2
Literature Survey

This chapter includes a detailed review of the existing works in the field of quality assessment and enhancement of fundus images. The first Section 2.1 includes a detailed review of fundus IQA algorithms which are divided into three categories as shown in Fig. 2.1. Section 2.2 includes the review of fundus enhancement methods. The observed limitations of state of the art in both of the research fields are discussed in Section 2.3. In the Section 2.4, a highlight of the contributions of the Thesis are provided. Finally, Section 2.5 summarizes the chapter with conclusions.

2.1 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 and Deep learning based methods, as shown in Fig. 2.1. A concise information about these algorithms, described below in chronological order, is provided in Tables 2.1, 2.2, and 2.3.

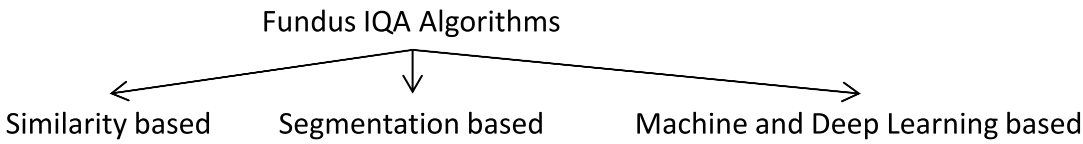

Figure 2.1 : Classification of fundus IQA algorithms

2.1.1 Similarity Based Methods

A few algorithms [Lee and Wang, 1999; Lalonde *et al.*, 2001] 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 thorough study of the related literature, Lee and Wang [1999] were the first to work on objectively assessing the quality of fundus images. 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*:

$$C = \sum_{i=0}^{255} K(i) * H(i). \tag{2.1}$$

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. 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. 2.3 shows the histogram of two different fundus images (shown in Fig. 2.2) 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, Lalonde *et al.* [2001] proposed 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.

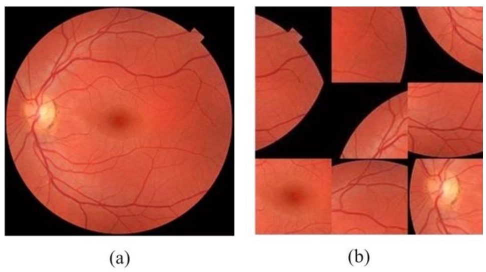

Figure 2.2 : Sample fundus images; (a) Good Quality, (b) Poor Quality.

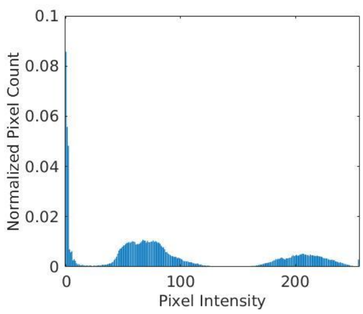

Figure 2.3 : Normalized histogram of both fundus images

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.

Table 2.1 : Summary of Similarity based fundus IQA algorithms. NS: Not Specified

Work	Quality Parameter	Categories of quality	# Images	Accuracy (%)
Lee and Wang [1999]	Intensity histogram	2	NS	NS
Lalonde *et al.* [2001]	Edge magnitude and	3	40	NS

- More reliable and efficient than NR-IQA methods.

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 [Dias *et al.*, 2014; Veiga *et al.*, 2014] 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.

2.1.2 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 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 Usher *et al.* [2003]. The authors have taken the pixel count of the blood vessels present in the image as the quality indicator; the larger the count the better the quality. In this work, the blood vessel extraction is achieved using matched filtering [Himaga *et al.*, 2002] 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.

$$Specificity = \frac{a}{a+c} \tag{2.2}$$

$$Sensitivity = \frac{b}{b+d} \tag{2.3}$$

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.

Fleming *et al.* [2006] addressed these issues, and presented a segmentation-based algorithm for fundus IQA. 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 *et al.* [2011] 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 Sinthanayothin *et al.* [1999]. Next, the segmentation of blood vessels is performed using a non-linear filtering based method termed as Tram-line algorithm [Hunter *et al.*, 2005]. 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.

Kohler *et al.* [2013] also presented a quality evaluation algorithm for fundus images, 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 [Zhu and Milanfar, 2010], a patch that can be modeled as:

$$p(x_k, y_k) = a_1(x_k - x_c)^2 + a_2(y_k - y_c)^2 \tag{2.4}$$

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 [Staal *et al.*, 2004] 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 PSNR and SSIM [Zhou Wang *et al.*, 2004] 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.

