Article

No-Reference Image Quality Assessment Using the Statistics of Global and Local Image Features

Domonkos Varga (ID)

Ronin Institute, Montclair, NJ 07043, USA; domonkos.varga@ronininstitute.org

Abstract: Methods of image quality assessment are widely used for ranking computer vision algorithms or controlling the perceptual quality of video and streaming applications. The ever-increasing number of digital images has encouraged the research in this field at an accelerated pace in recent decades. After the appearance of convolutional neural networks, many researchers have paid attention to different deep architectures to devise no-reference image quality assessment algorithms. However, many systems still rely on handcrafted features to ensure interpretability and restrict the consumption of resources. In this study, our efforts are focused on creating a quality-aware feature vector containing information about both global and local image features. Specifically, the research results of visual physiology indicate that the human visual system first quickly and automatically creates a global perception before gradually focusing on certain local areas to judge the quality of an image. Specifically, a broad spectrum of statistics extracted from global and local image features is utilized to represent the quality-aware aspects of a digital image from various points of view. The experimental results demonstrate that our method's predicted quality ratings relate strongly with the subjective quality ratings. In particular, the introduced algorithm was compared with 16 other well-known advanced methods and outperformed them by a large margin on 9 accepted benchmark datasets in the literature: CLIVE, KonIQ-10k, SPAQ, BIQ2021, TID2008, TID2013, MDID, KADID-10k, and GFIQA-20k, which are considered de facto standards and generally accepted in image quality assessment.

Keywords: no-reference image quality assessment; quality-aware features; multi-feature fusion

1. Introduction

Objective image quality assessment (IQA) aims at devising mathematical and computational models which can predict digital images' perceptual quality consistently with human judgment. This field is traditionally divided into three distinct areas in relation to the availability of reference (distortion-free) images for the IQA algorithms [1]. Specifically, full-reference methods possess all information about the reference images, while reduced-reference ones have some partial knowledge about them. On the other hand, no-reference IQA (NR-IQA) algorithms have only the distorted images in their input. In the literature, this third type of IQA is considered the most difficult one, since no information is available about the distortion-free counterpart of a distorted image whose quality has to be estimated [2]. Moreover, NR-IQA is also considered the most useful branch of IQA because access to reference images is impossible in many practical applications [3].

According to the research results of visual physiology, the human visual system (HVS) first quickly and automatically creates a global perception before gradually focusing on certain local areas to judge the quality of an image [4–7]. Therefore, we need to take both local and global information into account in order to more accurately imitate how the HVS perceives image quality. Based on the above observation, this study proposes a novel method by integrating the statistics of the global and local image features for NR-IQA. Unlike previous methods [5,7–9], the selected local and global features are not simply fused together or concatenated, since this process quickly results in a long feature

vector. Instead, a broad spectrum of statistics is extracted from them to represent the quality-aware aspects of a digital image from various points of view in a compact feature vector. Experimental results for nine benchmark IQA databases (CLIVE [10], KonIQ-10k [11], SPAQ [12], BIQ2021 [13], TID2008 [14], MDID [15], TID2013 [16], KADID-10k [17], and GFIQA-20k [18]) show that the proposed method is consistent with the subjective ratings, able to surpass the state of the art's efficiency, and balances effectiveness and computation time.

The rest of this study is organized as follows. Section 2 describes previous studies and papers on NR-IQA briefly. Section 3 provides the details of the proposed model. In Section 4, our experimental results are presented and analyzed. Finally, a conclusion is drawn and future works are stated in Section 5.

2. Literature Review

As already remarked, IQA methods are organized in the literature based on how much information is available about the reference images [19]. In the case of NR-IQA, no information about the reference images can be obtained.

Many NR-IQA methods rely on natural scene statistics (NSS) features. NSS builds on the assumption that natural scenes have some kind of statistical regularities which are influenced by noise or distortions [20,21]. For example, BLIINDS [22] and its later extension BLIINDS-II [23] decompose a digital image into the discrete cosine transform (DCT) domain. Next, the authors fit a generalized Gaussian distribution (GGD) on the extracted DCT coefficients. Moreover, the authors chose the parameters of the GGD as quality-aware features and used them in a regression process to estimate perceptual quality. Later, this method was extended to the domain of video quality assessment [23,24]. In contrast, BRISQUE [25] extracts NSS in the spatial domain. First, local mean subtraction and divisive normalization are applied to the image. Next, the NSS are extracted, which results in the mean subtracted contrast-normalized (MSCN) coefficients. A GGD model is fitted onto the MSCN coefficients, and its parameters are used as quality-aware features. Finally, the image quality is estimated with the help of a trained support vector regressor (SVR). BRISQUE [25] was developed further in [26] by introducing L moment-based robust parameter estimators [27]. In contrast, Bagade et al. [28] introduced the shape-adaptive discrete wavelet transform to capture quality-aware features more effectively and trained a neural network based on them.

In recent years, convolutional neural networks (CNNs) have been proven to be capable of learning image representation effectively and have provided excellent results in a number of computer vision tasks [29]. After the appearance of CNNs, many researchers have paid attention to different deep learning techniques for NR-IQA. However, many systems still rely on handcrafted features to ensure interpretability and restrict the consumption of resources [30,31].

The method of Kang et al. [32] was among the first ones which utilized a CNN for NR-IQA. Specifically, the authors decomposed training images into a set of 32×32 local contrast-normalized image patches. Based on them, a CNN with a depth of five and a linear regression layer was trained to predict the perceptual quality of image patches. Finally, the authors considered the arithmetic average of the image patches' quality ratings as the whole input image's estimated quality. This approach was developed further by the authors of [33] by applying 64×64 image patches and a CNN with a depth of seven and three fully connected layers. Similar to [32] and [33], Kim and Lee [34] presented a solution based on a CNN trained on image patches. The important difference compared with [32] and [33] was that the patches' target qualities were determined by applying a traditional FR-IQA metric [35] instead of simply taking the original image's quality. Due to the usage of a traditional FR-IQA method, this approach can be applied on images with artificial distortions and not authentic ones. In [36], the authors improved this approach by applying a nonlinear bilateral smoothing filtering technique and a nearest neighbor sampling approach [37]. Prior to deep feature extraction, Ryu [38] introduced a static

saliency detection module first to identify those regions which humans tend to pay more attention to. In contrast, Celona and Schettini [39] devised a deep architecture which handles images at different scales. Moreover, it was trained simultaneously by formulating NR-IQA as regression, classification, and pairwise ranking. Ryu [40] applied the ImageNet pretrained Inception-ResNet-V2 [41] CNN and SpinalNet [42] for perceptual quality estimation. Specifically, the image patches were introduced step by step and over regular intervals to this architecture. Aside from CNNs, some researchers have explored the usage of transformers [43] in NR-IQA [44]. For example, Ke et al. [45] introduced a multi-scale transformer-based architecture to manage digital images with different resolutions. The descriptions of other transformer-based methods can be found in [46–48].

For a general analysis of the current state of NR-IQA, we recommend the PhD dissertations of Jenadeleh [49] and Men [50] and the book by Xu et al. [51]

3. Proposed Method

A broad overview of the proposed NR-IQA method is depicted in Figure 1. As this illustration suggests, the statistics of the global and local image features are extracted from the images in the training database to train a machine learning model which is applied afterward in the testing stage. The process of feature extraction is shown in Figure 2. As already indicated, we apply a broad spectrum of statistics extracted from the different global and local features of an image. In this way, a more compact feature vector can be obtained rather than fusing global and local features together directly as in [7].
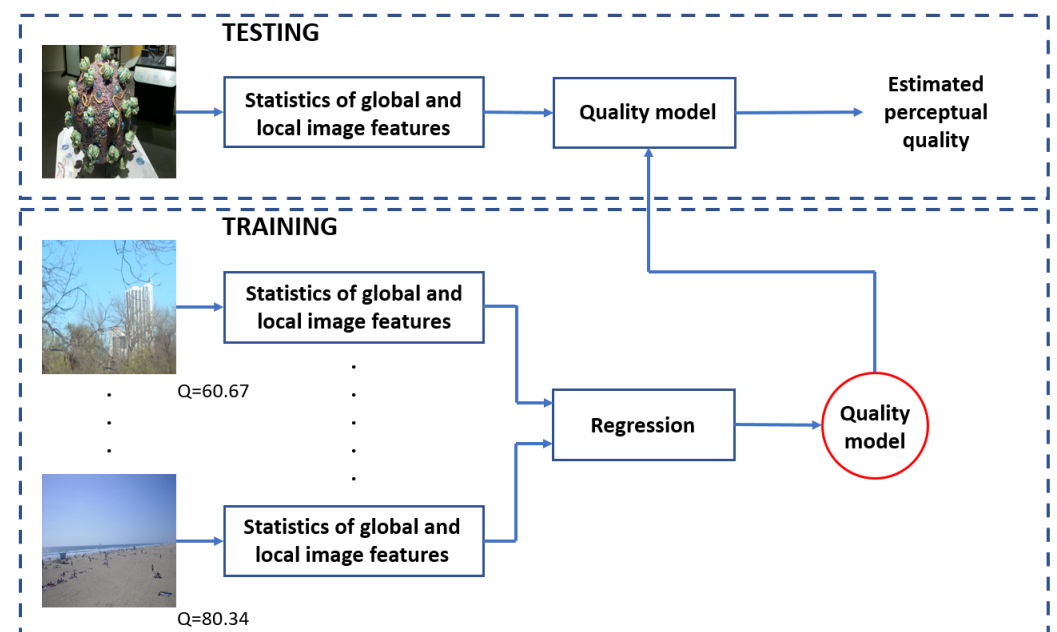

Figure 1. High-level flowchart of our proposed method which can be divided into training and test steps. In the training phase, the statistics of global and local image features extracted from the labeled training images are used to train the quality model, which is later used in the testing phase to assess previously unseen images.
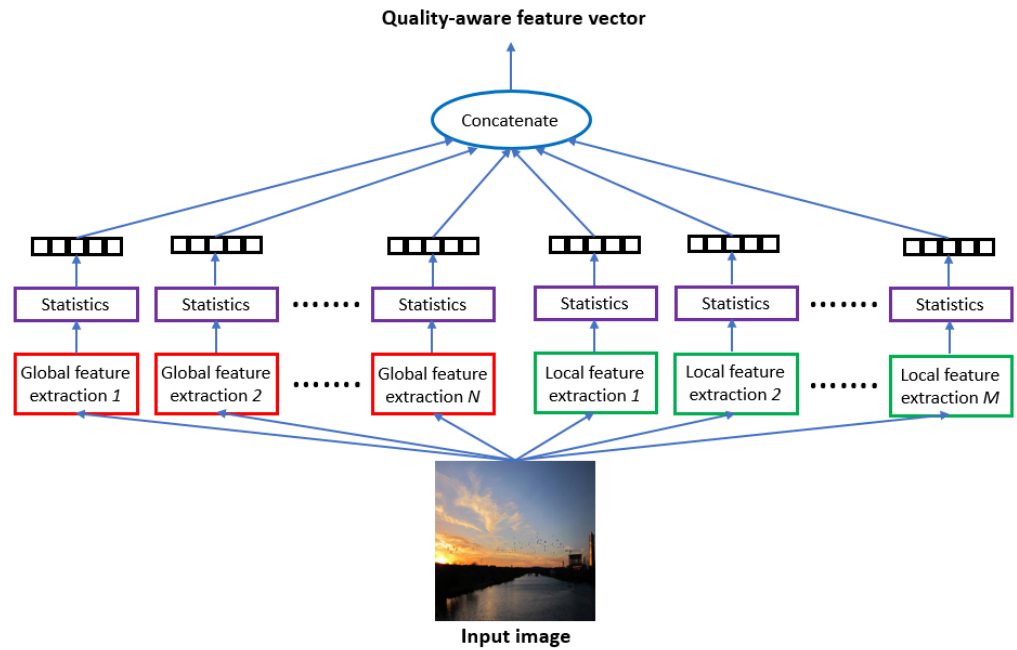
Figure 2. Process of feature extraction. To characterize an image from various points of view, different global and local feature vectors are extracted first. Next, statistics are extracted from them and fused together to avoid a long quality-aware feature vector.

