Multimedia Tools and Applications
https://doi.org/10.1007/s11042-019-07751-6

Retinal image quality assessment for diabetic retinopathy screening: A survey

Jiawen Lin [1,2,3] · Lun Yu [1] · Qian Weng [2] · Xianghan Zheng [2,3]

Received: 21 December 2018 / Revised: 12 March 2019 / Accepted: 9 May 2019
Published online: 30 May 2019

Abstract
Retinal image quality assessment (RIQA) is one of the key components in screening for diabetic retinopathy (DR). As one of the most serious complications of diabetes, DR has become a leading cause of blindness in adults globally. DR screening is essential to achieve early diagnosis so that effective treatment could be provided timely. However, the collected images of medically unsatisfactory quality always lead to failure of diagnosis and waste of ophthalmologists' precious time. Hence, the first step in a good DR screening program is verifying retinal images of good quality. In this paper, we provide a systematic review on automated assessment of retinal image quality for DR screening. Scheme and parameters for RIQA are firstly presented. Next, we provide detailed understanding of the existing RIQA techniques, algorithms and methodologies, including brief description and analysis of each existing state-of-art approaches and comparison between such methods. Datasets and evaluation metrics are also illustrated. Finally, several challenges and future research directions are summarized and discussed.

Keywords Diabetic retinopathy · Screening · Retinal image · Retinal image quality assessment

1 Introduction

The latest report from the International Diabetes Federation says that there are about 425million adults with diabetes over the world [5]. As the most common and serious complication of diabetes, diabetic retinopathy (DR) may attack over half of individuals with diabetes. The incidence of DR reach 98% in people with diabetes more than 15 years [48]. DR is a silent

✉ Jiawen Lin
rose_wei7@126.com

[1] College of Physics and Information Engineering, Fuzhou University, Fujian, China

[2] College of Mathematics and Computer Science, Fuzhou University, Fujian, China

[3] Fujian Key Laboratory of Network Computing and Intelligent Information Processing (Fuzhou University), Fujian, China

ocular disease causing the growing, but potentially avoidable, vision loss. Since the early stage of silent disease are usually asymptomatic, patients are always aware of the disease in the late stage after suffering severe vision damage. It is too late to prevent blindness. With the increasing number of adults with diabetes, DR has now become a leading cause of blindness globally [5]. Early diagnosis, accurate evaluation of DR and proper eye care are highly effective in preventing visual loss [8]. Consequently, regular DR screening based on the digital fundus photography technique is highly recommend to limit the increasing visual impairments [8]. It has been reported that early detection of DR with periodic screening could reduce the risk of blindness by over 50% [23]. Although DR screening benefits patients a lot, it is not practical to implement manual screening on a large worldwide population especially in rural areas and developing countries [23, 37].

The surveillance networks have been widely used recently [28], especially in the field of health monitoring. Developing health surveillance networks can helps to collect, process and disseminate health information effectively [32]. Here, telemedicine programs addressing DR can be deployed to guarantee diagnose and eye care to patients at any location [8]. Digital retinal images are collected by using the front end fundus camera from various regions, transmitted to the data center, and finally examined by experts to assess DR severity. These remote screening of DR promotes a more practical and less costly procedure. Reading fundus images is a highly skilled work. However, trained retinal image readers are always very expensive and in limited supply worldwide [47]. The whole process may be difficult and slow mainly due to the low ratio of ophthalmologists to patients. Thus, automated retinal image analysis is necessarily involved in DR screening, permitting more rapid diagnosis and maintaining cost effectiveness and accuracy [23, 26]. Moreover, DR screening with automation allows primary care physicians to devote more time for communication and education with patients [47].

Retinal images with 45 degree field of view (FOV) produced by digital fundus camera look like as those in Fig. 1. Figure 1 (a) illustrates a normal retinal image, where the optic disc is the brightest marker while the macular area is darkest. The central retinal artery and vein distribute with multiple branches supporting the retina. Microaneurysms, hemorrhage and hard exudation are the early pathological features of DR as shown in Fig. 1(b). Accurate detection of lesions from images would support effective referral. Unfortunately, not all collected retinal images are of sufficient quality [55] for reliable medical analysis. According to literature [3], 3%-30% of non-mydriatic retinal images were unsuitable for retinopathy grading. It was also pointed out that the proportion of inadequate quality images has been reported at 11% [47] and 20.8% [13] respectively for single field non-mydriatic images.

Some instances of bad quality retinal images are shown in Fig. 2. Out of focus, artefact, uneven illumination as well as loss of retinal structures are all make retinal images unsatisfied. Factors causing poor quality can be roughly divided into two aspects: inadequate imaging conditions (e.g. sufficient illumination, poor focus or improper operation), and patient related issues (e.g., head or eye movement, pupil dilation, wrong fixation or media opacity) [27]. Bad quality retinal images employing in DR screening may result in various disadvantages. Since only tiny visual differences may distinguish early disease symptoms such as retinal lesions from confounds, it is unable to evaluate such images medically. Misdiagnosis depend on the bad quality retinal images would contribute to delayed treatment, causing irreversible visual damages. Therefore, without identifying unacceptable retinal images immediately, either repeated screening or reviewing by ophthalmologists should be done. Ophthalmologists have to waste precious time attempting to interpret poor images. And it is inconvenient for a patient

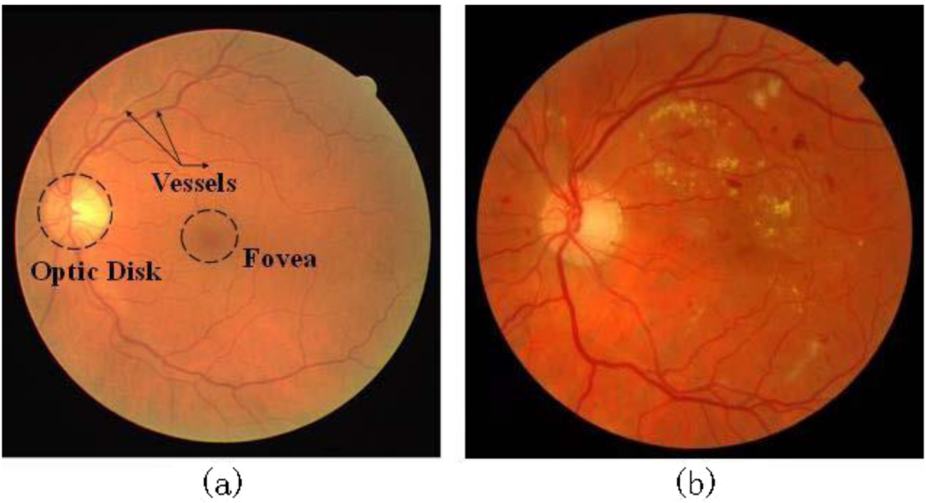

Fig. 1 Examples of good quality retinal images. (a) non-DR image, (b) image with early disease symptoms

back to the clinic for repeat retinal exams, especially when the image facility are in distant locations. Hence, in order to assure reliability of the following diagnosis, retinal image quality assessment (RIQA) becomes a preliminary preprocessing step in DR screening [47]. After determining by quality evaluation procedure, good quality fundus images are received for further analyzing, while bad quality images are discarded and re-imaging is required.

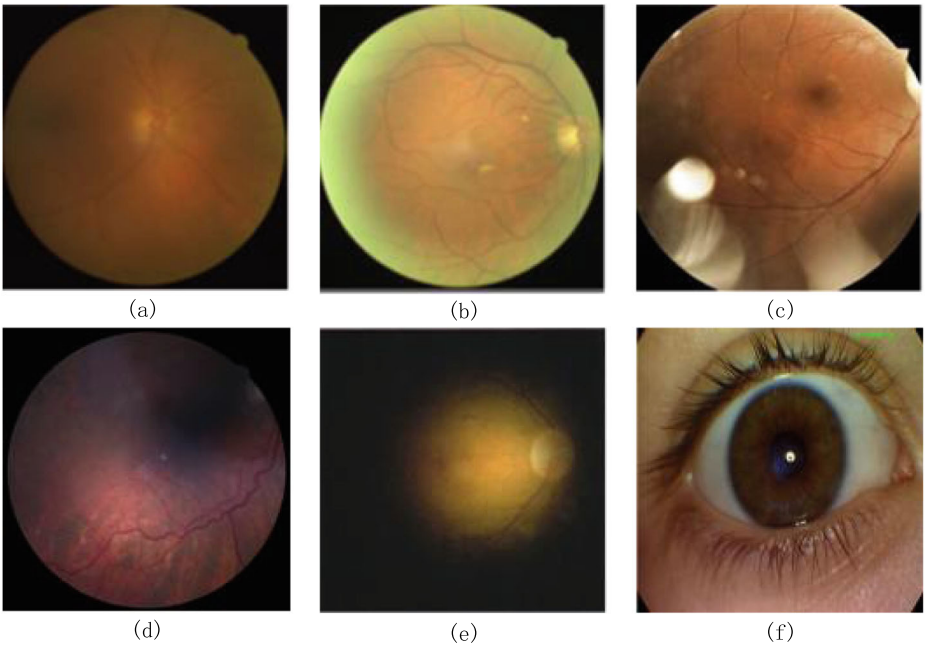

Fig. 2 Examples of bad quality retinal images. (a) out of focus, (b) uneven illumination, (c) artefact, (d) OD missing, (e) ungradable, (f) outlier

Despite that image quality assessment are well-known in the domain of image restoration [4, 39, 40, 54], deciding medical suitability of retinal image is a relativity new research field since that image quality is closely dependent on the disease types. In recent years, RIQA has become a very popular issue and gained wide attention. During the deployment of DR screening, researchers have summarized and continually improved the quality requirement of retinal images used in screening [7, 12, 13, 15]. Various techniques have been explored for evaluating retinal image quality automatically referring to kinds of parameters in the past two decades. RIQA algorithms have benefited many DR screening systems implemented all around the world [47, 49].

