAB R2022a. With an average execution time of 0.5049 s, the computation of SIC scores contributes approximately 0.44 s, while the calculations of the DR, the overall saturation and hue account for 0.06 s with almost equal contribution in between.

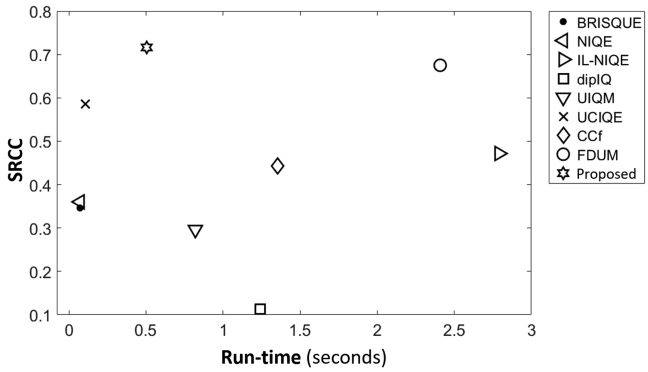


Fig. 8. Run-time versus SRCC scatter plots of both the proposed and compared methods on the UWIQA data set. The x-axis indicates the run-time and the y-axis indicates the SRCC.

Moreover, to have a better illustration of the performance and run-time efficiency, we plotted a 2-D plot of run-time versus SRCC of our proposed and compared methods in Fig. 8. As seen, our proposed method is the only method among the ones with reasonable performance (> 0.50), i.e., UCIQE, FDUM, and proposed, which has an efficient run-time and the best SRCC.

V. Conclusion

In this article, we proposed a novel BUIQA method by considering the proportional degradation and brightness/color richness (i.e., DR) of the image. Based on the experimental results, we observed the following: 1) the CIELAB color space is highly compatible with underwater IQA as it preserves the brightness/luminance information and the greed to red and blue to yellow color information thus yielding better results; 2) considering both the image's contour and saliency/GBVS maps in quality evaluation can significantly improve the performance; 3) fusion of the channel-based SIC and DR scores with the images overall saturation and hue makes the method capable of balancing the contribution of the luminance and chrominance channels on the overall score, resulting in better IQA performance for both raw and enhanced images. Eventually, the obtained results demonstrate the proposed method's accuracy, consistency, monotonicity, significance, and efficiency in comparison with current blind IQA methods.

Future work will develop more generic and robust image descriptors using DNNs, similar to [80], [81], to handle variations from enhancement, restoration, and color correction. Inspired by semi-supervised [82] and self-supervised learning strategies [83] and sensor modeling [84], we aim to reduce reliance on extensive annotations to account for wavelength-dependent attenuation, backscatter, and contrast loss, ensuring a comprehensive quality assessment of underwater images.

References

[1] K. Z. Mohd Azmi, A. S. Abdul Ghani, Z. Md Yusof, and Z. Ibrahim, "Natural-based underwater image color enhancement through fusion of swarm-intelligence algorithm," *Appl. Soft Comput.*, vol. 85, Dec. 2019, Art. no. 105810.

[2] 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.

[3] F. Ferreira, D. Machado, G. Ferri, S. Dugelay, and J. Potter, "Underwater optical and acoustic imaging: A time for fusion? a brief overview of the state-of-the-art," in *Proc. OCEANS MTS/IEEE Monterey*, 2016, pp. 1–6.

[4] H. Li, H. Wang, Y. Zhang, L. Li, and P. Ren, "Underwater image captioning: Challenges, models, and datasets," *ISPRS J. Photogramm. Remote Sens.*, vol. 220, pp. 440–453, Jan. 2025.

[5] P. Ren et al., "Oil spill drift prediction enhanced by correcting numerically forecasted sea surface dynamic fields with adversarial temporal convolutional networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 4701018.

[6] Y. Wang, J. Guo, H. Gao, and H. Yue, "UIEC^2-Net: CNN-based underwater image enhancement using two color space," *Signal Process. Image Commun.*, vol. 96, Aug. 2021, Art. no. 116250.

