



## Article

# Enhancing Underwater Image Quality Assessment with Influential Perceptual Features

Feifei Liu <sup>1</sup>, Zihao Huang <sup>2</sup>, Tianrang Xie <sup>3</sup>, Runze Hu <sup>3</sup> and Bingbing Qi <sup>3,\*</sup><sup>1</sup> Changshu Institute of Technology, Suzhou 215556, China; liuff@cslg.edu.cn<sup>2</sup> School of Business and Commerce, Southwest University, Chongqing 400715, China; apaobu@email.swu.edu.cn<sup>3</sup> Beijing Institute of Technology, Beijing 100811, China; hrz@bit.edu.cn (R.H.)

\* Correspondence: 6120230031@bit.edu.cn

**Abstract:** In the multifaceted field of oceanic engineering, the quality of underwater images is paramount for a range of applications, from marine biology to robotic exploration. This paper presents a novel approach in underwater image quality assessment (UIQA) that addresses the current limitations by effectively combining low-level image properties with high-level semantic features. Traditional UIQA methods predominantly focus on either low-level attributes such as brightness and contrast or high-level semantic content, but rarely both, which leads to a gap in achieving a comprehensive assessment of image quality. Our proposed methodology bridges this gap by integrating these two critical aspects of underwater imaging. We employ the least-angle regression technique for balanced feature selection, particularly in high-level semantics, to ensure that the extensive feature dimensions of high-level content do not overshadow the fundamental low-level properties. The experimental results of our method demonstrate a remarkable improvement over existing UIQA techniques, establishing a new benchmark in both accuracy and reliability for underwater image assessment.

**Keywords:** image quality assessment; vision transformer; low-level; feature selection



**Citation:** Liu, F.; Huang, Z.; Xie, T.; Hu, R.; Qi, B. Enhancing Underwater Image Quality Assessment with Influential Perceptual Features. *Electronics* **2023**, *12*, 4760. <https://doi.org/10.3390/electronics12234760>

Academic Editor: Krzysztof Okarma

Received: 31 October 2023

Revised: 15 November 2023

Accepted: 20 November 2023

Published: 23 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The burgeoning field of oceanic engineering heavily relies on the quality of underwater images for a multitude of applications ranging from marine biology research to underwater robotics. Underwater image quality assessment (UIQA) plays a pivotal role in ensuring the reliability and efficacy of these applications [1,2]. High-quality underwater images are essential for accurate data analysis and interpretation, making UIQA an indispensable aspect of oceanic exploration and monitoring.

Image quality assessment (IQA) fundamentally revolves around three principal categories: full-reference (FR) [3,4], reduced-reference (RR) [5,6], and no-reference (NR) [7,8]. FR IQA methods compare a degraded image with a high-quality reference image, while RR IQA methods require only partial information from the reference. However, in most underwater imaging scenarios, distortion-free reference images are not available, rendering FR and RR methods impractical. Consequently, NR IQA, which does not require any reference image, becomes particularly relevant for UIQA. This paradigm shift towards NR IQA is driven by its adaptability in dynamic, unpredictable underwater environments where obtaining reference images is often unfeasible.

Over the last decade, various IQA methods have been proposed and achieved impressive performance. For instance, Ma et al. [9] proposed the ASCAM-Former that introduces the channel-wised self-attention into IQA. The spatial and channel dependencies among features are characterized, rendering to a comprehensive quality evaluation. Nizami et al. [10] proposed an NR-IQA method that combines visual saliency and a convolution neural network (CNN), where the salient region of the image is identified first and then sent to

the CNN for quality evaluation. Rehman et al. [11] designed a network applicable for both natural scene images and screen content. The network is based on the patch salience, which is determined by the texture and edge properties of the image. Sendjasni et al. [12] proposed an NR-IQA model for 360 degree images using a perceptually weighted multi-channel CNN. Zhou et al. [13] further dealt with the 360 degree images through deep local and global spatiotemporal feature aggregation. HOU et al. [14] established an underwater IQA database of 60 degraded underwater images, which involves many unique distortions of underwater images, including blue-green scenes, greenish scenes, hazy scenes, and low-light scenes. Nizami et al. [15] proposed to utilize the discrete cosine transform (DCT) features for quality evaluation. Bouris et al. [16] developed a CNN-based IQA model for the quality evaluation of optic disc photographs, wherein a large-scale dataset is dedicatedly established. Nizami et al. [17] proposed an NR-IQA method that can extract quality-aware features in an elegant and efficient manner based on the Harris affine detector and scale invariant feature transformation. Zhou et al. [18] investigated the NR-IQA strategy for evaluating the perceptual quality of omnidirectional images (360 degree images), where an effective metric is proposed based on the low-level statistics and high-level semantics of omnidirectional images. Jiang et al. [19] proposed a full-reference solution for omnidirectional images by characterizing the degradation in a hierarchical manner.

Existing UIQA methods [20–23] predominantly focus on assessing low-level image attributes such as brightness, colorfulness, contrast, and sharpness. While these factors are crucial in determining the perceptual quality of underwater images, such methods often fall short by not considering the high-level semantic content that these images convey. The semantic aspect, encompassing the contextual and compositional elements of the image, is vital for a comprehensive understanding and utilization of the underwater visual data [24,25]. This oversight limits the scope and effectiveness of current UIQA techniques.

Moreover, while the advent of deep-learning techniques in IQA has significantly advanced the field, there remains a tendency in these methods to overlook the unique characteristics of underwater imagery, such as the specific distortions and color shifts caused by water medium. Deep-learning approaches often focus extensively on high-level feature extraction, potentially neglecting the intrinsic low-level properties specific to underwater images. This imbalance can lead to suboptimal performance when these models are applied to UIQA tasks [26], necessitating a more integrated approach that considers both low- and high-level image characteristics.

Our work aims to bridge this gap by developing an innovative approach that amalgamates both the low-level perceptual properties and abstract semantic features of underwater images, herein namely underwater image quality metric with low-level properties and selected high-level semantics (UIQM-LSH). Recognizing the disparity in feature dimensions between these two types of features, we employ the least-angle regression method (LARS) [27,28] for effective feature selection on the high-level semantic attributes. This approach ensures that the high-level features do not overshadow the crucial low-level properties, thus achieving a balanced and comprehensive feature representation for UIQA [29–31]. Our proposed UIQM-LSH stands out by adeptly integrating the diverse aspects of underwater images, providing a more nuanced and effective assessment of image quality. Through extensive experiments, UIQM-LSH demonstrates superior performance against existing UIQA or IQA approaches.