Therefore, existing work on RIQA is systematically reviewed in this paper. We describe the guidelines of retinal image quality evaluation, list the algorithms applied to verify quality of retinal images, discuss the pros and cons of these approaches, and illustrate challenges and suggestions for future research. Datasets which used for validating RIQA algorithms are also introduced. The remaining of this paper is structured as follows: Section 2 firstly gives a brief description of the quality standards of retinal images in DR screening and materials for analysis of performance are given in Section 3. Then, a review of existing methods employed for retinal image quality evaluation is presented in Section 4. Concise discussion of these methods and the future developments are illustrated in Section 5. Conclusions are finally drawn.

2 Guidelines of Retinal image Quality Assessment in DR Screening

The notion of RIQA is to define whether an image is of good quality and medically suitable for diagnosis or grading. Quality verification approaches should be explored following the guidelines of RIQA in DR screening. Thus, popular protocols of DR screening including factors that may impair retinal images are given.

The initial work about the grading protocol was performed in the Atherosclerotic Risk in Communities (ARIC) study [12] by the University of Wisconsin Madison. To meet the accuracy and convenience of screening, a fundus image with 45 degree field of view for each eye is required. Before evaluating the severity of abnormalities, graders were asked to assess the retinal image quality with five aspects covering focus and clarity, field definition, visibility of the optic disc, visibility of the macula and artefacts. Details of the quality classification are described in Table 1 briefly.

Since the quality grading protocol of retinal image by ARIC is proposed only for manual evaluation, some constraints are not so specific but subjective. Fleming et al. [13] defined the quality grading scheme of retinal image from two aspects, image clarity and filed definition, which are presented in Table 2 and Table 3 respectively.

Owing to the similarities in color and shape between vessels and microaneurysms, image clarity is measured by detecting vessels in macular area to estimate whether there are enough details to satisfy the requirement of early abnormalities of DR. As presented in Table 2, if third-generation branches of small vessels within one ODD around the macula area can be identified, pictures are graded of sufficient clarity. According to the Health Technology board for Scotland [7], images with 45 degree field of view should take a complete OD and temporal arcades, and the macular center should be at least two ODD from the edge of the field of view. The requirements were further refined by Fleming et al [13], as displayed in Table 3. Sufficient field definition means all the constraints list in the first two rows refer to Fig. 3 are meet.

Table 1 Description of retinal image quality classification in ARIC study [40]

Parameter	Quality grading details	Description
Focus and clarity	Good, fair, borderline, inadequate, cannot grade	This parameter is graded as inadequate when image is not clear enough for experts to detection arteriolar narrowing and subtle abnormalities.
Field definition	Good, fair, poor, cannot grade	It is measured on the correct position of the OD and the macula. The constraints are based on the metrics illustrated in Fig. 3. If optic disk is at least 1 Optic Disc Diameter (ODD) from nasal edge and fovea is contained, the image is decided qualified.
Visibility of the OD	Visible, obscured or missing	It is assessed as poor when disc is either obscured or partly missing.
Visibility of the macula	Visible, obscured or missing	It is assessed as poor when macular area is either obscured or missing.
Artefacts	Present, absent	The following artefacts are considered: aze, dust and dirt, uneven illumination over OD, macula or edge, and total blink

In 2017, experts in China recommend retinal diseases screening protocols [15] requiring carefully aligned images to include defined regions of the retina. According to the protocols, in a 45 degree macula centered image, the optic nerve head should be positioned in the midline, one disc diameter form the edge of the image field. When using the non-mydriatic fundus camera, shadows may appear in the macular area or the edge of image field because of pupil constriction at the moment of exposure. Images like Fig. 2(e) will lead to failure of diagnosis. Therefore, in addition to the clarity and field definition, the readable range of retinal image [15] is also taken into consideration in this study. The readable part should reach more than 80% of the whole image.

Overall, indicators mentioned above can be categorized into two classes: generic criteria and structural criteria. The former ones contain common image features, as focus, illumination, contrast and artefacts. And the latter ones are related to eye structures, including field definition, readable range, visibility of the optic disc and the macula, and the visibility of temporal arcades.

3 Materials for method development and performance analysis

3.1 Publicly available retinal image datasets

For the development and test of the RIQA algorithm, public databases and proprietary datasets are employed in studies presented in this survey. The frequently used public databases for

Table 2 Grading scheme for clarity of retinal image [13]

Quality grade	Description
Excellent	Small vessels are clearly visible and sharp within one ODD around the macula. The nerve fiber layer is visible.
Good	Either small vessels are clearly visible but not sharp within one ODD around the macula or the nerve fiber layer is not visible.
Fair	Small vessels are not clearly visible within one ODD around the macula but are of sufficient clarity to identify third generation branches within one optic disc diameter around the macula.
Inadequate	Third-generation branches within one ODD are invisible.

Table 3 Grading scheme for field definition of 45°retinal image [13]

Quality grade	Description of constraints
Excellent	Both the entire OD and macula are visible. The macula is in the center of the image horizontally and vertically.
Good	$D_{OD\_E} > 0.5$ ODD $D_{FOVEA\_E} > 2$ ODD $D_{ARCADE(S)} > 2.1$ ODD $D_{ARCADE(I)} > 2.1$ ODD $-5.7° < \theta_{OD\_FOVEA} < 24.7°$
Inadequate	Either a small-pupil artifact is present, or at least one of the macula, optic disc, superior temporal arcade, or inferior temporal arcade is incomplete

developing RIQA methods are described briefly in Table 4. Most of these public retinal image datasets are established from DR screening program. They consist of macula centered images with 45 degree field of view, corresponding to the standard mentioned above.

Only DRIMDB is specially set up for quality assessment. It includes 125 retinal images of good quality, 69 images of bad quality and 22 outlier images, which were labelled by ophthalmologists in advance. All the Other databases cannot be used for RIQA directly due to the lack of manual quality information. Most images in these datasets take good quality but some fundus images are unqualified. Hence, researchers always examined and categorized these images into different groups, such as 'accept', 'reject' or 'ungradable', before promoting quality evaluation of retinal images.

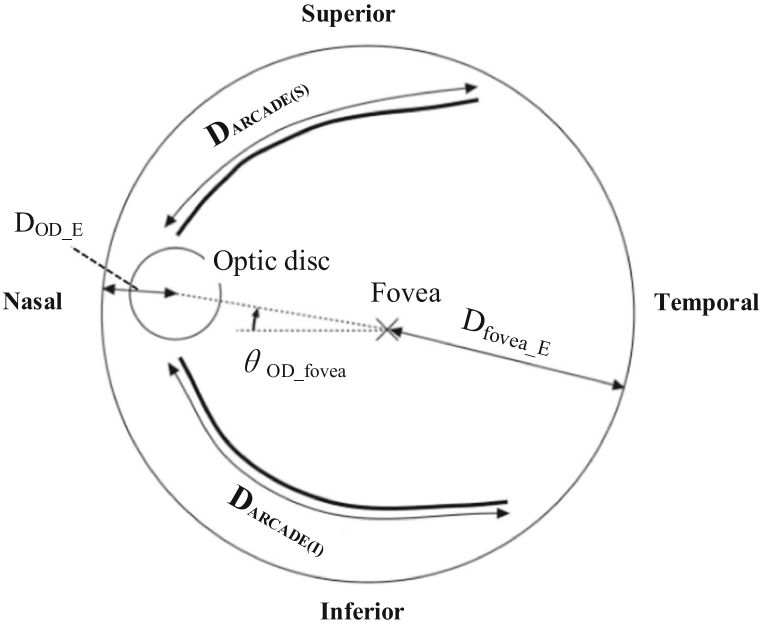

Fig. 3 The field definition metrics of retinal image with 45°field of view [13]. D_{OD_E} represents the shortest distance from the center point of OD to the edge of the circular image. D_{FOVEA_E} is the distance from the center of macular to the nearest point on the edge of the circular image. θ_{OD_fovea} means the angle of the line joining the center of the optic disc to the fovea. $D_{ARCADE(S)}$ and $D_{ARCADE\,(I)}$ are the length of the superior arcade and the one of inferior arcade respectively

Table 4 Brief description of some publicly available databases

Dataset	Year	Resolution	Number	Initial purpose
STARE [19]	2000	700 × 605	400	Blood vessel segmentation
DRIVE [36]	2004	565 × 584	40	Blood vessel segmentation
MESSIDOR [10, 30]	2004	1440 × 960, 2240 × 1488, 2304 × 1536	1200	Segmentation and DR grade
ROC [38]	2009	a wide range of image resolutions	100	Microaneurysms detection
HEI-MED [17]	2012	2196 × 1958	169	Exudate detection
Kaggle [44]	2015	a wide range of image resolutions	>80000	DR grade
DRIMDB [45]	2016	570 × 769	216	Quality grade

3.2 Performance metric

The quality verification model assigns each retinal image to a specific category as ophthalmologists do. Commonly, quantitative performance evaluation of the model is in terms of Specificity (*SP*), Sensitivity (*SN*), Accuracy (*Acc*), computational time (*T*) and area under the ROC Curve (*AUC*).

