Expert Systems

WILEY

Expert Systems

SURVEY ARTICLE

No-Reference Image Quality Assessment: Past, Present, and Future

Qingyu Mao[1] | Shuai Liu[2] | Qilei Li[3] | Gwanggil Jeon[4] | Hyunbum Kim[4] | David Camacho[5]

[1]College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China | [2]College of Applied Technology, Shenzhen University, Shenzhen, China | [3]School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK | [4]Department of Embedded Systems Engineering, Incheon National University, Incheon, South Korea | [5]School of Computer Systems Engineering, Universidad Politecnica de Madrid, Madrid, Spain

Correspondence: Gwanggil Jeon (ggjeon@outlook.com)

Received: 23 September 2024 | **Revised:** 27 November 2024 | **Accepted:** 8 January 2025

Funding: This work was supported by Research Project Support Program for Excellence Institute (2022, ESL) in Incheon National University.

Keywords: deep learning | neural network | no-reference image quality assessment

ABSTRACT

No-reference image quality assessment (NR-IQA) has garnered significant attention due to its critical role in various image processing applications. This survey provides a comprehensive and systematic review of NR-IQA methods, datasets, and challenges, offering new perspectives and insights for the field. Specifically, we propose a novel taxonomy for NR-IQA methods based on distortion scenarios and design principles, which distinguishes this work from previous surveys. Representative methods within each category are thoroughly examined, with a focus on their strengths, limitations, and performance characteristics. Additionally, we review 20 widely used NR-IQA datasets that serve as benchmarks for evaluating these methods, providing detailed information on the number of images, distortion types, and distortion levels for each dataset. Furthermore, we identify and discuss key challenges currently faced by NR-IQA methods, such as handling diverse and complex distortions, ensuring generalisation across datasets and devices, and achieving real-time performance. We also suggest potential future research directions to address these issues. In summary, this survey offers a comprehensive and systematic examination of NR-IQA methods, datasets, and challenges, offering valuable insights and guidance for researchers and practitioners working in the NR-IQA domain.

1 | Introduction

With the rapid evolution of information technology, vast amounts of multimedia data, particularly images, are being transmitted globally through various networks, making image quality a vital factor in effective communication. Digital images undergo multiple stages from acquisition to final display, including capture, storage, compression, transmission, and processing. Each stage can introduce varying levels of distortions (Chandler 2013; Zhai and Min 2020). Beyond influencing the visual experience of users, image quality is crucial in determining the performance of various computer vision tasks (Dodge and Karam 2016; Lin and Wang 2018). Therefore, it is imperative to design precise algorithms or evaluation metrics to accurately quantify image quality degradation across both daily life and industrial applications.

Image Quality Assessment (IQA) is a fundamental research problem in human visual perception and computer vision, aiming to develop objective metrics and methods to evaluate image quality (Wang et al. 2004; Chen et al. 2023, 2022). The human visual system (HVS) excels at recognising high-quality images, making subjective assessments the most reliable in IQA. However, manual assessment is resource-intensive and impractical for real-world applications. Objective IQA methods, developed to automatically predict visual quality by simulating

HVS, are classified into three types based on the availability of reference information: (1) Full-Reference IQA (FR-IQA), (2) Reduced-Reference IQA (RR-IQA), and (3) No-Reference IQA (NR-IQA). Although existing FR-IQA (Kim and Lee 2017; Zhang, Isola, et al. 2018; Ding et al. 2020; Lao et al. 2022) and RR-IQA methods (Wu et al. 2013; Liu et al. 2014; Xu et al. 2015; Wu et al. 2016) have achieved satisfactory results by leveraging reference information, their usability is typically limited, as many distorted images lack corresponding references in real-world scenarios. In contrast, NR-IQA methods are more widely applicable across various practical applications, as they do not rely on reference images for quality assessment.

Recent decades have seen the emergence of various NR-IQA methods, demonstrating significant progress. Consequently, there are several surveys systematically review these studies across various aspects. In 2015, Manap and Shao (2015) presented a survey of general-purpose NR-IQA methods, primarily reviewing contemporary hand-crafted approaches. They categorised these methods into two types: natural-scene statistics (NSS)-based and learning-based. This study also examined their performance and limitations, and presented research trends to address these limitations. After that, Xu, Jiang, and Min (2017) comprehensively reviewed the fundamental developments in NR-IQA methods, covering both distortion-specific and general-purpose approaches. They provided a detailed discussion on techniques for feature extraction and quality prediction within general-purpose approaches. Additionally, they evaluated performance on three IQA benchmark datasets and highlighted challenges in NR-IQA development. In their 2019 survey, Yang, Li, and Liu (2019) emphasised the progress in DNN-based NR-IQA methods, systematically analysing these approaches based on the role of DNNs and comparing their performance on four synthetic datasets and one authentic dataset. Moreover, they described emerging challenges and directions for future investigations.

In 2023, Yang, Sturtz, and Qingge (2023) reviewed NR-IQA methods, covering well-known traditional and recent DNN-based approaches. They evaluated these competitive methods on six public IQA datasets and suggested potential future research directions based on their findings. These NR-IQA surveys have examined the field from various perspectives, primarily focusing on methods applicable to natural images. However, they encounter challenges in thoroughly exploring image quality models due to three key factors. First, the proliferation of non-traditional image types has created unique challenges. The increasing prevalence of immersive and interactive media, such as omnidirectional images, light field images, and screen content images, presents distinct distortion types and viewing conditions that are not adequately addressed by methods originally developed for natural images. Second, NR-IQA is finding applications in numerous domains, including image super-resolution, multi-exposure fusion, and tone-mapping. Each of these fields introduces specific requirements and constraints for image quality assessment. Conducting a comprehensive review is essential to understand how NR-IQA methods are being adapted or need to be adapted for these diverse scenarios. Third, the rapid progress in deep learning (DL) is revolutionising NR-IQA methodologies. These advancements enable more sophisticated feature extraction and better leveraging of large-scale datasets. Furthermore, techniques like meta-learning, contrastive learning, and multi-modal learning have been utilised to boost the generalisation performance of NR-IQA methods. A thorough examination of latest DL-based approaches is crucial for understanding the current state-of-the-art (SOTA) and identifying future research directions. These factors highlight the urgent need for a comprehensive survey to systematically identify novel directions and key challenges in the rapidly evolving field of NR-IQA. Figure 1 presents a concise timeline of the development of NR-IQA methods.

In this review, we offer a comprehensive examination of NR-IQA, including hand-crafted-based methods, DL-based methods, and representative IQA datasets (e.g., synthetic and authentic distortion). Specifically, we categorise current NR-IQA methods into three groups based on their focus: (1) synthetic distortion-oriented methods, which focus on artificial degradations; (2) algorithm distortion-oriented methods, which deal with artefacts introduced by various image processing algorithms; and (3) authentic distortion-oriented methods, which are designed for assessing the quality of in-the-wild images. This categorization offers structured approach to analysing the diverse landscape of NR-IQA techniques, enabling a thorough examination of how different methods address various distortion types encountered in practical applications. We discuss the challenges faced by NR-IQA methods in real-world scenarios and highlight the recent advancements in the field. Additionally, we offer insights into potential future research directions. This survey aims to provide researchers and practitioners with a comprehensive understanding of the current NR-IQA landscape and inspire further innovations in this fundamental field.

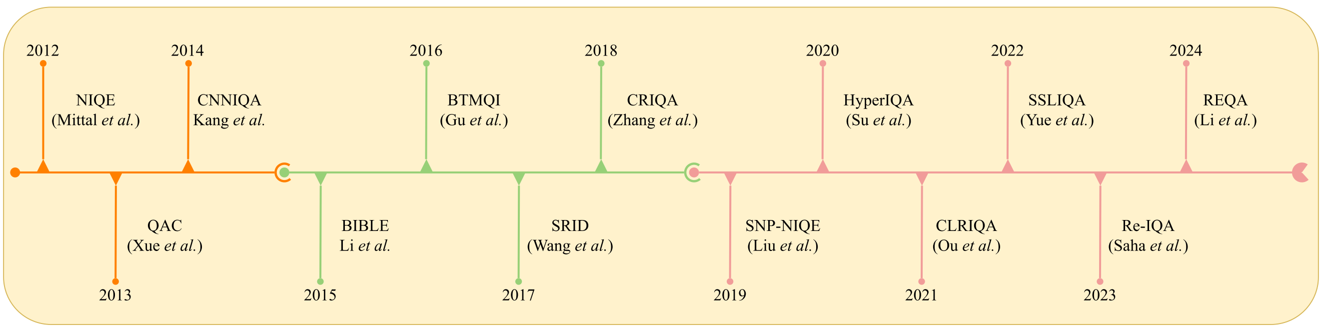


FIGURE 1 | A brief history of NR-IQA methods. These directions have shaped both current research and potential future developments. Synthetic distortion-oriented (in orange); algorithm distortion-oriented (in green); authentic distortion-oriented (in red).

The article is structured as follows: Section 2 provides a preliminary overview of NR-IQA. Sections 3–5 review both past and SOTA NR-IQA methods, categorising them into synthetic distortion-oriented, algorithm distortion-oriented, and authentic distortion-oriented approaches. Section 6.1 provides a detailed review of 20 widely used public datasets and shows some examples. Section 7 explores the key challenges and outlines future research directions in NR-IQA. Lastly, Section 8 concludes this article.

2 | Preliminaries

NR-IQA seeks to predict the visual quality of a distorted image I^d in the absence of its corresponding reference image I^r (Karray, Campilho, and Yu 2019). NR-IQA aims to obtain a quality score Q that aligns with human perceptual judgements. An NR-IQA model can be formally expressed as a function:

$$Q = f\left(I^d; \theta\right) \tag{1}$$

Here, $f(\cdot)$ represents the quality prediction function and θ represents the NR-IQA model parameters, learned from an IQA dataset containing distorted images and their corresponding subjective quality scores. Early NR-IQA methods mainly targeted the evaluation of images with synthetic distortions. These images are created by artificially introducing distortions such as Gaussian noise, salt-and-pepper noise, and JPEG compression artefacts to pristine or high-quality images. However, these methods often struggle to accurately evaluate real-world images, which exhibit more complex and diverse distortion characteristics. To overcome this limitation, researchers have shifted their focus towards developing NR-IQA methods specifically tailored for authentic distortions. The emergence of DL has greatly advanced progress in this area, enabling the development of more sophisticated and robust algorithms. Nevertheless, the scarcity of large-scale authentic distortion image quality assessment databases remains a significant challenge (Zhang, Ma, Yan, et al. 2018). Additionally, the complex interplay between image content and distortion types in real-world scenarios further compounds the difficulties faced by current algorithms.

NR-IQA methods can be categorised into three main groups based on the distortion types they evaluate (Yang, Sturtz, and Qingge 2023): (1) synthetic distortion-oriented methods, which focus on artificial degradations; (2) algorithm distortion-oriented methods, which deal with artefacts introduced by various image processing algorithms; and (3) authentic distortion-oriented methods, which are developed to assess quality of in-the-wild images. This article intends to offer a comprehensive overview of the NR-IQA research landscape, following the categorization illustrated in Figure 2. By examining the SOTA approaches in each category, we aim to showcase the current progress and challenges in NR-IQA. Particularly, this article emphasises NR-IQA methods that target authentic distortions, which are pivotal for unlocking the potential of image quality assessment in practical applications.

3 | Synthetic Distortion-Oriented NR-IQA

Synthetically distorted images refer to images whose visual quality has been degraded by artificially introducing noise or distortions. Common types of synthetic distortions include

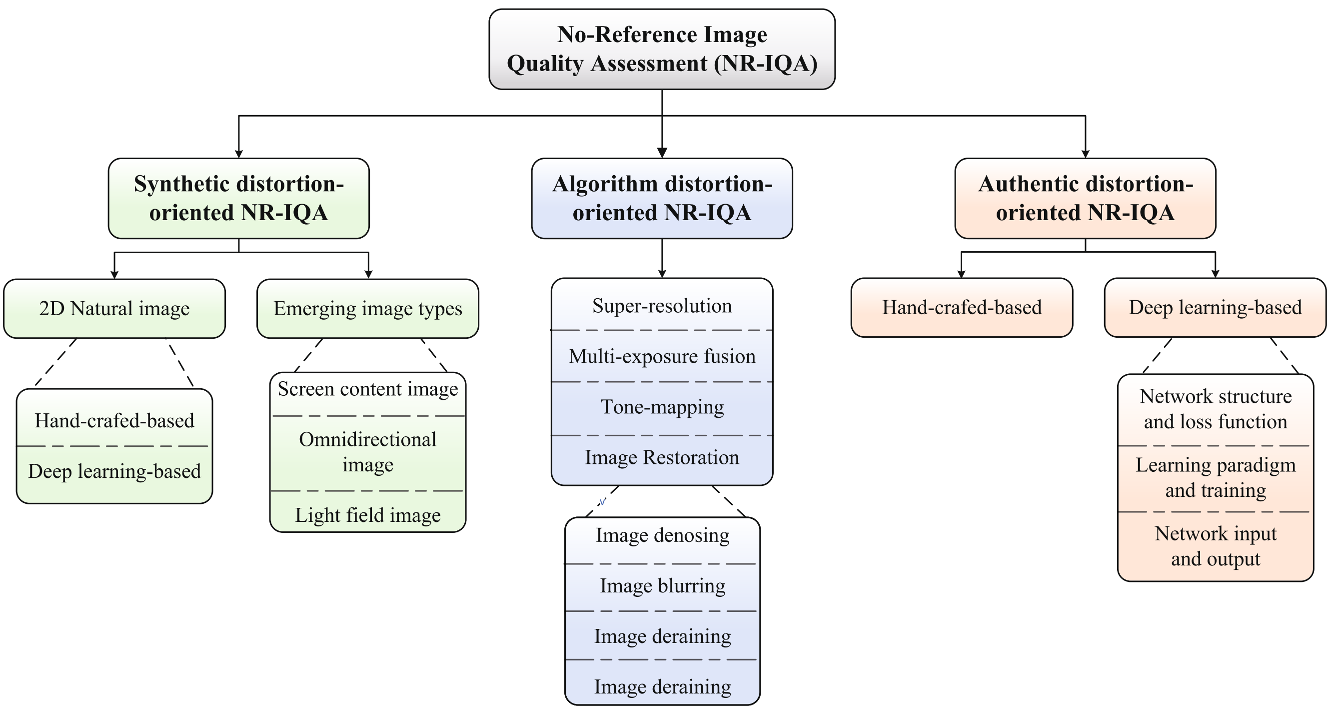


FIGURE 2 | The overall classification of NR-IQA methods. The second row divides the methods into three primary categories, which are further broken down into subcategories in the following row. The final row provides additional subdivisions based on specific image types, algorithm, and design principles.