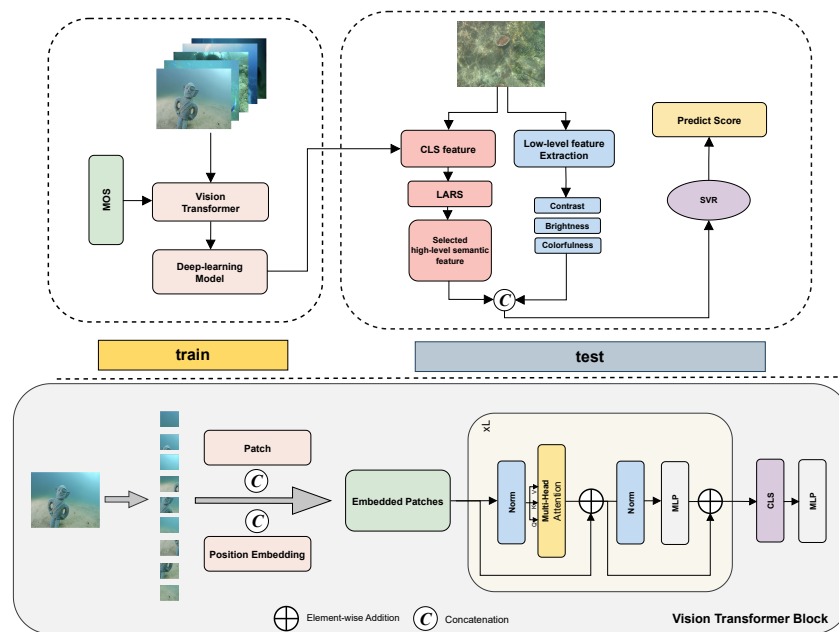
In summary, this work makes several new contributions:

- We have developed an innovative methodology, namely UIQM-LSH, which integrates both low-level image properties and high-level semantic features in underwater images. By considering both the intrinsic visual elements such as brightness, colorfulness, and sharpness, along with the contextual and compositional high-level semantic information, our method provides a nuanced understanding of underwater image quality, addressing a significant gap in existing UIQA methods.
- We implement the least-angle regression method for efficient feature selection in high-level semantics. This technique effectively balances the feature representation,

preventing high-level features from dominating the assessment. By fine-tuning the feature selection process, our approach ensures that both low-level and high-level features contribute appropriately to the overall image quality assessment, leading to a more accurate and reliable evaluation.

## 2. Proposed Method

In this section, we elaborate on the proposed UIQA metric, namely Underwater Image Quality Metric with Low-level properties and Selected High-level semantics (UIQM-LSH), which is capable of characterizing the perceptual properties of underwater images reliably and comprehensively. The proposed UIQM-LSH consists of three branches, i.e., (1) a low-level feature extraction branch for characterizing the low-level properties of underwater images, (2) a high-level semantic branch for extracting the image content-related features, and (3) a feature balancing branch to prevent the extracted features from being dominated by high-level features. The processed features and the corresponding MOS values are sent to the support vector machine (SVM) to derive the final quality score. Figure 1 illustrates the overall architecture of the UIQM-LSH. We will cover each of these modules in detail in the following sections.



**Figure 1.** The proposed UIQM-LSH architecture includes a training module and a testing module. In the training module, the deep-learning model is obtained according to the image and the corresponding subjective quality score (MOS) through the vision transformer block. The structure of the vision transformer block is shown below the dashed line. In the test module, CLS features are extracted from the training model and high-level semantic features are selected from the LARS module. The predicted quality score is obtained by the SVM model.

### 2.1. Low-Level Underwater Properties Characterization

#### 2.1.1. Color Cast

Underwater image quality is often impaired by color cast. The underwater images usually appear bluish or greenish. Color cast occurs due to unequal absorption of various colors of light underwater, specifically, red light suffers greater loss due to the absorption by the water. It is indicated in [32] that in the CIELAB color space, the chromaticity distribution on the ab chrominance coordinate plane of underwater images with color cast is relatively concentrated and has a single peak with a small degree of dispersion, and the mean of chromaticity is large. Conversely, images that do not exhibit color cast or have a less severe degree of color cast have a chromaticity distribution on the ab chrominance coordinate

plane that exhibits multiple peaks with a large degree of dispersion and a relatively small chromaticity mean. Based on the above results, the degree of color cast of underwater images can be characterized by calculating the chromaticity mean and variance in the  $ab$  chrominance space. The chromaticity means in  $a$  and  $b$  chrominance channel can be calculated as follows:

$$\bar{c}_a = \frac{\sum_{i=1}^M c_{ai}}{M}$$

$$\bar{c}_b = \frac{\sum_{i=1}^M c_{bi}}{M}$$

where  $\bar{c}_a$  and  $\bar{c}_b$ , respectively, represent the average chromaticity values of all pixel points in chrominance channels  $a$  and  $b$ .  $M$  is the amount of all pixels of the underwater image. The overall chromaticity mean can be denoted as  $\bar{c}$ :

$$\bar{c} = \sqrt{\bar{c}_a^2 + \bar{c}_b^2} \quad (1)$$

Then, we calculate the variances of the chromaticity distribution in chrominance channels  $a$  and  $b$

$$\sigma_a^2 = \frac{\sum_{i=1}^M (c_{ai} - \bar{c}_a)^2}{M}$$

$$\sigma_b^2 = \frac{\sum_{i=1}^M (c_{bi} - \bar{c}_b)^2}{M}$$

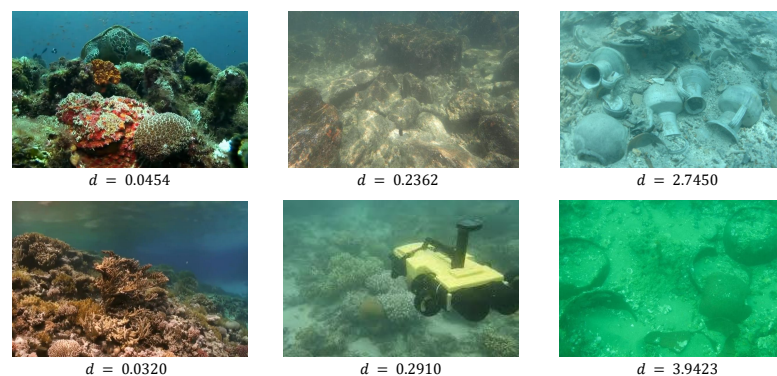
where  $\sigma_a^2$  and  $\sigma_b^2$  represent the variance of the chromaticity distribution of the underwater image in  $a$  and  $b$  chrominance channels, respectively. Then,  $\sigma_a^2$  and  $\sigma_b^2$  can be used to characterize the average chromaticity variance of the image  $\sigma^2$ :

$$\sigma^2 = \sqrt{\sigma_a^2 + \sigma_b^2} \quad (2)$$

Finally, we evaluate the color cast degree using the ratio of chromaticity mean to variance, denoted by  $d$ :

$$d = \frac{\bar{c}}{\sigma^2 + \tau} \quad (3)$$

where  $\tau$  is a small constant introduced to reduce instability, typically set to 0.0001 based on empirical values. The larger the ratio index  $d$  is, the greater the degree of color cast in the image is. Figure 2 shows some underwater images with color cast. Intuitively, as the degree of color cast in the underwater images increases, the value of the ratio index  $d$  correspondingly increases.