Sensitivity and specificity reflect the probability of correctly distinguished unqualified and acceptable retinal images respectively, whereas accuracy measures the overall performance of classification. Higher value of these metrics indicate better performance. Metrics used in various researches are summarized in Table 5. Here, *TP* and *TN* show that the identified results are consistent with the ophthalmologists' judgments, while *FP* and *FN* suggest an opposite result between the test result and the real situation.

4 Approaches to retinal image quality assessment

As depicted in Fig. 4, RIQA is mostly regarded as a classification problem, having three steps: preprocessing, feature extraction and classification. Preprocessing step for analyzing quality of a retinal image, always consist of selection of color channel, cropping images to define field of view (FOV) or to remove useless data, as well as applying enhancing methods. The core process of RIQA is feature extraction, with which various methods differ from each other. Researchers generate feature vectors according to the relative factors mentioned in Section 2. Consequently, retinal images are classified into predefined categories for medical diagnosis.

In this section, existing approaches to verify retinal image quality will be discussed in detail. Approaches are roughly divided into two groups according to the employed parameters: generic features based approaches and structural features involved approaches. The approaches

Table 5 Performance metrics commonly used for quality assessment of retinal images

Measurement	Description
Specificity (*SP*)	$\frac{TN}{TN+FP}$
Sensitivity (*SN*)	$\frac{TP}{TP+FN}$
Accuracy (*Acc*)	$\frac{TP+TN}{TP+FP+TN+FN}$
Area under the ROC Curve (*AUC*)	The ratio of the area under the ROC curve to the area of the whole image
Computational time (*T*)	The time required for the whole quality assessment process measured in seconds.

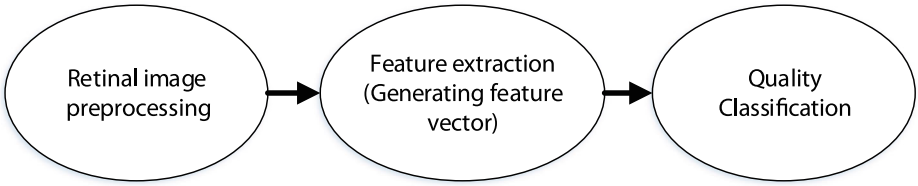

Fig. 4 General steps of RIQA

considering structural features can be further subdivided into segmentation based methods and the combinative methods avoiding segmentation procedures.

4.1 Generic features based approaches

Approaches to RIQA by using generic features are summarized in Table 6 with their results are presented in Table 7. These methods try to achieve a good quality classification with low time consumption only by applying simple image features list in ARIC study [12] but without anatomical segmentation in retinal images. The early work was mainly based on intensity histogram only, later, more and more factors have been taken into consideration.

The first method for fundus image quality assessment was proposed by Lee and Wang [25] in 1999. The authors defined an ideal template intensity histogram by analyzing 20 retinal images with excellent quality. Base on the convolution of the template intensity histogram, a quality index Q was obtained for each fundus image to measure their quality. The scale of the index Q range is 0 to 1, with the larger value meaning better quality. An image was thought to be ungradable when $Q = 0$.

Lalonde et al. [24] adopt a similar framework that a standard model of good quality should be defined with indicators derived from a set of qualified retinal images. However, it was found that the connection between histogram similarity and image quality is not so strong. Some images with poor quality would match the template histogram quite well. Thus, not only the distribution of the pixel gray value but also edge amplitude distribution were employed to characterize the focus and illumination of retinal image. Distribution of the edge magnitudes was analyzed base on the global histogram, while a region-based distribution of intensity was used instead of a global analysis.

Table 6 Generic features based methods

Authors	Year	Features Category	Dimensions of Feature Vector
Lee et al. [25]	1999	Contrast	**1**
Lalonde et al. [24]	2001	Focus, illumination	2
Bartling et al. [3]	2009	Sharpness, illumination	–
Davis et al. [9]	2009	Color, illumination, contrast	17
Dias et al. [11]	2014	Color, focus, contrast, illumination	14
Yao et al. [55]	2016	Color, Haralick textural features, Focus and sharpness: CPBD, center-surround features	113
Abdel-Hamid et al. [2]	2017	Sharpness, illumination	1
Yu et al. [58]	2017	Saliency map feature, features learned by a convolutional neural network (CNN)	5120

Table 7 Performance results of identification of unqualified retinal images by Generic features based methods

Authors	Dataset	Number of test images	*SP* (%)	*SN* (%)	*ACC* (%)	*AUC*	*T*(s)
Lalonde et al. [24] (2001)	A proprietary dataset	40	81	95	–	–	–
Bartling et al. [3] (2009)	A proprietary dataset	1000	–	–	–	–	–
Davis et al. [9] (2009)	MESSIDOR; A proprietary dataset	>2000	96	100	99	0.993	<1
Dias et al. [11] (2014)	DRIVE, ROC,STARE, MESSIDOR; Two proprietary datasets	2032	99.49	99.76	–	0.9970	11.53 (2240*1488); 1.61 (792*526); 0.31 (280*186)
Yao et al.[55] (2016)	A proprietary dataset	3224	88.73	93.08	91.38	0.9619	
Abdel-Hamid et al. [2] (2017)	DRIMDB, Two proprietary datasets	685	97.75	72	94.2	–	<0.4
Yu et al. [58] (2017)	Kaggle	2200	93.10	96.63	95.42	0.9819	–

Bartling et al. [3] developed a quantified quality evaluation algorithm only focus on sharpness and illumination of image. After being normalized in size (1024 pixels in width), the image was divided into small squares. For each square, sharpness and illumination were evaluated separately. Sharpness measurement started with a Laplacian operator [35] to exclude the squares containing insufficient content. The DWT2 using the eight-element Daubechies wavelet function [6] was then applied to the remaining sub-images. Sharpness level of each patch were determined based on the returned vandpass-filtered version. The mean value of all the squares sharpness value were calculated as the final sharpness result. A larger value represents higher sharpness. Bartling et al. analyzed illumination based on brightness and contrast. The mean intensity value and the standard deviation of each square were used to calculate the brightness and contrast value respectively. The addition of contrast and brightness were defined as illumination measurement. It was reported that 1000 retinal images were ultimately classified to one of four quality groups. The median kappa value achieved 0.64, showing great agreement with classification made by six experts.

Abdel-Hamid et al. [2] also measured retinal image quality depending on sharpness and illumination by a wavelet-based quality index. Wavelet entropy was computed as the sharpness numeric quality index by using the following equation.

$$Q_r = \frac{|En_H| + |En_V|}{|En_A|} \tag{1}$$

where En_A, En_A and En_A represent the horizontal, vertical wavelet Shannon entropies, and subbands within a certain wavelet, respectively.

Since uneven illumination would lead to an overall reduction of image quality, the index was then improved d by a homogeneity parameter from retinal image saturation channel. The saturation channel S_{retina} was created according to (2). The percentage of S_{retina} confined within the 75 middle intensity range of histogram was then defined as the homogeneity indicator P_{mid}, which was employed to get the modified quality score by (3).

$$S_{retina} = \begin{cases} \dfrac{(R-G) \times (R-B)}{R \times R}, & for\ R \neq 0 \\ 0, & for\ R = 0 \end{cases} \tag{2}$$

$$Q_{rm} = Q_r \times P_{mid} \tag{3}$$

The quality index was so computationally effective that the method was useful for a real-time DR screening system.

In other studies, more feature categories were involved. Davis et al. [9] provided a retinal image quality classification model relying on seventeen computationally simple features, which were extracted from channels in the CIELab space. Effective features were produced for seven interest regions of the retinal image, which is shown in Fig. 5. Properties used by Davis are related to three aspects of retinal image quality: color, luminance and contrast. Color measurement was performed based on the 1st and 3rd quartiles of each channel in the CIELab space. Mean intensity, skewness and kurtosis

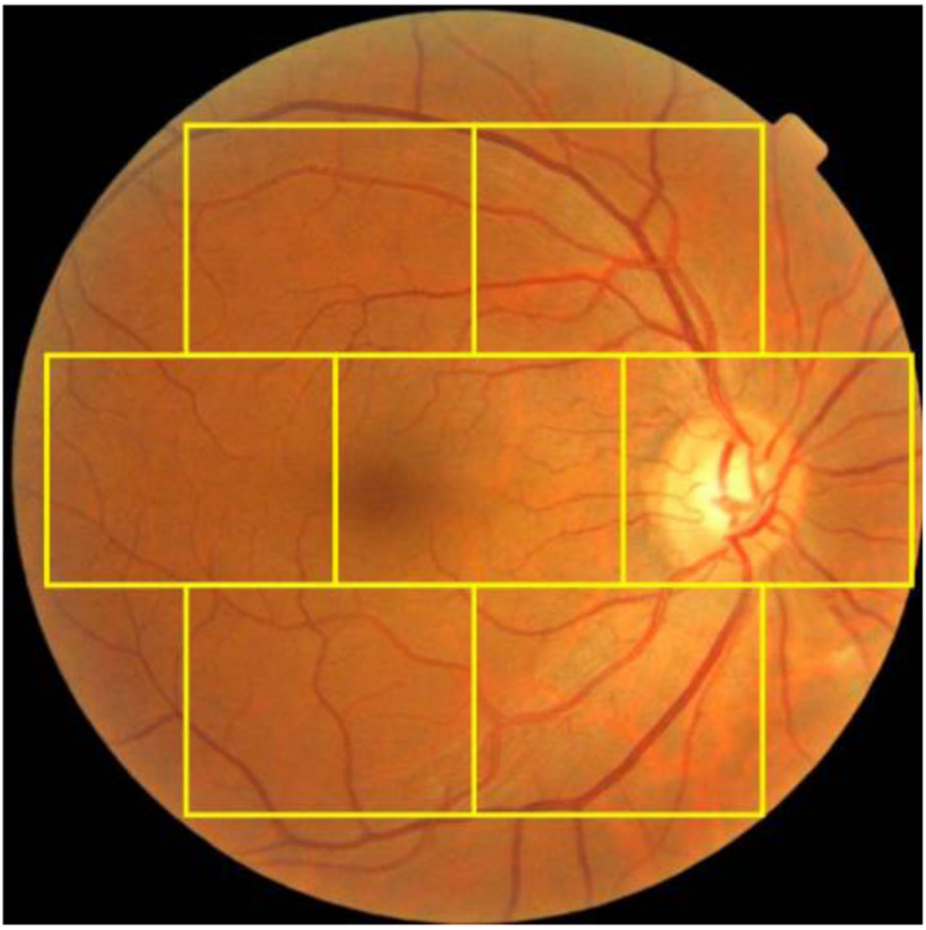

Fig. 5 Regions of interest where the features were extracted by Davis et al. [9]