3.1. Statistics of Global Image Features

In recent decades, many quality-aware feature vectors have been proposed which describe digital images globally. In this work, we chose seven different ones: BRISQUE [25], CurveletQA [52], GM-LOG-BIQA [53], GWH-GLBP [54], HIGRADE [55], OG-IQA [56], and SSEQ [57]. The main motivation was to chose feature vectors which had low computational complexity and characterized images from different points of view. Specifically, BRISQUE [25] derives features from the spatial domain, relying on mean subtracted normalized coefficients. CurveletQA [52] considers log histograms of the curvelet coefficients and the energy distributions of orientation and scale in the curvelet domain [58]. In contrast, GM-LOG-BIQA [53] and GWH-GLBP [54] characterize image quality by texture degradation. In particular, GM-LOG-BIQA [53] utilizes the normalized joint statistics of the gradient and Laplacian features, while GWH-GLBP [54] applies the local binary pattern [59] operator to quantify texture degradation. HIGRADE [55] describes log-derivative, spatial domain, and gradient domain scene statistics by fitting GGDs onto them. Next, the parameters of the GGDs were considered to be quality-aware features. OG-IQA [56] considers the histogram variances of the gradient, relative orientation, and relative gradient maps to be quality-aware features. SSEQ [57] uses the spatial entropies of block DCT coefficients. Aside from these global quality-aware features, the histogram of oriented gradients [60] of the input image's Y, Cb, and Cr color channels was also determined, since gradient orientations are seriously distorted in the presence of noise [61]. As illustrated in Figure 2, certain statistics are extracted from each global and local feature to compile a new, compact, quality-aware feature. Specifically, we considered the mean, the median, the standard deviation, the skewness, and the kurtosis of each feature. Skewness characterizes the asymmetry of the data around the mean. The skewness of a vector **x** can be determined as follows:

$$s = \frac{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^3}{\left(\sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} \right)^3}, \tag{1}$$

where x_i is the ith element of $\mathbf{x}$ and $\bar{x}$ denotes the mean of $\mathbf{x}$. Kurtosis characterizes how outlier-prone a distribution is and can be determined in the case of a vector $\mathbf{x}$ as follows:

$$k = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2}. \tag{2}$$

As a result, a vector $10 \times 5 = 50$ in length can be obtained from the statistics of the global features.

3.2. Statistics of Local Image Features

Motivated by our previous work [62], the statistics of local feature descriptors were used as quality-aware features. Specifically, several HVS-inspired filters (Bilaplacian, high-boost, and derivative) were applied first to the color channels (*YCbCr* color space) of the input image. Next, the features from the accelerated segment test (FAST) keypoints [63] were detected on all resulting feature maps. Moreover, each detected keypoint was described by a vector wiht a length of 25 containing the values of its 5×5 neighborhood. As already mentioned, *YCbCr*, which is defined by ITU-R BT.601 [64], is utilized in this paper because it is the preferred format for video broadcasting, with its efficient use of the channel bandwidth. Furthermore, the luma and chroma components are separate in *YCbCr* [65]. For conversion from to the *RGB* to *YCbCr* color space, the following matrix equation was used [66]:

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.2568 & 0.5041 & 0.0979 \\ -0.1482 & -0.2910 & 0.4392 \\ 0.4392 & -0.3678 & -0.0714 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{3}$$

The main motivation of Bilaplacian filters comes from the work of Ghosh et al. [67,67], who pointed out that the nature of retinal ganglion cells' extended classical receptive field can be described by a combination of three zero-mean Gaussians at three different scales, which corresponds to the Bilaplacian of the Gaussian filter. In this study, the following Laplacian kernels were applied:

$$\mathbf{L}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \ \mathbf{L}_2 = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}, \ \mathbf{L}_3 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -4 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \tag{4}$$

$$\mathbf{L}_4 = \begin{pmatrix} -2 & 1 & -2 \\ 1 & 4 & 1 \\ -2 & 1 & -2 \end{pmatrix}, \ \mathbf{L}_5 = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}. \tag{5}$$

As the terminology suggests, the convolution of two Laplacian kernels results in a Bilaplacian kernel:

$$\mathbf{L}_{ij}^2 = \mathbf{L}_i * \mathbf{L}_j. \tag{6}$$

In this paper, the $\mathbf{L}_{11}^2$, $\mathbf{L}_{22}^2$, $\mathbf{L}_{33}^2$, $\mathbf{L}_{44}^2$, $\mathbf{L}_{55}^2$, $\mathbf{L}_{13}^2$, and $\mathbf{L}_{24}^2$ kernels were used.

HVS is sensitive to the high-frequency regions of a natural scene [68], and a high-boost filter can be used to enhance the high-frequency components [69]. In this study, the following high-boost filter kernel was applied:

$$\mathbf{H} = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix}. \tag{7}$$

Aside from the filter given by Equation (7), $\mathbf{H}^2$, $\mathbf{H}^3$, and $\mathbf{H}^4$ were also applied, since image noise or distortion can occur at different scales.

First, Li et al. [54] applied derivative filters to extract statistical regularities from a natural scene for NR-IQA. In our study, we apply the following derivative kernel:

$$\mathbf{D}_1 = \begin{pmatrix} -1 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & -1 \end{pmatrix} * \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 1 \end{pmatrix}. \tag{8}$$

Moreover, $\mathbf{D}_2$, $\mathbf{D}_3$, $\mathbf{D}_4$, and $\mathbf{D}_5$ with sizes of 5×5, 7×7, and 11×11 were also used because image noise or distortion can occur at different scales.

To sum up, a $3 \times 7 \times 50 \times 5 = 5250$ dimensional vector from the Bilaplacian maps, $3 \times 4 \times 50 \times 5 = 3000$ dimensional vector from the high-boost maps, and $3 \times 5 \times 50 \times 5 = 3750$ dimensional vector from the derivative maps can be extracted. As in the case of global features, the mean, median, standard deviaton, skewness, and kurtosis were extracted from the local features, which resulted in a $3 \times 5 = 15$ dimensional quality-aware feature. When considering the statistics of global and local features together, a $50 + 15 = 65$ dimensional quality-aware feature can be compiled.

4. Results

This section introduces our obtained experimental results and their analysis. First, the applied benchmark IQA databases are introduced in Section 4.1. Afterward, in Section 4.2, the assessment conventions and evaluation metrics are specified. To explain the design choices of the proposed method, an ablation study is presented in Section 4.3. Finally, a comparison to the state of the art is presented in Section 4.4.

4.1. Datasets

To empirically corroborate the efficiency of the proposed method, nine publicly available benchmark databases which contain digital images of a quality annotated by human observers, were applied in our experiments. Since quality ratings obtained from humans may not directly express the visual quality due to other image characteristics (i.e., aesthetics) [70], a profound evaluation of NR-IQA algorithms must involve tests on multiple benchmark databases [71,72]:

- CLIVE [10]: This database contains 1162 individual high-resolution scenes with authentic distortions. Specifically, the images were captured by everyday users with different types of mobile devices. Furthermore, quality annotations were obtained from 8100 observers under a crowd-sourcing experiment.
- KonIQ-10k [11]: Similar to CLIVE [10], this database contains individual scenes exhibiting authentic distortions. The images were collected from a very large multimedia database [73] and assessed in a crowd-sourcing experiment. On the whole 10,073 digital images can be found in this database.
- SPAQ [12]: This database contains 11,125 individual high-resolution scenes captured by 66 smartphones. Furthermore, the images were evaluated under a laboratory environment by 600 observers.
- BIQ2021 [13]: This database contains 8000 images with authentic distortions collected by various imaging devices. Furthermore, the images were evaluated in a laboratory environment, calculating image quality scores obtained from at least 20 human observers.
- TID2008 [14]: This database contains 1700 distorted images which were generated from 25 reference images using 17 distortion types at 4 degradation levels.
- MDID [15]: This database contains 1600 distorted images which were produced from 20 reference images. From each reference image, 80 distorted images were generated by applying 4 different distortion types, such as Gaussian blur, a change in contrast, compression noise, and Gaussian noise. Furthermore, each distortion level was randomly chosen to be one of the four levels.
- TID2013 [16]: This database contains 3000 distorted images which were generated from 25 reference images, applying 25 different noise types at 5 different levels of degradation.

Quality annotations were obtained from approximately 524,000 ratings under a laboratory environment.

- KADID-10k [17]: This database contains 10,125 distorted images which were generated from 81 reference images using 25 distortion types (i.e., blurs, compression noise, color distortions, changes in brightness, changes in contrast, sharpening, and spatial distortions) at 5 different distortion levels. The authors collected approximately 30,400 quality ratings in a crowd-sourcing experiment [74] for the annotation of the distorted images.

- GFIQA-20k [18]: This database consists of 20,000 512×512 quality annotated face images of individuals sampled from the YFCC100M [73] multimedia database.

Table 1 provides an overview of the main attributes of the utilized benchmark datasets. In addition to this, the quality ratings' empirical distributions are depicted in Figures 3 and 4.