Gaussian noise, salt-and-pepper noise, and compression artefacts. Early studies primarily focused on analysing synthetic distortions, leading to the development of synthetic distortion-oriented NR-IQA approaches, which can be broadly categorised into hand-crafted-based and DL-based methods. Figure 3 provides a concise illustration of the basic flowchart for these two types of methods. This section also covers emerging topics, such as NR-IQA methods for screen content images, light field images, and omnidirectional images.

3.1 | Hand-Crafted-Based Methods

Hand-crafted-based methods rely on feature engineering, which involves handcrafted feature extraction combined with machine learning algorithms for quality prediction. Among these methods, Natural Scene Statistics (NSS) (Simoncelli and Olshausen 2001) is a widely used modelling approach in hand-crafted-based methods. NSS is based on the assumption that pristine images exhibit specific statistical regularities, which are altered by distortions. Numerous NSS-based NR-IQA methods have been developed based on this theory. Mittal, Moorthy, and Bovik (2012) introduced BRISQUE, which utilises local Mean Subtracted Contrast Normalised (MSCN) luminance coefficients. These coefficients are fitted to the generalised Gaussian distribution (GDD) and asymmetric generalised Gaussian distribution (AGGD) to extract features from distorted images (Liu et al. 2019). Subsequently, Mittal, Soundararajan, and Bovik (2012) expanded their approach by targeting the multivariate Gaussian (MVG) model and proposed a new method called NIQE. NIQE extracts NSS features from both distorted and natural images, fits them to MVG models, and calculates the distance between the models to assess image quality. Zhang, Zhang, and Bovik (2015) developed ILNIQE to enhance generalisation, extracting natural image statistical features from various perspectives, including local structure, orientation, and colour. These features are used to construct and fit a MVG model as reference one. Then, ILNIQE then evaluates distorted image quality by calculating the distances between patch MVG models and the reference model. Additionally, some algorithms extend beyond distance metrics and incorporate the probability values of the fitted distributions as quality features, such as Fang et al. (2014) and Saad, Bovik, and Charrier (2012). However, the visual features of some images may not consistently conform to predefined statistical regularities, which limits the applicability of these NSS-based modelling approaches (Ghadiyaram and Bovik 2017).

Unlike the above methods, some researchers have utilised dictionary learning to extract visual features from images for NR-IQA. Ye et al. (2012) introduced CORNIA, which employs K-Means clustering to construct a codebook from image patches and utilises soft assignment encoding and max pooling for effective quality representation. Similarly, Xue, Zhang, and Mou (2013) proposed a quality-aware clustering (QAC) model that first uses a full-reference quality assessment model (Zhang et al. 2011) to obtain quality scores for each patch of the input image. The QAC model then identifies cluster centres based on the quality variations among patches, using these centres as a codebook to evaluate the overall quality.

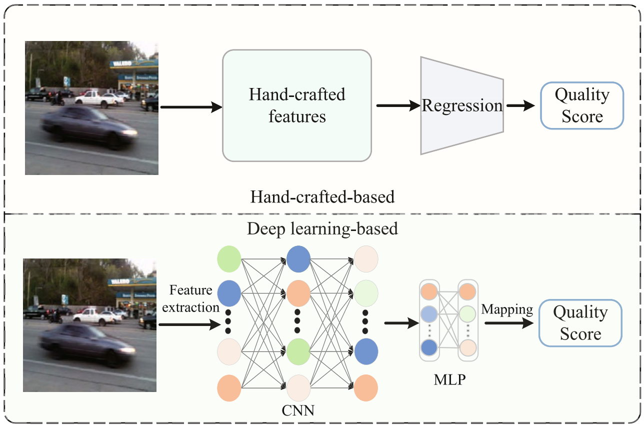


FIGURE 3 | The basic flowchart of hand-crafted-based and DL-based NR-IQA methods.

Furthermore, Xu et al. (2016) adopted a codebook construction approach similar to CORNIA and proposed HOSA, which estimates image quality by comparing the high-order statistics of local patch features with cluster centres. However, manually designed feature descriptors have limited capacity for representation, as they can only capture certain specific distortion types. Consequently, these hand-crafted-based methods are constrained in achieving better consistency with human perception.

3.2 | DL-Based Methods

DL-based methods eliminate the need for manually designed features and instead employ DNNs or other DL tools to extract image features and predict image quality directly, offering substantial improvements in feature representation and thereby outperforming traditional hand-crafted-based methods. For instance, Kang et al. (2014) introduced CNNIQA, a CNN-based NR-IQA model. The architecture includes a convolutional layer, a pooling layer, two fully connected layers for feature extraction, and a linear regression layer for quality prediction. Despite its relatively simple architecture, CNNIQA outperforms most traditional hand-crafted methods on the LIVE dataset (Sheikh, Sabir, and Bovik 2006). To enhance feature representation, Bianco et al. (2018) utilised a pre-trained CNN on image classification tasks for feature extraction and employed a support vector regression (SVR) model for quality prediction. Kim, Nguyen, and Lee (2018) proposed DIQA, a two-stage NR-IQA approach that first trains a CNN on normalised image patches to learn objective error maps and then fine-tunes the CNN to predict image quality. To overcome the limited training data, Liu, van de Weijer, and Bagdanov (2017) introduced RankIQA, which creates a large synthetic dataset of distorted images with relative quality labels and uses a Siamese network for quality-based image ranking. After ranking, one branch of the Siamese network is fine-tuned for quality prediction. Wu, Ma, et al. (2020) proposed the CaHDC model using cascaded CNNs. They first constructed a large-scale image quality assessment dataset using synthetic distortions and assigned pseudo-labels to all distorted images based on five classical full-reference metrics. The CaHDC model is then trained on this extensive dataset.

Unlike the above works, Lin and Wang (2018) integrated a generative adversarial network (GAN) with a CNN to proposed Hallucinated-IQA. Hallucinated-IQA comprises three components: quality-aware generative, quality discrimination, and a quality regression network. The generative network creates pseudo-reference images corresponding to the distorted ones, the discrimination network differentiates between real and pseudo-reference images, and the regression network evaluates the quality of distorted images based on the discrepancy between the pseudo-reference and distorted images. Meanwhile, Ren, Chen, and Wang (2018) also introduced RAN, a GAN-based NR-IQA method. Unlike Hallucinated-IQA, the RAN model leverages the comparison between restored and distorted images to predict quality more accurately. To better evaluate the quality of images with synthetic distortion, some researchers have explored incorporating multi-task learning into NR-IQA. For instance, Ma, Liu, et al. (2017) proposed MEON built on multi-task optimization. MEON employs a dual-sub-network architecture to tackle two related tasks simultaneously: predicting distortion types and assessing image quality. These sub-networks share parameters in their primary layers, improving both performance and efficiency. First, the distortion type prediction sub-network is trained on a distorted image dataset. Then, the quality prediction sub-network is jointly trained, benefiting from the learned distortion features to enhance predictive accuracy. Similarly, Yang, Jiang, et al. (2019) proposed SGDNet (Saliency-Guided Deep neural Network), which takes saliency prediction as an auxiliary task to augment image quality prediction. This model employs a shared feature extractor for both saliency and quality prediction tasks during the initial training stages, leveraging the interdependency of these tasks to improve overall assessment accuracy. The aforementioned deep learning-based methods demonstrate excellent performance in assessing synthetically distorted images, but their focus on 2D natural images means that they are not effective for other emerging image and distortion types.

3.3 | Emerging Topics

As multimedia technology has advanced, the diversity in image types and formats has significantly increased, prompting researchers to expand the scope of IQA beyond traditional 2D natural images. Current research in IQA now consider a variety of specialised image formats, such as screen content images (SCIs), light field image (LFIs), and omnidirectional image (OIs). Each of these formats presents unique challenges and necessitates tailored assessment techniques.

SCIs refer to images displayed on electronic devices, such as mobile phones and computer monitors, which mainly contain both pictorial and textual content. Popular screen content synthetic distortion image quality assessment datasets include SIQAD (Yang, Fang, and Lin 2015), SCID (Ni et al. 2017), and QACS (Wang et al. 2016; Shi et al. 2015). To effectively assess the quality of SCIs, several no-reference screen content image quality assessment (NR-SCIQA) methods have been developed. For example, Gu et al. (2017) extracted visual features from four key aspects: complexity, content statistics, global brightness, and detail sharpness, and then employed a quality regression model for prediction. Fang et al. (2017) leveraged the HVS's heightened sensitivity to brightness and texture to model and extract brightness statistical features and high-order reciprocal texture features in SCIs. Zheng et al. (2019) introduced an NR-SCIQA approach that segments SCIs into sharp and non-sharp edge regions using local standard deviation. Local features, such as entropy, contrast, and local phase consistency, are extracted from the sharp edge regions using the grey-level co-occurrence matrix, while brightness statistics across the entire image are used to derive global features. Chen, Shen, et al. (2018) proposed a naturalisation module that uses interpolation to minimise the characteristic differences between natural images and SCIs, bridging the gap between them. They also tailored a quality prediction model, PICNN, specifically for SCIs. To address the unsupervised domain transfer problem in NR-SCIQA, Chen et al. (2021) jointly trained the model by maximising the average difference, employing rank learning, and minimising mean squared error, enhancing the knowledge transfer from natural images (source domain) to SCIs (target domain) for quality prediction.

LFIs distinct from both natural images and SCIs, are captured by light field cameras that record comprehensive data on the angle and position of light rays. Existing representative datasets for light field image quality assessment include MPI-LFA (Kiran Adhikarla et al. 2017), VALID (Viola and Ebrahimi 2018), and Win5-LID (Shi, Zhao, and Chen 2019). LFIs contain rich scene information, making it more challenging to assess their quality. Therefore, relatively few NR-IQA methods have been proposed for LFIs. One pioneering approach is BELIF (Shi, Zhao, and Chen 2019), which uses tensor theory to process LFIs by transforming a raw LFI into a circular image tensor and applying Tucker decomposition for feature extraction. The extracted features, namely tensor spatial features and tensor structure change indices, assess spatial quality and angular consistency, respectively. Another innovative method is (Tian et al. 2020) that utilises multi-scale log-Gabor features from different planes to represent the quality of LFIs.

As a key form of VR media content, OIs significantly affect the visual experience of VR users. Xu, Zhou, and Chen (2020) proposed a viewport-oriented model called VGCN, which selects high-probability viewports as graph nodes and models the relationships between different viewports. The VGCN employs a graph convolution network for graph inference and integration to compute the overall image quality score. Duan et al. (2018) constructed an OI quality dataset that includes non-uniform distortions, user viewing conditions, and viewing behaviours. Utilising this dataset, they trained an NR-OIQA model consisting of a multi-scale feature extraction module and a module for perceptual quality prediction. Zhou et al. (2021) employed multi-task learning and introduced a distortion identification task to enhance the quality prediction of OIs, improving the model's performance.

4 | Algorithm Distortion-Oriented NR-IQA

Apart from common synthetic distortions, many image processing algorithms could produce new distortions during the process, which leads image quality degradation. Meanwhile, there are also numerous algorithms aimed at removing distortions in images to improve their quality. To analyse the characteristics of different image processing algorithms and evaluate the quality of images after algorithmic processing, algorithm distortion-oriented NR-IQA methods have been gradually developed. These methods primarily focus on image enhancement algorithms, such as super-resolution, multi-exposure fusion, tone mapping, and various image restoration algorithms. In the following section, we briefly introduce the NR-IQA methods designed for these algorithms.

4.1 | NR-IQA for Super-Resolution

Super-resolution (SR) focuses on reconstructing high-resolution (HR) images from low-resolution (LR) inputs, thereby enhancing image quality. Due to the differences in the techniques and processing methods employed by existing SR algorithms, the quality and visual effects of the reconstructed HR images vary considerably. To better analyse and compare different SR algorithms, it is essential to perform quality assessment on the images generated by these algorithms. To facilitate the quality assessment of SR images, researchers have constructed various SR image quality assessment datasets, including SRID (Wang et al. 2017), NID (Chen, Xu, et al. 2018), QADS (Zhou et al. 2019), SISR-Set (Shi, Wan, et al. 2019), SISAR (Zhao et al. 2021), and SUPE (Köhler et al. 2019). Based on these datasets, extensive research on SR images have been proposed. For example, Ma, Yang, et al. (2017) propose designing low-level statistical features in spatial and frequency domains to quantify super-resolved artefacts and learning a two-stage regression model from visual perceptual scores to predict quality scores. To further enhance the precision of SR image quality prediction, Zhang et al. (2019) adopted the same feature extraction approach as Ma, Yang, et al. (2017) and proposed a refined quality prediction model that cascades AdaBoost decision tree regression and ridge regression. Beron, Benitez-Restrepo, and Bovik (2020) proposed two NR-IQA models for SR images by selecting the most optimal quality-aware features. Unlike hand-crafted feature extraction methods, Fang et al. (2018) extract high-level representation features from each patch, utilising local information to predict SR image quality. Zhou et al. (2020) proposed DeepSRQ based on a dual-stream CNN. DeepSRQ takes distorted SR images and their texture structures as inputs, extracting discriminative features through the dual-stream architecture.