were used to evaluate image luminance. Moreover, the contrast in the image was characterized by the intensity variance within each region, spatial frequency, entropy and eight Haralick features [18]. The method was tested on over 2000 images from different fundus cameras against subjective quality evaluation of fundus images by ophthalmologists. It is reported that the technique cost less than a second, and obtained 100% sensitivity and 96% specificity in finding 'rejected' images.

Method proposed by Yao et al. [55] also aimed to enable real-timeRIQA. A total of 113 characters relating to six types of common properties, comprising statistical characteristics, entropy, texture, symmetry, frequency components and blur metric, were extracted to reveal the fundus image quality. Most features mentioned in [9] were taken into consideration here, such as texture features calculated based on the co-occurrence matrix [18]. Besides, high frequency component information based on Harr wavelet, blur metrics CPBD [33] were used to form the feature vector. Statistical characteristics and texture features were especially extracted for the central region, which contains macular area in center. A SVM classifier was finally trained to pick up medically unsuitable retinal image. Images from a DR project in northeastern area

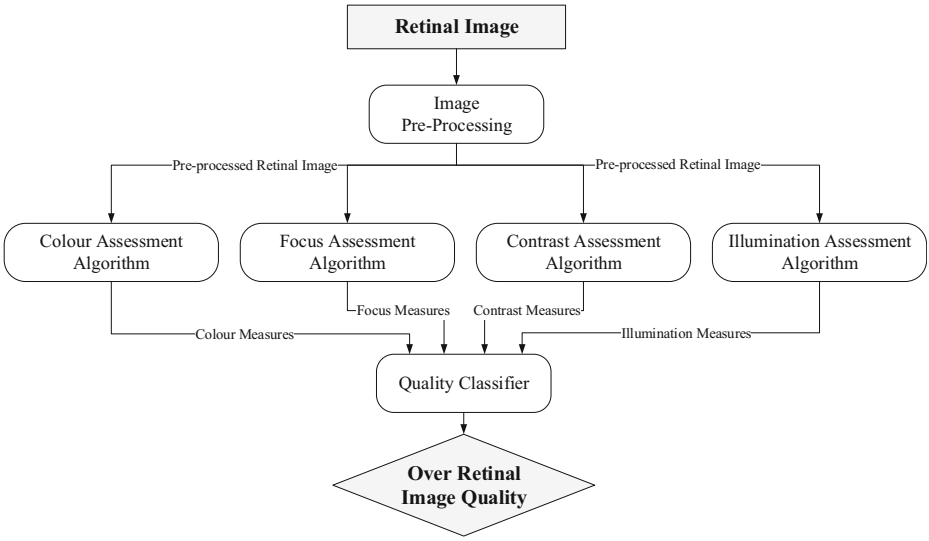

Fig. 6 Flowchart presented by Dias et al. [11]

of China were employed to verity the algorithm. The sensitivity of finding poor images was 0.93, and the accuracy was about 0.91.

As flowchart shown in Fig. 6, Dias et al. [11] assessed the quality of retinal image along the dimensions of color, focus, contrast and illumination. A Histogram back projection technique was involved in color, contrast and illumination measurements. Specific color map for each class of different aspects was calculated, composing of the relevant feature vectors. Focus assessment algorithm was established with the help of the Sobel operator and a multiscale analysis scheme. An input retinal image was estimated as ungradable or gradable by Feed-Forward Backpropagation Neural Network based on the measure values calculated by four algorithms respectively. The partial quality classification results were also provided. 2032 retinal images were selected from DRIVE, MESSIDOR, ROC and STARE public datasets to evaluate the algorithm, showing a specificity of 99.5% and a sensitivity of 99.8%.

Deep learning has been widely used in image processing and analysis recently. In 2017, Yu et al. [58] described their quality classification model based on the fused features from saliency map and a convolutional neural networks (CNN). Original retinal images are all resized to 256×256 as input images. A blurred saliency map was produced for each input image to extract feature reflecting the degree of visual attention to particular regions of the input. Following, Yu et al. applied a 8×8 window sliding on the saliency map with non-overlapping type, calculated the mean pixel value of each window and finally obtained a final saliency map feature vector of 1024 dimensions. Alexnet network was pre-trained on the Imagenet for initial weights. The training data is then used to achieve parameter fine tuning. The output of the final layer from the network is a vector with 4096 dimensions as a CNN-based feature. These feature vectors were normalized between 0 to 1repectively, and merged with 5120 dimensions totally. The multi-kernel SVM were taken to achieve final classification of retinal image quality. The algorithm would always failed when artefacts were present. A lot of qualified images were thought to be of bad quality.

Table 8 Structural features involved methods

Authors	Year	Features Category	Dimensions of Feature Vector
Usher et al. [50]	2003	Clarity: The area of detected vessels	1
Fleming et al. [13]	2006	Clarity: density and total length of detected macular vessels	1
		Field definition	5
Niemeijier et al.[37]	2006	Image structure clustering (ISC) features	5
		Color histogram	15
Giancardo et al. [16]	2008	Elliptical Local vessel density (ELVD)	28
		Color histogram	5
Paulus et al.[41]	2010	Clustering, sharpness, Haralick textural features	20
Hunter et al.[20]	2011	Clarity: Appearance of vessels within the macula region, contrast of the macular region	2
Pires et al.[43]	2012	Clarity	18
		Field definition	216
Fleming et al. [14]	2012	Vessel visibility	30
		Statistical features over CERs	432
Yu et al. [57]	2012	Vessel density	1
		Histogram features	7
		Haralick textural features	5
		local sharpness: CPBD	11
Katuwal et al. [22]	2013	Field definition	–
		Vessel density	
Yin et al. [56]	2014	Contrast and blur	1
		Entropy	1
		Structural features: blood vessel density and edge spread.	2
Sevik et al. [45]	2014	Shape: ISNT quadrant features, Zernike moments	10 (with optimization)
		Texture: Haralick features, edge histogram	
		Color: Intensity distribution, dominant color	
		Suitability metric	
Welikala et al. [53]	2016	Area, fragmentation and complexity based on measurements of segmented vessel map	3
Abdel-Hamid et al. [1]	2016	Sharpness, illumination, homogeneity, field definition, outliers	–
Feng et al. [46]	2018	Illumination, naturalness, structure indices	3

4.2 Structural features involved approaches

The methods list before only considered the general criteria of retinal images. However, important information of retina landmarks was ignored. Quality measurements derived from different image region may varied from each other. Hence, structural criteria were extracted for developing quality assessment model. The state-of-art approaches in which structural features were involved are list in Table 8 with their performance in Table 9. It is found that most approaches measure retinal image quality according to the guidelines proposed by Fleming [13]. Next, we subdivided these methods into two types dependent on whether or not segmentation procedures are required.

4.2.1 Segmentation based approaches

The first structural feature based method relying on segmentation was proposed by Usher [50] in 2003. Authors agreed the assumption that more vessels area represent better retinal image quality. Thus, retinal image clarity was described by calculating the area of detected vessels

Table 9 Performance results of identification of unqualified retinal images by structural features involved methods

Authors	Dataset	Number of test images	*SP* (%)	*SN* (%)	*Acc*(%)	*AUC*	*T*(s)
Usher et al. [50] (2003)	A proprietary dataset	1746	95	84.3	–	–	–
Fleming et al. [13] (2006)	A proprietary dataset	1039	89.4	99.1	–	–	–
Niemeijier et al. [37] (2006)	A proprietary dataset	1000	–	–	–	0.9968	About30
Giancardo et al. [16] (2008)	A proprietary dataset	84	–	–	100%(classified as good)	–	4
Paulus et al. [41] (2010)	A proprietary dataset	301	80	96.9	91.7	0.953	5.4
Hunter et al. [20] (2011)	A proprietary dataset	200	93	100	94	–	–
Pires et al. [43] (2012)	Two proprietary datasets	5776	–	–	90.8(clarity); 92.5 (field definition)	–	–
Fleming et al. [14] (2012)	A dataset used in [13]	2000	90	95.8	–	0.954	About 7
Yu et al. [57] (2012)	A proprietary dataset	1884	80	95.3	–	0.96	–
Katuwal et al. [22] (2013)	A proprietary dataset	88	–	–	–	–	–
Yin et al. [56] (2014)	A proprietary dataset	600	–	–	91.6	0.958	–
Sevik et al. [45] (2014)	DRiMDB DRIVE	216 40	–	97.6 100	98	–	>56
Welikala et al. [53] (2016)	A proprietary dataset	800	91.13	95.33	–	0.9829	13.57
Abedel-Hamid et al. [1](2016)	HRF, DRIMDB, MESSIDOR	>2692	–	–	–	0.927	2.5
Feng et al. [46] (2018)	Two datasets used in [43] DRIMDB A proprietary dataset MESSIDOR, DRIVE, STARE,	194, 4372	83.25 92.29	94.1 94.69	89.58 93.6	0.8869 0.9334	–

