**Journals**

The Institution of
Engineering and Technology

# Fundus Image Quality Assessment: Survey, Challenges, and Future Scope

*Aditya Raj[1] Anil Kumar Tiwari[1] Maria G Martini[2]*

[1] *Indian Institute of Technology, Jodhpur*
[2] *Kingston University, London, UK*
\* *E-mail: raj.2@iitj.ac.in, akt@iitj.ac.in, m.martini@kingston.ac.uk*

**Abstract:** Various ocular diseases, such as cataract, diabetic retinopathy, and glaucoma have affected a large proportion of the population worldwide. In ophthalmology, fundus photography is used for the diagnosis of such retinal disorders. Nowadays, the set-up of fundus image acquisition has changed from a fixed position to portable devices, making acquisition more vulnerable to distortions. However, a trustworthy diagnosis solely relies upon the quality of the fundus image. In recent years, fundus image quality assessment (IQA) has drawn much attention from researchers. This paper presents a detailed survey of the fundus IQA research. The survey covers a comprehensive discussion on the factors affecting the fundus image quality and the real-time distortions. The fundus IQA algorithms have been analyzed on the basis of the methodologies used and divided into three classes, namely: (i) Similarity-based, (ii) Segmentation-based, and (iii) Machine learning based. In addition, limitations of state of the art in this research field are also presented with the possible solutions. The objective of this paper is to provide a detailed information about the fundus IQA research with its significance, present status, limitations, and future scope. To the best of our knowledge, this is the first survey paper on the fundus IQA research.

## 1 Introduction

**Image quality assessment (IQA)** is the process of analyzing the quality of an image. Subjective IQA and Objective IQA are the two types of methods that are used for the IQA process. Subjective IQA is performed by human observers and it is assumed to be the most reliable method as humans are the end users in most of the multimedia applications. The types of subjective quality assessment methods and their guidelines are provided by the International Telecommunication Union (ITU) recommendation BT.500-13 [1] (although the focus of this one is television pictures) and ITU-T P.912 [2], addressing video quality assessment methods for recognition tasks. For medical images, a standard recommendation is still missing. For a review on the methods used in the medical imaging area, the reader can refer to [3].

Subjective IQA is a very costly and tedious process that makes it unsuitable to implement in real time. To overcome these limitations, objective IQA is being used. It is a process of predicting the quality of an image by means of mathematical models with an intention to produce results similar to subjective IQA process. In order to facilitate the challenge of developing an efficient IQA method, many IQA data-sets [4–11] have been created. These data-sets contain distorted images with their subjective ratings provided by human subjects. These subjective quality ratings have been used to evaluate the performance of IQA methods.

**Types of objective IQA methods:** Objective IQA methods can be categorized into three categories: (i) Full-Reference (FR) IQA methods, (ii) Reduced-Reference (RR) IQA methods, and (iii) No-Reference (NR) IQA methods. **FR-IQA**: Under this category, the unprocessed reference image is available and it is assumed to be of best quality. FR-IQA methods are intended to evaluate the statistical dissimilarities between the pixel values of the input image (i.e., distorted) and the reference image. Such methods evaluate a number of parameters and finally combine all the evaluated results into a single scalar value indicating the overall image quality. Peak signal-to-noise ration (PSNR) is one of the most commonly used metrics in different type of domains in multimedia (i.e., audio, image, and video). Some of the popular FR-IQA methods are Structural Similarity Index (SSIM) [12], Visual Information Fidelity (VIF) [13], Most Apparent Distortion (MAD) [11], FSIM [14], and GMSD [15],



**Fig. 1**: Advanced fundus image acquisition devices: (a) D-eye, (b) Plenoptic Ophthalmoscope.

etc. Next, the **RR-IQA** methods are designed to predict the quality of an image with little information about the reference image. RR-IQA methods are useful in evaluating the quality of multimedia transmitted over a communication channel. Some RR-IQA methods are WNISM [16], EPM [17], RIQMC [18], DNT-RR [19], SIRR [20], SDM [21], and FTB [22]. In contrast to FR-IQA and RR-IQA, the **NR-IQA** methods assess the image quality without any information about the reference image. Most of the NR-IQA methods are designed to predict the quality in presence of a specific distortion, such as compression, blur, etc. BLIINDS-II [23], DIIVINE [24], BRISQUE [25], NIQE [26], and CORNIA [27] are a few popular and highly cited NR-IQA methods. In addition, recently many convolutional neural network (CNN) based FR and NR IQA methods [28–36] have been published, which have shown standout performances over the data-sets mentioned above. The detailed information about the state-of-the-art in IQA research can be found in the following references [37–39].

**Applications of IQA:** Objectively assessing the visual quality of an image has been a research field of significant interest for the researchers over the years. It is growing exponentially due to its usefulness over a wide range of applications like performance evaluation and standardization of image acquisition devices, and various image processing algorithms like image restoration and image enhancement for various type of images. Different IQA algorithms are developed over a variety of images like natural [40, 41],

screen-content [42], document [43], tone-mapped [44], 3D stereo-scopic [45], Depth-image-based-rendering (DIBR) synthesized [46], medical [47] images, etc.

Among all, ensuring the quality of medical images is one of the most important application areas. The evolution of digital medical imaging enables an easier and more reliable diagnosis process. At the same time, it raises challenges like the selection of the required display device, the compression level, the accuracy level, and the reliability of computer-aided diagnosis. The overall focus of analyzing the entire medical imaging system is to ensure that the image quality enables diagnostic reliability. It is done either subjectively or objectively, by analyzing the necessary pathological information present in the image.

While in some cases IQA metrics developed for general purpose images and video have been applied to medical images [48], specific IQA algorithms have been developed so far for the different medical image modalities like magnetic resonance imaging (MRI), computed tomography (CT), ultrasound imaging, fundus images, etc. Fundus imaging is one of the important medical imaging techniques, that is used to monitor the health status of human eyes.
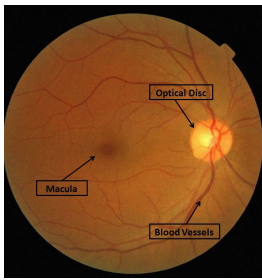
**Fig. 2**: Fundus Image

**Fundus Image Quality Assessment:** In the field of Ophthalmology, digital fundus photography is being used for early and improved diagnosis of various ocular disorders. Diagnosis through fundus images is done via inspection of morphological changes in the optical disc, macula, and blood vessels, as shown in Fig.2. Advancement in optics, computerized sensors, and picture handling led to the invention of sophisticated imaging devices like smartphone-based retinal imaging system D-EYE and the plenoptic ophthalmoscope, shown in Fig. 1(a) and (b). It allows an affordable way to capture, store and share the fundus images with minimum effort. However, such set-up of ophthalmic imaging is more open to quality impairments which may lead to false diagnosis. Furthermore, the effect of the deteriorating ratio of number-of-doctors to number-of-patients, can also be addressed with the advent of digital imaging and computer aided diagnosis (CAD) system used for various ocular diseases. CAD systems are very effective in diminishing the diagnostic oversights and consequently the false negative rates of ophthalmologists. The indicative significance of a fundus image relies on its visual quality perceived by an ophthalmologist. Regardless of whether it is manual or mechanized, to guarantee a dependable conclusion, the quality of the fundus images must be guaranteed. However, while dealing with huge amount of patient's data, the subjective quality evaluation of fundus images is quite tedious and hectic process. There is a certain need for objective quality evaluation of fundus images to mimic the quality prediction of the medical experts. For the last two decades, many research work has been published to automatically grade the fundus images.

This paper presents a detailed review of the fundus IQA algorithms developed to date with the listing of existing challenges in the field. The content of the paper is structured as follows: Section 2 discusses the comparison between the fundus images and natural images; Section 3 contains a brief discussion on factors affecting the retinal image quality, types of distortions, and the publicly available retinal image data-sets used by the researchers; Section 4 consists of a detailed survey of the retinal IQA algorithms; in Section 5 limitations in the state of the art and future scope of retinal IQA research

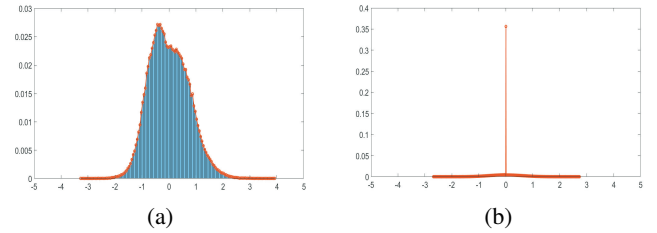**Fig. 3**: Example of Natural and Fundus images taken from [9, 49]

**Fig. 4**: Distribution of the naturalness values of the examples of Natural and Fundus images given in Fig.3.

has been discussed; finally Section 6 summarizes the paper including the important aspects and conclusions.

## 2 Need of new IQA methods for Fundus images

The only similarity between the fundus and natural images is that both are acquired from a digital camera. However, the statistical properties of fundus images vary largely from those of natural images. As mentioned in [50], the statistical behaviour of a digital image can be determined using its *naturalness property*. The *naturalness* of an image is derived in [51] by calculating naturalness value $\hat{I}(i, j)$ for pixel $I(i,j)$ of the image, as follows:

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + 1} \quad (1)$$

with $i \in \{1, 2, ...m\}$ and $j \in \{1, 2, ...n\}$, where $m$ and $n$ are the horizontal and vertical dimensions of the image $I$; $\mu(i, j)$ and $\sigma(i, j)$, estimating the local mean and contrast, are derived as follows:

$$\mu(i, j) = \sum_{k=-3}^{3} \sum_{l=-3}^{3} \omega_{k,l} I(i + k, j + l) \quad (2)$$

$$\sigma(i, j) = \sqrt{\sum_{k=-3}^{3} \sum_{l=-3}^{3} \omega_{k,l} [I(i + k, j + l) - \mu(i, j)]^2}. \quad (3)$$

Here $\omega_{k,l}$ is a circularly-symmetric 2D Gaussian weighting function. We have evaluated the *naturalness* property of various natural and fundus images. For illustration purpose, in Fig. 3 an example of natural and fundus image and in Fig. 4 the respective distribution of the naturalness values has been shown. It has been observed that the distribution of the naturalness values of natural images follows a Gaussian distribution. On the other hand, for fundus images the distribution curve is steep, indicating low naturalness. Hence, the IQA algorithms developed for natural images might not work adequately for the fundus images.

# 3 Factors affecting the quality of fundus images

The quality of a medical image can be considered to be appropriate if all the required pathological structures are clearly visible. The primary structures visible in fundus images of a healthy eye are optic disc, macula, and blood vessels, as shown in Fig.2. Following are the primary factors behind the degradation of the retinal image quality: (i) event of haze, (ii) presence of dust and dirt in the camera, (iii) dusty camera lenses, (iv) eye blink, and (v) event of occlusion by eyelashes. Uneven illumination, low contrast, blur, and light flares are the commonly seen distortions due to the above mentioned factors, as shown in Fig. 5. Artifacts generated by dust and dirt are clearly visible light spots scattered throughout the image, shown in Fig. 5(a). The event of total eye blink causes extremely poor quality fundus image as shown in Fig 5(b). An improper gamma and flash adjustment lead to color distortions like over-exposure, underexposure, and uneven illumination, as shown in Fig. 5(c), (d) and (e), respectively.