**Figure 2.** Underwater images with different degrees of color cast and the ratio index  $d$ .

### 2.1.2. Brightness

Due to the absorption of light by the water, underwater images often suffer from a lack of brightness. The overall brightness and the uniformity of brightness distribution are two key factors in evaluating the quality of underwater images. When evaluating the overall brightness, a method to calculate the gray-scale mean value of pixel points can be adopted. First, the underwater image obtained is converted into a gray-scale image  $G$ , and the gray-scale values of each pixel point in the image are obtained and the brightness mean value is calculated, which is recorded as  $\bar{g}$ .

$$\bar{g} = \frac{\sum_{i=1}^M g_i}{M} \quad (4)$$

where  $m$  represents the number of pixel points in the image and  $g_i$  represents the gray-scale values of pixel points in gray-scale image  $G$ . To evaluate the uniformity of brightness distribution in underwater images, it is necessary to statistic the brightness variations in different parts of the image. Here, the image is divided equally into  $n$  blocks geometrically, and the average brightness of each pixel point in each image block is calculated to obtain a set of brightness means containing  $n$  image blocks. Let the maximum element be  $g_{max}$ , which corresponds to the mean brightness of the brightest part of the image block. The minimum element is denoted by  $g_{min}$ . The brightness variations in underwater images can be calculated as:

$$h = \frac{g_{max} - g_{min}}{g_{max} + \eta} \quad (5)$$

where  $h$  represents the overall brightness variation degree of the image.  $\eta$  is a small constant introduced to reduce instability, typically set to 0.0001 based on empirical values. The smaller the value  $h$  is, the more uniform the brightness distribution of the underwater image is.

### 2.1.3. Contrast

Contrast is an important parameter for evaluating image quality. In this study, we evaluated the impact of contrast on underwater image quality from two aspects: local contrast and global contrast. When describing local contrast, we referred to the concept of contrast energy (CE) and calculated the contrast energy of the image in three chrominance channels  $gray, yb, rg$ , which can be linearly represented by the  $RGB$  chrominance channels [33]. Specifically, the calculation methods are as follows:  $gray = 0.299R + 0.587G + 0.114B$ ,  $yb = 0.5(R + G) - B$ , and  $rg = R - G$ . The calculation method for CE is as follows:

$$CE_k = \frac{\beta \cdot Q(k)}{Q(k) + \beta \cdot G} - e_k \quad (6)$$

$$Q(k) = \sqrt{(I_k \otimes d_h)^2 + (I_k \otimes d_v)^2} \quad (7)$$

where  $k$  represents one of the color channels  $\{gray, yb, \text{ or } rg\}$ .  $d_h$  and  $d_v$  are the horizontal and vertical second derivatives of the Gaussian function.  $\otimes$  represents convolution computation, and  $\beta$  is the maximum value of  $Q(k)$ .  $G$  represents contrast gain.  $n_k$  represents the noise threshold for the color channel  $k$ .

Evaluating the global contrast of underwater images can be achieved by analyzing the grayscale histogram of the image. Generally, images with high contrast have relatively flat grayscale histogram distribution with a wide range of grayscale values, while images with low contrast have more steeply shaped grayscale histograms with a smaller range of grayscale values. Based on the distribution characteristics of the grayscale histogram,

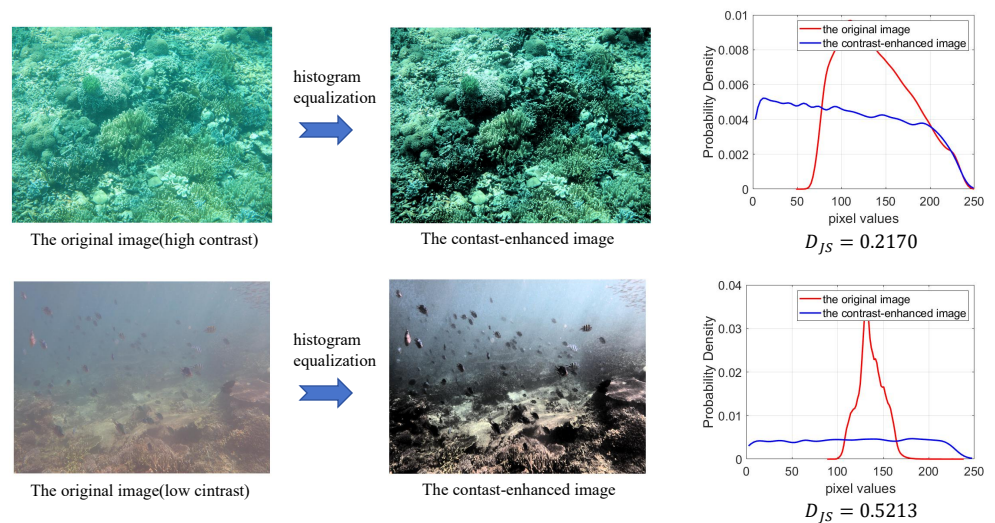
the global contrast of the image can be reflected by comparing the grayscale histogram of the image with the uniform distribution. Here, we calculate the Kullback–Leibler divergence (KL divergence) between the grayscale histogram distribution of the image and the uniform distribution to represent the flatness of the grayscale histogram. The smaller the KL divergence is, the closer the grayscale histogram is to a uniform distribution, and the higher the global contrast is. Let  $\mathbf{P}$  represent the grayscale histogram distribution of the underwater image and  $\mathbf{Q}$  represent the uniform distribution. The KL divergence between  $\mathbf{P}$  and  $\mathbf{Q}$  is calculated as follows,

$$D_{K_L}(\mathbf{P} \parallel \mathbf{Q}) = - \int P(x)\log Q(x)dx + \int P(x)\log P(x)dx \tag{8}$$

Due to the asymmetry of the KL divergence (i.e.,  $D_{JS}(\mathbf{P} \parallel \mathbf{Q}) = D_{KL}(\mathbf{P} \parallel \mathbf{Q})$ ), we adopt the symmetric JS divergence to characterize the flatness of the image grayscale histogram. The JS divergence can be derived from the KL divergence as follows,

$$D_{J_S}(\mathbf{P} \parallel \mathbf{Q}) = \frac{1}{2}[D_{K_L}(\mathbf{P} \parallel \mathbf{N}) + D_{K_L}(\mathbf{Q} \parallel \mathbf{N})] \tag{9}$$

where  $\mathbf{N} = \frac{\mathbf{P} + \mathbf{Q}}{2}$ . Finally, we use the J-S divergence to characterize the global contrast of the underwater images. In Figure 3, we plotted the results of enhancing the contrast of high-contrast and low-contrast images through histogram equalization. The results of the histogram distribution show that after histogram equalization, both the high-contrast original image and the low-contrast original image exhibit a more even histogram distribution, resulting in improved contrast. Additionally, we calculated the  $D_{JS}$  values for both original images. It is shown that the low-contrast image has a higher JS divergence compared to the high-contrast image.