[7] P. Sharma, I. Bisht, and A. Sur, "Wavelength-based attributed deep neural network for underwater image restoration," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 19, no. 1, pp. 1–23, Jan. 2023.

[8] H. F. Tolie, J. Ren, and E. Elyan, "DICAM: Deep inception and channel-wise attention modules for underwater image enhancement," *Neurocomputing*, vol. 584, Jun. 2024, Art. no. 127585.

[9] D. Berman, D. Levy, S. Avidan, and T. Treibitz, "Underwater single image color restoration using haze-lines and a new quantitative dataset," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 8, pp. 2822–2837, Aug. 2021.

[10] G. Sequeira, V. Mekkalki, J. Prabhu, S. Borkar, and M. Desai, "Hybrid approach for underwater image restoration and enhancement," in *Proc. Int. Conf. Emerg. Smart Comput. Informat.*, 2021, pp. 427–432.

[11] P. Guo, L. He, S. Liu, D. Zeng, and H. Liu, "Underwater image quality assessment: Subjective and objective methods," *IEEE Trans. Multimedia*, vol. 24, pp. 1980–1989, 2022.

[12] N. Yang, Q. Zhong, K. Li, R. Cong, Y. Zhao, and S. Kwong, "A reference-free underwater image quality assessment metric in frequency domain," *Signal Process. Image Commun.*, vol. 94, May. 2021, Art. no. 116218.

[13] H. F. Tolie, J. Ren, R. Chen, H. Zhao, and E. Elyan, "Blind sonar image quality assessment via machine learning: Leveraging micro-and macro-scale texture and contour features in the wavelet domain," *Eng. Appl. Artif. Intell.*, vol. 141, Feb. 2025, Art. no. 109730.

[14] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[15] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[16] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500–1512, Apr. 2012.

[17] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185–1198, May 2011.

[18] A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3378–3389, Aug. 2012.

[19] R. Soundararajan and A. C. Bovik, "RRED indices: Reduced reference entropic differencing for image quality assessment," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 517–526, Feb. 2012.

[20] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50–63, Jan. 2015.

[21] Q. Li, W. Lin, J. Xu, and Y. Fang, "Blind image quality assessment using statistical structural and luminance features," *IEEE Trans. Multimedia*, vol. 18, pp. 2457–2469, 2016.

[22] Y. Fang et al., "Blind quality assessment for tone-mapped images by analysis of gradient and chromatic statistics," *IEEE Trans. Multimedia*, vol. 23, pp. 955–966, 2021.

[23] S. Raimondo and C. Silvia, "Underwater image processing: State of the art of restoration and image enhancement methods," *Eurasip J. Adv. Signal Process.*, vol. 2010, pp. 1–14, Dec. 2010.

[24] D. L. Ruderman, "The statistics of natural images," *Netw.: Comput. Neural Syst.*, vol. 5, no. 4, pp. 517–548, Jul. 1994.

[25] Y. Zheng, W. Chen, R. Lin, T. Zhao, and P. Le Callet, "UIF: An objective quality assessment for underwater image enhancement," *IEEE Trans. Image Process.*, vol. 31, pp. 5456–5468, 2022.

[26] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.

[27] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.

[28] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.

[29] K. Ma, W. Liu, T. Liu, Z. Wang, and D. Tao, "dipIQ: Blind image quality assessment by learning-to-rank discriminable image pairs," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3951–3964, Aug. 2017.

[30] H. Lin, H. Men, Y. Yan, J. Ren, and D. Saupe, "Crowdsourced quality assessment of enhanced underwater images - A pilot study," in *Proc. 14th Int. Conf. Qual. Multimedia Exp.*, 2022, pp. 1–4.

[31] T. Leisti, J. Radun, T. Virtanen, R. Halonen, and G. Nyman, "Subjective experience of image quality: Attributes, definitions, and decision making of subjective image quality," in *Proc. SPIE*, vol. 7242, pp. 130–138, 2009.