In addition, due to errors generated by light refraction, a red color print occurs throughout the fundus image, which is also referred to as "leopard print" fundus image, shown in Fig. 6(a). Next, an improper cleaning of the camera lenses creates circular rings of light in the center of the image, shown in Fig. 6(b). It is important to mention that, although the visual appearance of both Fig. 6(a) and (b) images indicates inadequate quality, experts consider it as an acceptable quality image. In order to acquire an appropriate quality fundus image, the following are the important constituents:

1. Proper space between camera and eyes;
2. Appropriate gamma and flash adjustment;
3. Quality of sensors;
4. Compression;
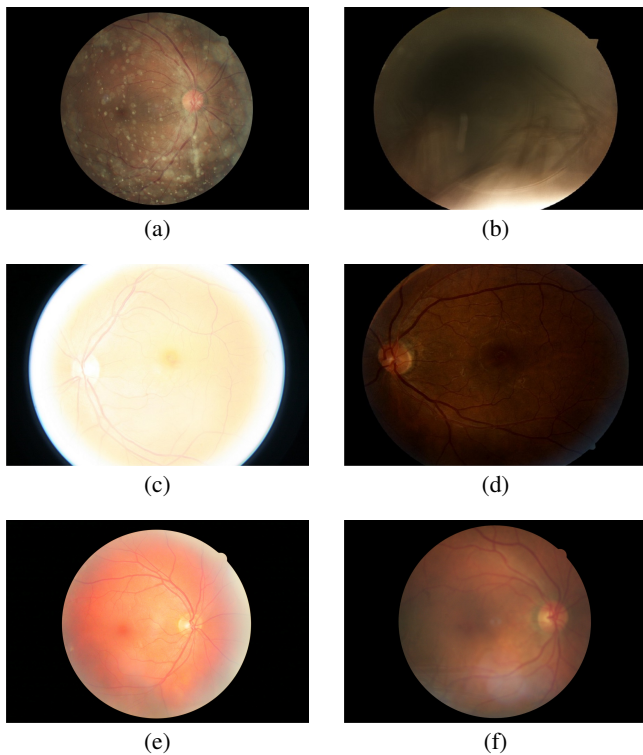5. Image resolution;
6. Color, contrast, and saturation.



**Fig. 5**: Samples of low quality fundus images with different distortions [49]: (a) Light flares, (b) Total eye blink, (c) Overexposed, (d) Underexposed, (e) Uneven illumination, (f) Blur.
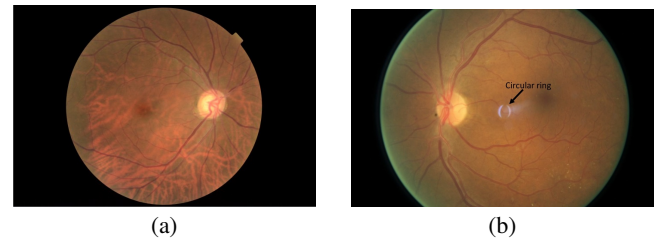


**Fig. 6**: Samples of fundus images with red color "leopard" print and circular ring [49]: (a) Leopard Print, (b) Circular Ring.

A proper setting or choice of the above-mentioned factors will be required for the desired quality of the image. Furthermore, in order to assess the fundus image quality, many different quality parameters are used, like illumination, blur, colour, contrast, and visibility of structures present in the image. In addition, a study [52] at the University of Wisconsin-Madison defined the following quality indicators for retinal images: focus and clarity, field definition, visibility of the macula, optical disc and blood vessels.

## 3.1 Fundus image data-sets

A concise details of *eighteen*(18) publicly available data-sets of retinal fundus images used by researchers is provided in Table 1. The table includes the following information about the data-sets: total number of images, spatial resolution, storage format, and the purpose of the data-set. It has been observed that most of the data-sets are built to facilitate the challenges of computer based ocular disease classification and blood vessel segmentation. There exist only one data-set, namely DRIMDB [53], for quality assessment purpose. It contains fundus images with two categories of quality: Good and Bad, without any detailed subjective quality score.

# 4 Survey of Fundus Image Quality Assessment Algorithms

On the basis of methodologies used, retinal IQA algorithms can be divided into three categories: (i) Similarity based methods, (ii) Segmentation based methods, and (iii) Machine learning based methods, as shown in Fig. 7. A concise information about these algorithms, described below in chronological order, is provided in Tables 2, 3, and 4.

## 4.1 Similarity Based Methods

A few algorithms [72, 73] reported in the literature use similarity comparison of some of the attributes of the target image with those of a set of good quality images. According to a thorough study of the related literature, Lee and Wang [72] were the first to work on objectively assessing the quality of fundus images (in 1999). Their proposed algorithm calculates the similarity measure between the intensity histogram of the target image and the template formed from a set of reference images. In order to have the reference template, the authors considered 20 high quality fundus images. The similarity metric ($C$) is calculated by performing a convolution operation between the intensity histograms of the reference template $K$ and the input image $H$:

$$C = \sum_{i=0}^{255} K(i) * H(i). \tag{4}$$

Here, $K(i)$ is the coefficient of the $i^{th}$ kernel of the template histogram, and $H(i)$ is the number of pixels with intensity value of $i$. A higher value of $C$ represents a higher correlation and similarity between $K$ and $H$. Since $K$ is obtained from high quality fundus images, a high value of $C$ indicates high quality of the target image.

**Table 1** Fundus image data-sets. DM: Diabetic Macular-Edema, M-A: Micro-Aneurysm, OD: Optical Disc, AMD: Age-Related Macular Degeneration, DR: Diabetic Retinopathy

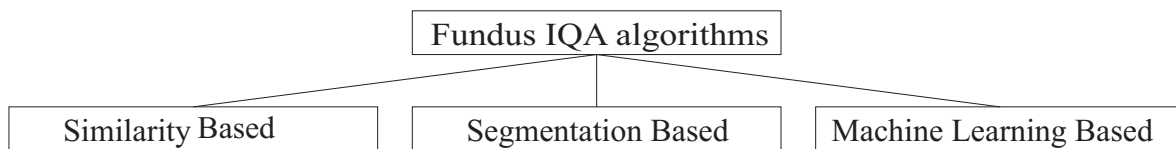| Database | Number of Images | Resolution | Format | Purpose |
|---|---|---|---|---|
| DRIMDB [53] | 216 | $760 \times 570$ | JPEG | Quality Assessment |
| HRF-DB [54] | 45 | $3504 \times 2336$ | JPEG | Vessel segmentation |
| DRIVE [55] | 40 | $768 \times 584$ | JPEG | Vessel segmentation |
| VICAVR [56] | 58 | $768 \times 584$ | JPEG | Vessel segmentation |
| STARE [57] | 400 | $700 \times 605$ | PPM | Vessel segmentation |
| REVIEW[58] | 14 | $2160 \times 1440$, $1360 \times 1024$, $3584 \times 2438$ | JPEG BMP | Vessel segmentation |
| CHASE [59] | 28 | $999 \times 960$ | PNG | Vessel segmentation |
| MESSIDOR[60] | 1200 | $1440 \times 960$, $2240 \times 1488$, $2304 \times 1536$ | TIFF | DR classification |
| ImageRet[61] [62] | 219 $1500 \times 1152$ | PNG | DR classification | |
| DRiDB [63] | 50 | $720 \times 576$ | BMP | DR classification |
| HEI-MED [64] | 169 | NA | JPEG | DM classification |
| ARIA [65] | 212 | $768 \times 576$ | TIFF | AMD classification |
| E-Optha[66] | 463 | $1440 \times 960$, $2048 \times 1360$, $2544 \times 1696$ | JPEG | M-A, Exudates detection |
| ROC [67] | 50 | $768 \times 576$, $1389 \times 1383$ | JPEG | M-A detection |
| RIM ONE [68] | 169 | NA | NA | OD segmentation |
| DRIONS-DB [69] | 110 | $600 \times 400$ | JPEG | OD segmentation |
| DRISHTI [70] | 101 | $2896 \times 1944$ | PNG | OD segmentation |
| ORIGA [71] | 650 | NA | NA | Glaucoma Analysis |



**Fig. 7**: Classification of fundus IQA algorithms

Since the histogram of an image is a global feature (it represents the number of pixels in the image with a specific value), it does not contain information about the location of the pixels. Hence, same histograms can be found for different images. Fig. 9 shows the histogram of two different fundus images (shown in Fig.8) where one is of good quality and the other one is of poor quality. Certainly, a higher value of $C$ may not always give a correct indication of the quality of the target image.

To address this, M. Lalonde, L. Gagnon, and M.C. Boucher [73] proposed (in 2001) a new similarity-based fundus IQA algorithm. The authors measured the similarity between the reference template and the target image on the basis of the following two parameters: distribution of edge magnitudes and local intensity distribution. The distribution of edge magnitude is derived by taking the squared distance between the edge magnitude histogram of the reference image template and the target image. The local intensity distribution is derived in four steps. The first step involves forming a reference grey-scale image using the set of high quality images. Second, the input image is sub-divided into uniform regions using a histogram splitting algorithm. In the third step, the histogram features are calculated for each sub-region in the target image as well as for the same sub-regions in the reference image. Finally, the summation of the squared difference between the respective mean of the histogram of each sub-region is calculated. These two derived features are used to determine the quality of the fundus image. For experiment purpose a set of *forty* (40) fundus images has been used and divided into three categories of quality: good, fair, and bad.

### 4.1.1 Advantages:

• This methodology resembles to the RR-IQA methods (a set of features extracted from the reference image is used for the quality estimation of target images) hence it can be useful in real time applications like telemedicine, where target fundus images have been transmitted over wireless networks.
• More reliable and efficient than NR-IQA methods.

### 4.1.2 Limitations:

• It is difficult to create a universal set of good quality fundus images as reference.
• It is difficult to derive an efficient and effective set of features to represent the quality class.
• Such methods are sensitive towards different types of distortions.
• The histogram features that have been used in [72, 73] do not include the structural characteristics of the fundus images.
• Such metrics do not effectively represent the Ophthalmologist's perception of fundus image quality.
• Less probability of efficient performance on cross data-set evaluation.

### 4.2 Segmentation based methods

Similarity-based methods use histogram features that do not explicitly include structural information of the image. Since structural information are of diagnostic importance, some work [74–78] have been proposed based on the segmentation of the structures present in the fundus image.

Segmentation based fundus IQA algorithms generally involve a two step process. The first step is the segmentation of structures and the second step involves its analysis, on the basis of certain parameters, in order to estimate the fundus image quality. The first segmentation based fundus IQA was proposed by D. Usher, M. Himaga, M. Dumskyj, and J. Boyce [74], in 2003. The authors have taken the pixel count of the blood vessels present in the image as
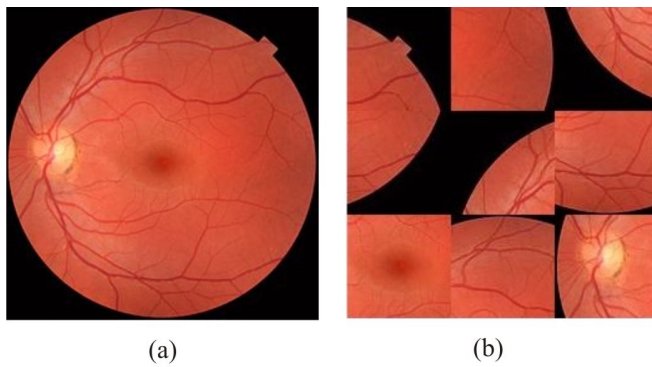
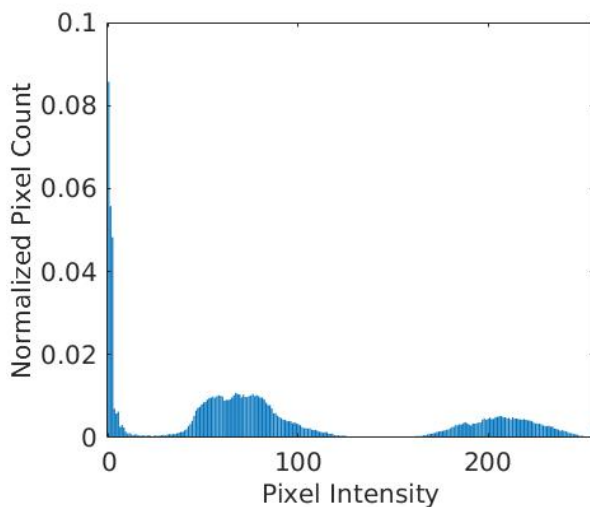**Fig. 8**: Sample fundus images; (a) Good Quality, (b) Poor Quality.