4.2 | NR-IQA for Tone-Mapping

Tone mapping is the process of converting high dynamic range (HDR) images into low dynamic range (LDR) images using specific operators, enabling HDR content to be better displayed on standard monitors. However, different tone-mapping operators introduce varying degrees of distortion when processing HDR images. Several datasets and algorithms have been developed to assess tone-mapped image quality. The main datasets include TMID (Yeganeh and Wang 2012) and ESPL_LIVE HDR (Kundu et al. 2017). In terms of algorithms, Gu et al. (2016) proposed BTMQI, which assesses tone-mapped images by measuring their information entropy and integrating natural statistics with structural fidelity. Kundu et al. (2017) proposed HIGRADE, which extracts spatial NSS and gradient features from tone-mapped images to evaluate their quality. BLIQUE-TMI (Jiang et al. 2017) models visual, local structural, and natural statistical information from tone-mapped images, employing an extreme learning machine for assessment. Chen et al. (2019) introduced a method based on luminance segmentation, which leverages the HVS's varying sensitivity to different luminance regions to quantify quality. VQGC (Fang, Yan, et al. 2020) utilises gradient and chromatic statistical information from tone-mapped images to predict quality. In DL-based approaches, Deep-TMI (Ravuri et al. 2019) first utilises a CNN to generate a quality map of the tone-mapped image, selects feature parameters using the AGGD model, and finally employs a SVR to predict quality scores from these features.

4.3 | NR-IQA for Image Restoration

Image restoration refers to the process of recovering and reconstructing a relatively high-quality image from a distorted one using image restoration algorithms. Common image restoration algorithms like denoising, deblurring, deraining, and dehazing. The inputs to these restoration algorithms differ, leading to significant variations in the quality of the restored images.

For denoised images, Kong et al. (2013) assumes that noise is independent of the pristine image and seeks to maximise the SSIM between the noise image and the estimated noise in both homogeneous and highly structured regions. The linear correlation coefficient between the two SSIM maps is then computed to quantify quality. Zhang, Cheng, and Hirakawa (2018) proposed CRIQA that uses a corrupted reference image instead of a high-quality one and achieves results consistent with full-reference methods.

For deblurred images, Narvekar and Karam (2011) proposed a method according to how humans perceive blur at varying contrast levels. This approach first uses a probabilistic model to estimate the likelihood of edge blur in the image, and then aggregates this data by cumulatively integrating the blur detection probabilities. Li et al. (2015) proposed BIBLE, which estimates Tchebichef moments from the gradients of the blurred image, then calculates the variance-normalised moment energy guided by a visual saliency model to assess image blur. Subsequently, based on BIBLE, they utilised NSS to model the multidimensional features of blurred images in both spatial and frequency domains (Li et al. 2017). To accurately measure image blur, Li, Li, et al. (2020) extracted multi-scale features and motion blur characteristics from blurred images, and fused these features through an attention module. Finally, they employed two stacked fully connected layers to obtain prediction scores.

To effectively analyse the quality of derained images, Wu et al. (2019) first constructed a derained image dataset and then proposed B-GFN based on bidirectional controlled fusion. B-GFN can adaptively extract and integrate multi-scale features to describe artefacts in derained images. Furthermore, Wu, Wang, et al. (2020) integrated global perception and local difference features into B-GFN and proposed the B-FEN model (bidirectional feature embedding network).

Considering the characteristics of dehazed images, Choi, You, and Bovik (2015) introduced FADE based on NSS theory and haze-aware statistical features. FADE estimates the perceived haze density globally and provides a local index for each image block. Min et al. (2018) established a dehazed image dataset for quality assessment and proposed a NR-IQA index called DHQI by extracting and fusing dehazing, structure preservation, and enhancement features.

To summarise, the necessity to evaluate the performance of various image restoration algorithms has driven the development of NR-IQA methods specifically designed for restored images. These methods typically extract features that capture the specific characteristics of each restoration algorithm and use them to predict the quality of the restored images. The availability of dedicated datasets for each restoration algorithm has further facilitated the advancement and validation of these methods. As image restoration algorithms continue to evolve, it is expected that more advanced and specialised NR-IQA methods will be proposed to meet the growing demands for accurate and reliable quality evaluation in various application scenarios.

5 | Authentic Distortion-Oriented NR-IQA

Unlike synthetic distortions and algorithm distortions, authentic distortions refer to various types and intensities of distortions introduced into images during acquisition due to factors such as the acquisition device, acquisition method, and acquisition environment. Authentically distorted images often come from various real-world scenes, making their content and distortion types more complex. Hence, quantifying the degree of degradation in these images is challenging. Early NR-IQA methods, such as BRISQUE (Mittal, Moorthy, and Bovik 2012), NIQE (Mittal, Soundararajan, and Bovik 2012), CORNIA (Ye et al. 2012), and HOSA (Xu et al. 2016), have good evaluation capabilities and interpretability for specific distortion types. However, these methods struggle to accurately assess the quality of authentically distorted images. To accurately and effectively assess the quality of various authentically distorted images, a growing number of researchers have started developing specific models or algorithms tailored for authentic distortions. Similarly, existing authentic distortion-oriented NR-IQA methods can also be divided into hand-crafted-based and DL-based. In this section, we begin with a brief overview of hand-crafted-based NR-IQA methods designed for authentic distortions.

The core idea of hand-crafted-based NR-IQA methods for authentically distorted images is to model image features through manually designed approaches. Specifically, Li et al. (2016) proposed a NR-IQA method called NRSL that uses statistical structure and luminance features. This method extracts perceptual structure features from the distorted image by utilising the distribution of local binary patterns. Additionally, it extracts luminance features by calculating the distribution of normalised luminance magnitudes. After extracting the structure and luminance features, a SVR model is employed to predict the quality of the image. Liu and Liu (2017) proposed WPDSE based on a model selection algorithm. WPDSE models image features from various aspects such as luminance, contrast, colour, and texture, and trains corresponding quality regression models. It then simplifies the overall model using a model selection algorithm and finally obtains the image quality by integrating the prediction results of all quality regression models. Zhang et al. (2020) used colour moments and log-Gabor layers to assess authentically distorted images. They first convert the distorted image to the HSV colour space to extract colour moments and then use log-Gabor filters to divide the image into four layers and extract texture features from each layer. Liu et al. (2020) proposed a non-subjective-aware NR-IQA method that designs a set of NSS features and HVS perception-related features to represent image quality. The method uses an MVG model learned on a collection of original images as a reference and infers the quality of a new image by computing the difference between the MVG model of the test image and the reference MVG model. Hu et al. (2021) used low-level attribute features, including brightness, saturation, contrast, noise, sharpness, and naturalness, along with high-level semantic features to describe the quality of authentically distorted images. By leveraging the complementarity of these low-level and high-level features, more accurate quality prediction is achieved when assessing image quality. Liu et al. (2019) proposed an unsupervised method called SNP-NIQE based on image structure, naturalness, and perceptual quality changes. SNP-NIQE captures structural changes by analysing deviations in image phase congruency and gradient distribution. It characterises naturalness changes using the distribution of MSCN coefficients and their adjacent pair products. Finally, the quality of the distorted image is represented by calculating the distance between the MVG models of the distorted image and natural images. DFE (Yang, An, and Shen 2022) employed a data-driven transform-based method to enhance features and combines hand-crafted features with learned features. This method extracts image structure information from the Karhunen–Loéve transform (KLT), phase congruency, and gradient magnitude coefficients, and then derives NSS features from local normalised coefficients. KLT is used as a feature enhancement process to improve the structure and NSS features. Finally, the distribution of transform coefficients across all frequency bands is modelled using the Weibull and generalised Gaussian distributions to achieve image quality perception. Due to the complexity of distortions in real-world scenarios, these methods often struggle to perform well on authentically distorted images.

Existing DL-based NR-IQA methods for authentically distorted images can be roughly divided into three categories according to their implementation approaches and the techniques used: (1) network structure and loss function; (2) learning paradigm and training data; (3) network input and output. Figure 4 provides an illustration of three design pipelines for DL-based authentic distortion-oriented NR-IQA methods.

5.1 | Network Structure and Loss Function

DL-based IQA methods generally train DL models through specified loss functions, extract image features end-to-end, and learn the map between features and quality scores. The representational capabilities of different DL models also vary, and it is necessary to perform fine-grained network structure design to effectively represent various authentically distorted images. For example, Bosse et al. (2017) proposed WaDIQaM

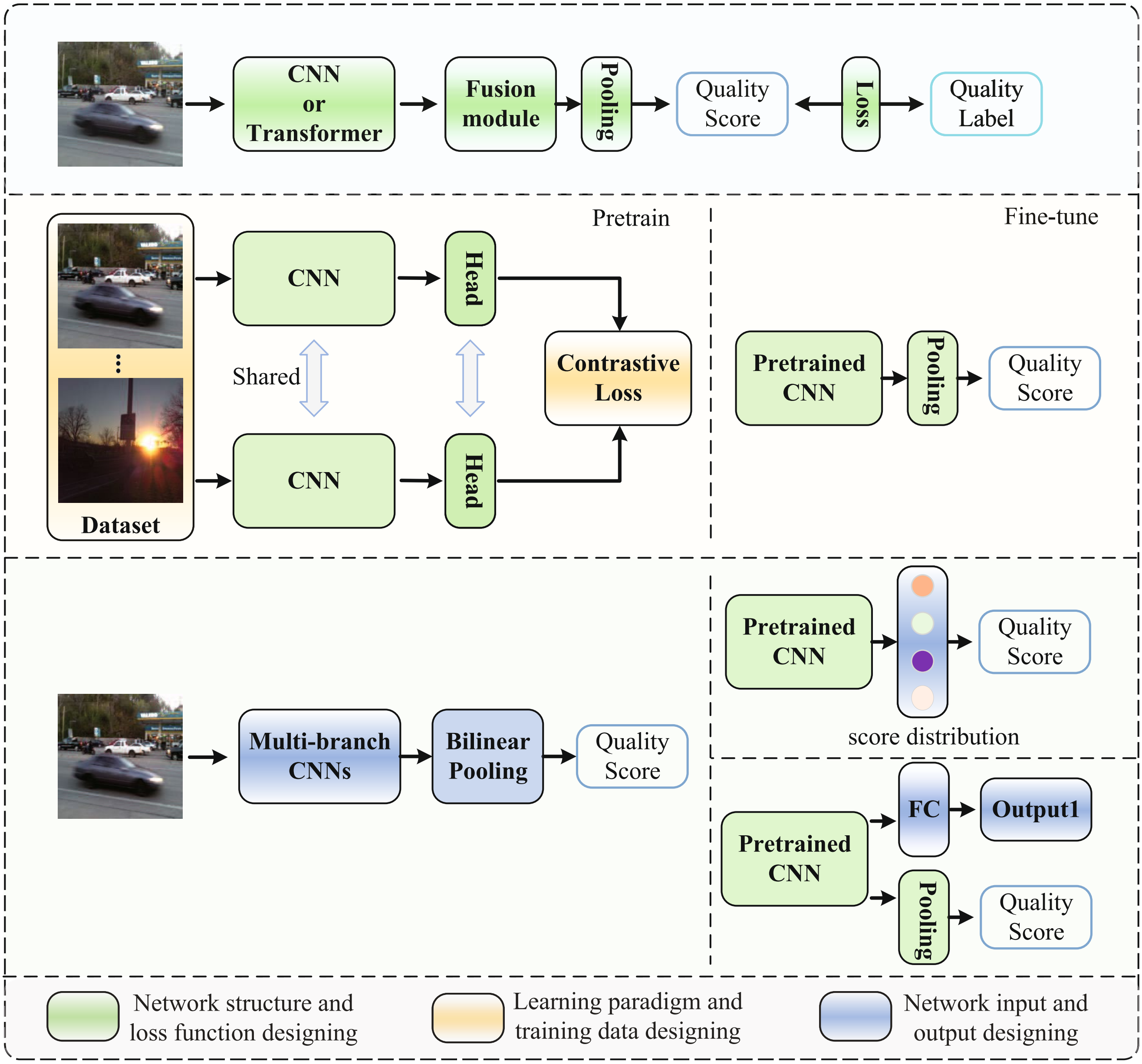


FIGURE 4 | A summary of three design pipelines for DL-based authentic distortion-oriented NR-IQA methods, which include network structures, loss functions, learning paradigms, and input–output strategies.

based on a weighted average and patch-wise joint optimization model. The feature extraction network of this model follows the structure of VGG16 and has been extended. WaDIQaM simultaneously realises FR-IQA and NR-IQA and can learn the relative importance of local to global quality. Sun et al. (2023) designed a step-wise feature extraction network that integrates multi-levels information extracted by the CNN into a comprehensive quality-aware representation in a hierarchical and step-wise manner. You and Korhonen (2021) applied the visual Transformer to the NR-IQA task and proposed the TRIQ model. TRIQ directly uses the encoder of the ViT model to process the visual features from CNN and designs an adaptive position encoding method to allow the model to assess images of arbitrary resolution. The MUSIQ (Ke et al. 2021) takes multi-scale images as input and designs hash-based two-dimensional spatial embedding and scale embedding to accommodate multi-scale feature representations, enabling the model to effectively capture image quality at different scale. Golestaneh, Dadsetan, and Kitani (2022) designed a hybrid feature extraction framework called TReS. TReS first utilises a CNN to extract local structural details, then uses a Transformer network to model the extracted global information, and finally the local and global features are integrated to assess image quality.