[32] K. Gu, J. Zhou, J.-F. Qiao, G. Zhai, W. Lin, and A. C. Bovik, "No-reference quality assessment of screen content pictures," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 4005–4018, Aug. 2017.

[33] H. F. Tolie and M. R. Faraji, "Screen content image quality assessment using distortion-based directional edge and gradient similarity maps," *Signal Process. Image Commun.*, vol. 101, Feb. 2022, Art. no. 116562.

[34] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6062–6071, Dec. 2015.

[35] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE J. Ocean. Eng.*, vol. 41, no. 3, pp. 541–551, Jul. 2016.

[36] J. Bednar and T. Watt, "Alpha-trimmed means and their relationship to median filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 1, pp. 145–153, Feb. 1984.

[37] K. Panetta, A. Samani, and S. Agaian, "Choosing the optimal spatial domain measure of enhancement for mammogram images," *Int. J. Biomed. Imag.*, vol. 2014, no. 1, Aug. 2014, Art. no. 937849.

[38] K. Panetta, S. Agaian, Y. Zhou, and E. J. Wharton, "Parameterized logarithmic framework for image enhancement," *IEEE Trans. Syst. Man Cybern., Part B*, vol. 41, no. 2, pp. 460–473, Apr. 2011.

[39] Y. Wang et al., "An imaging-inspired no-reference underwater color image quality assessment metric," *Comput. Electr. Eng.*, vol. 70, pp. 904–913, Aug. 2018.

[40] C. Guo et al., "Underwater ranker: Learn which is better and how to be better," in *Proc. AAAI Conf. Artif. Intell.*, Jun. 2023, vol. 37, no. 1, pp. 702–709.

[41] R. Wang, H. Yang, Z. Pan, B. Huang, and G. Hou, "Screen content image quality assessment with edge features in gradient domain," *IEEE Access*, vol. 7, pp. 5285–5295, 2019.

[42] C.-Y. Li, J.-C. Guo, R.-M. Cong, Y.-W. Pang, and B. Wang, "Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior," *IEEE Trans. Image Process.*, vol. 25, no. 12, pp. 5664–5677, Dec. 2016.

[43] H. Wang, S. Sun, X. Bai, J. Wang, and P. Ren, "A reinforcement learning paradigm of configuring visual enhancement for object detection in underwater scenes," *IEEE J. Ocean. Eng.*, vol. 48, no. 2, pp. 443–461, Apr. 2023.

[44] H. Wang et al., "Inspiration: A reinforcement learning-based human visual perception-driven image enhancement paradigm for underwater scenes," *Eng. Appl. Artif. Intell.*, vol. 133, Jul. 2024, Art. no. 108411.

[45] H. Wang, W. Zhang, and P. Ren, "Self-organized underwater image enhancement," *ISPRS J. Photogramm. Remote Sens.*, vol. 215, pp. 1–14, Sep. 2024.

[46] H. Wang, W. Zhang, L. Bai, and P. Ren, "Metalantis: A comprehensive underwater image enhancement framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5618319.

[47] H. Wang, K. Köser, and P. Ren, "Large foundation model empowered discriminative underwater image enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5609317.

[48] G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Hoboken, NJ, USA: Wiley, 2000, vol. 40.

[49] G. Sharma and R. Bala, *Digital Color Imaging Handbook*. Boca Raton, FL, USA: CRC Press, 2017.

[50] R. A. Kirsch, "Computer determination of the constituent structure of biological images," *Comput. Biol. Res.*, vol. 4, no. 3, pp. 315–328, Jun. 1971.

[51] Z. Ni, L. Ma, H. Zeng, C. Cai, and K. Ma, "Gradient direction for screen content image quality assessment," *IEEE Signal Process. Lett.*, vol. 23, pp. 1394–1398, Oct. 2016.

[52] Y. Fu, H. Zeng, L. Ma, Z. Ni, J. Zhu, and K. Ma, "Screen content image quality assessment using multi-scale difference of Gaussian," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 9, pp. 2428–2432, Sep. 2018.