Nugroho *et al.* [2014] 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 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 [Al-Rawi *et al.*, 2007]. 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 2.5. In total 47 images from the [Giancardo *et al.*, 2012] database have been used for the experiment purpose. The reported accuracy of the proposed work is 89.36%. The proposed contrast metric is

$$C = \left| \frac{1}{x}\sum_{i=1}^{x} I_{vi} - \frac{1}{y}\sum_{i=1}^{y} I_{bi} \right| \tag{2.5}$$

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.

Welikala *et al.* [2016] presented a vascular segmentation based automated retinal image quality assessment method for epidemiological studies. This work is presented with a prime objective of epidemiological studies. It is different from the perspective of defining the diagnostic reliability of retinal images. The authors mentioned that from a view of diagnostic suitability, the whole area of a retinal image is expected to be clean and distortion free. However, in the case of epidemiological studies, the prime focus is given to the vascular area present in the retinal image. As retinal vascular morphology is one of the important indicators of the overall healthy vascular system of the human body. It has the potential to early predict various epidemiological diseases like cardiovascular, diabetes, and other systemic diseases.

Furthermore, the authors have developed a retinal image analysis system called QUARTZ (quantitative analysis of retinal vessel topology and size). The QUARTZ system is used to segment the blood vessels. Thereafter, an assessment of the segmented vasculature is performed based on three global features: area, fragmentation, and complexity. The obtained features are used to further train the support vector machine classifier to classify the retinal images into two categories of quality: accept and reject. The experiments are conducted on the randomly selected 800 images from the UK Biobank data set.

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.

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 [Fleming *et al.*, 2006] reduces the effect of Gaussian noise. Therefore, it might not produce reliable and correct quality assessment results in presence of Gaussian noise.

Table 2.2 : Summary of Segmentation based fundus IQA algorithms. SN: Sensitivity, SP: Specificity, SC: Spearman's Correlation

Work	Quality Parameter	Categories of quality	# Images	Accuracy (%)
Usher *et al.* [2003]	Blood vessel density	2	1746	SN: 95, SP: 84.3
Fleming *et al.* [2006]	Blood vessel pixel count	2	1039	SN: 99.1, SP: 89.4
Hunter *et al.* [2011]	Visibility of blood vessels	5	200	SN: 100, SP: 93
Kohler *et al.* [2013]	Blood vessel pixel count	NA	58	SC: 0.89 with PSNR SC: 0.91 with SSIM
Nugroho *et al.* [2014]	Contrast of blood vessels	2	47	89.36
Welikala *et al.* [2016]	Blood vessel area fragmentation, complexity	2	800	SN: 95.33, SP: 91.13

2.1.3 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.

2.1.3.1 Feature Extraction Based on Structural Analysis

Niemeijer *et al.* [2006] 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.

Giancardo *et al.* [2008] 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 [Zana and Klein, 2001]. 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.

Paulus *et al.* [2010] 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 [Haralik, 2020]. 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 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 Pires *et al.* [2012]. The proposed work is influenced by the work of Fleming *et al.* [2006] 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) [Bay *et al.*, 2006]. 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.

Yu *et al.* [2012a] 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 multi-scale enhancement and second order entropy threshold [Yu *et al.*, 2012b]. 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 methods. Katuwal *et al.* [2013] 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 [Lalonde *et al.*, 2001; Hunter *et al.*, 2011; Kohler *et al.*, 2013] 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 produce a quality score for the fundus images rather than simply classifying the fundus images into categories of quality. Yin *et al.* [2014] 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 [Crete *et al.*, 2007], 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 use the 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.

2.1.3.2 Feature Extraction Based on Generic Image Statistics

To the best of our knowledge, Davis *et al.* [2009] 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 [Decenciere *et al.*, 2014] 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, Dias *et al.* [2014] 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 [1991]. 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 the 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, Veiga *et al.* [2014] presented a fuzzy classification based retinal IQA algorithm. 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 [Decenciere *et al.*, 2014] 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 Yao *et al.* [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, Wang *et al.* [2016] 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 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 [Sevik *et al.*, 2014; Staal *et al.*, 2004]) have been used for the 536 fundus images for the experiment purpose.