**Fig. 9**: Normalized histogram of both fundus images

the quality indicator; the larger the count the better the quality. In this work, the blood vessel extraction is achieved using matched filtering [79] followed by a region growing algorithm. In matched filtering, the input image is processed with two Gaussian kernels. One is intended to match regions of large blood vessels and another is intended to match the regions of small vessels. Further, a region growing algorithm is used over the results of these filtering processes in order to extract the blood vessels. Finally, the summation of the number of pixels belonging to the vessels is used as the quality score. For the performance evaluation of the algorithm, specificity and sensitivity, as given below, were used.

$$Specificity = \frac{a}{a+c}$$
$$Sensitivity = \frac{b}{b+d}$$
(5)

where $a$ and $b$ are the number of correctly classified good and poor quality fundus images respectively, $c$ and $d$ are the the number of wrongly classified good and poor quality fundus images respectively. On the basis of segmentation results over the set of 1746 fundus images, 84.3% sensitivity and 95% specificity have been reported. This was the first attempt in this direction with significant results. However, some important issues have been addressed in the subsequent research works.

Macula is an important part of fundus images and the blood vessels around it carry significant diagnostic information. The size of the vessels around the macula is comparatively very small and narrow. Hence, it has high chances of getting affected by any distortion. In addition, the absolute and relative position of various structures also play an important role while determining the fundus image quality. A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F.

Sharp [75] addressed these issues, and presented a segmentation-based algorithm in 2006. In this work, the overall image quality is determined by the following two parameters: (i) clarity, and (ii) field definition. The clarity feature is obtained on the basis of the visibility of the blood vessels around the macula region. The authors have segmented the blood vessels present around the macular region of the image. Furthermore, algorithm also approximates the field definition on the basis of the following parameters: location and diameter of the optical disc and visibility of the region within the 2 disk diameters (DD) around the fovea. The value of one DD is manually estimated by analyzing the optic disk diameter in good quality images, and set as 246 pixels. Overall 99.1% sensitivity, and 89.4% specificity is reported over a set of 1039 images. Although the blood vessel density around the macula provide a sufficient indication for the quality of fundus image, through this information it is difficult to capture the presence of blur in the image, as blood vessels can be visible even if they are blurred and may get added to the vessel pixel count.

Hunter, J. A. Lowell, M. Habib, *et al.* [76] addressed the difficulties with the blurred image. The algorithm, visibility of blood vessels near the fovea, that is in the macular region, is considered as the primary quality indicator. In order to examine the presence of blur, the contrast of the vessels with the background is calculated. The algorithm initially finds the location of the fovea by using an algorithm proposed by C. Sinthanayothin, J. Boyce, H. Cook, and T. Williamson [80]. Next, the segmentation of blood vessels is performed using a non-linear filtering based method termed as Tram-line algorithm [81]. A metric ($v$) quantifying the vascular information is calculated using the number of blood vessel pixels, their average distance from the fovea, and contrast with the local background. Further, the information of contrast around the region of the fovea is also quantified and used as second quality indicator ($k$). Finally, the overall quality metric is derived by taking the product of both $v$, and $k$ metrics. The authors categorized the fundus images into 5 categories of quality. The performance of the algorithm is evaluated over a data-set of 200 images and 100% sensitivity and 93% specificity has been reported.

T. Kähler, A. Budai, M. F. Kraus, *et al.* [77] also presented a quality evaluation algorithm for fundus images, in 2013, based on an assessment of blur by tracking the blood vessels. As the first step, the green channel of the fundus image is extracted and divided into a number of fixed $n \times n$ size patches. In the next step, all the anisotropic patches, discussed below, are selected and the singular value decomposition (SVD) of local gradient matrix from each anisotropic patch is calculated. As mentioned in [82], a patch that can be modeled as:

$$p(x_k, y_k) = a_1(x_k - x_c)^2 + a_2(y_k - y_c)^2$$
(6)

is called as a quadratic patch, where $p(x_k, y_k)$ is the pixel value of patch $p$ at location $k$, $(x_c, y_c)$ is the center point, and $a_1$ and $a_2$ decide the slope. A quadratic patch is called "anisotropic" patch only when $a_1 \neq a_2$. The probability of erroneous results while selecting anisotropic patches is reduced by using a proposed metric termed *vesselness measure*. This vesselness measure is derived by analyzing the blood vessel with the help of the Hessian matrix that is calculated for the green channel of the image. A local quality metric for each anisotropic patch is derived using the singular values obtained from the SVD. Finally, the global quality metric ($Q_v$) is derived by taking the addition of all the local metrics. The authors have created manually distorted fundus images from the DRIVE [55] data-set by modeling two distortions: (i) zero mean Gaussian noise, and (ii) blur using fixed size Gaussian filter. The *two* well known full-reference (FR) image quality metrics peak-signal-to-noise ratio (PSNR) and structural similarity index (SSIM) [12] are used to determine the noise levels. The final results are shown by deriving the Spearman's rank order correlation of 0.89 and 0.91 between the $Q_v$ and both PSNR and SSIM.

In 2014, H. A. Nugroho, T. Yulianti, N. A. Setiawan, and D. A. Dharmawan [78] presented a contrast assessment method in order to assess the fundus image quality. The algorithm calculates the contrast of the blood vessels as a quality parameter. The proposed

**Table 2** Summary of Similarity based fundus IQA algorithms. NS: Not Specified

| Work | Year | Quality Parameter | Categories of quality | # Images | Accuracy (%) | Pros | Cons |
|------|------|-------------------|------------------------|----------|--------------|------|------|
| [72] | 1999 | Intensity histogram | 2 | NS | NS | Less complex; Achieves high specificity as it measures similarity between characteristics of input image with good quality fundus image | Intensity histogram does not speak about the structural and local properties of the image. |
| [73] | 2001 | Edge magnitude and local intensity distribution | 3 | 40 | NS | Contains analysis of structural properties; Achieves high specificity | Not robust to distortions like high compression and blur |

algorithm starts with the pre-processing step which includes extraction of the green channel from the RGB image followed by image enhancement. In the next step, it segments the blood vessel area around the macular region using the match filtering method [83]. Finally, the algorithm calculates the proposed contrast metric by using the difference between the intensity values of pixels of blood vessels and background pixels using equation 7. In total 47 images from the [64] database have been used for the experiment purpose. The reported accuracy of the proposed work is 89.36%.

The proposed contrast metric is

$$C = |\frac{1}{x}\sum_{i=1}^{x} I_{vi} - \frac{1}{y}\sum_{i=1}^{y} I_{bi}| \qquad (7)$$

where $I_v$ and $I_b$ represent the blood vessel and background pixel intensity value respectively. $x$ and $y$ are the total number of selected pixels of blood vessels and background.

*4.2.1 Advantages:*

• These methods are based on the analysis of structural degradation in the image and effectively represent the doctor's approach for determining the fundus image quality.
• These methods can effectively perform over distortions like Color (Overexposed and Underexposed), Uneven Illumination, Additive Gaussian noise, and Blur.
• They achieves high specificity and sensitivity under fixed assumptions like fix shape, size, and location of the structures.

*4.2.2 Limitations:*

• The assumption behind segmentation based quality assessment is that poor segmentation results reflect poor fundus image quality. Here, segmentation algorithms work under the fix assumptions and criterion like fix shape, size, and location of the structures. Any changes to these parameters may lead to the decreased performance while cross-data set evaluation.
• Segmentation algorithms are expected to give good results even in presence of different noises and this in turn results into an erroneous quality assessment result. For example the Canny edge detection algorithm used in [75] reduces the effect of Gaussian noise. Therefore, it might not produce reliable and correct quality assessment results in presence of Gaussian noise.

*4.3 Machine learning based methods*

Machine learning (ML) based fundus IQA algorithms classify the fundus images into predefined categories of quality by learning from the samples. The process involves the following three steps: (i) Feature Extraction, (ii) Training and validation of the model, and (iii) Testing. ML based fundus IQA methods can be further classified into three categories based on the type of features extraction approach: (i) Feature Extraction Based on Structural Analysis, (ii) Feature Extraction Based on Generic Image Statistics, and (iii) Feature Extraction Based on Convolutional Neural Networks (CNN) Models. A brief introduction of these algorithms is provided in the subsequent subsections.

*4.3.1 Feature Extraction Based on Structural Analysis:* In 2006, M. Niemeijer, M. D. Abramoff, and B. van Ginneken [84]

presented the first machine learning based framework for fundus IQA by using the image structure clustering (ISC) method. The ISC method identifies the primary structures present in the fundus image by creating the clusters of the outputs received from a set of multi-scale filters. The authors have used a set of five rotation and translation invariant filter-bank at different scales to perform the ISC in fundus images. A total of five clusters have been computed with the input image using the filter-bank. Further, a set of features that contains the histogram of the ISC clustered pixels and the raw histogram of red R, G, and B planes was extracted from each cluster. This feature set is used to train four different classifiers: (i) support vector machines (SVM) with radial basis kernel, (ii) quadratic discriminant classifier, (iii) linear discriminant classifier, and (iv) k-nearest neighbor (k-NN) classifier. As mentioned in the result section, most astounding precision is accomplished by the SVM classification method with 99.68% accuracy. A total 1000 fundus images, taken from a proprietary data-set, have been used for both training and testing in order to divide the fundus images into two categories of quality: poor and good.

In 2008, L. Giancardo, M. D. Abramoff, E. Chaum, *et al.* [85] mentioned that one of the limitations of all the previously discussed works is their running time. To overcome this issue, the authors have presented a fast framework for fundus image quality estimation. The proposed algorithm incorporates both the approaches of segmentation and machine learning methodologies. Initially, the circular region of interest is localized and segmented from the fundus image by using a circular mask. This circular mask is obtained from the green channel of the image with the help of a region growing algorithm. Further, the vessel segmentation process is implemented by using a method based on mathematical morphology [86]. Further, local vessel density is calculated from the obtained segmented vessel area by dividing the image into local windows. A total of 18 local polar windows have been formed and the area of vessels for each window is calculated. Vessel density feature obtained from all the 18 local windows are used to train and test the classifiers in order to classify the fundus images into two categories of quality: Good and Poor. Classification is tested over two different classifiers: (i) SVM, and (ii) k-NN and the results reported are more favorable in case of SVM. The proposed system is tested over 82 fundus images with 100% sensitivity and 92% specificity.