[53] X. Zhang, X. Feng, W. Wang, and W. Xue, "Edge strength similarity for image quality assessment," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 319–322, Apr. 2013.

[54] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.

[55] S. K. Ghosal, A. Chatterjee, and R. Sarkar, "Image steganography based on kirsch edge detection," *Multimedia Syst.*, vol. 27, no. 1, pp. 73–87, Oct. 2021.

[56] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Oct. 2014.

[57] S. Wang, K. Gu, S. Ma, W. Lin, X. Liu, and W. Gao, "Guided image contrast enhancement based on retrieved images in cloud," *IEEE Trans. Multimedia*, vol. 18, pp. 219–232, 2016.

[58] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Advances in Neural Information Processing Systems*, B. Schö lkopf, J. Platt, and T. Hoffman, Eds., vol. 19. Cambridge, MA, USA: MIT Press, 2006.

[59] Y. Fang, W. Lin, B.-S. Lee, C.-T. Lau, Z. Chen, and C.-W. Lin, "Bottom-up saliency detection model based on human visual sensitivity and amplitude spectrum," *IEEE Trans. Multimedia*, vol. 14, no. 1, pp. 187–198, Feb. 2012.

[60] A. L. N. T. Da Costa and M. N. Do, "A retina-based perceptually lossless limit and a gaussian foveation scheme with loss control," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 3, pp. 438–453, Jun. 2014.

[61] K.-Q. Huang, Q. Wang, and Z.-Y. Wu, "Natural color image enhancement and evaluation algorithm based on human visual system," *Comput. Vis. Image Understanding*, vol. 103, no. 1, pp. 52–63, Jul. 2006.

[62] G. Yue, C. Hou, and T. Zhou, "Blind quality assessment of tone-mapped images considering colorfulness, naturalness, and structure," *IEEE Trans. Ind. Electron.*, vol. 66, no. 5, pp. 3784–3793, May 2019.

[63] M. R. Faraji, "Full-reference tone-mapped images quality assessment," *IET Image Process.*, vol. 15, no. 3, pp. 763–773, Dec. 2021.

[64] Y. Cui, M. Yu, G. Jiang, Z. Peng, and F. Chen, "Blind tone-mapped HDR image quality measurement by analysis of low-level and high-level perceptual characteristics," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 5022015.

[65] R. C. Gonzales and P. Wintz, *Digital Image Processing* 2nd ed., USA: Addison-Wesley Longman Publishing Co., Inc., 1987.

[66] J. Van de Weijer, T. Gevers, and A. D. Bagdanov, "Boosting color saliency in image feature detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 150–156, Jan. 2006.

[67] H. F. Tolie, M. R. Faraji, and X. Qi, "Blind quality assessment of screen content images via edge histogram descriptor and statistical moments," *Vis. Comput.*, vol. 40, no. 8, pp. 5341–5356, 2024.

[68] G. Hou, Y. Li, H. Yang, K. Li, and Z. Pan, "UID2021: An underwater image dataset for evaluation of no-reference quality assessment metrics," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 19, no. 4, pp. 1–24, Feb. 2023.

[69] Z. Wang, L. Shen, Z. Wang, Y. Lin, and Y. Jin, "Generation-based joint luminance-chrominance learning for underwater image quality assessment," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 3, pp. 1123–1139, Mar. 2023.

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

[71] Y. Lin, L. Shen, Z. Wang, K. Wang, and X. Zhang, "Attenuation coefficient guided two-stage network for underwater image restoration," *IEEE Signal Process. Lett.*, vol. 28, pp. 199–203, 2021.

[72] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Trans. Consum. Electron*, vol. 53, no. 2, pp. 593–600, May 2007.

[73] Y.-T. Peng and P. C. Cosman, "Underwater image restoration based on image blurriness and light absorption," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1579–1594, Apr. 2017.

[74] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 81–88.

[75] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and P. Bekaert, "Color balance and fusion for underwater image enhancement," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 379–393, Jan. 2018.