Shao *et al.* [2018] presented a retinal IQA method based on the idea similar to [Wang *et al.*, 2016]. All the steps involved in [Wang *et al.*, 2016] and the proposed method are the same except the features that 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.

2.1.3.3 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 [Kang *et al.*, 2014a; Kim and Lee, 2017; Kim *et al.*, 2018]. The first CNN based fundus IQA algorithm was proposed by Mahapatra *et al.* [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 [Sivaswamy *et al.*, 2015] 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 7×7 followed by max pooling of 4×4 to reduce each feature map into 36×36. In the second layer of the CNN, all 20 feature maps are convolved with 50 kernels of size 5×5 followed by again 4×4 max pooling. It generates 1000 (50×20) feature maps of size 8×8. 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, Yu *et al.* [2017] 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 256×256 resolution. Thereafter the saliency maps are obtained using the frequency-tuned salient region detection method presented in [Achanta *et al.*, 2009]. Further the saliency maps are reduced to 32×32 (1024×1) blocks by taking the mean value from every 8×8 non overlapping block. Next, the 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 4096×1 features vector. Finally, the features obtained from the saliency map (1024×1) and CNN network (4096×1) are fused to create a new and unique feature vector of size 5120×1. 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 [**?**] for the experiment purpose and achieved 95.42% accuracy. Similarly, Tennakoon *et al.* [2016] also presented a shallow CNN network with four convolution and two fully connected layers for two-class retinal quality classification. Recently, Zago *et al.* [2018] and Chalakkal *et al.* [2019] have used the virtues of pretrained model architectures (GoogLeNet [Szegedy *et al.*, 2015], AlexNet [Krizhevsky *et al.*, 2017], and ResNet [He *et al.*, 2016]) to classify fundus images into two categories.

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 and outperforms the conventional IQA methods.

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.

2.2 SURVEY OF THE FUNDUS IMAGE ENHANCEMENT METHODS

With the increasing need for telemedicine and advancement in portable image acquisition devices, fundus image enhancement research has attracted the researcher's attention in recent years. However, it is important to mention that in comparison to the IQA, the volume of retinal enhancement methods is significantly less. To the best of our knowledge, [Lin and Zheng, 2002] has proposed the first retinal image enhancement method to increase blood vessel segmentation efficiency. The proposed method estimates the background (i.e., the black area) of the image and subtracts it from the original image. Let, a grey scale retinal blood vessel image $I(x, y)$ with two components: blood vessels and

Table 2.3 : Summary of Machine learning based fundus IQA algorithms divided on the basis of feature extraction approach. GIS: Generic Image Statistics, DT: Decision tree, FC: Fuzzy classification, ANN: Artificial neural network, SP: Specificity, SN: Sensitivity, MF: Membersip function.

	Work	Method	Quality Parameter	Categories of quality	Images	Accuracy (%)
Structural Analysis Based	Niemeijer *et al.* [2006]	SVM	Image structure clustering	2	2000	99.86
	Giancardo *et al.* [2008]	SVM	Blood vessel density	2	84	SN: 100, SP: 92
	Paulus *et al.* [2010]	SVM	Structural and generic features	2	301	95.3
	Pires *et al.* [2012]	SVM	Field definition and blur	2	6696	96, 95.5
	Yu *et al.* [2012a]	PLS	Vessel density histogram, texture and sharpness	2	1884	96
	Katuwal *et al.* [2013]	SVM	Symmetry of blood vessels	5	88	60
	Yin *et al.* [2014]	SVM	Contrast, blur and blood vessel density	2	370	95.8
GIS Based	Davis *et al.* [2009]	PLS	Statistical features	2	2000	SN: 100, SP: 96
	Dias *et al.* [2014]	ANN	Color, focus,contrast and illumination	2	2032	99.87
	Veiga *et al.* [2014]	FC	Uneven illumination focus	2	1454	98
	Yao *et al.* [2016]	SVM	Uneven illumination and blur	2	3224	91.3
	Wang *et al.* [2016]	DT and SVM	Uneven illumination, color, blur and contrast	2	536	94.52
	Shao *et al.* [2018]	DT, DL and SVM	Uneven illumination, naturalness property and structural information	2	4372	92.39
CNN Based	Mahapatra *et al.* [2016]	CNN	High level features	2	101	99.87
	Tennakoon *et al.* [2016]	CNN	High Level Features	2	1852	98.27
	Yu *et al.* [2017]	CNN	Fusion of features extracted from CNN and saliency maps	2	5200	95.42
	Zago *et al.* [2018]	CNN	Pre-trained Models	2	1036	97.70
	Chalakkal *et al.* [2019]	CNN	Pre-trained Models	2	7007	91.75