In addition to methods that focus on designing the overall network structure, there are also research works that aim to improve the expressive power of features by designing feature fusion modules or attention modules. For example, since the HVS has different visual perceptions for various distortions, VIDGIQA (Guan et al. 2017) utilises a regression network to estimate the visual significance of different local regions within the image. The entire framework is jointly optimised based on distortion information and quality scores. Su et al. (2020) proposed HyperIQA, which decomposes quality assessment into three phases: content analysis, perceptual rule adaptation, and quality estimation. Initially, a CNN extracts both semantic and multi-scale content features from the image. These semantic features are passed into the hypernetwork, which dynamically adjusts perceptual rules to produce weights for the quality estimation module. Subsequently, the local distortion perception module aggregates the multi-scale content and semantic features into feature vectors. Finally, these vectors are utilised to predict the image quality. Li, Zhang, and Cosman (2021) introduced MMMNet, which leverages multi-scale and multi-hierarchy fusion, drawing on the characteristics of human visual system (HVS) attention. The network uses a multi-scale feature extraction module (MSFE) to hierarchically fuse saliency and quality features. Furthermore, a saliency detection task is incorporated to enhance the quality estimation process through saliency fusion, with both tasks being jointly optimised using a multi-task learning framework. Pan, Zhang, et al. (2022) proposed the DACNN model, which comprises three key modules: distortion perception, distortion fusion, and quality prediction. The distortion perception module is further divided into two components: synthetic and authentic distortion perception. The synthetic distortion perception network is pre-trained using a Siamese architecture, while the authentic distortion perception network is pre-trained on classification dataset. These networks extract features corresponding to synthetic and authentic distortions, respectively. The fusion module employs a weight-adaptive fusion strategy to combine the extracted distortion features, enabling the model to focus on the most relevant information. Finally, the fused features are used to estimate image quality. Chen et al. (2024) proposed TOPIQ that mimics the HVS perception process from global to local. This method first extracts multi-scale features and employs a gated local pooling module to remove redundant information. Then, TOPIQ employs self-attention module and cross-scale attention mechanisms in a top-down approach, progressively fusing global semantic features into local distortion features layer by layer. This enables the network to concentrate on distortion areas that are semantically more significant.

When making quality predictions, DL-based NR-IQA methods generally need to use max or average pooling to convert the extracted image features into feature vectors. However, these simple pooling methods cannot capture higher-order statistical information of the feature descriptors. Jiang et al. (2020) proposed Deep-DEN to address this issue. The Deep-DEN integrates dictionary encoding within a single trainable layer, which is used to extract higher-order statistical information of the features and then use the higher-order statistical information to assist in image quality prediction. To tackle the issue that GAP can only extract first-order statistics, Zhou, Xu, et al. (2022) employed second-order global covariance pooling (GCP) to aggregate feature maps. By combining GCP and GAP, they obtained a global representation that is more sensitive to distortions. Gu et al. (2019) proposed a learnable pooling mechanism that facilitates the simultaneous learning of local quality and importance weights while dynamically assigning visual priorities and predicting overall quality. Additionally, the network regulates training by penalising cases where the quality of more prominent regions significantly deviates from the overall score.

Loss functions are an indispensable part of DL models. During the training procedure, DL models need to be guided by loss functions. Choosing or designing appropriate loss functions for different tasks speed up model convergence and improve overall performance. Therefore, some researchers have designed specialised loss functions for IQA to enhance the evaluation capabilities of the models. For example, Wu et al. (2017) proposed a regression model with rank regularisation to tackle the issue that MSE loss fails to accurately rank image quality. This model is achieved by introducing ranking constraints into the maximum margin-based regression framework. Ou et al. (2021) proposed CLRIQA, which first models overexposure and underexposure as inverse functions based on the Weber-Fechner law using an imaging heuristic approach. Next, it simulates authentic distortions and produces ranked image samples via a fusion strategy and compression. A controllable list-wise ranking (CLR) loss function is also proposed, where ranking bounds are defined, and an adaptive margin is introduced to fine-tune the intervals between ranks. Finally, the CNN is trained using both the produced samples and the CLR loss. Golestaneh, Dadsetan, and Kitani (2022) measured the distances between images in every batch to capture their ranking relationships and proposed a ranking loss based on these comparisons. Additionally, they formulated a self-consistency loss by leveraging the consistency between images and their augmented versions. Li, Jiang, and Jiang (2020) proposed a Norm-in-Norm Loss by computing the discrepancy between normalised predictions and normalised labels. It has been theoretically proven that this loss function improves gradient stability and accelerates model convergence. Li and Huo (2024) proposed REQA that employs global and local features to perceive distortions at multi-scales. It gradually perceives distortions by introducing a feedback mechanism consistent with the HVS. Additionally, coarse and fine-grained losses are introduced to refine the perception. The coarse-grained ranking and gradient losses are employed to ensure consistency in both ranking and gradients between the predicted quality and the labels (Jiang et al. 2021). Simultaneously, the fine-grained multi-level tolerance loss follows a curriculum learning strategy for fine-grained prediction.

5.2 | Learning Paradigm and Training Data

Designing learning paradigms and training data refers to improving the model's perceptual ability for authentic distortions by adopting other learning paradigms and different training data construction methods. Many new learning paradigms, such as transfer learning, reinforcement learning, meta-learning, contrastive learning, and domain adaptation, have been successfully applied in various tasks. Researchers have investigated different learning paradigms for NR-IQA. For example, Zhu et al. (2020) proposed MetaIQA, which applies meta-learning to address the limited generalisation ability of existing IQA methods that rely on pre-trained CNNs for evaluating diverse distortions. This approach first acquires prior knowledge shared across different distortions, then transfers this knowledge to authentic distortions by fine-tuning the quality prior model. Tang et al. (2021) employed supervised contrastive learning to capture effective quality-aware representations. During training, sample pairs with shared quality labels are generated via data augmentation, and a quality-aware contrastive loss is applied. This loss function clusters samples with similar quality in the embedding space while pushing apart those with differing quality. Similarly, Madhusudana et al. (2022) adopted elf-supervised learning to obtain quality-aware representations, with distortion type and intensity prediction as auxiliary tasks. Contrastive losses are constructed to distinguish the type and intensity of synthetic distortions and the quality of authentically distorted images. A CNN is pre-trained on an unlabeled mixed dataset. In the quality assessment stage, the pre-trained CNN is frozen, and only the quality prediction model is fine-tuned to map features to labels. Saha, Mishra, and Bovik (2023) proposed a mixture of experts framework called Re-IQA and trained two independent CNN encoders based on this framework. Specifically, Re-IQA leverages the MoCo-v2 framework to train a quality-aware encoder for low-level quality features and a content-aware encoder for high-level content features. These features are then concatenated and fed into a single-layer regressor for quality prediction. Zhao et al. (2023) hypothesized that patches within a single distorted image should have similar quality, whereas patches from different distorted images should display distinct quality levels. Based on this premise, they introduced a quality-aware contrastive loss. Moreover, a complex image degradation scheme was designed to simulate the real degradation process of images. This scheme includes mixed distortion degradation, shuffling order, higher-order degradation, and skip operations, forming a degraded image space of approximately 2×10^7 in size. To fully exploit the quality-aware information concealed within the vast quantity of degraded images, a NR-IQA model called QPT was trained using a self-supervised contrastive learning approach. Agnolucci, Galteri, et al. (2024) developed ARNIQA that uses contrastive learning to increase the similarity between image patches affected by the same type of degradation. Wang, Chan, and Loy (2023) employed a prompt learning method based on the CLIP model. By designing a prompt pairing strategy to leverage the prior knowledge in the CLIP model, quality-aware and abstraction-aware perceptions of images were achieved, while realising zero-sample image quality assessment. Recent research (Huang et al. 2024; Agnolucci, Galteri, and Bertini 2024; Wu et al. 2023; Chen et al. 2025) has shifted towards the application of large multi-modal (LMM) models for IQA, as these models integrate visual and text information to capture complex relationships between image content and quality and achieve results that better align with human perception.

The data used for training DNNs also has a significantly influences on their performance and generalizability. To address the issue of insufficient samples in existing datasets, various approaches have been progressively developed, emphasising the design of training data. Zhang et al. (2021) presented a multi-dataset mixed training method and proposed UNIQUE. This method initially selects image pairs from various quality assessment datasets and computes the likelihood that the first image in each pair possesses superior quality. Subsequently, the proposed model undergoes optimization using a substantial number of image pairs by employing the fidelity loss. Simultaneously, the uncertainty estimation is regularised during the optimization process through the enforcement of a hinge constraint. Sun et al. (2023) proposed a step-structured feature extraction network, and trained the network on a mixed distortion dataset. With the mixed training strategy, the assessment capability and generalisation ability of the model are enhanced from two perspectives: improving the effectiveness of features and increasing the diversity of content and distortions in the training samples. Yue et al. (2022) proposed SSLIQA based on a dua-branch CNN with semi-supervised learning strategy, which transfer sample prediction consistency from the teacher CNN to the student CNN. Subsequently, a substantial number of unlabeled training samples are utilised to enhance the effectiveness and generalizability of the model.

5.3 | Network Input and Output

Designing the input and output of the network refers to improving the model's assessment ability by designing the input images, input branches, output features, or output tasks. For DNNs, in addition to the network structure, network modules, loss functions, learning paradigms, and training data, the input and output are also very important. Therefore, some researchers have focused on this direction from multiple perspectives to design NR-IQA methods. For instance, Zhang, Ma, Yan, et al. (2018) proposed DBCNN, which consists of two branches to separately address synthetic and authentic distortions. The synthetic branch is pre-trained on a large-scale dataset using distortion type and intensity classification. Meanwhile, the authentic branch leverages a pre-trained VGG16 network. After pre-training, the features from both branches are fused through bilinear pooling, and the model is fine-tuned on IQA datasets. Zhou, Lang et al. (2022) used three different CNNs to extract different image features, where a modified VGG19 network and a VGG16 network are used to extract bottom texture information and edge local information, respectively, and extract high-level semantic information with a ResNet50 network. Subsequently, a feature fusion module leveraging attention mechanisms is employed to integrate the three types of features for predicting quality scores. Zhang, Shao, and Li (2022) introduced a GAN-based approach consisting of a quality-aware network and a quality regression network to address the challenge of handling varying distortions with a single model. The quality-aware network simulates the distortion information, while the quality regression network learns the mapping from features to quality scores. Pan, Yuan, et al. (2022) proposed VCRNet, a visual compensation restoration network composed of two main components: a visual restoration network and a quality estimation network. To accurately reconstruct restored images, the visual restoration network integrates a visual compensation module and an asymmetric residual block (Lin and Wang 2018), and is trained using a hybrid loss function based on error maps. The quality estimation network then utilises the features extracted by the visual restoration network to predict image quality. Inspired by the HVS's global and local perceptual abilities, Yao et al. (2022) and Zhou et al. (2024) proposed OLN, which consists of a global distortion perception module and a local distortion observation module. The global distortion perception module obtains a global perception of image quality by classifying the distortion category and level of the image, while the local distortion observation module extracts local information of the image by simulating the observer's approach. During quality prediction,

TABLE 1 | Summary of basic statistics for 20 widely used IQA databases.

Category	Dataset	Reference	Distortion	Type	Level	Resolution
	LIVE (Sheikh, Sabir, and Bovik 2006)	30	779	5	5 or 4	768×512
	TID2008 (Ponomarenko et al. 2009)	25	1700	17	4	512×384
	TID2013 (Ponomarenko et al. 2015)	25	3000	24	5	512×384
	CSIQ (Larson and Chandler 2010)	30	866	6	5 or 4	512×512
	IRCCyN/IVC (Le Callet and Autrusseau 2005)	10	54	4	5	512×512
	MICT (Horita et al. 2011)	14	168	2	6	768×512
	A57 (Chandler and Hemami 2007)	3	54	6	3	512×512
Synthetic	WED (Ma et al. 2016)	4744	94,880	5	—	—
	KADID-10K (Lin, Hosu, and Saupe 2019)	81	10,125	25	5	512×384
	WIQ (De Simone et al. 2009)	7	80	—	—	512×512
	VCL@FER (Sazzad, Kawayoke, and Horita 2008)	23	552	4	6	683×512
	ESPL (Kundu and Evans 2015)	25	50	5	4	512×512
	LIVEMD (Jayaraman et al. 2012)	15	405	2	—	1280×720
	MDID2013 (Gu et al. 2014)	12	324	—	—	768×512, 1280×720
	MDID2016 (Sun, Zhou, and Liao 2017)	20	1600	—	—	512×384
	CID2013 (Virtanen et al. 2014)	0	480	—	—	1600×1200
	CLIVE (Ghadiyaram and Bovik 2015)	0	1162	—	—	500×500
Authentic	KonIQ-10K (Hosu et al. 2020)	0	10,073	—	—	1024×768
	SPAQ (Fang, Zhu, et al. 2020)	0	11,125	—	—	960×720–4608×3456
	FLIVE (Ying et al. 2020)	0	39,810	—	—	500×500–6144×4096

bilinear pooling is used to fuse the global perception and local information to achieve accurate perception of image quality.

The above methods are all designed for the input branches of the network. Next, we introduce methods designed for the input images. Ma et al. (2021) proposed AIGQA, which employs a GAN network to build an active inference module and incorporates both semantic consistency and structural integrity during optimization to predict the primary content of the image. Then, using a multi-stream CNN analyzes scene information, distortion types, and content degradation to ultimately estimate the quality score. Yin et al. (2022) proposed CVRKD-IQA using knowledge distillation technology, which comprises a full-reference teacher branch and a non-aligned reference student branch. First, the full-reference teacher model is trained, and then non-aligned reference images are introduced in the non-aligned reference branch, and knowledge distillation is employed to transfer the learned distribution differences to the student branch.

In addition to designing input images and input branches, some methods also focus on designing the network's output or output tasks. For example, Jiang et al. (2019) argued that the single scalar quality score output by existing models cannot reveal the subjective diversity of images that may receive multiple opinion scores. To tackle this issue, they used the empirical score distribution (ESD) to obtain a more informative vectorized label. By simultaneously predicting the quality score and ESD, the model's quality perception ability is improved. Yan, Bare, and Tan (2019) introduced NSS feature output in addition to the quality score output, and simultaneously performed quality score prediction and NSS feature prediction through multi-task learning. Gao et al. (2022) proposed an NR-IQA method using fuzzy theory to predict image quality score distributions. The method includes three stages: feature extraction with VGG16, feature fuzzification to model cognitive uncertainty, and fuzzy propagation to predict score distributions. To enhance the accuracy of predictions, the method introduces a loss function based on the cumulative distribution function and quantiles, which helps capture the subjective uncertainty in opinion scores.