(a)

(b)

(c)

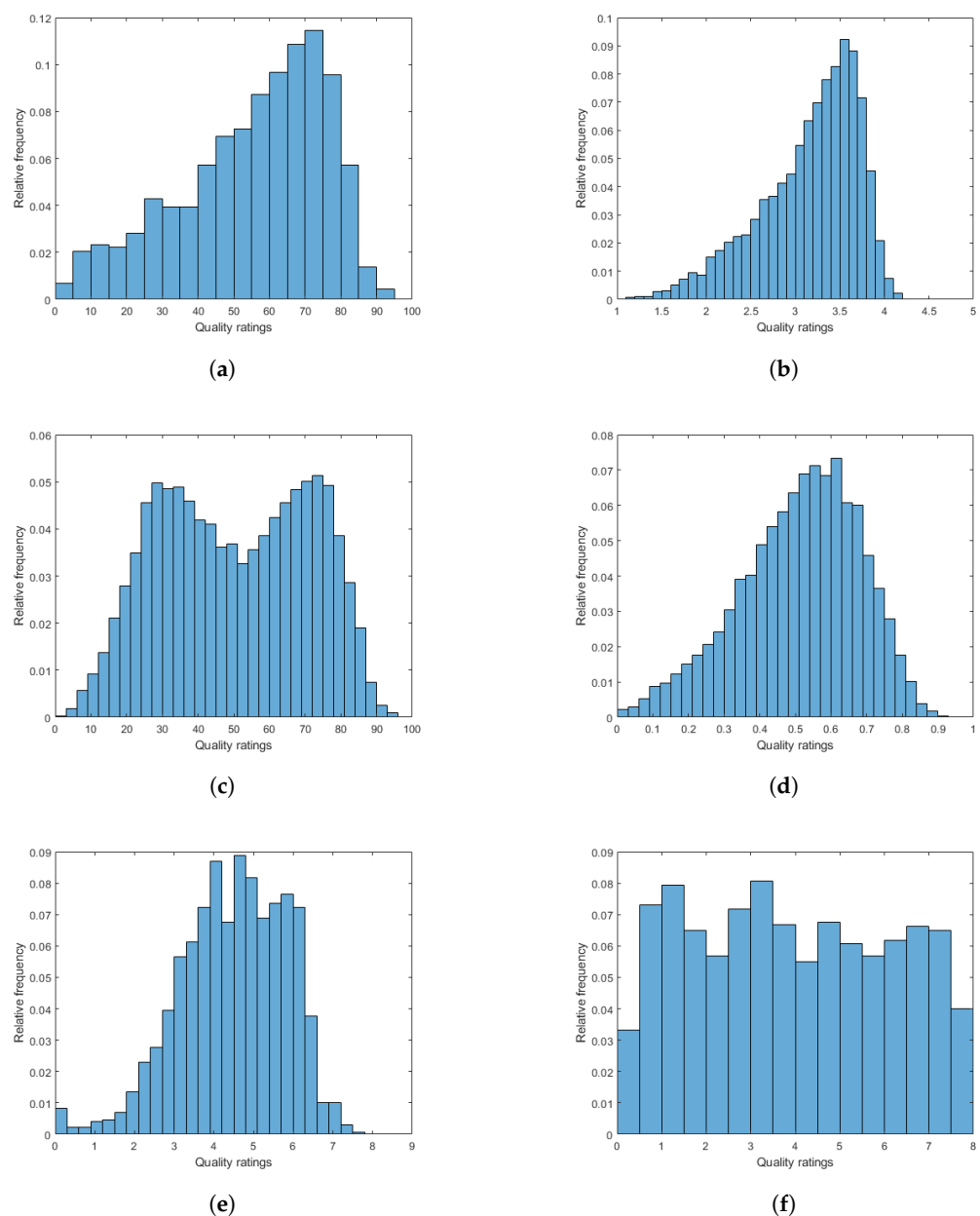(d)

(e)

(f)

Figure 3. The empirical distributions of quality ratings in the considered IQA benchmark databases: (**a**) CLIVE [10], (**b**) KonIQ-10k [11], (**c**) SPAQ [12], (**d**) BIQ2021 [13], (**e**) TID2008 [14], and (**f**) MDID [15].
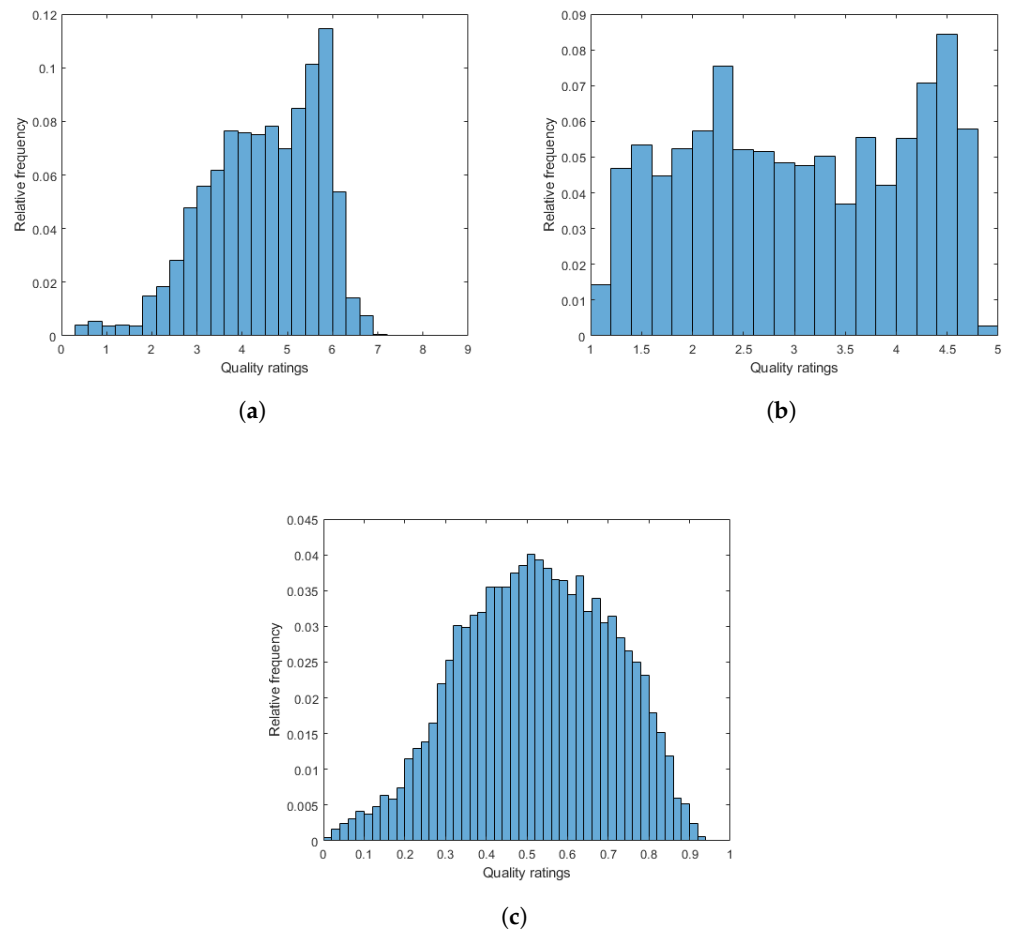
(a)

(b)

(c)

Figure 4. The empirical distributions of quality ratings in the considered IQA benchmark databases: (**a**) TID2013 [16], (**b**) KADID-10k [17], and (**c**) GFIQA-20k [18].

Table 1. Summary of the used benchmark IQA databases. The symbol # denotes the number of sign.

Database	No. of Reference Images	No. of Distorted Images	Distortion Type	Resolution	Environment
CLIVE [10]	-	1162	authentic	500×500	crowd-sourcing
KonIQ-10k [11]	-	10,073	authentic	1024×768	crowd-sourcing
SPAQ [12]	-	11,125	authentic	varied	laboratory
BIQ2021 [13]	-	8000	authentic	512×512	laboratory
TID2008 [14]	25	1700	artificial	512×384	laboratory
MDID [15]	20	1600	artificial	512×384	laboratory
TID2013 [16]	25	3000	artificial	512×384	laboratory
KADID-10k [17]	81	10,125	artificial	512×384	crowd-sourcing
GFIQA-20k [18]	-	20,000	authentic	512×512	laboratory

4.2. Evaluation Metrics and Protocol

In general, the NR-IQA methods' efficiencies are given by the intensity of the correlation between the estimated and ground truth quality ratings measured on a benchmark database. In the research field of NR-IQA, Pearson's linear correlation coefficient (PLCC), Spearman's rank order correlation coefficient (SROCC), and Kendall's rank order correlation coefficient (KROCC) are often applied as evaluation metrics.

PLCC between vectors $\mathbf{s} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ is defined as

$$PLCC = \frac{\sum_{i=1}^{n}(q_i - \bar{q})(s_i - \bar{s})}{\sqrt{\sum_{i=1}^{n}(q_i - \bar{q})^2 \sum_{i=1}^{n}(s_i - \bar{s})^2}}, \qquad (9)$$

where $\bar{q}$ and $\bar{s}$ are the arithmetic means of vectors **q** and **s**, respectively. Before the computation of PLCC, the predicted scores were mapped using a five-parameter nonlinear regression, as recommended in [75]. SROCC between vectors $\mathbf{s} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$ is described as follows:

$$SROCC = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{N(N^2-1)},$$

(10)

where d_i denotes the difference between the ranks of the *i*th entry in vectors **q** and **s**. KROCC is computed as follows:

$$KROCC = \frac{n_c - n_d}{\frac{1}{2}n(n-1)},$$

(11)

where n_c and n_d denote the number of concordant and discordant pairs between **s** and **q**, respectively.

In this paper, the median of the PLCC, SROCC, and KROCC values obtained was measured in 100 random train-test splits of a dataset. In particular, a database was randomly separated into a training (circa 80% of images) and a test (circa 20% of images) set. Next, PLCC, SROCC, and KROCC were computed on the test set. Databases with artificial distortions, such as TID2008 [14], MDID [15], TID2013 [16], and KADID-10k [17], were divided into training and test sets with respect to the reference images to avoid any semantic overlapping in these sets. On the other hand, databases with authentic distortions, such as CLIVE [10], KonIQ-10k [11], SPAQ [12], BIQ2021 [13], and GFIQA-20k [18], were divided with respect to all images, since they contained individual scenes.

Table 2 summarizes the main characteristics of the computer configuration which was applied in the experiments.

Table 2. Hardware configuration applied in experiments.

Computer model	Z590 D
CPU	Intel(R) Core(TM) i7-11700F CPU 2.50 GHz (8 cores)
Memory	31.9 GB
GPU	Nvidia GeForce RTX 3090

4.3. Ablation Study

To find a near-optimal regression algorithm for the proposed quality-aware feature extraction, a parameter study was carried out on CLIVE [10] using several different regression methods, such as Gaussian process regression (GPR) using the rational quadratic kernel function [76], radial basis function support vector regressor (RBF SVR) [77], generalized additive model (GAM) [78], extra tree [79], binary decision tree (BDT) [80], and LSBoost algorithm [81]. The obtained results are summarized in Figure 5. From these numerical results, we may draw the conclusion that GPR with the rational quadratic kernel function provided the highest outcomes, slightly outperforming the RBF SVR and significantly exceeding the other examined techniques. This is why GPR with the rational quadratic kernel function was applied in our proposed method, which is codenamed SGL-IQA in the following, emphasizing the fact that it relies on the statistics of the global and local image features. Figure 6 depicts scatter plots showing the ground truth versus predicted quality scores on a CLIVE [10] test set for each regression algorithm.