background. It can be modeled as:

$$I(x,y) = I_{bv}(x,y) + I_{bg}(x,y) \tag{2.6}$$

here $I_{bv}(x,y)$ and $I_{bg}(x,y)$ represent the blood vessel and background component respectively. Now, the blood vessel component can be derived by subtracting the background from the original image. The estimation of the background is achieved using the neighborhood pixel intensity information. Another work on gray scale retinal image enhancement was proposed by Feng *et al.* [2007]. The proposed method uses the contourlet transform method to estimate the noise present in the image by performing manipulations on the derived coefficient. The enhanced retinal image is then obtained using the inverse contourlet transform.

Shome and Vadali [2011] proposed a method for the contrast enhancement of the fundus images using contrast limited adaptive histogram equalization (CLAHE) method. Chen *et al.* [2016] proposed a method by first fusing the given fundus image with the background information and then applying on it the fourth-order differential equation and median filter. The mentioned background information is obtained by performing the normalized convolution operation to the given fundus images. In another work, Zhou *et al.* [2018] have presented contrast and luminosity adjustment-based enhancement of retinal images. The image's luminosity is adjusted using a luminance gain matrix derived by applying gamma correction over the value channel in HSV color space. Further, for contrast enhancement CLAHE Shome and Vadali [2011] method is applied over the luminosity channel of L*a*b* color space. A data-set of naturally distorted retinal images has been used for the performance evaluation of the proposed model.

With the similar approach, [Gupta and Tiwari, 2019] have presented a method using quantile based luminosity and contrast enhancement of fundus images. Initially, the image is divided into quantiles, followed by applying the adaptive gamma correction with the weighting distribution (AGCWD) method presented in Gupta and Tiwari [2016] for the enhancement task. Here, the quantile values were defined as the numeric values, which divides the input data into an equal proportion. In the AGCWD method, the gamma parameter is derived using the normalized probability density function (pdf) of the image histogram. You *et al.* [2019] have proposed a cycle generative adversarial networks (GAN) based model for the enhancement of fundus images. Cycle GAN is an unsupervised learning method using the GAN model that produces image output for each input image without paired training samples. The convolutional block attention module (CBAM) Woo *et al.* [2018] method is used along with the cycle GAN architecture to achieve the improved results. Authors have tested their model's performance over the images, taken from [Kaggle, 2015], that are artificially degraded using Gaussian and Perlin noise. Ghosh *et al.* [2019] proposed stacked deep convolutional denoising auto-encoders (SDCA) for fundus image denoising. SDCA is a stacked organization of multiple auto-encoders with shared layers. Patch-based training is adopted with retinal images distorted with Gaussian noise of different levels of standard deviation. In a similar approach, Biswas *et al.* [2020] used the virtues of variational autoencoders (VAE) for the restoration of retinal images.

2.3 LIMITATIONS

This section discusses the limitations that exists in state-of-the-art in fundus IQA and enhancement field separately.

2.3.1 Limitation of fundus IQA works

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.

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

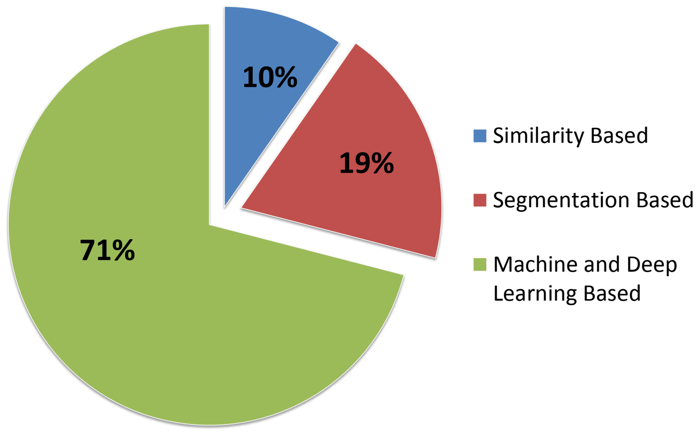

Figure 2.4 : Pie chart summarising the analysis of the fundus IQA algorithms in the literature.