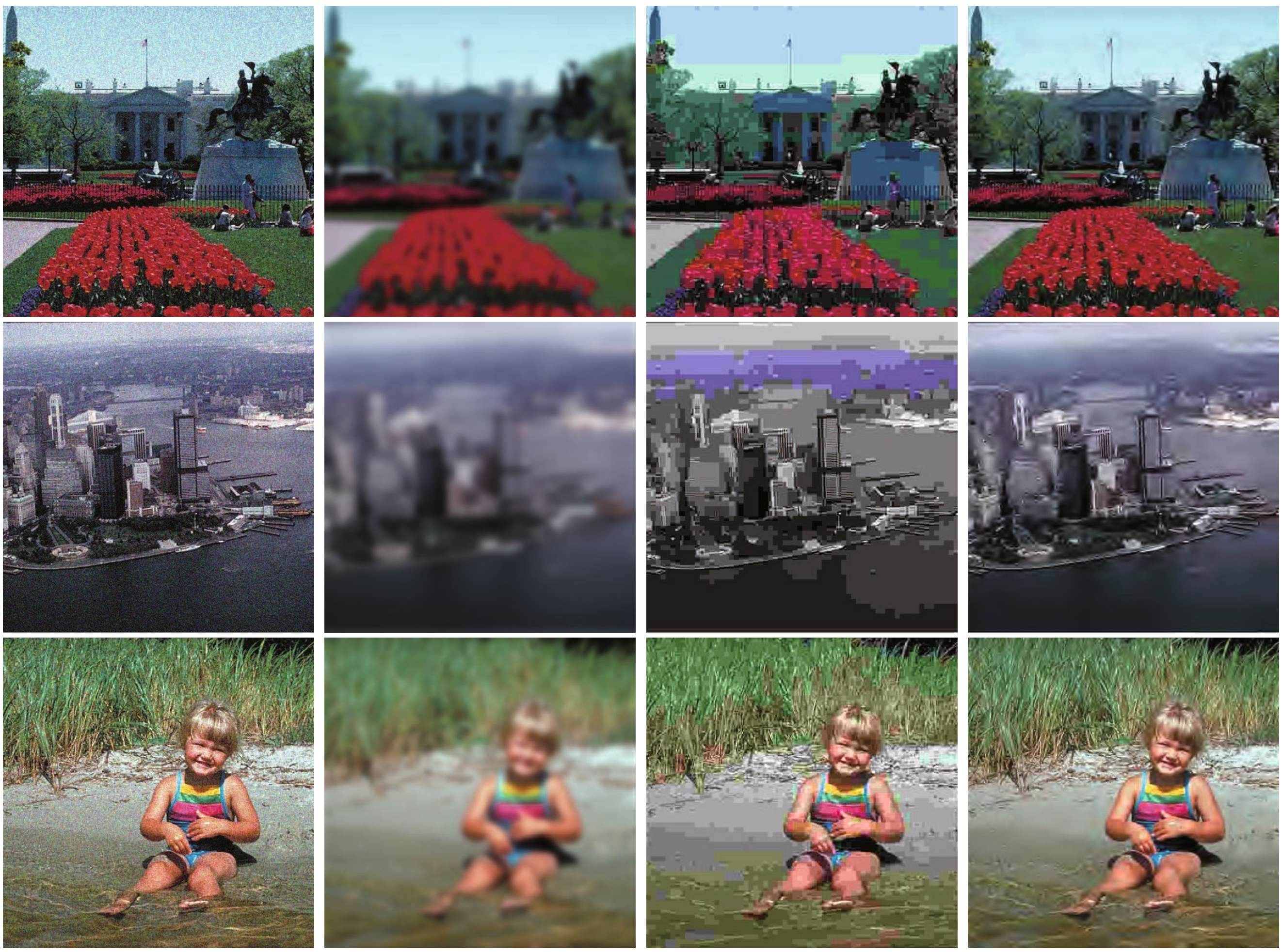
In summary, existing DL-based NR-IQA methods for authentically distorted images focus on addressing challenges through advancements in these three key areas. Developments in network structure and loss functions aim to enhance the ability of models to capture complex distortions and align better with human perception. Innovations in learning paradigms and the design of training datasets contribute to more robust and generalised models. Lastly, optimising network input and output ensures more accurate and interpretable quality predictions.

6 | NR-IQA Datasets and Metrics

6.1 | Datasets

Databases are essential for developing and evaluating NR-IQA algorithms. High-quality databases provide varying distortion types and levels, enabling researchers to train and test their algorithms effectively. A comprehensive database can help to construct more robust and generalised NR-IQA models. We present an in-depth review of 20 widely used databases in NR-IQA. These databases cover a broad spectrum of image content, distortion types, and subjective rating methodologies, catering to the diverse needs of researchers and practitioners. Table 1 summarises some representative datasets. Figures 5 and 6 show some examples of general synthetic and authentic distortions, respectively.

- **LIVE** (Sheikh, Sabir, and Bovik 2006) consists of 779 distorted images, derived from 29 high-quality images degraded by 5 types of distortions: JPEG2000 compression (JP2K), JPEG compression (JPEG), white noise (WN), Gaussian blur (GB), and fast fading (FF). Each distortion type is applied at 4 or 5 intensity levels. The resolution of most images is 768×512.
- **TID2008** (Ponomarenko et al. 2009) contains 25 reference images and 1700 distorted images across 17 distortion types, including additive JP2K, GB, JPEG, and Gaussian noise (GN). Each type is applied at 4 levels of severity. The resolution is 512×384.
- **TID2013** (Ponomarenko et al. 2015) is an extended version of the TID2008, which consists of 3000 images synthetically degraded from 25 reference images with 24 distortion types. The resolution is the same as that of TID2008.
- **CSIQ** (Larson and Chandler 2010) consists of 30 reference images and 866 distorted images, spanning 6 types of distortion: JPEG, JP2K, GB, global contrast decrements, and additive pink GN. Each type is applied at 4–5 levels of severity. The resolution is 512×512.
- **IRCCyN/IVC** (Le Callet and Autrusseau 2005) comprises 10 high-quality reference images, each with a resolution of 512×512 pixels. From these reference images, a total of 235 distorted images have been generated using four distinct



(a) White Noise (b) Gaussian Blur (c) JPEG Compression (d) JP2K Compression

FIGURE 5 | An illustration of four general synthetic distortions types. All samples are from CSIQ (Larson and Chandler 2010) dataset.

(a) Low Visibility (b) Over-exposure (c) Motion Blur (d) Lens Blur

FIGURE 6 | An illustration of four general authentic distortions types. All samples are from CLIVE (Ghadiyaram and Bovik 2015) dataset.

processes: JPEG, JP2, LAR coding, and blurring. Each distortion type have been optimised in order to uniformly cover the whole range of quality.

- **MICT** (Horita et al. 2011) consists 14 reference images and 168 distorted images. This dataset introduces two distortion types: JPEG and JP2K, each applied at six different compression levels. The resolution is 768 × 512.
- **A57** consists three reference images and 54 distorted images. This database have six distortion types with 3 levels: GB, JPEG, JP2K, quantization noise, JPEG transmission errors, and JPEG2000 transmission errors. The resolution is 512 × 512.
- **WED** (Ma et al. 2016) includes 4744 reference images and 94,880 distorted images corrupted by JPEG, JP2K, GB, and WN, with 5 levels. The database consists of images with diverse resolutions. Although the dataset does not include human opinion scores, the authors propose a set of novel test criteria to comprehensively evaluate the performance of IQA models.
- **KADID-10K** (Lin, Hosu, and Saupe 2019) contains 81 high-quality reference images and 10,125 distorted images, covering 25 distortion types such as compression artefacts, blurring, noise, and colour distortion. Each distortion type has five different severity levels. The resolution is 512 × 384.
- **WIQ** (De Simone et al. 2009) is a unique dataset designed specifically for evaluating NR-IQA algorithms in wireless image transmission. It contains seven reference images and 80 distorted images, captured using a wireless link simulator. The distortions in the WIQ database are caused by wireless channel errors, which sets it apart from other NR-IQA datasets that focus on compression and noise-related distortions. The resolution is 512 × 512.
- **VCL@FER** (Sazzad, Kawayoke, and Horita 2008) contains 23 reference images and 552 distorted images. The dataset covers four common distortion types with six different distortion levels. The resolution is 683 × 512.
- **ESPL** (Kundu and Evans 2015) contains 25 reference images and 500 distorted images, covering five distortion types: JPEG, JP2K, WN, GB, and FF, with four different distortion levels. The resolution is 512 × 512.
- **CID2013** (Virtanen et al. 2014) is a comprehensive NR-IQA benchmark dataset consisting of 474 reference images and 1786 distorted images. The CID2013 contains six distortion

types: GN, GB, JEPG, JP2K, contrast change (CC) and colour saturation change, with five different distortion levels. The resolution is 800 × 600.

- **LIVEMD** (Jayaraman et al. 2012) is an NR-IQA dataset for multi-distortion image quality assessment. The LIVEMD contains 15 high-quality images and 450 distorted images, covering two multiple distortion combinations with three distortion levels: blur+JPEG and blur+noise. The resolution is 1280 × 720.
- **MDID2013** (Gu et al. 2014) has 12 reference images and 324 distorted images. Each reference image is progressively degraded by GB, WN, and JPEG. The resolution is 768 × 512 or 1280 × 720.
- **MDID2016** (Sun, Zhou, and Liao 2017) 1600 distorted images generated by applying 5 types of synthetic distortions (i.e., WN, GB, JPEG, JP2K, and CC) to 20 reference images. Distortions are applied in a defined sequence: GB or CC is introduced first, followed by JPEG or JP2K compression, and finally, WN is added. The image resolution is 512 × 384.
- **CLIVE** (Ghadiyaram and Bovik 2015) includes 1162 authentically distorted images gathered from a variety of sources in real-world scenarios. The resolutions range from 500 × 500 pixels to 640 × 960 pixels, reflecting the diversity of images in the real world.
- **KonIQ-10K** (Hosu et al. 2020) comprises 10,073 authentically distorted images collected from the Internet. Unlike other datasets, KonIQ-10k covers a wide range of real-world distortions naturally present in various scenes. The image resolution is 1024 × 768.
- **SPAQ** (Fang, Zhu, et al. 2020) is a unique NR-IQA dataset designed to assess image quality in smartphone photography. It comprises 11,125 images captured using 66 different smartphone cameras, covering a broad range of real-world scenes, illumination conditions, and photography attributes. The resolutions range from 960 × 720 to 4608 × 3456 and representing the diversity of smartphone camera specifications and settings. SPAQ provides a comprehensive platform for evaluating NR-IQA algorithms in the context of smartphone photography, considering both technical and aesthetic aspects of image quality.
- **FLIVE** (Ying et al. 2020) is a large-scale, authentic distortion NR-IQA dataset sourced from the Flickr website. It consists of 39,810 images encompassing a broad spectrum of content, including landscapes, portraits, and objects, with resolutions ranging from 500 × 500 pixels to 6144 × 4096. The FLIVE dataset focuses on real-world distortions naturally introduced during the capture, processing, and transmission stages, such as motion blur, defocus blur, overexposure and compression artefacts.

6.2 | Metrics

The performance of NR-IQA models is typically evaluated from three perspectives: prediction accuracy, monotonicity, and consistency. These aspects correspond to three evaluation metrics (Antkowiak et al. 2000): the Pearson Linear Correlation Coefficient (PLCC), the Spearman Rank-Order Correlation Coefficient (SRCC), and the Root Mean Square Error (RMSE).

PLCC is employed to assess the prediction accuracy of IQA models. Before calculating PLCC, a nonlinear regression is typically performed on the objective and subjective scores. The logistic function used for nonlinear regression is defined as follows:

$$p(Q) = \beta_1 \left[\frac{1}{2} - \frac{1}{1 + e^{(\beta_2 (Q - \beta_3))}} \right] + \beta_4 Q^2 + \beta_5 \tag{2}$$

In Equation (2), Q represents the original objective quality score, p denotes the regressed objective quality score, and β_1, β_2, β_3, β_4 and β_5 are the model parameters. The PLCC is then calculated as:

$$PLCC = \frac{\sum_{i=1}^{N} \left(s_i - \bar{s}\right)\left(p_i - \bar{p}\right)}{\sqrt{\sum_{i=1}^{N} \left(s_i - \bar{s}\right)^2 \sum_{i=1}^{N} \left(p_i - \bar{p}\right)^2}} \tag{3}$$

In Equation (3), s_i and p_i represent the subjective quality score and the objective quality score of the i-th image, respectively. $\bar{s}$ and $\bar{p}$ denote the mean subjective quality score and the mean objective quality score, respectively. SRCC is used to measure the monotonicity of the prediction results of IQA models. Its calculation formula is expressed as:

$$SRCC = 1 - \frac{6 \sum_{i=1}^{N} d_i^2}{N\left(N^2 - 1\right)} \tag{4}$$

In Equation (4), N represents the total number of samples, and d_i denotes the difference between the subjective quality score rank and the objective quality score rank of the i-th image.

RMSE is used to evaluate the consistency of the predictions made by IQA models. Its calculation formula is expressed as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{N} \left(s_i - p_i\right)^2} \tag{5}$$

7 | NR-IQA Challenges and Future Directions

Thanks to the emergence of DL and the availability of specific datasets, NR-IQA has undergone substantial advancements in recent years. However, despite these advancements, several challenges remain to be addressed to make NR-IQA methods more robust, reliable, and practically applicable. These challenges span various aspects of image quality assessment, from the diversity and complexity of real-world distortions to the reliability and consistency of subjective quality ratings. Moreover, the lack of generalisation ability across different datasets and devices, the high computational complexity of existing methods, and the need for effective integration with other image processing tasks pose additional hurdles in the development and deployment of NR-IQA methods. In this section, an analysis of these challenges and potential future research directions are provided.