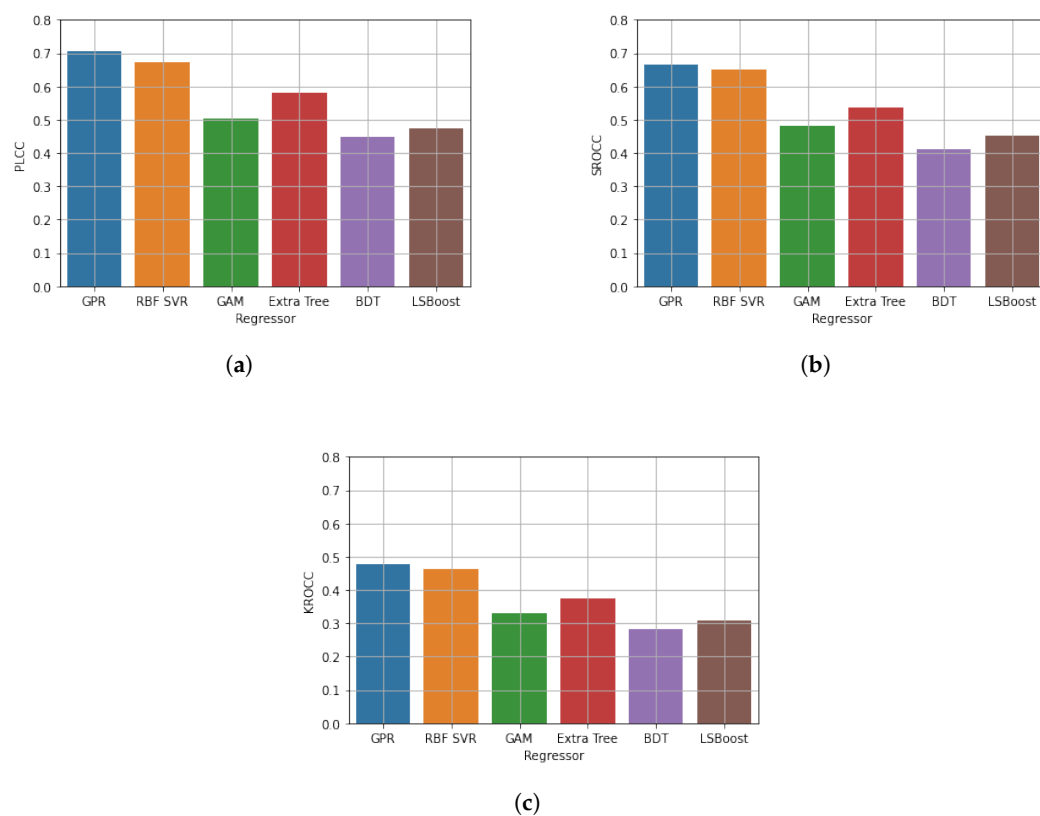(**a**)　　　　　　　　　　　　　　　　　　　　　(**b**)

(**c**)

Figure 5. Comparison of different regression techniques (GPR, RBF SVR, GAM, extra tree, BDT, and LSBoost) in terms of median (**a**) PLCC, (**b**) SROCC, and (**c**) KROCC. Measured on CLIVE [10] over 100 random train-test splits.

(**a**)　　　　　　　　　　　　　　　　　　　　　(**b**)
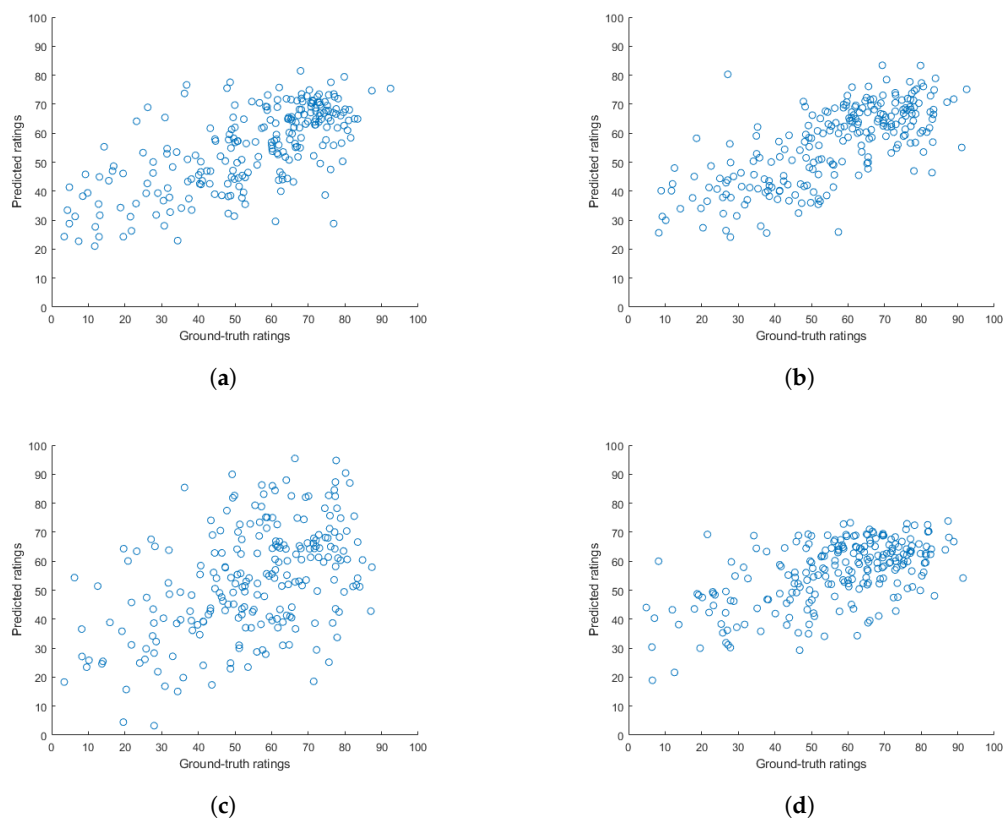
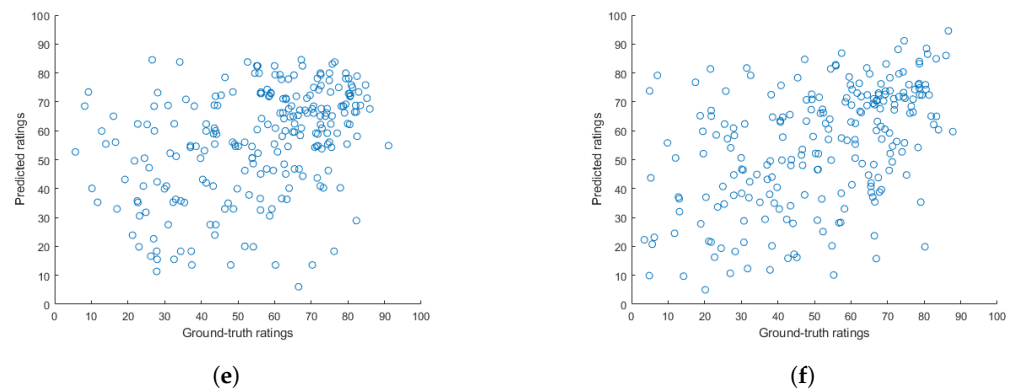

(**c**)　　　　　　　　　　　　　　　　　　　　　(**d**)

Figure 6. *Cont.*

Figure 6. Scatter plots of the ground truth versus the predicted quality scores on a CLIVE [10] test set for different regression algorithms: (**a**) GPR, (**b**) RBF SVR, (**c**) GAM, (**d**) extra tree, (**e**) BDT, and (**f**) LSBoost.

4.4. Comparison to the State of the Art

In this subsection, the proposed method is compared to the state of the art on nine benchmark databases (CLIVE [10], KonIQ-10k [11], SPAQ [12], BIQ2021 [13], TID2008 [14], MDID [15], TID2013 [16], KADID-10k [17], and GFIQA-20k [18]) using the evaluation metrics and protocol given in the previous subsection for the considered learning-based NR-IQA methods. On the other hand, opinion-unaware methods, such as IL-NIQE [82], NIQE [83], and PIQE [84], were assessed directly on the entire databases since they do not require any training steps. Furthermore, the source codes of all considered NR-IQA algorithms, i.e. BIQI (https://github.com/utlive/BIQI, accessed on 12 January 2023), BLIINDS-II (https://github.com/utlive/bliinds2, accessed on 12 January 2023), BMPRI (https://drive.google.com/file/d/1C_NxTLvnBOJDGhqqtixCkra0LIMP6loF/view, accessed on 12 January 2023), BRISQUE (https://github.com/utlive/brisque, accessed on 12 January 2023), CurveletQA (https://github.com/utlive/curveletqa, accessed on 12 January 2023), DIIVINE (https://github.com/utlive/DIIVINE, accessed on 12 January 2023), ENIQA (https://github.com/jacob6/ENIQA, accessed on 12 January 2023), GM-LOG-BIQA (http://ipl.xjtu.edu.cn/ftp/xqmou/GM-LOG-BIQA.zip, accessed on 12 January 2023), GWH-GLBP (https://github.com/JianjunXiang/PVRI/tree/master/GWH-GLBP-BIQA/GWH-GLBP-BIQA, accessed on 12 January 2023), IL-NIQE (http://www4.comp.polyu.edu.hk/~cslzhang/IQA/ILNIQE/ILNIQE.htm, accessed on 12 January 2023), NBIQA (https://github.com/GZHU-Image-Lab/NBIQA, accessed on 12 January 2023), NIQE (https://github.com/utlive/niqe, accessed on 12 January 2023), OG-IQA (https://github.com/utlive/og-iqa, accessed on 12 January 2023), PIQE (https://www.mathworks.com/help/images/ref/piqe.html, accessed on 12 January 2023), Robust BRISQUE (https://github.com/utlive/robustbrisque, accessed on 12 January 2023), and SSEQ (https://github.com/utlive/sseq, accessed on 12 January 2023), are publicly available online.

The results obtained on the individual databases are summarized in Tables 3–7. From the presented results, one can unambiguously see that the proposed SGL-IQA method was proven to be the best-performing algorithm on all considered benchmark databases which are de facto standards in NR-IQA. In Table 8, the results measured on the individual databases are summarized by taking their direct and weighted averages, where the weights correspond to the number of distorted images. From the numerical results, it can be seen that the proposed SGL-IQA gave higher performance indices in the case with weighted averages than in the case with direct averages. This suggests that the proposed method provides better performance on large databases than on smaller ones. For a visual competitive analysis between the proposed and other existing methods, radar graphs of the median SROCC values are depicted in Figure 7. As one can see in this figure, the proposed SGL-IQA method was able to achieve higher spikes than the other examined methods.

Table 3. Comparison to other state-of-the-art algorithms using CLIVE [10] and KonIQ-10k [11] databases. Median PLCC, SROCC, and KROCC values were measured over 100 random train-test splits. The best results are given in bold, the second-best results are underlined, and the third-best results are given in italics.