Few subjective inputs: According to a study held at the University of Wisconsin-Madison, the quality of a fundus image can be assessed using the following quality parameters: focus and clarity, field definition, visibility of the structures (i.e., macula, optical disc, and blood vessels). However, there exist only a few fundus IQA works that included a subjective opinion of a medical doctor about these quality parameters. As mentioned above, in Wang *et al.* [2016] the authors have included the subjective evaluation of the fundus images using three generic quality parameters. However, the assessment of structural properties is not included and generic parameters give global quality information. To get the information about the local quality of an image, the evaluation of structural parameters is essential. Also, the ratings were collected on a scale of only two numbers (0 and 1), which is too small to identify the erroneous subjective inputs. Further, only three medical doctors participated in the subjective assessment, which also limits the generalizability of the data-set. In order to get a better understanding of the perceptual quality of a fundus image, it is essential to collect subjective opinions for both generic and structural quality parameters.

Categories of quality and scope of enhancement: In the case of medical images, the IQA process aims to find out their diagnostic usefulness. Hence, fundus IQA methods are used to classify the images into different categories of quality. As shown in Fig. 2.4, most of the fundus IQA algorithms are developed using machine learning-based classification algorithms, with the aim to classify them into two categories of quality: Good and Poor. However, in real-time imaging scenarios, there also exists a type of fundus images that neither fall into good nor in the poor category. For example, the fundus images shown in Fig. 2.5 do contain visible artifacts, but still can be used for the diagnosis by the medical doctors. Hence, it cannot be put into "Poor" category of quality. At the same time, these images might lead to wrong diagnostic results from an automated diagnosis system; hence also should not be labeled as "Good".

Recently few methods aiming at enhancing the visual quality of fundus images were published. A fully automated diagnosis system requires an effective fundus IQA algorithm that can also determine the requirement of enhancement. A binary classification based IQA method may not be able to provide such information. Hence, there must exist one more category of quality indicating an "average" or "fair" quality fundus image.

2.3.2 Limitation of fundus image enhancement methods

The fundus image enhancement methods proposed to date can be broadly divided into two categories: (i) histogram equalization based methods and (ii) deep learning based methods. The

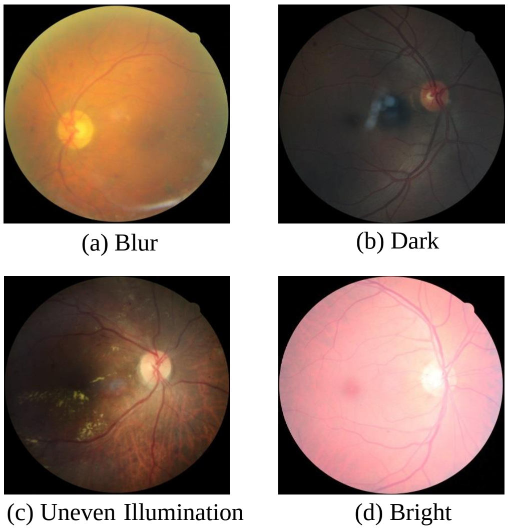

Figure 2.5 : Examples of average quality fundus images: (a) Blur, (b) Dark, (c) Uneven Illumination, and (d) Bright.

histogram equalization based methods try to enhance the image by redistributing the probability distribution of image intensities. These methods are computationally simple for implementation. However, they do not have any procedure to control the enhancement level that sometimes leads to produce over enhanced fundus images. For deep learning based methods, with the distorted images corresponding, clean images (ground truth) are typically required to compare how close the restored image resembles the ground truth one. State-of-the-art methods use additive white Gaussian noise [Ghosh *et al.*, 2019], multiplicative Gaussian noise [You *et al.*, 2019], and Impulse noise (salt and pepper) [Biswas *et al.*, 2020]. Since, such noises frequently occur, typically, in case of natural images. Use of these noises are prevalent to examine the performance of denoising and quality assessment methods developed for natural images. However, in fundus images, the existence of such noises is rare [Kaggle, 2015]. Therefore, enhancement methods developed for retinal images over such noises are of less significance and have limited scope in real conditions. Therefore, there is a requirement of an enhancement method that can perform efficiently over the distortions frequently appearing in retinal images.