1. **Diversity and complexity of real-world distortions**. Real-world images often experience multiple forms

distortions. These distortions can occur individually or in combination, leading to complex and diverse distortion patterns. Moreover, the severity and distribution of distortions may vary significantly across different images and devices. Existing NR-IQA methods often struggle to capture the intricate characteristics of these real-world distortions, as they are typically trained on datasets with limited distortion types and levels. Developing algorithms capable of efficiently modelling and evaluating the images with diverse and complex distortions remains a significant challenge. Future works should construct more comprehensive and representative datasets that cover a wider range of real-world distortions. This can be achieved by collecting images from social media platforms, online image databases, and real-world capture devices. Moreover, researchers should consider incorporating novel distortion types, such as those arising from emerging technologies like virtual and augmented reality, to ensure the relevance of NR-IQA methods in the swiftly evolving multimedia field. Regarding algorithm development, future efforts should investigate more sophisticated and flexible architectures, such as attention mechanisms, graph neural networks, and transformers, which have exhibited immense potential in capturing intricate and long-range dependencies in visual data. Additionally, incorporating prior knowledge of HVS with the statistical properties of natural images can help in developing more perceptually-aligned and interpretable NR-IQA models.

2. **Reliability and consistency of subjective quality ratings**. Subjective quality assessment serve as the foundation of NR-IQA, as it provides the ground truth for algorithm development and evaluation. However, obtaining reliable and consistent subjective quality ratings is a non-trivial task. For image quality, human perception is inherently subjective and can be swayed by various factors, including individual preferences, viewing conditions, and cultural backgrounds. Designing subjective quality assessment experiments that can effectively control these factors and minimise bias is crucial for collecting high-quality subjective data. Moreover, the choice of rating scales, such as discrete or continuous scales, and the number of rating levels can also impact the reliability and discriminatory power of subjective ratings. To improve the reliability and consistency of subjective quality ratings, future research efforts should prioritise the development of more sophisticated experimental designs and data collection protocols. This can include the use of adaptive rating scales, which can dynamically adjust the rating levels in accordance with the perceived quality range of the images being assessed. Moreover, incorporating anchor images with known quality levels can help in calibrating the subjective ratings and reducing inter-subject variability. Another promising direction is the development of crowdsourcing frameworks that can effectively aggregate the opinions of a vast number of diverse subjects while simultaneously ensuring the quality and consistency of the collected data. Furthermore, exploring the nature of physiological signals, such as eye tracking and brain activity, can provide additional insights into the subjective perception quality and help in developing more reliable and objective quality assessment methods.

3. **Cross-dataset and cross-device generalisation**. The efficacy of NR-IQA algorithms is often assessed using specific datasets that may exhibit limited diversity in image content, distortion types, and distortion degrees. As a result, algorithms that perform well on one dataset may fail to generalise to other datasets or real-world scenarios. This deficiency in generalisation capability impedes the practical applicability of NR-IQA methods across diverse real-world scenarios. Future research should concentrate on devising algorithms capable of learning more robust and transferable features to improve the generalisation ability. This can be achieved through techniques such as adversarial learning, where the model is insensitive to domain variations while preserving the quality-related information. Moreover, incorporating unsupervised and self-supervised learning approaches can help in learning more generic and task-agnostic features that can generalise well to new datasets and devices. Another promising direction is the development of meta-learning frameworks, which can adapt the model parameters to new domains with only a few examples, thus reducing the need for extensive fine-tuning. Additionally, future works should explore domain adaptation techniques, such as feature alignment and instance weighting, to bridge the gap between different datasets and devices.

4. **Computational efficiency and real-time performance**. Most NR-IQA algorithms, especially those DL-based methods, have high computational complexity due to their large model sizes and the need for extensive feature extraction. This high computational overhead hinders their deployment in real-time scenarios, such as live streaming services and real-time camera quality monitoring systems. To address the computational challenges of NR-IQA methods, upcoming research efforts should focus on the development of more streamlined and lightweight models. This can be achieved through techniques such as network pruning, where the redundant and less informative connections in the model are removed, resulting in a more compact and efficient architecture. Moreover, exploring the use of quantization and binarization techniques can mitigate the memory requirements and computational complexity of these models, thereby enhancing their suitability for deployment on devices with limited resources. Another promising direction is the development of adaptive inference frameworks, where the model can dynamically adjust its complexity based on the input image and the available computational resources. This can be achieved through techniques such as early exit and dynamic routing, which can selectively activate different model components based on the intricacy of the input image.

5. **Integration of quality assessment with other tasks**. NR-IQA is closely related to various low-level image processing tasks. The main goal of these tasks is to enhance the visual quality of images. By integrating quality assessment metrics as optimization targets or performance evaluation criteria, the effectiveness of these tasks can be significantly augmented. Nevertheless, incorporating NR-IQA methods into these tasks presents a complex challenge, as it necessitates a profound comprehension of the intricate

relationship between image quality and the particular task under consideration. To effectively integrate NR-IQA methods with other image processing tasks, future research should prioritise the development of more comprehensive and integrated learning architectures that facilitate seamless end-to-end processing. This can be achieved by jointly optimising the image processing and quality assessment components, allowing them to learn complementary and task-specific features. Moreover, exploring the use of quality-aware objective functions can help in guiding the image processing algorithms towards generating visually pleasing results that are consistent with human perception. Another promising direction is the development of multi-task processing frameworks, where the model can simultaneously learn to perform multiple related tasks. This can help in learning more robust and generalizable features that can benefit all the tasks involved. Additionally, future efforts should investigate attention mechanisms and feedback loops to enable more dynamic and interactive integration of quality assessment with image processing tasks.

8 | Conclusion

This survey provides a comprehensive overview of the current state of NR-IQA, focusing on methodologies, datasets, and challenges in the field. To provide a structured framework for comprehending the most advanced NR-IQA techniques, we have developed an original categorization system that classifies these methods according to their underlying design principles and the types of distortions they address. Additionally, we review 20 widely used NR-IQA datasets, detailing their size, distortion types, and levels, which serve as critical benchmarks for evaluating these methods. We also discuss key challenges faced by NR-IQA methods, including handling diverse and complex distortions, achieving robust generalisation across datasets and devices, and meeting real-time computational requirements. To address these challenges, we propose several future research directions, such as developing distortion-agnostic models, leveraging large-scale datasets with realistic distortions, and improving model efficiency.

Author Contributions

Conceptualization: Qingyu Mao, Shuai Liu and Gwanggil Jeon. Methodology: Qingyu Mao, Shuai Liu and David Camacho. Software: Qingyu Mao, Shuai Liu and David Camacho. Validation: Qingyu Mao, Shuai Liu and Qilei Li. Formal analysis: Qingyu Mao and Shuai Liu. Investigation: Qingyu Mao and Qilei Li. Resources: Gwanggil Jeon, Shuai Liu and Qilei Li. Writing: Qingyu Mao, Shuai Liu and Qilei Li. Supervision: Qilei Li, Gwanggil Jeon and Hyunbum Kim. Project administration: Gwanggil Jeon and Hyunbum Kim. All authors have reviewed and approved the published version of the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

Agnolucci, L., L. Galteri, and M. Bertini. 2024. "Quality-Aware Image-Text Alignment for Real-World Image Quality Assessment," arXiv Preprint arXiv:2403.11176.

Agnolucci, L., L. Galteri, M. Bertini, and A. del Bimbo. 2024. "Arniqa: Learning Distortion Manifold for Image Quality Assessment." In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 189–198. Waikoloa, HI: IEEE.

Antkowiak, J., T. J. Baina, F. V. Baroncini, et al. 2000. "Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment," March 2000.

Beron, J., H. D. Benitez-Restrepo, and A. C. Bovik. 2020. "Blind Image Quality Assessment for Super Resolution via Optimal Feature Selection." *IEEE Access* 8: 143201–143218.

Bianco, S., L. Celona, P. Napoletano, and R. Schettini. 2018. "On the Use of Deep Learning for Blind Image Quality Assessment." *Signal, Image and Video Processing* 12: 355–362.

Bosse, S., D. Maniry, K. R. Müller, T. Wiegand, and W. Samek. 2017. "Deep Neural Networks for No-Reference and Full-Reference Image Quality Assessment." *IEEE Transactions on Image Processing* 27, no. 1: 206–219.

Chandler, D. M. 2013. "Seven Challenges in Image Quality Assessment: Past, Present, and Future Research." *International Scholarly Research Notices* 2013, no. 1: 905685.

Chandler, D. M., and S. S. Hemami. 2007. "VSNR: A Wavelet-Based Visual Signal-To-Noise Ratio for Natural Images." *IEEE Transactions on Image Processing* 16, no. 9: 2284–2298.

Chen, B., H. Li, H. Fan, and S. Wang. 2021. "No-Reference Screen Content Image Quality Assessment With Unsupervised Domain Adaptation." *IEEE Transactions on Image Processing* 30: 5463–5476.

Chen, B., L. Zhu, C. Kong, H. Zhu, S. Wang, and Z. Li. 2022. "No-Reference Image Quality Assessment by Hallucinating Pristine Features." *IEEE Transactions on Image Processing* 31: 6139–6151.

Chen, C., J. Mo, J. Hou, et al. 2024. "Topiq: A Top-Down Approach From Semantics to Distortions for Image Quality Assessment." *IEEE Transactions on Image Processing* 33: 2404–2418.

Chen, F., H. Fu, H. Yu, and Y. Chu. 2023. "No-Reference Image Quality Assessment Based on a Multitask Image Restoration Network." *Applied Sciences* 13, no. 11: 6802.

Chen, J., L. Shen, L. Zheng, and X. Jiang. 2018. "Naturalization Module in Neural Networks for Screen Content Image Quality Assessment." *IEEE Signal Processing Letters* 25, no. 11: 1685–1689.

Chen, J., Y. Xu, K. Ma, H. Huang, and T. Zhao. 2018. "A Hybrid Quality Metric for Non-Integer Image Interpolation." In *Proceedings of the 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, 1–3. Sardinia, Italy: IEEE.

Chen, P., L. Li, X. Zhang, S. Wang, and A. Tan. 2019. "Blind Quality Index for Tone-Mapped Images Based on Luminance Partition." *Pattern Recognition* 89: 108–118.

Chen, Z., H. Qin, J. Wang, et al. 2025. *Promptiqa: Boosting the Performance and Generalization for No-Reference Image Quality Assessment via Prompts*, 247–264. Milano, Italy: Springer.

Choi, L. K., J. You, and A. C. Bovik. 2015. "Referenceless Prediction of Perceptual Fog Density and Perceptual Image Defogging." *IEEE Transactions on Image Processing* 24, no. 11: 3888–3901.

De Simone, F., M. Naccari, M. Tagliasacchi, F. Dufaux, S. Tubaro, and T. Ebrahimi. 2009. "Subjective Assessment of H. 264/AVC Video Sequences Transmitted Over a Noisy Channel." In *Proceedings of the 2009 International Workshop on Quality of Multimedia Experience*, 204–209. San Diego, CA: IEEE.

Ding, K., K. Ma, S. Wang, and E. P. Simoncelli. 2020. "Image Quality Assessment: Unifying Structure and Texture Similarity." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, no. 5: 2567–2581.

Dodge, S., and L. Karam. 2016. "Understanding How Image Quality Affects Deep Neural Networks." In *Proceedings of the 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, 1–6. Lisbon, Portugal: IEEE.

Duan, H., G. Zhai, X. Min, Y. Zhu, Y. Fang, and X. Yang. 2018. "Perceptual Quality Assessment of Omnidirectional Images." In *Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1–5. Florence, Italy: IEEE.

Fang, Y., K. Ma, Z. Wang, W. Lin, Z. Fang, and G. Zhai. 2014. "No-Reference Quality Assessment of Contrast-Distorted Images Based on Natural Scene Statistics." *IEEE Signal Processing Letters* 22, no. 7: 838–842.

Fang, Y., J. Yan, R. Du, et al. 2020. "Blind Quality Assessment for Tone-Mapped Images by Analysis of Gradient and Chromatic Statistics." *IEEE Transactions on Multimedia* 23: 955–966.

Fang, Y., J. Yan, L. Li, J. Wu, and W. Lin. 2017. "No Reference Quality Assessment for Screen Content Images With Both Local and Global Feature Representation." *IEEE Transactions on Image Processing* 27, no. 4: 1600–1610.

Fang, Y., C. Zhang, W. Yang, J. Liu, and Z. Guo. 2018. "Blind Visual Quality Assessment for Image Super-Resolution by Convolutional Neural Network." *Multimedia Tools and Applications* 77: 29829–29846.

Fang, Y., H. Zhu, Y. Zeng, K. Ma, and Z. Wang. 2020. "Perceptual Quality Assessment of Smartphone Photography." In *Proceedings of the Computer Vision and Pattern Recognition*, 3677–3686. Seattle, WA: IEEE.

Gao, Y., X. Min, Y. Zhu, J. Li, X. P. Zhang, and G. Zhai. 2022. "Image Quality Assessment: From Mean Opinion Score to Opinion Score Distribution." In *Proceedings of the 30th ACM International Conference on Multimedia*, 997–1005. Lisbon, Portugal: IEEE.

Ghadiyaram, D., and A. C. Bovik. 2015. "Massive Online Crowdsourced Study of Subjective and Objective Picture Quality." *IEEE Transactions on Image Processing* 25, no. 1: 372–387.

Ghadiyaram, D., and A. C. Bovik. 2017. "Perceptual Quality Prediction on Authentically Distorted Images Using a Bag of Features Approach." *Journal of Vision* 17, no. 1: 32.

Golestaneh, S. A., S. Dadsetan, and K. M. Kitani. 2022. "No-Reference Image Quality Assessment via Transformers, Relative Ranking, and Self-Consistency." In *Proceedings of the Computer Vision and Pattern Recognition (cs.CV)*, 1220–1230. New Orleans, LA: IEEE.

Gu, K., S. Wang, G. Zhai, et al. 2016. "Blind Quality Assessment of Tone-Mapped Images via Analysis of Information, Naturalness, and Structure." *IEEE Transactions on Multimedia* 18, no. 3: 432–443.

Gu, K., G. Zhai, X. Yang, and W. Zhang. 2014. "Hybrid No-Reference Quality Metric for Singly and Multiply Distorted Images." *IEEE Transactions on Broadcasting* 60, no. 3: 555–567.