**Figure 3.** Comparison of high- and low-contrast original underwater images with contrast-enhanced images and their histogram distributions with the  $D_{JS}$  values of the original images.

### 2.2. Vision Transformer

Image quality assessment is the perception of the content of an image by the human visual system (HVS), including a combination of global high-level quality perception and local low-level features [25,34]. In the previous section, we concentrated on evaluating the local features of underwater images. This section presents a global evaluation of image semantic quality based on the spatial domain. Specifically, we scan and sense the image quality in different spatial domains and obtain the global image quality after comparing and fusing the sensing results. In [35], we learnt that the quality relationship between different spatial domains of an image is significant in assessing the quality of the global image.

In recent years, due to the enhanced global perception capabilities of the transformer, it has found extensive applications in both the field of natural language processing and computer vision [36–38]. Therefore, we propose utilizing transformer modules to extract the quality relationships between adjacent spatial regions and obtain the global quality of an image.

When provided with an input image  $I$ , we first break down the image into  $N$  patches, each with a resolution of  $p * p$  and  $N = \frac{H_1 W_1}{p^2}$ .  $N$  is the total number of patches, and each image patch is represented as a vector containing that specific patch's pixel information. After flattening each patch, they are passed through a linear projection layer to map each patch into a  $D$ -dimension embedding. Next, we embed a learnable embedding called the Image-Quality-Collection cls token  $C_{IQC} \in \mathbb{R}^{1 \times D}$  to concatenate with the perceived underwater image information. Finally, we incorporate positional encoding into these  $N + 1$  embeddings to capture the spatial relationships within the underwater image.

#### Self-Attention Module

After obtaining the embedding vector  $E = \{C_{IQC}, E_1, \dots, E_N\} \in \mathbb{R}^{(N+1) \times D}$ , we use the multi-head self-attention (MHSA) block to obtain the self-attention scores. The MHSA block includes  $h$  heads, which means attending to different semantic information and feature representations while obtaining a dimension of  $d = \frac{D}{h}$ . Following the normalization layer, the self-attention layer generates projections for *query*( $Q$ ), *key*( $K$ ), and *value*( $V$ ) for  $Q_\ell, K_\ell, V_\ell \in \mathbb{R}^{(N+1) \times d}$ . For each head, the computation process of the self-attention layer is defined as follows:

$$A(Q_\ell, K_\ell, V_\ell) = \text{softmax}\left(\frac{Q_\ell K_\ell^T}{\sqrt{d_k}}\right) V_\ell, \quad \ell = 1 \dots h. \quad (10)$$

where  $d_k$  is the dimension of  $K$ . We compute the self-attention scores for each head separately and concatenate them to obtain the final self-attention map. The following is the calculation process for the transformer encoder.

$$\begin{aligned} \text{MHSA}(Q, K, V) &= \text{Concat}(A(Q_1, K_1, V_1), \dots, \\ &\quad A(Q_h, K_h, V_h))W \\ O_M &= \text{MHSA}(Q, K, V) + E \\ O_{Encoder} &= \text{MLP}(\text{LN}(O_M)) + O_M, \end{aligned}$$

where  $W$  refers to the weights of the linear projection layer and  $\text{LN}(\cdot)$  stands for the layer normalization.

#### 2.3. Least Angle Regression for Adaptive Feature Selection

The excessive number of high-level features we learn by cls token in the transformer will affect the role of low level in quality assessment. We here use the least angle regression (LARS) method to select those influential high-level features and thus reduce the dimension of the high-level features [39].

By constructing a first-order penalty function, the LARS method determines that the coefficients of some variables are zero to delete some invalid variables and realize the role of dimensionality reduction [40]. In the LARS model, we assume normalized data  $\bar{X} \in \mathbb{R}^{N \times M}$  and the regression coefficient vector  $\theta \in \mathbb{R}^M$  satisfying the conditions of

$$y = \bar{X}\theta \quad (11)$$

$$\text{subject to } \sum_{j=1}^m |\theta_j| \leq t, \quad (12)$$

where  $N$  represents the number of samples and  $M$  represents the number of features, making  $y \in \mathbb{R}^n$ , and  $t$  is a positive constraint value. The linear regression model to be solved is as follows:

$$\min S(\theta) = (y - \bar{X}\theta)^T(y - \bar{X}\theta) + \lambda\|\theta\|_1. \tag{13}$$

The idea of the LARS model to solve Equation (13) is to initialize all variable values of  $\bar{X}$  to zero first and make predictions about  $y$  by finding the independent variable  $x_i \in \bar{X}$  with the highest correlation with  $y$  [41,42]. The LARS model is then solved by finding the independent variable  $x_i$  with the highest correlation with the current residuals. Then, find an independent variable  $x_j \in \bar{X}$  belonging along the direction of this independent variable with the highest correlation with the current residuals. Then, proceed along the angular bisector of  $x_i$  with  $x_j$  until a third predictor variable is found. Therefore, the directions of the three predictor variables are equiangular, and so on, iterating until all the required variables are obtained. In the process of travelling, the correlation between the residuals and the selected variables will gradually decrease so that the irrelevant variables can be eliminated, thus achieving dimensionality reduction. For this purpose, let  $c = \{c_1, c_2, \dots, c_M\}^T$  be the vector of the correlation calculated as:

$$c = \bar{X}^T(y - \bar{X}\theta), \tag{14}$$

where the difference between  $y$  and  $\bar{X}\theta$  is named the residual vector. Each element in  $c$  embodies a correlation between the corresponding column in  $\bar{X}$  and its residual vector.

Now, we assume that  $k$  iterations have been performed and that the resulting regression vector is  $u_k$ . The set of  $k$  variables is selected to form a set  $X_A$  ( $A \subset \bar{X}$ ) and use more relevant variables in  $\bar{X}$  to populate set  $X_A$  with the number of iterations. Each column in set  $X_A$  consists of the column corresponding to the position of the maximum value of  $c$  indexed in the  $\bar{X}$  matrix, i.e.,  $|c_{max}| = \max_i|c_i|$ . At the same time, we define the following relationships:

$$\begin{aligned} G_A &= X_A^T X_A \\ A_A &= \left( l_A^T G_A^{-1} l_A \right)^{-1/2}, \end{aligned} \tag{15}$$

where  $l_A$  is the vector of ones whose length is the number of elements in set  $A$ . Additionally, we define:

$$u_A = X_A A_A G_A^{-1} l_A, \tag{16}$$

where  $X_A = \left( \dots, s_j x_j, \dots \right)_{j \in A}$ ,  $s_j = \text{sign}\{c_j\}$ .