It is well known that all the learning-based models, especially CNNs, require a significantly large amount of data-set. In the case of natural images, quality assessment, denoising, and enhance problems are being addressed using various publicly available data-sets [Ponomarenko *et al.*, 2015; Sheikh, 2005] containing a large amount of synthetically distorted images. Availability of these benchmark data-set lead to the development of many efficient IQA algorithms like SSIM, MAD, and CNN-IQA [Zhou Wang *et al.*, 2004; Larson and Chandler, 2010; Kang *et al.*, 2014a] etc. However, the medical image processing research field mostly suffers from delays due to the unavailability of the relevant labled data-sets. Creating such a large data-set of medical images is a challenging, laborious, and time-consuming process with heavy financial constraints. [Fu *et al.*, 2019a] made a commendable contribution by releasing a large data-set EYE-Q [Fu *et al.*, 2019a] to facilitate the fundus IQA problem. However, researchers are working on this problem for the last two decades but none of the works found prevalent due to the lack of a benchmark data-set. The release of EYE-Q data-set will definitely accelerate the research progress in the field. However, for the retinal image enhancement problem, there are no relevant data-sets available that can be used to train heavy deep learning based models. This problem leads researchers to work on the distortions, such as salt-and-pepper, and Gaussian noise, that are not strictly relevant to fundus images.

2.4 CONTRIBUTIONS

In order to address the above discussed limitations in the field of fundus IQA and enhancement, our contributions in this Thesis are as follows:

For Fundus IQA:

- A Fundus Image Quality Assessment (FIQuA) data-set of 1500 macula centered fundus images has been created, with three categories of quality: Good, Fair, and Poor. To get a clearer understanding of the ophthalmologists' perception, for each image in the data-set, subjective ratings in range of [0,10] have been collected for six quality parameters, both structural and generic. To increase the generalizability of the data-set, subjective assessment is carried out by fifteen accomplished ophthalmologists.

- A multivariate linear regression-based neural network model is proposed for the objective quality assessment of fundus images. The proposed model, trained with the help of the six subjective inputs, leads to achieving high classification accuracy.

For Fundus image enhancement:

- A total of five common degradations that appear in fair quality fundus images are identified as: (i) Uneven-illumination over macula (MUI), (ii) uneven illumination over border region (BUI), (iii) bright, (iv) dark and (v) haze.

- A total of *1000* good quality fundus images are randomly chosen as reference images from the EyeQ data-set. From these images, a data-set of *14000* degraded images resembling to the distortions mentioned above are created. Intuitively it is expected that the model trained to nullify these synthetic artifacts could potentially nullify the similar artifacts found in the naturally degraded fair quality images.

- A UNet based architecture Residual-Densely Connected UNet (RDC-UNet) is proposed for the task of enhancement. Using UNet as backbone architecture, we tried to exploit the capabilities of residual and dense connections for the enhancement task. The results obtained over synthetically degraded fundus images show that performance of the proposed model is significantly better than state-of-the-art methods.

- In case of natural images the type of distortion present is not known. Also, it is difficult to categories it prior to enhancement due to the presence of multiple such distortions at a time. To address this problem, an ensemble learning-based model is proposed. The model is built using the RDC-UNet trained individually for each of the five above mentioned distortions.

2.5 SUMMARY

- Most of the retinal IQA methods are developed using machine learning algorithms and divide the retinal images into two classes: good and poor.

- There two types of quality indicators for fundus images: (i) generic quality indicators such as illumination, colour, and contrast, and (ii) structural quality indicators that indicate the visibility of the structures

- The limitations exist in the path of retinal IQA research are: (a)lack of understanding of ophthalmologist's judging criteria, and (b) the shortcomings of binary classification based approaches on the use of image enhancement methods.

- The enhancement works reported earlier are developed for distortions, mostly caused by additive white Gaussian and salt-and-pepper noises. However, this poses a significant limitation about the applicability of these methods as occurrences of such distortions are least likely.

In the next chapter, we have presented the first contribution mentioned towards fundus IQA. It contains the peculiarities and specifications of the Fundus Image Quality Assessment (FIQuA) data-set.

...