[76] X. Fu, Z. Fan, M. Ling, Y. Huang, and X. Ding, "Two-step approach for single underwater image enhancement," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, 2017, pp. 789–794.

[77] 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, Apr. 2020.

[78] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," Aug. 2015. [Online]. Available: http://www.its.bldrdoc.gov/vqeg/vqeg-home.aspx

[79] Z. Ni, H. Zeng, L. Ma, J. Hou, J. Chen, and K.-K. Ma, "A Gabor feature-based quality assessment model for the screen content images," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4516–4528, Sep. 2018.

[80] Y. Xu et al., "Hyperspectral image super-resolution with CONVLSTM skip-connections," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5519016.

[81] Y. Xie, X. Hou, J. Ren, X. Zhang, C. Ma, and J. Zheng, "Binary quantization vision transformer for effective segmentation of red tide in multi-spectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 4202814.

[82] Z. Fang, J. Ren, J. Zheng, R. Chen, and H. Zhao, "Dual teacher: Improving the reliability of pseudo labels for semi-supervised oriented object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5602515.

[83] E. Zhang, H. Zong, X. Li, M. Feng, and J. Ren, "ICSF: Integrating inter-modal and cross-modal learning framework for self-supervised heterogeneous change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5501516.

[84] J. Cai, Z. Wei, and J. Luo, "ICS anomaly detection based on sensor patterns and actuator rules in spatiotemporal dependency," *IEEE Trans. Ind. Informat.*, vol. 20, no. 8, pp. 10647–10656, Aug. 2024.

Hamidreza Farhadi Tolie received the B.Sc. degree in information technology and the M.Sc. degree in computer science from the Institute for Advanced Studies in Basic Sciences, Zanjan, Iran, in 2018 and 2021, respectively, and the Ph.D. degree in artificial intelligence from Robert Gordon University, Aberdeen, Scotland, in 2025.

He is currently a Research Assistant with Warwick Manufacturing Group, University of Warwick, Coventry, U.K. His research interests include signal/image processing, image quality assessment/enhancement, and computer vision.

Jinchang Ren (Senior Member, IEEE) received the Ph.D. degree in electronic imaging from the University of Bradford, Bradford, U.K., in 2009.

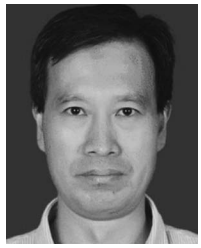
He is currently a Distinguished Professor with the School of Computer Science, Guangdong Polytechnic Normal University, China. His research interests include hyperspectral imaging, image processing, computer vision, Big Data analytics, and machine learning.

Jun Cai received the B.S. degree from Hunan Normal University, Changsha, China, in 2003, the M.S. degree from Jinan University, Guangzhou, China, in 2006, and the Ph.D. degree from Sun Yat-sen University, Guangzhou, in 2012, all in comunications and information systems.

He is currently a Professor with the School of Cyber Security, Guangdong Polytechnic Normal University, Guangzhou. His research interests include the research of complex networks, traffic modeling, anomaly detection, ICN, SDN, IIoT, and mobile network.

Rongjun Chen received the M.S. degree in control theory and control engineering from the Guangdong University of Technology, Guangzhou, China, in 2007, and the Ph.D. degree in communication and information system from Sun Yat-sen University, Guangzhou, in 2015.

He is currently a Professor with the School of Computer Science, Guangdong Polytechnic Normal University, Guangzhou. His research interests include image perception and processing, as well as the Internet of Things.



Huimin Zhao was born in Shaanxi, China, in 1966. He received the B.Sc. and M.Sc. degrees in signal processing from Northwestern Polytechnical University, Xi'an, China, in 1992 and 1997, respectively, and the Ph.D. degree in electrical engineering from Sun Yat-sen University, Guangzhou, China, in 2001.

He is currently a Professor with Guangdong Polytechnic Normal University, Guangzhou. His research interests include image, video, and information security technology.