	CLIVE [10]			KonIQ-10k [11]		
Method	**PLCC**	**SROCC**	**KROCC**	**PLCC**	**SROCC**	**KROCC**
BIQI [85]	0.519	0.488	0.329	0.688	0.662	0.471
BLIINDS-II [23]	0.473	0.442	0.291	0.574	0.575	0.414
BMPRI [86]	0.541	0.487	0.333	0.637	0.619	0.421
BRISQUE [25]	0.524	0.497	0.345	0.707	0.677	0.494
CurveletQA [52]	0.636	0.621	*0.421*	0.730	0.718	0.495
DIIVINE [87]	0.617	0.580	0.405	0.709	0.693	0.471
ENIQA [88]	0.596	0.564	0.376	*0.761*	*0.745*	0.544
GM-LOG-BIQA [53]	0.607	*0.604*	0.383	0.705	0.696	0.501
GWH-GLBP [54]	0.584	0.559	0.395	0.723	0.698	0.507
IL-NIQE [82]	0.487	0.415	0.280	0.463	0.447	0.306
NBIQA [89]	*0.629*	*0.604*	0.427	0.771	0.749	*0.515*
NIQE [83]	0.328	0.299	0.200	0.319	0.400	0.272
OG-IQA [56]	0.545	0.505	0.364	0.652	0.635	0.447
PIQE [84]	0.172	0.108	0.081	0.208	0.246	0.172
Robust BRISQUE [26]	0.522	0.484	0.330	0.718	0.668	0.477
SSEQ [57]	0.487	0.436	0.309	0.589	0.572	0.423
SGL-IQA	**0.704**	**0.667**	**0.478**	**0.815**	**0.794**	**0.596**

Table 4. Comparison to other state-of-the-art algorithms using SPAQ [12] and BIQ2021 [13] databases. Median PLCC, SROCC, and KROCC values were measured over 100 random train-test splits. The best results are given in bold, the second-best results are underlined, and the third-best results are given in italics.

	SPAQ [12]			BIQ2021 [13]		
Method	**PLCC**	**SROCC**	**KROCC**	**PLCC**	**SROCC**	**KROCC**
BIQI [85]	0.783	0.776	0.566	0.644	0.564	0.399
BLIINDS-II [23]	0.676	0.675	0.486	0.555	0.496	0.346
BMPRI [86]	0.739	0.734	0.506	0.633	0.494	0.345
BRISQUE [25]	0.726	0.720	0.518	0.669	0.555	0.392
CurveletQA [52]	0.793	0.774	0.503	0.698	0.630	0.453
DIIVINE [87]	0.774	0.756	0.514	0.684	0.617	0.442
ENIQA [88]	0.813	0.804	0.603	*0.703*	*0.634*	*0.456*
GM-LOG-BIQA [53]	0.786	0.782	0.572	0.699	0.617	0.443
GWH-GLBP [54]	0.801	*0.796*	0.542	0.664	0.602	0.431
IL-NIQE [82]	0.374	0.348	0.297	0.541	0.461	0.317
NBIQA [89]	*0.802*	0.793	0.539	0.718	0.642	0.463
NIQE [83]	0.264	0.310	0.206	0.301	0.356	0.240
OG-IQA [56]	0.726	0.724	*0.594*	0.403	0.371	0.253
PIQE [84]	0.211	0.156	0.091	0.255	0.213	0.142
Robust BRISQUE [26]	0.735	0.731	0.524	0.683	0.605	0.432
SSEQ [57]	0.745	0.742	0.549	0.603	0.528	0.369
SGL-IQA	**0.866**	**0.859**	**0.661**	**0.770**	**0.710**	**0.522**

Table 5. Comparison to other state-of-the-art algorithms using TID2008 [14] and MDID [15] databases. Median PLCC, SROCC, and KROCC values were measured over 100 random train-test splits carried out with respect to the reference images. The best results are given in bold, the second-best results are underlined, and the third-best results are given in italics.

	TID2008 [14]			MDID [15]		
Method	**PLCC**	**SROCC**	**KROCC**	**PLCC**	**SROCC**	**KROCC**
BIQI [85]	0.496	0.424	0.301	0.708	0.700	0.508
BLIINDS-II [23]	0.693	0.603	0.440	0.694	0.683	0.488
BMPRI [86]	0.770	*0.681*	0.500	0.770	0.755	*0.555*
BRISQUE [25]	0.602	0.451	0.317	0.632	0.624	0.442
CurveletQA [52]	0.607	0.539	0.393	0.692	0.682	0.484

Table 5. *Cont.*

DIIVINE [87]	0.546	0.537	0.381	0.727	0.724	0.518
ENIQA [88]	0.615	0.564	0.406	*0.771*	*0.766*	0.553
GM-LOG-BIQA [53]	0.698	<u>0.686</u>	<u>0.508</u>	0.640	0.650	0.450
GWH-GLBP [54]	0.368	0.375	0.258	0.724	0.753	0.551
IL-NIQE [82]	0.476	0.420	0.292	0.733	0.690	0.491
NBIQA [89]	*0.751*	0.676	*0.507*	<u>0.779</u>	<u>0.775</u>	<u>0.565</u>
NIQE [83]	0.195	0.186	0.125	0.460	0.420	0.285
OG-IQA [56]	0.663	0.583	0.426	0.754	0.718	0.515
PIQE [84]	0.427	0.327	0.235	0.280	0.253	0.175
Robust BRISQUE [26]	0.696	0.586	0.425	0.753	0.743	0.532
SSEQ [57]	0.704	0.640	0.470	0.747	0.738	0.537
SGL-IQA	**0.776**	**0.722**	**0.549**	**0.818**	**0.815**	**0.614**

Table 6. Comparison to other state-of-the-art algorithms using TID2013 [16] and KADID-10k [17] databases. Median PLCC, SROCC, and KROCC values were measured over 100 random train-test splits carried out with respect to the reference images. The best results are given in bold, the second-best results are underlined, and the third-best results are given in italics.

	TID2013 [16]			KADID-10k [17]		
Method	**PLCC**	**SROCC**	**KROCC**	**PLCC**	**SROCC**	**KROCC**
BIQI [85]	0.468	0.296	0.207	0.302	0.294	0.206
BLIINDS-II [23]	0.521	0.490	0.342	0.553	0.534	0.379
BMPRI [86]	*0.692*	0.583	0.422	0.555	0.534	0.382
BRISQUE [25]	0.565	0.411	0.289	0.426	0.398	0.276
CurveletQA [52]	0.560	0.471	0.337	0.471	0.442	0.316
DIIVINE [87]	0.521	0.487	0.340	0.429	0.436	0.307
ENIQA [88]	0.596	0.545	0.385	*0.637*	<u>0.641</u>	<u>0.466</u>
GM-LOG-BIQA [53]	0.662	*0.627*	*0.454*	0.590	0.570	0.415
GWH-GLBP [54]	0.315	0.357	0.245	0.302	0.285	0.196
IL-NIQE [82]	0.516	0.456	0.317	0.588	*0.630*	*0.453*
NBIQA [89]	<u>0.695</u>	<u>0.628</u>	<u>0.459</u>	<u>0.646</u>	0.615	0.446
NIQE [83]	0.263	0.277	0.184	0.302	0.338	0.228
OG-IQA [56]	0.564	0.452	0.321	0.527	0.447	0.314
PIQE [84]	0.491	0.364	0.255	0.289	0.237	0.201
Robust BRISQUE [26]	0.487	0.315	0.218	0.375	0.301	0.209
SSEQ [57]	0.615	0.520	0.373	0.454	0.434	0.304
SGL-IQA	**0.713**	**0.656**	**0.487**	**0.782**	**0.774**	**0.583**

Table 7. Comparison to other state-of-the-art algorithms using GFIQA-20k [18] database. Median PLCC, SROCC, and KROCC values were measured over 100 random train-test splits. The best results are given in bold, the second-best results are underlined, and the third-best results are given in italics.

Method	**PLCC**	**SROCC**	**KROCC**
BIQI [85]	0.794	0.790	0.599
BLIINDS-II [23]	0.685	0.674	0.491
BMPRI [86]	0.673	0.662	0.481
BRISQUE [25]	0.721	0.718	0.527
CurveletQA [52]	0.799	0.779	0.591
DIIVINE [87]	0.822	0.823	0.632
ENIQA [88]	<u>0.834</u>	<u>0.838</u>	<u>0.647</u>
GM-LOG-BIQA [53]	0.740	0.732	0.543
GWH-GLBP [54]	*0.831*	0.822	0.633
IL-NIQE [82]	0.728	0.714	0.518
NBIQA [89]	0.829	*0.830*	*0.640*
NIQE [83]	0.191	0.183	0.127
OG-IQA [56]	0.747	0.735	0.546
PIQE [84]	0.207	0.095	0.066
Robust BRISQUE [26]	0.817	0.816	0.625
SSEQ [57]	0.715	0.690	0.509
SGL-IQA	**0.866**	**0.862**	**0.663**

Table 8. PLCC , SROCC, and KROCC direct and weighted averages. The best results are given in bold, the second-best results are underlined, and the third-best results are given in italics.