In 2010, Paulus, J. Meier, R. Bock, J. Hornegger, and G. Michelson [87] presented a system for retinal IQA by combining both structural information and generic image quality statistics. The structural information includes visibility of optical disk and blood vessels, and generic quality indicators contains information about the illumination and contrast. The final feature set consists of three features: (i) clustering, (ii) sharpness, and (iii) Haralick texture features. Structural information is determined by the clustering method, in order to compute the clear differentiation between structures present in the image. It is determined by using the k-means clustering method. Total ten manually segmented images of $k$ clusters with fixed mean values have been used for initialization of cluster centers. Now, for each input image, the cluster size and difference between the values of each cluster mean is calculated. Further, generic quality features are quantified by the sharpness metric and Haralick features [88]. The sharpness metric is calculated by using the gradient magnitude of the image. Haralick feature metrics are computed from the co-occurrence matrix of the image that is intend to represent the texture features of the image. To evaluate the illumination and contrast features, the authors have utilized three Haralick metrics mostly known as texture metrics. Finally, all the above-mentioned feature set is

**Table 3** Summary of Segmentation based fundus IQA algorithms. SN: Sensitivity, SP: Specificity, SC: Spearman's Correlation

| Work | Year | Quality Parameter | Categories of quality | # Images | Accuracy (%) | Pros | Cons |
|------|------|-------------------|------------------------|----------|--------------|------|------|
| [74] | 2003 | Blood vessel density | 2 | 1746 | SN: 95, SP: 84.3 | Matched Filtering is easy and simple to implement, gives adequate performance in presence of Gaussian noise | Effective only under pre-defined assumptions, not effective for smaller or larger size blood vessels |
| [75] | 2006 | Blood vessel pixel count and field definition | 2 | 1039 | SN: 99.1, SP: 89.4 | Emphasis given to the Field definition and relative positioning of the structures present in the image | Assumes fix parameters for the locations of the structures hence not effective for cross data-set evaluation |
| [76] | 2011 | Visibility of blood vessels and field definition | 5 | 200 | SN: 100, SP: 93 | Effective performance for blur and gaussian noise | Sensitive towards different type of distortions like uneven illumination |
| [77] | 2013 | Blood vessel pixel count | NA | 58 | SC: 0.89 with PSNR SC: 0.91 with SSIM | Due to Patch-wise analysis this method is very effective in presence of uneven distortions additive Gaussian noise and Gaussian Blur | Method is tested over only two distortions, the naturalness property of images is low hence results obtained from SSIM and PSNR metrics may not be of much significance |
| [78] | 2014 | Contrast of blood vessels | 2 | 47 | 89.36 | Effective for Color and Contrast distortions | Effective only under pre-defined assumptions like size and shape |

used to train the SVM classifier in order to classify the fundus images into two classes: Good, and Poor. The proposed system is tested over 301 fundus images and achieves an accuracy of 95.3%.

Another work in this category was proposed by R. Pires, H. F. Jelinek, J. Wainer, and A. Rocha [89] in 2012. The proposed work is influenced by the work of Fleming et al. [75] that uses field definition of the fundus image as a primary quality indicator. The authors have inspected the quality of fundus images by analyzing the field definition and the level of blur present in the image. A set of 40 high quality fundus images have been selected as reference images. The verification of field definition is performed by analyzing the structural similarity between the reference image and the input image using the well-known SSIM method. Detection of blur is achieved by calculating a set of features, namely: (i) area descriptor, (ii) visual dictionary descriptor, (iii) blurring descriptor, (iv) sharpening descriptor, and (v) concatenation of blur and sharpness descriptor. Area descriptor estimates the area of blood vessels within the image. It is calculated using the well known Canny edge detection algorithm. Visual dictionary is built by detecting the stable point of interests in the image using a well known method namely Speeded Up Robust Feature (SURF) [90]. Further, in order to model the blur and sharpness measure, the authors used the input image as the reference image. The input image is blurred and sharpened progressively with different intensities and then the similarity measure between the input image and its transformed versions is calculated. The assumption behind the idea is that a poor quality image will be more similar to its distorted version rather than a good quality image. All the above derived set of feature vectors have been used to train and test the SVM classifier for generating the final results. Extensive experiments and results have been shown for the verification of field definition and blur detection with 96% and 95.5% accuracy.

In 2012, H. Yu, S. Barriga, C. Agurto et al. [91] presented another linear regression based retinal IQA method. In the proposed work the authors extracted various statistical features in order to train the regression model more efficiently. In addition, fundus images are divided into four categories of quality. The algorithm consists of two steps: (i) feature extraction and (ii) PLS regression. The feature extraction step involves the extraction of four different features: (i) vessel density, (ii) histogram, (iii) texture features, and (iv) local sharpness features. Vessel density feature is calculated by taking the ratio of the area of blood vessels over the area of the field of view. Blood vessels are segmented by using a method based on multiscale enhancement and second order entropy threshold [92]. Mean, variance, skewness, and kurtosis features are extracted for determining the histogram features. The texture features are derived using five Haralick texture features: (i) second order entropy, (ii) contrast, (iii) correlation, (iv) energy, and (v) homogeneity. Local sharpness

features are determined by using a well-known method named cumulative probability blur detection (CPBD). Each fundus image from the training data-set is assigned a quality score by the retinal expert, and graded into four categories of quality: high, medium, low, and reject. A linear relation is assumed between the derived features and the quality score. All the derived features are considered as the independent variable and quality score as the dependent variable. Finally, the PLS regression algorithm is implemented in order to estimate the overall quality. The proposed algorithm is tested over a proprietary data-set of 1884 fundus images and achieved 95% performance accuracy.

Another method proposed in this category consists of both segmentation and machine learning G. J. Katuwal, J. Kerekes, R. Ramchandran, C. Sisson, and N. Rao [93] proposed a retinal IQA algorithm for fundus images by analyzing the symmetry of retinal blood vessels. Initially, the stationary wavelet transform followed by median filtering, dilation, and circular masking, is used to extract the retinal blood vascular structure. Further, the image is horizontally divided into two equal parts, followed by dividing both halves into 10 equal size vertical windows. Now, with the help of segmented vasculature, the following four features are calculated: (i) global vessel density (GVD), (ii) local vessel density (LVD), (iii) difference between LVDs of top and bottom local windows, and (iv) difference between sum of LVDs in top half and bottom half. The GVD metric is the ratio of the number of blood vessel pixels and the total number of pixels present in the image. The LVD metric is similar to GVD, calculated individually for each window. Finally, all the derived feature vector set is used to train the SVM classifier. The proposed system divides the fundus images into 5 classes with reported performance accuracy of 60%. A proprietary data-set of 88 images has been used for the experiment.

Most of the methods discussed in this category are intended to divide the fundus images into two categories of quality. However, a few of the methods [73, 76, 77] attempted to classify the fundus images into more than two classes. The limitation of binary classification based retinal IQA approach is that it is unable to effectively model the doctor's perception of fundus image quality, as it draws a strict boundary between the two classes. An average quality fundus image that can be used for the diagnosis and closer to the boundary can be classified as poor quality image and vice versa. In both the conditions, the performance of the CAD systems will degrade. An IQA method that can provide a quality score via a number within a fixed range can provide better insights into doctor's judgment for the retinal image quality. The next work in this category is one of the few works that producea quality score for the fundus images rather than simply classifying the fundus images into categories of quality. F. Yin, D. W. K. Wong, A. P. Yow et al. [94] presented a retinal IQA algorithm named as automatic retinal interest evaluation system

(ARIES). The proposed algorithm is divided into three steps: (i) retinal image identification, (ii) confirmation of non-retinal images, and (iii) quality assessment. The first step involves the identification of the fundus images. Bag of visual words is used to train the SVM classifier to classify the fundus and non-fundus images. The second step is intended to suppress the effects of wrong classification results, as it is believed that a fundus image with bad quality might be wrongly classified as a non-fundus image. A reference fundus image template is created from a set of high-quality images. Then the SSIM values are calculated between each non-fundus image and the reference image. All the images with higher SSIM values are considered as fundus images. Next, the quality assessment process involves training of the SVM classifier with the following feature set: contrast ratio, blur ratio, entropy, blood vessel density. Contrast is calculated as the ratio of the mean intensity value and standard deviation of pixels for each color channel (R, G, B) individually. The blur metric is calculated by the method described in [95], that is based on the intensity range of the pixels. Next, blood vessel density is derived as the ratio of number of blood vessel pixels and total number of pixels in the image. Blood vessel pixels are extracted by using bottom hat filtering algorithm. The bottom-hat filtering method involves performing the morphological closing operation in the image followed by subtracting the original image from the result. Finally, all the derived features are used to train the SVM classifier. Another contribution of the work is that it does not directly usethe SVM classification results. Rather, the output of the SVM decision function is normalized to generate a quality score named as retinal image quality score (RQS). The value of RQS ranges from 0 to 1, where a higher value reflects the higher fundus image quality. The proposed system is trained and tested over 740 fundus images and achieved 95.4% accuracy.

### 4.3.2 *Feature Extraction Based on Generic Image Statistics:*

To the best of our knowledge, H. Davis, S. Russell, E. Barriga, M. Abramoff, and P. Soliz [96] represented the first retinal IQA algorithm that includes human perception for the fundus image quality in the form of subjective quality scores. A total of 400 artificially distorted images are created using Gaussian blur and intensity shift, from the images taken from the Messidor [60] data-set. All images are assigned a quality score by the ophthalmologists and divided into two classes of quality: (i) good and (ii) poor. The first step of the algorithm is to divide the image into seven equal size blocks for each channel of the two color models: RGB and CIE L*a*b* space. The CIE L*a*b space model is used because of its ability to comprehend the relation between change in color values and visual properties. Blur, overexposure, and underexposure are considered the primary artifacts to be observed in the work. They are mathematically demonstrated by deriving a set of six statistical properties of the pixels: mean, skewness, entropy, spatial frequency, and median. A linear relation is assumed between the features and the quality score. Features are considered to be as the independent variable and quality score as the dependent variable. Finally, the partial least square (PLS) linear regression model is trained to estimate the fundus image quality. The proposed system has reported an accuracy of 99.3%.

Based on four generic quality indicators: color, focus, contrast, and illumination, J. M. P. Dias, C. M. Oliveira, and L. A. da Silva Cruz [97] presented a retinal IQA algorithm (in 2014). The flow of the algorithm is as follows: Pre-processing, feature extraction, and fusion of features for final classification. The algorithm starts with a pre-processing step to exclude redundant background information and to retain only information of retinal structures by applying masking and cropping operation over the image. Feature computation includes an individual assessment of color, focus, contrast, and illumination features. Color assessment classifies the color of a retinal image into three categories; bright, dark, and normal. It is implemented by color indexing using the histogram back projection method presented by Swain and Ballard [104]. Three color maps for all three categories are obtained by the statistical analysis of 11 bright, 7 dark, and 232 normal images. Next, the focus assessment step involves classifying the image into the blurred, borderline or focused category. After converting the color image into grayscale, the focus is quantified by applying ther Sobel operator to the retinal

image followed by a multi-level focus analysis algorithm. Further, the contrast assessment algorithm classifies the retinal image into two classes: low and high. It is implemented by using color indexing, similar to the color assessment algorithm. Further, the illumination assessment is achieved by using the mean and variance properties present in the indexed image. Finally, all the extracted features are used as input to train three classifiers: Feed-forward back propagation neural network, radial basis function networks, and k-Nearest Neighbor. The most satisfactory results have been reported for the feed-forward neural networks classification method. A set of 2032 retinal images has been used for the experiment, that achieved sensitivity of 99.76%, and specificity of 99.49%.