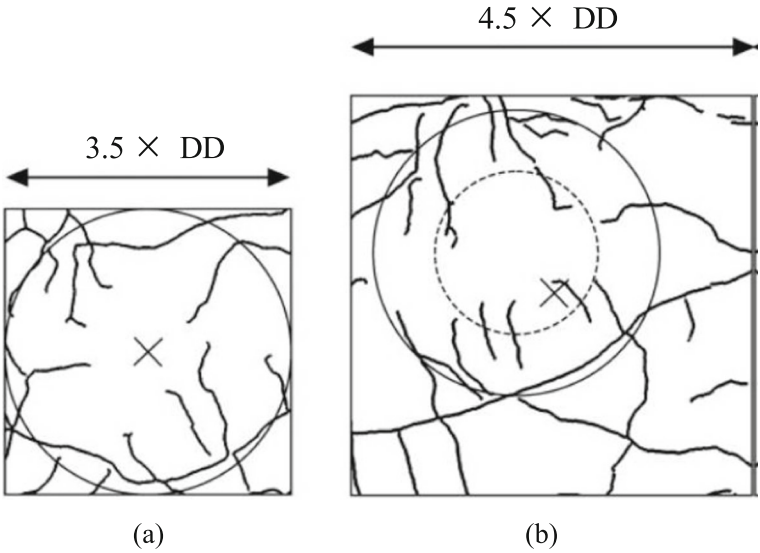

Fig. 7 Illustration of region for clarity assessment [13]. × represents detected fovea. (a) Detection region when fovea takes high contrast. (b) Detection region when fovea takes low contrast

over the whole retinal image. The algorithm reached a specificity of 95% and a sensitivity of around 85% in finding unsuitable image. However, lesions and abnormally thick vessels would make the algorithm failed.

Unlike the method proposed by Usher, Lowell et al. [41] evaluated image clarity based on vessels presence only in a circular region centered on the macula. This mind has been adopted by many subsequent studies [13, 20, 43].

Fleming et al. [13] assessed retinal image quality from two aspects: clarity and field definition. Image clarity was measured by the visibility of the vessels in a specific region, which is shown in Fig. 7. Vessels detection was applied over a circle of diameter 3.5 DD with the center on the fovea. Then total length of the detected vessels was then calculated. If the fovea took a low contrast, the size of region would expand to 4.5 DD × 4.5 DD. A retinal image was deemed to have adequate image clarity when the total length was greater than a threshold derived from the training set. Field definition was presented for the first time. A retinal image is defined as qualified when having adequately anatomical structures. Fleming et al. applied localization algorithm for both OD and macula, and calculated factors list in Table 3. As mentioned in Section 2, if all constraints were satisfied, a retinal image was assessed as having sufficient field definition. Retinal images of both adequate clarity and field definition were finally defined as acceptable images. Identification of unqualified images were achieved with specificity of 89.4% and sensitivity of 99.1%.

In Later work by Fleming et al. [14], image clarity was measured at selected regions also avoiding complicated analysis on the whole image. As Fig. 8 shown, several non-overlapping squares named clarity evaluation regions (CERs) were chosen, which center on pixels with local maxima of saliency.

Features to be evaluated for each CERs are list in Table 10, roughly divided into statistical feature set based on intensity and vessel analysis feature set based on segmentation. Here, $I^{(s)}$

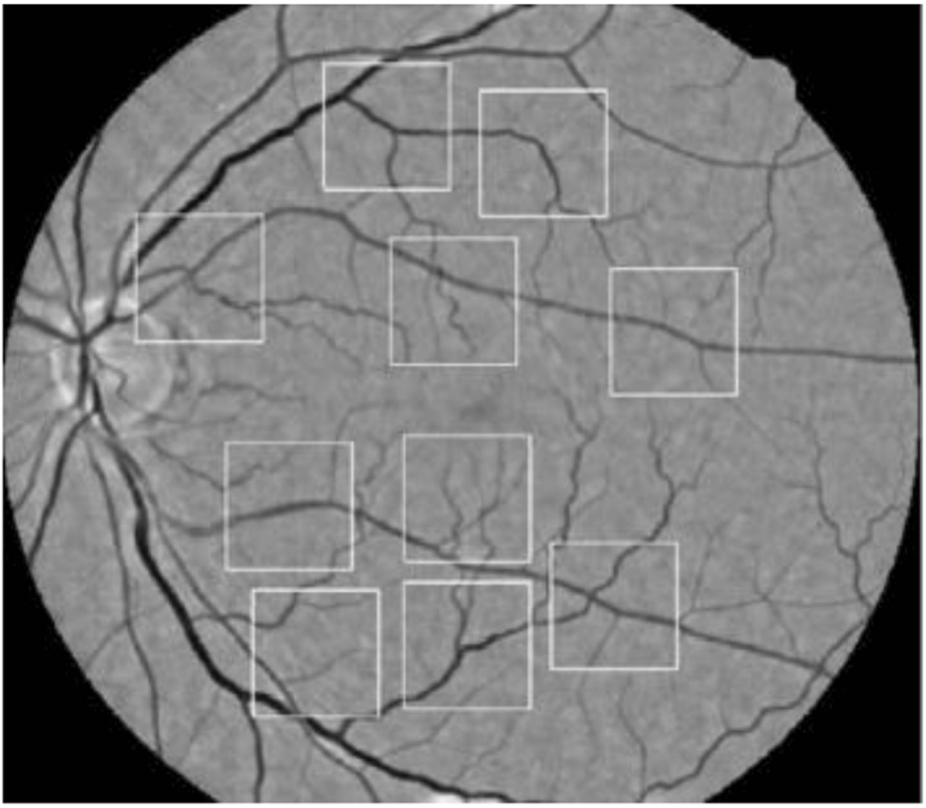

Fig. 8 Example of selected CERs [14]

means the preprocessed image and $G^{(s)}$ means the gradient magnitude at scale s. The overall clarity feature vector became available by combining measures of the individual CERs. Altogether there were 462 features composing the input vector for final classification.

Table 10 Features evaluated for each CER [14]

Name	Description	Notation
Staticstical feature set		
Mean gradient magnitude	The mean of $G^{(s)}$	$\mu_g^{(s)}$
Standard deviation of intensity	The standard deviation of $I^{(s)}$	$\sigma_i^{(s)}$
Standard deviation of gradient magnitude	The standard deviation of $G^{(s)}$	
Histogram of normalized intensity	The histogram of $I^{(s)}/\sigma_i^{(s)}$	$x_i^{(s)}(t)$
Histogram of gradient magnitude	The histogram of $G^{(s)}$	$x_g^{(s)}(t)$
Histogram of normalized gradient magnitude	The histogram of $G^{(s)}/\sigma_i^{(s)}$	$x_n^{(s)}(t)$
Vessel analysis set		
Mean positive response	The mean value of the positive pixels of V	$v_{mean}^{(s')}$
Total positive response	The sum of the positive pixels of V	$v_{sum}^{(s')}$
Mean absolute response	The mean value of $\|real\ (V)\|$	$v_{abs}^{(s')}$
Mean centerline value	The mean value of all pixel of U	$u_{mean}^{(s')}$
Centerline length	The number of positive pixels of U	$u_n^{(s')}$

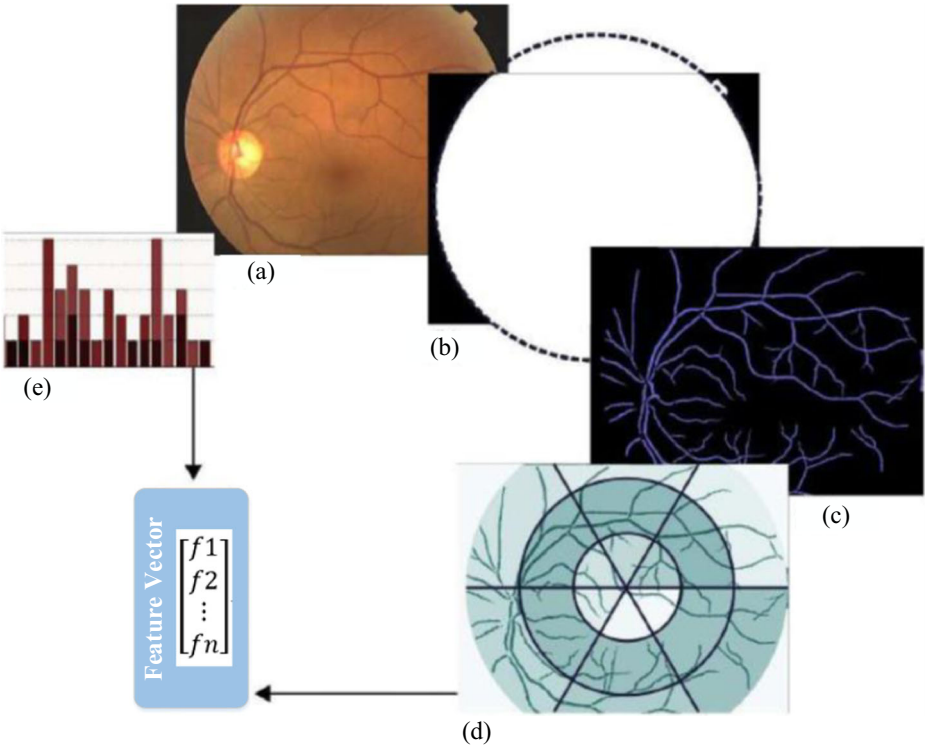

Fig. 9 Description of ELVD feature vector extraction [16]. (a) Original retinal image. (b) Fitted ellipse. (c) Segmented vessel map. (d) Local window in polar coordinates. (e) Color histogram