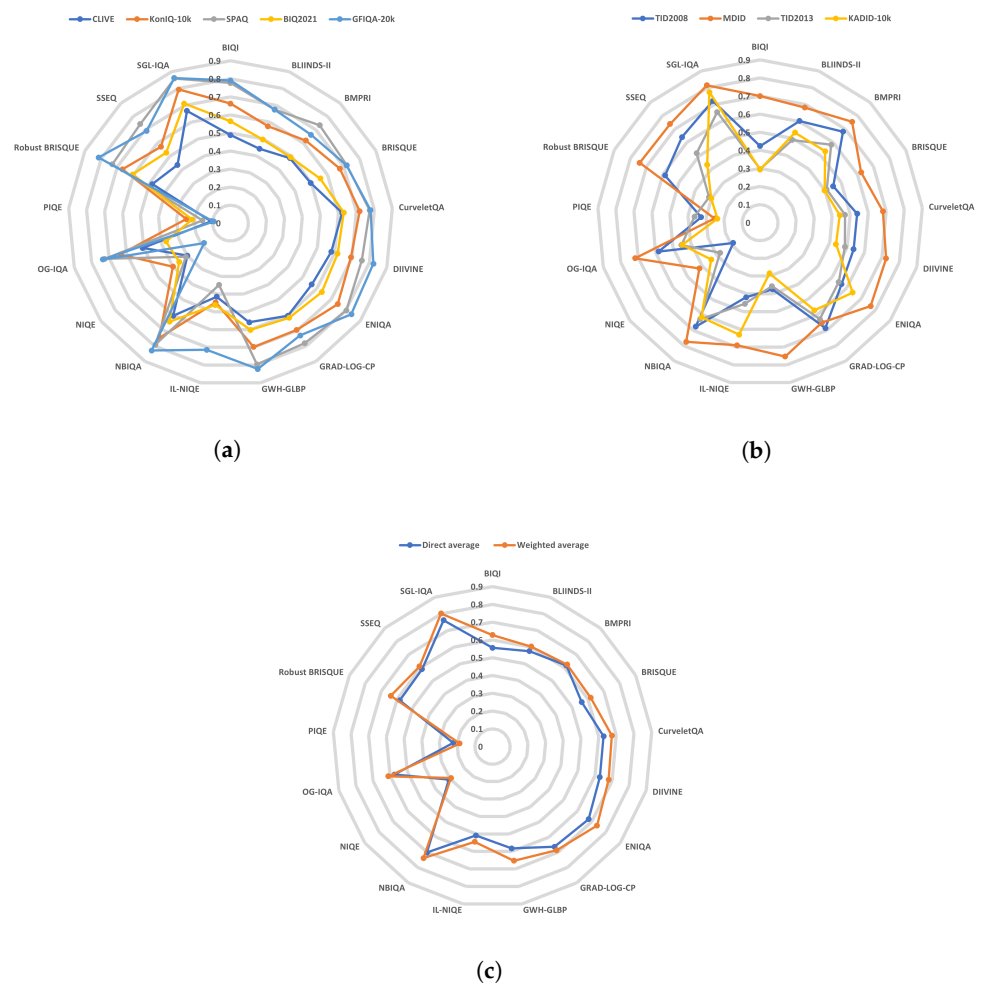| Method	Direct Average			Weighted Average		
	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC
BIQI [85]	0.600	0.555	0.398	0.655	0.627	0.459
BLIINDS-II [23]	0.603	0.575	0.409	0.621	0.603	0.433
BMPRI [86]	0.668	0.617	0.438	0.659	0.624	0.442
BRISQUE [25]	0.619	0.561	0.400	0.653	0.617	0.445
CurveletQA [52]	0.665	0.628	0.444	0.705	0.675	0.482
DIIVINE [87]	0.648	0.628	0.446	0.668	0.680	0.490
ENIQA [88]	*0.703*	*0.678*	*0.493*	*0.752*	*0.737*	<u>0.549</u>
GM-LOG-BIQA [53]	0.681	0.663	0.474	0.705	0.686	0.500
GWH-GLBP [54]	0.590	0.583	0.418	0.668	0.653	0.475
IL-NIQE [82]	0.545	0.509	0.363	0.565	0.545	0.396
NBIQA [89]	<u>0.736</u>	<u>0.701</u>	<u>0.507</u>	<u>0.762</u>	<u>0.738</u>	*0.537*
NIQE [83]	0.291	0.308	0.207	0.265	0.293	0.198
OG-IQA [56]	0.620	0.574	0.420	0.641	0.610	0.452
PIQE [84]	0.282	0.222	0.158	0.245	0.186	0.131
Robust BRISQUE [26]	0.643	0.583	0.419	0.681	0.640	0.469
SSEQ [57]	0.629	0.589	0.427	0.640	0.610	0.445
SGL-IQA	**0.790**	**0.762**	**0.573**	**0.821**	**0.802**	**0.608**

(**a**)

(**b**)

(**c**)

Figure 7. Radar graphs for SROCC comparison to other advanced NR-IQA methods based on (**a**) authentic distortions, (**b**) artificial distortions, and (**c**) overall performance on the considered benchmark databases.

To demonstrate that the performance differences between the proposed and other state-of-the-art methods on the considered benchmark datasets were significant, significance tests and a cross-database test were carried out. Specifically, Wilcoxon rank sum tests were applied [90] for justifying the significance of the results. Specifically, we applied the null hypothesis that two sets of SROCC values of two different NR-IQA methods were sampled from continuous distributions which had equal median values. Furthermore, we set the significance level to 5%. The outcomes of the Wilcoxon rank sum tests are summarized in Tables 9 and 10. In this table, the symbol '1' denotes that SGL-IQA provided significantly better results than the algorithm shown in the row on the database shown in the column. It can be seen that the proposed method provided significantly better results than those of the other considered state-of-the-art methods.

Table 9. Results of the two-sided Wilcoxon rank sum significance tests on CLIVE [10], KonIQ-10k [11], SPAQ [12], BID2021 [13], and TID2008 [14]. The symbol '1' denotes that SGL-IQA provided significantly better results than the method represented in the row on the database represented in the column.

Method	CLIVE [10]	KonIQ-10k [11]	SPAQ [12]	BID2021 [13]	TID2008 [14]
BIQI [85]	1	1	1	1	1
BLIINDS-II [23]	1	1	1	1	1
BMPRI [86]	1	1	1	1	1
BRISQUE [25]	1	1	1	1	1
CurveletQA [52]	1	1	1	1	1
DIIVINE [87]	1	1	1	1	1
ENIQA [88]	1	1	1	1	1
GM-LOG-BIQA [53]	1	1	1	1	1
GWH-GLBP [54]	1	1	1	1	1
NBIQA [89]	1	1	1	1	1
OG-IQA [56]	1	1	1	1	1
Robust BRISQUE [26]	1	1	1	1	1
SSEQ [57]	1	1	1	1	1

Table 10. Results of the two-sided Wilcoxon rank sum significance tests on MDID [15], TID2013 [16], KADID-10k [17], and GFIQA-20k [18]. The symbol '1' denotes that SGL-IQA provided significantly better results than the method represented in the row on the database represented in the column.

Method	MDID [15]	TID2013 [16]	KADID-10k [17]	GFIQA-20k [18]
BIQI [85]	1	1	1	1
BLIINDS-II [23]	1	1	1	1
BMPRI [86]	1	1	1	1
BRISQUE [25]	1	1	1	1
CurveletQA [52]	1	1	1	1
DIIVINE [87]	1	1	1	1
ENIQA [88]	1	1	1	1
GM-LOG-BIQA [53]	1	1	1	1
GWH-GLBP [54]	1	1	1	1
NBIQA [89]	1	1	1	1
OG-IQA [56]	1	1	1	1
Robust BRISQUE [26]	1	1	1	1
SSEQ [57]	1	1	1	1

In another test, the generalization abilities of the proposed SGL-IQA and the other state-of-the-art methods were examined. Specifically, a cross-database test was carried out where all machine learning-based algorithms were trained on the entire KonIQ-10k [11] database and subsequently tested on the entire CLIVE [10] database. The numerical results of this cross-database test are outlined in Table 11. From these results, it can be clearly seen that the proposed SGL-IQA method outperformed the other considered state-of-the-art methods by a large margin. The difference between the proposed and second-best-performing ones was approximately 0.09 in terms of PLCC and SROCC.

To give a comparison of the computational complexities of the proposed and other state-of-the-art methods, we measured the times for feature extractions on CLIVE [10] and KonIQ-10k [11]. The applied hardware configuration was already summarized in Table 2. In Table 12, our numerical results regarding the times for feature extraction are summarized. To illustrate the execution time and performance simultaneously, SROCC vs. time of feature extraction scatter plots based on the numerical results obtained from CLIVE [10] and KonIQ-10k [11] are presented in Figures 8 and 9. The demostrated numerical data suggest that SGL-IQA can give significantly higher performance metrics than the second- or third-best-performing methods at lower computational costs.

Table 11. Numerical results of the cross-database test. The considered algorithms were trained on the entire KonIQ-10k [11] database and tested on the entire CLIVE [10] database. The best results are given in bold, the second-best results are underlined, and the third-best results are given in italics.

Method	PLCC	SROCC	KROCC
BIQI [85]	0.477	0.424	0.289
BLIINDS-II [23]	0.107	0.090	0.063
BMPRI [86]	0.453	0.389	0.298
BRISQUE [25]	*0.509*	0.460	0.310
CurveletQA [52]	0.496	*0.505*	0.347
DIIVINE [87]	0.479	0.434	0.299
ENIQA [88]	0.428	0.386	0.272
GM-LOG-BIQA [53]	0.427	0.384	0.261
GWH-GLBP [54]	0.480	0.479	*0.328*
NBIQA [89]	0.503	0.509	0.284
OG-IQA [56]	0.442	0.427	0.289
Robust BRISQUE [26]	0.516	0.481	0.327
SSEQ [57]	0.270	0.256	0.170
SGL-IQA	**0.608**	**0.595**	**0.416**

Table 12. Feature extraction times (given in seconds) on CLIVE [10] and KonIQ-10k [11]. The best results are given in bold, the second-best results are underlined, and the third-best results are given in italics.

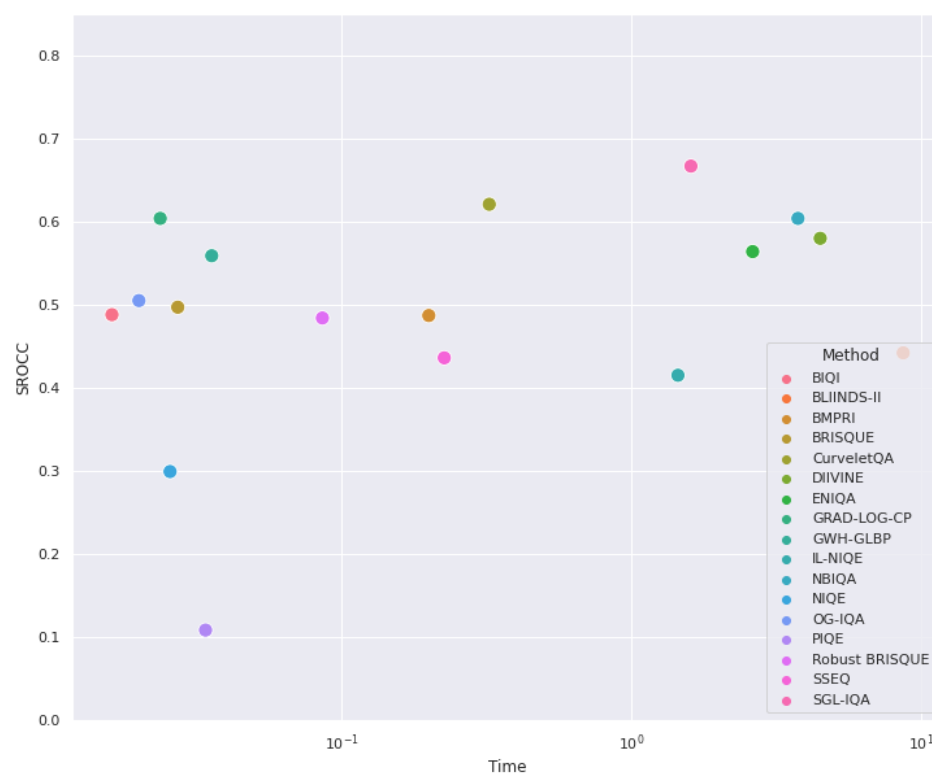| Method	CLIVE [10]	KonIQ-10k [11]
BIQI [85]	**0.016**	**0.041**
BLIINDS-II [23]	8.704	28.236
BMPRI [86]	0.200	0.487
BRISQUE [25]	0.027	0.077
CurveletQA [52]	0.324	0.937
DIIVINE [87]	4.500	12.896
ENIQA [88]	2.630	7.324
GM-LOG-BIQA [53]	*0.024*	*0.061*
GWH-GLBP [54]	0.036	0.101
IL-NIQE [82]	1.452	1.980
NBIQA [89]	3.770	12.580
NIQE [83]	0.026	0.063
OG-IQA [56]	0.020	0.053
PIQE [84]	0.034	0.072
Robust BRISQUE [26]	0.086	0.273
SSEQ [57]	0.227	0.732
SGL-IQA	1.610	4.783