Next, D. Veiga, C. Pereira, M. J. Ferreira, L. GonÃģalves, and J. Monteiro [98] presented a fuzzy classification based retinal IQA algorithm (also in 2014). The algorithm examines the image sharpness and field of view (FOV) area in order to differentiate between the low and good quality fundus images. Initially, the green channel of the fundus image is used to derive the noise mask and FOV mask. The noise mask determines the unevenly illuminated zones present in the image. The FOV mask is used for the segmentation of the area around the macula including optical disk. Next, both noise and FOV masks are compared to substantiate if their common area is greater than a predefined threshold. If the common area is less than the given threshold, then the image is considered as a poor quality image and the algorithm terminates. Otherwise, the process enters the next step that is focus evaluation. Focus analysis is done with three predefined focus operators: (i) wavelet-based, (ii) moment-based, and (iii) statistics-based. The output generated from the focus operators is fed as feature input to the fuzzy classifier to get the final result. A total of 1454 number of fundus images have been used for the experiment, out of which 1200 were taken from the [60] data-set and 254 from a proprietary data set. The reported accuracy of the proposed algorithm is 98%.

Another statistical quality parameters based retinal IQA method was proposed by the Z. Yao, Z. Zhang, L.-Q. Xu, Q. Fan, and L. Xu [99] in 2016. Primarily two quality parameters have been taken into consideration: (i) uneven illumination, and (ii) blur. In order to quantify these parameters the following features have been extracted: statistical characteristics of pixels, texture features, central statistical characteristics, symmetry, wavelet features, and blur metric features. Mean, standard deviation, skewness, kurtosis, and entropy parameters are calculated and used as statistical characteristics of the image. In order to model the texture features, first, the co-occurrence matrix is derived from the image. With the help of the co-occurrence matrix, the following features are derived: contrast, correlation, energy, and homogeneity, and used as texture features. A central region in the image containing the fovea part has been selected and all the above-mentioned features are calculated, which are termed as central statistical characteristics. The symmetry of the image is predicted by calculating the mean values of 9 squared regions selected in the image. Furthermore, the analysis of the blur component in the image is based on the idea that the presence of blur results in the loss of the high-frequency components in the image. Using the Harr wavelet transform, low and high-frequency components are separated from all the three color channels of the fundus images. All the statistical features discussed above are derived for each of the three high-frequency components. Finally, a well known method based on the cumulative probability of blur detection is used to extract the blur metric. The feature extraction step collectively forms a 113-dimensional feature vector that is used to train the SVM in order to classify the poor and good quality fundus images. The overall accuracy reported is 91.38%. All the experiments are carried out over a proprietary data-set of 3224 fundus images.

In the next work under this category the authors supported the importance of retinal IQA research with the fact that the portable and handy fundus imaging devices are more sensitive towards distortions. Based on the theory of Human Visual System (HVS) framework, in 2016 S. Wang, K. Jin, H. Lu *et al.* [100] presented a machine learning approach for quality prediction of portable fundus images. Initially, the quality scores are collected by the subjective evaluation from three ophthalmologists for a dataset of 536 images. It is important to note that the quality scores are collected for the following

**Table 4** Summary of Machine learning based fundus IQA algorithms divided on the basis of feature extraction approach. DT: Decision tree, FC: Fuzzy classification, ANN: Artificial neural network, SP: Specificity, SN: Sensitivity, MF: Membersip function.

| | Work | Year | Method | Quality Parameter | Categories of quality | Images | Accuracy (%) | Pros | Cons |
|---|---|---|---|---|---|---|---|---|---|
| Structural Analysis Based | [84] | 2006 | SVM | Image structure clustering | 2 | 2000 | 99.86 | Features extracted are rotation and translation invariant | Less effective for the detection of Color and Blur distortions |
| | [85] | 2008 | SVM | Blood vessel density | 2 | 84 | SN: 100, SP: 92 | Can effectively handle morphological changes in size blood vessels | Not effective for non-uniform distortions, tested over very small data-set |
| | [87] | 2010 | SVM | Structural and generic features | 2 | 301 | 95.3 | Method combines both structural and generic properties, k-mean clustering is computationally faster | k-means clustering is sensitive to data-set any changes to the data-set will lead to completely different results |
| | [89] | 2012 | SVM | Field definition and blur | 2 | 6696 | 96, 95.5 | Canny edge detection algorithm used for blood vessel segmentation is effective in presence of noise | Less efficient to assess the quality of fundus image in presence of Gaussian noise |
| | [91] | 2012 | PLS | Vessel density histogram, texture and sharpness | 2 | 1884 | 96 | Structural and generic properties are evaluated separately, included the subjective quality scores | Method is not evaluated over the artifacts generated due to high compression |
| | [93] | 2013 | SVM | Symmetry of blood vessels | 5 | 88 | 60 | Stationary wavelet transform (SWT used for the vessel segmentation is translation invariant | Evaluated over a very small data-set, not robust to color and blur artifacts |
| | [94] | 2014 | SVM | Contrast, blur and blood vessel density | 2 | 370 | 95.8 | In comparison to other methods this method quantifies the image quality on a fix scale, computationally efficient | The performance of Bottom hat filtering method used for the vessel segmentation relies upon the structuring element |
| Generic Image Statistics Based | [96] | 2009 | PLS | Statistical features | 2 | 2000 | SN: 100, SP: 96 | CIE L*a*b space model is effective for the extraction of color information | Not robust to distortions occurred due to compression and erroneous transmission |
| | [97] | 2014 | ANN | Color, focus, contrast and illumination | 2 | 2032 | 99.87 | Method analyzes the focus, color, contrast, and illumination individually, tested over comparatively large data-set | Not robust to the artifacts generated due to high compression and erroneous transmission |
| | [98] | 2014 | FC | Uneven illumination focus | 2 | 1454 | 98 | Method efficiently fuses the generic and structural properties, FC works efficiently with imprecise and incomplete data | MF used in FC works with "If and else" reasoning. However, it is not necessary that all the combined factors are equally important |
| | [99] | 2016 | SVM | Uneven illumination and blur | 2 | 3224 | 91.3 | Due to block-wise analysis method is effective in presence of non-uniform distortions | Method is not efficient in presence of refractive errors |
| | [100] | 2016 | DT and SVM | Uneven illumination, color, blur and contrast | 2 | 536 | 94.52 | Data-set includes subjective evaluation, method works on HVS based feature extraction methods | Method is not evaluated over high compression and transmission error distortions |
| | [101] | 2017 | DT, DL and SVM | Uneven illumination, naturalness property and structural information | 2 | 4372 | 92.39 | Gabor filters used for the optic disk segmentation are rotation, translation scale and illumination invariant | Classification accuracy is not up to the expected satisfactory level |
| CNN Based | [102] | 2016 | CNN | High level features extracted from CNN | 2 | 101 | 99.87 | Simple network architecture | Evaluated over a very small data-set |
| | [103] | 2017 | CNN | Fusion of features extracted from CNN and saliency maps | 2 | 5200 | 95.42 | Unique feature extraction, evaluated over real distorted images significantly large data-set is used | No significant change in the performance of the system while not using the features extracted from saliency maps |

three quality parameters on a scale of *two*: (i) uneven illumination, (ii) blur, and (iii) contrast. The proposed algorithm involves three major steps: (i) Pre-processing, (ii) HVS based feature extractions, and (iii) Machine learning. The pre-processing step separates the extraneous background information from the image by using a circular mask. Next, the feature extraction step analyzes the presence of the following three features in the image: (i) Multichannel sensation, (ii) Just noticeable blur (JNB), and (iii) Contrast sensitivity function (CSF). Multi-channel sensation parameter is modeled to discern the illumination and color features. Initially, the image is transformed from RGB space to HIS (H: hue, I: intensity, S: saturation) space. Further, two masks: illumination (MI1), and color (MI2) are produced using the thresholding method and combined to produce a single mask (MIROI). Finally, the multi-channel sensation parameter

is derived by taking the ratio of MIROI and MIs. Next, the JNB feature is derived for determining the level of blur present in the image by combining a well known cumulative probability blur detection (CPBD) method with a vessel density map feature. The vessel density map is derived using a morphological algorithm. Further, CSF is used to quantify the level of contrast present in the image. The final quality prediction is performed using two different methods: decision tree based method and machine learning based using the SVM. The decision tree-based structure simply compares the derived values of all the three parameters with the ground truth data and predicts the quality. In case of machine learning, all the extracted features are used to train and test the images using the SVM. It is reported that the SVM has achieved much better results in comparison with decision tree method. Two proprietary data-sets, namely: LOCAL1 and LOCAL2, and two public data-sets (DRIMDB and

DRIVE [53, 55]) have been used for the 536 fundus images for the experiment purpose.

F. Shao, Y. Yang, Q. Jiang, G. Jiang, and Y. S. Ho [101] presented a retinal IQA method based on the idea similar to [100]. All the steps involved in [100] and the proposed method are the same except that the features are used as quality parameters. To assess the quality, the authors have quantified three quality parameters: illumination, naturalness, and structure. The illumination property is examined by identifying three optimal threshold values in order to get the effects of dark, bright, and uneven illuminations. Next, the naturalness feature is based on the assumption that an image must look natural. In order to quantify the naturalness index (NI), the authors have trained the multivariate Gaussian regression model with high-quality fundus images. The NI of the input fundus image is determined by testing the image over the trained regression model. Finally, the location of the optical disc is used as structural information which is modeled with the help of Gabor filters. Fundus images have been classified into two classes: accept and reject. With the help of the above-calculated features, the authors have experimented two strategies for the quality prediction: (i) threshold based by using the decision tree, and (ii) learning based by using SVM and dictionary learning (DL). The results section reported that the algorithm performs best in the case of SVM and least in case of DL. A total of 4372 fundus images are used for the experiment, with reported sensitivity and specificity of 94.69%, and 92.29%, respectively.

*4.3.3   Feature Extraction Based on Convolutional Neural Networks (CNN) Models:*   All the previously reported machine learning based fundus IQA algorithms are based on the conventional hand-crafted feature learning methods. In recent years, the convolutional neural networks (CNN) based automated feature learning method outperforms conventional feature learning methods by a large performance gap. The automated feature learning has the ability to learn highly optimized features, that increases prediction accuracy. The literature shows that in recent years deep learning is successfully applied to the IQA framework for natural color images [105–107]. The first CNN based fundus IQA algorithm was proposed by D. Mahapatra, P. K. Roy, S. Sedai, and R. Garnavi [102] in 2016. The proposed CNN architecture classifies the fundus images into two classes: gradable, and ungradable. A CNN is trained with 101 fundus images obtained from Drishti data-set [70] and the trained network is tested on different data-sets of fundus images to divide the images. As 101 is a very limited sample to train a CNN, to overcome the issue the authors have divided the images into multiple overlapping patches of size 150x150 and labeled the same as the original image. Due to unavailability of ungradable fundus images, the authors manually created them by modeling three distortions: (i) Gaussian noise with mean zero and varying variance, (ii) Salt and pepper noise with varying noise density, and (iii) Speckle noise. In the training phase, in the first layer of CNN architecture, each patch is convolved with 20 kernels of size $7 \times 7$ followed by max pooling of $4 \times 4$ to reduce each feature map into $36 \times 36$. In the second layer of the CNN, all 20 feature maps are convolved with 50 kernels of size $5 \times 5$ followed by again $4 \times 4$ max pooling. It generates 1000 ($50 \times 20$) feature maps of size $8 \times 8$. Finally, the last layer is the logistic regression for generating the final output and stochastic gradient descent is performed with negative log likelihood as loss function. The proposed system achieved 100% sensitivity, and 99.8% specificity.