This method calculates  $a = \bar{X}^T u_A$  to determine the variables needed during the iteration. From there, the regression vector is updated:

$$\mu_k = \mu_{k-1} + \gamma_k u_A, \tag{17}$$

where  $\gamma_k$  can be calculated as

$$\gamma_k = \min_{j \in A^c} \left( \frac{C_{\max} - c_j}{A_A - a_j}, \frac{C_{\max} + c_j}{A_A + a_j} \right). \tag{18}$$

By choosing a smaller value of  $\gamma_k$ , we update subset  $X_A$  while calculating  $C'_{\max} = C_{\max} - \gamma_k A_A$ . The LARS model iterates until the values of  $c$  become negative or the desired number of variables is selected, thereby achieving dimensionality reduction.



#### 2.4. Quality Score Assessment

After obtaining the feature maps with LARS dimensionality reduction, we use SVR to map the feature maps to the final underwater image quality scores [8]. We want to learn a mapping that predicts the underwater quality score from the feature map  $x$ , defined as:

$$f(\mathbf{x}) = \langle \alpha, \mathbf{x} \rangle + b, \quad (19)$$

where  $\alpha$  and  $b$  indicate the model weights and bias, and  $\langle \cdot, \cdot \rangle$  is the inner product of the parameters. In the SVR model, we introduce two slack variables,  $\zeta_i$  and  $\zeta_i^*$ , to introduce soft intervals to construct the following optimization problem:

$$\begin{aligned} \min \quad & \frac{1}{2} \|\alpha\|^2 + \lambda \left( \sum_{i=1}^n \zeta_i + \sum_{i=1}^n \zeta_i^* \right) \\ \text{s.t.} \quad & \langle \alpha, \mathbf{x}_i \rangle + b - y_i \leq \varepsilon + \zeta_i \\ & y_i - \langle \alpha, \mathbf{x}_i \rangle - b \leq \varepsilon + \zeta_i^* \\ & \zeta_i, \zeta_i^* \geq 0, i = 1, 2, \dots, n \end{aligned} \quad (20)$$

where  $\lambda$  is a constant used to balance the accuracy of the model. In this model, our kernel function chosen is  $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$  to solve the above equation and finally obtain the predicted score.

### 3. Experiments Results And Discussion

#### 3.1. Dataset and Evaluation Protocols

In order to verify the effectiveness of our proposed UIQM-LSH, we conducted extensive validation experiments on the authentic underwater dataset, namely the underwater image quality assessment database (UWIQA). UWIQA contains 890 underwater images of different qualities [23] and is the largest public UIQA dataset at present. The labels, namely MOSs, are obtained by single-stimulus methodology, where each image is scored by a number of observers and the average value of these scores is considered the MOS. In Figure 4, we give example images of the UWIQA database. As can be seen, these images contain rich ocean features and a wide range of distortion types, thus ensuring a reliable evaluation of the proposed method.



**Figure 4.** Examples of underwater images in UWIQA database.

The IQA performance metrics commonly used for image quality evaluation are proposed by the Video Quality Experts Group (VQEG) [43], which recommends the use of the Spearman rank-order correlation coefficient (SRCC) and Kendall's rank-order correlation coefficient (KRCC) for the monotonicity of the prediction, Pearson's linear correlation coefficient (PLCC) for the accuracy of the prediction, and root mean square error (RMSE) for the consistency. These four metrics are used to assess the performance of the IQA method. Among them, SRCC and KRCC measure the degree to which the model's predicted scores vary with the mean opinion score (MOS). They are not affected by the magnitude of the change. SRCC and KRCC can be calculated as:

$$\begin{aligned}
 SRCC &= 1 - \frac{6 \sum_{n=1}^N d_n^2}{N(N^2 - 1)} \\
 KRCC &= \frac{N_c - N_d}{0.5 \times (N - 1) \times N}.
 \end{aligned}
 \tag{21}$$

Here,  $N$  denotes the total number of images in the dataset and  $d_n$  indicates the ranking difference between the subjective score (MOS) and the objective score for each pair.  $N_c$  and  $N_d$  denote the number of consistent and inconsistent images in the dataset. At the same time, PLCC and RMSE are used to measure the accuracy and consistency of the predicted scores by the model. PLCC measures the degree of linear correlation between the predicted and reference quality scores. In contrast, RMSE measures the average size of the difference between the predicted and reference quality scores. Both PLCC and RMSE are used to evaluate the performance of image quality assessment models. The PLCC and RMSE can be expressed as:

$$\begin{aligned}
 PLCC &= \frac{\sum_{n=1}^N (P_n - \hat{P})(M_n - \hat{M})}{\sqrt{\sum_{n=1}^N (P_n - \hat{P})^2 (M_n - \hat{M})^2}} \\
 RMSE &= \sqrt{\frac{1}{N} \sum (P_n - M_n)^2},
 \end{aligned}
 \tag{22}$$

where  $P_n$  and  $M_n$  denote the predicted objective score of the  $n$ -th picture and the subjective score of the MOS.  $\hat{P}$  and  $\hat{M}$  represent the mean values of the predicted objective score and the subjective MOS score, respectively. In the above metrics, the closer the values of SRCC, KRCC, and PLCC are to 1, and the closer the value of RMSE is to 0, the higher the performance of the model algorithm is indicated.

Before calculating these metrics, we need to map the objective prediction algorithm to the subjective MOS values using a nonlinear fitting function, which is formulated as:

$$O(P) = \beta_1 \left( \frac{1}{2} - \frac{1}{1 + \exp(\beta_2(P - \beta_3))} \right) + \beta_4 P + \beta_5,
 \tag{23}$$

where  $P$  and  $O(P)$  indicate the predicted objective and mapped scores,  $\beta_i$  for  $i = 1 \sim 5$  are the model parameters obtained in the fitting. The logistic function in Equation (23) is utilized to model the non-linearity inherent in the human subjective evaluation. It transforms the continuous predicted scores into values that better align with the categorical MOSs, thereby rendering to more reliable model evaluation.

### 3.2. Implementation Details

We implemented the proposed UIQM-LSH using Python and the Pytorch framework. In the transformer, each input image was randomly cropped into 10 image patches, each of which had the size of  $224 \times 224$ . In addition to random cropping, other data augmentation strategies, including flipping and rotation, were utilized. Each image patch was transformed into a number of patches with the size of 16. The number of transformer layers was set to 12, each of which contained 6 heads. We followed the standard train-test strategy, with 80% training data and testing 20% data. The training and testing data are not overlapped. The learning rate was set to  $10^{-4}$  with a decay factor of 10. The ADAM optimizer was utilized for training with a batch size of 16. The experiment was repeated 10 times, each with a different data splitting, in order to mitigate the performance bias. The averages of SRCC, KRCC, PLCC, and RMSE of the underwater image were reported.