Similarly, Hunter et al. [20] were likely to assess retinal image quality addressing vessels visibility and contrast issue within 1 ODD of the fovea. The segmentation algorithms of OD, fovea and blood vessels were required as basic and essential steps. Then, the appearance of vessels in macula area were quantified as below:

$$v = \sum_i \frac{\eta_i \alpha_i}{\gamma_i} \tag{4}$$

where in each vessel segment i, γ_i means the average distance of pixels from fovea, α_i means the contrast and η_i represents total pixels number. Besides, the contrast k was estimated. The overall quality metric μ was defined as:

$$\mu = vk \tag{5}$$

Retinal vessels were the main source of sharp edges in retinal images. Therefore, quality of a retinal image can be partly determined by quantifying the visibility of segmented vessel tree [16, 22, 57].

Giancardo et al. [16] estimated the quality of a retinal image based on elliptical local vessel density (ELVD) feature. Figure 9 describes the extraction of ELVD feature vector. A morphological vessel segmentation technique was first applied to get a complete vessel map. The initial ellipse obtained from non-iterative least squares based algorithm was then used to build a polar coordinate system. The area of vessel in each local window was computed and kept in the feature vector. Color histogram was added to the final feature vector.

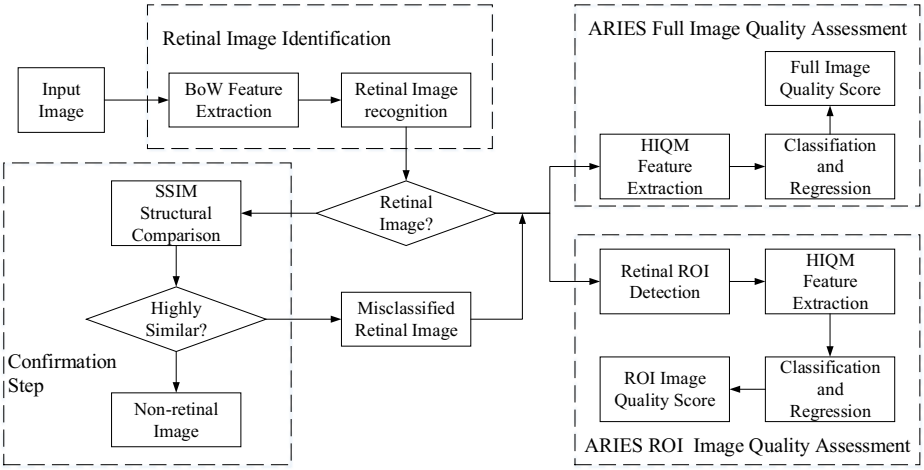

Fig. 10 Flowchart proposed by Yin et al. [56]

The quality evaluation method presented by Yu et al. [57] integrated vessel density, histogram features, textural features and local sharpness features. Inspiring from the previous work, seven statistical features of color, five features obtained from the co-occurrence matrix, and CPBD metric represented histogram features, textural information and local sharpness respectively. The proportion of the detected vessels area over the FOV was calculated as the vessel density feature. All these four features were sent to the classifier for distinguishing unqualified images.

In 2014, Yin et al. [56] assessed quality of a given image with three steps: retinal image identification, confirmation and final quality assessment. The flowchart of the RIQA system are provided in Fig. 10. Irrelevant images would be filter out in the first step, and confirmation step help to identified real fundus images requiring quality measurement. Contrast to considering low level features, authors extracted high level image quality measures (HIQM) to identify bad quality images, including totally three groups of features: contrast and blur, entropy and retina structural features. Thereinto, image structure features were characterized by blood vessel density (*BVD*) and the maximum edge spread (*ES*) through (6) and (7).

$$BVD = \frac{\sum_{i=1,j=1}^{m,n} V(i,j)}{m \times n} \tag{6}$$

$$ES = \frac{\max(l_1, l_2, l_3 \cdots l_k)}{\sqrt{m^2 + n^2}} \tag{7}$$

where m and n are the width and height of the retinal image respectively, V represents the segmented blood vessel map which was gained by applying bottom-hat filter to green channel of the image. The major axis length of every connected component in V is defined as l.

Also in 2014, Sevik et al. [45] presented a new idea about quality assessment. It kept in mind proposed by Fleming that a medically suitable image should take sufficient filed definition and clarity. OD detection and fovea detection using existing segmentation methods play basic and important role for following feature extraction. Based on previously segmented

image, the inferior, superior, nasal and temporal (ISNT) quadrant features [34] were regarded as the local information about vessel structure. Meanwhile, the Zernike moments [21] were extracted to describe the global characteristics. Haralick metrics were used to get global texture features once again, and the edge histogram described the texture locally. Color features were expressed by intensity histogram. One of the contributions in this paper is the proposed suitability metric, which is defined as follow:

$$S = |y_c - y_{OD}| \tag{8}$$

where y_c and y_{OD} represent the y component of centerline and the OD location. All the features combined the vector for retinal quality classification. Retinal image databases named DRIMDB was specially designed for RIQA in their study.

In order to generate suitable vessel data for epidemiological studies, Welikala et al. [53] designed a RIQA method based on three global features extracted from the segmented vessels, including area, fragmentation and complexity. By measuring the detected vessel map, area was obtained by counting the sum of all pixels in the map. And the number of connected component of the vessel tree represented fragmentation feature. Complexity was up to the total number of segments when original vessel map was divided and the small ones were removed.

In 2018, Shao et al. introduced a quality prediction model that focus on illumination, naturalness and structural features. For illumination evaluation, the degree of uneven illumination was calculated as below:

$$\sigma = \frac{1}{n \times n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\mu_{i,j} - \overline{\mu}\right)^2 \tag{9}$$

where $\overline{\mu}$ is the average luminance value of the whole image, $\mu_{i,\,j}$ represents the average luminance value of an un-overlapping patch with its size of 9×9. The naturalness score was pooled as the mean value of $\{q_i\}$, which is defined as below:

$$q_i = \sqrt{(\mu - \mu_i)^T \left(\frac{\sum + \sum_i}{2}\right)^{-1} (\mu - \mu_i)} \tag{10}$$

where q_i means the distance between $(\mu_i, \sum_i)$ and the pristine MVG model $(\mu, \sum)$. OD localization with vessels segmentation was adopted to evaluate the structure suitability. With these features, retinal images were assigned into 'accept' or 'reject'. The model achieved satisfied performances on several databases.

4.2.2 Combinative approaches without segmentation

Although the natural variance encountered in retinal images can be categorized by structural features extracted from previous segmentation, quality assessment may be so complicated and fail due to the incorrect identification of landmarks. Consequently, methods combining structural and generic features without previous segmentation were explored.

Niemejier et al. [37] hold that a qualified retinal image should take typical structures which could be predefined. Accordingly, a method named Image Structure Clustering (ISC) was provided to learn the structures of an image without any segmentation. By using a set of filters, response vectors were generated to characterize particular structures

in retinal image. After applying k-means clustering over all vectors, five clusters were obtained. Ultimately, taking the ISC representation together with the normalized histograms of the R, G and B color planes, totally twenty features were extracted for the classifier to identity bad quality images. The ISC technology can obtain information of local image structure and allow implicit segmentation of landmarks, which is widely applied to the following researches [41, 43].

Another new combinative approach considering generic criteria and structural criteria was proposed by Paulus et al. [41]. The authors did structure analysis similar to the ISC method, but applied k-means clustering on the pixel intensities instead of gradient information. The inter-cluster-differences were then computed to measure image contrast. Besides, sharpness metric were derived from local gradient information which can measure the distinctness between the components. For generic criteria assessment, three Haralick features were introduced. Entropy evaluated image sharpness, and energy described image homogeneity. Contrast was also involved. The method was tested with diverse feature combinations. The combination of clustering, sharpness and Haralick features help to achieve best performance when evaluating ungradable images.

Instead of focusing on a particular aberration, local structure was used to quantify the visual quality. Recently, visual similarity index was introduced to RIQA. Pires et al. [43] developed their quality verification algorithm following the rules proposed by Fleming. Blur detection as well as field definition were all taken into consideration. Authors explored four descriptors for blur detection, such as vessel area and visual dictionaries. As the major improvement, full-reference comparison of a given image with the ideal reference was operated to quantitatively measure field definition of the test image. Reference set contained 20 retinal images with good field definition and another 20 images would be abandoned. The average SSIM [52] value was calculated between the images and the reference set. As presented in [42], authors defined the retinal image quality derived from three factors: correlation between images,

contrast distortion and luminance distortion. Based on the MSSIM, a renormalization was conducted to assure the maximum value of the MSSIM metric corresponds to the best possible retinal image.