One major limitation that exists in the previously presented work in this category is that it is modeled over a very small data-set. Typically, a CNN models requires a large amount of data for learning. In view of this, F. Yu, J. Sun, A. Li *et al.* [103] presented a new CNN model to classify the fundus images into two classes of quality. The proposed model has two parallel steps for feature extraction: (i) Feature extraction from Saliency maps and (ii) Feature extraction from CNN model. Initially, every fundus image is resized to $256 \times 256$ resolution. Thereafter the saliency maps are obtained using the frequency-tuned salient region detection method presented in [108]. Further the saliency maps are reduced to $32 \times 32$ ($1024 \times 1$) blocks by taking the mean value from every $8 \times 8$ non overlapping block.Next, in CNN architecture contains a total of

five convolution layers and one fully connected layers. The resized fundus image is processed through the five convolution layers to generate a $4096 \times 1$ features vector. Finally, the features obtained from the saliency map ($1024 \times 1$) and CNN network ($4096 \times 1$) are fused to create a new and unique feature vector of size $5120 \times 1$. The obtained feature vector is further used with multilevel kernel SVM classifier to classify the fundus images into good and poor categories. A total of 5200 fundus images have been taken from the Kaggle dataset [49] for the experiment purpose and achieved 95.42% accuracy.

*4.3.4   Advantages:*

• Machine learning based fundus IQA methods can be modeled easily with different data sources.
• Once a model is trained, it can produce fast and real-time predictions.
• These models has ability to improve its accuracy and efficiency over the time without any human intervention.
• CNN based models showed best performance for IQA [28–36] and outperforms the conventional IQA methods.

*4.3.5   Limitations:*

• Most of the machine learning algorithms are data hungry.
• Finding a sufficiently large fundus image data-set for quality assessment purpose is a big challenge.
• Difficult to train an efficient model in absence of required data-set. Consequently, it reduces cross data-set performance.
• CNN based models are computationally expensive in comparison to other machine learning algorithms.

**Overall Observation**: It is important to mention here that all the quality parameters that are used in the development of the retinal IQA methods can be divided into two categories: (i) generic parameters, and (ii) structural parameters. Generic quality parameters determine the information regarding the illumination, contrast, and colour properties of the fundus image. These parameters are derived by analyzing the local and global properties of pixels in the image like histogram, mean, variance, and skewness etc. In addition, the structural information that indicates the visibility of the anatomical structures present in the fundus image is derived using segmentation algorithms. In most of the segmentation based retinal IQA methods, the authors have segmented the blood vessels and analyzed their properties like density, symmetry, or colour information. One limiting factor of retrieving the structural information through segmentation based methods is that it highly depends upon the accuracy of the segmentation algorithm. An error-prone segmentation leads to the wrong quality prediction result.

## 5   Challenges & Future scope

A considerable effort has been made by the researchers towards the development of fundus IQA algorithms. However, many fragments of stones of unresolved challenges and unanswered questions exist in the path, that need to be removed. In the subsequent sub-sections, we discuss some of the challenges in this field.

*5.1   Interpretation of ophthalmologist's perception*

Throughout the years of evolution of image acquisition and display devices, one factor that has not been changed is the HVS. Here, the expectation of ophthalmologists from a good quality fundus image also remains the same. Due to its persistent structural property, ophthalmologists assess the quality of a fundus image on the basis of some fixed quality parameters: visibility of blood vessels & optic disc, blur, color information, etc. Therefore, an efficient retinal IQA algorithm must incorporate the relation between the physical change in the quality parameters and the respective perceptual changes. In addition, this also gives rise to the challenge of determining the relative importance of the quality parameters. Here, relative importance
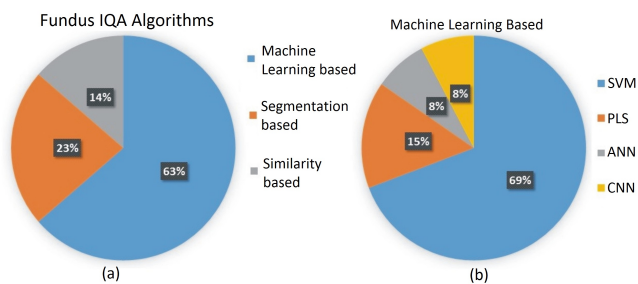
**Fig. 10**: Analysis of fundus IQA algorithms.
(a) Pie chart showing the statistics of the fundus IQA algorithms on the basis of methodologies used, (b) Pie chart showing the statistics of fundus IQA algorithms based on machine learning methods.



**Fig. 11**: Distortion generated due to erroneous transmission [49]

indicates the contribution of a quality parameter while determining the overall fundus image quality.

### 5.2 Number of classes

Another challenge is to determine the number of classes of fundus image quality. It can be analyzed from Table 2, 3, and 4 that most of the reported work divides the fundus images into binary classes of quality: Good and Poor. For manual diagnosis, it is not always required because quality prediction for the fundus image highly depends upon the experience level of the ophthalmologist. An experienced ophthalmologist may not always require the high-quality image and can diagnose effectively from a comparatively lower quality retinal image as well. For example, the image shown in Fig. 5(k) may be suitable for diagnosis by an expert ophthalmologist and might not be suitable for a relatively less experienced one. In case of CAD, an average quality image might be classified as a poor quality image causing an image recapturing. Definitely, it will lead to a waste of effort and time. Hence, in order to ignore the unnecessary wastage of efforts and resources there must exist one more category indicating the average quality fundus images. Although, few of the reported fundus IQA algorithms [73, 76, 77, 93] graded such images into more than two classes. However, none of them found prevalent in use due to the unavailability of comparative performance analysis.

### 5.3 Ground-truth information

As mentioned earlier, the main limitation in the field of fundus IQA research is the unavailability of comparative performance analysis. It is because no benchmark fundus image data-set exist for the quality assessment purpose. In literature, retinal images used for experiment purpose are mostly taken from different proprietary data-sets. As a result, the comparative analysis of the performance of retinal IQA algorithms is nearly missing in the literature. U. Sevik [53] presented a retinal image database for quality assessment purpose which contains two classes of retinal image quality: good and poor. However, the importance of the proposed data-set would have been significantly increased if it could include the subjective quality scores for each image.

### 5.4 Distortion Identification and Image enhancement

In order to get improved segmentation and disease classification results from fundus images, one viable solution is to use an appropriate image enhancement method. Image enhancement techniques are used to achieve improved visual appearance of the anatomical structures present in the fundus image. A fair amount of work has been reported [109–116] to facilitate the fundus image enhancement challenges. However, in an automated system, there is a need to have a tailor-made fundus IQA algorithm to determine if the given image needs enhancement. In contrast, from Fig. 10 it can be observed that a large proportion of the retinal IQA research work consists of binary classification systems using machine learning algorithms like SVM.
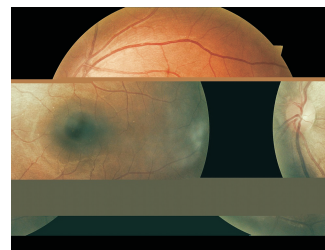
In case of only two classes of quality, if an image is classified as poor by the IQA algorithm, the question remains about the need of enhancement.

One challenge in the path of developing fundus image enhancement algorithms is to identify the type of distortion present in the image. It is because an image enhancement method is developed to normalize the effects of a specific distortion in the image. Therefore, an efficient method to identify the type of distortion present in the fundus image is highly required. In addition, retinal IQA algorithms developed to date have never been tested over the artifacts generated due to compression and erroneous transmission. One real time example of a distorted fundus image due to erroneous transmission is shown in Fig. 11. Such type of impaired fundus images greatly affects the performance accuracy of the CAD systems. In the era of telemedicine, the effects of the artifacts occurred due to compression and transmission must not be ignored while assessing the image quality.
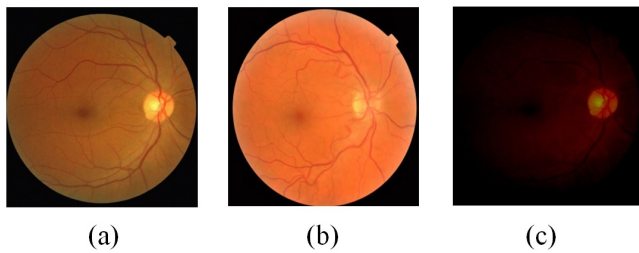
### 5.5 Suggestions

The previous subsections of this paper discuss the recent work and challenges in fundus IQA research. In this subsection, few suggestions to solve the challenges are provided. In order to address the first three challenges, a fundus IQA data-set can be created with the following peculiarities.

1. Collection of fundus images with all possible real time distortions.
2. Subjective quality assessment (SQA) of the images by various ophthalmologists with varying level of experiences.
3. Subjective quality scores collected for both structural and generic quality parameters like visibility of blood vessels (F1), visibility of optical disc (F2), color (F3), contrast (F4), and blur (F5).
4. In addition to the quality scores, a final grading of images should be done by the doctor into three classes: (i) Good, (ii) Average, and (iii) Bad.

In Fig. 12 we have shown an example of the suggested SQA process. It can be observed here that for each image we have quality scores for the five above mentioned quality parameters and an associated quality class provided by the doctors. In order to validate the data, a classification model can be trained using the quality scores obtained for the quality parameters (F1-F5) to classify the fundus images into the given three quality classes. After obtaining the optimally trained classification model, the associated coefficients or weights of the respective features can be easily derived. These coefficients will reflect the relative importance of quality parameters while determining the overall fundus image quality. This whole procedure will help to understand the ophthalmologist's perception of fundus image quality. In addition to that, the "Average" class of fundus image quality will lead a fundus IQA algorithm to automatically determine the need for enhancement.

## 6 Conclusion

Objectively assessing the quality of fundus images is an emerging research field. Nowadays, telemedicine is being widely used for providing the medical assistance to the ophthalmologists with minimum

| | F1 | F2 | F3 | F4 | F5 | Class |
|---|---|---|---|---|---|---|
| **(a)** | 9 | 8 | 9 | 9 | 9 | Good |
| **(b)** | 7 | 7 | 6 | 6 | 5 | Average |
| **(c)** | 2 | 1 | 4 | 4 | 2 | Bad |

**Fig. 12**: Example of the quality scores obtained from the suggested SQA process. Quality scores are assumed to be in the range [1,10], where a high value represents high quality.

efforts. However, there are many factors that can affect the quality of such images, including acquisition, compression, transmission, etc. Therefore, an automated IQA of fundus images is a necessary and important requirement for a reliable diagnosis. Through this paper, a comprehensive discussion over the state of the art in retinal IQA research is presented. In summary, most of the retinal IQA methods are developed using machine learning algorithms and divide the retinal images into two classes: good and poor. Overall, the retinal IQA challenge is solved as a binary classification problem. In addition to the binary classification, the IQA models have two types of quality indicators: (i) generic quality indicators such as illumination, colour, and contrast, and (ii) structural quality indicators that indicate the visibility of the structures.