Figure 8. SROCC vs. feature extraction time (given in seconds on a logarithmic scale) based on the results obtained on CLIVE [10].
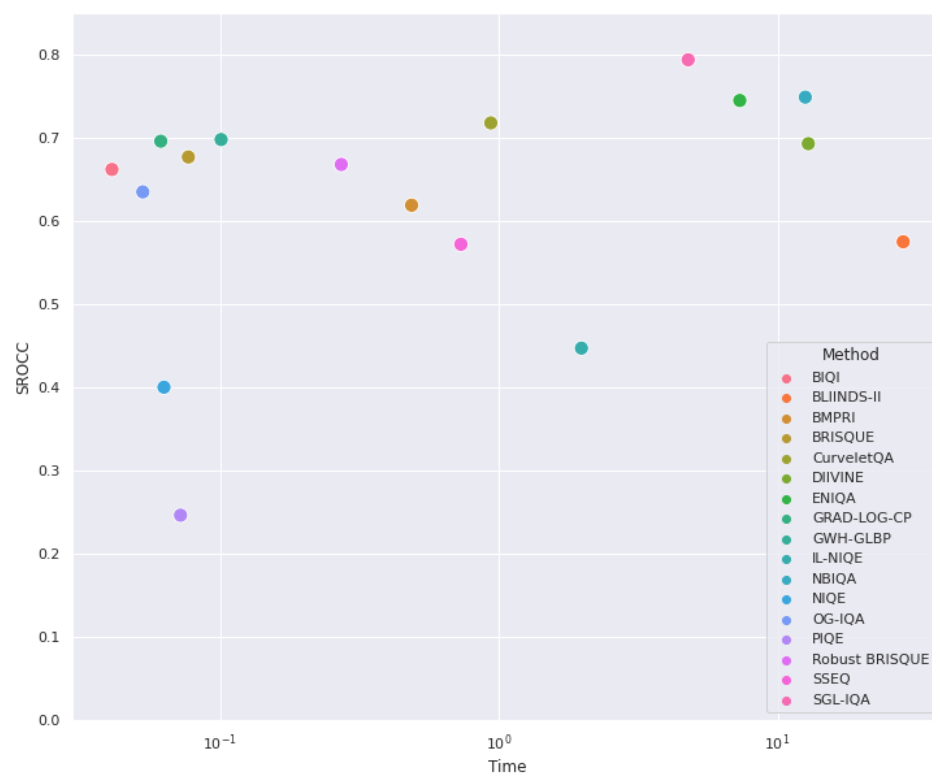

Figure 9. SROCC vs. feature extraction time (given in seconds on a logarithmic scale) based on the results obtained on KonIQ-10k [11].

5. Conclusions

In this paper, we made an effort to predict the perceptual quality of digital images relying on the rich statistics of global and local image features. Specifically, visual physiology's previous research results indicated that the human visual system first constructs a global perception about a visual scene. Next, it focuses on local areas to judge the perceptual image quality. Unlike previous studies, we did not simply fuse together certain global and local features since this process quickly produces a long feature vector, but a broad spectrum of statistics was extracted from them to characterize the quality-aware aspects of an image in a compact feature vector. The experimental results demonstrated that the proposed SGL-IQA method was able to significantly surpass the state of the art in terms of the correlation strength between the predicted and ground truth quality ratings evaluated on nine benchmark IQA databases. Furthermore, these findings were also reinforced in significance and cross-database tests.

To facilitate the reproduction of the presented results, the source code of the proposed SGL-IQA no-reference image quality assessment algorithm can be downloaded at: https://github.com/Skythianos/SGL-IQA (accessed on 2 March 2023).

Abbreviations

The following abbreviations are used in this manuscript:

BDT	binary decision tree
BIQ	blind image quality
CNN	convolutional neural network
DCT	discrete cosine transform
FAST	features from the accelerated segment test
GAM	generalized additive model
GGD	generalized Gaussian distribution
GPR	Gaussian process regression
HVS	human visual system
IQA	image quality assessment
KADID	Konstanz artificially distorted image quality database
KROCC	Kendall's rank order correlation coefficient
LIVE	laboratory for image and video engineering
MDID	multiply distorted image database

NR-IQA	no-reference video quality assessment
NSS	natural scene statistics
PLCC	Pearson's linear correlation coefficient
RBF	radial basis function
SPAQ	smartphone photography attribute and quality
SROCC	Spearman's rank order correlation coefficient
SVR	support vector regressor
TID	Tampere image database

References

1. Zhu, M.; Yu, L.; Wang, Z.; Ke, Z.; Zhi, C. A Survey on Objective Evaluation of Image Sharpness. *Appl. Sci.* **2023**, *13*, 2652. [CrossRef]

2. Afnan; Ullah, F.; Yaseen; Lee, J.; Jamil, S.; Kwon, O.J. Subjective Assessment of Objective Image Quality Metrics Range Guaranteeing Visually Lossless Compression. *Sensors* **2023**, *23*, 1297.

3. Wang, Z. Applications of objective image quality assessment methods [applications corner]. *IEEE Signal Process. Mag.* **2011**, *28*, 137–142. [CrossRef]

4. De Cesarei, A.; Loftus, G.R. Global and local vision in natural scene identification. *Psychon. Bull. Rev.* **2011**, *18*, 840–847. [CrossRef] [PubMed]

5. Bae, S.H.; Kim, M. A novel image quality assessment with globally and locally consilient visual quality perception. *IEEE Trans. Image Process.* **2016**, *25*, 2392–2406. [CrossRef] [PubMed]

6. Wang, H.; Qu, H.; Xu, J.; Wang, J.; Wei, Y.; Zhang, Z. Combining Statistical Features and Local Pattern Features for Texture Image Retrieval. *IEEE Access* **2020**, *8*, 222611–222624. [CrossRef]

7. Chang, H.W.; Du, C.Y.; Bi, X.D.; Chen, K.; Wang, M.H. LG-IQA: Integration of local and global features for no-reference image quality assessment. *Displays* **2022**, *75*, 102334. [CrossRef]

8. Shnayderman, A.; Gusev, A.; Eskicioglu, A.M. An SVD-based grayscale image quality measure for local and global assessment. *IEEE Trans. Image Process.* **2006**, *15*, 422–429. [CrossRef]

9. Yue, G.; Hou, C.; Gu, K.; Zhou, T.; Zhai, G. Combining local and global measures for DIBR-synthesized image quality evaluation. *IEEE Trans. Image Process.* **2018**, *28*, 2075–2088. [CrossRef]

10. Ghadiyaram, D.; Bovik, A.C. Massive online crowdsourced study of subjective and objective picture quality. *IEEE Trans. Image Process.* **2015**, *25*, 372–387. [CrossRef]

11. Lin, H.; Hosu, V.; Saupe, D. KonIQ-10K: Towards an ecologically valid and large-scale IQA database. *arXiv* **2018**, arXiv:1803.08489.

12. Fang, Y.; Zhu, H.; Zeng, Y.; Ma, K.; Wang, Z. Perceptual quality assessment of smartphone photography. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 3677–3686.

13. Ahmed, N.; Asif, S. BIQ2021: A Large-Scale Blind Image Quality Assessment Database. *arXiv* **2022**, arXiv:2202.03879.

14. Ponomarenko, N.; Lukin, V.; Zelensky, A.; Egiazarian, K.; Carli, M.; Battisti, F. TID2008—A database for evaluation of full-reference visual quality assessment metrics. *Adv. Mod. Radioelectron.* **2009**, *10*, 30–45.

15. Sun, W.; Zhou, F.; Liao, Q. MDID: A multiply distorted image database for image quality assessment. *Pattern Recognit.* **2017**, *61*, 153–168. [CrossRef]

16. Ponomarenko, N.; Jin, L.; Ieremeiev, O.; Lukin, V.; Egiazarian, K.; Astola, J.; Vozel, B.; Chehdi, K.; Carli, M.; Battisti, F.; et al. Image database TID2013: Peculiarities, results and perspectives. *Signal Process. Image Commun.* **2015**, *30*, 57–77. [CrossRef]

17. Lin, H.; Hosu, V.; Saupe, D. KADID-10k: A large-scale artificially distorted IQA database. In Proceedings of the 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), Berlin, Germany, 5–7 June 2019; pp. 1–3.

18. Su, S.; Lin, H.; Hosu, V.; Wiedemann, O.; Sun, J.; Zhu, Y.; Liu, H.; Zhang, Y.; Saupe, D. Going the Extra Mile in Face Image Quality Assessment: A Novel Database and Model. *arXiv* **2022**, arXiv:2207.04904.

19. Men, H.; Lin, H.; Jenadeleh, M.; Saupe, D. Subjective image quality assessment with boosted triplet comparisons. *IEEE Access* **2021**, *9*, 138939–138975. [CrossRef]

20. Sheikh, H.R.; Bovik, A.C.; De Veciana, G. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. Image Process.* **2005**, *14*, 2117–2128. [CrossRef]

21. Bovik, A.C. Automatic prediction of perceptual image and video quality. *Proc. IEEE* **2013**, *101*, 2008–2024.

22. Saad, M.A.; Bovik, A.C.; Charrier, C. A DCT statistics-based blind image quality index. *IEEE Signal Process. Lett.* **2010**, *17*, 583–586. [CrossRef]

23. Saad, M.A.; Bovik, A.C. Blind quality assessment of videos using a model of natural scene statistics and motion coherency. In Proceedings of the 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Grove, CA, USA, 4–7 November 2012; pp. 332–336.

24. Li, X.; Guo, Q.; Lu, X. Spatiotemporal statistics for video quality assessment. *IEEE Trans. Image Process.* **2016**, *25*, 3329–3342. [CrossRef]

25. 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]

26.   Mittal, A.; Moorthy, A.K.; Bovik, A.C. Making image quality assessment robust. In Proceedings of the 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Grove, CA, USA, 4–7 November 2012; pp. 1718–1722.

27.   Hosking, J.R. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc. Ser. B (Methodol.)* **1990**, *52*, 105–124. [CrossRef]

28.   Bagade, J.V.; Singh, K.; Dandawate, Y.H. No reference image quality assessment with shape adaptive discrete wavelet features using neuro-wavelet model. *Multimed. Tools Appl.* **2022**, *81*, 31145–31160. [CrossRef]

29.   Moutik, O.; Sekkat, H.; Tigani, S.; Chehri, A.; Saadane, R.; Tchakoucht, T.A.; Paul, A. Convolutional Neural Networks or Vision Transformers: Who Will Win the Race for Action Recognitions in Visual Data? *Sensors* **2023**, *23*, 734. [CrossRef]

30.   Afshar, P.; Mohammadi, A.; Plataniotis, K.N.; Oikonomou, A.; Benali, H. From handcrafted to deep-learning-based cancer radiomics: challenges and opportunities. *IEEE Signal Process. Mag.* **2019**, *36*, 132–160. [CrossRef]

31.   Fiaz, M.; Mahmood, A.; Javed, S.; Jung, S.K. Handcrafted and deep trackers: Recent visual object tracking approaches and trends. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 1–44. [CrossRef]

32.   Kang, L.; Ye, P.; Li, Y.; Doermann, D. Convolutional neural networks for no-reference image quality assessment. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1733–1740.