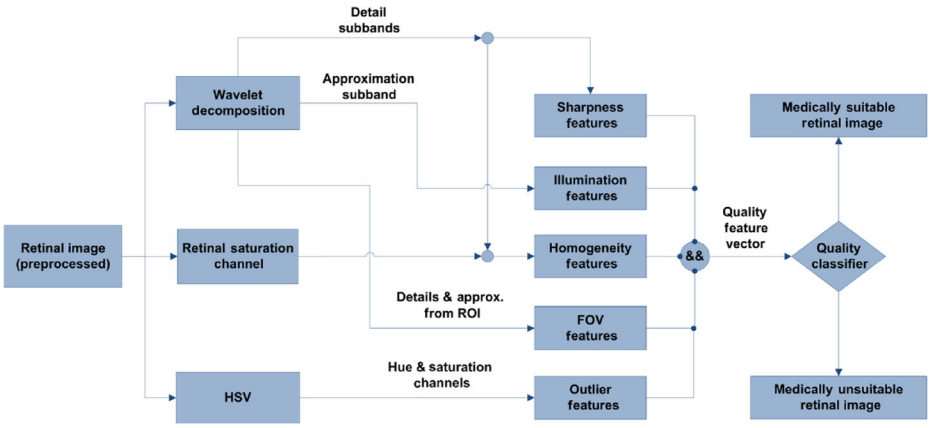

Fig. 11 Flowchart proposed by Abedel-Hamid et al. [1]

In 2016, Abedel-Hamid et al. [1] considered retinal image quality generally based on clarity and content. The flowchart is shown in Fig. 11. Following their previous opinions, homogeneity, sharpness and illumination over the entire image were measured by wavelet-based features. Moreover, authors checked whether the OD is correctly included in the image in order to achieve field definition judgment. The proposed sharpness and illumination features of some specific area were utilized.

5 Discussion and future developments

Many technologies are employed for generating features vectors. Most of color features are derived from intensity histogram. Haralick metrics are widely used to characterize the textural features of retinal images. Based on the co-occurrence matrix, three and five features were calculated by Paulus et al. [41] and Yu [57] respectively, while Davis et al. [9] and Yao [55] extract eight features for their later use. Wavelet transform (WT) has been introduced into RIQA recently. As reported in [55] and [3], WT decomposing images into subbands, is adopted for sharpness evaluation. Not only for sharpness measurement, Abdel-Hamid et al. [2, 56] also developed wavelet-based features for homogeneity and illumination assessment. Extraction of general features is common and simple, reducing computational complexity. Therefore, generic feature based approaches are suitable for platforms with limited computational power, and can be widely used in analyzing various medical images. Nevertheless, important structural characteristics of the fundus would be ignored if only general features were employed.

On the other hand, useful structural information which infers the correlation of different image regions, can help to get accurate recognition of retinal image quality. In order to distinguish medically suitable images, detection of OD and fovea is necessary for field definition [1, 13, 22, 43, 45]. The OD and fovea also help to define the interest areas for the following processing. Segmentation of blood vessels plays an important role in quantifying retinal image clarity. For example, vessels density of given regions is considered in [16], [22, 56, 57], while the length of detected vessels is calculated by Fleming [13] and Yin et al. [56]. However, high complexity of segmentation results in large time consumption. What's more, quality evaluation may fail when incorrect segmentation occur. Thus, combinative methods inspiring advantages of both generic and structural features, aim to achieve accurate and effective quality classification without pre-segmentation.

Most RIQA methods are only concerned with image quality but neglect non-retinal image identification. Since non-retinal images contain useless information for diagnosis, a negative impact on final quality assessment comes. Giancardo et al. [17] identified outlier images in their work, but only achieved detection rate by 80%. Later, Sevik et al. [45] proposed a better scheme for distinguishing medically suitable fundus images among qualified images. Also in 2014, Yin et al. [56] discarded irrelevant images by SVM classifier before extracting generic and structural features for quality assessment.

A comprehensive comparison of results of different RIQA methods was provided in previous section, but it is noteworthy that not direct conclusion can be arrived. That is because kinds of RIQA approaches were tested on different datasets unavailable, each of which includes retinal images with various resolutions and manual labeled by different experts. So far we have found that public dataset DRIMDB is set up for RIQA.

Generally, the previous methods show their strong power in finding qualified and medically suitable retinal images. However, studies on RIQA are rough and have several limitations. The following aspects still require further explore:

1) As the ultimate purpose of RIQA approach, researchers expect to embed a real-time and accurate quality prediction procedure into a DR screening system. Consequently, robustness and efficiency are the primary design criterion. The accuracy of quality prediction is highly dependent on the adopted features and classifiers. Generic and structural features should be still taken into account both and represented by more advanced techniques [51], such as sparse coding and so on. Deep learning tools are very hot in the field of computer vision now [29]. In the future, kinds of deep classifiers are more likely to be employed for RIQA.
2) For training a reliable retinal image quality classifier, massive images with label are essential. Thus, the most immediate requirement is publicly accessible datasets with varied retinal images for quality grading application [45]. It should contain much richer quality categories, including non-retinal images, and simulate different types of quality degradations, such as artifacts, noise or illumination. Each image in the dataset should be assigned to the corresponding category by several Ophthalmologists. Such publicly retinal image dataset will help researchers to develop and verify RIQA algorithms.
3) Not only RIQA but also image enhancement should be integrated within a DR screening so that it can serve the following automatic analysis well. Hence, as a subsequent process to the quality measurement, adaptive enhancement of an unsatisfied retinal image deserve more effort in the future.
4) Artifacts including fingerprints, dust and blobs caused by coughing or sneezing are often occurred in retinal images, leading to false diagnostics [31]. Artifacts should be picked out during image acquisition in order to avoid confusion between such artifacts and lesions. However, existing studies pay little attention to this problem. For a more accurate diagnosis, automatic artifact detection is worth to exploring. The results of images from a same fundus camera can also be used to check whether the equipment is in a normal condition.

6 Conclusion

In this paper, we make a systematic review of retinal image quality assessment. The frequently used guidelines and public databases for RIQA are introduced, and the state-of-art methods are revisited in detail. Particularly, we concentrate the features extracted for quality classification of these methods and further compare them through several evaluation metrics. After that, we briefly summarize the advantages and disadvantages of various approaches as well as the technologies used for feature extraction. In the future, more attention will be paid to developing public datasets, combinative features based methods and adaptive retinal image enhancement algorithms.

Acknowledgments This work is supported by the National Natural Science Foundation of China (No. 41801324), by the Natural Science Foundation of Fujian Province, China (No.2016 J0129), by the Educational Commission of Fujian Province of China (No.JAT160070).

References

1. Abdel-Hamid L, El-Rafei A, El-Ramly S, Michelson G, Hornegger J (2016) Retinal image quality assessment based on image clarity and content. J Biomed Opt 21(9):096007
2. Abdel-Hamid L, El-Rafei A, Michelson G (2017) No-reference quality index for color retinal images. Comput Biol Med 90:68–75
3. Bartling H, Wanger P, Martin L (2009) Automated quality evaluation of digital fundus photographs. Acta Ophthalmol 87(6):643–647
4. Chen F, Zeng X, Wang M (2015) Image denoising via local and nonlocal circulant similarity. J Vis Commun Image Rep 30:117–124
5. Cho N, Shaw J, Karuranga S, Huang Y, da Rocha Fernandes J, Ohlrogge A, Malanda B (2018) IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract 138:271–281
6. Cohen A, Daubechies I, Vial P (1993) Wavelets on the interval and fast wavelet transforms. Appl Comput Harm Anal 1(1):54–81
7. Cummings E, Facey K, Macpherson K, Morris A, Reay L, Slattery J (2002) Health technology assessement report. Glasgow, Scotland
8. Das T, Raman R, Ramasamy K (2015) Telemedicine in diabetic retinopathy: current status and future directions. Middle East Afr J Ophthalmol 22(2):174–178
9. Davis H, Russell S, Barriga E, Abramoff M, Soliz P (2009) Vision based, real-time retinal image quality assessment. IEEE International Symposium on Computer-Based Medical Systems, 1–6
10. Decenciere E, Zhang X, Cazuguel G, Lay B, Cochener B, Trone C, Gain P, Ordonez R et al (2014) Feedback on a publicly distributed image database: the Messidor database. Image Anal Stereol 33(3):231–234
11. Dias J, Oliveira CM, da Silva Cruz L (2014) Retinal image quality assessment using generic image quality indicators. Inform Fusion 19:73–90
12. F. P. R. C. Dept. of Ophthalmology & Visual Sciences of the University of Wisconsin-Madison (1995) ARIC grading protocol. http://eyephoto.ophth.wisc.edu/ResearchAreas.html. Accessed: 30th June 2016
13. Fleming AD, Philip S, Goatman KA, Olson JA, Sharp PF (2006) Automated assessment of diabetic retinal image quality based on clarity and field definition. Invest Ophthalmol Vis Sci 47(3):1120–1125
14. Fleming AD, Philip S, Goatman KA, Sharp PF, Olson JA (2012) Automated clarity assessment of retinal images using regionally based structural and statistical measures. Med Eng Phys 34(7):849–859
15. Fundus disease Group in Ophthalmology Branch of Chinese Medical Association (2017) Guidelines of retinal image acquisition and reading for diabetic retinopathy screening in China. Chin J Ophthalmol 53(12):890–896
16. Giancardo L, Abramoff MD, Chaum E, Karnowski T, Meriaudeau F, Tobin K (2008) Elliptical local vessel density: a fast and robust quality metric for retinal images. IEEE 30th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS): 3534–3537
17. Giancardo L, Meriaudeau F, Karnowski TP, Li Y, Garg S, Tobin KW Jr, Chaum E (2012) Exudate-based diabetic macular edema detection in fundus images using publicly available datasets. Med Image Anal 16(1):216–226
18. Haralick RM, Shanmugan K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybernet 6:610–621
19. Hoover A, Kouznetsova V, Goldbaum M (2000) Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans Med Imaging 19(3):203–210
20. Hunter A, Lowell JA, Habib M, Ryder B, Basu A, Steel D (2011) An automated retinal image quality grading algorithm. IEEE Annual International Conference of the Engineering in Medicine and Biology Society (EMBS): 5955–5958
21. Hwang SK, Kim WY (2006) A novel approach to the fast computation of zernike moments. Pattern Recogn 39(11):2065–2076
22. Katuwal GJ, Kerekes J, Ramchandran R, Sisson C, Rao N (2013) Automatic fundus image field detection and quality assessment. IEEE Western New York Image Processing Workshop (WNYIPW), 9–13.
23. Kawaguchi A, Sharafeldin N, Sundaram A, Campbell S, Tennant M, Rudnisky C (2018) Tele-Ophthalmology for Age-Related Macular Degeneration and Diabetic Retinopathy Screening: A Systematic Review and Meta-Analysis. Telemedicine and e-Health 24(4):301–308
24. Lalonde M, Gagnon L, Boucher MC et al (2001) Automatic visual quality assessment in optical fundus images. Proceedings of vision interface, Ottawa 32:259–264
25. Lee SC, Wang Y (1999) Automatic retinal image quality assessment and enhancement. Proc Int Soc Opt Photo 3661:1581–1590