Furthermore, it has also been identified that the following limitations exist in the path of retinal IQA research: (i)lack of understanding of ophthalmologist's judging criteria, (ii) the absence of a common data-set, and (iii) the shortcomings of binary classification based approaches on the use of image enhancement methods. However, it is realized that the absence of a common benchmark retinal IQA data-set is the root of all the limitations mentioned above. In order to elevate the progress in this field, it is highly required to have such data-set that can provide the ophthalmologist's perception of the quality of fundus images in terms of quality grading. Such data-set will provide a ground for the comparative performance analysis as well as a deciding criterion for the requirement of the enhancement in the image. Conclusively, the development of an efficient retinal IQA algorithm is a very important problem to be solved. A fair amount of work is done to solve this challenge but still, it is far from being solved.

## Acknowledgment

## 7 References

1 "Methodology for the subjective assessment of the quality of television pictures," 2012. [Online]. Available: [https://www.itu.int/dmspubrec/itu-r/rec/bt/R-REC-BT.500-11-200206-S!!PDF-E.pdf]
2 I. W. S. Carolyn G Ford, Mark A. McFarland, "Subjective video quality assessment methods for recognition tasks," *Proc.SPIE*, vol. 7240, pp. 7240 – 7240 – 11, 2009.
3 L. Lévêque, H. Liu, S. Baraković, J. B. Husić, M. Martini, M. Outtas, L. Zhang, A. Kumcu, L. Platisa, R. Rodrigues *et al.*, "On the subjective assessment of the perceived quality of medical images and videos," in *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*.   IEEE), 2018, pp. 1–6.
4 A. Ninassi, L. C. Patrick, and F. Autrusseau, "Pseudo no reference image quality metric using perceptual data hiding," in *Human Vision and Electronic Imaging XI*, vol. 6057.   International Society for Optics and Photonics, 2006, p. 60570G.
5 P. Le Callet and F. Autrusseau, "Subjective quality assessment irccyn/ivc database," 2005. [Online]. Available: http://www.irccyn.ec-nantes.fr/ivcdb/
6 H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Transactions on image processing*, vol. 15, no. 11, pp. 3440–3451, 2006.
7 Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
8 N. Ponomarenko, F. Battisti, K. Egiazarian, J. Astola, and V. Lukin, "Metrics performance comparison for color image database," in *Fourth international workshop on video processing and quality metrics for consumer electronics*, vol. 27, 2009.
9 N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "Tid2008-a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, no. 4, pp. 30–45, 2009.
10 U. Engelke, M. Kusuma, H.-J. Zepernick, and M. Caldera, "Reduced-reference metric design for objective perceptual quality assessment in wireless imaging," *Signal Processing: Image Communication*, vol. 24, no. 7, pp. 525–547, 2009.
11 E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, p. 011006, 2010.
12 Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
13 H. R. Sheikh and A. C. Bovik, "A visual information fidelity approach to video quality assessment," pp. 23–25, 2005.
14 L. Zhang, L. Zhang, X. Mou, D. Zhang *et al.*, "Fsim: a feature similarity index for image quality assessment," *IEEE transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.
15 W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684–695, 2014.
16 Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," in *Human Vision and Electronic Imaging X*, vol. 5666.   International Society for Optics and Photonics, 2005, pp. 149–160.
17 X. M. Min Zhang, Wufeng Xue, "Reduced reference image quality assessment based on statistics of edge," *Proc.SPIE*, vol. 7876, pp. 7876 – 7876, 2011.
18 K. Gu, G. Zhai, W. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE transactions on cybernetics*, vol. 46, no. 1, pp. 284–297, 2016.
19 Q. Li and Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation," *IEEE journal of selected topics in signal processing*, vol. 3, no. 2, pp. 202–211, 2009.
20 X. Min, K. Gu, G. Zhai, M. Hu, and X. Yang, "Saliency-induced reduced-reference quality index for natural scene and screen content images," *Signal Processing*, vol. 145, pp. 127–136, 2018.
21 K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new reduced-reference image quality assessment using structural degradation model," *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1095–1098, 2013.
22 M. Narwaria, W. Lin, I. V. McLoughlin, S. Emmanuel, and L.-T. Chia, "Fourier transform-based scalable image quality measure," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3364–3377, 2012.
23 M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," *IEEE transactions on Image Processing*, vol. 21, no. 8, pp. 3339–3352, 2012.
24 A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE transactions on Image Processing*, vol. 20, no. 12, pp. 3350–3364, 2011.
25 A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, 2012.
26 A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer." *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, 2013.
27 P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *2012 IEEE conference on computer vision and pattern recognition*.   IEEE, 2012, pp. 1098–1105.
28 S. Bosse, D. Maniry, T. Wiegand, and W. Samek, "A deep neural network for image quality assessment," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sep. 2016, pp. 3773–3777.
29 J. Kim and S. Lee, "Fully deep blind image quality predictor," *IEEE Journal of selected topics in signal processing*, vol. 11, no. 1, pp. 206–220, 2017.
30 J. Kim, A.-D. Nguyen, and S. Lee, "Deep cnn-based blind image quality predictor," *IEEE transactions on neural networks and learning systems*, no. 99, pp. 1–14, 2018.
31 L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1733–1740.
32 S. Bosse, D. Maniry, K.-R. Müller, T. Wiegand, and W. Samek, "Deep neural networks for no-reference and full-reference image quality assessment," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 206–219, 2018.
33 L. Kang, P. Ye, Y. Li, and D. Doermann, "A deep learning approach to document image quality assessment," in *Image Processing (ICIP), 2014 IEEE International Conference on*.   IEEE, 2014, pp. 2570–2574.
34 K. Gu, G. Zhai, X. Yang, and W. Zhang, "Deep learning network for blind image quality assessment," in *Image Processing (ICIP), 2014 IEEE International Conference on*.   IEEE, 2014, pp. 511–515.
35 W. Hou, X. Gao, D. Tao, and X. Li, "Blind image quality assessment via deep learning," *IEEE transactions on neural networks and learning systems*, vol. 26,

no. 6, pp. 1275–1286, 2015.

36 J. Guan, S. Yi, X. Zeng, W. Cham, and X. Wang, "Visual importance and distortion guided deep image quality assessment framework," *IEEE Transactions on Multimedia*, vol. 19, no. 11, pp. 2505–2520, Nov 2017.

37 P. Mohammadi, A. Ebrahimi-Moghadam, and S. Shirani, "Subjective and objective quality assessment of image: A survey," *arXiv preprint arXiv:1406.7799*, 2014.

38 D. M. Chandler, "Seven challenges in image quality assessment: past, present, and future research," *ISRN Signal Processing*, vol. 2013, 2013.

39 J. Kim, H. Zeng, D. Ghadiyaram, S. Lee, L. Zhang, and A. C. Bovik, "Deep convolutional neural models for picture-quality prediction: Challenges and solutions to data-driven image quality assessment," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 130–141, 2017.

40 G. Yue, C. Hou, K. Gu, N. Ling, and B. Li, "Analysis of structural characteristics for quality assessment of multiply distorted images," *IEEE Transactions on Multimedia*, vol. 20, no. 10, 2722–2732, Oct 2018.

41 G. Yue, C. Hou, T. Zhou, and X. Zhang, "Effective and efficient blind quality evaluator for contrast distorted images," *IEEE Transactions on Instrumentation and Measurement*, pp. 1–9, 2018.

42 H. Yang, Y. Fang, and W. Lin, "Perceptual quality assessment of screen content images," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4408–4421, 2015.

43 L. Kang, P. Ye, Y. Li, and D. Doermann, "A deep learning approach to document image quality assessment," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 2570–2574.

44 G. Yue, C. Hou, K. Gu, S. Mao, and W. Zhang, "Biologically inspired blind quality assessment of tone-mapped images," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2525–2536, March 2018.

45 G. Yue, C. Hou, Q. Jiang, and Y. Yang, "Blind stereoscopic 3d image quality assessment via analysis of naturalness, structure, and binocular asymmetry," *Signal Processing*, vol. 150, pp. 204–214, 2018.

46 G. Yue, C. Hou, K. Gu, T. Zhou, and G. Zhai, "Combining local and global measures for dibr-synthesized image quality evaluation," *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 2075–2088, April 2019.

47 P. C. Cosman, R. M. Gray, and R. A. Olshen, "Evaluating quality of compressed medical images: Snr, subjective rating, and diagnostic accuracy," *Proceedings of the IEEE*, vol. 82, no. 6, pp. 919–932, 1994.

48 M. Razaak, M. G. Martini, and K. Savino, "A study on quality assessment for medical ultrasound video compressed via hevc," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 5, pp. 1552–1559, 2014.

49 Kaggle, "Diabetic Retinopathy Detection ," https://www.kaggle.com/c/diabetic-retinopathy-detection, 2013, [Online; accessed 10-June-2017].

50 A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, March 2013.

51 D. L. Ruderman, "The statistics of natural images," *Network: computation in neural systems*, vol. 5, no. 4, pp. 517–548, 1994.

52 J. M. P. Dias, C. M. Oliveira, and L. A. da Silva Cruz, "Retinal image quality assessment using generic image quality indicators," *Information Fusion*, vol. 19, pp. 73–90, 2014.

53 U. Sevik, C. Kose, T. Berber, and H. Erdol, "Identification of suitable fundus images using automated quality assessment methods," *Journal of biomedical optics*, vol. 19, no. 4, p. 046006, 2014.

54 J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database," *IET Image Processing*, vol. 7, no. 4, pp. 373–383, 2013.

55 J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE transactions on medical imaging*, vol. 23, no. 4, pp. 501–509, 2004.

56 M. Ortega, N. Barreira, J. Novo, M. G. Penedo, A. Pose-Reino, and F. Gómez-Ulla, "Sirius: a web-based system for retinal image analysis," *International journal of medical informatics*, vol. 79, no. 10, pp. 722–732, 2010.

57 A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical imaging*, vol. 19, no. 3, pp. 203–210, 2000.

58 B. Al-Diri, A. Hunter, D. Steel, M. Habib, T. Hudaib, and S. Berry, "A reference data set for retinal vessel profiles," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. IEEE, 2008, pp. 2262–2265.

59 M. Fraz, P. Remagnino, A. Hoppe, A. R. Rudnicka, C. G. Owen, P. Whincup, and S. Barman, "Quantification of blood vessel calibre in retinal images of multi-ethnic school children using a model based approach," *Computerized Medical Imaging and Graphics*, vol. 37, no. 1, pp. 48–60, 2013.

60 E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, and J.-C. Klein, "Feedback on a publicly distributed database: the messidor database," *Image Analysis & Stereology*, vol. 33, no. 3, pp. 231–234, 2014.

61 T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "Diaretdb0: Evaluation database and methodology for diabetic retinopathy algorithms," *Machine Vision and Pattern Recognition Research Group, Lappeenranta University of Technology, Finland*, vol. 73, 2006.

62 T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "The diaretdb1 diabetic retinopathy database and evaluation protocol." in *BMVC*, 2007, pp. 1–10.

63 P. Prentasic, S. Loncaric, Z. Vatavuk, G. Bencic, M. Subasic, T. Petkovic, L. Dujmovic, M. Malenica-Ravlic, N. Budimlija, and R. Tadic, "Diabetic retinopathy image database (dridb): a new database for diabetic retinopathy screening programs research," in *Image and Signal Processing and Analysis (ISPA), 2013 8th*

64 L. Giancardo, F. Meriaudeau, T. P. Karnowski, Y. Li, S. Garg, K. W. Tobin, and E. Chaum, "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets," *Medical image analysis*, vol. 16, no. 1, pp. 216–226, 2012.