### 3.3. Prediction Performance Evaluation

The proposed method UIQM-LSH is compared with 13 mainstream UIQA or IQA methods, including BRISQUE [44], NFERM [45], DBCNN [46], HyperIQA [35], TReS [47], PaQ-2-PiQ [48], NIQE [49], NPQI [50], SNP-NIQE [20], UCIQE [21], UIQM [22], CCF [51],



### 3.4. Cross-Dataset Validation

Having testified the IQA methods on each database, we further evaluate their prediction performance through the cross-database validation methodology. The cross-database validation tests the generalization ability of IQA metrics, which is essential to ensure the reliability and effectiveness of the models. In implementation, we initially train an IQA model for each method on one database and then test the trained models on the other database without any fine-tuning or parameter adaptation. Here, we utilize another database, namely UID2021 [14], which contains 960 underwater images in total. The experimental results in terms of SRCC, KRCC, PLCC, and RMSE on both databases are presented in Table 3.

**Table 3.** Cross-database validation, where the best performances are highlighted in bold.

IQA Metrics	UWIQA (Trained on UID2021)				UID2021 (Trained on UWIQA)			
	SRCC	KRCC	PLCC	RMSE	SRCC	KRCC	PLCC	RMSE
DBCNN [46]	0.4851	0.3492	4904	0.1208	0.6089	0.4706	0.6237	1.4198
HyperIQA [35]	0.5516	0.4287	0.5625	0.1209	0.6703	0.5263	0.6679	1.4132
TReS [47]	0.3459	0.2346	0.3673	0.1341	0.5224	0.4097	0.5321	1.5339
UIQM-LSH (Pro.)	<b>0.6307</b>	<b>0.5777</b>	<b>0.6340</b>	<b>0.1309</b>	<b>0.7189</b>	<b>0.6645</b>	<b>0.7197</b>	<b>1.3933</b>

### 3.5. Ablation Study

The proposed UIQM-LSH leverages both low-level and high-level perceptual properties to characterize the quality of underwater images. It is therefore necessary to examine their individual contributions to the model performance. To this end, we first conduct the ablation analysis to examine the effectiveness of the designed three modules, i.e., low-level feature extraction, high-level feature extraction, and the LARS module. Experiments were conducted on the UWIQA database, and the experimental settings are the same as those described in Section 3.2. The experimental results in terms of the average of SRCC, KRCC, PLCC, and RMSE are reported in Table 4.

In Table 4, we can obtain several meaningful observations. Firstly, either low-level features or high-level features can achieve a moderate performance, and their combinations deliver the best prediction. Hence, it is evident that both types of features are effective and contribute to the overall model performance. Secondly, we can find out that the high-level features achieve better prediction accuracy than the low-level features. This is reasonable since the number of dimensions of high-level features is far larger than the low-level features, i.e., 768 high-level features vs. 8 low-level features. Such a large dimension can capture a wide variety of aspects of an image, such as the details about the objects presence, the scene context, and the interactions between elements. As a result, it can offer a richness and complexity of information that aligns closely with the multifaceted nature of the underwater image quality. However, the large number of high-level features may also result in the model being biased. As can be seen, the improvement of introducing the low-level features into original high-level features is much less than that into dimension reduced of high-level features. Thirdly, the LARS method can effectively promote the model performance further. It optimizes the high-level features, presenting a more balancing and effective feature fusion. The LARS method reduces redundancy and noise by selecting the most relevant high-level features and enhances the model's generalization capabilities. More importantly, it leads to a more interpretable and computationally efficient model.

**Table 4.** Performance evaluation of different types of features on UWIQA database. The high-level features are processed without or with the LARS method. The best performances are highlighted in bold.

Feature Types	LARS	UWIQA			
		SRCC	KRCC	PLCC	RMSE
Low-level perceptual features	✗	0.6748	0.5272	0.6383	0.1066
High-level perceptual features		0.6944	0.5916	0.6964	0.0928
Combination		0.7183	0.6237	0.7060	0.0915
Low-level perceptual features	✓	0.6748	0.5272	0.6383	0.1066
High-level perceptual features		0.7246	0.6299	0.7673	0.0801
Combination		<b>0.7580</b>	<b>0.6528</b>	<b>0.7890</b>	<b>0.0790</b>

The experiments in Table 4 were conducted from the module level. However, by introducing the low-level features and the LARS method, the proposed UIQM-LSH is more interpretable than the common black-box DL methods. We further analyze UIQM-LSH from the feature level. In particular, we ablate each low-level feature, i.e., brightness, colorfulness, and contrast, to examine their individual contribution. The experimental results are reported in Table 5. As can be observed, each type of feature is effective in characterizing the quality of underwater images. The combination of any two out of the three feature types presents a better performance than the single feature type. This finding indicates that the different types of features complement each other. By in-depth analysis of Table 5, the image contrast contributes most to the model performance. One of the main reasons is that the difference in either image brightness or colorfulness can impact the image contrast as well. Higher contrast enhances the visibility of objects and details, making images more visually appealing and informative, especially in situations where light is limited and scattered in water. On the other hand, the colorfulness feature may be less effective than contrast and brightness in estimating the quality of underwater images due to the limited color diversity in underwater environments. Though the color cast may impact the human subjective evaluation on the underwater image, a relatively higher quality score may be still assigned when the color-casted image is visually clear.

**Table 5.** Performance evaluation of low-level perceptual features on UWIQA database. The best performances are highlighted in bold.

Feature Types	UWIQA			
	SRCC	KRCC	PLCC	RMSE
Colorfulness	0.4192	0.3201	0.4396	0.1244
Brightness	0.5028	0.3777	0.5200	0.1266
Contrast	0.6325	0.5552	0.6172	0.1165
Colorfulness + Contrast	0.6418	0.5035	0.6196	0.1090
Colorfulness + Brightness	0.5859	0.4773	0.5936	0.1108
Contrast + Brightness	0.6578	0.5414	0.6198	0.1157
Contrast + Brightness + Colorfulness	<b>0.6748</b>	<b>0.5272</b>	<b>0.6383</b>	<b>0.1066</b>

#### 4. Conclusions

This paper deals with the challenging issues in the field of underwater quality assessment and proposes a novel method, namely underwater image quality metric with low-level properties and selected high-level semantics (UIQM-LSH). This approach bridges the gap in current UIQA methodologies by integrating both low-level image attributes and high-level semantic content. By implementing the least-angle regression method for feature selection, our model ensures a balanced representation of these diverse features, thus enhancing the accuracy and reliability of image quality assessments. The superior performance of UIQM-LSH over existing UIQA methods, as demonstrated through extensive

testing, highlights its potential as a more effective tool for various applications in oceanic engineering, marine biology, and underwater robotics. This work exhibits promising results to significantly enhance the understanding and exploration of underwater environments.