Gu, K., J. Zhou, J. F. Qiao, G. Zhai, W. Lin, and A. C. Bovik. 2017. "No-Reference Quality Assessment of Screen Content Pictures." *IEEE Transactions on Image Processing* 26, no. 8: 4005–4018.

Gu, J., G. Meng, S. Xiang, and C. Pan. 2019. "Blind Image Quality Assessment via Learnable Attention-Based Pooling." *Pattern Recognition* 91: 332–344.

Guan, J., S. Yi, X. Zeng, W. K. Cham, and X. Wang. 2017. "Visual Importance and Distortion Guided Deep Image Quality Assessment Framework." *IEEE Transactions on Multimedia* 19, no. 11: 2505–2520.

Horita, Y., K. Shibata, Y. Kawayoke, and Z. M. P. Sazzad. 2011. "MICT Image Quality Evaluation Database." http://mict.eng.u-toyama.ac.jp/mictdb.html.

Hosu, V., H. Lin, T. Sziranyi, and D. Saupe. 2020. "KonIQ-10k: An Ecologically Valid Database for Deep Learning of Blind Image Quality Assessment." *IEEE Transactions on Image Processing* 29: 4041–4056.

Hu, R., Y. Liu, K. Gu, X. Min, and G. Zhai. 2021. "Toward a No-Reference Quality Metric for Camera-Captured Images." *IEEE Transactions on Cybernetics* 53, no. 6: 3651–3664.

Huang, Z., Z. Zhang, Y. Lu, Z. J. Zha, Z. Chen, and B. Guo. 2024. "VisualCritic: Making LMMs Perceive Visual Quality Like Humans," arXiv Preprint arXiv:2403.12806.

Jayaraman, D., A. Mittal, A. K. Moorthy, and A. C. Bovik. 2012. "Objective Quality Assessment of Multiply Distorted Images." In *Proceedings of the 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, 1693–1697. Pacific Grove, CA: IEEE.

Jiang, H., G. Jiang, M. Yu, T. Luo, and H. Xu. 2021. "Multi-Angle Projection Based Blind Omnidirectional Image Quality Assessment." *IEEE Transactions on Circuits and Systems for Video Technology* 32, no. 7: 4211–4223.

Jiang, Q., W. Gao, S. Wang, et al. 2020. "Blind Image Quality Measurement by Exploiting High-Order Statistics With Deep Dictionary Encoding Network." *IEEE Transactions on Instrumentation and Measurement* 69, no. 10: 7398–7410.

Jiang, Q., Z. Peng, S. Yang, and F. Shao. 2019. "Authentically Distorted Image Quality Assessment by Learning From Empirical Score Distributions." *IEEE Signal Processing Letters* 26, no. 12: 1867–1871.

Jiang, Q., F. Shao, W. Lin, and G. Jiang. 2017. "BLIQUE-TMI: Blind Quality Evaluator for Tone-Mapped Images Based on Local and Global Feature Analyses." *IEEE Transactions on Circuits and Systems for Video Technology* 29, no. 2: 323–335.

Kang, L., P. Ye, Y. Li, and D. Doermann. 2014. "Convolutional Neural Networks for No-Reference Image Quality Assessment." In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, 1733–1740. Columbus, OH: IEEE.

Karray, F., A. Campilho, and A. Yu. 2019. "Image Analysis and Recognition," Proceedings of the 16th International Conference, ICIAR 2019, Waterloo, ON, Canada, August 27–29, 2019, Part I. 11662. Springer.

Ke, J., Q. Wang, Y. Wang, P. Milanfar, and F. Yang. 2021. "Musiq: Multi-Scale Image Quality Transformer." In *Proceedings of the 2021 International Conference on Computer Vision*, 5148–5157. Montreal, QC: IEEE.

Kim, J., and S. Lee. 2017. "Deep Learning of Human Visual Sensitivity in Image Quality Assessment Framework." In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1676–1684. Honolulu, HI: IEEE.

Kim, J., A. D. Nguyen, and S. Lee. 2018. "Deep CNN-Based Blind Image Quality Predictor." *IEEE Transactions on Neural Networks and Learning Systems* 30, no. 1: 11–24.

Kiran Adhikarla, V., M. Vinkler, D. Sumin, et al. 2017. "Towards a Quality Metric for Dense Light Fields." In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 58–67. Honolulu, HI: IEEE.

Köhler, T., M. Bätz, F. Naderi, A. Kaup, A. Maier, and C. Riess. 2019. "Toward Bridging the Simulated-To-Real Gap: Benchmarking Super-Resolution on Real Data." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42, no. 11: 2944–2959.

Kong, X., K. Li, Q. Yang, L. Wenyin, and M. H. Yang. 2013. "A New Image Quality Metric for Image Auto-Denoising." In *Proceedings of the 2013 IEEE International Conference on Computer Vision*, 2888–2895. Sydney, NSW: IEEE.

Kundu, D., and B. L. Evans. 2015. "Full-Reference Visual Quality Assessment for Synthetic Images: A Subjective Study." In *Proceedings* of the 2015 IEEE International Conference on Image Processing (ICIP), 2374–2378. Quebec City, QC: IEEE.

Kundu, D., D. Ghadiyaram, A. C. Bovik, and B. L. Evans. 2017. "Large-Scale Crowdsourced Study for Tone-Mapped HDR Pictures." *IEEE Transactions on Image Processing* 26, no. 10: 4725–4740.

Lao, S., Y. Gong, S. Shi, et al. 2022. "Attentions Help Cnns See Better: Attention-Based Hybrid Image Quality Assessment Network." In *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1140–1149. New Orleans, LA: IEEE.

Larson, E. C., and D. M. Chandler. 2010. "Most Apparent Distortion: Full-Reference Image Quality Assessment and the Role of Strategy." *Journal of Electronic Imaging* 19, no. 1: 11006.

Le Callet, P., and F. Autrusseau. 2005. "Subjective Quality Assessment Irccyn/Ivc Database, 2005".

Li, A., J. Li, Q. Lin, C. Ma, and B. Yan. 2020. "Deep Image Quality Assessment Driven Single Image Deblurring." In *Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME)*, 1–6. London, UK: IEEE.

Li, B., and F. Huo. 2024. "REQA: Coarse-To-Fine Assessment of Image Quality to Alleviate the Range Effect." *Journal of Visual Communication and Image Representation* 98: 104043.

Li, D., T. Jiang, and M. Jiang. 2020. "Norm-In-Norm Loss With Faster Convergence and Better Performance for Image Quality Assessment." In *Proceedings of the 28th ACM International Conference on Multimedia*, 789–797. Seattle, WA: IEEE.

Li, F., Y. Zhang, and P. C. Cosman. 2021. "MMMNet: An End-To-End Multi-Task Deep Convolution Neural Network With Multi-Scale and Multi-Hierarchy Fusion for Blind Image Quality Assessment." *IEEE Transactions on Circuits and Systems for Video Technology* 31, no. 12: 4798–4811.

Li, L., W. Lin, X. Wang, G. Yang, K. Bahrami, and A. C. Kot. 2015. "No-Reference Image Blur Assessment Based on Discrete Orthogonal Moments." *IEEE Transactions on Cybernetics* 46, no. 1: 39–50.

Li, L., Y. Yan, Z. Lu, J. Wu, K. Gu, and S. Wang. 2017. "No-Reference Quality Assessment of Deblurred Images Based on Natural Scene Statistics." *IEEE Access* 5: 2163–2171.

Li, Q., W. Lin, J. Xu, and Y. Fang. 2016. "Blind Image Quality Assessment Using Statistical Structural and Luminance Features." *IEEE Transactions on Multimedia* 18, no. 12: 2457–2469.

Lin, H., V. Hosu, and D. Saupe. 2019. "KADID-10k: A Large-Scale Artificially Distorted IQA Database." In *Proceedings of the 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*, 1–3. Berlin, Germany: IEEE.

Lin, K. Y., and G. Wang. 2018. "Hallucinated-IQA: No-Reference Image Quality Assessment via Adversarial Learning." In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 732–741. Salt Lake City, UT: IEEE.

Liu, D., Y. Xu, Y. Quan, and P. Le Callet. 2014. "Reduced Reference Image Quality Assessment Using Regularity of Phase Congruency." *Signal Processing: Image Communication* 29, no. 8: 844–855.

Liu, T. J., and K. H. Liu. 2017. "No-Reference Image Quality Assessment by Wide-Perceptual-Domain Scorer Ensemble Method." *IEEE Transactions on Image Processing* 27, no. 3: 1138–1151.

Liu, X., J. van de Weijer, and A. D. Bagdanov. 2017. "Rankiqa: Learning From Rankings for No-Reference Image Quality Assessment." In *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, 1040–1049. Venice, Italy: IEEE.

Liu, Y., K. Gu, X. Li, and Y. Zhang. 2020. "Blind Image Quality Assessment by Natural Scene Statistics and Perceptual Characteristics." *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 16, no. 3: 1–91.

Liu, Y., K. Gu, Y. Zhang, et al. 2019. "Unsupervised Blind Image Quality Evaluation via Statistical Measurements of Structure, Naturalness, and Perception." *IEEE Transactions on Circuits and Systems for Video Technology* 30, no. 4: 929–943.

Ma, C., C. Y. Yang, X. Yang, and M. H. Yang. 2017. "Learning a No-Reference Quality Metric for Single-Image Super-Resolution." *Computer Vision and Image Understanding* 158: 1–16.

Ma, J., J. Wu, L. Li, et al. 2021. "Blind Image Quality Assessment With Active Inference." *IEEE Transactions on Image Processing* 30: 3650–3663.

Ma, K., Z. Duanmu, Q. Wu, et al. 2016. "Waterloo Exploration Database: New Challenges for Image Quality Assessment Models." *IEEE Transactions on Image Processing* 26, no. 2: 1004–1016.

Ma, K., W. Liu, K. Zhang, Z. Duanmu, Z. Wang, and W. Zuo. 2017. "End-To-End Blind Image Quality Assessment Using Deep Neural Networks." *IEEE Transactions on Image Processing* 27, no. 3: 1202–1213.

Madhusudana, P. C., N. Birkbeck, Y. Wang, B. Adsumilli, and A. C. Bovik. 2022. "Image Quality Assessment Using Contrastive Learning." *IEEE Transactions on Image Processing* 31: 4149–4161.

Manap, R. A., and L. Shao. 2015. "Non-Distortion-Specific No-Reference Image Quality Assessment: A Survey." *Information Sciences* 301: 141–160.

Min, X., G. Zhai, K. Gu, X. Yang, and X. Guan. 2018. "Objective Quality Evaluation of Dehazed Images." *IEEE Transactions on Intelligent Transportation Systems* 20, no. 8: 2879–2892.

Mittal, A., A. K. Moorthy, and A. C. Bovik. 2012. "No-Reference Image Quality Assessment in the Spatial Domain." *IEEE Transactions on Image Processing* 21, no. 12: 4695–4708.

Mittal, A., R. Soundararajan, and A. C. Bovik. 2012. "Making a Completely Blinded Image Quality Analyzer." *IEEE Signal Processing Letters* 20, no. 3: 209–212.

Narvekar, N. D., and L. J. Karam. 2011. "A No-Reference Image Blur Metric Based on the Cumulative Probability of Blur Detection (CPBD)." *IEEE Transactions on Image Processing* 20, no. 9: 2678–2683.

Ni, Z., L. Ma, H. Zeng, J. Chen, C. Cai, and K. K. Ma. 2017. "ESIM: Edge Similarity for Screen Content Image Quality Assessment." *IEEE Transactions on Image Processing* 26, no. 10: 4818–4831.

Ou, F. Z., Y. G. Wang, J. Li, G. Zhu, and S. Kwong. 2021. "A Novel Rank Learning Based No-Reference Image Quality Assessment Method." *IEEE Transactions on Multimedia* 24: 4197–4211.

Pan, Z., F. Yuan, J. Lei, Y. Fang, X. Shao, and S. Kwong. 2022. "VCRNet: Visual Compensation Restoration Network for No-Reference Image Quality Assessment." *IEEE Transactions on Image Processing* 31: 1613–1627.

Pan, Z., H. Zhang, J. Lei, et al. 2022. "DACNN: Blind Image Quality Assessment via a Distortion-Aware Convolutional Neural Network." *IEEE Transactions on Circuits and Systems for Video Technology* 32, no. 11: 7518–7531.

Ponomarenko, N., L. Jin, O. Ieremeiev, et al. 2015. "Image Database TID2013: Peculiarities, Results and Perspectives." *Signal Processing: Image Communication* 30: 57–77.

Ponomarenko, N., V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti. 2009. "TID2008-A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics." *Advances of Modern Radioelectronics* 10, no. 4: 30–45.

Ravuri, C. S., R. Sureddi, S. V. R. Dendi, S. Raman, and S. S. Channappayya. 2019. "Deep No-Reference Tone Mapped Image Quality Assessment." In *Proceedings of the 2019 53rd Asilomar*

Conference on Signals, Systems, and Computers, 1906–1910. Pacific Grove, CA: IEEE.

Ren, H., D. Chen, and Y. Wang. 2018. "RAN4IQA: Restorative Adversarial Nets for No-Reference Image Quality Assessment." In *Proceedings of the The Thirty-Second AAAI Conferenc on Artificial Intelligence (AAAI-18)*, 1. New Orleans, LA: IEEE.

Saad, M. A., A. C. Bovik, and C. Charrier. 2012. "Blind Image Quality Assessment: A Natural Scene Statistics Approach in the DCT Domain." *IEEE Transactions on Image Processing* 21, no. 8: 3339–3352.

Saha, A., S. Mishra, and A. C. Bovik. 2023. "Re-Iqa: Unsupervised Learning for Image Quality Assessment in the Wild." In *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5846–5855. Vancouver, BC: IEEE.

Sazzad, Z. P., Y. Kawayoke, and Y. Horita. 2008. "No Reference Image Quality Assessment for JPEG2000 Based on Spatial Features." *Signal Processing: Image Communication* 23, no. 4: 257–268.