26. Li B, Li HK (2013) Automated analysis of diabetic retinopathy images: principles, recent developments, and emerging trends. Curr Diab Rep 13(4):453–459
27. Li JJ, Zhang L, Peng XY (2014) Some issues of diagnosis based on ocular fundus image reading in teleophthalmology. Ophthalmol Chin 23(4):217–220
28. Li Y, Ren W, Zhu T, Ren Y, Qin Y, Jie W (2018) RIMS: A Real-time and Intelligent Monitoring System for live-broadcasting platforms. Futur Gener Comput Syst 87:259–266
29. Lin KY, Wang G (2018) Hallucinated-IQA: No-reference image quality assessment via adversarial learning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 732-741
30. MESSIDOR (2004) Methods for Evaluating Segmentation and Indexing technique Dedicated to Retinal Ophthalmology. http://www.adcis.net/en/download-Third-Party/Messidor.html. Accessed: 25 Nov. 2018
31. Mora AD, Soares J, Fonseca JM (2014) A template matching technique for artifacts detection in retinal images. International Symposium on Image and Signal Processing and Analysis (ISPA), 717-722
32. Morse SS (2013) Public health disease surveillance networks. Microbiol Spectrum 2(1):OH-0002-2012
33. Narvekar ND, Karam LJ (2011) A no-reference image blur metric based on the cumulative probability of blur detection (CPBD). IEEE Trans Image Process 20(9):2678–2683
34. Nayak J, Acharya R, Bhat PS, Shetty N, Lim TC (2009) Automated diagnosis of glaucoma using digital fundus images. J Med Syst 33(5):337
35. Nayar SK, Nakagawa Y (1989) Shape from focus. IEEE Trans Pattern Anal Mach Intell 16(8):824–831
36. Niemeijer M, Staal J, Ginneken B, Loog M, Abramoff M (2004) Drive: digital retinal images for vessel extraction. Methods for Evaluating Segmentation and Indexing Techniques Dedicated to Retinal Ophthalmology. http://www.isi.uu.nl/Research/Databases/DRIVE/. Accessed: 13 Oct. 2017
37. Niemeijer M, Abramoff MD, Ginneken B (2006) Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening. Med Image Anal 10(6):888–898
38. Niemeijer M, Van Ginneken B, Cree MJ, Mizutani A, Quellec G, Zhang B, Hornero R et al (2010) Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs. IEEE Trans Med Imaging 29(1):185–195
39. Niu Y, Lin W, Ke X, Ke L (2017) Fitting-based optimization for image visual salient object detection. IET Comput Vis 11(2):161–172
40. Niu Y, Lin L, Chen Y, Ke L (2017) Machine learning-based framework for saliency detection in distorted images. Multimed Tools Appl 76(24):26329–26353
41. Paulus J, Meier J, Bock R, Hornegger J, Michelson G (2010) Automated quality assessment of retinal fundus photos. Int J Comput Assist Radiol Surg 5(6):557–564
42. Perez J, Espinosa J, Vazquez C, Mas D (2013) Retinal image quality assessment through a visual similarity index. J Mod Opt 60(7):544–550
43. Pires R, Jelinek HF, Wainer J, Rocha A (2012) Retinal image quality analysis for automatic diabetic retinopathy detection. IEEE 25th Conf Graph Patterns Images (SIBGRAPI): 229–236
44. Pratt H, Coenen F, Broadbent DM, Harding SP, Zheng Y (2016) Convolutional neural networks for diabetic retinopathy. Proc Comput Sci 90:200–205
45. Sevik U, Kose C, Berber T, Erdol H (2014) Identification of suitable fundus images using automated quality assessment methods. J Biomed Opt 19(4):046006
46. Shao F, Yang Y, Jiang Q, Jiang G, Ho YS (2018) Automated quality assessment of fundus images via analysis of illumination, naturalness and structure. IEEE Access 6:806–817
47. Sim DA, Keane PA, Tufail A, Egan CA, Aiello LP, Silva PS (2015) Automated retinal image analysis for diabetic retinopathy in telemedicine. Curr Diab Rep 15(3):1–9
48. Sopharak A, Uyyanonvara B, Barman S (2011) Automatic microaneurysm detection from non-dilated diabetic retinopathy retinal images using mathematical morphology methods. IAENG Int J Comput Sci 38(3):295–301
49. Soto-Pedre E, Navea A, Millan S, Hernaez-Ortega MC, Morales J, Desco MC, P'erez P (2015) Evaluation of automated image analysis software for the detection of diabetic retinopathy to reduce the ophthalmologists' workload. Acta Ophthalmol 93(1):e52–e56
50. Usher D, Himaga M, Dumskyj M, Boyce J (2003) Automated assessment of digital fundus image quality using detected vessel area. Proceedings of Medical Image Understanding and Analysis: 81–84
51. Wang S, Guo W (2017) Sparse multi-graph embedding for multimodal feature representation. IEEE Trans Multimedia 19(7):1454–1466
52. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
53. Welikala R, Fraz M, Foster P, Whincup P, Rudnicka AR, Owen GG, Strachan D et al (2016) Automated retinal image quality assessment on the uk biobank dataset for epidemiological studies. Comput Biol Med 71:67–76

54. Xia Y, Leung H, Kamel MS (2016) A discrete-time learning algorithm for image restoration using a novel l2-norm noise constrained estimation. Neurocomputing 198:155–170
55. Yao Z, Zhang Z, Xu LQ, Fan Q, Xu L (2016) Generic features for fundus image quality evaluation. IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), 1–6
56. Yin F, Wong DWK, Yow AP, Lee BH, Quan Y, Zhang Z, Gopalakrishnan K, Li R, Liu J (2014) Automatic retinal interest evaluation system (ARIES). IEEE 36th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), 162–165.
57. Yu H, Agurto C, Barriga S, Nemeth SC, Soliz P, Zamora G (2012) Automated image quality evaluation of retinal fundus photographs in diabetic retinopathy screening. IEEE southwest symposium on Image Analysis and Interpretation (SSIAI), 125-128
58. Yu F, Sun J, Li A, Cheng J, Wan C, Liu J (2017) Image quality classification for dr screening using deep learning. IEEE 39th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), 664–667

Jiawen Lin was born in Fuzhou, Fujian, China in 1985. She is working as a lecture at the College of Mathematics and Computer Science, Fuzhou University, China. She received the B.S. and M.S. degrees in computer science all from Fuzhou University, Fuzhou, China, in 2007 and 2010, respectively. Now she is a Ph.D candidate at Fuzhou University. Her current research interests include artificial intelligence, computer vision and medical image analysis.

Lun Yu is a Professor at the College of Physics and Information Engineering, Fuzhou University, China. And he is the winner of special government allowance. He served as the vice-chairman of the Chinese Society of Image and Graphics, the vice-chairman of the Fujian Society of Biomedical Engineering and the managing director of the Fujian Internet Association for a long time. His research interests include artificial intelligence, signal processing, medical image analysis and telemedicine.

Qian Weng was born in Jianou, Fujian, China in 1983. He is now a lecture at the College of Mathematics and Computer Sciences, Fuzhou University, China. He received the B.S. and M.S. degrees in computer science from Fuzhou University, Fuzhou, China, in 2006 and 2009, respectively. He is currently a Ph.D candidate at Fuzhou University. His current research interests include deep learning and remote sensing image analysis.

Xianghan Zheng is a professor in the College of Mathematics and Computer Sciences, Fuzhou University, China. He received his M.S. degrees of Distributed System and Ph.D of Information Communication Technology both from University of Agder, Norway, in 2007 and 2011, respectively. His current research interests include big data processing, cloud computing services and applications.