**Author Contributions:** Conceptualization, R.H. and B.Q.; methodology, F.L., Z.H., and T.X.; formal analysis, B.Q.; data curation, R.H. and B.Q.; writing—original draft preparation, F.L.; writing—review and editing, F.L. and R.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Key Research and Development Program of China (Grant No. 2022YFC2808002).

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ancuti, C.O.; Ancuti, C.; De Vleeschouwer, C.; Bekaert, P. Color balance and fusion for underwater image enhancement. *IEEE Trans. Image Process.* **2017**, *27*, 379–393. [[CrossRef](#)]
2. Gulliver, S.R.; Ghinea, G. Defining User Perception of Distributed Multimedia Quality. *ACM Trans. Multimed. Comput. Commun. Appl.* **2006**, *2*, 241–257. [[CrossRef](#)]
3. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
4. Zhang, L.; Shen, Y.; Li, H. VSI: A visual saliency-induced index for perceptual image quality assessment. *Image Process. IEEE Trans.* **2014**, *23*, 4270–4281. [[CrossRef](#)] [[PubMed](#)]
5. Soundararajan, R.; Bovik, A.C. RRED indices: Reduced reference entropic differencing for image quality assessment. *IEEE Trans. Image Process.* **2012**, *21*, 517–526. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, L.; Zhang, L.; Mou, X.; Zhang, D. FSIM: A feature similarity index for image quality assessment. *IEEE Trans. Image Process.* **2011**, *20*, 2378–2386. [[CrossRef](#)] [[PubMed](#)]
7. Zhang, Y.; Xie, Y.; Li, C.; Wu, Z.; Qu, Y. Learning All-In Collaborative Multiview Binary Representation for Clustering. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**. [[CrossRef](#)]
8. Hu, R.; Liu, Y.; Gu, K.; Min, X.; Zhai, G. Toward a no-reference quality metric for camera-captured images. *IEEE Trans. Cybern.* **2021**, *53*, 3651–3664. [[CrossRef](#)] [[PubMed](#)]
9. Ma, X.; Zhang, S.; Wang, Y.; Li, R.; Chen, X.; Yu, D. ASCAM-Former: Blind image quality assessment based on adaptive spatial & channel attention merging transformer and image to patch weights sharing. *Expert Syst. Appl.* **2023**, *215*, 119268.
10. Nizami, I.F.; Rehman, M.U.; Waqar, A.; Majid, M. Impact of visual saliency on multi-distorted blind image quality assessment using deep neural architecture. *Multimed. Tools Appl.* **2022**, *81*, 25283–25300. [[CrossRef](#)]
11. ur Rehman, M.; Nizami, I.F.; Majid, M. DeepRPN-BIQA: Deep architectures with region proposal network for natural-scene and screen-content blind image quality assessment. *Displays* **2022**, *71*, 102101. [[CrossRef](#)]
12. Sendjasni, A.; Larabi, M.C. PW-360IQA: Perceptually-Weighted Multichannel CNN for Blind 360-Degree Image Quality Assessment. *Sensors* **2023**, *23*, 4242. [[CrossRef](#)] [[PubMed](#)]
13. Zhou, W.; Chen, Z. Deep local and global spatiotemporal feature aggregation for blind video quality assessment. In Proceedings of the 2020 IEEE International Conference on Visual Communications and Image Processing (VCIP), Macau, China, 1–4 December 2020; pp. 338–341.
14. Hou, G.; Li, Y.; Yang, H.; Li, K.; Pan, Z. UID2021: An underwater image dataset for evaluation of no-reference quality assessment metrics. *ACM Trans. Multimed. Comput. Commun. Appl.* **2023**, *19*, 1–24. [[CrossRef](#)]
15. Nizami, I.F.; Rehman, M.u.; Majid, M.; Anwar, S.M. Natural scene statistics model independent no-reference image quality assessment using patch based discrete cosine transform. *Multimed. Tools Appl.* **2020**, *79*, 26285–26304. [[CrossRef](#)]
16. Bouris, E.; Davis, T.; Morales, E.; Grassi, L.; Salazar Vega, D.; Caprioli, J. A Neural Network for Automated Image Quality Assessment of Optic Disc Photographs. *J. Clin. Med.* **2023**, *12*, 1217. [[CrossRef](#)]
17. Nizami, I.F.; Majid, M.; Rehman, M.u.; Anwar, S.M.; Nasim, A.; Khurshid, K. No-reference image quality assessment using bag-of-features with feature selection. *Multimed. Tools Appl.* **2020**, *79*, 7811–7836. [[CrossRef](#)]
18. Zhou, W.; Wang, Z. Blind Omnidirectional Image Quality Assessment: Integrating Local Statistics and Global Semantics. *arXiv* **2023**, arXiv:2302.12393.
19. Jiang, Q.; Zhou, W.; Chai, X.; Yue, G.; Shao, F.; Chen, Z. A full-reference stereoscopic image quality measurement via hierarchical deep feature degradation fusion. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 9784–9796. [[CrossRef](#)]
20. Liu, Y.; Gu, K.; Zhang, Y.; Li, X.; Zhai, G.; Zhao, D.; Gao, W. Unsupervised blind image quality evaluation via statistical measurements of structure, naturalness, and perception. *IEEE Trans. Circuits Syst. Video Technol.* **2019**, *30*, 929–943. [[CrossRef](#)]
21. Yang, M.; Sowmya, A. An underwater color image quality evaluation metric. *IEEE Trans. Image Process.* **2015**, *24*, 6062–6071. [[CrossRef](#)]