65 Y. Zheng, M. H. A. Hijazi, and F. Coenen, "Automated âĂŸdisease / no diseaseâĂŹ grading of age-related macular degeneration by an image mining approach," *Investigative Ophthalmology Visual Science*, vol. 53, no. 13, p. 8310, 2012.

66 E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J.-C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno *et al.*, "Teleophta: Machine learning and image processing methods for teleophthalmology," *Irbm*, vol. 34, no. 2, pp. 196–203, 2013.

67 M. Niemeijer, B. Van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sánchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu *et al.*, "Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 185–195, 2010.

68 F. Fumero, S. Alayón, J. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "Rim-one: An open retinal image database for optic nerve evaluation," in *Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on*. IEEE, 2011, pp. 1–6.

69 E. J. Carmona, M. Rincón, J. García-Feijoó, and J. M. Martínez-de-la-Casa, "Identification of the optic nerve head with genetic algorithms," *Artificial Intelligence in Medicine*, vol. 43, no. 3, pp. 243–259, 2008.

70 J. Sivaswamy, S. Krishnadas, A. Chakravarty, G. Joshi, A. S. Tabish *et al.*, "A comprehensive retinal image dataset for the assessment of glaucoma from the optic nerve head analysis," *JSM Biomedical Imaging Data Papers*, vol. 2, no. 1, p. 1004, 2015.

71 Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng, and T. Y. Wong, "Origa-light: An online retinal fundus image database for glaucoma analysis and research," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE, 2010, pp. 3065–3068.

72 S. C. Lee and Y. Wang, "Automatic retinal image quality assessment and enhancement," in *Medical Imaging 1999: Image Processing*, vol. 3661. International Society for Optics and Photonics, 1999, pp. 1581–1591.

73 M. Lalonde, L. Gagnon, and M.-C. Boucher, "Automatic visual quality assessment in optical fundus images," in *Proceedings of vision interface*, vol. 32. Ottawa, 2001, pp. 259–264.

74 D. Usher, M. Himaga, M. Dumskyj, and J. Boyce, "Automated assessment of digital fundus image quality using detected vessel area," in *Proceedings of Medical Image Understanding and Analysis*. Citeseer, 2003, pp. 81–84.

75 A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp, "Automated assessment of diabetic retinal image quality based on clarity and field definition," *Investigative ophthalmology & visual science*, vol. 47, no. 3, pp. 1120–1125, 2006.

76 A. Hunter, J. A. Lowell, M. Habib, B. Ryder, A. Basu, and D. Steel, "An automated retinal image quality grading algorithm," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011, pp. 5955–5958.

77 T. Köhler, A. Budai, M. F. Kraus, J. Odstrčilik, G. Michelson, and J. Hornegger, "Automatic no-reference quality assessment for retinal fundus images using vessel segmentation," in *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*. IEEE, 2013, pp. 95–100.

78 H. A. Nugroho, T. Yulianti, N. A. Setiawan, and D. A. Dharmawan, "Contrast measurement for no-reference retinal image quality assessment," in *Information Technology and Electrical Engineering (ICITEE), 2014 6th International Conference on*. IEEE, 2014, pp. 1–4.

79 M. Himaga, D. Usher, and J. F. Boyce, "Retinal blood vessel extraction by using multi-resolution matched filtering and directional region growing segmentation," in *IAPR workshop on machine vision applications*, 2002, pp. 244–247.

80 C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson, "Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images," *British Journal of Ophthalmology*, vol. 83, no. 8, pp. 902–910, 1999.

81 A. Hunter, J. Lowell, R. Ryder, A. Basu, D. Steel *et al.*, "Tram-line filtering for retinal vessel segmentation," in *3rd European Medical and Biological Engineering Conference, November*. International Federation for Medical and Biological Engineering, 2005.

82 X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Transactions on image processing*, vol. 19, no. 12, pp. 3116–3132, 2010.

83 M. Al-Rawi, M. Qutaishat, and M. Arrar, "An improved matched filter for blood vessel detection of digital retinal images," *Computers in Biology and Medicine*, vol. 37, no. 2, pp. 262–267, 2007.

84 M. Niemeijer, M. D. Abràmoff, and B. van Ginneken, "Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening," *Medical image analysis*, vol. 10, no. 6, pp. 888–898, 2006.

85 L. Giancardo, M. D. Abràmoff, E. Chaum, T. Karnowski, F. Meriaudeau, and K. Tobin, "Elliptical local vessel density: a fast and robust quality metric for retinal images," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. IEEE, 2008, pp. 3534–3537.

86 F. Zana and J.-C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE transactions on image processing*, vol. 10, no. 7, pp. 1010–1019, 2001.

87 J. Paulus, J. Meier, R. Bock, J. Hornegger, and G. Michelson, "Automated quality assessment of retinal fundus photos," *International journal of computer assisted radiology and surgery*, vol. 5, no. 6, pp. 557–564, 2010.

88 "Haralik texture features." [Online]. Available: http://murphylab.web.cmu.edu/publications/boland/bolandnode26.html

89 R. Pires, H. F. Jelinek, J. Wainer, and A. Rocha, "Retinal image quality analysis for automatic diabetic retinopathy detection," in *Graphics, Patterns and Images (SIBGRAPI), 2012 25th SIBGRAPI Conference on*. IEEE, 2012, pp. 229–236.

90 H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*. Springer, 2006, pp. 404–417.

91 H. Yu, C. Agurto, S. Barriga, S. C. Nemeth, P. Soliz, and G. Zamora, "Automated image quality evaluation of retinal fundus photographs in diabetic retinopathy screening," in *Image analysis and interpretation (SSIAI), 2012 IEEE southwest symposium on*. IEEE, 2012, pp. 125–128.

92 H. Yu, S. Barriga, C. Agurto, G. Zamora, W. Bauman, and P. Soliz, "Fast vessel segmentation in retinal images using multi-scale enhancement and second-order local entropy," in *Medical Imaging 2012: Computer-Aided Diagnosis*, vol. 8315. International Society for Optics and Photonics, 2012, p. 83151B.

93 G. J. Katuwal, J. Kerekes, R. Ramchandran, C. Sisson, and N. Rao, "Automatic fundus image field detection and quality assessment," in *Image Processing Workshop (WNYIPW), 2013 IEEE Western New York*. IEEE, 2013, pp. 9–13.

94 F. Yin, D. W. K. Wong, A. P. Yow, B. H. Lee, Y. Quan, Z. Zhang, K. Gopalakrishnan, R. Li, and J. Liu, "Automatic retinal interest evaluation system (aries)," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 2014, pp. 162–165.

95 F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas, "The blur effect: perception and estimation with a new no-reference perceptual blur metric," in *Human vision and electronic imaging XII*, vol. 6492. International Society for Optics and Photonics, 2007, p. 64920I.

96 H. Davis, S. Russell, E. Barriga, M. Abramoff, and P. Soliz, "Vision-based, real-time retinal image quality assessment," in *Computer-Based Medical Systems, 2009. CBMS 2009. 22nd IEEE International Symposium on*. IEEE, 2009, pp. 1–6.

97 J. M. P. Dias, C. M. Oliveira, and L. A. da Silva Cruz, "Retinal image quality assessment using generic image quality indicators," *Information Fusion*, vol. 19, pp. 73–90, 2014.

98 D. Veiga, C. Pereira, M. J. Ferreira, L. Gonçalves, and J. Monteiro, "Quality evaluation of digital fundus images through combined measures," *Journal of Medical Imaging*, vol. 1, no. 1, p. 014001, 2014.

99 Z. Yao, Z. Zhang, L.-Q. Xu, Q. Fan, and L. Xu, "Generic features for fundus image quality evaluation," in *e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on*. IEEE, 2016, pp. 1–6.

100 S. Wang, K. Jin, H. Lu, C. Cheng, J. Ye, and D. Qian, "Human visual system-based fundus image quality assessment of portable fundus camera photographs," *IEEE transactions on medical imaging*, vol. 35, no. 4, pp. 1046–1055, 2016.

101 F. Shao, Y. Yang, Q. Jiang, G. Jiang, and Y.-S. Ho, "Automated quality assessment of fundus images via analysis of illumination, naturalness and structure," *IEEE Access*, vol. 6, pp. 806–817, 2018.

102 D. Mahapatra, P. K. Roy, S. Sedai, and R. Garnavi, "A cnn based neurobiology inspired approach for retinal image quality assessment," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 1304–1307.

103 F. Yu, J. Sun, A. Li, J. Cheng, C. Wan, and J. Liu, "Image quality classification for dr screening using deep learning," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2017, pp. 664–667.

104 M. J. Swain and D. H. Ballard, "Color indexing," *International journal of computer vision*, vol. 7, no. 1, pp. 11–32, 1991.

105 L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1733–1740.

106 J. Kim and S. Lee, "Fully deep blind image quality predictor," *IEEE Journal of selected topics in signal processing*, vol. 11, no. 1, pp. 206–220, 2017.

107 J. Kim, H. Zeng, D. Ghadiyaram, S. Lee, L. Zhang, and A. C. Bovik, "Deep convolutional neural models for picture-quality prediction: Challenges and solutions to data-driven image quality assessment," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 130–141, 2017.

108 R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Computer vision and pattern recognition, 2009. cvpr 2009. ieee conference on*. IEEE, 2009, pp. 1597–1604.

109 G. D. Joshi and J. Sivaswamy, "Colour retinal image enhancement based on domain knowledge," in *Computer Vision, Graphics & Image Processing, 2008. ICVGIP'08. Sixth Indian Conference on*. IEEE, 2008, pp. 591–598.

110 M. A. Fadzil, H. A. Nugroho, H. Nugroho, and I. L. Iznita, "Contrast enhancement of retinal vasculature in digital fundus image," in *Digital Image Processing, 2009 International Conference on*. IEEE, 2009, pp. 137–141.

111 K. Noronha, J. Nayak, and S. Bhat, "Enhancement of retinal fundus image to highlight the features for detection of abnormal eyes," in *TENCON 2006. 2006 IEEE Region 10 Conference*. IEEE, 2006, pp. 1–4.

112 T. Shimahara, T. Okatani, and K. Deguchi, "Contrast enhancement of fundus images using regional histograms for medical diagnosis," in *SICE 2004 Annual Conference*, vol. 1. IEEE, 2004, pp. 650–653.

113 H. Ab Rahim, A. S. Ibrahim, W. M. D. W. Zaki, and A. Hussain, "Methods to enhance digital fundus image for diabetic retinopathy detection," in *Signal Processing & its Applications (CSPA), 2014 IEEE 10th International Colloquium on*. IEEE, 2014, pp. 221–224.

114 C.-Y. Lu, B.-Z. Jing, P. P. Chan, D. Xiang, W. Xie, J. Wang, and D. S. Yeung, "Vessel enhancement of low quality fundus image using mathematical morphology and combination of gabor and matched filter," in *Wavelet Analysis and Pattern Recognition (ICWAPR), 2016 International Conference on*. IEEE, 2016, pp. 168–173.

115 U. Farooq and N. Y. Sattar, "Improved automatic localization of optic disc in retinal fundus using image enhancement techniques and svm," in *Control System, Computing and Engineering (ICCSCE), 2015 IEEE International Conference on*. IEEE, 2015, pp. 532–537.

116 M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color retinal image enhancement based on luminosity and contrast adjustment," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 3, pp. 521–527, 2018.