Sheikh, H. R., M. F. Sabir, and A. C. Bovik. 2006. "A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms." *IEEE Transactions on Image Processing* 15, no. 11: 3440–3451.

Shi, G., W. Wan, J. Wu, X. Xie, W. Dong, and H. R. Wu. 2019. "SISRSet: Single Image Super-Resolution Subjective Evaluation Test and Objective Quality Assessment." *Neurocomputing* 360: 37–51.

Shi, L., S. Zhao, and Z. Chen. 2019. "BELIF: Blind Quality Evaluator of Light Field Image With Tensor Structure Variation Index." In *Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP)*, 3781–3785. Taipei, Taiwan: IEEE.

Shi, S., X. Zhang, S. Wang, R. Xiong, and S. Ma. 2015. "Study on Subjective Quality Assessment of Screen Content Images." In *Proceedings of the 2015 Picture Coding Symposium (PCS)*, 75–79. Cairns: IEEE.

Simoncelli, E. P., and B. A. Olshausen. 2001. "Natural Image Statistics and Neural Representation." *Annual Review of Neuroscience* 24, no. 1: 1193–1216.

Su, S., Q. Yan, Y. Zhu, et al. 2020. "Blindly Assess Image Quality in the Wild Guided by A Self-Adaptive Hyper Network–Supplementary Material," Northwestern Polytechnical University: Xian, China.

Sun, W., X. Min, D. Tu, S. Ma, and G. Zhai. 2023. "Blind Quality Assessment for In-The-Wild Images via Hierarchical Feature Fusion and Iterative Mixed Database Training." *IEEE Journal of Selected Topics in Signal Processing* 17, no. 6: 1178–1192.

Sun, W., F. Zhou, and Q. Liao. 2017. "MDID: A Multiply Distorted Image Database for Image Quality Assessment." *Pattern Recognition* 61: 153–168.

Tang, J., Y. Fang, Y. Dong, et al. 2021. "Blindly Predict Image and Video Quality in the Wild." In *Proceedings of the Proceedings of the 3rd ACM International Conference on Multimedia in Asia*, 1–7. Virtual Event: IEEE.

Tian, Y., H. Zeng, J. Hou, J. Chen, and K. K. Ma. 2020. "Light Field Image Quality Assessment via the Light Field Coherence." *IEEE Transactions on Image Processing* 29: 7945–7956.

Viola, I., and T. Ebrahimi. 2018. "VALID: Visual Quality Assessment for Light Field Images Dataset." In *Proceedings of the 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, 1–3. Cagliari, Italy: IEEE.

Virtanen, T., M. Nuutinen, M. Vaahteranoksa, P. Oittinen, and J. Häkkinen. 2014. "CID2013: A Database for Evaluating No-Reference Image Quality Assessment Algorithms." *IEEE Transactions on Image Processing* 24, no. 1: 390–402.

Wang, G., L. Li, Q. Li, K. Gu, Z. Lu, and J. Qian. 2017. "Perceptual Evaluation of Single-Image Super-Resolution Reconstruction." In *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP)*, 3145–3149. Beijing, China: IEEE.

Wang, J., K. C. Chan, and C. C. Loy. 2023. "Exploring Clip for Assessing the Look and Feel of Images." In *Proceedings of the 2023 AAAI Conference on Artificial Intelligence*, 2555–2563. Washington, DC: IEEE.

Wang, S., K. Gu, X. Zhang, et al. 2016. "Subjective and Objective Quality Assessment of Compressed Screen Content Images." *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 6, no. 4: 532–543.

Wang, Z., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. 2004. "Image Quality Assessment: From Error Visibility to Structural Similarity." *IEEE Transactions on Image Processing* 13, no. 4: 600–612.

Wu, H., Z. Zhang, W. Zhang, et al. 2023. "Q-Align: Teaching LMMS for Visual Scoring via Discrete Text-Defined Levels," arXiv Preprint arXiv:2312.17090.

Wu, J., W. Lin, G. Shi, L. Li, and Y. Fang. 2016. "Orientation Selectivity Based Visual Pattern for Reduced-Reference Image Quality Assessment." *Information Sciences* 351: 18–29.

Wu, J., W. Lin, G. Shi, and A. Liu. 2013. "Reduced-Reference Image Quality Assessment With Visual Information Fidelity." *IEEE Transactions on Multimedia* 15, no. 7: 1700–1705.

Wu, J., J. Ma, F. Liang, W. Dong, G. Shi, and W. Lin. 2020. "End-To-End Blind Image Quality Prediction With Cascaded Deep Neural Network." *IEEE Transactions on Image Processing* 29: 7414–7426.

Wu, Q., H. Li, Z. Wang, et al. 2017. "Blind Image Quality Assessment Based on Rank-Order Regularized Regression." *IEEE Transactions on Multimedia* 19, no. 11: 2490–2504.

Wu, Q., L. Wang, K. N. Ngan, H. Li, and F. Meng. 2019. "Beyond Synthetic Data: A Blind Deraining Quality Assessment Metric Towards Authentic Rain Image." In *Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP)*, 2364–2368. Taipei, Taiwan: IEEE.

Wu, Q., L. Wang, K. N. Ngan, H. Li, F. Meng, and L. Xu. 2020. "Subjective and Objective De-Raining Quality Assessment Towards Authentic Rain Image." *IEEE Transactions on Circuits and Systems for Video Technology* 30, no. 11: 3883–3897.

Xu, J., P. Ye, Q. Li, H. Du, Y. Liu, and D. Doermann. 2016. "Blind Image Quality Assessment Based on High Order Statistics Aggregation." *IEEE Transactions on Image Processing* 25, no. 9: 4444–4457.

Xu, J., W. Zhou, and Z. Chen. 2020. "Blind Omnidirectional Image Quality Assessment With Viewport Oriented Graph Convolutional Networks." *IEEE Transactions on Circuits and Systems for Video Technology* 31, no. 5: 1724–1737.

Xu, S., S. Jiang, and W. Min. 2017. "No-Reference/Blind Image Quality Assessment: A Survey." *IETE Technical Review* 34, no. 3: 223–245.

Xu, Y., D. Liu, Y. Quan, and P. Le Callet. 2015. "Fractal Analysis for Reduced Reference Image Quality Assessment." *IEEE Transactions on Image Processing* 24, no. 7: 2098–2109.

Xue, W., L. Zhang, and X. Mou. 2013. "Learning Without Human Scores for Blind Image Quality Assessment." In *Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition*, 995–1002. Portland, OR: IEEE.

Yan, B., B. Bare, and W. Tan. 2019. "Naturalness-Aware Deep No-Reference Image Quality Assessment." *IEEE Transactions on Multimedia* 21, no. 10: 2603–2615.

Yang, C., P. An, and L. Shen. 2022. "Blind Image Quality Measurement via Data-Driven Transform-Based Feature Enhancement." *IEEE Transactions on Instrumentation and Measurement* 71: 1–12.

Yang, H., Y. Fang, and W. Lin. 2015. "Perceptual Quality Assessment of Screen Content Images." *IEEE Transactions on Image Processing* 24, no. 11: 4408–4421.

Yang, P., J. Sturtz, and L. Qingge. 2023. "Progress in Blind Image Quality Assessment: A Brief Review." *Mathematics* 11, no. 12: 2766.

Yang, S., Q. Jiang, W. Lin, and Y. Wang. 2019. "SGDNet: An End-To-End Saliency-Guided Deep Neural Network for No-Reference Image Quality Assessment." In *Proceedings of the 27th ACM International Conference on Multimedia*, 1383–1391. Nice, France: IEEE.

Yang, X., F. Li, and H. Liu. 2019. "A Survey of DNN Methods for Blind Image Quality Assessment." *IEEE Access* 7: 123788–123806.

Yao, X., Q. Cao, X. Feng, G. Cheng, and J. Han. 2022. "Learning to Assess Image Quality Like an Observer." *IEEE Transactions on Neural Networks and Learning Systems* 34, no. 11: 8324–8336.

Ye, P., J. Kumar, L. Kang, and D. Doermann. 2012. "Unsupervised Feature Learning Framework for No-Reference Image Quality Assessment." In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, 1098–1105. Providence, RI: IEEE.

Yeganeh, H., and Z. Wang. 2012. "Objective Quality Assessment of Tone-Mapped Images." *IEEE Transactions on Image Processing* 22, no. 2: 657–667.

Yin, G., W. Wang, Z. Yuan, et al. 2022. "Content-Variant Reference Image Quality Assessment via Knowledge Distillation." In *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22)*, 3134–3142. Vancouver, BS: IEEE.

Ying, Z., H. Niu, P. Gupta, D. Mahajan, D. Ghadiyaram, and A. Bovik. 2020. "From Patches to Pictures (PaQ-2-PiQ): Mapping the Perceptual Space of Picture Quality." In *Proceedings of the Computer Vision and Pattern Recognition*, 3575–3585. Seattle, WA: IEEE.

You, J., and J. Korhonen. 2021. "Transformer for Image Quality Assessment." In *Proceedings of the 2021 IEEE International Conference on Image Processing*, 1389–1393. Anchorage, AK: IEEE.

Yue, G., D. Cheng, L. Li, T. Zhou, H. Liu, and T. Wang. 2022. "Semi-Supervised Authentically Distorted Image Quality Assessment With Consistency-Preserving Dual-Branch Convolutional Neural Network." *IEEE Transactions on Multimedia* 25: 6499–6511.

Zhai, G., and X. Min. 2020. "Perceptual Image Quality Assessment: A Survey." *Science China Information Sciences* 63: 1–52.

Zhang, C., W. Cheng, and K. Hirakawa. 2018. "Corrupted Reference Image Quality Assessment of Denoised Images." *IEEE Transactions on Image Processing* 28, no. 4: 1732–1747.

Zhang, K., D. Zhu, J. Jing, and X. Gao. 2019. "Learning a Cascade Regression for No-Reference Super-Resolution Image Quality Assessment." In *Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP)*, 450–453. Taipei, Taiwan: IEEE.

Zhang, L., L. Zhang, and A. C. Bovik. 2015. "A Feature-Enriched Completely Blind Image Quality Evaluator." *IEEE Transactions on Image Processing* 24, no. 8: 2579–2591.

Zhang, L., L. Zhang, X. Mou, and D. Zhang. 2011. "FSIM: A Feature Similarity Index for Image Quality Assessment." *IEEE Transactions on Image Processing* 20, no. 8: 2378–2386.

Zhang, P., X. Shao, and Z. Li. 2022. "Cycleiqa: Blind Image Quality Assessment via Cycle-Consistent Adversarial Networks." In *Proceedings of the 2022 IEEE International Conference on Multimedia and Expo (ICME)*, 1–6. Taipei, Taiwan: IEEE.

Zhang, R., P. Isola, A. A. Efros, E. Shechtman, and O. Wang. 2018. "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric." In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 586–595. Salt Lake City, UT: IEEE.

Zhang, W., K. Ma, J. Yan, D. Deng, and Z. Wang. 2018. "Blind Image Quality Assessment Using a Deep Bilinear Convolutional Neural Network." *IEEE Transactions on Circuits and Systems for Video Technology* 30, no. 1: 36–47.

Zhang, W., K. Ma, G. Zhai, and X. Yang. 2021. "Uncertainty-Aware Blind Image Quality Assessment in the Laboratory and Wild." *IEEE Transactions on Image Processing* 30: 3474–3486.

Zhang, Z., Y. Fang, J. Yan, and R. Du. 2020. "No-Reference Quality Assessment for Realistic Distorted Images by Color Moment and Texture Features." In *Proceedings of the 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 342–347. Shenzhen, China: IEEE.

Zhao, K., K. Yuan, M. Sun, M. Li, and X. Wen. 2023. "Quality-Aware Pre-Trained Models for Blind Image Quality Assessment." In *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 22302–22313. Vancouver, BC, Canada: IEEE.

Zhao, T., Y. Lin, Y. Xu, W. Chen, and Z. Wang. 2021. "Learning-Based Quality Assessment for Image Super-Resolution." *IEEE Transactions on Multimedia* 24: 3570–3581.

Zheng, L., L. Shen, J. Chen, P. An, and J. Luo. 2019. "No-Reference Quality Assessment for Screen Content Images Based on Hybrid Region Features Fusion." *IEEE Transactions on Multimedia* 21, no. 8: 2057–2070.

Zhou, F., R. Yao, B. Liu, and G. Qiu. 2019. "Visual Quality Assessment for Super-Resolved Images: Database and Method." *IEEE Transactions on Image Processing* 28, no. 7: 3528–3541.

Zhou, M., S. Lang, T. Zhang, et al. 2022. "Attentional Feature Fusion for End-To-End Blind Image Quality Assessment." *IEEE Transactions on Broadcasting* 69, no. 1: 144–152.

Zhou, W., Q. Jiang, Y. Wang, Z. Chen, and W. Li. 2020. "Blind Quality Assessment for Image Superresolution Using Deep Two-Stream Convolutional Networks." *Information Sciences* 528: 205–218.

Zhou, Y., Y. Sun, L. Li, K. Gu, and Y. Fang. 2021. "Omnidirectional Image Quality Assessment by Distortion Discrimination Assisted Multi-Stream Network." *IEEE Transactions on Circuits and Systems for Video Technology* 32, no. 4: 1767–1777.

Zhou, Z., Y. Xu, Y. Quan, and R. Xu. 2022. "Deep Blind Image Quality Assessment Using Dual-Order Statistics." In *Proceedings of the 2022 IEEE International Conference on Multimedia and Expo (ICME)*, 1–6. Taipei, Taiwan: IEEE.

Zhou, Z., Z. Zhou, X. Tao, H. Chen, Z. Yu, and Y. Cao. 2024. "EARNet: Error-Aware Reconstruction Network for No-Reference Image Quality Assessment." *Expert Systems with Applications* 238: 122050.

Zhu, H., L. Li, J. Wu, W. Dong, and G. Shi. 2020. "MetaIQA: Deep Meta-Learning for No-Reference Image Quality Assessment." In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 14143–14152. Seattle, WA: IEEE.