22. Panetta, K.; Gao, C.; Agaian, S. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* **2015**, *41*, 541–551. [CrossRef]
23. Yang, N.; Zhong, Q.; Li, K.; Cong, R.; Zhao, Y.; Kwong, S. A reference-free underwater image quality assessment metric in frequency domain. *Signal Process. Image Commun.* **2021**, *94*, 116218. [CrossRef]
24. Zhang, Y.; Qu, Y.; Xie, Y.; Li, Z.; Zheng, S.; Li, C. Perturbed self-distillation: Weakly supervised large-scale point cloud semantic segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 11–17 October 2021; pp. 15520–15528.
25. Hu, R.; Liu, Y.; Wang, Z.; Li, X. Blind quality assessment of night-time image. *Displays* **2021**, *69*, 102045. [CrossRef]
26. He, C.; Li, K.; Zhang, Y.; Tang, L.; Zhang, Y.; Guo, Z.; Li, X. Camouflaged object detection with feature decomposition and edge reconstruction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 18–22 June 2023; pp. 22046–22055.
27. Hu, R.; Monebhurrin, V.; Himeno, R.; Yokota, H.; Costen, F. A general framework for building surrogate models for uncertainty quantification in computational electromagnetics. *IEEE Trans. Antennas Propag.* **2021**, *70*, 1402–1414. [CrossRef]
28. Hu, R.; Monebhurrin, V.; Himeno, R.; Yokota, H.; Costen, F. An adaptive least angle regression method for uncertainty quantification in FDTD computation. *IEEE Trans. Antennas Propag.* **2018**, *66*, 7188–7197. [CrossRef]
29. Lin, S.; Ji, R.; Yan, C.; Zhang, B.; Cao, L.; Ye, Q.; Huang, F.; Doermann, D. Towards optimal structured cnn pruning via generative adversarial learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 2790–2799.
30. Lin, S.; Ji, R.; Chen, C.; Tao, D.; Luo, J. Holistic cnn compression via low-rank decomposition with knowledge transfer. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *41*, 2889–2905. [CrossRef] [PubMed]
31. Lin, S.; Ji, R.; Li, Y.; Wu, Y.; Huang, F.; Zhang, B. Accelerating Convolutional Networks via Global & Dynamic Filter Pruning. In Proceedings of the IJCAI. Stockholm, Stockholm, Sweden, 13–19 July 2018; Volume 2, p. 8.
32. Li, F.; Wu, J.; Wang, Y.; Zhao, Y.; Zhang, X. A color cast detection algorithm of robust performance. In Proceedings of the 2012 IEEE Fifth International Conference on Advanced Computational Intelligence (ICACI), Nanjing, China, 18–20 October 2012; pp. 662–664.
33. Choi, L.K.; You, J.; Bovik, A.C. Referenceless Prediction of Perceptual Fog Density and Perceptual Image Defogging. *IEEE Trans. Image Process.* **2015**, *24*, 3888–3901. [CrossRef]
34. Fu, H.; Liu, G.; Yang, X.; Wei, L.; Yang, L. Two Low-Level Feature Distributions Based No Reference Image Quality Assessment. *Appl. Sci.* **2022**, *12*, 4975. [CrossRef]
35. Su, S.; Yan, Q.; Zhu, Y.; Zhang, C.; Ge, X.; Sun, J.; Zhang, Y. Blindly assess image quality in the wild guided by a self-adaptive hyper network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 3667–3676.
36. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
37. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.
38. Qin, G.; Hu, R.; Liu, Y.; Zheng, X.; Liu, H.; Li, X.; Zhang, Y. Data-Efficient Image Quality Assessment with Attention-Panel Decoder. *arXiv* **2023**, arXiv:2304.04952.
39. Hu, R.; Monebhurrin, V.; Himeno, R.; Yokota, H.; Costen, F. A statistical parsimony method for uncertainty quantification of FDTD computation based on the PCA and ridge regression. *IEEE Trans. Antennas Propag.* **2019**, *67*, 4726–4737. [CrossRef]
40. Efron, B.; Hastie, T.; Johnstone, I.; Tibshirani, R. Least angle regression. *Ann. Statist.* **2004**, *32*, 407–499. [CrossRef]
41. Márquez-Vera, M.A.; Ramos-Velasco, L.; López-Ortega, O.; Zúñiga-Peña, N.; Ramos-Fernández, J.; Ortega-Mendoza, R.M. Inverse fuzzy fault model for fault detection and isolation with least angle regression for variable selection. *Comput. Ind. Eng.* **2021**, *159*, 107499. [CrossRef]
42. Blatman, G.; Sudret, B. Adaptive sparse polynomial chaos expansion based on least angle regression. *J. Comput. Phys.* **2011**, *230*, 2345–2367. [CrossRef]
43. Video Quality Experts Group. Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment. *ITU-T Standards Contribution COM 2000*. pp. 9–80. Available online: [https://www.academia.edu/2102517/FINAL\\_REPORT\\_FROM\\_THE\\_VIDEO\\_QUALITY\\_EXPERTS\\_GROUP\\_ON\\_THE\\_VALIDATION\\_OF\\_OBJECTIVE\\_MODELS\\_OF\\_VIDEO\\_QUALITY\\_ASSESSMENT\\_March](https://www.academia.edu/2102517/FINAL_REPORT_FROM_THE_VIDEO_QUALITY_EXPERTS_GROUP_ON_THE_VALIDATION_OF_OBJECTIVE_MODELS_OF_VIDEO_QUALITY_ASSESSMENT_March) (accessed on 13 November 2023).
44. Mittal, A.; Moorthy, A.K.; Bovik, A.C. No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.* **2012**, *21*, 4695–4708. [CrossRef] [PubMed]
45. Gu, K.; Zhai, G.; Yang, X.; Zhang, W. Using free energy principle for blind image quality assessment. *IEEE Trans. Multimed.* **2014**, *17*, 50–63. [CrossRef]
46. Zhang, W.; Ma, K.; Yan, J.; Deng, D.; Wang, Z. Blind image quality assessment using a deep bilinear convolutional neural network. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *30*, 36–47. [CrossRef]

47. Golestaneh, S.A.; Dadsetan, S.; Kitani, K.M. No-reference image quality assessment via transformers, relative ranking, and self-consistency. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2022; pp. 1220–1230.
48. Ye, P.; Kumar, J.; Kang, L.; Doermann, D. Unsupervised feature learning framework for no-reference image quality assessment. In Proceedings of the 2012 IEEE conference on computer vision and pattern recognition, Providence, RI, USA, 16–21 June 2012; pp. 1098–1105.
49. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **2013**, *20*, 209–212. [[CrossRef](#)]
50. Liu, Y.; Gu, K.; Li, X.; Zhang, Y. Blind image quality assessment by natural scene statistics and perceptual characteristics. *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)* **2020**, *16*, 1–91. [[CrossRef](#)]
51. Wang, Y.; Li, N.; Li, Z.; Gu, Z.; Zheng, H.; Zheng, B.; Sun, M. An imaging-inspired no-reference underwater color image quality assessment metric. *Comput. Electr. Eng.* **2018**, *70*, 904–913. [[CrossRef](#)]
52. Sun, S.; Yu, T.; Xu, J.; Lin, J.; Zhou, W.; Chen, Z. GraphIQA: Learning Distortion Graph Representations for Blind Image Quality Assessment. *IEEE Trans. Multimed.* **2021**, *25*, 2912–2925. [[CrossRef](#)]
53. Gu, K.; Zhai, G.; Lin, W.; Yang, X.; Zhang, W. Learning a blind quality evaluation engine of screen content images. *Neurocomputing* **2016**, *196*, 140–149. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.