33.   Wiedemann, O.; Hosu, V.; Lin, H.; Saupe, D. Disregarding the big picture: Towards local image quality assessment. In Proceedings of the 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), Sardinia, Italy, 29–31 May 2018; pp. 1–6.

34.   Kim, J.; Lee, S. Fully deep blind image quality predictor. *IEEE J. Sel. Top. Signal Process.* **2016**, *11*, 206–220. [CrossRef]

35.   Hore, A.; Ziou, D. Image quality metrics: PSNR vs. SSIM. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 2366–2369.

36.   Muthusamy, D.; Sathyamoorthy, S. Feature Sampling based on Multilayer Perceptive Neural Network for image quality assessment. *Eng. Appl. Artif. Intell.* **2023**, *121*, 106015. [CrossRef]

37.   Bhatia, N.; Vandana. Survey of nearest neighbor techniques. *arXiv* **2010**, arXiv:1007.0085.

38.   Ryu, J. A Visual Saliency-Based Neural Network Architecture for No-Reference Image Quality Assessment. *Appl. Sci.* **2022**, *12*, 9567. [CrossRef]

39.   Celona, L.; Schettini, R. Blind quality assessment of authentically distorted images. *JOSA A* **2022**, *39*, B1–B10. [CrossRef] [PubMed]

40.   Ryu, J. Improved Image Quality Assessment by Utilizing Pre-Trained Architecture Features with Unified Learning Mechanism. *Appl. Sci.* **2023**, *13*, 2682. [CrossRef]

41.   Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.

42.   Kabir, H.D.; Abdar, M.; Khosravi, A.; Jalali, S.M.J.; Atiya, A.F.; Nahavandi, S.; Srinivasan, D. Spinalnet: Deep neural network with gradual input. *IEEE Trans. Artif. Intell.* **2022**, 1–13. [CrossRef]

43.   Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. [CrossRef]

44.   Zhu, H.; Zhou, Y.; Shao, Z.; Du, W.L.; Zhao, J.; Yao, R. ARET-IQA: An Aspect-Ratio-Embedded Transformer for Image Quality Assessment. *Electronics* **2022**, *11*, 2132. [CrossRef]

45.   Ke, J.; Wang, Q.; Wang, Y.; Milanfar, P.; Yang, F. Musiq: Multi-scale image quality transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 5148–5157.

46.   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.

47.   Wang, J.; Fan, H.; Hou, X.; Xu, Y.; Li, T.; Lu, X.; Fu, L. MSTRIQ: No Reference Image Quality Assessment Based on Swin Transformer with Multi-Stage Fusion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1269–1278.

48.   You, J.; Korhonen, J. Attention integrated hierarchical networks for no-reference image quality assessment. *J. Vis. Commun. Image Represent.* **2022**, *82*, 103399. [CrossRef]

49.   Jenadeleh, M. Blind Image and Video Quality Assessment. Ph.D. Thesis, University of Konstanz, Konstanz, Germany, 24 October 2018.

50.   Men, H. Boosting for Visual Quality Assessment with Applications for Frame Interpolation Methods. Ph.D. Thesis, University of Konstanz, Konstanz, Germany, 11 April 2022.

51.   Xu, L.; Lin, W.; Kuo, C.C.J. *Visual Quality Assessment by Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2015.

52.   Liu, L.; Dong, H.; Huang, H.; Bovik, A.C. No-reference image quality assessment in curvelet domain. *Signal Process. Image Commun.* **2014**, *29*, 494–505. [CrossRef]

53.   Xue, W.; Mou, X.; Zhang, L.; Bovik, A.C.; Feng, X. Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features. *IEEE Trans. Image Process.* **2014**, *23*, 4850–4862. [CrossRef]

54.   Li, Q.; Lin, W.; Fang, Y. No-reference quality assessment for multiply-distorted images in gradient domain. *IEEE Signal Process. Lett.* **2016**, *23*, 541–545. [CrossRef]

55. Kundu, D.; Ghadiyaram, D.; Bovik, A.C.; Evans, B.L. No-reference image quality assessment for high dynamic range images. In Proceedings of the 2016 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 6–9 November 2016; pp. 1847–1852.

56. Liu, L.; Hua, Y.; Zhao, Q.; Huang, H.; Bovik, A.C. Blind image quality assessment by relative gradient statistics and adaboosting neural network. *Signal Process. Image Commun.* **2016**, *40*, 1–15. [CrossRef]

57. Liu, L.; Liu, B.; Huang, H.; Bovik, A.C. No-reference image quality assessment based on spatial and spectral entropies. *Signal Process. Image Commun.* **2014**, *29*, 856–863. [CrossRef]

58. Ma, J.; Plonka, G. The curvelet transform. *IEEE Signal Process. Mag.* **2010**, *27*, 118–133. [CrossRef]

59. Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [CrossRef]

60. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005; Volume 1, pp. 886–893.

61. Oszust, M. No-reference image quality assessment with local gradient orientations. *Symmetry* **2019**, *11*, 95. [CrossRef]

62. Varga, D. A Human Visual System Inspired No-Reference Image Quality Assessment Method Based on Local Feature Descriptors. *Sensors* **2022**, *22*, 6775. [CrossRef] [PubMed]

63. Rosten, E.; Drummond, T. Fusing points and lines for high performance tracking. In Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05), Beijing, China, 17–21 October 2005; Volume 2, pp. 1508–1515.

64. ITU Radiocommunication Assembly. *Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios*; International Telecommunication Union: Geneva, Switzerland; CCIR: Cambridge, UK, 2011.

65. Chai, D.; Bouzerdoum, A. A Bayesian approach to skin color classification in YCbCr color space. In Proceedings of the 2000 TENCON Proceedings, Intelligent Systems and Technologies for the New Millennium (Cat. No. 00CH37119), Kuala Lumpur, Malaysia, 24–27 September 2000; Volume 2, pp. 421–424.

66. Poynton, C.A. *A Technical Introduction to Digital Video*; John Wiley & Sons: Hoboken, NJ, USA, 1996.

67. Ghosh, K.; Sarkar, S.; Bhaumik, K. A possible mechanism of zero-crossing detection using the concept of the extended classical receptive field of retinal ganglion cells. *Biol. Cybern.* **2005**, *93*, 1–5. [CrossRef] [PubMed]

68. Anoop, B.; Joseph, J.; Williams, J.; Jayaraman, J.S.; Sebastian, A.M.; Sihota, P. A prospective case study of high boost, high frequency emphasis and two-way diffusion filters on MR images of glioblastoma multiforme. *Australas. Phys. Eng. Sci. Med.* **2018**, *41*, 415–427. [CrossRef]

69. Alirezanejad, M.; Saffari, V.; Amirgholipour, S.; Sharifi, A.M. Effect of locations of using high boost filtering on the watermark recovery in spatial domain watermarking. *Indian J. Sci. Technol.* **2014**, *7*, 517. [CrossRef]

70. Dai, Y. Building CNN-Based Models for Image Aesthetic Score Prediction Using an Ensemble. *J. Imaging* **2023**, *9*, 30. [CrossRef]

71. Babnik, Z.; Štruc, V. Assessing bias in face image quality assessment. In Proceedings of the 2022 30th European Signal Processing Conference (EUSIPCO), Belgrade, Serbia, 29 August–2 September 2022; pp. 1037–1041.

72. Fabbrizzi, S.; Papadopoulos, S.; Ntoutsi, E.; Kompatsiaris, I. A survey on bias in visual datasets. *Comput. Vis. Image Underst.* **2022**, *223*, 103552. [CrossRef]

73. Thomee, B.; Shamma, D.A.; Friedland, G.; Elizalde, B.; Ni, K.; Poland, D.; Borth, D.; Li, L.J. YFCC100M: The new data in multimedia research. *Commun. ACM* **2016**, *59*, 64–73. [CrossRef]

74. Saupe, D.; Hahn, F.; Hosu, V.; Zingman, I.; Rana, M.; Li, S. Crowd workers proven useful: A comparative study of subjective video quality assessment. In Proceedings of the QoMEX 2016: 8th International Conference on Quality of Multimedia Experience, Lisbon, Portugal, 6–8 June 2016.

75. Sheikh, H.R.; Sabir, M.F.; Bovik, A.C. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Trans. Image Process.* **2006**, *15*, 3440–3451. [CrossRef]

76. Seeger, M. Gaussian processes for machine learning. *Int. J. Neural Syst.* **2004**, *14*, 69–106. [CrossRef] [PubMed]

77. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [CrossRef]

78. Lou, Y.; Caruana, R.; Gehrke, J. Intelligible models for classification and regression. In Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China, 12–16 August 2012; pp. 150–158.

79. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [CrossRef]

80. Loh, W.Y. Regression tress with unbiased variable selection and interaction detection. *Stat. Sin.* **2002**, *12*, 361–386.

81. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

82. Zhang, L.; Zhang, L.; Bovik, A.C. A feature-enriched completely blind image quality evaluator. *IEEE Trans. Image Process.* **2015**, *24*, 2579–2591. [CrossRef]

83. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a "completely blind" image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [CrossRef]

84. Venkatanath, N.; Praneeth, D.; Bh, M.C.; Channappayya, S.S.; Medasani, S.S. Blind image quality evaluation using perception based features. In Proceedings of the 2015 Twenty First National Conference on Communications (NCC), Mumbai, India, 27 February–1 March 2015; pp. 1–6.

85. Moorthy, A.; Bovik, A. A modular framework for constructing blind universal quality indices. *IEEE Signal Process. Lett.* **2009**, *17*, 7.

86. Min, X.; Zhai, G.; Gu, K.; Liu, Y.; Yang, X. Blind image quality estimation via distortion aggravation. *IEEE Trans. Broadcast.* **2018**, *64*, 508–517. [CrossRef]

87. Moorthy, A.K.; Bovik, A.C. Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Trans. Image Process.* **2011**, *20*, 3350–3364. [CrossRef] [PubMed]
88. Chen, X.; Zhang, Q.; Lin, M.; Yang, G.; He, C. No-reference color image quality assessment: From entropy to perceptual quality. *EURASIP J. Image Video Process.* **2019**, *2019*, 77. [CrossRef]
89. Ou, F.Z.; Wang, Y.G.; Zhu, G. A novel blind image quality assessment method based on refined natural scene statistics. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; pp. 1004–1008.
90. Fix, E.; Hodges, J.L., Jr. Significance probabilities of the Wilcoxon test. *Ann. Math. Stat.* **1955**, *26*, 301–312. [CrossRef]